Quantum learning in bosonic and fermionic systems

Antonio Anna Mele
Freie Universität Berlin

30 April 2026

Abstract

These notes are written to guide the 1.5-hour blackboard lecture on quantum learning in bosonic and fermionic systems that I will give at the EPFL workshop Quantum learning, sensing and verification. We review the basic formulation of tomography for qudits, and then discuss the same problem for bosonic and fermionic systems. We introduce Gaussian states, reviewing their mathematical definitions and basic properties, and then discuss the problem of Gaussian state tomography. Finally, we explore tomography beyond the Gaussian setting, focusing on states with a small amount of magic, such as t-doped Gaussian states and their analogues in the fermionic, bosonic, and Clifford/stabilizer settings.

Contents

1 Bosonic and fermionic preliminaries

In this section, we recall the basic Hilbert-space and operator formalism for bosonic and fermionic systems. The purpose is to make explicit the parallel between bosonic modes, fermionic modes and qudits.

A qudit of local dimension d is described by the Hilbert space

$$\mathcal{H}_d = \mathrm{span}\{|0\rangle, |1\rangle, \ldots, |d-1\rangle\}. \tag{1}$$

A single bosonic mode can be regarded as the infinite-dimensional analogue of a qudit. Its Hilbert space is

$$\mathcal{H}_{\mathrm{B},1} = \mathrm{span}\{|0\rangle, |1\rangle, |2\rangle, \ldots\}. \tag{2}$$

The vector $|k\rangle$ is called the k-particle Fock state. Hence, an n-mode bosonic system is described by

$$\mathcal{H}_{\mathrm{B},n} = \mathcal{H}_{\mathrm{B},1}^{\otimes n} = \mathrm{span}\{|k_1, \ldots, k_n\rangle : k_i \in \mathbb{N}\}. \tag{3}$$

Thus each bosonic mode can contain an arbitrary number of particles.

A single fermionic mode, instead, has only two possible occupation numbers. Its Hilbert space is

$$\mathcal{H}_{\mathrm{F},1} = \mathrm{span}\{|0\rangle, |1\rangle\}. \tag{4}$$

For n fermionic modes, we write

$$\mathcal{H}_{\mathrm{F},n} = \mathrm{span}\{|x_1, \ldots, x_n\rangle : x_i \in \{0,1\}\}. \tag{5}$$

Therefore,

$$\mathcal{H}_{\mathrm{F},n} \simeq (\mathbb{C}^2)^{\otimes n} \tag{6}$$

as a Hilbert space.

For bosons, the annihilation and creation operators a_j and $a_j^\dagger$ act on the Fock basis as

$$a_j |k_1, \ldots, k_j, \ldots, k_n\rangle = \sqrt{k_j}\, |k_1, \ldots, k_j - 1, \ldots, k_n\rangle, \tag{7}$$

$$a_j^\dagger |k_1, \ldots, k_j, \ldots, k_n\rangle = \sqrt{k_j + 1}\, |k_1, \ldots, k_j + 1, \ldots, k_n\rangle. \tag{8}$$

They satisfy the canonical commutation relations

$$[a_i, a_j^\dagger] = \delta_{ij} I, \qquad [a_i, a_j] = [a_i^\dagger, a_j^\dagger] = 0. \tag{9}$$

The number operator of mode j is

$$N_j = a_j^\dagger a_j, \tag{10}$$

and satisfies

$$N_j |k_1, \ldots, k_n\rangle = k_j |k_1, \ldots, k_n\rangle. \tag{11}$$

For fermions, it is useful to give an explicit representation of the creation and annihilation operators on qubits. This is provided by the Jordan–Wigner transformation. Let

$$\sigma^+ = |1\rangle\langle 0|, \qquad \sigma^- = |0\rangle\langle 1|, \qquad Z = |0\rangle\langle 0| - |1\rangle\langle 1|. \tag{12}$$

The fermionic annihilation and creation operators are represented as

$$f_j = Z_1 \cdots Z_{j-1} \sigma_j^-, \qquad f_j^\dagger = Z_1 \cdots Z_{j-1} \sigma_j^+. \tag{13}$$

Equivalently, on the occupation-number basis,

$$f_j |x_1, \ldots, x_j, \ldots, x_n\rangle = (-1)^{\sum_{\ell<j} x_\ell} x_j |x_1, \ldots, 0, \ldots, x_n\rangle, \tag{14}$$

$$f_j^\dagger |x_1, \ldots, x_j, \ldots, x_n\rangle = (-1)^{\sum_{\ell<j} x_\ell} (1 - x_j) |x_1, \ldots, 1, \ldots, x_n\rangle. \tag{15}$$

The string $Z_1 \cdots Z_{j-1}$ is responsible for the correct fermionic signs. The operators satisfy the canonical anti-commutation relations

$$\{f_i, f_j^\dagger\} = \delta_{ij} I, \qquad \{f_i, f_j\} = \{f_i^\dagger, f_j^\dagger\} = 0. \tag{16}$$

In particular,

$$f_j^2 = 0, \qquad (f_j^\dagger)^2 = 0. \tag{17}$$

This is the algebraic form of the Pauli exclusion principle: a fermionic mode cannot be occupied by more than one particle. The fermionic number operator is

$$n_j = f_j^\dagger f_j, \tag{18}$$

and satisfies

$$n_j \left|x_1, \ldots, x_n\right\rangle = x_j \left|x_1, \ldots, x_n\right\rangle. \tag{19}$$

For bosonic systems, it is often convenient to pass from creation and annihilation operators to quadrature operators. For each mode j, define

$$x_j = \frac{a_j + a_j^\dagger}{\sqrt{2}}, \qquad p_j = \frac{a_j - a_j^\dagger}{i\sqrt{2}}. \tag{20}$$

Equivalently,

$$a_j = \frac{x_j + ip_j}{\sqrt{2}}, \qquad a_j^\dagger = \frac{x_j - ip_j}{\sqrt{2}}. \tag{21}$$

Collecting all quadratures into the vector

$$R = (x_1, p_1, \ldots, x_n, p_n)^T, \tag{22}$$

the canonical commutation relations take the form

$$[R_k, R_\ell] = i\Omega_{k\ell} I, \qquad \Omega = \bigoplus_{j=1}^{n} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{23}$$

For fermionic systems, the corresponding self-adjoint operators are the Majorana operators. For each mode j, define

$$\gamma_{2j-1} = f_j + f_j^\dagger, \qquad \gamma_{2j} = -i(f_j - f_j^\dagger). \tag{24}$$

They satisfy

$$\gamma_\mu^\dagger = \gamma_\mu, \qquad \{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} I. \tag{25}$$

Thus bosonic systems are naturally described by quadratures satisfying commutation relations, whereas fermionic systems are naturally described by Majorana operators satisfying anti-commutation relations.

