System identification and metrology

for quantum Markov processes

Mădălin Guță

School of Mathematical Sciences University of Nottingham

L. Bouten, N. Yamamoto M. van Horssen, C. Catana, K. Macieszczak, J. Kiukas J.P. Garrahan, I. Lesanovsky



Quantum Control Enginering: Mathematical Principles and Applications



Isaac Newton Institute, Cambridge, July 2014

Open system in the input-output formalism



- system interacts with environment in a Markov fashion
- Unitary dynamics: quantum stochastic differential equations ^{1 2}
- Goal: comprehensive theory of the output process and applications
 - ► Central Limit Theorem for fluctuations → statistical theory of system identification
 - ▶ Large Deviations → thermodynamics of quantum trajectories³
 - ▶ Ergodicity of quantum trajectories → filtering theory (stability, purification)
 - Applications in quantum metrology and quantum control

¹ K. R. Parthasarathy, An introduction to quantum stochastic calculus, Birkhäuser 1992

² C. W. Gardiner and P. Zoller, *Quantum Noise*, Springer 2004

³J.P. Garrahan and I. Lesanovsky, Phys. Rev. Lett. 2010

Linear / non-linear input-output quantum systems



Linear system \longrightarrow see presentation by N. Yamamoto⁴



Non-linear system: atom maser \longrightarrow this presentation

⁴M.G. and N. Yamamoto, Proc. IEEE CDC 2013

Outline

- Discrete time quantum Markov chains
 - output state, ergodicity
 - Central Limit Theorem for time averages
- Identifiability
 - which dynamical parameters can be identified ?
- Estimation precision
 - what is the quantum/classical Fisher information ?
 - are the errors normally distributed ?
- Quantum enhanced metrology
 - which Markov systems exhibit Heisenberg scaling ?
 - Is the Markov setting suitable for generating metrologically useful states ?

Successive unitary interactions of "memory" with "quantum noise units"



- output state is closely related to finitely correlated⁵ and matrix product states⁶
- quantum analogue of the 'classical' hidden Markov chains
- standard tools⁷ for switching between discrete and continuous time

⁷thanks to R. Hudson's work and others

⁵ M. Fannes, B. Nachtergale and R. Werner, Commun. Math. Phys. 1992

⁶D. Perez-Garcia, F. Verstraete, M. Wolf and I. Cirac, Quantum Inf. Comput. 2007



Dynamics depends only on the isometry

$$\begin{split} V : \mathbb{C}^D &\to & \mathbb{C}^D \otimes \mathbb{C}^k \\ V : |\varphi\rangle &\mapsto & U |\varphi\rangle \otimes |\chi\rangle = \sum_i K_i |\varphi\rangle \otimes |i\rangle \end{split}$$

■ System + output state as MPS

$$|\Psi_{V,\varphi}(n)\rangle \quad := \quad V^{(n)}\cdots V^{(1)}|\varphi\rangle = \sum_{(i_1,\dots,i_n)} K_{i_n}\cdots K_{i_1}|\varphi\rangle \otimes |i_n\otimes\cdots\otimes i_1\rangle$$

System's transition operator

$$T_V(\rho) = \sum_i K_i \rho K_i^*$$

Quantum Perron-Frobenius Theorem

■ If *S* is an irreducible CP map (no invariant subspace)

- The spectral radius r(S) is a non-degenerate eigenvalue of S
- the corresponding eigenvector is a strictly positive matrix
- the eigenvalues on the unit circle form a group
- If T is a transition operator (CP & TP) then
 - $\blacktriangleright \ r(T) = 1$
 - unique full rank stationary state: $T(\rho_{ss}) = \rho_{ss}$
- If T is primitive (irreducible and aperiodic) then
 - $|\lambda| < 1$ for all remaining eigenvalues
 - convergence to stationary state

 $T^n(\sigma) = (|\rho_{ss}\rangle \langle \mathbf{1}| + \mathsf{Rest})^n \, |\sigma\rangle = (|\rho_{ss}\rangle \langle \mathbf{1}| + \mathsf{Rest}^n) \, |\sigma\rangle \longrightarrow \rho_{ss}$

