

# Chasing shadows with Gottesman–Kitaev–Preskill codes

Jonathan Conrad<sup>1,2</sup>, Jens Eisert<sup>2,3</sup>, and Steven T. Flammia<sup>4,5</sup>

<sup>1</sup>Institute of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne CH-1015, Switzerland

<sup>2</sup>Dahlem Center for Complex Quantum Systems, Physics Department, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany

<sup>3</sup>Helmholtz-Zentrum Berlin für Materialien und Energie, Hahn-Meitner-Platz 1, 14109 Berlin, Germany

<sup>4</sup>Department of Computer Science, Virginia Tech, Alexandria, USA

<sup>5</sup>Phasecraft Inc., Washington DC, USA

13/01/2026

We consider the task of performing shadow tomography of a logical subsystem defined via the Gottesman–Kitaev–Preskill (GKP) error correcting code. Our protocol does not require the input state to be a code state but is implemented by appropriate twirling of the measurement channel, such that the encoded logical tomographic information becomes encoded in the classical shadow. We showcase this protocol for measurements natural in continuous variable (CV) quantum computing. For heterodyne measurement, the protocol yields a probabilistic decomposition of any input state into Gaussian states that simulate the encoded logical information of the input relative to a fixed GKP code where we prove bounds on the Gaussian compressibility of states in this setting. For photon parity measurements, our protocol is equivalent to a Wigner sampling protocol for which we develop the appropriate sampling strategies. Finally, by randomizing over the reference GKP code, we show how Wigner samples of any input state relative to a random GKP codes can be used to estimate any sufficiently bounded observable.

## 1 Introduction

Recent years have seen steady progress in experimental realizations of quantum error correction. On the one hand, experiments towards qubit-based quantum error correction have demonstrated impressive control of large quantum systems consisting of hundreds of physical qubits to encode and process encoded logical information [1], as well the ability to encode quantum information into a single quantum harmonic oscillator beyond break-even via the *Gottesman–Kitaev–Preskill* (GKP) code from the realm of bosonic quantum error correction [2–4]. On the other hand, quantum experiments are being conducted to understand practical capabilities of present noisy quantum devices, e.g., through variational quantum algorithms [5], or to benchmark the readily accessible “quantumness” [6] through randomized sampling experiments. Aside from experimental progress, the design of error mitigation and smart post-processing techniques yields valuable insights into the design of

Jonathan Conrad: [jonathan.conrad@epfl.ch](mailto:jonathan.conrad@epfl.ch)

future experiments and has specifically developed into a quest for learning properties of quantum states from randomly accessible snapshots.

Any such effort makes sense only, however, if the anticipated state preparations or protocols are being implemented with high levels of accuracy. To ensure this, one usually has to resort to techniques of benchmarking, certification or tomographic recovery [7]. An interesting technique in this realm is the so-called classical shadow tomography protocol [8–10] that demonstrates just how little classical information needs to be extracted from any quantum state to reproduce the expectation values of a bounded number of suitably bounded observables. This family of protocols reconstructs a description of the input state as perceived by a random observable. The so obtained “average-case description” of the state becomes appropriate for many observables, so that the choice on what observables to focus on can be made later.

On the technical level, the (classical) shadow tomography protocol combines two ingredients: the fact that a channel-twirl of a *positive operator valued measure* (POVM) capturing a generalized measurement projects it onto a channel with the fixed structure of a depolarizing channel and the existence of strong statistical anticoncentration bounds for medians of means estimation. The protocol proceeds by implementing a Clifford channel-twirl of a POVM that outputs samples over the reconstructed pointer states – which are stabilizer states – such that expectation values of observables over such samples match the expectation value of the depolarized input state and can be classically processed to yield the targeted expectation value. We will refer to such a group twirl involving a POVM a POVM twirl. The classical post-processing necessary is informed by the structure of the depolarizing channel that resulted from the projection and is efficiently possible due to the stabilizer state structure of the samples. Medians of means estimation then yields a process to combine samples and prove bounds on the necessary sample complexity. It has been recognized in refs. [11, 12] that the projective nature of the channel twirl allows one to render the protocol robust to errors in the POVM, as any noisy version of the POVM would simply be projected onto depolarizing channel with amended parameters, which can be accounted for in post-processing. In fact, the projected POVM is simply one instance of a very well-structured informationally complete (generalized) POVM [13–15] can be used to implement the protocol.

Bosonic quantum systems also offer a wide variety of POVMs beyond those exactly expressible in finite systems, such as heterodyne, homodyne, photon-counting, and photon parity measurements. Despite the richness of these POVMs, they can be similarly tamed into an effective channel with a simple structure by appropriate twirling. Recent work [16, 17] has shown that twirling techniques are applicable for continuous variable quantum systems by only focusing only on certain energy-constrained subspaces. Prior work has treated this constraint as necessary due to the inaccessibility of random operations on this infinite Hilbert space.

In this work, we develop shadow tomography protocols by focusing on logical subsystems prescribed by the GKP code. This also yields an effective finite subspace of the infinite dimensional CV Hilbert space and we show how effective shadow tomography protocols can be derived that reproduce logical expectation values of operators relative to the chosen GKP-codes. On the technical level, this is executed by twirling a CV-POVM over a set of random *logical* Clifford gates, which has the effect that the logical action of the POVM becomes projected onto a depolarizing channel. This effect is revealed when the pointer states output by the protocol are evaluated in accordance to a decoder associated to the code. This structure reveals an interesting interplay between the physical structure of the system and its logical content. We identify different applications of the GKP-shadow

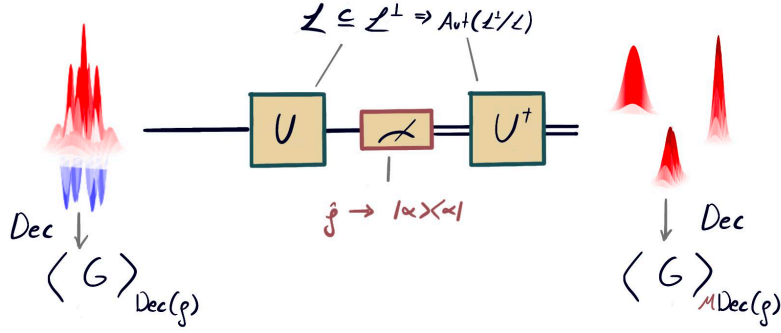


Figure 1: Illustration of Gaussian decomposition of arbitrary CV states via twirled heterodyne measurements. Relative to a GKP code described by a lattice  $\mathcal{L} \subseteq \mathcal{L}^\perp$  is shadow tomography protocol yields a probabilistic decomposition of the input state into Gaussian states that reproduce logical expectation values up to a logical depolarization  $\mathcal{M}$ .

tomography toolbox developed here by considering different choices of bosonic POVMs and finally show how a general shadow protocol for CV states can be obtained by combining our GKP-shadow tomography tools with a random choice of GKP codes.

