Tomography of state and parameters in open quantum systems

Pierre Six
CAS - MINES ParisTech
QUANTIC Team - INRIA

works as a PhD student of Pierre Rouchon

Wednesday, the 25th of January, 2017

MINES ParisTech

Subject of this talk

Traditionally, quantum state tomography is based on **single projective measurements** [Paris et Řeháček, 2004]
Several techniques exist, including:

- *MaxLike principle* [Lvovsky et Raymer, RMP 2009]

$$\hat{\rho} = \arg\max_{\rho} \left(\prod_{\nu=1}^{n} \mathrm{Tr}\left(\rho \Pi_\nu\right)^{m_\nu} \right)$$

- *MaxEnt principle* [Bužek, Lecture Notes on Physics **649**, 2004]

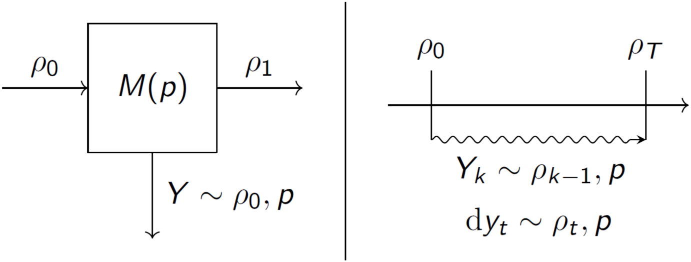
$\hat{\rho}$ most entropic state verifying $\mathrm{Tr}\left(\hat{\rho}\Pi_\nu\right) = \pi_\nu, \forall \nu \leq n$

- *Compressed Sensing* [Gross *et al.*, PRL 2010]

$\hat{\rho}$ minimizing $||\hat{\rho}||_{\mathrm{tr}}$ and verifying $\mathrm{Tr}\left(\hat{\rho}\Pi_\nu\right) = \pi_\nu, \forall \nu \leq n$

Subject of this talk

Previous techniques don't take into account the **dynamics** of the system



Goal of my thesis: state and parameter quantum tomography **along** some **discrete-time and diffusive quantum trajectories** in finite-dimensionnal systems

Contributions of this talk

Based on **quantum trajectories**

- for diffusive systems, a discretization scheme preserving the **positivity** of state
- methods for state quantum tomography using **MaxLike principle**
- introduction to process quantum tomography via **particle filters**, and a result of **stability** of this estimation w.r.t. initial state
- **validations** using outcome trajectories produced by **supraconducting qubit**

Discretization scheme for diffusive systems

Discrete-time systems: a brief reminder

In the following, ρ evolves in a Hilbert space $\mathcal{H}$ **of dimension d**

Definition (Kraus applications, *Nielsen and Chuang*, 2000)

A discrete-time measurement on $\mathcal{H}$ is described by the **complete set of partial Kraus applications** $(\boldsymbol{K}_y)_{y\leq r}$, verifying:

$$r \leq d^2$$

$$\forall y \leq r, \exists l_y \text{ finite and } (M_{yl})_{l\leq l_y} \text{ such that}$$

$$\forall \rho \in \mathcal{L}(\mathcal{H}), \quad \boldsymbol{K}_y(\rho) = \sum_{l=1}^{l_y} M_{yl}\rho M_{yl}^\dagger$$

The set $(\boldsymbol{K}_y)_{y\leq r}$ is complete as:

$$\sum_{y=1}^{r}\sum_{l=1}^{l_y} M_{yl}^\dagger M_{yl} = I_{\mathcal{H}}$$

Discrete-time systems: a brief reminder

- Denoting ρ_- and ρ_+ the states before and after the measurement, and y its outcome:

$$\rho_+ = \frac{\boldsymbol{K}_y(\rho_-)}{\mathrm{Tr}\left(\boldsymbol{K}_y(\rho_-)\right)}$$

$$\mathbb{P}\left(y = \tilde{y} | \rho_- = \tilde{\rho}\right) = \mathrm{Tr}\left(\boldsymbol{K}_{\tilde{y}}(\tilde{\rho})\right)$$

- **Measurement imperfections**: measurement of $1 \leq z \leq r_z$ instead of y

$$\rho_+ = \frac{\sum_{\tilde{y}=1}^{r} \eta_{z\tilde{y}} \boldsymbol{K}_{\tilde{y}}(\rho_-)}{\mathrm{Tr}\left(\sum_{\tilde{y}=1}^{r} \eta_{z\tilde{y}} \boldsymbol{K}_{\tilde{y}}(\rho_-)\right)}$$

$$\mathbb{P}\left(z = \tilde{z} | \rho_- = \tilde{\rho}\right) = \sum_{\tilde{y}=1}^{r} \eta_{\tilde{z}\tilde{y}} \mathrm{Tr}\left(\boldsymbol{K}_{\tilde{y}}(\tilde{\rho})\right)$$

where $\eta_{\tilde{z}\tilde{y}} = \mathbb{P}\left(z = \tilde{z} | y = \tilde{y}\right)$.

Discrete-time systems: a brief reminder

Introducing $\eta_{\tilde{z}\tilde{y}}$ and unitary evolutions in $\boldsymbol{K}_{\tilde{y}}$

Definition (Belavkin filter)

Denote $t_0 < t_1 < t_2 < \ldots < t_T$, with t_0 the initial time and $(t_k)_{k\geq 1}$ the instants during which a **discrete-time measurement occurs** on the system. We associate to each of them a complete Kraus set $(\boldsymbol{K}_{k,y})_{y<r_k}$.

Then, the state evolves from ρ_0 **according to the Belavkin filter**:

$$\rho_k = \frac{\boldsymbol{K}_{k,y_k}(\rho_{k-1})}{\mathrm{Tr}\left(\boldsymbol{K}_{k,y_k}(\rho_{k-1})\right)}$$

$$\mathbb{P}\left(y_k = \tilde{y} | \rho_{k-1} = \tilde{\rho}\right) = \mathrm{Tr}\left(\boldsymbol{K}_{k,\tilde{y}}(\tilde{\rho})\right)$$

where $\rho_k = \rho(t_k)$.

We recall that $\boldsymbol{K}_y(\rho) = \sum_{l=1}^{l_y} M_{yl}\rho M_{yl}^\dagger$.