2 Quantum state tomography

We now recall the standard formulation of quantum state tomography. Let $\mathcal{H}$ be a D-dimensional Hilbert space, and let $\mathcal{D}(\mathcal{H})$ denote the set of density operators on $\mathcal{H}$. A tomography problem is specified by a subset

$$\mathcal{S} \subseteq \mathcal{D}(\mathcal{H}), \tag{26}$$

which represents the prior information on the unknown state.

The learner is given N copies of an unknown state $\rho \in \mathcal{S}$, namely the joint state

$$\rho^{\otimes N}. \tag{27}$$

A tomography protocol performs a measurement on $\rho^{\otimes N}$, followed by classical post-processing, and outputs a classical description of an estimate $\widehat{\rho}$. Unless stated otherwise, the measurement is allowed to be collective across the N copies.

Bosons	**Fermions**
One mode is the infinite-dimensional analogue of a qudit.	One mode is the occupation-number analogue of a qubit.
$\mathcal{H}_{\mathrm{B},n} = \mathrm{span}\{\|k_1,\ldots,k_n\rangle : k_i \in \mathbb{N}\}$.	$\mathcal{H}_{\mathrm{F},n} = \mathrm{span}\{\|x_1,\ldots,x_n\rangle : x_i \in \{0,1\}\}$.
Each mode has arbitrary occupation number.	Each mode has occupation number 0 or 1.
Creation and annihilation operators satisfy commutation relations.	Creation and annihilation operators satisfy anti-commutation relations.
Canonical variables: quadratures x_j, p_j.	Canonical variables: Majorana operators $\gamma_{2j-1}, \gamma_{2j}$.

Table 1: Basic dictionary between bosonic and fermionic modes.

Definition 2.1 (Quantum state tomography of a class of states)**.** Let $\varepsilon, \delta \in (0,1)$. A tomography protocol learns the class $\mathcal{S}$ with accuracy ε and failure probability δ if, for every $\rho \in \mathcal{S}$, its output $\widehat{\rho}$ satisfies

$$\Pr\left[\frac{1}{2}\|\widehat{\rho} - \rho\|_1 \leq \varepsilon\right] \geq 1 - \delta. \tag{28}$$

The probability is taken over the measurement outcome and over any internal randomness of the protocol.

Definition 2.2 (Sample complexity)**.** The sample complexity of tomography for the class $\mathcal{S}$, denoted by

$$N_{\mathrm{tom}}(\mathcal{S}, \varepsilon, \delta), \tag{29}$$

is the smallest integer N for which there exists a tomography protocol using $\rho^{\otimes N}$ that learns $\mathcal{S}$ with accuracy ε and failure probability δ. When δ is fixed to a constant, we write

$$N_{\mathrm{tom}}(\mathcal{S}, \varepsilon) \tag{30}$$

and suppress the dependence on δ.

The optimal sample complexity is known for general finite-dimensional state tomography. Let

$$\mathcal{S}_{\mathrm{mixed}}(D) = \mathcal{D}(\mathbb{C}^D) \tag{31}$$

be the class of all mixed states on a D-dimensional Hilbert space. Then

$$N_{\mathrm{tom}}(\mathcal{S}_{\mathrm{mixed}}(D), \varepsilon) = \Theta\left(\frac{D^2}{\varepsilon^2}\right). \tag{32}$$

On the other hand, let

$$\mathcal{S}_{\mathrm{pure}}(D) = \{|\psi\rangle\langle\psi| : |\psi\rangle \in \mathbb{C}^D,\ \||\psi\rangle\|_2 = 1\} \tag{33}$$

be the class of pure states. Then

$$N_{\mathrm{tom}}(\mathcal{S}_{\mathrm{pure}}(D), \varepsilon) = \Theta\left(\frac{D}{\varepsilon^2}\right). \tag{34}$$

These bounds are due to O'Donnell and Wright [1] and Haah, Harrow, Ji, Wu, and Yu [2]. (Recently it was also shown a simple-to-prove algorithm for mixed-state tomography as well [3] based on the so-called random purification channel, which shows how mixed-state tomography reduces to pure-state tomography. This will have applications also for bosonic and fermionic Gaussian state learning due to its general group-theoretic working principle.)

As a consequence, full tomography of arbitrary n-qubit states is exponentially expensive in n. Indeed, in this case $D = 2^n$, and therefore tomography of general mixed states requires $\Theta\left(\frac{4^n}{\varepsilon^2}\right)$ copies.

The same conclusion applies to arbitrary n-modes fermionic states, since they have an Hilbert space isomorphic to that of n qubits.

2.1 Quantum state tomography in continuous-variable systems

We now discuss quantum state tomography for continuous-variable systems. In the bosonic setting, an n-mode system is described by the infinite-dimensional Hilbert space

$$\mathcal{H}_{\mathrm{B},n} = \mathrm{span}\{|k_1, \ldots, k_n\rangle : k_i \in \mathbb{N}\}. \tag{35}$$

Thus, in contrast to an n-qubit system, there is no finite local dimension. Consequently, without any further restriction on the class of states, tomography is impossible: no finite number of copies can suffice to learn an arbitrary state in $\mathcal{D}(\mathcal{H}_{\mathrm{B},n})$ to fixed trace-distance accuracy.

The physically relevant way to make the problem meaningful is to impose an energy constraint. Let

$$\widehat{N} = \sum_{j=1}^{n} a_j^\dagger a_j \tag{36}$$

be the total number operator. On the Fock basis,

$$\widehat{N}\,|k_1, \ldots, k_n\rangle = \left(\sum_{j=1}^{n} k_j\right) |k_1, \ldots, k_n\rangle\,. \tag{37}$$

For $E > 0$, define the class of n-mode pure states with mean photon number at most E per mode as

$$\mathcal{S}_{\mathrm{CV}}^{\mathrm{pure}}(n, E) = \left\{|\psi\rangle\langle\psi| : |\psi\rangle \in \mathcal{H}_{\mathrm{B},n},\ \||\psi\rangle\|_2 = 1,\ \mathrm{Tr}\left[|\psi\rangle\langle\psi|\,\widehat{N}\right] \le nE\right\}. \tag{38}$$

Theorem 2.3 (Tomography of energy-constrained pure CV states [4])**.** *For the class $\mathcal{S}_{\mathrm{CV}}^{\mathrm{pure}}(n, E)$, the sample complexity of tomography in trace distance satisfies*

$$N_{\mathrm{tom}}\left(\mathcal{S}_{\mathrm{CV}}^{\mathrm{pure}}(n, E), \varepsilon\right) = \widetilde{\Theta}\left(\frac{E^n}{\varepsilon^{2n}}\right). \tag{39}$$

