Efficient Hamiltonian, structure and trace distance learning of Gaussian states

Marco Fanizza[1,2], joint work with

Cambyse Rouzé[1]

Daniel Stilck França[2]

[1]Inria Saclay, Institut Polytechnique de Paris

[2]University of Copenhagen

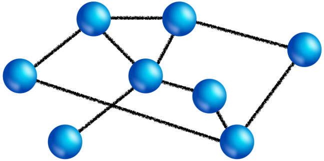
arXiv:2411.03163

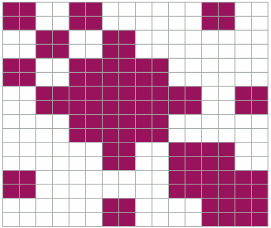
Learning the Hamiltonian of Gaussian states

- Learn Hamiltonian $\mathcal{H}$ of a CV system of m modes from N copies of Gibbs state $\rho = \frac{e^{-\beta\mathcal{H}}}{\mathrm{Tr}[e^{-\beta\mathcal{H}}]}$
- Access to **thermal equilibrium** states encoding dynamics
- In many-body physics interactions are local: interaction **graph** $G = ([m], \mathsf{E})$ of degree $\Delta - 1$
- Mixed Gaussian states are Gibbs states of quadratic Hamiltonian in $R = (X_1, P_1, ..., X_m, P_m)$,

$$\beta\mathcal{H} = \sum_{i,j=1}^{2m} (R_i - t_i) H_{ij} (R_j - t_j)$$

$\rightarrow$ learn Hamiltonian matrix H





Background: Hamiltonian learning from Gibbs states

Qudit systems:

- Recent work on computational and sample efficient algorithms for known graphs, but no efficient graph learning algorithm[1][2][3][4][5].
- $N \geq \mathrm{poly}(m, \epsilon^{-2^{O(\beta)}})$ copies and $N^{O(1)}$ time suffice[4]
- $O(\log m)$ in general and $O\left(\frac{e^{(\mathrm{poly}(\beta))}}{\beta^2 \epsilon^2} \log m \, \mathrm{poly}(\log \frac{1}{\epsilon})\right)$ for lattices[5]

Classical Random Markov fields/Gaussian Graphical Models:
$O(\frac{\log m}{\epsilon^2})$ for both Hamiltonian and graph learning

[1] Anshu, Arunachalam, Kuwahara, Soleimanifar 2021
[2] Haah, Kothari, Tang 2021
[3] Rouzé, Stilck França, Onorati, Watson 2024
[4] Bakshi, Liu, Moitra, Tang 2024
[5] Chen, Anshu, Nguyen 2025

Learning with continuous variable systems

Systems with quadratic CV Hamiltonians:

- Electromagnetic fields (free space, fibre, cavity)
- Optomechanical systems
- Trapped ions (vibrational degrees of freedom)
- Cold atoms in optical lattices

...and analog simulators of lattice bosonic Hamiltonians

- Periwal et al. 2021: 18 sites, 10^4 Rubidium atoms/site, magnetization field
- Youssefi et al. 2022: 10 sites of a superconducting optomechanical systems
- Senanian et al. 2023: Photonic synthetic lattices with up to 10^5 sites

Learning with continuous variable systems

Rapidly developing fields about finite sample size guarantees:

- Trace distance learning[6]
- Gaussian trace distance bounds[7] [8] [9]
- Gaussian testing[10]
- Gaussian unitary learning[11]
- Hamiltonian learning from time evolution[12] [13]

[6] Mele, Mele, Bittel, Eisert, Giovannetti, Lami, Leone, Oliviero 2024
[7] Bittel, Mele, Mele, Tirone, Lami 2024
[8] Bittel, Mele, Eisert, Mele 2025
[9] Holevo 2024a,2024b
[10] Girardi, Witteveen, Mele, Bittel, Oliviero, Gross, Walter, 2025
[11] Fanizza, Iyer, Lee, Mele, Mele, 2025
[12] Li, Tong, Gefen, Ni, Ying 2024
[13] Möbus, Bluhm, Caro, Werner, Rouzé 2023

Our contributions

We establish the following:

- Sample optimal scaling in precision for trace distance
- Efficient Gaussian Hamiltonian learning (known graph)
- Efficient graph learning (threshold κ)