Evolution of diffusive systems

Stochastic Master Equation [Barchielli et Gregoratti, 2009]

Evolution of **state** ρ, with respect to diffusive measurement processes $(\mathrm{d}y_t^\nu)_{\nu\leq n}$

$$\mathrm{d}\rho_t = \left[-\frac{i}{\hbar}\left[H_t, \rho_t\right] + \sum_{\nu=1}^{n} \mathcal{D}_\nu\left(\rho_t\right)\right]\mathrm{d}t + \sum_{\nu=1}^{n}\sqrt{\eta_\nu}\left[\mathcal{M}_\nu\left(\rho_t\right) - \mathrm{Tr}\left(\mathcal{M}_\nu\left(\rho_t\right)\right)\rho_t\right]\mathrm{d}W_t^\nu$$

and the **measurement processes** $\mathrm{d}y_t^\nu$:

$$\mathrm{d}y_t^\nu = \sqrt{\eta_\nu}\mathrm{Tr}\left(\mathcal{M}_\nu\left(\rho_t\right)\right)\mathrm{d}t + \mathrm{d}W_t^\nu$$

Model **experimentally validated** on a superconducting qubit [Campagne-Ibarcq, PRX **6**]

Evolution of diffusive systems

$$
\begin{aligned}
\mathrm{d}\rho_t &= \left[-\frac{i}{\hbar}\left[H_t, \rho_t\right] + \sum_{\nu=1}^{n} \mathcal{D}_\nu\left(\rho_t\right) \right] \mathrm{d}t \\
&\quad + \sum_{\nu=1}^{n} \sqrt{\eta_\nu}\left[\mathcal{M}_\nu\left(\rho_t\right) - \mathrm{Tr}\left(\mathcal{M}_\nu\left(\rho_t\right)\right)\rho_t\right] \mathrm{d}W_t^\nu \\
\mathrm{d}y_t^\nu &= \sqrt{\eta_\nu}\,\mathrm{Tr}\left(\mathcal{M}_\nu\left(\rho_t\right)\right)\mathrm{d}t + \mathrm{d}W_t^\nu
\end{aligned}
$$

- ρ_t operator of the **finite-dimension** Hilbert space $\mathcal{H}$, **density matrix**, i.e. $\mathrm{Tr}\left(\rho_t\right) = 1$, $\rho_t = \rho_t^\dagger$ and $\rho_t \geq 0$
- H_t **Hamiltonien operator**, zero-trace and Hermitian
- $\mathcal{D}_\nu$ **Lindblad superoperator**:

$$
\mathcal{D}_\nu\left(\rho\right) = L_\nu \rho L_\nu^\dagger - \left(L_\nu^\dagger L_\nu \rho + \rho L_\nu^\dagger L_\nu\right)/2
$$

- $\mathcal{M}_\nu$ **measurement superoperator**:

$$
\mathcal{M}_\nu\left(\rho\right) = L_\nu \rho + \rho L_\nu^\dagger
$$

Evolution of diffusive systems

$$
\begin{aligned}
\mathrm{d}\rho_t &= \left[-\frac{i}{\hbar} \left[H_t, \rho_t \right] + \sum_{\nu=1}^{n} \mathcal{D}_\nu \left(\rho_t \right) \right] \mathrm{d}t \\
&\quad + \sum_{\nu=1}^{n} \sqrt{\eta_\nu} \left[\mathcal{M}_\nu \left(\rho_t \right) - \mathrm{Tr}\left(\mathcal{M}_\nu \left(\rho_t \right) \right) \rho_t \right] \mathrm{d}W_t^\nu \\
\mathrm{d}y_t^\nu &= \sqrt{\eta_\nu} \mathrm{Tr}\left(\mathcal{M}_\nu \left(\rho_t \right) \right) \mathrm{d}t + \mathrm{d}W_t^\nu
\end{aligned}
$$

$(\mathrm{d}W_t^\nu)_\nu$ independent **Wiener processes** (independent and stationary increments, continuous path), verifying:

$$
\mathrm{d}W_t^\nu \sim N\left(0, \sqrt{\mathrm{d}t}\right), \quad \nu \leq n
$$

$\eta_\nu \in [0;1]$: **detection efficiency** of measurement ν

Discretization scheme

The Stochastic Master Equation is not **linear**.

Definition (Discretization scheme)

We provide a **discretization scheme** of the SME :

$$\rho_t = \mathbb{E}\left(\rho_t \middle| \rho_{t-\Delta_t}, \Delta \mathbf{y}_t\right) = \frac{\boldsymbol{K}_{t,\Delta \mathbf{y}_t}\left(\rho_{t-\Delta_t}\right)}{\mathrm{Tr}\left(\boldsymbol{K}_{t,\Delta \mathbf{y}_t}\left(\rho_{t-\Delta_t}\right)\right)}$$

- Δ_t: **sampling time** of diffusive signals
- $\Delta \mathbf{y}_t = (\Delta y_t^\nu)_{\nu \leq n}$ and $\Delta y_t^\nu = \int_{s=t-\Delta_t}^{s=t} \mathrm{d}y_s^\nu$ the **sampled measurements**
- $\rho \mapsto \boldsymbol{K}_{t,\Delta \mathbf{y}_t}(\rho)$ **linear** application with **Kraus application** structure:

$$\boldsymbol{K}_{t,\Delta \mathbf{y}}(\rho) = (N_{t,\Delta \mathbf{y}})\rho(N_{t,\Delta \mathbf{y}})^\dagger + \sum_{\nu=1}^{n} G_\nu \rho G_\nu^\dagger \Delta_t$$

Discretization scheme

This scheme has two advantages:

- it preserves the **positivity of the state** along a measurement trajectory
- it is associated with a **probability law**:

$$\mathbb{P}\Big(\forall \nu \leq n, \Delta y^{\nu} \leq \Delta y_t^{\nu} \leq \Delta y^{\nu} + \mathrm{d}y^{\nu} \Big| \rho_{t-\Delta_t} = \rho\Big) = \mathrm{Tr}\left(\boldsymbol{K}_{t,\Delta\mathbf{y}}(\rho)\right) \times \left(\prod_{\nu=1}^{n} \mathrm{d}y^{\nu}\right)$$

Discretization scheme

Expression of the scheme

The **discretized Kraus application** $\boldsymbol{K}_{t,\Delta\mathbf{y}}$ writes, for any ρ and $\Delta\mathbf{y}$:

$$\boldsymbol{K}_{t,\Delta\mathbf{y}}(\rho) = \frac{\exp\left(-\frac{||\Delta\mathbf{y}||^2}{2}\right)}{(2\pi\Delta t)^{\frac{N}{2}}}\tilde{\boldsymbol{K}}_{t,\Delta\mathbf{y}}(\Sigma_t\rho\Sigma_t^\dagger)$$

where:

