Higher moment theory and learnability of bosonic states

Joseph T. Iosue, Y. Wang, I. Datta, S. Ghosh, C. Oh, B. Fefferman, A. V. Gorshkov

arXiv:2510.01610

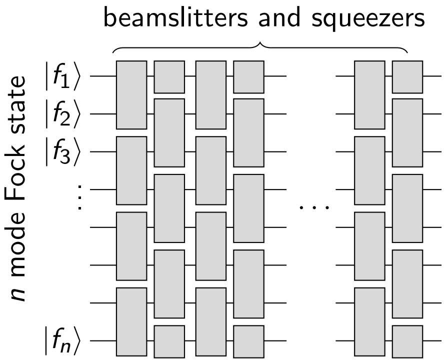
UT Austin QI seminar, March 2, 2026

Introduction

- Learning (properties of) quantum states as efficiently as possible is a fundamental task in quantum information
- There is a lot of recent work on learning continuous variable (CV) states (eg. Mele, Mele, et al. 2024; Zhao, Liao, et al. 2024; Bittel, Mele, et al. 2025; Fanizza, Iyer, et al. 2025)
- Most *sample and time efficient* learning algorithms work for **Gaussian** or **near-Gaussian** states
- Learning a generic n-mode CV state requires $\sim$ (energy)^n samples (Mele, Mele, et al. 2024)

Our goal

- Our goal is to find interesting, **highly non-Gaussian** classes of CV states with *sample and time efficient* learning algorithms
- Motivated by an open problem proposed in Aaronson and Grewal 2023, we start with **BosonSampling states**



Overview

Background I

- A bosonic system of n modes is described by the creation and annihilation operators $a_1, \ldots, a_n, a_1^\dagger, \ldots, a_n^\dagger$
- Equivalently, by position and momentum operators $\boldsymbol{r} = (x_1, \ldots, x_n, p_1, \ldots, p_n)$
- These obey the CCR $[r_i, r_j] = \mathrm{i}\Omega_{ij}$, with

$$\Omega = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}$$

- A unitary transformation must perserve the CCR

$$[\mathcal{U} r_i \mathcal{U}^\dagger, \mathcal{U} r_j \mathcal{U}^\dagger] = \mathcal{U}[r_i, r_j]\mathcal{U}^\dagger = \mathrm{i}\mathcal{U}\Omega_{ij}\mathcal{U}^\dagger = \mathrm{i}\Omega_{ij}$$

Background II

- A unitary transformation is called *Gaussian* if it acts affinely on $\boldsymbol{r}$

$$\mathcal{U}_{S,\boldsymbol{d}}^\dagger r_i \mathcal{U}_{S,\boldsymbol{d}} = d_i + \sum_j S_{ij} r_j$$

- To preserve the CCR, S must satisfy $S\Omega S^T = \Omega$

$$S \in \mathrm{Sp}(2n, \mathbb{R}) := \left\{ S \in \mathrm{GL}(2n, \mathbb{R}) \mid S\Omega S^T = \Omega \right\}$$

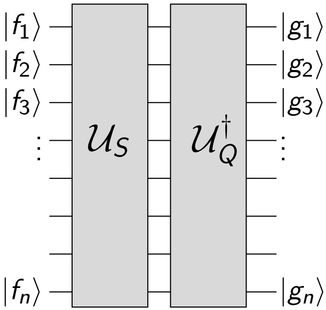
- Easy to see the composition $\mathcal{U}_{S,\boldsymbol{d}} \mathcal{U}_{R,\boldsymbol{e}} = \mathcal{U}_{SR,R\boldsymbol{d}+\boldsymbol{e}}$
- Thus, **the group of Gaussian unitaries is** $\mathrm{Sp}(2n, \mathbb{R}) \ltimes \mathbb{R}^{2n}$

Problem statement (Aaronson and Grewal 2023)

- Let $|\boldsymbol{f}\rangle$ denote an unknown Fock state on n modes, $\boldsymbol{f} = (f_1, \ldots, f_n) \in \mathbb{N}_0^n$
- Let $\mathcal{U}_S$ denote a Gaussian unitary specified by an unknown matrix $S \in \mathrm{Sp}(2n, \mathbb{R})$
- Let $|\psi\rangle = \mathcal{U}_S |\boldsymbol{f}\rangle$

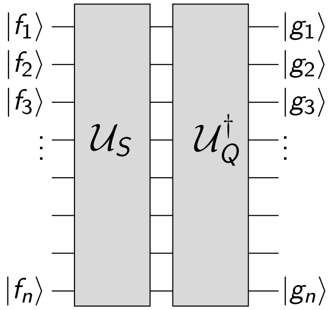
Problem statement (Aaronson and Grewal 2023)

- Let $|\boldsymbol{f}\rangle$ denote an unknown Fock state on n modes, $\boldsymbol{f} = (f_1, \ldots, f_n) \in \mathbb{N}_0^n$
- Let $\mathcal{U}_S$ denote a Gaussian unitary specified by an unknown matrix $S \in \mathrm{Sp}(2n, \mathbb{R})$
- Let $|\psi\rangle = \mathcal{U}_S |\boldsymbol{f}\rangle$
- Given N copies of ψ, we want to learn ψ to ϵ precision
- Find a $\boldsymbol{g}$ and Q such that $|\langle \boldsymbol{g}| \mathcal{U}_Q^\dagger \mathcal{U}_S |\boldsymbol{f}\rangle| \geq 1 - \varepsilon$



Problem statement (Aaronson and Grewal 2023)

- Let $|\boldsymbol{f}\rangle$ denote an unknown Fock state on n modes, $\boldsymbol{f} = (f_1, \ldots, f_n) \in \mathbb{N}_0^n$
- Let $\mathcal{U}_S$ denote a Gaussian unitary specified by an unknown matrix $S \in \mathrm{Sp}(2n, \mathbb{R})$
- Let $|\psi\rangle = \mathcal{U}_S |\boldsymbol{f}\rangle$
- Given N copies of ψ, we want to learn ψ to ϵ precision
- Find a $\boldsymbol{g}$ and Q such that $|\langle \boldsymbol{g}| \mathcal{U}_Q^\dagger \mathcal{U}_S |\boldsymbol{f}\rangle| \geq 1 - \varepsilon$



- How large must N be? *Want it be be* $\leq \mathrm{poly}(\mathrm{energy})$
- What is the measurement scheme?
- How hard is the classical postprocessing? *Want it be be* $\leq \mathrm{poly}(n)$

Analogous fermionic problem

- The analogous fermionic problem was solved in Aaronson and Grewal 2023
- Crucial difference: fermionic Fock states are Gaussian, bosonic Fock states are highly non-Gaussian

Moments

- We will solve this problem with *moments*

$$
\begin{aligned}
\Lambda^{(2t)}_{i_1,\dots,i_t;j_1,\dots,j_t} &= \langle\psi|\, r_{i_1}\dots r_{i_t} r_{j_1}\dots r_{j_t}\,|\psi\rangle\,, \\
(\Lambda^{(2t)}_0)_{i_1,\dots,i_t;j_1,\dots,j_t} &= \langle \boldsymbol{f}|\, r_{i_1}\dots r_{i_t} r_{j_1}\dots r_{j_t}\,|\boldsymbol{f}\rangle\,.
\end{aligned}
$$

Moments

- We will solve this problem with *moments*

$$\Lambda^{(2t)}_{i_1,\ldots,i_t;j_1,\ldots,j_t} = \langle\psi|\, r_{i_1}\ldots r_{i_t} r_{j_1}\ldots r_{j_t}\,|\psi\rangle\,,$$
$$(\Lambda^{(2t)}_0)_{i_1,\ldots,i_t;j_1,\ldots,j_t} = \langle\boldsymbol{f}|\, r_{i_1}\ldots r_{i_t} r_{j_1}\ldots r_{j_t}\,|\boldsymbol{f}\rangle\,.$$

- How do you measure moments? For the purposes of this talk, the upshot is that poly$(1/\varepsilon)$ **copies suffice to learn the moments to ε precision** (Welsch, Vogel, et al. 1999)

