Quantum state tomography including measurement duration, imperfections and decoherence

PRACQSYS, 20-25 July 2015, Sydney.

Pierre Rouchon
QUANTIC Research Team (INRIA/ENS/Mines)
Mines ParisTech, PSL Research University

Outline

Quantum state tomography based on POVM $\sum_j \pi_j = \boldsymbol{I}$

- Tomography of ρ via N independent measurements $\boldsymbol{Y}$ associated to POVM: probability $\text{Tr}\left(\rho\pi_j\right)$ of each measurement outcome j given by π_j; for $\boldsymbol{N_j}$ the number of j outcomes, $\boldsymbol{Y} \equiv (\boldsymbol{N_j})$ with $\sum_j \boldsymbol{N_j} = N$, the number of measurements.
- Several estimation methods:

  MaxEnt: ρ_{ME} maximizes $-\text{Tr}\left(\rho\log(\rho)\right)$ under the constraints $|\text{Tr}\left(\rho\pi_j\right) - \boldsymbol{N_j}/N| \leq \epsilon$ (Bužek et al, Ann. Phys. 1996).

  Compress Sensing: ρ_{CS} minimizes $\text{Tr}\left(\rho\right)$ under the constraints $|\text{Tr}\left(\rho\pi_j\right) - \boldsymbol{N_j}/N| \leq \epsilon$ (Gross et al PRL2010)

  MaxLike: ρ_{ML} maximizes the likelihood function, $\rho \mapsto \mathbb{P}(\boldsymbol{Y} \mid \rho) = \prod_j \left(\text{Tr}\left(\rho\pi_j\right)\right)^{\boldsymbol{N_j}}$ (see, e.g., Lvovsky/Raymer RMP 2009)

  Bayesian Mean: $\rho_{BM} \propto \int \rho\mathbb{P}(\boldsymbol{Y} \mid \rho)\mathbb{P}_0(\rho)d\rho$ where $\mathbb{P}_0$ is some prior distribution $\mathbb{P}_0(\rho)d\rho$ (see, e.g., Blume-Kohout NJP2010).

  Low rank, high dimensional systems: see, e.g, key contributions of Robert Kosut and also of Madalin Guta.

Filtering: estimation of the quantum state ρ_t at time $t > 0$ from the measurement trajectory $[0, t[\ni \tau \mapsto \boldsymbol{y_\tau}$ and the initial state ρ_0; see Belavkin semilar contributions (links with Monte-Carlo quantum-trajectories).

State tomography: estimation of the initial state $\boldsymbol{\rho = \rho_0}$ from a collection of N measurement trajectories: $\boldsymbol{Y} = \left(\boldsymbol{y_t^{(n)}}\right)$ with $n \in \{1, \ldots, N\}$ and $t \in [0, T]$.

Process tomography: estimation of a parameter p from a known initial state ρ_0 and a collection of N measurement trajectories $\boldsymbol{Y}$.

Focus on quantum state tomography: decoherence, exp. imperfections during the measurement duration T can be included via the adjoint state E already introduced in quantum smoothing [1]
Talk contribution: how to compute the likelihood function $\mathbb{P}\left(\boldsymbol{Y}/\rho_0\right)$ from the stochastic master equation governing filtering.

[1] Tsang PRL 2009, Gammelmark/Julsgaard/Mølmer PRL 2013, Guevara/Wiseman 2015...

Four features:

1. Bayes law: $\mathbb{P}(\mu'/\mu) = \mathbb{P}(\mu/\mu')\mathbb{P}(\mu') \;/\; \left(\sum_{\nu'} \mathbb{P}(\mu/\nu')\mathbb{P}(\nu')\right)$,
2. Schrödinger equations defining unitary transformations.
3. Randomness, irreversibility and dissipation induced by the measurement of observables with degenerate spectra.
4. Entanglement and tensor product for composite systems.