- $\tilde{\boldsymbol{K}}_{t,\Delta\mathbf{y}}(\rho) = (M_{t,\Delta\mathbf{y}})\rho(M_{t,\Delta\mathbf{y}})^\dagger + \sum_{\nu=1}^{n}(1-\eta_\nu)L_\nu\rho L_\nu^\dagger\Delta t$
- $M_{t,\Delta\mathbf{y}} = I_{\mathcal{H}} + \sum_{\nu=1}^{n}\eta_\nu L_\nu \Delta y^\nu - C_t\Delta t$
- $C_t = \frac{i}{\hbar}H_t + \sum_{\nu=1}^{n}\frac{L_\nu^\dagger L_\nu}{2}$
- $\Sigma_t = \left[I_{\mathcal{H}} + C_t^\dagger C_t(\Delta t)^2\right]^{-\frac{1}{2}}$

Thanks to Σ_t, $\int_{\mathbb{R}^n}\mathrm{Tr}\left(\boldsymbol{K}_{t,\Delta\mathbf{y}}(\rho)\right)\mathrm{d}\Delta y^1\ldots\mathrm{d}\Delta y^n = 1$ for any ρ and a **strictly positive** Δt

Discretization scheme

Theorem (Convergence in distribution of the discretization scheme)

The discretization scheme defined by:

$$\rho_t = \frac{\boldsymbol{K}_{t,\Delta \mathbf{y}_t}(\rho_{t-\Delta_t})}{Tr(\boldsymbol{K}_{t,\Delta \mathbf{y}_t}(\rho_{t-\Delta_t}))},$$

$$\mathrm{d}\mathbb{P}\left(\Delta \mathbf{y}_t = \Delta \mathbf{y} \middle| \rho_{t-\Delta_t} = \rho\right) = Tr(\boldsymbol{K}_{t,\Delta \mathbf{y}}(\rho))$$

converges in distribution *towards the SME, for Δt tending to the infinitesimal $\mathrm{d}t$.*

Sketch of proof: we show two assumptions

- same **first-order generators** for both equations: $\forall f$ scalar C^2-function and $\forall \rho$ density matrix, the expectancy of

  $$\frac{f(\rho_t) - f(\rho_{t-\Delta t})}{\Delta t} \text{ over } \Delta \mathbf{y}_t, \text{ assuming } \rho_{t-\Delta t} = \rho$$

  converges when $\Delta t \to \mathrm{d}t$ to the value obtained with SME.
- **tightness hypothesis**

Discretization scheme: first-order generators

Concerning the SME, recalling that:

$$\mathrm{d}\rho_t = \mu(\rho_t, t)\mathrm{d}t + \sum_{\nu=1}^{n} \sigma_\nu(\rho_t, t)\mathrm{d}W_t^\nu$$

the first-order generator is computed using the **Itō differential calculus rules**:

$$Af(\rho) = \mathbb{E}_{\mathrm{d}\mathbf{y}_t}\left(\frac{f(\rho_t) - f(\rho_{t-\mathrm{d}t} = \rho)}{\mathrm{d}t}\right) =$$

$$\nabla f\big|_\rho .\mu(\rho_t, t) + \frac{1}{2}\,\nabla^2 f\Big|_\rho \cdot \left(\sum_{\nu=1}^{n}[\sigma_\nu(\rho, t), \sigma_\nu(\rho, t)]\right)$$

Discretization scheme: first-order generators

Concerning the discretization scheme, we simply use **expansion series**:

$$
Bf(\rho) = \lim_{\Delta t \to \mathrm{d}t} \mathbb{E}_{\Delta \mathbf{y}} \left(\frac{f(\rho_t) - f(\rho_{t-\Delta t} = \rho)}{\Delta t} \right) =
$$

$$
\lim_{\Delta t \to \mathrm{d}t} \mathbb{E}_{\Delta \mathbf{y}} \left(\frac{f\left(\frac{\boldsymbol{K}_{t,\Delta \mathbf{y}}(\rho)}{\mathrm{Tr}\left(\boldsymbol{K}_{t,\Delta \mathbf{y}}(\rho)\right)} \right) - f(\rho_{t-\Delta t} = \rho)}{\Delta t} \right)
$$

- the term $\exp\left(-\frac{||\Delta \mathbf{y}||^2}{4}\right) / (2\pi \Delta t)^{\frac{N}{4}}$ **vanishes**
- terms of the order Δy_t^ν, $(\Delta y_t^\nu)^2$ and $\mathrm{d}t$: we consider $\tilde{\boldsymbol{K}}_{t,\Delta \mathbf{y}}(\rho)$ whose additive structure fits well with expansion series

Quantum state tomography using MaxLike principle

MaxLike principle

Definition (Maximum Likelihood)

Using realisations $X_1, \ldots, X_N$ successively obtained according to the probability law

$$X \mapsto f(X \,\middle|\, \theta = \overline{\theta}),$$

the MaxLike estimation of $\overline{\theta}$ consists in choosing $\hat{\theta}_{ML}$ which maximizes on the set Θ the **conditional probabilities** of X_i :

$$\hat{\theta}_{ML} = \underset{\theta \in \Theta}{\arg\max}\, f\left(X_1, \ldots, X_N \middle| \theta\right)$$

$$= \underset{\theta \in \Theta}{\arg\max} \prod_{n=1}^{N} f\left(X_n \middle| \theta\right)$$

f is called **likelihood function**.

Consistency of Maximum Likelihood

The MaxLike estimation **can be biased** :

$$\hat{\theta}_{ML}(N) = \mathbb{E}_{X_1,\dots,X_N}\left(\hat{\theta}_{ML}\left(X_1,\dots,X_N\right)\right) \neq \overline{\theta},$$

but, if Θ is a compact space, $\overline{\theta} \in \mathrm{int}\left(\Theta\right)$ (among other properties), this estimation is **consistent** :

$$\lim_{N\to+\infty} \hat{\theta}_{ML}(N) = \overline{\theta} \quad \text{in probability}$$

Definition (Estimation variance)

Estimation variance $\sigma^2_{ML}(N)$ of Maximum Likelihood is defined by:

$$\sigma^2_{ML}(N) = \mathbb{E}_{X_1,\dots,X_N} \\ \left(\left(\hat{\theta}_{ML}\left(X_1,\dots,X_N\right) - \hat{\theta}_{ML}(N)\right) . \left(\hat{\theta}_{ML}\left(X_1,\dots,X_N\right) - \hat{\theta}_{ML}(N)\right)^T\right)$$

Maximum Likelihood properties

Cramér-Rao bound

- For any **unbiased** estimator C, if the likelihood f is a C^2-function w.r.t. θ, and $\mathbb{E}_X$ are $\partial^2/\partial\theta^2$ invertible :

$$\sigma_C^2(N) \geq \frac{\left(-\mathbb{E}_X\left(\frac{\partial^2}{\partial\theta^2}\log\left(f\left(X\middle|\overline{\theta}\right)\right)\right)\right)^{-1}}{N}$$

  This bound **does not depend** on the estimator C.

