Random tensors and
related topics
02/10/2024

Learning finitely correlated states: stability of the spectral reconstruction

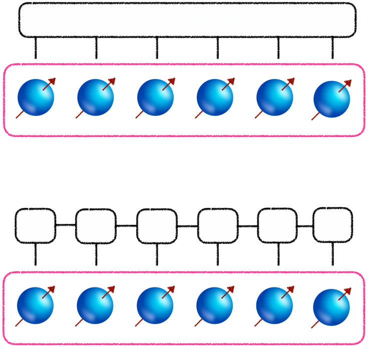
Marco Fanizza[1,2], Niklas Galke[1], Josep Lumbreras[2], Cambyse Rouzé[3], Andreas Winter[1,4]

arXiv:2312.07516

[1]University of Copenhagen
[2]Universitat Autònoma de Barcelona
[3]Centre for Quantum Technologies, National University of Singapore
[4]INRIA, Télécom Paris
[5]ICREA

Learning states on a qudit chain

- Task: given $\omega^{\otimes n}$, find $\hat{\omega}$ s.t. $\frac{\|\hat{\omega}-\omega\|_1}{2} \leq \epsilon$ with high probability
- Local dimension $d_{\mathcal{A}}$
- General case, t sites: $n = \Theta(\frac{d_{\mathcal{A}}^{2t}}{\epsilon^2})$ (optimal tomography)
- Promised structure: marginals of finitely correlated states (special case of matrix product operators), hidden dimension $\leq m$
- New result: $O(\mathrm{poly}(t, \epsilon^{-1}, d_{\mathcal{A}}, m, \text{additional parameters}))$



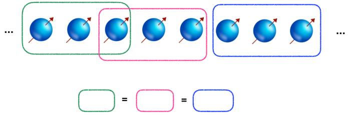
Finitely correlated states

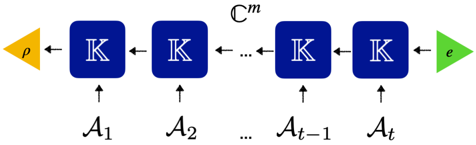
- Let ω be a translation invariant state (stationary quantum source).
  ω_t = marginals of segments of length t
- ω is an FCS if there exist $m \in \mathbb{N}$ and a realization
  - $\rho \in \mathbb{C}^m$,
  - $e \in \mathbb{C}^m$,
  - $\mathbb{K} : \mathcal{L}(\mathcal{H}_{\mathcal{A}}) \to \mathbb{M}_{m,m}(\mathbb{C})$,

  s.t.

  $$\langle j_1, ...j_t|\omega_t|i_1, ...i_t\rangle = \rho \mathbb{K}_{|i_1\rangle\langle j_1|} ... \mathbb{K}_{|i_t\rangle\langle j_t|} e$$

  for any $t \in \mathbb{N}$.





[Fannes, Nachtergaele, Werner 1992]

Model reconstruction from marginals

Exact reconstruction

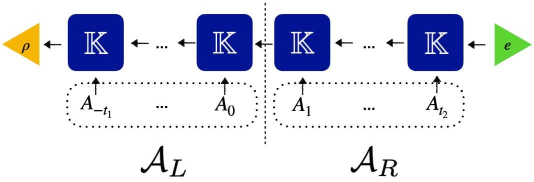
- Viable model parameters $\rho, \mathbb{K}, e$ can be reconstructed from ω_{t^*}, t^* large enough (at most $2m+1$)
- The exact reconstruction algorithm uses ω_{t^*} and linear algebra [Baumgratz, Gross, Cramer, Plenio 2013]

Approximate reconstruction (LearnFCS)

- Obtain approximate knowledge of ω_{t^*} from tomography
- Estimates of ω_{t^*} can be plugged in (a modification of) the exact reconstruction algorithm, with outputs $\hat{\rho}, \hat{\mathbb{K}}, \hat{e}$
- Compute expectation values of estimate $\hat{\omega}_t$ as

$$\mathrm{Tr}[\hat{\omega}_t(A_1 \otimes \ldots \otimes A_t)] = \hat{\rho}\hat{\mathbb{K}}_{A_1} \ldots \hat{\mathbb{K}}_{A_t}\hat{e}.$$