$\Rrightarrow$ **Discrete-time models**[2]

Take a set of operators $\boldsymbol{M}_\mu$ satisfying $\sum_\mu \boldsymbol{M}_\mu^\dagger \boldsymbol{M}_\mu = \boldsymbol{I}$ and a left stochastic matrices $(\eta_{\boldsymbol{y_t},\mu})$. Consider the following Markov process of state ρ (density op.) and measured output $\boldsymbol{y}$:

$$\boldsymbol{\rho_{t+1}} = \frac{\boldsymbol{K_{y_t}(\rho_t)}}{\mathrm{Tr}\big(\boldsymbol{K_{y_t}(\rho_t)}\big)}, \text{ with proba. } \mathbb{P}_{\boldsymbol{y_t}}(\rho_t) = \mathrm{Tr}\left(\boldsymbol{K}_{\boldsymbol{y_t}}(\rho_t)\right)$$

with $\boldsymbol{K}_{\boldsymbol{y}}(\rho) = \sum_{\mu=1}^{m} \eta_{\boldsymbol{y},\mu} \boldsymbol{M}_\mu \rho \boldsymbol{M}_\mu^\dagger$. It is associated to the Kraus map (ensemble average, quantum channel)

$$\mathbb{E}\left(\rho_{t+1}|\rho_t\right) = \boldsymbol{K}(\rho_t) = \sum_{\boldsymbol{y}} \boldsymbol{K}_{\boldsymbol{y}}(\rho_t) = \sum_{\mu} \boldsymbol{M}_\mu \rho_t \boldsymbol{M}_\mu^\dagger.$$

[2] see, e.g., the book of Haroche/Raimond and the publications around the LKB photon box.

Continuous/discrete-time Stochastic Master Equation (SME)

Discrete-time models: Markov chains $\rho_{t+1} = \frac{\boldsymbol{K}_{\boldsymbol{y_t}}(\rho_t)}{\text{Tr}\left(\boldsymbol{K}_{\boldsymbol{y_t}}(\rho_t)\right)}$, with

$\boldsymbol{K}_{\boldsymbol{y_t}}(\rho_t) = \sum_{\mu=1}^{m} \eta_{\boldsymbol{y_t},\mu} \boldsymbol{M}_\mu \rho_t \boldsymbol{M}_\mu^\dagger$, and proba. $\mathbb{P}_{\boldsymbol{y_t}}(\rho_t) = \text{Tr}\left(\boldsymbol{K}_{\boldsymbol{y_t}}(\rho_t)\right)$.
Ensemble averages correspond to Kraus linear maps

$$\mathbb{E}\left(\rho_{t+1}|\rho_t\right) = \boldsymbol{K}(\rho_t) = \sum_{\boldsymbol{y}} \boldsymbol{K}_{\boldsymbol{y}}(\rho_t) = \sum_{\mu} \boldsymbol{M}_\mu \rho_t \boldsymbol{M}_\mu^\dagger \quad \text{with} \quad \sum_{\mu} \boldsymbol{M}_\mu^\dagger \boldsymbol{M}_\mu = \boldsymbol{I}$$

Continuous-time models: stochastic differential systems (see, e.g., Barchielli/Gregoratti, 2009)

$$d\rho_t = \left(-\tfrac{i}{\hbar}[\boldsymbol{H}, \rho_t] + \sum_\nu \boldsymbol{L}_\nu \rho_t \boldsymbol{L}_\nu^\dagger - \tfrac{1}{2}(\boldsymbol{L}_\nu^\dagger \boldsymbol{L}_\nu \rho_t + \rho_t \boldsymbol{L}_\nu^\dagger \boldsymbol{L}_\nu)\right) dt \\ + \sum_\nu \sqrt{\eta_\nu}\left(\boldsymbol{L}_\nu \rho_t + \rho_t \boldsymbol{L}_\nu^\dagger - \text{Tr}\left((\boldsymbol{L}_\nu + \boldsymbol{L}_\nu^\dagger)\rho_t\right)\rho_t\right) dW_{\nu,t}$$