Proof sketch. We first provide a proof sketch of the upper bound. The idea is that the energy constraint allows us to truncate the infinite-dimensional Hilbert space effectively to a finite-dimensional subspace. For $K \in \mathbb{N}$, let $\Pi_{\le K} := \sum_{k_1+\cdots+k_n \le K} |k_1, \ldots, k_n\rangle\langle k_1, \ldots, k_n|$ be the projector onto the subspace with total photon number at most K. Its dimension is $D_K := \mathrm{Tr}[\Pi_{\le K}] = \binom{n+K}{n}$. If ρ satisfies $\mathrm{Tr}\left[\rho\widehat{N}\right] \le nE$, then Markov's inequality gives $\mathrm{Tr}[(I - \Pi_{\le K})\rho] \le \mathrm{Tr}\left[\rho\widehat{N}\right]/K \le nE/K$. Thus, choosing $K \simeq nE/\varepsilon^2$, we get $\mathrm{Tr}[\Pi_{\le K}\rho] \ge 1 - O(\varepsilon^2)$. By the gentle measurement lemma, the normalized truncation $\rho_{\le K} := \Pi_{\le K}\rho\Pi_{\le K}/\,\mathrm{Tr}[\Pi_{\le K}\rho]$ satisfies $\frac{1}{2}\|\rho - \rho_{\le K}\|_1 \le O(\varepsilon)$. Hence, up to trace-distance error $O(\varepsilon)$, every energy-constrained state

is supported on a subspace of dimension $D_K = \binom{n+K}{n} = K^{O(n)} = \widetilde{O}(E^n/\varepsilon^{2n})$. Applying the finite-dimensional pure-state tomography upper bound on this effective subspace gives

$$N_{\mathrm{tom}}\left(\mathcal{S}_{\mathrm{CV}}^{\mathrm{pure}}(n,E),\varepsilon\right) \leq \widetilde{O}\left(\frac{E^n}{\varepsilon^{2n}}\right).$$

We now sketch the lower bound. The idea is to embed a large finite-dimensional pure-state tomography problem inside the energy-constrained Hilbert space. Let K be a cutoff and define $\mathcal{H}_K^\circ := \mathrm{Ran}(\Pi_{\leq K}) \cap (|0\rangle^{\otimes n})^\perp$. Its dimension is $d_K := \dim \mathcal{H}_K^\circ = \binom{n+K}{n} - 1 = \widetilde{\Theta}(K^n)$. Choose a packing $\mathcal{U} \subset \mathcal{H}_K^\circ$ of unit vectors with $\log|\mathcal{U}| = \Omega(d_K)$ and with pairwise overlaps bounded away from one. For every $u \in \mathcal{U}$, define

$$|\psi_u\rangle = \sqrt{1-\alpha}\,|0\rangle^{\otimes n} + \sqrt{\alpha}\,|u\rangle,$$

where $\alpha \simeq \varepsilon^2$. These states satisfy the energy constraint: since $|u\rangle \in \mathrm{Ran}(\Pi_{\leq K})$, we have $\langle u|\,\widehat{N}\,|u\rangle \leq K$, and since $|u\rangle \perp |0\rangle^{\otimes n}$, the cross terms vanish. Therefore $\langle\psi_u|\,\widehat{N}\,|\psi_u\rangle = \alpha\,\langle u|\,\widehat{N}\,|u\rangle \leq \alpha K$. Choosing $K \simeq nE/\alpha$, with suitable constants, gives $\langle\psi_u|\,\widehat{N}\,|\psi_u\rangle \leq nE$, so all these states belong to $\mathcal{S}_{\mathrm{CV}}^{\mathrm{pure}}(n,E)$.

The states are also separated in trace distance. Indeed, for $u \neq v$, $\langle\psi_u|\psi_v\rangle = 1 - \alpha + \alpha\,\langle u|v\rangle$. Since the vectors in $\mathcal{U}$ have overlaps bounded away from one,

$$\frac{1}{2}\||\psi_u\rangle\langle\psi_u| - |\psi_v\rangle\langle\psi_v|\|_1 = \sqrt{1 - |\,\langle\psi_u|\psi_v\rangle\,|^2} \geq c\sqrt{\alpha}.$$

Taking $\alpha \simeq \varepsilon^2$, the states are separated by trace distance of order ε. Hence any tomography protocol with error $O(\varepsilon)$ must distinguish a packing of size $|\mathcal{U}|$. The Holevo–Fano lower bound gives $N \geq \widetilde{\Omega}(d_K)$. Since $K \simeq nE/\varepsilon^2$, we get $d_K = \widetilde{\Theta}(K^n) = \widetilde{\Theta}(E^n/\varepsilon^{2n})$, and therefore

$$N_{\mathrm{tom}}\left(\mathcal{S}_{\mathrm{CV}}^{\mathrm{pure}}(n,E),\varepsilon\right) \geq \widetilde{\Omega}\left(\frac{E^n}{\varepsilon^{2n}}\right).$$

Combining the upper and lower bounds proves the claim. □

The theorem shows that the energy constraint makes continuous-variable tomography possible, but still extremely inefficient in general. Even for pure states and constant energy per mode, the copy complexity scales exponentially in the number of modes and has the characteristic ε^{-2n} dependence on the trace-distance error. This is in sharp contrast with finite-dimensional pure-state tomography, where the dependence on the trace-distance error is only ε^{-2}.

2.2 Preliminaries: Bosonic and fermionic Gaussian states

We now introduce bosonic and fermionic Gaussian states. In both settings, Gaussian states are associated with quadratic Hamiltonians and are completely specified by their so-called covariance matrices. Throughout this subsection we restrict to the zero-first-moment setting.

Let $R = (x_1, p_1, \ldots, x_n, p_n)^T$ be the bosonic quadrature vector, satisfying

$$[R_j, R_k] = i\Omega_{jk} I. \tag{40}$$

A mixed bosonic Gaussian state is a state of the form

$$\rho = \frac{e^{-H}}{\mathrm{Tr}[e^{-H}]}, \tag{41}$$

where H is quadratic in the quadrature operators. Such a Hamiltonian can be written as

$$H = \frac{1}{2} R^T A R \tag{42}$$

Bosons	**Fermions**
$R = (x_1, p_1, \ldots, x_n, p_n)^T, \quad [R_j, R_k] = i\Omega_{jk} I$	$\gamma_1, \ldots, \gamma_{2n}, \quad \{\gamma_j, \gamma_k\} = 2\delta_{jk} I$
$V_{jk}(\rho) = \mathrm{Tr}[\rho\{R_j, R_k\}]$	$\Gamma_{jk}(\rho) = -\frac{i}{2}\,\mathrm{Tr}[\rho[\gamma_j, \gamma_k]]$
$V(\lvert 0\rangle\langle 0\rvert^{\otimes n}) = I_{2n}$	$\Gamma(\lvert 0\rangle\langle 0\rvert^{\otimes n}) = \bigoplus_{j=1}^n \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
$U_S^\dagger R U_S = SR, \quad S\Omega S^T = \Omega$	$U_O^\dagger \gamma_j U_O = \sum_{k=1}^{2n} O_{jk}\gamma_k, \quad O \in \mathrm{SO}(2n)$
$V(U_S \rho U_S^\dagger) = SV(\rho)S^T$	$\Gamma(U_O \rho U_O^\dagger) = O\Gamma(\rho)O^T$
$V(\rho) = S\left(\bigoplus_{j=1}^n \nu_j I_2\right) S^T, \quad \nu_j \geq 1$	$\Gamma(\rho) = O\left(\bigoplus_{j=1}^n \begin{pmatrix} 0 & \lambda_j \\ -\lambda_j & 0 \end{pmatrix}\right) O^T, \quad O \in \mathrm{O}(2n), \quad 0 \leq \lambda_j \leq 1$

Table 2: Basic dictionary for bosonic and fermionic Gaussian states.

for a suitable real symmetric matrix A.