- Maximum Likelihood is **asymptotically efficient** :

$$\sigma_{ML}^2(N) \sim_{N\to+\infty} \frac{\left(-\mathbb{E}_X\left(\frac{\partial^2}{\partial\theta^2}\log\left(f\left(X\middle|\overline{\theta}\right)\right)\right)\right)^{-1}}{N} \quad \text{in prob.}$$

  when $\mathbb{E}_X$ and $\partial^3/\partial\theta^3$ are locally invertible around $\overline{\theta}$.

Likelihood function of measurement trajectories

Likelihood of a measurement trajectory $(y_1, \dots, y_T)$, **supposing that** $\rho_0 = \hat{\rho}$?

- Bayes' law gives:

$$\mathbb{P}\left((y_k)_{k\leq T} \middle| \hat{\rho}\right) = \prod_{k=1}^{T} \mathbb{P}\left(y_k \middle| \hat{\rho}, (y_s)_{s\leq k-1}\right)$$

- **Approximated quantum filter** initialized with $\hat{\rho}$:

$$\hat{\rho}_0 = \hat{\rho}, \quad \hat{\rho}_k = \frac{\boldsymbol{K}_{y_k}(\hat{\rho}_{k-1})}{\mathrm{Tr}\left(\boldsymbol{K}_{y_k}(\hat{\rho}_{k-1})\right)},$$

$$\mathbb{P}\left(y_k \middle| \hat{\rho}, (y_s)_{s\leq k-1}\right) = \mathrm{Tr}\left(\boldsymbol{K}_{y_k}(\hat{\rho}_{k-1})\right)$$

- **Linearity** of Kraus applications

Likelihood function of measurement trajectories

Likelihood of a measurement trajectory $(y_1, \ldots, y_T)$, supposing that $\rho_0 = \hat{\rho}$:

$$\mathbb{P}\left((y_k)_{k\leq T} \middle| \hat{\rho}\right) = \mathrm{Tr}\left(\boldsymbol{K}_{y_T} \circ \ldots \circ \boldsymbol{K}_{y_1}(\hat{\rho})\right)$$

Adjoint Kraus applications denoted $\boldsymbol{K}^*$ [**Quantum Smoothing** - Tsang 2009, **Past Quantum States** - Gammelmark *et al.*, 2013] :

$$\forall \text{ operators } A \text{ and } B, \quad \mathrm{Tr}\left(A\boldsymbol{K}_y(B)\right) = \mathrm{Tr}\left(\boldsymbol{K}_y^*(A)B\right)$$

Thus, the **likelihood of a trajectory** is given by:

$$\mathbb{P}\left((y_k)_{k\leq T} \middle| \hat{\rho}\right) = \mathrm{Tr}\left(\boldsymbol{K}_{y_1}^* \circ \ldots \circ \boldsymbol{K}_{y_T}^*(I)\,\hat{\rho}\right)$$

Likelihood function of measurement trajectories: conclusion

Thus, when a physicist have N quantum trajectories at his disposal, he associates to any trajectory an **adjoint state** $E^{(n)}$:

$$\boxed{E^{(n)} = \frac{\boldsymbol{K}^*_{y_1^{(n)}} \circ \ldots \circ \boldsymbol{K}^*_{y_T^{(n)}}(I)}{\mathrm{Tr}\left(\boldsymbol{K}^*_{y_1^{(n)}} \circ \ldots \circ \boldsymbol{K}^*_{y_T^{(n)}}(I)\right)}}$$

The log-likelihood of the measurement data is then given by:

$$\log\left(\mathbb{P}\left(\left(\mathbf{y}^{(n)}\right)_{n\leq N} \middle| \hat{\rho}\right)\right) = \sum_{n=1}^{N} \log \mathrm{Tr}\left(E^{(n)}\hat{\rho}\right) + C^{(n)}$$

where $C^{(n)}$ is independent from $\hat{\rho}$.

From now on, we use the function **g** :

$$\boxed{\hat{\rho} \mapsto g\left(\hat{\rho}\right) = \sum_{n=1}^{N} \log \mathrm{Tr}\left(E^{(n)}\hat{\rho}\right)}$$

Optimization

Maximization problem

The goal is to find :

$$\hat{\rho}_{ML} = \underset{\hat{\rho}\in\mathcal{D}_{\mathcal{H}}}{\arg\max}\, g\left(\hat{\rho}\right) = \underset{\hat{\rho}\in\mathcal{D}_{\mathcal{H}}}{\arg\max} \left(\sum_{n=1}^{N} \log \operatorname{Tr}\left(E^{(n)}\hat{\rho}\right)\right)$$

- $\mathcal{D}_{\mathcal{H}}$ the set of density operators on $\mathcal{H}$ is **convex and compact**
- the log-likelihood g is **strongly concave** (iff $\mathrm{Vect}(E^{(n)})$ describes the set of Hermitian operators)

$=>$ *classical optimization by gradient descent*

$=>$ *asymptotic development of Cramér-Rao bound* ***around*** $\hat{\rho}_{ML}$

The gradient is easy to implement and can be commputed quickly:

$$\nabla g|_{\hat{\rho}} = \sum_{n=1}^{N} \frac{E^{(n)}}{\operatorname{Tr}\left(E^{(n)}\hat{\rho}\right)}$$

Optimization

Optimality conditions

The density matrix $\overline{\rho}$ **maximizes** the likelihood function iff [Rouchon and Six, arXiv:1607.00948]:

- $\mathrm{Tr}\left(E^{(n)}\overline{\rho}\right) > 0$ for $1 \leq n \leq N$
- $\left[\overline{\rho}, \nabla g|_{\overline{\rho}}\right] = \overline{\rho}\, \nabla g|_{\overline{\rho}} - \nabla g|_{\overline{\rho}}\overline{\rho} = 0$
- there exists $\overline{\lambda} > 0$ such that $\overline{\lambda P} = \overline{P}\, \nabla g|_{\overline{\rho}}$ and $\nabla g|_{\overline{\rho}} \leq \overline{\lambda} I_{\mathcal{H}}$, where $\overline{P}$ is the **orthogonal projector on the range** of $\overline{\rho}$