driven by Wiener processes $dW_{\nu,t}$, with measurements $\boldsymbol{dy_{\nu,t}}$,
$\boldsymbol{dy_{\nu,t}} = \sqrt{\eta_\nu}\, \text{Tr}\left((\boldsymbol{L}_\nu + \boldsymbol{L}_\nu^\dagger)\,\rho_t\right) dt + dW_{\nu,t}$, detection efficiencies $\eta_\nu \in [0, 1]$ and Lindblad-Kossakowski master equations ($\eta_\nu \equiv 0$):

$$\frac{d}{dt}\rho = -\tfrac{i}{\hbar}[\boldsymbol{H}, \rho] + \sum_\nu \boldsymbol{L}_\nu \rho \boldsymbol{L}_\nu^\dagger - \tfrac{1}{2}(\boldsymbol{L}_\nu^\dagger \boldsymbol{L}_\nu \rho + \rho \boldsymbol{L}_\nu^\dagger \boldsymbol{L}_\nu)$$

The Belavkin quantum filter

$$d\rho_t = \left(-\frac{i}{\hbar}[\boldsymbol{H}, \rho_t] + \sum_\nu \boldsymbol{L}_\nu \rho_t \boldsymbol{L}_\nu^\dagger - \frac{1}{2}(\boldsymbol{L}_\nu^\dagger \boldsymbol{L}_\nu \rho_t + \rho_t \boldsymbol{L}_\nu^\dagger \boldsymbol{L}_\nu) \right) dt + \sum_\nu \sqrt{\eta_\nu} \left(\boldsymbol{L}_\nu \rho_t + \rho_t \boldsymbol{L}_\nu^\dagger - \operatorname{Tr}\left((\boldsymbol{L}_\nu + \boldsymbol{L}_\nu^\dagger)\rho_t \right) \rho_t \right) \boldsymbol{dW_{\nu,t}}$$

with $\boldsymbol{dW_{\nu,t}} = \boldsymbol{dy_{\nu,t}} - \sqrt{\eta_\nu}\operatorname{Tr}\left((\boldsymbol{L}_\nu + \boldsymbol{L}_\nu^\dagger)\,\rho_t \right) dt$ given by the measurement signal $\boldsymbol{dy_{\nu,t}}$, is always a stable filtering process.[3]
Using Itō rules, it can be written as a "discrete-time" Markov model[4]

$$\rho_{t+dt} = \boldsymbol{K}_{\boldsymbol{dy_t}}(\rho_t) / \operatorname{Tr}(\boldsymbol{K}_{\boldsymbol{dy_t}}(\rho_t))$$

with "partial Kraus maps"

$$\boldsymbol{K}_{\boldsymbol{dy_t}}(\rho_t) = \boldsymbol{M}_{\boldsymbol{dy_t}} \rho_t \boldsymbol{M}_{\boldsymbol{dy_t}}^\dagger + \textstyle\sum_\nu (1-\eta_\nu) \boldsymbol{L}_\nu \rho_t \boldsymbol{L}_\nu^\dagger dt$$

$$\boldsymbol{M}_{\boldsymbol{dy_t}} = \boldsymbol{I} + \left(-\frac{i}{\hbar}\boldsymbol{H} - \frac{1}{2}\left(\textstyle\sum_\nu \boldsymbol{L}_\nu^\dagger \boldsymbol{L}_\nu \right) \right) dt + \textstyle\sum_\nu \sqrt{\eta_\nu}\, \boldsymbol{dy_{\nu,t}} \boldsymbol{L}$$

where the probability of outcome $\boldsymbol{dy_t} = (dy_{\nu,t})$ reads:

$$\mathbb{P}\left(\boldsymbol{dy_t} \in \textstyle\prod_\nu [\xi_\nu, \xi_\nu + d\xi_\nu] \;/\; \rho_t \right) = \operatorname{Tr}(\boldsymbol{K}_{\boldsymbol{\xi}}(\rho_t)) \textstyle\prod_\nu e^{-\xi_\nu^2/2dt} \frac{d\xi_\nu}{\sqrt{2\pi dt}}$$

[3] H. Amini et al., Russian J. of Math. Physics, 2014, 21, 297-315.