For example, when the protocol is executed using heterodyne measurement as POVM, the fact that GKP Clifford gates are represented by Gaussian unitary operations implies that the protocol outputs an ensemble of *Gaussian states*, which contain the same logical information (relative to the chosen GKP code) as the given input state. Here the key upshot of our combination of techniques is that the statistical methods used in classical shadow tomography allow for the derivation of rigorous bounds on the number of such Gaussian states needed to faithfully retain the logical content of the input. While the bounds we derive scale exponentially in the system size, the key point of this result is that they are obtained *without* knowledge of an analytical expression for the input state or strong assumptions of its physical structure. In the context of engineering GKP states for quantum computation this is particularly relevant, since many different analytical approximations to GKP states are used throughout the literature [18] and its specific physical structure finally depends on varying engineering details [2, 4].

This protocol yields an experimental *black-box* procedure to convert an arbitrary physical input state into a convex combination of Gaussian states. As Gaussian states are easy to simulate classically [19], we expect this technique to be of value in assessing the performance of quantum computation and error correction using real GKP states. An illustration of this process is provided in fig. 1. In the concrete application with photon parity detectors we consider, our protocol descends to one that samples the Wigner function of a given quantum state at random points according to a well-tailored distribution and we show how the toolbox we have developed allows one to derive sample complexity bounds for this protocol to estimate a number of arbitrary observables within certain bounds, unconditional on properties of the input state. This is achieved by combining the logical GKP shadow tomography protocol with a random choice of GKP code. Here the core technical ingredient is the existence of a Haar measure over the space of symplectic lattices and simple expressions for averages of functions of lattices over this measure. At the bottom, this is a randomized protocol to sample the Wigner function of an arbitrary CV state where the intricate way in which we choose where to sample allows us to rigorously bound the required sample complexity to estimate given CV observables to high confidence. By averaging over GKP codes, this protocol effectively interpolates between tomography of the *logic* encoded in a physical system, and its *physics*.

The key point of this work, however, is to highlight the intersection between continuous variable physics and randomized tomographic methods originally derived for discrete variable systems. As becomes apparent in the course of our presentation, thinking about classical shadow tomography through the lens of GKP codes helps to refine our general understanding of the nuances of the classical shadow protocol while, vice versa, we obtain experimental handles to learn relevant aspects of a physical CV state using methods from random coding theory. Next to the presentation of the concrete results our purpose here is hence a pedagogical one: we hope that, through the lens of GKP codes, our explorations help the curious reader to develop a more refined understanding of the interesting intersection between physics, logic and everything in between.

## 1.1 Overview

This article is structured as follows. We begin by a review of the basics of quantum harmonic oscillators and the structure of GKP codes in sec. 2. In sec. 3, we give a broad overview on twirling and discuss its various incarnations in state purification protocols, dynamical decoupling, noise mitigation and shadow tomography. This section is meant to provide a pedagogical ground-up introduction to the utility of twirling and the role of random operations for quantum experiments. We discuss extensions of these tools to the realm of continuous variable systems and explain how the infinitude of the associated groups of Gaussian unitaries can be suitably regularized. Finally, in sec. 4 we apply the developed tools to design and prove bounds for logical shadow tomography protocols relative to GKP codes, where we also examine the behavior of GKP shadow tomography with random GKP codes in sec. 5. We show this yields a protocol that allows one to estimate *arbitrary* CV observables in a regularized manner. The core contribution of this section is the introduction of new techniques for estimating and bounding the performance of full tomography of a CV quantum system. We close with a brief discussion and open questions for future work. For a better flow of presentation, the detailed proofs of most statements made throughout this manuscript are found in the appendix.

## 2 Preliminaries

In this section we review basic knowledge on continuous variable quantum physics and quantum error correction with the GKP code relevant to understand this manuscript. The following material is brief and introductory, and we refer the reader to the established literature, see, e.g., refs. [19–21] and references therein, for further details.

### 2.1 Quantum harmonic oscillators

Bosonic quantum error correction studies the robust embedding of discrete quantum information into a system of multiple *quantum harmonic oscillators* (QHO), each of which can be described by an infinite dimensional Hilbert space  $\mathcal{H} = \text{span} \{|n\rangle\}_{n=0}^{\infty}$  where  $|n\rangle$  denote the well known Fock state vectors whose labels correspond to the *eigenvalues* of the number operator  $\hat{n} = \hat{a}^\dagger \hat{a}$  and  $\hat{a} = (\hat{q} + i\hat{p})/\sqrt{2}$  denotes the annihilation operator.  $\hat{q}$  and  $\hat{p}$  are the position and momentum operators, the canonical coordinates, whose improper eigenstates yield a basis for the underlying Hilbert space. The associated phase space inherits a non-trivial geometry from the canonical commutation relations (we will set  $\hbar = 1$  throughout), and is most naturally studied in the Heisenberg frame, i.e., in terms of the transformation behaviour of operators on this space. On a system of  $n$  QHOs

– which we will refer to as having  $n$  modes – we define a generalized quadrature operator  $\hat{\mathbf{x}} = (\hat{q}_1, \hat{q}_2, \dots, \hat{p}_{n-1}, \hat{p}_n)^T$  such that the canonical commutation relations are captured by the anti-symmetric symplectic form

$$J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \quad (1)$$

where  $I_n$  denotes the  $n \times n$  identity matrix. Unless explicitly needed, we will omit the index  $n$  from the symplectic form and simply denote it by  $J$ .

Analogous to the Pauli-operators for qubit-systems, the Heisenberg-Weyl operators for this infinite dimensional Hilbert space are given by displacement operators

$$D(\boldsymbol{\xi}) = \exp \left\{ -i\sqrt{2\pi}\boldsymbol{\xi}^T J \hat{\mathbf{x}} \right\} \quad (2)$$

for  $\boldsymbol{\xi} \in \mathbb{R}^{2n}$  being elements of phase space. These displacement operators satisfy the *Weyl relations*

$$D(\boldsymbol{\xi}) D(\boldsymbol{\eta}) = e^{-i\pi\boldsymbol{\xi}^T J \boldsymbol{\eta}} D(\boldsymbol{\xi} + \boldsymbol{\eta}) \quad (3)$$

for  $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^{2n}$ .<sup>1</sup> They form a basis for operators and are Hilbert-Schmidt orthogonal as  $\text{Tr} [D^\dagger(\boldsymbol{\xi}) D(\boldsymbol{\eta})] = \delta^{(2n)}(\boldsymbol{\xi} - \boldsymbol{\eta})$ , such that states can be represented by their Wigner function

$$W_\rho(\mathbf{x}) = \int_{\mathbb{R}^{2n}} d\boldsymbol{\eta} e^{-i2\pi\mathbf{x}^T J \boldsymbol{\eta}} \text{Tr} [D(\boldsymbol{\eta}) \rho]. \quad (4)$$

Displacement operators represent the unitary time evolution induced by Hamiltonians linear in the quadrature operators that implement the transformation

$$D(\boldsymbol{\xi})^\dagger \hat{\mathbf{x}} D(\boldsymbol{\xi}) = \hat{\mathbf{x}} + \sqrt{2\pi}\boldsymbol{\xi} \quad (5)$$

and commute and multiply as

$$\begin{aligned} D(\boldsymbol{\xi}) D(\boldsymbol{\eta}) &= e^{-i\pi\boldsymbol{\xi}^T J \boldsymbol{\eta}} D(\boldsymbol{\xi} + \boldsymbol{\eta}), \\ &= e^{-i2\pi\boldsymbol{\xi}^T J \boldsymbol{\eta}} D(\boldsymbol{\eta}) D(\boldsymbol{\xi}). \end{aligned} \quad (6)$$

It is these properties that make them a natural set to choose stabilizer groups from.

