Quantum tomography based on quantum trajectories

Quantum Control Theory:
Mathematical Aspects and Physical Applications
TUM-IAS, Garching, April 3-5, 2017

Pierre Rouchon
Centre Automatique et Systèmes, Mines ParisTech, PSL Research University
Quantic Research Team, Inria

Quantum tomography versus quantum filtering

Likelihood function calculations via adjoint states

- Discrete time case
- Continuous time case

MaxLike estimations with experimental data

- Process tomography for QND measurement of photons
- State tomography of a quantum Maxwell demon

Fisher information and low-rank MaxLike estimates

Appendix: asymptotics for multi-dimension Laplace integrals and boundary corrections

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 $\text{Tr}\,(\rho\pi_j)$ of each measurement outcome j given by π_j; for $\boldsymbol{N_j}$ the number of j outcomes, $\boldsymbol{Y \equiv (N_j)}$ with $\sum_j \boldsymbol{N_j} = N$, the number of measurements.
- Several estimation methods:

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

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

MaxLike: ρ_{ML} maximizes the likelihood function, $\rho \mapsto \mathbb{P}(\boldsymbol{Y} \mid \rho) = \prod_j \left(\text{Tr}\,(\rho\pi_j)\right)^{\boldsymbol{N_j}}$ (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, PhD thesis "Efficient and Robust Methods for Quantum Tomography" of Charles Heber Baldwin, University of New Mexico, December 2016.

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 $\rho_0 = \boldsymbol{\rho}$ 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 $\mathbf{p}$ from a known initial state $\boldsymbol{\rho}$ and a collection of N measurement trajectories $\boldsymbol{Y}$.

This talk: MaxLike estimation with decoherence and measurement imperfections (PhD thesis of Pierre Six, November 2016):

1. How to compute the likelihood function $\mathbb{P}\big(\boldsymbol{Y}/\boldsymbol{\rho}, \mathbf{p}\big)$ and its gradient from the stochastic master equation governing filtering (P. Six et al. PRA 2016).
2. For state estimation: variance computation based on asymptotic expansions of Laplace integrals for low rank MaxLike estimates (P. Six /PR, chapter in Lecture Notes in Control and Information Sciences no 473, April 2017).

Log-likelihood function $f(p) = \log\left(\mathbb{P}(\boldsymbol{Y} \mid p)\right)$ admits a unique maximum at p_{ML} ($\nabla f(p_{ML}) = 0$) with a negative definite Hessian ($\nabla^2 f(p_{ML}) < 0$).

f coming from N independent realisations: $f(p) \equiv N\bar{f}(p)$ with asymptotics for $N \mapsto +\infty$ of the Laplace integrals connecting

- Bayesian Mean p_{BM} and MaxLike estimation p_{ML}:

$$p_{BM} = \frac{\int p\, e^{N\bar{f}(p)} \mathbb{P}_0(p) dp}{\int e^{N\bar{f}(p)} \mathbb{P}_0(p) dp} = p_{ML} + O(1/N).$$

  with any smooth prior distribution $\mathbb{P}_0(p)dp$

- Bayesian variance and Fisher information $\overline{F}_{ML} = -\nabla^2 \bar{f}(p_{ML})$:

$$\frac{\int \|p - p_{ML}\|^2\, e^{N\bar{f}(p)} \mathbb{P}_0(p) dp}{\int e^{N\bar{f}(p)} \mathbb{P}_0(p) dp} = \text{Tr}\left(\left(\overline{F}_{ML}\right)^{-1}\right)/(2N) + O(1/N^2).$$

Confidence intervals based on $-\nabla^2 f(p_{ML})$.

Outline

Quantum tomography versus quantum filtering

Likelihood function calculations via adjoint states
- Discrete time case
- Continuous time case

MaxLike estimations with experimental data
- Process tomography for QND measurement of photons
- State tomography of a quantum Maxwell demon

Fisher information and low-rank MaxLike estimates

Appendix: asymptotics for multi-dimension Laplace integrals and boundary corrections

Four features[1]:

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**

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 $\boldsymbol{y}$:

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

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)

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

[1] See the book of S. Haroche and J.M. Raimond.

