THESE

pour l'obtention du grade de Docteur, délivré par

l'ECOLE NORMALE SUPERIEURE DE LYON

Ecole Doctorale N°52
PHAST - Physique et Astrophysique

Discipline : PHYSIQUE (Physique)

Soutenue publiquement le 29 novembre 2024, par :

Hector HUTIN

Mesure et contrôle assisté par réseau de neurones de modes micro-ondes quantiques avec des circuits supraconducteurs

Measurement and neural network-assisted control of quantum microwave modes with superconducting circuits

---

Devant le jury composé de :

Pascal DEGIOVANNI, Directeur de recherche en EPST, ENS de Lyon	Président
Denis VION, Personnalité scientifique HDR, Commisariat à l'Energie Atomique (CEA), Saclay	Rapporteur
Valentina PARIGI, Professeure des universités, Sorbonne Université	Rapporteure
Mazyar MIRRAHIMI, Directeur de recherche en EPST, INRIA Paris	Examinateur
Audrey BIENFAIT, Chargée de recherche en EPST - HDR, ENS de Lyon	Examinatrice
Irfan SIDDIQI, Professeur, University of California, Berkeley	Examinateur
Uri VOOL, Chercheur, Max Planck Institute for Chemical Physics of Solids	Examinateur
Benjamin HUARD, Professeur des universités, ENS de Lyon	Directeur de thèse



Hector Hutin
Quantum Circuit group - ENS de Lyon

MEASUREMENT AND NEURAL NETWORK-ASSISTED CONTROL OF QUANTUM MICROWAVE MODES WITH SUPERCONDUCTING CIRCUITS

MEASUREMENT AND NEURAL NETWORK-ASSISTED CONTROL OF QUANTUM MICROWAVE MODES WITH SUPERCONDUCTING CIRCUITS

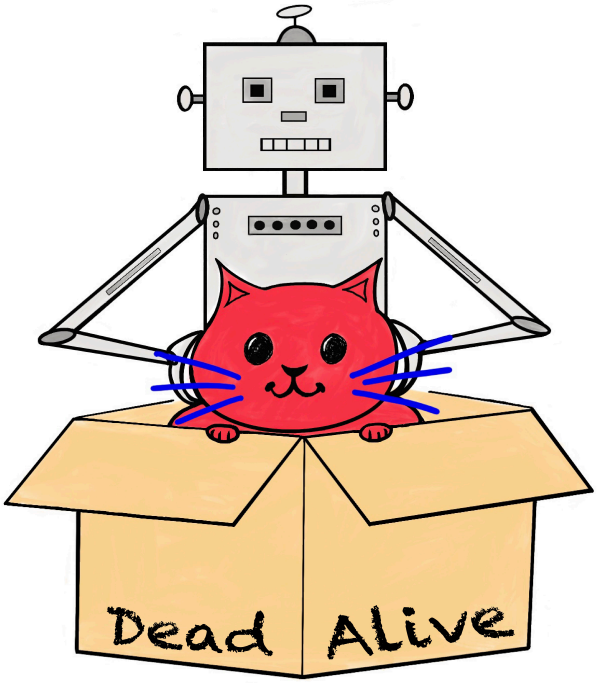
HOW TO COUNT APPLES AND PREPARE CATS

HECTOR HUTIN

Under the supervision of Benjamin HUARD & Audrey BIENFAIT

ENS de Lyon

November 2024



Robot putting Schrödinger's cat in a box
Julie Hutin BG

Cover
Design - fé studio



ABSTRACT

The association of a quantum bit and a harmonic mode dispersively coupled is one of the most studied systems in the field of quantum electrodynamics of superconducting circuits. It offers one of the most reliable current methods for measuring the state of a quantum bit in a non-destructive manner and has enabled numerous quantum information experiments, studying the information obtained through this measurement and the dynamics of the qubit when subjected to it. It also provides the ability to prepare and manipulate the quantum state of the harmonic mode via the qubit, allowing information to be stored and processed in this mode, which is fundamental to building a quantum processor. This thesis presents two experiments based on this architecture.

The first demonstrates how to use qubit fluorescence to non-destructively read the number of photons contained in a microwave cavity in a single shot, that is, before the cavity has relaxed. This makes it possible to measure photon number trajectories and observe them leave the cavity one by one. The analysis of this experiment relies on a fluorescence analysis in terms of propagating modes and addresses a number of questions about the information that these modes may contain.

The second is a demonstration of the use of a neural network for quantum state preparation in the cavity. Here, the goal is to accelerate the optimization of control electromagnetic pulses at the frequencies of the cavity and qubit to create a quantum state in the cavity with high fidelity, and this for a broad class of states. To evaluate the efficiency of this new optimization protocol over this class of states, an optimal method to assess the fidelity of the prepared states using a Wigner function measurement protocol was experimentally implemented for the first time.

RÉSUMÉ

L'association d'un bit quantique et d'un mode harmonique couplé dispersivement est l'un des systèmes les plus étudiés dans le domaine de l'électrodynamique quantique de circuits supraconducteurs. Il offre un des moyens actuels les plus fiables pour mesurer l'état d'un bit quantique de manière non-destructive, et a permis nombre d'expériences d'information quantique étudiant l'information obtenue via cette mesure et la dynamique du qubit lorsqu'il y est soumis. Il offre également la possibilité de préparer et de manipuler l'état quantique du mode harmonique grâce au qubit, ce qui permet de stocker et de manipuler de l'information dans ce mode. C'est notamment crucial pour des applications telles que la construction d'un ordinateur quantique. Dans cette thèse sont présentées deux expériences reposant sur cette architecture.

La première démontre comment utiliser la fluorescence du qubit pour lire le nombre de de photons contenus dans le mode de manière non-destructive et en un seul coup, c'est-à-dire avant que les photons soient dissipés. Cela permet de mesurer des trajectoires du nombre de photons, et de les voir quitter la cavité un par un. L'analyse de cette expérience repose sur une analyse de la fluorescense en termes de modes propageants, et permet d'aborder un certain nombre de questionnements sur l'information que peuvent contenir ces modes.

La seconde constitue une démonstration de l'utilisation d'un réseau de neurones dans le cadre de la préparation d'états quantiques dans le mode. Il s'agit ici d'accélérer l'optimisation des impulsions électromagnétiques de contrôle du mode et du qubit pour créer un état quantique dans le mode avec une grande fidélité, et ce pour une grande classe d'états. Pour évaluer l'efficacité de ce nouveau protocole d'optimisation sur l'ensemble de cette classe, une manière optimale d'évaluer la fidélité des états préparés à partir d'un protocole de mesure de la function de Wigner du mode a été pour la première fois implémentée expérimentalement.

REMERCIEMENTS

Lorsque j'ai commencé mon premier stage dans le groupe Circuits Quantiques, j'ai fait l'expérience de vertiges de plusieurs natures différentes. D'abord, ce groupe constituait la plus grand concentration de connaissances physiques que j'avais jamais vue. Je me suis mis à côtoyer des êtres aux connaissances aux limites à première vue insondables, capables de discuter de physique en toute situation avec une grande profondeur. Ces êtres m'ont appris à faire de la physique quantique à coups de fer à souder avec un enthousiasme proprement stupéfiant. De là vient la deuxième sorte de vertige : le travail que j'ai effectué dans ce laboratoire a été permis par une somme de connaissances délirante. Que ces expériences soient faisables (et cela tient parfois à une petite soudure), et les résulats prévisibles par cette théorie absurdement abstraite qu'est la physique quantique, continue de me fasciner.

Benjamin, Audrey, merci de m'avoir laissé la possibilité de goûter à ce monde lors de mon premier stage. Merci de m'avoir laissé y revenir, pour un deuxième stage (certains pensent encore qu'il ne s'est jamais terminé...), et pour l'intégralité de cette thèse. Benjamin, la physique y était certes pour quelque chose, mais le temps que tu m'as accordé, et l'enthousiasme avec lequel tu m'expliquais la physique alors que je n'étais encore "que" stagiaire, ont été pour beaucoup dans mon envie de revenir pour une thèse. Tu m'as laissé une liberté qui m'as parfois (encore) donné un petit peu le vertige, mais tu n'as jamais cessé d'être présent, prêt à discuter de mes petites découvertes. Toujours avec cet enthousiasme qui a su, à plusieurs reprises, me relancer alors que je commençais à être un peu blasé par ce que je faisais. Bref, tu as été un parfait papa de thèse. Audrey, tu en as été la maman, toujours présente, que ce soit pour m'aider à câbler le frigo (voire le faire à ma place quand il me manquait une main), pour me faire prendre du recul quand j'avais la tête dans le sac, ou pour me donner de précieux conseils sur mon après-thèse. Avoir pu te cotoyer au quotidien est une chance incroyable.

Mon apprentissage a commencé avec toi Réouven. Grâce à toi, j'ai pu commencer à faire de la physique quantique expérimentale avant même de mettre un premier pied physiquement dans le labo. Ton savoir encyclopédique nourri par ta curiosité apparemment sans limite m'ont donné d'emblée un excellent modèle. Tu disposes d'un temps apparemment illimité dès qu'il s'agit d'aider quelqu'un à résoudre un problème, ce dont j'ai profité de très nombreuses fois. Mon deuxième professeur, et mon mentor de thèse a été toi Antoine (ou plutôt Tonio). J'ai eu droit à une formation à grande vitesse sur mon sujet de thèse qui m'a fait prendre conscience de la virtuosité d'un doctorant de 3e année en prise de donnée et en analyse. Cela m'a souvent donné mal au crâne (en partie à cause de Mathematica, mais j'ai réglé ce problème depuis), signe que les informations rentrent à flux tendu. J'ai toujours religieusement suivi l'essentiel des procédures que tu m'as enseignées en nanofabrication. Aucun transmon ne m'a pété entre les mains durant cette thèse, au point que je ne suis plus convaincu aujourd'hui de l'utilité de l'ionisateur d'air. Jérémy, le duo que tu formais avec Tonio contribuait grandement à l'ambiance du groupe lorsque je suis arrivé et a grandement participé à mon envie de revenir. Merci pour tes conseils. Qu'un "ancien" me dise que ma thèse se passerait bien m'a bien aidé à prendre confiance. Et elle s'est effectivement très bien passée.

Je ne peux pas être exhaustif dans la liste de ce que les membres de ce groupe m'ont appris. Tous ont leur puits de sciences qu'ils sont prêts à partager à la moindre occasion. Alexis, discuter avec toi a été d'une richesse incroyable, que ce soit pour m'expliquer la matière condensée ou de deviser sur la signification du mot "comprendre". J'écoute toujours ton frère de temps en temps. Ambroise, tu as pris sa suite, pour me parler jonction Josephson biaisée en tension ou triangle vocalique. Yannick, tu m'as extirpé un peu du "shut up and calculate" classique du physicien blasé. Rémy, tes conseils étaient toujours précieux. Je n'oublierai pas que c'est toi qui m'as fait comprendre que non, le dieu des TLS n'était pas contre moi, mon transmon était juste mal foutu.

Impossible de parler de ces quasiment quatre années sans parler du basement crew, Arne "Babahr" Bahr et Antoine "tOny tHe tIGeR" Marquet. Ça a commencé avec une vue commune sur les Alpes, et a fini avec une vue commune sur la cour de l'ENS. Entre-temps, nous

avons trouvé le moyen de faire marcher les instruments à la menace. Seul inconvénient, l'embarras systématique à chaque visite du labo... Arne, j'ai appris une étonnante quantité de choses sur la spéléologie, la plongée, la technologie des recycleurs, les effets physiologoiques de la plongée à grande profondeur, la forge... Tes hobbies sont les plus étranges que j'aie jamais vus, mais super intéressants. Désolé de ne pas avoir continué avec autant d'enthousiasme les concours d'apnée. J'ai beaucoup réfléchi, c'est quelque chose de singulièrement désagréable et inutile. Mais c'est vrai que c'est marrant. Ah, et j'ai aussi appris de nouvelles insultes en allemand. Malgré cela, ta patience et ta minutie sont des plus impressionnantes. Antoine, crossfitter fou, d'un côté j'ai hâte de te mettre la pilule à la Saintexpress, de l'autre j'ai peur que tu me traces sans aucun entraînement spécifique alors que je me défonce depuis deux ans. Je me souviens de tes hauts et de tes bas devant ta manip, surtout un d'un moment de pure folie devant tes résultats quand ils ont commencé à devenir absurdes, des partages de mangas/manwhas. J'espère que tu n'as pas trop souffert du fouillis que je laisse systématiquement sur mon bureau. A tous les deux, merci pour toutes ces discussion, ces fous rires. Merci pour ce temps passé avec vous. J'ai enfin fini mon stage.

Merci également à tous ceux que j'ai pu côtoyer depuis que je suis arrivé : Matteo, Joel, Camille, Cyril, Yutian, Sylvain, Sergey, Chengzhi, Linda, Biswajit. J'ai apprécié chaque discussion avec vous. Le groupe s'est considérablement agrandi depuis mon arrivée, avec de nouveaux thésards : Simon, Adria, Tristan, Maxime, Evan. L'ambiance est bonne et vous avez faim de physique, j'espère que vous allez bien vous amuser.

Les membres de ce groupe ne sont pas les seuls physiciens que j'ai croisé pendant cette thèse. Tout d'abord Mazyar, je n'oublie pas que mon premier contact avec ce domaine a eu lieu avec toi, dans ton cours à Polytechnique. Ensuite, Pierre, Pierre (Rouchon et Guilmin) et Antoine (Tilloy) merci pour ces discussions lorsque je m'arrachais les cheveux sur la théorie de mon expérience, et pour tes calculs Pierre (Rouchon). Thank you Pavlo Bilous, Florian Marquardt, Tom Dvir, Yonathan Cohen, Tirth Shah and Satya Bade for this fruitful collaboration in the ARTEMIS project and all the interesting discussions.

Merci évidemment au laboratoire de physique, Jean-Christophe Géminard, Nicolas Plihon, Eric Freyssingeas. Merci à Erika et Fatiha, vous êtes redoutablement efficaces dans une bonne humeur permanente. Fatiha, merci pour ta patience devant mon indécision crasse et ma phobie administrative.

Finally, I thank the members of the jury for having accepted to evaluate this work.

J'ai très bien vécu cette thèse car je suis extrêmement bien entouré. Par mes parents d'abord, vous avez toujours été là, m'avez toujours soutenu. Je ne mesure toujours pas la chance que j'ai de vous avoir comme parents. Au passage, merci Maman pour cette chouette illustration ! Merci à mes frères et soeurs, c'était sympa d'être tout seul mais je me serais bien emmerdé à la longue. J'ai hâte de voir ce que vous allez devenir. Merci à mes grands-parents, à mes tantes, oncles, parrain, marraine, cousins, cousines, qui m'accompagnez depuis toujours. Vous constituez mes fondations, sans lesquelles tout serait bien plus fragile.

Je suis extrêmement bien entouré par mes amis ensuite. Au commencement, il y eut les scouts Noé, Benoît et Florent. Vous m'avez poursuivi alors que je m'enterrais en prépa puis fuyais à Paris, je ne vous remercierai jamais assez pour cela. Le groupe s'est agrandi pour former mes amis "quasi de Lyon" indispensable à ma santé mentale, avec Solène, Coline, Clémence, Edwige, Margot, Marie, Ambre, Marie et Jean. En parallèle, il y a la bande de la prépa, avec Titus, Constance, Lucas et Alexis. Enfin, il y a mes amis d'école, groupe composé de la section escrime, avec Louise, Aymeric, Adrien, Lou, Maël, Diego, Nicolas, Clément, Guilhem, Quentin, William, Marine, et Cyril, ainsi que de Théophile, Titus et Constance encore. Malgré la distance, je suis vraiment très heureux de vous avoir dans ma vie. Mention spéciale pour le sous-groupe des quasi-tous futurs docteurs. Le moi de 17 ans aurait bien du mal à comprendre comment j'en suis arrivé à une telle ribambelle de prénoms importants dans ma vie.

Justine enfin, t'avoir à mes côtés me donne une force extraordinaire.

CONTENTS

I INTRODUCTION

1 INTRODUCTION 3
- 1.1 Monitoring the number of photons in a cavity 5
- 1.2 Neural network-assisted state preparation 10
- 1.3 Outline 12

II QUANTUM MEASUREMENTS WITH SUPERCONDUCTING CIRCUITS

2 QUANTUM MICROWAVE MODES 15
- 2.1 Circuits Quantum Electrodynamics 16
  - 2.1.1 Quantum LC oscillator 16
  - 2.1.2 The transmon qubit 19
- 2.2 Representation of quantum states 23
  - 2.2.1 Bloch ball 25
  - 2.2.2 Phase-space representations of a harmonic oscillator 28
- 2.3 Propagating modes 42
  - 2.3.1 Quantization of the transmission line 43
  - 2.3.2 Mode structure of the transmission line 48
  - 2.3.3 Time and frequency operators 52
- 2.4 Coupling stationary and propagating modes 54
  - 2.4.1 Quantum Langevin equations 54
  - 2.4.2 Examples 56

3 QUANTUM MEASUREMENT OF MICROWAVE MODES 61
- 3.1 Quantum Measurement of Superconducting circuits . . 61
  - 3.1.1 Homodyne detection 62
  - 3.1.2 Heterodyne detection 63
  - 3.1.3 Between homodyne and heterodyne 67
  - 3.1.4 Photon detector 70
  - 3.1.5 Voltage measurement 72
- 3.2 Quantum amplifiers 75
  - 3.2.1 Beamsplitter 76

3.2.2 Phase-preserving amplifier 78
3.2.3 Phase-sensitive amplifier 81
3.2.4 Quantum efficiency 82
3.3 Dispersive readout of a qubit 87
3.3.1 Pointer states 88
3.3.2 Dephasing rate 90
3.3.3 Measurement rate 92
3.3.4 Measurement rate: homodyne detection 94
3.3.5 Homodyne detection does not recover all the information 96
3.3.6 Using the information: measurement back-action 98
3.4 Conclusion 102

III PHOTOCOUNTING

4 MONITORING THE PHOTON NUMBER OF A CAVITY 105
4.1 Monitoring the energy of a cavity by observing the emission of a repeatedly excited qubit 107
4.2 Device and measurement setup 117
4.2.1 Device fabrication 117
4.2.2 Measurement setup 117
4.2.3 Comparison with previous experiment 119
4.3 Single-tone photon counting for pre- and post-selection 120
4.3.1 Single shot measurement with a single tone . . 120
4.3.2 Pre- and post selection using the qubit 122
4.4 Computing the theoretical measurement rates 124
4.5 Data processing 128
4.5.1 Quantum model of the recorded voltage 128
4.5.2 Data processing flow 129
4.5.3 Reconstruction of the qubit fluorescence signal 131
4.5.4 Relevance of the demodulation procedure . . . 133
4.5.5 Measurement rate estimation 133
4.5.6 Simulation 134
4.5.7 Bayesian filter for the photon number 135
4.5.8 Average time to reach a given confidence on the photon number measurement 136
4.6 Dephasing rate 137
4.6.1 Measuring the dephasing rate 137

4.6.2 Simulation of the dephasing rates 139
4.6.3 Predicted dephasing rate 140
4.7 Deviation to Quantum Nondemolition 141
4.8 Dependence of the cavity lifetime on photon number . 143
4.9 Supplementary sections 144
4.9.1 Wigner function of prepared Fock states 145
4.9.2 Neural-network-assisted record analysis 145
5 INFORMATION AND MEASUREMENT RATES 153
5.1 Modeling the experiment 154
5.1.1 Heterodyne measurement 154
5.1.2 Frequency-resolved photon detector 160
5.1.3 Accessible information and Holevo bound . . . 162
5.2 Mutual information and measurement rate 164
5.2.1 Frequency-resolved limit 164
5.2.2 Other regimes 168
5.3 Discriminating more than two frequencies 171
5.3.1 Information rate 173
5.3.2 Measurement time 175
5.4 Continuous case 176
5.4.1 Quantum Fisher Information 177
5.4.2 Frequency-resolved photon detector 180
5.4.3 Heterodyne measurement 182
5.4.4 Homodyne measurement 185
5.4.5 Qubit case: summary 191
5.5 Conclusions 192

IV NEURAL-NETWORK ASSISTED STATE PREPARATION

6 CAT STATE PREPARATION USING NEURAL NETWORKS 197
6.1 Introduction 197
6.2 Neural Network 201
6.3 Experimental results 204
6.4 Conclusion 210
6.5 Acknowledgments 211
6.6 B-spline basis set 212
6.7 Neural network training 213
6.8 Neural network performance 214
6.9 Accounting for decoherence 216

6.10 Device and measurement setup 218
6.10.1 Device fabrication 218
6.10.2 Measurement setup 220
6.11 Calibration and heralding 221
6.11.1 Qubit readout and reset 222
6.11.2 TLS mitigation 222
6.11.3 Heralding the cavity in the vacuum state 228
6.11.4 Pulse calibration 230
6.12 State preparation and fidelity measurement 230
6.12.1 Fidelity estimation 230
6.12.2 Experimental implementation 233
6.12.3 Wigner functions of various states 236
6.13 Optimization using GRAPE followed by Krotov 238
6.13.1 Numerical optimization 238
6.13.2 Theoretical infidelities 240
6.13.3 Experimental implementation of GRAPE . . . 241
6.14 Error analysis 241
6.14.1 Error budget 242
6.14.2 Effect of T_2 fluctuations 243

V APPENDIX

A QUANTUM INFORMATION TOOLS 247
A.1 Quantum communication 247
A.1.1 Classical channel 247
A.1.2 Quantum channel 248
A.1.3 Quantum channel with global measurements . . 250
A.1.4 Pure qubit case 251
A.2 Quantum metrology 253
A.2.1 Classical Fisher Information 253
A.2.2 Link with the mutual information 254
A.2.3 Quantum Fisher Information 256

B MULTIMODE PHASE-SPACE REPRESENTATION 257
B.1 Multimode Wigner function 257
B.2 Multimode Q function 257

C FABRICATION TECHNIQUES 259
C.1 Optical lithography 259
C.1.1 Cleaning 259

C.1.2 Coating 260
C.1.3 Exposure 260
C.1.4 Development 261
C.1.5 Etching 261
C.1.6 Dicing 262
C.2 Electronic lithography 262
C.2.1 Cleaning 263
C.2.2 Coating 265
C.2.3 Exposure 265
C.2.4 Development 266
C.2.5 Evaporation 266
C.2.6 Lift off 267
C.2.7 Junction resistance 267
C.3 Aluminum cavity 268

VI BIBLIOGRAPHY

BIBLIOGRAPHY 273

LIST OF FIGURES

Figure 1.1 Scheme of a mechanical analogy of the photon counting experiment 4
Figure 1.2 Twelve photons number trajectories taken with the same initial conditions using the fluorescence of a dispersively coupled qubit 6
Figure 1.3 Experimental and theoretical measurement and dephasing rates for different detectors monitoring the fluorescence of the qubit 7
Figure 1.4 Fisher information F for the homodyne and heterodyne detection, for a Frequency-Resolved Photon detector (FRPD) and the Quantum Fisher Information (QFI) as a function of $p(1)$ the excitation probability of the qubit after each kick 9
Figure 1.5 Optimal sampling of Wigner phase space and fidelity estimation of experimental cat states . 11
Figure 2.1 The LC oscillator 17
Figure 2.2 A non-linear oscillator 20
Figure 2.3 From the Cooper box to the transmon regime 21
Figure 2.4 The Bloch sphere 26
Figure 2.5 Experimental Wigner and Husimi Q functions of a few Fock states 34
Figure 2.6 Comparison between the Wigner function and the Husimi Q function of coherent and incoherent superposition of two coherent states . . 39
Figure 2.7 P distribution, Wigner distribution, Husimi Q function and generalized Q function of a coherent state 41
Figure 2.8 Alternative way of obtaining the generalized Q function from the Wigner function 42
Figure 2.9 Telegraph model for a portion of transmission line 44

Figure 2.10 Scheme of the transmission line as a discrete collection of harmonic oscillators 51
Figure 2.11 Capacitive coupling of a harmonic oscillator to a transmission line 56
Figure 2.12 Scheme of the release of a state by a linear cavity 58
Figure 3.1 The difference between homodyne and heterodyne detection lies in the mode probed 68
Figure 3.2 Scheme of a demodulation procedure and its outcomes 71
Figure 3.3 Scheme of the measurement of voltage and current measurement of a coaxial transmission line 73
Figure 3.4 Scheme of a microwave detection setup 76
Figure 3.5 Scheme of a beamsplitter 77
Figure 3.6 Phase-preserving amplification and measurement principle 80
Figure 3.7 Phase-sensitive amplification and measurement principle 83
Figure 3.8 Scheme of the equivalent measurement chain characterized by the quantum efficiency 84
Figure 3.9 Scheme of the dispersive readout principle . . 91
Figure 3.10 Dispersive readout of a qubit in phase space . 95
Figure 3.11 Comparison between ideal homodyne measurement and ideal measurement for the dispersive readout of a qubit 97
Figure 4.1 Experiment scheme and principle 108
Figure 4.2 Frequency comb and dynamics of the qubit under its drive 111
Figure 4.3 Photon number tracking experiment 113
Figure 4.4 Measurement and dephasing rates 114
Figure 4.5 Schematic of the setup (wiring) 120
Figure 4.6 Pogo pin and filter performances 121
Figure 4.7 Single-frequency drive photon counting 123
Figure 4.8 Pre-and post-selection on the photon number 124
Figure 4.9 Pre-and post-selected average qubit fluorescence 125
Figure 4.10 Gaussian approximation of the outcomes $\tilde{m}_n$. 127

Figure 4.11 Scheme of the quantum model of the detection of the fluorescence 129
Figure 4.12 Diagram of the signal processing 130
Figure 4.13 Conditional probability distribution of the orthonormalized records r_n 132
Figure 4.14 Average time needed to reach various confidence thresholds as a function of the average number of photons 136
Figure 4.15 Dephasing rate measurement and simulation . 138
Figure 4.16 Measurement-induced decay rate of the cavity 141
Figure 4.17 Measurements of the decay rate of the cavity . 142
Figure 4.18 Measurement of the average decay of the cavity with the tomograhy qubit 144
Figure 4.19 Measured Wigner functions of Fock states $|n\rangle$ for n between 0 and 9, after a heralding sequence on heterodyne measurement of the qubit emission 146
Figure 4.20 Predictions obtained using the neural network and the demodulation procedure 148
Figure 4.21 Average predictions obtained using the neural network and the demodulation procedure . . . 149
Figure 4.22 Fidelity and mutual information obtained using the neural network and the demodulation procedure 150
Figure 5.1 Scheme of the communication protocol between Alice and Bob 159
Figure 5.2 Communication scheme with a Frequency-Resolved Photon Detector 161
Figure 5.3 Scheme of the equivalent communication protocol between Alice and Bob for two possible number of photons in the cavity 166
Figure 5.4 Mutual information between the number of photons and a heterodyne measurement record as a function of the kick amplitude on the qubit and the quantum efficiency 167

Figure 5.5 Holevo bound and mutual information as a function of the qubit excitation for the heterodyne detection, the FRPD, and an optimal detector, for different photon number resolutions 169
Figure 5.6 Measured and theoretical measurement rates for heterodyne, FRPD and ideal measurement 170
Figure 5.7 Measured and theoretical information rate when discriminating the photon number 0 from the photon number k with heterodyne detection, for $\theta = \pi/2$ and $\eta = 0.17$ 172
Figure 5.8 Information rates with 8 possible photon numbers as a function of the drive amplitude . . . 173
Figure 5.9 Experimentally determined generalized measurement rate Γ'_m as a function of the number of possible photon numbers m 175
Figure 5.10 Experimentally determined measurement rate matrix $\mathbf{MR}$ 176
Figure 5.11 Influence of the direction of squeezing on the Quantum Fisher Information 180
Figure 5.12 Fisher information as a function of the quantum efficiency and the kick amplitude on the qubit 185
Figure 5.13 Scheme of the transformation of the two-mode Wigner function for a small detuning between the homodyne detection frequency and the signal frequency 190
Figure 5.14 FI for the homodyne and heterodyne detection, for an FRPD, and the QFI as a function of $p(1)$. 192
Figure 6.1 Principle of the experiment 198
Figure 6.2 Data processing, training pipeline and simulation example 202
Figure 6.3 Experimental Wigner functions for uniform and optimal sampling 205
Figure 6.4 Experimental and theoretical infidelity as a function of α and φ 207
Figure 6.5 B-spline basis of size $n = 11$ and degree $k = 3$ 213
Figure 6.6 Neural network training progress with batches 215

Figure 6.7 Simulated infidelity of cat states obtained by driving the system in the way suggested by the trained neural network 216
Figure 6.8 Simulated infidelity of cat states with $\alpha \in (0, 4)$ and $\varphi = 0$ obtained using the original scenario (purple curve), a larger B-spline basis set (orange curve) and in absence of decoherence (brown curve). 217
Figure 6.9 Schematic of the measurement setup 219
Figure 6.10 Scheme of the device 220
Figure 6.11 Readout and reset of the qubit 223
Figure 6.12 Ramsey measurement of the qubit 224
Figure 6.13 Heralding the TLS 225
Figure 6.14 TLS state measurement 227
Figure 6.15 Distinguishability of the TLS states 228
Figure 6.16 Cavity vacuum heralding procedure 229
Figure 6.17 Cavity displacement calibration 231
Figure 6.18 Comparison between optimal and square sampling on fidelity estimation 234
Figure 6.19 State Preparation and Measurement (SPM) sequence 235
Figure 6.20 Sensitivity of the fidelity estimation to the T_2 time of the qubit 236
Figure 6.21 Wigner functions of various cat states 237
Figure 6.22 GRAPE & Krotov-optimized control of the cavity and qubit 238
Figure 6.23 Comparison between the GRAPE & Krotov method and the NN method 239
Figure 6.24 Influence of the measurement imperfections on the experimental fidelity estimation 242
Figure A.1 Scheme of a classical communication channel . 247
Figure A.2 Scheme of a quantum communication channel 249
Figure A.3 Scheme of a global quantum communication strategy 250
Figure A.4 Optimal measurement to discriminate two quantum states 252
Figure C.1 Optical lithography procedure 260

Figure C.2 Josephson junction fabrication procedure . . . 263
Figure C.3 Pictures of the influence of the Purestrip on the cleaning procedure 264

LIST OF TABLES

Table 4.1 Table of circuit parameters of the photon counting experiment 118
Table 4.2 Summary of the differences with [126] 122
Table 4.3 Yield table of the pre- and post-selection after a displacement of $\alpha \approx 3.4$ when the kick angle θ is set to 0. 126
Table 6.1 Table of circuit parameters of the state preparation experiment 220

Γ_m	Measurement rate. Appears on pages 7, 93, 94, 96, 101, 114–116, 119, 122, 133, 134, 136, 137, 170–173.
Γ_d	Dephasing rate. Appears on pages 7, 91–94, 96, 100, 101, 170.
$P_{\hat{\rho},\hat{A}}$	Glauber-Sudarshan P representation of the mode represented by the annihilation operator $\hat{A}$ in the state $\hat{\rho}$, defined in Eq. (2.41). Appears simplified as $P_{\hat{\rho}}$ if understandable in the context. Appears on pages 26, 30–32, 36–38, 40, 41, 65, 77.
$W_{\hat{\rho},\hat{A}}$	Wigner function of the mode represented by the annihilation operator $\hat{A}$ in the state $\hat{\rho}$, defined in Sec. 2.2.2.2. Appears simplified as $W_{\hat{\rho}}$ or W_f if understable in the context. Appears on pages 31–34, 36, 37, 39–42, 65, 77, 78, 81, 83–85, 187–190.
$W^{\zeta}_{\hat{\rho}}$	Wigner function of a mode in the squeezed state $\hat{\mathcal{S}}(\zeta)\hat{\rho}\hat{\mathcal{S}}(-\zeta)$, defined in Eq. (2.59). Appears on pages 34, 35, 40.
$Q_{\hat{\rho},\hat{A}}$	Husimi Q function of the mode represented by the annihilation operator $\hat{A}$ in the state $\hat{\rho}$, defined in Sec. 2.2.2.3. Appears simplified as $Q_{\hat{\rho}}$ or Q_f if understandable in the context. Appears on pages 35–38, 41, 42, 65, 69, 74, 78–81, 83, 156–158, 165, 182, 183.
$Q^{\hat{\sigma}}_{\hat{\rho}}$	Generalized Husimi Q function, parameterized by the density matrix $\hat{\sigma}$, of a mode in the state $\hat{\rho}$, defined in Sec. 2.2.2.4. Appears on pages 39, 40.

$Q^{\zeta}_{\hat{\rho},\hat{A}}$	Generalized Husimi Q function, parameterized by the density matrix $\hat{\sigma} = \hat{\mathcal{S}}(\zeta)\,\lvert 0\rangle\langle 0\rvert\,\hat{\mathcal{S}}(-\zeta)$ of a squeezed vacuum, of the mode represented by the annihilation operator $\hat{A}$ in the state $\hat{\rho}$, defined in Sec. 2.2.2.4. Appears simplified as $Q^{\zeta}_{\hat{\rho}}$ ot Q^{ζ}_{f} if understable in the context. Appears on pages 40–42, 83.
$\hat{A}_f$	Annihilation operator of the propagating mode defined by the normalized function $f \in L^2(\mathbb{R},\mathbb{C})$. Defined in Eq. (2.108). Appears on pages 48, 49, 57, 59, 62, 64–69, 74, 76–79, 81, 83–85, 89, 92, 155, 156, 158, 160, 163, 182.
$\hat{\tau}$	Time operator, defined on $L^2(\mathbb{R},\mathbb{C})$ in Eq. (2.121). Appears on pages 52, 178, 179, 182–184, 186–189.
$\hat{\omega}$	Frequency operator, defined on $L^2(\mathbb{R},\mathbb{C})$ in Eq. (2.122). Appears on page 52.
$\mathcal{T}_{\Delta\omega}$	Frequency translation operator, defined on $L^2(\mathbb{R},\mathbb{C})$ in Eq. (2.123). Appears on pages 52, 177, 186, 187.
$\mathcal{T}_{\Delta\tau}$	Time translation operator, defined on $L^2(\mathbb{R},\mathbb{C})$ in Eq. (2.124). Appears on page 52.
$\hat{\boldsymbol{\tau}}$	Quantum time operator, defined on the space of propagating states in Eq. (2.129). Appears on pages 52, 53, 178, 186.
$\hat{\boldsymbol{\omega}}$	Quantum frequency operator, defined on the space of propagating states in Eq. (2.130). Appears on pages 52, 53.
$\hat{\mathcal{T}}_{\Delta\omega}$	Quantum frequency translation operator, defined on the space of propagating states in Eq. (2.131). Appears on pages 53, 177, 186, 187.

$\hat{\mathcal{T}}_{\Delta\tau}$	Quantum time translation operator, defined on the space of propagating states in Eq. (2.132). Appears on page 53.
$\hat{x}(t)$	Time quadrature operator at time t. Appears on pages 62–64, 66, 67, 71.
$\hat{x}_f$	Operator obtained by demodulating the time quadrature operator $\hat{x}(t)$ by the function $f \in L^2(\mathbb{R}, \mathbb{C})$. Appears on pages 62–65, 67–69, 73, 74, 158.
x	Measurement record. $x(t)$ is the measurement outcome obtained at time t by measuring the time quadrature operator $\hat{x}(t)$. Appears on pages 62, 65, 71, 154–159, 164–167, 182, 189.
x_f	Measurement outcome obtained by demodulating the measurement record x by the function $f \in L^2(\mathbb{R}, \mathbb{C})$. Appears on pages 64, 65, 70, 71, 75, 78, 79, 81–83, 156–159, 165, 166, 182, 183, 186, 189, 190.
$I(a : b)$	Mutual information shared by the random variables a and b, defined in Eq. (A.2). Appears on pages 93, 95, 96, 154, 155, 157–159, 164, 165, 167, 171, 174, 248, 249.
I_{acc}	Accessible information, defined in Eq. (A.5). Appears on pages 93, 94, 163, 249–251, 253.
Γ_i	Information rate. Appears on pages 173–175.
F_Q	Quantum Fisher information, defined in Eq. (A.28). Appears on pages 178–180, 256.
F	Fisher information, defined in Eq. (A.18). Appears on pages 181–183, 185, 189, 191, 254, 255.
$\mathbf{W}$	Multimode Wigner function, defined in Eq. (B.2). Appears on pages 187–190, 257, 258.

$\mathbf{Q}$	Multimode Husimi Q function, defined in Sec. B.2. Appears on page 258.

Part I

INTRODUCTION

1

INTRODUCTION

With the formulation of Heisenberg's uncertainty relations in 1927 [1] came the revolutionary idea of a fundamental (quantum) limit on the information an observer could obtain from a measurement. It was first formulated as a limit in the resolution of the position and the momentum of a particle, and generalized to every pair of observables that do not commute. Since then, quantum limits have been formulated in many different contexts. In this thesis, we will use tools developed in the context of quantum measurements, quantum communication, and quantum metrology to answer questions naturally raised when implementing a measurement: what information do we retrieve? How much is it possible to obtain? How could we get more?

We can give a more concrete idea of how such questions arise with an example. Suppose that we have a basket of apples and that we want to count them. Let us consider special apples that cannot be simply looked at. Instead, we can weigh them. A solution consists in hanging the basket to a string, which is going to resonate at a frequency that depends on the number of apples. To count them, we hit the string with a hammer and listen to the sound it makes. Depending on the listeners' familiarity with music and the weight of the apples, one hit might be enough to determine this number. Formally, we encoded the information of apples into the frequency of the sound emitted by the string. This solution is not as simple as putting the basket on a scale but has the merit of making explicit the weighing process, which is necessary to fully model a measurement in quantum physics.

We can now think about what would happen if we hit the string less and less strongly or if we keep lowering the weight of the apples. The listener will have more and more trouble identifying the frequency with enough precision. We can switch to more efficient measurement tools, gather more signal by guiding the sound waves, cool down the signal to mitigate the noise... However, at some point, if we hit it

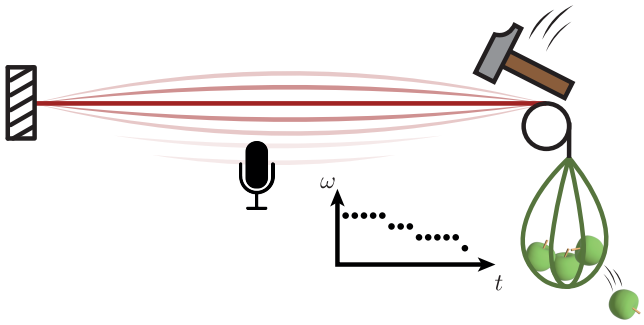


Figure 1.1: Scheme of a mechanical analogy of the photon counting experiment. The frequency of the sound emitted by a string hit with a hammer allows us to count the number of apples in a basket hanged to it. This analogy illustrates the experiment of Chap. 4.

weakly enough, the information carried by the sound should simply not be enough to know the number of apples in one go.

However, classical physics gives no fundamental bound to the amount of noise added when performing a measurement. It would thus be possible to have none, which poses a fundamental issue: at zero temperature, without any noise in the measurement, we could perfectly know the number of apples, whatever the strength of the kicks. The information content of the pulse would then go from zero, if there is no kick, to 100 % of the wanted information for any kick strength. As a consequence, the concept of the intrinsic information content of the emitted sound has no meaning in a classical theory without noise. This hiatus is filled by a quantum description of the experiment. When implementing an experimental measurement apparatus, the following questions arise naturally:

1. What information do we get from the measurement (the *mutual information* between the number of apples and the measurement record)?

2. What information is available in the sound (the intrinsic *accessible information* of emitted sound)?

3. Would we recover all the information if the measurement could fully exploit all the emitted signals?

4. How would other measurement apparatuses compare?

1.1 MONITORING THE NUMBER OF PHOTONS IN A CAVITY

This analogy describes well the experiment conducted in Chap. 4. Instead of apples, we count photons in a superconducting resonator. The string is replaced by a dispersively coupled qubit and emits an electromagnetic wave into a waveguide instead of sound after each kick in the form of a sharp electromagnetic pulse. This experiment is conducted in a dilution refrigerator, which sets the thermal noise on the signals to approximately zero. Furthermore, since this wave is emitted by the relaxation of a qubit, it contains at most one photon, enabling us to fully explore the limit of weak hammer hits and providing non-classical subtleties in the measured signal. The measurement performed on this wave is a well-known heterodyne detection. A series of 12 measured trajectories of photon number is shown in Fig. 1.2. After populating the cavity with a short coherent pulse, the qubit is kicked every 95 ns for 1 ms. The estimation of the photon number takes the form of a probability distribution $\mathbb{P}_t(n)$ for each number of photons n between 0 and 9. This distribution is initialized to $\mathbb{P}_0(n) = 1/10$ at $t = 0$ and updated using the Bayes rule every 2 μs, which corresponds to 21 hammer kicks. For most of these trajectories, after each photon loss, we see that the estimation manages to converge again to the number of photons immediately below, showing successful monitoring of the number of apples-photons.

Using standard information theory tools, we extract the measurement rate, characterizing the speed at which the information about the number of photons is obtained. Doing this for each kick amplitude, which is expressed as the angle θ traveled by the qubit on the Bloch sphere at each kick, we plot the measurement rate obtained with our experimental technique (red error bars in Fig. 1.3a) as a function of θ. The Fig. 1.2 was obtained for $\theta = \pi/2$. We compare it to a theoretical model (blue) and a simulation (orange) for a similar quantum efficiency, which characterizes the portion of the emitted signal that is exploited by the measurement. We obtain a good agreement, at least for low kick amplitude. This provides the answer to the first question.

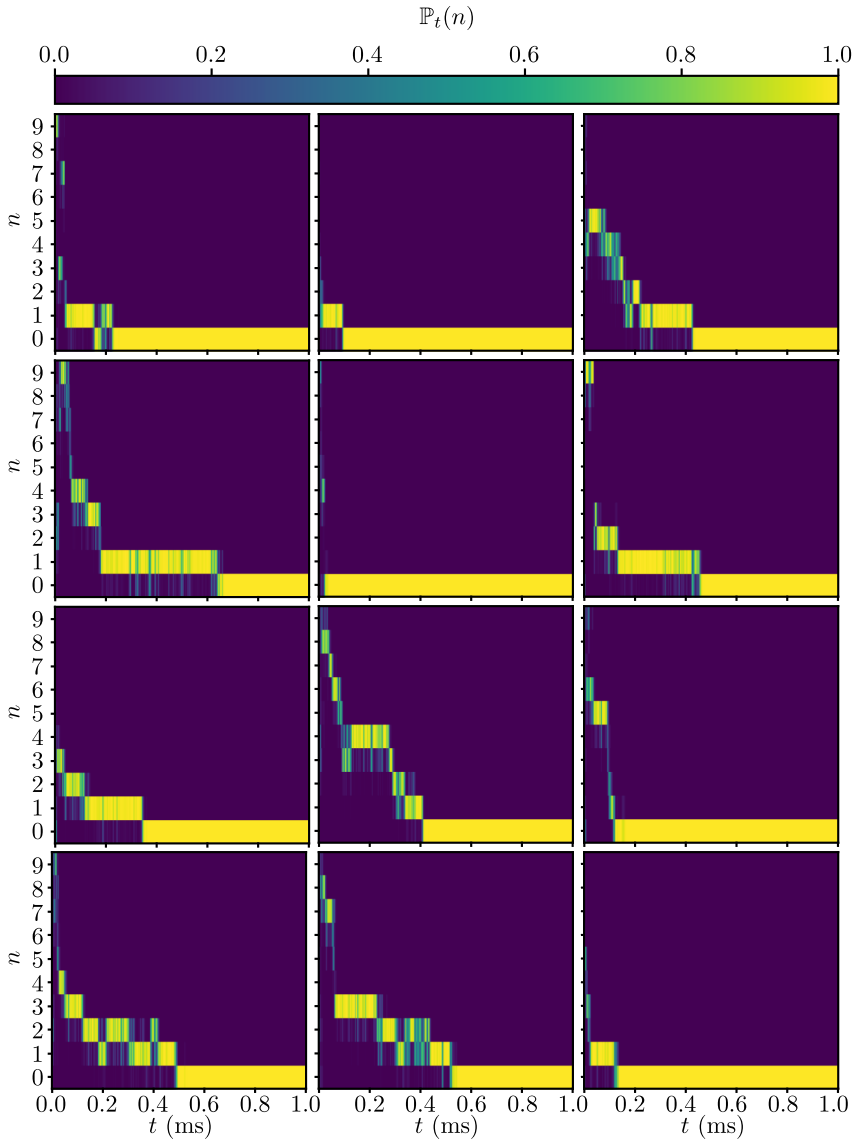


Figure 1.2: Twelve photons number trajectories taken with the same initial conditions using the fluorescence of a dispersively coupled qubit (see Chap. 4).

The answer to the second question exploits one of the key features of quantum mechanics. The information about the system that leaked in the environment is directly observable in the system itself: the

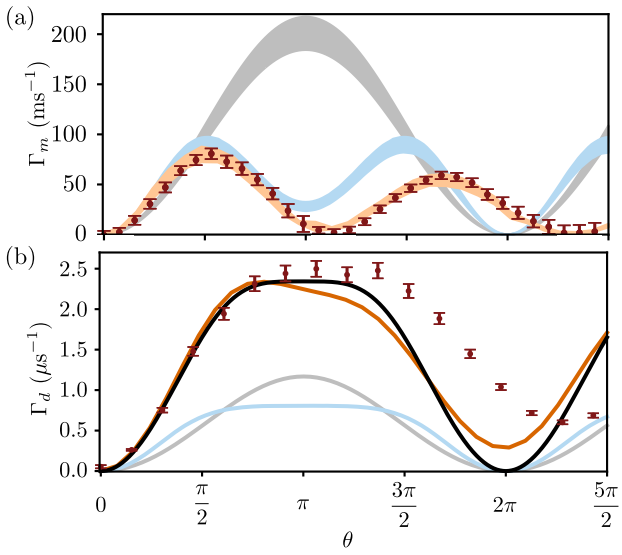


Figure 1.3: Experimental and theoretical measurement and dephasing rates for different detectors monitoring the fluorescence of the qubit. (a) Red dots: observed average measurement rate Γ_m as a function of drive amplitude, expressed as the qubit expected rotation angle θ per kick. Orange shadow: measurement rate obtained using a stochastic master equation with detection efficiency η spanning the range $[0.17, 0.20]$. Blue shadow: theoretical bound for instantaneous kicks and heterodyne measurement with efficiencies $\eta \in [0.17, 0.20]$. Gray shadow: theoretical measurement rate for instantaneous kicks obtained with a Frequency-Resolved Photon Detector (FRPD) with efficiencies $\eta \in [0.17, 0.20]$. (b) Red dots: observed cavity dephasing rate Γ_d as a function of θ. Orange: simulated measurement induced dephasing rate. Black: theoretical accessible information rate with instantaneous kicks. Blue: theoretical maximal ($\eta = 1$) measurement rate obtained by heterodyne detection and instantaneous kicks. Grey: theoretical maximal ($\eta = 1$) measurement rate obtained using an FRPD and instantaneous kicks. These results are presented in Secs. 4.1 and 5.2.2.

dephasing induced by the entanglement of this system with its environment. When interested in peculiar information, such as the number of photons, we can perform a measurement on the environment. In the

case where there are two possible photon numbers, the measurement rate, defined as the speed at which information is gathered, is bounded by the dephasing rate, defined as the rate at which the off-diagonal components of the density matrix converge to zero due to its continuous entanglement with the environment. The dephasing rate of the cavity as a function of θ is shown in Fig. 1.3b, both experimentally (red error bars) and theoretically (black for a theoretical model, orange for the simulation). The curve in blue shows the answer to the third question: if our measurement could fully exploit the totality of the emission of the qubit (which means unit quantum efficiency), it would still not reach the maximal measurement rate (the dephasing rate). For $\theta = \pi$, it only extracts around a third of the available information in the qubit emission. This is analogous to a standard dispersive qubit readout performed with a homodyne detection with unit quantum efficiency, but whose phase is non-optimal, which is the case discussed in Sec. 3.3.6. This answers our third question. Finally, we (theoretically) considered the most direct type of frequency measurement we could imagine, using a frequency-resolved photon detector (FRPD), whose principle is quite simple: perform a direct frequency measurement on every detected photon in the qubit emission. The performance of such a detector is shown in grey in Fig. 1.3a for the same quantum efficiency as the experiment, and in grey in Fig. 1.3b for unit quantum efficiency. We see that it can outperform the heterodyne detection with the experiment settings, but that it does not reach the maximum measurement rate either. That is part of the answer to the fourth question.

Once these questions are answered for the experiment, another one naturally comes: How do these rates depend on the experimental parameters? The Chap. 5 deals with the number-resolved regime as well as the opposite regime, where the qubit linewidths for each photon number have an overlap close to unity.

Interestingly, the latter regime can be tackled by shifting from a quantum communication point of view (how many bits of information about the photon number are sent by the qubit, how many bits are received by the experimentalist?) to a quantum parameter estimation problem (how precise can we be in the estimation of the photon number?), by considering the number of photons in the cavity (and

so, the frequency of the qubit) as a continuous parameter. The quantity of interest becomes the Fisher information [2]. The main result (Fig. 1.4) is the theoretical comparison of four different measurement schemes as a function of the excitation probability of the qubit $p(1)$: optimal homodyne detection, heterodyne detection, FRPD, and the optimal measurement reaching the quantum Fisher Information (QFI). Crucially, optimal homodyne detection works as well as the quantum bound when the emitted qubit state can be assimilated to a coherent state ($p(1) \ll 1$), as it is the case for a dispersive qubit readout, but gets less and less optimal until it collapses to 0 when a Fock state $|1\rangle$ ($p(1) = 1$) is emitted. However, it is still more efficient than the two other schemes up to $p(1) \simeq 0.8$. Interestingly, in this regime, the heterodyne detection always performs better than the FRPD, thanks to its ability to exploit the state coherence.

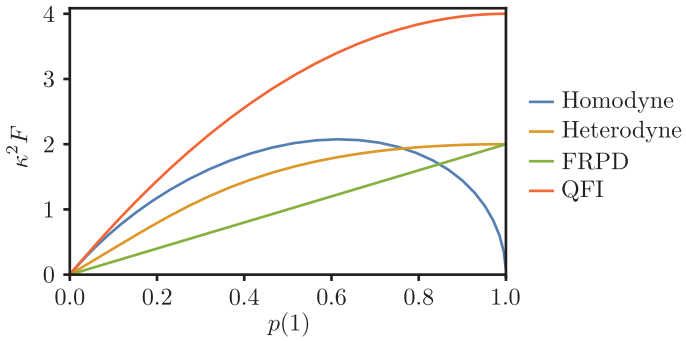


Figure 1.4: Fisher information F for the homodyne and heterodyne detection, for a Frequency-Resolved Photon detector (FRPD) and the Quantum Fisher Information (QFI) as a function of $p(1)$ the excitation probability of the qubit after each kick (see Sec. 5.4.5).

This experiment can then be viewed in this thesis as a means to explore various aspects of the field of quantum information, and it sets an interesting starting point to explore the links between quantum measurements, quantum communication and quantum metrology.

1.2 NEURAL NETWORK-ASSISTED STATE PREPARATION

The second experiment presented here is a demonstration of neural network-assisted preparation of a family of states, called the Schrödinger cat states, in a cavity. These states are parameterized by a real number α and a phase φ. Once trained, a neural network associates output pulses to each pair (α, φ). These pulses, when sent to the cavity and its dispersively coupled qubit, prepare the cavity in the target state and the qubit in the ground state. Compared to standard optimal control methods such as Gradient Ascent Pule Engineering (GRAPE) [3] or Krotov [4] methods, this neural network-assisted method has the key advantage of providing near-optimal control pulses for a wide range of parameters after being trained only once. This can lead to potential applications needing fast changes of parameters, such as feedback, adaptive techniques, or error correction.

The experimental work first consisted in calibrating the experiment to prepare these states with a fidelity $\mathcal{F}(\alpha, \varphi)$ to the target state as high as possible. Second, we had to estimate this fidelity as fast and precisely as possible. Indeed, we had to estimate it for many parameters, and the long lifetime of the cavity imposes a long reset time between two experiments. This leads to a quantum metrology-like question: how to estimate $\mathcal{F}(\alpha, \varphi)$ with the best accuracy and with as few samples as possible? We experimentally implement the optimal fidelity estimation strategy from the cavity Wigner function, which was established in [5]. It consists in performing parity measurement after displacement by a randomly picked amplitude β along the law $p(\beta)$, which is proportional to the absolute value of the Wigner function of the target state W_α^φ. The best estimator of $\mathcal{F}(\alpha, \varphi)$ is then, up to a normalizing factor, simply given by a weighted average of all the parity measurement outcomes. The Fig. 1.5a shows $p(\beta)$ for an even cat state of size 2, which corresponds to $\alpha = 2$ and $\varphi = 0$. One can reconstruct a Wigner function by pixel-averaging the parity measurements. This is shown in Fig. 1.5b. It contrasts with the square sampling that we usually see when performing Wigner tomography. The measured fidelity for 20 values of α in $[0.1, 2]$ and 9 values of φ in $[0, \pi]$ using this technique, and their theoretical counterpart are shown in Fig. 1.5c.

Each fidelity estimation was obtained with 10^5 samples, and the whole map took 8 days.

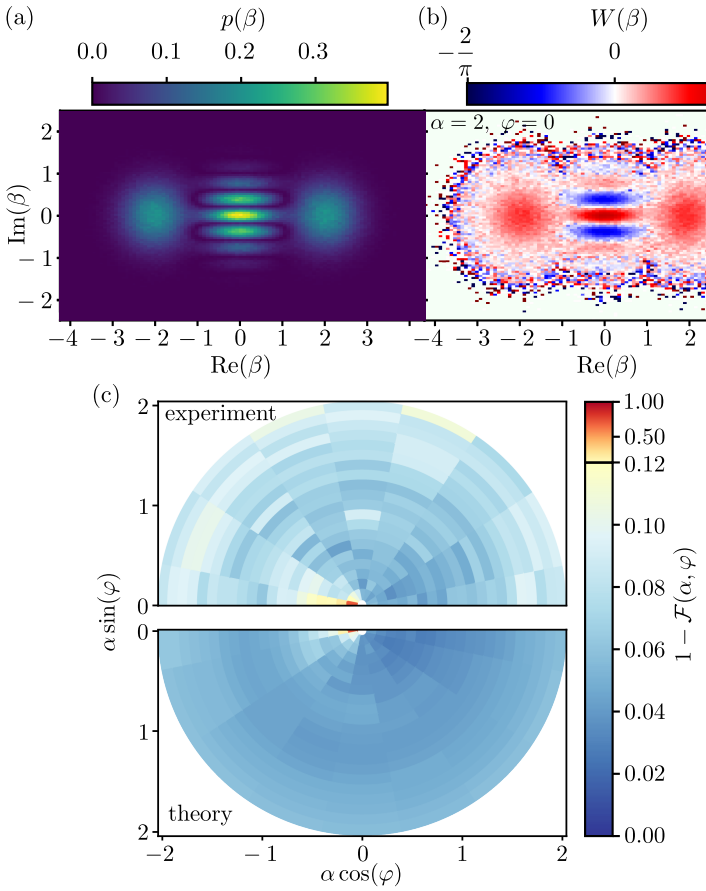


Figure 1.5: Optimal sampling of phase space and fidelity estimation of experimental cat states. (a) Optimal sampling density for a target even cat state of size 2. (b) Measured Wigner functions of the cavity with the sampling shown in (a) for $\alpha = 2$, and $\varphi = 0$, obtained by averaging the parity in each pixel (c) Bottom: simulated infidelity for a sampling of the whole half-disk (reflected) in the parameter space $(\alpha\cos(\varphi), \alpha\sin(\varphi))$. Top: experimental infidelities obtained with 10^5 shots sampled optimally.

1.3 OUTLINE

Chap. 2 is dedicated to introducing the theoretical tools to describe two-level systems and electromagnetic modes. The main tools to visualize the state of harmonic modes are presented, with special stress on Husimi and generalized Husimi Q functions. We also present a quantum optics formalism that allows us to describe the propagating signals as quantum states decomposed on a countable number of propagating modes. Their interaction with usual quantum systems is derived as well. The Chap. 3 is dedicated to linking this description of the propagating modes to the outcomes of the standard measurement setups in the field of superconducting circuits, implementing homodyne and heterodyne detection of these propagating modes. This framework is then used to analyze the standard qubit dispersive readout. The notions of quantum efficiency, dephasing, and measurement rates are introduced. Chap. 4 is, up to a few new appendices, a reproduction of the paper [6] describing the experimental realization of the monitoring of a cavity photon number. Chap. 5 uses the framework introduced in Chap. 2 and Chap. 3 to further explain a few theoretical points of this paper. It then uses the experiment as a starting point to study possible generalizations of the experiment and new related theoretical questions. Finally, Chap. 6 is a reproduction of our preprint experimentally demonstrating the use of a neural network to prepare a family of states in a high-Q cavity using a dispersively coupled qubit [7].

Part II

QUANTUM MEASUREMENTS WITH SUPERCONDUCTING CIRCUITS

2

QUANTUM MICROWAVE MODES

At its core, cavity quantum electrodynamics (CQED) aims to study the interaction between confined light and matter. Linear cavities exhibit resonant modes, which can be thought of as "photon boxes" into which light can be stored and interact with other systems. The first platform that demonstrated coherent interactions between a few photons and single degrees of freedom of matter was the cavity quantum electrodynamics with atoms [8], to which the photons in the cavity coupled via the dipolar interaction. Superconducting qubits were shown to act as artificial atoms [9], which founded the field of circuit quantum electrodynamics. The flexibility of this approach as well as the strong couplings enabled by this platform gave birth to a lot of diverse experiments, both witnessing the fundamental aspects of quantum physics and paving the way for more concrete applications: quantum sensing [10], quantum communication [11], quantum computing [12], quantum thermodynamics [13]...

The goal of this chapter is first to introduce two ubiquitous physical systems in circuit quantum electrodynamics: the resonator and the transmon. A more general description of the procedure used here to quantize these two simple circuits is made in [14]. We will also introduce the theoretical tools used to model these systems, visualize their states, and build intuition about their properties and dynamics. We will finally present the model of the transmission line, which is central to modeling the quantum signals driving or emitted by these systems, as well as the quantum Langevin equation, which describes the interaction between a quantum system and a transmission line.

2.1 CIRCUITS QUANTUM ELECTRODYNAMICS

2.1.1 *Quantum LC oscillator*

"Photon boxes" or harmonic oscillators take two main forms in superconducting circuits. The first is that of a linear resonant mode in a distributed superconducting circuit, *i.e.* a circuit whose characteristic size is that of the wavelength of the mode, like $\lambda/2$ or $\lambda/4$ resonators. Such modes are associated to a frequency $\omega_c/2\pi$ and a characteristic impedance Z_c setting the ratio between the amplitudes of the voltage oscillations and the current oscillations in the walls of the box (or the ratio between the electric and magnetic field oscillation amplitudes in the interior of the box). It can thus be modeled by an inductor and a capacitor in parallel, forming an LC oscillator, with frequency $1/\sqrt{LC} = \omega_c$ and characteristic impedance $Z_c = \sqrt{L/C}$ (see Fig. 2.1a). This gives the second way of building a harmonic oscillator, which is by building this very circuit: a linear inductance L and a linear capacitance C in parallel. This is called a lumped resonator.

From the voltage $V(t)$ and the current $I(t)$ in the LC oscillator, we can then define the flux Φ and charge Q:

$$\begin{cases} \Phi(t) = \displaystyle\int_{-\infty}^{t} V(t')\mathrm{d}t' \\ Q(t) = \displaystyle\int_{-\infty}^{t} I(t')\mathrm{d}t'. \end{cases} \tag{2.1}$$

The state of the oscillator is completely determined by the values of Φ and Q, corresponding to a point in the plane Φ and Q called the (classical) phase space (see Fig. 2.1b). The Lagrangian of this system reads

$$\mathcal{L} = \frac{C}{2}\dot{\Phi}^2 - \frac{1}{2L}\Phi^2 \tag{2.2}$$

where Φ and Q are conjugate variables:

$$\frac{\partial \mathcal{L}}{\partial \dot{\Phi}} = C\dot{\Phi} = Q. \tag{2.3}$$

From this Lagrangian, we can obtain the Hamiltonian of the system using the Legendre transform:

$$H = \frac{\partial \mathcal{L}}{\partial \dot{\Phi}} - \mathcal{L} = \frac{1}{2C}Q^2 + \frac{1}{2L}\Phi^2. \tag{2.4}$$

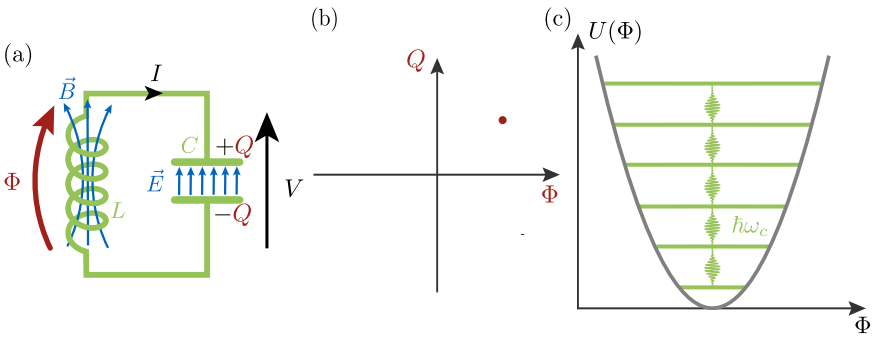


Figure 2.1: The LC oscillator. (a) Schematic of the LC oscillator (b) Representation of the state of the oscillator as a dot (here in red) in the classical phase space. (c) Energy spectrum of the quantum harmonic oscillator represented in the quadratic potential $U(\Phi)$. The levels are evenly spaced.

This is the Hamiltonian of a classical harmonic oscillator, with Φ and Q characterizing a single degree of freedom. They play the same role as the position and the momentum in a spring-mass system, the mechanical analog of the LC oscillator. To complete the analogy, one is free to associate Φ and Q respectively to position and momentum or the other way around. However, as we will see later, the main (and almost only) non-linear element we can realize in the domain is a Josephson junction, an inductive element. Thus, it is easier to think of the variable Φ as the position variable. Replacing the linear inductor with a Josephson junction is then analogous to changing the potential $U(\Phi)$ in the Hamiltonian for the Josephson potential.

We can now promote these conjugate variables into conjugate quantum operators satisfying the canonical commutation rule:

$$\begin{cases} \Phi \to \hat{\Phi} \\ Q \to \hat{Q} \\ [\hat{\Phi}, \hat{Q}] = i\hbar. \end{cases} \tag{2.5}$$

The Hamiltonian now reads the same in terms of these newly defined operators:

$$\hat{H} = \frac{1}{2C}\hat{Q}^2 + \frac{1}{2L}\hat{\Phi}^2. \tag{2.6}$$

To diagonalize this Hamiltonian, we introduce the annihilation and creation operators, which are hermitian conjugate of each other:

$$\begin{cases} \hat{a} = \dfrac{1}{\sqrt{2\hbar Z_c}}(\hat{\Phi} + iZ_c\hat{Q}) \\ \hat{a}^\dagger = \dfrac{1}{\sqrt{2\hbar Z_c}}(\hat{\Phi} - iZ_c\hat{Q}). \end{cases} \tag{2.7}$$

These operators verify the canonical commutation relations for bosonic operators $[\hat{a}, \hat{a}^\dagger] = \mathbb{1}$. They are dimensionless operators which carry the essence of the physics of quantum harmonic oscillators. The operators $\hat{\Phi}$ and $\hat{Q}$ can then be expressed as a function of the bosonic annihilation and creation operators:

$$\begin{cases} \hat{\Phi} = \Phi_{\text{zpf}}(\hat{a} + \hat{a}^\dagger) \\ \hat{Q} = iQ_{\text{zpf}}(\hat{a}^\dagger - \hat{a}). \end{cases} \tag{2.8}$$

Here, $\Phi_{\text{zpf}} = \sqrt{\frac{\hbar Z_c}{2}}$ and $Q_{\text{zpf}} = \sqrt{\frac{\hbar}{2Z_c}}$ are called the flux and charge *zero-point fluctuations*. They link the physically measurable operators $\hat{\Phi}$ and $\hat{Q}$ to the bosonic field quadratures $\hat{a} + \hat{a}^\dagger$ and $i(\hat{a}^\dagger - \hat{a})$. Their ratio $\frac{\Phi_{\text{zpf}}}{Q_{\text{zpf}}} = Z_c$ is the characteristic impedance of the resonator, and their product $\Phi_{\text{zpf}}Q_{\text{zpf}} = \frac{\hbar}{2}$ sets the minimal uncertainty $\Delta\Phi\Delta Q$ in Heisenberg's inequality.

Using the creation and annihilation operators, the Hamiltonian now reads

$$\hat{H} = \hbar\omega_c(\hat{a}^\dagger\hat{a} + \frac{\mathbb{1}}{2}). \tag{2.9}$$

The ground state $|0\rangle$, called the vacuum state, can be found by solving the equation $\hat{a}\,|\psi\rangle = 0$ in the flux representation, and the eigenstates $\{|n\rangle\}_{n\in\mathbb{N}}$ of the Hamiltonian, called the Fock states, can be obtained via successive applications of the creation operator $\hat{a}^\dagger$ to the vacuum state: $|n\rangle = \frac{1}{\sqrt{n!}}\hat{a}^{\dagger n}\,|0\rangle$. These states also verify $\hat{a}\,|n+1\rangle = \sqrt{n+1}\,|n\rangle$ and $\hat{a}^\dagger\hat{a}\,|n\rangle = n\,|n\rangle$, from which we can deduce the spectrum of the Hamiltonian $\{\hbar\omega(n + \frac{1}{2})\}_{n\in\mathbb{N}}$. This spectrum is discrete, with levels evenly spaced by $\hbar\omega$ (see Fig. 2.1c). This quantum of electromagnetic energy is called a *photon*. The harmonic oscillator can then be thought of as a box of photons. A Fock state $|n\rangle$ is then a state with a defined number of photons, and any pure state $|\psi\rangle$ can be written as a coherent

superposition of these number states: $|\psi\rangle = \sum_n c_n |n\rangle$, with the c_n being complex numbers such that $\sum_n |c_n|^2 = 1$. In particular, coherent states, which will be presented in Sec. 2.2.2, are an example of states obtained by a coherent superposition of Fock states. Unlike the Fock states, which have no classical counterpart, these states are obtained naturally when performing classical operations on a harmonic oscillator. As for a classical harmonic oscillator, the state of a quantum harmonic oscillator can be represented in phase space, but in several different ways, which will be introduced in Sec. 2.2.2.

2.1.2 *The transmon qubit*

Harmonic oscillators driven with classical sources are not enough to explore their full Hilbert space. The only accessible states are coherent states, and their statistical mixturesFor example, it is impossible to create a Fock state with only a cavity driven with classical signals. Intuitively, this is due to the fact that the levels of the harmonic oscillator are evenly spaced: any incoming light drives all the transitions simultaneously. A solution to obtain other quantum states then consists in modifying the harmonic potential to change this level distribution. In the domain of the superconducting circuits, this is done using a Josephson junction. This element consists of two superconducting electrodes spearated by an insulating barrier, and its energy potential depends on the superconducting phase difference operator $\hat{\varphi} \in [0, 2\pi[$ between both sides of the junction:

$$U(\hat{\varphi}) = -E_J \cos(\hat{\varphi}) \tag{2.10}$$

where E_J is the Josephson energy of the junction. The phase $\hat{\varphi}$ is also linked to a flux across the junction $\hat{\Phi} = \varphi_0 \hat{\varphi}$, with $\varphi_0 = \frac{\hbar}{2e}$ the reduced flux quantum. This flux plays the same role as the magnetic flux for an inductive element, but it is not associated to a magnetic field: the energy is stored in the phase difference between the two sides of the junction. The Josephson junction is then a purely inductive element with a non-quadratic potential. To build a non-harmonic oscillator, we connect it to a capacitor C. The circuit then consists of a superconducting island separated from the ground by a Josephson junction on one side and a capacitor on the other side. A schematic

of the circuit is shown in Fig. 2.2. The electromagnetic environment needs to be taken into account as well, as it generally applies an offset voltage V_g on the device through a gate capacitance C_g. The charge on the island reads $\hat{Q} = -2\hat{n}e$ with $\hat{n}$ the number of Cooper pairs, and the total electrostatic energy $E(\hat{n}) = 4E_C(\hat{n} - n_g)^2$ where $E_C \simeq \frac{e^2}{2C}$ is the characteristic charging energy of the capacitor, and $n_g \simeq \frac{C_g V_g}{2e}$ the offset charge taking into account the effect of V_g. This is valid in the regime where $C_g \ll C$, which is almost always true, as circuits are mostly designed to minimize the coupling to the environment. The

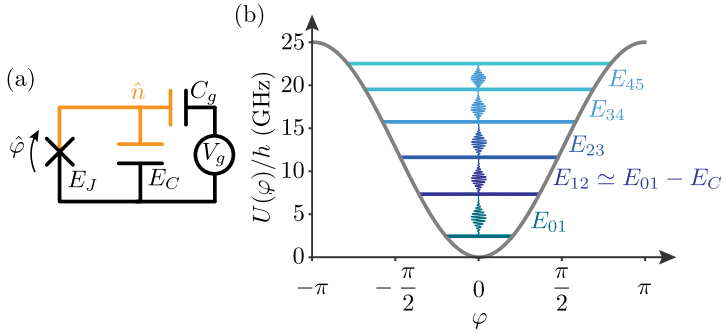


Figure 2.2: A non-linear oscillator. (a) Schematic of the device. The Josephson junction is symbolized by the cross. (b) Energy otential and spectrum at $n_g = 0$ of a transmon with $E_J/E_C = 50$ and $\sqrt{8E_J E_C}/h = 5$ GHz.

total Hamiltonian thus reads

$$H = E(\hat{n}) + U(\hat{\varphi}) = 4E_C(\hat{n} - n_g)^2 - E_J \cos(\hat{\varphi}). \tag{2.11}$$

Here, $\hat{n}$ and $\hat{\varphi}$ are linked by the commutation relation $[\hat{n}, \hat{\varphi}] = i$. We can now diagonalize this Hamiltonian and plot its spectrum $\{E_n(n_g)\}$ as a function of n_g. The result is shown in Fig 2.3 for four different values of the ratio E_J/E_C. These four plots are computed [15] for the same $E_J E_C$ product, such that $\sqrt{8E_J E_C}/h = 5$ GHz. We can see that the levels are not evenly spaced: with a drive at the right frequency, we can then address only the transition between the ground state and the first excited state. We get a periodicity in n_g of 1, which can be seen as the fact that the Hamiltonian is invariant if we set $n_g \to n_g + 1$ and $\hat{n} \to \hat{n} + 1$. As $\hat{n}$ has integer eigenvalues, this is the smallest increment

we can add to $\hat{n}$. The first devices built in the domain were in the regime $E_J/E_C < 10$ and named Cooper pair boxes. They were the first demonstrated superconducting qubits [16, 17].

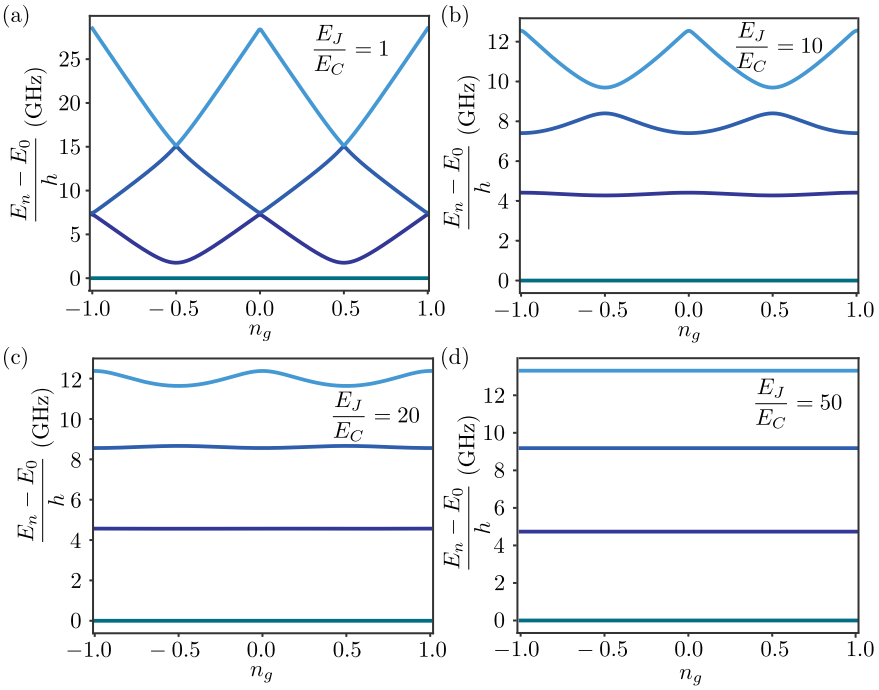


Figure 2.3: From the Cooper box to the transmon regime. (a) to (d): spectrum of the device for increasing ratio E_J/E_C. A qubit is obtained by selectively adressing the transition between the ground state and the first excited state.

At zero temperature, the performances of a qubit can be measured by two quantities. The first one is the decay time T_1, which is the caracteristic time during which the qubit remains excited after it was prepared in the excited state. The second one is the coherence time T_2, which characterizes how long the qubit keeps its phase φ after being prepared in the state $(|0\rangle + e^{i\varphi}|0\rangle)/\sqrt{2}$. The coherence time T_2 of these first Cooper pair box qubits, although measurable, was of the order of 1 ns [17]. This coherence time was limited by the circuit used to measure the qubit quantum state. Based on the Cooper pair box, the quantronium was developed to introduce a single shot readout and

mitigate the sensitivity to charge noise by introducing a sweet spot, which allowed them to reach a few hundred ns T_2.

However, these coherence times were typically not as good as they could be compared to the decay time T_1 of these systems. Indeed, it is less than 1 μs, whereas the T_1 of the order of 2 μs were demonstrated [18]. This is due to the strong dependence of its spectrum on the offset voltage V_g. This voltage is typically very noisy, which leads to a fluctuating frequency of the first transition of the device, thus inducing a lot of decoherence [19].

The transmon was invented to mitigate this problem: as it can be seen in Fig. 2.3, the higher the ratio E_J/E_C, the smaller the charge dispersion. We call the qubit a transmon in the regime where $E_J/E_C \gtrsim 10$. There, the charge dispersion ϵ_n, defined as $E_n(n_g = 1/2) - E_n(n_g = 0)$, reads [20]

$$\epsilon_n = (-1)^n E_C \frac{2^{4m+5}}{m!}\sqrt{\frac{2}{\pi}}\left(\frac{E_J}{2E_C}\right)^{\frac{m}{2}+\frac{3}{4}} e^{-\sqrt{8E_J/E_C}}. \tag{2.12}$$

The charge dispersion is thus exponentially suppressed as a function of $\sqrt{8E_J/E_C}$. In practice, $E_J/E_C > 20$ is enough to mitigate the charge noise and reach decent coherence time ($T_2 \geq 10$ μs), and $E_J/E_C > 40$ is necessary to be almost completely insensitive to the full charge dispersion. With this new device, decay times and coherence times increased to now reach more than $T_2 = 100$ μs [21], and $T_1 = 500$ μs [22]. Other designs of qubits, such as the fluxonium [23], demonstrated even better results, but the transmon is still widely used because of its performances that are good enough for a wide range of experiments and its relatively easy fabrication.

In this regime, $n_g = 0$is irrelevant for the first excited states and can be taken equal to 0. We can write the Hamiltonian in the form

$$\hat{H} = \left[\frac{1}{2C}\hat{Q}^2 + \frac{1}{2L_J}\hat{\Phi}^2\right] - \left[\cos\left(\frac{\hat{\Phi}}{\phi_0}\right) + \frac{1}{2L_J}\hat{\Phi}^2\right], \tag{2.13}$$

where $\hat{Q} = -2e\hat{n}$ and $L_J = \frac{\varphi_0^2}{2E_J}$ is the characteristic inductance of the Josephson junction. This Hamiltonian is then the sum of a linear part (first part in Eq. (2.13)) and a non-linear part (second part). The characteristic impedance $Z = \sqrt{\frac{L_J}{C}} = 2\sqrt{\frac{E_C}{E_J}}$ then sets the ratio of the

characteristic fluctuations of flux and charge. A high enough E_J/E_C ratio means that $\Delta\Phi/\phi_0 \ll 1$: the flux is then localized at the bottom of the potential. Writing annihilation operators as in Eq. (2.7), we have

$$\hat{b} = \frac{1}{\sqrt{2\hbar Z}}(\hat{\Phi} + iZ_c\hat{Q}) = \frac{1}{2}\Big(\frac{E_J}{2E_C}\Big)^{\frac{1}{4}}(\hat{\varphi} - 2i\sqrt{\frac{2E_C}{E_J}}\hat{n}). \tag{2.14}$$

In the ground state, $\langle\hat{\varphi}\rangle = 0$ and $\Delta\varphi = 2\Big(\frac{E_J}{2E_C}\Big)^{-\frac{1}{4}} \simeq 0.2\pi$ for $\frac{E_J}{E_C} = 50$: the phase is mostly localized around the origin. Writing $\varphi = \Delta\varphi(\hat{b}+\hat{b}^\dagger)$, we can expand the cosine potential up to the fourth order in $\hat{\varphi}$:

$$\hat{H}/\hbar \simeq \sqrt{8E_JE_C}(\hat{b}^\dagger\hat{b} + \frac{1}{2}) - E_J\frac{1}{4!}\Delta\varphi^4(\hat{b} + \hat{b}^\dagger)^4. \tag{2.15}$$

We can here develop the term in $(\hat{b}+\hat{b}^\dagger)^4$ and perform a rotating wave approximation (RWA) to keep only the resonant terms. Rearranging these terms in normal order gives

$$\hat{H}/\hbar = \sqrt{8E_JE_C}(\hat{b}^\dagger\hat{b} + \frac{1}{2}) - E_C\hat{b}^\dagger\hat{b} - \frac{E_C}{2}\hat{b}^{\dagger 2}\hat{b}^2. \tag{2.16}$$

The non-linearity thus renormalizes the frequency of the transmon by E_C/h. The transition frequencies $f_{01} = (E_1 - E_0)/h$ and $f_{12} = (E_2 - E_1)/h$ thus differ by the so-called anharmonicity $\alpha = E_C/h$. Typically, $\alpha \simeq 200$ MHz. This quantity sets the timescale of the fastest gates that can be performed on a qubit, as the maximum bandwidth of the pulses that can be sent on resonance to operate it [24]: if the pulse is too short, then it overlaps spectrally with the f_{12} transition, which excites the transmon into its second excited state, outside of the computational subspace. For a fixed frequency, E_C cannot be too large either to stay in the transmon regime. This is another reason why other more anharmonic devices, such as the fluxonium, are explored: higher anharmonicity potentially means faster gates.

2.2 REPRESENTATION OF QUANTUM STATES

In quantum mechanics, states are either an element $|\psi\rangle$ of the Hilbert space $\mathcal{H}$, for pure states, or statictical mixtures of these states, which can be represented by a density matrix $\hat{\rho}$, defined on the operator

space of $\mathcal{H}$,.The most straightforward way to represent the state is then by decomposing a pure state $|\psi\rangle$ on an orthonormal basis $\{|\psi_n\rangle\}$ of $\mathcal{H}$:

$$|\psi\rangle = \sum_n \psi_n |\psi_n\rangle . \tag{2.17}$$

The set of $\psi_n = \langle\psi_n|\psi\rangle$ are a representation of $|\psi\rangle$ in the basis $\{|\psi_n\rangle\}$. We can perform the same kind of decomposition for a density matrix $\hat{\rho}$

$$\hat{\rho} = \sum_{nm} \rho_{nm} |\psi_n\rangle\langle\psi_m| . \tag{2.18}$$

Here, the Hermitian property of the density matrix imposes $\rho^*_{nm} = \rho_{mn}$, its normalization $\mathrm{Tr}(\hat{\rho}) = 1$ that $\sum_n \rho_{nn} = 1$, and its positivity that its eigenvalues are positive.

Although this matrix representation contains all the information about the state, it is often more practical to work with graphical representations. The most ubiquitous one for two-level systems (TLS) is the Bloch sphere representation. Any state can indeed be represented (up to a global phase) by a unique vector inside the unit ball of $\mathbb{R}^3$. This representation is very useful, as any unitary operation performed on the state can be viewed as rotations. Dynamics in the Bloch sphere, governed by Bloch equations, are completely analogous to that of a classical dissipative system. This renders qubit dynamics a lot more intuitive.

The Bloch sphere can theoretically be generalized to any number of dimensions, but it becomes a lot less convenient for any number higher than 2, as the dimension of the sphere becomes larger than 3, which is the human limit for graphical representations. For a harmonic oscillator, the Bloch sphere representation and its infinite number of dimensions then happens on an infinite-dimensional ball.

Alternative representations of states for the harmonic oscillator exist. The most widely used are the so-called phase-space representations, which we will here introduce for the density matrix of a harmonic oscillator. They are *quasiprobability* distributions, which, despite not being actual probability distributions, give intuition about the distribution of actual observable outcomes. They are especially useful as means to visualize the dynamics of quantum states. Besides, their features

can give intuition about what we can define as the "quantumness" of a state. Finally, as we will discuss more in chapter 3, they give direct access to the probability distribution of measurement outcomes obtained routinely with superconducting circuits.

2.2.1 *Bloch ball*

It is convenient to start defining the Bloch sphere with pure states first before generalizing to mixed states on the whole ball. A state of a TLS (or qubit) can be written in an orthonormal basis $\{|1\rangle, |0\rangle\}$

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle \tag{2.19}$$

with a, b two complex numbers. Using that $|\langle\psi|\psi\rangle|^2 = 1$ gives $|a|^2 + |b|^2 = 1$, which means that we can parameterize a and b by an angle $\theta/2 \in [0, \pi/2]$ and two phases $\varphi_a, \varphi_b \in [0, 2\pi[$. The state $|\psi\rangle$ becomes:

$$|\psi\rangle = e^{i\varphi_a}\cos(\theta/2)\,|0\rangle + e^{i\varphi_b}\sin(\theta/2)\,|1\rangle \tag{2.20}$$

$$= e^{i\varphi_a}\big(\cos(\theta/2)\,|0\rangle + e^{i(\varphi_b-\varphi_a)}\sin(\theta/2)\,|1\rangle\big). \tag{2.21}$$

As in quantum mechanics, a global phase factor such as $e^{i\varphi_a}$ does not change any observable, we can safely take $\varphi_a = 0$. We now parameterized our state by $\theta \in [0, \pi]$ and $\varphi = \varphi_b - \varphi_a \in [0, 2\pi[$, the same way the unit sphere of $\mathbb{R}^3$ is parameterized. We can then associate to each $|\psi\rangle$ a unit vector $\vec{u}_{|\psi\rangle}$ called the Bloch vector, parameterized by θ and φ on the so-called Bloch sphere. The state $|0\rangle$ is then on the south pole of the Bloch sphere, while $|1\rangle$ is on the north pole. In general, the coordinates of the Bloch vector (x, y, z) reads:

$$(x, y, z) = \big(\langle\psi|\,\hat{\sigma}_x\,|\psi\rangle, \langle\psi|\,\hat{\sigma}_y\,|\psi\rangle, \langle\psi|\,\hat{\sigma}_z\,|\psi\rangle\big). \tag{2.22}$$

We introduce here the three Pauli operators $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ whose respective matrix representations are the Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ in the basis $\{|1\rangle, |0\rangle\}$

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.23}$$

This can be generalized to density matrices in a very simple and convenient way. To $\hat{\rho} = \sum_n P_n \, |\psi_n\rangle\langle\psi_n|$, we associate

$$\vec{u}_{\hat{\rho}} = \sum_n P_n \vec{u}_{|\psi_n\rangle}. \tag{2.24}$$

This association does not depend on the choice of the decomposition of $\hat{\rho}$. The Bloch vector of a density matrix is then the convex sum of the Bloch vectors of its components. Eq. (2.22) generalizes to:

$$(x, y, z) = \big(\mathrm{Tr}(\hat{\rho}\hat{\sigma}_x), \mathrm{Tr}(\hat{\rho}\hat{\sigma}_y), \mathrm{Tr}(\hat{\rho}\hat{\sigma}_z)\big) \tag{2.25}$$

and in turn, (x, y, z) gives the decomposition of $\hat{\rho}$ on the Pauli operators:

$$\hat{\rho} = \frac{1}{2}\big(\mathbb{1} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z\big). \tag{2.26}$$

We see that Bloch vectors of density matrices can live inside of the sphere, in the ball. There is an equivalence between $\hat{\rho}$ being a mixed state and $\vec{u}_{\hat{\rho}}$ being strictly inside the sphere.

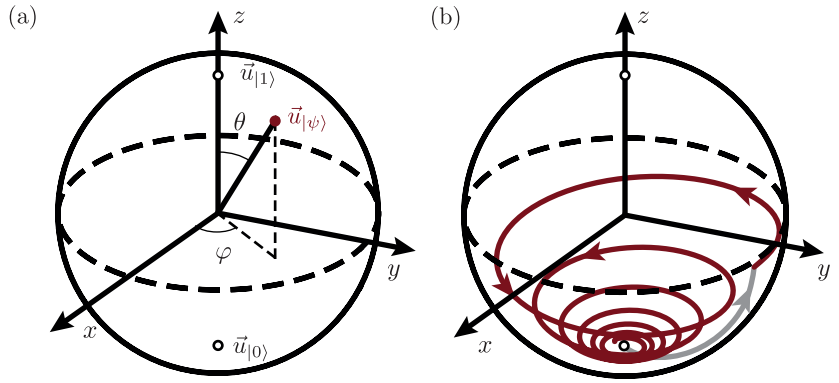


Figure 2.4: The Bloch sphere. (a) Bloch sphere schematic. (b) Example of trajectory on the Bloch sphere during a brief excitation (grey: rotation by an angle $\theta = \pi/2$) followed by relaxation of the qubit (red) without any pure dephasing of the qubit.