Moments

- We will solve this problem with *moments*

$$\Lambda^{(2t)}_{i_1,\ldots,i_t;j_1,\ldots,j_t} = \langle\psi| \, r_{i_1}\ldots r_{i_t} r_{j_1}\ldots r_{j_t} \, |\psi\rangle \,,$$
$$(\Lambda^{(2t)}_0)_{i_1,\ldots,i_t;j_1,\ldots,j_t} = \langle \boldsymbol{f}| \, r_{i_1}\ldots r_{i_t} r_{j_1}\ldots r_{j_t} \, |\boldsymbol{f}\rangle \,.$$

- How do you measure moments? For the purposes of this talk, the upshot is that poly$(1/\varepsilon)$ **copies suffice to learn the moments to ε precision** (Welsch, Vogel, et al. 1999)

- For the rest of this section, we will assume that *the moments are known perfectly*

Moment transformations

- Recall $S \in \mathrm{Sp}(2n, \mathbb{R})$ acts as
$$\mathcal{U}_S^\dagger r_i \mathcal{U}_S = \sum_j S_{ij} r_j$$
- How do moments transform under **arbitrary** Gaussian unitaries?

$$\begin{aligned}
\Lambda^{(2t)}_{i_1,\ldots,i_t;j_1,\ldots,j_t} &= \langle \boldsymbol{f} | \mathcal{U}_S^\dagger r_{i_1} \ldots r_{i_t} r_{j_1} \ldots r_{j_t} \mathcal{U}_S | \boldsymbol{f} \rangle \\
&= \sum_{k,\ell} S_{i_1,k_1} \ldots S_{i_t,k_t} S_{j_1,\ell_1} \ldots S_{j_1,\ell_1} \langle \boldsymbol{f} | r_{k_1} \ldots r_{k_t} r_{\ell_1} \ldots r_{\ell_t} | \boldsymbol{f} \rangle \\
&= (S^{\otimes t} \Lambda_0^{(2t)} S^{T \otimes t})_{i_1,\ldots,i_t;j_1,\ldots,j_t}
\end{aligned}$$

Passive moment transformations

- Recall $W \in U(n)$ acts as

$$\mathcal{U}_W^\dagger a_i \mathcal{U}_W = \sum_j W_{ij} a_j$$

- How do moments transform under **passive** Gaussian unitaries?

$$\sigma^{(2t)}_{i_1,\ldots,i_t;j_1,\ldots,j_t} = \langle\psi| a_{i_1} \ldots a_{i_t} a^\dagger_{j_1} \ldots a^\dagger_{j_t} |\psi\rangle$$

$$(\sigma_0^{(2t)})_{i_1,\ldots,i_t;j_1,\ldots,j_t} = \langle\boldsymbol{f}| a_{i_1} \ldots a_{i_t} a^\dagger_{j_1} \ldots a^\dagger_{j_t} |\boldsymbol{f}\rangle$$

$$\sigma^{(2t)} = W^{\otimes t} \sigma_0^{(2t)} W^{\dagger\otimes t}$$

Reducing the problem

- By using **second moments** and linear algebra, we can reduce the problem to when the Gaussian unitary is **passive** and the Fock state is $|b, \ldots, b\rangle$
- If we were dealing with Gaussian states, this would be sufficient

Reducing the problem

- By using **second moments** and linear algebra, we can reduce the problem to when the Gaussian unitary is **passive** and the Fock state is $|b, \ldots, b\rangle$
- If we were dealing with Gaussian states, this would be sufficient
- But for us **second moments are now useless** — easy to check that for $\boldsymbol{f} = (b, \ldots, b)$, $\sigma_0^{(2)} = (b+1)\mathbb{I}$
- Therefore, $\sigma^{(2)} = W\sigma_0^{(2)}W^\dagger = \sigma_0^{(2)}$, so we learn nothing about W by measuring second moments

Reducing the problem

- By using **second moments** and linear algebra, we can reduce the problem to when the Gaussian unitary is **passive** and the Fock state is $|b, \ldots, b\rangle$
- If we were dealing with Gaussian states, this would be sufficient
- But for us **second moments are now useless** — easy to check that for $\boldsymbol{f} = (b, \ldots, b)$, $\sigma_0^{(2)} = (b+1)\mathbb{I}$
- Therefore, $\sigma^{(2)} = W\sigma_0^{(2)}W^\dagger = \sigma_0^{(2)}$, so we learn nothing about W by measuring second moments
- So now we use **fourth moments**, $\sigma^{(4)}$

Learning algorithm

Given the fourth moments $\sigma^{(4)} = W^{\otimes 2}\sigma_0^{(4)} W^{\dagger\otimes 2}$ for the state $\mathcal{U}_W \left|b, \ldots, b\right\rangle$, determine W

Learning algorithm

Given the fourth moments $\sigma^{(4)} = W^{\otimes 2}\sigma_0^{(4)} W^{\dagger\otimes 2}$ for the state $\mathcal{U}_W \left|b,\ldots,b\right\rangle$, determine W

- Because $\sigma_0^{(4)}$ is a matrix on $(\mathbb{C}^n)^{\otimes 2}$, we can represent it in bra-ket notation using the standard basis $\left|1\right\rangle, \ldots, \left|n\right\rangle$ of $\mathbb{C}^n$

$$\sigma_0^{(4)} = (b+1)^2(\mathbb{I} + U_{\mathrm{SWAP}}) - b(b+1)\sum_{i=1}^{n} \left|i,i\right\rangle\!\left\langle i,i\right|,$$

Learning algorithm

Given the fourth moments $\sigma^{(4)} = W^{\otimes 2}\sigma_0^{(4)} W^{\dagger\otimes 2}$ for the state $\mathcal{U}_W \left|b,\ldots,b\right\rangle$, determine W

- Because $\sigma_0^{(4)}$ is a matrix on $(\mathbb{C}^n)^{\otimes 2}$, we can represent it in bra-ket notation using the standard basis $\left|1\right\rangle,\ldots,\left|n\right\rangle$ of $\mathbb{C}^n$

$$\sigma_0^{(4)} = (b+1)^2(\mathbb{I} + U_{\mathrm{SWAP}}) - b(b+1)\sum_{i=1}^{n} \left|i,i\right\rangle\left\langle i,i\right|,$$

- Denote the i^{th} column vector of W by $\left|w_i\right\rangle$. Given access to $\sigma^{(4)}$, we can thus form the matrix

$$\begin{aligned} A &= \frac{1}{b(b+1)}\left((b+1)^2(\mathbb{I} + U_{\mathrm{SWAP}}) - \sigma^{(4)}\right) \\ &= \sum_{i=1}^{n}(\left|w_i\right\rangle \otimes \left|w_i\right\rangle)(\left\langle w_i\right| \otimes \left\langle w_i\right|), \end{aligned}$$

Learning algorithm

Given access to a matrix $A = \sum_{i=1}^{n} (|w_i\rangle \otimes |w_i\rangle)(\langle w_i| \otimes \langle w_i|)$, learn $|w_i\rangle$ (up to phases and permutations)

Learning algorithm

Given access to a matrix $A = \sum_{i=1}^{n} (|w_i\rangle \otimes |w_i\rangle)(\langle w_i| \otimes \langle w_i|)$, learn $|w_i\rangle$ (up to phases and permutations)

- Unitarily diagonalize A and take the $+1$ eigenvalues — we learn the eigenspace $\mathrm{span}\{|w_i\rangle \otimes |w_i\rangle \mid i = 1, \ldots, n\}$

Learning algorithm

Given access to a matrix $A = \sum_{i=1}^{n} (|w_i\rangle \otimes |w_i\rangle)(\langle w_i| \otimes \langle w_i|)$, learn $|w_i\rangle$ (up to phases and permutations)

- Unitarily diagonalize A and take the $+1$ eigenvalues — we learn the eigenspace $\mathrm{span}\{|w_i\rangle \otimes |w_i\rangle \mid i = 1, \ldots, n\}$
- A numerical diagonalization routine will return the vectors

$$|\tilde{w}_i\rangle = \sum_{j=1}^{n} U_{ij} |w_j\rangle \otimes |w_j\rangle = \sum_{j=1}^{n} |U_{ij}| e^{\mathrm{i}\phi_{ij}} |w_j\rangle \otimes |w_j\rangle$$