Task	Sample compl.
Trace distance	$\mathcal{O}\left(\frac{m^3}{\epsilon^2}\log(\frac{m}{\delta})\right)$
Hamiltonian	$\mathcal{O}\left(\log\left(\frac{m}{\delta}\right)\frac{1}{\epsilon^{2+\gamma}}\right) \quad \forall\gamma > 0$
(lattices)	$\mathcal{O}\left(\log\left(\frac{m}{\delta}\right)\frac{1}{\epsilon^2}\mathrm{poly}(\log\frac{1}{\epsilon})\right)$
Graph	$\mathcal{O}\left(\log\left(\frac{m}{\delta}\right)\frac{1}{\kappa^{2+\gamma}}\right) \quad \forall\gamma > 0$

Non-entangled (heterodyne) measurements and classical post-processing, requiring $O(\mathrm{poly}(mN))$ time for Hamiltonian learning.
Implicit dependence on $\|H\|_\infty$ (equivalent to $\beta\Delta$) and $\|H^{-1}\|_\infty$, see paper.

Classical vs Quantum Gaussian Hamiltonian learning

Classical Gaussian distributions

- $P(x) \sim e^{-(x-t)^\top \Theta (x-t)}$
- $\Sigma_{ij} := \mathbb{E}[(x_i - t_i)(x_j - t_j)]$
- $\Theta = (2\Sigma)^{-1}$

Quantum Gaussian states

- $\rho \sim e^{-(R-t)^\top H (R-t)}$
- $V_{ij} := \mathbb{E}[\frac{1}{2}\{R_i - t_i,\, R_j - t_j\}]$
- $[R_j, R_k] = i\Omega_{jk}$
- $H = \frac{1}{2}\log\left(\frac{2i\Omega V + I}{2i\Omega V - I}\right) i\Omega$

Classical Gaussian Hamiltonian learning

- **Task**: Learn Hamiltonian Θ from N sample from Gaussian distribution

$$P(x) = \frac{e^{-x^\top \Theta x}}{\int_{\mathbb{R}^m} e^{-x^\top \Theta x} \mathrm{d}^m x},$$

  where $\Theta \in \mathbb{M}_m(\mathbb{R}), \Theta > 0, \Theta = \Theta^\top$.

- Covariance matrix:

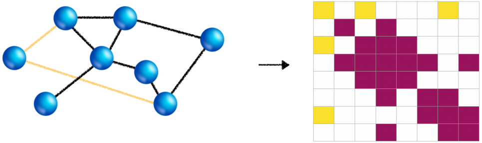
$$\Sigma_{ij} := \mathbb{E}[x_i x_j] = \frac{(\Theta^{-1})_{ij}}{2}$$

- Straightforward strategy: estimate Σ as $\hat{\Sigma}$, invert it to get $\hat{\Theta}$.
  Error propagation from $A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}$:

$$\|\Theta - \hat{\Theta}\|_\infty \leq \frac{1}{2}\|\Sigma^{-1}\|_\infty \|\hat{\Sigma}^{-1}\|_\infty \|\Sigma - \hat{\Sigma}\|_\infty$$

Classical Gaussian Hamiltonian learning, using sparsity

- Assumption: G has degree $\Delta - 1 \rightarrow \; \Theta$ is Δ-sparse.



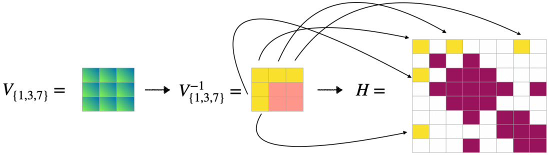
- **Conditional independence**: $P(x_1|x_2,..,x_8) = P(x_1|x_3,x_7)$
- We can obtain H_{j1} inverting the local V for vertices $\{1,3,7\}$.
  Marginal over vertices $\{1,3,7\}$ (still Gaussian):

$$P_{\{1,3,7\}}(x_1,x_3,x_7) \sim \int_{\mathbb{R}^{(m-3)}} e^{-\sum_{j\in\{1,3,7\}}(x_1\Theta_{1j}x_j + x_j\Theta_{j1}x_1)+\dots}$$

$$\sim e^{-\sum_{j\in\{1,3,7\}}(x_1\Theta_{1j}x_j + x_j\Theta_{j1}x_1) - \sum_{i,j\in\{1,3,7\},i\neq 1,j\neq 1}(x_i\Theta'_{ij}x_j)}$$

Classical Gaussian Hamiltonian learning, using sparsity

- **Sparse inverse from local inversions:**



- Schur's complement explanation: for $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \mathbb{C}^{m_1+m_2} \times \mathbb{C}^{m_1+m_2}$,

$$M^{-1} = \begin{pmatrix} N \equiv (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{pmatrix}.$$

First row of B and first column of C are zero $\rightarrow$ $A_{1j} = (N^{-1})_{1j}$, $A_{j1} = (N^{-1})_{j1}$