Starting from a pure state, $\hat{\rho}$ generically becomes mixed because of dissipation during its evolution, as the TLS gets entangled with its unmonitored environment. When the interaction with the environment

is Markovian, *i.e* when it has no memory of the past interactions, the most widely used equation governing both Hamiltonian evolution and dissipation is the Lindblad equation. In its most general form, it reads:

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}(t)] + \sum_n \Gamma_n \mathbb{D}_{\hat{A}_n}[\hat{\rho}(t)], \tag{2.27}$$

where $\hat{H}$ is the Hamiltonian of the system, $\mathbb{D}_{\hat{A}_n}[\hat{\rho}(t)] = \hat{A}_n \hat{\rho} \hat{A}_n^\dagger - \frac{1}{2}\{\hat{A}_n^\dagger \hat{A}_n \hat{\rho} + \hat{\rho} \hat{A}_n^\dagger \hat{A}_n\}$ are the dissipators associated to the jump operators $\hat{A}_n$, and the Γ_n are the associated jump rates. In the case of a qubit coupled to an outside electromagnetic field at thermal equilibrium at temperature T, we consider three main dissipation channels: excitation at rate $\Gamma_\uparrow = n_{\mathrm{th}}\Gamma_0$, associated to the jump operator $\hat{\sigma}_+ = \frac{1}{2}(\hat{\sigma}_x + i\hat{\sigma}_y) = |1\rangle\langle 0|$, de-excitation at rate $\Gamma_\downarrow = (1 + n_{\mathrm{th}})\Gamma_0$ associated to $\hat{\sigma}_- = \frac{1}{2}(\hat{\sigma}_x - i\hat{\sigma}_y) = |0\rangle\langle 1|$, and pure dephasing at rate $\Gamma_\varphi/2$, associated to $\hat{\sigma}_z$. $\Gamma_1 = \Gamma_\uparrow + \Gamma_\downarrow$ is the decay rate of the qubit, and n_{th} its thermal population of the qubit, *i.e.* its average number of excitations in thermal equilibrium with the environment. Detailed balance relates the temperature T to n_{th} through $\Gamma_\uparrow = e^{-\frac{\hbar\omega}{\mathrm{k_B}T}}\Gamma_\downarrow$, where $\mathrm{k_B}$ is the Boltzmann constant. Although the temperature is never zero for real systems, experimentalists work hard to make n_{th} as low as possible and negligible in practice. Setting $n_{\mathrm{th}} = 0$, the Lindblad equation then reads

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}(t)] + \Gamma_1 \mathbb{D}_{\hat{\sigma}_-}[\hat{\rho}(t)] + \frac{\Gamma_\varphi}{2}\mathbb{D}_{\hat{\sigma}_z}[\hat{\rho}(t)], \tag{2.28}$$

where we write $\hat{H}$ in the most general form for a qubit:

$$\hat{H}/\hbar = \frac{\Omega_x}{2}\hat{\sigma}_x + \frac{\Omega_y}{2}\hat{\sigma}_y + \frac{\delta}{2}\hat{\sigma}_z. \tag{2.29}$$

Ω_x and Ω_y are referred to as Rabi frequencies and δ to the drive detuning. This is the general Hamiltonian of a qubit considered in a Bloch ball frame rotating at a frequency detuned by δ from its resonant frequency. Multiplying Eq. (2.28) by $\hat{\sigma}_x, \hat{\sigma}_y$ and $\hat{\sigma}_z$ and taking the trace gives the optical Bloch equations for x, y and z:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \wedge \begin{pmatrix} \Omega_x \\ \Omega_y \\ \delta \end{pmatrix} - \begin{pmatrix} (\Gamma_1/2 + \Gamma_\varphi)x \\ (\Gamma_1/2 + \Gamma_\varphi)y \\ \Gamma_1(1+z) \end{pmatrix}. \tag{2.30}$$

The dynamics of the Bloch vector are then decomposed into two parts. The Hamiltonian part is the one materialized by the vector product. It consists of rotations in the Bloch ball. This part is strictly equivalent to that of a magnetic moment under a magnetic field $\vec{B} \propto \Omega_x \vec{e}_x + \Omega_y \vec{e}_y + \delta_z \vec{e}_z$ such that the Larmor pulsation ω is $\omega = \sqrt{\Omega_x^2 + \Omega_y^2 + \delta^2}$.

The second part, representative of the dissipation and dephasing, behaves like classical dissipation: the Bloch vector is linearly dampened towards the bottom of the Bloch sphere at a rate $\Gamma_1 = \Gamma_\downarrow$ and towards the z axis at a rate $\Gamma_2 = \Gamma_1/2 + \Gamma_\varphi$.

This representation thus allows classical intuition to work, which is very useful to picture quantum operations on qubits. An example of trajectory in the Bloch sphere is shown in Fig. 2.4. The qubit starts in the ground state $|0\rangle$ and is strongly driven during a short time, rotating around an axis close to the x-axis, During this time, we have $\Omega_x, \Omega_y \gg \delta, \Gamma_1, \Gamma_\varphi$ (grey arrow). Ω_x and Ω_y are then turned off, letting the qubit free to rotate at an angular speed δ while decaying back to $|0\rangle$ under the effect of the dissipation. This is the kind of trajectory experienced by the qubit in the experiment presented in Chap. 4.

2.2.2 *Phase-space representations of a harmonic oscillator*

As stated in the introduction of this chapter, there exist several different representations for the state of a harmonic oscillator. We present here the three main phase-space representations: the Glauber-Sudarshan P representation, the Wigner function W, and the Husimi Q function. Each of them represents, in some sense, the quantum version of the classical phase space. These three representations are defined using the coherent states and the displacement operators. Coherent states are called quasi-classical states, as they are the states that resemble the most classical states of a harmonic oscillator. A coherent state α is defined as

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{2.31}$$

It is the eigenvector of eigenvalue α of $\hat{a}$. It is called a gaussian state, as all its phase-space representations are gaussian functions. Fig. 2.7

shows schematics of each of them. It is obtained by applying the displacement operator $\hat{\mathcal{D}}(\alpha)$ to the vacuum state:

$$\hat{\mathcal{D}}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}} \tag{2.32}$$

$$|\alpha\rangle = \hat{\mathcal{D}}(\alpha)\,|0\rangle\,. \tag{2.33}$$

The displacement operator is the most classical operation we can apply on quantum states. This is even the only quantum operation we can implement on a harmonic oscillator using only classical resources [25]. Thus, if the oscillator starts in the vacuum state, the only states we can create are coherent states. Fock states, although the most simple states regarding the photon distribution, are inaccessible.

A first geometric property of the displacement operator is that it conserves the vector structure of $\mathbb{C}$:

$$\hat{\mathcal{D}}(\alpha)\hat{\mathcal{D}}(\beta) = e^{(\alpha\beta^* - \alpha^*\beta)/2}\hat{\mathcal{D}}(\alpha+\beta). \tag{2.34}$$

Thus, up to a global phase factor $^{(\alpha\beta^* - \alpha^*\beta)/2}$, the displacement operator $\hat{\mathcal{D}}(\alpha)$ behaves like a translation in phase space along a vector given by α.

Another useful operator is the squeezing operator $\hat{\mathcal{S}}(\zeta)$. It is defined as

$$\hat{\mathcal{S}}(\zeta) = e^{\frac{1}{2}(\zeta^*\hat{a}^2 - \zeta\hat{a}^{\dagger 2})}. \tag{2.35}$$

This operator is unitary and is of great importance, especially in the field of quantum metrology [26, 27], and for quantum measurement in general. As we will see in Chap. 3, any amplification procedure is intrinsically a squeezing procedure. This operator is called the "squeezing" operator, as its effect is to divide the variance of one quadrature by a factor $e^{|\zeta|}$ at the expense of the other which is multiplied by the same factor. The quadrature amplified (and the one deamplified) is set by $\theta = \mathrm{Arg}(\zeta)$. This operation becomes purely geometrical when working with Wigner functions (see Fig. 2.7 shows the Wigner function of a squeezed vacuum state). The properties of these operators, as well as that of the P, Q, and Wigner functions, are derived in [28]. We will use three main properties of the squeezed states in what follows. The first one is that

$$\hat{\mathcal{S}}(\zeta)^\dagger = \hat{\mathcal{S}}(-\zeta). \tag{2.36}$$

The inverse of a squeezing operation is thus the opposite squeezing operation. The second is that

$$\hat{\mathcal{S}}(-\zeta)\hat{a}\hat{\mathcal{S}}(\zeta) = \cosh(r)\hat{a} - \sinh(r)e^{i\theta}\hat{a}^\dagger \tag{2.37}$$

with $\zeta = re^{i\theta}$. This expression defines a transform on the annihilation and creation operators which is called a Bogoliubov transformation. The new creation operator $\hat{\mathcal{S}}(-\zeta)\hat{a}^\dagger\hat{\mathcal{S}}(\zeta)$ builds squeezed Fock states. The third is that

$$\hat{\mathcal{S}}(-\zeta)\hat{\mathcal{D}}(\alpha)\hat{\mathcal{S}}(\zeta) = \hat{\mathcal{D}}(\beta) \tag{2.38}$$

with

$$\beta = \cosh(r)\alpha + \sinh(r)e^{i\theta}\alpha^*. \tag{2.39}$$

This defines the commutation rules with the displacement operator, and the change of variables $\alpha \to \beta$ used to transform the Wigner function under the squeezing operator (see Sec. 2.2.2.2).

2.2.2.1 *The Glauber-Sudarshan P representation*

The coherent states form an *overcomplete* set of states. Thus, there exist a closure relation

$$\mathbb{1} = \frac{1}{\pi}\int_{\mathbb{C}} |\beta\rangle\langle\beta|\, \mathrm{d}^2\beta. \tag{2.40}$$

We use here the notation $\mathrm{d}^2\beta = \mathrm{d}\,\mathrm{Re}(\beta)\mathrm{d}\,\mathrm{Im}(\beta)$. The factor $1/\pi$ accounts for the overcompleteness: this ensemble is non-orthonormal and redundant, thus not normalized. This overcompletness comes with another interesting property: we can decompose any density matrix $\hat{\rho}$ of the harmonic oscillator in a diagonal form on the coherent states as

$$\hat{\rho} = \int_{\mathbb{C}} P_{\hat{\rho}}(\beta)\, |\beta\rangle\langle\beta|\, \mathrm{d}^2\beta. \tag{2.41}$$

The function P is unique and is called the Glauber-Sudarshan P function and It constitutes our first phase-space representation. It is defined on a complex plane and takes real values (in the sense of the distributions). It is a *quasiprobability* distribution: although normalized to unity ($\int_{\mathbb{C}} P_{\hat{\rho}}(\alpha)\mathrm{d}^2\alpha = 1$), it can take negative values.

As for the Wigner function, such negative value the signtature of a quantum state with no classical counterpart. Furthermore it can generate moments the same way a probability distribution would. In particular it directly generates any *normally ordered* moments of the annihilation and creation operators [29, 30]. For any $n, m \in \mathbb{N}$

$$\langle \hat{a}^{\dagger n} \hat{a}^m \rangle = \mathrm{Tr}\Big(\hat{\rho} \hat{a}^{\dagger n} \hat{a}^m\Big) = \int_{\mathbb{C}} \beta^{*n} \beta^m P_{\hat{\rho}}(\beta) \mathrm{d}^2\beta. \tag{2.42}$$

We can easily get the average value of $\hat{a}$ as $\langle \hat{a} \rangle = \int_{\mathbb{C}} \beta P_{\hat{\rho}}(\beta) \mathrm{d}^2\beta$. The average number of photons $\langle \hat{a}^\dagger \hat{a} \rangle$ is then directly given by $\langle \hat{a}^\dagger \hat{a} \rangle = \int_{\mathbb{C}} |\beta|^2 P_{\hat{\rho}}(\beta) \mathrm{d}^2\beta$.

Displacement operators are very simple to implement on the P function. It is indeed easy to show that the P function of $\hat{\rho}' = \hat{\mathcal{D}}(\alpha)\hat{\rho}\hat{\mathcal{D}}(-\alpha)$ is given by $P_{\hat{\rho}'}(\beta) = P_{\hat{\rho}}(\beta - \alpha)$. The displacement operator thus translates the P function.

The inverse transformation from $\hat{\rho}$ to P is given by [29]:

$$P_{\hat{\rho}}(re^{i\theta} = \beta) = \sum_{n,m} \langle n|\hat{\rho}|m\rangle \frac{\sqrt{n!m!}}{2\pi r (n+m)!} e^{r^2 - i(n-m)\theta} \Big[\Big(-\frac{\partial}{\partial r} \Big)^{n+m} \delta(r) \Big]. \tag{2.43}$$

P thus contains the same information as the density matrix. However, for pure states, this is a highly singular function. For a coherent state $\hat{\rho} = |\alpha\rangle\langle\alpha|$, $P_{|\alpha\rangle}(\beta) = \delta(\beta - \alpha)$, with δ the 2D Dirac function. For a pure Fock state $|n\rangle$, it becomes even more singular: $P_{|n\rangle}(\beta) = \frac{n!}{2\pi r (2n)!} e^{r^2} \Big[\big(\frac{\partial}{\partial r}\big)^{2n} \delta(r) \Big]$ which involves derivatives of the Dirac function. These reasons make it a quantity that is not convenient to visualize and hard to manipulate.

2.2.2.2 *The Wigner function*

The most commonly used representation in the field of superconducting circuits is the Wigner function. It is defined by the Wigner transform of the density matrix, which reads

$$W_{\hat{\rho}}(\beta) = \frac{2}{\pi} \mathrm{Tr}\Big(\hat{\mathcal{D}}(-\beta)\hat{\rho}\hat{\mathcal{D}}(\beta)\hat{\mathcal{P}}\Big), \tag{2.44}$$

where $\hat{\mathcal{P}} = (-1)^{\hat{a}^\dagger \hat{a}}$ is the parity operator. The value of $W_{\hat{\rho}}(\beta)$ is then the average parity of $\hat{\rho}$ displaced by $-\beta$. This expression is especially

useful in experiments, as having a qubit dispersively coupled to a cavity allows us to directly implement this measurement. This measurement process is called *Wigner tomography*.

The Wigner function is linked to the P function by a very simple relation:

$$W_{\hat{\rho}}(\beta) = (P_{\hat{\rho}} * W_{|0\rangle})(\beta) \tag{2.45}$$

where $W_{|0\rangle}(\beta) = \frac{2}{\pi} e^{-2|\beta|^2}$ is the Wigner function of the vacuum state.

It is also a quasiprobability distribution: $\int_{\mathbb{C}} W_{\hat{\rho}}(\beta) \mathrm{d}^2\beta = \mathrm{Tr}(\hat{\rho}) = 1$, and it can be negative. However, it is already much more convenient than the P function, as it is a smooth function in general (for states with finite energy). Furthermore, negativity in the Wigner function is fundamentally linked to the "quantumness" of a state, *i.e.* the ability to exhibit non-classical behavior. For example, in continuous-variables quantum computing, states that exhibit Wigner negativity are necessary to get any speed-up compared to classical algorithms in the field of continuous-variable quantum computing [31].

Its moments are the *symmetrically ordered* moments of the creation and annihilation operators [30]:

$$\langle \left(\hat{a}^n \hat{a}^{\dagger m}\right)_{\mathcal{S}} \rangle = \int_{\mathbb{C}} \beta^n \beta^{*m} W_{\hat{\rho}}(\beta) \mathrm{d}^2\beta \tag{2.46}$$

Here, $(.)_{\mathcal{S}}$ stands for the symmetric ordering of the operators. As an example, one can get the average number of photons $\langle \hat{a}^\dagger \hat{a} \rangle$ from the Wigner function: the first step is to write it in a symmetric way: $\hat{a}^\dagger \hat{a} = \frac{1}{2}(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger - 1)$. Then

$$\langle \hat{a}^\dagger \hat{a} \rangle = \langle \left(\hat{a}^\dagger \hat{a}\right)_{\mathcal{S}} \rangle - \frac{1}{2} = \int_{\mathbb{C}} |\beta|^2 W_{\hat{\rho}}(\beta) \mathrm{d}^2\beta - \frac{1}{2}, \tag{2.47}$$

and as for the P representation, we also have

$$\langle \hat{a} \rangle = \int_{\mathbb{C}} \beta W_{\hat{\rho}}(\beta) \mathrm{d}^2\beta. \tag{2.48}$$

.

The Wigner function is a tool that goes beyond visualizing the states, as the Wigner transform can also be applied on operators the

same way as for density matrices. For any operator $\hat{O}$, we define its Wigner transform as

$$W_{\hat{O}}(\beta) = \frac{2}{\pi}\,\mathrm{Tr}\Big(\hat{\mathcal{D}}(-\beta)\hat{O}\hat{\mathcal{D}}(\beta)\hat{\mathcal{P}}\Big). \tag{2.49}$$

The expectation value of $\hat{O}$ can then be obtained using

$$\langle\hat{O}\rangle = \mathrm{Tr}\Big(\hat{\rho}\hat{O}\Big) = \pi\int_{\mathbb{C}} W_{\hat{\rho}}(\beta)W_{\hat{O}}(\beta)\mathrm{d}^2\beta. \tag{2.50}$$

Wigner functions of frequently used operators are listed below:

$$W_{\hat{a}}(\beta) = \frac{\beta}{\pi} \tag{2.51}$$

$$W_{\hat{a}^\dagger\hat{a}}(\beta) = \frac{1}{\pi}(|\beta|^2 - \frac{1}{2}) \tag{2.52}$$

$$W_{|n\rangle\langle n|}(\beta) = W_{|n\rangle}(\beta) = (-1)^n\frac{2}{\pi}e^{-2|\beta|^2}L_n(4|\beta|^2), \tag{2.53}$$

where L_n is the n-th Laguerre polynomial. We also get $W_{|n\rangle\langle m|}$ by

$$W_{|n\rangle\langle m|}(x+ip) = \frac{1}{\pi}\int e^{-ipy}\Psi_n(x+\frac{y}{2})\Psi_m^*(x-\frac{y}{2})\mathrm{d}y, \tag{2.54}$$

where Ψ_n is the wavefunction of the Fock state $|n\rangle$:

$$\Psi_n(x) = \langle x|n\rangle = (\frac{2}{\pi})^{\frac{1}{4}}\frac{1}{\sqrt{2^n n!}}\mathcal{H}_n(\sqrt{2}x)e^{-x^2} \tag{2.55}$$

with $\mathcal{H}_n$ the n-th Hermite polynomial. The $W_{|n\rangle\langle m|}$ are then all we need to reconstruct a density matrix in the Fock state basis from the Wigner tomography, using that $\langle n|\hat{\rho}|m\rangle = \mathrm{Tr}(\hat{\rho}\,|n\rangle\langle m|)$ and Eq. (2.49)

Another theoretical way of reconstructing the density matrix (or an operator) from the Wigner function is by using the Weyl transform:

$$\hat{O} = \iiiint W_{\hat{O}}(x+ip)e^{ik_x(\hat{x}-x)+ik_p(\hat{p}-p)}\mathrm{d}x\mathrm{d}p\mathrm{d}k_x\mathrm{d}k_y. \tag{2.56}$$

This formula, less convenient, is not used in practice.

Finally, for a pure state $|\psi\rangle$, we recover an actual probability density by taking the marginal of the Wigner function along any quadrature. In particular:

$$\int_{\mathbb{R}} W_{|\psi\rangle}(x+ip)\mathrm{d}p = |\,\langle x|\psi\rangle\,|^2 \tag{2.57}$$

$$\int_{\mathbb{R}} W_{|\psi\rangle}(x+ip)\mathrm{d}x = |\,\langle p|\psi\rangle\,|^2. \tag{2.58}$$

Here, $|x\rangle$ and $|p\rangle$ have to be understood as the "eigenvectors" of the $\hat{x} = (\hat{a} + \hat{a}^\dagger)/2$ and $\hat{p} = (\hat{a}^\dagger - \hat{a})/2i$ the two adimensional quadrature operators. Note that they verify here $[\hat{x}, \hat{p}] = i/2$ [1].

Experimental tomography of Wigner functions of Fock states $|0\rangle$, $|1\rangle$, $|5\rangle$ and $|9\rangle$ are shown Fig 2.5. Details about the way these Wigners were obtained are given in Sec. 4.3.2 and 4.9.1.

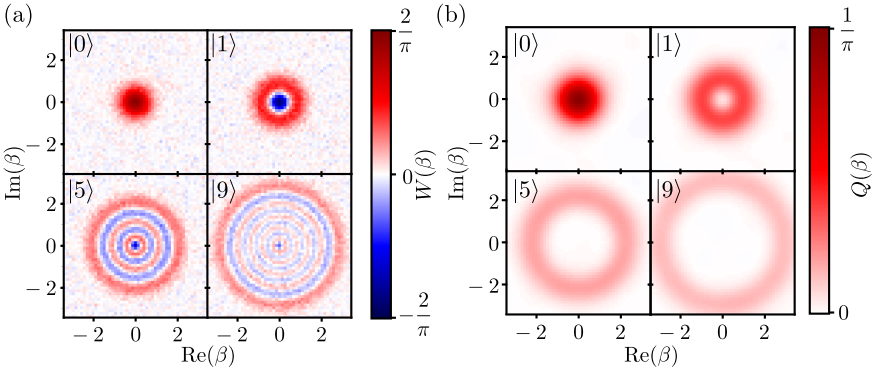


Figure 2.5: Experimental Wigner and Husimi Q functions of a few Fock states. (a) Measured Wigner functions of Fock states $|0\rangle$, $|1\rangle$, $|5\rangle$ and $|9\rangle$ that were prepared by heralding on the fluorescence of a dispersively coupled qubit. (b) Corresponding Husimi Q functions for the same states obtained using Eq. (2.67).

The displacement and squeezing operators correspond to simple geometrical operations on the Wigner functions. It is straightforward to see that the effect of the displacement operator $\hat{\mathcal{D}}(\alpha)$ is, as for the P function, to displace the Wigner function by α: $W_{\hat{\rho}}(\beta) \xrightarrow{\hat{\mathcal{D}}(\alpha)} W_{\hat{\rho}}(\beta - \alpha)$. The effect of a squeezing operator $\hat{\mathcal{S}}(\zeta)$ on the Wigner function is also a change of coordinates:

$$W_{\hat{\rho}}(\alpha) \xrightarrow{\hat{\mathcal{S}}(\zeta)} W^{\zeta}_{\hat{\rho}} = W_{\hat{\rho}}\big(\alpha \cosh(r) + \alpha^* e^{i\theta} \sinh(r)\big) \tag{2.59}$$

1 Other conventions for the definition of these operators exist, in particular $\hat{x} = (\hat{a} + \hat{a}^\dagger)/\sqrt{2}$ and $\hat{p} = (\hat{a}^\dagger - \hat{a})/\sqrt{2}i$, leading to $[\hat{x}, \hat{p}] = i$, which looks more natural, given that $[\hat{\Phi}, \hat{Q}] = i\hbar$. This is the choice made in [32]. We prefer our convention, as it gives the simple relation $\hat{a} = \hat{x} + i\hat{p}$

This amounts to transforming $\hat{a}$ and $\hat{a}^\dagger$ in Eq. (2.44) following

$$\hat{a} \to \hat{\mathcal{S}}(-\zeta)\hat{a}\hat{\mathcal{S}}(\zeta) = \cosh(r)\hat{a} - \sinh(r)e^{i\theta}\hat{a}^\dagger \tag{2.60}$$

$$\hat{a}^\dagger \to \hat{\mathcal{S}}(-\zeta)\hat{a}^\dagger\hat{\mathcal{S}}(\zeta) = \cosh(r)\hat{a}^\dagger - \sinh(r)e^{-i\theta}\hat{a}. \tag{2.61}$$

We can illustrate this transformation by applying it on the vacuum state. This gives a squeezed vacuum state:

$$W^{\zeta}_{|0\rangle}(x+ip) = \frac{2}{\pi}e^{-2e^{2r}(\cos\left(\frac{\theta}{2}\right)x+\sin\left(\frac{\theta}{2}\right)p)^2}e^{-2e^{-2r}(\cos\left(\frac{\theta}{2}\right)p-\sin\left(\frac{\theta}{2}\right)x)^2} \tag{2.62}$$

which has a very simple geometrical interpretation: the original Wigner function is compressed by e^r in the direction given by $ei^{\frac{\theta}{2}}$ and expanded by e^r in the orthogonal direction. This is the general way squeezing acts on Wigner function. For $\theta = 0$, it gives

$$W^{\zeta}_{|0\rangle}(x+ip) = \frac{2}{\pi}e^{-2e^{2r}x^2}e^{-2e^{-2r}p^2}. \tag{2.63}$$

It is then straightforward to compute $\Delta\hat{x}^2 = \langle\hat{x}^2\rangle = e^{-2r}/2$ and $\Delta\hat{p}^2 = \langle\hat{p}^2\rangle = e^{2r}/2$. These states still saturate the Heisenberg inequality $\Delta\hat{x}\Delta\hat{p} \geq 1/2$ but have reduced fluctuations in one quadrature. This way of reducing the uncertainty in one direction is a standard tool in quantum metrology [33]. These states are especially useful for the detection of gravitational waves [26, 27].

2.2.2.3 *The Husimi Q function*

The expression of the third phase-space representation is the simplest: this is the average of the density matrix for coherent states [34, 35]:

$$Q_{\hat{\rho}}(\alpha) = \frac{1}{\pi}\langle\alpha|\hat{\rho}|\alpha\rangle. \tag{2.64}$$

We can rewrite this expression as

$$Q_{\hat{\rho}}(\alpha) = \langle\alpha|\hat{\rho}|\alpha\rangle = \mathrm{Tr}\Big(|0\rangle\langle 0|\,\hat{\mathcal{D}}(-\alpha)\hat{\rho}\hat{\mathcal{D}}(\alpha)\Big). \tag{2.65}$$

We can interpret this expression as the probability of finding 0 photons in the state displaced by α. Knowing that the Wigner function of a

coherent state $W_{|\alpha\rangle}$ is $W_{|\alpha\rangle}(\beta) = W_{|0\rangle}(\beta - \alpha)$, we can use Eq. (2.50) applied to $\hat{O} = \hat{\mathcal{D}}(\alpha)\,|0\rangle\langle 0|\,\hat{\mathcal{D}}(-\alpha)$:

$$Q_{\hat{\rho}}(\alpha) = \int_{\mathcal{C}} W_{|0\rangle}(\beta - \alpha) W_{\hat{\rho}}(\beta) \mathrm{d}^2\beta. \tag{2.66}$$

Thus, Q and W are all obtained from P by a convolution with the Wigner function of the vacuum state:

$$Q_{\hat{\rho}} = W_{\hat{\rho}} * W_{|0\rangle} = P_{\hat{\rho}} * W_{|0\rangle} * W_{|0\rangle} = P_{\hat{\rho}} * Q_{|0\rangle}. \tag{2.67}$$

Plus, we now see that this representation does not miss any information compared to the Wigner function and the P function, as the convolution with a Gaussian function is a reversible operation. However, this comes with an increased sensitivity to noise, which makes it hard to use in actual experiments for state reconstruction (see Fig. 2.5).

On top of furnishing another geometrical representation of the states, its importance lies in the fact that it gives the distribution the outcome of a heterodyne measurement, which is the most widely used type of measurement in superconducting circuits (see Sec. 3.1.2). The fact that Q is positive also gives an intuition about the allowed negative regions of the Wigner function: their extent needs to be small enough so that convolving it with $W_{|0\rangle}$, whose typical spatial extension is $1/2$, makes them disappear. For pure states, the presence of negative regions in the Wigner function is equivalent to the presence of zeros of the Q function [36].

Another more constructive way of writing the Q function exists. Let $\hat{\rho} = |\psi\rangle\langle\psi|$ be a pure state, and

$$|\psi\rangle = \sum_k \psi_k |k\rangle = \sum_k \frac{\psi_k}{\sqrt{k!}} \hat{a}^{\dagger k} |0\rangle \tag{2.68}$$

its decomposition on the Fock basis. We can define the stellar function $F^\star$ [37] as

$$F^\star_{|\psi\rangle}(\alpha) = \sum_k \frac{\psi_k}{\sqrt{k!}} \alpha^k. \tag{2.69}$$

In this notation, we simply have $|\psi\rangle = F^{\star}_{|\psi\rangle}(\hat{a}^\dagger)\,|0\rangle$. Eq. (2.64) then gives

$$Q_{|\psi\rangle}(\alpha) = \frac{1}{\pi}\langle\alpha|F^{\star}_{|\psi\rangle}(\hat{a}^\dagger)\,|0\rangle\langle 0|\,F^{\star\star}_{|\psi\rangle}(\hat{a})|\alpha\rangle \tag{2.70}$$

$$= e^{-|\alpha|^2}\frac{1}{\pi}|F^{\star}_{|\psi\rangle}(\alpha^*)|^2. \tag{2.71}$$

This expression is especially useful when the ψ_n are known, as it gives a straightforward way of writing the Q function from the decomposition in the Fock basis. It generalizes to density matrices by summing up the Q functions in the same way as pure density matrices: if $\hat{\rho} = \sum_n P_n\,|\psi_n\rangle\langle\psi_n|$, then $Q = \sum_n P_n Q_{|\psi_n\rangle}$. For example, if the oscillator is in a thermal state of temperature T defined as

$$\hat{\rho}_{\text{th}} = (1-\nu)\sum_n \nu^n\,|n\rangle\langle n| \tag{2.72}$$

where $\nu = e^{-\frac{\hbar\omega_c}{\mathrm{k_B}T}}$ is the Boltzmann factor, and $\omega_c/2\pi$ the frequency of the mode. The Husimi Q function then reads

$$Q_{\hat{\rho}_{\text{th}}}(\alpha) = \frac{1-\nu}{\pi}e^{(\nu-1)|\alpha|^2}. \tag{2.73}$$

The Q function of a thermal state is thus simply widened by a factor $1/\sqrt{1-\nu}$ compared to that of the vacuum state:

$$Q_{\hat{\rho}_{\text{th}}}(\alpha) = \sqrt{1-\nu}\,Q_{|0\rangle}(\alpha\sqrt{1-\nu}). \tag{2.74}$$

Writing $Q_{\hat{\rho}_{\text{th}}} = W_{|0\rangle} * W_{\hat{\rho}_{\text{th}}}$, we can deconvolve $Q_{\hat{\rho}_{\text{th}}}$ by the Wigner function of the vacuum state (which is easy as all the distributions involved are Gaussian) to also get the Wigner function of a thermal state:

$$W_{\hat{\rho}_{\text{th}}}(\alpha) = \frac{2(1-\nu)}{\pi(1+\nu)}e^{-\frac{2(1-\nu)}{1+\nu}|\alpha|^2}. \tag{2.75}$$

Similarly, $W_{\hat{\rho}_{\text{th}}}$ is then obtained by scaling the Wigner function of the vacuum:

$$W_{\hat{\rho}_{\text{th}}}(\alpha) = \sqrt{\frac{1-\nu}{1+\nu}}W_{|0\rangle}\Big(\sqrt{\frac{1-\nu}{1+\nu}}\alpha\Big). \tag{2.76}$$

The variance $V_{\hat{\rho}_{\rm th}}$ of this Wigner function thus reads

$$V_{\hat{\rho}_{\rm th}} = \frac{1+\nu}{2(1-\nu)} = \frac{1}{2} + \frac{\nu}{1-\nu} = \frac{1}{2} + n_{\rm th}, \tag{2.77}$$

where $n_{\rm th} = \frac{\nu}{1-\nu}$ is the average number of photons in the mode. Deconvolving again by the Wigner function of the vacuum state, we can even obtain the P distribution $P_{\hat{\rho}_{\rm th}}$ of $\rho_{\rm th}$ as a Gaussian function with variance $n_{\rm th}$.

The moments of the Q function are the *anti-normally* ordered moments of $\hat{a}$ and $\hat{a}^\dagger$ [30]:

$$\langle \hat{a}^n \hat{a}^{\dagger m} \rangle = \int_{\mathbb{C}} \beta^n \beta^{*m} Q_{\hat{\rho}}(\beta) \mathrm{d}^2\beta \tag{2.78}$$

In particular, we have $\langle \hat{a} \rangle = \int_{\mathbb{C}} \beta Q_{\hat{\rho}}(\beta) \mathrm{d}^2\beta$ and $\langle \hat{a}^\dagger \hat{a} \rangle = \int_{\mathbb{C}} (|\beta|^2 - 1) Q_{\hat{\rho}}(\beta) \mathrm{d}^2\beta$. Examples of experimentally measured Q functions for Fock states $|0\rangle$, $|1\rangle$, $|5\rangle$ and $|9\rangle$ are given Fig. 2.5. They are obtained from the corresponding Wigner functions using Eq. (2.67). Interestingly, the experimental noise is almost invisible compared to the noise on the Wigner function. We can illustrate how hard it is to experimentally perform a proper state reconstruction from the Q function by comparing the Q function of distant states. The first one is an odd cat state $|C_\alpha^-\rangle \propto |\alpha\rangle - |-\alpha\rangle$, which is a superposition of two opposite coherent states. The second is given by $\hat{\rho}_\alpha = \frac{1}{2}(|\alpha\rangle\langle\alpha| + |-\alpha\rangle\langle-\alpha|)$, which is a classical mixture of the same coherent states. We see in Fig. 2.6 for $\alpha = 2$ that the two Wigner functions exhibit very different features: the Wigner function of $|C_\alpha^-\rangle$ exhibits negative regions, which is not the case for $\hat{\rho}_\alpha$. However, these striking differences almost disappear when comparing the Q functions on a linear scale. We can distinguish some very faint structure between the two Gaussian components for $|C_\alpha^-\rangle$, but that's pretty much everything. To understand where the key differences between the two Q functions lie, one needs to visualize them in log scale (right panel Fig. 2.6). The Q function of $|C_\alpha^-\rangle$ exhibits zeros on the vertical axis, which is not the case for $\hat{\rho}_\alpha$. Actually, it has an infinite number of such zeros, one for each negative region in the Wigner function [38]. These zeros are present in a region where the Q function is already very close to zero and do not survive any experimental noise, which renders very hard faithful experimental discrimination between the pure $|C_\alpha^-\rangle$ and 1-bit entropic state $\hat{\rho}_\alpha$.

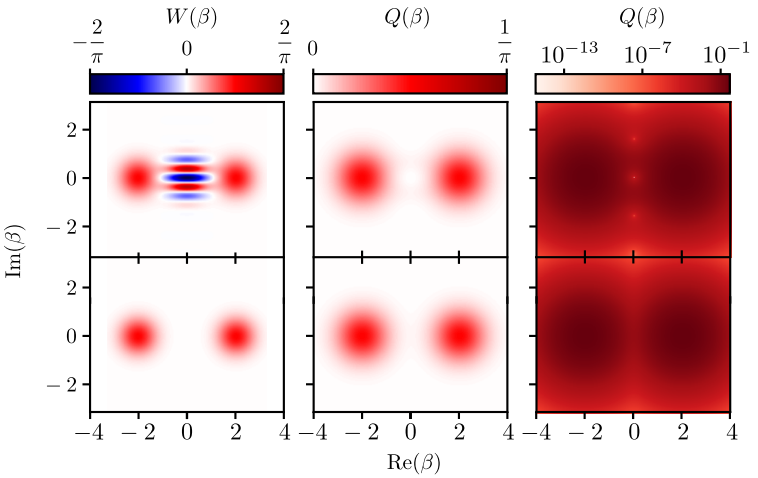


Figure 2.6: Comparison between the Wigner function and the Husimi Q function of coherent and incoherent superposition of two coherent states. Top, from left to right: Wigner functions for an odd cat state $\left|C_2^-\right\rangle$, its Q function in linear scale and its Q function in logarithmic scale. Bottom, from left to right: Wigner functions for a mixture of coherent states $\hat{\rho}_\alpha = \frac{1}{2}(|\alpha\rangle\langle\alpha| + |-\alpha\rangle\langle-\alpha|)$, its Q function in linear scale and its Q function in logarithmic scale.

2.2.2.4 *Generalized Q functions*

The notion of Q function can be generalized by replacing the vacuum state $|0\rangle\langle 0|$ in Eq. (2.65) by an arbitrary density matrix $\hat{\sigma}$:

$$Q_{\hat{\rho}}^{\hat{\sigma}}(\alpha) = \frac{1}{\pi}\,\mathrm{Tr}\Big(\hat{\sigma}\hat{\mathcal{D}}(-\alpha)\hat{\rho}\hat{\mathcal{D}}(\alpha)\Big). \tag{2.79}$$

It is straightforward to generalize Eq. (2.67) to

$$Q_{\hat{\rho}}^{\hat{\sigma}} = W_{\hat{\rho}} * W_{\hat{\sigma}} \tag{2.80}$$

where $W_{\hat{\sigma}}$ is the Wigner function of the state $\hat{\sigma}$.

As both $\hat{\sigma}$ and $\hat{\mathcal{D}}(-\alpha)\hat{\rho}\hat{\mathcal{D}}(\alpha)$ are density matrices, this is the trace of the product of two positive operators, which is always positive. This property can be seen as the fact that the average value of an observable with only positive outcome (such as a density matrix) is always positive. This gives a much stronger property to the Wigner

function: its convolution with any Wigner function representing a quantum state is positive.

For example, as in [39], generalized $Q^{\hat{\sigma}}$ functions for $\hat{\sigma} = |n\rangle\langle n|$ can be measured in a cavity using a dispersively coupled qubit. With these states, the measurement can be understood intuitively: it is the probability of finding n photons in the state displaced by α, which provides an experimental protocol to measure $Q^{\hat{\sigma}}$ functions.

The most interesting choice studied in this thesis consists in taking $|\sigma\rangle\langle\sigma| = \hat{\mathcal{S}}(\zeta)\,|0\rangle\langle 0|\,\hat{\mathcal{S}}(-\zeta)$ a squeezed vacuum state. This function is crucial to understanding the statistics of the measurement outcomes discussed in Chap. 3, in particular in the experimental implementation of homodyne measurements.

The equation (2.80) simply gives

$$Q^{\zeta}_{\hat{\rho}} = W_{\hat{\rho}} * W^{\zeta}_{|0\rangle}. \tag{2.81}$$

Instead of taking the convolution with a symmetric Gaussian function, we use a squeezed one. An example is shown in Fig. 2.7a. In black, the P distribution $P_{|\alpha\rangle}$ of a coherent state $|\alpha\rangle = 1 + i$. In red, its Wigner function $W_{|\alpha\rangle}$. It is a Gaussian distribution with quadrature standard deviations $\Delta x_W = \Delta P_W = \Delta_0 = 1/2$, with Δ_0 that of the vacuum state. In blue, its Husimi Q distribution, obtained by convolution with the Wigner function of the vacuum state $W_{|0\rangle}$. The result is a Gaussian function with quadrature standard deviations $\Delta x_Q = \Delta P_Q = \sqrt{\Delta x_W^2 + \Delta_0^2} = 1/\sqrt{2}$. In green, its generalized Q function obtained by convolution of the Wigner function $W_{\hat{\mathcal{S}}(\zeta)|0\rangle}$ of a squeezed vacuum state, plotted in fainter red. In this example, the squeezing factor ζ is $\zeta = \log(2)$. $Q^{\zeta}_{\hat{\rho}}$ is a Gaussian distribution, whose quadrature standard deviations are $\Delta x_{Q^{\zeta}_{\hat{\rho}}} = \sqrt{\Delta x_W^2 + (\Delta_0 e^{-\zeta})^2} = \sqrt{5}/4$ and $\Delta P_{Q^{\zeta}_{\hat{\rho}}} = \sqrt{\Delta P_W^2 + (\Delta_0 e^{\zeta})^2} = \sqrt{5}/2$.

The properties given by Eq. (2.57) and (2.58) can be thought of as a limit where r goes to infinity in Eq. (2.81) for $\theta = 0$ and $\theta = \pi$: as one quadrature gets infinitely squeezed, the Husimi Q function becomes essentially a 1D distribution on the other quadrature, given by the modulus square of the wavefunction expressed in this quadrature.

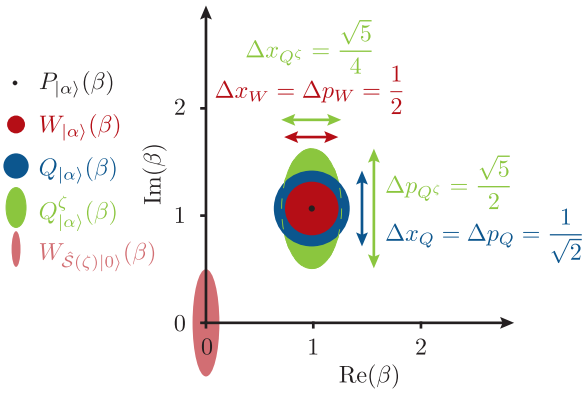


Figure 2.7: P distribution $P_{|\alpha\rangle}$, Wigner function $W_{|\alpha\rangle}$, Q function $Q_{|\alpha\rangle}$ and generalized Q function $Q^{\zeta}_{|\alpha\rangle}$ of a coherent state $|\alpha\rangle = |1+i\rangle$ with $\zeta = \log(2)$. The latter is obtained by convolving $W_{|\alpha\rangle}$ with the Wigner function of the squeezed vacuum $W_{\hat{\mathcal{S}}(\zeta)|0\rangle}$.

A second way of picturing this transformation gives an intuition of what is measured with a phase-sensitive amplifier, as developed in Sec. 3.2.3. The equation (2.79) gives, for a squeezed vacuum state:

$$Q^{\zeta}_{\hat{\rho}}(\beta) = \frac{1}{\pi}\,\mathrm{Tr}\Big(\hat{\mathcal{S}}(\zeta)\,|0\rangle\langle 0|\,\hat{\mathcal{S}}(-\zeta)\hat{\mathcal{D}}(-\beta)\hat{\rho}\hat{\mathcal{D}}(\beta)\Big). \tag{2.82}$$

Using the invariance of the trace by permutation, inserting $\hat{\mathcal{S}}(\zeta)\hat{\mathcal{S}}(-\zeta) = \mathbb{1}$ and using Eq. (2.38) gives

$$\begin{aligned} &Q^{\zeta}_{\hat{\rho}}(\beta) = \\ &\frac{1}{\pi}\mathrm{Tr}\Big(|0\rangle\langle 0|\hat{\mathcal{S}}(-\zeta)\hat{\mathcal{D}}(-\beta)\hat{\mathcal{S}}(\zeta)\hat{\mathcal{S}}(-\zeta)\hat{\rho}\hat{\mathcal{S}}(\zeta)\hat{\mathcal{S}}(-\zeta)\hat{\mathcal{D}}(\beta)\hat{\mathcal{S}}(\zeta)\Big) \\ &\qquad = \mathrm{Tr}\Big(|0\rangle\langle 0|\,\hat{\mathcal{D}}(-\gamma)\hat{\mathcal{S}}(-\zeta)\hat{\rho}\hat{\mathcal{S}}(\zeta)\hat{\mathcal{D}}(\gamma)\Big) \\ &\qquad = Q_{\hat{\mathcal{S}}(-\zeta)\hat{\rho}\hat{\mathcal{S}}(\zeta)}(\gamma), \end{aligned} \tag{2.83}$$

with

$$\gamma = \cosh(r)\beta + \sinh(r)e^{i\theta}\beta^*. \tag{2.84}$$

The inverse change of variable is given by

$$\beta = \cosh(r)\gamma - \sinh(r)e^{i\theta}\gamma^*. \tag{2.85}$$

$Q^{\zeta}_{\hat{\rho}}$ can thus geometrically be obtained by squeezing the Wigner function of $\hat{\rho}$ by $-\zeta$, (which gives the Wigner function of $\hat{\mathcal{S}}(-\zeta)\hat{\rho}\hat{\mathcal{S}}(\zeta)$), convolving with the Wigner function of the vacuum state (which gives the Q function of $\hat{\mathcal{S}}(-\zeta)\hat{\rho}\hat{\mathcal{S}}(\zeta)$), and scaling it back as a function of β with the change of variable $\gamma \to \beta$ with $\beta = \cosh(r)\gamma - \sinh(r)e^{i\theta}\gamma^*$. The procedure is pictured in Fig. 2.8. Thus, the Husimi Q function of the (inversely) squeezed states gives, up to a coordinate change, the same distribution as the generalized Q function.

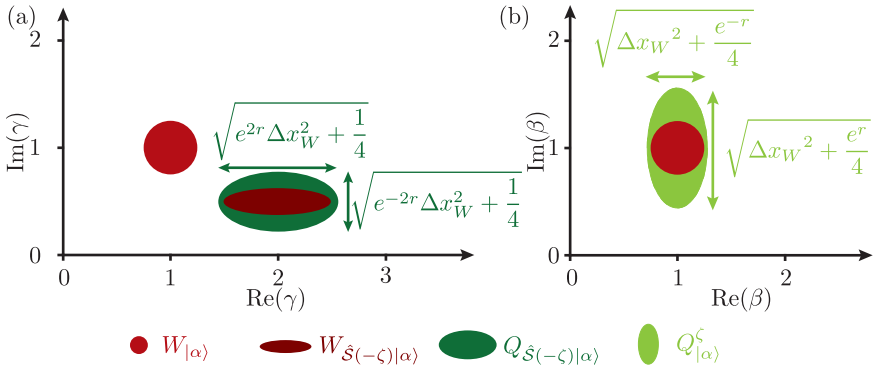


Figure 2.8: Alternative way of obtaining the generalized Q function $Q^{\zeta}_{|\alpha\rangle}$ from the Wigner function $W_{|\alpha\rangle}$ with $\alpha = 1$ and $\zeta = \log(2)$ as an example. (a) First steps of the transformation. The state is squeezed with parameter $-\zeta$, which gives the Husimi Q function $Q_{\hat{\mathcal{S}}(-\zeta)|\alpha\rangle}$ of this squeezed state. (b) Last step: the change of coordinates $\beta = \cosh(r)\gamma - \sinh(r)e^{i\theta}\gamma^*$ gives $Q^{\zeta}_{|\alpha\rangle}$ from $Q_{\hat{\mathcal{S}}(-\zeta)|\alpha\rangle}$. We plot again $W_{|\alpha\rangle}$ for reference.

2.3 PROPAGATING MODES

Resonant modes are the main systems studied in cavity QED. Controlling and measuring them requires sending them signals and measuring what they emit back. In the domain of superconducting circuits, this is done with transmission lines, generally in the form of coaxial cables or coplanar waveguides. As these signals are non-classical in general, a fully quantum description of the transmission line is needed. The transmission line is a peculiar medium through which the photons

propagate and can be thought of as the reduction to one dimension and one polarization degree of freedom of the free electromagnetic field. It leads to one-dimensional quantum optics, whose results can be generalized to more degrees of freedom. The full quantization procedure described here for the transmission line can be found in [40–42]. We will here highlight the main steps leading to this quantization. We will then present the mode structure of the transmission line, which will be the basis of the analysis done in the next chapters. A very detailed quantum optics review on the modal structure of light can be found in [43].

2.3.1 *Quantization of the transmission line*

To quantize the transmission line, the approach is similar to that of the LC oscillator. We can build a telegraph model for the transmission line by seeing it as the continuum limit of a series of small LC resonators, with inductance per unit length l and capacitance per unit length c. This line is characterized by its characteristic impedance $Z_c = \sqrt{\frac{l}{c}}$ and traveling wave velocity $v = \sqrt{\frac{1}{lc}}$.

Following the infinitesimal model of Fig. 2.9, we define the local flux $\phi(x,t)$ and the charge density $q(x,t)$ relatively to the local current $I(x,t)$ and voltage $V(x,t)$:

$$\begin{aligned} \phi(t) &= \int_{-\infty}^{t} V(x,t')\mathrm{d}t' \\ q(t) &= \int_{-\infty}^{t} \frac{\partial I}{\partial x}(x,t')\mathrm{d}t'. \end{aligned} \tag{2.86}$$

The voltage at a node can be written as the sum of two counter-propagating voltage waves:

$$V(x,t) = V^{\rightarrow}(x,t) + V^{\leftarrow}(x,t). \tag{2.87}$$

The right-and left-propagating voltage waves obey

$$v\partial_t V^{\rightleftarrows}(x,t) \pm \partial_x V^{\rightleftarrows}(x,t) = 0, \tag{2.88}$$

where v is the light velocity in the line. The right- and left propagating voltages are then entirely determined by either all the $V^{\rightleftarrows}(0,t)$ or

$V^{\rightleftarrows}(x,0)$: they propagate at speed v without any dispersion. We can write

$$V^{\rightleftarrows}(x,t) = V^{\rightleftarrows}(x=0, t \mp \frac{x}{v}) = V^{\rightleftarrows}(x \mp vt, t=0) \equiv V^{\rightleftarrows}(\tau^{\mp}) \tag{2.89}$$

where we defined here $\tau^{\mp} = t \mp \frac{x}{v}$ the only degree of freedom of $V^{\rightleftarrows}(x,t)$.

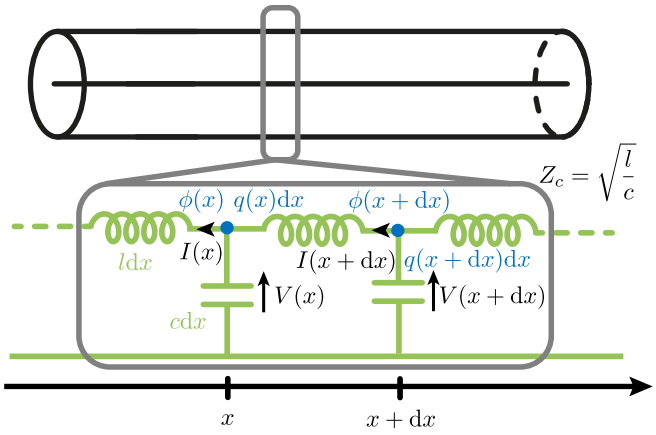


Figure 2.9: Telegraph model for a portion of transmission line.

In the canonical quantization of the transmission line, $q(x,t)$ and $\phi(x,t)$ are promoted to operators verifying the commutation relation

$$\begin{aligned} &\phi(x,t) \to \hat{\phi}(x,t) \\ &q(x,t) \to \hat{q}(x,t) \\ &[\hat{\phi}(x,t), \hat{q}(x',t)] = i\hbar\delta(x-x'). \end{aligned} \tag{2.90}$$

We can now express $V^{\rightleftarrows}$ as operators $\hat{V}^{\rightleftarrows}$, and define the standard creation and annihilation field operators of the transmission line in frequency

$$\hat{a}^{\rightleftarrows}[\omega] = \sqrt{\frac{2}{\hbar|\omega|Z_c}}\hat{V}^{\rightleftarrows}[\omega] \tag{2.91}$$

where $\hat{V}^{\rightleftarrows}[\omega] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\hat{V}^{\rightleftarrows}(\tau)e^{i\omega t}\mathrm{d}\tau$. The only non-zero commutation relations for $\hat{a}^{\rightleftarrows}[\omega]$ are:

$$[\hat{a}^{\rightarrow}[\omega], \hat{a}^{\rightarrow}[\omega']^{\dagger}] = \mathrm{sign}[\omega]\delta(\omega - \omega') \tag{2.92}$$

$$[\hat{a}^{\leftarrow}[\omega], \hat{a}^{\leftarrow}[\omega']^{\dagger}] = \mathrm{sign}[\omega]\delta(\omega - \omega'). \tag{2.93}$$

It verifies $\hat{a}^{\rightarrow}[-\omega] = \hat{a}^{\rightarrow}[\omega]^{\dagger}$, which can be interpreted as the fact that annihilating a photon[2] with negative frequency amounts to creating a real photon with positive frequency. With these operators, we can write the Hamiltonian of the line as tha Hamiltonian of a collection of harmonic oscillators:

$$\hat{H} = \int_0^{+\infty} \frac{\hbar\omega}{2}\Big(\big\{\hat{a}^{\rightarrow}[\omega], \hat{a}^{\rightarrow}[\omega]^{\dagger}\big\} + \big\{\hat{a}^{\leftarrow}[\omega], \hat{a}^{\leftarrow}[\omega]^{\dagger}\big\}\Big)\mathrm{d}\omega. \tag{2.94}$$

The brackets $\{.\}$ here stand for the anti-commutator. The voltage operator as a function of $\hat{a}^{\rightarrow}[\omega]$ reads

$$\hat{V}^{\rightleftarrows}(\tau) = \sqrt{\frac{Z_c\hbar}{4\pi}}\Bigg(\int_0^{+\infty}\sqrt{\omega}\hat{a}^{\rightleftarrows}[\omega]e^{-i\omega\tau}\mathrm{d}\omega + \int_0^{+\infty}\sqrt{\omega}\hat{a}^{\rightleftarrows}[\omega]^{\dagger}e^{i\omega\tau}\mathrm{d}\omega\Bigg). \tag{2.95}$$

We can define the time domain counterpart of $\hat{a}^{\rightarrow}[\omega]$ as

$$\hat{a}^{\rightleftarrows}(\tau) = \frac{1}{\sqrt{2\pi}}\int_0^{+\infty}\hat{a}^{\rightleftarrows}[\omega]e^{-i\omega\tau}\mathrm{d}\omega \tag{2.96}$$

where the integral is taken only on positive frequencies, *i.e.* the domain where $\hat{a}^{\rightleftarrows}[\omega]$ is an annihilation operator: only positive frequency (or energy) excitations can exist in the line. This is the time-domain bosonic annihilation field operator, and its commutation relations read

$$[\hat{a}^{\rightleftarrows}(\tau), \hat{a}^{\rightleftarrows}(\tau')^{\dagger}] = \frac{1}{2}\Big(\frac{1}{i\pi}\mathrm{p.v.}\Big(\frac{1}{\tau - \tau'}\Big) + \delta(\tau - \tau')\Big). \tag{2.97}$$

The bosonic counterpart of the voltage is the real quadrature of $\hat{a}^{\rightleftarrows}(\tau)$ defined as

$$\hat{x}^{\rightleftarrows}(\tau) = \big(\hat{a}^{\rightleftarrows}(\tau) + \hat{a}^{\rightleftarrows}(\tau)^{\dagger}\big). \tag{2.98}$$

2 We loosely call photons the excitations of the line, even if it can also be described as plasmons.

This field is conceptually different from the voltage $\hat{V}^{\rightleftarrows}$. As the latter allows to directly write the Hamiltonian in the time domain

$$\hat{H} = \frac{1}{Z_c}\int_{-\infty}^{+\infty}\left(\hat{V}^{\rightarrow}(\tau)^2 + \hat{V}^{\leftarrow}(\tau)^2\right)\mathrm{d}\tau, \tag{2.99}$$

the one described by $\hat{x}^{\rightleftarrows}$ is linked to the bosonic field, which describes the transmission line in terms of mode occupation. The photon number operator reads

$$\begin{aligned}\hat{N} &= \int_{-\infty}^{+\infty}\left(\hat{a}^{\rightarrow}(\tau)^\dagger\hat{a}^{\rightarrow}(\tau) + \hat{a}^{\leftarrow}(\tau)^\dagger\hat{a}^{\leftarrow}(\tau)\right)\mathrm{d}\tau \\ &= \int_{0}^{+\infty}\left(\hat{a}^{\rightarrow}[\omega]^\dagger\hat{a}^{\rightarrow}[\omega] + \hat{a}^{\leftarrow}[\omega]^\dagger\hat{a}^{\leftarrow}[\omega]\right)\mathrm{d}\omega.\end{aligned} \tag{2.100}$$

We can note here that because of the factor $\hbar\omega$ in the definition of $\hat{a}^{\rightleftarrows}[\omega]$ in Eq. (2.91), states that have finite energy can have an infinite number of photons, in particular close to zero frequency: a constant voltage pulse has a diverging number of photons.

The two fields $\hat{x}^{\rightleftarrows}$ and $\hat{V}^{\rightleftarrows}$ are related in a non-trivial way: we cannot avoid going through a Fourier transform, scale it by ω, and coming back to the time domain to get $\hat{x}(\tau)^{\rightleftarrows}$ from $\hat{V}^{\rightleftarrows}(\tau)$. It means that they are not related causally [44], *i.e.* we theoretically need the values for all times of $\hat{V}^{\rightleftarrows}$ to compute each $\hat{x}(\tau)^{\rightleftarrows}$. This non-causal relation between these two fields leads to paradoxical ascertainments: one cannot both localize the energy and the number of photons in time. As shown in [45], if a microwave pulse has a perfectly localized voltage shape, as it is typically the case in the experiment since they have a finite duration, the bosonic field quadrature $\hat{x}^{\rightleftarrows}$ cannot be localized. In particular, it means that one can theoretically detect photons after the end of a voltage pulse, or even more troubling, *before* it started. This comes from the fact that the notion of position wavefunction for a photon is not well defined in general [46]. The very notion of time-resolved photon detector is thus ill-defined.

These effects are non-negligible only if the bandwidth of the signals is comparable to their average frequency. In most of the experiments performed in quantum optics, the signals are comprised in a bandwidth

$\Delta\omega$ centered around a frequency ω_0 such that $\Delta\omega \ll \omega_0$. In this regime, we can approximate Eq. (2.95) by integrating only inside the bandwidth

$$\begin{aligned}\hat{V}^{\rightleftarrows}(\tau) \simeq \sqrt{\frac{Z_c\hbar\omega_0}{4\pi}}\Bigg(&\int_{\omega_0-\frac{\Delta\omega}{2}}^{\omega_0+\frac{\Delta\omega}{2}} \hat{a}^{\rightleftarrows}[\omega]e^{-i\omega\tau}\mathrm{d}\omega \\ &+ \int_{\omega_0-\frac{\Delta\omega}{2}}^{\omega_0+\frac{\Delta\omega}{2}} \hat{a}^{\rightleftarrows}[\omega]^\dagger e^{i\omega\tau}\mathrm{d}\omega\Bigg).\end{aligned} \tag{2.101}$$

In the context of the Wigner-Weisskopf approximation [40], we can integrate again over the whole set of frequencies (equivalent to taking the bandwidth $\Delta\omega$ go to infinity again) to get

$$\hat{V}^{\rightleftarrows}(\tau) \simeq \sqrt{\frac{\hbar\omega_0 Z_c}{2}}\hat{x}^{\rightleftarrows}(\tau). \tag{2.102}$$

The voltage is now directly proportional to the bosonic quadrature. We have to keep in mind that this is true only in the bandwidth set by $\Delta\omega$. In standard microwave setups, the center frequency is typically defined by the local oscillator and the bandwidth by the mixers, analog-to-digital converter (ADC), and digital-to-analog converter (DAC) or simply filters. The mixers convert the signals from an intermediate frequency to a radio frequency and vice versa, so that the pulses generated and measured fit in the bandwidth of the DAC and ADC. Secs. 4.2.2, 6.10.2 detail a typical microwave setup. In these sections, Figs. 4.5 and 6.9 show typical wiring schemes.

Furthermore, it is often convenient to work in a rotating frame at frequency $\omega_0/2\pi$, which amounts to performing the transformation $\hat{a}^{\rightleftarrows}(\tau) \to \hat{a}^{\rightleftarrows}(\tau)e^{i\omega_0\tau}$ and $\hat{a}^{\rightleftarrows}[\omega] \to \hat{a}^{\rightleftarrows}[\omega+\omega_0]$. In this frame, $\hat{a}^{\rightleftarrows}(\tau)$ reads

$$\begin{aligned}\hat{a}^{\rightleftarrows}(\tau) &= \frac{1}{\sqrt{2\pi}}\int_{-\omega_0}^{+\infty} \hat{a}^{\rightleftarrows}[\omega]e^{-i\omega\tau}\mathrm{d}\omega \\ &\simeq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \hat{a}^{\rightleftarrows}[\omega]e^{-i\omega\tau}\mathrm{d}\omega.\end{aligned} \tag{2.103}$$

It is then common to approximate ω_0 to infinity. In this approximation, the zero frequency is set to ω_0. We thus recover

$$\hat{a}^{\rightleftarrows}[\omega] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \hat{a}^{\rightleftarrows}(\tau)e^{i\omega\tau}\mathrm{d}\tau \tag{2.104}$$

$$\hat{a}^{\rightleftarrows}(\tau) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \hat{a}^{\rightleftarrows}[\omega]e^{-i\omega\tau}\mathrm{d}\omega. \tag{2.105}$$

$\hat{a}^{\rightleftarrows}(\tau)$ and $\hat{a}^{\rightleftarrows}[\omega]$ are thus proper Fourier transform of each other and have the expected commutation relations

$$[\hat{a}^{\rightleftarrows}[\omega], \hat{a}^{\rightleftarrows}[\omega']^\dagger] = \delta(\omega - \omega') \tag{2.106}$$
$$[\hat{a}^{\rightleftarrows}(\tau), \hat{a}^{\rightleftarrows}(\tau')^\dagger] = \delta(\tau - \tau'). \tag{2.107}$$

Within this approximation, all frequencies are equivalent, and the symmetry around $\omega = 0$ does not exist anymore. It allows to recover the notion of the wavefunction of a photon [47], which gives a lot of intuition about the way quantum signals propagate in the line.

2.3.2 *Mode structure of the transmission line*

We now stick to only one direction of propagation and note $\hat{a} = \hat{a}^{\rightleftarrows}$. All derivations made here can be applied to either direction of propagation.

We built $\hat{a}(\tau)$ and $\hat{a}[\omega]$ the time and frequency domain annihilation field operator. We can think of $\hat{a}(\tau)^\dagger$ as the operator "creating" a photon at time τ and of $\hat{a}[\omega]^\dagger$ as the operator creating a photon of frequency ω. However, in practice, a photon is created within a certain linewidth: let $f(\tau)$ be an element of $L^2(\mathbb{R}, \mathbb{C})$, the space of complex functions whose modulus square is integrable on $\mathbb{R}$. We can build

$$\hat{A}_f = \int_{-\infty}^{+\infty} f(\tau)^* \hat{a}(\tau) \mathrm{d}\tau = \int_{-\infty}^{+\infty} f[\omega]^* \hat{a}[\omega] \mathrm{d}\omega \tag{2.108}$$

where $f[\omega] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(\tau) e^{i\omega\tau} \mathrm{d}\tau$ is the Fourier transform of f.

If $\int_{-\infty}^{+\infty} |f(\tau)|^2 \mathrm{d}\tau = 1$, then $\hat{A}_f$ is a standard bosonic operator:

$$[\hat{A}_f, \hat{A}_f^\dagger] = \mathbb{1} \tag{2.109}$$
$$[\hat{A}_f, \hat{A}_f] = 0 \tag{2.110}$$

We can build $\hat{A}_f^\dagger |\mathrm{vac}\rangle$ a single-photon state in the transmission line, and think of $f(\tau)$ as the "wavefunction" of the photon [47]. This works precisely thanks to the narrow-band approximation performed before. In particular, $|f(\tau)|^2$ can be thought of as the probability density of detecting a photon at time τ with a detector placed at $x = 0$, or conversely, the probability of detecting a photon at $x = v\tau$ at $t = 0$, which gives an intuition of what a propagating mode is. However, this

concept has to be manipulated with care, as it is hard to extend it to more than one photon [40].

The value $f(\tau)$ is then expressed in $\sqrt{\text{Hz}}$ and represents one propagating mode of the transmission line. A multiphoton state with n photons can be built by multiple applications of $\hat{a}^\dagger(\tau)$ on the vacuum state:

$$|\psi\rangle = \iint ... \int f(t_1, t_2, ..., t_n)\hat{a}^\dagger(t_1)\hat{a}^\dagger(t_2)...\hat{a}^\dagger(t_n)\mathrm{d}t_1\mathrm{d}t_2...\mathrm{d}t_n, \tag{2.111}$$

or alternatively in the frequency domain:

$$|\psi\rangle = \iint ... \int f[\omega_1, \omega_2, ..., \omega_n]^*\hat{a}^\dagger[\omega_1]\hat{a}^\dagger[\omega_2]...\hat{a}^\dagger[\omega_n]\mathrm{d}\omega_1\mathrm{d}\omega_2...\mathrm{d}\omega_n, \tag{2.112}$$

where $f[\omega_1, \omega_2, ..., \omega_n]$ is given by the multidimensional Fourier transform of f. Although f could theoretically be any distribution such that $|\psi\rangle$ is a normalized state, we stick here to $f \in L^2(\mathbb{R}^n, \mathbb{C})$, which is enough to capture any state that can be produced experimentally. f can be decomposed on the basis of $L^2(\mathbb{R}^n, \mathbb{C})$ comprised of product elements of an orthonormal basis $\{g_n\}$ of $L^2(\mathbb{R}, \mathbb{C})$:

$$f(t_1, t_2, ..., t_n) = \sum_{(k_1,k_2,...,k_n)\in\mathbb{N}^n} \nu_{(k_1,k_2,...,k_n)} \prod_n g_{k_1}(t_1)g_{k_2}(t_2)...g_{k_n}(t_n). \tag{2.113}$$

Here, $\nu_{(k_1,k_2,...,k_n)}$ are discrete complex coefficients. We can now factorize the integrals to write

$$|\psi\rangle = \left(\sum_{(k_1,k_2,...,k_n)\in\mathbb{N}^n} \nu_{(k_1,k_2,...,k_n)} \prod_{l=1}^{n} \hat{A}^\dagger_{g_{k_l}} \right) |\text{vac}\rangle \,. \tag{2.114}$$

Here, we have that for $u, v \in L^2(\mathbb{R}, \mathbb{C})$, $\hat{A}_u$ and $\hat{A}_v$ verify:

$$\begin{aligned} [\hat{A}_u, \hat{A}_v^\dagger] &= \Big(\int_{-\infty}^{+\infty} u(t)^* v(t)\mathrm{d}t \Big)\mathbb{1} = (u \cdot v)\mathbb{1} \\ [\hat{A}_u, \hat{A}_v] &= 0 \end{aligned} \tag{2.115}$$

Thus, if $f \cdot g = 0$, then $\hat{A}_f$ and $\hat{A}_g$ correspond to independent propagating modes. In Eq. (2.114), $|\psi\rangle$ is then expressed as a function of the creation operators of n independent modes.

In general, any state $|\Psi\rangle$ is an infinite superposition of such multi-photon states $|\psi_n\rangle$ with n photons each:

$$|\Psi\rangle = \sum_n \Psi_n |\psi_n\rangle \tag{2.116}$$

The study of the modes of the transmission line can thus be reduced to that of a discrete number of modes. Formally, it amounts to writing the Hilbert space $\mathcal{H}$ of the transmission line as

$$\mathcal{H} = \bigotimes_{n=0}^{\infty} \mathcal{H}_n \tag{2.117}$$

where each $\mathcal{H}_n$ is the Hilbert space of the mode represented by g_n. Thus, any state can be written as

$$|\Psi\rangle = \sum_{k_1,k_2,\ldots} \Psi_{k_1,k_2,\ldots} |k_1\rangle \otimes |k_2\rangle \otimes \ldots \tag{2.118}$$

These properties are summarized in Fig. 2.10. As an example, the modes of the transmission line can be decomposed in the time-frequency Hermite-Gauss basis, well known in optics. In the spatial domain, they can be used in quantum metrology for optical source separation [48–50], and allow beating the Rayleigh criterion. In the time-frequency domain, they are a basis used to demonstrate quantum communication protocols [51, 52]. In Fig. 2.10, they are shifted in frequency to illustrate the fact that the signals of interest are centered around a nonzero frequency.

The expansion given Eq. (2.114) depends on the chosen basis and can be tedious to write in a basis that is not adapted. However, given $\hat{\rho}$ a (possibly mixed) state of the transmission line, there exists a way to find the sparsest way to represent the state. This is known in statistics as the Karhunen-Loève expansion, and it uses the first-order correlation function defined as [53, 54]:

$$\mathcal{G}_1(t_1, t_2) = \langle \hat{a}^\dagger(t_1)\hat{a}(t_2)\rangle. \tag{2.119}$$

We can diagonalize this function and sort the eigenvalues in decreasing order:

$$\mathcal{G}_1(t_1, t_2) = \sum_k n_k \nu_k(t_1)^* \nu_k(t_2), \tag{2.120}$$

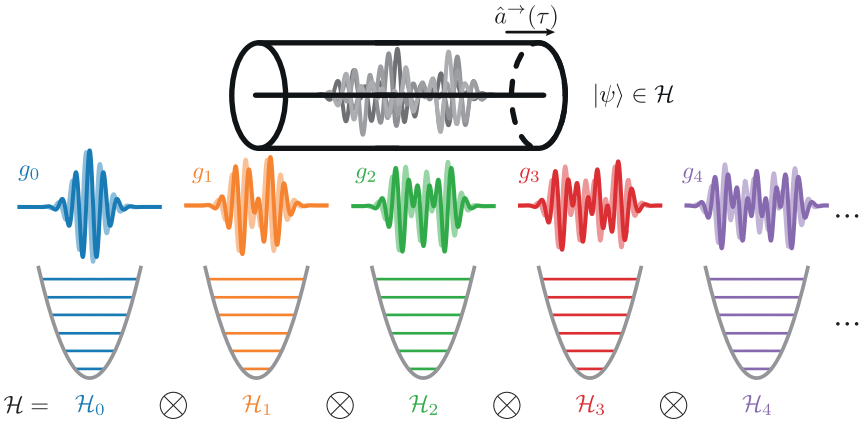


Figure 2.10: Description of the right-propagating transmission line as a discrete collection of harmonic oscillators, each of them describing a propagating mode, represented by g_k, of the transmission line. Here, the g_k are elements of the Hermite-Gauss basis. Real parts are represented by solid lines, and imaginary parts are represented by fainter colors. The Hilbert space of the system $\mathcal{H}$ reads as the tensor product of the Hilbert spaces $\mathcal{H}_k$ of each of these modes.

with $n_0 \geq n_1 \geq n_2 \geq ... \geq 0$ the eigenvalues, and $\{\nu_k\}$ the corresponding eigenmodes. The value n_k is the average number of photons in the mode defined by ν_k. We call a state that verifies $n_{k\geq 1} > 0$ a monomode state. Coherent states and single-photon states are examples of such states. A state that has $n_1 \neq 0$ cannot be expressed as a function of a single mode and is called a multimode state. It comes out of this analysis that the multimode character of the states of the transmission line is not related to a peculiar choice of basis but is rather intrinsic. The light scattered by a non-linear device is generally multimode. A few examples are studied in Sec. 2.4.2.

2.3.3 *Time and frequency operators*

The space of the mode functions is then the space of normalized elements of $L^2(\mathbb{R}, \mathbb{C})$, which is a Hilbert space. Two important operators in this space are the time operator and frequency operator, defined as

$$(\hat{\tau} f)(\tau) = \tau f(\tau) \tag{2.121}$$
$$(\hat{\omega} f)[\omega] = \omega f[\omega]. \tag{2.122}$$

These two operators are the generators of frequency and time translation operators defined as

$$\mathcal{T}_{\Delta\omega} = e^{-i\Delta\omega\hat{\tau}} \tag{2.123}$$
$$\mathcal{T}_{\Delta\tau} = e^{i\hat{\omega}\Delta\tau} \tag{2.124}$$

that act on $L^2(\mathbb{R}, \mathbb{C})$ the following way:

$$(\mathcal{T}_{\Delta\omega} f)[\omega] = f[\omega - \Delta\omega] \tag{2.125}$$
$$(\mathcal{T}_{\Delta\omega} f)(\tau) = f(\tau) e^{-i\Delta\omega\tau} \tag{2.126}$$

and

$$(\mathcal{T}_{\Delta\tau} f)(\tau) = f(\tau - \Delta\tau) \tag{2.127}$$
$$(\mathcal{T}_{\Delta\tau} f)[\omega] = f[\omega] e^{i\omega\Delta\tau}. \tag{2.128}$$

We can also check that $[\hat{\tau}, \hat{\omega}] = \mathbb{1}$. The algebra on the functions of the modes is the same as the standard algebra for the position and momentum of a single (massive) particle. The mode functions can then be seen as wavefunctions for a single photon [47]. Their modulus square $|f|^2(\tau)$ gives the probability of detecting the photon at time τ, and conversely $|f|^2[\omega]$ the probability of detecting it with a frequency ω.

The action of time and frequency operators on the states of the transmission line is described by their corresponding quantum operators $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{\omega}}$ [55]

$$\hat{\boldsymbol{\tau}} = \int_{-\infty}^{+\infty} \tau \hat{a}^\dagger(\tau) \hat{a}(\tau) \mathrm{d}\tau \tag{2.129}$$
$$\hat{\boldsymbol{\omega}} = \int_{-\infty}^{+\infty} \omega \hat{a}^\dagger[\omega] \hat{a}[\omega] \mathrm{d}\omega. \tag{2.130}$$

Applying it on single photon states, one gets the observables associated to the arrival of the photon respectively on a time-resolved and frequency-resolved photon detector. They are also the generators of the frequency and time translation quantum operators

$$\hat{\mathcal{T}}_{\Delta\omega} = e^{-i\Delta\omega\hat{\boldsymbol{\tau}}} \tag{2.131}$$

$$\hat{\mathcal{T}}_{\Delta\tau} = e^{i\hat{\boldsymbol{\omega}}\Delta\tau}. \tag{2.132}$$

The effect of $\hat{\mathcal{T}}_{\Delta\tau}$ on a multiphoton state $|\psi\rangle$ defined as in Eq. (2.111) is such that

$$\begin{aligned}\hat{\mathcal{T}}_{\Delta\tau}|\psi\rangle = \iint ... \int f(t_1 - \Delta\tau, t_2 - \Delta\tau, ..., t_n - \Delta\tau) \\ \times \hat{a}^\dagger(t_1)\hat{a}^\dagger(t_2)...\hat{a}^\dagger(t_n)\mathrm{d}t_1\mathrm{d}t_2...\mathrm{d}t_n\end{aligned} \tag{2.133}$$

which corresponds to the expected effect consisting of translating the wavepacket in time. Conversely, we can see the effect of $\hat{\mathcal{T}}_{\Delta\omega}$ in the frequency domain as

$$\begin{aligned}\hat{\mathcal{T}}_{\Delta\omega}|\psi\rangle = \iint ... \int f[\omega_1 - \Delta\omega, \omega_2 - \Delta\omega, ..., \omega_n - \Delta\omega] \\ \times \hat{a}^\dagger[\omega_1]\hat{a}^\dagger[\omega_2]...\hat{a}^\dagger[\omega_n]\mathrm{d}\omega_1\mathrm{d}\omega_2...\mathrm{d}\omega_n\end{aligned} \tag{2.134}$$

which is the operation corresponding to frequency conversion.

The operators $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{\omega}}$, when restricted to single-photon states, follow the canonical commutation relation $[\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\omega}}] = \mathbb{1}$. Thus, they generate the same algebra as the usual algebra obtained with the standard position and momentum operators. This formalism is developed in the work of Descamp and coworkers [55] to explore the resources offered by single photon propagating modes in quantum metrological tasks. This approach is quite similar to the Wigner function approach that was described in [56] for single electrons in quantum Hall edge channels. It illustrates the fact that single photons can be treated similarly to single electrons (up to their different fermionic and bosonic nature) despite the theoretical difficulties arising when trying to define a proper wavefunction for the photon. Some results obtained in [55] will be used in Sec. 5.4.

2.4 COUPLING STATIONARY AND PROPAGATING MODES

Quantum systems need to be opened to the environment in order both to be manipulated and to be probed. In the context of the Lindblad equation, the environment is traced out, which allows us to recover an average quantum trajectory for the system. However, it gives up the description of the environment, which is necessary to understand what is measured and what back-action is performed on the system. The Quantum Langevin equation offers a very convenient description of the interaction between a high-quality oscillator and a bath. This equation is central in our work, as it allows us to establish that any state prepared in a linear cavity ends up completely transferred into one mode of the transmission line. The framework that we just presented, coupled with the theory of measurements presented in Chap. 3, allows us to fully interpret the experiment presented in Chap. 4.

2.4.1 *Quantum Langevin equations*

The simplest way to open a closed system, such as a qubit or a cavity, is to couple it to a transmission line. Standard coupling schemes are either capacitive or inductive. Fig. 2.11 shows a typical capacitive coupling between a transmission line and a harmonic oscillator. The goal here is to give the main steps leading to the Quantum Langevin Equation, which is an equation in the Heisenberg picture for $\hat{a}$ connecting it to $\hat{a}_{\text{in}}$ and $\hat{a}_{\text{out}}$.

The Hamiltonian of the system {oscillator + transmission line} reads

$$\hat{H} = \hat{H}_{\text{osc}} + \hat{H}_{\text{line}} + \hat{H}_{\text{int}} \tag{2.135}$$

with

$$\hat{H}_{\text{line}} = \int_0^{+\infty} \hbar\omega \big(\hat{a}_{\text{in}}^\dagger[\omega]\hat{a}_{\text{in}}[\omega] + \hat{a}_{\text{out}}^\dagger[\omega]\hat{a}_{\text{out}}[\omega]\big)\mathrm{d}\omega, \tag{2.136}$$

with $\hat{a}_{\text{in}}$ and $\hat{a}_{\text{out}}$ the ingoing and outgoing time annihilation operators of the transmission line. They correspond to the $\hat{a}^{\rightleftarrows}$ of Sec. 2.3.1. The Hamiltonian of the system is left unspecified, as this equation also applies to a wide range of non-linear systems. The interaction is, in general, a linear coupling between one quadrature of the oscillator and

one quadrature of one of each mode of the line (including both field directions "in" and "out"). Up to a phase, we can take as in [40] the example of a coupling $g(\omega)$, between quadratures of $\hat{a} - \hat{a}^\dagger$, $\hat{a}_{\text{in}} + \hat{a}_{\text{in}}^\dagger$ and $\hat{a}_{\text{out}} + \hat{a}_{\text{out}}^\dagger$, that we here take real.

$$\hat{H}_{\text{int}} = i\hbar \int_{-\infty}^{+\infty} g(\omega)(\hat{a} - \hat{a}^\dagger)(\hat{a}_{\text{out}}[\omega] + \hat{a}_{\text{in}}[\omega] + \hat{a}_{\text{out}}^\dagger[\omega] + \hat{a}_{\text{in}}^\dagger[\omega])\mathrm{d}\omega. \tag{2.137}$$

The first step here consists in making a rotating wave approximation. $\hat{H}_{\text{int}}$ becomes

$$\hat{H}_{\text{int}} = i\hbar \int_{-\infty}^{+\infty} g(\omega)\hat{a}(\hat{a}_{\text{out}}^\dagger[\omega] + \hat{a}_{\text{in}}^\dagger[\omega]) - \hat{a}^\dagger(\hat{a}_{\text{out}}[\omega] + \hat{a}_{\text{in}}[\omega])\mathrm{d}\omega. \tag{2.138}$$

The second step is to make another Markov approximation: the coupling $g(\omega)$ is assumed to be constant over the relevant range of frequencies of the system: $g(\omega) = \frac{\sqrt{\kappa}}{2}$. The Heisenberg equation for $\hat{a}$ thus reads

$$\partial_t \hat{a}(t) = \frac{i}{\hbar}[\hat{H}_{\text{sys}}, \hat{a}] - \frac{\sqrt{\kappa}}{2}(\hat{a}_{\text{in}}(t) + \hat{a}_{\text{out}}(t)). \tag{2.139}$$

Boundary conditions at $x = 0$ give:

$$\hat{a}_{\text{out}} - \hat{a}_{\text{in}} = \sqrt{\kappa}\hat{a} \tag{2.140}$$

which gives

$$\partial_t \hat{a}(t) = \frac{i}{\hbar}[\hat{H}_{\text{sys}}, \hat{a}] - \frac{\kappa}{2}\hat{a}(t) - \sqrt{\kappa}\hat{a}_{\text{in}}(t) \tag{2.141}$$

the Quantum Langevin Equation written as a function of the incoming field $\hat{a}_{\text{in}}$. It can alternatively read as a function of the outgoing modes

$$\partial_t \hat{a}(t) = \frac{i}{\hbar}[\hat{H}_{\text{sys}}, \hat{a}] + \frac{\kappa}{2}\hat{a}(t) - \sqrt{\kappa}\hat{a}_{\text{out}}(t). \tag{2.142}$$

In the capacitive case described in Fig. 2.11, the coupling strength κ can be found from a (classical) equivalent circuit in the regime where

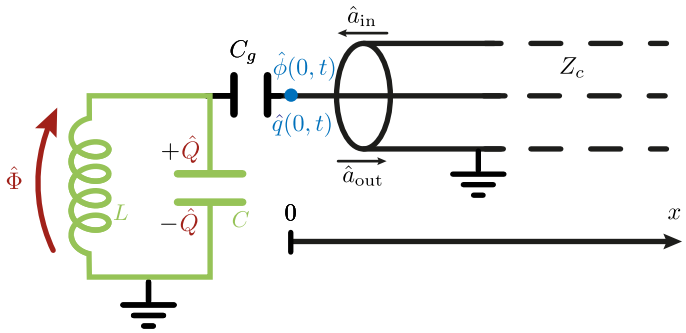


Figure 2.11: Capacitive coupling of a harmonic oscillator to a transmission line.

$Z_c C_g \ll 1$ and $C \gg C_g$, where the capacitance and the transmission line are replaced by a resistor with resistance $R_0 = \frac{1}{Z_c C_g^2 \omega_0^2}$ with $\omega_0 = \frac{1}{LC}$ the frequency of the resonator. In this regime, we can get the coupling strength κ from [42]:

$$\kappa = \frac{1}{R_0 C} = \frac{Z_c C_g^2 \omega_0^2}{C}. \tag{2.143}$$

2.4.2 *Examples*

Let us consider a simple case. At $t = 0$, a state $|\psi\rangle$ is prepared in a linear cavity. As in Eq. (2.9), the Hamiltonian of the system reads

$$\hat{H}_{\text{sys}} = \hbar\omega_0(\hat{a}^\dagger \hat{a} + \frac{\mathbb{1}}{2}). \tag{2.144}$$

For $t > 0$, the system is left free to decay. The Quantum Langevin Equation gives

$$\partial_t \hat{a}(t) = -i\omega_0 \hat{a}(t) - \frac{\kappa}{2}\hat{a}(t) - \sqrt{\kappa}\hat{a}_{\text{in}}(t). \tag{2.145}$$

Solving this ordinary differential linear equation for $\hat{a}$ gives

$$\hat{a}(t) = e^{-(\frac{\kappa}{2}+i\omega_0)t}\hat{a}(0) - \sqrt{\kappa}\int_0^t e^{-(\frac{\kappa}{2}+i\omega_0)(t-s)}\hat{a}_{\text{in}}(s)\mathrm{d}s. \tag{2.146}$$

Denoting $\hat{\Lambda}_{\text{in}}(t) = \sqrt{\kappa}\int_0^t e^{(\frac{\kappa}{2}+i\omega_0)s}\hat{a}_{\text{in}}(s)\mathrm{d}s$, $\hat{\Lambda}_{\text{out}}(t) = \sqrt{\kappa}\int_0^t e^{(-\frac{\kappa}{2}+i\omega_0)s}\hat{a}_{\text{out}}(s)\mathrm{d}s$ and plugging this expression into Eq. (2.140) gives

$$\hat{a}_{\text{out}}(t) = \sqrt{\kappa}e^{-(\frac{\kappa}{2}+i\omega_0)t}\left(\hat{a}(0) - \hat{\Lambda}_{\text{in}}(t)\right) + \hat{a}_{\text{in}}(t) \tag{2.147}$$

$$= \sqrt{\kappa}e^{-(\frac{\kappa}{2}+i\omega_0)t}\left(\hat{a}(0) - \hat{\Lambda}_{\text{in}}(t) + \hat{\Lambda}'_{\text{in}}(t)/\kappa\right) \tag{2.148}$$

Demodulating by $e^{(-\frac{\kappa}{2}+i\omega_0)s}$ and integrating gives for $\hat{\Lambda}_{\text{out}}(t)$:

$$\sqrt{\kappa}\int_0^t e^{(-\frac{\kappa}{2}+i\omega_0)s}\hat{a}_{\text{out}}(s)\mathrm{d}s = (1-e^{-\kappa t})\hat{a}(0) + \int_0^t e^{-\kappa s}\left(\hat{\Lambda}'_{\text{in}}(s) - \kappa\hat{\Lambda}_{\text{in}}(s)\right)\mathrm{d}s \tag{2.149}$$

$$\hat{\Lambda}_{\text{out}}(t) = (1-e^{-\kappa t})\hat{a}(0) + \left[e^{-\kappa s}\hat{\Lambda}_{\text{in}}(s)\right]_0^t \tag{2.150}$$

$$\hat{\Lambda}_{\text{out}}(t) = (1-e^{-\kappa t})\hat{a}(0) + e^{-\kappa t}\hat{\Lambda}_{\text{in}}(t) \xrightarrow[t\to\infty]{} \hat{a}(0). \tag{2.151}$$

We can thus define

$$\hat{A} = \sqrt{\kappa}\int_0^\infty e^{(-\frac{\kappa}{2}+i\omega_0)t}\hat{a}_{\text{out}}(t)\mathrm{d}t. \tag{2.152}$$

the bosonic creation operator of a photon in the temporal mode of shape $e^{-\frac{\kappa}{2}t}$ and frequency $\omega_0/2\pi$. Formally, we have $\hat{A} = \hat{A}_f$ as defined in Eq. (2.108) where $f(t) = \sqrt{\kappa}e^{(-\frac{\kappa}{2}-i\omega_0)t}$.

We can either check that all the orthogonal outgoing modes are defined as a function of $\hat{a}_{\text{in}}$ only (see Appendix B), assumed to be in the vacuum state, or invoke the fact that this system conserves the number of photons, as we can see from Eq. (2.138). It means that the state of the cavity has been transferred to the mode defined by f. Let $|\psi(0)\rangle = F^\star_{|\psi\rangle}(\hat{a}^\dagger)|0\rangle$ be the initial state of the cavity. The free decay thus corresponds to creating a state $|\Psi\rangle = F^\star_{|\psi\rangle}(\hat{A}^\dagger)|vac\rangle$ in the transmission line. An illustration of this process is shown in Fig. 2.12.

This state is then monomode. Actually, a linear cavity with dynamical coupling $\kappa(t)$ and frequency $\omega_0(t)$ can emit its initial state in an arbitrary mode of the transmission line [57]. This also works the other way around: one can perfectly "catch" an incoming mode and store it in a linear cavity [58, 59]. Under the right conditions, one can use

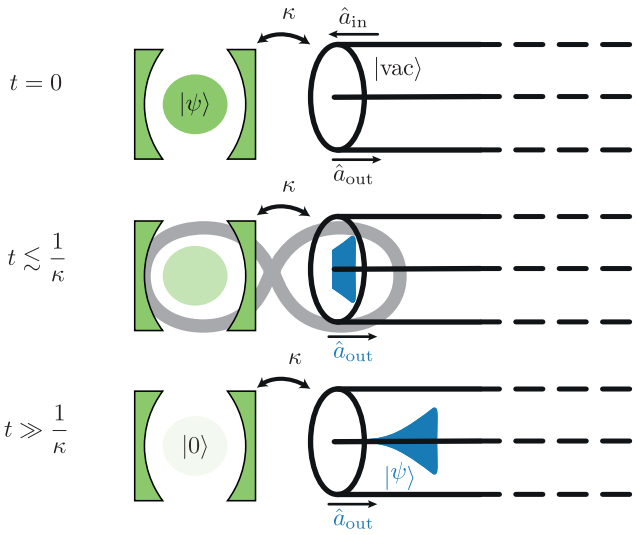


Figure 2.12: Scheme of the release of a state by a linear cavity. At $t = 0$, the state is prepared in the cavity. At short time $t \lesssim \frac{1}{\kappa}$, the state becomes an entangled state between the cavity and a single mode of the line. At $t \gg \frac{1}{\kappa}$, the state has been fully transferred to a mode of the line with envelope $e^{-\frac{\kappa}{2}\tau}$.

a tunable cavity to store the content of an incoming mode into spin ensembles [60–62], which is a way to build- and communicate with- a quantum memory in a quantum network [63, 64].