Key observation: if T_{ϵ} is a small perturbation of primitive $T \Longrightarrow$ dominant eigenvalue λ_{ϵ} is smooth and determines the asymptotics

" $T_{\epsilon}^n \approx \lambda_{\epsilon}^n$ "



⁸D. E. Evans and R. Hoegh-Krohn, J. London Math. Soc, 1978; M. Sanz et al, IEEE Trans. Inform. Th., 2010

Quantum Markov chains: sequential output measurements



- Observable $A = \sum_i a_i |i\rangle \langle i|$ is measured on each unit \longrightarrow outcomes A_1, A_2, \ldots
- Basic statistic: time (empirical) average $S_n(A) := \frac{1}{n} \sum_{i=1}^n A_i$
- Moment generating function

$$\phi(s) := \mathbb{E}\left(e^{snS_n(A)}\right) = \operatorname{Tr}\left(T_s^n(\rho_{in})\right)$$

Deformed transition operator $T_s : M(\mathbb{C}^D) \to M(\mathbb{C}^D)$ (CP, non-TP):

$$T_{s}:
ho \longmapsto \sum_{i} e^{sa_{i}} K_{i}
ho K_{i}^{*}$$

Theorem (Central Limit)

Let T be primitive. Then 1) time averages converge to stationary means: $S_n(A) \longrightarrow \mathbb{E}_{ss}(A)$

2) fluctuations are normal

$$\mathbb{F}_n(A) := \sqrt{n} \left(S_n(A) - \mathbb{E}_{ss}(A) \right) \xrightarrow[n \to \infty]{\mathcal{L}} N(0, V(A))$$

with variance

$$V(A) = \begin{cases} \mathbb{E}_{ss} \left(A^2 \right) + 2\mathbb{E}_{ss} \left(A \otimes (\mathrm{Id} - T)^{-1}(B) \right), & B = \langle \psi | U^* (\mathbf{1} \otimes A) U | \psi \rangle \\ \frac{d^2 \log \lambda_s}{ds^2} \Big|_{s=0}, & \lambda_s = \text{dominant eigenvalue of } T_s. \end{cases}$$

Remarks

1) A similar CLT holds for time averages of multiple-outcomes functions $f(A_1, \ldots, A_r)$, and it involves an extended deformed transition operator T_f which keeps track of previous outcomes ⁹

2) A similar CLT holds for the total counts and integrated homodyne current in continuous-time measurements $^{\rm 10}$

⁹M. van Horssen and M.G., arXiv:1407.5082

¹⁰C. Catana, L. Bouten and M.G., arXiv:1407.5131

¹¹B. Kummerer and H. Maaseen, J. Phys. A: Math. Gen. 2003; M.G. Phys. Rev. A, 2011

CLT follows from the convergence of characteristic functions

$$\mathbb{E}\left(e^{is\mathbb{F}_n(A)}\right) = \operatorname{Tr}\left(T^n_{is/\sqrt{n}}(\rho_{in})\right) \xrightarrow[n \to \infty]{} e^{-s^2V(A)/2}$$

A) Use non-zero spectral gap and second order expansion

$$T_{is/\sqrt{n}} = T_0 + \frac{1}{\sqrt{n}}T_1 + \frac{1}{n}T_2 + \text{Rest}$$

to show

$$T_{is/\sqrt{n}}^{n}(\mathbf{1}) = \left(1 - \frac{s^{2}V(A)}{2n}\right)^{n} \mathbf{1} + o(1) \to \exp(-s^{2}V(A)/2)\mathbf{1}$$

B) Follows (together with large deviations) from

$$\operatorname{Tr}\left(T_{is/\sqrt{n}}^{n}(\rho_{in})\right) \approx e^{n\log\lambda_{is/\sqrt{n}}} \approx \exp\left(-\frac{s^{2}}{2} \left.\frac{d^{2}\log\lambda_{s}}{ds^{2}}\right|_{s=0}\right)$$

Outline

- Discrete time quantum Markov chains
 - output state, ergodicity
 - Central Limit Theorem for time averages
- Identifiability
 - which dynamical parameters can be identified ?
- Estimation precision
 - what is the quantum/classical Fisher information ?
 - are the errors normally distributed ?
- Quantum enhanced metrology
 - which Markov systems exhibit Heisenberg scaling ?
 - Is the Markov setting suitable for generating metrologically useful states ?