Exact reconstruction



$$\omega_{t_2-t_1+1} \to \Omega^{(-t_1,t_2)} : \mathcal{A}_R \to \mathcal{A}_L^*, \quad \Omega^{(-t_1,t_2)}(A_r)[A_l] = \mathrm{Tr}[\omega_{t_2-t_1+1}(A_l \otimes A_r)]$$

$$\omega_{t_2-t_1+2} \to \Omega_A^{(-t_1,t_2)} : \mathcal{A}_R \to \mathcal{A}_L^*, \quad \Omega_A^{(-t_1,t_2)}(A_r)[A_l] = \mathrm{Tr}[\omega_{t_2-t_1+2}(A_l \otimes A \otimes A_r)]$$

$$\omega_{t_2} \to \Omega^{(t_2)} \in \mathcal{A}_R^*, \quad \Omega^{(t_2)}(A_r) = \mathrm{Tr}[\omega_{t_2}(A_r)]$$

$$\omega_{t_1+1} \to \Omega^{(-t_1)} \in \mathcal{A}_L^*, \quad \Omega^{(-t_1)}(A_l) = \mathrm{Tr}[\omega_{t_1+1}(A_l)]$$

- ω is FCS iff the rank of $\Omega^{(-t_1,t_2)}$ stays bounded as t_1, t_2 grow.
- $m :=$ minimal dimension of a realization $=$ maximum rank of $\Omega^{(-t_1,t_2)}$ realizations of minimal dimension are called regular.

Observable realization and approximate reconstruction

Observable realization

Let ω be an FCS with a regular realization of dimension m.
If $\Omega^{(-s+1,s)}$ has rank m, with singular value decomposition $\Omega^{(-t_1,t_2)} = UDO$, then

- $e := U^\top \Omega^{(s)} \in \mathbb{C}^m$,
- $\rho := \Omega^{(-s+1)}(U^\top \Omega^{(-s+1,s)})^+ \in \mathbb{C}^m$,
- $\mathbb{K}_A := U^\top \Omega_A^{(-s+1,s)}(U^\top \Omega^{(-s+1,s)})^+ \in \mathbb{M}_{m,m}(\mathbb{C})$.

is a regular realization.

LearnFCS: plugin estimates of Ωs and truncate singular value decomposition.

Approximation theorem

Definition

We define $\Sigma(m, s, \eta)$ as the class of FCSs with regular realization of dimension $m' \leq m$, $\mathrm{rank}(\Omega^{(-s+1,s)}) = m'$, and with $\sigma_{m'}(\Omega^{(-s+1,s)}) \geq \eta$.

Theorem

Let $\omega \in \Sigma(m, s, \eta)$. Let $\hat{\omega}_s, \hat{\omega}_{2s}, \hat{\omega}_{2s+1}$ be estimates of $\omega_s, \omega_{2s}, \omega_{2s+1}$ respectively, such that $D_{HS}(\hat{\omega}_s, \omega_s)$, $D_{HS}(\hat{\omega}_{2s}, \omega_{2s})$, $D_{HS}(\hat{\omega}_{2s+1}, \omega_{2s+1})$ are smaller than $\frac{\epsilon\eta^3}{20tm\sqrt{d_{\mathcal{A}}}}$. Then, $\hat{\omega}_t$ obtained from LearnFCS satisfies

$$\frac{\|\hat{\omega}_t - \omega_t\|_1}{2} \leq \epsilon. \quad (1)$$

Sample complexity bounds

Set n be the number of copies of ω_{2s+1} provided to the learner. LearnFCS outputs $\hat{\omega}_t$, $\|\hat{\omega}_t - \omega_t\|_1 \leq 2\epsilon$ when n large enough with

- local measurements
  (via [Guta, Kahn, Kueng, Tropp 2020]):

$$n = \tilde{O}\left(\frac{t^2 \min(m^2, d_B^2) d_A^{6s+4}}{\epsilon^2 \eta^6}\right). \tag{2}$$