Computation of the likelihood function via the adjoint state (1)

- Denote by $\mathbb{P}_n(\boldsymbol{\rho}, \mathbf{p})$ the probability of getting measurement trajectory n, $(\boldsymbol{y}_t^{(n)})_{t=0,\ldots,T}$, knowing the initial state $\rho_0^{(n)} = \boldsymbol{\rho}$ and parameter $\mathbf{p}$.

- Since $\rho_{t+1}^{(n)} = \frac{\boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_t^{(n)}}\left(\rho_t^{(n)}\right)}{\operatorname{Tr}\left(\boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_t^{(n)}}\left(\rho_t^{(n)}\right)\right)}$ with $\operatorname{Tr}\left(\boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_t^{(n)}}\left(\rho_t^{(n)}\right)\right)$ the probability of having detected $\boldsymbol{y}_t^{(n)}$ knowing $\rho_t^{(n)}$ and $\mathbf{p}$, a direct use of Bayes law yields

$$\mathbb{P}_n(\boldsymbol{\rho}, \mathbf{p}) = \prod_{t=0}^{T} \operatorname{Tr}\left(\boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_t^{(n)}}\left(\rho_t^{(n)}\right)\right) = \operatorname{Tr}\left(\boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_T^{(n)}} \circ \ldots \circ \boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_0^{(n)}}(\boldsymbol{\rho})\right).$$

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

- With adjoint map $\boldsymbol{K}_{\boldsymbol{y}}^{\mathbf{p}*}$ ($\forall A, B,\ \mathrm{Tr}\left(\boldsymbol{K}_{\boldsymbol{y}}^{\mathbf{p}}(A)B\right) \equiv \mathrm{Tr}\left(A\boldsymbol{K}_{\boldsymbol{y}}^{\mathbf{p}*}(B)\right)$):

$$\mathbb{P}_n(\boldsymbol{\rho}, \mathbf{p}) = \mathrm{Tr}\left(\boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_T^{(n)}} \circ \ldots \circ \boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_0^{(n)}}(\boldsymbol{\rho}) \quad \boldsymbol{I}\right) = \mathrm{Tr}\left(\boldsymbol{\rho} \quad \boldsymbol{K}^{\mathbf{p}*}_{\boldsymbol{y}_0^{(n)}} \circ \ldots \circ \boldsymbol{K}^{\mathbf{p}*}_{\boldsymbol{y}_T^{(n)}}(\boldsymbol{I})\right).$$

- Normalized adjoint quantum filter[2] $E_t^{(n)} = \dfrac{\boldsymbol{K}^{\mathbf{p}*}_{\boldsymbol{y}_t^{(n)}}\left(E_{t+1}^{(n)}\right)}{\mathrm{Tr}\left(\boldsymbol{K}^{\mathbf{p}*}_{\boldsymbol{y}_t^{(n)}}\left(E_{t+1}^{(n)}\right)\right)}$ with

  $E_{T+1}^{(n)} = \boldsymbol{I}/\ \mathrm{Tr}(\boldsymbol{I})$, we get

$$\mathbb{P}_n(\boldsymbol{\rho}, \mathbf{p}) = \prod_{t=T}^{0} \mathrm{Tr}\left(\boldsymbol{K}^{\mathbf{p}*}_{\boldsymbol{y}_t^{(n)}}\left(E_{t+1}^{(n)}\right)\right) \mathrm{Tr}\left(\boldsymbol{\rho} E_0^{(n)}\right) \triangleq g_n(\boldsymbol{Y}, \mathbf{p})\ \mathrm{Tr}\left(\boldsymbol{\rho} E_0^{(n)}\right).$$

- A simple expression of the gradients:

$$\nabla_{\boldsymbol{\rho}} \log \mathbb{P}_n = \frac{E_0^{(n)}}{\mathrm{Tr}\left(\boldsymbol{\rho} E_0^{(n)}\right)}, \quad \nabla_{\mathbf{p}} \log \mathbb{P}_n \cdot \delta\mathbf{p} = \sum_{t=0}^{T} \frac{\mathrm{Tr}\left(E_{t+1}^{(n)}\ \nabla_{\mathbf{p}} \boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_t^{(n)}}\left(\rho_t^{(n)}\right) \cdot \delta\mathbf{p}\right)}{\mathrm{Tr}\left(E_{t+1}^{(n)}\ \boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y}_t^{(n)}}\left(\rho_t^{(n)}\right)\right)},$$

[2] M. Tsang. Time-symmetric quantum theory of smoothing. PRL 2009.