- The unitary U is totally arbitrary and unpredictable

Learning algorithm

Given access to the vectors $|\tilde{w}_i\rangle = \sum_{j=1}^{n} |U_{ij}| e^{\mathrm{i}\phi_{ij}} |w_j\rangle \otimes |w_j\rangle$ with U, ϕ unknown and arbitrary, learn $|w_j\rangle$ (up to phases and permutations)

Learning algorithm

Given access to the vectors $|\tilde{w}_i\rangle = \sum_{j=1}^{n} |U_{ij}| e^{\mathrm{i}\phi_{ij}} |w_j\rangle \otimes |w_j\rangle$ with U, ϕ unknown and arbitrary, learn $|w_j\rangle$ (up to phases and permutations)

- Schmidt decomposition theorem: once we have $|\tilde{w}_i\rangle$, the $|U_{ij}|$ are unique up to reordering
- So we can perform the Schmidt decomposition (via SVD) to learn all the $|w_i\rangle$ up to phases

---

Subtlety: see Slide 42

Learning algorithm

Given access to the vectors $|\tilde{w}_i\rangle = \sum_{j=1}^{n} |U_{ij}| e^{\mathrm{i}\phi_{ij}} |w_j\rangle \otimes |w_j\rangle$ with U, ϕ unknown and arbitrary, learn $|w_j\rangle$ (up to phases and permutations)

- Schmidt decomposition theorem: once we have $|\tilde{w}_i\rangle$, the $|U_{ij}|$ are unique up to reordering
- So we can perform the Schmidt decomposition (via SVD) to learn all the $|w_i\rangle$ up to phases
- We define V to be the unitary matrix whose columns are precisely these learned vectors
- By construction, V is equal to W up to a permutation of its columns and global phases applied to the columns

---

Subtlety: see Slide 42

Learning algorithm

- In other words, we have learned the matrix $V = W\Phi P$, where Φ is some arbitrary diagonal unitary matrix, and P is some arbitrary permutation matrix
- It follows that the state $\mathcal{U}_V \left| b \ldots b \right\rangle$ is the same as $\mathcal{U}_W \left| b \ldots b \right\rangle$ up to an irrelevant global phase, thereby completing the learning algorithm

Error bounds

- Thus far, we have assumed perfect knowledge of the moments
- In practice, because we have a finite number of samples to measure, there will be an error

$$\sigma^{(2t)\prime} = \sigma^{(2t)} + \varepsilon_{2t} E^{(2t)} \qquad \|E^{(2t)}\| \leq 1$$
$$\Lambda^{(2t)\prime} = \Lambda^{(2t)} + \varepsilon_{2t} F^{(2t)} \qquad \|F^{(2t)}\| \leq 1$$

- How do we deal with this?
- Lots of perturbation theory and linear algebra
- A few interesting parts

Passive theorem

Theorem

- *Let* $|\psi\rangle = \mathcal{U}_W |\boldsymbol{f}\rangle$ *for an unknown unitary* $W \in \mathrm{U}(n)$ *specifying an arbitrary* **passive Gaussian unitary** *and an arbitrary unknown Fock state* $|\boldsymbol{f}\rangle$
- *Suppose our measurements* $\sigma^{(2)\prime}, \sigma^{(4)\prime}$ *of the moment matrices* $\sigma^{(2)}, \sigma^{(4)}$ *satisfy* $\|\sigma^{(2t)\prime} - \sigma^{(2t)}\| \leq \varepsilon_{2t}$
- *Then we can* **efficiently** *find a* $V \in \mathrm{U}(n)$ *and* $\boldsymbol{g}$ *such that*

$$|\langle \boldsymbol{f}| \mathcal{U}_W^\dagger \mathcal{U}_V |\boldsymbol{g}\rangle| \geq 1 - \frac{\gamma \|\boldsymbol{f}\|_\infty n}{1 - \gamma \|\boldsymbol{f}\|_\infty n}$$

with

$$\gamma = \varepsilon_2 \left(32\sqrt{5} n^2 (3 \|\boldsymbol{f}\|_\infty^2 + 5 \|\boldsymbol{f}\|_\infty + 2) + 4n\right) + 2\sqrt{5}\varepsilon_4 n$$

General theorem

Theorem

- *Let* $|\psi\rangle = \mathcal{U}_S |\boldsymbol{f}\rangle$ *for an unknown symplectic matrix* $S \in \mathrm{Sp}(2n, \mathbb{R})$ *specifying an* **arbitrary Gaussian unitary** *and an arbitrary Fock state* $|\boldsymbol{f}\rangle$
- *Suppose our measurements* $\Lambda^{(2)\prime}, \Lambda^{(4)\prime}$ *of the moment matrices* $\Lambda^{(2)}, \Lambda^{(4)}$ *satisfy* $\|\Lambda^{(2t)\prime} - \Lambda^{(2t)}\| \leq \varepsilon_{2t}$
- *Then we can efficiently find a* $Q \in \mathrm{Sp}(2n, \mathbb{R})$ *and* $\boldsymbol{g}$ *such that*

$$|\langle \boldsymbol{f}| \mathcal{U}_S^\dagger \mathcal{U}_Q |\boldsymbol{g}\rangle| \geq 1 - \mathcal{O}(\gamma \mathrm{e}^s n \|\boldsymbol{f}\|_\infty)$$

with

$$\gamma = \mathcal{O}\left(\varepsilon_2^{1/8} n^{3+1/2} \|S\|^{6+1/4} \|\boldsymbol{f}\|_\infty^5 + \varepsilon_4 n \|S\|^5 \|\boldsymbol{f}\|_\infty^{2+1/2}\right)$$

Number of measurements

- Recall that moments can be measured to precision ε with poly$(1/\varepsilon)$ measurements (Welsch, Vogel, et al. 1999)
- From this and our theorems:

Theorem

Let $|\psi\rangle = \mathcal{U}_S |\boldsymbol{f}\rangle$ for an unknown symplectic matrix $S \in \mathrm{Sp}(2n, \mathbb{R})$ specifying an arbitrary Gaussian unitary and an arbitrary Fock state $|\boldsymbol{f}\rangle$.

Using $\mathrm{poly}(n, \|\boldsymbol{f}\|_\infty, \|S\|)$ *copies of $|\psi\rangle$, with probability at least $1 - \frac{1}{\mathrm{poly}(n)}$ our algorithm learns $|\psi\rangle$ to fidelity at least $1 - \frac{1}{\mathrm{poly}(n)}$*

Number of measurements

- Recall that moments can be measured to precision ε with $\mathrm{poly}(1/\varepsilon)$ measurements (Welsch, Vogel, et al. 1999)
- From this and our theorems:

Theorem

Let $|\psi\rangle = \mathcal{U}_S |\boldsymbol{f}\rangle$ for an unknown symplectic matrix $S \in \mathrm{Sp}(2n, \mathbb{R})$ specifying an arbitrary Gaussian unitary and an arbitrary Fock state $|\boldsymbol{f}\rangle$.

Using $\mathrm{poly}(n, \|\boldsymbol{f}\|_\infty, \|S\|)$ copies of $|\psi\rangle$, with probability at least $1 - \frac{1}{\mathrm{poly}(n)}$ our algorithm learns $|\psi\rangle$ to fidelity at least $1 - \frac{1}{\mathrm{poly}(n)}$

Number of measurements

- Recall that moments can be measured to precision ε with $\mathrm{poly}(1/\varepsilon)$ measurements (Welsch, Vogel, et al. 1999)
- From this and our theorems:

Theorem

Let $|\psi\rangle = \mathcal{U}_S |\boldsymbol{f}\rangle$ for an unknown symplectic matrix $S \in \mathrm{Sp}(2n, \mathbb{R})$ specifying an arbitrary Gaussian unitary and an arbitrary Fock state $|\boldsymbol{f}\rangle$.