[4] PR, J. Ralph PRA2015; see also PR Int. Congress of Mathematicians at Seoul 2014, and PhD of Ph. Campagne-Ibracq at ENS-Paris, 2015.

- Denote by $\mathbb{P}_n(\rho)$ the probability of getting measurement trajectory n, $(\boldsymbol{y_t^{(n)}})_{t=0,\dots,T}$, knowing the initial state $\rho_0^{(n)} = \rho$.
- Since $\rho_{t+1}^{(n)} = \frac{\boldsymbol{K}_{\boldsymbol{y_t^{(n)}}}\left(\rho_t^{(n)}\right)}{\text{Tr}\left(\boldsymbol{K}_{\boldsymbol{y_t^{(n)}}}\left(\rho_t^{(n)}\right)\right)}$ with $\text{Tr}\left(\boldsymbol{K}_{\boldsymbol{y_t^{(n)}}}\left(\rho_t^{(n)}\right)\right)$ the probability of having detected $\boldsymbol{y_t^{(n)}}$ knowing $\rho_t^{(n)}$, a direct use of Bayes law yields $\mathbb{P}_n(\rho) = \prod_{t=0}^{T} \text{Tr}\left(\boldsymbol{K}_{\boldsymbol{y_t^{(n)}}}\left(\rho_t^{(n)}\right)\right)$. Some elementary computations show that:

$$\mathbb{P}_n(\rho) = \text{Tr}\left(\boldsymbol{K}_{\boldsymbol{y_T^{(n)}}} \circ \ldots \circ \boldsymbol{K}_{\boldsymbol{y_0^{(n)}}}(\rho)\right).$$

- The adjoint map $\boldsymbol{K}_{\boldsymbol{y}}^*$ of $\boldsymbol{K}_{\boldsymbol{y}}$ is defined by $\text{Tr}\left(A\boldsymbol{K}_{\boldsymbol{y}}(B)\right) \equiv \text{Tr}\left(\boldsymbol{K}_{\boldsymbol{y}}^*(A)B\right)$ for all Hermitian operators A and B. Thus

$$\mathbb{P}_n(\rho) = \text{Tr}\left(\boldsymbol{K}_{\boldsymbol{y_T^{(n)}}} \circ \ldots \circ \boldsymbol{K}_{\boldsymbol{y_0^{(n)}}}(\rho) \quad \boldsymbol{I}\right) = \text{Tr}\left(\rho\, \boldsymbol{K}_{\boldsymbol{y_0^{(n)}}}^* \circ \ldots \circ \boldsymbol{K}_{\boldsymbol{y_T^{(n)}}}^*(\boldsymbol{I})\right).$$

Computation of the likelihood function via the adjoint state (2)

- The normalized adjoint quantum filter, $E_t^{(n)} = \frac{\boldsymbol{K}^*_{\boldsymbol{y}_t^{(n)}}\left(E_{t+1}^{(n)}\right)}{\operatorname{Tr}\left(\boldsymbol{K}^*_{\boldsymbol{y}_t^{(n)}}\left(E_{t+1}^{(n)}\right)\right)}$ with $E_{T+1}^{(n)} = \boldsymbol{I}/\operatorname{Tr}(\boldsymbol{I})$, defines a family of Hermitian and non-negative operators $(E_t^{(n)})$ on unit trace depending only on the measurement data $\boldsymbol{Y}$.
- We have $\boldsymbol{K}^*_{\boldsymbol{y}_0^{(n)}} \circ \ldots \circ \boldsymbol{K}^*_{\boldsymbol{y}_T^{(n)}}(\boldsymbol{I}) = g_n(\boldsymbol{Y}) E_0^{(n)}$ with $\frac{1}{g_n(\boldsymbol{Y})} = \operatorname{Tr}\left(\boldsymbol{K}^*_{\boldsymbol{y}_0^{(n)}} \circ \ldots \circ \boldsymbol{K}^*_{\boldsymbol{y}_T^{(n)}}(\boldsymbol{I})\right)$ independent of ρ.
- Thus $\mathbb{P}_n(\rho) = g_n(\boldsymbol{Y}) \operatorname{Tr}\left(\rho E_0^{(n)}\right)$ and