MaxLike tomography based on N trajectories data $\boldsymbol{Y} = \left(\boldsymbol{y_t^{(n)}}\right)$

From $\mathbb{P}_n(\boldsymbol{\rho}, \mathbf{p}) = g_n(\boldsymbol{Y}, \mathbf{p}) \operatorname{Tr}\left(\boldsymbol{\rho} E_0^{(n)}\right)$ we have

$$\mathbb{P}(\boldsymbol{\rho}, \mathbf{p}) \triangleq \prod_{n=1}^{N} \mathbb{P}_n(\boldsymbol{\rho}, \mathbf{p}) = \left(\prod_{n=1}^{N} g_n(\boldsymbol{Y}, \mathbf{p})\right) \left(\prod_{n=1}^{N} \operatorname{Tr}\left(\boldsymbol{\rho} E_0^{(n)}\right)\right).$$

- MaxLike **state tomography:** $\mathbf{p}$ is known and $\boldsymbol{\rho}_{ML}$ maximizes

$$\boldsymbol{\rho} \mapsto \sum_{n=1}^{N} \log\left(\operatorname{Tr}\left(\boldsymbol{\rho} E_0^{(n)}\right)\right)$$

  a concave function on the convex set of density operators ρ:
  a well structured convex optimization problem.

- MaxLike **process tomography:** $\boldsymbol{\rho}$ is known and $\mathbf{p}_{ML}$ maximizes $\mathbf{p} \mapsto f(\mathbf{p}) = \log \mathbb{P}(\boldsymbol{\rho}, \mathbf{p})$ those gradient is given by

$$\nabla_{\mathbf{p}} f(\mathbf{p}) \cdot \delta\mathbf{p} = \sum_{n=1}^{N} \sum_{t=0}^{T} \frac{\operatorname{Tr}\left(E_{t+1}^{(n)} \ \nabla_{\mathbf{p}} \boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y_t^{(n)}}}\left(\rho_t^{(n)}\right) \cdot \delta\mathbf{p}\right)}{\operatorname{Tr}\left(E_{t+1}^{(n)} \ \boldsymbol{K}^{\mathbf{p}}_{\boldsymbol{y_t^{(n)}}}\left(\rho_t^{(n)}\right)\right)},$$

  The Hessian $\nabla^2_{\mathbf{p}} f$ can be computed similarly (Fisher information).

Discrete-time models: Markov chains $\rho_{t+1} = \frac{K_{y_t}(\rho_t)}{\text{Tr}\left(K_{y_t}(\rho_t)\right)}$, with
$K_{y_t}(\rho_t) = \sum_{\mu=1}^{m} \eta_{y_t,\mu} M_\mu \rho_t M_\mu^\dagger$, and proba. $\mathbb{P}_{y_t}(\rho_t) = \text{Tr}\left(K_{y_t}(\rho_t)\right)$.
Ensemble averages correspond to Kraus linear maps

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

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

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

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

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

Continuous/discrete-time diffusive SME

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]

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

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

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

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

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

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

[4] PR, J. Ralph PRA2015; see also PhD thesis of Ph. Campagne-Ibracq (2015) and of P. Six (2016).