Using poly(energy) copies of $|\psi\rangle$, with probability at least $1 - \frac{1}{\mathrm{poly}(n)}$ our algorithm learns $|\psi\rangle$ to fidelity at least $1 - \frac{1}{\mathrm{poly}(n)}$

Overview

Gaussian states

Definition

A Gaussian state is a thermal or ground state of a Hamiltonian that is quadratic in the quadrature operators

Theorem

Gaussian states are fully specified by their first two moments

Proof.

Let ρ be an arbitrary state and let ρ_G be a Gaussian state whose first and second moments match those of ρ. Then $S(\rho_G) \geq S(\rho)$ with equality iff $\rho = \rho_G$. $\square$

Gaussian states are maximum entropy states subject to constraints on their first two moments

G_t states

Definition

a G_t state is a thermal (or ground) state of a (non-degenerate) Hamiltonian that is degree t in the quadrature operators

- A G_t state is fully specified by its first t moments
- The proof of this in the mixed state case is totally analogous to the Gaussian ($t = 2$) proof
- How about the pure state case?

Pure G_t states

- Suppose ψ is a G_t state; it is the ground state of $H = \sum_i \alpha_i \hat{M}_i$ where $\hat{M}_i$ are degree $\leq t$ operators

Pure G_t states

- Suppose ψ is a G_t state; it is the ground state of $H = \sum_i \alpha_i \hat{M}_i$ where $\hat{M}_i$ are degree $\leq t$ operators

$$\begin{aligned}\langle\psi| H |\psi\rangle &= \min_{\phi} \sum_i \alpha_i \langle\phi| \hat{M}_i |\phi\rangle \\ &= \min_{m_1, m_2, \ldots} \sum_i \alpha_i m_i \quad \text{st } \exists \phi, \langle\phi| \hat{M}_i |\phi\rangle = m_i\end{aligned}$$

Pure G_t states

- Suppose ψ is a G_t state; it is the ground state of $H = \sum_i \alpha_i \hat{M}_i$ where $\hat{M}_i$ are degree $\leq t$ operators

$$\begin{aligned} \langle\psi| H |\psi\rangle &= \min_{\phi} \sum_i \alpha_i \langle\phi| \hat{M}_i |\phi\rangle \\ &= \min_{m_1, m_2, \ldots} \sum_i \alpha_i m_i \quad \text{st } \exists\phi, \langle\phi| \hat{M}_i |\phi\rangle = m_i \end{aligned}$$

- Because H is non-degenerate, any state with the matching moments is ψ

Relevance to the learning algorithm

- A Fock state $|\boldsymbol{f}\rangle$ is a G_4 state; it is the ground state of

$$H = \sum_i (\hat{a}_i^\dagger \hat{a}_i - f_i)^2$$

- The set of G_t states is closed under the action of Gaussian unitaries (because they act affinely)
- Therefore, $\mathcal{U}_S |\boldsymbol{f}\rangle$ is a G_4 state
- The **first four moments suffice** to learn the state

Learning G_t states

- There is some way of learning a G_t state by measuring only the first t moments (for bosons, fermions, and qudits)
- If t is constant, there is only poly(n) numbers to measure — can this give efficient algorithms to learn arbitrary constant t G_t states?

Learning G_t states

- There is some way of learning a G_t state by measuring only the first t moments (for bosons, fermions, and qudits)
- If t is constant, there is only poly(n) numbers to measure — can this give efficient algorithms to learn arbitrary constant t G_t states?
- Main difficulty: if the moments are known to ε precision, what is the error $f(\varepsilon)$ in the reconstructed state?
- For our algorithm we showed $f(\varepsilon) \sim \text{poly}(1/\varepsilon)$; can we show this to be generally true?

Overview

Summary

- By going beyond second moments, we found a sample and time efficient algorithm to learn *BosonSampling states*
- Solves open problem left by Aaronson and Grewal 2023

Summary

- By going beyond second moments, we found a sample and time efficient algorithm to learn *BosonSampling states*
- Solves open problem left by Aaronson and Grewal 2023
- We construct an infinite set of Gaussian invariants for generic bosonic states using classical invariant theory of the action of symplectic group on moments, relevant for state conversion (Chabaud, Markham, et al. 2020; Hahn, Ferrini, et al. 2024)

Summary

- By going beyond second moments, we found a sample and time efficient algorithm to learn *BosonSampling states*
- Solves open problem left by Aaronson and Grewal 2023
- We construct an infinite set of Gaussian invariants for generic bosonic states using classical invariant theory of the action of symplectic group on moments, relevant for state conversion (Chabaud, Markham, et al. 2020; Hahn, Ferrini, et al. 2024)
- Lots of possible extensions!

Extensions

- Our algorithm generalizes to a slightly larger class of states, and fails interestingly in other places
- Certain classes of states *hide information* from lower moments
- Invariants and generalizations of Wick's theorem

Extensions

- Our algorithm generalizes to a slightly larger class of states, and fails interestingly in other places
- Certain classes of states *hide information* from lower moments
- Invariants and generalizations of Wick's theorem
- Can we make our algorithm have more favorable scaling with energy?
    - Our algorithm requires $N \sim \mathrm{poly}(n, \|\boldsymbol{f}\|_\infty, \|S\|)$ samples

Extensions

- Our algorithm generalizes to a slightly larger class of states, and fails interestingly in other places
- Certain classes of states *hide information* from lower moments
- Invariants and generalizations of Wick's theorem
- Can we make our algorithm have more favorable scaling with energy?
  - Our algorithm requires $N \sim \mathrm{poly}(n, \|\boldsymbol{f}\|_\infty, \|S\|)$ samples
  - Applying (Bittel, Mele, et al. 2025), I *think* we get $N \sim \mathrm{poly}(n, \|\boldsymbol{f}\|_\infty, \log \|S\|)$

Additional slides

Overview

Constant Fock theorem

Theorem

Suppose $\boldsymbol{f} = (b, \ldots, b)$ *for some nonnegative integer* b *and let* $|\psi\rangle = \mathcal{U}_W |\boldsymbol{f}\rangle$ *for an unknown unitary* $W \in \mathrm{U}(n)$ *specifying an arbitrary passive Gaussian unitary. If our measurement* $\sigma^{(4)\prime}$ *of the moment matrix* $\sigma^{(4)}$ *for* $|\psi\rangle$ *satisfies* $\left\|\sigma^{(4)\prime} - \sigma^{(4)}\right\| \leq \varepsilon$, *then we can efficiently find a* $V \in \mathrm{U}(n)$ *such that*

$$\|V - W\Phi P\| \leq \frac{4\sqrt{5}\varepsilon n}{b(b+1)}$$

for some (irrelevant) diagonal unitary matrix Φ *and permutation matrix* P. *In particular,*

$$|\langle b^n| \mathcal{U}_W^\dagger \mathcal{U}_V |b^n\rangle| \geq 1 - \frac{4\sqrt{5}\varepsilon n^2/(b+1)}{1 - 4\sqrt{5}\varepsilon n^2/(b+1)}$$

as long as $\varepsilon \leq \frac{b+1}{4\sqrt{5}n^2}$.