$$\mathbb{P}(\boldsymbol{Y}/\rho) = g(\boldsymbol{Y}) \prod_{n=1}^{N} \operatorname{Tr}\left(\rho E_0^{(n)}\right)$$

where $g(\boldsymbol{Y}) = \prod_{n=1}^{N} g_n(\boldsymbol{Y})$.

MaxLike quantum-state tomography revisited

- MaxLike tomography based on POVM π_j: ρ_{ML} maximizes

$$\mathbb{P}(\boldsymbol{Y} \mid \rho) = \prod_j \left(\operatorname{Tr}(\rho \pi_j) \right)^{\boldsymbol{N_j}} = \prod_n \operatorname{Tr}\left(\rho \pi_{\boldsymbol{j_n}}\right)$$

  with $\boldsymbol{Y} \equiv (\boldsymbol{N_j})$ derived from $\boldsymbol{j_n}$, the measurement outcome number $n = 1, \ldots, N$.

- MaxLike tomography based on the adjoint states: ρ_{ML} maximizes

$$\mathbb{P}(\boldsymbol{Y} \mid \rho) = g(\boldsymbol{Y}) \prod_n \operatorname{Tr}\left(\rho E^{(n)}\right)$$

  where $E^{(n)} = E_0^{(n)}$ is the adjoint state at $t = 0$ associated to measurement trajectory $(\boldsymbol{y_t^{(n)}})$ number n.

Convex optimization problem: the set $\mathcal{D}$ of density operators is convex; the log-likelihood function $f : \mathcal{D} \ni \rho \mapsto \log\left(\mathbb{P}(\boldsymbol{Y} \mid \rho)\right)$ is concave (see Robert Kosut talk ...)

Gradient and Hessian of the log-likelihood function

We have

$$f(\rho) \triangleq \log\left(\mathbb{P}(\boldsymbol{Y} \mid \rho)\right) = \log(g(\boldsymbol{Y})) + \sum_{n=1}^{N} \log\left(\operatorname{Tr}\left(\rho E^{(n)}\right)\right).$$

The gradient of f,

$$\nabla f_\rho = \sum_{n=1}^{N} \frac{E^{(n)}}{\operatorname{Tr}\left(E^{(n)}\rho\right)}.$$

and its Hessian $\nabla^2 f$ (self-adjoint super-operator)

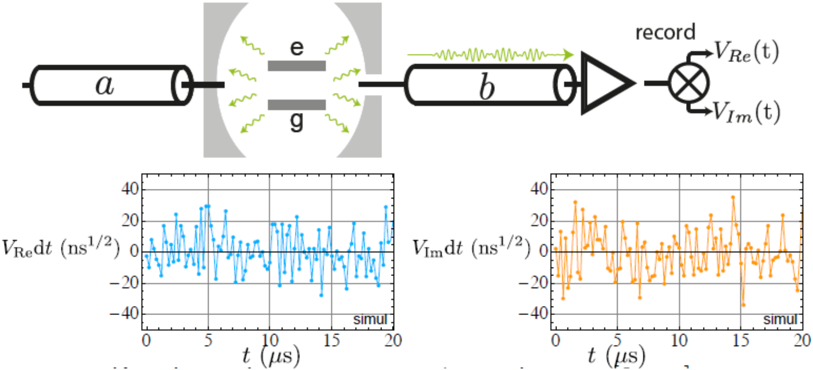
$$\xi \mapsto \nabla^2 f_\rho(\xi) = -\sum_{n=1}^{N} \frac{\operatorname{Tr}\left(E^{(n)}\xi\right)}{\operatorname{Tr}^2\left(E^{(n)}\rho\right)} E^{(n)}.$$