- global but single-copy measurements:
  (via [Qin, Jameson, Gong, Wakin, Zhu 2023]),

$$n = \tilde{O}\left(\frac{s^3 t^2 m^2 \min(m^2, d_B^2) d_A^3}{\epsilon^2 \eta^6}\right). \tag{3}$$

Error propagation: proof overview

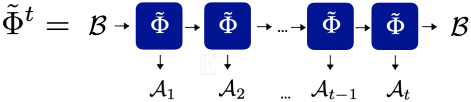
- $\tilde{\rho}, \tilde{\mathbb{K}}, \tilde{e}$ empirical (exact) realization, depend on measurement results

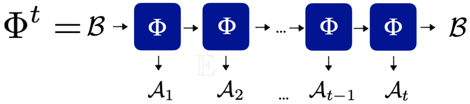
$$\omega_t = \tilde{\rho}\tilde{\mathbb{K}}^t\tilde{e} \qquad \hat{\omega}_t = \hat{\rho}\hat{\mathbb{K}}^t\hat{e}$$

$$\|\hat{\omega}_t - \omega_t\|_1 \leq \underbrace{\|(\hat{\rho} - \tilde{\rho})\tilde{\mathbb{K}}^t\tilde{e}\|_1}_{\text{(I)}} + \underbrace{\|(\hat{\rho} - \tilde{\rho})(\hat{\mathbb{K}}^t\hat{e} - \tilde{\mathbb{K}}^t\tilde{e})\|_1}_{\text{(II)}} + \underbrace{\|\tilde{\rho}(\tilde{\mathbb{K}}^t\tilde{e} - \hat{\mathbb{K}}^t\hat{e})\|_1}_{\text{(III)}} \qquad (4)$$

- Bound each term using telescoping sums, triangle inequalities, submultiplicativity of (cb) operator norms (for operator systems)
- Connect with Hilbert-Schmidt error of the estimates from tomography

Error propagation: heuristic from quantum maps





- If $\Phi, \tilde{\Phi}$ are channels
$$\|\Phi^t(\rho) - \tilde{\Phi}^t(\rho)\|_1 \leq t\|\Phi - \tilde{\Phi}\|_\diamond. \tag{5}$$
- In general,
$$\|\Phi^t(\rho) - \tilde{\Phi}^t(\rho)\|_1 \leq (\|\Phi\|_\diamond + \|\Phi - \tilde{\Phi}\|_\diamond)^t - \|\Phi\|_\diamond^t. \tag{6}$$
from triangle inequalities, submultiplicativity and induction.
- Similar argument valid also for approximations of the regular realizations and operator systems cb norms. Which cb norms should we use?

Solution: cb norm for operator systems

- Convex cone: $C \subseteq \mathbb{R}^m \subseteq \mathbb{C}^m$.
- Positivity: $v \geq_C 0$ $(= v \in C$, generalizes $A \geq 0)$
- Norm: $\|v\|_e = \{\min \lambda \geq 0 : \pm v \leq_C \lambda e\}$ (generalizes $\|A\|_\infty$)
- Operator system: Family of cones $\{C_n\}$ in $\mathbb{C}^m \otimes \mathbb{M}_n(\mathbb{C})$ (with certain properties)
- Operator systems $\{C_n\}$ for $\mathbb{C}^{m_1}$, $\{C'_n\}$ for $\mathbb{C}^{m_2}$
  $\rightarrow$ notions of cb norm, complete positivity for maps $\mathbb{C}^{m_1} \rightarrow \mathbb{C}^{m_2}$.

Operator system structure of ω

- $x \in \mathbb{C}^m$ can be written as $\mathbb{K}^t(X \otimes e)$ for some t, $X \in \mathbb{M}_{d^t_{\mathcal{A}}, d^t_{\mathcal{A}}}(\mathbb{C})$.
- $C_n := \{x \in \mathbb{C}^m \otimes \mathbb{M}_n(\mathbb{C})\,; \exists t \in \mathbb{N}, X \geq 0 \,\text{s.t.}\, x = (\mathrm{id}_{\mathbb{M}_n(\mathbb{C})} \otimes \mathbb{K}^t)(X \otimes e)\}$.
- With these choices $\mathbb{K}$ is **completely positive and unital**, therefore $||\mathbb{K}||_{cb} = 1$.
- Telescoping sum trick gives the desired bound.