Maximum Likelihood: estimation variance

$\hat{\rho}_{ML} \in \textbf{int}\left(\mathcal{D}_{\mathcal{H}}\right)$ is equivalent to $\hat{\rho}_{ML}$ **invertible** matrix

Variance of estimation: the invertible case

When $\hat{\rho}_{ML}$ is **invertible**, the estimation variance of $\operatorname{Tr}(\rho A)$ is denoted $\sigma^2_{ML}(A)$ and verifies:

$$\sigma^2_{ML}(A) \sim_{N \mapsto +\infty} \left(- \nabla^2 g \Big|_{\hat{\rho}_{ML}} \right)^{-1} (A_{\parallel}, A_{\parallel})$$

where

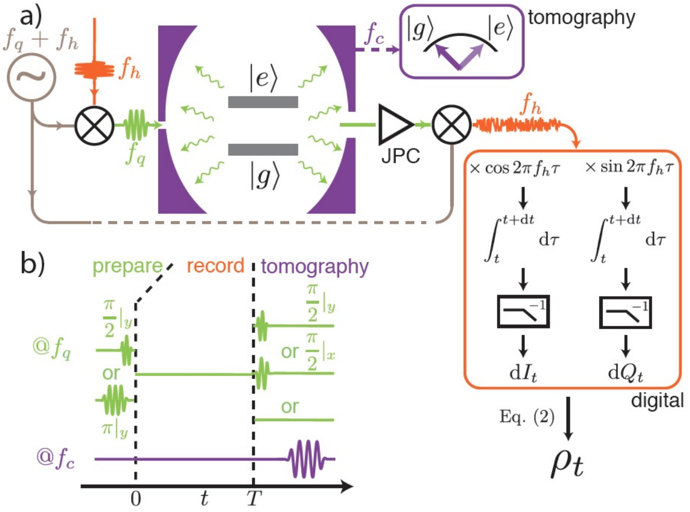
$$\nabla^2 g \Big|_{\hat{\rho}_{ML}} = - \sum_{n=1}^{N} \frac{E_{\parallel}^{(n)} \otimes E_{\parallel}^{(n)}}{E^{(n)} \hat{\rho}_{ML}}$$

tensor of order 2 on $\mathcal{D}_{\mathcal{H}}$.
$E_{\parallel}^{(n)} = E^{(n)} - \operatorname{Tr}\left(E^{(n)}\right) I / \operatorname{Tr}(I)$ the **projection** of $E^{(n)}$ on the subspace of zero-trace operators

Results for **non-invertible matrices** to be published

Superconducting qubit from LPA group

Two-energy-level system [Campagne-Ibarcq et *al.*, PRX]



Experimental validations

- $\mathcal{H} = \mathbb{C}^2$, qubit state in the **Bloch sphere**
- 3 channels of measurement
  - $L_I = \sigma_- / \sqrt{2T_1}$, $\eta_I = \eta > 0$
  - $L_Q = i\sigma_- / \sqrt{2T_1}$, $\eta_Q = \eta > 0$
  - $L_\phi = \sigma_z / \sqrt{2T_\phi}$, $\eta_\phi = 0$
- no Hamiltonian operator H_t

$$
\begin{aligned}
\mathrm{d}\rho_t = {} & \left[\mathcal{D}_I(\rho_t) + \mathcal{D}_Q(\rho_t) + \mathcal{D}_\phi(\rho_t)\right]\mathrm{d}t \\
& + \sqrt{\eta}\left[\mathcal{M}_I(\rho_t) - \mathrm{Tr}\left(\mathcal{M}_I(\rho_t)\right)\rho_t\right]\mathrm{d}W_t^I \\
& + \sqrt{\eta}\left[\mathcal{M}_Q(\rho_t) - \mathrm{Tr}\left(\mathcal{M}_Q(\rho_t)\right)\rho_t\right]\mathrm{d}W_t^Q
\end{aligned}
$$

$$
\mathrm{d}y_t^I = \sqrt{\frac{\eta}{2T_1}}\mathrm{Tr}\left(\sigma_x\rho_t\right)\mathrm{d}t + \mathrm{d}W_t^I
$$

$$
\mathrm{d}y_t^Q = \sqrt{\frac{\eta}{2T_1}}\mathrm{Tr}\left(\sigma_y\rho_t\right)\mathrm{d}t + \mathrm{d}W_t^Q
$$

Experimental validations

We have at our disposal $N = 3.10^6$ outcome trajectories consisting in $T = 47$ measurements on I and Q channels.
Sampling time: $\Delta_t = 200$ ns, while:

- $T_1 \approx 4.15$ μs the characteristic time of decoherence due to the **measurement**
- $T_\phi \approx 35$ μs the characteristic time of decoherence due to the **dephasing**

We aim to prepare ρ_0 on $(1, 0, 0)$, but **slight imperfections in the preparation process**.

Experimental validations

We also estimate:

$$\bar{\rho}_t = \frac{1}{N}\sum_{n=1}^{N} \bar{\rho}_t^{(n)}$$

the **mean of state values** at time t, using:

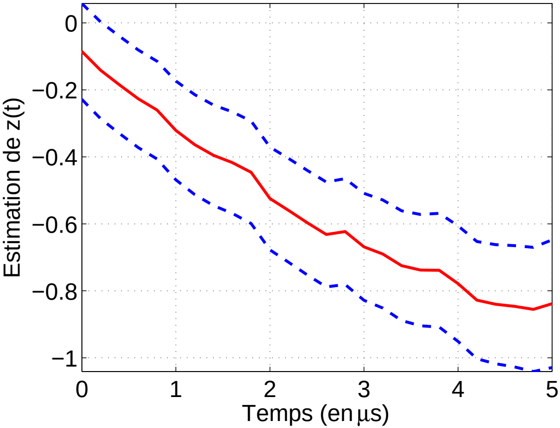
$$\hat{\rho}_{t,ML} = \underset{\hat{\rho}\in\mathcal{D}_{\mathcal{H}}}{\arg\max} \prod_{n=1}^{N} \mathrm{d}\mathbb{P}\left(\Delta\mathbf{y}_{t+1}^{(n)}, \ldots, \Delta\mathbf{y}_{T}^{(n)} \middle| \rho_t^{(n)} = \hat{\rho}\right)$$