Unitary evolution via Hamiltonians strictly quadratic in the quadrature operators, also termed *Gaussian* unitary transformations [19, 22], implement symplectic transformations

$$U = e^{-\frac{i}{2}\hat{\mathbf{x}}^T C \hat{\mathbf{x}}}, \quad C = C^T, \quad (7)$$

$$U^\dagger \hat{\mathbf{x}} U = S \hat{\mathbf{x}}, \quad S = e^{CJ}, \quad (8)$$

where  $S \in \text{Sp}_{2n}(\mathbb{R}) = \{S \in \mathbb{R}^{2n \times 2n} : S^T J S = J\}$  is a real symplectic matrix which follows from unitarity of  $U$  and we have

$$D(\boldsymbol{\xi}) U_S = U_S D(S^{-1}\boldsymbol{\xi}), \quad (9)$$

such that it also holds that

$$W_{U_S \rho U_S^\dagger}(\mathbf{x}) = W_\rho(S\mathbf{x}). \quad (10)$$

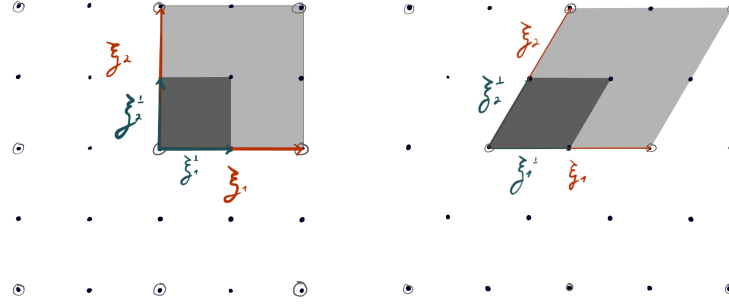


Figure 2: The square  $\mathbb{Z}^2$  (l.) and hexagonal  $A_2$  (r.) GKP codes each encoding a qubit. The logical displacement amplitudes are marked in turquoise and stabilizer displacements are marked in red.

## 2.2 GKP codes and their Cliffords

The GKP code [3, 23] – or rather the family of it – is a quantum error correcting code defined to embed discrete quantum information into a system of  $n$  quantum harmonic oscillators by identifying a code space symmetric under the stabilizer group

$$\mathcal{S} = \langle D(\xi_1) \dots D(\xi_{2n}) \rangle = \left\{ e^{i\phi_M(\xi)} D(\xi), \xi \in \mathcal{L} \right\}, \quad (11)$$

which is isomorphic to a full rank lattice  $\mathcal{L} = \mathbb{Z}^{2n} M$  with generator matrix

$$M = (\xi_1, \xi_2, \dots, \xi_{2n})^T. \quad (12)$$

The symplectic dual lattice

$$\mathcal{L}^\perp = \left\{ \mathbf{x} \in \mathbb{R}^{2n} : \mathbf{x}^T J \xi \in \mathbb{Z} \forall \xi \in \mathcal{L} \right\} \quad (13)$$

labels the centralizer of the GKP stabilizer group, such that the GKP stabilizer group is abelian if and only if it is isomorphic to a *weakly symplectically self-dual* lattice

$$\mathcal{L} \subseteq \mathcal{L}^\perp \Leftrightarrow M = AM^\perp, \quad (14)$$

where the right hand side describes the sublattice structure by identifying how basis vectors of  $\mathcal{L}$  are described by (integer) linear combinations of basis vectors of  $\mathcal{L}^\perp$  as given by the the symplectic Gram matrix  $A = MJM^T$  when the dual basis is chosen via some canonical choice [3]. The phases  $\Phi_M(\xi) = \pi \mathbf{a}^T A \mathbf{a}$ ,  $\mathbf{a} = M^{-T} \xi$  in eq. (11) are determined by the basis in which the stabilizer generators are fixed to a +1 eigenvalue and are trivial when the symplectic Gram matrix  $A$  is even [23].

A special class of GKP codes, called *scaled* GKP codes, is obtained from rescaling a symplectic self dual lattice  $\mathcal{L}_0 = \mathcal{L}_0^\perp \mapsto \mathcal{L} = \sqrt{d} \mathcal{L}_0 : \mathcal{L} \subseteq \mathcal{L}^\perp$ , via the square root of the desired local dimension  $d \in \mathbb{N}$  and gives rise to the well-known GKP codes that encode a qubit ( $d = 2$ ) into a single oscillator via the square- or the hexagonal lattice with bases

$$M_{\mathbb{Z}^2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad M_{A_2} = \frac{1}{\sqrt{2\sqrt{3}}} \begin{pmatrix} 2 & 0 \\ 1 & \sqrt{3} \end{pmatrix}. \quad (15)$$

<sup>1</sup>The Weyl relations are actually a way of rigorously capturing the canonical commutation relations without having to resort to unbounded operators.

These GKP codes have been widely explored in the literature: They afford a distance (given by the length of the shortest non-trivial logical displacement)  $\Delta(\sqrt{2}\mathbb{Z}^2) = 2^{-\frac{1}{2}}$  and  $\Delta(\sqrt{2}A_2) = 3^{-\frac{1}{4}}$ . We depict their structure in fig. 2, where it can also be seen that the lattices are respectively symmetric under  $\pi/2$  and  $\pi/3$  rotations  $R_{\pi/2}$  and  $R_{\pi/3}$  which correspond to the logical Hadamard  $\hat{H}$  gate for the square GKP code and the Hadamard-phase gate  $\hat{H}\hat{S}^\dagger$  for the hexagonal GKP code.

The identification as logical Clifford gates is made through their property as symplectic lattice automorphisms, of which the general structure has been explored in refs. [24–27]. Another such symplectic automorphism is given by the transvection  $S = I + \mathbf{e}_1\mathbf{e}_2^T$ , where  $\mathbf{e}_i$  are unit vectors, that yields a logical phase gate  $\hat{S}$  for the square GKP code. In general, for scaled GKP codes all symplectic automorphisms are given by the symplectic matrices  $S = M^T U^T M^{-T}$ , where  $M$  is the generator for the lattice basis and  $U \in \text{Sp}_{2n}(\mathbb{Z}_d)$  labels the logical action of the corresponding non-trivial Clifford gate. We refer to ref. [24] for an in-depth discussion. In the appendix, we show that any such  $U \in \text{Sp}_{2n}(\mathbb{Z}_d)$  (and consequently, by conjugation with  $M^T$ , any corresponding real symplectic automorphism  $S$ ) can be generated by a length  $O(dn^2)$  sequence of elementary local matrices in  $\text{Sp}_{2n}(\mathbb{Z}_d)$  that correspond to the qudit versions of the usual Hadamard, phase and CNOT gates. Throughout this manuscript, by (logical) Clifford group we mean the group of automorphisms of logical Pauli operators under conjugation while the prefix *trivial* refers to those elements that implement a projectively trivial automorphism. We refer to ref. [24, 25] for a more in-depth discussion of these groups and their connection to the respective groups of lattice automorphisms.