Outline

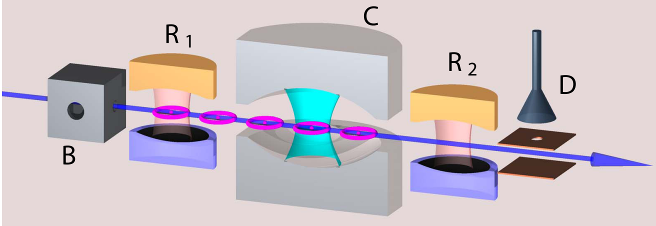
Quantum tomography versus quantum filtering

Likelihood function calculations via adjoint states
- Discrete time case
- Continuous time case

MaxLike estimations with experimental data
- Process tomography for QND measurement of photons
- State tomography of a quantum Maxwell demon

Fisher information and low-rank MaxLike estimates

Appendix: asymptotics for multi-dimension Laplace integrals and boundary corrections



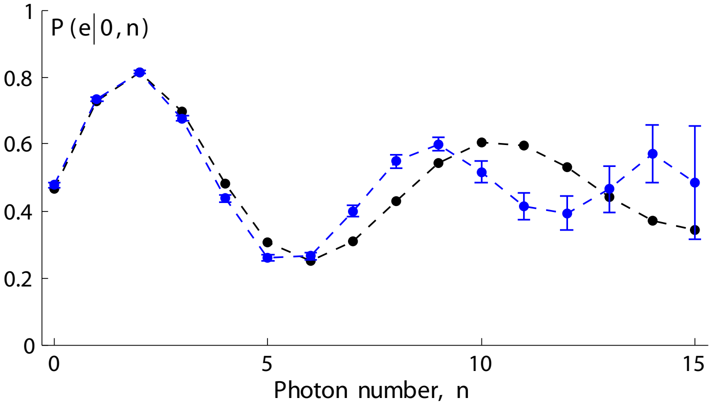
- The probability $\mathbb{P}(y \mid \phi_R, n)$ to get $y \in \{g, e\}$ knowing the Ramsey angle ϕ_R and the number of photon(s) $n \in \{0, 1, 2, \ldots\}$:

  $$\mathbb{P}(y \mid \phi_R, n) = 1 + \epsilon_y \big(A + B_c(n) \cos\phi_R + B_s(n) \sin\phi_R \big) \text{ with } \epsilon_{e/g} = \pm 1.$$

  depends on the parameters $\mathbf{p} = (B_c(n), B_s(n))_{n \in \{0,1,\ldots,\}}$.

- The Kraus maps $\boldsymbol{K}_y^{\mathbf{p}}$ based on known cavity decay and thermal photons.

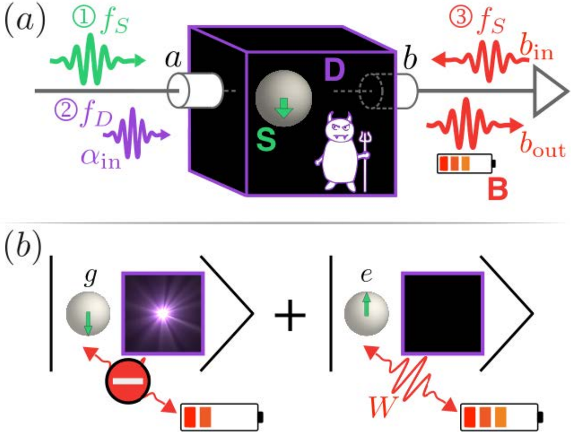
A priori calibration[5] (black dots) versus MaxLike (blue dots)



MaxLike estimation of 32 parameters **p** based on $N = 8000$ trajectories of $T = 6000$ outcome measurements.

[5] T. Rybarczyk, B. Peaudecerf, M. Penasa, S. Gerlich, B. Julsgaard, K. Mølmer, S. Gleyzes, M. Brune, J. M. Raimond, S. Haroche, and I. Dotsenko. Forward-backward analysis of the photon-number evolution in a cavity. PRA 2015.

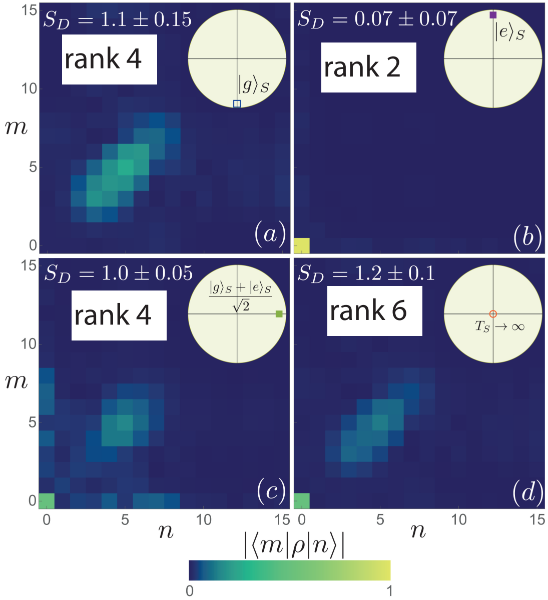
A quantum Maxwell demon experiment arXiv:1702.01917v1