Reciprocally, if the incoming wavepacket is not monomode, then it is impossible to load the whole content of the wavepacket into such a cavity. A simple example of creation of such a multimode state is given in [53]. They compare the case of a linear cavity whose frequency is chirped in time to that of a non-linear cavity, exhibiting a non-zero self-Kerr coefficient. They show, by diagonalizing the first-order correlation function $\mathcal{G}_1$ defined Eq. (2.119), that the first case is still monomode, while the latter is intrinsically multimode.

Another example of multimode light is the light scattered by a qubit. In [65], Masters and coworkers study the light scattered by a qubit that is coherently driven. In such experimental setups, the light scattered by the qubit has three distinct spectral components (the Mollow triplet [66]) centered around the drive frequency. By removing

the spectral component around the drive frequency, which corresponds to the coherent emission, they managed to experimentally demonstrate that the resulting light consists of pairs of frequency-entangled photons with distinct spectral features. We can write the state $|\psi\rangle$ of such a pair as

$$|\psi\rangle = \iint f[\omega_1, \omega_2]\hat{a}^\dagger[\omega_1]\hat{a}^\dagger[\omega_2]\mathrm{d}\omega_1\mathrm{d}\omega_2\,|\mathrm{vac}\rangle\,. \tag{2.153}$$

Entanglement in frequency then means that we cannot factorize $f[\omega_1, \omega_2]$, which makes it impossible to write $|\psi\rangle$ in the form $|\psi\rangle = \iint f_1[\omega_1]f_2[\omega_2]\hat{a}^\dagger[\omega_1]\hat{a}^\dagger[\omega_2]\mathrm{d}\omega_1\mathrm{d}\omega_2\,|\mathrm{vac}\rangle = \hat{A}^\dagger_{f_1}\hat{A}^\dagger_{f_2}\,|\mathrm{vac}\rangle$. As a single mode two-photon state is necessarily in the form $\hat{A}^{\dagger 2}_{f_0}\,|\mathrm{vac}\rangle$, we can conclude that the state is indeed multimode.

Note that, in general, multimode and entanglement are independent notions. For example, if $\hat{a}$ and $\hat{b}$ represent two independent modes, $\frac{\hat{a}^\dagger+\hat{b}^\dagger}{\sqrt{2}}$ is also a bosonic operator. Applied on the tensor product of the vacuum state for the two modes $|0\rangle \otimes |0\rangle$, it creates a Bell state $|\psi\rangle = \frac{|1\rangle\otimes|0\rangle+|0\rangle\otimes|1\rangle}{\sqrt{2}}$, which is a maximally entangled state in the initial basis. Writing the state in the basis of the two orthogonal modes represented by $\frac{\hat{a}^\dagger+\hat{b}^\dagger}{\sqrt{2}}, \frac{\hat{a}^\dagger-\hat{b}^\dagger}{\sqrt{2}}$, we recover that $|\psi\rangle = |1\rangle \otimes |0\rangle$, with no entanglement. Conversely, multimode light can perfectly be non-entangled in a certain basis. As an example, populating two orthogonal modes with exactly one photon each gives a product state $|\psi\rangle = |1\rangle \otimes |1\rangle$ in the right basis. The notion of *mode-intrisic* or *mode-independent* entanglement, or *passive separability* characterizes states that are entangled in whatever mode basis [67]. Such states are crucial to obtain a quantum advantage in bosonic quantum computation [68]. A more general and detailed discussion about entangled and multimode states of propagating light can be found in [43].

3

QUANTUM MEASUREMENT OF MICROWAVE MODES

Typical quantum measurements in the field of superconducting circuits are essentially built on two key elements: voltmeters and amplifiers. Voltage measurement is the most common measurement performed at room temperature. The typical microwave setup essentially consists of a digital-to-analog converter (DAC) to generate the controls of the system and an analog-to-digital converter (ADC), which is essentially a very fast voltmeter. Mixers and local oscillators are only here for up- and down-conversion of the signals. The ADC, together with quantum amplifiers, allow us to implement quantum homodyne and heterodyne measurements, which are ubiquitous in quantum optics, without the use of photon detectors. The goal of this chapter is to offer a description of the full experimental measurement pipeline used in the field of superconducting circuits and then to apply it to the most standard measurement of the field, which is the dispersive readout of a qubit.

3.1 QUANTUM MEASUREMENT OF SUPERCONDUCTING CIRCUITS

The goal here is to introduce the three main types of measurement used in quantum optics: homodyne and heterodyne detection, and photon detection. We will see that homodyne and heterodyne can be considered as the same measurement performed in different regimes. We will see that the standard measurements performed by room temperature electronics can be viewed as heterodyne measurements, whose outcome statistics are given by the Q function of the modes of the line.

3.1.1 *Homodyne detection*

A homodyne measurement at frequency ω_m of a propagating field is defined as the measurement of

$$\hat{x}(\tau) = e^{i(\omega_m \tau - \varphi)}\hat{a}(\tau) + e^{-i(\omega_m \tau - \varphi)}\hat{a}^\dagger(\tau). \tag{3.1}$$

with φ and arbitrarily chosen phase. Given f_0 a real-valued normalized function of $L^2(\mathbb{R}, \mathbb{C})$, and $f(\tau) = f_0(\tau)e^{-i\omega_0\tau}$, we can define:

$$\hat{x}_{f_0} = \int_{-\infty}^{+\infty} f_0(\tau)\hat{x}(\tau)\mathrm{d}\tau = e^{-i\varphi}\hat{A}_f + e^{i\varphi}\hat{A}_f^\dagger. \tag{3.2}$$

A homodyne measurement then measures the quadrature given by $e^{i\varphi}$ of each of the modes which can be written in the same form as f, *i.e.* that have a hermitian symmetry with respect to ω_0: $f[\omega_0 - \omega] = f[\omega_0 + \omega]^*$. This amounts to measuring all the modes, as we can always find a basis $\{g_n\}$ of $L^2(\mathbb{R}, \mathbb{C})$ that has this property[1]. The $\hat{x}_{g_n}$ then fully characterize the measurement process. A scheme of such decomposition is shown in Fig. 3.2.

As an example, if the line is in the vacuum state, then the outcome x_{g_n} of the $\hat{x}_{g_n}$ are uncorrelated Gaussian variables with the same density of probability p_0 given by $p_0(x) = \frac{2}{\pi}e^{-2x^2}$. The outcome $x(\tau)$ of the measurement of $\hat{x}(\tau)$ is then a white noise, which can be expressed using a Wiener process $\mathrm{d}W$:

$$x(\tau) = \frac{\mathrm{d}W}{\mathrm{d}\tau}. \tag{3.3}$$

In stochastic calculus, we can characterize a Wiener process by its integrals against L^2 functions [69]. In general, the transform $\hat{x} \to \{\hat{x}_{g_n}\}$ converts the continuous stochastic process corresponding to the homodyne measurement into a discrete number of random variables. The

1 This property is the same as the one that allows to decompose any operator on a two-level system on the Pauli basis $\{\mathbb{1}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$ of hermitian operators. We can see it from the fact that any function f can be decomposed as the sum of a hermitian part and an anti-hermitian part: $f = \frac{f+f^*}{2} + \frac{f-f^*}{2}$. An anti-hermitian function is a hermitian function in disguise: $\frac{f-f^*}{2} = \frac{(if)+(if)^*}{2i}$, where $\frac{(if)+(if)^*}{2}$ is a hermitian function. Thus, f can be decomposed as a complex sum of two hermitian functions.

Karhunen-Loève expansion is a particular case of such an expansion, for which the basis $\{g_n\}$ is optimal in some sense [70].

We can then distinguish two cases. The first one is when the modes of interest exhibit a hermitian symmetry in frequency around ω_0. This is typically the case for a signal made of the fluorescence of a linear cavity mode, example given Sec. 2.4.2, if $\omega_m = \omega_0$: the Fourier transform of $f(\tau) = \sqrt{\kappa} e^{-\frac{\kappa}{2}\tau - i\omega_0\tau}$ is $f[\omega] = \frac{\sqrt{\kappa}}{\frac{\kappa}{2} + i(\omega - \omega_0)}$ which is symmetric with respect to ω_0. In this case, setting $\omega_m = \omega_0$ and $f_0(\tau) = e^{-\frac{\kappa}{2}\tau}$ allows to sample the quadrature $e^{-i\varphi}\hat{a} + e^{i\varphi}\hat{a}^\dagger$ of the state prepared in the cavity. This is the usual way homodyne measurement is understood: most of the modes that are probed are symmetric with respect to a frequency ω_0. Choosing $\omega_m = \omega_0$ amounts to performing a homodyne detection of all these modes. This corresponds to the situation described in Fig. 3.1a. The action on these modes is thus that of a projective measurement of $\hat{x}_{f_0}$.

The shape of the populated itinerant modes can be experimentally obtained using a Karhunen-Loève expansion. Given a sufficiently high number of realizations of the experiment giving an ensemble $\{x^k(t)\}$ of time traces, it is possible to build the correlation function $\langle \hat{x}(t_1)\hat{x}(t_2)\rangle$. As for the first-order correlation function $\mathcal{G}_1$ defined in Eq. (2.119), we can diagonalize it by writing

$$\langle \hat{x}(t_1)\hat{x}(t_2)\rangle = \sum_n \lambda_k g_k(t_1) g_k(t_2) \tag{3.4}$$

where $\lambda_0 \geq \lambda_1 ... > 0$ are its eigenvalues sorted in decreasing order, and the $\{g_k\}$ the corresponding eigenvector functions. They are an optimal basis for analyzing the homodyne record [71].

3.1.2 *Heterodyne detection*

The second case is obtained when ω_m is far from ω_0 compared to the bandwidth of the signals of interest: this is the case of heterodyne measurement, pictured in Fig. 3.1b. The resulting signal, which is the outcome of the measurement of $\hat{x}$ is thus oscillating at a frequency $(\omega_0 - \omega_m)/2\pi$, supposed to be large compared to the bandwidth of the signal. In this regime, we can define two quadratures of the signal,

accessible by demodulating the signal at the frequency of the carrier. We can thus demodulate the signal by $f(\tau)e^{i\omega_m\tau}$ which gives

$$\begin{aligned}\hat{x}_f &= \int_{-\infty}^{+\infty} f^*(\tau)e^{-i\omega_m\tau}\hat{x}(\tau)\mathrm{d}\tau \\ &= \int_{-\infty}^{+\infty} f^*(\tau)e^{-i\varphi}\hat{a}(\tau)\mathrm{d}\tau + \int_{-\infty}^{+\infty} f^*(\tau)e^{-2i\omega_m\tau+\varphi}\hat{a}^\dagger(\tau)\mathrm{d}\tau \\ &= e^{-i\varphi}\hat{A}_f + e^{i\varphi}\hat{A}^\dagger_{(f^{-2\omega_m})^*}\end{aligned} \tag{3.5}$$

where $f^{-2\omega}(\tau) = f(\tau)e^{2i\omega\tau}$. If f is sufficiently localized around ω_0 (or if ω_m is sufficiently far from ω_0), then $f \cdot f^{-2\omega_m} \simeq 0$. Heterodyne detection thus implies two modes: the mode of interest plus another independent mode. In this case, this mode is another mode of the line at a different frequency. Assuming that all the modes outside the bandwidth of interest are in the vacuum state, we can derive the moments of $\hat{x}_f$, which allows us to identify the probability distribution p of its outcomes. First, we can take $\varphi = 0$, as it only amounts to a rotation of p, and write

$$\hat{x}_f = \hat{A}_f + \hat{B}^\dagger_f. \tag{3.6}$$

This corresponds to the most general form for $\hat{x}_f$. Indeed, some experimental implementations, and in particular the experiments performed with lasers, implement the heterodyne measurement as two homodyne measurements performed after a beamsplitter operation [72]. In this case, $\hat{B}_f$ is introduced by the beamsplitter and simply represents a mode that comes from the other side of the beamsplitter. In general, this mode is also in the vacuum state.

Second, we can notice that $\hat{x}_f$ is a non-hermitian operator, but which commutes with its hermitian conjugate: $[\hat{x_f}, \hat{x_f}^\dagger] = 0$. We can thus think of $\hat{x}_f$ as being a complex observable, made of two real commuting observables $\mathrm{Re}(\hat{x}_f)$ and $\mathrm{Im}(\hat{x}_f)$. These two observables can be thought of as the two homodyne measurements involved when implementing the heterodyne measurement with a beamsplitter.

As a consequence, when computing its moments M_{nm} as

$$M_{nm} = \langle \hat{x_f}^n \hat{x}_f^{\dagger m}\rangle = \langle (\hat{A}_f + \hat{B}^\dagger_f)^n (\hat{A}_f + \hat{B}^\dagger_f)^{\dagger m}\rangle, \tag{3.7}$$

the ordering of $\hat{x}_f$ and $\hat{x}_f^\dagger$ does not change their value. We choose to develop this expression, where all the terms are in normal order for $\hat{A}_f$ and in anti-normal order for $\hat{B}_f$. Since the mode represented by $\hat{B}_f$ is in the vacuum state, any term involving $\hat{B}_f^{\dagger j}\hat{B}_f^k$ for $j \geq 0$ or $k \geq 0$ vanishes, leaving only

$$\langle \hat{x_f}^n \hat{x_f}^{\dagger m} \rangle = \langle \hat{A}_f^n \hat{A}_f^{\dagger m} \rangle. \tag{3.8}$$

The moments of $\hat{x_f}$ are thus the *anti-normally* ordered moments of $\hat{A}_f$, which is the characteristic of the Q function $Q_{\hat{\rho},\hat{A}_f}$ of the mode represented by $\hat{A}_f$ (see Eq. (2.78)). Thus, $p = Q_{\hat{\rho},\hat{A}_f}$. We note $P_{\hat{\rho},\hat{A}_f}$ (resp. $W_{\hat{\rho},\hat{A}_f}$ and $Q_{\hat{\rho},\hat{A}_f}$) the P distribution (resp. the Wigner function and the Husimi Q function) of the mode represented by the annihilation operator $\hat{A}_f$ in the state $\hat{\rho}$. If the mode of interest is not ambiguous, we stick to the notation used so far: $P_{\hat{\rho}}$, $W_{\hat{\rho}}$, and $Q_{\hat{\rho}}$.

The heterodyne measurement of the modes can be modeled as a Positive Operator-Valued Measure (POVM). Each outcome α is associated with a measurement operator $\hat{M}_\alpha$ such that:

$$\hat{M}_\alpha = \frac{1}{\sqrt{\pi}} |\alpha\rangle\langle\alpha| . \tag{3.9}$$

The closure relation on the coherent states given in Eq. (2.40) gives $\int_{\mathbb{C}} \hat{M}_\beta^\dagger \hat{M}_\beta \mathrm{d}^2\beta = \mathbb{1}$. The heterodyne measurement can thus be thought of as a POVM on the basis of the coherent states. The back-action corresponding to the measurement outcome α is thus to leave the state in the coherent state $|\alpha\rangle$ [73].

As for the homodyne measurement, given $\{g_n\}$ a basis of the modes of the transmission line, the measurement outcomes $x(t)$ are fully characterized by the outcomes x_{g_n} of $\hat{x}_{g_n}$. The joint law of multiple outcomes is given by the corresponding multidimensional Husimi Q function (see Appendix B). A scheme of this demodulation procedure is presented in Fig. 3.2.

The heterodyne measurement record can alternatively be presented as a complex measurement record. Usually, in quantum optics, it comes from the fact that the experimental setup consists in performing two homodyne detections after a beamsplitter operation, which directly gives the real part and the imaginary part of the heterodyne mea-

surement record. With microwave modes, the complex measurement record is obtained as follows. We replace $\hat{x}(t)$ by a discretized version:

$$\hat{x}(t) \to \frac{1}{\delta t}\int_t^{t+\delta t} e^{i(\omega_m-\omega_0)\tau}\hat{x}(\tau)\mathrm{d}\tau. \tag{3.10}$$

which corresponds to the action of a mixer which down-converts a signal at $\omega_0 - \omega_m$ to zero frequency (the homodyne detection already performs a conversion from ω_0 to $\omega_0 - \omega_m$), followed by an ADC with a finite bandwidth set by $\frac{1}{\delta t}$. Under the assumption that the modes of interest are within a bandwidth κ around ω_0, the new demodulated signal $\hat{x}(t)$ keeps all the information available if $\kappa \ll 1/\delta t$. Writing $f_t(\tau) = \frac{1}{\sqrt{\delta t}}e^{-i(\omega_0-\omega_m)\tau}$ for $t < \tau < t + \delta t$, and $f_t(\tau) = 0$ otherwise, the condition $\frac{1}{\delta t} \ll |\omega_0 - \omega_m|$ gives $|f \cdot f^*| \ll 1$, and we can apply Eq. (3.6) to write

$$\hat{x}(t) = \frac{1}{\sqrt{\delta t}}\hat{A}_f(t) + \frac{1}{\sqrt{\delta t}}\hat{B}_f^\dagger(t) \tag{3.11}$$

We then have

$$[\frac{1}{\sqrt{\delta t}}\hat{A}_{f_t}, \frac{1}{\sqrt{\delta t}}\hat{A}^\dagger_{f_{t'}}] = \frac{1}{\delta t}\delta_{t,t'} \underset{\delta t\to 0}{\sim} \delta(t-t') \tag{3.12}$$

$$[\frac{1}{\sqrt{\delta t}}\hat{B}_{f_t}, \frac{1}{\sqrt{\delta t}}\hat{B}^\dagger_{f_{t'}}] = \frac{1}{\delta t}\delta_{t,t'} \underset{\delta t\to 0}{\sim} \delta(t-t') \tag{3.13}$$

$$[\frac{1}{\sqrt{\delta t}}\hat{A}_{f_t}, \frac{1}{\sqrt{\delta t}}\hat{B}^\dagger_{f_{t'}}] = 0 \tag{3.14}$$

Note that $\delta_{t,t'}$ is the Kronecker operator applied to t and t', which is equal to 1 if $t = t'$ and 0 otherwise. We suppose here that time is discretized in time bins of width δt. We can now make the coarse-graining approximation that $\frac{1}{\sqrt{\delta t}}\hat{A}_{f_t}$ approximates well $\hat{a}(t)$ and that $\frac{1}{\sqrt{\delta t}}\hat{B}_{f_t}$ is an approximation of $\hat{b}(t)$ an independent time bosonic operator, as long as $\kappa \ll \frac{1}{\delta t} \ll |\omega_0 - \omega_m|$. Thus we can make the substitution

$$\frac{1}{\sqrt{\delta t}}\hat{A}_{f_t} \to \hat{a}(t) \tag{3.15}$$

$$\frac{1}{\sqrt{\delta t}}\hat{B}_{f_t} \to \hat{b}(t), \tag{3.16}$$

with $[\hat{b}(t), \hat{b}^\dagger(t')] = [\hat{a}(t), \hat{a}^\dagger(t')] = \delta(t-t')$. It gives

$$\hat{x}(t) \simeq \hat{a}(t) + \hat{b}^\dagger(t) \tag{3.17}$$

which is now complex, and is the standard way of writing a heterodyne measurement record.

Using this complex measurement record, we have $\langle \hat{x}(t) \rangle = \langle \hat{a}(t) \rangle$, and the first order correlation function, defined in Eq. (2.119)

$$\begin{aligned} \mathcal{G}_1(t_1, t_2) &= \langle \hat{a}^\dagger(t_1) \hat{a}(t_2) \rangle \\ &= \langle \hat{x}(t_1)^\dagger \hat{x}(t_2) \rangle - \langle \hat{b}(t_1) \hat{b}^\dagger(t_2) \rangle, \end{aligned} \tag{3.18}$$

which was already defined in Sec. 2.3.2. Using that the field described by $\hat{b}$ is in the vacuum state, it finally gives

$$\mathcal{G}_1(t_1, t_2) = \langle \hat{x}(t_1)^\dagger \hat{x}(t_2) \rangle - \delta(t_1 - t_2), \tag{3.19}$$

As explained in Sec. 2.3.2, sorting the eigenvectors of $\mathcal{G}_1$ by decreasing eigenvalues gives a preferential basis to analyze the heterodyne measurement record, as it matches the mode structure of the measured propagating state. In the case where the state of the line is a coherent state, this state can be expressed as a monomode state. The shape of the mode populated is given by $\langle \hat{x}(t) \rangle$. This much simpler way of finding the optimal demodulation function works when the state is monomode, defined by the function f and such that $\langle \hat{A}_f \rangle \neq 0$. This is the approach used in Chap. 4.

3.1.3 *Between homodyne and heterodyne*

It seems that we can continuously go from the homodyne measurement to the heterodyne measurement by shifting ω_m from ω_0 to ω away from ω_0. A natural question is then: What happens in the in-between case? This case is pictured in Fig. 3.1c.

We can start again from Eq. (3.5), and take $\omega_m = 0$ for simplicity, which corresponds to working in the frame rotating at ω_m. Furthermore, we do the replacement $f \to e^{-i\varphi} f$, which corresponds to demodulating with $e^{i\varphi} f^*$ instead of only f^*. In this frame, $\hat{x}_f$ reads

$$\hat{x}_f = \hat{A}_f + e^{2i\varphi} \hat{A}^\dagger_{f^*}. \tag{3.20}$$

The condition for heterodyne measurement simply reads $f \cdot f^* = 0$, which we now assume is not necessarily true (see Fig. 3.1). We can

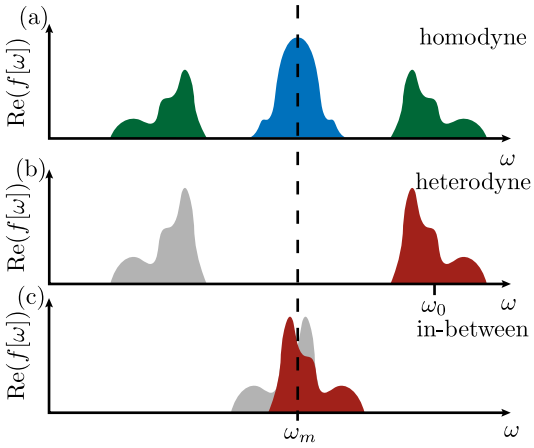


Figure 3.1: The difference between homodyne and heterodyne detection lies in the mode probed. (a) Two examples of modes probed by homodyne detection. For both blue and green modes, the real part of the spectral function $f[\omega]$ is symmetric with respect to ω_m. Their imaginary part is anti-symmetric with respect to ω_m. (b) Example of a mode probed by heterodyne detection. The red mode is the mode of interest and is "mixed" with the symmetric grey mode (generally in the vacuum state) for the measurement. (c) Example of an in-between case where the mode is not symmetric. The red mode (left) is mixed with the grey mode, which has a finite overlap with the red one, giving a "squeezed" heterodyne measurement.

decompose f^* as the sum of a function colinear to f and a function g orthogonal to f:

$$f^* = (f \cdot f^*)f + \sqrt{1 - |f^* \cdot f|^2}g, \tag{3.21}$$

with

$$g = \frac{f^* - (f \cdot f^*)f}{\sqrt{1 - |f^* \cdot f|^2}}. \tag{3.22}$$

Using this decomposition, $\hat{x}_f$ reads

$$\hat{x}_f = \hat{A}_f + e^{2i\varphi}(f \cdot f^*)\hat{A}_f^\dagger + e^{2i\varphi}\sqrt{1 - |f^* \cdot f|^2}\hat{A}_g^\dagger. \tag{3.23}$$

Assuming that the mode corresponding to g is in the vacuum state, we can replace $e^{2i\varphi}\hat{A}_g^\dagger$ by a generic creation operator $\hat{B}^\dagger$ whose mode

is in the vacuum state. We can divide both sides by $\sqrt{1-|f^*\cdot f|^2}$ which gives

$$\frac{\hat{x}_f}{\sqrt{1-|f^*\cdot f|^2}} = \lambda\hat{A}_f + \mu e^{i\varphi_0}\hat{A}_f^\dagger + \hat{B}^\dagger \tag{3.24}$$

where $\lambda = 1/\sqrt{1-|f^*\cdot f|^2}$ and $\mu = e^{2i\varphi}(f\cdot f^*)/\sqrt{1-|f^*\cdot f|^2}$. We can check that $|\lambda|^2 - |\mu|^2 = 1$, which means that we can finally parameterize these factors as $\lambda = \cosh(r)$, $\mu = -\sin(r)e^{i\theta}$, and define $\zeta = re^{i\theta}$:

$$\cosh(r)\hat{x}_f = \cosh(r)\hat{A}_f - \sinh(r)e^{i\theta}\hat{A}_f^\dagger + \hat{B}^\dagger = \hat{A}_f^\zeta + \hat{B}^\dagger, \tag{3.25}$$

where $\hat{A}_f^\zeta = \hat{\mathcal{S}}(-\zeta)\hat{A}_f\hat{\mathcal{S}}(\zeta)$ is the squeezed creation operator of the mode. Thus, up to the last change of coordinates made explicit in Eq. (2.83), the measurement outcomes of the observable $\cosh(r)\hat{x}_f$ are distributed along the law given by $Q^\zeta_{\hat{\rho},\hat{A}_f}$ the generalized Husimi Q function of the mode represented by f. Notably, it corresponds to squeezing the state of the mode represented by f before measuring the standard Husimi Q function. As we will see in Sec. 3.2.3, this corresponds to the action of a phase-sensitive amplifier before a heterodyne measurement.

We recover that the squeezing factor is 0 for $(f\cdot f^*) = 0$, which is the case of heterodyne measurement

The other limit is a bit more tricky, as $\zeta \to \infty$ for $|(f\cdot f^*)| \to 1$, which corresponds to an infinite squeezing. This can be interpreted as one quadrature of the Wigner function getting convolved with a Gaussian function with infinite width, which amounts to an integration on this quadrature. The other quadrature is convolved by a Gaussian function with zero width, which amounts to being perfectly conserved. The statistics of the second quadrature are thus given by the corresponding marginal of the Wigner function, and we recover the homodyne measurement case.

To summarize, as illustrated in Fig. 3.1, this measurement can be viewed alternatively as the homodyne measurement of all the hermitian modes of the line or as the heterodyne measurement of all the modes on one side of ω_m providing that all the corresponding modes on the other side are in the vacuum state. The word homodyne or heterodyne thus depends on whether the modes of interest are in

one regime or the other. The intermediate regime can be viewed as a heterodyne measurement where the target mode has been squeezed, recovering homodyne measurement as an infinitely squeezed heterodyne measurement. An illustration of this is shown in Fig. 3.7.

We can finally describe the demodulation procedure as in Fig. 3.2: Demodulating the signal on a basis $\{g_n\}$ of $L^2(\mathbb{R}, \mathbb{C})$, one gets samples each distributed according to the generalized Husimi Q function of the corresponding mode of the transmission line. If we are well in the heterodyne regime, then $\zeta = 0$, and we sample the Q functions of the modes. The joint law of the x_{g_n} is given by the corresponding multidimensional Husimi Q function (see Appendix B). If the g_n are symmetric in frequency with respect to ω_m ($g_n[\omega_m - \omega] = g_n[\omega_m + \omega]^*$), then $|\zeta| = +\infty$ and we sample one quadrature of each of the modes. The joint law of the outcomes x_{g_n} is given by the corresponding multidimensional Wigner function (see Appendix B).

3.1.4 *Photon detector*

Time-resolved photon detection is obtained by measuring the observable $\hat{a}^\dagger(\tau)\hat{a}(\tau)$, which obeys the eigenvalue relation $\hat{a}^\dagger(\tau)\hat{a}(\tau)\hat{a}^\dagger(\tau')\,|\mathrm{vac}\rangle = \delta_{\tau\tau'}\hat{a}^\dagger(\tau')\,|\mathrm{vac}\rangle$. It corresponds to counting the number of photons in an infinitely narrow region in time. For a physical state of the transmission line, the modes are spread in time. This implies that most of the time, 0 photons are found. As the length of the time interval goes to zero, the probability of finding more than one photon in the same interval vanishes, and we are left with a discrete set of time values for which we detected 1 photon. The outcomes of this operator are thus 1 ("click") or 0 (no "click") for each time τ. We recover a measurement of the total number of photons when integrating it over all times:

$$\hat{N} = \int_{-\infty}^{+\infty} \hat{a}^\dagger(\tau)\hat{a}(\tau)\mathrm{d}\tau. \tag{3.26}$$

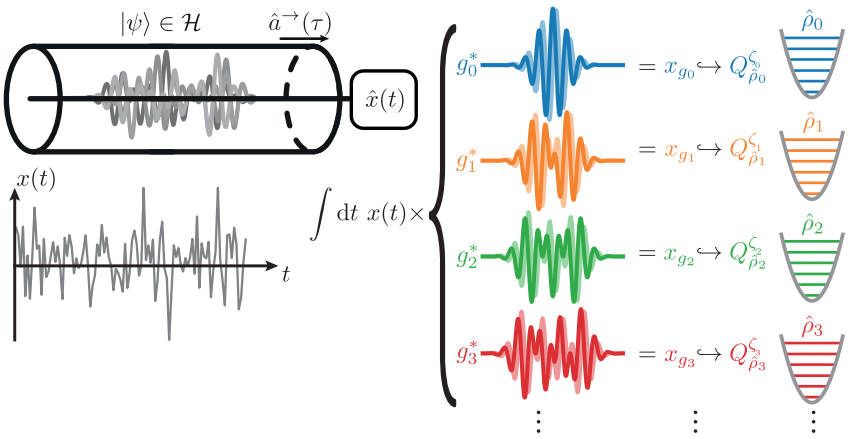


Figure 3.2: Scheme of a demodulation procedure and its outcomes. The temporal outcomes $x(t)$ of a homodyne measurement $\hat{x}(t)$ at frequency $\omega_m/2\pi$ of the right-propagating modes of the line are integrated against the conjugate functions of a basis $\{g_n\}$ of $L^2(\mathbb{R}, \mathbb{C})$, here, the Hermite-Gauss basis with finite carrier frequency. Real parts are represented by solid lines, and imaginary parts are represented by fainter colors. Each resulting outcome x_{g_n} follows the law given by the generalized Husimi Q function of the corresponding mode, with a squeezing factor ζ given by the spectral properties of the functions g_n with respect to ω_m (see Sec. 3.1.3).

We can also build a frequency-resolved photon detector by measuring $\hat{a}^\dagger[\omega]\hat{a}[\omega]$, which obeys the eigenvalue relation $\hat{a}^\dagger[\omega]\hat{a}[\omega]\hat{a}^\dagger[\omega'] \left|\mathrm{vac}\right\rangle = \delta_{\omega\omega'}\hat{a}^\dagger[\omega'] \left|\mathrm{vac}\right\rangle$, and recover similarly

$$\hat{N} = \int_{-\infty}^{+\infty} \hat{a}^\dagger[\omega]\hat{a}[\omega]\mathrm{d}\omega. \tag{3.27}$$

Photon detectors are one of the building blocks of quantum optics above the infrared domain [74]. They can be used to herald propagating non-classical states [75–77]. They are key to applications such as quantum computing [78, 79]. They also allow us to perform more fundamental experiments, such as the demonstration of Bell inequalities violation [80] or the demonstration of the commutation relations of photonic creation and annihilation operators [81].

Photon detectors in the microwave range are more challenging to build owing to the five orders of magnitude lower energy of the photons, which requires working at temperatures in the millikelvin range. Recent progress in superconducting qubit-based photon detectors [82, 83] enabled the detection of individual electron spins [84] and the detection and manipulation of individual nuclear spins [85].

3.1.5 *Voltage measurement*

We will here derive an idealized model of measurement in the field of superconducting circuits. One standard tool in RF measurement modeling is the ideal voltmeter, placed in parallel with a resistor $R = Z_c$ at zero temperature. Crucially, a correct quantum description of the resistor is a semi-infinite transmission line with characteristic impedance given by R (see Fig.3.3). The voltmeter thus measures the observable

$$\hat{V}(\tau) = \hat{V}(\tau)^{\rightarrow} + \hat{V}^{\leftarrow}(\tau) = \sqrt{\frac{\hbar\omega_0 Z_c}{2}}(\hat{a}^{\rightarrow}(\tau) + \hat{a}^{\rightarrow\dagger}(\tau) + \hat{a}^{\leftarrow}(\tau) + \hat{a}^{\leftarrow\dagger}(\tau)), \tag{3.28}$$

which is the sum of the (rightward) propagating voltages coming from the transmission line on the left and the (leftward) propagating voltages coming from the resistor. Typical RF measurements involve demodulating the signal at frequency ω_0 during a certain time t.

$$\int_0^t e^{i\omega_0\tau}\hat{V}(\tau)\mathrm{d}\tau = \int_0^t \sqrt{\frac{\hbar\omega_0 Z_c}{2}} e^{i\omega_0\tau}(\hat{a}^{\rightarrow}(\tau) + \hat{a}^{\rightarrow\dagger}(\tau) + \hat{a}^{\leftarrow}(\tau) + \hat{a}^{\leftarrow\dagger}(\tau))\mathrm{d}\tau. \tag{3.29}$$

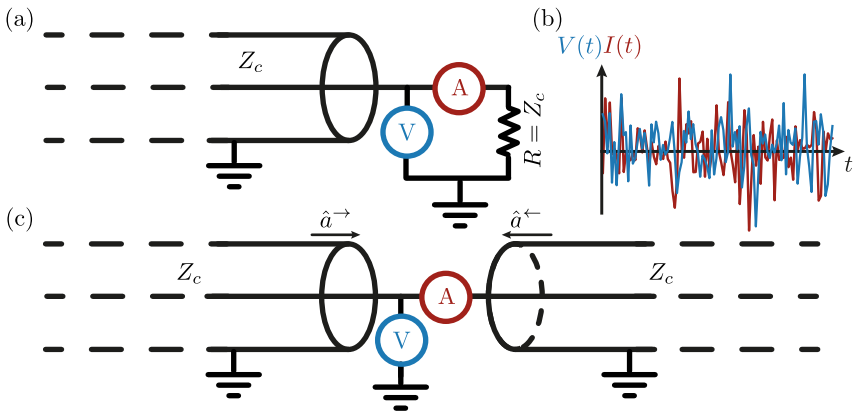


Figure 3.3: Scheme of the measurement of voltage and current measurement of a coaxial transmission line. (a) Ideal voltage and current measurement scheme. (b) Typical records of voltage and current measurement. They can be modeled by continuous stochastic processes. (c) An equivalent circuit obtained from the modeling of the resistor by a semi-infinite transmission line.

As for the heterodyne case, we can demodulate by a complex function $f(\tau)$ which defines an observable

$$\begin{aligned}\hat{V}_f &= \int_{-\infty}^{+\infty} \sqrt{\frac{\hbar\omega_0 Z_c}{2}} f^*(\tau)\big(\hat{a}^{\rightarrow}(\tau) + \hat{a}^{\rightarrow\dagger}(\tau) \\ &\qquad + \hat{a}^{\leftarrow}(\tau) + \hat{a}^{\leftarrow\dagger}(\tau)\big)\mathrm{d}\tau \\ &= \sqrt{\frac{\hbar\omega_0 Z_c}{2}}\big(\hat{A}_f^{\rightarrow} + \hat{A}_{f^*}^{\rightarrow\dagger} + \hat{A}_f^{\leftarrow} + \hat{A}_{f^*}^{\leftarrow\dagger}\big) \\ &= \sqrt{\frac{\hbar\omega_0 Z_c}{2}}\big(\hat{x}_f^{\rightarrow} + \hat{x}_f^{\leftarrow}\big)\end{aligned} \tag{3.30}$$

where $\hat{x}_f^{\rightarrow}$ and $\hat{x}_f^{\leftarrow}$ are defined as in Eq. (3.5) for the right- and left-propagating field for $\varphi = 0$. Here, ω_m is zero, as the voltmeter operates in the laboratory frame.

Under the same approximations made so far, we have $f \cdot f^* \ll 1$. We are then in the regime of heterodyne detection. Thus, assuming that the left-propagating modes are in the vacuum state (the resistor has zero temperature) and that the modes outside the bandwidth of interest in the right-propagating modes are also in the vacuum state,

we get that $\hat{V}_f$ is proportional to the sum of two independent complex observables $\hat{x}_f^{\rightarrow}$ and $\hat{x}_f^{\leftarrow}$ that commute. The outcomes of $\hat{x}_f^{\rightarrow}$ follow the law given by the Q function $Q_{\hat{\rho},\hat{A}_f}$ of the mode of interest, and $\hat{x}_f^{\leftarrow}$ that of the vacuum state $Q_{|0\rangle}$. The probability density p of their sum is thus

$$p = Q_{\hat{\rho},\hat{A}_f} * Q_{|0\rangle}. \tag{3.31}$$

An idealized voltage measurement at zero temperature is thus distributed along the Q function of the incoming modes, plus some noise due to the vacuum fluctuations of the resistor. As we will see later, this corresponds to a heterodyne measurement with quantum efficiency $\eta = 1/2$. One can get rid of this additional noise by adding an ammeter on the other side of the voltmeter. As we can see from [42], it amounts to measuring the current that reads

$$\begin{aligned}\hat{I}(\tau) = \frac{1}{Z_c}\Big(\hat{V}^{\rightarrow}(\tau) - \hat{V}^{\leftarrow}(\tau)\Big) = \sqrt{\frac{\hbar\omega_0}{2Z_c}}(\hat{a}^{\rightarrow}(\tau) + \hat{a}^{\rightarrow\dagger}(\tau) \\ -\hat{a}^{\leftarrow}(\tau) - \hat{a}^{\leftarrow\dagger}(\tau)).\end{aligned} \tag{3.32}$$

We can check that $[\hat{V}_f, \hat{I}_f] = 0$, thus allowing us to measure both simultaneously. Demodulating against f gives:

$$\begin{aligned}\hat{I}_f &= \int_{-\infty}^{+\infty} \sqrt{\frac{\hbar\omega_0}{2Z_c}} f_0(\tau) e^{i\omega_0\tau}\big(\hat{a}^{\rightarrow}(\tau) + \hat{a}^{\rightarrow\dagger}(\tau) \\ &\qquad - \hat{a}^{\leftarrow}(\tau) - \hat{a}^{\leftarrow\dagger}(\tau)\big)\mathrm{d}\tau \\ &= \sqrt{\frac{\hbar\omega_0}{2Z_c}}\big(\hat{A}_f^{\rightarrow} + \hat{A}_{f^*}^{\rightarrow\dagger} - \hat{A}_f^{\leftarrow} - \hat{A}_{f^*}^{\leftarrow\dagger}\big) \\ &= \sqrt{\frac{\hbar\omega_0}{2Z_c}}\big(\hat{x}_f^{\rightarrow} - \hat{x}_f^{\leftarrow}\big).\end{aligned} \tag{3.33}$$

We can then reconstruct $\hat{x}_f^{\rightarrow}$ as

$$\hat{x}_f^{\rightarrow} = \sqrt{\frac{2}{Z_c\hbar\omega_0)}}\hat{V}_f + \sqrt{\frac{2Z_c}{\hbar\omega_0}}\hat{I}_f, \tag{3.34}$$

which describes a heterodyne measurement. Strikingly, the state of the resistor does not matter anymore as we can separately measure right- and left-propagating modes.

Hence, there is no way to perform direct homodyne measurement on the field with a voltmeter, as we are stuck to $\omega_m = 0$. Furthermore, in practice, the detectors themselves are noisy, and the signal is polluted by modes in a thermal state $T \simeq 300$ K on the way to the detector. For $\omega_0/2\pi$ usually between 1 and 10 GHz, it corresponds to convolving with a much broader Q function given by Eq. (2.73). At 5 GHz, the Boltzmann factor ν is equal to $\nu = 1 - 8 \times 10^{-4}$, which corresponds to a Q function with a standard deviation 35 times larger than for the vacuum state. The signal, in general made of a few photons, is thus drowned in thermal fluctuations, which highlights the necessity to amplify the signals before measuring them.

3.2 QUANTUM AMPLIFIERS

We presented homodyne and heterodyne measurements and showed that the use of a perfect voltmeter (ADC) naturally implements a heterodyne measurement on the input modes. In practice, this measurement, performed at room temperature, typically adds a lot of noise compared to the zero-point fluctuations. Furthermore, losses introduced by the lines and microwave components, and the finite temperature of the lines also introduce noise. A solution to mitigate the effect of this noise is to use quantum amplifiers, which we will introduce in the following. The typical microwave measurement chain is then described in Fig. 3.4. The noise added can be modeled by beamsplitters along the line, which mix the input modes with modes that are assumed to be in a thermal state. The noise introduced on the measurement record is then a Gaussian noise. The goal of this chapter is to understand the statistics of the integrated outcomes x'_f obtained by the ADC when using quantum amplifiers, taking into account these noise sources.

Along the measurement chain, the signal is attenuated, mixed with spurious modes, and amplified, which can be cumbersome to take into account in the statistics. In what follows, the goal is to establish the statistics of the measurement outcomes x_f obtained from the experimental measurement outcome x'_f *referred to the input state*. It means rescaling x'_f so that, if the input mode defined by f is in a coherent state $|\alpha\rangle$, the measurement outcome x_f is such that its average

$\overline{x}_f$ is such that $\overline{x}_f = \alpha$. This allows to treat losses, amplification, and thermal noises in a unified way using a single parameter η called the *quantum efficiency*.

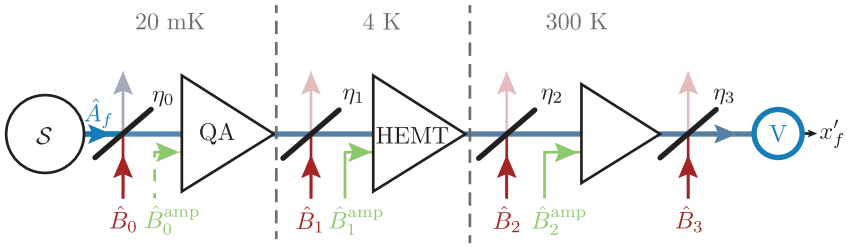


Figure 3.4: Scheme of a microwave detection setup. (a) Typical microwave detection setup comprises a quantum amplifier (QA), followed by an amplifier at the 4 K stage of the fridge (generally a HEMT) and room-temperature amplifiers, before being measured with a voltmeter (ADC), usually after frequency conversion of the signal. Losses on the line are represented by beamsplitters with transparencies η_k and mix the probed modes with modes represented by $\hat{B}_k$, which are in a thermal state. Imperfect amplification also adds noise, represented by the $\hat{B}_k^{\text{amp}}$.

3.2.1 *Beamsplitter*

Losses along the line can be modeled as the action of a beamsplitter. The effect of a beamsplitter with transparency η on the input state is as follows: the input bosonic operator $\hat{A}_f$ is mixed with a mode $\hat{B}$, which gives the mode $\hat{A}'_f$ that reads

$$\hat{A}_f \to \hat{A}'_f = \sqrt{\eta}\hat{A}_f + \sqrt{1-\eta}\hat{B}, \tag{3.35}$$

where $\hat{B}$ is a bosonic annihilation operator of an idler mode. Assuming that this mode represented by $\hat{B}$ is in the vacuum state, the normally ordered moment $\langle \hat{A}'^{\dagger m}_f \hat{A}'^{n}_f \rangle$ of $\hat{A}'_f$ read

$$\langle \hat{A}'^{\dagger m}_f \hat{A}'^{n}_f \rangle = \sqrt{\eta}^{n+m} \langle \hat{A}_f^{\dagger m} \hat{A}_f^{n} \rangle. \tag{3.36}$$

The P distribution $P_{\hat{\rho}',\hat{A}'_f}$ of the mode $\hat{A}'_f$ is thus $P_{\hat{\rho},\hat{A}_f}$ that of the mode $\hat{A}_f$ scaled by $\sqrt{\eta}$:

$$P_{\hat{\rho}',\hat{A}'_f}(\beta) = \frac{1}{\eta} P_{\hat{\rho},\hat{A}_f}\left(\frac{\beta}{\sqrt{\eta}}\right) = P^{\eta}_{\hat{\rho},\hat{A}_f}(\beta). \tag{3.37}$$

The Wigner function $W_{\hat{\rho}',\hat{A}'_f}(\beta)$ of this new mode is obtained by

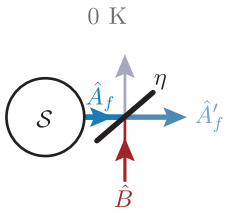


Figure 3.5: Scheme of a beamsplitter. The input mode represented by $\hat{A}_f$ is mixed with a mode in the vacuum state, represented by $\hat{B}$, which forms a new bosonic operator $\hat{A}_f$'.

convolution with the Wigner function of the vacuum state

$$W_{\hat{\rho}',\hat{A}'_f} = P_{\hat{\rho}',\hat{A}'_f} * W_{|0\rangle}. \tag{3.38}$$

In order to refer this Wigner function to the input Wigner function $W_{\hat{\rho},\hat{A}_f}$, we rescale this Wigner function by $\frac{1}{\sqrt{\eta}}$, which we can write

$$W^{1/\eta}_{\hat{\rho}',\hat{A}'_f} = P^{1/\eta}_{\hat{\rho}',\hat{A}'_f} * W^{1/\eta}_{|0\rangle} = P_{\hat{\rho},\hat{A}_f} * W_{|0\rangle} * p_{\mathcal{N}} = W_{\hat{\rho},\hat{A}_f} * p_{\mathcal{N}} \tag{3.39}$$

where $p_{\mathcal{N}}$ is a Gaussian function with variance σ_0^2 such that

$$\sigma_0^2 + \frac{1}{2} = \frac{1}{2\eta} \tag{3.40}$$

obtained using $W^{1/\eta}_{|0\rangle} = W_{|0\rangle} * p_{\mathcal{N}}$, which gives

$$\sigma_0^2 = \frac{1-\eta}{2\eta} \tag{3.41}$$

$$\eta = \frac{1}{1+2\sigma_0^2}. \tag{3.42}$$

Referred to the input mode, a beamsplitter with transparency η at zero temperature thus corresponds to convolving the Wigner function by a Gaussian function of variance $\sigma_0^2 = \frac{1-\eta}{2\eta}$.

3.2.2 *Phase-preserving amplifier*

In general, a noisy measurement with Gaussian noise can be modeled in the following way. The voltmeter at room temperature is giving rescaled integrated outcomes x_f following the probability density p_f such that

$$p_f = Q_{\hat{\rho},\hat{A}_f} * p_{\mathcal{N}}. \tag{3.43}$$

where $p_{\mathcal{N}}$ stands for an arbitrary Gaussian noise density of probability. This noise is set by the microwave components, but it is at least given by the room temperature noise, so that, to get p_f close to $Q_{\hat{\rho},\hat{A}_f}$, one needs to amplify the Q function. As we saw, the Q function of the mode $\hat{A}_f$ gives the probability distribution of the measurement of $\hat{A}_f + \hat{B}^\dagger$ where $\hat{B}$ is an independent mode assumed to be in the vacuum state. A solution would then ideally consist of applying the transformation $\hat{A}_f \to \hat{A}'_f = \sqrt{G}\hat{A}_f$ to amplify the Q function by $\sqrt{G}$. As $[\sqrt{G}\hat{A}_f, \sqrt{G}\hat{A}_f^\dagger] = G$, the commutation relations do not hold anymore for $G \neq 1$. This transform amounts to uniformly scaling the Wigner function in phase space. One way to understand why this is non-physical is to see that any negative regions would be stretched as well. Convolving with the Wigner function of the vacuum state could then give negative regions, which is forbidden for $Q_{\hat{\rho}'} = W_{\hat{\rho}'} * W_0$.

In order to make it work, we need at least to introduce a third bosonic operator $\hat{B}^{\text{amp}}$, and apply the transformation

$$\hat{A}_f \to \hat{A}'_f = \sqrt{G}\hat{A}_f + \sqrt{G-1}\hat{B}^{\text{amp}\dagger}. \tag{3.44}$$

This transformation is realized using two-mode squeezing. The two-mode squeezing operator $\hat{\mathcal{S}}_2$ is defined on two independent modes represented by the annihilation operators $\hat{a}_1$ and $\hat{a}_2$ as

$$\hat{\mathcal{S}}_2(\zeta) = e^{\frac{1}{2}(\zeta^* \hat{a}_1 \hat{a}_2 - \zeta \hat{a}_1^\dagger \hat{a}_2^\dagger)}. \tag{3.45}$$

Its action on $\hat{a}_1$ and $\hat{a}_2$ reads [86]

$$\hat{\mathcal{S}}_2(-\zeta)\hat{a}_1\hat{\mathcal{S}}_2(\zeta) = \cosh(r)\hat{a}_1 - \sinh(r)e^{i\theta}\hat{a}_2^\dagger \tag{3.46}$$

$$\hat{\mathcal{S}}_2(-\zeta)\hat{a}_2\hat{\mathcal{S}}_2(\zeta) = \cosh(r)\hat{a}_2 - \sinh(r)e^{i\theta}\hat{a}_1^\dagger, \tag{3.47}$$

with $\zeta = re^{i\theta}$. Note that we recover the single-mode squeezing operator defined in Eq. (2.35) by setting $\hat{a}_1 = \hat{a}_2$.

Setting $\hat{a}_1 = \hat{A}_f$, $\hat{a}_2 = \hat{B}^{\text{amp}}$, $\theta = \pi$ and $G = \cosh^2(r)$, we obtain Eq. (3.44). Assuming that the mode $\hat{B}^{\text{amp}}$ is in the vacuum state, we get the action of an ideal phase-insensitive (or phase-preserving) amplifier. We can show with similar arguments as those given in Sec. 3.1.2 to get Eq. (3.8) that the anti-normally ordered moments M_{nm} of $\hat{A}'_f$ read

$$\begin{aligned} M_{nm} &= \langle \hat{A}_f'^n \hat{A}_f'^{\dagger m} \rangle \\ &= \langle \left(\sqrt{G}\hat{A}_f + \sqrt{G-1}\hat{B}^{\text{amp}\dagger}\right)^n \left(\sqrt{G}\hat{A}_f + \sqrt{G-1}\hat{B}^{\text{amp}\dagger}\right)^{\dagger m} \rangle \\ &= \sqrt{G^{n+m}} \langle \hat{A}_f^n \hat{A}_f^{\dagger m} \rangle, \end{aligned} \tag{3.48}$$

which corresponds to scaling the Husimi Q function by $\sqrt{G}$. In terms of coordinates of the Q functions, the transformation given by Eq. (3.44) then corresponds to

$$Q_{\hat{\rho}}(\alpha) \to Q_{\hat{\rho}'} = \frac{1}{G} Q_{\hat{\rho}}\left(\frac{\alpha}{\sqrt{G}}\right). \tag{3.49}$$

This is illustrated in Fig 3.6. The Q function $Q_{\hat{\rho}'}$ (dark blue), associated to the Wigner function (red) of a coherent state $|\alpha\rangle = |1+i\rangle$, is simply scaled by a factor $\sqrt{G}$, and gives the law of the measurement records x'_f obtained with the perfect ADC. For G large enough, any added noise on the amplified Q function is negligible (convolving with $p_{\mathcal{N}}$ has almost no effect), and the rescaled measurement outcomes x_f (to refer it to the input mode) follow a law that is simply given by the Husimi Q function of the mode.

For G big enough, the Husimi Q function is then amplified enough to avoid distortion by the noise added after the amplifier, which allows for a faithful heterodyne measurement of the input modes. This is called phase-insensitive or phase-preserving amplifier, as this type of amplification scales the signal irrespectively of its phase and does not modify the average phase. It comes with the cost that this is the Husimi Q function that is perfectly amplified and not the Wigner function. This amounts to adding 1/4 to the variance of each quadrature of the Wigner function before amplification, as is the case for heterodyne

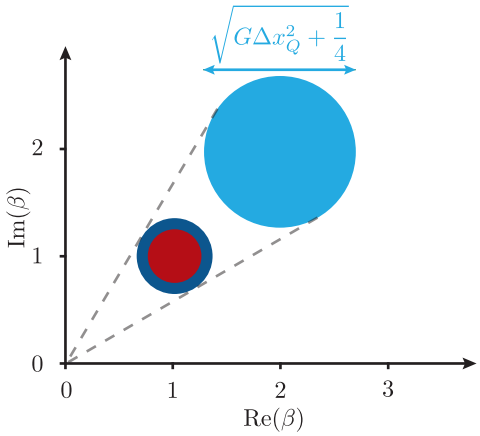


Figure 3.6: Phase-preserving amplification and measurement principle. Red: Wigner function of a coherent state $|\alpha\rangle = |1 + i\rangle$. Dark blue: its Husimi Q function $Q_{|\alpha\rangle}$. Light blue: Husimi Q function of state obtained with a phase-preserving amplifier with gain $G = 4$.

detection compared to homodyne detection. If we are only interested in one of them, using a phase-preserving amplifier amounts to adding some noise at the input of the detector, introduced by the idler mode represented by $\hat{B}^{\text{amp}}$. This can be thought of as a consequence of the Heisenberg inequality: one cannot measure both quadratures of a mode with infinite precision. Amplifying both amounts to choose to measure both of them in some sense, hence coming with additional noise. Here, this noise is minimal: this is known as the quantum limit for phase-preserving amplifiers. Building such amplifiers is key to faithful measurement and feedback control of quantum systems.

Note that the operation performed by the amplifier is unitary and thus can be reversed: it renders the input state more robust to noise when performing heterodyne measurement (by the ADC), but the measurement does not happen at this point. As an example, in [87], Flurin and coworkers use two phase-preserving amplifiers to demonstrate how a phase-preserving amplifier entangles the signal mode and the idler mode via two-mode squeezing. The analysis was done using a second amplifier placed at the two outputs of the first one. In particular, they show a setting allowing us to reverse the action of the

first amplifier with the second one, leaving the signal and idler mode both in the vacuum state after amplification and de-amplification.

The measurement operation happens after the amplifier. The description of the measurement process is still an open question in quantum physics, and it is out of the scope of this work to answer it. In this work, it is assumed to be performed by the room-temperature ADC, modeled as an ideal voltmeter.

3.2.3 *Phase-sensitive amplifier*

The previous amplification scheme alone does not allow for homodyne measurement. As explained in Sec. 3.1.3, obtaining a homodyne measurement from a heterodyne measurement requires single-mode squeezing. As a matter of fact, it corresponds to taking $\hat{B}^{\text{amp}} = \hat{A}_f$ or a rotated version $\hat{B}^{\text{amp}} = e^{-i\theta}\hat{A}_f$ in Eq. (3.44). This gives

$$\hat{A}_f \to \hat{A}'_f = \sqrt{G}\hat{A}_f + \sqrt{G-1}e^{i\theta}\hat{A}_f^\dagger. \tag{3.50}$$

The commutation relations are preserved during this transformation. This corresponds to transforming $\hat{A}_f$ according to the single-mode squeezing operator $\hat{\mathcal{S}}(-\zeta)$ with $\zeta = re^{i\theta} = \log\left(\sqrt{G}\right)e^{i\theta}$ (see Eq. (2.37)). The voltmeter outcome x'_f then follows the Husimi Q function of this scaled operator, which corresponds to the in-between case discussed in Sec. 3.1.3. An example is shown in Fig. 3.7 with a coherent state $|\alpha\rangle = |1+i\rangle$ and $\zeta = \log 2$. From the Wigner function of $|\alpha\rangle$ in red, we see that phase-sensitive amplification enhances one quadrature (x) and lowers the other (p) of the Wigner function $W_{|\alpha\rangle}$to get a squeezed Wigner function $W_{\hat{\mathcal{S}}(-\zeta)|\alpha\rangle}$. This amplification happens *before* convolving with the Wigner function of the vacuum state to obtain the Husimi Q function $Q_{\hat{\mathcal{S}}(-\zeta)|\alpha\rangle}$ (dark green), which gives the law of x'_f. The amplified quadrature thus "survives" the convolution, while the de-amplified one is drowned. To refer the measurement outcome to the input mode, we can rescale the measurement records x'_f to get $x_f = \cosh(r)x'_f - \sinh(r)e^{i\theta}x'^*_f$, which follows the law given

by the generalized Q function $Qs^{\zeta}_{|\alpha\rangle}$ (Eq. (2.83)) whose quadrature standard deviations read

$$\Delta x_{Q^\zeta} = \sqrt{\Delta x_W^2 + \frac{1}{4G}} \tag{3.51}$$

$$\Delta p_{Q^\zeta} = \sqrt{\Delta p_W^2 + \frac{G}{4}}. \tag{3.52}$$

We reach here a similar conclusion to Sec. 2.2.2.4: for sufficiently large G, we get measurement records that follow a law that is very close to the marginal of the Wigner function for the x quadrature and something very noisy with no information on the p quadrature. A large G value also allows us to overcome any additional noise. The effect on the amplified quadrature would be negligible, and the one on the de-amplified one harmless as this quadrature is already almost pure vacuum noise from the convolution with the vacuum Wigner function. Homodyne detection is thus achieved by a large gain phase-sensitive amplifier followed by heterodyne measurement. The name "phase-sensitive" here refers to the fact that incoming signals are either amplified or de-amplified depending on their phase. In the example shown in Fig. 3.7, a coherent state $|\alpha\rangle = |i\rangle$ would be completely de-amplified, whereas $|\alpha\rangle = |1\rangle$ would be entirely amplified.

This amplifying operation is also unitary and thus can be reversed: it only renders one quadrature of the input mode less sensitive to noise when performing heterodyne measurement (by the ADC).

3.2.4 *Quantum efficiency*

In practice, the measurement outcomes can be polluted at any stage of the setup due to the coupling of the elements of the line to unwanted degrees of freedom (see Fig. 3.4). We will take into account here three contributions to this noise. The first one is due to losses, modeled by a beamsplitter of transparency η_k, which couples the line to modes represented by the $\hat{B}_k$. The second one is the finite temperature of the elements in the line, corresponding to having the modes represented by $\hat{B}_k$ in a thermal state. The last one is due to the noise of the various amplifiers on the line, represented by the operators $\hat{B}_k^{\text{amp}}$, which can be in a thermal state. At room temperature, amplification typically

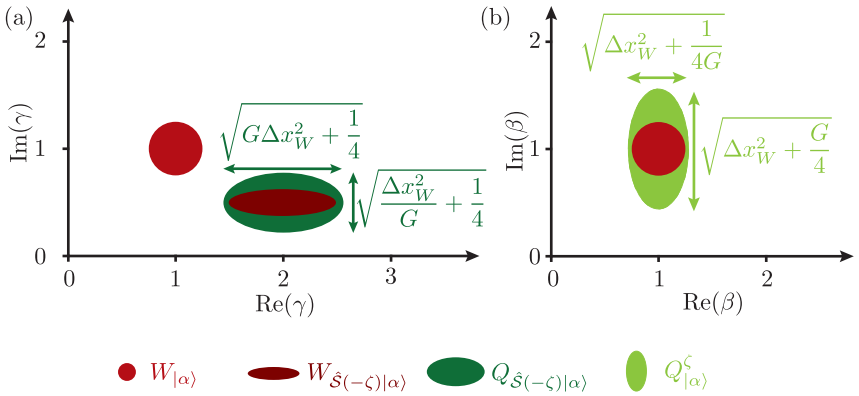


Figure 3.7: Phase-sensitive amplification and measurement principle. (a) Action of a phase-sensitive amplifier. The state is squeezed, which transforms $W_{|\alpha\rangle}$ into $W_{\hat{S}(-\zeta)|\alpha\rangle}$. The corresponding measurement record x'_f are distributed following $Q_{\hat{S}(-\zeta)|\alpha\rangle}$. (b) Rescaling the measurement outcome x'_f following the change of variable $\beta \to \gamma = \cosh(r)\gamma - \sinh(r)e^{i\theta}$ gives $x_f = \cosh(r)x'_f - \sinh(r)e^{i\theta}x_f^{'*}$, whose distribution follows the generalized Husimi Q function $Q^{\zeta}_{|\alpha\rangle}$.

adds much more noise than the quantum limit, and real voltmeters are imperfect. For this reason, the gain of the first amplifier is crucial to neglect all the noise added after it. In practice, a gain of $G = 20$ dB is considered enough when using a cryogenic High Electron Mobility Transistor (HEMT) as a second amplifier.

A lot of other contributions exist ($1/f$ noise, drifts in the setup...), but we identified the most significant ones. They lead to an additional Gaussian noise on the measurement outcomes. The added noise $\hat{\rho}$ can be taken into account by convolving the Q function $Q_{\hat{\rho},\hat{A}_f}$ of the input mode represented by f with a Gaussian probability density $p_{\mathcal{N}}$. The measurement records referred to the input then follow the law given by the equivalent Q function $Q_{\hat{\rho}',\hat{A}'_f}$:

$$Q_{\hat{\rho}',\hat{A}'_f} = Q_{\hat{\rho},\hat{A}_f} * p_{\mathcal{N}}. \tag{3.53}$$

The Q function is obtained by convolving the Wigner function with that of the vacuum state. To compute the effect of noise, it is thus equivalent to apply the effect of noise directly on the Wigner function.

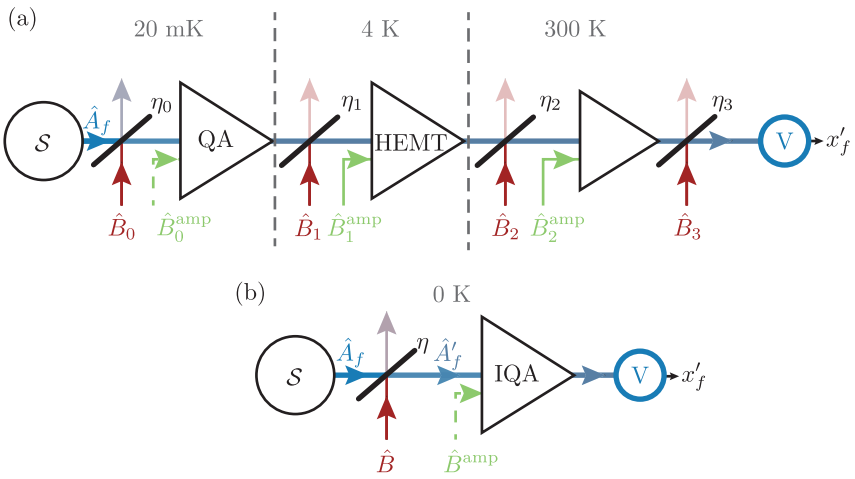


Figure 3.8: Scheme of the equivalent measurement chain characterized by the quantum efficiency. The full measurement chain (a) can be modeled as (b) a single beamsplitter with transparency η before an ideal quantum amplifier (IQA) and voltage measurement. The amplifier can be phase-sensitive or insensitive. In the latter case, it involves the mode represented by $\hat{B}^{\text{amp}}$, which is supposed to be in the vacuum state.

By convolving the Wigner function $W_{\hat{\rho},\hat{A}_f}$ of the mode with $p_{\mathcal{N}}$, it gives the noisy Wigner function $W_{\hat{\rho}',\hat{A}'_f}$ (again, referred to the input)

$$W_{\hat{\rho}',\hat{A}'_f} = W_{\hat{\rho},\hat{A}_f} * p_{\mathcal{N}}. \tag{3.54}$$

The state $\hat{\rho}'$ is a fictional state whose measurement statistics, when measured with ideal homodyne or heterodyne detection, are the same as the measurement statistics of the noisy measurement chain up to a rescaling factor.

In the case where the measurement chain is composed of phase-preserving amplifiers only, we can assume that the added noise is modelled by an uncorrelated Gaussian function of variance σ_0^2, and we write $p_{\mathcal{N}}(\beta) = \frac{1}{2\pi\sigma_0^2} e^{-\frac{|\beta|^2}{2\sigma_0^2}}$. Applying the results of Sec. 3.2.1, the full measurement chain referred to the input state is equivalent to a simplified chained composed of a single beamsplitter with transparency $\eta = \frac{1}{1+2\sigma_0^2}$ followed by a perfect phase-preserving amplifier with an

idler mode in the vacuum state and a perfect voltage measurement. This is summarized in Fig. 3.4b. The transparency η is called the quantum efficiency, and it fully characterizes imperfect measurements based on heterodyne measurement. An alternative characterization is the noise temperature. As the Wigner function of a thermal state is also a Gaussian function, one can write, if the input mode is in the vacuum state ($W_{\hat{\rho},\hat{A}_f} = W_{|0\rangle}$):

$$W_{\hat{\rho}',\hat{A}'_f} = W_{\hat{\rho},\hat{A}_f} * p_{\mathcal{N}} = W_{\hat{\rho}_{\text{th}}} \tag{3.55}$$

where $W_{\hat{\rho}_{\text{th}}}$ is the Wigner function of a mode at temperature T. We can write the condition of equality between the variances

$$\frac{1}{2} + \sigma_0^2 = \frac{1}{2} + n_{\text{th}}. \tag{3.56}$$

Thus, $n_{\text{th}} = \sigma_0^2$. Knowing that $\sigma_0^2 = \frac{1-\eta}{2\eta}$, it gives

$$n_{\text{th}} = \frac{1-\eta}{2\eta} \tag{3.57}$$

$$\eta = \frac{1}{2n_{\text{th}}+1} = \frac{1-\nu}{1+\nu} \tag{3.58}$$

$$\nu = \frac{1-\eta}{1+\eta}, \tag{3.59}$$

where $\nu = e^{-\frac{\hbar\omega}{k_{\text{B}}T}}$ is the Boltzmann factor at the frequency $\omega/2\pi$ of the mode. The quantum efficiency η can be obtained from the statistics of the measurement of a single quadrature in the case of an input coherent state. We can define the signal to noise ratio SNR on the quadrature $\hat{x} = (\hat{a} + \hat{a}^\dagger)/2$ as

$$\text{SNR} = \frac{\langle \hat{x} \rangle^2}{\langle \hat{x}^2 \rangle - \langle \hat{x} \rangle^2}. \tag{3.60}$$

We compute this SNR both for the statistics of the Wigner function of a coherent state $|\alpha\rangle$ and for the statistics of the measurement outcomes when the input state is $|\alpha\rangle$. The quantum efficiency η can then be defined as the ratio between the two. However, for a heterodyne measurement, the measurement statistics are given by the Q function, which is the convolution of the Wigner function by that of the vacuum state. Therefore, the reduction of SNR is at least the same as the one

induced by a beam splitter of transparency 1/2 and connected to a vacuum bath. Effectively, the apparent quantum efficiency is thus at most 1/2. This can alternatively be seen by the fact that a heterodyne measurement can be modeled as two homodyne measurements performed on both sides of a balanced beamsplitter [72]: the measurement of each quadrature looks like it is performed at half the quantum efficiency. Thus, the quantum efficiency is often said to be limited to 1/2 for phase-preserving amplifiers with this alternative definition of η. Consequently, the minimal noise temperature is such that $n_{\mathrm{th}} = 1/2$, which is often referred to as the "half photon of noise" necessarily added by phase preserved amplifiers [40].

When the first amplifier is phase-sensitive, $p_{\mathcal{N}}$ could in principle be asymmetric. However, these amplifiers are typically used when only one signal quadrature is recorded through homodyne detection. We can then carry our analysis with a symmetric $p_{\mathcal{N}}$ whose variance is extrapolated from the SNR of the measurement of this quadrature.

For typical microwave setups, the quantum efficiency of the amplification chain is $\eta \simeq 0.2$, which corresponds to $n_{\mathrm{th}} = 2$ photons. At 5 GHz, it corresponds to a noise temperature of 0.6 K. The records of quantum efficiency for homodyne measurement are up to 80 % [88, 89], and for heterodyne measurement of around 70 % [90], which goes down to 50 % [90, 91] when taking into account the losses before the amplifier. These losses are typically due to isolators that have to be placed between the experiment and the amplifier to prevent any reflection of the amplifier back on the experiment. n the experiment presented in Chap. 4, the quantum efficiency of the amplification chain is $\eta \simeq 0.17$. We attribute this relatively low value compared to the state of the art to the microwave components between the experiment and the TWPA (see Fig. 4.5): two circulators, a diplexer and a directional coupler. Each of these components has insertion losses that contribute to attenuate the signal before it is amplified, thus leading to a lowered overall quantum efficiency. The TWPA itself is expected to limit η to a factor around 0.5 [90][2].

2 Note that this factor is exclusively due to the losses in the TWPA, and has nothing to do with the above discussion on the quantum efficiency of a phase preserving amplifier. If the TWPA had no losses, it could reach unit efficiency in our definition. It just happens to be close to 0.5.

Note that this characterization by the quantum efficiency does not work for a photon detector. The performance of such a measurement needs another figure of merit, which is the dark count rate [82]. We can see it simply, considering two cases, one with a beamsplitter of transparency 1/2, one with a beamsplitter with transparency 3/4, but mixing with an idler mode at thermal equilibrium with $n_{\mathrm{th}} = 1$ photon. Given the previous results, homodyne and heterodyne detection have the same properties (up to a scaling factor), hence the characterization by a quantum efficiency $\eta = 1/2$. However, in the second case, the beamsplitter introduces thermal photons in the line, which causes parasitic false positive clicks by the photon detector. In this case, the quantum efficiency characterizes the proportion of the signal that is lost, and the dark count the proportion of parasitic thermal noise.

3.3 DISPERSIVE READOUT OF A QUBIT

The dispersive readout of a qubit is a way to read out the state of a qubit in a QND way using a dispersively coupled cavity coupled to a transmission line that we can probe. QND stands for Quantum Non-Demolition, which characterizes measurement schemes that do not destroy the measured system. The basic theory of the dispersive readout of a qubit has been extensively studied, notably in [40, 92, 93]. What we propose here is to analyze it from the point of view of a propagating mode measurement. Indeed, we will see that the cavity is an encoder of the information into the transmission line directly. We will use this simple case to introduce the key notions of dephasing and measurement rates. In particular, we will see that homodyne measurement with the right phase is an optimal measurement when considering *time local* measurements, but can be outperformed by a *time global* measurement. We will then further detail the differences between homodyne and heterodyne measurements in terms of extracted and destroyed information.

3.3.1 *Pointer states*

The dispersive measurement of a qubit can be described as follows. A linear cavity is coupled to a qubit in the dispersive regime. The Hamiltonian of the system reads

$$\hat{H}/\hbar = \frac{\omega_q}{2}\hat{\sigma}_z + \omega_0 \hat{a}^\dagger \hat{a} - \chi \left|e\right\rangle\!\left\langle e\right| \hat{a}^\dagger \hat{a} \tag{3.61}$$

where $\left|e\right\rangle\!\left\langle e\right| = \hat{\sigma}_z + \left|g\right\rangle\!\left\langle g\right| = \frac{\hat{\sigma}_z + \mathbb{1}}{2}$ is the projector on the excited state of the qubit, and $\left|g\right\rangle\!\left\langle g\right|$ on the ground state. Therefore, the cavity has a frequency $\omega_0 = \omega_0/2\pi$ if the qubit is in the ground state and a frequency $\omega_1 = (\omega_0 - \chi)/2\pi$ if the qubit is excited. Connecting the cavity to a transmission line with a rate κ, the response to an incoming pulse changes depending on the state of the qubit, thus allowing us to extract information about that state. The goal of this part is to quantify the information content of the pulse. The quantum Langevin equation for the cavity reads

$$\partial_t \hat{a} = \frac{i}{\hbar}[\hat{H}, \hat{a}] - \frac{\kappa}{2}\hat{a} - \sqrt{\kappa}\hat{a}_{\text{in}}, \tag{3.62}$$

which gives

$$\partial_t \hat{a} = -i(\omega_0 - \chi \left|e\right\rangle\!\left\langle e\right|)\hat{a}(t) - \frac{\kappa}{2}\hat{a}(t) - \sqrt{\kappa}\hat{a}_{\text{in}}(t). \tag{3.63}$$

The pulses sent to the cavity are coherent pulses. Since the cavity is supposed to be linear, the state of the transmission line, as well as that of the cavity, are thus coherent states: a semi-classical model is sufficient. We can solve it by replacing the operator $\hat{a}(t)$ and $\hat{a}_{\text{in}}(t)$ by their eigenvalues $\alpha(t)$ and $\alpha_{\text{in}}(t)$. This gives

$$\partial_t \alpha = -i(\omega_0 - \chi \left|e\right\rangle\!\left\langle e\right|)\alpha(t) - \frac{\kappa}{2}\alpha(t) - \sqrt{\kappa}\alpha_{\text{in}}(t). \tag{3.64}$$

This is a linear ordinary differential equation. At this point, we keep the projector $\left|e\right\rangle\!\left\langle e\right|$ in a slight abuse of the notation, but it should be seen as a scalar taking the value 1 if the qubit is in the excited, and 0 otherwise. Assuming that the cavity is initially in the vacuum state (for infinite negative times), the solution reads

$$\alpha(t) = -\sqrt{\kappa}\int_{-\infty}^{t} e^{-(\frac{\kappa}{2} + i(\omega_0 - \left|e\right\rangle\!\left\langle e\right|\chi))(t-s)}\alpha_{\text{in}}(s)\mathrm{d}s. \tag{3.65}$$

The input-output relations then give

$$\begin{aligned}\alpha_{\text{out}}(t) &= \alpha_{\text{in}}(t) + \sqrt{\kappa}\alpha(t) \\ &= \alpha_{\text{in}}(t) - \kappa \int_{-\infty}^{t} e^{-(\frac{\kappa}{2}+i(\omega_0-|e\rangle\langle e|\chi))(t-s)}\alpha_{\text{in}}(s)\mathrm{d}s.\end{aligned} \tag{3.66}$$

Performing a readout then starts by sending a pulse at frequency ω between the times $t = 0$ and $t = T$. We take $\alpha_{\text{in}}(t) = \alpha_{\text{in}}^0 e^{-i\omega t}$ between $t = 0$ and $t = T$ and $\alpha_{\text{in}}(t) = 0$ elsewhere. This gives

$$\begin{aligned}\alpha_{\text{out}}(t) =& \alpha_{\text{in}}^0 e^{-i\omega t}\big(1 + \frac{\kappa}{-\kappa/2 + i(\omega - \omega_0 + \chi\,|e\rangle\langle e|)}\big) \\ &- \frac{\alpha_{\text{in}}^0 \kappa e^{-(\frac{\kappa}{2}+i(\omega_0-|e\rangle\langle e|\chi))t}}{-\kappa/2 + i(\omega - \omega_0 + \chi\,|e\rangle\langle e|)}.\end{aligned} \tag{3.67}$$

We choose to neglect the transient regime, given by the second term of the equation. This approximation corresponds to the limit of long pulses, compared to the decay time $1/\kappa$ of the cavity. Neglecting the same way the transient occurring after turning off the pulse, we end up with

$$\alpha_{\text{out}}(t) = \alpha_{\text{in}}^0 e^{-i\omega t}\left(1 + \frac{\kappa}{-\kappa/2 + i(\omega - \omega_0 + \chi\,|e\rangle\langle e|)}\right)\mathbb{1}_{[0,T]}(t). \tag{3.68}$$

Defining the annihilation operator $\hat{A}_f$ of the mode defined by f:

$$f(t) = \frac{e^{-i\omega t}}{\sqrt{T}}\mathbb{1}_{[0,T]}(t), \tag{3.69}$$

we see that the readout operation corresponds to populating the mode defined by f with a coherent state whose dimensionless amplitude $\alpha_{\text{out}}^{g/e}$ depends on the qubit state:

$$\alpha_{\text{out}}^{g} = \sqrt{T}\alpha_{\text{in}}^0\big(1 + \frac{\kappa}{-\kappa/2 + i(\omega - \omega_0)}\big) = \sqrt{T}(\alpha_{\text{in}}^0 + \sqrt{\kappa}\alpha_g) \tag{3.70}$$

if the qubit is in the ground state and

$$\alpha_{\text{out}}^{e} = \sqrt{T}\alpha_{\text{in}}^0\big(1 + \frac{\kappa}{-\kappa/2 + i(\omega - \omega_0 + \chi)}\big) = \sqrt{T}(\alpha_{\text{in}}^0 + \sqrt{\kappa}\alpha_e) \tag{3.71}$$

if it is in the excited state. Here, $\alpha_{e/g}$ are the amplitude of the intracavity field in the steady state for each qubit state. We call these states the *pointer states*, as they are the ones that we directly measure by homodyne or heterodyne measurement to determine the state of the qubit. The cavity thus acts as an encoder, directly encoding the state of the qubit into the transmission line.

3.3.2 *Dephasing rate*

The amount of information available depends on how well we can distinguish these pointer states at the end of the pulse, which is given by the absolute value of the overlap of these two states:

$$|\langle \alpha^g_{\text{out}} | \alpha^e_{\text{out}} \rangle| = e^{-\frac{|\alpha^g_{\text{out}} - \alpha^e_{\text{out}}|^2}{2}} \tag{3.72}$$

with

$$\begin{aligned} |\alpha^g_{\text{out}} - \alpha^e_{\text{out}}|^2 &= \kappa T |\alpha_g - \alpha_e|^2 \\ &= \frac{16 T |\alpha^0_{\text{in}}|^2 \kappa^2 \chi^2}{(\kappa^2 + 4(\omega - \omega_0)^2)(\kappa^2 + 4(\omega - \omega_0 + \chi)^2)} \end{aligned} \tag{3.73}$$

This overlap thus follows a simple exponential decay with time.

We now note $|\alpha^g_{\text{out}}(T)\rangle$ and $|\alpha^e_{\text{out}}(T)\rangle$ the pointer states of the outgoing field after an input pulse of duration T. Remember that we neglected the transient, which amounts to considering that the cavity ends up in the vacuum state $|0\rangle$ right after the end of the pulse. We can then factor out its state. Neglecting the AC-Stark shift of the qubit due to the population of the cavity and working in a rotating frame for the qubit, the readout operation thus corresponds to the linear map

$$|g\rangle \otimes |\text{vac}\rangle \to |g\rangle \otimes |\alpha^g_{\text{out}}(T)\rangle \tag{3.74}$$

$$|e\rangle \otimes |\text{vac}\rangle \to |e\rangle \otimes |\alpha^e_{\text{out}}(T)\rangle\,. \tag{3.75}$$

Fig. 3.9 illustrates how the state of the qubit is encoded into the state of the outgoing field.

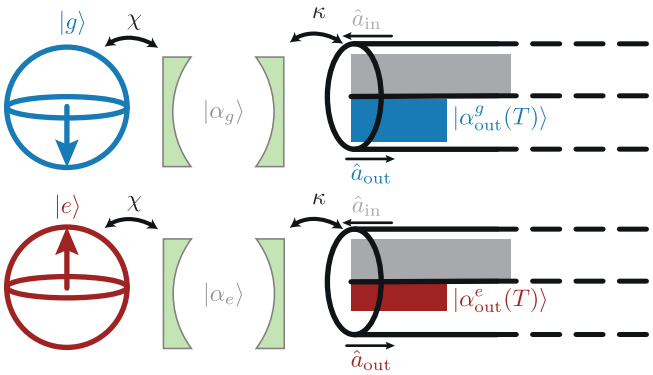


Figure 3.9: Scheme of the dispersive readout principle. The qubit is coupled to a cavity, itself coupled to a transmission line. The cavity is in a coherent state $|\alpha_{g/e}\rangle$ that depends on the qubit state, generating two possible coherent states in the line $\left|\alpha^{g/e}_{\text{out}}\right\rangle$

To illustrate the action of this map, we suppose that the qubit is initially in the $|+X\rangle = \frac{|g\rangle+|e\rangle}{\sqrt{2}}$. The initial state of the system reads

$$|\psi(0)\rangle = \frac{|g\rangle + |e\rangle}{\sqrt{2}} \otimes |\text{vac}\rangle\,, \tag{3.76}$$

and at time T:

$$|\psi(T)\rangle = \frac{1}{\sqrt{2}}\big(\,|g\rangle \otimes |\alpha^g_{\text{out}}(T)\rangle + |e\rangle \otimes |\alpha^e_{\text{out}}(T)\rangle\,\big). \tag{3.77}$$

The state of the qubit is thus entangled with the outgoing field. The density matrix of the qubit in the basis $\{|e\rangle\,, |g\rangle\}$, obtained by tracing out the outgoing field thus reads

$$\begin{aligned}\hat{\rho}(T) &= \frac{1}{2}\begin{pmatrix} 1 & \langle\alpha^g_{\text{out}}(T)|\alpha^e_{\text{out}}(T)\rangle \\ \langle\alpha^e_{\text{out}}(T)|\alpha^g_{\text{out}}(T)\rangle & 1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 & e^{i\phi(T)}e^{-\kappa T\frac{|\alpha_e-\alpha_g|^2}{2}} \\ e^{-i\phi(T)}e^{-\kappa T\frac{|\alpha_e-\alpha_g|^2}{2}} & 1 \end{pmatrix}\end{aligned} \tag{3.78}$$

where $\phi(T)$ is the phase of the scalar product $\langle\alpha^e_{\text{out}}(T)|\alpha^g_{\text{out}}(T)\rangle$. We can now define the dephasing rate as

$$\Gamma_d = \frac{\kappa|\alpha_g - \alpha_e|^2}{2}. \tag{3.79}$$

It is the rate at which the off-diagonal elements of the qubit decay towards zero when the qubit is measured. In the Lindblad equation (2.27), it corresponds to the action of the dissipator $\frac{\Gamma_d}{2}\mathbb{D}_{\hat{\sigma}_z}$. This is a quantity to maximize to read out the state of a qubit as fast as possible before it has time to decay. The easiest way consists in using more power. The quantity $|\alpha_{\text{in}}^0|^2 = \langle \hat{a}_{\text{in}}^\dagger(t)\hat{a}_{\text{in}}(t)\rangle$ is indeed the photon flux of the incoming pulse. However, this cannot be set too high, as populating the readout too much can trigger undesired interactions between the readout cavity and the qubit and lead to a measurement that is not QND. This behavior limits the fidelity of the current best dispersive readouts of transmons [94–96], and is extensively studied to mitigate as much as possible its effects [94–101]. Below this power limit, it has been shown in [102] that for a fixed value of χ and $\overline{n}$ the number of photons inside the resonator (which depends on α_{in}^0, ω and κ), the equation (3.79) is maximized for $\kappa = \chi$ and $\omega = \omega_0 - \chi/2$, which means probing in-between the two possible frequencies of the readout resonator. The regime $\chi \simeq \kappa$ is thus the most efficient and has to be targeted in the design of the experiment.

3.3.3 *Measurement rate*

To understand how we can define the measurement rate, we can first discretize $\hat{A}_f$ by dividing it into n timepieces:

$$\begin{aligned}\hat{A}_f &= \int_0^T \frac{1}{\sqrt{T}} e^{i\omega t}\hat{a}_{\text{out}}(t)\mathrm{d}t = \frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}\int_{\frac{kT}{n}}^{\frac{(k+1)T}{n}} \sqrt{\frac{n}{T}} e^{i\omega t}\hat{a}_{\text{out}}(t)\mathrm{d}t \\ &= \frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}\hat{A}_{f_k},\end{aligned} \tag{3.80}$$

with

$$f_k(t) = \sqrt{\frac{n}{T}} e^{-i\omega t}\mathbb{1}_{[\frac{kT}{n},\frac{(k+1)T}{n}]}. \tag{3.81}$$

Naming $t_k = \frac{kT}{n}$ and $\delta t = \frac{T}{n}$ and taking the limit $n \gg \omega T$, we have

$$\hat{A}_{f_k} \simeq e^{i\omega t_k}\int_{t_k}^{t_{k+1}} \frac{1}{\sqrt{\delta t}}\hat{a}_{\text{out}}(t)\mathrm{d}t \simeq e^{i\omega t_k}\hat{a}_{\text{out}}(t_k)\sqrt{\delta t}. \tag{3.82}$$

First, this gives an intuition about what $\hat{a}_{\text{out}}(t)$ is: we can think of $\hat{a}_{\text{out}}(t)\sqrt{\mathrm{d}t}$ as the annihilation operator of the mode localized between t and $t+\mathrm{d}t$ in the infinitesimal limit. Second, it gives another simple picture of the outgoing state. We can either think of it as a "long" mode hosting a big coherent state or as a collection of "short" (in the sense of their time extent) modes hosting small coherent states. The size of each of the coherent states is indeed directly given by Eq. (3.70) and Eq. (3.71) applied to $T=\delta t$, which gives amplitudes $\alpha^{g/e}_{\text{out,k}} = \alpha^{g/e}_{\text{out}}/\sqrt{N} = \sqrt{\delta t}(\alpha^0_{\text{in}} + \sqrt{\kappa}\alpha_{g/e})$ for each of these modes. The state $\left|\psi^{g/e}\right\rangle_{\text{out}}$ of the transmission line can thus be written

$$\left|\psi^{g/e}\right\rangle_{\text{out}} = \left|\alpha^{g/e}_{\text{out}}\right\rangle_f \otimes |\text{vac}\rangle\,, \tag{3.83}$$

or alternatively

$$\left|\psi^{g/e}\right\rangle_{\text{out}} = \bigotimes_{k=0}^{n-1} \left|\sqrt{\delta t}(\alpha^0_{\text{in}} + \sqrt{\kappa}\alpha_{g/e})\right\rangle_{f_k} \otimes |\text{vac}\rangle\,. \tag{3.84}$$

We choose here to note the vacuum state for all the other modes of the line by the same ket $|\text{vac}\rangle$. This allows us to interpret the situation as follows: the cavity "shoots" a series of small coherent states, streaming a continuous flow of information into the line. The measurement rate is defined as the information rate of this flow:

$$\Gamma_m = \frac{\delta I}{\delta t} \tag{3.85}$$

with δI defined as the mutual information between the state of the qubit and the measurement outcome of one of these modes. The measurement rate is thus strongly dependent on the measurement apparatus. The maximal achievable information rate ${\Gamma_m}^{\max}$ is given by the accessible information δI_{acc} (see Appendix A.1.2) in a time δt. It can be formulated in the context of communications as sending classical binary information (the state $|0\rangle$ or $|1\rangle$ of the qubit) with a qubit (the state of the line in $\text{Span}\left(\left|\alpha^{g/e}_{\text{out}}\right\rangle_f\right)$), as in Appendix A.1.4. Computing δI_{acc} is done by writing the overlap between the two pointer states as

$$\cos\frac{\theta}{2} = |\,\langle\alpha^g_{\text{out}}(\delta t)|\alpha^e_{\text{out}}(\delta t)\rangle\,| = e^{-\Gamma_d \delta t}. \tag{3.86}$$

The angle θ is the angle they form in the Bloch sphere. Expanding it at short times δt gives

$$\cos\frac{\theta}{2} \simeq 1 - \Gamma_d \delta t \tag{3.87}$$

$$\cos^2\frac{\theta}{2} \simeq 1 - 2\Gamma_d \delta t \tag{3.88}$$

$$\sin\frac{\theta}{2} \simeq \sqrt{2\Gamma_d \delta t} \simeq \frac{\theta}{2}. \tag{3.89}$$

Using the results of the Appendix A.1.4, we get

$$\delta I_{\text{acc}} \simeq \frac{\theta^2}{8} = \Gamma_d \delta t. \tag{3.90}$$

Thus, the maximal measurement rate is such that

$${\Gamma_m}^{\max} = \Gamma_d. \tag{3.91}$$

This is very convenient, as it allows for a very simple expression for the quantum efficiency (see 3.3.4), but we have to keep in mind that both quantities have a very different meaning: on the one hand, a dephasing rate, which is the rate at which the overlap between the two pointer states vanishes. On the other hand, an information rate expressed in nats/s.