### 3 Twirling theory

We now turn to discussing constructions and applications of *random* logical GKP Clifford gates. While random Clifford gates is a widely and well studied topic for qubit-based systems, for the GKP code the question of how to define a measure over the – now infinite – trivial- and non-trivial group becomes more nuanced. In this context, as a versatile tool, twirls become important as averages over groups. They here appear both as twirls on the logical- and on the physical level and will be the key technical ingredient in the following developments.

Constructions of random (trivial or non-trivial) Clifford gates find widespread applications from state preparation to error mitigation to benchmarking and most notably in many recent works on shadow tomography. The common ground of these applications is that the random Clifford gates are used to implement various incarnations of group projectors – also phrased *twirling* – which we summarize in fig. 3. As exploited in ref. [28] for the square GKP code to Floquet-engineer a GKP Hamiltonian, once a suitable measure for one type of twirl is found, it is easy to translate it into the different incarnations. In the following we discuss the twirling theory for states, channels and POVMs. Although listed in fig. 3, we do not discuss the Hamiltonian twirl but merely mention its equivalence to the state twirl and refer to the relevant literature, see refs. [28, 29].

#### 3.1 Twirling states, channels and POVMs

In this section, we first review how the different incarnations of twirling work on a qubit level before discussing its logical analogue using the GKP code.



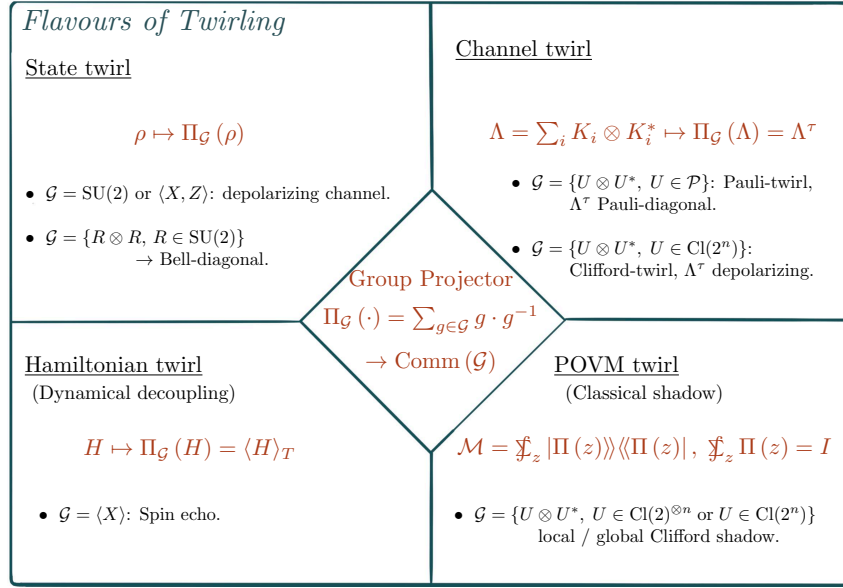


Figure 3: Different notions of twirling discussed in the main text.

**State twirling.** On the (logical) qubit level, a state twirl over a group  $\mathcal{G} \subseteq \mathcal{U}(N)$  maps

$$\rho \mapsto \Pi_{\mathcal{G}}(\rho) = \int_{\mathcal{G}} d\mu^{\mathcal{G}}(U) U \rho U^\dagger, \quad (16)$$

where the integration is taken over the group of all unitaries and we specify the uniform measure  $\mu^{\mathcal{G}}(U)$  over unitaries in  $\mathcal{G}$  (and zero outside). It can be verified that the corresponding group projector indeed is a projector  $\Pi_{\mathcal{G}}^2 = \Pi_{\mathcal{G}}$  and that it maps onto the *commutant*  $\{U \in \mathcal{U}(N) : [U, g] = 0 \forall g \in \mathcal{G}\}$  which is spanned by the *irreducible representations* (irreps) of  $\mathcal{G}$ . Note that in eq. (16) phases attached to each group element  $g$  do not matter and we in fact only need that the operators used in (16) form a projective unitary representation of the group. The group projector associated to the Pauli group  $\mathcal{P} = \langle \hat{X}, \hat{Z} \rangle$  can hence be written as

$$\Pi_{\mathcal{P}} = \Pi_{\langle \hat{X} \rangle} \circ \Pi_{\langle \hat{Z} \rangle} \quad (17)$$

and projects any input state onto a state that is invariant under both random bit- and phase-flips. The only state satisfying this property is completely mixed  $\rho = I/2$ , aligned with the fact that the Pauli group forms a unitary 1-design – i.e., group twirls involving single powers  $g, g^\dagger$  as in eq. (16) act on a state in the same way as a twirl over the whole unitary group  $\mathcal{G} = \mathcal{U}(N)$  which has only the identity  $I$  in its commutant.

The same idea also works for subgroups, giving rise to twirls that project onto suitable symmetric subspaces. For example, we can consider state twirls for a group  $\mathcal{G} \subseteq \mathcal{U}(N)$  that acts as

$$\rho \mapsto \Pi_{\mathcal{G}}(\rho) = \int_{\mathcal{G}} d\mu^{\mathcal{G}}(U) (U \otimes U) \rho (U \otimes U)^\dagger, \quad (18)$$

where the integration is taken over the Haar measure over  $\mathcal{G}$ . This is a group projector onto the commutant  $\{O \in L(\mathcal{H} \otimes \mathcal{H}) : [O, g \otimes g] = 0 \forall g \in \mathcal{G}\}$ , where  $L(\mathcal{H} \otimes \mathcal{H})$  denotes the space of all linear operators on  $\mathcal{H} \otimes \mathcal{H}$ . This form of the twirl has been applied in protocols to distill entangled states, where the random application of correlated Pauli operators ( $\mathcal{G} = \mathcal{P}$ ) yields entangled (Werner) states [30, 31].



**Channel twirling.** In a similar way as one can think of state twirls, it makes sense to consider channel twirls. A channel twirl  $\tau_{\mathcal{G}}$  maps a quantum channel, written in its  $\chi$ -matrix representation with  $P_{\alpha}$  a Pauli operator with index  $\alpha$

$$\mathcal{C}(\cdot) = \sum_{\alpha, \beta} \chi_{\alpha, \beta} P_{\alpha} \cdot P_{\beta} \quad (19)$$

onto a “symmetrized” version of itself. Specifically, it is defined as the map

$$\tau_{\mathcal{G}} \circ \mathcal{C} = \int_{\mathcal{U}(N)} d\mu(U) U \mathcal{C}(U^{\dagger} \cdot U) U^{\dagger}. \quad (20)$$

In vectorized notation where  $\hat{\mathcal{C}}$  acts on state vectors  $|\rho\rangle\rangle \mapsto \hat{\mathcal{C}}|\rho\rangle\rangle$ , this is recast in the recognizable form

$$\tau_{\mathcal{G}} \circ \hat{\mathcal{C}} = \int_{\mathcal{U}(N)} d\mu(U) \mathcal{U} \hat{\mathcal{C}} \mathcal{U}^{\dagger}, \quad \mathcal{U} := U \otimes U^* \quad (21)$$

and again for the uniform measure  $\mu^{\mathcal{G}}$  on  $\mathcal{G}$  we can recognize that the channel is projected onto the commutant of the group representation  $R(\mathcal{G} \otimes \mathcal{G}^*) = \{R(g) \otimes R(g)^* \mid g \in \mathcal{G}\}$ .