The system identification problem



- Suppose that the dynamics depends on an unknown parameter θ , so $U = U_{\theta}$
- Task: estimate the parameter by measuring the output state
- Methods: Bayesian/extended filter¹², maximum likelihood, compressed sensing¹³
- Here we are interested in statistical aspects
 - which parameters are identifiable?
 - what is the accuracy / Fisher information?
 - what is the structure of the output state?

¹²H. Mabuchi Quant. Semiclass. Optics 1996; J. Gambetta and H. M. Wiseman Phys. Rev. A 2001;

S. Gammelmark and K. Mølmer Phys. Rev. A 2013

¹³M. Cramer et al, Nat. Commun. 2010

Definition

Two primitive chains with isometries V_1 and V_2 are called equivalent if for all n,

$$\rho_{V_1}^{out}(n) = \rho_{V_2}^{out}(n).$$

Theorem (equivalence classes)

Two primitive chains with isometries V_1 and V_2 are equivalent if and only if there exists a phase $e^{i\phi}$ and a unitary $W: \mathbb{C}^D \to \mathbb{C}^D$ such that

 $V_2 = e^{i\phi}(W \otimes \mathbf{1})V_1W^*$

or equivalently

$$K_i^{V_2} = e^{i\phi} W K_i^{V_1} W^*, \qquad i = 1, \dots, k.$$

Remarks

1) Theorem 1 is a quantum extension of the 'classical' result by Petrie^{14} on equivalence classes of ergodic hidden Markov chains

2) similar result holds in continuos-time: $L_i^{V_2} = W L_i^{V_1} W^*$ and $H^{V_2} = W H^{V_1} W^* + cI$

3) similar result holds for (passive) linear systems ¹⁵ [see presentation by N. Yamamoto]

¹⁴T. Petrie, Annals of Math. Statistics, 1969

¹⁵M.G. and N. Yamamoto, *IEEE Proceedings 52nd CDC* 2013

¹⁶M.G. and J. Kiukas, arXiv:1402.3535

Define the "off-diagonal transition operator"

$$T_{12}:\rho\mapsto \sum_{i=1}^D K_i^{V_1}\rho K_i^{V_2*}$$

Overlap of the two system-output states

$$\langle \Psi_{V_2,\varphi}(n)|\Psi_{V_1,\varphi}(n)\rangle = \operatorname{Tr}\left(T_{1,2}^n(|\varphi\rangle\langle\varphi|)\right) \approx \lambda_{1,2}^n$$

Two alternatives:

A) $|\lambda_{12}| = 1 \implies K_i^{V_2} = e^{i\phi}WK_i^{V_1}W^* \implies$ equivalent systems

 $\mathsf{B}) \ |\lambda_{12}| < 1 \implies \text{ overlap decays exponentially } \implies \text{ non-equivalent systems}$

Outline

- Discrete time quantum Markov chains
 - output state, ergodicity
 - Central Limit Theorem for time averages
- Identifiability
 - which dynamical parameters can be identified ?

Estimation precision

- what is the quantum/classical Fisher information ?
- are the errors normally distributed ?
- Quantum enhanced metrology
 - which Markov systems exhibit Heisenberg scaling ?
 - Is the Markov setting suitable for generating metrologically useful states ?

Statistical problem:

estimate parameter $\theta \in \mathbb{R}$ by measuring ensemble of n i.i.d. systems with state

$$|\psi_{\theta}\rangle = e^{i\theta G} |\psi_0\rangle, \qquad \langle \psi_0 | G |\psi_0\rangle = 0.$$

Local asymptotic normality a.k.a. Holstein-Primakov:

In an "uncertainty neighbourhood" of size $n^{-1/2}, \, {\rm the \ overlaps}$ of joint states are approximately Gaussian

$$\left\langle \psi_{\theta_0+v/\sqrt{n}}^{\otimes n} \middle| \psi_{\theta_0+u/\sqrt{n}}^{\otimes n} \right\rangle = \left\langle \psi_0 \left| e^{i(u-v)G/\sqrt{n}} \right| \psi_0 \right\rangle^n \longrightarrow e^{\frac{(u-v)^2F}{8}} = \left\langle \sqrt{F/2v} \middle| \sqrt{F/2u} \right\rangle$$