(a) After preparation in a thermal or quantum state the system S (superconducting qubit) state is recorded into the demon's quantum memory D (microwave cavity) via a pulse that populates the cavity mode only if the qubit is in the ground state. This information is used to extract work which charges a battery with one extra photon: system S emits this photon only when the demon's cavity is empty. The memory reset is performed by cavity relaxation.

(b) When the system starts in a quantum superposition of the demon and system are entangled after the record step.

Tomography of the demon after the work extraction step



Computations are based on a truncation to 20 photons

How to define the confidence intervals for low rank ρ_{ML}?

Outline

Quantum tomography versus quantum filtering

Likelihood function calculations via adjoint states
- Discrete time case
- Continuous time case

MaxLike estimations with experimental data
- Process tomography for QND measurement of photons
- State tomography of a quantum Maxwell demon

Fisher information and low-rank MaxLike estimates

Appendix: asymptotics for multi-dimension Laplace integrals and boundary corrections

Asymptotic when ρ_{ML} is on the boundary

- To bypass boundary problem, consider Bayesian estimate instead of MaxLike ones

$$\rho_{BM} = \frac{\int \rho \, \mathbb{P}(\boldsymbol{Y} \mid \rho)\mathbb{P}_0(\rho)d\rho}{\int \mathbb{P}(\boldsymbol{Y} \mid \rho)\mathbb{P}_0(\rho)d\rho}$$

  with some prior distribution $\mathbb{P}_0(\rho)d\rho$.

- When the likelihood $\exp(f(\rho)) \equiv \mathbb{P}(\boldsymbol{Y} \mid \rho)$ is concentrated ($f = \boldsymbol{N}\bar{f}$ with $N \gg 1$) around its maximum ρ_{ML} that lies on the boundary (ρ_{ML} not full rank), how to compute the first terms of an asymptotic expansion versus $\boldsymbol{N}$ of

$$\int \mathrm{Tr}\,(\rho A)^r \exp(\boldsymbol{N}\bar{f}(\rho))\mathbb{P}_0(\rho)d\rho$$

  for any operator A and exponent r and for some prior distribution $\mathbb{P}_0(\rho)d\rho$ (e.g., Gausssian unitary ensemble).

- Since all functions are analytic such an asymptotic expansion versus $\boldsymbol{N}$ always exists: Integration by parts, Watson's lemma, Laplace's method, stationary phase, steepest descents, Hironaka's resolution of singularities [6], "singular learning" [7]

[6] An important reference: V.I. Arnold, S.M. Gusein-Zade, and A.N. Varchenko. Singularities of Differentiable Maps, Vol. II. Birkhäuser, Boston, 1985

[7] S. Watanabe: Algebraic Geometry and Statistical Learning Theory, Cambridge University Press, 2009.

Assume that ρ_{ML} is an argument of the maximum of

$$f : \mathcal{D} \ni \rho \mapsto \sum_{\mu \in \mathcal{M}} \log \left(\text{Tr} \left(\rho E^{(\mu)} \right) \right) \in [-\infty, 0]$$

over $\mathcal{D}$ (the set of density operators, $E^{(\mu)} \in \mathcal{D}$.). Then necessarily, ρ_{ML} satisfies the following conditions:

- $\text{Tr}\left(\rho_{ML} E^{(\mu)}\right) > 0$ for each $\mu \in \mathcal{M}$;
- $\left[\rho_{ML} \, , \, \nabla f|_{\rho_{ML}}\right] = 0$, where $\nabla f|_{\rho_{ML}} = \sum_{\mu \in \mathcal{M}} \frac{E^{(\mu)}}{\text{Tr}\left(\rho_{ML} E^{(\mu)}\right)}$ is the gradient of f at ρ_{ML} for the Frobenius scalar product;
- there exists $\lambda_{ML} > 0$ such that $\lambda_{ML} P_{ML} = P_{ML} \, \nabla f|_{\rho_{ML}}$ and $\nabla f|_{\rho_{ML}} \leq \lambda_{ML} I$, where P_{ML} is the orthogonal projector on the range of ρ_{ML} and I is the identity operator.