Operator systems from FCS

Connection with Hilbert-Schmidt errors

Equivalence with Euclidean norm on $\mathbb{C}^m \otimes \mathbb{M}_n(\mathbb{C})$:

$$\sigma_m(\Omega)\|\cdot\|_{e_n} \leq \|\cdot\|_{n,2} \leq \sqrt{n}\|\cdot\|_{e_n}$$

- It follows that $\|\tilde{\mathbb{K}} - \hat{\mathbb{K}}\|_{\mathbf{1}_{d_A}\otimes e\to e,cb} \leq m\frac{\sqrt{d_A}}{\sigma_m(\Omega)}\|\tilde{\mathbb{K}} - \hat{\mathbb{K}}\|_{2\to 2}$,
- and $\|\tilde{\mathbb{K}} - \hat{\mathbb{K}}\|_{2\to 2} \leq 4\left(\frac{\|\Omega-\hat{\Omega}\|_2}{\sigma_m(\Omega)^2} + \frac{\|\Omega_{(\cdot)}-\hat{\Omega}_{(\cdot)}\|_2}{3\sigma_m(\Omega)}\right)$ via matrix perturbation theory

Extensions

- Quantum model: bounds from **quotient operator system**, finite amplification bound for cb norm of maps into matrices.
- The **non-translation-invariant case** can be addressed with essentially the same algorithm and proof for injective tensor networks.
  In this case, one needs that all the estimations of the necessary marginals are accurate, increasing the sample and computational complexity by $\mathrm{poly}(t)$.
- The algorithm is **robust to noise** and allows to learn states that are close to FCS. Potential $\mathrm{poly}(t)$ algorithm for 1D Gibbs states: efficient FCS approximation, but $\sigma_m(\Omega)$ is difficult to control. See also [Gondolf, Scalet, Ruiz-de-Alarcon, Alhambra, Capel 2024] for a different approach.

Conclusions

Summary

- FCSs can be learned with sample complexity polynomial in system size.
- Theory of operator systems proved useful for learning problems.

Future directions

- Finitely correlated channels, online learning
- Beyond 1D: graphs, 2D lattices
- Exhibit genuinely quantum FCSs which do not have models with quantum memory.

Thank you!

Related work

Learning HMM with *spectral* algorithms (vs maximum likelihood)

- [Hsu, Kakade, Zhang 2008]: first analysis of the reconstruction algorithm for HMMs.
- [Siddiqi, Boots, Gordon 2009], [Balle, 2013] relaxed some assumptions.
- No analysis for quantum or general models.

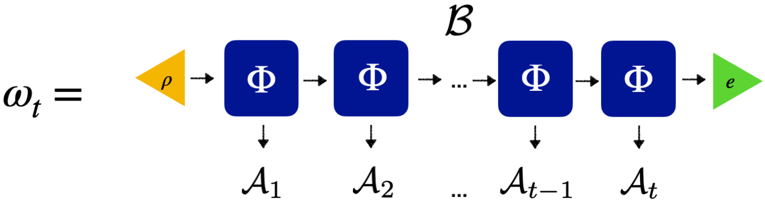
Learning FCSs

- [Baumgratz, Gross, Cramer, Plenio 2013]: reconstruction algorithm proposed (no error analysis).
- [Holzapfel, Cramer, Datta, Plenio 2018]: analysis of reconstruction for a bipartite state, not sufficient for trace distance guarantees.
- [Qin, Jameson, Gong, Wakin, Zhu 2023]: random measurements, analysis with Hilbert-Schmidt error.