3.3.4 *Measurement rate: homodyne detection*

Fig. 3.10 illustrates the measurement of the qubit based on homodyne detection. In Fig. 3.10a, the two Wigner functions of the mode of the line corresponding to the two possible states of the qubit are shown. The optimal detection phase, to get the best Signal-to-Noise Ratio (SNR), is materialized by the grey dashed line. The corresponding statistics of measurement outcomes are shown in Fig. 3.10b. The optimal measurement phase φ is such that $d(T) = |\alpha^g_{\text{out}} - \alpha^e_{\text{out}}|(T)$. We define the SNR as in [40]:

$$\text{SNR}(T) = \frac{d(T)^2}{4\sigma^2}. \tag{3.92}$$

In the case of an ideal homodyne measurement, we have $\sigma = 1/2$. Using Eq. (3.73) and Eq. (3.79) gives

$$\text{SNR}(T) = 2\Gamma_d T. \tag{3.93}$$

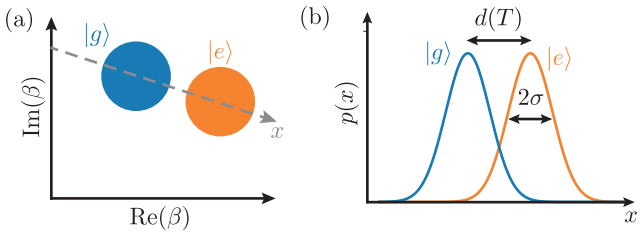


Figure 3.10: Dispersive readout of a qubit in phase space. (a) Scheme of a homodyne measurement applied to the dispersive readout of a qubit. The Wigner functions of the outgoing field conditioned on the state of the qubit are pictured in blue ($|g\rangle$) and orange ($|e\rangle$). The grey dashed line indicates the quadrature x of the homodyne measurement that maximizes the SNR. (b) Corresponding distribution of the outcomes of the homodyne measurement, for g and e.

The SNR thus conveniently linearly increases with the integration time.

We can compute the mutual information obtained by a homodyne detection at short time δt by expanding the formula Eq. (A.2) for the mutual information δI between the state of the qubit and the measurement outcome, in the case where the initial probability of qubit excitation $\mathbb{P}(e) = 1/2$. It reads

$$\delta I = H(x) - \frac{1}{2}(H(x|e) + H(x|g)) \tag{3.94}$$

where x is the measurement outcome, $H(x)$ its Shannon entropy. $H(x|e)$ and $H(x|g)$ are the Shannon entropy of x knowing the state of the qubit. Here, $H(x|e) = H(x|g) = \frac{1}{2}\log 2\pi\sigma^2 + \frac{1}{2}$ is the differential entropy of a Gaussian functions with variance σ^2. $H(x)$ is the entropy of the corresponding mixture of two Gaussian functions. In general, there is no closed-form formula for this quantity. However, we can perform an expansion for short times δt and approximate this mixture by a single Gaussian function with the same variance $\sigma^2 + \frac{d(\delta t)^2}{4}$. $H(x)$ becomes

$$H(x) \simeq \frac{1}{2}\log\left(\sigma^2 + \frac{d(\delta t)^2}{4}\right) + \frac{1}{2}(\log(2\pi) + 1) \tag{3.95}$$

The constant terms then cancel out in the expression of δI, which gives

$$\begin{aligned}\delta I &\simeq \frac{1}{2}\log\left(\sigma^2 + \frac{d(\delta t)^2}{4}\right) - \frac{1}{2}\log(\sigma^2) \\ &\simeq \frac{d(\delta t)^2}{8\sigma^2} \\ &\simeq \frac{\mathrm{SNR}(\delta t)}{2}.\end{aligned} \tag{3.96}$$

By definition (Eq. (3.85)), $\delta I = \Gamma_m \delta t$. For an ideal homodyne measurement, $\sigma = 1/2$, which gives $\Gamma_m = \Gamma_d = {\Gamma_m}^{\max}$. An ideal homodyne measurement thus reaches the maximal measurement rate.

For a finite quantum efficiency η, we established in Sec. 3.2.4 that the Wigner function is convolved with a Gaussian function of standard deviation $\sigma_0 = \sqrt{\frac{1-\eta}{2\eta}}$, which amounts to having $\sigma = \frac{1}{2\sqrt{\eta}}$, and thus

$$\mathrm{SNR}(T) = 2\eta\Gamma_d T. \tag{3.97}$$

The SNR is thus scaled by η. This gives all the ingredients to experimentally measure the quantum efficiency of a homodyne detection performed with the optimal phase: Γ_m is given by the SNR as a function of time, and Γ_d can be measured with Ramsey experiments [103].

3.3.5 *Homodyne detection does not recover all the information*

Until now, we focused on the measurement rate, which is a pertinent quantity to characterize the information obtained from time local measurements, which corresponds to the limit $T \to 0$. In this regard, an ideal homodyne measurement is optimal, as it reaches the maximum obtainable information. However, one can wonder whether this holds for time *global* measurement. The question is: how does a homodyne measurement compare to the optimal measurement performed on the pointer state *after an arbitrary time* T*?* As a matter of fact, a homodyne measurement of duration T performs slightly worse than an ideal measurement performed at time T on the mode. This is shown in Fig. 3.11, where the mutual information obtained with a homodyne measurement is compared to the ideal measurement using Eq. (A.15)

as a function of $|\langle\alpha^e_{\text{out}}(T)|\alpha^g_{\text{out}}(T)\rangle|$. The information ratio between homodyne and ideal goes down to 0.92 for $|\langle\alpha^e_{\text{out}}(T)|\alpha^g_{\text{out}}(T)\rangle| \simeq 0.43$. The limit discussed in Sec. 3.3.4 corresponds to $|\langle\alpha^e_{\text{out}}(T)|\alpha^g_{\text{out}}(T)\rangle| \to 1$, where this ratio goes to 1, as it can be seen in Fig. 3.11.

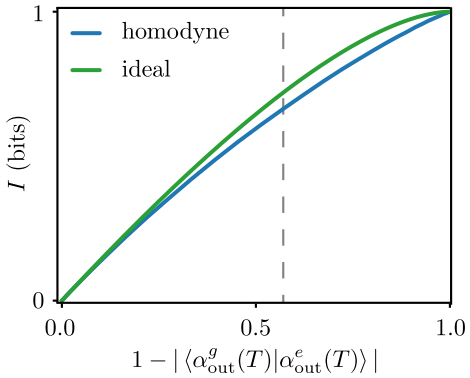


Figure 3.11: Comparison between perfect homodyne measurement and ideal measurement for the dispersive readout of a qubit. We plot the mutual information between the state of the qubit and the measurement outcomes for homodyne (blue) and ideal (green) measurement as a function of the overlap of the pointer states. The minimum ratio between the two mutual information is materialized by the gray dashed line.

It is an example demonstrating the advantage of communication using global measurements over local measurements (see Appendix A.1.3). We write here again the equations Eq. (3.83) and Eq. (3.84):

$$\left|\psi^{g/e}\right\rangle_{\text{out}} = \left|\alpha^{g/e}_{\text{out}}\right\rangle_f \otimes |\text{vac}\rangle \tag{3.98}$$

$$= \bigotimes_{k=0}^{n-1} \left|\sqrt{\delta t}(\alpha^0_{\text{in}} + \sqrt{\kappa}\alpha_{g/e})\right\rangle_{f_k} \otimes |\text{vac}\rangle\,. \tag{3.99}$$

The homodyne detection extracts the best out of the measurement of each local mode represented by f_k, as an ideal measurement on each of them would do. However, a joint measurement performed on all of them manages to extract even more information. This is counterintuitive, as we are dealing with coherent pointer states, which exhibit no entanglement in any basis. The full density matrix $\hat{\rho}_{AB}$

does not exhibit any entanglement either, as it is a mixture of product states.

As to finding how to implement this optimal measurement, it is not easy. In [104], Han and coworkers study various schemes that aim to discriminate between two coherent states. They analyze the information obtained, destroyed, and left in the system for each of them. However, none of these schemes achieve the upper bound. Furthermore, they are implemented on stationary modes, which are easier to handle than propagating ones.

The work [105] of Strandberg and coworkers can help us think about this situation from another point of view. They demonstrate an implementation of a heterodyne detection of a stationary mode using a qubit via stroboscopic SWAP interactions with a qubit followed by a quadrature measurement of this qubit. The intuition behind this is the fact that each of the states $\left|\sqrt{\delta t}(\alpha_{\text{in}}^0 + \sqrt{\kappa}\alpha_{g/e})\right\rangle_{f_k}$ has a vanishing number of photons for $\delta t \to 0$, so each of them can be hosted by a qubit.

Suppose now that, as for a qubit readout, we want to distinguish two possible coherent states in the standing mode. In the experiment, there is only one interacting qubit to perform the qubitdyne detection. We then have no other choice than to measure the qubit after each interaction and reset it for the next one. By measuring the right quadrature of the qubit each time, we recover the optimal homodyne measurement. To recover the time global measurement that would outperform the optimal homodyne measurement, we would need a collection of a large number of qubits: it would then constitute a quantum memory. After each of them interacted with the stationary mode, a global measurement performed on the qubits would allow us to beat the homodyne measurement. In the limit where the field interacted with a large number of qubits, we could recover the ideal global measurement that beats the optimal homodyne detection.

3.3.6 *Using the information: measurement back-action*

We established that information was available in the emitted field of the readout cavity. The goal of this part is to understand how to

make use of this information and to compute the back-action of the measurement on the qubit.

We start first with the optimal homodyne detection for a pure state of the qubit. The qubit is initially in the state $|\psi(0)\rangle_q = a\,|g\rangle + b\,|e\rangle$, with a and b two complex coefficients. The outgoing field represented by f starts in the vacuum state. We thus have, at $t = 0$:

$$|\psi(0)\rangle = |\psi(0)\rangle_q \otimes |0\rangle_f \,. \tag{3.100}$$

After a pulse of duration T, the state $|\psi(T)\rangle$ of the system {qubit, mode f} reads:

$$|\psi(T)\rangle = a\,|g\rangle \otimes |\alpha^g_{\text{out}}\rangle_f + b\,|e\rangle \otimes |\alpha^e_{\text{out}}\rangle_f \,. \tag{3.101}$$

Homodyne measurement with $\varphi = 0$ corresponds to measuring $\hat{x}$. We can write the wavefunction $\Psi_{g/e}(x)$ of $|\alpha^g_{\text{out}}\rangle$ and $|\alpha^e_{\text{out}}\rangle$ as

$$\Psi_{g/e}(x) = \left\langle x \middle| \alpha^{g/e}_{\text{out}} \right\rangle = \left(\frac{2}{\pi}\right)^{\frac{1}{4}} e^{-(x-\alpha^{g/e}_{\text{out}})^2}. \tag{3.102}$$

Using these wavefunctions, $|\psi(T)\rangle$ reads

$$|\psi(T)\rangle = \int_{-\infty}^{+\infty} a\,|g\rangle \otimes \Psi_g(x')\,\big|x'\big\rangle + b\,|e\rangle \otimes \Psi_e(x')\,\big|x'\big\rangle\,\mathrm{d}x'. \tag{3.103}$$

Applying the measurement back-action corresponding to the outcome x obtained at time T on the system gives, just after the measurement at time T^+:

$$\begin{aligned} |\psi(T)\rangle \underset{x}{\to} \big|\psi(T^+)\big\rangle &\propto a\,|g\rangle \otimes \Psi_g(x)\,|x\rangle + b\,|e\rangle \otimes \Psi_e(x)\,|x\rangle \\ &= \big(a\,|g\rangle \otimes \Psi_g(x) + b\,|e\rangle \otimes \Psi_e(x)\big) \otimes |x\rangle \,. \end{aligned} \tag{3.104}$$

The back-action gives a product state. Factoring it out gives the following transform to the qubit between time $t = 0$ and $t = T^+$:

$$|\psi(0)\rangle_q \underset{x}{\to} \big|\psi(T^+)\big\rangle_q \propto \Psi_g(x)a\,|g\rangle + \Psi_e(x)b\,|e\rangle \,. \tag{3.105}$$

It corresponds to a quantum measurement of the qubit defined by the operators $\hat{\Pi}_x$:

$$\hat{\Pi}_x = \frac{\Psi_e(x) + \Psi_g(x)}{2}\mathbb{1} + \frac{\Psi_e(x) - \Psi_g(x)}{2}\hat{\sigma}_z \tag{3.106}$$

with $\int \hat{\Pi}_x \mathrm{d}x = \mathbb{1}$. Applied to the density matrix, it gives

$$\begin{aligned}\hat{\rho}_q \underset{x}{\to} \hat{\rho}_q(T^+) \propto & \left(\frac{\Psi_e(x)+\Psi_g(x)}{2}\right)^2 \hat{\rho}_q \\ & + \left(\frac{\Psi_e(x)-\Psi_g(x)}{2}\right)^2 \hat{\sigma}_z \hat{\rho}_q \hat{\sigma}_z \\ & + \frac{\Psi_e(x)^2-\Psi_g(x)^2}{2}\left(\hat{\rho}_q\hat{\sigma}_z + \hat{\sigma}_z\hat{\rho}_q\right).\end{aligned} \tag{3.107}$$

The Stochastic Master Equation for homodyne measurement is obtained by taking the limit of short times [106]. It allows to reconstruct continuous quantum trajectories of the qubit state [93, 107, 108].

We can discuss here two particular cases. The first is when the homodyne measurement is along the gray dashed line in Fig. 3.10. It corresponds to taking α^g_{out} and α^e_{out} real. In this case, we see from Eq. (3.105) that the phase of the superposition of the qubit is unchanged. The only thing that changes is the excitation probability of the qubit. This corresponds to the ideal homodyne case discussed in Sec. 3.3.4 and maximizes the local information rate about the state of the qubit. It is thus expected that the corresponding back-action exclusively modifies this information.

The other case is when the measurement is performed along the orthogonal direction. It corresponds to having α^g_{out} and α^e_{out} imaginary. In this case, we have $|\Psi_g(x)| = |\Psi_e(x)|$ for all x: measuring $\hat{x}$ does not modify the excitation probability of the qubit, but rather the phase of the superposition. As a consequence, it simply yields more information about the phase of the superposition. In the continuum limit, it makes this phase diffuse at a rate Γ_d [107–112].

We can now understand what happens in a heterodyne measurement: it is a detuned homodyne measurement. It means that the axis of the measurement rotates at the frequency of the detuning. As a consequence, it is as if we spent half of the time measuring in the "right" direction, giving information about the qubit state, and half of the time measuring in the "wrong" direction, making the Bloch vector diffuse on the equator of the Bloch sphere. With this intuitive picture, we feel that the information rate about the qubit excitation is divided by 2. We can then interpret the inefficiency of the heterodyne measurement as being due to an excess of back-action, as half of the time is invested in measuring a quantity that does not bring

any information about the qubit excitation probability. We stress the fact that this conclusion that the heterodyne is inherently inefficient is only true if we are looking for information about the excitation probability of the qubit. As the qubit remains in a pure state after an ideal heterodyne measurement, it still gives a complete knowledge of its full quantum state. The information about the quantum state is preserved.

This illustrates the two ways one can define the quantum efficiency. The first one, with the SNR (Eq. (3.97)), characterizes the speed at which information about the qubit excitation is obtained (the measurement rate) by comparing it to the maximal achievable rate (the dephasing rate) with the same quantum resource (the pointer states): $\eta = \Gamma_m/\Gamma_d$. From this point of view, the quantum efficiency of a heterodyne detection is limited to 1/2, and that of a homodyne detection can vary as a function of the quadrature measured. The second point of view, which was presented in Sec. 3.2.4, characterizes the amount of quantum resource that was lost during the measurement. This characterization thus only characterizes the measurement apparatus and does not even bother about the existence of the qubit, nor does it depend on what we want to learn about the system. In this sense, it is the most general characterization of a homodyne-or heterodyne-based detection setup. In this thesis, quantum efficiencies are given following this point of view.

If we are only interested in the excitation probability of the qubit, we can just look at the probability distribution of its states, which is given by the diagonal elements of the density matrix. It gives

$$\mathbb{P}_0(g) = \langle g| \hat{\rho}_q |g\rangle \underset{x}{\to} \mathbb{P}_{T^+}(g) = \langle g| \hat{\rho}_q(T^+) |g\rangle \propto \mathbb{P}_0(g)\Psi_g(x)^2 \tag{3.108}$$

$$\mathbb{P}_0(e) = \langle e| \hat{\rho}_q |e\rangle \underset{x}{\to} \mathbb{P}_{T^+}(e) = \langle e| \hat{\rho}_q(T^+) |e\rangle \propto \mathbb{P}_0(e)\Psi_e(x)^2 \tag{3.109}$$

It corresponds, up to normalization, to a Bayes actuation of the probability:

$$\mathbb{P}_0(g/e) \underset{x}{\to} \mathbb{P}(x|g/e)\mathbb{P}_0(g/e) \tag{3.110}$$

where $\mathbb{P}(x|g/e)$ is the probability of obtaining x knowing that the qubit was in the state g or e. This formula generalizes to any kind

of measurement record, even noisy ones, and it shows how we can experimentally obtain the update rule of the state probabilities of a qubit: we only need the distribution of outcomes conditioned to each state. More generally, x can be multidimensional, as in the experiment presented in Chap. 4. If we have an experimental characterization of the law of x knowing each possible parameter, then such Bayesian actuation is possible. Note that this approach does not take into account any coherence of the state: this gives a classical stochastic evolution of the information we have about the state of a classical bit.

3.4 CONCLUSION

In this chapter, we modeled the way measurements are performed on the propagating modes in typical experiments involving superconducting circuits. We presented the notions of homodyne and heterodyne detection and linked the integrated measurement outcome statistics to the Wigner, Husimi Q function, and generalized Q function of the propagating modes. We explained the roles of quantum phase-sensitive and phase-insensitive amplifiers in the measurement. We derived their effect on the propagating modes and what it implies for the measurement outcomes obtained with a voltmeter, as in a typical microwave experiment. We introduced the notion of quantum efficiency, which encapsulates all the imperfections of the measurement apparatus.

In a second part, we analyzed the well-known case of the dispersive readout of a qubit in the formalism of the propagating modes and introduced the notions of dephasing and measurement rates. We highlighted that the notion of measurement rate was limited to *time local* measurements. In particular, we showed that a *time global* measurement could outperform the best time local measurements for the same total duration. We then showed how to compute the back-action of a homodyne measurement on the qubit and retrieved the Bayes rule for the actuation of the qubit excitation probability. This Bayes rule is generalized for an arbitrary number of states and an arbitrary noise and will be used in a generalized way in Chap. 4 to compute photon number trajectories.

Part III

PHOTOCOUNTING

4

MONITORING THE PHOTON NUMBER OF A CAVITY

This chapter contains the article about the monitoring of a cavity photon number accepted for publication in PRL [6]. This experiment has to be understood in the context of quantum trajectories and quantum jumps. Quantum trajectories of a qubit under weak dispersive measurement of its excitation [107–111], transverse component [113], or homodyne or heterodyne measurement of its fluorescence [114–116] have been demonstrated in several experimental works. Dispersive measurement and heterodyne detection of the fluorescence of the qubit can even be combined [112]. In this work, we use the fluorescence emitted by a qubit to infer the number of photons in a dispersively coupled cavity. The quantum trajectory of the qubit is not studied, yet probing its fluorescence allows us to observe quantum jumps of the cavity photon number. Quantum jumps, first measured [117–119] in atomic ions and later in superconducting circuits [120], constitute one of the key features of quantum mechanics. Their detection is key in feedback protocols, in particular in quantum error correction.

Quantum jumps are the object of study in numerous experiments. Notably, it was demonstrated in [121] that quantum jumps, often thought to be unpredictable and unreversible, are continuous and could even be reversed in the middle of the process by exploiting the non-hermitian dynamics generated by a well-chosen environment monitoring.

As for quantum jumps of light, they have been experimentally evidenced using Rydberg atoms [122] and superconducting circuits [123], and full photon number trajectories were measured in [124] and used in a feedback protocol to stabilize Fock states [125]. These measurement apparatuses have all in common the fact that they use a dispersively coupled qubit to store the results of stroboscopic measurements of the light field. In [126], Antoine Essig and coworkers experimentally show that it is also possible to measure photon number

trajectories by using the qubit in a fundamentally different way: instead of storing the result of the measurements in its state, it can be used to encode the information about the photon number into the states of a transmission line via the photon number dependent properties of its fluorescence in the dispersive regime. I will here summarize the main results of [126].

Experimentally, the qubit was driven at all the possible frequencies corresponding to every possible photon number, but a more involved theoretical scheme was introduced, involving an infinite frequency comb. This way of understanding the experiment considerably simplified the interpretation of the qubit dynamics and allowed us to understand the experiment as a succession of the same atomic experiment, which consists in repeatedly exciting the qubit using very short pulses and letting the qubit decay in the line. In the number-resolved regime where the dispersive shift χ is much larger than the decay rate $1/T_q$ of the qubit, a Gedanken experiment was developed, making use of an array of diplexers and photon detectors. It was then shown that the characteristic time to measure the number of photons was simply given by T_q. Strikingly, this measurement time does not depend on the maximal number of photons $N_{\max}$. In comparison, a heterodyne measurement scheme was demonstrated to reach a measurement time proportional to $\log(N_{\max})$.

This first version of the experiment [126], despite an achievable information rate theoretically higher than in the second version of the experiment that is presented here, exhibited a lifetime of the cavity too short for a single-shot readout of the photon number. In this version, this issue was solved thanks to the use of a 3D aluminum cavity with a high quality factor and a filter at the output of the qubit to protect the cavity from decaying through it. This allows us to reach here a measurement time that beats the decay rate of the cavity by an order of magnitude, hence providing a single shot photon counter. The main result of this chapter is the photon number trajectories shown in Figs. 4.3, 1.2 and in [127], demonstrating that the quantum jumps corresponding to the cavity losing photons can be monitored.

Sec. 4.1 reproduces the article [6], Secs. 4.2 to 4.8 reproduce the supplementary materials of this article [128]. Sec. 4.9 provides unpublished results showing Wigner functions of Fock states prepared by

post-selection and an analysis of simulated measurement records by a neural network.

4.1 MONITORING THE ENERGY OF A CAVITY BY OBSERVING THE EMISSION OF A REPEATEDLY EXCITED QUBIT

Resolving the number of photons in an electromagnetic mode is at the core of many quantum information protocols [129–132]. Most of them require quantum nondemolition (QND) measurements for measurement based feedback or heralding. In the microwave domain, such measurements can be performed using dispersively coupled Rydberg atoms [122] or superconducting qubits [133]. Predetermined [124, 134] or adaptive [135–139] measurement sequences can monitor the photon number in time and even detect quantum jumps. Each measurement step yields at most a single bit of information about the photon number. By adapting each step, it is possible to reach this upper bound and photocount in a number of cycles that scales logarithmically with the maximal photon number [137–139]. Recently, another qubit-based detector was introduced, which is able to track the photon number using a train of identical qubit pulses forming a frequency comb [126]. Consequently the frequency of the qubit fluorescence encodes the photon number at any time. While a proof-of-principle experiment demonstrated signals proportional to the photon number [126], the measurement rate was insufficient compared to the cavity lifetime for single-shot extraction. Here, we demonstrate photon number tracking in a 3D cavity using a frequency comb driving a dispersively coupled qubit. We experimentally compare the photon number measurement rate of our scheme based on heterodyne detection of the qubit fluorescence to the rate at which the environment could extract information. This QND photon number monitor, with a fixed drive and detection scheme, could simplify feedback schemes and quantum error correction for bosonic codes [140] or qudits [137, 141].

The detection principle can easily be grasped with a classical analogy (Fig. 4.1a). Consider a basket (cavity) filled with apples (photons) that can escape. The number n of apples can be determined by hitting a string (qubit) from which the basket hangs. The string oscillates at a frequency depending on n and recording the emitted sound (heterodyne

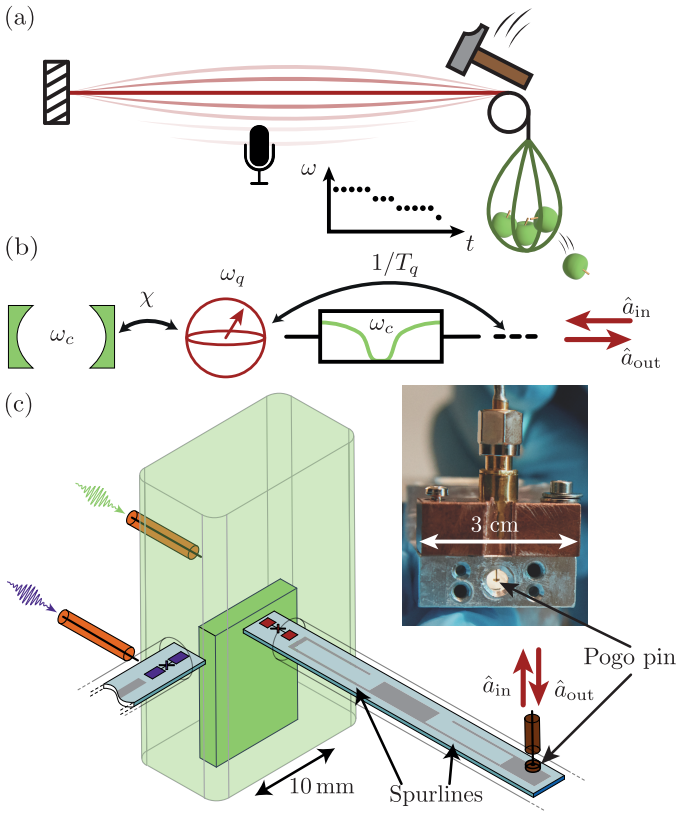


Figure 4.1: Experiment scheme and principle. (a) Acoustic analogy of the experiment. In order to monitor the number of apples in a basket, it is attached to a string. By repeatedly hitting the string, recording the emitted sound reveals the number n at any time. (b) The heterodyne detection of the driven qubit (red) emission into the line leads to the monitoring of the photon number of the dispersively coupled cavity (green). A notch filter centered on ω_c (box) prevents the cavity from decaying through the qubit. (c) The cavity is a high purity aluminum $\lambda/4$ coaxial resonator. Two transmon qubits on sapphire chips stick into the cavity. The filter on the (red) qubit drive line is composed of two tantalum spurlines and galvanically connected to a transmission line using a pogo pin (photograph in inset). Wigner tomography of the cavity is performed using the auxiliary transmon qubit (purple) and its dedicated readout resonator (not shown).

measurement of fluorescence signal) reveals the apple number. Hitting

repeatedly the string leads to the monitoring of n: in the frequency domain, it corresponds to driving the string with a comb.

Experimentally, the basket is an aluminum $\lambda/4$ coaxial cavity [142] at frequency $\omega_c/2\pi = 4.573$ GHz. The qubit is a transmon at $\omega_q/2\pi = 6.181$ GHz, dispersively coupled to the cavity with a frequency shift $-\chi/2\pi = -5.25$ MHz per photon. Photon number tracking requires to operate in the photon number resolved regime $\chi > \Gamma_2$, with $\Gamma_2/2\pi \approx 3.5$ MHz the qubit coherence rate [126]. To optimize the information rate, we maximize the qubit emission rate $1/T_q \leq 2\Gamma_2$ into the measurement line under this constraint.

To protect the cavity from decaying through the qubit, we use a notch filter at the cavity frequency (Fig. 4.1b). The filter circuit is composed of two on-chip spurlines in series (Fig. 4.1c). The qubit and the notch filter are patterned out of a tantalum film on the same sapphire chip, which is inserted into the cavity [128]. A galvanic connection is ensured between the measurement line and the on-chip filter using an SMA microwave connector terminated by a pogo pin [1], which is a pin connected to a spring (Fig. 4.1c). This filter design leads to a high coupling rate $1/T_q = (23 \pm 3 \text{ ns})^{-1}$ between the qubit and the measurement line, while preserving a cavity lifetime T_c larger than 200 μs for a single photon. The heterodyne detection benefits from the large bandwidth of a traveling wave parametric amplifier (TWPA [90]) that covers many χ. An auxiliary transmon qubit and its readout resonator are used to perform direct Wigner tomography of the cavity state [128, 143–145].

The qubit is driven with a frequency comb of amplitude Ω and peaks at $\omega_q + k\Delta\omega$ where k spans all integers between $-K$ and K. In the lab frame, the qubit drive Hamiltonian thus reads

$$\hat{H}_d = -\hbar\frac{\Omega}{2}\sum_{k=-K}^{K}\cos\left[(\omega_q + k\Delta\omega)t\right]\hat{\sigma}_y. \tag{4.1}$$

In the limit of an infinite Dirac comb ($K \to \infty$), it becomes a series of Dirac peaks in the time domain with a period $2\pi/\Delta\omega$. In the frame

1 POGO-PIN-19.0-1 by Emulation Technology

rotating at $\omega_q - \chi\hat{c}^\dagger\hat{c}$, and under the rotating wave approximation, it gives

$$\hat{H}_d = \hbar\frac{\pi\Omega}{\Delta\omega}\sum_{l=-\infty}^{\infty}\delta\left(t - \frac{2\pi l}{\Delta\omega}\right)\hat{\sigma}(t), \tag{4.2}$$

where $\hat{\sigma}(t) = \sin\left(\hat{c}^\dagger\hat{c}\chi t\right)\hat{\sigma}_x - \cos\left(\hat{c}^\dagger\hat{c}\chi t\right)\hat{\sigma}_y$ and $\hat{c}^\dagger\hat{c}$ is the photon number in the cavity. The dynamics of the qubit Bloch vector thus consists in periodic kicks every $2\pi/\Delta\omega$ by an angle $\theta = 2\pi\Omega/\Delta\omega$. The natural choice is to have one peak per possible qubit frequency ($\Delta\omega = \chi$). The rotation axis of the kicks would then be the same for any photon number since $\varphi = 2\pi\hat{c}^\dagger\hat{c}\chi/\Delta\omega$ is a multiple of 2π. However, the period between two kicks would then be $2\pi/\chi \approx 190$ ns, which is much longer than T_q. To limit idle times in the qubit fluorescence signal, we choose a twice larger peak spacing $\Delta\omega = 2\chi$, which doubles the information rate (Fig. 4.2a). Consequently, φ is equal to 0 mod 2π for even photon numbers and π mod 2π for odd photon numbers. Therefore the kick direction flips with each kick for odd photon numbers.

In the experiment, we choose a finite number $2K + 1 = 21$ peaks in the comb. In the frequency domain the drive is the product of a square window of width $21\chi/2\pi$ and the infinite comb. Consequently, in the time domain, the resulting waveform is the convolution of the infinite comb with a sinc function. This width sets the timescale of each qubit kick to $\pi/21\chi \approx 5$ ns, which is much shorter than T_q. Additionally, this choice guarantees that the qubit frequency remains well within the bandwidth of the comb, regardless of the desired photon number ranging from 0 to $N_{\max} = 9$. To further minimize boundary effects, we position the comb center frequency at $\omega_q - 4\chi$, which corresponds to the qubit frequency associated with 4 photons in the cavity.

The average predicted dynamics of the qubit is shown in Fig. 4.2b,c when the cavity has 0 (blue) or 1 (green) photon. At intervals of π/χ, the qubit state undergoes a kick lasting approximately 5 ns, followed by relaxation as it fluoresces into the measurement line. In contrast with even photon numbers, the rotation axis flips at every kick for odd photon numbers. Heterodyne detection of the fluorescence field measures the two quadratures $(\hat{a}_{\text{out}}^\dagger + \hat{a}_{\text{out}})/2$ and $i(\hat{a}_{\text{out}} - \hat{a}_{\text{out}}^\dagger)/2$. The emitted field amplitude can be expressed as $\hat{a}_{\text{out}} = \hat{a}_{\text{in}} - \hat{\sigma}_-/\sqrt{T_q}$, where $\langle\hat{a}_{\text{in}}\rangle$ is the driving comb and $\hat{\sigma}_- = (\hat{\sigma}_x - i\hat{\sigma}_y)/2$ is the qubit

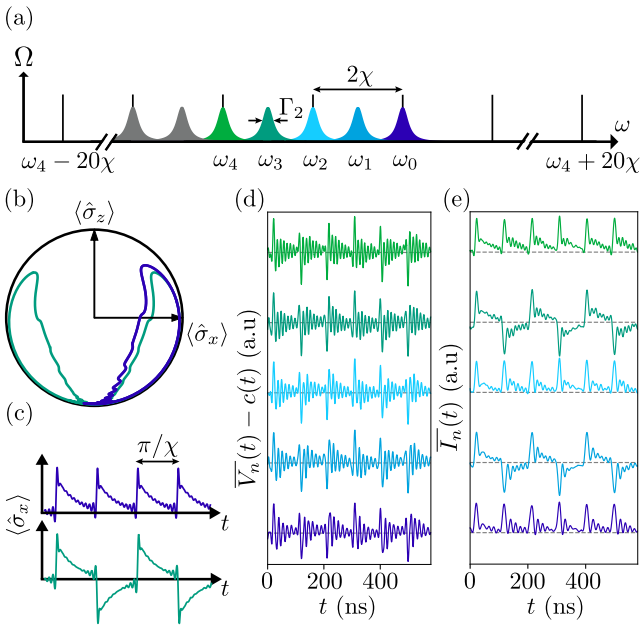


Figure 4.2: Frequency comb and dynamics of the qubit under its drive. (a) Black peaks: driving comb amplitude Ω Eq. (4.1) in the frequency domain. Colored Lorentzian shapes: qubit frequency spectra corresponding to cavity photon numbers from 0 to 6. (b) Simulated average trajectories of the qubit in the Bloch sphere for $n = 0$ (blue) and $n = 1$ (green). The kick angle is set to $\theta = \pi\Omega/\chi = 3\pi/4$, and as in the experiment, $\chi T_q = 0.76$. (c) Corresponding average values of $\langle\hat{\sigma}_x\rangle$ as a function of time. (d) Averaged measured record $\overline{V_n}(t)$ when the cavity is prepared in a Fock state $|n\rangle$ from $n = 0$ to 4 from bottom to top, and for a comb amplitude such that $\theta = \pi\Omega/\chi \approx \pi/2$. The curves are offset by the much larger driving comb signal $c(t)$. (e) Corresponding quadrature $\overline{I_n}(t)$ of the emitted fluorescence signal reconstructed (see text) at each qubit frequency $\omega_q - n\chi$.

lowering operator [146]. The average dynamics of the qubit coherences (Fig. 4.2c) can thus be directly observed in the heterodyne signal.

We first prepare a Fock state $|n\rangle$ using a coherent excitation on the cavity followed by heralding using the qubit emission under a drive at a single tone $\omega_q - n\chi$ [126, 128, 147]. We then apply the comb. The qubit fluorescence is amplified and downconverted by a local oscillator at $\omega_q + \omega_{\mathrm{IF}}$, with $\omega_{\mathrm{IF}} = 2\pi \times 66$ MHz. The amplified fluorescence signal

is recorded as a voltage $V(t)$ using an analog-to-digital-converter. The average $\overline{V_n}(t)$ of these records under the heralding of n photons is shown in Fig. 4.2d offset by the contribution of the reflected driving comb $c(t) \propto \mathrm{Re}(\langle \hat{a}_{\mathrm{in}} \rangle)$ [128]. These signals can be processed to reveal the evolution of $\langle \hat{\sigma}_x \rangle$ and $\langle \hat{\sigma}_y \rangle$ in the qubit frame when there are n photons. To do so, we extract their analytic representations [128] and demodulate them at $\omega_{\mathrm{IF}} + n\chi$ to obtain two average quadratures $\overline{I_n}(t)$ and $\overline{Q_n}(t)$. The traces of $\overline{I_n}(t)$ are shown in Fig. 4.2e for $n = 0$ to 4 and match the expected evolution of $\langle \hat{\sigma}_x \rangle$. The kicks and the subsequent decays are visible. The kick direction alternates for odd numbers of photons as expected. The remaining oscillations in the reconstructed signal may be due to an imperfect subtraction of the driving comb $c(t)$, or to a distortion of the driving comb or output signal by the measurement setup.

Decoding the measurement record $V(t)$ in order to infer the photon number is a task similar to quantum sensing using continuous measurement [55, 148–155]. Here we use the average records $\overline{V_n}$ as demodulation weight functions, and define $N_{\max}+1$ measurement outcomes represented as a vector $\vec{m}$ whose components are

$$m_n(t) = \int_{t-\tau}^{t} V(t')\overline{V_n}(t')\mathrm{d}t', \tag{4.3}$$

where the integration time $\tau = 2$ μs is chosen much shorter than the cavity lifetime T_c and multiple ($21\times$) of π/χ. As a demonstration, we excite the cavity with a coherent state with more than 20 photons on average using a strong resonant pulse, then drive the qubit with the frequency comb and record $\vec{m}$ as a function of time t. Quantum jumps on a single realization can already be visualized by a simple data processing. We perform a time independent linear transform [128] $\vec{r}(t) = \mathbf{G}^{-1}\vec{m}(t)$ so that, on average, $r_n(t) = 1$ for n photons in the cavity, while all the other components $r_{k\neq n}$ vanish. Concretely, $\mathbf{G}$ is the Gram matrix of the average records $\overline{V_n}(t)$ so that $\mathbf{G}_{nm} = \int_0^\tau \overline{V_n}(t)\overline{V_m}(t)\mathrm{d}t$. The evolution of $\vec{r}$ is shown in Fig. 4.3b for one realization of the experiment. A faint red trace emerges from the noise, which reveals the successive losses of single photons in the cavity that here decays from 9 to 0 photons over 1 ms. To predict the number n of photons at any time of the evolution, the probability distribution

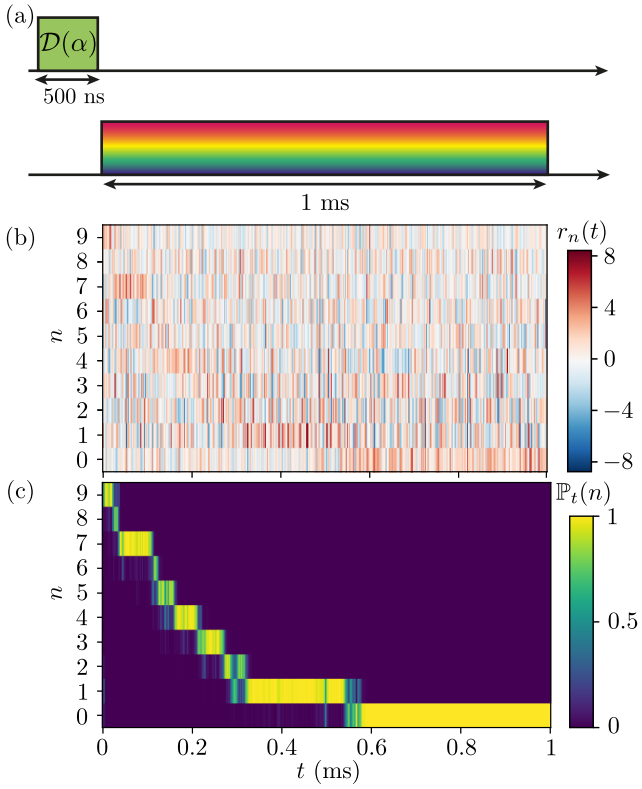


Figure 4.3: Photon number tracking experiment. (a) Pulse sequence for the observation of quantum jumps. (b) Evolution of the measurement outcomes r_n as a function of time t. The measurement is performed on one realization using a driving comb whose amplitude corresponds to an expected kick rotation angle $\theta = \pi/2$. (c) Evolution of the inferred probability distribution $\mathbb{P}_t(n)$ of the photon number n using the outcomes in (b). 1000 realizations in the same conditions are available in [127].

of n is updated conditionally on the outcome $\vec{m}(t)$ through Bayesian update at every time step $j\tau$. It first requires determining the likelihood $\mathbb{P}(\vec{m}|n)$ conditioned on the cavity being in the Fock state $|n\rangle$. $\mathbb{P}(\vec{m}|n)$ can be approximated by a Gaussian function of $\vec{m}$ owing to the small measurement efficiency $\eta = 0.17 \pm 0.02$. Therefore, we characterize its distribution by the measured mean $\langle\vec{m}\rangle_{|n\rangle}$ and covariance matrix of $\vec{m}$ for each $|n\rangle$ only [128]. Using this procedure, along with accounting for photon loss during each time step $j\tau$, the noisy measurement

outcomes $\vec{m}(j\tau)$ of Fig. 4.3b lead to the probability distribution $\mathbb{P}_{j\tau}(n)$ shown in Fig. 4.3c for the same realization. Note that we assume no prior information ($\mathbb{P}_0(n) = 1/10$), but this choice has anyway no impact on the quantum trajectory after a few τ. With many realizations, we extract the average photon number decay. Interestingly, it is not exponential, which indicates a subtle interplay between cavity dissipation and qubit dynamics [128].

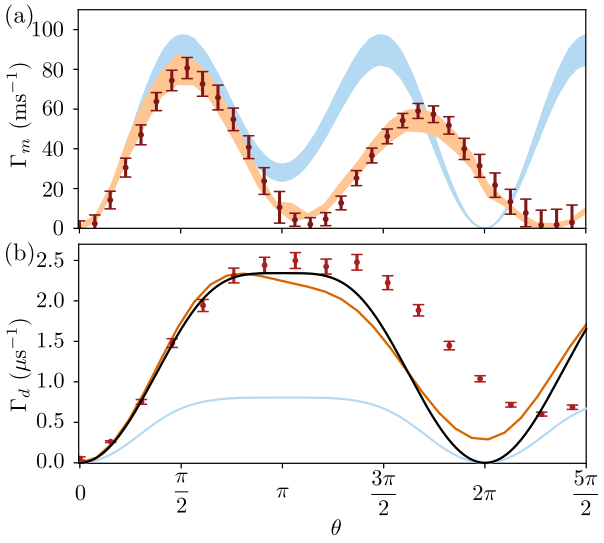


Figure 4.4: Measurement and dephasing rates. (a) Red dots: observed average measurement rate Γ_m as a function of drive amplitude, expressed as the qubit expected rotation angle θ per kick. Orange shadow: measurement rate obtained using a stochastic master equation with detection efficiency η spanning the range $[0.17, 0.20]$. Blue shadow: theoretical bound for an infinite comb and heterodyne measurement with efficiencies $\eta \in [0.17, 0.20]$. (b) Red dots: observed cavity dephasing rate Γ_d as a function of θ. Orange: simulated measurement induced dephasing rate. Black: theoretical accessible information rate. Blue: theoretical maximal ($\eta = 1$) measurement rate obtained by heterodyne detection.

We now determine the measurement rate Γ_m of the photon number. Formally, it is the time derivative of the mutual information between

the photon number n and the outcome $\vec{m}$ at $t = 0$ [40, 92]. In the weak measurement regime $\tau\Gamma_m \ll 1$, it can be approximated by

$$\begin{aligned}\tau\Gamma_m = &- \int \mathbb{P}(\vec{m}) \log \mathbb{P}(\vec{m}) \mathrm{d}\vec{m} \\ &+ \sum_n \mathbb{P}_0(n) \int \mathbb{P}(\vec{m}|n) \log \mathbb{P}(\vec{m}|n) \mathrm{d}\vec{m},\end{aligned} \tag{4.4}$$

where $\mathbb{P}(\vec{m}) = \sum_n \mathbb{P}_0(n)\mathbb{P}(\vec{m}|n)$. As shown in Fig. 4.3c, at most two photon numbers are likely at any time. We thus choose the prior $\mathbb{P}_0(q) = \mathbb{P}_0(q+1) = 1/2$, and average over all q values to compute a measurement rate Γ_m [128]. Plugging in Eq. (4.4) the measured distributions $\mathbb{P}(\vec{m}|n)$ at various driving amplitudes Ω, Γ_m is obtained as a function of the kick angle $\theta = \pi\Omega/\chi$ (Fig. 4.4a). It is maximal when the qubit is kicked to states corresponding to the largest $\langle\hat{\sigma}_x\rangle$. It can be intuitively understood since heterodyne detection probes the quadratures of the emitted fluorescence signal, which are on average proportional to the coherence $\langle\hat{\sigma}_x - i\hat{\sigma}_y\rangle$. Notably, the rate Γ_m is much larger than the cavity decay rate $1/T_c$ (more than 16 times), which is well in the single shot measurement regime. On average, this measurement scheme allows us to reach more than 95% average confidence in 20 μs [128]. Our complete model [128] reproduces the observed rate Γ_m using simulated measurement records by a stochastic master equation with detection efficiency $\eta = 0.185 \pm 0.015$ as a fit parameter, in agreement with the independently measured η.

It is interesting to compare this measurement rate to the rate at which information about the photon number leaks into the environment, *i.e.* the cavity dephasing rate Γ_d. We compute it as the added decay rate of $\mathrm{Tr}\{\hat{\rho}(t)\hat{a}\}$ on top of the natural decoherence rate [128].

We use the auxiliary transmon qubit (purple in Fig. 4.1c) to perform Wigner tomography on the cavity state and extract $\mathrm{Tr}\{\hat{\rho}(t)\hat{a}\} = \int \alpha W_t(\alpha) \mathrm{d}^2\alpha$, with $W_t(\alpha)$ the Wigner function of $\hat{\rho}(t)$. The dependence of Γ_d on kick angle is shown as red dots in Fig. 4.4b. Strikingly, its maximum is reached at $\theta \approx \pi$, where the qubit has the largest energy to emit, and thus the most information to leak out. We note that the model that successfully predicts the measurement rate (orange) underestimates the measurement induced dephasing rate at large drive amplitudes, indicating that the driving comb leads to stronger decoherence than anticipated.

The dephasing rate Γ_d is about 20 times larger than Γ_m at $\theta = \pi/2$ (its maximum). Indeed our measurement setup does not recover the full information available because of a limited detection efficiency and the very use of heterodyne detection. To better understand this information loss, we use a simplified model where the comb is infinite. It reproduces the measured Γ_m with $0.17 < \eta < 0.20$ (blue in Fig. 4.4a) for small angles θ, but not at larger angles where the finite duration of the kicks reduces the actual rotation angle. However, the dephasing rate is larger than the measurement rate of an ideal ($\eta = 1$) heterodyne measurement (blue in Fig. 4.4b). Accordingly, with the same model, the upper bound on the measurement rate for any detection scheme – accessible information rate – (black in Fig. 4.4b) is close to the measured Γ_d for small angles θ and up to about 3 times larger than what the best heterodyne detector could do: Even an ideal heterodyne detector would destroy up to about 2/3 of the accessible information. The experiment provides here a textbook example of destroyed information by a measurement apparatus, here the heterodyne detector [156]. In contrast to low detection efficiency, heterodyne measurement with $\eta = 1$ would reveal information even for $\theta = \pi$. Indeed, while the average heterodyne signal is zero, its cumulants reveal the photon number. The signal-to-noise ratio on the cumulants of order l scales as η^l so that, in the experiment, the average is the main source of information, hence the minimum at $\theta = \pi$.

In conclusion, our superconducting circuit and signal processing demonstrate the possibility to monitor photon numbers in a cavity with a fixed driving. Information about the photon number is extracted up to 16 times faster than the cavity decay rate. The detector requires 20 μs of measurement on average to reach 95 % fidelity between 0 and 9 photons. A circuit using a harmonic oscillator or qudit instead of the qubit as an encoder could enable even faster measurement rate in future devices. A simple model quantitatively explains the dephasing and measurement rates for small drive amplitudes. As we look ahead to more integrated amplifiers, it would be interesting to observe how the driving amplitude that maximizes the measurement rate evolves with increased efficiency. Additionally, the pogo-pin and spurline filters offer a convenient architecture for achieving galvanic coupling with quantum circuits within a long lived microwave cavity. Interesting open

questions remain to be explored such as the origin of the dependence of Fock state decay rates on the comb amplitude. This seems to be a dual effect to the readout problem of superconducting qubit [157–164].

4.2 DEVICE AND MEASUREMENT SETUP

4.2.1 *Device fabrication*

The circuit is composed of one 3D $\lambda/4$ coaxial cavity resonator, into which two samples are inserted. The first sample comprises the monitoring transmon and its notch filter, the second contains the tomography transmon qubit with its readout resonator and Purcell filter. These samples are made of a 200 nm thick film of sputtered Tantalum on a 430 μm thick sapphire substrate (deposited by Star Cryoelectronics, Santa Fe, USA). The Josephson junctions of both transmons are standard Dolan bridge e-beam evaporated Al/AlOx/Al junctions [103].

4.2.2 *Measurement setup*

The readout resonator, the tomography qubit, and the qubit are driven on resonance by pulses that are generated using an OPX from Quantum Machines®. It has a sampling rate of 1 GS/s. AWG driving pulses are modulated at a frequency 125 MHz for readout, 110 MHz for tomography qubit, 70 MHz for cavity and 68 MHz for the monitoring qubit. They are up-converted using I-Q mixers for the readout resonator, monitoring qubit and cavity, regular mixers for the tomography qubit, with continuous microwave tones produced by three channels of an AnaPico® APUASYN20-4 for the readout resonator, cavity and qubit, Agilent® E8257D for tomography qubit. The fourth channel of AnaPico® APUASYN20-4 is used for the TWPA pump. An AnaPico® APSIN12G is used to provide a strong detuned drive on the qubit in order to calibrate the reflection coefficient [165].

The two reflected signals from the readout and qubit are combined with a diplexer and then amplified with a TWPA provided by Lincoln Labs [90]. The follow-up amplification is performed by a HEMT amplifier from Low Noise Factory® at 4 K and by two room-temperature

Table 4.1: Table of circuit parameters

Table of circuit parameters			
Circuit parameter	Symbol	Hamiltonian term	Value
Cavity frequency	$\omega_c/2\pi$	$\hbar\omega_c\hat{a}^\dagger\hat{a}$	4.573 GHz
Qubit frequency	$\omega_q/2\pi$	$\hbar\omega_q\hat{\sigma}_z/2$	6.181 GHz
Tomography qubit frequency	$\omega_t/2\pi$	$\hbar\omega_t\hat{\sigma}_z^t/2$	3.459 GHz
Readout frequency	$\omega_r/2\pi$	$\hbar\omega_r\hat{r}^\dagger\hat{r}$	7.875 GHz
Cavity-qubit cross Kerr rate	$\chi/2\pi$	$-\hbar\chi\hat{a}^\dagger\hat{a}\hat{\sigma}_z$	5.25 MHz
Cavity-tomography qubit cross-Kerr rate	$\chi_{ct}/2\pi$	$-\hbar\chi_{ct}\hat{a}^\dagger\hat{a}\hat{\sigma}_z^t/2$	593 kHz
Cavity self-Kerr rate	$\chi_{cc}/2\pi$	$-\hbar\chi_{cc}\hat{a}^{\dagger 2}\hat{a}^2$	9 kHz
Circuit parameter	Symbol	Dissipation operator	Value
Qubit decay time	T_q	$1/T_q\mathbb{D}_{\hat{\sigma}_-}$	22 ns
Cavity decay time	T_c	$1/T_c\mathbb{D}_{\hat{a}}$	200 μs
Tomography qubit decay time	T_t	$1/T_t\mathbb{D}_{\hat{\sigma}_-^t}$	3.6 μs
Readout decay time	$T_{1,r}$	$1/T_{1,r}\mathbb{D}_{\hat{r}}$	415 ns
Cavity dephasing time	$T_{c,\phi}$	$2/T_{c,\phi}\mathbb{D}_{\hat{a}^\dagger\hat{a}}$	36 μs
Tomography qubit dephasing time	$T_{q,\phi}$	$1/2T_{q,\phi}\mathbb{D}_{\hat{\sigma}_z^t}$	12.3 μs

amplifiers. The two signals are down-converted using image reject mixers before digitization by the input ports of the OPX. We tune the TWPA pump frequency (6.079 GHz) and power in order to reach a total quantum efficiency of $\eta = 17 \pm 1$ % at $\omega_q/2\pi + 10$ MHz. The efficiency was extracted from the measured mean and variance of the demodulated quadratures using the qubit as a reference for calibrating the gain of the amplification chain (see section 6 in Ref. [126]). However, we observed on posterior experiments that this efficiency can fluctuate by almost 20 % between $\omega_q/2\pi$ and $\omega_q/2\pi - 9\chi$.

The frequency comb is generated by numerically computing the temporal shape corresponding to a frequency comb. The use of an I-Q mixer allows to use negative intermediate frequencies to expand the comb on both sides of the source frequency.

4.2.3 *Comparison with previous experiment*

We summarize in table 4.2 the main differences with the reference [126] that allowed us to reach single shot photon number measurement, extract a measurement rate and observe quantum jumps. One big difference is the use of a 3D cavity that extended the resonator decay by 2 orders of magnitude. Another key difference is the use of the analog of a Purcell filter for the cavity, galvanically connected in a 3D design. The filter allowed us to preserve the two orders of magnitude of improvement on the resonator lifetime despite a similar coupling rate between qubit and cavity. The expected measurement rate is slightly lower, because the dispersive coupling rate χ between the qubit and the cavity could have been larger to benefit from the strong qubit emission rate into the measurement line. But this is compensated by a far the longer decay time for the cavity.

In terms of results, the main figure of merit of the experiment is $\Gamma_m T_c$, which was improved from 0.5 to 16, reaching the single shot regime. Additionally, the experimental technique changed, as we use a much wider comb experimentally, as it was proposed (but not implemented) in [126].

The expected measurement rate for the reference [126] was not computed but is here obtained via the same model than for the blue curve of Fig. 4.4a in the main text.

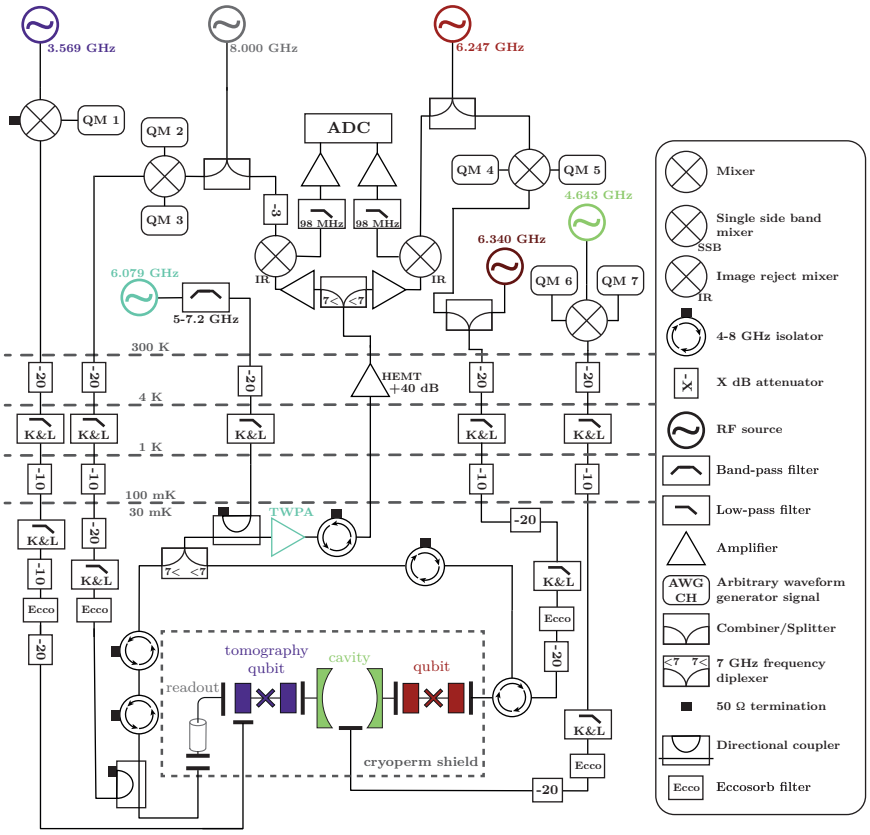


Figure 4.5: Schematic of the setup. Each element has its RF source, whose color is matched. Room-temperature isolators are not represented for the sake of clarity.

4.3 SINGLE-TONE PHOTON COUNTING FOR PRE- AND POST-SELECTION

4.3.1 *Single shot measurement with a single tone*

In this section, we focus on the case where the qubit is driven with a single tone to determine whether the cavity contains a certain number of photons [126, 147]. This configuration is reciprocal to the usual dispersive readout of a qubit. We introduce the reflection coefficient as the ratio $\overline{r} = \langle \hat{a}_{\text{out}} \rangle / \langle \hat{a}_{\text{in}} \rangle$ in the steady state. The denominator is calibrated by applying a strong ac Stark shift drive to the qubit

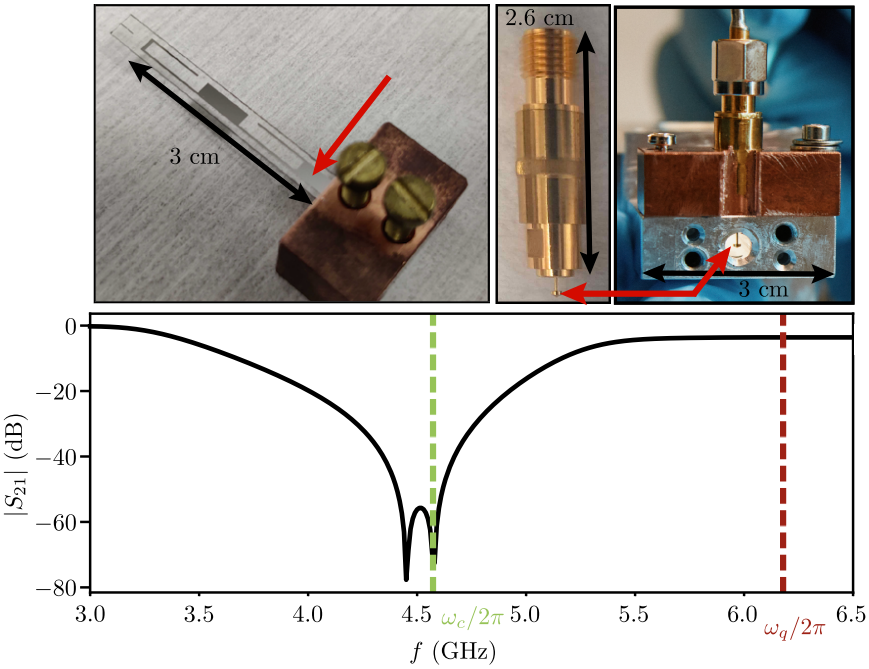


Figure 4.6: Pogo pin and filter performances. Top panel: pictures of, from right to left, the qubit and the on-chip filter on its support, the pogo pin, and the pogo pin inserted in the tunnel in which the chip is inserted. Red arrows show the pogo pin and the contact point on the sample. Bottom panel: frequency profile of the spurline filters, from an HFSS simulation. The actual frequency of the cavity is represented in green, that of the qubit in dark red.

(Fig. 4.7a), so that it is detuned far from ω_q. We generalize the definition of a reflection coefficient r to the case of single measurement records. Therefore, after a displacement of the cavity, we can build histograms of r as a function of the frequency of the tone (Fig. 4.7b). For a given photon number, the average reflection coefficient spans a circle in the complex plane as the probe frequency is changed [103]. We introduce a phase φ relative to the center $r_0 = 0.3 + 0 \times i$ of that circle (see Fig. 4.7b). We recover the usual 2π phase shift of the reflection coefficient on a harmonic oscillator in the angular distribution $\mathbb{P}(\varphi)$. Each number of photons leads to a unique phase shift as can be observed in Fig. 4.7c. In particular, when we drive at $\omega_q - n\chi$,

Table 4.2: Summary of the differences with [126]

	Work [126]	This work
Architecture	2D	3D
Cavity filter	no	yes
Qubit decay rate	42 ns^{-1}	22 ns^{-1}
Dispersive shift $\chi/2\pi$	4.9 MHz	5.25 MHz
Cavity decay time T_c	3.8 μs	240 μs
Bandwidth of the comb	9χ	40χ
Expected maximal measurement rate Γ_m	130 ms^{-1}	80 ms^{-1}
$\Gamma_m T_c$	0.5	16

the histogram peak corresponding to Fock state $|n\rangle$ is centered on a π phase shift. Thus, even though the histograms of various photon numbers overlap, it is possible to pre/post-select this Fock state by setting a threshold on the real value of one averaged single record. The tighter the threshold, the higher the fidelity but the smaller the selection yield as we discuss now.

4.3.2 *Pre- and post selection using the qubit*

The protocol we use to pre- and post-select a given number of photons is shown in Fig. 4.8a. It is performed by sending a 20 μ-drive at $\omega_q - n\chi$ before and after the measurement. For each pulse, we determine the associated reflection coefficient. The measurement is heralded to Fock state $|n\rangle$ if both measurements give reflection coefficients falling below threshold. The thresholds we use are shown in Fig. 4.9a,c as dashed red lines on the measured histograms at the post-selection step for each desired photon number n. The value of the threshold r_0 was chosen so that the likelihood ratio between {the number of photons n being equal to the selected number of photons k} and {n being equal to k'} is higher than 100 for every target photon number k(i.e., for any selected frequency $f - k\chi$) and for any $k' \neq k$:

$$\frac{\mathbb{P}(\text{Re}(r) < r_0 | n = k)}{\mathbb{P}(\text{Re}(r) < r_0 | n = k')} > 100.$$

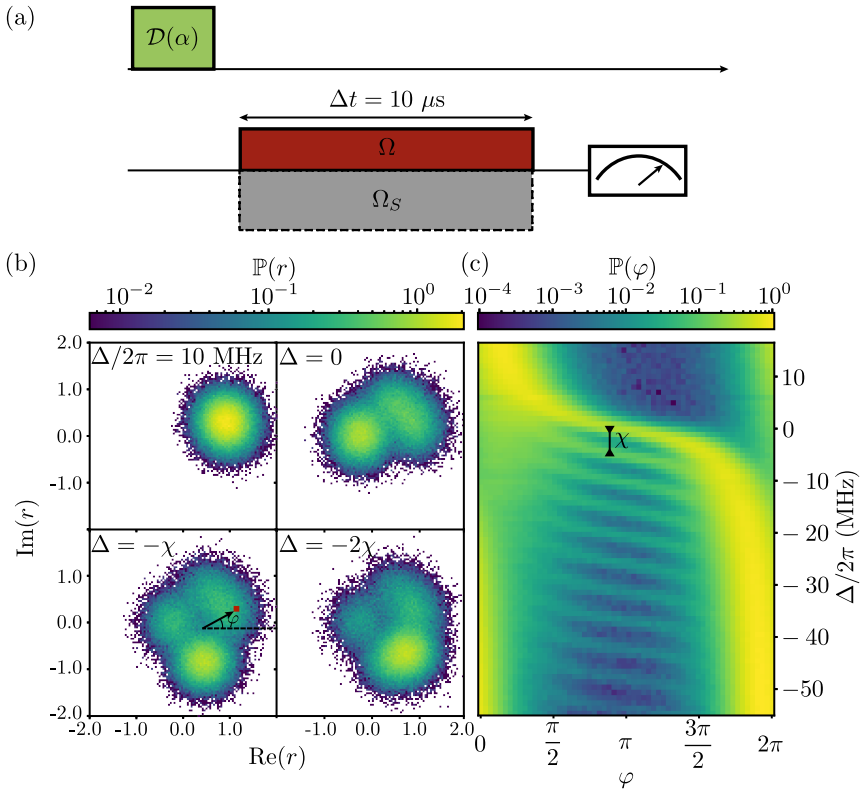


Figure 4.7: Single-frequency drive photon counting. (a) Pulse sequence. The cavity is driven on resonance so as to excite it with large number of photons, and a series of 10 μs readout pulse (red) with amplitude Ω and detuning $\Delta = \omega - \omega_q$ are applied on the qubit while the reflected signal is recorded to infer a reflection coefficient r. The same protocol is then performed while applying a strong ac Stark shift drive (gray, dashed). It detunes the qubit far from the frequencies of interest, which provides a reference signal to properly calibrate the reflection coefficient. (b) Histograms of the reflection coefficient r constructed from this procedure. We can clearly distinguish the components coming from the different number of photons. (c) Distribution of the argument φ of the complex number r computed from (b) as a function of the detuning Δ. χ appears as the frequency translation parameter of the resonance of the qubit.

There is a trade-off between preparation fidelity by heralding and its yield. We choose a drastic pre-selection threshold on Re(r) that

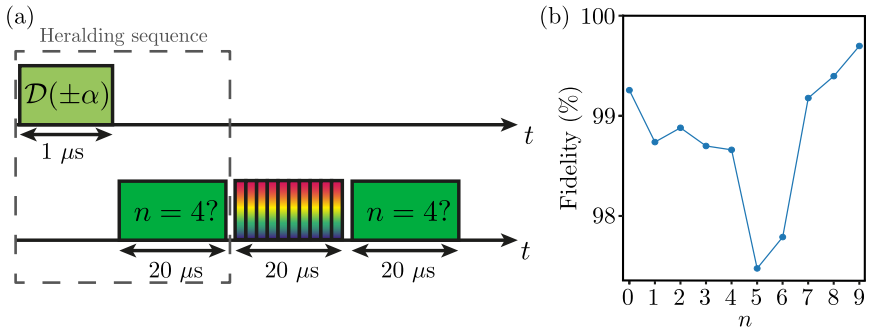


Figure 4.8: Pre-and post-selection on the photon number. (a) Pre- and post-selection pulse sequence. The cavity is alternatively displaced by $\pm\alpha = 3.4$ (which is believed to improve slightly the pre-selection yield without reset compared to $+\alpha$ always). A first pre-selection 20 µs pulse is then applied to the qubit. This heralding sequence (dashed box) is repeated until the measured $\mathrm{Re}(r)$ gets below a threshold. A 20 µs comb pulse is then applied to the qubit. In practice, it results from the concatenation of 10 identical sequences of 2 µs. We then apply a post-selection pulse on the qubit. (b) Estimated pre-selection fidelity as a function of the Fock state. The selection yields associated to the threshold we used (red dashed lines in Fig. 4.9) are given in Table 4.3. These fidelities are determined by computing the probability to get the right number of photons after applying the threshold. For this, a prior probability distribution for the number of photons is needed, which we deduced from the yield statistics in Table 4.3.

ensures preparation fidelities of at least 97 % by the pre-selection pulse alone as shown in Fig. 4.8b while the selection yields (below 1 %) are reported in Table 4.3.

4.4 COMPUTING THE THEORETICAL MEASUREMENT RATES

Our model for the measurement focuses on what happens after a single kick of the qubit state. Indeed, the experiment can be decomposed into a repetition of the same sequence. Every π/χ, the qubit is prepared in a certain state $\hat{\rho}_q$ that depends on the amplitude of the comb, and decays. Since in our case, $\pi/\chi \simeq 4T_q$, we use a model where the qubit fully relaxes to its ground state between each kick. We can thus

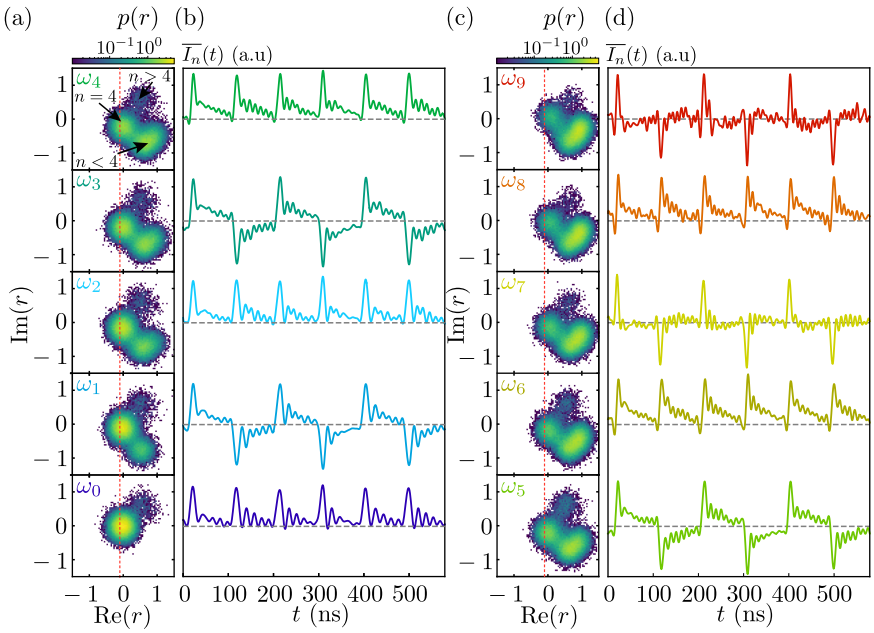


Figure 4.9: Pre-and post-selected average qubit fluorescence. (a) and (c) Histograms of the measured quadratures during the second (post-selection) pulse on the qubit for each number n of photons corresponding to a driving tone and heterodyne measurement at ω_n. The upper threshold on $\mathrm{Re}(r)$ for the pre- and post-selection readout is materialized by the dashed red lines. There are more than one Gaussian component due to the decay between the pre- and post-selection pulses. (b) and (d) Reconstructed average fluorescence of the qubit for each number of photons. Note that the figure (b) is identical to Fig. 2e.

assume the kick places the qubit in a pure state $\hat{\rho}_q = |\psi\rangle\langle\psi|$ with $|\psi\rangle = \cos(\theta/2)\,|0\rangle + \sin(\theta/2)\,|1\rangle$, where $\theta = 2\pi\Omega/2\chi$ is the kick angle (see main text).

Following section F of the supplemental material of Ref. [40], the measurement rate of the photocounting can be extracted from the mutual information between the qubit frequency and a measurement record, for the most entropic prior (all states assumed equally likely initially). We compute it in the case where the qubit can only have one out of two frequencies separated by χ corresponding to a given photon number or one more.

Table 4.3: Yield table of the pre- and post-selection after a displacement of $\alpha \approx 3.4$ when the kick angle θ is set to 0.

Fock state fraction	Pre-selection fraction (%)	Post-selection fraction (%)	Typical number of samples $2N_n$ after post-selection
0	0.5	33.3	200580
1	0.7	26.0	157370
2	0.7	19.7	119370
3	0.6	14.1	85800
4	0.5	11.1	67150
5	0.4	7.4	44990
6	0.3	4.2	25650
7	0.2	3.2	19330
8	0.2	1.9	11360
9	0.1	0.9	5580

We model the experiment as follows. The bipartite system {cavity, qubit} starts in the initial state $\hat{\rho}(0) = \hat{\rho}_c \otimes \hat{\rho}_q$. The full Stochastic Master Equation (SME), neglecting the pure dephasing of the qubit, of the system reads

$$\begin{aligned} d\hat{\rho} = &-i\left[\omega_c \hat{a}^\dagger \hat{a} + \tfrac{1}{2}\omega_q \hat{\sigma}_{\mathbf{z}} - \tfrac{1}{2}\chi \hat{a}^\dagger \hat{a}\hat{\sigma}_{\mathbf{z}},\ \hat{\rho}\right] \mathrm{d}t \\ &+ \left(\frac{1}{T_q}\mathbb{D}_{\hat{\sigma}_-}(\hat{\rho}) + \frac{1}{T_c}\mathbb{D}_{\hat{a}}(\hat{\rho})\right) \mathrm{d}t \\ &+ \sqrt{\frac{\eta}{T_q}}\Big(\hat{\sigma}_-\hat{\rho} - \mathrm{Tr}\{\hat{\sigma}_-\hat{\rho}\}\hat{\rho}\Big)\mathrm{d}W^* \\ &+ \sqrt{\frac{\eta}{T_q}}\Big(\hat{\rho}\hat{\sigma}_+ - \mathrm{Tr}\{\hat{\rho}\hat{\sigma}_+\}\hat{\rho}\Big)\mathrm{d}W \end{aligned} \tag{4.5}$$

where $\mathrm{d}W$ is now a complex Wiener process such that $\mathrm{d}W\mathrm{d}W^* = \mathrm{d}t$. The complex measurement record with efficiency $0 < \eta < 1$ reads

$$\mathrm{d}y_t = \sqrt{\eta/T_q}\,\mathrm{Tr}\{\hat{\rho}\hat{\sigma}_-\}\mathrm{d}t + \mathrm{d}W.$$

It can be proven that the information contained in the measurement record y_t about the photon number in the cavity is preserved by demodulating the measurement record as follows

$$\tilde{m}_k = \frac{1}{\sqrt{T_q}} \int_0^{+\infty} e^{i(\omega_q - k\chi)t - \frac{t}{2T_q}} \mathrm{d}y_t, \tag{4.6}$$

for all possible photon numbers k.

We can take into account the finite efficiency η of the detection as a beam-splitter that combines a fraction of the field emitted by the qubit with vacuum noise. Assuming the cavity is in state $|n\rangle$, the outcome $\tilde{m}_n$ is a stochastic variable that reads

$$\tilde{m}_n = \sqrt{\eta}\alpha + \sqrt{1-\eta}\gamma, \tag{4.7}$$

where α and γ are stochastic variables whose probability densities are respectively given by the Husimi Q functions of the qubit state $\hat{\rho}_q$ and of the vacuum state. Owing to the low measurement efficiency $\eta \approx 0.2$, the distribution function of $\tilde{m}_n$ is almost Gaussian as can be seen in Fig. 4.10 in the cut along the real axis. For the quantum efficiency $\eta \leq 0.2$ of the experiment, we see that the two curves are very close, thus justifying the Gaussian approximation of the measurement outcomes.

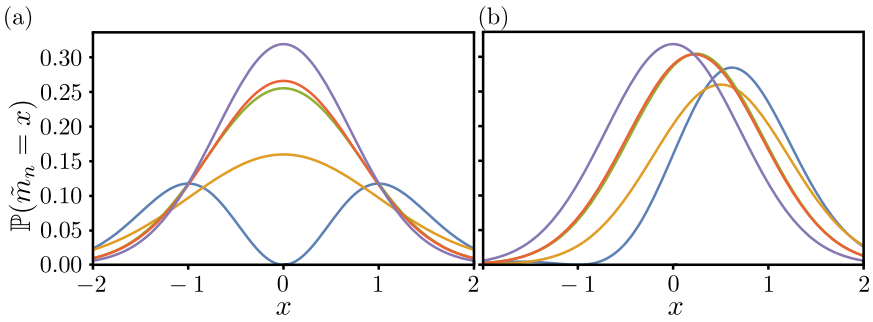


Figure 4.10: Gaussian approximation of the outcomes $\tilde{m}_n$. Calculated probability distributions of $\mathbb{P}(\tilde{m}_n = x)$ as a function of x when the cavity is in $|n\rangle$ for (a) $|\psi\rangle = |1\rangle$ and (b) $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Blue: $\eta = 1$. Orange: its Gaussian approximation with same mean and covariance matrix (on the whole complex plane $\tilde{m}_n \in \mathbb{C}$). Green: $\eta = 0.2$. Red: its Gaussian approximation. Purple: case where $|\psi\rangle = |0\rangle$.

The other measurement outcomes $\tilde{m}_{k\neq n}$ are obtained with a wrong demodulation frequency $\omega_q - k\chi$. However there is still a finite overlap $\frac{\Gamma_q}{\Gamma_q+i(k-n)\chi}$ between the decaying modulation functions at n and k, where $\Gamma_q = 1/T_q$. The outcome $\tilde{m}_{k\neq n}$ can thus be written as a combination of $\tilde{m}_n$ and of a stochastic variable β_k distributed as the Husimi Q function of the vacuum state with weights corresponding to that overlap.

$$\tilde{m}_{k\neq n} = \frac{\Gamma_q}{\Gamma_q + i(k-n)\chi}\tilde{m}_n + \frac{i(k-n)\chi}{\Gamma_q + i(k-n)\chi}\beta_k \tag{4.8}$$

Knowing that α, β_k, γ are independent variables, one can compute $\mathbb{P}(\vec{\tilde{m}}|n)$. Knowing that the repetition time is $\tau = \frac{\pi}{\chi}$, we use equation (4) in the main text to predict the measurement rates for $\eta = 0.17, 0.2$ (blue in Fig. 4a) and $\eta = 1$ (blue in Fig. 4b). For larger kick angles, we observe a discrepancy between the experimental data and the model in Fig. 4a. The x-axis corresponds to the ideal angle θ reached by the qubit for an infinite frequency comb. However, both in the experiment and in the simulation, the finite duration of the kicks lowers the actually reached rotation angle below θ. This is why the blue curve behaves as expected while the experiment and simulations appear distorted in θ.

4.5 DATA PROCESSING

This section provides details on all the steps involved in the data processing used in the letter.

4.5.1 *Quantum model of the recorded voltage*

The qubit couples the incoming modes $\hat{a}_{\text{in}}(t)$ to the outgoing modes $\hat{a}_{\text{out}}(t)$ following the input-output relation

$$\hat{a}_{\text{out}}(t) = \hat{a}_{\text{in}}(t) - \sqrt{\frac{1}{T_q}}\hat{\sigma}_-(t). \tag{4.9}$$

The outgoing modes $\hat{a}_{\text{out}}(t)$ get amplified and downconverted by a local oscillator at a frequency $\omega_{LO}/2\pi = 6.247$ GHz. The noise

added by the amplifiers can be modeled as a finite temperature of input modes $\hat{b}_{\text{in}}(t)$, which are assumed to be in a Gibbs state. This model (Fig. 4.11) leads to the following expression for the voltage $V(t)$ that is measured by the analog-to-digital converter (ADC) at room temperature, where G encompasses the conversion factor between the $\sqrt{\text{Hz}}$ of the propagating mode amplitudes and Volts of the ADC records. $V(t)$ is the measurement outcome of the observable

$$\begin{aligned}\hat{V}(t) = \sqrt{G}\Big(\hat{a}_{\text{out}}(t)e^{i\omega_{\text{LO}}t} + \hat{a}^{\dagger}_{\text{out}}(t)e^{-i\omega_{\text{LO}}t}\Big) \\ + \sqrt{G-1}\Big(\hat{b}^{\dagger}_{\text{in}}(t)e^{i\omega_{\text{LO}}t} + \hat{b}_{\text{in}}(t)e^{-i\omega_{\text{LO}}t}\Big).\end{aligned} \tag{4.10}$$

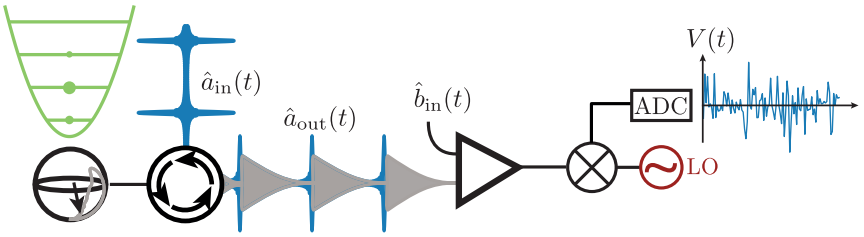


Figure 4.11: Scheme of the quantum model of the detection of the fluorescence.

4.5.2 *Data processing flow*

From the measured voltages, we get several relevant quantities that are discussed in the main text. Here, we show how they are related. Each Fock state $|n\rangle$ from 0 to 9 is prepared a large number of times (see Table. 4.3) by pre- and post-selection, leading to as many measurement records $V_n(t)$ of the voltage $\hat{V}(t)$ over 10 slices of 2 μs.

For any measurement record $V(t)$, we define the measurement outcome $\vec{m}$ whose components are

$$m_j = \int_0^{2\ \mu\text{s}} V(t)\overline{V_j}(t)\text{d}t, \tag{4.11}$$

which corresponds to $m_j(2\ \mu\text{s})$ in Eq. (4.3) of the main text.

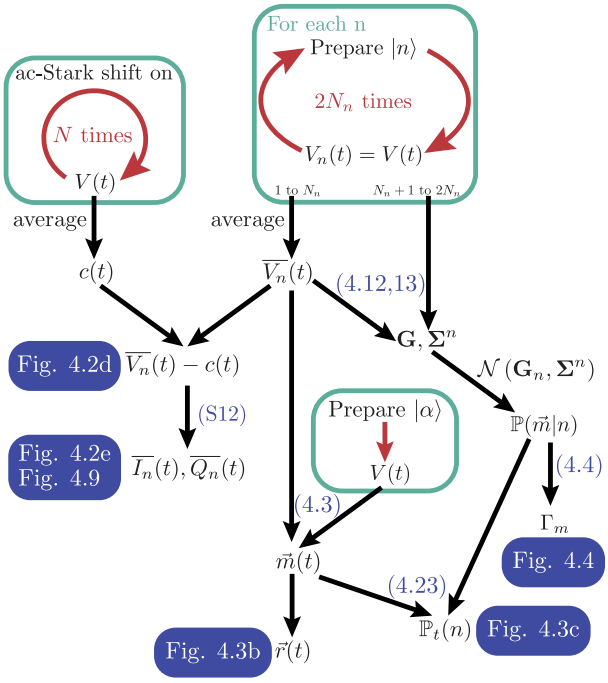


Figure 4.12: Diagram of the signal processing (see text).

For each photon number n, we compute the mean $\overline{V_n}(t)$ of the $V_n(t)$ whose expectation is $\langle \hat{V}(t) \rangle_{|n\rangle}$. We can then calculate the Gram matrix (matrix of scalar products)

$$\mathbf{G}_{nm} = (\overline{V_n}|\overline{V_m}) = \int_0^{2\ \mu s} \overline{V_n}(t)\overline{V_m}(t)dt. \tag{4.12}$$

Since the number of samples is finite, the mean $\overline{V_n}$ is a stochastic variable. In order to avoid biasing the Gram matrix with the particular realization of $\overline{V_n}$, we compute each scalar product between averages $(\overline{V_n}|\overline{V_m})$ by taking half the samples for the left part and the other half for the right part (see Fig. 4.12). We also compute the covariance matrices $\mathbf{\Sigma}^n$ of the probability distribution of the measurement outcomes $\vec{m}^n$ obtained from $V_n(t)$ when there are n photons. Its matrix elements read

$$\mathbf{\Sigma}^n_{jk} = \mathrm{Cov}\{(V_n|\overline{V}_j), (V_n|\overline{V_k})\} = \overline{(m^n_j - \mathbf{G}_{jn})(m^n_k - \mathbf{G}_{kn})}. \tag{4.13}$$

As for the Gram matrix, we compute each scalar product $(V_n|\overline{V}_j)$ by taking half the samples for the left part and the other half for the right part.

As explained in Fig. 4.10, the distribution probability $\mathbb{P}(\vec{m}|n)$ of $\vec{m}^n$ can be approximated by a Gaussian law for our low efficiency $\eta \simeq 0.2$. Therefore, we use the experimentally determined Gram and covariance matrices to reconstruct this law $\mathcal{N}(\mathbf{G}_n, \mathbf{\Sigma}^n)$. It is used in the Bayesian update of the photon number and for the computation of the measurement rate. For the estimation of $\mathbb{P}(\vec{m}|n)$ used in the Bayesian filter for the photon number tracking in Fig. 3, we pre- and post-selected 60600×10 realizations following the protocol in Fig. 4.8. However, since we require many drive amplitudes for Fig. 4, we selected only 6600×10 realizations for each amplitude in that figure. We there chose to remove the numbers 8 and 9 from the analysis as the insufficient number of samples induces very large uncertainties on the measurement rate.

From the m_n, we can compute the r_n (as in the main text) and look at their probability distribution, which is shown Fig.4.13 for r_5. The average value of r_5 is 1 when there are 5 photons in the cavity (brown), and 0 otherwise. All the records that are conditioned to other numbers of photons have the same Gaussian law, centered around 0. Because of the noise of the measurement, the value of the records may take values as large as 8 or as low as -8, which can also be seen in Fig. 3b. This is analogous to the measurement record of a dispersive readout of a qubit. Note here that all the r_n are correlated in general, so looking at the marginals is not enough to extract all the information about the number of photons. In particular, while this figure gives a hint about how well we can distinguish the Fock state 5 from the others in 2 μs, it only gives a lower bound on the Signal-to Noise Ratio.

4.5.3 *Reconstruction of the qubit fluorescence signal*

In order to avoid being blinded by the drive, we remove the comb contribution (which is measured by driving the qubit off resonance using ac-Stark effect)

$$c(t) = \sqrt{G}\Big(\langle \hat{a}_{\text{in}}(t)\rangle e^{i\omega_{\text{LO}}t} + \langle \hat{a}^\dagger_{\text{in}}(t)\rangle e^{-i\omega_{\text{LO}}t}\Big) \tag{4.14}$$

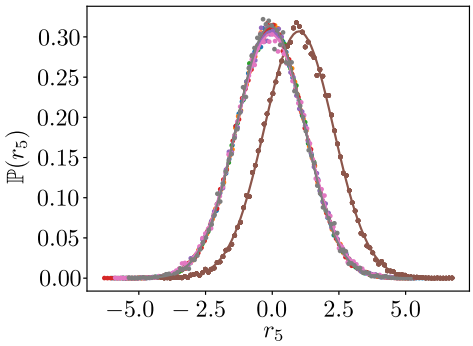


Figure 4.13: Conditional probability distribution of the orthonormalized records r_n. Brown dots: measured probability distribution for the record r_5 after 2 μs of integration when 5 photons are prepared in the cavity. Dots with other colors: measured probability distribution for the record r_5 when there are k photons in the cavity, for k between 0 and 9, $k \neq 5$, and the same integration time. Solid line: Gaussian fits.

to the average measurement records $\overline{V_n}(t)$ (Fig. 2d in the main text).

Following Eqs. (4.10) and (4.14), we express the expectation value of $\overline{V_n}(t) - c(t)$ as

$$\overline{V_n}(t) - c(t) = -\sqrt{\frac{G}{T_q}}\left(\langle\hat{\sigma}_-(t)\rangle_{|n\rangle} e^{i\omega_{\mathrm{LO}}t} + \langle\hat{\sigma}_+(t)\rangle_{|n\rangle} e^{-i\omega_{\mathrm{LO}}t}\right). \quad (4.15)$$

Here, $\langle\hat{\sigma}_-(t)\rangle_{|n\rangle}$ oscillates at $(\omega_q - n\chi)$, so $\langle\hat{\sigma}_-(t)\rangle_{|n\rangle} e^{i\omega_{\mathrm{LO}}t}$ oscillates at $\omega_{\mathrm{LO}} - \omega_q + n\chi = \omega_{\mathrm{IF}} + n\chi$. To reconstruct the fluorescence of the qubit in the frame rotating at its frequency, we bring these oscillating signals to zero frequency. This is performed using their Hilbert transform $\mathcal{H}$ to reconstruct their analytic form $-\sqrt{\frac{G}{T_q}}\langle\hat{\sigma}_-(t)\rangle_{|n\rangle} e^{i\omega_{\mathrm{LO}}t}$ (here with negative frequencies). This allows to perform a sideband-free numerical demodulation to recover the average coherence of the qubit. It yields

$$\overline{I_n}(t) - i\overline{Q_n}(t) = e^{-i(\omega_{\mathrm{IF}} + n\chi)t + i\phi_n}\left(\overline{V_n}(t) - c(t) + i\mathcal{H}[\overline{V_n} - c](t)\right) \quad (4.16)$$

where ϕ_n is chosen such that the most significant quadrature, i.e the one carrying the most fluorescence, is $\overline{I_n}$. These quadratures are plotted in Fig. 2.e in the main text and Fig. 4.9. Note that the demodulation frequency is $(\omega_{\mathrm{IF}} + n\chi)/2\pi$ and not $-n\chi$, which originates from the fact that the qubit is driven with a lower sideband.