These conditions are also sufficient and characterize the unique maximum when, additionally, the vector space spanned by the $E^{(\mu)}$'s coincides with the set of Hermitian matrices.

For any Hermitian operator A, its Bayesian mean and variance read:

$$I_A(N) = \frac{\int_{\mathcal{D}} \mathrm{Tr}\,(\rho A)\, e^{Nf(\rho)}\, \mathbb{P}_0(\rho)\, d\rho}{\int_{\mathcal{D}} e^{Nf(\rho)}\, \mathbb{P}_0(\rho)\, d\rho}, \quad V_A(N) = \frac{\int_{\mathcal{D}} \left(\mathrm{Tr}\,(\rho A) - I_A(N)\right)^2 e^{Nf(\rho)}\, \mathbb{P}_0(\rho)\, d\rho}{\int_{\mathcal{D}} e^{Nf(\rho)}\, \mathbb{P}_0(\rho)\, d\rho}.$$

Denote by ρ_{ML} the unique maximum of f on $\mathcal{D}$ and by P_{ML} the orthogonal projector on its range. In addition to the necessary and sufficient geometric conditions above, assume that $\ker\left(\lambda_{ML} I - \nabla f|_{\rho_{ML}}\right) = \ker(I - P_{ML})$.

$$I_A(N) = \mathrm{Tr}\,(A\rho_{ML}) + O(1/N), \quad \boldsymbol{V_A(N) = \mathrm{Tr}\left(A_{\|}\ (F_{ML})^{-1}(A_{\|})\right)/N} + O(1/N^2)$$

where $B_{\|}$ is an orthogonal projection

$$B_{\|} = B - \frac{\mathrm{Tr}\,(BP_{ML})}{\mathrm{Tr}\,(P_{ML})} P_{ML} - (I - P_{ML})B(I - P_{ML});$$

and where $\boldsymbol{F_{ML}}$ is a linear super-operator, corresponds to the Hessian at ρ_{ML} of some restriction of f and **generalizes the Fisher information matrix**:

$$\boldsymbol{F_{ML}}(X) = \sum_{\mu} \frac{\mathrm{Tr}\left(XE_{\|}^{(\mu)}\right)}{\mathrm{Tr}^2\,(\rho_{ML} E^{(\mu)})} E_{\|}^{(\mu)} + \left(\lambda_{ML} I - \nabla f|_{\rho_{ML}}\right) X \rho_{ML}^{+} + \rho_{ML}^{+} X \left(\lambda_{ML} I - \nabla f|_{\rho_{ML}}\right)$$

with ρ_{ML}^{+} the Moore-Penrose pseudo-inverse of ρ_{ML}.

- Low-rank approximations and efficient numerical schemes for computations of ρ_{ML}, the adjoint states $E^{(n)}$, ...
- Asymptotics when the log-likelihood function is not strongly concave, when $\ker\left(\lambda_{ML} I - \nabla f|_{\rho_{ML}}\right) \neq \ker(I - P_{ML})$ ...
- Process tomography: log-likelihood function not concave ...
- Parameter estimation along quantum trajectories (in real-time) ...
- Thematic quarter at Institut Henri Poincaré in Paris next Spring 2018 gathering experimental physicists and theoreticians.