result from the following second order expansion:

$$f(\rho + \delta\rho) - f(\rho)$$
$$= \sum_{n=1}^{N} \left(\frac{\operatorname{Tr}\left(E^{(n)}\delta\rho\right)}{\operatorname{Tr}\left(E^{(n)}\rho\right)} - \frac{1}{2} \frac{\operatorname{Tr}^2\left(E^{(n)}\delta\rho\right)}{\operatorname{Tr}^2\left(E^{(n)}\rho\right)} \right) + o\left(\operatorname{Tr}\left(\delta\rho^2\right) \right)$$

.

Circuit QED: the LPA qubit with fluorescence measurements[5]



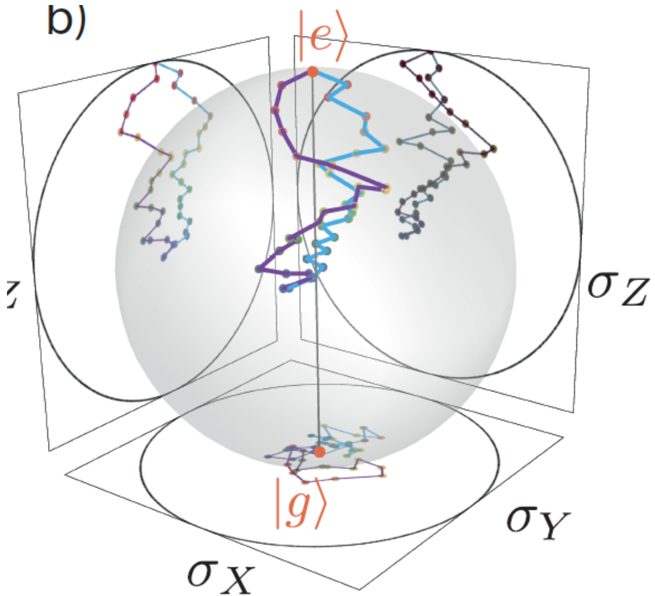
No drive $\boldsymbol{H} = 0$, $\nu \in \{1, 2, 3\}$

Two fluorescence measurements $\boldsymbol{L}_1 = \sqrt{\frac{1}{2T_1}}\boldsymbol{\sigma}_{\text{-}}$ and $\boldsymbol{L}_2 = i\boldsymbol{L}_1$ with $T_1 = 4\ \mu s$ and efficiencies $\eta_1 = \eta_2 \approx 1/4$.

Dephasing channel $\boldsymbol{L}_3 = \sqrt{\frac{1}{2T_\phi}}\boldsymbol{\sigma_z}$ with $T_\phi = 35\ \mu s$ $(\eta_3 = 0)$.

[5] Ph. Campagne-Ibracq et al., Phys. Rev. Lett., 2014, 112, 180402.

Two quantum filtering trajectories (experimental data)



MaxLike tomography from $\rho \approx (\boldsymbol{I} + \boldsymbol{\sigma_z})/2$ (experimental data)

- $N = 3000$ trajectories of length $T = \frac{5}{2} T_1$, with $dt = \frac{1}{20} T_1$
- Two measurements of efficiency $\eta = \frac{1}{4}$:

$$dy_1 = \sqrt{\frac{\eta}{2T_1}} \,\mathrm{Tr}\left(\rho \boldsymbol{\sigma_x}\right) + dW_1, \quad dy_2 = \sqrt{\frac{\eta}{2T_1}} \,\mathrm{Tr}\left(\rho \boldsymbol{\sigma_y}\right) + dW_2$$