Passive theorem

Theorem

Let $|\psi\rangle = \mathcal{U}_W |\boldsymbol{f}\rangle$ *for an unknown unitary* $W \in \mathrm{U}(n)$ *specifying an arbitrary passive Gaussian unitary and an arbitrary Fock state* $|\boldsymbol{f}\rangle$*. If our measurements* $\sigma^{(2)\prime}, \sigma^{(4)\prime}$ *of the moment matrices* $\sigma^{(2)}, \sigma^{(4)}$ *satisfy* $\|\sigma^{(2t)\prime} - \sigma^{(2t)}\| \leq \varepsilon_{2t}$*, then we can efficiently find a* $V \in \mathrm{U}(n)$ *and* $\boldsymbol{g}$ *such that* $\|V - W\Phi P\| \leq \gamma$*, with*

$$\gamma = \varepsilon_2 \left(32\sqrt{5}n^2(3f_{\max}^2 + 5f_{\max} + 2) + 4n\right) + 2\sqrt{5}\varepsilon_4 n$$

for some diagonal unitary matrix Φ *and a permutation matrix* P*, and* $f_{\max} = \max_i f_i$*. Specifically,* $\boldsymbol{g}$ *is some permutation of* $\boldsymbol{f}$ *and* P *performs this permutation along with other (irrelevant) permutations within blocks of equal* g_i*. In particular,*

$$|\langle \boldsymbol{f}| \mathcal{U}_W^\dagger \mathcal{U}_V |\boldsymbol{g}\rangle| \geq 1 - \frac{\gamma f_{\max} n}{1 - \gamma f_{\max} n}$$

as long as $\gamma f_{\max} n < 1$*.*

General theorem

Theorem

Let $|\psi\rangle = \mathcal{U}_S |\boldsymbol{f}\rangle$ *for an unknown symplectic matrix* $S \in \mathrm{Sp}(2n, \mathbb{R})$ *specifying an arbitrary Gaussian unitary and an arbitrary Fock state* $|\boldsymbol{f}\rangle$*. If our measurements* $\Lambda^{(2)\prime}, \Lambda^{(4)\prime}$ *of the moment matrices* $\Lambda^{(2)}, \Lambda^{(4)}$ *satisfy* $\|\Lambda^{(2t)\prime} - \Lambda^{(2t)}\| \leq \varepsilon_{2t}$*, then we can efficiently find a* $Q \in \mathrm{Sp}(2n, \mathbb{R})$ *and* $\boldsymbol{g}$ *such that* $\|Q - S\Phi P\| \leq \gamma$*, where*

$$\gamma = \mathcal{O}\left(\varepsilon_2^{1/8} \mathrm{e}^{25s/4} n^{3+1/2} f_{\max}^5 + \varepsilon_4 \mathrm{e}^{5s} n f_{\max}^{2+1/2}\right)$$

for some symplectic matrices Φ *and* P *that implement global phases and mode permutations,* $f_{\max} = \max_i f_i$*, and* s *is the maximum magnitude of squeezing in* S *(that is,* e^s *is the largest singular value of* S*). Specifically,* $\boldsymbol{g}$ *is some permutation of* $\boldsymbol{f}$ *and* P *performs this permutation along with other (irrelevant) permutations within blocks of equal* g_i*. In particular,*

$$|\langle \boldsymbol{f}| \mathcal{U}_S^\dagger \mathcal{U}_Q |\boldsymbol{g}\rangle| \geq 1 - \mathcal{O}(\gamma \mathrm{e}^s n f_{\max}) .$$

Overview

5 Additional slides
- Theorem statements
- Algorithm
- Error analysis
- Generalization of Wick's theorem

Reducing the problem

- Easy to check that $\mathrm{Re}\,\Lambda_0^{(1)} = \begin{pmatrix} \frac{1}{2}\mathbb{I} + P_{\boldsymbol{f}} & 0 \\ 0 & \frac{1}{2}\mathbb{I} + P_{\boldsymbol{f}} \end{pmatrix}$ where $P_{\boldsymbol{f}} = \mathrm{diag}\{f_1, \ldots, f_n\}$
- Suppose we know $\Lambda^{(2)}$ and $\Lambda^{(4)}$ perfectly. Note that $\Lambda^{(2)} = S\Lambda_0^{(1)}S^T$
- We can *symplectically diagonalize* $\mathrm{Re}\,\Lambda^{(2)}$ to yield

$$\mathrm{Re}\,\Lambda^{(2)} = \mathrm{Re}\,S\Lambda_0^{(1)}S^T = \mathrm{Re}\,R\Lambda_0^{(1)}R^T$$

Reducing the problem

- Easy to check that $\mathrm{Re}\,\Lambda_0^{(1)} = \begin{pmatrix} \frac{1}{2}\mathbb{I} + P_{\boldsymbol{f}} & 0 \\ 0 & \frac{1}{2}\mathbb{I} + P_{\boldsymbol{f}} \end{pmatrix}$ where $P_{\boldsymbol{f}} = \mathrm{diag}\{f_1, \ldots, f_n\}$
- Suppose we know $\Lambda^{(2)}$ and $\Lambda^{(4)}$ perfectly. Note that $\Lambda^{(2)} = S\Lambda_0^{(1)}S^T$
- We can *symplectically diagonalize* $\mathrm{Re}\,\Lambda^{(2)}$ to yield

$$\mathrm{Re}\,\Lambda^{(2)} = \mathrm{Re}\,S\Lambda_0^{(1)}S^T = \mathrm{Re}\,R\Lambda_0^{(1)}R^T$$

- Note that $R \neq S$, however $R^{-1}S$ is passive, $R^{-1}S \in \mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n) \cong \mathrm{U}(n)$

Reducing the problem, putting it together

- $\Lambda^{(2)} = S\Lambda_0^{(2)}S^T$ and $\Lambda^{(4)} = (S \otimes S)\Lambda_0^{(4)}(S \otimes S)^T$
- Assume we know $\Lambda^{(2)}$ and $\Lambda^{(4)}$ perfectly (but S and $\boldsymbol{f}$ are unknown)

Reducing the problem, putting it together

- $\Lambda^{(2)} = S\Lambda_0^{(2)}S^T$ and $\Lambda^{(4)} = (S \otimes S)\Lambda_0^{(4)}(S \otimes S)^T$
- Assume we know $\Lambda^{(2)}$ and $\Lambda^{(4)}$ perfectly (but S and $\boldsymbol{f}$ are unknown)
- From $\Lambda^{(2)}$, we learn R such that $R^{-1}S$ is a passive Gaussian unitary
- Therefore, $R^{-1}\Lambda^{(2)}R^{-1T}$ and $(R^{-1} \otimes R^{-1})\Lambda^{(4)}(R^{-1} \otimes R^{-1})^T$ correspond to the state $\mathcal{U}_W \left|\boldsymbol{f}\right\rangle$ for a passive W
- If we can solve the passive problem to learn W, then we reapply R and we will have $S \sim RW$

Reducing the passive problem

- Now we just need to solve the passive problem
- Assume we know the moments $\sigma^{(2)}$ and $\sigma^{(4)}$ for $\mathcal{U}_W \left|\boldsymbol{f}\right\rangle$ where $\mathcal{U}_W$ is a passive Gaussian unitary

Reducing the passive problem

- Now we just need to solve the passive problem
- Assume we know the moments $\sigma^{(2)}$ and $\sigma^{(4)}$ for $\mathcal{U}_W \left| \boldsymbol{f} \right\rangle$ where $\mathcal{U}_W$ is a passive Gaussian unitary
- This means that $W \in \mathrm{U}(n)$ and $\sigma^{(2t)} = W^{\otimes t} \sigma_0^{(2t)} W^{\dagger \otimes t}$

Reducing the passive problem

- Now we just need to solve the passive problem
- Assume we know the moments $\sigma^{(2)}$ and $\sigma^{(4)}$ for $\mathcal{U}_W \left|\boldsymbol{f}\right\rangle$ where $\mathcal{U}_W$ is a passive Gaussian unitary
- This means that $W \in \mathrm{U}(n)$ and $\sigma^{(2t)} = W^{\otimes t}\sigma_0^{(2t)}W^{\dagger\otimes t}$
- Easy to check that $\sigma_0^{(1)} = \mathbb{I} + P_{\boldsymbol{f}}$
- If we unitarily diagonalize, we will find a U such that $\sigma^{(2)} = W\sigma_0^{(1)}W^{\dagger} = U\sigma_0^{(1)}U^{\dagger}$

Reducing the passive problem

- Now we just need to solve the passive problem
- Assume we know the moments $\sigma^{(2)}$ and $\sigma^{(4)}$ for $\mathcal{U}_W \left|\boldsymbol{f}\right\rangle$ where $\mathcal{U}_W$ is a passive Gaussian unitary
- This means that $W \in \mathrm{U}(n)$ and $\sigma^{(2t)} = W^{\otimes t}\sigma_0^{(2t)} W^{\dagger\otimes t}$
- Easy to check that $\sigma_0^{(1)} = \mathbb{I} + P_{\boldsymbol{f}}$
- If we unitarily diagonalize, we will find a U such that $\sigma^{(2)} = W\sigma_0^{(1)} W^\dagger = U\sigma_0^{(1)} U^\dagger$
- Note $U \neq W$, but $U^\dagger W$ is block diagonal, $\left[U^\dagger W, P_{\boldsymbol{f}}\right] = 0$