Similarly, let $\gamma_1, \ldots, \gamma_{2n}$ be the Majorana operators of an n-mode fermionic system, satisfying

$$\{\gamma_j, \gamma_k\} = 2\delta_{jk} I. \tag{43}$$

A mixed fermionic Gaussian state is a state of the form

$$\rho = \frac{e^{-H}}{\mathrm{Tr}[e^{-H}]}, \tag{44}$$

where H is quadratic in the Majorana operators. Equivalently,

$$H = \frac{i}{4} \sum_{j,k=1}^{2n} h_{jk}\gamma_j\gamma_k, \tag{45}$$

where $h \in \mathbb{R}^{2n \times 2n}$ is real and antisymmetric.

Gaussian unitaries are defined analogously as unitaries generated by quadratic Hamiltonians. In the bosonic case, a Gaussian unitary is specified by a symplectic matrix S, satisfying $S\Omega S^T = \Omega$. The corresponding unitary U_S acts on quadratures as

$$U_S^\dagger R U_S = SR. \tag{46}$$

Equivalently, component-wise,

$$U_S^\dagger R_j U_S = \sum_{k=1}^{2n} S_{jk} R_k. \tag{47}$$

In the fermionic case, a Gaussian unitary is specified by an orthogonal matrix $O \in \mathrm{O}(2n)$. The corresponding unitary U_O acts on Majorana operators as

$$U_O^\dagger \gamma_j U_O = \sum_{k=1}^{2n} O_{jk}\gamma_k. \tag{48}$$

Pure Gaussian states are obtained by applying Gaussian unitaries to the vacuum. Thus, in the bosonic case,

$$|\psi\rangle = U_S \,|0\rangle^{\otimes n}, \tag{49}$$

while in the fermionic case,

$$|\psi\rangle = U_O |0\rangle^{\otimes n}. \tag{50}$$

These are also called states prepared by matchgate circuits or fermionic linear optics.

We now define the covariance matrices. For a possibly non-Gaussian bosonic state ρ with zero first moment, the covariance matrix is the real symmetric matrix

$$V_{jk}(\rho) = \operatorname{Tr}\left[\rho\{R_j, R_k\}\right], \qquad j,k \in [2n]. \tag{51}$$

It satisfies the uncertainty relation

$$V(\rho) + i\Omega \geq 0. \tag{52}$$

The bosonic vacuum has covariance matrix

$$V\left(|0\rangle\langle 0|^{\otimes n}\right) = I_{2n}. \tag{53}$$

Moreover, if $\rho' = U_S \rho U_S^\dagger$, then one can verify that

$$V(\rho') = SV(\rho)S^T. \tag{54}$$

If ρ is bosonic Gaussian with zero first moment, then $V(\rho)$ uniquely specifies the state.

For a possibly non-Gaussian fermionic state ρ, the covariance matrix is the real antisymmetric matrix

$$\Gamma_{jk}(\rho) = -\frac{i}{2}\operatorname{Tr}\left[\rho[\gamma_j, \gamma_k]\right], \qquad j,k \in [2n]. \tag{55}$$

It satisfies

$$\Gamma(\rho)^T = -\Gamma(\rho), \qquad \|\Gamma(\rho)\|_\infty \leq 1. \tag{56}$$

The fermionic vacuum has covariance matrix

$$\Gamma\left(|0\rangle\langle 0|^{\otimes n}\right) = \bigoplus_{j=1}^{n} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{57}$$

Moreover, if $\rho' = U_O \rho U_O^\dagger$, then

$$\Gamma(\rho') = O\Gamma(\rho)O^T. \tag{58}$$

If ρ is fermionic Gaussian, then $\Gamma(\rho)$ uniquely specifies the state.

We now recall two standard matrix normal forms. First, Williamson's theorem says that every real positive definite matrix $V = V^T > 0$ can be written as

$$V = S\left(\bigoplus_{j=1}^{n} \nu_j I_2\right) S^T, \tag{59}$$

where S is symplectic and $\nu_j > 0$. The numbers ν_j are the symplectic eigenvalues. If V is a bosonic covariance matrix, then the uncertainty relation implies $\nu_j \geq 1$.

Second, every real antisymmetric matrix $A = -A^T$ admits the skew-normal form

$$A = O\left(\bigoplus_{j=1}^{n} \begin{pmatrix} 0 & \lambda_j \\ -\lambda_j & 0 \end{pmatrix}\right) O^T, \tag{60}$$

where $O \in \mathrm{O}(2n)$ and $\lambda_j \geq 0$. If $A = \Gamma$ is a fermionic covariance matrix, then $0 \leq \lambda_j \leq 1$.

We now apply these normal forms to Gaussian states. In the bosonic case, Williamson's theorem implies that every valid covariance matrix can be written as

$$V = S\left(\bigoplus_{j=1}^{n} \nu_j I_2\right) S^T, \tag{61}$$

where S is symplectic and $\nu_j \geq 1$. Hence every zero-mean bosonic Gaussian state can be written as

$$\rho = U_S\left(\tau_{\frac{\nu_1-1}{2}} \otimes \cdots \otimes \tau_{\frac{\nu_n-1}{2}}\right) U_S^\dagger, \tag{62}$$

where τ_μ is the one-mode thermal state with mean photon number μ:

$$\tau_\mu = \frac{1}{\mu+1} \sum_{k=0}^{\infty} \left(\frac{\mu}{\mu+1}\right)^k |k\rangle\langle k| . \tag{63}$$

The state is pure if and only if $\nu_1 = \cdots = \nu_n = 1$. A pure bosonic Gaussian covariance matrix has the form

$$V = SS^T \tag{64}$$

for some symplectic matrix S.