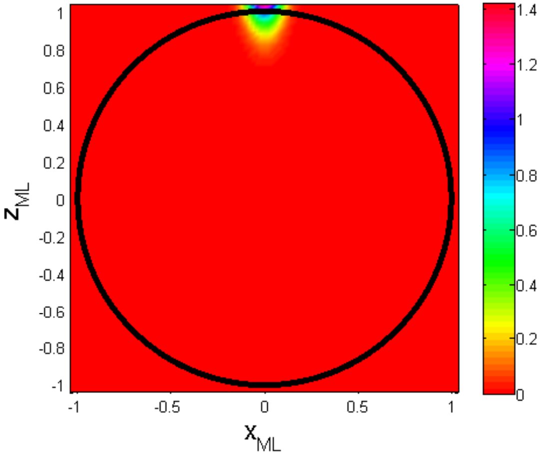
  For each trajectory, the data corresponds to 2×50 real values ($dt = 200\ ns$).
- The resulting ρ_{ML}

$$x_{ML} = 0.0134, \quad y_{ML} = 0.0213, \quad z_{ML} = 0.9997$$

  is pure since it satisfies $x_{ML}^2 + y_{ML}^2 + z_{ML}^2 = 1$.
- Gradient and Hessian of the log-likelihood function f

$$\nabla f_{\rho_{ML}} = \begin{pmatrix} 0.12 \\ 0.20 \\ 9.22 \end{pmatrix}, \quad \nabla^2 f_{\rho_{ML}} = \begin{pmatrix} -241.1 & 3.6 & 0.4 \\ 3.6 & -235.7 & 1.4 \\ 0.4 & 1.4 & -23.5 \end{pmatrix}$$

Experimental likelihood function corresponding to $\rho \approx (\boldsymbol{I} + \boldsymbol{\sigma_z})/2$



$N = 3000$ fluorescence trajectories of length $2T_1$.
Cross section passing through the center of Bloch sphere
z_{ML}-axis aligned with ρ_{ML}, close to z-axis, x_{ML}-axis close to x-axis.

MaxLike tomography from $\rho \approx (\boldsymbol{I} + \frac{2}{3}\boldsymbol{\sigma_z})/2$ (experimental data)

- $N = 3000$ trajectories of length $T = 2T_1$, with $dt = \frac{1}{20}T_1$.
- Two measurements of efficiency $\eta = \frac{1}{4}$:

$$dy_1 = \sqrt{\frac{\eta}{2T_1}}\,\mathrm{Tr}\left(\rho\boldsymbol{\sigma_x}\right) + dW_1, \quad dy_2 = \sqrt{\frac{\eta}{2T_1}}\,\mathrm{Tr}\left(\rho\boldsymbol{\sigma_y}\right) + dW_2$$

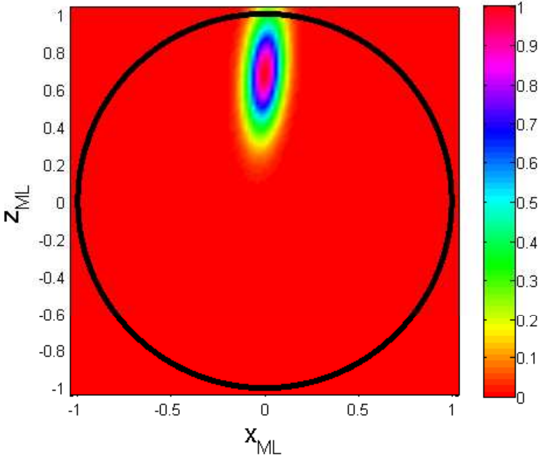
  For each trajectory, the data corresponds to 2×40 real values ($dt = 200$ ns).
- The resulting ρ_{ML}