4.5.4 *Relevance of the demodulation procedure*

In this section, we justify how well the demodulation we perform following Eq. (4.3) (main text) preserves the information about the photon number that is contained in the measurement records. Indeed, using the average measurement record $\overline{V_n}(t)$ as a demodulation function does not straightforwardly appear equivalent to Eq. (4.6) that states that using the demodulation function $e^{i(\omega_q - n\chi)t - \frac{t}{2T_q}}$ preserves all the information about the photon number.

Actually, the demodulation function $e^{i(\omega_q - n\chi)t - \frac{t}{2T_q}}$ is proportional to $\mathrm{Tr}\left(\overline{\hat{\rho}_n(t)}\hat{\sigma}_-\right)$, where $\overline{\hat{\rho}_n(t)}$ is the solution of a simple Lindblad master equation describing the evolution of a qubit at frequency $\omega_q - n\chi$ and with a relaxation time T_q. Therefore a valid way to demodulate the measurement record in the experiment would be to measure the average down-converted and pre/post-selected complex amplitude $e^{i\omega_{\mathrm{LO}}t}\,\mathrm{Tr}\left(\overline{\hat{\rho}_n(t)}\hat{\sigma}_-\right)$ first, and to use it to demodulate the measurement record. Formally, it corresponds to $\overline{V_n}(t) - c(t) + i\mathcal{H}[\overline{V_n} - c](t)$ defined in Eq. (4.16), where $\mathcal{H}$ is the Hilbert transform. However, using $\overline{V_n}(t)$ alone does not lead to any visible loss of information between the two methods. As a matter of fact, demodulating by $\overline{V_n}(t) - c(t) + i\mathcal{H}[\overline{V_n} - c](t)$, $\overline{V_n}(t) - c(t)$ alone or $\overline{V_n}(t)$ does gather the same amount of information, up to experimental uncertainties.

Note that this simplification is only possible thanks to the relatively low quantum efficiency η. Were the efficiency η close to 1, we would be able to gather information in all cumulants of the probability distribution of $\tilde{m}$ as well (see Fig. 4.10), which would require to keep the full expression $\overline{V_n}(t) - c(t) + i\mathcal{H}[\overline{V_n} - c](t)$.

4.5.5 *Measurement rate estimation*

Once $\mathbf{G}$ and $\mathbf{\Sigma}$ known, the measurement rate can be computed using

$$\begin{aligned}\tau\Gamma_m = I(n:\vec{m}) = &-\int \mathbb{P}(\vec{m})\log\mathbb{P}(\vec{m})\mathrm{d}\vec{m} \\ &+\sum_n \mathbb{P}_0(n)\int \mathbb{P}(\vec{m}|n)\log\mathbb{P}(\vec{m}|n)\mathrm{d}\vec{m},\end{aligned} \tag{4.17}$$

for any prior $\mathbb{P}_0$ for which there exists a number q such that $\mathbb{P}_0(q) = \mathbb{P}_0(q+1) = 1/2$, and where $I(n : \vec{m})$ is the mutual information of n and $\vec{m}$.

A subtlety arises from the fact that this expression is only valid for times τ such that $\tau\Gamma_m \ll 1$. For the largest measurement rates we obtained in the experiment, we cannot use the full integration time of 2 μs. However it is straightforward to compute what this expression gives if τ were as small as the period of the pulse train driving the qubit, which is here $\pi/\chi = \tau/21$. For this computation, we then rescale the Gram matrix into $\mathbf{G}' = \mathbf{G}/\sqrt{21}$.

Note that we chose to integrate over 2 μs because of a technical reason. It is simply the smallest multiple of the pulse train period that is also a multiple of the inverse sampling rate of the OPX instrument (1 ns).

The mutual information obtained in this way is actually biased. Indeed, even if all the $\mathbf{G}_n$ and $\mathbf{\Sigma}^n$ are equal ($I(n : \vec{m}) = 0$), their experimental value are slightly different owing to noise, thus rendering $I(n : \vec{m})$ positive. To estimate this bias, we measured $I(n : \vec{m}_0)$ for zero driving amplitude, and assumed that it should be zero exactly, neglecting the thermal emission of the qubit. We observed that $I(n : \vec{m}_0)$ scales linearly with the inverse of the number of samples. We then used the computed $I(n : \vec{m}_0)$ at zero amplitude as a calibration of the systematic errors for all the other amplitudes. For each driving amplitude, the expected positive systematic error inversely scales with the number of measured samples given in Table 4.3. The error bars displayed in Fig. 4 correspond to this systematic error.

4.5.6 *Simulation*

In order to simulate the experiment, we numerically solve the following Stochastic Master Equations (SME), for each n between 0 and 7.

$$\mathrm{d}\hat{\rho}(t) = -\frac{i}{\hbar}[\hat{H}_n, \hat{\rho}]\mathrm{d}t + \frac{1}{T_q}\mathbb{D}_{\hat{\sigma}_-}(\hat{\rho}(t))\mathrm{d}t + \sqrt{\frac{\eta}{T_q}}\mathcal{H}_{\hat{\sigma}_-}(\hat{\rho}(t))\mathrm{d}W(t), \tag{4.18}$$

with

$$\hat{H}_n = \hbar\left(\omega_{IF} + n\chi\right)|e\rangle\langle e| \tag{4.19}$$

$$\mathbb{D}_{\hat{A}}(\hat{\rho}) = \hat{A}\hat{\rho}\hat{A}^\dagger - \frac{1}{2}\Big(\hat{A}^\dagger\hat{A}\hat{\rho} + \hat{\rho}\hat{A}^\dagger\hat{A}\Big) \tag{4.20}$$

$$\mathcal{H}_{\hat{A}}(\hat{\rho}) = \hat{\rho}\hat{A}^\dagger + \hat{A}\hat{\rho} - \mathrm{Tr}\Big\{\hat{\rho}\hat{A}^\dagger + \hat{A}\hat{\rho}\Big\}. \tag{4.21}$$

The homodyne measurement record reads

$$\mathrm{d}y_t = \sqrt{\frac{\eta}{T_q}}\,\mathrm{Tr}(\hat{\sigma}_x\hat{\rho}(t))\mathrm{d}t + \mathrm{d}W(t), \tag{4.22}$$

where $\mathrm{d}W$ is a Wiener process with zero mean and variance $\mathrm{d}W^2 = \mathrm{d}t$. The record models the outcome of a *homodyne* detection of the fluorescence of the qubit for a known number of photons n in the cavity, with a quantum efficiency of η. The intermediate frequency is $\omega_{IF}/2\pi = 66$ MHz.

The orange curve of Fig. 4a results from calculating 10 000 quantum trajectories solutions of these SME for each amplitude and frequency, for $\eta = 0.17$ and 0.2 and applied the same procedure as for the experimental records for $V(t) = \sqrt{G}\mathrm{d}y_t/\mathrm{d}t$. The only difference is that the average signal $\overline{V}$ was directly obtained as the solution $\hat{\rho}(t)$ to the Lindblad equations, given by Eq. (4.18) when $\eta = 0$, and computing $\sqrt{\frac{\eta G}{T_q}}\,\mathrm{Tr}(\hat{\sigma}_x\hat{\rho}(t))$.

4.5.7 *Bayesian filter for the photon number*

Tracking the photon number corresponds to actuating the probability of having n photons in the cavity $\mathbb{P}_t(n)$ conditioned on the measurement record V from time 0 to t. Given $\mathbb{P}_{j\tau}(n)$ and the measured $\vec{m}(j\tau+\tau)$, we compute $\mathbb{P}_{j\tau+\tau}(n)$ as follows:

- During $\Delta t = 2$ μs, apply the dissipation operator $1/T_c\mathbb{D}_{\hat{a}}$ on the diagonal density matrix of the cavity, whose diagonal elements are the $\mathbb{P}_{j\tau}(n)$ coefficients. We denote the resulting diagonal elements of the density matrix as $\mathbb{P}'_{j\tau}(n)$, and use a cavity lifetime $T_c = 200$ μs,
- Actuate it with the corresponding $\vec{m}(j\tau+\tau)$ following

$$\mathbb{P}_{j\tau+\tau}(n) = \mathbb{P}'_{j\tau}(n)\mathbb{P}(\vec{m}(t)|n)/Z, \tag{4.23}$$

where Z is a normalization constant ensuring $\sum_n \mathbb{P}_{j\tau+\tau}(n) = 1$.

4.5.8 *Average time to reach a given confidence on the photon number measurement*

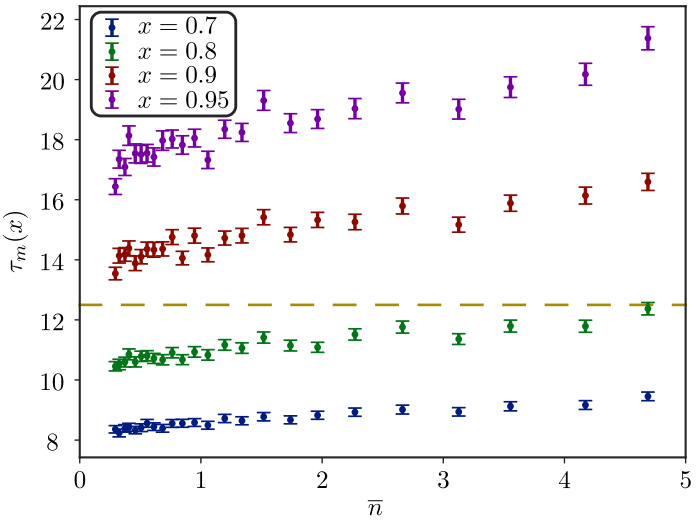


Figure 4.14: Average time needed to reach various confidence thresholds as a function of the average number of photons. Dots: average measurement time $\tau_m(x)$ required so that one of the photon number probabilities $P_n(t)$ exceeds the confidence $x = 70$ % (blue), 80 % (green), 90 % (red) or 95 % (purple). The times are represented as a function of the mean photon number at time 0. Error bars represent the uncertainty coming from the finite number 1100 of realizations that were processed. Dashed line: Measured $1/\Gamma_m$ under the same conditions of driving ($\theta = \pi/2$ as in Fig. 3).

In addition to the measurement rate, it is possible to extract the average time it takes for the measurement procedure to reach a certain confidence $x < 1$ in the number of photons (Fig. 4.14). Using the same set of measurement records from which the realization in Fig. 3 is taken, and starting at an arbitrary time t_0 with an agnostic probability vector $\mathbb{P}_{t_0}(n) = 1/10$, we process the data with the Bayesian update *without including the dissipation*, until the time $t_0 + \tau$ at which the estimator reaches a confidence x for one of the numbers of photons.

We then define $\tau_m(x) = \overline{\tau}(x)$, the average time that is required to reach the level of confidence x. Using the average expected number of photons $\overline{n}$ (see Fig.4.17) as a function of time t_0, we can plot the average time $\tau_m(x)$ as a function of the mean initial photon number $\bar{n}$ in Fig. 4.14. Note that $\tau_m(x)$ differs from $1/\Gamma_m$, which is the time needed to reach a fidelity of 92% (Signal-to-Noise Ratio = 2) when thresholding for a standard qubit readout without decay. But these two timescales give the same order of magnitude for the measurement time.

The increase of τ_m as a function of the average number of photons can be understood as the fact that the effect of the dissipation increases with the number of photons, so the cavity is increasingly more likely to lose a photon during the measurement time, which increases the time needed by the estimator to converge, as it has to rebuild confidence when a photon is lost.

4.6 DEPHASING RATE

4.6.1 *Measuring the dephasing rate*

To extract the dephasing rate of the cavity in Fig. 4b, we use the technique of Ref. [126].

Performing Wigner tomography via the tomography qubit, we extract $|\langle\hat{a}(t)\rangle|$ after preparing a coherent state $\langle\hat{a}(0)\rangle = -1.11$ in the cavity. In order to cover the full range of dephasing rates, we perform the tomography for multiples of $\pi/\chi \simeq 95$ ns (period of the comb), with an uneven spacing to cover fast and slow dynamics alike. An example of such dynamics with an exponentially decaying amplitude is shown Fig. 4.15a. We use an exponential fit to extract the dephasing rate γ_ϕ. Note that, at small qubit driving amplitude, the decay of $|\langle\hat{a}\rangle|$ is driven by the self-Kerr effect leading to a non-exponential behavior. Fig. 4.15d shows the measured Wigner function without drive, 15.2 μs after the displacement.

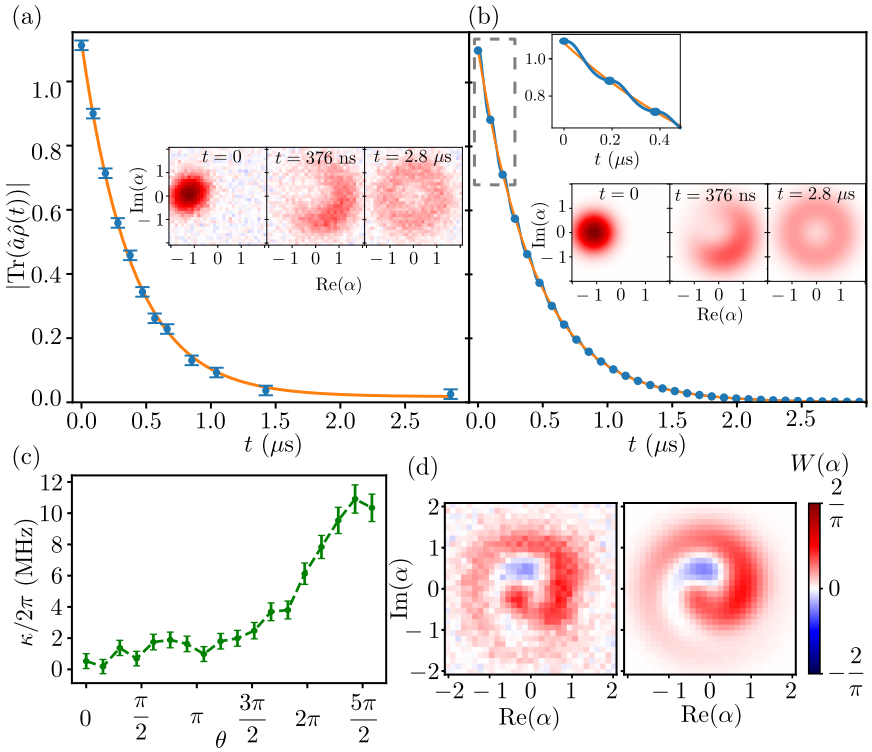


Figure 4.15: Dephasing rate measurement and simulation. (a) Blue dots: measured $|\,\mathrm{Tr}\{\hat{a}\hat{\rho}(t)\}|$ as a function of time for a comb amplitude leading to a qubit kick angle $\theta = \pi$. Solid orange line: fitted exponential decay at a rate γ_ϕ. Inset: timeframes of the measured Wigner functions. (b) Blue line and dots: simulated $|\,\mathrm{Tr}\{\hat{a}\hat{\rho}(t)\}|$ as a function of time for $\theta = \pi$ using Eq. (4.26). Solid orange line: fitted exponential decay from which the predicted dephasing rate is extracted. Inset: timeframes of the simulated Wigner functions. A zoom on the earliest times shows the dephasing dynamics during a period of the comb drive. (c) Decay rate κ of the cavity extracted from the decay of the mean photon number Eq. (4.24). (d) Measured (left) and simulated (right) Wigner functions at time $t = 15.2$ μs without driving the qubit. The simulation is performed using Eq. (4.28) where the self-Kerr rate is $\chi_{cc}/2\pi = 10 \pm 1$ kHz, and pure dephasing rate is $\frac{1}{T_{c,\phi}} = (36 \pm 6\ \mu\mathrm{s})^{-1}$.

,

If the measurement was QND, we could obtain the measurement induced dephasing rate by subtracting the measured dephasing rate at zero drive $1/T_{c,\phi}$ to the dephasing rate γ_ϕ at finite drive. However, as explained below, we observe a yet unexplained increase in the cavity decay rate for large qubit drive that forbids such a simple approach. We therefore proceed in two steps.

First, to obtain pure dephasing rates $\Gamma_\phi(\theta)$, we subtract the contribution related to the decay rate κ of the cavity to the measured γ. This decay rate is determined using the tomography at various times. For each measured Wigner function $W_{\hat{\rho}(t)}(\alpha)$, we compute the mean photon number

$$\overline{n}(t) = \mathrm{Tr}\Big\{\hat{a}^\dagger \hat{a}\hat{\rho}(t)\Big\} \simeq \sum_{k=0}^{7} k\mathbb{P}_k(\hat{\rho}(t)), \tag{4.24}$$

where $\mathbb{P}_k(\hat{\rho}) = \int W_{\hat{\rho}}(\alpha) W_{|k\rangle\langle k|}(\alpha)\mathrm{d}^2\alpha$ with $W_{|k\rangle\langle k|}(\alpha)$ the Wigner map of $|k\rangle\langle k|$ (see (B4) in [126]). The resulting fitted decay rate κ is shown in Fig. 4.15c as a function of the qubit kick angle θ.

Second, in order to extract the measurement induced dephasing rate Γ_d plotted in Fig. 4b, we subtract the measured pure dephasing rate without qubit driving $1/T_{c,\phi} = (36 \pm 6\ \mu\mathrm{s})^{-1}$ to the pure dephasing rate $\Gamma_\phi(\theta)$.

For clarity, the final expression for $\Gamma_d(\theta)$ reads

$$\Gamma_d(\theta) = \gamma(\theta) - \kappa/2 - 1/T_{c,\phi}. \tag{4.25}$$

4.6.2 *Simulation of the dephasing rates*

To predict the expected dephasing rate, we use the Lindblad equation of the qubit-cavity coupled system in the interaction picture

$$\frac{\mathrm{d}\hat{\rho}(t)}{\mathrm{d}t} = -\frac{i}{\hbar}[\hat{H}_{\mathrm{int}} + \hat{H}_{\mathrm{d}}, \hat{\rho}] + \frac{1}{T_q}\mathbb{D}_{\hat{\sigma}_-}(\hat{\rho}(t)), \tag{4.26}$$

with

$$\hat{H}_{\mathrm{int}} = -\hbar\frac{\chi}{2}(\mathbb{1} + \hat{\sigma}_z)\hat{a}^\dagger\hat{a}, \tag{4.27}$$

and $\hat{H}_d$ defined in Eq. (4.1).

Here, we neglect any other dephasing source, as well as the decay rate of the cavity. We use this simulation to compute $\mathrm{Tr}\{\hat{a}\hat{\rho}(t)\}$. Taking only one point per period of the comb. It fits well with an exponential decay (see Fig. 4.15b).

The dephasing rate at zero drive $1/T_{c,\phi}$ can be precisely determined by fitting the dynamics of the cavity state when starting in a coherent state. The Linblad Master equation we use takes into account the self-Kerr effect as well as pure dephasing without drive:

$$\frac{\mathrm{d}\hat{\rho}(t)}{\mathrm{d}t} = -\frac{i}{\hbar}[\hat{H}_{\mathrm{kerr}}, \hat{\rho}] + \frac{2}{T_{c,\phi}}\mathbb{D}_{\hat{a}^\dagger\hat{a}}(\hat{\rho}(t)), \tag{4.28}$$

where

$$\hat{H}_{\mathrm{kerr}} = -\hbar\chi_{cc}\hat{a}^{\dagger 2}\hat{a}^2. \tag{4.29}$$

We find that the measured evolution of the cavity state is well reproduced when $\chi_{cc}/2\pi = 10 \pm 1$ kHz and $1/T_{c,\phi} = (36 \pm 6\ \mu\mathrm{s})^{-1}$ (see Fig. 4.15d).

4.6.3 *Predicted dephasing rate*

We start with the same model as in Sec. 4.4 where the qubit is kicked once and has infinite time to decay. We can model the experiment as follows. Irrespective of the state of the cavity, the qubit is prepared in a state $|\psi\rangle = \cos\theta/2\,|0\rangle + e^{i\phi}\sin\theta/2\,|1\rangle$. It emits a state $|\phi\rangle_n$ in the transmission line that depends on the number of photons n, which we can write

$$|\phi\rangle_n = \Big(\cos(\theta/2)\mathbb{1} + e^{i\phi}\sin(\theta/2)\sqrt{\frac{1}{T_q}}\int_0^{+\infty} e^{-i(\omega_q - n\chi)t - \frac{t}{2T_q}}\hat{a}^\dagger_{\mathrm{out}}(t)\mathrm{d}t\Big)|\mathrm{vac}\rangle \tag{4.30}$$

where $|\mathrm{vac}\rangle$ is the vacuum state of the outgoing transmission line. In the simple case where $n = 0$ or $n = 1$, the cavity can be treated as a qubit as in Ref. [40]. Each period $\tau = \frac{\pi}{\chi}$ of the train of pulses leads to a decay of the off-diagonal matrix elements of the cavity density

matrix by a factor $|\,{}_0\langle\phi|\phi\rangle_1\,|$. The dephasing rate Γ_d (black solid line in Fig. 4b) is thus given by

$$e^{-\Gamma_d \tau} = |\,{}_0\langle\phi|\phi\rangle_1\,|$$

where

$${}_0\langle\phi|\phi\rangle_1 = \cos^2(\theta/2) + \frac{\sin^2\theta/2}{1 - i\chi T_q}.$$

4.7 DEVIATION TO QUANTUM NONDEMOLITION

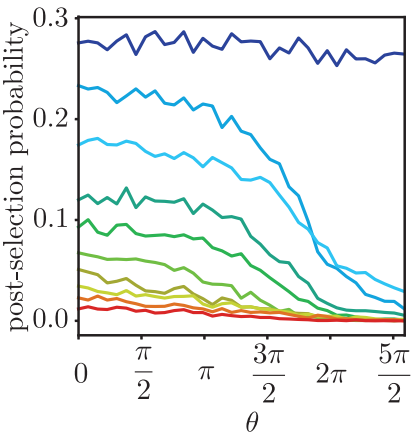


Figure 4.16: Measurement-induced decay rate of the cavity. Probability of measuring Re(r) below threshold during the post-selection pulse of the sequence in Fig. 4.8a as a function of the qubit rotation angle θ associated to the comb. Each line corresponds to a different selected photon number n from 0 (dark blue) to 9 (red). The selection thresholds are the same as in Table 4.3

.

In order to better characterize how the driving amplitude of the qubit affects the number of photons in the cavity (deviation to QNDness of the photon number tracking), we perform several experiments. The drive amplitude is given either as a Rabi frequency Ω for single tone driving or an angle of rotation θ during a qubit kick.

A direct signature of non-QNDness can be seen in Fig. 4.16. It shows the yield of the post-selection for each n (each color) as a

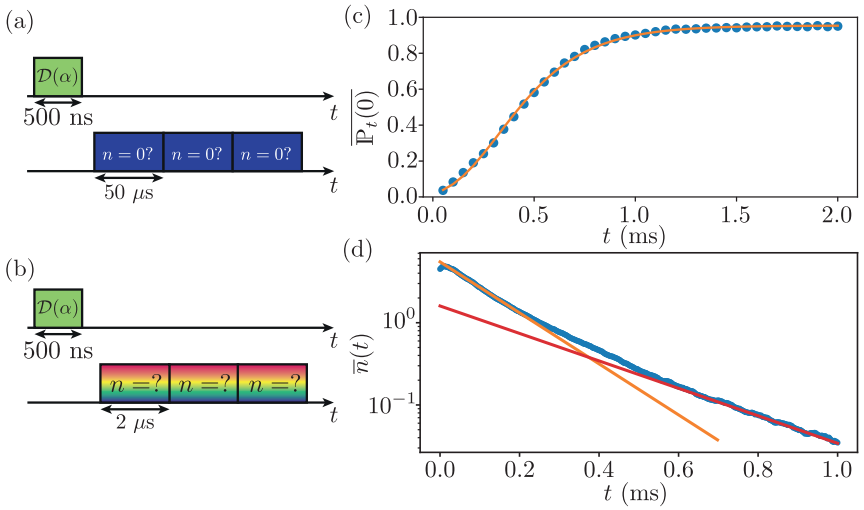


Figure 4.17: Measurements of the decay rate of the cavity. (a) Pulse sequence tracking the occupation of the vacuum state. (b) Pulse sequence tracking the number of photons. (c) Blue dots: probability $\overline{\mathbb{P}_t(n=0)}$ that the cavity is in vacuum as a function of time. The measurement is obtained by single-tone readout of the number of photons after a strong ($|\alpha|^2 > 20$) displacement of the memory. The fit (orange) to Eq. (4.31) gives $T_c = 230$ μs. The absence of a plateau at early times despite the large displacement could be attributed to an increasing decay rate for higher number of photons. (d) Blue dots: reconstructed average number of photons using 1100 records of the tracked photon number. Exponential fits with $T_c = 130$ μs (orange) and $T_c = 250$ μs (red) are superimposed.

function of the rotation angle θ. The yield decreases with driving amplitude, showing that driving the qubit changes the transition rates between Fock states. The variation is more pronounced for larger driving amplitudes ($\theta > \pi$). As a matter of fact, the curves for $n = 1$ and $n = 2$ cross, which indicates a decay rate that changes with n.

In Fig. 4.15c, we show the observed cavity decay rate κ as a function of the angle θ. It is obtained by computing the number of photons as a function of time from the Wigner tomography. A small increase of κ seems to occur as θ grows but still an order of magnitude smaller than the measurement rate for $\theta = \pi/2$. However a sharp rise in κ occurs for $\theta \approx 3\pi/2$.

4.8 DEPENDENCE OF THE CAVITY LIFETIME ON PHOTON NUMBER

We also performed a measurement of the average probability of having 0 photons $\overline{\mathbb{P}_t(n=0)}$ as a function of time (Fig. 4.17a,c). This is obtained with a single frequency tone on the qubit at $\Omega = 5.4$ MHz after a cavity displacement. Fitting it with the decay of a coherent state

$$\overline{\mathbb{P}_t(n=0)} = |\langle\sqrt{n_0}e^{-t/T_c}|0\rangle|^2 = e^{-n_0 e^{-t/T_c}} \tag{4.31}$$

extracts the decay time T_c of one photon in the cavity.

We do not observe a plateau at initial times (small effective n_0 in Eq. (4.31)) despite the large variation of displacement amplitudes we tested. We believe that this could be due to a nonlinear decay rate of the cavity owing the presence of the transmon and its Purcell filter.

Using the quantum trajectories obtained with the comb drive, we can also reconstruct the average number of photons as a function of time. It does not decay exponentially. This non exponential cavity decay was also observed using the tomography qubit on a different run. As a side remark, the slight increase of the estimated mean number of photons at small times is well understood. Indeed the comb signal is only analyzed within the range from 0 to 9 photons so that a wrong prior estimates the average number of photons at 4.5 initially. As the number of photons decreases and enters the detecting window $[\![0, 9]\!]$, the update tends to increase the average number of photons at short times.

We do not have a simple explanation for this behaviour. We believe that higher order conversion processes between the cavity and the multiplexing qubit could be involved in these observations. As a matter of fact, the frequency of the qubit is higher than that of the cavity, which could more easily trigger the effects described in [161]. However, it would be in a very unusual regime, as the cavity has a long decay time and the qubit a low one in our work.

In a previous cool down of the same device, we measured the decay of the cavity for a wide range of numbers of photons using the tomography qubit. The procedure consists in first displacing the cavity to a coherent state with about 25 average photons. The tomography

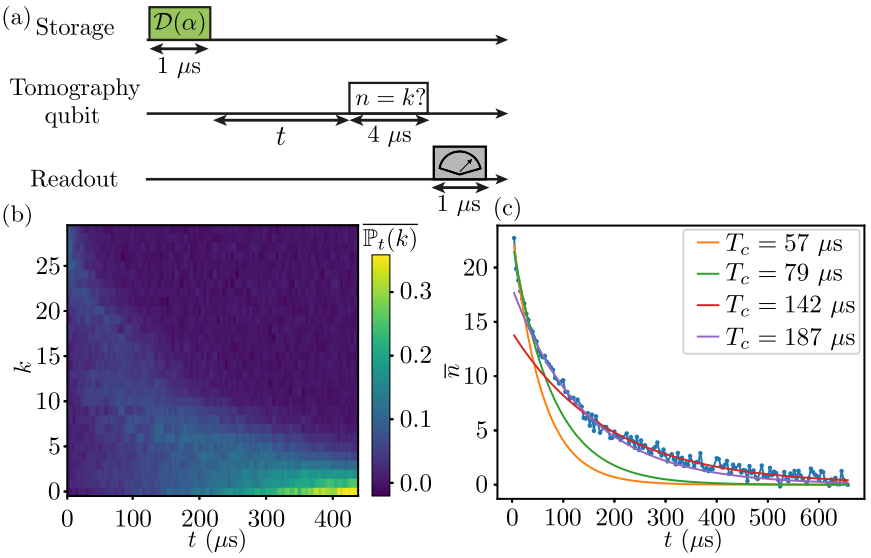


Figure 4.18: Measurement of the average decay of the cavity with the tomograhy qubit. (a) Pulse sequence for the measurement of the probability $\overline{\mathbb{P}_t(k)}$ that there are k photons in the resonator. (b) Measured $\overline{\mathbb{P}_t(k)}$ as a function of time t and photon number k after a displacement pulse of about 25 mean photons. (c) Measured decay of the average photon number $\overline{n} = \sum_{k=0}^{30} k\overline{\mathbb{P}_t(k)}$ as a function of time.

qubit is then driven with a selective π pulse (at a frequency $\omega_t - k\chi_{ct}$, see Table 4.1), conditioned on the number of photons being equal to a certain k. The state of the tomography qubit is then read out using its readout resonator. From this, we can reconstruct the average occupation of each Fock state $\overline{\mathbb{P}_t(k)}$ in the cavity at any time t after the displacement (Fig. 4.18b). The behaviour of $\overline{\mathbb{P}_t(k)}$ is non exponential for high number of photons, and seems to drop faster for more than 10 photons. Looking at the average number of photons on Fig. 4.18b, we fit the curve portions with several exponential functions, and see that the decay rate of the cavity increases with the average number of photons.

4.9 SUPPLEMENTARY SECTIONS

4.9.1 *Wigner function of prepared Fock states*

In a previous cooldown of the experiment, we could perform a Wigner tomography of the pre-selected Fock states using the procedure described in Sec. 4.3.2. For each number of photons between 0 and 9, the heralding sequence shown in Fig. 4.8a was performed, before a displaced parity measurement sequence to reconstruct the Wigner function of the post-selected state. The resulting Wigner functions are shown in Fig. 4.19.

4.9.2 *Neural-network-assisted record analysis*

One of the potential follow-up ideas of the project was to use a neural network embedded in a Field-Programmable Gate Array (FPGA) to analyze the time traces and to perform low-latency feedback control on the cavity to prepare stabilize Fock states, as in [125]. Such neural network-based feedback control was already demonstrated using weak measurements of a qubit [166, 167]. However, this necessitates specialized hardware that was not available at this time. Still, we could perform a few prospective tests without this hardware. It was already demonstrated that a neural network could reconstruct quantum trajectories of a qubit under a weak measurement [168, 169]. In this section, we will demonstrate that a neural network can be trained to analyze the fluorescence record of the qubit driven by a comb to build the vector of probabilities in Fig. 4.3 for a time trace of 2 μs.

To do so, we generated $N_{\text{samples}} = 608000$ stochastic records in the same conditions as in Sec. 4.5.6 for each possible qubit frequency associated to n between 0 and 9. The quantum efficiency is set to $\eta = 0.15$ (slightly lower than in the experiment) and the decay rate of the qubit $1/(2\pi T_q) = 6$ MHz, which corresponds to $T_q = 27$ ns, (slightly longer than in the experiment). For each trace generated for a number n of photons, the actual photon number distribution is $\mathbb{P}^n_{\text{label}}(k) = \delta_{kn}$. This distribution is used as a target for the training of the NN.

This approach addresses photon counting as a classification problem. The network learns to classify the traces into 10 categories, each corresponding to a photon number. It then falls in the category of

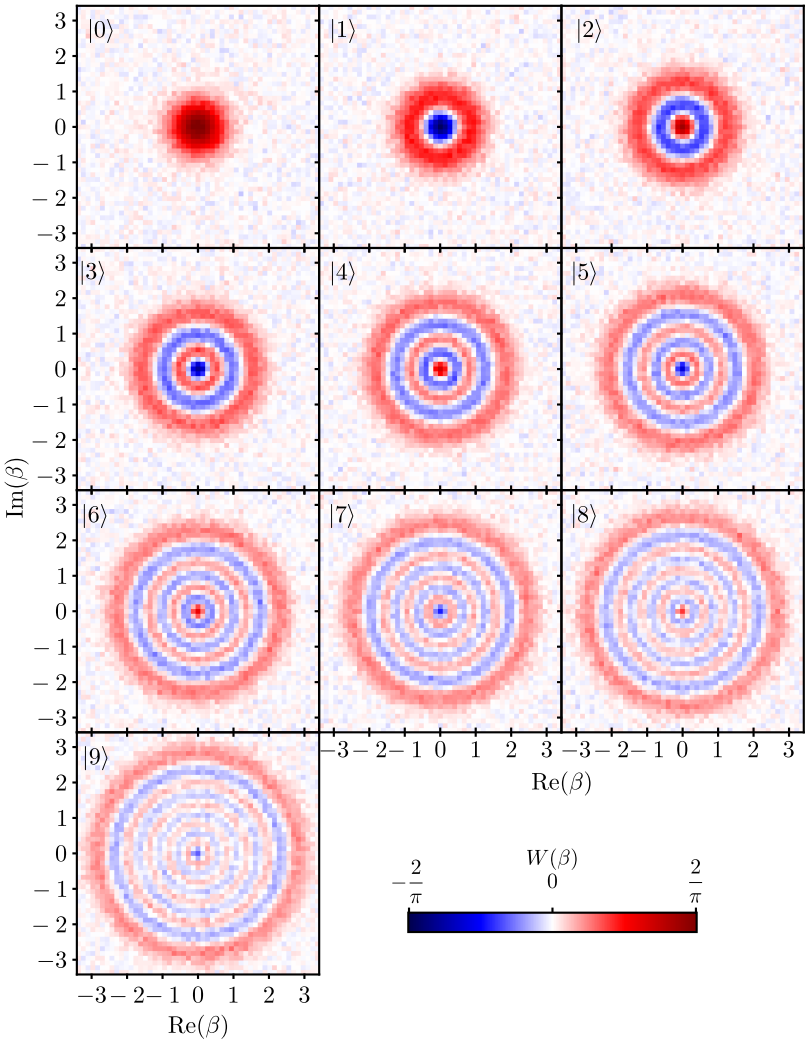


Figure 4.19: Measured Wigner functions of Fock states $|n\rangle$ for n between 0 and 9, after a heralding sequence on heterodyne measurement of the qubit emission.

supervised learning, as the network will be presented with the expected answer during the training. Classification is a common task in machine learning, especially in image recognition. It also applies to quantum tasks, for example, to detect non-classical states [170, 171] and entanglement [172].

The network is the simplest possible network we could build: it is constituted of a single fully-connected layer, linking the 2001 input time values to each of the 10 outputs constituting the output of the network. The chosen activation function is the softmax function so that the output of the network is a proper probability distribution. It is build using the Keras package [173] of the library Tensorflow [174].

Keeping $N_{\text{test}} = 10000$ trajectories for each n to evaluate the network after the training, we trained the network for 12 epochs, using an Adam optimizer [175] and the categorical cross-entropy as a loss function between the distribution $\mathbb{P}_{\text{NN}}$ outputted by the network and $\mathbb{P}_{\text{label}}$. Because of the very high number of samples, the network converged after only 2 epochs.

The network is then tested on a set $\{\mathrm{d}y_t^{j,n}\}$ of $N_{\text{test}} \times 10$ homodyne records where $n \in [\![0, 9]\!]$ stands for the number of photons prepared and $j \in [\![0, N_{\text{test}} - 1]\!]$ is the sample number. Each of these records is then associated with the distributions $\mathbb{P}^{j,n}_{\text{demod}}(k)$ obtained from the demodulation procedure described in this chapter, and $\mathbb{P}^{j,n}_{\text{NN}}(k)$ from the trained neural network. The result on a subset of the test data is shown in Fig. 4.20. We see that both the demodulation procedure (Fig. 4.20)a and the network (Fig. 4.20b) seem to extract some information about the photon number, as we can distinguish the pattern of Fig. 4.20c. Note that it does not manage to reach the precision of the targets $\mathbb{P}^{j,n}_{\text{label}}(k) = \delta_{kn}$ presented in Fig. 4.20c as the information in the record is intrinsically not sufficient to reach this level of precision.

Strikingly, we see that most features in Fig. 4.20a are replicated in Fig. 4.20b, even though they appear random. It indicates that the predictions $\mathbb{P}_{\text{demod}}$ and $\mathbb{P}_{\text{NN}}$ are very similar. This was expected from the fact that this analysis is near-optimal (see Sec. 5.1.1). This seems to indicate that the network finds a similar analysis, which we could have investigated by trying to "open the black box". We have two possible ideas at this stage: looking at the weights of the network to compare them to the demodulation functions, and trying to find out which inputs allow to obtain the most peaked functions, and compare to the known structure of the pulses.

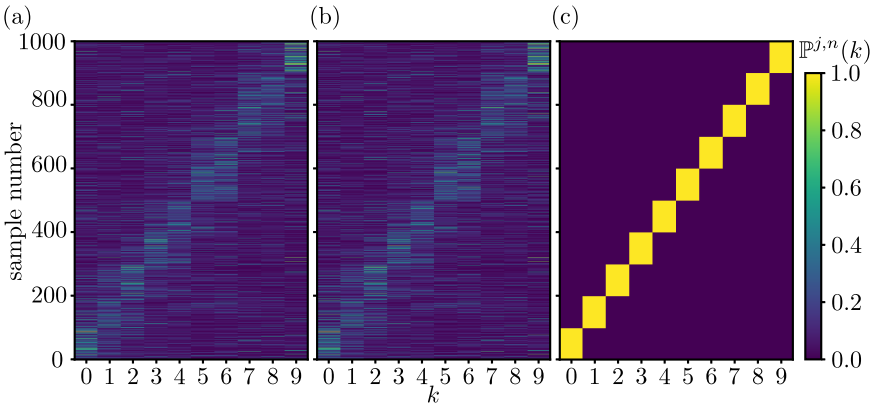


Figure 4.20: Comparison between the predictions obtained using the neural network and the demodulation procedure. (a) Prediction $\mathbb{P}_{\text{demod}}(k)$ from the demodulation-based analysis of the heterodyne measurement record. (b) Prediction $\mathbb{P}_{\text{NN}}(k)$ from the neural network-based analysis of the heterodyne measurement record. (c) Label distribution $\mathbb{P}_{\text{label}}$, indicating the cavity photon number for each trace. The sample number is linked to n and j by the relations sample number $= 10n + j$ and $j =$ sample number mod 10.

To illustrate how close to each other our two predictors are, we compare the average probabilities

$$\overline{\mathbb{P}}^{n}_{\text{demod}}(k) = \frac{1}{N_{\text{test}}} \sum_{j} \mathbb{P}^{j,n}_{\text{demod}}(k) \tag{4.32}$$

$$\overline{\mathbb{P}}^{n}_{\text{NN}}(k) = \frac{1}{N_{\text{test}}} \sum_{j} \mathbb{P}^{j,n}_{\text{NN}}(k), \tag{4.33}$$

and plotted them in Fig. 4.21: they are barely distinguishable.

To push the comparison further, we compute the prediction fidelities of the two methods. We translate each photon number distribution $\mathbb{P}$ to a photon number prediction (by taking the argmax of the distribution), and compute the probability that this prediction is right, *i.e.* that the trace was generated with this parameter. Defining

$$r^{j,n}_{\text{demod}} = \begin{cases} 1 \text{ if } \underset{k}{\text{argmax}}(\mathbb{P}^{i,n}_{\text{demod}}) = n \\ 0 \text{ otherwise,} \end{cases} \tag{4.34}$$

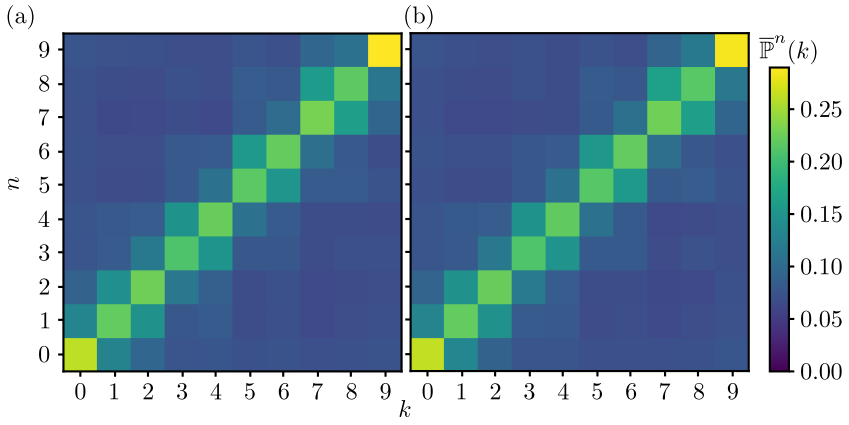


Figure 4.21: Comparison between the average predictions obtained using the neural network and the demodulation procedure. (a) Average predictor probability $\mathbb{P}^n_{\text{demod}}(k)$ for each photon number obtained with the demodulation technique. (b) Average predictor probability $\mathbb{P}^n_{\text{NN}}(k)$ for each photon number obtained with the neural network.

and $r^{j,n}_{\text{NN}}$ in the same way, the estimated fidelity $\mathcal{F}^{j,n}_{\text{demod}}$ when there are n photons reads

$$\mathcal{F}^n_{\text{demod}} = \frac{10}{N_{\text{test}}} \sum_j r^{j,n}. \tag{4.35}$$

The fidelity for the neural network $\mathcal{F}^n_{\text{NN}}$ is defined similarly with the $\{r^{j,n}_{\text{NN}}\}$. The result is shown in Fig. 4.22a. Both schemes are well above 1/10 fidelity, which shows their predictive power. We see that both methods yield very similar results, the demodulation method performing slightly better than the neural network for each n. The better fidelity for $n = 0$ and $n = 9$ can simply be explained by the fact that these two states have only one close neighbor, on contrary to all the others, which have two of them ($n - 1$ and $n + 1$).

The two predictors, however, have slightly different behaviors when trying to extract a mutual information between the number of photons k and the outcome of each predictor. For the demodulation-based predictor, we can write this mutual information I_{demod} as

$$I_{\text{demod}} = H_0 - \frac{1}{10} \sum_n H^n, \tag{4.36}$$

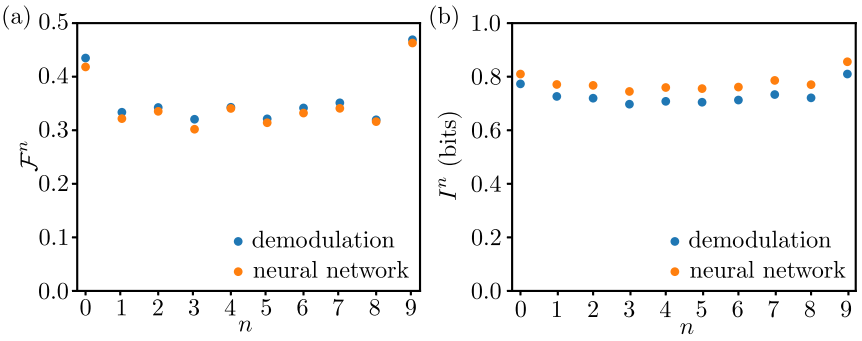


Figure 4.22: Fidelity and mutual information obtained using the neural network and the demodulation procedure. (a) Fidelity obtained for the demodulation technique and the neural network for each number of photons. (b) Mutual information, expressed in bits, inferred from the statistics of the predictions using the demodulation technique and the neural network.

where H_0 is the entropy about the photon number, set to $\log(10)$ as there are 10 possible outcomes with equal probabilities. H^n is the conditional entropy of the predictor knowing that the photon number is n. This mutual information characterizes how much the entropy about the photon number is reduced thanks to the predictor.

The mutual information I_{demod} thus decomposes into mutual informations I^n_{demod} conditioned on the photon number being n

$$I_{\text{demod}} = \frac{1}{10}\sum_n I^n_{\text{demod}}, \tag{4.37}$$

where

$$I^n_{\text{demod}} = H_0 - H^n. \tag{4.38}$$

We can estimate H^n as

$$H^n = -\frac{10}{N_{\text{test}}}\sum_j \sum_k \mathbb{P}^{j,n}_{\text{demod}}(k)\log\Big(\mathbb{P}^{j,n}_{\text{demod}}(k)\Big). \tag{4.39}$$

The same procedure is used to compute I^n_{NN} and I_{demod} from $\mathbb{P}^{j,n}_{\text{demod}}(k)$.

The result is shown in Fig. 4.22b. Strikingly, the neural network exhibits significantly higher mutual information than the demodulation

process. We have to be careful with this observation, as a higher mutual information, as computed, does not necessarily implies that the network extracts more information. It could for example output very peaked distributions that are completely wrong.

We interpret this higher apparent mutual information by the fact that the neural network is trained to reach the very peaked distribution $\mathbb{P}_{\text{label}}$ even though it does not have access to enough information to reach it. It could be that it is encouraged during the learning part to output peaked distributions, even though they do not necessarily reflect the information that could be extracted from the raw measurement record.

We thus demonstrated that neural networks can be used to process such time traces to readout the state of the cavity. This was done in a model-free setting, *i.e.* without having access to the model of the experiment during the training. This principle could for example be applied to process readout traces in a superconducting quantum computer. Indeed, in processors with a lot of qubits, optimizing the space on the chip is crucial. There is a tradeoff between the density of elements and the parasitic couplings (crosstalks between them), which renders the readout task more and more complex with the number of qubits. Machine-learning-based solutions were already proposed to adress this type of issue [176–178]. Such a model-free approach is peculiarily useful in these cases where the model may not be known with precision, or would require measuring a prohibitive amount of parameters.

5

INFORMATION AND MEASUREMENT RATES

In Ref. [126], Antoine Essig and coworkers already measured the dephasing rate in the first version of the experiment presented in Chap. 4. They also introduced the idea of using an infinite frequency comb and theoretically studied the accessible information as a function of time after one relaxation of the qubit (in the case $\theta = \pi$) in the number-resolved regime ($\chi \gg \Gamma_1$). In this case, they demonstrated that the measurement time could be independent of the maximal number of photons $N_{\max}$ with an appropriate measurement setup. They also built a model of the measurement records obtained with a heterodyne detector and established that, under a Gaussian approximation of the record statistics, the measurement time is proportional to $\log(N_{\max})$, where $N_{\max}$ is the maximal number of photons allowed in the cavity.

The goal of this chapter is to carry further the analysis of the information content of the fluorescence emitted by the qubit. In Fig 4.4, we compare the measurement rate obtained experimentally to the dephasing rate (Fig 4.4b orange and black) and to the maximal measurement rate that a heterodyne measurement could reach. We will first develop the theoretical framework that allowed us to produce the latter curve. In a second part, we will compare the heterodyne scheme used in the experiment to an optimal scheme maximizing the measurement rate and to a scheme using a Frequency-Resolved Photon Detector (FRPD), which is close to the original Gedanken experiment [126]. In a third part, we will generalize the notion of measurement rate to take into account the fact that our photon counter has to discriminate between more than two possible outcomes. Finally, we will transform the problem of discriminating between discrete frequencies into a continuous frequency estimation problem and make a theoretical comparison between heterodyne, homodyne, FRPD, and optimal measurements. Interestingly, we will establish cases where the homodyne measurement performs infinitely worse than a heterodyne detection and FRPD.

5.1 MODELING THE EXPERIMENT

5.1.1 *Heterodyne measurement*

As explained in Sec. 3.3.3, in order to derive a measurement rate for heterodyne detection of the propagating field, we need to compute the mutual information $I(k:x)$ between the number of photons k and the heterodyne record x.

Let us consider the following communication scheme, pictured in Fig. 5.1a. Alice communicates the word $k \in [\![0, m-1]\!]$ with probability $\mathbb{P}(k)$ to Bob by preparing the outgoing mode represented by functions f_k in the state $\hat{\rho}_k$, and Bob performs a heterodyne measurement on it, which gives him a record x. We are in the context developed in Appendix A.1.1, and we can use the formula given by Eq. (A.2) applied to $a = k$ and $b = x$ to compute $I(k:x)$. This is a more general case than for the experiment of Chap. 4, but it is handled just the same. We write $I(k:x)$ as a functional integral over all the possible measurement records x

$$\begin{aligned} I(k:x) = &- \int p(x) \log p(x) \mathcal{D}x \\ &+ \sum_k \mathbb{P}(k) \int p_{f_k}(x) \log p_{f_k}(x) \mathcal{D}x, \end{aligned} \tag{5.1}$$

where $p_{f_k}(x)$ is the conditional probability of finding the measurement record x knowing that the state represented by f_k was chosen and

$$p(x) = \sum_k \mathbb{P}(k) p_{f_k}(x) \tag{5.2}$$

is the overall probability of finding the measurement record x. We can write

$$\begin{aligned} I(k:x) = &- \int \Big(\sum_k \mathbb{P}(k) p_{f_k}(x) \Big) \log \left(\sum_k \mathbb{P}(k) p_{f_k}(x) \right) \mathcal{D}x \\ &+ \sum_k \mathbb{P}(k) \int p_{f_k}(x) \log(p_{f_k}(x)) \mathcal{D}x. \end{aligned} \tag{5.3}$$

Factoring out the $\mathbb{P}(k)$ gives

$$I(k:x) = \sum_k \mathbb{P}(k) \left[- \int p_{f_k}(x) \log \left(\sum_l \mathbb{P}(l) p_{f_l}(x) \right) \mathcal{D}x + \int p_{f_k}(x) \log(p_{f_k}(x)) \mathcal{D}x \right] \tag{5.4}$$

And finally, assembling the log functions in the integral:

$$I(k:x) = \sum_k \mathbb{P}(k) \left[- \int p_{f_k}(x) \log \left(\sum_l \mathbb{P}(l) \frac{p_{f_l}(x)}{p_{f_k}(x)} \right) \mathcal{D}x \right]. \tag{5.5}$$

This expression reads more simply in terms of expectation values:

$$I(k:x) = \sum_k \mathbb{P}(k) \mathbb{E} \left[- \log \left(\sum_l \mathbb{P}(l) \frac{p_{f_l}(x)}{p_{f_k}(x)} \right) \middle| k \right], \tag{5.6}$$

where $\mathbb{E}[X|k]$ stands for the expectation value of X conditioned on the choice k of Alice. We now only need the probability ratios p_{f_l}/p_{f_k}, which can be obtained from a SME. This enables a Monte-Carlo approach to compute $I(k:x)$: we could generate measurement records and compute their likelihood step by step from the SME and estimate the mutual information from it. What we are going to show here is that there is a way to do the same kind of estimation but by directly generating integrated measurement records, which is much more efficient than whole stochastic trajectories. This approach is less general than the SME, as it only works for single modes (even though we believe that generalizing this derivation to a finite number of modes is possible). This allowed us to compute this mutual information for a wide range of parameters.

Let the outgoing field be in the state $|\psi_{\text{out}}\rangle = \sum_l \psi_l \hat{A}_f^{\dagger l} |\text{vac}\rangle = |\psi\rangle_f \otimes |\text{vac}\rangle$, where f is a normalized element of $L^2(\mathbb{R}, \mathbb{C})$, and $x(t)$ be the complex heterodyne record. Formally, the probability $p_f(x)$ to measure this measurement record would read:

$$p_f(x) = \prod_t p_f(x(t) | \{x(t' < t)\}) \tag{5.7}$$

where $p_f(x(t)|\{x(t' < t)\})$ is the probability to measure $x(t)$ at time t knowing all the previous results at times $t' < t$. This expression is exact if the number of time values considered is finite, which is true in practice: the measurement record $x(t)$ at time t is integrated over the inverse bandwidth δt of the detector. In the following, we will stick to integral expressions for convenience, but they have to be understood as discrete sums in the limit where the time step goes to zero and the time extent to infinity. In this setting, we consider that the time resolution δt is small enough and the time extent of the measurement large enough so that f is well approximated by its discrete version.

We can decompose x against an orthonormal basis $\{g_k\}_k$ of complex functions such that $g_0 = f$ by demodulating against them. It gives the integrated records x_{g_k} that read

$$x_{g_k} = \int g_k^*(t)x(t)\mathrm{d}t. \tag{5.8}$$

Note here that this basis is finite of cardinal n for a finite number n of time values.

We saw in Sec. 3.1.2 that x_{g_k} follows the law given by the Husimi Q function of the mode represented by g_k. As in this mode basis, the state is a product state, the x_{g_k} are statistically independent, following the law given by the Husimi Q function of the k-th mode $Q_{\hat{\rho}_k,\hat{A}_{g_k}}$. To simplify the notation in this chapter, we will note Q_f the Q function of the mode defined by f in the state $\hat{\rho}$. The probability $p_f(x)$ thus reads

$$p_f(x) = \prod_{k\geq 0} Q_{g_k}(x_{g_k}) = Q_f(x_f) \prod_{k\geq 1} Q_{|0\rangle}(x_{g_k}) \tag{5.9}$$

with $Q_{|0\rangle}(x_{g_k}) = \frac{1}{\pi}e^{-|x_{g_k}|^2}$ the Q function of the vacuum state (we assume the temperature to be zero). It gives

$$p_f(x) = \prod_{k\geq 0} Q_{g_k}(x_{g_k}) = \frac{Q_f(x_f)}{\pi^{n-1}} e^{-\sum_{k\geq 1}|x_{g_k}|^2} \tag{5.10}$$

We now have to use the fact that the L^2 norm $\|x\|$ of x is given by

$$\|x\|^2 = \int |x(t)|^2 \mathrm{d}t = \sum_k |x_{g_k}|^2 \tag{5.11}$$

Finally, this gives

$$\boxed{p_f(x) = Q_f(x_f)\frac{e^{|x_f|^2}}{\pi^{n-1}}e^{-\|x\|^2} = \frac{Q_f(x_f)}{Q_{|0\rangle}(x_f)}\frac{e^{-\|x\|^2}}{\pi^n}.} \tag{5.12}$$

We decomposed $p_f(x)$ into a part that depends on x_f and a part that depends on the norm of the full measurement record x only, and *does not depend on the state prepared*. Note that in the continuum limit, this expression goes to zero. In a real experiment, it means that it can be very small, which could potentially pose problems if it reaches the machine precision. This quite simple expression for p_f shows its usefulness when we want to compare two different scenarios where the populated mode is either represented by f_0 or by f_1. For a measurement record x then the *likelihood ratio* $p_{f_0}(x)/p_{f_1}(x)$ between the two reads:

$$\frac{p_{f_0}(x)}{p_{f_1}(x)} = \frac{Q_{f_0}(x_{f_0})Q_{|0\rangle}(x_{f_1})}{Q_{f_1}(x_{f_1})Q_{|0\rangle}(x_{f_0})}. \tag{5.13}$$

To compare the likelihood and perform a Bayesian update on the two possible scenarios, we then only need x_{f_0} and x_{f_1}, and what we call the *pseudo-likelihoods* $\mathcal{L}(f_0|x)$ and $\mathcal{L}(f_1|x)$, defined as

$$\boxed{\mathcal{L}(f|x) = \frac{Q_f(x_f)}{Q_{|0\rangle}(x_f)}.} \tag{5.14}$$

In other words, one only needs to demodulate the signal by f_0 and f_1, and the results contain all the information stored in the record. x_{f_0} and x_{f_1} are said to be a *sufficient statistic* [179] to discriminate the two scenarios. This is not a likelihood in the statistical sense because $\int \mathcal{L}(f|x)\mathcal{D}x \neq 1$, but it is proportional to it. Note that a similar pseudo-likelihood is accessible from the SME.

We can now express the mutual information using $\mathcal{L}$ as

$$I(k:x) = \sum_k \mathbb{P}(k)\left[-\int p_{f_k}(x)\log\left(\sum_l \mathbb{P}(l)\frac{\mathcal{L}(f_l|x)}{\mathcal{L}(f_k|x)}\right)\mathcal{D}x\right] \tag{5.15}$$

$$= \sum_k \mathbb{P}(k)\mathbb{E}\left[-\log\left(\sum_l \mathbb{P}(l)\frac{\mathcal{L}(f_l|x)}{\mathcal{L}(f_k|x)}\right)\middle| k\right] \tag{5.16}$$

All we need to know now are the joint laws of all the x_{f_l} conditioned on k to be able to estimate $I(k:x)$ by a Monte Carlo method.

To do so, we assume that the populated mode is the one represented by f_k, and we write that x_{f_l} is the result of the measurement of $\hat{x}_{f_l} = \hat{A}_{f_l} + \hat{B}^\dagger_{f_l}$, where the mode represented by $\hat{B}_{f_l}$ is in the vacuum state. We write it as a function of x_{f_k}:

$$\hat{x}_{f_l} = \hat{A}_{f_l} + \hat{B}^\dagger_{f_l} = (f_l \cdot f_k)\hat{A}_{f_k} + \sqrt{1 - |f_l \cdot f_k|^2}\hat{A}_{f'_{kl}} + \hat{B}^\dagger_{f_l} \tag{5.17}$$

where

$$f'_{kl} = \frac{f_l - (f_l \cdot f_k)}{\sqrt{1 - |f_l \cdot f_k|^2}} f_k. \tag{5.18}$$

By construction, $f_k \cdot f'_{kl} = 0$, so the corresponding annihilation operators $\hat{A}_{f_k}$ and $\hat{A}_{f'_{kl}}$ commute. The functions f_k and f_{kl} thus represent independent modes. We can decompose $\hat{B}_{f_l}$ in the same way (on two independent modes in the vacuum state), which gives

$$\begin{aligned}\hat{x}_{f_l} =& (f_l \cdot f_k)\hat{A}_{f_k} + \sqrt{1 - |f_l \cdot f_k|^2}\hat{A}_{f'_{kl}} \\ &+ (f_l \cdot f_k)\hat{B}^\dagger_{f_k} + \sqrt{1 - |f_l \cdot f_k|^2}\hat{B}^\dagger_{f'_{kl}} \\ =& (f_l \cdot f_k)\big(\hat{A}_{f_k} + \hat{B}^\dagger_{f_k}\big) + \sqrt{1 - |f_l \cdot f_k|^2}\big(\hat{A}_{f'_{kl}} + \hat{B}^\dagger_{f'_{kl}}\big)\end{aligned} \tag{5.19}$$

Thus, we can write $\hat{x}_{f_l}$ as a linear combination of $\hat{x}_{f_k}$ and an independent observable $\hat{b}_{kl}$ that is effectively a heterodyne observable on a mode in the vacuum state. Thus, we can write

$$x_{f_l} = (f_l \cdot f_k)x_{f_k} + \sqrt{1 - |f_l \cdot f_k|^2}b_{kl} \tag{5.20}$$

where b_{kl} is a random variable statistically independent on x_{f_k} that follows the Gaussian law given by $Q_{|0\rangle}$. The correlations between the b_{kl} can be obtained with the same trick, which gives

$$\mathbb{E}[b_{kl}b^*_{kl'}] = f'_{kl} \cdot f'_{kl'} \tag{5.21}$$

$$\mathbb{E}[b_{kl}b_{kl'}] = 0 \tag{5.22}$$

This completely characterizes the joint law of the x_{f_l} conditioned to the k-th mode being populated. In Monte-Carlo simulations, to generate a sample $\{x_{f_k}\}_{k\in[\![0,m-1]\!]}$, we sample x_k from Q_f, and the b_{kl}

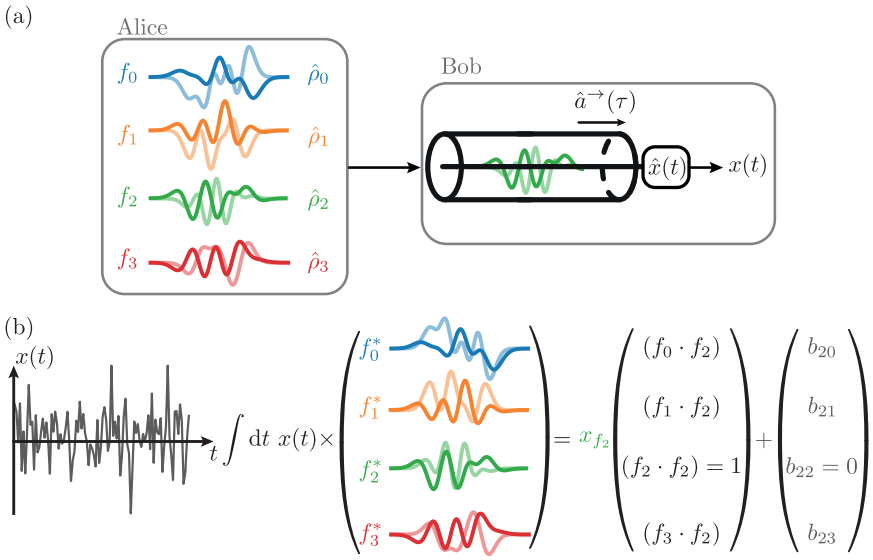


Figure 5.1: Scheme of the communication protocol between Alice and Bob. (a) Alice chooses one mode function f_k among $m = 4$ to prepare in the state ρ_k. Here, $k = 2$. Bob performs a heterodyne measurement and gets a record $x(t)$. (b) Bob integrates its records against the complex conjugates of the mode functions used by Alice. It gives him m records x_{f_l}. Since $k = 2$ here, they can be written as the sum of two vectors, one whose coefficients are given by $x_{f_2}(f_l \cdot f_2)$, and the other with stochastic coefficients b_{2l} following a multidimensional complex Gaussian law, whose covariance matrix is given by the Eq. (5.21) and Eq. (5.22).

from a multidimensional Gaussian law with zero mean, and covariance matrix given by the $\mathbb{E}[b_{kl}b^*_{kl'}]$. This is summarized in Fig. 5.1b. The procedure we follow to compute $I(k:x)$ is then a Monte-Carlo procedure: we generate the x_{f_l} conditioned on k, compute the average value of $-\log\left(\sum_k \mathbb{P}(k)\frac{\mathcal{L}(f_l|x)}{\mathcal{L}(f_k|x)}\right)$ for each of them, then perform their sum over k weighted by $\mathbb{P}(k)$. This method is an efficient way of computing $I(k:x)$, as it can be written as a $2m$-dimensional integral of a mixture with non-Gaussian components. The case $m = 2$ already corresponds to an integral in dimension 4, and we went up to $m = 10$, which corresponds to dimension 20. Note that this procedure allows us to compute the mutual information for a heterodyne measurement

for any (possibly mixed) state preparation in any single mode of the transmission line, with as many codewords as wanted. It could even be generalized to more modes, using the multidimensional Husimi Q function introduced in Appendix B, as soon as the number of populated modes is finite.

In the experiment described in Chap. 4, we are interested in the peculiar case where $f_k(t) = \sqrt{\kappa} e^{-i(\omega_0 - k\chi)t - \frac{\kappa}{2}t}$ for $t > 0$, and $f_k(t) = 0$ otherwise. It gives

$$f_l \cdot f_k = \frac{\kappa}{\kappa + i(k-l)\chi}. \tag{5.23}$$

We can write

$$\hat{A}_{f_l} = \frac{\kappa}{\kappa + i(k-l)\chi} \hat{A}_{f_k} + \frac{i(k-l)\chi}{\kappa + i(k-l)\chi} \hat{A}_{f'_{kl}}. \tag{5.24}$$

Note that here we replaced $\sqrt{1 - |f_l \cdot f_k|^2} = \frac{|k-l|\chi}{\sqrt{\kappa^2 + \left((k-l)\chi\right)^2}}$ by the factor $\frac{i(k-l)\chi}{\kappa + i(k-l)\chi}$ by changing the phase in the definition of f'_{kl} which is more convenient in what follows.

Using Eq. (5.24) and the following commutation relations (Eq. (2.115))

$$[\hat{A}_{f_l}, \hat{A}^\dagger_{f_k}] = (f_l \cdot f_k)\mathbb{1} \tag{5.25}$$

$$[\hat{A}_{f_l}, \hat{A}^\dagger_{f'_k l}] = 0, \tag{5.26}$$

we obtain the covariance matrix between the b_{kl}, given by the coefficients $(f'_{kl} \cdot f'_{kl'})$:

$$(f'_{kl} \cdot f'_{kl'})\mathbb{1} = [\hat{A}_{f'_{kl}}, \hat{A}^\dagger_{f'_{kl}}] = \frac{\kappa}{\kappa + i(l'-l)\chi}\mathbb{1} = [\hat{A}_{f_l}, \hat{A}^\dagger_{f_{l'}}] \tag{5.27}$$

5.1.2 *Frequency-resolved photon detector*

A second measurement setup that we can imagine uses a frequency-resolved photon detector (FRPD). This is very close to the idea that was at the origin of the photon counting project [126]. A possible experimental implementation uses time-resolved photon detectors, to which we are more used. It consists of an array of multiplexers

with a very narrow linewidth, each of them having a time-resolved photon detector behind it. This kind of architecture, out of reach in the microwave domain, exists in the optical domain. It is even possible to multiplex temporal or spatial modes, which allows us to perform measurements on each mode independently. It has applications in quantum communication [51, 52] and in quantum metrology for separation estimation of optical sources of light [48–50].

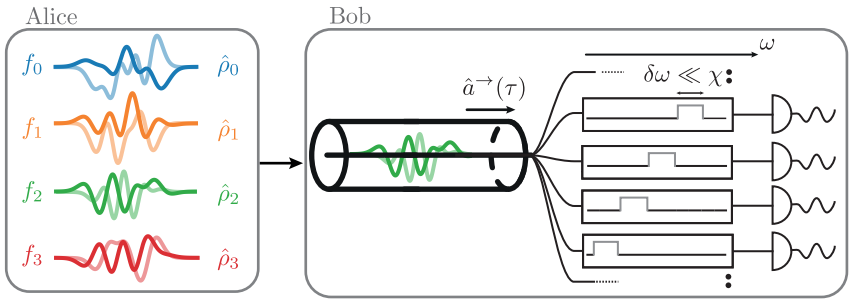


Figure 5.2: Communication scheme with a Frequency-Resolved Photon Detector (FRPD). Alice prepared a mode represented by f_k in the state $\hat{\rho}_k$, which goes through an array of photon detectors that are each positioned at the output of a bandpass filter. Each detector's click gives a measure of the frequency of the wavepacket.

In the continuum limit, this setup measures $\hat{a}^\dagger_{\text{out}}[\omega]\hat{a}_{\text{out}}[\omega]\mathrm{d}\omega$ for every ω. In the case where there is exactly one photon in the wavepacket, we note $p_f(\omega)\mathrm{d}\omega$ the probability density that the photon detector at frequency ω clicked when the mode represented by f was populated by the Fock state $|1\rangle$. We have:

$$p_f(\omega)\mathrm{d}\omega = \langle\psi|\,\hat{a}^\dagger_{\text{out}}[\omega]\hat{a}_{\text{out}}[\omega]\mathrm{d}\omega\,|\psi\rangle = |f[\omega]|^2\mathrm{d}\omega. \tag{5.28}$$

In our case (Fig. 5.2), we have:

$$p_{f_k}(\omega)\mathrm{d}\omega = \frac{1}{2\pi}\frac{\kappa}{\kappa^2 + 4(\omega-\omega_k)^2}\mathrm{d}\omega. \tag{5.29}$$

We note $I_1(k:\omega)$ the mutual information corresponding to this case of a single Fock state $|1\rangle$ in the mode represented by f. In the case where the propagating mode contains n photons, each of them independently produces a click on a photon detector, so that detecting

n photons corresponds to repeating n times the experiment with one photon only. The outcome is a list $\{\omega_j\}_{j\in[\![0,n-1]\!]}$ of all the frequencies associated to the photon detectors that clicked. The corresponding mutual information $I_n(k:\{\omega_j\})$ reads

$$I_n(k:\{\omega_j\}) = nI_1. \tag{5.30}$$

In the case where the photon number is not determined, we can compute an average mutual information $\overline{I}(k)$ as

$$\overline{I}(k) = \sum_n I_n(k:\{\omega_j\}) = \langle\hat{n}\rangle I_1 \tag{5.31}$$

with $\langle\hat{n}\rangle$ the average number of photons in the mode. Note that this is valid only if the photon number statistics given by $\langle n|\hat{\rho}_k|n\rangle$ is the same for all k, *i.e.* only if the number of detected photons carries no information about k.

5.1.3 *Accessible information and Holevo bound*

The communication scheme described in the previous sections is a *quantum communication channel* (see Appendix A.1), as the information is transmitted using quantum states. When fixing the distribution $\mathbb{P}$ used for the codewords of Alice, we can compute two different bounds on the classical information that Alice can send through this channel. These two quantities are defined in detail in Appendix A.1.

The first one is the accessible information. It sets the maximal information that can be transmitted using multiple *local* uses of this quantum channel. Here, the locality means that Alice cannot entangle the states she sends across several uses of the channel, and Bob cannot make joint measurements on the states received.

The second one is the Holevo bound. It sets the maximal information that can be transferred through the channel when allowing entanglement to Alice and joint measurement to Bob. As this case is more permissive, the Holevo bound is larger than the accessible information. The difference between the two is called the *quantum discord* and constitutes a purely quantum resource that can be used to beat classical schemes [180], such as in quantum radar experiments [10].

5.1.3.1 *Accessible information*

We compute the accessible information in the binary case ($k = 0$ or 1) where Bob receives pure states. We can apply the results derived in Appendix A.1.4 by identifying our two states $|\psi_0\rangle$ and $|\psi_1\rangle$ to the states $|b_0\rangle$ and $|b_1\rangle$ in the appendix. $|\psi_0\rangle$ and $|\psi_1\rangle$ read

$$|\psi_0\rangle = F^\star_{f_0}(\hat{A}^{\dagger n}_{f_0})\,|\text{vac}\rangle = \sum_n \frac{\psi^0_n}{\sqrt{n!}}\hat{A}^{\dagger n}_{f_0}\,|\text{vac}\rangle\,, \tag{5.32}$$

$$|\psi_1\rangle = F^\star_{f_1}(\hat{A}^{\dagger n}_{f_1})\,|\text{vac}\rangle = \sum_n \frac{\psi^1_n}{\sqrt{n!}}\hat{A}^{\dagger n}_{f_1}\,|\text{vac}\rangle\,. \tag{5.33}$$

As explained in Sec. A.1.4, all we need to get the accessible information is their overlap $|\,\langle\psi_0|\psi_1\rangle\,|$:

$$\langle\psi_0|\psi_1\rangle = \langle\text{vac}|\sum_{n,m}\frac{\psi^{0*}_n\psi^1_m}{\sqrt{n!m!}}\hat{A}^n_{f_0}\hat{A}^{\dagger m}_{f_1}\,|\text{vac}\rangle\,. \tag{5.34}$$

Only the terms that conserve the number of photons are non-zero, which gives

$$\langle\psi_0|\psi_1\rangle = \langle\text{vac}|\sum_n\frac{\psi^{0*}_n\psi^1_n}{n!}\hat{A}^n_{f_0}\hat{A}^{\dagger n}_{f_1}\,|\text{vac}\rangle \tag{5.35}$$

so

$$\langle\psi_0|\psi_1\rangle = \sum_n (f_0\cdot f_1)^n\psi^{0*}_n\psi^1_n. \tag{5.36}$$

If the two modes are prepared in the same state, we note $\psi^0_n = \psi^1_n = \psi_n$ for all n, which gives

$$\langle\psi_0|\psi_1\rangle = \sum_n (f_0\cdot f_1)^n|\psi_n|^2 \tag{5.37}$$

where $|\psi_n|^2$ is the photon number probability of the state[1].

We can thus compute the overlap between any single-mode states of the transmission line. The optimal measurement to apply and the corresponding maximal mutual information I_{acc} are given by the procedure described in A.1.4, as a function of this overlap.

1 Note that it reads $\langle\psi_0|\psi_1\rangle = G[(f_0\cdot f_1)]$ where G is the probability-generating function of the photon number distribution.

5.1.3.2 *Holevo bound*

We also compute the Holevo bound in the case where Bob receives a pure state. If Bob receives the states $|\psi_k\rangle$ for $k < m$, Eq. (5.37) gives the Gram matrix $\mathbf{G}$ defined as

$$G_{kl} = \{\langle\psi_k|\psi_l\rangle\}_{kl}. \tag{5.38}$$

This matrix is positive definite if **the family $\{|\psi_k\rangle\}$ is a free family, which we consider true in the following**. The Cholesky decomposition gives T an upper triangular matrix such that $\mathbf{G} = T^\dagger T$, which directly gives the coordinates of all the $|\psi_k\rangle$ in an orthonormal basis.

As we are considering pure states here, the Holevo bound χ reads

$$\chi = S\Big(\sum_k \mathbb{P}(k)\,|\psi_k\rangle\langle\psi_k|\Big), \tag{5.39}$$

directly computable knowing the $\mathbb{P}(k)$ and T. S is the von Neumann entropy.

5.2 MUTUAL INFORMATION AND MEASUREMENT RATE

5.2.1 *Frequency-resolved limit*

We can now compute the mutual information $I(k:x)$ between the number of photons in the cavity k (or the frequency of the qubit) and the heterodyne record x for a single kick of the qubit in the experiment presented in Chap. 4. It can be described as follows: at $t = 0$, the qubit is prepared in a state $|\psi\rangle = \cos(\theta/2)\,|g\rangle + \sin(\theta/2)\,|e\rangle$ which is then released in the transmission line. Once again, this case can be mapped onto the quantum communication scheme between Alice and Bob described above by taking $f_k(t) = \sqrt{\kappa}e^{-i(\omega_0 - k\chi)t - \frac{\kappa}{2}t}$ for $t > 0$, and $f_k(t) = 0$ otherwise.

In this section, we place ourselves in the limit where $\chi \gg \kappa$, which is the optimal case, as the two line modes are then orthogonal. We also take $m = 2$ to stick with the way measurement rates are computed in Chap. 4. It is illustrated in Fig. 5.3 for $\theta = \pi$. In this case, $|f_l \cdot f_k| = |\frac{\kappa}{\kappa + (k-l)i\chi}| \ll 1$: all the functions f_k can be considered as orthogonal. Therefore, if Bob demodulates the record x with the same function

f_k as the one chosen by Alice, he will obtain a record that follows the statistics given by the Husimi Q function of the corresponding propagating mode represented by f_k. If he demodulates with another f_l, he will obtain the statistics given by the Husimi Q function of the vacuum state.

The example presented in Fig. 5.3a illustrates the case where a Fock state $|2\rangle$ or a Fock state $|3\rangle$ is prepared in the cavity. It corresponds to Alice having the choice between the two words $k = 2$ and $k = 3$. Let us assume that $k = 2$. In the peculiar case where $\theta = \pi$, the emitted state in the transmission line is a Fock state $|1\rangle$. By demodulating by f_2 and f_3, Bob either obtains outcomes obeying the statistics given by $Q_{|0\rangle}$ or $Q_{|1\rangle}$ (see Fig. 5.3b). Here, the only case that would allow Bob to unambiguously tell which word k Alice sent is the case where Bob obtains exactly $x_{f_2} = 0$. Indeed, $Q_{|1\rangle}(0) = 0$ and $Q_{|0\rangle} = 1/\pi$, so obtaining $x_{f_2} = 0$ means the mode represented by f_2 is not in the Fock state $|1\rangle$, so it is the mode f_3 that is in this state. As this situation has a probability of zero to happen exactly, Bob can never ascertain the word Alice communicates. As a consequence, this communication protocol sends less than one bit of information per realization.

We can now compute the mutual information $I = I(k : x)$ for all θ. From $|\psi\rangle = \cos(\theta/2)\,|g\rangle + \sin(\theta/2)\,|e\rangle$, we use Eq. (2.70) to get the Q function of the mode chosen by Alice, represented by f_k:

$$Q_{f_k}(\beta) = \frac{1}{\pi} e^{-|\beta|^2} \left(\cos\left(\frac{\theta}{2}\right)^2 + \cos\left(\frac{\theta}{2}\right) \sin\left(\frac{\theta}{2}\right) (\beta + \beta^*) + \sin\left(\frac{\theta}{2}\right)^2 |\beta|^2 \right). \tag{5.40}$$

Convolving with $p_{\mathcal{N}}$ (see Eq. (3.54)) with $\sigma^2 = \frac{1-\eta}{2\eta}$ gives the Husimi Q function for a finite quantum efficiency η. Fig. 5.4a shows the mutual information I in the frequency-resolved case $\chi \gg \kappa$ as a function of θ and η. Fig. 5.4b shows cuts of this plot along the lines $\eta = 0.1, 0.25, 0.4, 0.55, 0.7, 0.85$ and 1, which correspond to increasing values for the mutual information. The optimal angle θ_{opt} that maximizes I increases with η, and is plotted in Fig. 5.4c. Interestingly, it remains close to $\theta = \pi/2$ up to high quantum efficiencies compared to the state of the art [88–91]: it crosses $\theta_{\text{opt}} = 3\pi/4$ only for $\eta \simeq 0.9$,

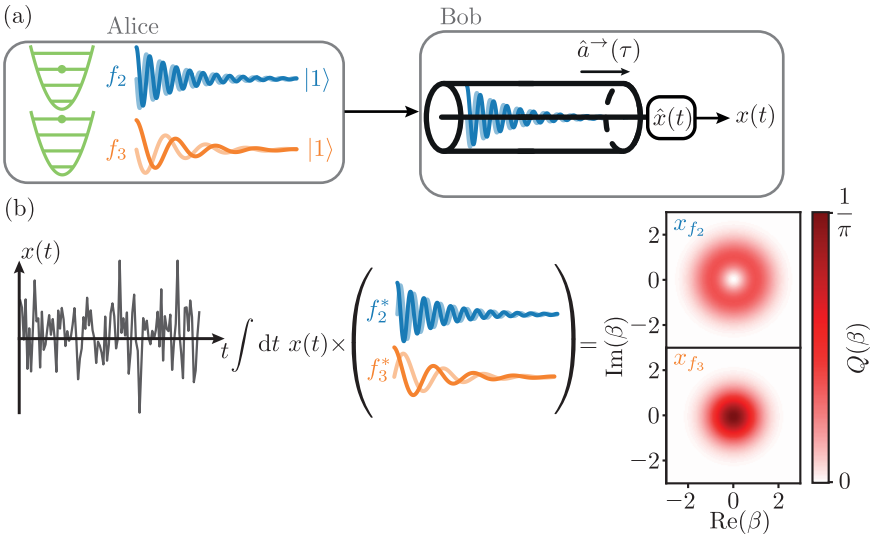


Figure 5.3: Scheme of the equivalent communication protocol between Alice and Bob in the case $m = 2$. (a) Alice chooses between two frequencies for the propagating mode. In the experiment, it would, for example, correspond to two possible photon numbers 2 and 3 in the cavity. Here, she chooses $k = 2$, which prepares a single photon state in the mode defined by f_2. Bob performs a heterodyne measurement and gets a record $x(t)$. (b) Bob can integrate its records against the complex conjugate of the mode functions f_2 and f_3. It gives him 2 complex outcomes x_{f_2} and x_{f_3}. Since $k = 2$ here, x_{f_2} is sampled from the Husimi Q function of the Fock state $|1\rangle$. Since $\chi \gg \kappa$, x_{f_3} is sampled from the Husimi Q function of the vacuum state.

and reaches $\theta_{\text{opt}} = \pi$ very sharply for $\eta = 1$. The noise in the plot is due to the remaining statistical noise due to the use of a Monte Carlo method. We see that the number of bits transmitted saturates at around $I_{\max} = 0.3$ bits. We can link these curves to the following expected behavior for the information as a function of the quantum efficiency. When going through a beamsplitter with transparency η, the average signal is attenuated by a factor $\sqrt{\eta}$. However, the variance of this average signal, which gives the noise power, is attenuated by η. In fact, moments of order n are attenuated by a factor $\sqrt{\eta}^n$. We can think of it this way: getting information from the heterodyne

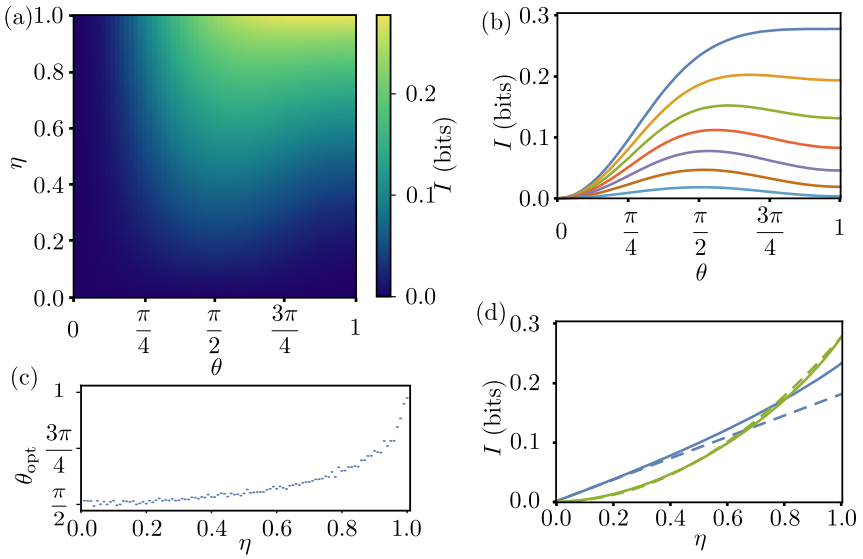


Figure 5.4: Mutual information between the number of photons and a heterodyne measurement record as a function of the kick amplitude on the qubit and the quantum efficiency. (a) Map of the mutual information $I = I(k : x)$ between the number of photons $k \in \{0; 1\}$ and the measurement record x of a heterodyne measurement in the frequency-resolved case $\chi \gg \kappa$ as a function of θ and η. (b) Cuts of the color map along the lines $\eta = 0.1, 0.25, 0.4, 0.55, 0.7, 0.85$ and 1 corresponding to increasing mutual information values. (c) Optimal angle θ_{opt} as a function of η. (d) Mutual information I as a function of the quantum efficiency for $\theta = \pi/2$ (blue solid line) and $\theta = \pi$ (green solid line). The blue dashed line is a linear fit for small η, and the green dashed line shows the function $I = I_{\text{max}}\eta^2$.

measurement records consists in discriminating the Husimi Q function of the mode and the Q function of the vacuum state. For small θ, the qubit state resembles a small coherent state of amplitude $\theta/2$, thus carrying most of the information in the average amplitude. The Husimi Q function is thus a Gaussian function, and when getting the measurement outcomes, we want to know whether their average is 0 (vacuum) or $\theta/2$: the information lies in the average. A finite quantum efficiency η, which is equivalent to putting a beamsplitter on the line,

will thus transform the problem into discriminating 0 from $\sqrt{\eta}\theta/2$. The SNR, defined as in Eq. (3.92), is thus proportional to η.

When θ gets close to π (and the state $\hat{\rho}_k$ closer to $|1\rangle\langle 1|$), this information in the average of the outcome gets smaller and smaller as the average signal vanishes. The information then lies in the higher-order moments of the Q distribution, which vanish faster than the average as a function of η. We illustrate this in Fig. 5.4d by plotting the mutual information as a function of the quantum efficiency for $\theta = \pi/2$ (blue solid line) and $\theta = \pi$ (green solid line). We fitted these curves with a straight line for $\theta = \pi/2$ at short values (blue dashed line) and plotted the function $I = I_{\max}\eta^2$ (green dashed line), which almost coincides with the solid green line. $I_{\max}$ is the mutual information attained for $\eta = 1$. Interestingly, the mutual information is almost linear in η when it is mostly encoded in the average and quadratic when there is no average signal. In the latter case, this is coherent with the intuition that most of the information lies in the variance at small η: the Husimi Q function can be seen as a Gaussian function with a slightly bigger variance than that of the vacuum state in this regime.

An FRPD works better in this regime $\chi \gg \kappa$: when loaded with 1 photon, the frequency click unambiguously discriminates between the two cases, as the two possible frequency distributions p_{f_0} and p_{f_1} defined in Eq. (5.29) do not overlap. Each realization thus brings one bit of information. This thus saturates the number (1) of bits transmittable. The accessible information and the Holevo bound thus amount to one bit of information each.

5.2.2 *Other regimes*

We further explore the full range of values for χ/κ and compare the heterodyne detection, the FRPD, the accessible information, and the Holevo bound for each of these cases. The result as a function of the qubit excitation probability $p(e) = \cos(\theta/2)^2$ is shown in Fig. 5.5. The previous case $\chi \gg \kappa$ is shown in Fig. 5.5a. We see that the heterodyne measurement performs worse than all the others but that the FRPD does achieve the accessible information bound only for $\theta = \pi$, or $p(e) = 1$. This can be understood as the fact that this measurement is

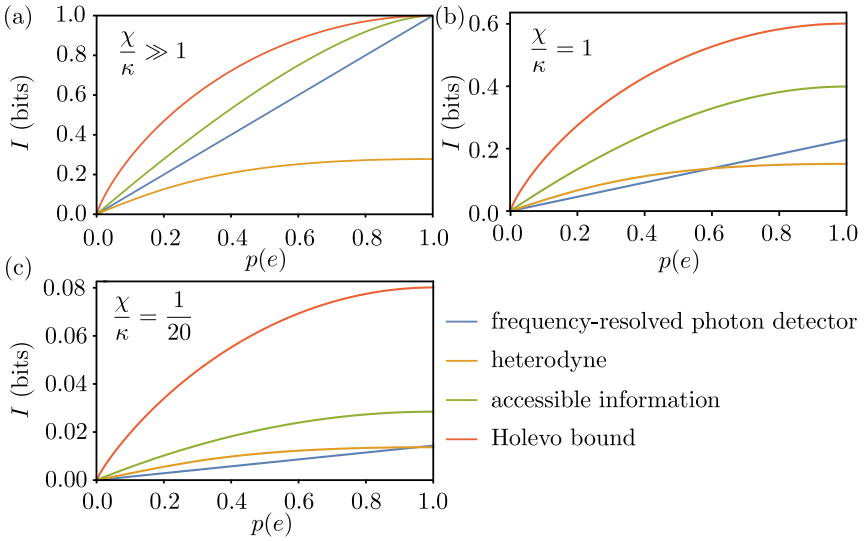


Figure 5.5: (a) to (c): Holevo bound and mutual information as a function of $p(e) = \cos(\theta/2)^2$ for the heterodyne detection, the FRPD, and an optimal detector, for different values of χ/κ.

completely insensitive to the coherence of the state, thus dismissing a useful resource. Alternatively, we can think about the fact that the FRPD gives additional information: the total number of photons in the line, which does not help in discriminating the two scenarios. As a consequence, this information is associated to an excess back-action on the state, which lowers the efficiency of this measurement, as in Sec. 3.3.6. The Holevo bound exceeds the accessible information everywhere and saturates to 1, as expected.