Subclass: states generated by a finite quantum memory

A class of FCSs:

- Quantum systems $\mathcal{B}$, $\mathcal{A}$, with $\dim \mathcal{B} = d_{\mathcal{B}}$, $\dim \mathcal{A} = d_{\mathcal{A}}$,
- Quantum channel $\Phi : \mathcal{B} \to \mathcal{B}\mathcal{A}$, $\mathcal{E} = \Phi^\dagger : \mathcal{B}\mathcal{A} \to \mathcal{B}$ adjoint map
- ρ state of $\mathcal{B}$, $\mathrm{Tr}_{\mathcal{A}}[\Phi(\rho)] = \rho$, $e = \mathbb{1}_{\mathcal{B}}$



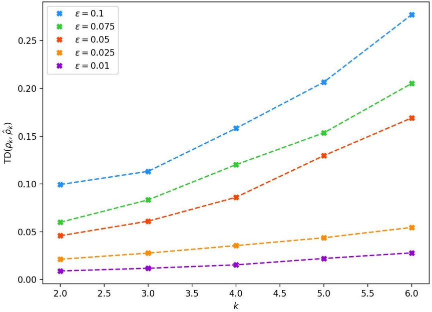
- $\mathrm{Tr}[\omega_t(A_1 \otimes ... \otimes A_t)] = \mathrm{Tr}_{\mathcal{B}}[\rho \mathcal{E}_{A_1} ... \mathcal{E}_{A_t}(\mathbb{1}_{\mathcal{B}})]$

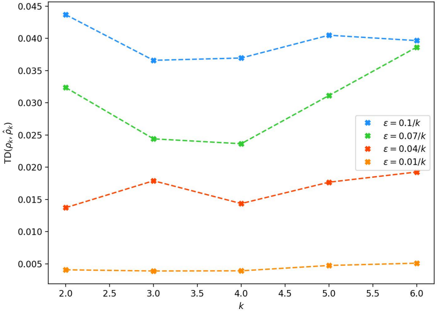
Numerical results

The ground state of the AKLT Hamiltonian

$$H_{\text{AKLT}} = \sum_i \frac{1}{2}\mathbf{S}_i \cdot \mathbf{S}_{i+1} + \frac{1}{6}\left(\mathbf{S}_i \cdot \mathbf{S}_{i+1}\right)^2 + \frac{1}{3}, \tag{7}$$

is an FCS with a quantum realization. Trace distance error of the reconstruction:





Solution: hidden system as a GPT

Generalized probabilistic theory (GPT)

Mathematical blueprint of theories describing preparations and measurements

- Not all FCSs have finite-dimensional quantum models [F., Lumbreras, Winter 2024]...
- ... but all FCSs can be seen as generated by sequential maps on a memory system described by a GPT.
- Generalization of cb norms from GPT structure.

GPT dictionary

- Convex cone: $C \subseteq V$, s.t. $\alpha x + \beta y \in C$, for all $\alpha, \beta \geq 0$, $x, y \in C$.
- Dual cone of C: $C^* = \{f \in V^* | f(x) \geq 0 \, \forall x \in C\}$.

	Quantum	GPT
positive observables	$A \geq 0$	$v \geq_C 0$ ($= v \in C$)
identity observable	$\mathbb{1}$	$e \geq_C 0$ s.t. $\forall x \in V \; \exists \lambda \geq 0 : \lambda e \geq_C x$
states	$\rho \geq 0, \mathrm{Tr}[\rho] = 1$	$f \geq_{C'}, C' \subseteq C^*, f(e) = 1$
measurements	POVM	$\{x_i\}_{i=1}^m, x_i \geq_C 0, \sum_{i=1}^m x_i = e$
state transformations	positive maps	cone-preserving maps
norms	$\|A\|_\infty$, $\|\rho\|_1$	$\|v\|_e = \{\min \lambda \geq 0 : \pm v \leq \lambda e\}$. $\|f\|_{e,*} = \{\max |f(v)| : v \in V, \|v\|_e \leq 1\}$.