April 16th to July 13th, 2018

Organized by:

Etienne Brion, Université Paris-Sud, ENS Paris-Saclay, CNRS
Eleni Diamanti, Université Pierre et Marie Curie & CNRS
Alexei Ourjoumtsev, Collège de France & CNRS
Pierre Rouchon, Mines ParisTech & Inria

Institut Henri Poincaré

11 rue Pierre et Marie Curie
75231 Paris Cedex 05
France

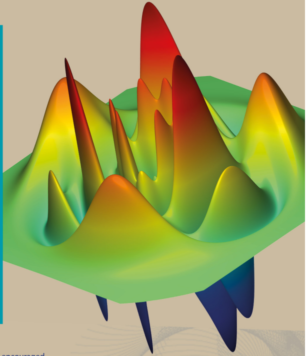
Measurement and control of quantum systems: theory and experiments

CIRM Pre-school at Marseille
Modeling and control of open quantum systems
April 16th - 20th, 2018

Observability and estimation in quantum dynamics
May 15th to 17th, 2018

Quantum control and feedback: foundations and applications
June 5th to 7th, 2018

PRACQSYS 2018: Principles and Applications of Control in Quantum Systems
July 2nd to 6th, 2018



Program coordinated by the Centre Emile Borel at IHP
Participation of Postdocs and PhD Students is strongly encouraged
Scientific program at: **https://sites.google.com/view/mcqs2018/home**

Registration is free however mandatory at : **www.ihp.fr**
Deadline for financial support : **September 15th, 2017**
Contact : mcqs2018@ihp.fr

Sylvie Lhermitte : CEB Manager

Supported also by:

Theorem (interior max) $\mathcal{I}_g(N) = \int_{z\in(-1,1)^n} g(z) \exp\left(Nf(z)\right) \, \mathrm{d}z$
with f and g analytic functions of z on a compact neighbourhood of $\overline{\mathcal{D}}$, the closure of $\mathcal{D}$. Assume that f admits a unique maximum on $\overline{\mathcal{D}}$ at $z = 0$ with $\left.\frac{\partial^2 f}{\partial z^2}\right|_0$ negative definite.
If $g(0) \neq 0$, we have the following dominant term in the asymptotic expansion of $\mathcal{I}_g(N)$ for large N:

$$\mathcal{I}_g(N) = \left(\frac{g(0)\ (2\pi)^{n/2}\ e^{Nf(0)} N^{-n/2}}{\sqrt{\left|\det\left(\left.\frac{\partial^2 f}{\partial z^2}\right|_0\right)\right|}} \right) + O\Big(e^{Nf(0)} N^{-n/2-1}\Big).$$

If $g(0) = 0$, with $\left.\frac{\partial g}{\partial z}\right|_0 = 0$, then we have:

$$\mathcal{I}_g(N) = \left(\frac{\mathrm{Tr}\left(- \left.\frac{\partial^2 g}{\partial z^2}\right|_0 \left(\left.\frac{\partial^2 f}{\partial z^2}\right|_0\right)^{-1}\right)\ (2\pi)^{n/2}}{2\sqrt{\left|\det\left(\left.\frac{\partial^2 f}{\partial z^2}\right|_0\right)\right|}} \right) e^{Nf(0)} N^{-n/2-1} + O\left(e^{Nf(0)} N^{-n/2-2}\right).$$

Corollary (interior max): Assume that f admits a unique maximum on $\overline{\mathcal{D}}$ at $z = 0$ with $\left.\frac{\partial^2 f}{\partial z^2}\right|_0$ negative definite. Then we have the following asymptotic for any analytic function $g(z)$:

$$\mathcal{M}_g(N) \triangleq \frac{\int_{z\in(-1,1)^n} g(z) \exp\left(Nf(z)\right) \mathrm{d}z}{\int_{z\in(-1,1)^n} \exp\left(Nf(z)\right) \mathrm{d}z} = g(0) + O(N^{-1})$$

We have also:

$$\mathcal{V}_g(N) \triangleq \frac{\int_{z\in(-1,1)^n} \left(g(z) - \mathcal{M}_g(N)\right)^2 \exp\left(Nf(z)\right)\ \mathrm{d}z}{\int_{z\in(-1,1)^n} \exp\left(Nf(z)\right)\ \mathrm{d}z}$$

$$= \frac{\mathrm{Tr}\left(-\left.\frac{\partial^2 g}{\partial z^2}\right|_0 \left(\left.\frac{\partial^2 f}{\partial z^2}\right|_0\right)^{-1}\right)}{2N} + O\big(N^{-2}\big).$$