Interestingly, when lowering the ratio χ/κ, the heterodyne measurement starts to perform better than the FRPD for small qubit excitation probabilities $p(e)$. At $\chi/\kappa = 1$ (Fig. 5.5b), it performs worse only for $p(e) \gtrsim 0.6$. The FRPD performs now worse than the accessible information, indicating that the coherence of the state becomes a crucial resource. The Holevo bound is now never reached by the accessible information. At $\kappa = 20\chi$, the FRPD's best performance is half the accessible information for $p(e) = 1$, as for the heterodyne detection, which performs always better than the FRPD for $p(e) < 1$. Strikingly, the Holevo bound is much larger than the accessible information. As

a matter of fact, we observed that the ratio between Holevo bound and accessible information diverges[2] for $\chi/\kappa \to 0$.

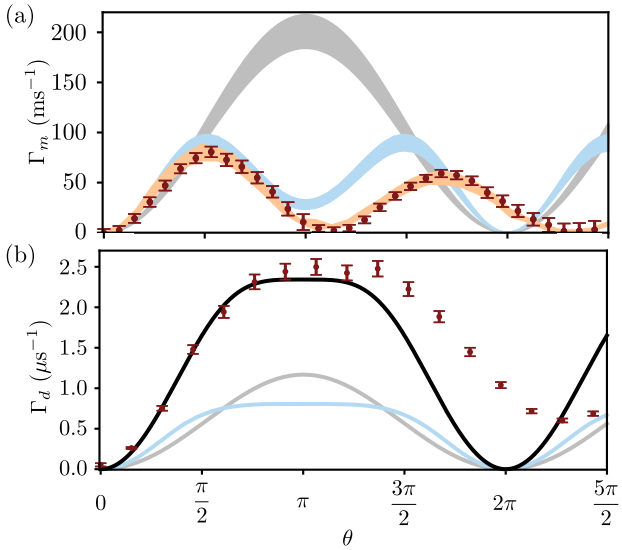


Figure 5.6: Measured and theoretical measurement rates for heterodyne, FRPD and ideal measurement. (a) Red dots: observed average measurement rate Γ_m as a function of drive amplitude, expressed as the qubit expected rotation angle θ per kick. Blue shadow: theoretical bound for an infinite comb and heterodyne measurement with efficiencies $\eta \in [0.17, 0.20]$. Grey shadow: theoretical bound for an infinite comb using an FRPD with efficiency $\eta \in [0.17, 0.20]$. (b) Red dots: observed cavity dephasing rate Γ_d as a function of θ. Black: theoretical accessible information rate. Blue: theoretical maximal ($\eta = 1$) measurement rate obtained by heterodyne detection. Grey: theoretical maximal ($\eta = 1$) measurement rate obtained with an FRPD.

Finally, we show again the Fig. 4.4 but including this time the FRPD, which gives the Fig. 5.6. This regime of the experiment is close to the regime $\chi \simeq \kappa = \frac{1}{T_1}$. The FRPD would perform better than a heterodyne detector around $\theta = \pi$ in both cases $\eta \simeq 0.17$ and $\eta = 1$,

2 This divergence between communication rates is reminiscent of the Heisenberg scaling in quantum metrology [181].

whereas the heterodyne detector performs better in the regions where the information lies mostly in the qubit average quadratures.

To compute a measurement rate from the mutual information I at $\eta = 1$ in this case, it is not enough to compute $\frac{I}{\tau}$ as in Eq. (4.17). Indeed, the mutual information becomes too large compared to $\log(2)$ (the initial entropy of Alice's codewords) for this first-order development to be valid. As a consequence, we corrected the measurement rate by writing:

$$\Gamma_m = \Gamma_m^{\mathrm{eq}} \tag{5.41}$$

where Γ_m^{eq} is defined as the equivalent measurement rate obtained for a dispersive qubit measurement performed with $\eta = 1$ and an *optimal measurement*, defined in section Sec. 3.3.5, performed after a time $\tau = \frac{\pi}{\chi}$, which would give the same mutual information I.

It is important to note that this procedure, applied on the optimal measurement, allows us to recover exactly the dephasing rate as computed in Chap. 4. The dephasing rate is thus attained for a *time global* measurement (see Sec. 3.3.5) on the emitted mode over the repetition time. However, this emitted state is now entangled when expressed in a time basis, which forbids decomposing it on a smaller time grid, as it was done in Sec. 3.3.3. An optimal time local measurement could then very well miss some information and, in the end, exhibit a measurement rate lower than the dephasing rate defined this way. Interesting questions thus remain open: does a time local measurement reach the dephasing rate, as it is the case for a standard dispersive qubit readout? If not, what is the best a local measurement could do?

5.3 DISCRIMINATING MORE THAN TWO FREQUENCIES

This section aims at discussing how to generalize the notion of measurement rate to a case with more than two outcomes. Up to this point, we have measured the performance of our photon counter solely by its capacity to distinguish between two consecutive photon numbers. This is justified by three reasons. First by the fact that the frequencies that are the hardest to distinguish are consecutive frequencies. On average, the numbers that are far from the actual photon number

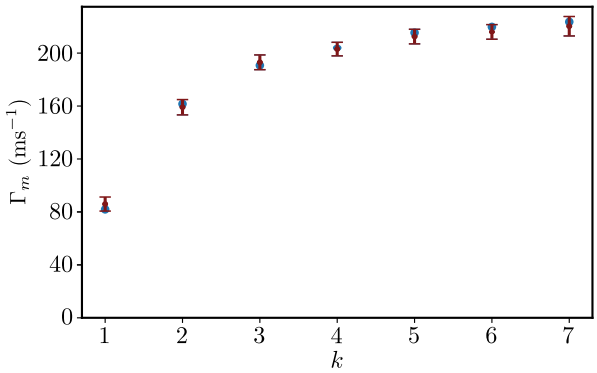


Figure 5.7: Red dots: observed average information rate Γ_m for $\theta = \pi/2$ when discriminating the photon number 0 from the photon number k, as a function of k. Blue: theoretical information rate for a quantum efficiency $\eta = 0.17$.

are ruled out quickly. Fig. 5.7 shows the observed measurement rate when discriminating the Fock state $|0\rangle$ from a Fock state $|k\rangle$ as a function of k. The case $k = 2$ exhibits a photon number twice as high as the case $k = 1$. As a consequence, the last thing to do is almost always to discriminate between a few photon numbers. Second, once the Bayesian estimator converges to a number of photons, the Bayesian filter only has to detect single photon losses, and as a consequence, it mostly has to decide between two consecutive frequencies. Finally, it allows us to apply the theory of the qubit readout to this case with more than two possible outcomes, and it gives an experimental way to compare the measurement rate to a dephasing rate obtained by Wigner tomography.

However, the corresponding binary measurement rate does not quantify the information rate that can be transmitted over the full bandwidth of this detection setup. It does not answer a crucial question: how fast can the estimator converge at short times when the entropy is maximal?

5.3.1 *Information rate*

In this regard, Eq. (3.85) needs to be treated with care. Using this formula gives the amount of information that we can transmit with this communication channel. It thus purely gives an information rate. We computed this information rate Γ_i both theoretically and experimentally for $m = 8$ possible frequencies as a function of θ and compared it once again against the accessible information and to the FRPD. We plotted the result in the Fig. 5.8. What we can notice is that it does not change the curves qualitatively but quantitatively. We have

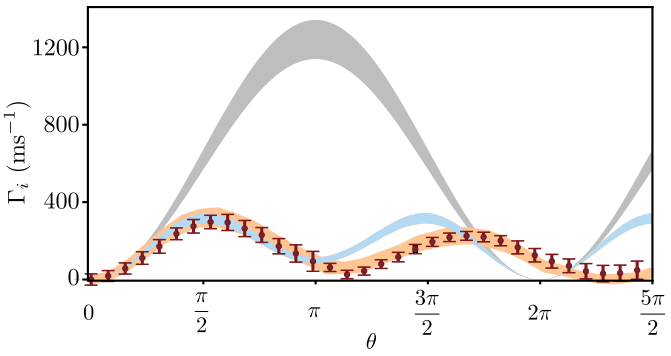


Figure 5.8: Information rates with 8 possible photon numbers. Red dots: observed average information rate Γ_i as a function of drive amplitude, expressed as the qubit expected rotation angle θ per kick. Blue shadow: theoretical bound for an infinite comb and heterodyne measurement with efficiencies $\eta \in [0.17, 0.20]$. Grey shadow: theoretical bound for an infinite comb using an FRPD with efficiency $\eta \in [0.17, 0.20]$.

approximately $\Gamma_i \simeq 3\Gamma_m$ for the heterodyne measurement. This is expected, as in the theoretical modeling, Alice is sending 3 bits instead of 1. We still quantitatively reproduce the experiment for $\theta \leq \pi$. The amount of information transmittable with the heterodyne communication scheme typically scales with $\log m$. Quite unexpectedly, $\Gamma_i \simeq 6\Gamma_m$ for the FRPD, which is at first really surprising, and seems to violate the rules of information transfers. The reason for this behavior is subtle. This is due to the fact that when adding a new possible number of photons (a codeword in the quantum communication language),

the associated Lorentzian distribution has an overlap with the other distribution of the other possible number of photons that goes to zero as a function of the difference in number of photons. Asymptotically, if we need a frequency difference of $M\chi$ to differentiate reliably the two associated lorentzians, the amount of transferrable information per detected photon is at least $I = \log(m/M) = \log(m) - \log(M) \sim \log(m)$. By contrast, with the heterodyne measurement scheme, even with an infinitely large frequency difference, the probability distributions of the measurement record for two different numbers of photons still have a finite overlap. In this case, the amount of transferrable information seems to scales like $\log(m)$ but with a certain factor $x \in [0, 1[$ corresponding to the fraction of information transferrable: $I \sim x \log(m)$. We observed this with up to $m = 8$ experimentally in Fig. 5.9, where we defined $\Gamma'_m = \Gamma_i / \log(m)$ (see next Sec. 5.3.2).

However, when simulating larger values for m, this scaling changes, and Γ_i saturates to a certain value $\Gamma_i^{\max}$. This is surprising at first because it seems that we should be able to transfer more information when increasing the number of possible number of photons m: if Alice has more codewords, she should be able to transfer more information to Bob. As discussed in [126] with a toy model in Appendix F1, the codewords become more and more unreliably transferred, and inferring the right number of photons takes also a measurement time (or a number of repetitions) that is logarithmic in m. We observed this behavior whatever the quantum efficiency or the kick angle. It is linked to the fact that all the conditional probability distributions have a non-vanishing overlap. Asymptotically, the probability to make a mistake in identifying the correct number of photons in one realization then goes to 1.

This first way of generalizing the measurement rate thus gives a nice quantum communication point of view. However, it is not satisfactory to define a characteristic measurement time. Indeed, we do not expect it to be shorter for a higher number of photons but rather to increase with the number of possible numbers of photons m.

5.3.2 *Measurement time*

A simple toy model to recover such an increasing measurement time is to look at the characteristic time T_m needed for the mutual information to reach the initial entropy H_0 about the number of photons. It corresponds to define T_m by the equality

$$\Gamma_i T_m = H_0. \tag{5.42}$$

For $H_0 = \log(m)$, we simply recover $T_m = \log(m)/\Gamma_i$, which gives the expected logarithmic scaling in the number of possibilities. The measurement rate can then be defined as

$$\Gamma'_m = \frac{1}{T_m} = \frac{\Gamma_i}{\log(m)}. \tag{5.43}$$

Note that in the case $m = 2$, it gives a higher measurement rate

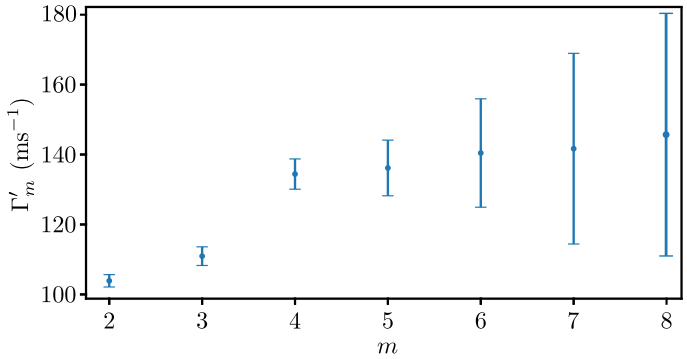


Figure 5.9: Experimentally determined generalized measurement rate Γ'_m as a function of the number of possible photon numbers m (corresponding to photon numbers between 0 and $m-1$). The large error bar size for $m \geq 6$ is due to the lack of samples.

by a factor $1/\log(2)$. This solution looks satisfactory, but in our case, the information rate seems to also logarithmically with m. The measurement time thus seems asymptotically constant, as shown in Fig. 5.9. To recover the expected logarithmic scaling demonstrated in [126], we need to look at higher values for m. As Γ_i saturates, T_m indeed scales asymptotically with $\log(m)$.

On a final remark, this newly defined measurement time does not correspond to a dephasing rate. The dephasing rate is indeed only defined between two possible photon numbers. In the end, what fully characterizes the measurement rate is the matrix of the measurement rates **MR** corresponding to the pair of states $\{(|k\rangle, |l\rangle)\}_{kl}$. This matrix is pictured in Fig. 5.10. For a measurement saturating the accessible information bound, this matrix would correspond to the dephasing rate between the states $\{(|k\rangle, |l\rangle)\}_{kl}$.

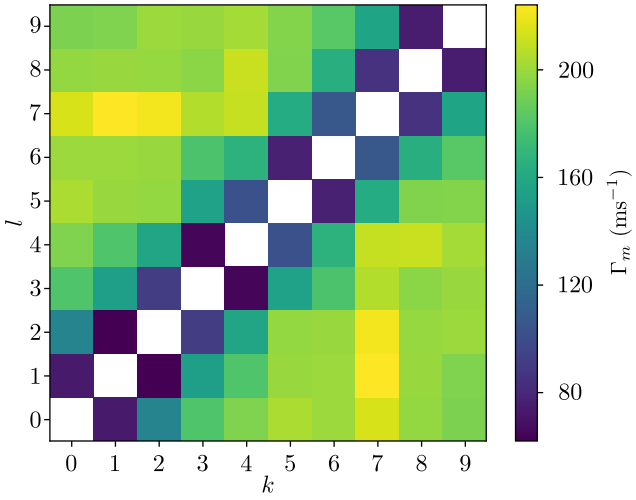


Figure 5.10: Experimentally determined measurement rate matrix **MR**.

5.4 CONTINUOUS CASE

The problem tackled here can be thought of as discriminating which mode of the transmission line was populated. In the case of the experiment of Chap. 4, it amounts to identifying the frequency at which the wavepacket has been emitted. We can formulate this more formally like this: we have a mode, described by the function f, which can be translated in frequency by an offset $\Delta\omega$ that we want to determine. The goal of the experiment is to discriminate between a finite number of values for $\Delta\omega$, but we can now turn it into a *parameter estimation* problem: if $\Delta\omega$ can take continuous values, how fast can we estimate it? We then turn from a communication point of view

to a metrology point of view. The quantities to compute to quantify the precision in the estimation of $\Delta\omega$ are the Fisher information (FI) and the Quantum Fisher Information (QFI) (see Appendix A.2). We can also interpret this continuous case as a binary case in the limit where the two wavefunctions f_0 and f_1 are infinitely close to each other ($|f_0 \cdot f_1| \simeq 1$). More quantitatively, in the case of balanced binary information sent using a noisy classical channel, the mutual information is proportional to a Fisher information (see Appendix A.2.2).

We can model the problem as follows. We consider a state $|\psi\rangle = |\psi\rangle_f \otimes |\text{vac}\rangle$ that we can shift in frequency. We receive a state

$$|\psi(\Delta\omega)\rangle = \hat{\mathcal{T}}_{\Delta\omega} |\psi\rangle \tag{5.44}$$

$$= |\psi\rangle_{f_{\Delta\omega}} \otimes |\text{vac}\rangle \tag{5.45}$$

where we noted $f_{\Delta\omega} = \mathcal{T}_{\Delta\omega} f$, with $\mathcal{T}_{\Delta\omega}$ the translation operator in the frequency domain defined in Eq. (2.123). We will now compute the FI about $\Delta\omega$ associated to a heterodyne or homodyne measurement, FRPD, and compare it to the QFI.

This problem can be seen in the more general setting of estimating a parameter encoded in the modal structure of a light beam. Indeed, the transform could be any unitary operation applied on the mode function f, and the state may not be pure. In this regard, a more general framework is developed in Ref. [182] to identify quantum limits on the estimation of such parameters.

5.4.1 *Quantum Fisher Information*

The problem, as stated, closely aligns with the topic addressed in Ref. [55]. We adopt the framework presented therein but focus on the Fisher information concerning frequency rather than time. Given the symmetric relationship between time and frequency, we can modify

equation (2) from [55] accordingly. The Quantum Fisher Information (QFI), $F_Q(\Delta\omega)$, is expressed as:

$$\begin{aligned} F_Q(\Delta\omega) &= 4(\Delta\hat{\boldsymbol{\tau}})^2 \\ &= 4\Big(\langle\psi(\Delta\omega)|\, \hat{\boldsymbol{\tau}}^2\, |\psi(\Delta\omega)\rangle - \langle\psi(\omega)|\, \Delta\hat{\boldsymbol{\tau}}\, |\psi(\Delta\omega)\rangle^2 \Big) \\ &= 4\Big(\langle\hat{n}\rangle \Delta\hat{\tau}^2 + \Delta\hat{n}^2 \overline{\tau}^2 \Big), \end{aligned} \tag{5.46}$$

where $\langle\hat{n}\rangle$ is the average number of photons in the state $|\psi(\Delta\omega)\rangle$, $\Delta\hat{n}^2$ its variance, $\overline{\tau} = \int_{-\infty}^{+\infty} t|f(t)|^2 \mathrm{d}t$ the average time of the wavepacket, defined as the average value of the operator $\hat{\tau}$ on the function f, and $\Delta\hat{\tau}^2 = \int_{-\infty}^{+\infty} (t-\overline{\tau})^2 |f(t)|^2 \mathrm{d}t$ its variance. The time operator $\hat{\tau}$ is defined in Eq. (2.121), and its quantum operator counterpart $\hat{\boldsymbol{\tau}}$ in Eq. (2.129).

$F_Q(\Delta\omega)$ is a sum of two terms. The first term, proportional to $\langle\hat{n}\rangle$, shows the contribution of the average population of the mode. The one proportional to $\Delta\hat{n}^2$ is the contribution of the coherence of the state. For pure states, the larger $\Delta\hat{n}^2$, the more defined the phase, as they are conjugate quantities. It is then possible to have a scaling that is proportional to $\overline{n}^2$: for a state of the form $|\psi\rangle = \frac{|0\rangle + |N\rangle}{\sqrt{2}}$, we have $\Delta\hat{n}^2 = N^2/4 = \overline{n}^2$. Using this kind of state makes maximum use of the resources (the number of photons) and allows to reach the so-called Heisenberg scaling [55]. In this term proportional to $\Delta\hat{n}^2$, $\overline{\tau}$ is the typical time allowed to measure the frequency: longer time then means better frequency resolution. Note that this expression assumes that the experiment started at time $t = 0$. It implies that the temporal wavepackets are defined for positive times only.

An interesting property of this expression is that it only depends on the photon number statistics of the state. In particular, it means that changing the phases of its coordinated in the Fock basis does not change the Fisher information.

We can already look at different examples. In the case of a mode prepared by a decaying linear cavity, we have $f(t) = \sqrt{\kappa} e^{-(\frac{\kappa}{2}+i\omega_0)t}$. It gives:

$$\overline{\tau} = \frac{1}{\kappa} \tag{5.47}$$

$$\Delta\hat{\tau}^2 = \frac{1}{\kappa^2}. \tag{5.48}$$

Hence, a very simple expression for the QFI is

$$F_Q = \frac{4}{\kappa^2}\big(\langle\hat{n}\rangle + \Delta\hat{n}^2\big). \tag{5.49}$$

We can compute it in a few simple cases. For a coherent state of amplitude α prepared in the mode of interest, the average number of photons is $\langle\hat{n}\rangle = |\alpha|^2$ and the variance also reads $\Delta\hat{n}^2 = |\alpha|^2$, as the distribution is Poissonian. It gives:

$$F_Q = \frac{8|\alpha|^2}{\kappa^2}. \tag{5.50}$$

We can also prepare a displaced vacuum squeezed state $|\alpha, \zeta\rangle$, defined as

$$|\alpha, \zeta\rangle = \hat{\mathcal{D}}(\alpha)\hat{\mathcal{S}}(\zeta)\,|0\rangle\,. \tag{5.51}$$

For $\zeta = r$ with r real, it gives

$$\begin{aligned} F_Q = \frac{1}{\kappa^2}\Big(&-3 + 2\cosh(2r) + \cosh(4r) \\ &+ 8\cosh(r)(e^{-3r}\,\mathrm{Re}(\alpha)^2 + e^{3r}\,\mathrm{Im}(\alpha)^2)\Big) \end{aligned} \tag{5.52}$$

This expression is divided into two parts. The first part does not depend on α and thus remains present even for $\alpha = 0$. It comes from the pure squeezing of the state. The other one that depends on α can be enhanced if α is imaginary and $r > 0$ (or α real and $r < 0$), which means that the amplified quadrature is aligned with α. Or it can be diminished if the amplified quadrature is orthogonal to α. This can be intuited from the fact that the first case has a much more defined phase than the other, in the sense that its angular spread is smaller in phase space. This is pictured in Fig. 5.11. The optimum for fixed squeezing and displacement amplitudes is reached when its

phase is the most defined. As we will see in Sec. 5.4.4, it comes from the fact that the frequency translation can be seen at first order as a rotation in a two-mode phase space. A more defined phase then means higher distinguishability. This behavior is thus very close to usual phase estimation problems.

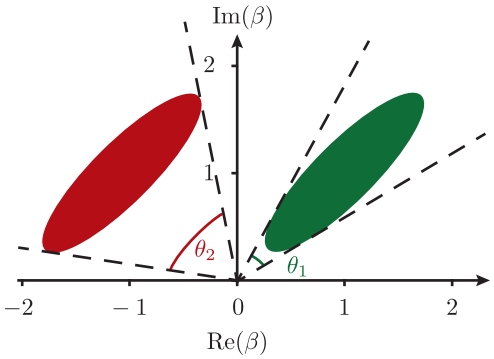


Figure 5.11: Influence of the direction of squeezing on the QFI. We represent two displaced vacuum states with the same absolute amplitudes $|\alpha_1| = |\alpha_2| = 1$. The first one (left, red) is associated to a better QFI than the second one (right, green). This can be understood by looking at the angular spreads θ_1 and θ_2, which are significantly different.

We can finally compute the QFI in the case where the mode is emitted by a qubit for which the occupation probability of the Fock states $|0\rangle$ and $|1\rangle$ are $p(0) = \cos(\theta/2)^2$ and $p(1) = \sin(\theta/2)^2$. The QFI reads

$$F_Q = \frac{4}{\kappa^2}\big(p(1) + p(0)p(1)\big). \tag{5.53}$$

This situation would correspond to the experiment performed in Chapter 4 if the number of photons was a continuous parameter. We will compare it to the FI obtained with an FRPD, a heterodyne, and a homodyne detection.

5.4.2 *Frequency-resolved photon detector*

The photon counter case is the simplest. We can first compute it for a single photon state. We will then use the additivity of the Fisher

information to generalize it to arbitrary states. It corresponds to computing the Fisher information associated to the parameter $\Delta\omega$ of the probability density $\mathbb{P}(\omega)$ that reads

$$\mathbb{P}(\omega) = |f[\omega - \Delta\omega]|^2. \tag{5.54}$$

In the case of a linearly decaying cavity at frequency ω_0, we write again Eq. (5.29) as:

$$\mathbb{P}(\omega) = \frac{1}{2\pi} \frac{\kappa}{\kappa^2 + 4(\omega - \omega_0 + \Delta\omega)^2} \tag{5.55}$$

which defines a Cauchy distribution. The expression of the FI associated to the location parameter of a Cauchy distribution is well known in statistical inference. Here, it is given by:

$$F(\Delta\omega) = \frac{1}{2}\frac{4}{\kappa^2} = \frac{2}{\kappa^2}. \tag{5.56}$$

Using the additivity of the FI, we can generalize to an arbitrary state by writing an average FI $\overline{F}(\Delta\omega)$:

$$\overline{F}(\Delta\omega) = \frac{2\langle\hat{n}\rangle}{\kappa^2} \tag{5.57}$$

with $\langle\hat{n}\rangle$ the average number of photons in the prepared mode. This gives the average FI obtained using an FRPD[3].

For a coherent state, the FI thus reads

$$\overline{F}(\Delta\omega) = \frac{2|\alpha|^2}{\kappa^2}, \tag{5.58}$$

which is four times smaller than the QFI. Once again, it can be understood by the fact that this measurement does not exploit the state coherence.

As for a displaced squeezed vacuum state, the FI reads

$$\overline{F}(\Delta\omega) = \frac{1}{\kappa^2}\left(2\cosh(2r) - 2 + e^{-2r}\,\mathrm{Re}(\alpha)^2 + e^{2r}\,\mathrm{Im}(\alpha)^2\right). \tag{5.59}$$

Here also, it is significantly lower than the QFI.

3 In this peculiar case of a Cauchy distribution, the notion of Fisher Information makes only sense after a large number of samples and when using efficient estimators such as a maximum likelihood estimator [183].

5.4.3 *Heterodyne measurement*

To work out the heterodyne case, we can simply use Eq. (A.18), replacing p_λ by the pseudo-likelihood $\mathcal{L}$ defined in Eq. (5.14):

$$F(\Delta\omega_0) = \mathbb{E}\Big[\Big(\frac{\partial \log(\mathcal{L}(\Delta\omega|x))}{\partial\Delta\omega}\Big)^2|\Delta\omega\Big], \tag{5.60}$$

where x is the measurement record. From the Sec. 5.1.1, we have both the expression for $\mathcal{L}(\Delta\omega|x)$ as a function of $x_{f_{\Delta\omega}}$, and the law of $x_{f_{\Delta\omega}}$ knowing that the only mode that is not empty in the line is the mode defined by $f_{\Delta\omega}$. Thus, we have its law as a function of $\Delta\omega$. In what follows, we denote $x(\Delta\omega) = x_{f_{\Delta\omega}}$ for simplicity.

Recalling that $\mathcal{L}(\Delta\omega|x) = Q(x(\Delta\omega))/Q_{|0\rangle}(x(\Delta\omega))$, what we need now is the law of $\frac{\partial x(\Delta\omega)}{\partial\Delta\omega} = x'(\Delta\omega)$. To determine it, we write

$$x(\Delta\omega) = \int_{-\infty}^{+\infty} e^{i\Delta\omega t} f^*(t)x(t)\mathrm{d}t. \tag{5.61}$$

Taking the derivative with respect to $\Delta\omega$ gives

$$\begin{aligned}\frac{\partial x(\Delta\omega)}{\partial\Delta\omega} = x'(\Delta\omega) &= \int_{-\infty}^{+\infty} ite^{i\Delta\omega t} f^*(t)x(t)\mathrm{d}t \\ &= i\int_{-\infty}^{+\infty} (\hat{\tau} f^*_{\Delta\omega})(t)x(t)\mathrm{d}t.\end{aligned} \tag{5.62}$$

We can thus write $x'(\Delta\omega)$ as another random variable, coming from the measurement of the mode defined by $\hat{\tau} f_{\Delta\omega}$.

Applying the same techniques as in Sec. 5.1.1, we can decompose $\hat{\tau} f_{\Delta\omega}$ into a part colinear to f, and an orthogonal part. We write

$$g_{\Delta\omega} = \frac{1}{\sqrt{\hat{\tau} f_{\Delta\omega} \cdot \hat{\tau} f_{\Delta\omega}}} \hat{\tau} f_{\Delta\omega} \tag{5.63}$$

the normalized version of $\hat{\tau} f_{\Delta\omega}$. Decomposing $g_{\Delta\omega}$ into a part parallel to f and a part orthogonal gives

$$\begin{aligned} g_{\Delta\omega} =& (g_{\Delta\omega} \cdot f_{\Delta\omega}) f_{\Delta\omega} \\ &+ \sqrt{1 - |g_{\Delta\omega} \cdot f_{\Delta\omega}|^2} \frac{g_{\Delta\omega} - (g_{\Delta\omega} \cdot f_{\Delta\omega}) f_{\Delta\omega}}{\sqrt{1 - |g_{\Delta\omega} \cdot f_{\Delta\omega}|^2}}. \end{aligned} \tag{5.64}$$

Note that we have $(\hat{\tau} f_{\Delta\omega} \cdot f_{\Delta\omega}) = (\hat{\tau} f \cdot f) = \overline{\tau}$. We can then write

$$\hat{A}_{g_{\Delta\omega}} = (g_0 \cdot f)\hat{A}_{f_{\Delta\omega}} + \sqrt{1 - |g_0 \cdot f|^2}\hat{B}_{\Delta\omega} \tag{5.65}$$

where $\hat{B}_{\Delta\omega}$ is the bosonic operator associated to the function $\frac{g_{\Delta\omega}-(g_{\Delta\omega}\cdot f_{\Delta\omega})f_{\Delta\omega}}{\sqrt{1-|g_{\Delta\omega}\cdot f_{\Delta\omega}|^2}}$, which is assumed to be in the vacuum state. In the end, again following the same steps as in Sec. 5.1.1:

$$x'(\Delta\omega) = i\frac{(\hat{\tau}f\cdot f)}{(g_0\cdot f)}\Big((g_0\cdot f)x(\Delta\omega) + \sqrt{1-|g_0\cdot f|^2}b\Big) \tag{5.66}$$

where b is a Gaussian random variable, independent of x_f, which follows the probability density given by $Q_{|0\rangle}$. Note that this derivation can be easily generalized to any kind of unitary operation $\hat{\mathcal{U}}(\lambda) = e^{-i\lambda\hat{O}}$ performed on the mode function f. The previous equation reads in general

$$x'(\lambda) = i\frac{(\hat{O}f\cdot f)}{(g_0\cdot f)}\Big((g_0\cdot f)x_\lambda + \sqrt{1-|g_0\cdot f|^2}b\Big) \tag{5.67}$$

where g_0 is defined the same way from $\hat{O}f$:

$$g_0 = \frac{1}{\sqrt{\hat{O}f\cdot\hat{O}f}}\hat{O}f. \tag{5.68}$$

We then have all the needed tools to derive the FI in the heterodyne case. We have only two independent random variables: x_f, whose law is given by the Husimi Q function of the state of the mode prepared, and b, which is a simple Gaussian variable. The Fisher information thus reads as a four-dimensional integral over two complex variables:

$$F(\Delta\omega_0) = \iint\Big(\frac{\partial\log(\mathcal{L}(\Delta\omega|x))}{\partial\Delta\omega}\Big)^2 Q_f(x)Q_{|0\rangle}(b)\mathrm{d}x\mathrm{d}b \tag{5.69}$$

where x in the integral stands for the values of $x(\Delta\omega)$. Defining

$$h(x(\Delta\omega)) = \log(\mathcal{L}(\Delta\omega|x)), \tag{5.70}$$

we can compute its derivative with respect to $\Delta\omega$ using the chain rule:

$$\frac{\partial}{\partial\Delta\omega}h(x(\Delta\omega)) = x'(\Delta\omega)\cdot\Big(\frac{\partial}{\partial\,\mathrm{Re}(x(\Delta\omega))} + i\frac{\partial}{\partial\,\mathrm{Im}(x(\Delta\omega))}\Big)(h(\Delta\omega)) \tag{5.71}$$

where $(\cdot)$ here stands for the complex scalar product defined as $z_1 \cdot z_2 = \mathrm{Re}(z_1^* z_2)$. It can be written in a nicer form using the Wirtinger derivative. Defining

$$h'(x) = \frac{1}{2}\Big(\frac{\partial}{\partial \mathrm{Re}(x)} - i\frac{\partial}{\partial \mathrm{Re}(x)}\Big)h(x), \tag{5.72}$$

it gives

$$\frac{\partial}{\partial \Delta\omega}h(x(\Delta\omega)) = x'(\Delta\omega)h'(x(\Delta\omega)) + x'(\Delta\omega)^* h'(x(\Delta\omega))^*. \tag{5.73}$$

In our peculiar case, we have

$$f(t) = \sqrt{\kappa}e^{-\left(\frac{\kappa}{2}+i\omega_0\right)t} \text{ and} \tag{5.74}$$

$$f_{\Delta\omega}(t) = \sqrt{\kappa}e^{-\left(\frac{\kappa}{2}+i(\omega_0+\Delta\omega)\right)t}. \tag{5.75}$$

We can write

$$(\hat{\tau}f \cdot \hat{\tau}f) = \frac{2}{\kappa^2} \tag{5.76}$$

$$(\hat{\tau}f \cdot f) = \frac{1}{\kappa}, \tag{5.77}$$

which gives

$$x'(\Delta\omega) = \frac{i}{\kappa}\big(x(\Delta\omega) + b\big). \tag{5.78}$$

We can now numerically evaluate the Fisher information for any pure state preparation in the mode f. As for the mutual information, we computed it in the qubit case with finite quantum efficiency $\eta \in [0, 1]$ and $\theta \in [0, \pi]$. The result is shown in Fig. 5.12. Strikingly, the FI behaves just the same as the mutual information. We did the same figures as in Fig. 5.4, and we see that the plots are qualitatively the same. We could even compare the optimal angle θ_{opt} and see that they are the same, up to the statistical noise in the estimation of the mutual information. The same discussion about the moments of the heterodyne measurements in Sec. 5.2.1 carrying the information as a function of η can be conducted here with Fig. 5.12d. For a Fock state $|1\rangle$ prepared in the mode, the heterodyne measurement performs as well as the FRPD and half as well as an ideal measurement achieving the QFI. This is the same as the $\kappa \ll \chi$ regime studied in 5.2.2.

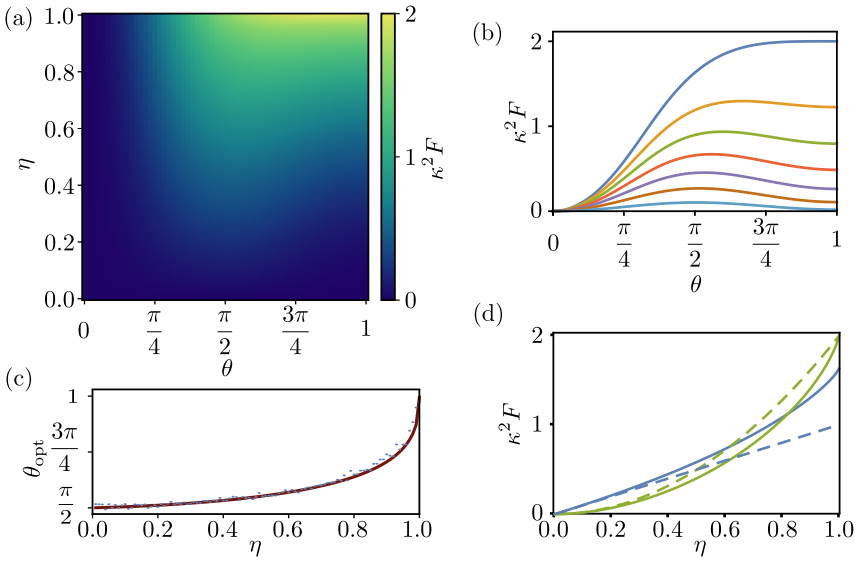


Figure 5.12: Fisher information as a function of the quantum efficiency and the kick amplitude on the qubit. (a) Map of the Fisher information F of a heterodyne measurement as a function of θ and η. (b) Cuts (a) along the lines $\eta = 0.1, 0.25, 0.4, 0.55, 0.7, 0.85$ and 1 corresponding to increasing Fisher information. (c) Dark red solid line: Optimal angle θ_{opt} as a function of η for the Fisher information. Blue dots: θ_{opt} for the frequency-resolved mutual information (see Fig. 5.4. (d) Fisher information F as a function of the quantum efficiency for $\theta = \pi/2$ (blue solid line) and $\theta = \pi$ (green solid line). The blue dashed line is a linear fit for small η, and the green dashed line shows the function $\kappa^2 F = 2\eta^2$.

5.4.4 *Homodyne measurement*

We now derive the FI for a homodyne measurement performed on a mode that is hermitian symmetric with respect to ω_m. For simplicity, we consider that $\omega_m = 0$. The condition that the mode is symmetric translates into $f^* = f$: it is a real-valued function. What we want to compute is $F(\Delta\omega = 0)$, *i.e.* the FI when the measurement frequency ω_m is already equal to ω_0. As for the QFI, it makes sense when considering the asymptotic behavior of an adaptive scheme where ω_m is adapted at each iteration to the best estimation of $\Delta\omega$ so far.

The goal here is to express a quasi-likelihood function for the measurement record the same way as in Sec. 5.1.1. To do so, we will see that we can integrate the measurement record against well-chosen normalized real functions g_n and compute the probability density of the outcomes x_{g_n}. As we saw in Sec. 3.1.1, it corresponds to sampling the Wigner functions of the modes represented by the elements g_n. Therefore, we have to understand how to transform the Wigner function $\boldsymbol{W}_0$ of the full transmission line under the frequency translation $\hat{\mathcal{T}}_{\delta\omega}$. We will do it at first order in $\delta\omega$, which allows us to express the frequency translation as a rotation in a subspace of dimension 2 only. This leads to the definition of a simple pseudo-likelihood that we can use to compute the Fisher information.

For a small frequency translation by $\delta\omega$, we can linearize the unitary operator $\hat{\mathcal{T}}_{\delta\omega}$ around $\Delta\omega = 0$, which gives:

$$\hat{\mathcal{T}}_{\delta\omega} = \mathbb{1} - i\delta\omega\hat{\boldsymbol{\tau}} \tag{5.79}$$

The corresponding action on f reads

$$\mathcal{T}_{\delta\omega} f = f - i\delta\omega\hat{\tau} f. \tag{5.80}$$

It is sufficient to study the effect of $\mathcal{T}_{\delta\omega}$ on $\mathrm{Span}(\{\hat{\tau}^n f\})$. From now on, we will work in this subspace.

In order to compute the action of $\mathcal{T}_{\delta\omega}$ in this space, we first build $\{g_n\}$ the orthonormal basis obtained by applying the Gram-Schmidt algorithm on the family $\{\hat{\tau}^n f\}_{n\in\mathbb{N}}$. By construction, we have the following properties:

$$g_0 = f \tag{5.81}$$

$$g_n \cdot g_m = \delta_{mn} \tag{5.82}$$

$$(\hat{\tau} g_m) \cdot (g_{m+k}) = (g_m) \cdot (\hat{\tau} g_{m+k}) = 0 \text{ if } k \geq 2. \tag{5.83}$$

The latter property comes from the fact that $(\hat{\tau} g_m) \in \mathrm{Span}(\{\hat{\tau}^n f\}_{n\leq m+1})$ by construction. In general, this family has an infinite cardinal number, which we will consider true in the following.

The action of $\mathcal{T}_{\delta\omega}$ on $\mathrm{Span}(\{\hat{\tau}^n f\})$ can then simply be expressed by a change of basis

$$g_n \to \mathcal{T}_{\delta\omega} g_n. \tag{5.84}$$

We can check that $\{\mathcal{T}_{\delta\omega} g_n\}$ remains an orthonormal basis.

$\mathcal{T}_{\delta\omega}$ is a rotation in the vector space spanned by $\{g_n\}$. In this basis, its matrix elements are of the form

$$(g_m \cdot \mathcal{T}_{\delta\omega} g_n) \tag{5.85}$$

which, at first order in $\delta\omega$ becomes

$$g_m \cdot (\mathcal{T}_{\delta\omega} g_n) = -i\delta\omega g_m \cdot (\hat{\tau} g_n). \tag{5.86}$$

Using Eq. (5.83), the rotation matrix $\mathbf{R}_{\delta\omega}$ describing $\mathcal{T}_{\delta\omega}$ in the basis $\{g_n\}$ finally reads, at first order in $\delta\omega$:

$$\mathbf{R}_{\delta\omega} \simeq \mathbb{1} - i\delta\omega \times \begin{pmatrix} \big(g_0 \cdot (\hat{\tau} g_0)\big) & \big(g_0 \cdot (\hat{\tau} g_1)\big) & 0 & 0 & \dots \\ \big(g_1 \cdot (\hat{\tau} g_0)\big) & \big(g_1 \cdot (\hat{\tau} g_1)\big) & \big(g_1 \cdot (\hat{\tau} g_2)\big) & 0 & \dots \\ 0 & \big(g_2 \cdot (\hat{\tau} g_1)\big) & \big(g_2 \cdot (\hat{\tau} g_2)\big) & \big(g_2 \cdot (\hat{\tau} g_3)\big) & \dots \\ 0 & 0 & \big(g_3 \cdot (\hat{\tau} g_2)\big) & \big(g_3 \cdot (\hat{\tau} g_3)\big) & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{5.87}$$

At first order, it thus only mixes the mode g_m with the mode g_{m+1} and g_{m-1}.

This matrix gives the coordinate transform to express how the full Wigner function of the line $\mathbf{W}_0$ transforms under the frequency translation $\hat{\mathcal{T}}_{\delta\omega}$. It simply reads

$$\mathbf{W}_0(\boldsymbol{\beta}) \xrightarrow{\hat{\mathcal{T}}_{\delta\omega}} \mathbf{W}_{\delta\omega} = \mathbf{W}_0(\mathbf{R}_{\delta\omega}^{-1}\boldsymbol{\beta}), \tag{5.88}$$

with $\boldsymbol{\beta} = (\beta_0, \beta_1, ...)$ its coordinates (see Appendix. B). At this point, the Wigner is an infinite-dimensional quasidistribution. To work it out, we use the fact that the only state that is not in the vacuum state is the one represented by f. As a consequence, under the rotation, and at first order in $\delta\omega$, the only two modes that are not in the vacuum state are the ones represented by g_0 and g_1. The Wigner function reads

$$\mathbf{W}_0(\boldsymbol{\beta}) = W_f(\beta_0) \prod_{n\geq 1} W_{|0\rangle}(\beta_n) \tag{5.89}$$

before applying the rotation. Applying the rotation, it reads

$$\mathbf{W}_{\delta\omega}(\boldsymbol{\beta}) = \mathbf{W}_0(\boldsymbol{R}_{\delta\omega}^{-1}\boldsymbol{\beta}) = \mathbf{W}_{\delta\omega}^{01}(\beta_0, \beta_1) \prod_{n\geq 2} W_{|0\rangle}(\beta_n) + o(\delta\omega), \tag{5.90}$$

where $\boldsymbol{W}_{\delta\omega}^{01}$ is the two-dimension Wigner function of the modes represented by g_0 and g_1. Thus, we are only interested in the coordinates β_0 and β_1 of $\mathbf{W}_0(\mathbf{R}_{\delta\omega}^{-1})$. As all the terms involving β_n for $n \geq 2$ can be factored out as $W_{|0\rangle}(\beta_n)$, the change of coordinate $\boldsymbol{\beta} \to \mathbf{R}_{\delta\omega}^{-1}\boldsymbol{\beta}$ for the Wigner function $\boldsymbol{W}_0$ is equivalent to the following change of coordinates **in this peculiar case where only the state represented by g_0 is not the vacuum state**:

$$\mathbf{R}'_{\delta\omega} \simeq \mathbb{1} - i\delta\omega \begin{pmatrix} \begin{pmatrix} g_0 \cdot (\hat{\tau} g_0) \\ g_1 \cdot (\hat{\tau} g_0) \end{pmatrix} & \begin{pmatrix} g_0 \cdot (\hat{\tau} g_1) \\ g_1 \cdot (\hat{\tau} g_1) \end{pmatrix} & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots & \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{5.91}$$

This can be understood by a geometrical argument: when only the mode represented by f is populated, the Wigner function is invariant under the action of $\boldsymbol{R}_{\delta\omega}$ almost everywhere, except in the manifold $\mathrm{Span}(g_0, g_1)$. We thus restrict our description to this space and write $\mathbf{S}_{\delta\omega}$ the restriction of $\mathbf{R}'_{\delta\omega}$ to this subspace as:

$$\mathbf{S}_{\delta\omega} = \mathbb{1} - i\delta\omega \begin{pmatrix} \begin{pmatrix} g_0 \cdot (\hat{\tau} g_0) \\ g_1 \cdot (\hat{\tau} g_0) \end{pmatrix} & \begin{pmatrix} g_0 \cdot (\hat{\tau} g_1) \\ g_1 \cdot (\hat{\tau} g_1) \end{pmatrix} \end{pmatrix}. \tag{5.92}$$

The expression of $\boldsymbol{W}_{\delta\omega}^{01}$ is thus

$$\boldsymbol{W}_{\delta\omega}^{01} = \boldsymbol{W}_0^{01}(\boldsymbol{S}_{\delta\omega}^{-1}(\beta_0, \beta_1)) \tag{5.93}$$

The transformation of $\boldsymbol{W}_0^{01} \to \boldsymbol{W}_{\delta\omega}^{01}$ in this subspace is pictured in Fig. 5.13 for a coherent state $|\alpha\rangle$ prepared in the mode defined by $f = g_0$ (and the vacuum state in g_1), shown in (a) and (c). Translating

the mode in frequency corresponds to rotating the Wigner function in a four-dimensional space, which we can visualize by plotting the Wigner function of the same modes defined by g_0 and g_1 ((b) and (d)). The distinguishability of the states depends on the distances in phase space, given by $\delta\omega|g_0 \cdot \hat{\tau} g_0| = \delta\omega\overline{\tau}$ ($\times 2$ between $+\delta\omega$ and $-\delta\omega$) in the mode defined by $f = g_0$, and $\delta\omega|g_1 \cdot \hat{\tau} g_0|$ ($\times 2$ between $+\delta\omega$ and $-\delta\omega$) in the mode defined by g_1. Note that since the two modes are in a coherent state, the state $|\psi_{01}\rangle$ describing the two modes is separable and reads (still at first order in $\delta\omega$)

$$|\psi_{01}\rangle \simeq |(1 - i\delta\omega\overline{\tau})\alpha\rangle \otimes \left|-i\delta\omega\big(g_1 \cdot (\hat{\tau} g_0)\big)\alpha\right\rangle, \tag{5.94}$$

and we can factorize $\boldsymbol{W}_{\delta\omega}^{01}$ into

$$\begin{aligned} \boldsymbol{W}_{\delta\omega}^{01}(\beta_0, \beta_1) &= W_{g_0}(\beta_0) W_{g_1}(\beta_1) \\ &= W_{|(1-i\delta\omega)\alpha\rangle}(\beta_0) W_{|(-i\delta\omega)\alpha\rangle}(\beta_1). \end{aligned} \tag{5.95}$$

This is not the case when working with non-coherent states.

Coming back to the measurement record, we can integrate it against $f^* = g_0^*$ and g_1^*, which gives measurement outcomes x_{g_0} and x_{g_1}. The law of these two measurement records is given by the joint probability density $p_{\delta\omega}(x_{g_0}, x_{g_1})$ that reads

$$p_{\delta\omega}(x_{g_0}, x_{g_1}) = \iint \mathbf{W}_{\delta\omega}^{01}(x_{g_0} + ip_0, x_{g_1} + ip_1) \mathrm{d}p_0 \mathrm{d}p_1. \tag{5.96}$$

We can now apply the same arguments as in Sec. 5.1.1 and write a pseudo-likelihood function

$$\mathcal{L}(\delta\omega|x) = \mathcal{L}(\delta\omega|x_{g_0}, x_{g_1}) = \frac{p_{\delta\omega}(x_{g_0}, x_{g_1})}{p_{\mathrm{vac}}(x_{g_0}, x_{g_1})}, \tag{5.97}$$

where p_{vac}

$$p_{\mathrm{vac}} = \iint W_{|0\rangle}(x_{g_0} + ip_0) W_{|0\rangle}(x_{g_1} + ip_1) \mathrm{d}p_0 \mathrm{d}p_1 \tag{5.98}$$

is defined the same way as p_0 but with the two modes defined by g_0 and g_1 in the vacuum state. We now have an explicit way to compute the FI:

$$F(0) = \iint \left(\frac{\partial \log\Big(\mathcal{L}(\delta\omega|x_{g_0}, x_{g_1})\Big)}{\partial\delta\omega} \right)^2 p_0(x_{g_0}, x_{g_1}) \mathrm{d}x_{g_0} \mathrm{d}x_{g_1} \tag{5.99}$$

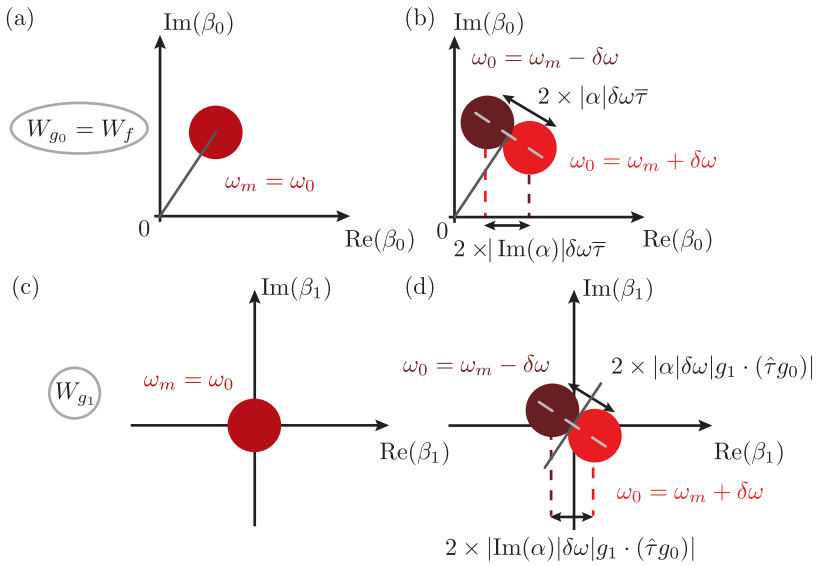


Figure 5.13: Scheme of the transformation of the two-mode Wigner function for a small detuning $\delta\omega$ between the homodyne detection frequency and the signal frequency. (a) and (c): Wigner functions W_f and W_{g_1} of the modes defined by $f = g_0$ and g_1 when the mode defined by f is in a coherent state $|\alpha\rangle$. (b) and (d) Wigner functions of the same modes after a time translation by $+\delta\omega$ (lighter red) or $-\delta\omega$ (darker red). The direction of the displacement in phase space (light grey dashed line) is orthogonal to the direction set by α (dark grey). A homodyne detection has measurement outcomes distributed following the marginals of $\mathbf{W}^{01}_{\delta\omega}$ as a function of Re(β_0) and Re(β_1).

where $p_0(x_{g_0}, x_{g_1})$ factorizes as

$$p_0(x_{g_0}, x_{g_1}) = \int W_f(x_{g_0} + ip_0)\mathrm{d}p_0 \int W_{|0\rangle}(x_{g_1} + ip_1)\mathrm{d}p_1. \tag{5.100}$$

We can compute it for a few examples in the case $f(t) = \sqrt{\kappa}e^{-\frac{\kappa}{2}t}$ for $t \geq 0$ and 0 otherwise. In this case, the g_n are expressed as a function of the Laguerre polynomials L_n of order n:

$$g_n(t) = \begin{cases} \sqrt{\kappa}L_n(\kappa t)e^{-\frac{\kappa}{2}t} \text{ if } t \geq 0 \\ 0 \text{ otherwise.} \end{cases} \tag{5.101}$$

$\boldsymbol{S}_{\delta\omega}$ reads

$$\mathbf{R}'_{\delta\omega} \simeq \mathbb{1} - i\delta\omega \begin{pmatrix} 1/\kappa & -1/\kappa \\ -1/\kappa & 3/\kappa \end{pmatrix}. \tag{5.102}$$

We can now apply the formula Eq. (5.99) for various states prepared in the mode f. We can first take a look at coherent states. Preparing a state $|\alpha\rangle$ in the mode defined by f gives a Fisher information

$$F(0) = 8\frac{\mathrm{Im}(\alpha)^2}{\kappa^2}. \tag{5.103}$$

Interestingly, when α is imaginary, $F(0)$ is equal to the QFI. This is expected, as the direction of displacement in phase space (Fig. 5.13) is the same as that of the homodyne measurement. This is similar to the ideal dispersive readout of a qubit discussed in Sec. 3.3.4, for which all the available information is retrieved. However, for α real, it vanishes. It can be understood by the fact that the marginal distributions of $\mathrm{Re}(\beta_0)$ and $\mathrm{Re}(\beta_1)$ do not change at first order, as the rotation in phase space corresponds to a vertical displacement. This dependence in $\mathrm{Im}(\alpha)$ is pictured in Fig. 5.13.

It does not mean that an estimator would not converge, but rather that it would converge more slowly than in the central limit theorem. This kind of critical point has been discussed in [184]. The convergence, instead of scaling like $1/N$ for the variance, scales at most like $1/\sqrt{N}$.

For a displaced squeezed state, it gives

$$F(0) = 8\frac{e^{3r}\cosh(r)\,\mathrm{Im}(\alpha)^2}{\kappa^2}. \tag{5.104}$$

It is thus equal to only one term of the QFI in Eq. (5.52), the one proportional to $\mathrm{Im}(\alpha)$. The homodyne measurement is blind to the component along $\mathrm{Re}(\alpha)$ and to the part that is intrinsic to a squeezed state (the first part in the QFI). We see once again that the homodyne detection needs a nonzero average value along the imaginary axis to extract information.

5.4.5 *Qubit case: summary*

We can now compare the FI associated to each measurement and the QFI in the qubit case (as in the experiment of Chap. 4). The homodyne

measurement was taken in the optimal direction. Fig. 5.14 shows them as a function of $p(1)$. For small $p(1)$, we see that the homodyne measurement reaches the QFI, as expected from the fact that the state can be assimilated to a very small coherent state. The heterodyne measurement reaches only half of it, and the FRPD only a fourth. For increasing $p(1)$, the homodyne measurement performs worse and worse until it reaches 0 for $p(1) = 1$. Once again, we see that the homodyne measurement can work well only if the Wigner function of the state has a non-zero average imaginary part, which is not the case for $p(1) = 1$. This can be interpreted in the same way as for the case of a coherent state: when the measurement frequency $\omega_m/2\pi$ is aligned with the frequency of the mode of interest, if there is no average imaginary part, the likelihood is the same whether we are slightly detuned by $\delta\omega$ or by $-\delta\omega$. This degeneracy disappears as soon as we set $\Delta\omega \neq 0$. Increasing $|\Delta\omega|$ would set the measurement in-between homodyne and heterodyne, with Fisher information continuously shifting from the blue curve to the yellow curve in Fig. 5.14.

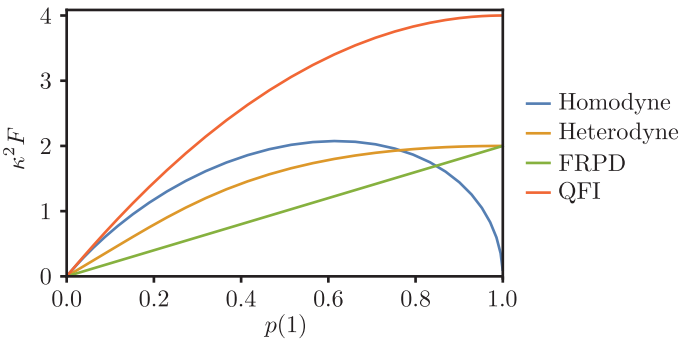


Figure 5.14: FI for the homodyne and heterodyne detection, for an FRPD, and the QFI as a function of $p(1)$.

5.5 CONCLUSIONS

To give more theoretical insights about the experiment presented in Chap. 4, we used the framework presented in Chap. 2 and Chap. 3 to compare the measurement based on heterodyne detection to the use of a frequency-resolved photon detector, and to the accessible

information. We could theoretically explore the full range of possible parameters χ and Γ_1 that the experiment could have taken.

In a second part, we used the experiment as a base to discuss the notion of measurement rate when there are more than two possible outcomes. In particular, we distinguished the measurement rate, which characterizes the speed at which an estimator converges toward the right outcome, and the information rate, which is simply the rate at which information is transmitted via the measurement setup. As these two values could be unified in one "measurement rate" in the case of only two possible outcomes, it is no longer the case when considering more than two outcomes.

In a third part, we shifted the point of view to understand the experiment of Chap. 4 as a particular case of parameter estimation task. Generalizing it to a case where the parameter to estimate can take continuous values, we could compare heterodyne and homodyne detection, a frequency-resolved photon detector, and an ideal optimal measurement. It reveals which properties of the propagating states are best exploited by each of these measurements.

Part IV

NEURAL-NETWORK ASSISTED STATE PREPARATION

6

CAT STATE PREPARATION USING NEURAL NETWORKS

A part of this PhD work was dedicated to finding applications of machine learning to quantum experiments in the context of the European project ARTEMIS. This project was originally born from the idea of finding complex feedback procedures to stabilize Fock states using the measurement scheme developed in Chap. 4 [185]. It was a very ambitious project that we could not realize fully. However, we could still demonstrate in Sec. 4.9.2 that a neural network can process measurement records just as well as an optimal procedure. In addition, the following article [7], currently under review, demonstrates neural network-assisted cat state preparation and is reproduced here in full.

6.1 INTRODUCTION

Quantum information processing demands exquisite control over quantum systems. While direct optimization methods have demonstrated success in quantum control, they often necessitate access to gradient information [3, 186–188], and require to be run for each task. Much more flexibility is promised by the novel powerful tools emerging from the domain of machine learning and artificial intelligence [189–191]. Model-free reinforcement learning considers the experiment or a simulation as a black box and trains the network on it [5, 192–203]. A few experiments have demonstrated this strategy for optimized qubit state readout and initialization [166, 167, 204], generation of unitaries [205], demonstrating dynamical decoupling [206], improved measurements [207, 208], quantum state tomography [209], Wigner negativity detection [170], Bose-Einstein condensate preparation [210], boson sampling validation [211], quantum control in trapped ions [212], or optimized quantum error correction [213]. Yet many experiments can be modeled completely with only a few *a priori* undetermined or evolving parameters, employing a differentiable simulation. This

can make it feasible to take gradients directly through simulations, the most efficient version of model-based reinforcement learning [193, 214–216]. Neural-network-based variants of this approach have yet to be demonstrated in experiments. In this domain, it has recently been shown theoretically that one might train a network once for a whole family of tasks and then use it to rapidly generate control sequences for any desired tasks selected from that family [217], making the whole process even more efficient, exploiting the general concept of transfer learning (learning for some tasks and benefiting from this for other tasks).

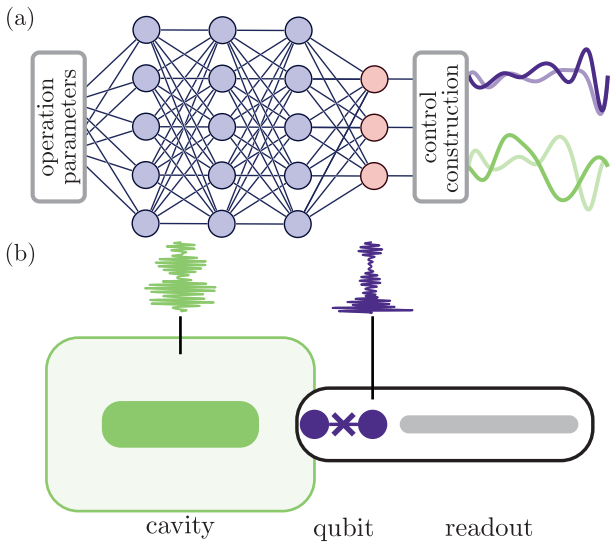


Figure 6.1: Principle of the experiment. (a) A neural network takes the input parameters and outputs optimized control pulses characterizing a quantum operation.(b) These pulses drive a cavity (green) and its dispersively coupled qubit (purple) in order to prepare a desired quantum state. A readout resonator is used to measure the qubit state and perform Wigner tomography of the cavity.

In this work, we demonstrate model-based reinforcement learning for the preparation of quantum states in a cavity. We focus on arbitrary states within a specific class, namely two-component cat states in a cavity. This choice stems from the potential of cat qubits for quantum information processing, particularly in quantum sensing and

error correction [140, 218–221]. Breakthroughs in cat qubit experimentation [222–231] have prompted both academic and commercial development of quantum processors leveraging this technology. We designed and trained a neural network to directly generate control pulses simultaneously applied to the cavity and a dispersively coupled qubit (Fig. 6.1a). These control sequences are decomposed into a basis set of a few smooth basis functions, ensuring compatibility with existing real-time control hardware. We describe how to implement such a controller in an experiment and show how to estimate the preparation fidelity as fast as possible. We show that the network, once trained, is able to generate control sequences about five orders of magnitude faster than an established gradient-based technique. The rapidity of the control generation offered by the neural network could also prove instrumental in the case of real-time feedback protocols, needed in other tasks [5, 167, 195, 197, 200–202, 232].

In this work, we demonstrate model-based reinforcement learning for the preparation of quantum states in a cavity. We focus on arbitrary states within a specific class, namely two-component cat states in a cavity. This choice stems from the significant promise of such states for various quantum information processing applications such as quantum error correction and sensing [140]. We designed and trained a neural network to directly generate control pulses simultaneously applied to the cavity and a dispersively coupled qubit (Fig. 6.1a). These control sequences are decomposed into a basis set of a few smooth basis functions, ensuring compatibility with existing real-time control hardware. We describe how to implement such a controller in an experiment and show how to estimate the preparation fidelity as fast as possible. We show that the network, once trained, is able to generate control sequences about five orders of magnitude faster than an established gradient-based technique. The rapidity of the control generation offered by the neural network could also prove instrumental in the case of real-time feedback protocols, needed in other tasks [5, 167, 195, 197, 200–202, 232].

A schematics of the experimental system is shown in Fig. 6.1b (see also Sec. 6.10). The cavity whose state preparation is optimized by the neural network is the fundamental mode of a high-Q coaxial $\lambda/4$ resonator made of aluminum (green). Its frequency is $\omega_c/2\pi =$

4.628 GHz and its decay time is $T_c = 225\ \mu$s. The cavity is dispersively coupled with a shift $\chi/2\pi = 238.5$ kHz to a transmon qubit (purple) resonating at $\omega_q/2\pi = 3.235$ GHz, and whose decay time is $T_1 = 35\ \mu$s and decoherence time T_2 varies between 20 and 60 μs. The transmon qubit is also coupled to a readout mode composed of a $\lambda/2$ resonator connected to the readout line through a Purcell bandpass filter. Both the qubit and the cavity can be directly driven through distinct microwave ports.

In the interaction picture, the Hamiltonian of the qubit and cavity reads

$$\hat{H}(t)/\hbar = -\chi \hat{a}^\dagger \hat{a} |1\rangle\langle 1|_q + \left[\varepsilon_c(t)\hat{a}^\dagger + \varepsilon_q(t)\hat{\sigma}_+ + \text{h.c.}\right] , \tag{6.1}$$

where $\varepsilon_c(t)$ and $\varepsilon_q(t)$ are the complex-valued time-dependent external fields driving the cavity and the qubit in their rotating frame, respectively. $\hat{a}$ is the annihilation operator of the cavity, $|1\rangle_q$ is the first excited state of the transmon and $\hat{\sigma}_+ = \hat{\sigma}_-^\dagger = |1\rangle\langle 0|_q$ is the qubit raising operator. The system relaxation is included perturbatively in the simulations, as discussed in Sec. 6.9.

The preparation sequence is initiated when the system is in its ground state $|\psi(0)\rangle = |0\rangle_q \otimes |0\rangle_c$. Both qubit and cavity are then driven during a fixed time T in order to obtain a target state in the cavity. In this work we focus on target cat states of the form

$$|C_\alpha^\varphi\rangle \propto |\alpha\rangle + e^{-i\varphi}\,|-\alpha\rangle \tag{6.2}$$

where $|\pm\alpha\rangle$ are coherent states of amplitude α and φ is a phase. Additionally, we require the qubit to come back to the ground state, so that the full target state is

$$|\mathcal{T}_\alpha^\varphi\rangle = |0\rangle_q \otimes |C_\alpha^\varphi\rangle_c \, . \tag{6.3}$$

Under the action of $\hat{H}(t)$ in presence of decoherence, the density matrix of the system $\hat{\rho}(t)$ evolves from $\hat{\rho}(0) = |\psi(0)\rangle\langle\psi(0)|$ to $\hat{\rho}(T)$. The fidelity to the target state is defined as $\mathcal{F}(\alpha,\varphi) = \langle\mathcal{T}_\alpha^\varphi|\,\hat{\rho}(T)\,|\mathcal{T}_\alpha^\varphi\rangle$. The optimization problem we set to solve consists in identifying the two complex (and thus four real) control functions $\varepsilon_c(t)$ and $\varepsilon_q(t)$ that maximize the fidelity $\mathcal{F}(\alpha,\varphi)$ after a driving time T for every $(\alpha,\varphi) \in \mathcal{S} = [0,4] \times [0,\pi]$.

6.2 NEURAL NETWORK

To this end, we employ a neural network (NN). The NN input consists of the real parameters α and φ characterizing the target state $|\mathcal{T}_\alpha^\varphi\rangle$. The NN output is a set of 9 expansion coefficients for each of the 4 real control fields in a B-spline basis (see Sec. 6.6). The encoding of the pulses is thus very light, which potentially enables much faster communication with the controller. As an example, loading one full pulse sequence of 2 μs on our OPX necessitates 4×2000 points, which takes approximately 450 ms, whereas loading only the 4×9 corresponding parameters takes 24 ms, limited by the network communication time, which is already an order of magnitude faster.

In order to apply the reinforcement learning approach, the NN is included in the data processing pipeline shown in Fig. 6.2a. Each real control field is constructed by summing up the B-splines weighted by the output coefficients of the NN. In order to avoid the computational overhead associated with the implementation of a full density matrix $\hat{\rho}(t)$ simulation, the coefficients are passed to a simpler Schrödinger equation solver, which computes the system state evolved under these control fields $|\psi(\alpha, \varphi; t)\rangle$. The fidelity $\mathcal{F}(\alpha, \varphi)$ can be approximated by

$$\mathcal{F}(\alpha, \varphi) \simeq \left|\langle \mathcal{T}_\alpha^\varphi | \psi(\alpha, \varphi; t = T) | \mathcal{T}_\alpha^\varphi | \psi(\alpha, \varphi; t = T)\rangle\right|^2 - \Delta F \ , \tag{6.4}$$

where ΔF is a first-order correction to the fidelity due to the decoherence (T_1, T_2, T_c, see Sec. 6.9). We then use the infidelity as the loss function that has to be minimized:

$$\mathcal{L}(\alpha, \varphi) = 1 - \mathcal{F}(\alpha, \varphi) \ . \tag{6.5}$$

All these operations are differentiable, so gradients of the loss with respect to the NN parameters can be computed through the complete pipeline comprising the network, control construction and simulation, enabling thus gradient-based NN training. The NN is trained on target cat states $|C_\alpha^\varphi\rangle$ sampled randomly within our parameter space $\mathcal{S}$. After a successful training, the NN is able to generate control fields which produce the cat state for *any* $(\alpha, \varphi) \in \mathcal{S}$ in a preparation time $T = 2$ μs, which was selected to be close to π/χ.

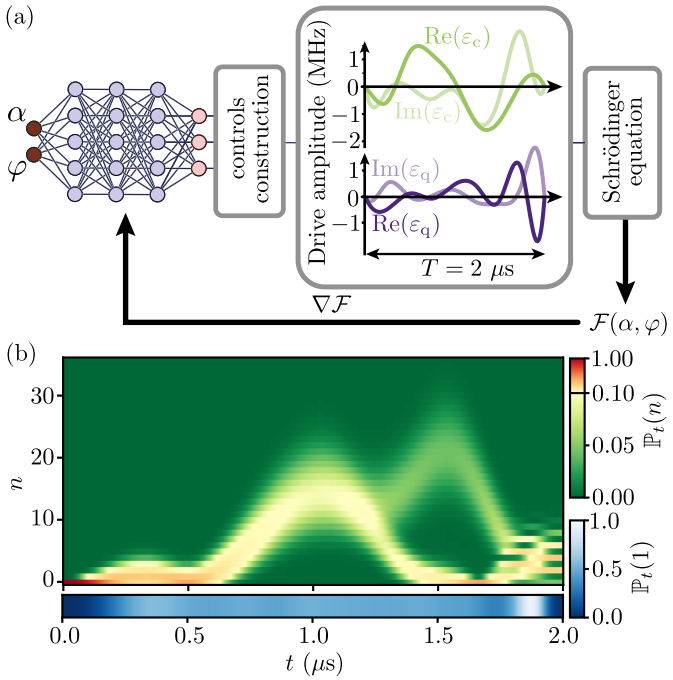


Figure 6.2: Data processing, training pipeline and simulation example. (a) Schematic representation of the data processing and training pipeline. Gradients can be taken through the network, control construction and simulation to update the network weights while optimizing the fidelity $\mathcal{F}$ via gradient ascent. The output of the network is a vector of 36 real numbers which is then converted into control pulses $\varepsilon_c(t)$ and $\varepsilon_q(t)$ with the goal of preparing a cat state $|C_\alpha^\varphi\rangle \propto |\alpha\rangle + e^{-i\varphi}|-\alpha\rangle$ with $\varphi \in [0, 2\pi]$ and $\alpha \in [0, 4]$. (b) Simulated probability distribution of the photon number (top) and qubit excitation (bottom) as a function of time for the control fields shown in (a), which are generated by the neural network when $\alpha = 2$ and $\varphi = 0$. The fidelity of the predicted final state to $|C_\alpha^\varphi\rangle$ is $\mathcal{F}(2, 0) = 94$ %.

We stress that the described machine learning method to train our NN is model-based reinforcement learning and should not be confused with supervised learning. The method belongs to the reinforcement learning branch since no "correct" answers are available for the searched control signals, as would be the case in supervised learning. Moreover, for a given target quantum state, the optimal control pulses are gener-

ally speaking not unique and can differ significantly [217]. The typical reinforcement learning components are identified in our approach as follows. The agent applies control signals (action) and observes the fidelity between the constructed quantum state and the target state (reward). The NN encodes the agent's choice of a particular control pulse (policy), which is iteratively improved based on the rewards. In this work, we employ the physical model which enables us to improve the policy based on the reward gradients that are directly computed through the full processing pipeline including both the physical simulation as well as the neural network (Fig. 6.2a). We note that in our approach the policy is deterministic and yields directly the control signals. In contrast, the standard model-free policy gradient [233] needs to use less efficient stochastic policies for exploration of the action space in order to estimate the reward gradients.

The numerical modeling is performed on a time grid with 40 discretization intervals at the NN training stage and 200 intervals at the subsequent testing stage. The highest included Fock state $N_{\max}$ for the cavity (determining also the output size of the pipeline in Fig. 6.2a) is varied during the NN training between $N_{\max} = 20$ and $N_{\max} = 60$, and fixed to the value $N_{\max} = 70$ at the testing stage. We made sure that this choice is sufficient for accurate modeling of all quantum states considered in this work. We use a NN of the usual dense (deep forward, multilayer perceptron) architecture with 3 hidden layers of the size 30, 60 and 30 neurons, respectively (in total approx. 5000 trainable parameters). It is trained using the Adam algorithm [175] on batches sampled from $\mathcal{S}$. The batch size is gradually increased with the training stages. We use uniform sampling for φ, whereas α's are sampled with a probability linearly growing with larger values. This accounts for the fact that the NN training for the more interesting larger amplitudes α is more difficult. Additionally, at the NN training stage both α and φ intervals are slightly extended beyond the edges in order to avoid the presence of boundary effects in $\mathcal{S}$. The pipeline discussed in the text above and shown in Fig. 6.2a was implemented in Python using the JAX [234] and FLAX [235] libraries. We additionally benchmark our computations by performing the same system modeling including dissipation and decoherence with the QuTiP package [32,

236], to extract fidelities. In appendices 6.7 and 6.8 we provide further details on the NN training and performance study, respectively.

An example of the result of such simulations is shown in Fig. 6.2b, where the evolution of the cavity photon number distribution and qubit excitation is plotted as a function of time when the neural network optimizes the creation of $|\mathcal{C}_2^0\rangle$. The qubit and cavity get quickly entangled leading to $\mathbb{P}_t(1_\mathrm{q}) \approx 1/2$ for the qubit as the number of photons in the cavity rises. The cavity goes through states with large photon numbers and a superposition of low and large numbers arises around 1.6 μs. At the end, the parity of the photon number seems to flip as well as the qubit state. Since large photon numbers are known to lead to dynamics that are not captured by a simple model [96, 161], it is desirable to keep them small. Here, no constraint was imposed on the maximal photon number during training, but the chosen B-spline basis with broad basis functions indirectly limits that number. Indeed finely controlling a system that evolves at a speed χn in phase space requires pulses with a bandwidth of at least χn. The bandwidth of the B-spline scales inversely with their number. The chosen basis thus embeds a hidden cost naturally preventing the network from finding solutions with too high photon numbers. Tuning the number of elements in the basis then allows to control this limitation, that can be pushed to prepare cat states with a higher number of photons.

6.3 EXPERIMENTAL RESULTS

We now test the NN controller on the experiment for several targeted cat states $|C_\alpha^\varphi\rangle$. After the cavity and the qubit are prepared in their ground states (see Sec. 6.11), we apply the pulse sequence generated by the neural network on the cavity and qubit drive lines. A Wigner tomography is then performed using the same qubit in order to directly measure the Wigner function $W_\alpha^\varphi(\beta)$ of the cavity after the pulse sequence [144, 237]. It consists in mapping the displaced parity operator $\hat{\Pi}(\beta) = \hat{D}(\beta)e^{i\pi\hat{a}^\dagger\hat{a}}\hat{D}(-\beta)$ onto the qubit state before reading it out. The Wigner function is then given by $W_\alpha^\varphi(\beta) = 2\langle\hat{\Pi}(\beta)\rangle/\pi$. We introduce $W_\mathrm{exp}(\beta)$, which is obtained by averaging the measurement

outcomes of the observable $2\hat{\Pi}(\beta)/\pi$. In Fig. 6.3a, it is averaged over 1000 realizations per pixel (with 21×51 pixels in total) for $|C_1^0\rangle$.

From the Wigner function $W_\alpha^\varphi(\beta)$, we can compute the fidelity to the target cat state $|C_\alpha^\varphi\rangle$ as (see Sec. 6.12.1)

$$\mathcal{F}(\alpha,\varphi) = \pi \int_{\mathbb{C}} W_\alpha^\varphi(\beta) W_{|C_\alpha^\varphi\rangle\langle C_\alpha^\varphi|}(\beta) \mathrm{d}\beta \tag{6.6}$$

with $W_{|C_\alpha^\varphi\rangle\langle C_\alpha^\varphi|}(\beta) = \frac{2}{\pi}\mathrm{Tr}(\hat{\Pi}(\beta)|C_\alpha^\varphi\rangle\langle C_\alpha^\varphi|)$ the Wigner function of the target state. Based on this formula, we can build an experimental estimator of $\mathcal{F}(\alpha,\varphi)$ (details in Sec. 6.12.1).

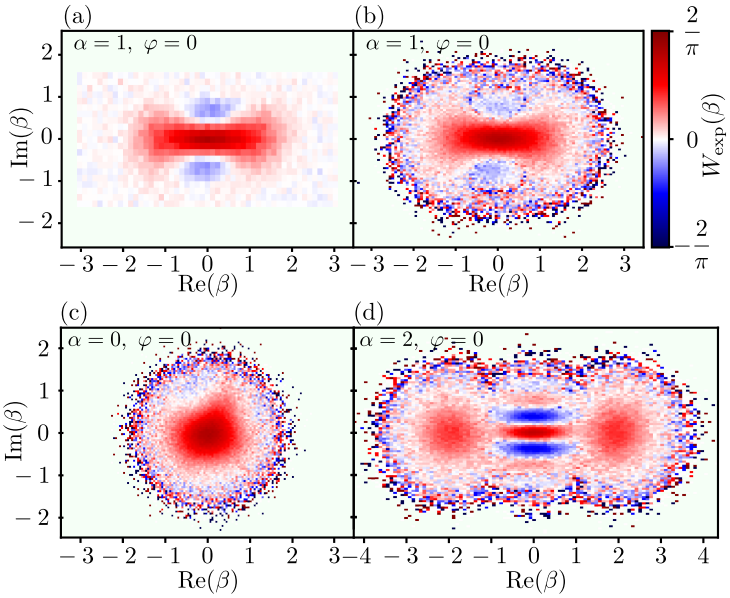


Figure 6.3: Experimental Wigner functions for uniform and optimal sampling. (a) Measured Wigner function $W_{\mathrm{exp}}(\beta)$ of the cavity for $\alpha = 1$ and $\varphi = 0$ and uniform sampling. (b) Same preparation evaluated with optimal sampling. (c) and (d) Measured Wigner functions of the cavity with optimal sampling for $\alpha = 0$ and 2, and $\varphi = 0$.

When performing state tomography of an unknown state, it makes sense to sample uniformly $W_\alpha^\varphi(\beta)$. However, it is possible to obtain a theoretically unbiased estimation of the preparation fidelity to a known target state with maximal precision using another sampling [5] (in our case typically half the uncertainty compared to uniform sampling).

In Fig. 6.3b, we show the average measurement outcome $W_{\mathrm{exp}}(\beta)$ of $2\hat{\Pi}(\beta)/\pi$ for $N = 10^6$ samples distributed in β according to the optimal probability density $p_{\mathrm{opt}}(\beta) \propto |W_{|C_\alpha^\varphi\rangle\langle C_\alpha^\varphi|}(\beta)|$ for the target state $|\mathcal{C}_1^0\rangle$. Regions with low value for the target Wigner function appear very noisy owing to the small number of samples per pixel. For Fig. 6.3b, one finds $\mathcal{F}(1,0) = 94.7\ \% \pm 0.4\ \%$. Details about this estimation and the errors introduced by the decay during Wigner tomography and by the double use (preparation and measurement) of the qubit can be found in Sec. 6.12.2. The estimated fidelities for two other target states in Fig. 6.3c and d are $\mathcal{F}(0,0) = 94.3\ \% \pm 0.2\ \%$ and $\mathcal{F}(2,0) = 91.1\ \% \pm 0.3\ \%$. We deliberately focused on training the NN for large values of α, which explains the low fidelity when preparing the vacuum state.

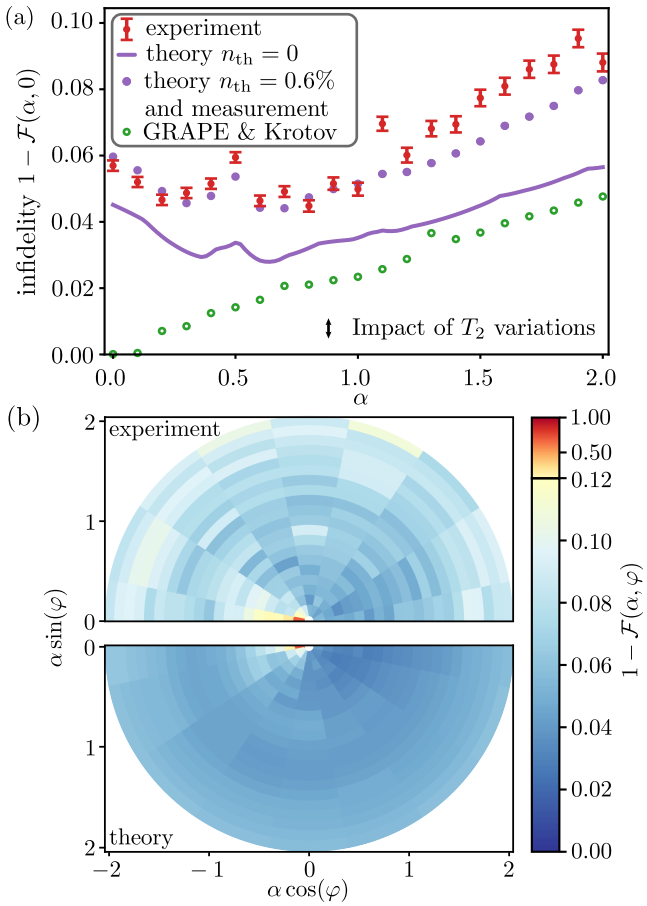


Figure 6.4: Experimental and theoretical infidelity as a function of α and φ. (a) Red dots: measured infidelity $1 - \mathcal{F}(\alpha, 0)$ for 21 values of α. Each point is obtained by averaging 10^6 samples. Purple line: simulated infidelity $1 - \mathcal{F}(\alpha, 0)$ for the same parameters. Blue dots: infidelities obtained by simulating also the measurement procedure and taking into account the finite thermal population n_{th} of the cavity. Green open circles: simulated infidelity using pulses optimized by GRAPE followed by the Krotov method [3, 4, 236]. Double arrow: typical amount of infidelity change due to possible T_2 variations between 20 μs and 60 μs over the 8 days of measurement (see Sec. 6.14.2). (b) Bottom: simulated infidelity for a sampling of the whole half-disk (reflected) in the parameter space $(\alpha \cos(\varphi), \alpha \sin(\varphi))$. Top: experimental infidelities obtained with 10^5 shots sampled optimally. The measurement of these 20×9 infidelities took about 8 days as well.

We use this measurement protocol to characterize the performance of the neural network in the preparation of many quantum states. In Fig. 6.4a (red dots), we show the estimated preparation infidelities $1-\mathcal{F}(\alpha,\varphi)$ as a function of α for $\varphi=0$ ($N=10^6$ samples per parameter). For comparison, we simulate the evolution of the density matrix $\hat{\rho}(\alpha,\varphi,t)$ under the same drives and compute predicted infidelities $1-\mathcal{F}(\alpha,\varphi)=1-\pi\int_{\mathbb{C}} W_{\hat{\rho}(\alpha,\varphi,T)}(\beta)W_{|C_\alpha^\varphi\rangle\langle C_\alpha^\varphi|}(\beta)\mathrm{d}\beta$ (purple line in Fig. 6.4a). Both the experiment and the prediction exhibit the same qualitative behavior. For the smallest amplitudes, the infidelity decreases before increasing steadily. At low amplitudes, the small sampling (proportional to α) of the network training dominates the infidelity of the preparation and it gets better as α rises. Around $\alpha\approx 0.5$, the infidelity grows with α because it is harder and harder to faithfully prepare a cat state with higher number of photons both theoretically and experimentally. As seen in Fig. 6.2b, the number of photons during preparation exceeds the final one. We trained the same neural network on a dissipation-free model of the same system (see appendix 6.8 and Fig. 6.8 for details and an experimental test), and obtained about 2 to 6 times smaller infidelities than with dissipation (purple solid line in Fig. 6.4a). This indicates that the dissipation is the main mechanism limiting the theoretical fidelity of the prepared states.

In order to illustrate the trade-off between preparation infidelity and duration of pulse generation, we computed the infidelity one should obtain using a standard direct optimization method. The green open circles in Fig. 6.4a represent the theoretical infidelities obtained using 2 μs long control pulses optimized by GRAPE (GRadient Ascent Pulse Engineering) followed by the Krotov method [3, 4, 236]. Note that in contrast to the 9 coefficients we use to parameterize the control pulses with the NN, GRAPE & Krotov uses 2000 coefficients (one per ns). Interestingly, the infidelity curve has the same slope as the one obtained with the NN (solid purple line) in the region $\alpha\gtrsim 0.7$. The GRAPE & Krotov performs better by about a constant infidelity offset of around 0.01 in this region. However, generating the control pulses with GRAPE & Krotov for the 21 values of α plotted in Fig. 6.4 a took about 6 hours on 14 cores i9 Intel CPU. This significantly exceeds the time of approximately 0.5 hours needed to train the NN on a

GPU. More importantly, once trained, the NN can generate control signals for any cat state from the parameter space about five orders of magnitude faster than using GRAPE & Krotov: generating the 21 control pulses takes 0.3 s on an i9 Intel CPU. Note that the generation of pulses by the NN could be considerably accelerated using dedicated hardware (GPU, FPGA, TPU...).

We also measure the preparation infidelities for several phases with $N = 10^5$ samples per parameter pair $(\alpha, \varphi) \in [0.1, 2] \times [0, \pi]$ (Fig. 6.4b upper panel). The predicted infidelities for the same parameter set are shown in the bottom panel of Fig. 6.4b. From this plot, it appears that the neural network behaves similarly for any phase φ except at low amplitudes α. Beyond scarce sampling at low α, this can be understood since the cat state varies rapidly with φ next to π. At low amplitude α, cat states $|C_\alpha^\varphi\rangle$ are close to $|0\rangle$ except in the vicinity of $\varphi = \pi$ where $|C_\alpha^\varphi\rangle$ is close to the Fock state $|1\rangle$. The large infidelity at $\varphi = \pi$ illustrates that a neural network performs poorly in the neighborhood of singularities since it inherently yields solutions which are continuous with respect to the input parameters.

Overall, we observe larger infidelities in the experiment than in the prediction. We attribute this mismatch to two main phenomena (see Sec. 6.14). The first one is an underestimation of the fidelity by the measurement protocol. Indeed the qubit is not always in the ground state at the end of the pulse sequence, which can artificially lower the measured fidelity, since the same qubit is then re-used for the Wigner tomography. Likewise, but with a lesser impact, the parity measurement procedure has a finite duration $t_{\text{par}} = \frac{\pi}{\chi}$, during which the cavity decay probability is non negligible. The second reason for a deviation between experimental and theoretical fidelities is an imperfect state initialization. After the qubit and cavity reset, there is a remaining $n_{\text{th}} = 0.6$ % thermal population in the cavity (see Sec. 6.11.3), which is directly responsible for the same amount of fidelity decrease $\Delta\mathcal{F}_{\text{th}} \simeq n_{\text{th}}$. This can be understood simply: when the cavity (or the qubit) is in the Fock state $|1\rangle$ before the pulse sequence, the system is in a state that has zero overlap with the expected initial state (ground state). Thus, at the end of the unitary evolution, it will also have zero overlap with the expected final state, thus exhibiting zero fidelity to the target state. On average, the fidelity

is then lowered by $\Delta\mathcal{F}_{\text{th}} \simeq n_{\text{th}}$. Simulating the whole measurement procedure and taking into account the remaining thermal population n_{th} of the cavity, we are able to capture the measured infidelities to a higher degree of accuracy (purple dots in Fig. 6.4a).