The symplectic matrix S admits an Euler decomposition

$$S = O_1 Z O_2, \tag{65}$$

where O_1 and O_2 are orthogonal symplectic matrices and

$$Z = \bigoplus_{j=1}^{n} \begin{pmatrix} z_j^{-1} & 0 \\ 0 & z_j \end{pmatrix}, \qquad z_j \geq 1. \tag{66}$$

The squeezing of S is defined as

$$\|S\|_\infty = \|Z\|_\infty = \max_j z_j. \tag{67}$$

If $Z = I$, then the squeezing is equal to 1. In this case S is both symplectic and orthogonal. The group

$$\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n) \tag{68}$$

is compact and is isomorphic to $\mathrm{U}(n)$. The corresponding bosonic Gaussian unitaries are so-called "particle-number preserving" and are called passive Gaussian unitaries.

In the fermionic case, every covariance matrix admits the normal form

$$\Gamma = O\left(\bigoplus_{j=1}^{n} \begin{pmatrix} 0 & \lambda_j \\ -\lambda_j & 0 \end{pmatrix}\right) O^T, \tag{69}$$

where $O \in \mathrm{O}(2n)$ and $0 \leq \lambda_j \leq 1$. The corresponding fermionic Gaussian state is

$$\rho = U_O\left(\rho_{\lambda_1} \otimes \cdots \otimes \rho_{\lambda_n}\right) U_O^\dagger, \tag{70}$$

where

$$\rho_\lambda = \frac{1+\lambda}{2} |0\rangle\langle 0| + \frac{1-\lambda}{2} |1\rangle\langle 1| . \tag{71}$$

A fermionic Gaussian state is pure if and only if $\lambda_1 = \cdots = \lambda_n = 1$. Equivalently, a fermionic Gaussian state is pure if and only if

$$\Gamma^2 = -I. \tag{72}$$

In the general bosonic case, a Gaussian unitary also includes a displacement. It has the form

$$G_{r,S} = D_r U_S, \qquad D_r = e^{-ir^T \Omega R}, \tag{73}$$

where $r \in \mathbb{R}^{2n}$. It acts on the quadrature operators as

$$G_{r,S}^\dagger R_j G_{r,S} = \sum_{k=1}^{2n} S_{jk} R_k + r_j I, \qquad j \in [2n]. \tag{74}$$

Equivalently,

$$G_{r,S}^\dagger R G_{r,S} = SR + rI. \tag{75}$$

In these notes we mostly suppress r, since we restrict to the zero-first-moment setting.

Thus, in the zero-first-moment setting, Gaussian state tomography reduces to learning a covariance matrix: V for bosons and Γ for fermions. In the general bosonic case one must also learn the displacement vector. This is the basic reason why Gaussian state tomography can be efficient, in contrast with tomography of arbitrary bosonic or fermionic states.

2.3 Trace-distance bounds and Gaussian state tomography

Gaussian states are specified by their covariance matrices. For tomography, however, one needs a quantitative stability statement: if two covariance matrices are ε-close, how does this covariance error propagate to the trace norm $\|\cdot\|_1$ between the corresponding Gaussian states?

Let ρ_Γ and $\rho_{\Gamma'}$ denote fermionic Gaussian states with Majorana covariance matrices Γ and Γ'. For pure fermionic Gaussian states [5], one has

$$\|\rho_\Gamma - \rho_{\Gamma'}\|_1 \le \frac{1}{2}\|\Gamma - \Gamma'\|_2. \tag{76}$$

For mixed fermionic Gaussian states, one has

$$\|\rho_\Gamma - \rho_{\Gamma'}\|_1 \le \frac{1}{2}\|\Gamma - \Gamma'\|_1. \tag{77}$$

Let ρ_V and ρ_W denote the zero-mean bosonic Gaussian states with covariance matrices V and W. For pure bosonic Gaussian states [6], one has

$$\|\rho_V - \rho_W\|_1 = O\left(\left\|V^{-1/2}(V-W)W^{-1/2}\right\|_2\right). \tag{78}$$

For mixed bosonic Gaussian states [7], one has

$$\|\rho_V - \rho_W\|_1 = O\left(\mathrm{Tr}\left[\left(V^{-1} + W^{-1}\right)|V - W|\right]\right), \tag{79}$$

where $|A| = (A^\dagger A)^{1/2}$.

2.4 Tomography of Gaussian states

The trace-distance bounds above have a direct algorithmic consequence. To learn a Gaussian state in trace distance, it is enough to estimate its covariance matrix in the appropriate norm and then output the Gaussian state with the estimated covariance matrix.

Setting	Trace-norm bound
Bosonic pure Gaussian states [6]	$\|\rho_V - \rho_W\|_1 = O\big(\big\|V^{-1/2}(V-W)W^{-1/2}\big\|_2\big)$
Bosonic mixed Gaussian states [8]	$\|\rho_V - \rho_W\|_1 = O\big(\mathrm{Tr}\big[(V^{-1}+W^{-1})\|V-W\|\big]\big)$
Fermionic pure Gaussian states [5]	$\|\rho_\Gamma - \rho_{\Gamma'}\|_1 \le \frac{1}{2}\|\Gamma - \Gamma'\|_2$
Fermionic mixed Gaussian states [5]	$\|\rho_\Gamma - \rho_{\Gamma'}\|_1 \le \frac{1}{2}\|\Gamma - \Gamma'\|_1$

Table 3: Trace-norm bounds for Gaussian states in terms of covariance matrices.

2.4.1 Learning fermionic Gaussian states

We first discuss fermionic Gaussian states. For pure fermionic Gaussian states, the bound $\|\rho_\Gamma - \rho_{\Gamma'}\|_1 \le \frac{1}{2}\|\Gamma - \Gamma'\|_2$ shows that trace-distance tomography reduces to estimating the covariance matrix in Frobenius norm. The currently standard sample- and time-efficient way to do this covariance estimation is: one partitions the Majorana two-point observables into commuting families and measures them jointly, or equivalently uses matchgate-shadow-type measurements [9]. This gives $N = O(n^3/\varepsilon^2)$ copies for learning pure fermionic Gaussian states in trace distance [5].

This scaling is not information-theoretically optimal. The pure fermionic Gaussian manifold has $O(n^2)$ real parameters, and one can prove the lower bound $N = \Omega(n^2/\varepsilon^2)$. Moreover, this lower bound is achievable by statistically optimal but less practical procedures. For pure fermionic Gaussian states, covariant multi-copy measurements over the Gaussian manifold achieve $N = O(n^2/\varepsilon^2)$ copies [10]. One can also obtain an optimal information-theoretic procedure by using global Clifford classical-shadow measurements [11] and then optimizing over a covering net of pure Gaussian states.

For (general) mixed fermionic Gaussian states, the trace-distance bound is $\|\rho_\Gamma - \rho_{\Gamma'}\|_1 \le \frac{1}{2}\|\Gamma - \Gamma'\|_1$. Using the norm conversion $\|\Gamma - \Gamma'\|_1 \le \sqrt{2n}\|\Gamma - \Gamma'\|_2$, it is enough to estimate Γ in Frobenius norm to accuracy $O(\varepsilon/\sqrt{n})$. Plugging this into the covariance-estimation algorithm gives $N = O(n^4/\varepsilon^2)$ copies [5]. At the information-theoretic but experimentally and computationally challening level, however, the fermionic analogue of the random purification channel reduces mixed fermionic Gaussian tomography to pure fermionic Gaussian tomography, giving the optimal scaling $N = \Theta(n^2/\varepsilon^2)$ also for mixed fermionic Gaussian states [10].