Reducing the passive problem

- Now we just need to solve the passive problem
- Assume we know the moments $\sigma^{(2)}$ and $\sigma^{(4)}$ for $\mathcal{U}_W \,|\boldsymbol{f}\rangle$ where $\mathcal{U}_W$ is a passive Gaussian unitary
- This means that $W \in \mathrm{U}(n)$ and $\sigma^{(2t)} = W^{\otimes t} \sigma_0^{(2t)} W^{\dagger \otimes t}$
- Easy to check that $\sigma_0^{(1)} = \mathbb{I} + P_{\boldsymbol{f}}$
- If we unitarily diagonalize, we will find a U such that $\sigma^{(2)} = W \sigma_0^{(1)} W^\dagger = U \sigma_0^{(1)} U^\dagger$
- Note $U \neq W$, but $U^\dagger W$ is block diagonal, $\left[U^\dagger W, P_{\boldsymbol{f}}\right] = 0$
- The passive Gaussian unitary $U^\dagger W$ does not mix modes with different initial Fock numbers

Reducing the passive problem, putting it together

Example

Suppose that $\boldsymbol{f} = (1,1,2,2,2)$; then $U^\dagger \sigma^{(2)} U$ and $U^{\dagger\otimes 2}\sigma^{(4)} U^{\otimes 2}$ correspond to a state

$$(\mathcal{U}_{W_1}\,|1,1\rangle) \otimes (\mathcal{U}_{W_2}\,|2,2,2\rangle), \qquad U^\dagger W = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix}$$

Reducing the passive problem, putting it together

Example

Suppose that $\boldsymbol{f} = (1, 1, 2, 2, 2)$; then $U^\dagger \sigma^{(2)} U$ and $U^{\dagger \otimes 2} \sigma^{(4)} U^{\otimes 2}$ correspond to a state

$$(\mathcal{U}_{W_1} |1, 1\rangle) \otimes (\mathcal{U}_{W_2} |2, 2, 2\rangle), \qquad U^\dagger W = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix}$$

Thus, if we can solve the learning problem for each state $\mathcal{U}_{W_i} |f_i, \ldots, f_i\rangle$, then we can reapply U to learn W

Reducing the passive problem, putting it together

Example

Suppose that $\boldsymbol{f} = (1,1,2,2,2)$; then $U^\dagger \sigma^{(2)} U$ and $U^{\dagger\otimes 2}\sigma^{(4)} U^{\otimes 2}$ correspond to a state

$$(\mathcal{U}_{W_1}\left|1,1\right\rangle) \otimes (\mathcal{U}_{W_2}\left|2,2,2\right\rangle), \qquad U^\dagger W = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix}$$

Thus, if we can solve the learning problem for each state $\mathcal{U}_{W_i}\left|f_i,\dots,f_i\right\rangle$, then we can reapply U to learn W

- We have therefore reduced learning $\mathcal{U}_S\left|\boldsymbol{f}\right\rangle$ to learning $\mathcal{U}_W\left|b,\dots,b\right\rangle$ for unknown $W \in \mathrm{U}(n)$ and known $b \in \mathbb{N}_0$

Applying higher moments

- We have so far only used second moments ($\Lambda^{(2)}$ and $\sigma^{(2)}$)
- Note that Λ is only necessary for active Gaussian unitaries; σ contains all the information for passive
- If we were dealing with Gaussian states, this would be sufficient

Applying higher moments

- We have so far only used second moments ($\Lambda^{(2)}$ and $\sigma^{(2)}$)
- Note that Λ is only necessary for active Gaussian unitaries; σ contains all the information for passive
- If we were dealing with Gaussian states, this would be sufficient
- But for us second moments are now useless — easy to check that for $\boldsymbol{f} = (b, \ldots, b)$, $\sigma_0^{(2)} = (b+1)\mathbb{I}$
- Therefore, $\sigma^{(2)} = W\sigma_0^{(2)}W^\dagger = \sigma_0^{(2)}$, so we learn nothing about W by measuring second moments

Applying higher moments

- We have so far only used second moments ($\Lambda^{(2)}$ and $\sigma^{(2)}$)
- Note that Λ is only necessary for active Gaussian unitaries; σ contains all the information for passive
- If we were dealing with Gaussian states, this would be sufficient
- But for us second moments are now useless — easy to check that for $\boldsymbol{f} = (b, \ldots, b)$, $\sigma_0^{(2)} = (b+1)\mathbb{I}$
- Therefore, $\sigma^{(2)} = W\sigma_0^{(2)}W^\dagger = \sigma_0^{(2)}$, so we learn nothing about W by measuring second moments
- So now we use fourth moments, $\sigma^{(4)}$

Learning algorithm

Given the fourth moments $\sigma^{(4)} = W^{\otimes 2}\sigma_0^{(4)} W^{\dagger\otimes 2}$ for the state $\mathcal{U}_W |b, \ldots, b\rangle$, determine W

Learning algorithm

Given the fourth moments $\sigma^{(4)} = W^{\otimes 2}\sigma_0^{(4)} W^{\dagger\otimes 2}$ for the state $\mathcal{U}_W |b, \ldots, b\rangle$, determine W

- Because $\sigma_0^{(4)}$ is a matrix on $(\mathbb{C}^n)^{\otimes 2}$, we can represent it in bra-ket notation using the standard basis $|1\rangle, \ldots, |n\rangle$ of $\mathbb{C}^n$

$$\sigma_0^{(4)} = (b+1)^2(\mathbb{I} + U_{\mathrm{SWAP}}) - b(b+1)\sum_{i=1}^{n} |i,i\rangle\langle i,i| ,$$

Learning algorithm

Given the fourth moments $\sigma^{(4)} = W^{\otimes 2} \sigma_0^{(4)} W^{\dagger \otimes 2}$ for the state $\mathcal{U}_W |b, \ldots, b\rangle$, determine W

- Because $\sigma_0^{(4)}$ is a matrix on $(\mathbb{C}^n)^{\otimes 2}$, we can represent it in bra-ket notation using the standard basis $|1\rangle, \ldots, |n\rangle$ of $\mathbb{C}^n$

$$\sigma_0^{(4)} = (b+1)^2(\mathbb{I} + U_{\mathrm{SWAP}}) - b(b+1) \sum_{i=1}^{n} |i, i\rangle\langle i, i| ,$$

- Denote the i^{th} column vector of W by $|w_i\rangle$. Given access to $\sigma^{(4)}$, we can thus form the matrix

$$\begin{aligned} A &= \frac{1}{b(b+1)} \left((b+1)^2(\mathbb{I} + U_{\mathrm{SWAP}}) - \sigma^{(4)} \right) \\ &= \sum_{i=1}^{n} (|w_i\rangle \otimes |w_i\rangle)(\langle w_i| \otimes \langle w_i|), \end{aligned}$$

Learning algorithm

Given access to a matrix $A = \sum_{i=1}^{n} (|w_i\rangle \otimes |w_i\rangle)(\langle w_i| \otimes \langle w_i|)$, learn $|w_i\rangle$ (up to phases and permutations)

Learning algorithm

Given access to a matrix $A = \sum_{i=1}^{n} (|w_i\rangle \otimes |w_i\rangle)(\langle w_i| \otimes \langle w_i|)$, learn $|w_i\rangle$ (up to phases and permutations)

- Unitarily diagonalize A and take the $+1$ eigenvalues — we learn the eigenspace $\text{span}\{|w_i\rangle \otimes |w_i\rangle \mid i = 1, \ldots, n\}$

Learning algorithm

Given access to a matrix $A = \sum_{i=1}^{n} (|w_i\rangle \otimes |w_i\rangle)(\langle w_i| \otimes \langle w_i|)$, learn $|w_i\rangle$ (up to phases and permutations)