The measured fidelities seem to vary a bit more than expected from the statistical error bars in Fig. 6.4a, which is not the case of the theoretical predictions (purple solid line and purple dots). We explain this behavior by the variability of the qubit decoherence time T_2 during the experiment. While the value of T_2 for the theoretical predictions in Fig. 6.4a was fixed to 42 μs, we have observed variations from 20 μs to 60 μs during the experiment. Since each of the two panels in Fig. 6.4 required 8 days of measurement records, it is likely that T_2 varied substantially from one point to the next. As shown in Sec. 6.14.2, our measurement can be explained by variations of T_2 that are compatible with our observations.

6.4 CONCLUSION

In summary, the key findings of our work are the following. (i) Fast generation of optimized pulses for quantum state preparation using reinforcement learning: We introduce a new approach that can generate control sequences for quantum states 5 orders of magnitude faster than existing methods. (ii) Neural network-based control: Our method utilizes a neural network to directly generate optimized control pulses for a class of tasks, eliminating the need to run an optimization algorithm for each individual task. (iii) Experimental demonstration on cat states: We present an experiment on a superconducting circuit showing the fidelity of cat state preparation, a promising candidate for quantum error correcting codes. (iv) Experimental and theoretical comparison between GRAPE and neural network control: We implemented both our method and also GRAPE in order to provide a comparison between the fidelities and speed of both techniques. (v) Demonstration of optimal sampling for fidelity estimation: We perform an optimal sampling of the Wigner function of a microwave cavity for the task of estimating the preparation fidelity, and compare it to the traditional technique of uniform sampling.

There are multiple potential follow-up applications: one may promote the model parameters as trainable, and directly use the experimental measurements as input to learn the correct model parameters from the experiment. Alternatively, one may train for a whole range of model parameters in advance and adapt them to the experiment on the fly, e.g. when parameters are drifting. Since the network structure is light and easy to handle, it can form an efficient building block for complex network-based real-time feedback control of quantum systems. In that setting the moderate size of the network, efficient communication and information compression are crucial for low-latency data processing, whether the network is hosted on the FPGA controller itself or on a dedicated hardware.

In future work on similar challenges, one can imagine variations in the computational pipeline that connects the input task parameters to the final control pulse shape. First of all, one may replace B-splines by other suitable smooth parametrizations, such as Fourier decompositions. Second, the neural network itself could take as an additional input the time variable and be used to produce the entire pulse shape by querying it at all required values of the time variable on a grid, at the expense of longer inference times. Third, one can at least imagine to replace the neural network by some alternative architecture, such as kernel machines, or even use symbolic regression approaches for interpretability (representing the output as a formula whose number of terms is kept limited).

In the future, applying the scheme demonstrated here may be key for potential applications requiring fast change of pulse parameters (feedback, sensing, adaptive techniques, error correction etc.).

6.5 ACKNOWLEDGMENTS

This research was supported by the QuantERA grant ARTEMIS, by ANR under the grant ANR-22-QUA1-0004 and by German Federal Ministry of Education and Research under the grant 13N16360 within the program “from basic research to market”. We also thank the Munich Quantum Valley, which is supported by the Bavarian state government with funds from the Hightech Agenda Bayern Plus. We acknowledge IARPA and Lincoln Labs for providing a Josephson Traveling-Wave

Parametric Amplifier. We thank Antoine Essig, Omar Fawzi, Thomas Fösel, Clément Panais, and Arthur Strauss for useful discussions.

6.6 B-SPLINE BASIS SET

Instead of working with the control fields directly, we consider their expansion in a B-spline basis in the form described in Ref. [238] in the context of computational atomic physics applications. A crucial advantage of such an approach is a possibility to pre-load the B-splines on the hardware and send only the expansion coefficients of control signals for different cat states. A particular B-spline basis is built on a fixed knot sequence $t_1 \leq t_2 \leq \ldots \leq t_{n+k+1}$ and is characterized by the basis size n and the degree k of the B-splines. Each k-degree B-spline $B_{i,k}(t)$ with $i = 1, \ldots, n$ is a piecewise polynomial of degree k inside the interval $[t_i, t_{i+k+1})$ and vanishes outside this interval. They are constructed recursively as follows:

$$B_{i,0}(t) = \begin{cases} 1, & \text{if } t \in [t_i, t_{i+1}) \ , \\ 0, & \text{otherwise} \ , \end{cases} \tag{6.7}$$

$$B_{i,k}(t) = \frac{t - t_i}{t_{i+k} - t_i} B_{i,k-1}(t) + \frac{t_{i+k+1} - t}{t_{i+k+1} - t_{i+1}} B_{i+1,k-1}(t) \ . \tag{6.8}$$

We follow Ref. [238] and choose for the $k+1$ leftmost knots $t_1 = \ldots = t_{k+1} = 0$ and for the $k+1$ rightmost knots $t_{n+1} = \ldots = t_{n+k+1} = T$. The rest $n - k - 1$ knots are distributed uniformly between 0 and T.

In this work we stick to a B-spline basis set of size $n = 11$ and degree $k = 3$ shown in Fig. 6.5 for the excitation time interval $T = 2$ μs. We exclude the first and last B-splines which are the only ones having non-zero values at the interval edges. In this way, we restrict our driving signals to start and end with zero amplitude. The problem of the control field optimization reduces to searching for a total of 9×4 optimal expansion coefficients for the 2 complex (qubit + cavity), and thus 4 real driving fields. Apart from a compact representation of the control signals, B-splines offer also a practical advantage of being non-zero only in restricted intervals. In particular, for the chosen basis set only up to 4 B-splines are non-zero at each point of the driving time

interval. This ensures compatibility with existing real-time quantum control hardware. For instance, the OPX from Quantum Machines can generate up to 18 signals at the same time, which would allow to use 4 B-splines for 4 control fields at the same time.

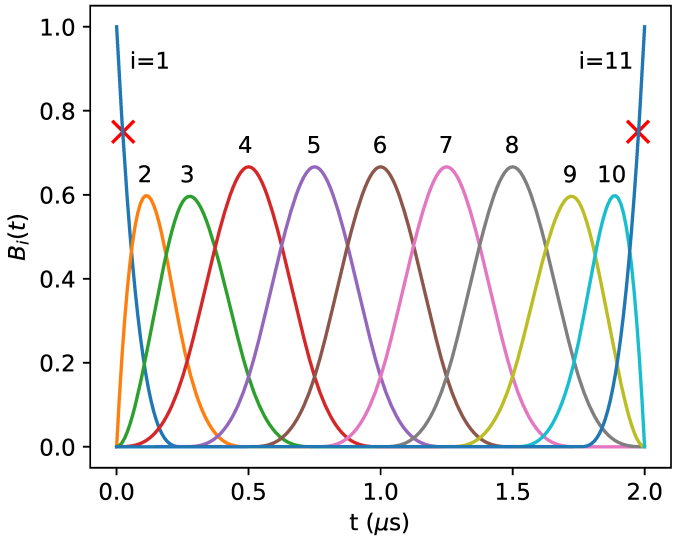


Figure 6.5: B-spline basis of size $n = 11$ and degree $k = 3$. In this work the first and last B-splines are excluded in order to restrict the driving signals to those starting and ending at zero.

6.7 NEURAL NETWORK TRAINING

Here we show details of the neural network training performed in the way described in Section 6.2. Whereas we experimentally test preparation of cat states with parameters $\alpha \in (0, 2)$ and $\varphi \in (0, \pi)$, the neural network training is performed on a wider and more challenging parameter space $\alpha \in (-4, 4)$ and $\varphi \in (0, \pi)$. Due to the equality $|C_\alpha^{2\pi-\varphi}\rangle = e^{i\varphi}|C_{-\alpha}^{\varphi}\rangle$ obtained directly from the definition of cat states (6.2), the neural network is equivalently able to predict control signals for $\alpha \in (0, 4)$ and $\varphi \in (0, 2\pi)$.

In Fig. 6.6 we demonstrate the training procedure performed in a few stages by gradually increasing the "complexity" of the parameter space and the numerical model of the system. Concretely:

- we increase the α-interval $(0, \alpha_{\max}^{(\text{train})})$ from $\alpha_{\max}^{(\text{train})} = 2$ to $\alpha_{\max}^{(\text{train})} = 4$;
- the number of Fock states included in the cavity numerical modeling is increased from $N_{\text{cav}}^{\max} = 20$ to $N_{\text{cav}}^{\max} = 60$;
- since the volatility of the training loss is larger for a larger parameter space, we increase also the batch size at later training stages.

At the starting stage (first 1000 batches) shown by the red-shaded area, the training often stagnates and the batch-averaged loss $\langle \mathcal{L}(\alpha, \varphi) \rangle_{\text{batch}}$ does not decrease, see the upper panel of Fig. 6.6. This is a local minimum in which the neural network learns not to drive the qubit at all, whereas the cavity remains in a coherent state. We overcome this trap by adding an incoherent excitation process from ground to excited state with a rate $\Gamma_{\uparrow} = (5\ \mu\text{s})^{-1}$ at this initial stage. We stress that this pumping is not present in the experiment and is used here only to tackle the mentioned training problem. Numerically this is achieved within the perturbative framework for inclusion of relaxation channels described in Sec. 6.9. While such a trapping occurred at the beginning, training proceeded smoothly afterwards.

6.8 NEURAL NETWORK PERFORMANCE

In order to check the neural network performance, we apply it first to a selection of size 3200 from the parameter space $\alpha \in (0, 4)$ and $\varphi \in (0, 2\pi)$. In Fig. 6.7 we show the simulated fidelity of the target states obtained by driving the system as suggested by the neural network. The parameters (α, φ) are encoded in polar coordinates. In Fig. 6.8 (purple curve) we show additionally the results for $\varphi = 0$. We observe indeed that the obtained fidelity is worse for larger α-s justifying sampling with linearly growing density.

We demonstrate now how the final fidelity changes by increasing the number of the B-splines in the basis set. As an example, we consider a basis with parameters $n = 25$, $k = 3$ (see Appendix 6.6) without the left and right B-splines consisting of 23 functions. Since the NN outputs now 92 values instead of 36, we increased the size of the last

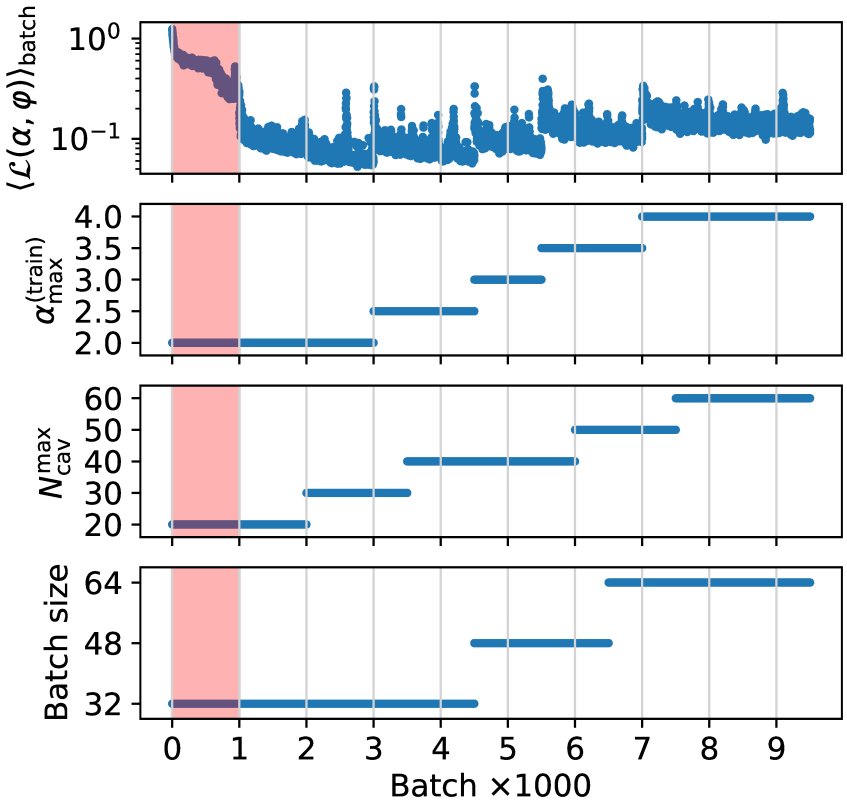


Figure 6.6: Neural network training progress with batches. In the upper panel we show the batch-averaged loss $\langle \mathcal{L}(\alpha, \varphi) \rangle_{\text{batch}}$, whereas the other panels show the gradual change of the parameters $\alpha_{\text{max}}^{(\text{train})}$, $N_{\text{cav}}^{\text{max}}$, and the batch size. See text for details.

hidden layer from 30 to 90. We train this NN and demonstrate its performance by modeling preparation of cat states with $\alpha \in (0, 4)$ and $\varphi = 0$. The results are shown in Fig. 6.8 (orange curve). For states with $\alpha > 2$ (which were not experimentally tested in this work) the increase of the B-spline basis leads to a better NN performance. For smaller α, however, the quality of the constructed states drops. Getting a larger fidelity for all α requires to optimize the sampling of α during training.

We switch now back to the original B-spline basis of size 9 and the original NN architecture, and study the influence of decoherence on

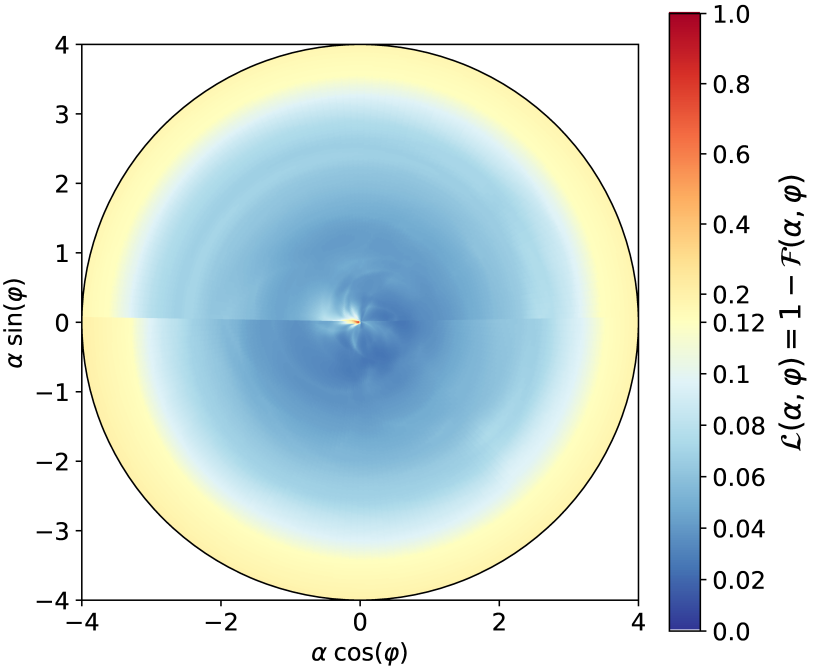


Figure 6.7: Simulated infidelity of cat states obtained by driving the system in the way suggested by the trained neural network. The parameter space (α, φ) is shown in polar coordinates. The color indicates the infidelity $\mathcal{L}(\alpha, \varphi) = 1 - \mathcal{F}(\alpha, \varphi)$.

the NN performance. We model an idealistic situation without decoherence and repeat the NN training following the scenario described in Appendix 6.7. The NN is then applied to prepare cat states with $\alpha \in (0, 4)$ and $\varphi = 0$ also in absence of the decoherence channels. The results shown in Fig. 6.8 (brown curve) indicate approximately 2 to 6 times smaller infidelity than in the original realistic situation suggesting that dissipation is the main source of infidelity.

6.9 ACCOUNTING FOR DECOHERENCE

In general, the decoherence effects cannot be addressed using the Schrödinger equation, and need switching to an open quantum system approach such as the Lindblad equation for the density matrix. In this

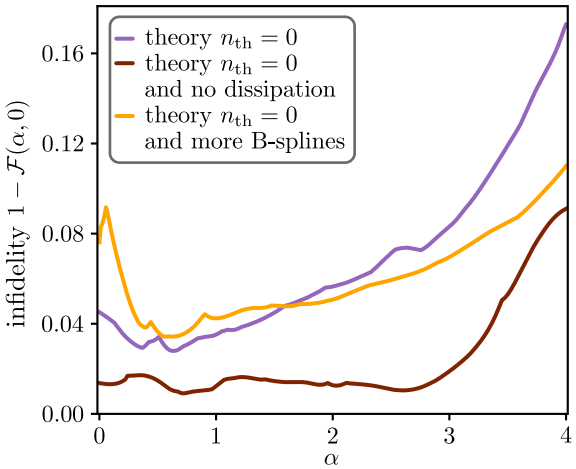


Figure 6.8: Simulated infidelity of cat states with $\alpha \in (0, 4)$ and $\varphi = 0$ obtained using the original scenario (purple curve), a larger B-spline basis set (orange curve) and in absence of decoherence (brown curve).

work, however, we are not interested in the whole system dynamics upon inclusion of decoherence, but only in the correction of the final state fidelity. Therefore, instead of solving the more complicated Lindblad equation, we still solve the Schrödinger equation and correct the fidelity for the decoherence using first-order perturbation theory.

We derive here the aforementioned correction. The Lindblad equation reads

$$\frac{\partial \hat{\rho}}{\partial t} = -i[\hat{H}, \hat{\rho}] + \sum_i \frac{1}{\tau_i} D_i[\hat{\rho}] \,, \tag{6.9}$$

where the summation is performed over all decoherence channels with respective characteristic times τ_i. For each decoherence channel

$$D_i[\hat{\rho}] = \hat{a}_i \hat{\rho} \hat{a}_i^\dagger - \frac{1}{2} \hat{a}_i^\dagger \hat{a}_i \hat{\rho} - \frac{1}{2} \hat{\rho} \hat{a}_i^\dagger \hat{a}_i \,, \tag{6.10}$$

where $\hat{a}_i$ is the corresponding jump operator. In our case, for instance, the following decoherence channels (governed by the jump operators shown in the parenthesis) are present:

- cavity dissipation (photon annihilation operator $\hat{a}$);

- qubit dissipation ($\hat{\sigma}_-$);
- qubit dephasing ($\hat{\sigma}_z/\sqrt{2}$).

The density matrix at time t in absence of decoherence is the pure state density matrix

$$\hat{\rho}^{(0)}(t) = |\psi(t)\rangle \langle\psi(t)| \ . \tag{6.11}$$

The first-order correction due to the i-th decoherence channel is

$$\hat{\rho}_i^{(1)}(t) = \frac{1}{\tau_i} \int_0^t dt' \ \hat{U}^{(0)}(t',t) \ D_i[\hat{\rho}^{(0)}(t)] \ \hat{U}^{(0)\dagger}(t',t) \tag{6.12}$$

leading to the change in the final state fidelity at $t = T$ with respect to the target state $|\mathcal{T}\rangle$

$$\Delta F_i = \Big\langle \mathcal{T}| \ \hat{\rho}_i^{(1)}(t=T) \ |\mathcal{T}\Big| \mathcal{T}| \ \hat{\rho}_i^{(1)}(t=T) \ |\mathcal{T}\Big\rangle \ , \tag{6.13}$$

We make at this point an additional assumption that the obtained state without decoherence $\psi(t = T)$ is very close to the target state $|\mathcal{T}\rangle$ and write

$$\Delta F_i \approx \Big\langle \psi(t=T)| \ \hat{\rho}_i^{(1)}(t=T) \ |\psi(t=T)\Big|\psi(t=T)| \ \hat{\rho}_i^{(1)}(t=T) \ |\psi(t=T)\Big\rangle \ . \tag{6.14}$$

In this approximation, the evolution operator $\hat{U}(t', T)$ connects the intermediate system state at time t' and the target system state. From here we obtain finally:

$$\begin{aligned} \Delta F_i \approx \frac{1}{\tau_i} \int_0^T dt' \big[\ & \big|\big\langle \psi(t')|\hat{a}_i|\psi(t')\big|\psi(t')|\hat{a}_i|\psi(t')\big\rangle\big|^2 \\ & - \Big\langle \psi(t')|\hat{a}_i^\dagger\hat{a}_i|\psi(t')\Big|\psi(t')|\hat{a}_i^\dagger\hat{a}_i|\psi(t')\Big\rangle \big]. \end{aligned} \tag{6.15}$$

6.10 DEVICE AND MEASUREMENT SETUP

6.10.1 *Device fabrication*

The system is composed of one 3D $\lambda/4$ coaxial cavity resonator in 99.99 % pure Aluminum, into which a chip, containing the transmon

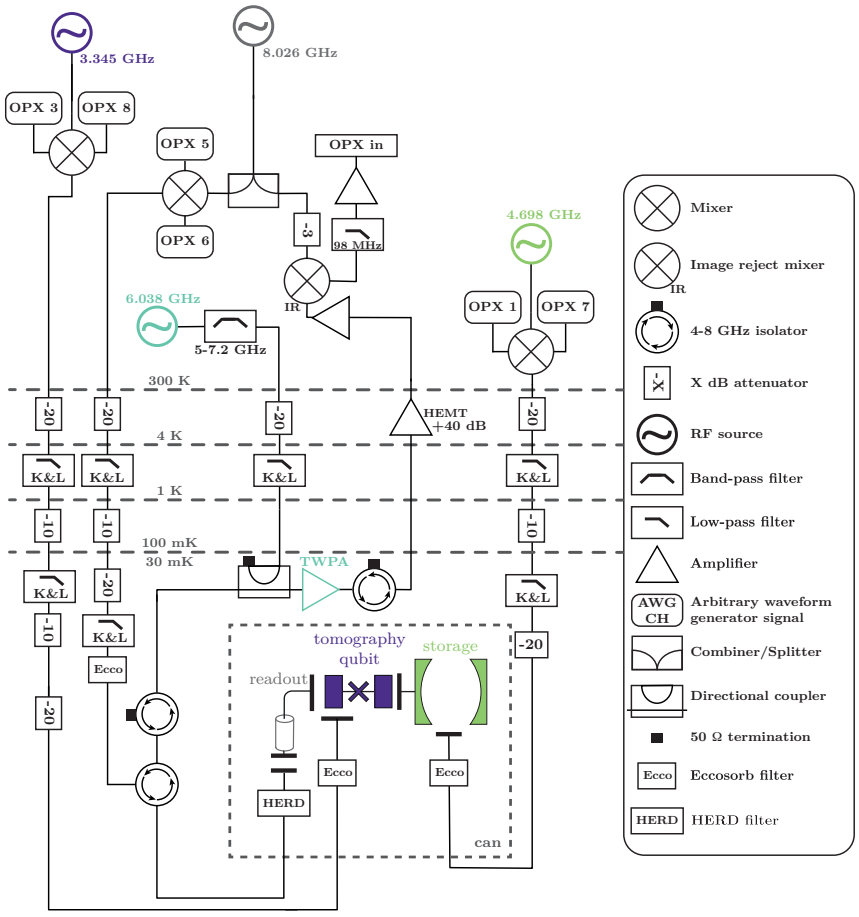


Figure 6.9: Schematic of the measurement setup. A color matched RF source is dedicated to each element. Room-temperature isolators are not represented for the sake of clarity.

qubit with its readout resonator and Purcell filter, is inserted (see Fig. 6.10). This chip is made of an etched 200 nm thick film of sputtered Tantalum on a 430 μm thick sapphire substrate (deposited by Star Cryoelectronics, Santa Fe, USA). The Josephson junctions of both transmons are standard Dolan bridge e-beam evaporated Al/AlOx/Al junctions [103].

Table 6.1: Table of circuit parameters

Table of circuit parameters			
Circuit parameter	Symbol	Hamiltonian term	Value
Cavity frequency	$\omega_c/2\pi$	$\hbar\omega_c\hat{a}^\dagger\hat{a}$	4.628 GHz
Qubit frequency	$\omega_q/2\pi$	$\hbar\omega_q\|1\rangle\langle 1\|_q$	3.235 GHz
Readout frequency	$\omega_r/2\pi$	$\hbar\omega_r\hat{r}^\dagger\hat{r}$	7.960 GHz
Cavity-qubit cross Kerr rate	$\chi/2\pi$	$-\hbar\chi\hat{a}^\dagger\hat{a}\|1\rangle\langle 1\|_q$	238.5 kHz
Circuit parameter	Symbol	Dissipation operator	Value
Qubit decay time	T_q	$1/T_q\mathbb{D}_{\hat{\sigma}_-}$	35 μs
Cavity decay time	T_c	$1/T_c\mathbb{D}_{\hat{a}}$	225 μs
Readout decay time	$T_{1,r}$	$1/T_{1,r}\mathbb{D}_{\hat{r}}$	80 ns
Qubit dephasing time	$T_{q,\varphi}$	$1/2T_{q,\varphi}\mathbb{D}_{\hat{\sigma}_z}$	175 μs

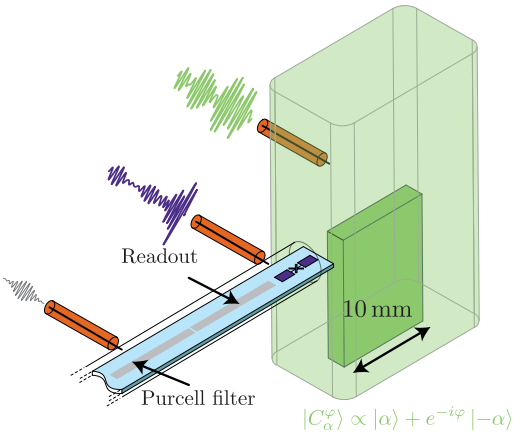


Figure 6.10: Scheme of the device.

6.10.2 *Measurement setup*

The readout resonator, the cavity and the qubit are driven on resonance by pulses that are generated using an OPX from Quantum Machines®.

It has a sampling rate of 1 GS/s. The generated pulses are modulated at a frequency 68 MHz for readout, $\omega_{\mathrm{q}}^{\mathrm{IF}}/2\pi = 110$ MHz for the qubit, $\omega_{\mathrm{c}}^{\mathrm{IF}}/2\pi = 70$ MHz for cavity. They are up-converted using I-Q mixers with continuous microwave tones produced by three channels of an AnaPico® APUASYN20-4 for the readout resonator and cavity, Agilent® E8257D for the qubit.

The reflected signals from the readout is amplified with a TWPA provided by Lincoln Labs [90] powered by a pump tone at 6.038 GHz. The follow-up amplification is performed by a HEMT amplifier from Low Noise Factory® at 4 K and by room-temperature amplifiers. The signal is down-converted using image reject mixers before digitization by the input ports of the OPX.

The samples are placed inside a can comprising a layer of lead, gold-plated copper and a layer of cryoperm. The inside of the can is coated with an absorptive mixture of 86 % Stycast 2850 FT, 7 % catalyst 23 LV and 7 % carbon powder.

For the cavity, the controls pipeline is as follows. We define $I_{\mathrm{c}}(t)$ and $Q_{\mathrm{c}}(t)$ the time dependent voltages to send on resonance to the cavity. These controls are then converted into $I_{\mathrm{c}}^{\mathrm{IF}}(t) = \cos\big(\omega_{\mathrm{c}}^{\mathrm{IF}}t\big)I_{\mathrm{c}}(t) - \sin\big(\omega_{\mathrm{c}}^{\mathrm{IF}}t\big)Q_{\mathrm{c}}(t)$ and $Q_{\mathrm{c}}^{\mathrm{IF}}(t) = \sin\big(\omega_{\mathrm{c}}^{\mathrm{IF}}t\big)I_{\mathrm{c}}(t) + \cos\big(\omega_{\mathrm{c}}^{\mathrm{IF}}t\big)Q_{\mathrm{c}}(t)$ and sent on DAC OPX 5 and 6. These signals are then mixed with a local oscillator at frequency $\omega_{\mathrm{c}}^{\mathrm{LO}}$ with an IQ mixer in a lower sideband setting to give a RF voltage $V^{\mathrm{RF}}(t) \propto \cos\big((\omega_{\mathrm{c}}^{\mathrm{LO}} - \omega_{\mathrm{c}}^{\mathrm{IF}})t\big)I_{\mathrm{c}}(t) + \sin\big((\omega_{\mathrm{c}}^{\mathrm{LO}} - \omega_{\mathrm{c}}^{\mathrm{IF}})t\big)Q_{\mathrm{c}}(t)$ at the output of the IQ mixer. The pulsations $\omega_{\mathrm{c}}^{\mathrm{LO}}$ and $\omega_{\mathrm{c}}^{\mathrm{IF}}$ are chosen such that $\omega_{\mathrm{c}}^{\mathrm{LO}} - \omega_{\mathrm{c}}^{\mathrm{IF}} = \omega_{\mathrm{c}}$. Through attenuation in the lines and the coupling of the qubit to these lines, this translates to the control amplitudes $\varepsilon(t) = \xi_{\mathrm{c}}(I_{\mathrm{c}}(t) + iQ_{\mathrm{c}}(t))$ on the cavity. Details about the calibration of ξ_{c} can be found in 6.11.4. The qubit is driven similarly with control voltages $I_{\mathrm{q}}(t)$ and $Q_{\mathrm{q}}(t)$ from OPX channels 3 and 8.

6.11 CALIBRATION AND HERALDING

This section aims at giving technical details about the system initialization and measurement. The initialization part was especially complex, as we needed to make sure that both cavity and qubit are in the ground state. This is done using standard measurement-based

feedback techniques. On top of this standard reset protocol, the presence of a Two-Level System (TLS) made the frequency of the qubit jitter between two frequencies at long timescales, which obviously deteriorates the fidelity of the state preparation if not mitigated. Here, we study this TLS dynamics and desmonstrate a heralding protocol for the qubit Fig. 6.13, demonstrating close to total mitigation of the TLS detrimental impact.

6.11.1 *Qubit readout and reset*

We readout the qubit with a square pulse of length 2.2 μs and measure the quadrature I that encodes the qubit information. Single-shot fidelity is around 98 %. The reset of the qubit consists in a feedback loop using two thresholds represented in Fig. 6.11a. Exceeding the first one (red dashed line) is the ending condition, allowing to herald the qubit in its ground state $|0\rangle$ with more than 99.9 % fidelity. The second threshold (gray dashed line) triggers a π pulse (gate X) if the qubit is more likely to be in the excited state, (*i.e* I below threshold). Getting below a third threshold (yellow dashed line in Fig. 6.11a) allows to herald the qubit in the excited state with the same fidelity, and is used for the reset of the cavity (see section 6.11.3).

6.11.2 *TLS mitigation*

Probing the frequency of the qubit is done with a standard Ramsey pulse sequence. The signal is obtained by subtracting the results of two Ramsey sequences performed with opposite parity for the second $\pi/2$ pulse (see Fig.6.12a). We see that the qubit has two possible frequencies (see Fig. 6.12b). We attribute this to the presence of a spurious Two Level System (TLS) dispersively coupled to the qubit, with a dispersive shift $\delta/2\pi = 32$ kHz. This TLS has a thermal occupation of about 40 %, and its state switches over the course of the full experiment, which makes the frequency of the qubit jitter between $\omega_q/2\pi$ and $\omega_q/2\pi - \delta/2\pi$. The experimental temporal Ramsey signal fits well with a model including two frequencies with the same decay time T_2, which drifted between 20 and 60 μs over a few hours (see Fig. 6.12b).

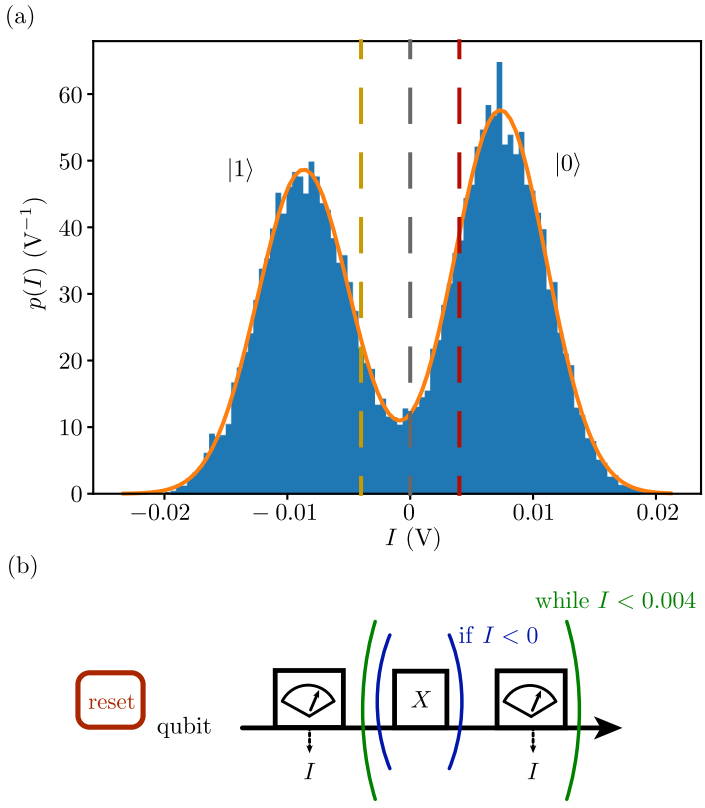


Figure 6.11: Readout and reset of the qubit. (a) Blue: Histogram of 20 000 single shot readouts of the qubit in its thermal state combined with 20 000 more after an approximate π pulse. Dashed lines represent thresholds for qubit and cavity reset operations (see text). (b) Reset scheme. The qubit is read until the measurement record exceeds the threshold $I = 0.004$ V. If the qubit is more likely to be in the excited state, a π pulse is applied before the readout.

The frequency jumps introduced by this TLS are detrimental to the fidelity of the prepared state, as the state preparation needs a precise calibration of the frequency of the qubit. We mitigated this effect using a heralding procedure. Since the TLS state cannot be inferred from the qubit state in one shot, we chose to take a decision using a repetition of 30 times the same Ramsey sequence on resonance with the frequency

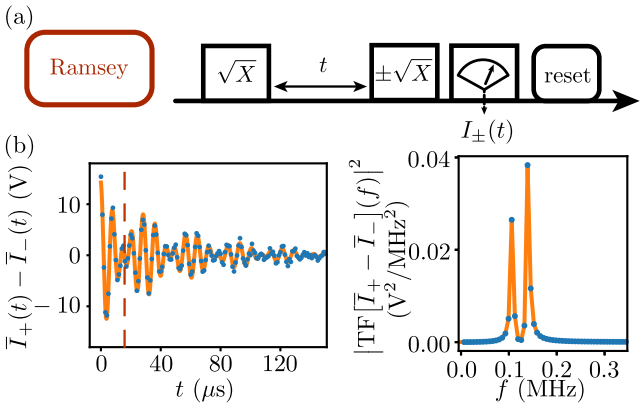


Figure 6.12: Ramsey measurement of the qubit. (a) Ramsey pulse sequence. The $\pi/2$ ($\sqrt{X}$) pulses are detuned from the qubit frequency by around 140 kHz, and are separated by a waiting time t. The rotation axis of the second pulse is reversed at each repetition. (b) Blue dots: Signal averaged over a 2000 repetitions as a function of t. Orange solid line: fit with the sum of two oscillations detuned by $\delta/2\pi = 32$ kHz and decaying with a time $T_2 = 51$ μs. (c) Blue dots: Spectral density obtained by a Fourier transform of (b). Solid orange line: Spectral density of the fit function in (b).

that the qubit takes when the TLS is in the ground state (identified by the highest peak in Fig. 6.12), with a delay $\tau_{\mathrm{TLS}} = \frac{\pi}{\delta} = 15.5$ μs. The pulse sequence is schematized in Fig. 6.13a. To each of the 30 measurement outcomes $\{I_i\}_{1\leq i\leq 30}$, we associate

$$r_i = \begin{cases} 1 \text{ if } \mathrm{I_i} < 0 \text{ V} \\ 0 \quad \text{else.} \end{cases} \tag{6.16}$$

From this, we compute $R = \sum_{i=1}^{30} r_i$. The heralding condition is then fixed to $R \geq 20$. The Figure 6.13a shows the pulse sequence used to perform a standard Ramsey measurement after this heralding sequence. Figures 6.13b and c show the resulting decaying Ramsey signal and its Fourier transform. The beating in Fig. 6.12b disappears when using the heralding, as well as the second peak in the spectrum of Fig. 6.12c, whose expected position is materialized by the vertical red dashed line in 6.13c. This demonstrates a successful mitigation of the TLS effects.

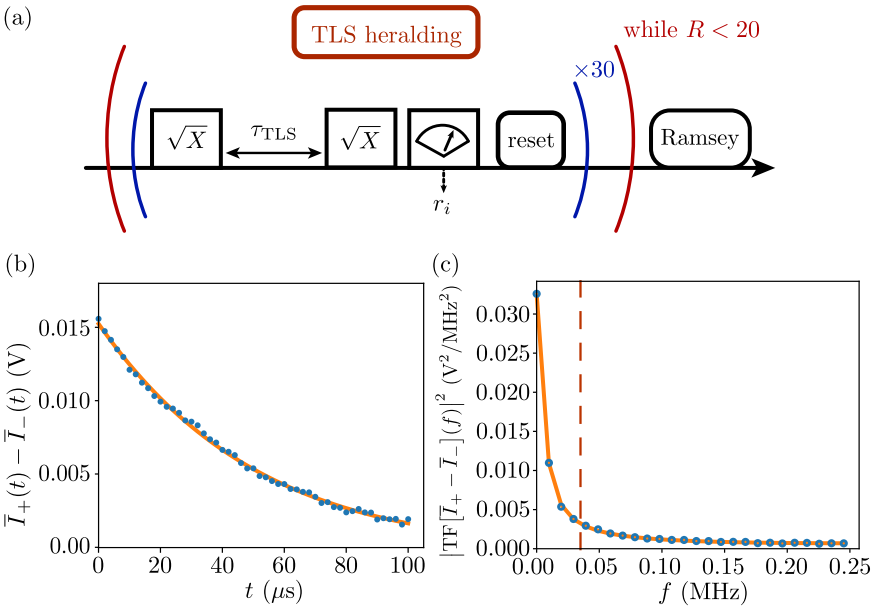


Figure 6.13: Heralding the TLS. (a) Scheme of the TLS heralding procedure. (b) On-resonance Ramsey measurement performed using the TLS heralding procedure. Blue dot: experimental signal. Orange solid line: exponential fit with $T_2 = 50\ \mu s$. (c) Blue dot and orange solid line: spectral density of (b) transform. Vertical dashed line: position of the expected second peak if it were not suppressed by the heralding procedure.

Further analysis of the TLS dynamics, and details about the determination of the heralding procedure is shown in the following.

6.11.2.1 *Dynamics of the TLS and high fidelity detection of its state*

We first studied the dynamics of the frequency jumps of the qubit, by performing 10^5 times the same Ramsey measurement on resonance with the upper frequency in 6.12b with a delay τ_{TLS}. The procedure is the same as for the parity measurement on a dispersively coupled harmonic oscillator [145]. The pulse sequence is shown in Fig. 6.14a. Without qubit decoherence, this would allow us to read out the state of the TLS in a single shot.

With the associated measurement signals $\{I_i\}$, we could first investigate the switching time of this TLS. To do this, we computed the correlations between the measurement records I_i and I_{i+k}. We define $c_k = \mathrm{Cov}(\{I_i\}, \{I_{i+k}\})$, and plot this quantity as a function of k (Fig. 6.14b). We observe a first decrease at short times $k < 100$, corresponding to a correlation time shorter than 1 ms. This we attributed to the cavity thermal population. The much longer drift that follows is attributed to the frequency jumps. It is not clear to us why this does not really follow an exponential decay. We believe that the correlations at long times of the TLS state cannot be properly captured by our measurement because it is too short (10^5 shots). Fitting the beginning of the drifts gives a characteristic time of around 400 ms, which gives plenty of time to perform post-selection. In order to have a good fidelity, we need to accumulate statistics over several measurements, as τ_{TLS} cannot be neglected compared to the decoherence time T_2 of the qubit. This measurement cannot be single shot.

6.11.2.2 *Heralding on the TLS state*

In order to understand how we can herald the state of the TLS, we perform a running average on the $\{I_i\}$ over $n = 30$ values of I_i defining

$$\overline{I}_i^{30} = \frac{1}{30} \sum_{k=-15}^{14} I_{i+k}. \tag{6.17}$$

The histograms of the $\{\overline{I}_i^{30}\}$ are shown Fig. 6.14c in blue. If we artificially cancel the effect of any correlation in the TLS dynamics on the measurement record I_i by shuffling the values I_i, we obtain the orange histogram in 6.14c. Since both histograms differ, it implies that the measurement records contain extractable information about the TLS.

We plot the histograms of the measurement record I_i post-selected on the condition $\overline{I}_i^{30} > -3.3$ mV (Fig. 6.15a blue histogram) and $\overline{I}_i^{30} < -3.3$ mV (Fig. 6.15a orange histogram). The value of the threshold is indicated in dashed line Fig. 6.14c.

We can now define the measurement records

$$r_i = \begin{cases} 1 \text{ if } \mathrm{I_i} < 0 \text{ V} \\ 0 \quad \text{else} \end{cases} \tag{6.18}$$

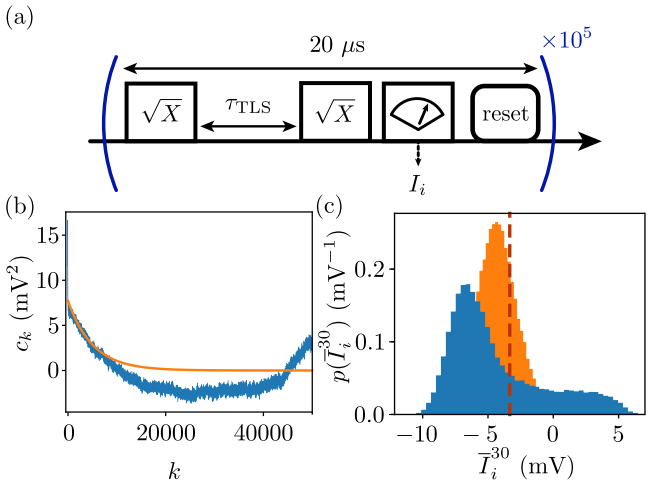


Figure 6.14: TLS state measurement. (a) Pulse sequence of the TLS state measurement. (b) Blue: Measured covariance c_k as a function of k. Orange: Exponential fit for $k < 10000$. The characteristic fitted time is 400 ms. (c) Blue: Probability density of the running average of the $\{I_i\}$. Orange: Same quantity, but with over $\{I_{\sigma(i)}\}$, where σ is a random permutation. Red dashed line: Threshold used for Fig. 6.15.

and infer a probability $p(r_i = 1|\text{TLS in ground state}) = p_0 = 0.80$ and $p(r_i = 1|\text{TLS excited}) = p_1 = 0.45$. The inferred corresponding conditional probability laws of R $p(R = n|\text{TLS excited})$ and $p(R = n|\text{TLS in ground state})$ are binomial laws with parameters 30 and $p_{0/1}$. Histograms of these laws are shown in orange (for p_0) and blue (for p_1) in Fig. 6.15b. We deduce from this histogram that $p(R \geq 20|\text{TLS in ground state}) = 0.0013$, while $p(R \geq 20|\text{TLS excited}) = 0.975$, justifying the heralding threshold presented in Fig. 6.13a.

Note that the heralding fidelity is actually a bit lower than this, as I_i enters in the computation of $\overline{I}_i$, which introduces a bias. This bias is sufficiently small so that the heralding procedure still works, as shown in Fig. 6.13.

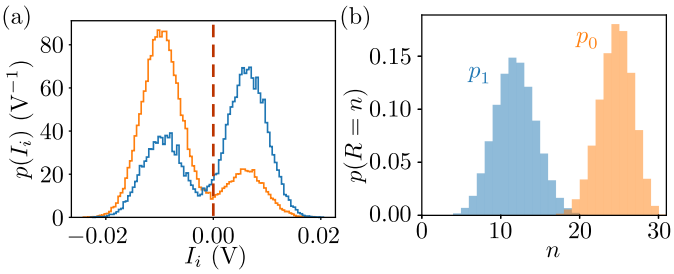


Figure 6.15: Distinguishability of the TLS states. (a) Probability densities of I_i post-selected on the value of $\overline{I}_i^{30} > -3.3$ mV (blue) and $\overline{I}_i^{30} \leq -3.3$ mV.(b) Estimated probability density for the value of R conditionned on the TLS state.

6.11.3 *Heralding the cavity in the vacuum state*

The cavity is reset to the vacuum state before each State Preparation and Measurement (SPM) sequence. The complete sequence thus contains an upgraded version of the TLS heralding sequence that includes a Gaussian 20 μs long selective π pulse on the qubit at $\omega_q/2\pi - 0\chi/2\pi$ at the end (see Fig. 6.16a). The corresponding linewidth of this truncated Gaussian pulse is 50 kHz, well below $\chi/2\pi$, which ensures the selectivity of this gate. The readout of the qubit gives $r_{\text{cav}} = 1$ if the recorded quadrature goes below the threshold (dashed yellow in Fig. 6.11). The full heralding sequence then consists in a while loop that exits when both the cavity and the TLS are found in their ground state.

In order to test the performance of this procedure in resetting the cavity population, a Ramsey measurement is performed after each heralding sequence. The result is shown in Fig. 6.16b. When the cavity is occupied by a thermal state with n_{th} average photons, the Ramsey signal is supposed to be proportional to $e^{-t/T_2}(1 - n_{\text{th}} + n_{\text{th}} \cos \chi t)$. In blue, the measured signal and fit show a cavity population of $n_{\text{th}} = 0.6$ % and fitted T_2 is here 40 μs after the heralding procedure, to be compared to the orange curve, corresponding to the same fit but with the cavity rigorously in the vacuum. This average photon number n_{th} is attributed to the imperfect QNDness of the qubit readout.

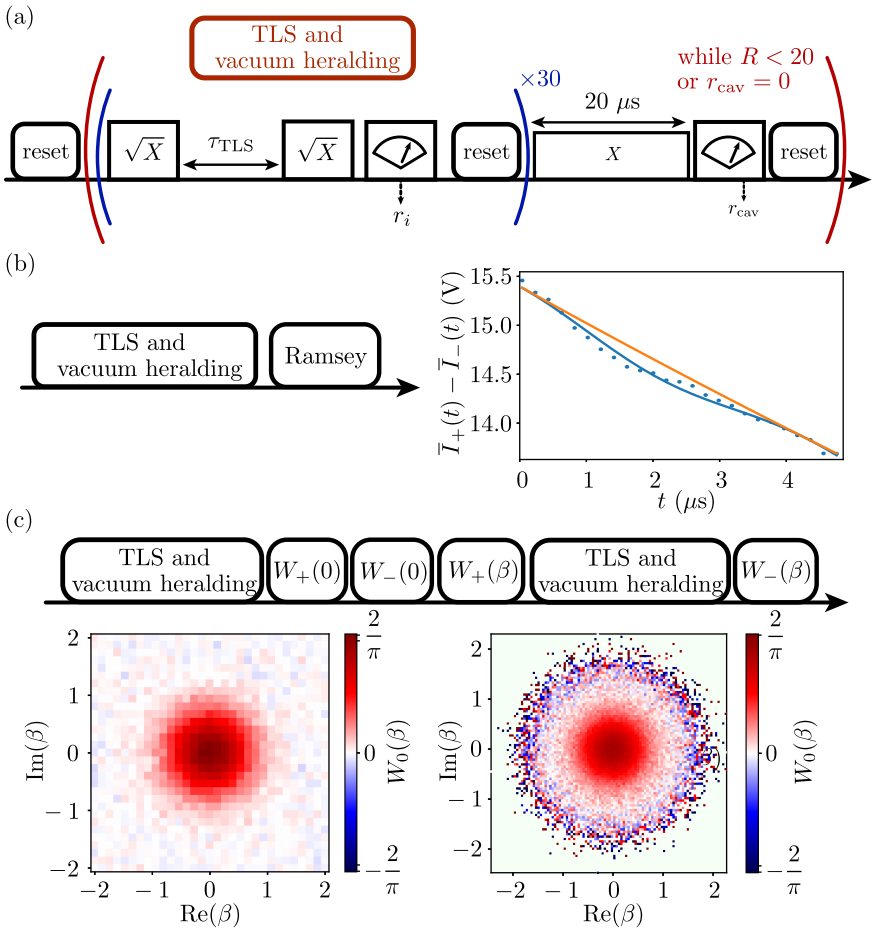


Figure 6.16: Cavity vacuum heralding procedure. (a) Pulse sequence for the TLS and vacuum heralding procedure. (b) top: Pulse sequence and Ramsey measurement on the qubit. Blue points: measured quadrature. Solid blue line: fit with a thermal population of 0.6 % using Eq. (6.19). Orange: expected signal for a pure vacuum. (c) top: high-level pulse sequence for the measurement of the Wigner of the vacuum. Bottom left: Wigner measurement with square uniform sampling. Bottom right: Wigner measurement with optimal sampling.

We also measure the Wigner function W_0 of the cavity after this reset Fig. 6.16c with a SPM sequence showed Fig 6.19(a), without playing the NN pulses. The left one is a Wigner tomography with 31×31 pixels

averaged 1000 times each, and gives a fidelity $\mathcal{F}_0^{\text{square}} = 99.4\,\% \pm 0.5\,\%$ to the vacuum. The right one is a measurement of $2\hat{\Pi}(\beta)/\pi$ using the optimal sampling p_{opt} detailed in the text. The estimated fidelity is $\mathcal{F}_0^{\text{opt}} = 98.9\,\% \pm 0.15\,\%$. We attribute the apparent disagreement with the experiment to a remaining population of the qubit after the reset possibly due to a slightly destructive behavior of the readout.

6.11.4 *Pulse calibration*

The control pulses consist in two pairs of voltages signals $(I_{\text{c}}(t), Q_{\text{c}}(t))$ and $(I_{\text{q}}(t), Q_{\text{q}}(t))$ each upconverted by the OPX controller and a mixer so that the drive amplitudes in the Hamiltonian read $\varepsilon_{\text{c}}(t) = \xi_{\text{c}}(I_{\text{c}}(t) + iQ_{\text{c}}(t))$ and $\varepsilon_{\text{q}}(t) = \xi_{\text{q}}(I_{\text{q}}(t) + iQ_{\text{q}}(t))$. Determining ξ_{q} and ξ_{c} is required for a proper implementation of the pulse sequences generated by the NN.

Calibrating ξ_{q} is done using a standard measurement of Rabi oscillations. Calibrating ξ_{c} is done using a so-called populated Ramsey sequence, which is a Ramsey sequence at resonance with the qubit, which is performed after a displacement of amplitude β on the cavity (see 6.17a). The measured Ramsey signal reads [239]

$$\begin{aligned} \overline{I}_+(t) - \overline{I}_-(t) \propto \exp\left[-\frac{t}{T_2}\right] \\ \times \exp\left[|\beta|^2(\cos(\chi t) - 1)\right] \cos\left(|\beta|^2 \sin(\chi t)\right). \end{aligned} \tag{6.19}$$

The curve Fig. 6.17b fitted with this expression indicates a displacement amplitude $\beta = 2.1$ after a 200 ns long displacement pulse of amplitudes $I_{\text{c}}(t) = 0.03$ V and $Q_{\text{c}}(t) = 0$ which gives $\xi_{\text{c}}/2\pi = 55.2$ MHz/V, as well as a proper measurement for χ. The fitted coherence time in the presence of cavity occupation is $T_2 = 17.5\ \mu$s. What limits the T_2 in this case is the dephasing induced by the population of the cavity.

6.12 STATE PREPARATION AND FIDELITY MEASUREMENT

6.12.1 *Fidelity estimation*

The goal of this section is to define the fidelity estimators used to construct the Figure 6.4 of the main text. For any target state $|\mathcal{T}\rangle =$

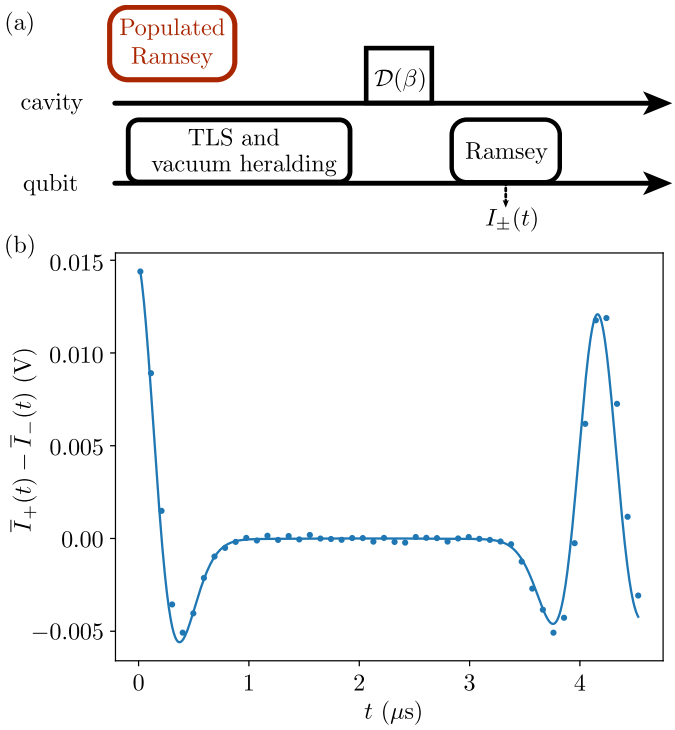


Figure 6.17: Cavity displacement calibration. (a) Pulse sequence of the populated Ramsey measurement scheme. (b) Blue dots: measured signal for a 200 ns long displacement pulse of amplitudes $I_c(t) = 0.03$ V and $Q_c(t) = 0$. Solid line: fit using Eq. (6.19) with $T_2 = 17.5$ μs and $\beta = 2.1$.

$|\psi_c\rangle\,|0\rangle$, the fidelity of the prepared state $\hat{\rho}$ of the cavity and qubit bipartite system is defined as

$$\mathcal{F} = \langle \mathcal{T}|\hat{\rho}|\mathcal{T}\rangle. \tag{6.20}$$

Using the decomposition $\hat{\rho} = \hat{\rho}^c_{00}\,|0\rangle\langle 0| + \hat{\rho}^c_{11}\,|1\rangle\langle 1| + \hat{\rho}^c_{01}\,|0\rangle\langle 1| + \hat{\rho}^c_{10}\,|1\rangle\langle 0|$, we have

$$\mathcal{F} = \langle\psi_c|\hat{\rho}^c_{00}|\psi_c\rangle = p_0\langle\psi_c|\hat{\rho}^c_0|\psi_c\rangle, \tag{6.21}$$

where $\hat{\rho}^c_0 = \frac{\hat{\rho}^c_{00}}{\mathrm{Tr}(\hat{\rho}^c_{00})}$ is a properly normalized density matrix for the cavity, and $p_0 = \mathrm{Tr}(\hat{\rho}^c_{00})$ is the probability of finding the qubit in the ground state. This fidelity is thus the product of the probability of

finding the qubit in the ground state and the fidelity of the cavity conditioned to the qubit being in its ground state:

$$\mathcal{F} = p_0 \mathcal{F}^c. \tag{6.22}$$

The probability p_0 is directly obtained by measuring the qubit at the end of the pulse sequence. Measuring $\mathcal{F}^c$ can be done with Wigner tomography: we use the fact that

$$\mathcal{F}^c = \pi \int_{\mathbb{C}} W_{|\psi_c\rangle\langle\psi_c|}(\beta) W_{\hat{\rho}_0^c}(\beta) \mathrm{d}\beta. \tag{6.23}$$

where $W_{|\psi_c\rangle\langle\psi_c|}(\beta)$ is the Wigner function of the target state, and $W_{\hat{\rho}_0^c}$ that of the state represented by $\hat{\rho}_0^c$.

Let us assume that β is picked at random in the complex plane following a law $p(\beta)$. We can rewrite this expression of the fidelity as the expectation

$$\mathcal{F}^c = \pi \mathbb{E}\left[\frac{W_{|\psi_c\rangle\langle\psi_c|}(\beta)}{p(\beta)} W_{\hat{\rho}_0^c}(\beta)\right]. \tag{6.24}$$

This expression still holds when replacing the function $W_{\hat{\rho}_0^c}(\beta)$ by a random variable $\tilde{W}_{\mathrm{exp}}(\beta)$ representing the outcome of the measurement of the observable $2\hat{\Pi}(\beta)/\pi$ with mean value $W_{\hat{\rho}_0^c}(\beta)$. Experimentally, the measurement record is indeed typically a binary random number in $\{\frac{2}{\pi}, -\frac{2}{\pi}\}$. A widely used choice for $p(\beta)$ is a square uniform distribution

$$p_{\mathrm{square}}(\beta) = \begin{cases} \dfrac{1}{\Delta x \Delta y} & \text{if } |\mathrm{Re}(\beta)| \leq \dfrac{\Delta x}{2} \\ & \text{and } |\mathrm{Im}(\beta)| \leq \dfrac{\Delta y}{2} \\ 0 & \text{otherwise} \end{cases} \tag{6.25}$$

approximated by averaging $\tilde{W}_{\mathrm{exp}}(\beta)$ on a grid of pixels. In this case, for each β, $W_{\mathrm{exp}}(\beta)$ is the mean value of $\tilde{W}_{\mathrm{exp}}(\beta_i)$. We choose $n_x n_y$ evenly spaced pixels on a grid of size $\Delta x \times \Delta y$ centered around 0 in the complex plane. This is the most natural way of reconstructing the Wigner function, as it is visually exhaustive. However, it is not the most efficient in terms of the number of samples. In [5], it is demonstrated that the most efficient sampling strategy is given by

$$p_{\mathrm{opt}}(\beta) = \frac{|W_{|\psi_c\rangle\langle\psi_c|}(\beta)|}{\| W_{|\psi_c\rangle\langle\psi_c|} \|_1} \tag{6.26}$$

with $\| W_{|\psi_c\rangle\langle\psi_c|} \|_1 = \int_{\mathbb{C}} |W_{|\psi_c\rangle\langle\psi_c|}(\beta)| \mathrm{d}\beta$. The procedure then consists in sampling N points $\{\beta_i\}$ in the complex plane according to the law p_{opt}, and in recording one value $\tilde{W}_{\mathrm{exp}}(\beta_i)$ for each of these points. We can then build the estimators $\tilde{\mathcal{F}} = p_0 \tilde{\mathcal{F}^c}$ of fidelity for both strategies. In general, we have

$$\tilde{\mathcal{F}^c} = \frac{\pi}{N} \sum_i \frac{W_{|\psi_c\rangle\langle\psi_c|}(\beta_i)}{p(\beta_i)} \tilde{W}_{\mathrm{exp}}(\beta_i). \tag{6.27}$$

Therefore, in the case of the uniform sampling, the fidelity estimator reads

$$\tilde{\mathcal{F}^c}_{\mathrm{square}} = \frac{\pi \Delta x \Delta y}{N} \sum_i W_{|\psi_c\rangle\langle\psi_c|}(\beta_i) \tilde{W}_{\mathrm{exp}}(\beta_i). \tag{6.28}$$

In contrast, in the case of the optimal sampling it reads

$$\tilde{\mathcal{F}^c}_{\mathrm{opt}} = \frac{\pi \| W_{|\psi_c\rangle\langle\psi_c|} \|_1}{N} \sum_i \frac{W_{|\psi_c\rangle\langle\psi_c|}(\beta_i)}{|W_{|\psi_c\rangle\langle\psi_c|}(\beta_i)|} \tilde{W}_{\mathrm{exp}}(\beta_i). \tag{6.29}$$

Note that the expression (6.28) of $\tilde{\mathcal{F}}_{\mathrm{square}}$ is formally equivalent to computing a discrete version of Eq. (6.23) between the target Wigner function and the experimental one. The estimated fidelities are compared for both strategies in Fig.6.18 where the improvement in the statistical error can be seen in the error bars. Some values differ between both methods, which gives a sense of the uncertainty originating from the low frequency drifts in the contrast $\tilde{c}$ due to drifts in the qubit coherence time T_2 (see Sec. 6.14.2).

6.12.2 *Experimental implementation*

The schematics of the State Preparation and Tomography (SPM) pulse sequence is presented in Fig. 6.19a and b. Two parity measurements with opposite polarity are performed for each amplitude β_i. Wigner measurement of the vacuum at $\beta = 0$ are interleaved to calibrate the possibly drifting contrast of the Wigner measurement.

Before each sequence, a heralding sequence on the TLS and on the cavity emptiness is performed to start with a state as pure as possible before applying the pulse sequence. For each α and φ, the SPM sequence is performed for $N = 10^5$ times for the optimal strategy,

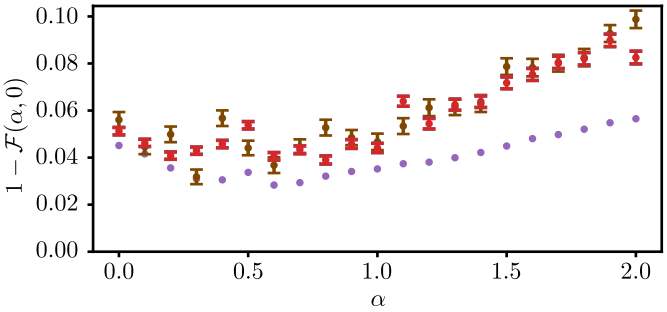


Figure 6.18: Comparison between optimal and square sampling on fidelity estimation. Red dots: Fidelity measured with optimal sampling. Brown dots: Fidelity measured with square sampling. The error bars are wider for about the same number of samples. Purple dots: theoretical expectation from the solid line in Fig. 6.4a.

and $N = 21 \times 51 \times 10^2$ times for the square strategy. Knowing that the cavity is on average in a thermal state with $n_{\text{th}} = 0.006$, we now define the experimental contrast $\tilde{c}$ and Wigner estimator $\tilde{W}_{\text{exp}}(\beta_i)$:

$$\tilde{W}^{\pm}_{\text{exp}}(\beta_i) = \frac{2}{\pi} r^{\pm}_c(\beta_i) \tag{6.30}$$

$$\tilde{W}_{\text{exp}}(\beta_i) = \tilde{W}^{+}_{\text{exp}}(\beta_i) - \tilde{W}^{-}_{\text{exp}}(\beta_i) \tag{6.31}$$

Note that with this way of sampling the Wigner function, $\tilde{W}_{\text{exp}}(\beta_i)$ is a random number that can take discrete values in $\{-\frac{2}{\pi}, 0, \frac{2}{\pi}\}$.

We now define the estimator $\tilde{\mathcal{F}}$ of the fidelity $\mathcal{F}$:

$$\tilde{c} = \frac{1}{(1-2n_{\text{th}})N} \sum_{i=1}^{N} (r^{+}_{\text{contrast,i}} - r^{-}_{\text{contrast,i}}) \tag{6.32}$$

$$\tilde{\mathcal{F}} = \frac{\pi}{\tilde{c}N} \sum_i \frac{W_{|\psi_c\rangle\langle\psi_c|}(\beta_i)}{p(\beta_i)} \tilde{W}_{\text{exp}}(\beta_i) \tag{6.33}$$

This renormalization by the contrast $\tilde{c}$ is necessary to take into account imperfections of the measurement of $\hat{\Pi}(\beta)$, in particular the finite coherence time of the qubit and the $\pi/2$ pulses and readout fidelities. It is also crucial to interleave the contrasts measurement and the fidelity estimation in a SPM sequence, as the contrast typically drifts by 10 % over several realization of fidelity measurement (typically

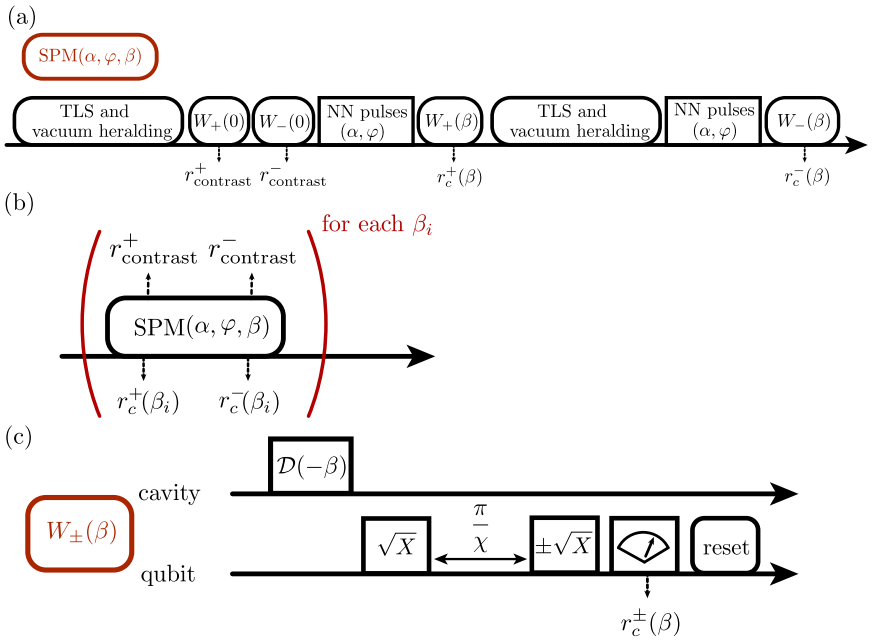


Figure 6.19: State Preparation and Measurement (SPM) sequence. (a) High-level diagram of the SPM sequence. It lasts 8 ms for each realization, which is dominated by the cavity reset time. (b) High-level diagram of the Wigner sampling scheme. (c) Pulse sequence for the Wigner measurement

comprising $N = 10^5$ SPM sequences). These drifts are the main reason why Fig. 6.18 red and brown do not always agree. The contrast $\tilde{c}$ typically ranges from 0.8 to 0.9.

Since we do not measure the qubit at the end of the NN drive pulses, and do not reset it, doing the Wigner tomography right away biases the estimator $\tilde{\mathcal{F}}$ towards

$$\mathcal{F}' = \langle\psi_c|\hat{\rho}^c_{00}|\psi_c\rangle - \langle\psi_c|\hat{\rho}^c_{11}|\psi_c\rangle. \tag{6.34}$$

This can be simply understood: if we post-select the qubit to be in $|1\rangle$ at the beginning of the Wigner pulse sequence, then the probabilities for it to be in $|0\rangle$ or $|1\rangle$ at the end are flipped compared to the situation where it starts in $|0\rangle$. The signal $\tilde{W}_{\mathrm{exp}}(\beta_i)$ becomes an estimator for

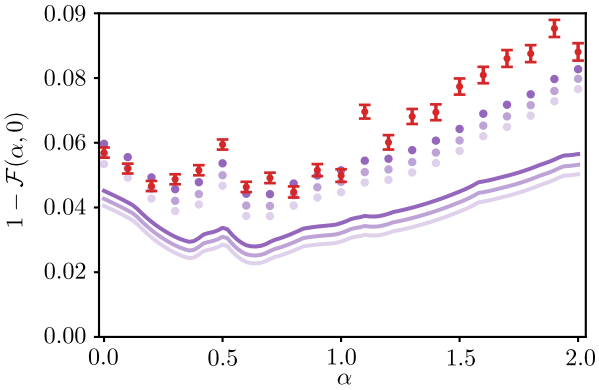


Figure 6.20: Sensitivity of the fidelity estimation to the T_2 time of the qubit. (a) Red dots: measured infidelity $1-\mathcal{F}(\alpha,0)$. Each point is obtained by averaging 10^6 samples. Purple solid line: simulated infidelity $1-\mathcal{F}(\alpha,0)$ for the same parameters, for $T_2 = 42$ μs, $T_2 = 50$ μs and $T_2 = 60$ μs (dark to bright). Purple points: infidelities obtained by simulating also the experimental fidelity measurement procedure and taking into account the finite thermal population of the cavity, for $T_2 = 42$ μs, $T_2 = 50$ μs and $T_2 = 60$ μs (dark to bright).

$-W_{\hat{\rho}_1^c}$, where we defined similarly $\hat{\rho}_1^c = \frac{\hat{\rho}_1^c}{\mathrm{Tr}(\hat{\rho}_1^c)}$. In the end, our estimator $\tilde{W}_{\mathrm{exp}}(\beta_i)$ becomes such that:

$$\mathbb{E}[\tilde{W}_{\mathrm{exp}}(\beta_i)] = p_0 W_{\hat{\rho}_0^c}(\beta_i) - (1-p_0) W_{\hat{\rho}_1^c}(\beta_i) \tag{6.35}$$

$$\mathbb{E}[\tilde{\mathcal{F}}] = \mathcal{F}' = \mathcal{F} - \mathcal{F}_1, \tag{6.36}$$

where $\mathcal{F}_1 = \langle\psi_t|\, \hat{\rho}_1^c (1-p_0)\, |\psi_t\rangle$.

The bias of this estimator thus depends on whether $\hat{\rho}_1^c$ is close to $|\psi_c\rangle$ or not.

6.12.3 *Wigner functions of various states*

A series of additional experimental tomography of several states is shown Fig 6.21. For the square sampling strategy (Fig. 6.21a), each pixel is averaged 1000 times. To reconstruct Wigner functions from the optimal strategy (Fig. 6.21b), we build an histogram of the $\{\beta_i\}$ with 100×100 bins and average the values of the $\{\tilde{W}_{\mathrm{exp}}(\beta_i)\}$ corresponding

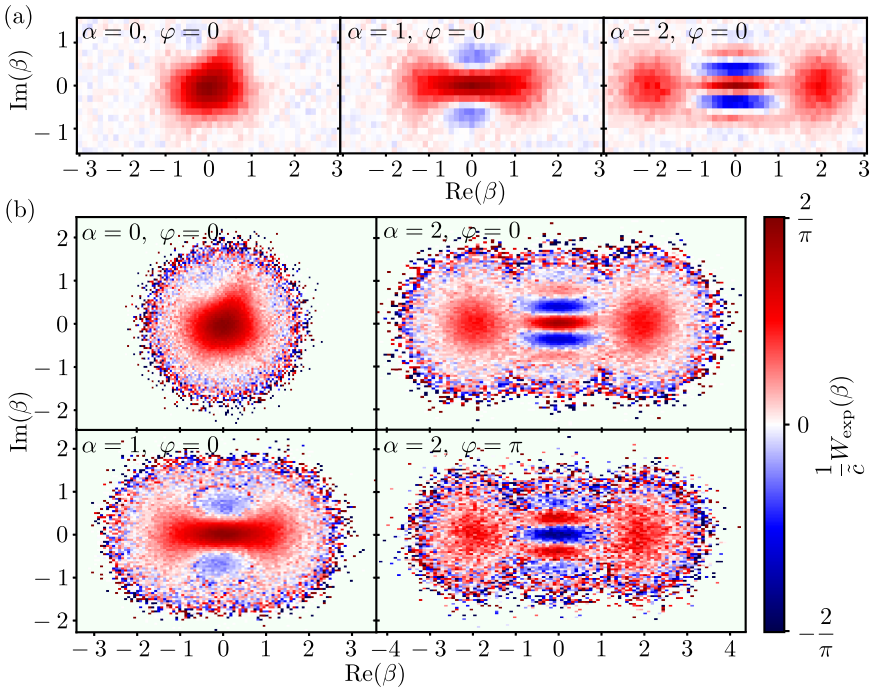


Figure 6.21: Wigner functions of various cat states. (a) Renormalized reconstructed Wigner functions for several parameters (indicated in the top left) of the neural network, with square sampling. Estimated fidelities are $\mathcal{F}(0,0) = 93.8\ \% \pm 0.3\ \%$, $\mathcal{F}(2,0) = 89.5 \pm 0.4\ \%$. (b) Reconstructed Wigner functions for several parameters of the neural network, with the optimal sampling. $N = 10^6$ samples were used for $\varphi = 0$, $N = 10^5$ for $\varphi = \pi$. estimated fidelities are $\mathcal{F}(0,0) = 94.3\ \% \pm 0.2\ \%$, $\mathcal{F}(1,0) = 95.0\ \% \pm 0.2\ \%$, $\mathcal{F}(2,0) = 91.1\ \% \pm 0.3\ \%$ and $\mathcal{F}(2,\pi) = 91.5 \pm 0.8\ \%$.

to each bin to get $W_{\mathrm{exp}}(\beta_i)$ inside each bin. The resulting Wigner is then very noisy when $W_{|\psi_c\rangle\langle\psi_c|}$ is close to zero. The large Signal-to-Noise Ratio (SNR) thus concentrates on the regions where the target Wigner function has the highest values.

In order to more closely visualize the fidelity of this Wigner function, we renormalize it by $\tilde{c}$. For visualization purposes, we kept the colorbar between $-\frac{2}{\pi}$ and $\frac{2}{\pi}$, even though some pixels contain the average of only a few $\{\tilde{W}_{\mathrm{exp}}(\beta_i)\}/\tilde{c} \in \{-\frac{2}{\pi\tilde{c}}, 0, \frac{2}{\pi\tilde{c}}\}$, which can be out of range.

Note that this does not happen on well averaged pixels. This allows to take into account the imperfections of the measurement apparatus to fully recover the contrast of the experimental Wigner function.

6.13 OPTIMIZATION USING GRAPE FOLLOWED BY KROTOV

In this section, we provide details on the numerical optimization of the control pulses using the more traditional technique denominated as GRAPE followed by Krotov [3, 4, 236], and about the experimental implementation of these pulses.

6.13.1 *Numerical optimization*

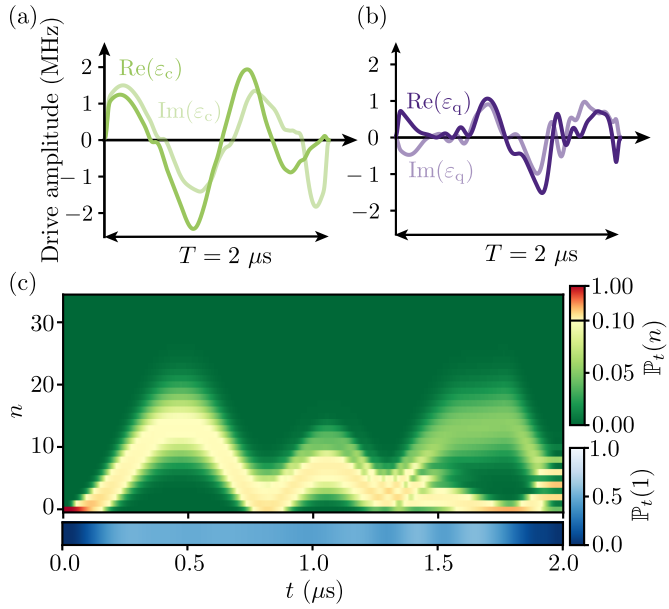


Figure 6.22: GRAPE & Krotov-optimized control of the cavity and qubit. (a,b) Control pulses generated using the GRAPE & Krotov method for $\alpha = 2$ and $\varphi = 0$. (c) Simulated probability distribution of the photon number (top) and qubit excitation (bottom) as a function of time for the control fields shown in (a) and (b). The fidelity of the predicted final state to $|C_\alpha^\varphi\rangle$ is $\mathcal{F}(2, 0) = 95.7$ %.

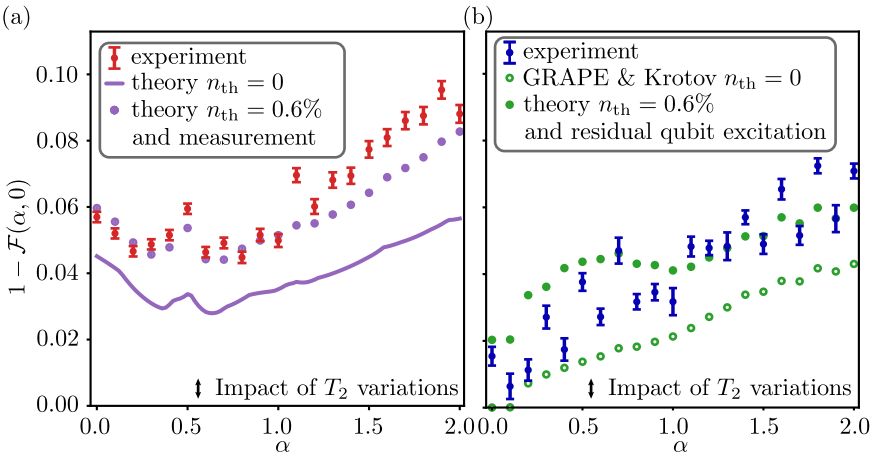


Figure 6.23: Comparison between the GRAPE & Krotov method and the NN method. (a) Reproduction of Fig. 6.4 showing the predicted and measured infidelities when generating the pulses with the NN. (b) Green open circles: simulated infidelity using pulses optimized by GRAPE followed by the Krotov method [3, 4, 236]. Green dots: infidelities obtained by simulating also the residual qubit excitation during the measurement procedure and taking into account the finite thermal population $n_{\rm th}$ of the cavity. Blue dots with error bars: measured infidelity $1 - \mathcal{F}(\alpha, 0)$ for 21 values of α using the control sequences provided by GRAPE followed by the Krotov method. The error bars are statistical.

GRAPE and Krotov methods are two optimization algorithms that aim at minimizing the same loss function $\mathcal{L}$, which is the infidelity of Eq. (6.5). We respectively implemented them in python using the packages QuTiP 4.7.6 [236] and Krotov 1.3.0 [4]. This implementation of GRAPE only works on Hamiltonian evolution but is easy to parallelize. Feeding the output of GRAPE as an initial guess for the Krotov method allows to take into account decoherence. In practice, we combine the two for generating the four control sequences $\mathrm{Re}(\varepsilon_{\rm c}(t))$, $\mathrm{Im}(\varepsilon_{\rm c}(t))$, $\mathrm{Re}(\varepsilon_{\rm q}(t))$, and $\mathrm{Im}(\varepsilon_{\rm q}(t))$ as follows:

1. Run GRAPE algorithm based on the unitary evolution set by the Hamiltonian (6.1) on a grid of 163 time steps uniformly spanning a time interval of 1956 ns. The algorithm is initialized with sinusoidal functions (one of the few allowed initialization

in the optimize_pulse_unitary function of control.pulseoptim in QuTiP 4.7.6) ;

2. Upsample the resulting GRAPE optimization results into a denser time grid of 2000 time steps over the total duration time $T = 2$ μs. Then smooth the pulses by convolving them with a Gaussian function (standard deviation of 11 ns). This process has the embedded benefit of flattening the beginning and end of the pulses and to limit the bandwidth of the control signals to an experimentally feasible range;

3. Run Krotov's method based on the Lindblad master equation on the 2000 time steps using the smoothed GRAPE pulses as the initial guess.

Note that the dimension of the Hilbert space we use grows with cat state amplitude α. The dimension $N_{\max}$ is the nearest integer to $16 + 4\alpha + 2\alpha^2$.

An example of control pulses is shown in Fig. 6.22a,b for the optimized preparation of $|C_\alpha^\varphi\rangle$ with $\alpha = 2$ and $\varphi = 0$. They present more abrupt changes in time than the pulses that are generated by the neural network (Fig. 6.2), which can be explained by the freedom offered by the tuning of every 2000 time steps as opposed to the 9 coefficients of the B-spline decomposition. The photon number distribution is shown in Fig. 6.22c.

6.13.2 *Theoretical infidelities*

We solve the Lindblad master equation (6.9) using the solution of GRAPE & Krotov obtained for each amplitude α. The computed infidelity at time T is shown in Fig. 6.4a.

Since we implement this optimization method experimentally on a different cool-down, we rerun everything for the slightly different parameters of the second cool-down ($\chi/2\pi = 237.5$ kHz, $T_q = 38$ μs, $T_{q,\varphi} = 70.5$ μs and $T_c = 220$ μs) and plot the predicted fidelities in Fig. 6.23b as a function of α (green open circles).

To compare with the experimental results, we compute how the estimation of infidelity is affected by the Wigner tomography errors.

If we take into account the residual excitation probability of the qubit and the thermal occupation of the cavity $n_{\mathrm{th}} = 0.6$ % as in Fig. 6.24, we obtain an overestimation of the infidelity as shown in Fig. 6.23 (green dots). This overestimation is similar in amplitude to the one we find for the control pulses generated by the NN (purple dots compared to purple solid line in Fig. 6.4a).

6.13.3 *Experimental implementation of GRAPE*

We implement the 21 control sequences generated by GRAPE & Krotov on a different cool-down for the same device as the one used with the NN. The procedure for the fidelity estimation is the same as the one described in Sec. 6.12.2, except for the computation of $\tilde{c}$. It is now estimated by measuring an evenly distributed grid of 11×11 complex amplitudes $\beta = x + iy$ for the Wigner function $W_0(\beta)$ where $-1 \leq x, y \leq 1$, obtained when the cavity is heralded in the vacuum. In contrast, the formula given in Eq. (6.32) only relies on the point $\beta = 0$.

The resulting measured infidelity is plotted as blue dots as a function of α in Fig. 6.23b. The error bars represent statistical errors. The number of averaging of each point varies owing to frequent temperature rises of the refrigerator triggered by another experiment. The fidelity increases with amplitude α similarly as the prediction. However, qubit frequency drifts and thermal occupation of the cavity varied often during the run leading to a spread of the measured fidelities beyond the statistical uncertainty.

Overall, Fig. 6.23a,b show that the experimentally observed fidelities are close to the predicted infidelities we should observe, owing to measurement errors. In Fig. 6.23, the actual state fidelities at the end of the preparation are represented by the solid purple line for the NN and by the open green circles for GRAPE & Krotov. Both differ by about 0.01 for large amplitudes α.

6.14 ERROR ANALYSIS

6.14.1 *Error budget*

In order to understand the difference between the prediction of the state preparation fidelities (purple in Fig. 6.4a) and the measured ones (red in Fig. 6.4a), we numerically solve a Lindblad master equation of the whole measurement sequence that takes into account the three phenomena leading to infidelity losses $\Delta\mathcal{F}_{\text{th}}$ due to the remaining thermal population of the cavity before the pulse sequence, $\Delta\mathcal{F}_{\text{dissip}}$ due to dissipation during the parity measurement sequence, and $\Delta\mathcal{F}_{\text{ex}}$ due to the remaining qubit excited population before the parity measurement sequence. Simulations removing separately these three effects reveal their respective contributions. Results are shown Fig. 6.24. The thermal population n_{th} systematically leads to $\Delta\mathcal{F}_{\text{th}} \simeq n_{\text{th}} = 0.6$ %, and $\Delta\mathcal{F}_{\text{par}}$ varies monotonically from 0.6 % for small cat states to 2 % for the largest values. It is negative for small alphas, as relaxation towards the vacuum actually improves the fidelity. $\Delta\mathcal{F}_{\text{ex}}$ is of the order of 2 %. Its non trivial variations depend on the way the NN was trained.

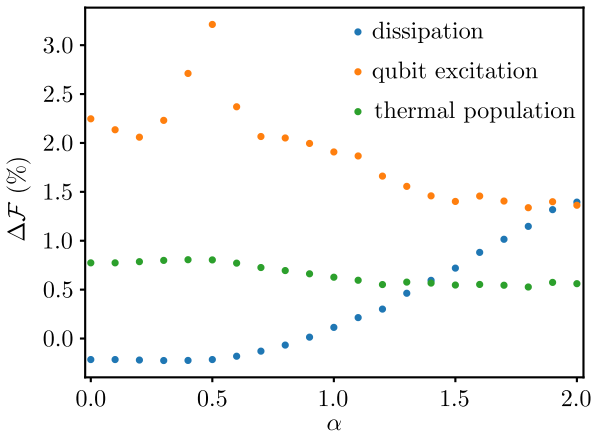


Figure 6.24: Influence of the measurement imperfections on the experimental fidelity estimation. Blue dots: contribution to the infidelity $\Delta\mathcal{F}_{\text{dissip}}$ due to the dissipation during the parity measurement. Orange dots: contribution due to the residual qubit excitation at the end of the pulse sequence. Green dots: effect of the thermal population of the cavity $n_{\text{th}} = 0.6$ %.

6.14.2 *Effect of T_2 fluctuations*

The effect of T_2 fluctuations are shown in Fig. 6.20. Simulations are run in the same way as Fig. 6.4(a), but T_2 of 42 (dark), 50 (medium) and 60 μs (bright) are considered. The mismatch between theory and experiment could thus be explained by variations in T_2. However, we still see that the experiment has a tendency to perform worse for larger alphas compared to the prediction. We attribute this either to imprecision in the measured value of χ, or higher order processes taking place during the preparation. The consequences of these two reasons become more important when the cavity is loaded with more photons during the preparation. As the intermediate number of photons in the cavity typically increases for larger values of α, we expect a larger deviation for larger α.

Part V

APPENDIX

A

QUANTUM INFORMATION TOOLS

This appendix aims to introduce the main quantum information tools used in this thesis.

A.1 QUANTUM COMMUNICATION

A.1.1 *Classical channel*

A classical communication channel is depicted in Fig. A.1. The principle is as follows: n codewords (here 0 to $n-1$) are used to convey a message through a noisy channel $\mathcal{N}$. As a result, when Alice sends the word a, Bob receives an outcome b with a probability conditioned on the codeword $p_a(b)$. This gives an overall probability distribution for b:

$$p(b) = \sum_a p_a(b)\mathbb{P}(a) \tag{A.1}$$

The amount of information that can be transferred with a single use

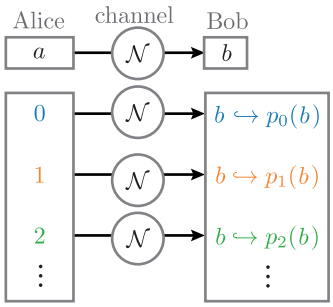


Figure A.1: Scheme of a classical communication channel. Alice communicates with Bob through a noisy channel using letters a (here numbers) with a prior probability distribution $\mathbb{P}(a)$. Bob receives outcomes that follow a distribution that depends on the letters Alice sent, which allows him to infer the most likely message.

of this channel is computed using the mutual information $I(a:b)$, which depends on the probability distribution $\mathbb{P}(a)$ of the codewords:

$$I(a:b) = H(b) - \sum_a \mathbb{P}(a) H(b|a) \tag{A.2}$$

where $H(b) = -\sum_b p(b)\log(p(b))$ is the Shannon entropy of the distribution of b, which depends on $\mathbb{P}(a)$, and $H(b|a) = -\sum_b p(b|a)\log(p(b|a))$ the entropies of the distributions $p(b|a)$ knowing the value of a. This expression is simple: it is the total entropy of b lowered by the average entropy of b when a is known.