•
$$F = 4 \text{Var}(G) = 4 \left\| \frac{d\psi_{\theta}}{d\theta} \right\|^2$$
 is the quantum Fisher information

 $\blacktriangleright~\left|\sqrt{F/2}u\right\rangle$ is a one-mode coherent state with displacement $\left(\sqrt{F/2}u,0\right)$

Remark

LAN holds for mixed states &multi-dimensional models, and has an operational interpretation in terms of mutual approximation through quantum channels ¹⁷

¹⁷J. Kahn and M.G., Commun. Math. Phys., 2009

System identification: system+output setup¹⁹

- Primitive Markov chain $V = V_{\theta}$ with $\theta \in \mathbb{R}$ unknown parameter
- 'Localise' θ as $\theta = \theta_0 + \frac{u}{\sqrt{n}}$ by using adaptive measurements
- Assume complete access to system + output state $|\Psi_{u,\varphi}(n)\rangle \equiv |\Psi_{V_{\theta},\varphi}(n)\rangle$

Theorem (LAN for quantum Markov chains)

The output model $|\Psi_{u,\varphi}(n)
angle$ converges 'weakly' to the coherent state model $\left|\sqrt{F/2}\,u
ight
angle$

$$\lim_{n \to \infty} \langle \Psi_{v,\varphi}(n) | \Psi_{u,\varphi}(n) \rangle = \left\langle \sqrt{F/2} \, v \, \middle| \, \sqrt{F/2u} \right\rangle = \exp(-F(u-v)^2/8)$$

ightarrow F is the quantum Fisher information 'per sample'

 \rightarrow optimal estimator satisfies asymptotic normality

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} N(0, F^{-1})$$

Remark: the above Theorem is a quantum extension of 'classical' results on HMC ¹⁸

¹⁸P.J. Bickel, Y. Ritov, and T. Ryden Ann. Statist (1998)

¹⁹M.G. and J. Kiukas, arXiv:1402.3535

Overlap can be reduced to dominant eigenvalue of a deformed transition operator

$$\langle \Psi_{\boldsymbol{v},\varphi}(n)|\Psi_{\boldsymbol{u},\varphi}(n)\rangle = \operatorname{Tr}\left(T^n_{\frac{\boldsymbol{u}}{\sqrt{n}},\frac{\boldsymbol{v}}{\sqrt{n}}}(|\varphi\rangle\langle\varphi|)\right) \approx \exp(n\log\lambda_{\frac{\boldsymbol{u}}{\sqrt{n}},\frac{\boldsymbol{v}}{\sqrt{n}}})$$

• Expanding in
$$\frac{u}{\sqrt{n}}, \frac{v}{\sqrt{n}}$$
, and setting $\frac{\partial \lambda a, b}{\partial a}\Big|_{a=b=0} = 0$
 $\langle \Psi_{v,\varphi}(n) | \Psi_{u,\varphi}(n) \rangle \longrightarrow \exp\left(\frac{1}{2} \left. \frac{\partial^2 \log \lambda_{a,b}}{\partial a \partial b} \right|_{a=b=0} \right)$

so that

$$F = -4 \left. \frac{\partial^2 \log \lambda_{a,b}}{\partial a \partial b} \right|_{a=b=0}$$

Similar methods have been used in²⁰, ²¹

²⁰M.Cozzini, R. Ionicioiu, and P. Zanardi, Phys. Rev. B, 2007

²¹S. Gammelmark and K. Mølmer, Phys. Rev. Lett., 2014

Quantum Fisher information = Markovian variance of the generator $G := i \dot{U} U^*$

 $F = 4\operatorname{Var}_V(G) = 4(G, G)_V$

Fluctuations operator: for $X \in M_D \otimes M_k$ with $\mathbb{E}_{ss}(X) = 0$

$$\mathbb{F}_n(X) = \frac{1}{\sqrt{n}} \sum_{i=1}^n X(i), \qquad X(i) = U(i)^* X_i U(i)$$