A Clifford twirl on a channel can be understood as a combination of twirls

$$\Pi_{\text{Cl}(N)} = \Pi_{\text{Cl}(N)/P(N)} \circ \Pi_{P(N)} \quad (22)$$

over the trivial- and non-trivial Clifford groups [32, 33], where the channel twirl enforces that every pair of Pauli operators  $P_{\alpha} \otimes P_{\beta}^*$  commute with every operator of the form  $P \otimes P^*$ ,  $P \in \mathcal{P}(N)$ . Since Pauli operators only commute up to phases, this twirl effectively enforces that the channel only has “diagonal” Pauli elements for which  $\alpha = \beta$  in eq. (19). Subsequently, the twirl over the non-trivial Clifford elements enforces permutation symmetry on the non-trivial Pauli elements indexed in  $\chi_{\alpha, \alpha}$ ,  $\alpha \neq 0$ . For a single qubit, the desired measure over non-trivial Cliffords hence should satisfy (up to phases)

$$\mu_{\text{Cl}(2)/\mathcal{P}}(I \neq P \in \mathcal{P} \mapsto X, Y, Z) = \frac{1}{3}, \quad (23)$$

and by assuming the input channel to be trace preserving, the twirl produces a channel

$$\Pi_{\text{Cl}(N)} \circ \mathcal{C}(\rho) = \chi_{0,0} \rho + \frac{1 - \chi_{0,0}}{3} \sum_{\alpha > 0} P_{\alpha} \rho P_{\alpha}. \quad (24)$$

Writing

$$\sum_{\alpha > 0} P_{\alpha} \rho P_{\alpha} = \sum_{\alpha} P_{\alpha} \rho P_{\alpha} - \rho = 4\Pi_{\mathcal{P}}(\rho) - \rho, \quad (25)$$

we obtain, knowing that the Pauli state-twirl produces a completely mixed state  $\Pi_{\mathcal{P}}(\rho) = 2^{-1}I$ ,

$$\Pi_{\text{Cl}(N)} \circ \mathcal{C}(\rho) = \frac{4\chi_{0,0} - 1}{3} \rho + \frac{2(1 - \chi_{0,0})}{3} I. \quad (26)$$

A channel twirl over the Clifford group  $\Pi_{\text{Cl}(2)}$  that produces an output of such a form – i.e., it is in vectorized notation a linear combination of the trivial channel  $I \otimes I^*$  and the fully depolarizing channel  $|I\rangle\rangle\langle\langle I|$  – is also called a Clifford 2-design [34]. There are several ways to implement a twirl over either the full or non-trivial Clifford group. One straightforward way to define a non-trivial Clifford twirl is to pick one of the 6 elements in  $\text{Sp}_2(2)$  at random (corresponding to the matrices with 1 entries on the diagonal and anti-diagonal and one of the four choices of  $\mathbb{Z}_2^{2 \times 2}$  matrices with one 0 entry). While this

strategy is the simplest, it requires to exhaustively enumerate all elements of the non-trivial Clifford group does not generalize easily to a multi-qubit setting where it becomes desirable to approximate such twirl via a random walk over  $\text{Sp}_{2n}(\mathbb{Z})$  or, more commonly, to directly sample random Clifford circuits (involving Pauli gates) gate-by-gate to set up a good approximation to a (Haar-)random Clifford circuit. On a single qubit, the first strategy effectively samples from one of the 24 elements in the set

$$S_0 = \{CP, C \in \text{Cl}(2)/\mathcal{P}, P \in \mathcal{P}\} \subseteq \mathbb{C}^{2 \times 2} \quad (27)$$

which, in fact, turns out to be larger than necessary. From the discussion above we have learned that the main function of the non-trivial Pauli twirl is to set up a random permutation of the Pauli operators. This function is already fulfilled by the 3 element cyclic subgroup  $\langle \hat{H}\hat{S}^\dagger \rangle$ , so that only a total of 12 elements are necessary to consider to build a single qubit Clifford twirl. In fact, it is easy to verify that the set

$$S = \langle \hat{H}\hat{S}^\dagger \rangle \mathcal{P} = \{CP \mid C \in \langle \hat{H}\hat{S}^\dagger \rangle, P \in \mathcal{P}\} \quad (28)$$

is a Clifford 2-design. This is verified either by checking that it correctly twirls a Pauli-diagonal channel into a depolarizing channel from its transitivity over the Pauli operators or by verifying an equivalent condition proven in ref. [34], which is that the *frame potential* evaluates to a value of  $\mathcal{F} = 2$ , proving that one indeed encounters an exact unitary 2-design. In fact, the group stated above has the minimal cardinality of  $(d^2 - 1)d^2$  shown to yield a Clifford 2-design [34, 35], which, in our construction, comes from the fact that  $\langle \hat{H}\hat{S}^\dagger \rangle$  is the minimal Pauli transitive subgroup. For any dimension, a Pauli transitive subgroup has at least  $d^2 - 1$  elements, such that a Clifford design of the type proposed above has at least  $(d^2 - 1)d^2$  elements, matching the bound conjectured ref. [34] and later proven to hold in ref. [35].

**POVM twirling.** A special application of the channel twirl – and at the same time a most important one for the purposes of this work – is found when it is applied to a POVM reflecting a generalized measurement. This setting has been recently popularized by showing how it can be used for *shadow tomography* [8, 9, 11, 36] giving rise to various *noisy-intermediate-scale-quantum* (NISQ) friendly applications. On a single mode consider the POVM representing a computational basis measurement

$$\mathcal{M}_Z = \sum_{z \in \mathbb{Z}_2} |\Pi_z\rangle\rangle \langle\langle \Pi_z|, \Pi_z = |z\rangle \langle z|. \quad (29)$$

The corresponding  $\chi$  matrix representation has  $\chi_{0,0} = 1/2$ , such that we can compute

$$\mathcal{M} := \Pi_{\text{Cl}(2)} \circ \mathcal{M}_Z = (\rho + I)/3. \quad (30)$$