The mutual information has an interesting property, which is that it is additive: the information conveyed when sending (independently) m words $a_1, a_2 ... a_m$ and receiving outcomes $b_1, b_2 ... b_m$ is

$$I_m\big((a_1, a_2, ..., a_m) : (b_1, b_2 .. b_m)\big) = mI(a:b). \tag{A.3}$$

A.1.2 *Quantum channel*

The simplest way to extend the classical picture is illustrated in Fig. A.2. The codewords a are associated to orthogonal states $|\psi_a\rangle \in \mathcal{H}_A$, that are sent by Alice through a noisy quantum channel $\hat{\mathcal{N}}$. When Alice wants to send the word a, Bob receives a corresponding density matrix ρ_a on a Hilbert space $\mathcal{H}_B$ and can perform a measurement on it, which gives him an outcome b. The corresponding density matrix of the system reads

$$\hat{\rho}_{AB} = \sum_a \mathbb{P}(a)\, |\psi_a\rangle\langle\psi_a| \otimes \rho_a. \tag{A.4}$$

This is then essentially a classical communication channel carried by quantum states. The information carried is then computed the same way as in the previous section. Here, Bob is free to choose the measurement (a POVM) he wants to perform on the state. The information carried by the corresponding channel thus depends on the choice of measurement. As an example, if the codewords are $|\psi_0\rangle = |g\rangle$ and $|\psi_1\rangle = |e\rangle$ the two states of a qubit, which are transferred perfectly ($\rho_0 = |g\rangle\langle g|$, $\rho_1 = |e\rangle\langle e|$), Bob can choose to measure $\hat{\sigma}_z$, with outcomes 1 or -1. The information is perfectly transferred, and Alice is able to send $\log 2$ nats (or one bit) of information.

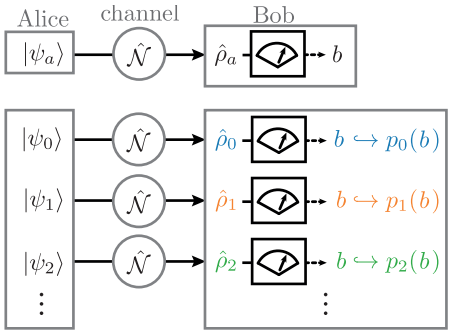


Figure A.2: Scheme of a quantum communication channel. Alices sends classical information in the form of quantum states indexed by a through a quantum channel $\hat{\mathcal{N}}$. Bob performs a quantum measurement (POVM) on the density matrices he gets, which gives him outcomes b.

However, if Bob chooses to measure $\hat{\sigma}_x$, the distribution of the outcomes will be the same (50 % chance for each) for whatever state Alice sent. Thus, the mutual information is zero, and no information has been transferred: the information has been destroyed, and Bob has no way of recovering it.

A third case that can be considered is when Bob performs a weak measurement of $\hat{\sigma}_z$ on the qubit. He will get a little information, or a "hint" of the state of the qubit, but will not be able to discriminate the two cases fully. However, further measurements may still give him more information: some information still remains in the state of the qubit. In [104], Han and coworkers formalize this decomposition into the information available to Bob into the information he extracts from his measurement, the information he destroys, and the information remaining in the state.

The information available to Bob is obtained by performing the measurement that maximizes the mutual information and is called the accessible information $I_{\text{acc}}(A : B)$. Given a set of measurement operators $\Pi = \{\hat{\Pi}_0, \hat{\Pi}_1, ...\}$ such that $\sum_k \hat{\Pi}_k^\dagger \hat{\Pi}_k = \mathbb{1}$, we note $I = I(a : b_\Pi)$ when b_Π is the outcome of the measurement.

$$I_{\text{acc}} = \max_{\Pi}(I(a : b_\Pi)). \tag{A.5}$$

As for the classical case, when repeating this procedure m times, the total transferred information will be m times the information for each step.

A.1.3 *Quantum channel with global measurements*

Using m times the same channel is equivalent to considering a bigger Hilbert space $\mathcal{H}_A^{\otimes m}$ and $\mathcal{H}_B^{\otimes m}$. Alice sends product states made of her initial set of states, and Bob performs a measurement comprised of *local* POVM on each of the ρ_i. We can now tweak the rules of the game

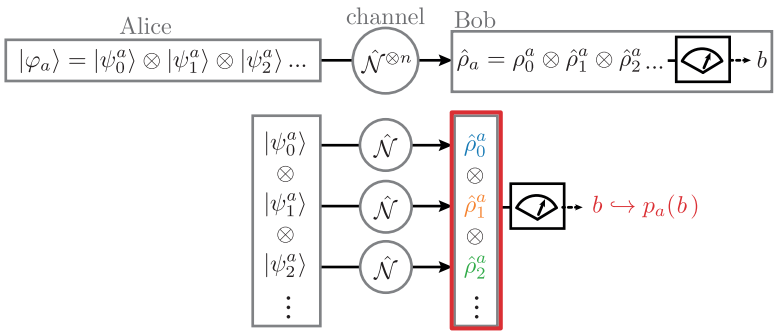


Figure A.3: Scheme of a global quantum communication strategy. Alice sends multiple states $|\psi_k^a\rangle$ into the quantum channel. This forms a product state $|\psi_0^a\rangle \otimes |\psi_a^a\rangle \otimes |\psi_2^a\rangle \ldots$. Bob can perform a *joint* measurement on the state $\hat{\rho}_k^a$ he gets. This allows us to beat the local measurement strategy shown in Fig. A.2.

a bit. What if Alice still uses product states (possibly different from the original ones), but now Bob can use whatever *global* measurement on this bigger Hilbert space? This is the situation described in Fig. A.3. The Holevo-Schumacher-Westmoreland theorem [240, 241] states that the information that we can carry this way can be larger than mI_{acc}. It is equal to m times the Holevo bound χ [242], that reads:

$$\chi = S\big(\sum_a \mathbb{P}(a)\hat{\rho}_a\big) - \sum_a \mathbb{P}(a) S\big(\hat{\rho}_a\big) \tag{A.6}$$

with $S(\hat{\rho}) = -\hat{\rho}\log(\hat{\rho})$ the von Neumann entropy of $\hat{\rho}$. This quantity was first derived as a bound for classical information transmittable through a quantum channel by Holevo [242], before being shown to

be asymptotically reached when allowing many copies for Alice (still restricting to product states) and joint measurements for Bob with the Holevo-Schumacher-Westmoreland theorem [240, 241].

The Holveo bound is the *quantum* mutual information of the state describing this quantum communication scheme. The quantum mutual information $I_{\mathrm{q}}(A:B)$ is defined as

$$I_{\mathrm{q}}(A:B) = S(\hat{\rho}_A) + S(\hat{\rho}_B) - S(\hat{\rho}_{AB}) = \chi \tag{A.7}$$

where $\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho}_{AB})$ and $\hat{\rho}_B = \mathrm{Tr}_A(\hat{\rho}_{AB})$ are the reduced density matrices of each subsystem. It describes the amount of quantum information that is shared between the subsystems A and B, and constitutes another possible generalization of mutual information to quantum systems, along with the accessible information.

We always have $I_{\mathrm{acc}} \leq I_{\mathrm{q}}$, and in general, $I_{\mathrm{acc}} < I_{\mathrm{q}}$. The difference is called the quantum discord [243] and provides a purely quantum additional resource that can allow us to beat classical schemes at certain tasks. The quantum radar [10, 244] is an example of such purely quantum advantage arising from this difference. Another example is shown in Sec 3.3.5 in the context of the dispersive readout of a qubit.

It can also be shown that

$$\chi \leq -\sum_a \mathbb{P}(a) \log \mathbb{P}(a). \tag{A.8}$$

It means that Alice cannot communicate more information than with a classical noiseless channel.

Additional quantum effects and advantages arise when Alice can use entangled states and when allowing shared entanglement between Alice and Bob. For example, the latter case allows Alice to send 2 bits of information with a single qubit. This protocol is called *superdense coding*, and examples can be found in Gyongyosi's work [245].

A.1.4 *Pure qubit case*

Quantum discord is a quantity that is hard to compute in general because computing I_{acc} is an NP-complete optimization problem [246], as one needs to find the optimal POVM. However, the case where Bob receives a qubit is simpler and has closed-form formulas in the

simplest cases [247, 248]. We will here consider the simple case where Bob gets pure qubit states $|b_0\rangle$ and $|b_1\rangle$ from Alice, which sends them with maximal entropy: $\mathbb{P}(0) = \mathbb{P}(1) = \frac{1}{2}$. This case is characterized by the overlap $|\langle b_0|b_1\rangle| = \cos\theta/2$ between the two states. Here, θ is the angle between their respective two Bloch vectors $\vec{u}_{|b_0\rangle}$ and $\vec{u}_{|b_1\rangle}$. Without loss of generality, up to a change of phase, we can write

$$|b_0\rangle = \cos\left(\frac{\pi-\theta}{4}\right)|g\rangle + \sin\left(\frac{\pi-\theta}{4}\right)|e\rangle \tag{A.9}$$

$$|b_1\rangle = \cos\left(\frac{\pi+\theta}{4}\right)|g\rangle + \sin\left(\frac{\pi+\theta}{4}\right)|e\rangle\,. \tag{A.10}$$

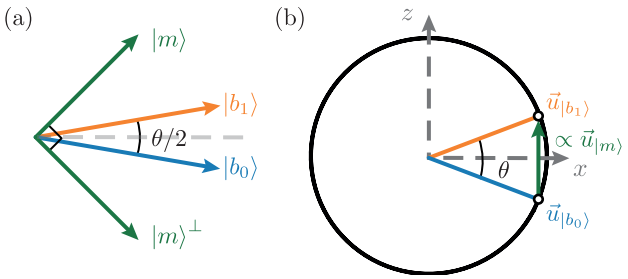


Figure A.4: Optimal measurement to discriminate two quantum states. (a) Representation of $|b_0\rangle$ and $|b_1\rangle$ in the Hilbert space. The optimal observable $|m\rangle\langle m|$ to measure is given by $|m\rangle$ and $|m\rangle^\perp$ as pictured. (b) Representation of $\vec{u}_{|b_0\rangle}$ and $\vec{u}_{|b_1\rangle}$ in the plane (xy) of the Bloch sphere. They form an angle θ, and the optimal measurement is given by $\vec{u}_{|m\rangle} \propto \vec{u}_{|b_1\rangle} - \vec{u}_{|b_0\rangle}$. Here, it corresponds to the measurement of $\hat{\sigma}_z$, as $\vec{u}_{|e\rangle}$ is parallel to the axis of the Bloch sphere.

This situation is pictured in Fig. A.4a in the Hilbert space. In [247], Fuschs shows that the optimal measurement is obtained with the observable $|m\rangle\langle m|$ given by the orthogonal kets $|m\rangle$ and $|m\rangle^\perp$ such that they are symmetric with respect to the bisector of $|b_0\rangle$ and $|b_1\rangle$ (grey dashed line). The geometrical construction is simpler in the Bloch sphere, pictured Fig. A.4. The ket $|m\rangle$ is given by its Bloch

vector $\vec{u}_{|m\rangle} \propto \vec{u}_{|b_1\rangle} - \vec{u}_{|b_0\rangle}$ in the Bloch sphere. Here, it corresponds to measuring $|e\rangle\langle e|$, which gives $b = 0$ or 1. We write

$$p(b=0) = p(b=1) = p(b) = \frac{1}{2} \tag{A.11}$$

$$p(b=0|a=0) = p(1|0) = \frac{1+\sin\left(\frac{\theta}{2}\right)}{2} \tag{A.12}$$

$$p(b=1|a=0) = p(0|1) = \frac{1-\sin\left(\frac{\theta}{2}\right)}{2}, \tag{A.13}$$

which gives the following accessible information:

$$\begin{aligned} I_{\text{acc}}(A:B) = &- p(0)\log(p(0)) - p(1)\log(p(1)) \\ &+ \mathbb{P}(0)\Big(p(0|0)\log(p(0|0)) + p(1|0)\log(p(1|0))\Big) \\ &+ \mathbb{P}(1)\Big(p(0|1)\log(p(0|1)) + p(1|1)\log(p(1|1))\Big) \end{aligned} \tag{A.14}$$

and finally:

$$\begin{aligned} I_{\text{acc}}(A:B) = &\frac{1}{2}\Bigg((1+\sin(\theta/2))\log(1+\sin(\theta/2)) \\ &+ (1-\sin(\theta/2))\log(1-\sin(\theta/2))\Bigg) \end{aligned} \tag{A.15}$$

To get the formula Eq. (3.90), we take the limit $\theta \ll 1$, which gives, at second order in θ

$$I_{\text{acc}}(A:B) \simeq \frac{\theta^2}{8}. \tag{A.16}$$

A.2 QUANTUM METROLOGY

A.2.1 *Classical Fisher Information*

The Fisher information is an important quantity in parameter estimation problems. We want to estimate λ the parameter of a probability distribution p_λ from n *independent* stochastic outcomes b_k, which follow the law given by p_λ. We can build an estimator $\tilde{\lambda}_n(\{b_k\})$ from these outcomes and assume that it is unbiased, *i.e.* that $\mathbb{E}[\tilde{\lambda}_n(\{b_k\})|\lambda_0] = \lambda_0$.

Here, $\mathbb{E}[X|\lambda_0]$ stands for the expectation value of X assuming the value λ_0 for λ. The precision of this estimator is given by its variance $\delta\tilde{\lambda}^2$. If p_λ is sufficiently regular, the Cramer-Rao bound gives a lower bound on the variance of this estimator [249]:

$$\delta\tilde{\lambda}_n^2 \geq \frac{1}{nF(\lambda)} \tag{A.17}$$

where $F(\lambda)$ is the Fisher information. It is defined as

$$F(\lambda_0) = \mathbb{E}\Big[\Big(\frac{\partial \log(p_\lambda(b))}{\partial\lambda}\Big)^2|\lambda_0\Big] = \sum_b \Big(\frac{\partial p_\lambda}{\partial\lambda}|_{\lambda=\lambda_0}(b)\Big)^2/p_{\lambda_0}(b). \tag{A.18}$$

Under additional regularity conditions [249], it also reads

$$F(\lambda_0) = -\mathbb{E}\Big[\frac{\partial^2 \log(p_\lambda(b))}{\partial\lambda^2}|\lambda_0\Big]. \tag{A.19}$$

As the mutual information (see Appendix A.1.1), the Fisher information is additive: iterating n independent times the same experiment gives a Fisher information that is n times that of a single experiment.

The existence of an estimator that saturates the Cramer-Rao bound is not guaranteed. However, a maximum likelihood estimator $\tilde{\lambda}_n^{\mathrm{m}}$ asymptotically reaches this lower bound:

$$\mathbb{E}[\tilde{\lambda}_n^{\mathrm{m}}|\lambda_0] \xrightarrow{n\to\infty} \lambda_0 \tag{A.20}$$

$$n\delta\tilde{\lambda}_n^{\mathrm{m}^2} \xrightarrow{n\to\infty} \frac{1}{F(\lambda_0)}. \tag{A.21}$$

The Fisher information then gives the accessible asymptotic precision in the parameter estimation.

A.2.2 *Link with the mutual information*

We can link the Fisher information to the mutual information obtained in a classical communication setting. Up to now, we supposed that Bob obtains the outcome b with probability $p_0(b)$ if Alice sends the bit $a = 0$, and with probability $p_1(b)$ if she sends the bit $a = 1$. Let us replace $p_1(b)$ by an aribtrary distribution $p_\lambda(b)$ parameterized by $\lambda \in \mathbb{R}$, such that $p_{\lambda=0} = p_0$. In the case where $\mathbb{P}(0) = \mathbb{P}(1) = \frac{1}{2}$,

we can establish a result very close to a known result in information theory relating the Kullback-Leibler divergence of the distribution p_λ with respect to p_0 with the Fisher information $F(\lambda_0 = 0)$ for small λ[1].

We write the mutual information $I_\lambda(a : b)$ as

$$I_\lambda(a : b) = -\sum_b \frac{p_0(b) + p_\lambda(b)}{2} \log\left(\frac{p_0(b) + p_\lambda(b)}{2}\right) + \frac{1}{2}\sum_b p_0(b)\log(p_0(b)) + \frac{1}{2}\sum_b p_\lambda(b)\log(p_\lambda(b)). \tag{A.24}$$

The first and second derivative for $\lambda = 0$ of this expression gives

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} I_\lambda(a : b)\big|_{\lambda=0} = 0 \tag{A.25}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} I_\lambda(a : b)\big|_{\lambda=0} = \frac{1}{4}\sum_b \left(\frac{\partial p_\lambda}{\partial \lambda}\big|_{\lambda=0}(b)\right)^2 / p_0(b) = \frac{1}{4}F(0) \tag{A.26}$$

where $F(0)$ is the Fisher information associated to the parameter λ of the distributions p_λ. For λ close to zero, we can expand $I_\lambda(a : b)$ up to the order 2, which gives

$$I_\lambda(a : b) \simeq \frac{F(0)}{8}\lambda^2. \tag{A.27}$$

The Fisher information thus becomes a relevant quantity to compute when working in a low information regime in a classical communication scheme.

1 The Kullback-Leibler divergence of the distribution p_λ with respect to p_0 is defined as

$$\mathrm{KL}(p_\lambda || p_0) = \sum_b p_\lambda(b)\log\left(\frac{p_\lambda}{p_0}\right). \tag{A.22}$$

If p_λ is sufficiently regular with respect to λ, we can link the Kullback-Leibler divergence to the Fisher information $F(\lambda = 0)$ associated to the parameter λ of the distributions by expanding it in second order with respect to λ [250]:

$$\mathrm{KL}(p_\lambda || p_0) \simeq \frac{F(0)}{2}\lambda^2. \tag{A.23}$$

A.2.3 *Quantum Fisher Information*

We can translate this parameter estimation problem to a parameter estimation involving quantum states. Given $\hat{O}$ an observable, $\hat{\mathcal{U}}(\lambda) = e^{-i\lambda\hat{O}}$ the unitary operation parameterized by λ it generates, and $|\psi_\lambda\rangle = \hat{\mathcal{U}}(\lambda)\,|\psi\rangle$ the probe states, evolved according to this unitary evolution. The question we want to answer is: if we have access to a system in state $|\psi_\lambda\rangle$, how precise can we be on the estimation of λ? In other words, what is the best measurement to perform to get the highest Fisher information? This is the Quantum Fisher Information (QFI). It is defined as

$$F_Q(\lambda_0) = \max_{\Pi} F_\Pi(\lambda_0) \tag{A.28}$$

where Π defines a quantum measurement and F_Π the Fisher information associated to its outcomes. In this case, where the probe states are pure, the QFI has a very convenient form [251]:

$$F_Q(\lambda_0) = 4\Delta\hat{O}^2_{|\psi_{\lambda_0}\rangle} = 4\Delta\hat{O}^2_{|\psi\rangle} \tag{A.29}$$

where

$$\Delta\hat{O}^2_{|\psi\rangle} = \langle\psi|\,\hat{O}^2\,|\psi\rangle - \langle\psi|\,\hat{O}\,|\psi\rangle^2 \tag{A.30}$$

is the variance of $\hat{O}$ on the state $|\psi\rangle$. Note that the measurement that achieves this bound is dependent on λ_0, so it is also an asymptotically reachable bound that needs to use the previous estimations of the parameter to get closer and closer to the optimal measurement.

B

MULTIMODE PHASE-SPACE REPRESENTATION

The P function, the Wigner function, and the Husimi Q function can be generalized to n modes. This is done with two modes using the characteristic functions in [28], but can be generalized to n modes, as done in [43, 252]. We present here a generalized version of some formulas given in Sec. 2.2.2.

B.1 MULTIMODE WIGNER FUNCTION

Given n modes, represented by their bosonic operators $\{\hat{a}_k\}$, we call $\hat{\mathcal{D}}_k$ the displacement operator of the k-th mode. We generalize the displacement operator as a function of $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_{n-1})$ one the system by

$$\hat{\mathcal{D}}(\boldsymbol{\beta}) = \prod_{k=0}^{n-1} \hat{\mathcal{D}}(\beta_k) \tag{B.1}$$

We can then generalize the formula given by Eq. (2.44) by [253]

$$\mathbf{W}(\boldsymbol{\beta}) = \frac{2^n}{\pi^n} \operatorname{Tr}\left(\hat{\mathcal{D}}(-\boldsymbol{\beta})\hat{\rho}\hat{\mathcal{D}}(\boldsymbol{\beta})\hat{\mathcal{P}}\right) \tag{B.2}$$

Here, $\hat{\mathcal{P}}$ is the parity operator, which reads

$$\hat{\mathcal{P}} = (-1)^{\sum \hat{a}_k^\dagger \hat{a}_k}. \tag{B.3}$$

B.2 MULTIMODE Q FUNCTION

From the generalized displacement operator, we can define a generalized coherent state by applying them on the vacuum state, defined as

$$|\text{vac}\rangle = \bigotimes_k |0\rangle_k \tag{B.4}$$

where $|0\rangle_k$ stands for the vacuum state of the k-th mode. It gives:

$$|\boldsymbol{\beta}\rangle = \hat{\mathcal{D}}(\boldsymbol{\beta})\,|\mathrm{vac}\rangle = \bigotimes_k |\beta_k\rangle_k \tag{B.5}$$

where now $|\beta_k\rangle_k$ is the state of the k-th mode in a coherent state of amplitude β_k.

The multimode Q function is then obtained by a generalization of Eq. (2.64) as

$$\mathbf{Q}(\alpha) = \frac{1}{\pi^n}\,\langle\boldsymbol{\beta}|\,\hat{\rho}\,|\boldsymbol{\beta}\rangle\,, \tag{B.6}$$

which defines a multidimensional probability distribution. Alternatively, it can also be obtained as the generalization of Eq. (2.67) by performing a multidimensional convolution between the Wigner function and the vacuum Wigner function $\mathbf{W}_0$:

$$\mathbf{Q} = \mathbf{W} * \mathbf{W}_0. \tag{B.7}$$

C

FABRICATION TECHNIQUES

We detail in this appendix the different experimental processes used to fabricate the samples used in Chap. 4 and Chap. 6. These samples are made of 200 nm of tantalum sputtered on a sapphire substrate grown using a Heat Exchanger Method (HEM). A first optical lithography is performed on 2' wafers to create the circuit, which consists of the capacitor of the transmon, the readout resonator, and the Purcell resonator for the tomography qubit part. For the photon counting part, it consists of the capacitor of the transmon and the filter. A second step, the electronic lithography, is performed to create the Josephson junctions of the transmons. The 3D cavity also goes through a cleaning step after machining, which we will also detail in a third part.

C.1 OPTICAL LITHOGRAPHY

The main steps of our optical lithography process are summarized in Fig. C.1. The first one consists in depositing a layer of optical resist on top of the tantalum. This layer is then exposed where the tantalum needs to be removed. The exposed resist can be selectively dissolved by a developer, which exposes the tantalum. Finally, exposing the sample to an etching agent removes the exposed tantalum, whereas the resist protects the unexposed regions.

C.1.1 *Cleaning*

The cleaning is done in acetone or NMP (or remover PG). The wafer is then rinsed 20 s in isopropanol (propan-2-ol), abbreviated IPA, and finally dried using nitrogen on a cleanroom wipe.

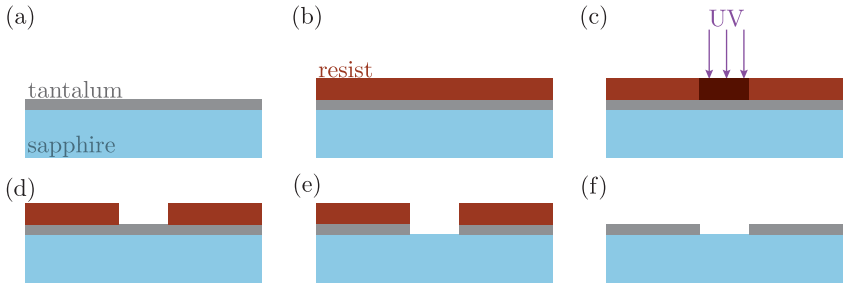


Figure C.1: Optical lithography procedure. (a) A layer of 200 nm of tantalum is sputtered on the sapphire substrate. This was done by STAR Cryoelectronics company. (b) The wafer is coated with a layer of optical resist. (c) The resist is exposed to UV light, which alters its structure. (d) The resist is developed, which dissolves the altered parts and exposes the tantalum. (e) The exposed tantalum is wet etched, drawing the desired pattern in the tantalum. (f) The resist is removed

C.1.2 *Coating*

The wafer is heated up to 115°C for 1 min to remove residual water. The wafer is then coated in an S1813 photoresist with a spincoater. The program is as follows:

- 500 rpm for 5 s, with an acceleration of 500 rpm/s ;
- 2000 rpm for 55 s, with an acceleration of 4000 rpm/s

The wafer is then baked 1 min at 115°C on a hot plate.

C.1.3 *Exposure*

The exposure was first done using a Smart Print from Microlight3D. It uses a projector and a digital micromirror device (DMD), which acts as a configurable mask to expose the desired areas. It then exposes the whole wafer by subdividing it into zones that are exposed one by one. While it can lead to so-called stitching errors when adjacent areas do not reconnect properly, it was not an issue regarding our samples, as the typical size of the details is sufficiently large. The right

exposure intensity is determined using a dose test, which consists in developing the same pattern exposed with increasing UV doses. The test is then visually examined to find the dose that renders the pattern with the highest precision. The value of the optimal exposure time was around 4s, which can vary over several months. Once again, the desired features were sufficiently large so that the time of exposure could be away from the optimal time by 50 % without significantly impacting the quality of the lithography.

The group acquired a μMLA Heisenberg laser lithography system during my PhD thesis. Compared to the Smart Print system, this system has a higher precision, and the stitching issues disappear. The procedure is essentially the same as with the Smart Print. This machine was used for the fabrication of the qubit used in Chap. 6, which is nominally the same qubit as the tomography qubit of Chap. 4.

C.1.4 *Development*

The development program is as follows:

- development in MF319 for 50 s ;
- rinsing in DI water for 20 s ;
- dry with a nitrogen gun.

The MF319 is stored in a fridge at 4°C. It is important to keep the MF319 as cold as possible, as it slows down the reaction. The development is then less sensitive to the development time, which makes it more stable.

C.1.5 *Etching*

The etching program is as follows:

- Transene Tantalum etchant 111 at room temperature for 17 s ;
- DI water for 60 s ;
- dry with a nitrogen gun.

The result of the lithography is very sensitive to the duration of the etching step. It is regularly updated, as external parameters like the temperature can significantly change the optimal time. If we stop the etching too soon, it does not have the time to remove perfectly the tantalum in the exposed regions. If we stop it too late, it will start to etch unwanted regions, as the resist is also etched, although at a slower rate.

C.1.6 *Dicing*

The wafer is then cleaned in acetone and coated again in resist. It is finally diced in the Nanolyon clean room. The samples are now ready for electronic lithography.

C.2 ELECTRONIC LITHOGRAPHY

The Josephson junction (JJ) has a typical size of 200×200 nm^2. To properly fabricate it, we need a precision of 10s of nanometers, too small to be handled by optical lithography, which uses light at 365 nm wavelength. The exposure is then done in a Scanning Electron Microscope (SEM) using electronic beam (e-beam) resists, which is much more precise and allows to reach the desired precision. The lithography procedure is summarized in Fig. C.2. The junctions are fabricated using a Dolan bridge technique [254] to deposit two layers of aluminum (superconducting) separated by an aluminum oxide (insulating) layer. The sample is coated with a double layer of polymethylglutarimide (PMGI) and Polymethyl methacrylate (PMMA). The latter is a positive e-beam resist, which means that the exposed regions are weakened and will be dissolved by the developer. After exposure and development of the PMMA layer, the PMGI becomes accessible in these regions. Immersing the sample in a second developer allows us to dissolve the PMGI in these regions and to create an *undercut*: the PMGI is removed under the PMMA in the vicinity of these regions, which crucially creates a bridge. Aluminum is then deposited in two steps, following two different angles, separated by an oxidation step, which creates the junction in AlOx. Its area is fixed by

the two widths W_1 and W_2, and the thickness by the oxygen pressure and the duration of the oxidation step. We will detail the recipe used in the fabrication of the junctions.

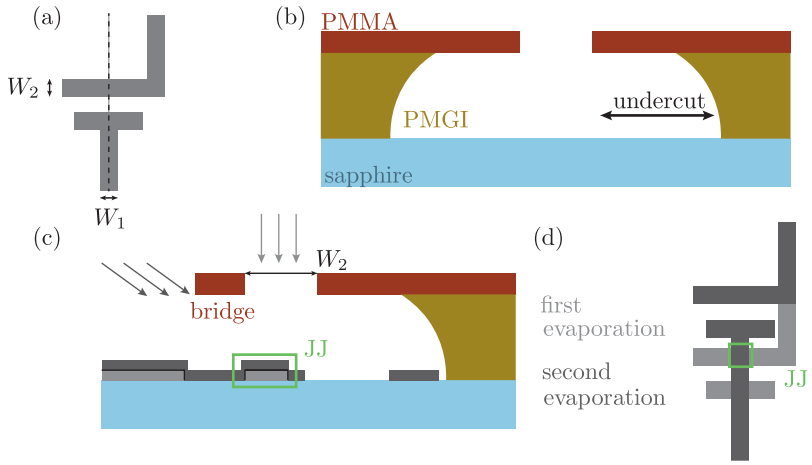


Figure C.2: Josephson junction fabrication procedure. (a) Shape of the pattern exposed during the lithography. The width W_1 and W_2 sets the area of the JJ. (b) Scheme of the layers of PMMA and PMGI after exposure. The PMGI is removed on a larger region under the PMMA, creating the undercut. (c) Aluminum deposition. A first evaporation is done with an angle of 0° (light gray) and a second with an angle of 30° (dark gray, the angle has been exaggerated on the figure) after an oxidation step. The oxidized zone is materialized as a solid black line. The zone containing the JJ is highlighted by the green rectangle. (d) Top view of the two evaporations. The JJ is created in the overlap region.

C.2.1 *Cleaning*

The cleaning part of the lithography process is the part that evolved the most in the past three years. It started from the recipe given in [103], using toluene, acetone, methanol, IPA and a stabilized version of the piranha solution (a mixture of sulfuric acid and hydrogen peroxide) called Purestrip. Using this protocol during the first year, we finally abandoned it, as the Purestrip seemed to deteriorate the samples (see

Fig. C.3). It started to be visible after a few cycles of fabrication performed on the same sample, but we highly suspect that it already deteriorated the quality of the sample at the first cycle. Only after we abandoned this step could we obtain a decay time of the cavity $T_c > 200$ μs with a sapphire sample inserted, whereas it was stuck to a few tens of μs before.

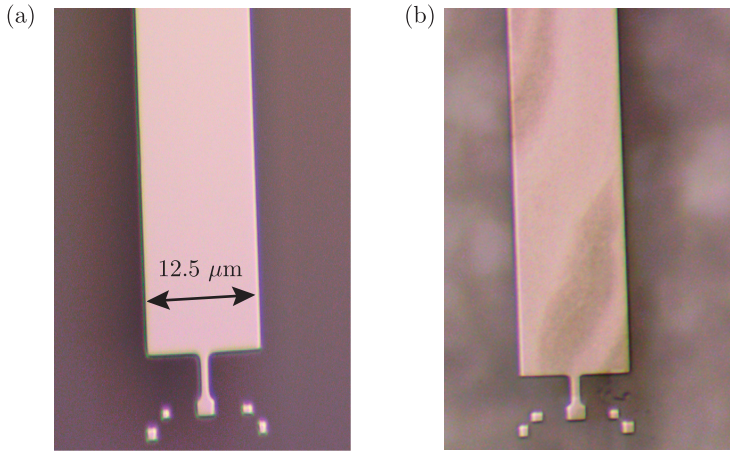


Figure C.3: Pictures of the influence of the Purestrip on the cleaning procedure. Clean (a) versus dirty (b) qubit capacitor.

The use of Buffered Oxide Etch (BOE), introduced in the group around one year ago, further improved the quality factor of the qubit: whereas the T_1 of the tomography qubit is around 3.6 μs in Chap. 4 with the old cleaning process, it reaches 38 μs in Chap. 6, with the same design. This is not the only factor, as the shielding of the experiment also improved during this time, but it participated a lot. The recipe presented here is the last one, followed for the qubit of Chap. 6.

If a lithography was performed before on the sample, the sample is first immersed in a potassium hydroxide solution with 2 % mass concentration for a few minutes to remove the aluminum. The following steps are as follows:

- NMP at 60°C for 30 minute ;
- rinse in IPA for 20 s ;
- dry with nitrogen gun ;

- BOE (7:1) for 15 min ;
- DI water for 60 s ;
- dry with a nitrogen gun.

The BOE ratio of 7:1 sets the ratio between the two water solutions it is made of; one constituted of 40 % of ammonium fluoride and the other of 49 % hydrofluoric acid. It is important to lower as much as possible the time between the BOE and the cooldown of the qubit to limit the size of the reconstituted oxide layer after this step.

C.2.2 *Coating*

The coating step starts with a minute on the hot plate (at least 115°C) to remove any residual humidity on the sample. Two steps are then followed. First, the PMGI:

- 500 rpm for 5 s, acceleration 500 rpm/s, then 2000 rpm for 55 s, acceleration 4000 rpm/s ;
- bake 5 min at 200°C.

We then let the sample cool down for at least a minute before the second step with the PMMA:

- 500 rpm for 5 s, acceleration 500 rpm/s, then 4000 rpm for 55 s, acceleration 4000 rpm/s ;
- bake 15 min at 180°C.

C.2.3 *Exposure*

The exposure step is done in a ZEISS Supra 55 VP SEM after the deposition of 10 nm of aluminum to evacuate the electrons. The settings are:

- voltage 30 keV ;
- working distance around 7 mm ;
- aperture 7.5 μm (current around 12 pA) ;
- 380 μC cm^{-2}.

C.2.4 *Development*

The development steps are listed below:

- aluminum removal: KOH solution (2 % mass concentration) for 1 min ;
- deionized (DI) water for 20 s
- dry with nitrogen gun ;
- PMMA development: stir in a MIBK-IPA (1:3 in volume) solution for 1 minute ;
- rinse in IPA beaker for 20 s ;
- dry with nitrogen gun ;
- inspect the sample with the microscope ;
- PMGI development: stir in MF319 taken straight out of the fridge for 35 s ;
- rinse in DI water beaker for 20 s
- dry with nitrogen gun ;

C.2.5 *Evaporation*

The evaporation steps are the same as in [103]. Using a Plassys evaporator:

- insert the sample in the evaporator ;
- pump the chamber at least overnight so that it reaches a pressure below 1×10^{-7} mbar ;
- ion milling for 30 s (400 V, 22 mA, Ar, 0° tilt) ;
- ion milling for 30 s (400 V, 22 mA, Ar, 30° tilt) ;
- titanium evaporation at a rate of 0.2 nm/s for 2 min, while the sample is protected by a shutter to lower the oxygen pressure ;

- 20 nm aluminum evaporation at rate 0.5 nm/s at 0° tilt ;
- static oxidation for 25 min with pressure between 1 and 5 mbar ;
- 40 nm aluminum evaporation at rate 0.5 nm/s at 30° tilt ;
- static oxidation for 10 min at 10 mbar.

C.2.6 *Lift off*

The lift-off is performed in NMP at 60°C for 30 min. After 30 min, the sonication is turned on for a few minutes until the aluminum on top of the resist is completely removed. The sample is then quickly moved to another NMP beaker to get rid of the aluminum bits that remain. It is then rinsed in IPA for 20 s and dried with nitrogen.

C.2.7 *Junction resistance*

After the fabrication process, we measure the resistance of the junction. The normal-state resistance R(0) at 0 K is linked to the critical current of the junction through the Ambegaokar–Baratoff formula [255]:

$$R(0) = \frac{\pi\Delta}{2eI_0}, \tag{C.1}$$

where $\Delta \simeq 180$ μeV is the superconducting gap of the aluminum. As a function of the Josephson energy E_J, it reads:

$$R(0) = \frac{\pi\Delta\varphi_0}{2eE_J}. \tag{C.2}$$

The resistance measurement is performed at room temperature, which yields a resistance $R(300)$ at 300 K that is typically 10 to 20 % smaller than $R(0)$. To adjust the frequency of the qubit, we can either play on the area W_1W_2 or on the oxidation pressure P and time t. We indeed have

$$R(0) \propto \frac{\sqrt{Pt}}{W_1W_2}. \tag{C.3}$$

The relationship $R(0) \propto \sqrt{Pt}$ is considered less accurate than $R(0) \propto 1/W_1W_2$. In our case, we fabricated the samples in batches of 2 to

4. A good practice is then to make the main adjustments with the junction geometry and to evaporate the samples one by one or two by two to be able to tweak the oxidation parameters if the first sample resistances are off.

During the development of this fabrication process by Antoine Essig, a lot of junctions were destroyed by electrostatic discharges (see Appendix A in [103]). The solution chosen consists in fabricating a wire that connects the two pads of the qubit capacitor. This line is cut using the tip of a probe at the probe station under an ionized airflow. The airflow is maintained until the qubit is inserted into the cavity.

C.3 ALUMINUM CAVITY

The 3D cavity is machined in a block of pure aluminum at 99.99 %. The cleaning step is crucial to reach a large bare Q factor. After receiving the cavity, the following steps are followed:

- wash with dishwashing liquid ;
- rinse in DI water ;
- acetone in sonicator for 10 min ;
- IPA in sonicator for 10 min ;
- dry with nitrogen gun ;
- type A aluminum etchant at 50°C for 45 min ;
- rinse in DI water ;
- acetone in sonicator for 5 min ;
- rinse in IPA ;
- dry with nitrogen gun ;
- bake 30 min at 50°C ;

The most important part is the etching part, which removes a few nanometers of contaminated aluminum. Note that after several cooldowns, the Q factor of the cavity decreases. This procedure allowed

to reach a decay time T_1 of 1 photon in the cavity of around 1 ms at around 4.5 GHz, which corresponds to a quality factor above 4×10^6. This was obtained for the cavity "bare", *i.e.* without anything else coupled to it.

Part VI

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Werner Heisenberg. "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik." In: *Zeitschrift für Physik* 43 (1927), pp. 172–198 (Cited on page 3).

[2] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Advances in quantum metrology." In: *Nature photonics* 5 (2011), pp. 222–229 (Cited on page 9).

[3] Navin Khaneja et al. "Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms." In: *Journal of magnetic resonance (San Diego, Calif. : 1997)* 172 (Feb. 2005), pp. 296–305 (Cited on pages 10, 197, 207, 208, 238, 239).

[4] Michael Goerz et al. "Krotov: A Python implementation of Krotov's method for quantum optimal control." In: *SciPost Physics* 7 (2019), p. 080 (Cited on pages 10, 207, 208, 238, 239).

[5] V. V. Sivak et al. "Model-Free Quantum Control with Reinforcement Learning." In: *Physical Review X* 12 (2022), p. 011059. arXiv: 2104.14539 (Cited on pages 10, 197, 199, 205, 232).

[6] H. Hutin et al. "Monitoring the Energy of a Cavity by Observing the Emission of a Repeatedly Excited Qubit." In: *Phys. Rev. Lett.* 133 (2024), p. 153602 (Cited on pages 12, 105, 106).

[7] Hector Hutin et al. "Preparing Schrödinger Cat States in a Microwave Cavity Using a Neural Network." In: *PRX Quantum* 6 (2025), p. 010321 (Cited on pages 12, 197).

[8] Serge Haroche. "Nobel Lecture: Controlling photons in a box and exploring the quantum to classical boundary." In: *Reviews of Modern Physics* 85 (2013), p. 1083 (Cited on page 15).

[9] A Wallraff et al. "Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics." In: *Nature* 431 (2004), pp. 162–167 (Cited on page 15).

[10] R. Assouly et al. "Quantum advantage in microwave quantum radar." In: *Nature Physics* 19 (2023), pp. 1418–1422 (Cited on pages 15, 162, 251).

[11] C. Eichler, D. Bozyigit, and A. Wallraff. "Characterizing quantum microwave radiation and its entanglement with superconducting qubits using linear detectors." In: *Physical Review A* 86 (Jan. 2012), p. 032106 (Cited on page 15).

[12] Frank Arute et al. "Quantum supremacy using a programmable superconducting processor." In: *Nature* 574 (Oct. 2019), pp. 505–510 (Cited on page 15).

[13] Nathanaël Cottet and Benjamin Huard. "Maxwell's Demon in Superconducting Circuits." In: *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*. Ed. by Felix Binder et al. Cham: Springer International Publishing, 2018, pp. 959–981 (Cited on page 15).

[14] Uri Vool and Michel Devoret. "Introduction to quantum electromagnetic circuits." In: *International Journal of Circuit Theory and Applications* 45 (2017), pp. 897–934. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cta.2359 (Cited on page 15).

[15] Peter Groszkowski and Jens Koch. "Scqubits: a Python package for superconducting qubits." In: *Quantum* 5 (2021), p. 583 (Cited on page 20).

[16] V Bouchiat et al. "Quantum coherence with a single Cooper pair." In: *Physica Scripta* T76 (Jan. 1998), pp. 165–170 (Cited on page 21).

[17] Y. Nakamura, Yu. A. Pashkin, and J. S. Tsai. "Coherent control of macroscopic quantum states in a single-Cooper-pair box." In: *Nature* 398 (Apr. 1999), pp. 786–788 (Cited on page 21).

[18] D Vion et al. "Manipulating the quantum state of an electrical circuit." In: *Science (New York, N.Y.)* 296 (May 2002), pp. 886–9 (Cited on page 22).

[19] G Ithier et al. "Decoherence in a superconducting quantum bit circuit." In: *Physical Review B* 72 (Jan. 2005), p. 134519 (Cited on page 22).

[20] Jens Koch et al. "Charge-insensitive qubit design derived from the Cooper pair box." In: *Physical Review A—Atomic, Molecular, and Optical Physics* 76 (2007), p. 042319 (Cited on page 22).

[21] P. Krantz et al. "A quantum engineer's guide to superconducting qubits." In: *Applied Physics Reviews* 6 (June 2019), p. 021318 (Cited on page 22).

[22] Chenlu Wang et al. "Towards practical quantum computers: transmon qubit with a lifetime approaching 0.5 milliseconds." In: *npj Quantum Information* 8 (2022), p. 3 (Cited on page 22).

[23] Aaron Somoroff et al. "Millisecond coherence in a superconducting qubit." In: *Physical Review Letters* 130 (2023), p. 267001 (Cited on page 22).

[24] Daoquan Zhu et al. "Quantum computing with superconducting circuits in the picosecond regime." In: *Physical Review Applied* 16 (2021), p. 014024 (Cited on page 23).

[25] Roy J Glauber. "Coherent and incoherent states of the radiation field." In: *Physical Review* 131 (1963), p. 2766 (Cited on page 29).

[26] Maggie Tse et al. "Quantum-enhanced advanced LIGO detectors in the era of gravitational-wave astronomy." In: *Physical Review Letters* 123 (2019), p. 231107 (Cited on pages 29, 35).

[27] Fausto Acernese et al. "Increasing the astrophysical reach of the advanced virgo detector via the application of squeezed vacuum states of light." In: *Physical review letters* 123 (2019), p. 231108 (Cited on pages 29, 35).

[28] Stephen Barnett and Paul M Radmore. *Methods in theoretical quantum optics*. Vol. 15. Oxford University Press, 2002 (Cited on pages 29, 257).

[29] George E. C. Sudarshan. "Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams." In: *Physical Review Letters* 10 (1963), p. 277 (Cited on page 31).

[30] Marlan O Scully and M Suhail Zubairy. *Quantum optics*. Cambridge university press, 1997 (Cited on pages 31, 32, 38).

[31] Andrea Mari and Jens Eisert. "Positive Wigner functions render classical simulation of quantum computation efficient." In: *Physical review letters* 109 (2012), p. 230503 (Cited on page 32).

[32] J.R. Johansson, P.D. Nation, and Franco Nori. "QuTiP: An open-source Python framework for the dynamics of open quantum systems." In: *Computer Physics Communications* 183 (2012), pp. 1760–1772 (Cited on pages 34, 203).

[33] Marco Barbieri. "Optical quantum metrology." In: *PRX Quantum* 3 (2022), p. 010202 (Cited on page 35).

[34] Wolfgang P Schleich. *Quantum optics in phase space*. John Wiley & Sons, 2011 (Cited on page 35).

[35] S Haroche and J Raimond. "Exploring the Quantum: Atoms, Cavities, and Photons." In: *Oxford Graduated Text* (Jan. 2006) (Cited on page 35).

[36] N Lütkenhaus and Stephen M Barnett. "Nonclassical effects in phase space." In: *Physical Review A* 51 (1995), p. 3340 (Cited on page 36).

[37] Ulysse Chabaud, Damian Markham, and Frédéric Grosshans. "Stellar representation of non-Gaussian quantum states." In: *Physical Review Letters* 124 (2020), p. 063605 (Cited on page 36).

[38] Ulysse Chabaud. *Continuous Variable Quantum Advantages and Applications in Quantum Optics*. 2021. arXiv: 2102.05227 [quant-ph] (Cited on page 38).

[39] Gerhard Kirchmair et al. "Observation of quantum state collapse and revival due to the single-photon Kerr effect." In: *Nature* 495 (Mar. 2013), pp. 205–209 (Cited on page 40).

[40] A. A. Clerk et al. "Introduction to quantum noise, measurement, and amplification." In: *Reviews of Modern Physics* 82 (Apr. 2010), pp. 1155–1208 (Cited on pages 43, 47, 49, 55, 86, 87, 94, 115, 125, 140).

[41] Baleegh Abdo et al. "Full Coherent Frequency Conversion between Two Propagating Microwave Modes." In: *Physical Review Letters* 110 (Apr. 2013), p. 173902 (Cited on page 43).

[42] Emmanuel Flurin. "The Josephson Mixer, a Swiss army knife for microwave quantum optics." PhD thesis. Ecole Normale Superieure, Paris, 2014 (Cited on pages 43, 56, 74).

[43] C. Fabre and N. Treps. "Modes and states in quantum optics." In: *Rev. Mod. Phys.* 92 (2020), p. 035005 (Cited on pages 43, 59, 257).

[44] Stéphane Virally and Bertrand Reulet. "Unidimensional time-domain quantum optics." In: *Physical Review A* 100 (2019), p. 023833 (Cited on page 46).

[45] Evangelos Varvelis, Debjyoti Biswas, and David P DiVincenzo. "The photonic content of a transmission-line pulse." In: *Proceedings of the National Academy of Sciences* 121 (2024), e2314846121 (Cited on page 46).

[46] A. I. Akhiezer, V. B. Berestetskii, and R. A. Shaffer. "Quantum Electrodynamics." In: *American Journal of Physics* 33 (Nov. 1965), pp. 976–976 (Cited on page 46).

[47] MV Fedorov et al. "Spontaneous emission of a photon: Wave-packet structures and atom-photon entanglement." In: *Physical Review A—Atomic, Molecular, and Optical Physics* 72 (2005), p. 032110 (Cited on pages 48, 52).

[48] Mankei Tsang, Ranjith Nair, and Xiao-Ming Lu. "Quantum theory of superresolution for two incoherent optical point sources." In: *Physical Review X* 6 (2016), p. 031033 (Cited on pages 50, 161).

[49] Clémentine Rouvière et al. "Ultra-sensitive separation estimation of optical sources." In: *Optica* 11 (2024), pp. 166–170 (Cited on pages 50, 161).

[50] Ilya Karuseichyk et al. "Exploiting separation-dependent coherence to boost optical resolution." In: *Physical Review A* 109 (2024), p. 043524 (Cited on pages 50, 161).

[51] Benjamin Brecht et al. "Photon temporal modes: a complete framework for quantum information science." In: *Physical Review X* 5 (2015), p. 041017 (Cited on pages 50, 161).

[52] Laura Serino et al. "Realization of a multi-output quantum pulse gate for decoding high-dimensional temporal modes of single-photon states." In: *PRX quantum* 4 (2023), p. 020306 (Cited on pages 50, 161).

[53] Maryam Khanahmadi et al. "Multimode character of quantum states released from a superconducting cavity." In: *Physical Review Research* 5 (2023), p. 043071 (Cited on pages 50, 58).

[54] Michael G Raymer and Ian A Walmsley. "Temporal modes in quantum optics: then and now." In: *Physica Scripta* 95 (2020), p. 064002 (Cited on page 50).

[55] Eloi Descamps et al. "Quantum metrology using time-frequency as quantum continuous variables: Resources, sub-shot-noise precision and phase space representation." In: *Physical Review Letters* 131 (2023), p. 030801 (Cited on pages 52, 53, 112, 177, 178).

[56] Dario Ferraro et al. "Wigner function approach to single electron coherence in quantum Hall edge channels." In: *Physical Review B—Condensed Matter and Materials Physics* 88 (2013), p. 205303 (Cited on page 53).

[57] Wolfgang Pfaff et al. "Controlled release of multiphoton quantum states from a microwave cavity memory." In: *Nature Physics* 13 (June 2017), pp. 882–887 (Cited on page 57).

[58] Yi Yin et al. "Catch and Release of Microwave Photon States." In: *Physical Review Letters* 110 (Mar. 2013), p. 107001 (Cited on page 57).

[59] R. Dassonneville et al. "Number-Resolved Photocounter for Propagating Microwave Mode." In: *Physical Review Applied* 14 (Oct. 2020), p. 044022 (Cited on page 57).

[60] V Ranjan et al. "Multimode storage of quantum microwave fields in electron spins over 100 ms." In: *Physical Review Letters* 125 (2020), p. 210505 (Cited on page 58).

[61] Brian Julsgaard et al. "Quantum memory for microwave photons in an inhomogeneously broadened spin ensemble." In: *Physical review letters* 110 (2013), p. 250503 (Cited on page 58).

[62] James O'Sullivan et al. "Random-access quantum memory using chirped pulse phase encoding." In: *Physical Review X* 12 (2022), p. 041014 (Cited on page 58).

[63] J. Cirac et al. "Quantum State Transfer and Entanglement Distribution among Distant Nodes in a Quantum Network." In: *Physical Review Letters* 78 (Apr. 1997), pp. 3221–3224 (Cited on page 58).

[64] H. J. Kimble. "The quantum internet." In: *Nature* 453.7198 (2008), pp. 1023–1030 (Cited on page 58).

[65] Luke Masters et al. "On the simultaneous scattering of two photons by a single two-level atom." In: *Nature Photonics* 17 (2023), pp. 972–976 (Cited on page 58).

[66] B. Mollow. "Power Spectrum of Light Scattered by Two-Level Systems." In: *Physical Review* 188 (Dec. 1969), pp. 1969–1975 (Cited on page 58).

[67] Jan Sperling et al. "Mode-independent quantum entanglement for light." In: *Physical Review A* 100 (2019), p. 062129 (Cited on page 59).

[68] Ulysse Chabaud and Mattia Walschaers. "Resources for bosonic quantum computational advantage." In: *Physical Review Letters* 130 (2023), p. 090602 (Cited on page 59).

[69] Ali S Üstünel. *An introduction to analysis on Wiener space*. Springer, 2006 (Cited on page 62).

[70] Stéphane Mallat. *A wavelet tour of signal processing*. Elsevier, 1999 (Cited on page 63).

[71] Olivier Morin, Claude Fabre, and Julien Laurat. "Experimentally accessing the optimal temporal mode of traveling quantum light states." In: *Physical review letters* 111 (2013), p. 213602 (Cited on page 63).

[72] H M Wiseman. "SU(2) distribution functions and measurement of the fluorescence of a two-level atom." In: *Quantum and Semiclassical Optics: Journal of the European Optical Society Part B* 7 (Aug. 1995), p. 569 (Cited on pages 64, 86).

[73] Carlton M. Caves et al. "Quantum limits on phase-preserving linear amplifiers." In: *Physical Review A* 86 (Dec. 2012), p. 063802 (Cited on page 65).

[74] Robert H Hadfield. "Single-photon detectors for optical quantum information applications." In: *Nature photonics* 3 (2009), pp. 696–705 (Cited on page 71).

[75] Alexander I Lvovsky et al. "Quantum state reconstruction of the single-photon Fock state." In: *Physical Review Letters* 87 (2001), p. 050402 (Cited on page 71).

[76] Alexei Ourjoumtsev et al. "Generating Optical Schrödinger Kittens for Quantum Information Processing." In: *Science* 312 (2006), pp. 83–86 (Cited on page 71).

[77] Alessandro Zavatta, Valentina Parigi, and Marco Bellini. "Experimental nonclassicality of single-photon-added thermal light states." In: *Physical Review A—Atomic, Molecular, and Optical Physics* 75 (2007), p. 052106 (Cited on page 71).

[78] Pieter Kok et al. "Linear optical quantum computing with photonic qubits." In: *Reviews of modern physics* 79 (2007), pp. 135–174 (Cited on page 71).

[79] Robert Raussendorf and Hans J Briegel. "A one-way quantum computer." In: *Physical review letters* 86 (2001), p. 5188 (Cited on page 71).

[80] Alain Aspect, Jean Dalibard, and Gérard Roger. "Experimental Test of Bell's Inequalities Using Time- Varying Analyzers." In: *Physical Review Letters* 49 (Dec. 1982), pp. 1804–1807 (Cited on page 71).

[81] Valentina Parigi et al. "Probing quantum commutation rules by addition and subtraction of single photons to/from a light field." In: *Science* 317 (2007), pp. 1890–1893 (Cited on page 71).

[82] Raphaël Lescanne et al. "Irreversible Qubit-Photon Coupling for the Detection of Itinerant Microwave Photons." In: *Physical Review X* 10 (May 2020), p. 021038 (Cited on pages 72, 87).

[83] Léo Balembois et al. *Cyclically operated Single Microwave Photon Counter with* 10^{-22} $\mathrm{W}/\sqrt{\mathrm{Hz}}$ *sensitivity*. 2024. arXiv: 2307.03614 [quant-ph] (Cited on page 72).

[84] Zhiren Wang et al. "Single-electron spin resonance detection by microwave photon counting." In: *Nature* 619 (2023), pp. 276–281 (Cited on page 72).

[85] J. Travesedo et al. *All-microwave spectroscopy and polarization of individual nuclear spins in a solid*. 2024. arXiv: 2408.14282 [quant-ph] (Cited on page 72).

[86] Christopher C Gerry and Peter L Knight. *Introductory quantum optics*. Cambridge university press, 2023 (Cited on page 78).

[87] E. Flurin et al. "Generating Entangled Microwave Radiation Over Two Transmission Lines." In: *Physical Review Letters* 109 (Oct. 2012), p. 183901 (Cited on page 80).

[88] A. Eddins et al. "High-Efficiency Measurement of an Artificial Atom Embedded in a Parametric Amplifier." In: *Physical Review X* 9 (2019) (Cited on pages 86, 165).

[89] Florent Lecocq et al. "Efficient qubit measurement with a nonreciprocal microwave amplifier." In: *Physical Review Letters* 126 (2021), p. 020502 (Cited on pages 86, 165).

[90] C Macklin et al. "A near – quantum-limited Josephson traveling-wave parametric amplifier." In: *Science* 350 (2015), p. 307 (Cited on pages 86, 109, 117, 165, 221).

[91] M Renger et al. "Beyond the standard quantum limit for parametric amplification of broadband signals." In: *npj Quantum Information* 7 (2021), p. 160 (Cited on pages 86, 165).

[92] A Clerk, Girvin, and A Stone. "Quantum-limited measurement and information in mesoscopic detectors." In: *Physical Review B* 67 (Apr. 2003), p. 165324 (Cited on pages 87, 115).

[93] Kater W Murch, Rajamani Vijay, and Irfan Siddiqi. "Weak measurement and feedback in superconducting quantum circuits." In: *Superconducting Devices in Quantum Optics* (2016), pp. 163–185 (Cited on pages 87, 100).

[94] Morten Kjaergaard et al. "Superconducting qubits: Current state of play." In: *Annual Review of Condensed Matter Physics* 11 (2020), pp. 369–395 (Cited on page 92).

[95] Joachim Cohen et al. "Reminiscence of Classical Chaos in Driven Transmons." In: *PRX Quantum* 4 (2023), p. 020312 (Cited on page 92).

[96] Marie Frédérique Dumas et al. "Measurement-Induced Transmon Ionization." In: *Phys. Rev. X* 14 (2024), p. 041023 (Cited on pages 92, 204).

[97] Theodore Walter et al. "Rapid high-fidelity single-shot dispersive readout of superconducting qubits." In: *Physical Review Applied* 7 (2017), p. 054020 (Cited on page 92).

[98] S. Touzard et al. "Gated Conditional Displacement Readout of Superconducting Qubits." In: *Physical Review Letters* 122 (Feb. 2019), p. 080502 (Cited on page 92).

[99] Y. Sunada et al. "Fast Readout and Reset of a Superconducting Qubit Coupled to a Resonator with an Intrinsic Purcell Filter." In: *Physical Review Applied* 17 (Apr. 2022), p. 044016 (Cited on page 92).

[100] François Swiadek et al. *Enhancing Dispersive Readout of Superconducting Qubits Through Dynamic Control of the Dispersive Shift: Experiment and Theory*. 2023. arXiv: 2307.07765 [quant-ph] (Cited on page 92).

[101] Cristóbal Lledó et al. "Cloaking a qubit in a cavity." In: *Nature Communications* 14 (2023), p. 6313 (Cited on page 92).

[102] Jay Gambetta et al. "Quantum trajectory approach to circuit QED: Quantum jumps and the Zeno effect." In: *Physical Review A* 77 (Jan. 2008), p. 012112 (Cited on page 92).

[103] Antoine Essig. "Photon counting with a multiplexed dispersive readout." PhD thesis. Ecole Normale Superieure de Lyon, 2021 (Cited on pages 96, 117, 121, 219, 263, 266, 268).

[104] Rui Han, Gerd Leuchs, and Markus Grassl. "Residual and Destroyed Accessible Information after Measurements." In: *Physical Review Letters* 120 (Apr. 2018), p. 160501 (Cited on pages 98, 249).

[105] Ingrid Strandberg et al. "Digital Homodyne and Heterodyne Detection for Stationary Bosonic Modes." In: *Phys. Rev. Lett.* 133 (2024), p. 063601 (Cited on page 98).

[106] Kurt Jacobs and Daniel A Steck. "A straightforward introduction to continuous quantum measurement." In: *Contemporary Physics* 47 (2006), pp. 279–303 (Cited on page 100).

[107] K W Murch et al. "Observing single quantum trajectories of a superconducting quantum bit." In: *Nature* 502 (Oct. 2013), pp. 211–214 (Cited on pages 100, 105).

[108] Steven J. Weber et al. "Quantum trajectories of superconducting qubits." In: *Comptes Rendus Physique* 17 (2016), pp. 766–777 (Cited on pages 100, 105).

[109] M Hatridge et al. "Quantum Back-Action of an Individual Variable-Strength Measurement." In: *Science* 339 (Jan. 2013), p. 178 (Cited on pages 100, 105).

[110] G. de Lange et al. "Reversing Quantum Trajectories with Analog Feedback." In: *Physical Review Letters* 112 (Feb. 2014), p. 080501 (Cited on pages 100, 105).

[111] Shay Hacohen-Gourgy et al. "Quantum dynamics of simultaneously measured non-commuting observables." In: *Nature* 538 (Oct. 2016), pp. 491–494 (Cited on pages 100, 105).

[112] Quentin Ficheux et al. "Dynamics of a qubit while simultaneously monitoring its relaxation and dephasing." In: *Nature Communications* 9 (Dec. 2018), p. 1926 (Cited on pages 100, 105).

[113] U. Vool et al. "Continuous quantum nondemolition measurement of the transverse component of a qubit." In: *Physical Review Letters* 117 (2016), p. 133601 (Cited on page 105).

[114] Philippe Campagne-Ibarcq et al. "Observing Quantum State Diffusion by Heterodyne Detection of Fluorescence." In: *Phys. Rev. X* 6 (Jan. 2016), p. 11002 (Cited on page 105).

[115] Philippe Campagne-Ibarcq et al. "Using Spontaneous Emission of a Qubit as a Resource for Feedback Control." In: *Physical Review Letters* 117 (Aug. 2016), p. 060502 (Cited on page 105).

[116] M Naghiloo et al. "Mapping quantum state dynamics in spontaneous emission." In: *Nature Communications* 7 (May 2016), p. 11527 (Cited on page 105).

[117] Warren Nagourney, Jon Sandberg, and Hans Dehmelt. "Shelved optical electron amplifier: Observation of quantum jumps." In: *Physical Review Letters* 56 (1986), pp. 2797–2799 (Cited on page 105).

[118] T Sauter et al. "Observation of quantum jumps." In: *Physical Review Letters* 57 (1986), pp. 1696–1698 (Cited on page 105).

[119] J. Bergquist et al. "Observation of Quantum Jumps in a Single Atom." In: *Physical Review Letters* 57 (1986), pp. 1699–1702 (Cited on page 105).

[120] Vijay, Slichter, and Siddiqi. "Observation of Quantum Jumps in a Superconducting Artificial Atom." In: *Physical Review Letters* 106 (Jan. 2011), p. 110502 (Cited on page 105).

[121] Z. K. Minev et al. "To catch and reverse a quantum jump mid-flight." In: *Nature* 570 (2019), pp. 200–204 (Cited on page 105).

[122] Sébastien Gleyzes et al. "Quantum jumps of light recording the birth and death of a photon in a cavity." In: *Nature* 446 (Mar. 2007), pp. 297–300 (Cited on pages 105, 107).

[123] L Sun et al. "Tracking photon jumps with repeated quantum non-demolition parity measurements." In: *Nature* 511 (July 2014), pp. 444–448 (Cited on page 105).

[124] Christine Guerlin et al. "Progressive field-state collapse and quantum non-demolition photon counting." In: *Nature* 448 (Aug. 2007), pp. 889–93 (Cited on pages 105, 107).

[125] C Sayrin et al. "Real-time quantum feedback prepares and stabilizes photon number states." In: *Nature* 477 (Sept. 2011), pp. 73–77 (Cited on pages 105, 145).

[126] Antoine Essig et al. "Multiplexed Photon Number Measurement." In: *Physical Review X* 11 (Aug. 2021), p. 031045 (Cited on pages 105–107, 109, 111, 119, 120, 122, 137, 139, 153, 160, 174, 175).

[127] Hector Hutin. *Photon number trajectories for "Monitoring the energy of a cavity by observing the emission of a repeatedly excited qubit"*. June 2024 (Cited on pages 106, 113).

[128] Hector Hutin et al. *Supplementary Material for "Monitoring the energy of a cavity by observing the emission of a repeatedly excited qubit"*. 2024. arXiv: 2402.05046 [quant-ph] (Cited on pages 106, 109, 111–115).

[129] E. Knill, R. Laflamme, and G. J. Milburn. "A scheme for efficient quantum computation with linear optics." In: *Nature* 409 (Jan. 2001), pp. 46–52 (Cited on page 107).

[130] Christoph Simon et al. "Quantum Repeaters with Photon Pair Sources and Multimode Memories." In: *Physical Review Letters* 98 (May 2007), p. 190503 (Cited on page 107).

[131] Scott Aaronson and Alex Arkhipov. "The computational complexity of linear optics." In: *Optics InfoBase Conference Papers* 9 (2014), pp. 143–252 (Cited on page 107).

[132] Craig S. Hamilton et al. "Gaussian Boson Sampling." In: *Physical Review Letters* 119 (Oct. 2017), p. 170501 (Cited on page 107).

[133] DI Schuster et al. "Resolving photon number states in a superconducting circuit." In: *Nature* 445 (2007), pp. 515–518 (Cited on page 107).

[134] B R Johnson et al. "Quantum non-demolition detection of single microwave photons in a circuit." In: *Nature Physics* 6 (Sept. 2010), pp. 663–667 (Cited on page 107).

[135] B. Peaudecerf et al. "Quantum feedback experiments stabilizing Fock states of light in a cavity." In: *Physical Review A* 87 (Apr. 2013), p. 042320 (Cited on page 107).

[136] B. Peaudecerf et al. "Adaptive Quantum Nondemolition Measurement of a Photon Number." In: *Physical Review Letters* 112 (Feb. 2014), p. 080401 (Cited on page 107).

[137] Christopher S. Wang et al. "Efficient Multiphoton Sampling of Molecular Vibronic Spectra on a Superconducting Bosonic Processor." In: *Physical Review X* 10 (June 2020), p. 021060 (Cited on page 107).

[138] R. Dassonneville et al. "Fast High-Fidelity Quantum Nondemolition Qubit Readout via a Nonperturbative Cross-Kerr Coupling." In: *Physical Review X* 10 (Feb. 2020), p. 011045 (Cited on page 107).

[139] Jacob C. Curtis et al. "Single-shot number-resolved detection of microwave photons with error mitigation." In: *Physical Review A* 103 (Feb. 2021), p. 023705 (Cited on page 107).

[140] Weizhou Cai et al. "Bosonic quantum error correction codes in superconducting quantum circuits." In: *Fundamental Research* 1 (Jan. 2021), pp. 50–67 (Cited on pages 107, 199).

[141] Jaewoo Joo et al. "Logical measurement-based quantum computation in circuit-QED." In: *Scientific Reports* 9 (Nov. 2019), p. 16592 (Cited on page 107).

[142] Matthew Reagor et al. "Quantum memory with millisecond coherence in circuit QED." In: *Physical Review B* 94 (July 2016), p. 014506 (Cited on page 109).

[143] L. Lutterbach and L. Davidovich. "Method for Direct Measurement of the Wigner Function in Cavity QED and Ion Traps." In: *Physical Review Letters* 78 (Mar. 1997), pp. 2547–2550 (Cited on page 109).

[144] P. Bertet et al. "Direct Measurement of the Wigner Function of a One-Photon Fock State in a Cavity." In: *Physical Review Letters* 89 (Oct. 2002), p. 200402 (Cited on pages 109, 204).

[145] Brian Vlastakis et al. "Deterministically Encoding Quantum Information Using 100-Photon Schrodinger Cat States." In: *Science* 342 (Nov. 2013), pp. 607–610 (Cited on pages 109, 225).

[146] C. W. Gardiner and M. J. Collett. "Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation." In: *Physical Review A* 31 (June 1985), pp. 3761–3774 (Cited on page 111).

[147] Mario F. Gely et al. "Observation and stabilization of photonic Fock states in a hot radio-frequency resonator." In: *Science* 363 (Mar. 2019), pp. 1072–1075 (Cited on pages 111, 120).

[148] Jay Gambetta and H M Wiseman. "State and dynamical parameter estimation for open quantum systems." In: *Physical Review A* 64 (Sept. 2001), p. 42105 (Cited on page 112).

[149] Søren Gammelmark and Klaus Mølmer. "Bayesian parameter inference from continuously monitored quantum systems." In: *Phys. Rev. A* 87 (2013), p. 032115 (Cited on page 112).

[150] Søren Gammelmark and Klaus Mølmer. "Fisher Information and the Quantum Cramér-Rao Sensitivity Limit of Continuous Measurements." In: *Phys. Rev. Lett.* 112 (2014), p. 170401 (Cited on page 112).

[151] Catalin Catana, Luc Bouten, and Mădălin Guţă. "Fisher informations and local asymptotic normality for continuous-time quantum Markov processes." In: *Journal of Physics A: Mathematical and Theoretical* 48 (Aug. 2015), p. 365301 (Cited on page 112).

[152] Alexander Holm Kiilerich and Klaus Mølmer. "Bayesian parameter estimation by continuous homodyne detection." In: *Physical Review A* 94 (Sept. 2016), p. 32103 (Cited on page 112).

[153] Francesco Albarelli et al. "Restoring Heisenberg scaling in noisy quantum metrology by monitoring the environment." In: *Quantum* 2 (Dec. 2018), p. 110 (Cited on page 112).

[154] Francesco Albarelli et al. "Continuous Measurements for Advanced Quantum Metrology." In: *Proceedings* 12 (2019) (Cited on page 112).

[155] Theodoros Ilias et al. "Criticality-Enhanced Quantum Sensing via Continuous Measurement." In: *PRX Quantum* 3 (Mar. 2022), p. 010354 (Cited on page 112).

[156] Rui Han, János A Bergou, and Gerd Leuchs. "Near optimal discrimination of binary coherent signals via atom–light interaction." In: *New Journal of Physics* 20 (2018), p. 043005 (Cited on page 116).

[157] Daniel Sank et al. "Measurement-Induced State Transitions in a Superconducting Qubit: Beyond the Rotating Wave Approximation." In: *Phys. Rev. Lett.* 117 (2016), p. 190503 (Cited on page 117).

[158] Raphaël Lescanne et al. "Escape of a Driven Quantum Josephson Circuit into Unconfined States." In: *Phys. Rev. Appl.* 11 (2019), p. 014030 (Cited on page 117).

[159] Alexandru Petrescu, Moein Malekakhlagh, and Hakan E. Türeci. "Lifetime renormalization of driven weakly anharmonic superconducting qubits. II. The readout problem." In: *Phys. Rev. B* 101 (2020), p. 134510 (Cited on page 117).

[160] Ross Shillito et al. "Dynamics of Transmon Ionization." In: *Phys. Rev. Appl.* 18 (Sept. 2022), p. 34031 (Cited on page 117).

[161] Mostafa Khezri et al. "Measurement-induced state transitions in a superconducting qubit: Within the rotating-wave approximation." In: *Physical Review Applied* 20 (Nov. 2023), p. 054008 (Cited on pages 117, 143, 204).

[162] Michiel Burgelman et al. "Structurally Stable Subharmonic Regime of a Driven Quantum Josephson Circuit." In: *Physical Review Applied* 18 (Dec. 2022), p. 064044 (Cited on page 117).

[163] Ted Thorbeck et al. "Readout-Induced Suppression and Enhancement of Superconducting Qubit Lifetimes." In: *Phys. Rev. Lett.* 132 (2024), p. 090602 (Cited on page 117).

[164] Andreas Bengtsson et al. "Model-Based Optimization of Superconducting Qubit Readout." In: *Phys. Rev. Lett.* 132 (2024), p. 100603 (Cited on page 117).

[165] J. Stevens et al. "Energetics of a Single Qubit Gate." In: *Physical Review Letters* 129 (Sept. 2022), p. 110601 (Cited on page 117).

[166] Easwar Magesan et al. "Machine learning for discriminating quantum measurement trajectories and improving readout." In: *Physical review letters* 114 (2015), p. 200501 (Cited on pages 145, 197).

[167] Kevin Reuer et al. "Realizing a deep reinforcement learning agent for real-time quantum feedback." In: *Nature Communications* 14 (2023), p. 7138 (Cited on pages 145, 197, 199).

[168] G Koolstra et al. "Monitoring fast superconducting qubit dynamics using a neural network." In: *Physical Review X* 12 (2022), p. 031017 (Cited on page 145).

[169] E. Flurin et al. "Using a Recurrent Neural Network to Reconstruct Quantum Dynamics of a Superconducting Qubit from Physical Observations." In: *Physical Review X* 10 (Jan. 2020), p. 011006 (Cited on page 145).

[170] Valeria Cimini et al. "Neural networks for detecting multimode wigner negativity." In: *Physical Review Letters* 125 (2020), p. 160504 (Cited on pages 146, 197).

[171] Valentin Gebhart and Martin Bohmann. "Neural-network approach for identifying nonclassicality from click-counting data." In: *Physical Review Research* 2 (2020), p. 023150 (Cited on page 146).

[172] Xiaoting Gao et al. "Correlation-Pattern-Based Continuous Variable Entanglement Detection through Neural Networks." In: *Phys. Rev. Lett.* 132 (2024), p. 220202 (Cited on page 146).

[173] François Chollet et al. *Keras*. `https://keras.io`. 2015 (Cited on page 147).

[174] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015 (Cited on page 147).

[175] Diederik Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization." In: *International Conference on Learning Representations* (Dec. 2015) (Cited on pages 147, 203).

[176] Satvik Maurya et al. "Scaling qubit readout with hardware efficient machine learning architectures." In: *Proceedings of the 50th Annual International Symposium on Computer Architecture*. 2023, pp. 1–13 (Cited on page 151).

[177] Liangyu Chen et al. "Transmon qubit readout fidelity at the threshold for quantum error correction without a quantum-limited amplifier." In: *npj Quantum Information* 9 (2023), p. 26 (Cited on page 151).

[178] Benjamin Lienhard et al. "Deep-Neural-Network Discrimination of Multiplexed Superconducting-Qubit States." In: *Phys. Rev. Appl.* 17 (1 2022), p. 014024 (Cited on page 151).

[179] George Casella and Roger Berger. *Statistical inference*. CRC Press, 2024 (Cited on page 157).

[180] Mile Gu et al. "Observing the operational significance of discord consumption." In: *Nature Physics* 8 (2012), pp. 671–675 (Cited on page 162).

[181] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Quantum-Enhanced Measurements: Beating the Standard Quantum Limit." In: *Science* 306 (2004), pp. 1330–1336 (Cited on page 170).

[182] Manuel Gessner, Nicolas Treps, and Claude Fabre. "Estimation of a parameter encoded in the modal structure of a light beam: a quantum theory." In: *Optica* 10 (2023), pp. 996–999 (Cited on page 177).

[183] ZD Bai and JC Fu. "On the maximum-likelihood estimator for the location parameter of a cauchy distribution." In: *Canadian Journal of Statistics* 15 (1987), pp. 137–146 (Cited on page 181).

[184] Matteo Bottai. "Confidence regions when the Fisher information is zero." In: *Biometrika* 90 (2003), pp. 73–84 (Cited on page 191).

[185] Riccardo Porotti et al. "Deep Reinforcement Learning for Quantum State Preparation with Weak Nonlinear Measurements." In: *Quantum* 6 (2022), p. 747. arXiv: 2107.08816 (Cited on page 197).

[186] V. F. Krotov. "Global Methods in Optimal Control Theory." In: *Progress in Systems and Control Theory ((PSCT,volume 17))*. 2018, pp. 74–121 (Cited on page 197).

[187] Patrick Doria, Tommaso Calarco, and Simone Montangero. "Optimal Control Technique for Many-Body Quantum Dynamics." In: *Physical Review Letters* 106 (2011), p. 190501 (Cited on page 197).

[188] József Somlói, Vladimir A. Kazakov, and David J. Tannor. "Controlled dissociation of I2 via optical transitions between the X and B electronic states." In: *Chemical Physics* 172 (1993), pp. 85–98 (Cited on page 197).

[189] Giuseppe Carleo et al. "Machine learning and the physical sciences." In: *Reviews of Modern Physics* 91 (2019), p. 45002. arXiv: 1903.10563 (Cited on page 197).

[190] Anna Dawid et al. *Modern applications of machine learning in quantum sciences*. 2023. arXiv: 2204.04198 [quant-ph] (Cited on page 197).

[191] Mario Krenn et al. "Artificial intelligence and machine learning for quantum technologies." In: *Physical Review A* 107 (2023), p. 010101. arXiv: 2208.03836 (Cited on page 197).

[192] Moritz August and José Miguel Hernández-Lobato. *Taking gradients through experiments: LSTMs and memory proximal policy optimization for black-box quantum control*. 2018. arXiv: 1802.04063 [cs.LG] (Cited on page 197).

[193] Xiao-Ming Zhang et al. "Automatic spin-chain learning to explore the quantum speed limit." In: *Physical Review A* 97 (2018), p. 052333. arXiv: 1802.09248 (Cited on pages 197, 198).

[194] Zheng An and D. L. Zhou. "Deep reinforcement learning for quantum gate control." In: *EPL (Europhysics Letters)* 126 (2019), p. 60002. arXiv: 1902.08418 (Cited on page 197).

[195] Thomas Foesel et al. "Reinforcement Learning with Neural Networks for Quantum Feedback." In: *Physical Review X* 8 (2018), p. 031084. arXiv: 1802.05267 (Cited on pages 197, 199).

[196] Murphy Yuezhen Niu et al. "Universal quantum control through deep reinforcement learning." In: *npj Quantum Information* 5 (2019), p. 33 (Cited on page 197).

[197] Jelena Mackeprang, Durga B. Rao Dasari, and Jörg Wrachtrup. "A reinforcement learning approach for quantum state engineering." In: *Quantum Machine Intelligence* 2 (2020), p. 5 (Cited on pages 197, 199).

[198] Yuan Hang Zhang et al. "Topological Quantum Compiling with Reinforcement Learning." In: *Physical Review Letters* 125 (2020), p. 170501. arXiv: 2004.04743 (Cited on page 197).

[199] Ryan Sweke et al. "Reinforcement learning decoders for fault-tolerant quantum computation." In: *Machine Learning: Science and Technology* 2.2 (2020), p. 025005 (Cited on page 197).

[200] Zhikang T. Wang, Yuto Ashida, and Masahito Ueda. "Deep Reinforcement Learning Control of Quantum Cartpoles." In: *Physical Review Letters* 125 (2020), p. 100401. arXiv: 1910.09200 (Cited on pages 197, 199).

[201] Sangkha Borah et al. "Measurement-Based Feedback Quantum Control with Deep Reinforcement Learning for a Double-Well Nonlinear Potential." In: *Physical Review Letters* 127 (2021), p. 190403. arXiv: 2104.11856 (Cited on pages 197, 199).

[202] Riccardo Porotti et al. "Deep Reinforcement Learning for Quantum State Preparation with Weak Nonlinear Measurements." In: *Quantum* 6 (2022), p. 747. arXiv: 2107.08816 (Cited on pages 197, 199).

[203] Li-Li Ye et al. *Entanglement engineering of optomechanical systems by reinforcement learning*. 2024. arXiv: 2406.04550 [quant-ph] (Cited on page 197).

[204] Neel R. Vora et al. *ML-Powered FPGA-based Real-Time Quantum State Discrimination Enabling Mid-circuit Measurements*. 2024. arXiv: 2406.18807 [quant-ph] (Cited on page 197).

[205] Yuval Baum et al. "Experimental Deep Reinforcement Learning for Error-Robust Gate-Set Design on a Superconducting Quantum Computer." In: *PRX Quantum* 2 (2021), p. 040324. arXiv: 2105.01079 (Cited on page 197).

[206] Pai Peng et al. "Deep Reinforcement Learning for Quantum Hamiltonian Engineering." In: *Physical Review Applied* 18 (2022), p. 024033. arXiv: 2102.13161 (Cited on page 197).

[207] Alessandro Lumino et al. "Experimental Phase Estimation Enhanced by Machine Learning." In: *Physical Review Applied* 10 (2018), p. 044033. arXiv: 1712.07570 (Cited on page 197).

[208] Jia-Hao Cao et al. "Detection of Entangled States Supported by Reinforcement Learning." In: *Physical Review Letters* 131 (2023), p. 073201 (Cited on page 197).

[209] Egor S Tiunov et al. "Experimental quantum homodyne tomography via machine learning." In: *Optica* 7 (2020), pp. 448–454 (Cited on page 197).

[210] Shuai-Feng Guo et al. "Faster State Preparation across Quantum Phase Transition Assisted by Reinforcement Learning." In: *Physical Review Letters* 126 (2021), p. 060401. arXiv: 2011.11987 (Cited on page 197).

[211] Iris Agresti et al. "Pattern recognition techniques for boson sampling validation." In: *Physical Review X* 9 (2019), p. 011013 (Cited on page 197).

[212] Ming-Zhong Ai et al. "Experimentally realizing efficient quantum control with reinforcement learning." In: *Science China Physics, Mechanics & Astronomy* 65 (2022), p. 250312 (Cited on page 197).

[213] V. V. Sivak et al. "Real-time quantum error correction beyond break-even." In: *Nature* 616 (2023), pp. 50–55. arXiv: 2211.09116 (Cited on page 197).

[214] Mohamed Abdelhafez et al. "Universal gates for protected superconducting qubits using optimal control." In: *Physical Review A* 101 (2020), p. 22321. arXiv: 1908.07637 (Cited on page 198).

[215] Nelson Leung et al. "Speedup for quantum optimal control from automatic differentiation based on graphics processing units." In: *Physical Review A* 95 (2017), p. 042318. arXiv: 1612.04929 (Cited on page 198).

[216] Riccardo Porotti, Vittorio Peano, and Florian Marquardt. "Gradient-Ascent Pulse Engineering with Feedback." In: *PRX Quantum* 4 (2023), p. 030305. arXiv: 2203.04271 (Cited on page 198).

[217] Frédéric Sauvage and Florian Mintert. "Optimal Control of Families of Quantum Gates." In: *Phys. Rev. Lett.* 129 (2022), p. 050507 (Cited on pages 198, 203).

[218] Mazyar Mirrahimi et al. "Dynamically protected cat-qubits: a new paradigm for universal quantum computation." In: *New Journal of Physics* 16.4 (2014), p. 045014 (Cited on page 199).

[219] Jérémie Guillaud and Mazyar Mirrahimi. "Repetition Cat Qubits for Fault-Tolerant Quantum Computation." In: *Phys. Rev. X* 9 (2019), p. 041053 (Cited on page 199).

[220] Christopher Chamberland et al. "Building a Fault-Tolerant Quantum Computer Using Concatenated Cat Codes." In: *PRX Quantum* 3 (2022), p. 010329 (Cited on page 199).

[221] Diego Ruiz et al. *LDPC-cat codes for low-overhead quantum computing in 2D*. 2024. arXiv: 2401.09541 [quant-ph] (Cited on page 199).

[222] Nissim Ofek et al. "Extending the lifetime of a quantum bit with error correction in superconducting circuits." In: *Nature* 536 (2016), pp. 441–445 (Cited on page 199).

[223] Z. Leghtas et al. "Confining the state of light to a quantum manifold by engineered two-photon loss." In: *Science* 347 (2015), pp. 853–857. eprint: https://www.science.org/doi/pdf/10.1126/science.aaa2085 (Cited on page 199).

[224] S. Touzard et al. "Coherent Oscillations inside a Quantum Manifold Stabilized by Dissipation." In: *Phys. Rev. X* 8 (2018), p. 021005 (Cited on page 199).

[225] S. Rosenblum et al. "Fault-tolerant detection of a quantum error." In: *Science* 361 (2018), pp. 266–270. eprint: https://www.science.org/doi/pdf/10.1126/science.aat3996 (Cited on page 199).

[226] A. Grimm et al. "Stabilization and operation of a Kerr-cat qubit." In: *Nature* 584 (2020), pp. 205–209 (Cited on page 199).

[227] Jeffrey M. Gertler et al. "Protecting a bosonic qubit with autonomous quantum error correction." In: *Nature* 590.7845 (2021), pp. 243–248 (Cited on page 199).

[228] Raphaël Lescanne et al. "Exponential suppression of bit-flips in a qubit encoded in an oscillator." In: *Nature Physics* 16 (2020), pp. 509–513 (Cited on page 199).

[229] Daisuke Iyama et al. "Observation and manipulation of quantum interference in a superconducting Kerr parametric oscillator." In: *Nature Communications* 15.1 (2024), p. 86 (Cited on page 199).

[230] U. Réglade et al. "Quantum control of a cat qubit with bit-flip times exceeding ten seconds." In: *Nature* 629 (2024), pp. 778–783 (Cited on page 199).

[231] A. Marquet et al. "Harnessing two-photon dissipation for enhanced quantum measurement and control." In: *Phys. Rev. Appl.* 22 (3 2024), p. 034053 (Cited on page 199).

[232] Matteo Puviani et al. *Boosting the Gottesman-Kitaev-Preskill quantum error correction with non-Markovian feedback*. 2023. arXiv: 2312.07391 [quant-ph] (Cited on page 199).

[233] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: A Bradford Book, 2018 (Cited on page 203).

[234] James Bradbury et al. *JAX: composable transformations of Python+NumPy programs*. Version 0.3.13. 2018 (Cited on page 203).

[235] Jonathan Heek et al. *Flax: A neural network library and ecosystem for JAX*. Version 0.8.5. 2023 (Cited on page 203).

[236] J.R. Johansson, P.D. Nation, and Franco Nori. "QuTiP 2: A Python framework for the dynamics of open quantum systems." In: *Computer Physics Communications* 184 (Apr. 2013), pp. 1234–1240 (Cited on pages 203, 207, 208, 238, 239).

[237] Luiz Davidovich. "Sub-Poissonian processes in quantum optics." In: *Reviews of Modern Physics* 68 (1996), pp. 127–173 (Cited on page 204).

[238] W. R. Johnson. *Atomic Structure Theory: Lectures on Atomic Physics*. Springer, New York, 2007 (Cited on page 212).

[239] L Bretheau et al. "Quantum dynamics of an electromagnetic mode that cannot have N photons." In: *Science* 348 (2015), p. 776 (Cited on page 230).

[240] Benjamin Schumacher and Michael D Westmoreland. "Sending classical information via noisy quantum channels." In: *Physical Review A* 56 (1997), p. 131 (Cited on pages 250, 251).

[241] A Holevo. "The capacity of the quantum channel with general signal states." In: *Information Theory, IEEE Transactions on* 44 (1998), p. 269 (Cited on pages 250, 251).

[242] Alexander Semenovich Holevo. "Bounds for the quantity of information transmitted by a quantum communication channel." In: *Problemy Peredachi Informatsii* 9 (1973), pp. 3–11 (Cited on page 250).

[243] Harold Ollivier and Wojciech H Zurek. "Quantum discord: a measure of the quantumness of correlations." In: *Physical review letters* 88 (2001), p. 017901 (Cited on page 251).

[244] Shabir Barzanjeh et al. "Microwave Quantum Illumination." In: *Physical Review Letters* 114 (Feb. 2015), p. 080503 (Cited on page 251).

[245] Laszlo Gyongyosi, Sandor Imre, and Hung Viet Nguyen. "A survey on quantum channel capacities." In: *IEEE Communications Surveys & Tutorials* 20 (2018), pp. 1149–1205 (Cited on page 251).

[246] Yichen Huang. "Computing quantum discord is NP-complete." In: *New journal of physics* 16 (2014), p. 033027 (Cited on page 251).

[247] Christopher A. Fuchs. *Distinguishability and Accessible Information in Quantum Theory*. 1996. arXiv: `quant-ph/9601020 [quant-ph]` (Cited on page 252).

[248] L. B. Levitin. "Optimal Quantum Measurements for Two Pure and Mixed States." In: *Quantum Communications and Measurement* (1995), pp. 439–448 (Cited on page 252).

[249] Erich L Lehmann and George Casella. *Theory of point estimation*. Springer Science & Business Media, 2006 (Cited on page 254).

[250] Solomon Kullback. *Information theory and statistics*. Courier Corporation, 1997 (Cited on page 255).

[251] Samuel L Braunstein, Carlton M Caves, and Gerard J Milburn. "Generalized uncertainty relations: theory, examples, and Lorentz invariance." In: *annals of physics* 247 (1996), pp. 135–173 (Cited on page 256).

[252] Mattia Walschaers et al. "Statistical signatures of multimode single-photon-added and-subtracted states of light." In: *Physical Review A* 96 (2017), p. 053835 (Cited on page 257).

[253] Todd Tilma et al. "Wigner functions for arbitrary quantum systems." In: *Physical review letters* 117 (2016), p. 180401 (Cited on page 257).

[254] GJ Dolan. "Offset masks for lift-off photoprocessing." In: *Applied Physics Letters* 31 (1977), pp. 337–339 (Cited on page 262).

[255] Vinay Ambegaokar and Alexis Baratoff. "Tunneling between superconductors." In: *Physical review letters* 10 (1963), p. 486 (Cited on page 267).