Classical Gaussian Hamiltonian learning, using sparsity

- Estimates $\hat{V}_{ij} = \frac{1}{N}\sum_{t=1}^{N} x_i^{(t)} x_j^{(t)}$.
  For any $\delta \in (0,1)$, $\epsilon \in (0,1/2)$ and if $N = \Omega(\frac{1}{\epsilon^2}\log\left(\frac{m+1}{\delta}\right))$ then
  $|\hat{V}_{i,j} - V_{i,j}| \leq \epsilon \qquad \forall i,j \in [m]$.
- Naive inverse: invert $\hat{V} \rightarrow \hat{\Theta}$
- Local inverse: construct $\hat{H}$ from local inverses of $\hat{V}$
- Advantage of local inverse:
  $\|V - \hat{V}\|_\infty \leq m\epsilon \rightarrow \ |\Theta_{ij} - \hat{\Theta}_{ij}| \leq \mathcal{O}(m\epsilon)$,
  $\|V_{\{1,3,7\}} - \hat{V}_{\{1,3,7\}}\|_\infty \leq \Delta\epsilon \rightarrow |\Theta_{ij} - \hat{\Theta}_{ij}| \leq \mathcal{O}(\Delta\epsilon)$

Classical Gaussian Hamiltonian and structure learning, using sparsity

Classical algorithms estimate both Hamiltonian and graph strucutre efficiently, assuming sparsity

- Graphical LASSO [Friedman et al. 2008, Yuan, Lin 2007]: maximizing the l_1-regularized log-likelihood. Gives sample complexity upper bound $O(\frac{\Delta^2}{\alpha^2\epsilon^2}\log m)$ sample complexity [Ravikumar et al. 2011], where α is a parameter encoding a certain incoherence condition of the precision matrix.

- Other approaches: one row at a time, via LASSO [Meinshausen, Bühlmann 2006], Danzig [Yuan 2010], or l_1-constrained optimization (CLIME) [Cai et al. 2011], still sensitive to condition number.

- Graph selection: [Misra et al, 2020] shows a sample complexity $O(\frac{\Delta \log m}{\kappa^2})$, with κ being a lower bound on the relative strenghts and no condition number dependence, and matching lower bounds in [Wang et al., 2010]. The tradeoff is a worse scaling in computational complexity.

(Demistifying) Hamiltonian-Covariance relations

- V to H

$$H = \frac{1}{2}\log\left(\frac{2i\Omega V + I}{2i\Omega V - I}\right) i\Omega = \frac{1}{2}\log\left(I + \frac{2}{i\Omega(2V - i\Omega)}\right) i\Omega.$$

$$= \int_0^\infty \frac{1}{2i\Omega V + \frac{t-1}{t+1} I} \frac{dt}{(t+1)^2} i\Omega$$

$$= (2V - i\Omega)^{-1} i\Omega \int_0^\infty \frac{1}{I + \frac{2t}{t+1}(2V - i\Omega)^{-1} i\Omega} \frac{dt}{(t+1)^2} i\Omega$$

- Error propagation bounds can be derived for approximation of V or $(2V - i\Omega)^{-1}$.

Quantum Gaussian trace distance learning

- Estimate V as $\hat{V}$ via heterodyne
  (sample from Gaussian distribution with covariance matrix $V + I$)
- Let $D(\rho\|\sigma) = \mathrm{Tr}[\rho \log \rho] - \mathrm{Tr}[\rho \log \sigma]$
- If $|V_{ij} - \hat{V}_{ij}| \leq \epsilon$, then $|H_{ij} - \hat{H}_{ij}| = O(m\epsilon)$ and

$$\begin{aligned} \|\rho - \hat{\rho}\|_1 &\leq \sqrt{2D(\rho\|\hat{\rho}) + 2D(\hat{\rho}\|\rho)} = \sqrt{\mathrm{Tr}[(\hat{H} - H)(V - \hat{V})]} \\ &= O(m^{3/2}\epsilon) \end{aligned}$$

Previous bounds[14] were $\mathcal{O}(\sqrt{\epsilon})$. See also concurrent work [15]

[14] Mele et al. 2024, Holevo 2024a, 2024b

[15] Bittel et al. 2024

Quantum Gaussian trace distance learning

Theorem (Learning Gaussian states in trace distance (informal))

Let $\rho(t, H)$ be a Gaussian state on m modes, with $H = SDS^\top$. Then, for $1 > \epsilon, \delta > 0$, it suffices to measure

$$N = \mathcal{O}(\epsilon^{-2} m^3 \ln(m\delta^{-1}) \mathrm{poly}(\|S\|_\infty, (e^{2\|D\|_\infty} - 1)(1 - e^{-2\|D^{-1}\|_\infty^{-1}})^{-1}, \max_i |t_i|))$$

copies of ρ with heterodyne to obtain an estimate of ρ up to trace distance ϵ with success probability at least $1 - \delta$.