Thus, for fermions, covariance estimation gives the current practical algorithms: $O(n^3/\varepsilon^2)$ for pure states and $O(n^4/\varepsilon^2)$ for mixed states. The statistically optimal scaling is $\Theta(n^2/\varepsilon^2)$, but whether is possible obtaining this scaling with a time-efficient and practical (e.g., single-copy) algorithm remains open.

2.4.2 Learning bosonic Gaussian states

The bosonic case can seem more subtle. In contrast to fermions, the Hilbert space is infinite-dimensional, so a priori one expects the sample complexity to depend on an energy bound E.

For Gaussian states, however, the relative form of the trace-distance bounds shown above suggests that this energy dependence should not be fundamental. The relevant error is not an absolute covariance error such as $\left\|V - \widehat{V}\right\|$, but a relative covariance error adapted to the scale of the state. This intuition is made precise in Ref. [8]. There, it is shown that general bosonic Gaussian states can be learned with only a doubly logarithmic dependence on the energy.

Specifically, the sample complexity is

$$N = \widetilde{O}\left(\frac{n^3}{\varepsilon^2} + n \log\log E\right). \tag{80}$$

The algorithm is based on Gaussian measurements and "an adaptive unsqueezing procedure": one first obtains a rough estimate of the covariance matrix, uses it to approximately "undo the squeezing" of the unknown state, and then repeats the procedure on a better-conditioned state. In this way, highly squeezed directions are progressively brought to a regime where standard covariance estimation is efficient.

For pure bosonic Gaussian states, the optimal scaling is

$$N = \widetilde{\Theta}\left(\frac{n^2}{\varepsilon^2}\right). \tag{81}$$

This can be obtained by combining the energy-independent tomography approach of Ref. [8] with the pure-state tomography results of Ref. [6]. Thus pure bosonic Gaussian states can be learned sample-optimally in a practical manner using single-copy Gaussian measurements. For general mixed bosonic Gaussian states, the best efficient upper bound is currently

$$N = \widetilde{O}\left(\frac{n^3}{\varepsilon^2} + n \log\log E\right). \tag{82}$$

Moreover, the n^3/ε^2 scaling is optimal among Gaussian protocols [6]. Whether arbitrary, possibly non-Gaussian, measurements can improve this to the information-theoretic scaling $\Omega(n^2/\varepsilon^2)$ for general mixed bosonic Gaussian states remains open.

2.4.3 Effectively energy-independent Gaussian state tomography

Let us explain the mechanism behind the almost energy-independent scaling. The starting point is the trace-distance bound [8]

$$\left\|\rho_V - \rho_{\widehat{V}}\right\|_1 \leq O\left(\mathrm{Tr}\left[\left(V^{-1} + \widehat{V}^{-1}\right)|V - \widehat{V}|\right]\right). \tag{83}$$

This says that the right notion of covariance error is relative error estimation. Ideally, one would like an estimate $\widehat{V}$ such that

$$\left|\widehat{V} - V\right| \preceq \eta V. \tag{84}$$

Indeed, in that case one gets

$$\mathrm{Tr}\left[V^{-1}|\widehat{V} - V|\right] \leq \eta\,\mathrm{Tr}\left[V^{-1}V\right] = 2n\eta. \tag{85}$$

Thus a relative covariance estimate with $\eta = O(\varepsilon/n)$ is enough to control the trace distance, independently of the energy or squeezing.

This is exactly what happens in classical covariance estimation. If one had direct samples from the underlying classical Gaussian probability distribution $\mathcal{N}(0, V)$

$$X_1, \ldots, X_N \sim \mathcal{N}(0, V), \tag{86}$$

then empirical covariance estimation naturally gives relative control of V, with no dependence on the condition number (minimum eigenvalue) of V. The problem is that, in the quantum setting, standard measurements do not directly give samples directly from $\mathcal{N}(0, V)$, but something close

to it. For example, heterodyne measurement on a zero-mean Gaussian state with covariance V gives samples from

$$X \sim \mathcal{N}\left(0, \frac{V+I}{2}\right). \tag{87}$$

Thus the measured covariance is

$$\Sigma = \frac{V+I}{2}, \tag{88}$$

(the identity term corresponds to the so-called vacuum noise added by the measurement).

So from heterodyne data we can estimate Σ in relative error:

$$|\widehat{\Sigma} - \Sigma| \preceq \eta\Sigma. \tag{89}$$

If we set $\widehat{V} = 2\widehat{\Sigma} - I$, then

$$|\widehat{V} - V| = 2|\widehat{\Sigma} - \Sigma| \preceq 2\eta\Sigma = \eta(V+I). \tag{90}$$

Plugging this into the relative trace-distance expression gives the schematic bound

$$\operatorname{Tr}\left[V^{-1}|\widehat{V} - V|\right] \lesssim \eta \operatorname{Tr}\left[V^{-1}(V+I)\right] \tag{91}$$

$$= \eta\left(2n + \operatorname{Tr}\left[V^{-1}\right]\right). \tag{92}$$

The bad term is $\operatorname{Tr}[V^{-1}]$. For a highly squeezed pure Gaussian state, $V = SS^T$, and if

$$S = O_1\left(\bigoplus_{j=1}^{n}\begin{pmatrix} z_j^{-1} & 0 \\ 0 & z_j \end{pmatrix}\right) O_2, \tag{93}$$

then the eigenvalues of V contain z_j^{-2} and z_j^2. Hence

$$\left\|V^{-1}\right\|_\infty = \max_j z_j^2, \qquad \operatorname{Tr}\left[V^{-1}\right] \le 2n \max_j z_j^2. \tag{94}$$

This is where the squeezing, and therefore the energy dependence, enters the naive heterodyne approach.

The adaptive idea is close in spirit to what one often does in quantum metrology or Hamiltonian learning: first learn a rough approximation, and then use this rough information to choose a better measurement. The first estimate does not need to be accurate enough for tomography. Its role is only to identify the large squeezing directions approximately. One then changes coordinates accordingly, learns again in the new coordinates, and repeats. Schematically, the idea is

$$\text{rough estimate} \longrightarrow \text{partial unsqueezing} \longrightarrow \text{better conditioned state} \longrightarrow \text{rough estimate again.} \tag{95}$$

After a few rounds, the unknown state has been transformed to a regime where ordinary covariance estimation is well conditioned. Only at the final stage we spend the samples needed for the target accuracy ε.