The twirled POVM is a depolarizing channel and is (as a matrix, not physically as a quantum channel) invertible with  $\mathcal{M}^{-1}(X) = 3X - I$ . Furthermore, its action on state vectors  $|\rho\rangle\rangle$

$$\mathcal{M}|\rho\rangle\rangle = \frac{1}{|\text{Cl}(2)|} \sum_{C \in \text{Cl}(2)} C |\Pi_z\rangle\rangle \langle\langle \Pi_z| C^\dagger |\rho\rangle\rangle \quad (31)$$

can be interpreted as protocol to decompose arbitrary quantum state vectors  $|\rho\rangle\rangle$  into stabilizer state vectors  $C|\Pi_z\rangle\rangle$ . Note that, as we have done earlier as well, we use the calligraphic symbols for unitary channels in place of  $\mathcal{C} = C \otimes C^*$ . An experimental protocol to reconstruct (properties of) the state is hence identified by realizing that measuring in

the computational basis after applying random Clifford gates to an input state allows for a reconstruction

$$|\rho\rangle\rangle = \mathcal{M}^{-1} \mathbb{E} [C_i |\Pi_{z_i}\rangle\rangle], \quad (32)$$

where the Cliffords  $C_i$  are picked uniformly from the Clifford group and  $z$  is determined by the Born rule  $z_i \sim \langle\langle \Pi_z | \mathcal{C}^\dagger | \rho \rangle\rangle$ . Similarly, expectation values of observables can also be estimated as

$$\langle\langle O | \rho \rangle\rangle = \langle\langle \mathcal{M}^{-1}(O) | \mathbb{E} [C_i |\Pi_{z_i}\rangle\rangle], \quad (33)$$

where we also used that  $\mathcal{M}^{-1}$  is self-adjoint. This protocol is particularly appealing in practical NISQ-era questions for two reasons. 1. the projective nature of the channel twirl projects noisy versions of the measurement  $\mathcal{M}_Z$  onto  $\mathcal{M}$  (possibly with adapted coefficients [11]) and 2. a relatively small numbers of samples from the distribution over  $(C, z)$  for a given  $\rho$  allows to estimate expectation values of exponentially many well-bounded observables to high confidence [9]. Here, the samples  $(C, z)$  can be generated in a quantum experiment without yet having decided on the observable  $O$ . Due to the design property of the twirl the partial tomographic data obtained this way suffices to estimate selected observables in purely classical post-processing. We refer to the statistical ensemble of samples  $(C, z)$  colloquially as the *shadow* of the state.

For applications in quantum computation with bosonic quantum error correction, such as using the GKP code, the shadow tomography protocol is particularly interesting since, typically, measurements of logical observables can only be carried out indirectly using more naturally accessible measurements such as homodyne detection, heterodyne detection, or photon counting and do not admit a simple and direct physical measurement procedure. The projection property (1.) of the logical Clifford-twirl allows us to naturally use more accessible measurements that may be badly aligned with the observables of interest and prescribe how to classical post-process the results to estimate the logical observable at hand. Another important implication of this property is that classical shadow protocols designed this way are inherently robust to imperfections in the POVM, as the imperfections are projected together with the POVM onto the effective channel and only a potential change of the effective channel parameters needs to be accounted for. This perspective was leveraged in ref. [11] already for discrete variable shadow tomography. Furthermore, the effective decomposition of the state (32) obtained from the shadow protocol is interesting as it may give rise to new representations of states in phase space that capture core logical information. After discussing how a Clifford channel-twirl can be set up for the GKP code, we will see how the shadow protocol gives rise to an approximation of GKP states using Gaussian states by logical Clifford twirling a heterodyne measurement. We expect this representation to be particularly useful in the development of new simulation methods for GKP error correction.

### 3.2 Displacement twirling

In this subsection, we turn to discussing twirls that can be realized by implementing appropriate displacements in the physical Hilbert space. This is a core technical ingredient for the following developments and we discuss its the application of displacement twirling to states and channels.

In ref. [28], approximate twirls over groups of displacement operators distributed over lattices  $\mathcal{L}^\perp$  associated to the GKP code have been constructed by approximating the uniform measure over the (infinite) lattice  $\mathcal{L}^\perp$  via a random walk over a generating set given by the rows of  $M^\perp = (\xi_1^{\perp T}, \dots, \xi_{2n}^{\perp T})^T$ . Concretely, we define a random walk from

the joint distribution of  $N' = 2N$  half-steps  $\pm \xi_i/2$ , each of which are selected with  $1/2$  probability at each step. Define for  $i = 1, \dots, 2n$  the associated (discrete) measure

$$\mu'_i(\mathbf{x}) = \frac{1}{2}\delta\left(\mathbf{x} - \xi_i^\perp/2\right) + \frac{1}{2}\delta\left(\mathbf{x} + \xi_i^\perp/2\right), \quad (34)$$

so that we obtain the measure corresponding to  $m$  steps of the random walk as the  $2m$ -fold convolution  $\mu_i^{(*m)} := \mu_i'^{(*2m)}$ .

**State twirling.** Applying this twirl to a state, we obtain that

$$\rho = \int_{\mathbb{R}^2} d\alpha \rho(\alpha) D(\alpha) \mapsto \int_{\mathbb{R}^2} d\mu_i^{(*m)}(\gamma) D(\gamma) \rho D^\dagger(\gamma) = \int_{\mathbb{R}^2} d\alpha [\nu_i(\alpha)]^m \rho(\alpha) D(\alpha) \quad (35)$$

modifies the characteristic function of the state with the  $m$ -th power of the symplectic Fourier transform of the measure

$$\nu_i(\alpha) = \int_{\mathbb{R}^2} d\mu_i(\gamma) e^{-i2\pi\gamma^T J\alpha} = \cos^2\left(\pi\left(\xi_i^\perp\right)^T J\alpha\right). \quad (36)$$

In the limit  $m \rightarrow \infty$ , this suppresses all contributions  $\alpha$  except for those in the symplectic dual of  $\xi_i^\perp$ . We define the joint measure over all generators in  $M^\perp$  to be the joint random walk given by the  $2n$ -fold convolution

$$\mu_{M^\perp} = \mu_1 * \mu_2 * \dots * \mu_{2n} \quad (37)$$

which has the Fourier transform

$$\nu_{M^\perp}(\alpha) = \prod_{i=1}^{2n} \cos^2\left(\pi\left(\xi_i^\perp\right)^T J\alpha\right). \quad (38)$$

The total effect of this twirl is that in the limit  $m \rightarrow \infty$  only logically trivial contributions  $\alpha \in \mathcal{L}$  will survive while all other contributions are exponentially suppressed and the state becomes logically fully depolarized. An alternative view is that in this limit eq. (35) converges to the group projector of the group generated by the displacement in  $\mathcal{L}^\perp$ . The commutant of this group is the stabilizer group generated by displacements in  $\mathcal{L}$ . Non-trivial displacements of  $\mathcal{L}^\perp$  are not in this commutant such that it cannot carry logical information. While twirling a state over  $\mathcal{L}^\perp$  does not appear to bear any interesting applications outside deliberate logical depolarization of the state <sup>2</sup> note that the above outlined method of state twirling is not restricted to using generators in  $M^\perp$ . A stabilizer twirl using generator  $M$  can equally be used to (approximately) project the state onto one one with a characteristic function supported only on  $\mathcal{L}^\perp$ . In conjunction with a twirl over a set of Gaussian unitaries representing a set of logical Clifford gates such as the logical  $\hat{H}$  for the square GKP code or  $\hat{H}\hat{P}^\dagger$ , we expect this technique to be useful for the measurement-less preparation of *magic states* [39], analogous to previous proposal for entanglement distillation [31] as well as and the preparation of entangled GKP states analogous to the procedure in refs. [30, 31].