General probabilistic theories entangled with quantum systems

- $V = \mathbb{C}^m$.
- Cone $C_1 \subseteq \mathbb{R}^m \subseteq \mathbb{C}^m$ with unit e defines a GPT.
- Positive observables of a general-quantum composite system are cones $\mathcal{C} = \{C_n\}_{n\in\mathbb{N}}$
$$C_n \subseteq (\mathbb{C}^m \otimes \mathbb{M}_{n,n}(\mathbb{C}))_h,$$
with units $e_n := e \otimes \mathbf{1}_n$
- Consistency of complete positivity on the quantum system: for each $k \times n$-matrix $M \in \mathbb{M}_{k,n}$ it holds that $M^\dagger C_k M \subseteq C_n$.
- $(V, \mathcal{C}, e)$ is called an **operator system**.
- A linear map $\mathbb{K} : V \to V$ is completely positive if $(\mathrm{id}_n \otimes \mathbb{K})(C_n) \subseteq C_n$ for every n.
- A map $\mathbb{K} : V \to V'$ between two operator systems (with cones $\mathcal{C}, \mathcal{C}'$) is completely positive if $(\mathrm{id}_n \otimes \mathbb{K})(C_n) \subseteq C'_n$ if for every n.

Norms for operator systems

- $(V, \mathcal{C}, e)$ provides norms $||X||_{e_n} := \min\{\lambda : \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix} \leq \lambda \begin{pmatrix} e_n & 0 \\ 0 & e_n \end{pmatrix}\}$
- Completely bounded norm of $\mathbb{K} : \mathbb{M}_d(\mathbb{C}) \otimes V \to V$:

$$||\mathbb{K}||_{cb} = \sup_n ||\mathrm{id}_{\mathbb{M}_{n,n}(\mathbb{C})} \otimes \mathbb{K}||_{\mathbf{1}_d \otimes e_n \to e_n} \tag{8}$$

- If $\mathbb{K}$ is completely positive and $\mathbb{K}e = e$ (unital), then $||\mathbb{K}||_{cb} = 1$.
- Unital completely positive maps are the "physical" maps between operator systems.

Operator systems from FCS

Positivity of ω implies convex geometry structure.
Recall observable realization $(\mathbb{C}^m, e, \rho, \mathbb{K})$

- $x \in \mathbb{C}^m$ can be written as $\mathbb{K}^t(X \otimes e)$ for some t, $X \in \mathbb{M}_{d^t_{\mathcal{A}}, d^t_{\mathcal{A}}}(\mathbb{C})$.
- $$C_1 := \{x \in \mathbb{C}^m : \exists t \in \mathbb{N}, X \geq 0 \, \text{s.t.} \; x = \mathbb{K}^t(X \otimes e)\},$$
  unit $e = U^\top \Omega(\mathbf{1}) \in C_0$, $\rho \in C_1^*$, $\rho e = 1$.
- $$C_n := \{x \in \mathbb{C}^m \otimes \mathbb{M}_n(\mathbb{C}) \, ; \exists t \in \mathbb{N}, X \geq 0 \, \text{s.t.} \, x = (\mathrm{id}_{\mathbb{M}_n(\mathbb{C})} \otimes \mathbb{K}^t)(X \otimes e)\}.$$
- With these choices $\mathbb{K}$ is **completely positive and unital**, therefore $||\mathbb{K}||_{cb} = 1$.
- Telescoping sum trick gives the desired bound.

Operator systems from FCS

Connection with Hilbert-Schmidt errors

Equivalence with Euclidean norm on $\mathbb{C}^m \otimes \mathbb{M}_{n,n}(\mathbb{C})$:

$$\sigma_m(\Omega)\|\cdot\|_{e_n} \leq \|\cdot\|_{n,2} \leq \sqrt{n}\|\cdot\|_{e_n}$$

- It follows that $\|\tilde{\mathbb{K}} - \hat{\mathbb{K}}\|_{\mathbf{1}_{d_A}\otimes e \to e, cb} \leq m\frac{\sqrt{d_A}}{\sigma_m(\Omega)}\|\tilde{\mathbb{K}} - \hat{\mathbb{K}}\|_{2\to 2}$,
- and $\|\tilde{\mathbb{K}} - \hat{\mathbb{K}}\|_{2\to 2} \leq 4\left(\frac{\|\Omega-\hat{\Omega}\|_2}{\sigma_m(\Omega)^2} + \frac{\|\Omega_{(\cdot)}-\hat{\Omega}_{(\cdot)}\|_2}{3\sigma_m(\Omega)}\right)$ via matrix perturbation theory