Comparison with the estimation of $\bar{x}_t$ and $\bar{y}_t$ by **averaging** the measurement signals:

$$< \mathrm{d}y_t^I > = \sqrt{\frac{\eta}{2T_1}} x_t \mathrm{d}t$$

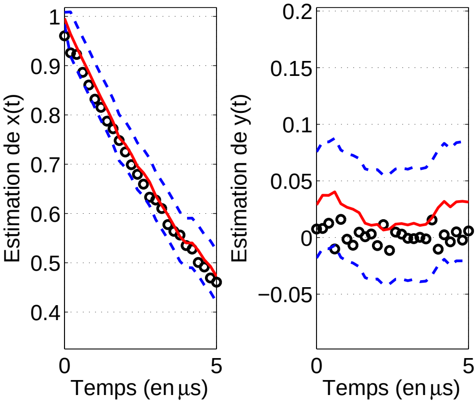
$$< \mathrm{d}y_t^Q > = \sqrt{\frac{\eta}{2T_1}} y_t \mathrm{d}t$$

Experimental validations on Z



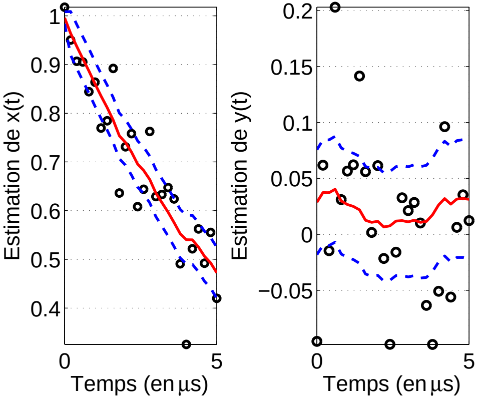
- red: **Maximum Likelihood** estimation of $z_{ML}(t)$, obtained using $N = \mathbf{4.10^4}$ outcome trajectories
- blue: **confidence interval** $z_{ML}(t) \pm 2\sigma_{ML}^2(\sigma_z)$

Experimental validations on X and Y



- red: **Maximum Likelihood** estimation of $x_{ML}(t)$ and $y_{ML}(t)$, obtained using $N = \mathbf{4.10^4}$ outcome trajectories
- blue: **confidence interval** $x_{ML}(t) \pm 2\sigma^2_{ML}(\sigma_x)$ (id. for y)
- points: **averaging** estimation using $N = \mathbf{3.10^6}$ outcome traj.

Experimental validations on X and Y



- red: **Maximum Likelihood** estimation of $x_{ML}(t)$ and $y_{ML}(t)$, obtained using $N = \mathbf{4.10^4}$ outcome trajectories
- blue: **confidence interval** $x_{ML}(t) \pm 2\sigma^2_{ML}(\sigma_x)$ (id. for y)
- points: **averaging** estimation using $N = \mathbf{4.10^4}$ outcome traj.

Quantum process tomography using MaxLike principle and particle filters

Introduction to process tomography

The problem is to estimate the **exact parameter value** $\overline{p}$ that generated the outcome trajectories $\mathbf{y}^{(n)}$ thanks to the Belavkin filter

$$\overline{\rho}_0^{(n)} = \overline{\rho}, \quad \overline{\rho}_k^{(n)} = \frac{\boldsymbol{K}^{\overline{p}}_{k^{(n)}, y_k^{(n)}}\left(\overline{\rho}_{k-1}^{(n)}\right)}{\mathrm{Tr}\left(\boldsymbol{K}^{\overline{p}}_{k^{(n)}, y_k^{(n)}}\left(\overline{\rho}_{k-1}^{(n)}\right)\right)}$$

$$\mathbb{P}\left(y_k^{(n)} = y | \overline{p}, \overline{\rho}_{k-1}^{(n)}\right) = \mathrm{Tr}\left(\boldsymbol{K}^{\overline{p}}_{k^{(n)}, y}(\overline{\rho}_{k-1}^{(n)})\right)$$

- exact values $\overline{\rho}$ and $\overline{p}$ may not be known
  => consider the **approximate filter** using ρ and p
- stability is measured using **fidelity** pseudo-metric F:

$$F\left(\rho, \rho'\right) = \sqrt{\sqrt{\rho}\rho'\sqrt{\rho}} \in [0; 1]$$

with $F\left(\rho, \rho'\right) = 1$ equivalent to ρ and ρ' describing the same physical state.

A first stability result

Theorem (Stability w.r.t. initial condition, *Rouchon et al.*, IEEE, 2011)

Exact filter $(\overline{\rho}, \overline{p})$	*Approximate filter* $(\rho, \overline{p})$
$\overline{\rho}_0 = \overline{\rho}$	$\rho_0^{\overline{p}} = \rho$
$\overline{\rho}_k = \dfrac{\boldsymbol{K}^{\overline{p}}_{k,y_k}(\overline{\rho}_{k-1})}{Tr\left(\boldsymbol{K}_{k,y_k}\left(\overline{\rho}^{\overline{p}}_{k-1}\right)\right)}$	$\rho_k^{\overline{p}} = \dfrac{\boldsymbol{K}^{\overline{p}}_{k,y_k}\left(\rho^{\overline{p}}_{k-1}\right)}{Tr\left(\boldsymbol{K}^{\overline{p}}_{k,y_k}\left(\rho^{\overline{p}}_{k-1}\right)\right)}$

The y_k are generated by the exact filter.
*Then, at each time k, the fidelity between exact and approximate states **increases in expectancy**:*

$$\boxed{\mathbb{E}_{y_k}\left(F\left(\overline{\rho}_k, \rho_k^{\overline{p}}\right)\middle| y_1, \dots, y_{k-1}\right) \geq F\left(\overline{\rho}_{k-1}, \rho_{k-1}^{\overline{p}}\right)}$$

A quantum filter tends to forget its initial condition.

Problematics

*Is it possible to **estimate** $\overline{p}$ initializing the filter with **approximate** initial state ρ ?*
*Is **Maximum-Likelihood Principle** relevant for that ?*

- Device with $(\overline{\rho}, \overline{p})$ generated trajectory $(y_1, \ldots, y_T)$.
- Experimenter knows that $\overline{p} \in \{a; b\}$.
- Experimenter has to test filters (ρ, a) and (ρ, b), with ρ an approximate initial state.
- The measurements are the **only** information available on $\overline{p}$.