$$x_{ML} = -0.0465, \quad y_{ML} = 0.0625, \quad z_{ML} = 0.6787$$

  is mixed since it satisfies $x_{ML}^2 + y_{ML}^2 + z_{ML}^2 < 1$.
- Gradient and Hessian of the log-likelihood function f:

$$\nabla f_{\rho_{ML}} = 0, \quad \nabla^2 f_{\rho_{ML}} = \begin{pmatrix} -231.6 & -2.6 & 1.7 \\ -2.6 & -227.5 & -0.4 \\ 1.7 & -0.4 & -21.6 \end{pmatrix}$$

Experimental likelihood function corresponding to $\rho \approx (\boldsymbol{I} + \frac{2}{3}\boldsymbol{\sigma_z})/2$



$N = 3000$ fluorescence trajectories of length $\frac{3}{2}T_1$.
Cross section passing through the center of Bloch sphere
z_{ML}-axis aligned with ρ_{ML}, close to z-axis, x_{ML}-axis close to x-axis.

- Another possible estimation is given by $\rho_{BM} \propto \int \rho \mathbb{P}(\boldsymbol{Y} \mid \rho)\mathbb{P}_0(\rho)d\rho$ with some prior distribution $\mathbb{P}_0(\rho)d\rho$.
- With Bloch variables (x, y, z) and $\mathbb{P}_0(\rho)d\rho \propto dxdydz$ we have,

$$x_{BM} = \frac{\iiint_{x^2+y^2+z^2 \le 1} x e^{f(x,y,z)} dxdydz}{\iiint_{x^2+y^2+z^2 \le 1} e^{f(x,y,z)} dxdydz}, \quad y_{BM} = \ldots$$

- With the normalization $f = \bar{N}\bar{f}$ with $\bar{N} > 0$ large , we have approximation of x_{BM} via the asymptotics

$$\iiint_{x^2+y^2+z^2 \le 1} g(x,y,z) e^{\bar{N}\bar{f}(x,y,z)} dxdydz = \frac{e^{\bar{N}\bar{f}(x_{ML},y_{ML},z_{ML})}}{\bar{N}^2} \left(c_0 + \frac{c_1}{\bar{N}} + \frac{c_2}{\bar{N}^2} + \ldots \right)$$

where c_0, c_1, c_2 ...depend on the derivatives of $\bar{f}$ and g at (x_{ML}, y_{ML}, z_{ML}) [6]

[6] For an elementary theory see Bleistein, N. Handelsman, R. : Asymptotic Expansions of Integrals. Dover, 1986.
For the general recent theory called Singular Learning see Watanabe S., Algebraic Geometry and Statistical Learning Theory, Cambridge University Press, 2009.
See also Shaowei Lin, Algebraic Methods for Evaluating Integrals Bayesian Statistics, PhD thesis, University of California, Berkeley, 2011.

1. Another validation on the experimental data of the LKB photon box underlying Rybarczyk et al: Past quantum state analysis of the photon number evolution in a cavity, to appear in PRA.

   Compensation of photon life-time $1/\kappa$ comparable with the time during the QND measurement of photons.

2. In the near future: application to Wigner tomography of a cavity field based on the measurement protocol used, e.g., in Leghtas et al.: Confining the state of light to a quantum manifold by engineered two-photon loss; Science, 2015, 347, 853-857.

   Compensation for initial thermal state of the probe qubit, qubit measurement errors, cavity field damping and several nonlinear Kerr effects.

3. Extension to parameter estimation (quantum process tomography) where the adjoint state simplifies the gradient computation of the log-likelihood function.

4. Correction to low-rank ρ_{ML} via $\rho_{BM} \propto \int \rho \mathbb{P}(\boldsymbol{Y} \mid \rho)\mathbb{P}_0(\rho) d\rho$ and its approximate computation via asymptotics techniques developed for **multidimensional integrals of Laplace type**.