Quantum Gaussian Hamiltonian learning, approximate sparsity

$H = f((2V - i\Omega)^{-1})$, with $(2V - i\Omega)^{-1}$ **approximately sparse** matrix:

$$(2V - i\Omega)^{-1} = -\frac{I - e^{+2Hi\Omega}}{2}(i\Omega) = \frac{1}{2}\sum_{n=1}^{\infty}\frac{(2Hi\Omega)^n}{n!}(i\Omega)$$

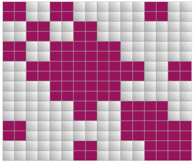
$$= \frac{1}{2}\sum_{n=1}^{l}\frac{(2Hi\Omega)^n}{n!}(i\Omega) + E\,,$$

with $\|E\|_\infty$ decreasing faster than exponentially in truncation degree.

$H = S^\top D S$ (normal form)

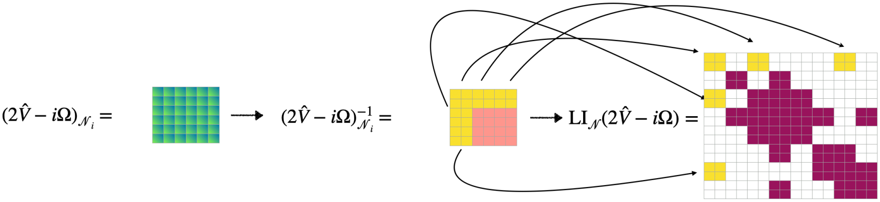
$$\|E\|_\infty \leq \frac{\|S\|_\infty^2}{2}\left(\frac{(2\|D\|_\infty)^{l+1}e^{2\|D\|_\infty}}{(l+1)!}\right)$$

$$l \sim \frac{\log\frac{1}{\epsilon}}{\log\log\frac{1}{\epsilon}} \rightarrow \|E\|_\infty \sim \epsilon$$

$(2V - i\Omega)^{-1} =$ 

Quantum Gaussian Hamiltonian learning, approximate sparsity

- $\mathcal{N} = \{\mathcal{N}_i(l)\}_{i\in[m]}$ set of neighborhoods of radius l
- **Quantum local inversion:** For $i = 1, ..., m$:



- $\|\mathrm{LI}_{\mathcal{N}}(2\hat{V} - i\Omega) - (2V - i\Omega)^{-1}\|_\infty = O(\Delta^l \|E\|_\infty + \Delta^{2l}\zeta)$, where ζ is the entry-wise error of $\hat{V}$.
- $\hat{H} = \mathrm{LI}_{\mathcal{N}}(2\hat{V} - i\Omega)i\Omega \int_0^\infty \frac{I}{I + \frac{t}{t+1} 2\,\mathrm{LI}_{\mathcal{N}}(2\hat{V} - i\Omega)i\Omega} \frac{dt}{(t+1)^2} i\Omega$

Quantum Gaussian Hamiltonian Learning

We can take $l = \left\lfloor 2\Delta d_{\max} e \exp\left(W\left(\frac{\log \frac{C}{\varepsilon'}}{2\Delta d_{\max} e}\right)\right)\right\rfloor$, $\zeta = C' \frac{\epsilon}{\Delta^{2l}}$ and obtain

Theorem (Gaussian Hamiltonian learning)

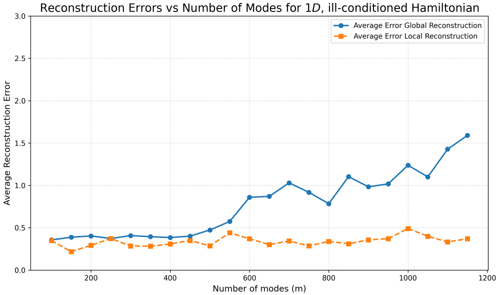
Let ρ be a GS with Hamiltonian H of maximal degree Δ. Then it suffices to take

$$N = \mathcal{O}\left(\frac{1}{\epsilon^{2+\gamma}} \log\left(\frac{m}{\delta}\right)\right) \quad \forall \gamma > 0 \tag{1}$$

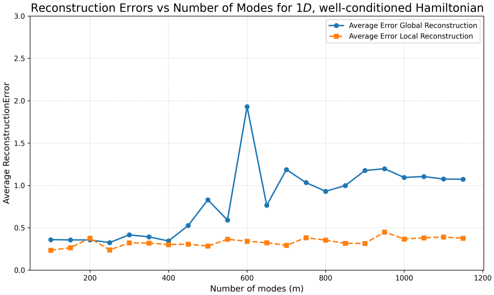
copies of ρ suffice to obtain an estimate $\hat{H}$ satisfying $\|H - \hat{H}\|_\infty \leq \epsilon$ with probability at least $1 - \delta$.