Probability filter

We denote: $\left| \begin{array}{l} \pi_k^a = \mathbb{P}\left(\overline{p} = a \middle| \rho, y_1, \ldots, y_k\right) \\ \pi_k^b = \mathbb{P}\left(\overline{p} = b \middle| \rho, y_1, \ldots, y_k\right) \end{array} \right.$

These probabilities iterate with the following parallel filters, via **Bayes' law**:

Approximate Filter (ρ, a)	Approximate Filter (ρ, b)
$\rho_0^a = \rho, \quad \pi_0^a = \pi^a$	$\rho_0^b = \rho, \quad \pi_0^b = \pi^b$
$\rho_k^a = \dfrac{\boldsymbol{K}^a_{k,y_k}\left(\rho^a_{k-1}\right)}{\mathrm{Tr}\left(\boldsymbol{K}^a_{k,y_k}\left(\rho^a_{k-1}\right)\right)}$	$\rho_k^b = \dfrac{\boldsymbol{K}^b_{k,y_k}\left(\rho^b_{k-1}\right)}{\mathrm{Tr}\left(\boldsymbol{K}^b_{k,y_k}\left(\rho^b_{k-1}\right)\right)}$
$\pi_k^a = \dfrac{\pi^a_{k-1}\mathrm{Tr}\left(\boldsymbol{K}^a_{k,y_k}\left(\rho^a_{k-1}\right)\right)}{C_k}$	$\pi_k^b = \dfrac{\pi^b_{k-1}\mathrm{Tr}\left(\boldsymbol{K}^b_{k,y_k}\left(\rho^b_{k-1}\right)\right)}{C_k}$

C_k such that $\pi_k^a + \pi_k^b = 1$.

Set $\overline{p} = a$. What can we say about π_k^a and ρ_k^a ?

Extended density matrices

Extended density matrices Ξ_k are parallel density matrices that enclose the uncertainty on the parameter value:

$$\Xi_k = \left(\begin{array}{c|c} \pi_k^a \rho_k^a & 0 \\ \hline 0 & \pi_k^b \rho_k^b \end{array} \right).$$

(ρ_k^a, π_k^a) and $\left(\rho_k^b, \pi_k^b\right)$ are respectively the states of probability filters (ρ, a) and (ρ, b).
Ξ_k fulfils all the **properties of a density matrix**.

Extended Belavkin filter

Thus, the extended state Ξ_k is iterated thanks to the following formula:

$$\Xi_0 = \left(\begin{array}{c|c} \pi_0^a \rho & 0 \\ \hline 0 & \pi_0^b \rho \end{array}\right),$$

$$\Xi_k = \frac{\left(\begin{array}{c|c} \pi_{k-1}^a \boldsymbol{K}_{k,y_k}^a \left(\rho_{k-1}^a\right) & 0 \\ \hline 0 & \pi_{k-1}^b \boldsymbol{K}_{k,y_k}^b \left(\rho_{k-1}^b\right) \end{array}\right)}{\pi_{k-1}^a \mathrm{Tr}\left(\boldsymbol{K}_{k,y_k}^a \left(\rho_{k-1}^a\right)\right) + \pi_{k-1}^b \mathrm{Tr}\left(\boldsymbol{K}_{k,y_k}^b \left(\rho_{k-1}^b\right)\right)}$$

This is actually an **extended approximate Belavkin filter**, as it is equivalent to $\Xi_k = \boldsymbol{K}_{k,y_k}^{\text{ext}}\left(\Xi_{k-1}\right) / \mathrm{Tr}\left(\boldsymbol{K}_{k,y_k}^{\text{ext}}\left(\Xi_{k-1}\right)\right)$, with $\boldsymbol{K}_y^{\text{ext}}$ a partial Kraus map containing the extended matrices built from the ones of $\boldsymbol{K}_y^a$ and $\boldsymbol{K}_y^b$.

Exact extended filter

The exact filter corresponding to the real device can be extended as well. Its initial state writes:

$$\overline{\Xi}_0 = \left(\begin{array}{c|c} \overline{\rho} & 0 \\ \hline 0 & 0 \end{array} \right), \text{ as } \overline{p} = a.$$

Its state is iterated thanks to the **same** extended Kraus map $\left(\boldsymbol{K}_y^{\text{ext}}\right)$:

$$\overline{\Xi}_k = \boldsymbol{K}^{\text{ext}}_{k,y_k}\left(\overline{\Xi}_{k-1}\right) / \mathrm{Tr}\left(\boldsymbol{K}^{\text{ext}}_{k,y_k}\left(\overline{\Xi}_{k-1}\right)\right)$$

which ensures that:

$$\overline{\Xi}_k = \left(\begin{array}{c|c} \overline{\rho}_k & 0 \\ \hline 0 & 0 \end{array} \right),$$

with $\overline{\rho}_k$ generated by the **exact filter** $(\overline{\rho}, \overline{p})$.

Stability results

Theorem (Stability to initial condition, *Six et al.*, CDC 2015)

Exact filter ($\overline{\rho}$, *a and b*)

$$\Xi_0 = \left(\begin{array}{c|c} \overline{\rho} & 0 \\ \hline 0 & 0 \end{array}\right),$$

$$\Xi_k = \frac{\boldsymbol{K}^{ext}_{k,y_k}\left(\Xi_{k-1}\right)}{Tr\left(\boldsymbol{K}^{ext}_{k,y_k}\left(\Xi_{k-1}\right)\right)}$$

Approximate filter (ρ, *a and b*)

$$\Xi_0 = \left(\begin{array}{c|c} \pi^a_0 \rho & 0 \\ \hline 0 & \pi^b_0 \rho \end{array}\right)$$

$$\Xi_k = \frac{\boldsymbol{K}^{ext}_{k,y_k}\left(\Xi_{k-1}\right)}{Tr\left(\boldsymbol{K}^{ext}_{k,y_k}\left(\Xi_{k-1}\right)\right)}$$

The y_k are generated by the ***exact filter****.*
Thus, $\pi^a_k F\left(\overline{\rho}_k, \rho^a_k\right)$ is a ***sub-martingale****.*

$$\boxed{\mathbb{E}_{y_k}\left(\pi^a_k F\left(\overline{\rho}_k, \rho^a_k\right) \middle| \overline{\Xi}_{k-1}, \Xi_{k-1}\right) \geq \pi^a_{k-1} F\left(\overline{\rho}_{k-1}, \rho^a_{k-1}\right)}$$

Furthermore, if the initial condition $\rho = \overline{\rho}$, it becomes:

$$\boxed{\mathbb{E}_{y_k}\left(\pi^a_k \middle| \overline{\Xi}_{k-1}, \Xi_{k-1}\right) \geq \pi^a_{k-1}}$$