Let us make this more concrete. A heterodyne measurement on a zero-mean Gaussian state with covariance V produces a classical Gaussian outcome with covariance

$$\Sigma = \frac{V+I}{2}. \tag{96}$$

The extra I is the vacuum noise added by performing the standard Gaussian "heterodyne" measurement. More generally, one can perform a Gaussian measurement in which the vacuum reference state is replaced by an auxiliary Gaussian state with covariance M. In this case the classical outcome has covariance

$$\Sigma_M = \frac{V + M}{2}. \tag{97}$$

Thus M is the covariance of the auxiliary Gaussian state used in the measurement. It controls the noise scale added by the measurement.

The ideal situation would be to choose M comparable to the unknown covariance V. Indeed, if

$$cV \preceq M \preceq CV \tag{98}$$

for universal constants $c, C > 0$, then

$$\frac{1+c}{2}V \preceq \Sigma_M \preceq \frac{1+C}{2}V. \tag{99}$$

In this case, a relative estimate of Σ_M gives a relative estimate of V. For example, if

$$\left|\widehat{\Sigma}_M - \Sigma_M\right| \preceq \eta\Sigma_M, \tag{100}$$

and if we output

$$\widehat{V} = 2\widehat{\Sigma}_M - M, \tag{101}$$

then

$$\begin{aligned} \left|\widehat{V} - V\right| &= 2\left|\widehat{\Sigma}_M - \Sigma_M\right| && (102)\\ &\preceq 2\eta\Sigma_M && (103)\\ &\preceq (1+C)\eta V. && (104) \end{aligned}$$

So the covariance error is relative to V, and no squeezing-dependent factor appears.

Of course, V is unknown, so one cannot choose $M \simeq V$ at the beginning. The protocol therefore bootstraps. It first obtains a crude estimate $\widehat{V}_0$. From this estimate one extracts an approximate squeezing transformation S_0 from its Williamson decomposition, and then applies an approximate inverse squeezing S_0^{-1}. In the new coordinates, the effective covariance matrix is

$$V_1 = S_0^{-1} V S_0^{-T}. \tag{105}$$

If S_0 captures the large squeezing directions even roughly, then V_1 is better conditioned than V. We then repeat the same idea: estimate V_1 roughly, partially unsqueeze again, and obtain an even better conditioned covariance matrix V_2. Thus the protocol produces a sequence

$$V \longmapsto V_1 \longmapsto V_2 \longmapsto \cdots \longmapsto V_L, \tag{106}$$

where the effective squeezing decreases at each stage. After

$$L = O(\log\log E) \tag{107}$$

rounds, the covariance matrix is sufficiently well conditioned, and a final covariance-estimation step gives the desired trace-distance guarantee.

In summary, the protocol is a bootstrap: learn roughly, rotate and unsqueeze a bit, learn roughly again, and only at the end estimate with the final precision.

Operationally, this can be implemented with standard Gaussian experimentally available tools like passive linear optics and homodyne detection [8].

2.5 Tomography of t-doped bosonic/fermionic Gaussian states and stabilizer states

Gaussian states are efficiently learnable because they have a compact description. The next question is what happens if the state is not exactly Gaussian, but differs from a Gaussian state by a small number of non-Gaussian gates. This is the t-doped setting [4, 12–15].

Definition 2.4 (t-doped Gaussian unitary)**.** Fix a class of Gaussian unitaries and a class of non-Gaussian gates, called magic gates. A unitary U_t is t-doped Gaussian if it can be written as

$$U_t = G_t W_t \cdots G_1 W_1 G_0, \tag{108}$$

where $G_0, \ldots, G_t$ are Gaussian unitaries and $W_1, \ldots, W_t$ are magic gates. A t-doped Gaussian state is a state of the form

$$|\psi\rangle = U_t \, |0\rangle^{\otimes n} . \tag{109}$$

The meaning of the magic gates depends on the model. In the bosonic setting, Gaussian unitaries are generated by Hamiltonians at most quadratic in the quadratures, while the magic gates are local non-Gaussian unitaries. In the simplest version, these are single-mode non-Gaussian gates, for instance gates generated by polynomial Hamiltonians of degree larger than two. In the fermionic setting, Gaussian unitaries are generated by quadratic Majorana Hamiltonians, while the magic gates are κ-local Majorana gates. A typical fermionic magic gate has the form

$$W = \exp\left(i\theta \, \gamma_{\mu_1} \cdots \gamma_{\mu_\kappa}\right), \tag{110}$$

where $\mu_1, \ldots, \mu_\kappa \in [2n]$. In most physical examples one has $\kappa = 4$, namely quartic Majorana interactions. Thus a (t, κ)-doped fermionic Gaussian state is obtained from arbitrary fermionic Gaussian dynamics and at most t κ-local non-Gaussian Majorana gates.

There is also a completely analogous stabilizer version. One replaces Gaussian unitaries by Clifford unitaries, and the magic gates are non-Clifford gates, such as T-gates [13–15]. Thus the three models are

$$\text{bosonic Gaussian circuits} + t \text{ local non-Gaussian gates}, \tag{111}$$

$$\text{fermionic Gaussian circuits} + t \text{ local Majorana gates}, \tag{112}$$

and

$$\text{Clifford circuits} + t \text{ non-Clifford gates}. \tag{113}$$

The reason these classes remain efficiently learnable for low t is a magic-compression theorem: all the non-Gaussian or non-Clifford part can be moved to a small subsystem.

For bosons with single-mode non-Gaussian gates, every t-doped Gaussian state can be written as

$$|\psi\rangle = G \left(u_{2t} \otimes I_{n-2t}\right) |0\rangle^{\otimes n} , \tag{114}$$

where G is a bosonic Gaussian unitary and u_{2t} is an arbitrary unitary acting only on $2t$ modes. Equivalently,

$$|\psi\rangle = G \left(|\phi\rangle \otimes |0\rangle^{\otimes n-2t} \right), \tag{115}$$

where $|\phi\rangle$ is a non-Gaussian $2t$-mode state. Thus the non-Gaussianity is confined to $2t$ effective modes after a Gaussian change of basis [4].