- Unitarily diagonalize A and take the $+1$ eigenvalues — we learn the eigenspace $\mathrm{span}\{|w_i\rangle \otimes |w_i\rangle \mid i = 1, \ldots, n\}$
- A numerical diagonalization routine will return the vectors

$$|\tilde{w}_i\rangle = \sum_{j=1}^{n} U_{ij} |w_j\rangle \otimes |w_j\rangle = \sum_{j=1}^{n} |U_{ij}| e^{\mathrm{i}\phi_{ij}} |w_j\rangle \otimes |w_j\rangle$$

- The unitary U is totally arbitrary and unpredictable

Learning algorithm

Given access to a vector $|\tilde{w}_i\rangle = \sum_{j=1}^{n} |U_{ij}| e^{\mathrm{i}\phi_{ij}} |w_j\rangle \otimes |w_j\rangle$ with U, ϕ unknown and arbitrary, learn $|w_j\rangle$ (up to phases and permutations)

Learning algorithm

Given access to a vector $|\tilde{w}_i\rangle = \sum_{j=1}^{n} |U_{ij}| e^{\mathrm{i}\phi_{ij}} |w_j\rangle \otimes |w_j\rangle$ with U, ϕ unknown and arbitrary, learn $|w_j\rangle$ (up to phases and permutations)

- Schmidt decomposition theorem: once we have $|\tilde{w}_i\rangle$, the $|U_{ij}|$ are unique up to reordering
- So we can perform the Schmidt decomposition (via SVD) to learn all the $|w_i\rangle$ up to phases

Subtlety: see Slide 42

Learning algorithm

Given access to a vector $|\tilde{w}_i\rangle = \sum_{j=1}^{n} |U_{ij}| e^{\mathrm{i}\phi_{ij}} |w_j\rangle \otimes |w_j\rangle$ with U, ϕ unknown and arbitrary, learn $|w_j\rangle$ (up to phases and permutations)

- Schmidt decomposition theorem: once we have $|\tilde{w}_i\rangle$, the $|U_{ij}|$ are unique up to reordering
- So we can perform the Schmidt decomposition (via SVD) to learn all the $|w_i\rangle$ up to phases
- We define V to be the unitary matrix whose columns are precisely these learned vectors
- By construction, V is equal to W up to a permutation of its columns and global phases applied to the columns

---

Subtlety: see Slide 42

Learning algorithm

From Slide 14

Given access to a vector $|\tilde{w}_i\rangle = \sum_{j=1}^{n} U_{ij} |w_j\rangle \otimes |w_j\rangle$ with U, ϕ unknown and arbitrary, learn $|w_j\rangle$ (up to phases and permutations)

- I had said we just perform the Schmidt decomposition
- This only gives us the true $|w_i\rangle$ if there is no degeneracy in the $|U_{ij}|$
- Draw $O \in \mathrm{U}(n)$ Haar randomly, define $|c_i\rangle = \sum_j O_{ij} |\tilde{w}_j\rangle = \sum_j (OU)_{ij} |w_j\rangle \otimes |w_j\rangle$
- OU is also Haar random, so with probability 1 the Schmidt basis is unique up to phases
- Alternatively, we show that by taking the intersection of the Schmidt decompositions of all $|\tilde{w}_i\rangle$, you uniquely get each $|w_i\rangle$

Learning algorithm

- In other words, we have learned the matrix $V = W\Phi P$, where Φ is some arbitrary diagonal unitary matrix, and P is some arbitrary permutation matrix
- It follows that the state $\mathcal{U}_V |b \dots b\rangle$ is the same as $\mathcal{U}_W |b \dots b\rangle$ up to an irrelevant global phase, thereby completing the learning algorithm

Summary

Task

Using the moment matrices $\Lambda^{(2)}$ and $\Lambda^{(4)}$. learn $Q, \boldsymbol{g}$ such that $\mathcal{U}_Q \left|\boldsymbol{g}\right\rangle = \mathcal{U}_S \left|\boldsymbol{f}\right\rangle$

Summary

Task

Using the moment matrices $\Lambda^{(2)}$ and $\Lambda^{(4)}$. learn $Q, \boldsymbol{g}$ such that $\mathcal{U}_Q \left|\boldsymbol{g}\right\rangle = \mathcal{U}_S \left|\boldsymbol{f}\right\rangle$

- **Step 1:** Symplectic diagonlization of covariance matrix $\Lambda^{(2)}$, learn R such that $R^{-1}S$ is passive, thereby represented by $W \in \mathrm{U}(n)$
- Create $\sigma^{(2)}$ and $\sigma^{(4)}$ from the moments $R^{-1}\Lambda^{(2)}R^{-1T}$ and $(R^{-1})^{\otimes 2}\Lambda^{(4)}(R^{-1T})^{\otimes 2}$

Summary

Task

Using the moment matrices $\Lambda^{(2)}$ and $\Lambda^{(4)}$. learn $Q, \boldsymbol{g}$ such that $\mathcal{U}_Q |\boldsymbol{g}\rangle = \mathcal{U}_S |\boldsymbol{f}\rangle$

- **Step 1:** Symplectic diagonlization of covariance matrix $\Lambda^{(2)}$, learn R such that $R^{-1}S$ is passive, thereby represented by $W \in \mathrm{U}(n)$
- Create $\sigma^{(2)}$ and $\sigma^{(4)}$ from the moments $R^{-1}\Lambda^{(2)}R^{-1T}$ and $(R^{-1})^{\otimes 2}\Lambda^{(4)}(R^{-1T})^{\otimes 2}$
- **Step 2:** Block diagonalize $\sigma^{(2)}$ to learn U such that $U^\dagger W$ is block diagonal with $P_{\boldsymbol{g}}$
- For each block, create $\sigma^{(2,i)}$ from that relevant block of $U^{\dagger\otimes 2}\sigma^{(4)}U^{\otimes 2}$

Summary

Task

Using the moment matrices $\Lambda^{(2)}$ and $\Lambda^{(4)}$. learn $Q, \boldsymbol{g}$ such that $\mathcal{U}_Q \left|\boldsymbol{g}\right\rangle = \mathcal{U}_S \left|\boldsymbol{f}\right\rangle$

- **Step 1:** Symplectic diagonlization of covariance matrix $\Lambda^{(2)}$, learn R such that $R^{-1}S$ is passive, thereby represented by $W \in \mathrm{U}(n)$
- Create $\sigma^{(2)}$ and $\sigma^{(4)}$ from the moments $R^{-1}\Lambda^{(2)}R^{-1\,T}$ and $(R^{-1})^{\otimes 2}\Lambda^{(4)}(R^{-1\,T})^{\otimes 2}$
- **Step 2:** Block diagonalize $\sigma^{(2)}$ to learn U such that $U^\dagger W$ is block diagonal with $P_{\boldsymbol{g}}$
- For each block, create $\sigma^{(2,i)}$ from that relevant block of $U^{\dagger\otimes 2}\sigma^{(4)}U^{\otimes 2}$
- **Step 3:** Apply the learning algorithm to each $\sigma^{(2,i)}$, therefore learning $V = V_1 \oplus \cdots \oplus V_k$
- Let Q be the symplectic matrix corresponding to the passive Gaussian unitary UV
- Set $Q \to RQ$

Overview

Error in symplectic diagonalization

- A positive definite matrix M can be symplectically diagonalized $M = SDS^T = RDR^T$
- Let $M' = M + E = R'D'R'^T$
- One can bound $\|D - D'\|$ in terms of $\|E\|$
- Of course it does not make sense to bound $\|S - R'\|$ or $\|R - R'\|$
- Eg $\|S - R\|$ is not bound by $\|E\|$
- However, $SS^T = RR^T$; $\left\|SS^T - R'R'^T\right\|$ can be bound
- Using the bound on $\left\|SS^T - R'R'^T\right\|$, we need to bound how far $R'^{-1}S$ is from a passive Gaussian unitary
- Bound how far $(R'^{-1})^{\otimes 2}\Lambda^{(4)\prime}(R'^{-T})^{\otimes 2}$ is from error-free passive moments, send into the passive moments algorithm error analysis