---

<sup>2</sup>Except perhaps for error mitigation methods where tunable noise is desired as in zero-noise extrapolation [37, 38].

**Channel twirling.** Acting on a quantum channel (see ref. [28])

$$\mathcal{C} = \int_{\mathbb{R}^{2n}} d\alpha \int_{\mathbb{R}^{2n}} d\beta c(\alpha, \beta) D(\alpha) \otimes D^*(\beta) \quad (39)$$

with chi-function  $c(\alpha, \beta)$ , the  $m$ -fold displacement channel twirl using our measure  $\mu_M^\perp$  implements the action on the chi-function

$$c(\alpha, \beta) \mapsto [\nu_{M^\perp}(\alpha - \beta)]^m c(\alpha, \beta). \quad (40)$$

Similar to the above, this channel twirl approximately projects the channel onto a channel where non-stabilizer coherences are suppressed, i.e., contributions in the chi-function  $c(\alpha, \beta)$ , for which  $\alpha - \beta \notin \mathcal{L}$  become exponentially suppressed as  $m \rightarrow \infty$ . We visualize the factor  $\nu_{M^\perp}$  for the square- and hexagonal GKP code in fig. 4.

For finite strength  $m$  of the twirl, the error can be bounded as follows. Let

$$\bar{\nu}_{\mathcal{L}}(\Delta) := \lim_{m \rightarrow \infty} \nu_{M^\perp}(\Delta)^m, \quad (41)$$

in this limit the function is independent of the choice of dual generating set  $M^\perp$ , which is why the index has been replaced by  $\mathcal{L}$ . For any finite  $m$ , we have

$$\nu_{M^\perp}(\Delta)^m = \bar{\nu}_{\mathcal{L}^\perp}(\Delta) + \nu_{M^\perp}(\Delta)^m [1 - \bar{\nu}_{\mathcal{L}^\perp}(\Delta)]. \quad (42)$$

To bound the error term, observe that the contribution  $[1 - \bar{\nu}_{\mathcal{L}^\perp}(\Delta)]$  is only non-zero when  $\Delta \notin \mathcal{L}$ . Let  $\mathbf{x} = \text{CVP}(\Delta, \mathcal{L})$  be the closest vector in  $\mathcal{L}$  to  $\Delta$  and let  $\delta = \Delta - \mathbf{x}$  be the corresponding minimal vector between  $\Delta$  and the lattice. Assuming  $\delta$  is small, we can bound each cosine term

$$\cos^2\left(\pi \left(\xi_i^\perp\right)^T J \Delta\right) = e^{-\pi^2 \|(\xi_i^\perp)\|^2 \|\delta\|^2} + O(\|\delta\|^4), \quad (43)$$

such that, in total, we have for  $\Delta \notin \mathcal{L}$  close to the lattice

$$\nu_{M^\perp}(\Delta)^m = e^{-\pi^2 \|M^\perp\|_F^2 \|\delta\|^2 m} + O(\|\delta\|^{4m}), \quad (44)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm and  $\delta = \Delta - \text{CVP}(\Delta, \mathcal{L})$  is the minimal distance between  $\Delta$  and  $\mathcal{L}$ .

Note that in the above construction we have decided to work with the twirl induced by the described random walk because it yielded a particularly nice analytic form for the characteristic function

$$\nu_{M^\perp}(\mathbf{x}) = \int_{\mathbb{R}^{2n}} d\mu_{M^\perp}(\gamma) e^{-i2\pi \gamma^T J \mathbf{x}}. \quad (45)$$

There is no other particularly good reason to use this parametrization. In general, one may also choose an arbitrary regularizer  $R(\gamma)$  to regularize a uniform distribution over the lattice  $\mathcal{L}^\perp$  as

$$d\mu(\gamma) = d\gamma R(\gamma) \sum_{\xi^\perp \in \mathcal{L}^\perp} \delta^{(2n)}(\gamma - \xi^\perp). \quad (46)$$

By the properties of the Fourier transform the result will be provided by the convolution of the asymptotic characteristic with the Fourier-transform of the regularization  $\hat{R}$ , which will be more localized the more homogeneous  $R$  is,

$$\nu(\Delta) = (\hat{R} * \bar{\nu}_{\mathcal{L}}(\Delta)). \quad (47)$$

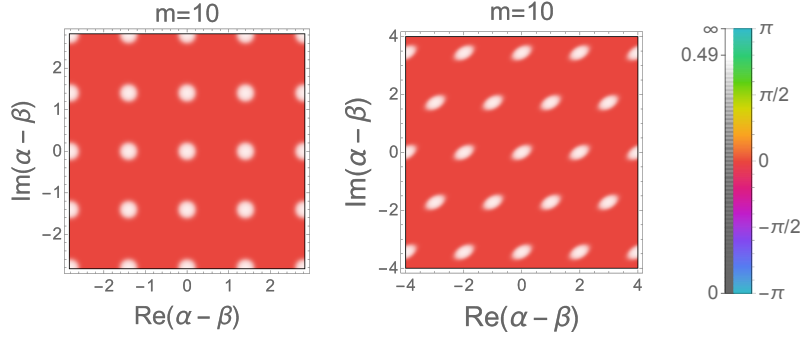


Figure 4: Figure for  $\nu_{M^\perp}$  for square and hexagonal GKP codes.

### 3.3 Gaussian unitary twirling

In addition to diagonalizing a state- or channel in the displacement (or Pauli-) basis, the diagonal components can also further be symmetrized. This was discussed previously by twirling over the Clifford group for (qubit-) Pauli-operators. Here we follow the same rationale to symmetrize the logical components of a state or channel relative to a prescribed GKP code. It is convenient the logical Clifford group for the GKP code can be implemented using only Gaussian unitary operations, since this leads to a relatively simple analysis of the physical incarnation of the relevant twirl. It will turn out later that some physical measurement channels possess natural symmetries that remedy the necessity to symmetrize them further via a twirl. We will exploit this feature later in the explicit examples using heterodyne measurements and photon-parity measurements. We implement a random logical Clifford operation as a Gaussian unitary twirl parametrized by a symmetric measure over the symplectic matrices  $\mu(S) = \mu(S^{-1})$ , which implements the map on the displacement twirled chi-function

$$\begin{aligned} [\nu_{M^\perp}(\alpha - \beta)]^m c(\alpha, \beta) &\mapsto \int d\mu(S) \left[ \nu_{M^\perp}(S^{-1}(\alpha - \beta)) \right]^m c(S^{-1}\alpha, S^{-1}\beta) \\ &= \int d\mu(S) [\nu_{M^\perp S^T}(\alpha - \beta)]^m. \end{aligned} \quad (48)$$

Although such random unitary modifications will never truly project a channel onto one that acts solely within code space, our goal is to modify it such that for any logical input state and any channel, decoded logical readout will make it appear as if the twirled channel was a full logically depolarizing. I.e., we target that the projection of the channel onto code space to behave like a logically depolarizing channel.