Markov covariance inner product: for $X, Y \in M_D \otimes M_k$

$$(X,Y)_V := \lim_{n \to \infty} \frac{1}{n} \langle \mathbb{F}_n(X^*) \mathbb{F}_n(Y) \rangle = \mathbb{E}_{ss} \left[X^*Y + X^* \left(\mathcal{R} \circ \mathcal{E}(Y) \otimes \mathbf{1} \right) + \left(\mathcal{R} \circ \mathcal{E}(X^*) \otimes \mathbf{1} \right) Y \right]$$

where

•
$$\mathcal{E}: M_D \otimes M_k \to M_D$$
 is the conditional expectation $\mathcal{E}(X) = V^*XV$

$$\mathbf{\mathbb{P}} \mathcal{R} = (\mathrm{Id} - T)^{-1} |_{\mathcal{D}} \text{ with } \mathcal{D} := \{Y : \mathrm{Tr}(\rho_{ss}Y) = 0\}$$

- Assume that dynamics is in the stationary regime
- Assume access to output state $\rho_{u}^{out}(n) := \rho_{V_{a}}^{out}(n)$

Theorem (LAN for quantum Markov chains)

The output model $\rho_u^{out}(n)$ converges strongly to the coherent state model $|\sqrt{F/2} u\rangle$: there exists quantum channels T_n and S_n such that

$$\lim_{n \to \infty} \sup_{\|\boldsymbol{u}\| < C} \left\| T_n(\rho_u^{out}(n)) - \left| \sqrt{F/2} \, \boldsymbol{u} \right\rangle \left\langle \sqrt{F/2} \, \boldsymbol{u} \right| \right\|_1 = 0$$
$$\lim_{n \to \infty} \sup_{\|\boldsymbol{u}\| < C} \left\| \rho_u^{out}(n) - S_n\left(|\sqrt{F/2} \, \boldsymbol{u}\rangle \langle \sqrt{F/2} \, \boldsymbol{u} | \right) \right\|_1 = 0$$

 \rightarrow F is the quantum Fisher information 'per sample'

 \rightarrow "output contains all information"

²²M.G. and J. Kiukas, arXiv:1402.3535

Classical Fisher information of time averages



• Unknown dynamics V_{θ} with parameter $\theta = \theta_0 + u/\sqrt{n}$

Time average $S_n = \frac{1}{n} \sum_{i=1}^n A_i$ captures deviations from mean $\mu_{\theta_0} = \mathbb{E}_{\theta_0}(A)$

$$\sqrt{n}(S_n - \mu_{\theta_0}) \xrightarrow[n \to \infty]{\mathcal{L}} N\left(\frac{d\mu}{d\theta}u, V(A)\right)$$

• Classical Fisher information = signal to noise ratio (in terms of dom. eigenv. $\lambda_{s,\theta}$ of $T_{s,\theta}$)

$$I^{A}(\theta_{0}) = \frac{\left(\frac{d\mu}{d\theta}\right)^{2}}{V(A)} = \frac{\left(\frac{\partial^{2}\lambda_{s,\theta}}{\partial s\partial \theta}\Big|_{s=0,\theta=\theta_{0}}\right)^{2}}{\frac{\partial^{2}\lambda_{s,\theta}}{\partial^{2}s^{2}}\Big|_{s=0}}$$

Both quantum and classical Fisher informations rely on an underlying CL behavior

$$I^{A} = \frac{\left(\frac{d\mu}{d\theta}\right)^{2}}{\operatorname{Var}_{V}(A)} \leq F = 4\operatorname{Var}_{V}(G)$$

Open questions

- Which measurement achieves the quantum Fisher information?
- Is there an "invariance principle" for quantum Markov chains ?