Error in passive diagonalization

- U and W unitarily diagonalize the ideal $\sigma^{(2)}$, U' diagonalizes $\sigma^{(2)\prime}$
- Need to determine how far away $U'^{\dagger}W$ is from *any* block diagonal matrix
- This can be bound by $\|[U'^{\dagger}W, \sigma_0^{(2)}]\|$
- Then bound how far $U'^{\dagger\otimes 2}\sigma^{(4)\prime}U'^{\otimes 2}$ is from error-free block diagonal moments, send into the constant Fock algorithm error analysis

Error in constant Fock algorithm

- Track error in matrix diagonalization
- Track error in Schmidt decomposition

Overview

Wick's theorem

- Let $\langle A \rangle \equiv \mathrm{Tr}[\rho A]$
- If ρ is a Gaussian state with zero first moments, then Wick's theorem says eg

$$\langle x_1 x_2 x_3 x_4 \rangle = \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle + \langle x_1 x_3 \rangle \langle x_2 x_4 \rangle + \langle x_1 x_4 \rangle \langle x_2 x_3 \rangle$$

G_t states

- Now let ρ be a G_4 state
- Consider the following "algorithm" to compute sixth moments from the first four moments

G_t states

- Now let ρ be a G_4 state
- Consider the following "algorithm" to compute sixth moments from the first four moments
  - Given the first four moments, learn the corresponding unique state, then compute the sixth moments

G_t states

- Now let ρ be a G_4 state
- Consider the following "algorithm" to compute sixth moments from the first four moments
  - Given the first four moments, learn the corresponding unique state, then compute the sixth moments
- Can we do better? Some generalization of Wick's theorem?

“Deriving” Wick’s theorem

- Because centered Gaussian states are fully specified by their second moments $\sigma^{(4)}$, their fourth moment must be a function of the second, $\Sigma^{(4)}(\sigma^{(4)})$

"Deriving" Wick's theorem

- Because centered Gaussian states are fully specified by their second moments $\sigma^{(4)}$, their fourth moment must be a function of the second, $\Sigma^{(4)}(\sigma^{(4)})$

- Of course it must be an equivariant function, $\Sigma^{(4)}(S^{\otimes 2}\sigma^{(4)}) = S^{\otimes 4}\Sigma^{(4)}(\sigma^{(4)})$

"Deriving" Wick's theorem

- Because centered Gaussian states are fully specified by their second moments $\sigma^{(4)}$, their fourth moment must be a function of the second, $\Sigma^{(4)}(\sigma^{(4)})$
- Of course it must be an equivariant function, $\Sigma^{(4)}(S^{\otimes 2}\sigma^{(4)}) = S^{\otimes 4}\Sigma^{(4)}(\sigma^{(4)})$
- Such functions are of the form

$$\Sigma^{(4)}(\sigma^{(4)}) = \sum_{\sigma \in S_4} f_\sigma(\sigma^{(4)}) P_\sigma \sigma^{(4)} \otimes \sigma^{(4)}$$

P_σ are permutations, and $f_\sigma(S^{\otimes 2}\sigma^{(4)}) = f_\sigma(\sigma^{(4)})$

“Deriving” Wick’s theorem

- Because centered Gaussian states are fully specified by their second moments $\sigma^{(4)}$, their fourth moment must be a function of the second, $\Sigma^{(4)}(\sigma^{(4)})$
- Of course it must be an equivariant function, $\Sigma^{(4)}(S^{\otimes 2}\sigma^{(4)}) = S^{\otimes 4}\Sigma^{(4)}(\sigma^{(4)})$
- Such functions are of the form

$$\Sigma^{(4)}(\sigma^{(4)}) = \sum_{\sigma \in S_4} f_\sigma(\sigma^{(4)}) P_\sigma \sigma^{(4)} \otimes \sigma^{(4)}$$

  P_σ are permutations, and $f_\sigma(S^{\otimes 2}\sigma^{(4)}) = f_\sigma(\sigma^{(4)})$
- For $\langle x_1 x_2 x_3 x_4 \rangle$, must be permutation symmetric, so that

$$\langle x_1 x_2 x_3 x_4 \rangle = f(\sigma^{(4)}) \left(\langle x_1 x_2 \rangle \langle x_3 x_4 \rangle + \langle x_1 x_3 \rangle \langle x_2 x_4 \rangle + \langle x_1 x_4 \rangle \langle x_2 x_3 \rangle \right)$$

- How do we find $f(\sigma^{(4)})$?

"Deriving" Wick's theorem II

- I don't know how to determine it from this argument alone
- For $\sigma^{(4)}$, the only invariants are the symplectic eigenvalues
- For a pure state, these are all 1
- Thus, this simple argument *almost* gives us Wick's theorem for pure Gaussian states:

$$\langle x_1 x_2 x_3 x_4 \rangle \propto \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle + \langle x_1 x_3 \rangle \langle x_2 x_4 \rangle + \langle x_1 x_4 \rangle \langle x_2 x_3 \rangle$$

Generalizing

- Can we generalize this type of argument to G_t states beyond $t = 2$?
- The difficulty appears to be determining the invariant functions
- For $t = 2$, there are very few invariants and they are easily understood (symplectic eigenvalues)
- For larger t, there are more (this is something we characterize)

Guess

- The guess that comes from this argument is: eg for G_4 states,

$$\begin{aligned}\langle x_1 x_2 x_3 x_4 x_5 x_6 \rangle = \Big[&\\ & f_1(\sigma^{(4)}, \Sigma^{(3)}, \Sigma^{(4)}) \left(\langle x_1 x_2 \rangle \langle x_3 x_4 \rangle \langle x_5 x_6 \rangle + \text{permutations} \right) \\ & + f_2(\sigma^{(4)}, \Sigma^{(3)}, \Sigma^{(4)}) \left(\langle x_1 x_2 x_3 x_4 \rangle \langle x_5 x_6 \rangle + \text{permutations} \right) \\ & + f_3(\sigma^{(4)}, \Sigma^{(3)}, \Sigma^{(4)}) \left(\langle x_1 x_2 x_3 \rangle \langle x_4 x_5 x_6 \rangle + \text{permutations} \right) \\ & \Big],\end{aligned}$$

References I

- Iosue, Joseph T., Yu-Xin Wang, et al. (Oct. 2025). *Higher moment theory and learnability of bosonic states*. arXiv: 2510.01610 [quant-ph].
- Mele, Francesco Anna, Antonio Anna Mele, et al. (May 2024). *Learning quantum states of continuous variable systems*. arXiv: 2405.01431 [quant-ph].
- Zhao, Xiaobin, Pengcheng Liao, et al. (Nov. 2024). *Complexity of quantum tomography from genuine non-Gaussian entanglement*. arXiv: 2411.01609 [quant-ph].
- Bittel, Lennart, Francesco A. Mele, et al. (Aug. 2025). *Energy-independent tomography of Gaussian states*. arXiv: 2508.14979 [quant-ph].
- Fanizza, Marco, Vishnu Iyer, et al. (Oct. 2025). *Efficient learning of bosonic Gaussian unitaries*. arXiv: 2510.05531 [quant-ph].

References II

- Aaronson, Scott and Sabee Grewal (July 2023). "Efficient Tomography of Non-Interacting-Fermion States". In: *18th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2023)*. Ed. by Omar Fawzi and Michael Walter. Vol. 266. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 12:1–12:18. arXiv: 2102.10458 [quant-ph].
- Welsch, Dirk-Gunnar, Werner Vogel, et al. (May 1999). "Homodyne Detection and Quantum-State Reconstruction". In: ed. by E. Wolf. Vol. 39. Progress in Optics. Elsevier, pp. 63–211. arXiv: 0907.1353 [quant-ph].
- Chabaud, Ulysse, Damian Markham, et al. (2020). "Stellar Representation of Non-Gaussian Quantum States". In: *Phys. Rev. Lett.* 124 (6), p. 063605. arXiv: 1907.11009 [quant-ph].
- Hahn, Oliver, Giulia Ferrini, et al. (Oct. 2024). *Assessing non-Gaussian quantum state conversion with the stellar rank*. arXiv: 2410.23721 [quant-ph].