Numerical examples

1D-Hamiltonians $H = (2+c)I - |0\rangle\langle 0| - (\sum_i |i\rangle\langle i+1| + h.c.)$,
local inversion vs plug-in, 10^4 samples, 5 repetitions



(a) Ill-conditioned Hamiltonian $(c = 0)$



(b) Well-conditioned Hamiltonian $(c = 0.1)$

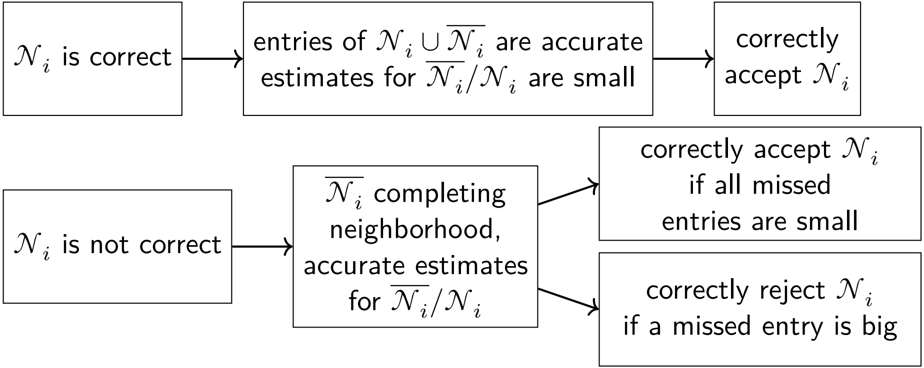
Graph Learning for Gaussian States

Promise: $0 < \kappa \leq \min_{(i,j)\in \mathsf{E}} \max_{\delta_1,\delta_2\in\{0,1\}} |H_{2i-\delta_1,2j-\delta_2}|$
Algorithm sketch:

- Iterate over all neighborhoods $\mathcal{N}_i$ and adversarial neighborhoods $\overline{\mathcal{N}}_i$, $|\mathcal{N}_i| = |\overline{\mathcal{N}}_i| = \Delta^l$
- For each choice $\mathcal{N}_i$ and adversarial neighborhoods $\overline{\mathcal{N}}_i$, perform local inversion with $\mathcal{N}_i \cup \overline{\mathcal{N}}_i$. If in the inverse there are never large entries on the blocks (i,j), $j \in \overline{\mathcal{N}}_i/\mathcal{N}_i$, we accept $\mathcal{N}_i$
- Compute $\mathrm{LI}_{\mathcal{N}}(2\hat{V} - i\Omega)$ and thus $\hat{H}$
- Remove spurious edges (below threshold)

Graph Learning for Gaussian States

Why it works:



- With the accepted neighborhood we obtain a good estimate of the Hamiltonian. We obtain a logarithmic scaling in the number of modes because we only need to invert locally.

Graph Learning for Gaussian States

Theorem (Learning the graph of GS)

Let H be a Hamiltonian with graph G of degree Δ and edge set E, the condition $0 < \kappa \leq \min_{(i,j)\in\mathsf{E}} \max_{\delta_1,\delta_2\in\{0,1\}} \left|H_{2i-\delta_1,2j-\delta_2}\right|$. Then it suffices to take

$$N = \mathcal{O}\left(\frac{1}{\kappa^{2+\gamma}}\log\left(\frac{m}{\delta}\right)\right) \quad \forall\gamma > 0 \tag{2}$$

copies of ρ suffice to learn the graph G with probability of success at least $1-\delta$.

Conclusions

Summary:

- Sample complexity upper bounds for Hamiltonian learning for bosonic Gaussian states
- Efficient learning algorithms with practical measurement schemes (heterodyne)
- Optimal scaling in precision for trace distance estimation

Future directions:

- Lower bounds
- Improve dependence on ϵ, Δ, $\|H\|_\infty$ and $\|H^{-1}\|_\infty$
- Graph learning with relative strength promise
- Fermionic states