Experimental validations

- $\mathcal{H} = \mathbb{C}^2$, qubit state in the **Bloch sphere**
- 3 channels of measurement
  - $L_I = \sigma_- / \sqrt{2T_1}$, $\eta_I = \eta > 0$
  - $L_Q = i\sigma_- / \sqrt{2T_1}$, $\eta_Q = \eta > 0$
  - $L_\phi = \sigma_z / \sqrt{2T_\phi}$, $\eta_\phi = 0$
- no Hamiltonian operator H_t

$$
\begin{aligned}
\mathrm{d}\rho_t = {} & \left[\mathcal{D}_I(\rho_t) + \mathcal{D}_Q(\rho_t) + \mathcal{D}_\phi(\rho_t)\right]\mathrm{d}t \\
& + \sqrt{\eta}\Big[\mathcal{M}_I(\rho_t) - \mathrm{Tr}\left(\mathcal{M}_I(\rho_t)\right)\rho_t\Big]\mathrm{d}W_t^I \\
& + \sqrt{\eta}\Big[\mathcal{M}_Q(\rho_t) - \mathrm{Tr}\left(\mathcal{M}_Q(\rho_t)\right)\rho_t\Big]\mathrm{d}W_t^Q
\end{aligned}
$$

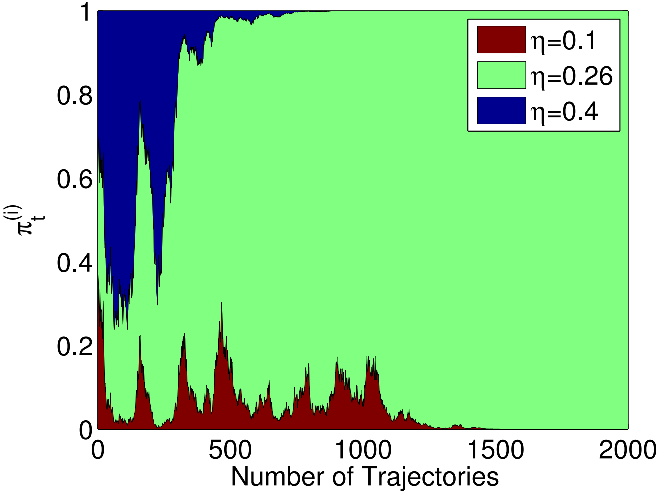
$$
\begin{aligned}
\mathrm{d}y_t^I &= \sqrt{\frac{\eta}{2T_1}}\mathrm{Tr}\left(\sigma_x\rho_t\right)\mathrm{d}t + \mathrm{d}W_t^I \\
\mathrm{d}y_t^Q &= \sqrt{\frac{\eta}{2T_1}}\mathrm{Tr}\left(\sigma_y\rho_t\right)\mathrm{d}t + \mathrm{d}W_t^Q
\end{aligned}
$$

Experimental validations

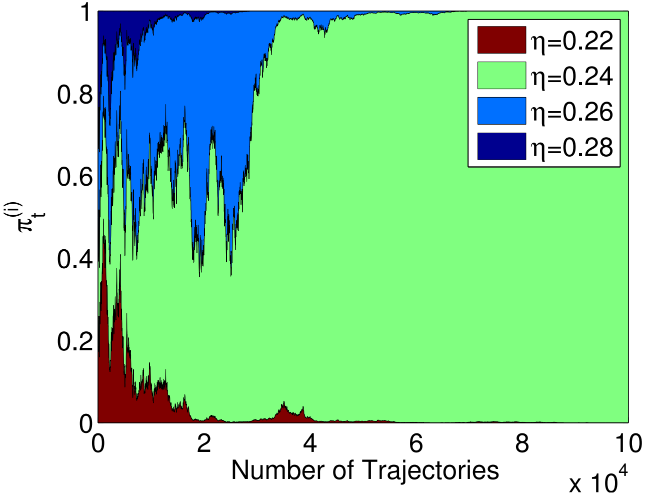
Now, we want to estimate the **detection efficiency** η.

- Preliminary calibration shows that $\overline{\eta} = 0.26 \pm 0.02$
- Still use of $N = 3.10^6$ outcome trajectories of $T = 50$ measurements on I and Q channels
- Same sampling period and measurement/decoherence damping times
- Initial state around $\left(I_{\mathcal{H}} + \sigma_x\right)/2$ to maximize the information on η

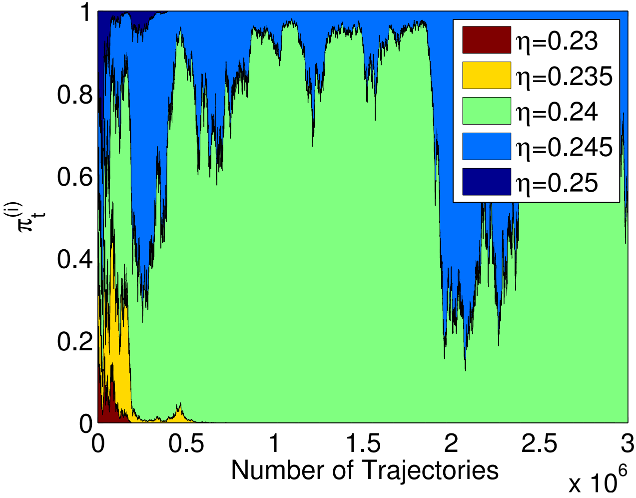
Experimental Validation on a Superconducting Qubit



Experimental Validation on a Superconducting Qubit

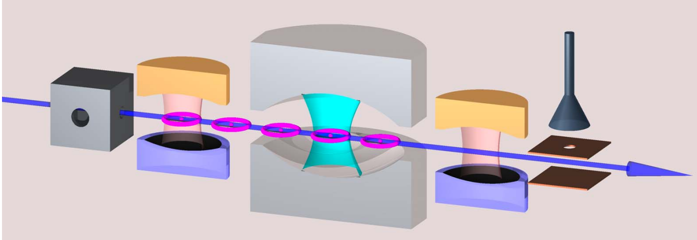


Experimental Validation on a Superconducting Qubit



Outline of other works

- direct optimization for **multi-dimensional** parametric estimation: use of **adjoint method**



For LKB photon box, simultaneous estimation of 32 parameters in measurement matrices

- results on **MaxLike estimation variance** when $\hat{\rho}_{ML}$ **is not invertible**
- link with the Bayesian Mean Estimation