Theorem (boundary max):
$\mathcal{I}_g(N) = \int_{x\in(0,1)} \int_{z\in(-1,1)^n} x^m g(x,z) \exp\left(Nf(x,z)\right) \,\mathrm{d}x\, \mathrm{d}z$ with f and g analytic functions of (x,z) on a compact neighbourhood of $\overline{\mathcal{D}}$, the closure of $\mathcal{D}$. Assume that f admits a unique maximum on $\overline{\mathcal{D}}$ at $(x,z)=(0,0)$, with $\frac{\partial^2 f}{\partial z^2}\Big|_{(0,0)}$ negative definite and $\frac{\partial f}{\partial x}\Big|_{(0,0)} < 0$. If $g(0,0)\neq 0$, we have the following dominant term in the asymptotic expansion of $\mathcal{I}_g(N)$ for large N:

$$\mathcal{I}_g(N) = \left(\frac{g(0,0)\; m!\; (2\pi)^{n/2}\; e^{Nf(0,0)} N^{-m-n/2-1}}{\sqrt{\left|\det\left(\frac{\partial^2 f}{\partial z^2}\Big|_{(0,0)}\right)\right|}\;\left(-\frac{\partial f}{\partial x}\Big|_{(0,0)}\right)^{m+1}}\right) + O\Big(e^{Nf(0,0)}N^{-m-n/2-2}\Big).$$

If $g(0,0)=0$, with $\frac{\partial g}{\partial x}\Big|_{(0,0)} = 0$ and $\frac{\partial g}{\partial z}\Big|_{(0,0)} = 0$, then we have:

$$\mathcal{I}_g(N) = \left(\frac{\mathrm{Tr}\left(-\frac{\partial^2 g}{\partial z^2}\Big|_{(0,0)}\left(\frac{\partial^2 f}{\partial z^2}\Big|_{(0,0)}\right)^{-1}\right)\; m!\; (2\pi)^{n/2}}{2\sqrt{\left|\det\left(\frac{\partial^2 f}{\partial z^2}\Big|_{(0,0)}\right)\right|}\;\left(-\frac{\partial f}{\partial x}\Big|_{(0,0)}\right)^{m+1}}\right) e^{Nf(0,0))} N^{-m-n/2-2}$$

$$+\, O\left(e^{Nf(0,0))}N^{-m-n/2-3}\right).$$

Corollary (boundary max): Assume that f admits a unique maximum on $\overline{\mathcal{D}}$ at $(x,z)=(0,0)$, with $\left.\frac{\partial^2 f}{\partial z^2}\right|_{(0,0)}$ negative definite and $\left.\frac{\partial f}{\partial x}\right|_{(0,0)}<0$. Then, we have the following asymptotic for any analytic function $g(x,z)$:

$$\mathcal{M}_g(N) \triangleq \frac{\int_{x\in(0,1)}\int_{z\in(-1,1)^n} x^m g(x,z)\exp\left(Nf(x,z)\right)\,\mathrm{d}x\,\mathrm{d}z}{\int_{x\in(0,1)}\int_{z\in(-1,1)^n} x^m \exp\left(Nf(x,z)\right)\,\mathrm{d}x\,\mathrm{d}z} = g(0,0)+O(N^{-1})$$

We have also:

$$\mathcal{V}_g(N) \triangleq \frac{\int_{x\in(0,1)}\int_{z\in(-1,1)^n} x^m \Big(g(x,z)-\mathcal{M}_g(N)\Big)^2 \exp\left(Nf(x,z)\right)\,\mathrm{d}x\,\mathrm{d}z}{\int_{x\in(0,1)}\int_{z\in(-1,1)^n} x^m \exp\left(Nf(x,z)\right)\,\mathrm{d}x\,\mathrm{d}z}$$

$$= \frac{\mathrm{Tr}\left(-\left.\frac{\partial^2 g}{\partial z^2}\right|_{(0,0)}\left(\left.\frac{\partial^2 f}{\partial z^2}\right|_{(0,0)}\right)^{-1}\right)}{2N} + O(N^{-2}).$$