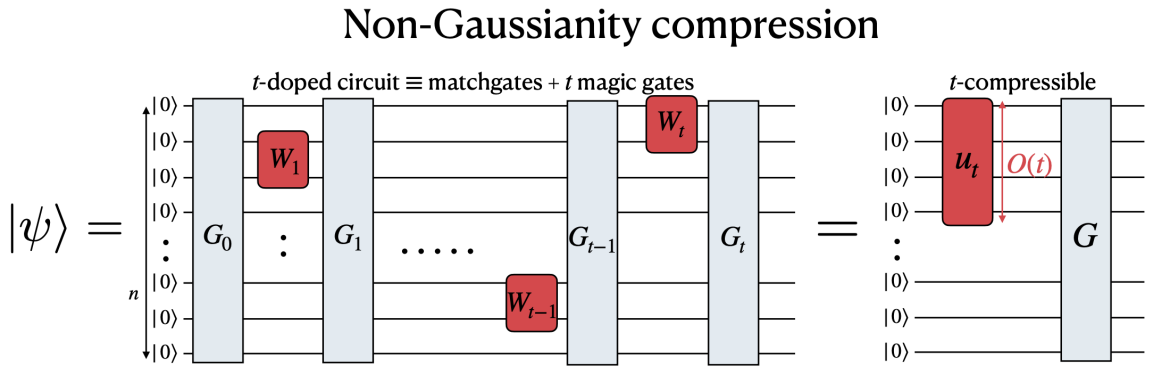

Figure 1: Schematic picture of t-doped Gaussian (or Clifford) circuit, and magic-compression for t-doped states. The non-Gaussian or non-Clifford part can be moved to a small subsystem after a Gaussian or Clifford change of basis. The tomography algorithm must first learn to decode this Gaussian or Clifford change of basis, then undoing it, and at the end perform state-tomography only on the small $O(t)$-subsystem.

For fermions, the analogous statement is that every (t, κ)-doped fermionic Gaussian state can be written as

$$|\psi\rangle = G\left(|\phi\rangle \otimes |0\rangle^{\otimes n-m}\right), \qquad m = O(\kappa t), \tag{116}$$

where G is a fermionic Gaussian unitary and $|\phi\rangle$ is an arbitrary state on m fermionic modes.

For stabilizer states with t non-Clifford gates, the same form holds with the Gaussian G replaced by Clifford C.

The learning algorithm follows from this compression form. First, learn the "compressing" Gaussian G or Clifford C. For bosonic and fermionic Gaussian states this is done by covariance estimation followed by classical post-processing: one estimates the covariance matrix and extracts (via Williamson or normal form decomposition) the Gaussian change of basis which isolates the non-Gaussian subsystem. For stabilizer states, the analogous step is to learn the Clifford frame, for instance by Bell sampling or random Clifford measurements. Second, undo the learned Gaussian or Clifford unitary. After this step, the state is supported on a small subsystem of size $m = O(t)$, tensored with vacuum or stabilizer degrees of freedom. One then performs ordinary tomography only on this small subsystem.

Therefore the cost is polynomial in n and exponential only in the amount of magic. In the fermionic case, for fixed κ, the trace-distance tomography algorithm has sample and time complexity

$$\mathrm{poly}\left(n, 2^t, \frac{1}{\varepsilon}\right). \tag{117}$$

Thus the algorithm is efficient whenever $t = O(\log n)$, and this is tight: beyond this number $t = \omega(\log n)$ of magic-gates, any tomography algorithm must be time-inefficient up to common crypto-assumptions [12]. Similar scaling holds in the bosonic and Clifford settings as well [4, 13–15].

References

[1] Ryan O'Donnell and John Wright. Efficient quantum tomography. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, STOC '16, pages 899–912. Association for Computing Machinery, 2016. doi: 10.1145/2897518.2897544.

[2] Jeongwan Haah, Aram W. Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. *IEEE Transactions on Information Theory*, 63(9): 5628–5641, 2017. doi: 10.1109/TIT.2017.2719044.

[3] Angelos Pelecanos, Jack Spilecki, Ewin Tang, and John Wright. Mixed state tomography reduces to pure state tomography, 2025. URL https://arxiv.org/abs/2511.15806.

[4] Francesco A. Mele, Antonio A. Mele, Lennart Bittel, Jens Eisert, Vittorio Giovannetti, Ludovico Lami, Lorenzo Leone, and Salvatore F. E. Oliviero. Learning quantum states of continuous-variable systems. *Nature Physics*, 21(12):2002–2008, 2025. doi: 10.1038/s41567-025-03086-2.

[5] Lennart Bittel, Antonio Anna Mele, Jens Eisert, and Lorenzo Leone. Optimal trace-distance bounds for free-fermionic states: Testing and improved tomography. *PRX Quantum*, 6: 030341, 2025. doi: 10.1103/PRXQuantum.6.030341.

[6] Senrui Chen, Francesco Anna Mele, Marco Fanizza, Alfred Li, Zachary Mann, Hsin-Yuan Huang, Yanbei Chen, and John Preskill. Towards sample-optimal learning of bosonic Gaussian quantum states. *arXiv preprint arXiv:2603.18136*, 2026.

[7] Lennart Bittel, Francesco Anna Mele, Antonio Anna Mele, Salvatore Tirone, and Ludovico Lami. Optimal estimates of trace distance between bosonic Gaussian states and applications to learning. *Quantum*, 9:1769, 2025. doi: 10.22331/q-2025-06-12-1769.

[8] Lennart Bittel, Francesco A. Mele, Jens Eisert, and Antonio A. Mele. Energy-independent tomography of Gaussian states. *arXiv preprint arXiv:2508.14979*, 2025.

[9] Kianna Wan, William J. Huggins, Joonho Lee, and Ryan Babbush. Matchgate shadows for fermionic quantum simulation. *Communications in Mathematical Physics*, 404(2): 629–700, October 2023. ISSN 1432-0916. doi: 10.1007/s00220-023-04844-0. URL http://dx.doi.org/10.1007/s00220-023-04844-0.

[10] Michael Walter and Freek Witteveen. A random purification channel for arbitrary symmetries with applications to fermions and bosons, 2025. URL https://arxiv.org/abs/2512.15690.

[11] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020. ISSN 1745-2481. doi: 10.1038/s41567-020-0932-7. URL http://dx.doi.org/10.1038/s41567-020-0932-7.

[12] Antonio Anna Mele and Yaroslav Herasymenko. Efficient learning of quantum states prepared with few fermionic non-gaussian gates. *PRX Quantum*, 6(1), January 2025. ISSN 2691-3399. doi: 10.1103/prxquantum.6.010319. URL http://dx.doi.org/10.1103/PRXQuantum.6.010319.

[13] Lorenzo Leone, Salvatore F. E. Oliviero, and Alioscia Hamma. Learning t-doped stabilizer states. *Quantum*, 8:1361, May 2024. ISSN 2521-327X. doi: 10.22331/q-2024-05-27-1361. URL http://dx.doi.org/10.22331/q-2024-05-27-1361.

[14] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. Efficient learning of quantum states prepared with few non-clifford gates. *Quantum*, 9:1907, November 2025. ISSN 2521-327X. doi: 10.22331/q-2025-11-06-1907. URL http://dx.doi.org/10.22331/q-2025-11-06-1907.

[15] Dominik Hangleiter and Michael J. Gullans. Bell sampling from quantum circuits. *Physical Review Letters*, 133(2), 2024. ISSN 1079-7114. doi: 10.1103/physrevlett.133.020601. URL http://dx.doi.org/10.1103/PhysRevLett.133.020601.