Conjecture: general CLT

Observable $X \in M(\mathbb{C}^D) \otimes M(\mathbb{C}^k)^{\otimes r}$ "localised" in system and noise units. There is a CCR algebra with canonical coordinates B(X) and Gaussian state Φ such that

$$\Phi(B(X)B(Y)) = \lim_{n \to \infty} \langle \mathbb{F}_n(X) | \mathbb{F}_n(Y) \rangle$$
$$[B(X), B(Y)] = \lim_{n \to \infty} \langle [\mathbb{F}_n(X), \mathbb{F}_n(Y)] \rangle \mathbf{1}$$



Jaynes-Cummings interaction between a two-level atom and a cavity

$$U: |1\rangle \otimes |k\rangle \mapsto \cos\left(\phi\sqrt{k+1}\right) |1\rangle \otimes |k\rangle + \sin\left(\phi\sqrt{k+1}\right) |0\rangle \otimes |k+1\rangle$$

Coarse grained cavity dynamics for Poisson distributed input atoms with rate N_{ex}

$$\frac{d\rho}{dt} = \mathcal{L}(\rho) = \sum_{i=1}^{4} \left(L_i \rho L_i^* - \frac{1}{2} \{ L_i^* L_i, \rho \} \right)$$

with jump operators

- $L_1: |k\rangle \mapsto \sqrt{N_{ex}} \sin(\phi \sqrt{k+1}) |k+1\rangle$ (excitation absorbed from atom)
- $L_2: |k\rangle \mapsto \sqrt{N_{ex}} \cos(\phi \sqrt{k+1}) |k\rangle$ (atom remains in excited state)

•
$$L_3: \ket{k} \mapsto \sqrt{k(
u+1)} \ket{k-1}$$
 (photon emitted in the bath)

•
$$L_4: |k
angle \mapsto \sqrt{(k+1)
u} |k+1
angle$$
 (photon absorbed from the bath)

²³H.-J. Briegel, B.-G. Englert, N. Sterpi, and H. Walther, Phys. Rev. A 1994

The many Fisher informations of the atom maser ²⁴,²⁵



²⁴C. Catana, M van Horssen, M.G., Phil. Trans. Royal Soc. A (2012)

 $^{^{25}}$ C.Catana, T. Kypraios and M.G. arXiv:1311.4091

Outline

- Discrete time quantum Markov chains
 - output state, ergodicity
 - Central Limit Theorem for time averages
- Identifiability
 - which dynamical parameters can be identified ?
- Estimation precision
 - what is the quantum/classical Fisher information ?
 - are the errors normally distributed ?
- Quantum enhanced metrology
 - which Markov systems exhibit Heisenberg scaling ?
 - Is the Markov setting suitable for generating metrologically useful states ?

Simplest example of Markov dynamics with degenerate stationary states ("phases")

$$\begin{split} V: \quad \mathbb{C}^2 &\to \mathbb{C}^2 \otimes \mathbb{C}^2 \\ V: \quad |0\rangle &\mapsto \quad |0\rangle \otimes |0\rangle \\ V: \quad |1\rangle &\mapsto e^{i\theta} |1\rangle \otimes |1\rangle \end{split}$$

 \blacksquare Output + system state exhibits Heisenberg scaling for initial state $|0\rangle+|1\rangle/\sqrt{2}$

$$|\Psi^{n}(\theta)\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle \otimes |0\rangle^{\otimes n} + e^{in\theta}|1\rangle \otimes |1\rangle^{\otimes n}\right)$$

Heisenberg scaling for Markov systems with degenerate stationary phases

■ If system has 2 stationary "phases" then O+S state is a "massive superposition"

$$|\Psi_{\theta}(n)\rangle = \frac{1}{\sqrt{2}} \left(|\Psi_{\theta}^{(0)}(n)\rangle + |\Psi_{\theta}^{(1)}(n)\rangle \right)$$

Fisher information = Var(G) w.r.t. a mixture of Gaussian distributions

 $F \approx n^2 (g_1 - g_2)^2$



 Exploit this for quantum enhanced metrology with open systems near dynamical phase transitions ²⁶ ²⁷

²⁶C. Catana and M.G., Phys. Rev. A 2014

²⁷K. Macieszczak, J.P. Garrahan, I. Lesanovsky and M.G., in preparation

- Stationary (primitive) quantum Markov chains can be characterised completely up to unitary "change of coordinates" by measuring the output
- The output state is asymptotically Gaussian with quantum Fisher information equal to the "Markov variance of the generator"
- Multiple phases chains can exhibit Heisenberg scaling
- Future work
 - enhanced metrology & dynamical phase transitions
 - general quantum Markov CLT
 - use of feedback in system identification
 - design better input states