Note that for the explicit examples we consider here  $n = 1$  is small enough to simply enumerate the corresponding logical groups  $\text{Sp}_2(\mathbb{Z}_d)$  and sample directly from those sets. Nevertheless, we briefly outline how the twirl works in larger systems where the size of  $\text{Sp}_2(\mathbb{Z}_d)$  grows super-exponentially. The idea is to choose a generating set  $G$  for the symplectic automorphism group  $\mathcal{G} = \text{Aut}^S(\mathcal{L}^\perp)$  and generate random group elements by choosing random sequence from this generating set. It is a result from the theory of random matrices [40, 41] that a close-to uniform distribution over the target group is obtained by performing at least  $k$  steps, once  $k \geq |G| \text{diam}^2(\mathcal{G}, G)$ , where  $|G|$  is the size of the generating set and the group diameter  $\text{diam}(\mathcal{G}, G)$  expresses the minimal number of generators needed to express any element in the group  $\mathcal{G}$ .

For  $\mathcal{G} = \text{Sp}_{2n}(\mathbb{Z}_d)$  with generating set given by elementary generalized Hadamard, phase and CNOT gates (see appendix), which in their real representation are given by

symplectic transvections, this evaluates to a bound  $k \geq O(d^2 n^6)$ . This bound can be slightly improved to  $k \geq O(d n^6)$  when adding the  $r = 0, \dots, d-1$  powers of the elementary gates. We hence obtain that a logical twirl over the full (non-trivial) logical Clifford can be realized by random sequences of  $k \geq O(d^2 n^6)$  elementary symplectic transvections.

## 4 GKP logical shadows

In the previous section we have discussed how a general CV channel  $L$  can be twirled using displacement operators and Gaussian unitary operations such that it effectively yields a logical depolarizing channel. We here turn to showing how these insights and the established machinery can be turned into a scheme devising GKP logical shadows for the efficient measurement of expectation values of observables on the logical level. The key to this analysis is to develop an understanding of how the physical twirl manifests itself on the logical level, where the conversion from a physical representation of the state to its logical information content is facilitated by a decoder.

### 4.1 From physical to logical twirls

In the following we discuss the how the physical structure of a twirl manifests on the logical level. Understanding this connection allows us to build logical shadow tomography protocols based on physical measurements statistics. This approach is more versatile and effective than projecting a physically reconstructed state onto the code space of a chosen code, an approach chosen in ref. [42], but not suitable for the GKP code.

For any physical channel  $\mathcal{C}$ , denote by  $\mathcal{C}^\tau$  the channel twirled using logical displacements from  $\mathcal{L}^\perp$  and Gaussian unitaries represented by  $\text{Aut}^S(\mathcal{L}^\perp)$  as discussed earlier. We define a *decoder*  $\text{Dec}$  to be an surjective map from physical states to logical, error-free states  $\text{Dec}|\rho\rangle\rangle = |\bar{\rho}\rangle\rangle$ . The decoder is required to commute with noiseless logical channels  $\text{Dec}\mathcal{C} = \mathcal{C}\text{Dec}$ , where  $\mathcal{C}$  are logical Clifford operations (appropriately represented by a Gaussian unitary channel) and reduce error channels  $\mathcal{R}$  compactly supported on displacements from the Voronoi cell

$$\mathcal{V}(\mathcal{L}^\perp) := \left\{ \mathbf{x} \in \mathbb{R}^{2n}, \|\mathbf{x}\| \leq \|\mathbf{x} - \boldsymbol{\xi}^\perp\| \forall \boldsymbol{\xi}^\perp \in \mathcal{L}^\perp \right\} \quad (49)$$

to the logical identity channel  $\Pi = \Pi_{\mathcal{L}} \otimes \Pi_{\mathcal{L}}^*$ ,  $\Pi\mathcal{R} = \Pi$ , where  $\Pi_{\mathcal{L}}$  is the code space projector.

Applying the decoder to the twirled channel, we obtain for any input state  $\rho$

$$\begin{aligned} \text{Dec}\mathcal{C}^\tau|\rho\rangle\rangle &= \alpha\text{Dec}|\rho\rangle\rangle + \beta|\Pi_{\mathcal{L}}\rangle\rangle \\ &= \alpha|\bar{\rho}\rangle\rangle + \beta|\Pi_{\mathcal{L}}\rangle\rangle := \widetilde{\mathcal{M}}|\bar{\rho}\rangle\rangle, \end{aligned} \quad (50)$$

that the decoded twirled channel effectively behaves like a logically depolarizing channel. This decomposition results from the fact that for any logical Clifford operation we have that  $\text{Dec}\mathcal{C} = \mathcal{C}\text{Dec}$ , such that the decoder in particular also commutes with the logical twirl. On the (decoded) logical level it is already known that a Clifford twirl projects on a channel of the above form of a depolarizing channel. The depolarizing channel is invertible as  $\widetilde{M}^{-1}(X) = (X - \beta\Pi_{\mathcal{L}})/\alpha$  as long as  $\alpha \neq 0$ . Let  $\mathcal{C} = \sum_i |\Pi_i\rangle\rangle\langle\langle\Pi_i|$  be the channel corresponding to a projective POVM (see also a formulation using the theory of measurement frames [13]). Equation (50) takes the form

$$\widetilde{\mathcal{M}}|\bar{\rho}\rangle\rangle = \text{Dec} \sum_i \int_{\text{GU}(n;\mathcal{L})} d\mu(U) \mathcal{U}|\Pi_i\rangle\rangle\langle\langle\Pi_i|\mathcal{U}^\dagger|\rho\rangle\rangle, \quad (51)$$



where the unitaries are drawn from a distribution over Gaussian unitary operations representing logical Clifford gates obtained from concatenating the displacement- and non-trivial logical Clifford GU ( $n; \mathcal{L}$ ), as discussed in the previous sections. Since  $\widetilde{\mathcal{M}}$  commutes as logical channel with  $\text{Dec}$ , given a physical state  $|\rho\rangle\rangle$ , we can obtain the expectation value of any logical observable  $O$  via the following shadow protocol.

1. Sample a set of  $\{\mathcal{U}_i\}_{i=1}^N$  from GU ( $n; \mathcal{L}$ ) and perform the POVM for each sample. The probabilities of obtaining outcome pointers  $|\Pi_i\rangle\rangle$  are each given by the Born probability  $\langle\langle \Pi_i | \mathcal{U}_i^\dagger | \rho \rangle\rangle$ .
2. Reconstruct the shadow as empirical expectation value over the output states

$$|S\rangle\rangle = N^{-1} \sum_{i=1}^N \mathcal{U}_i |\Pi_i\rangle\rangle. \quad (52)$$

3. Compute the logical expectation value

$$\langle\langle O | \widetilde{M}^{-1} | S \rangle\rangle = \langle\langle \widetilde{M}^{-1}(O) | S \rangle\rangle. \quad (53)$$

Since

$$\text{Dec} \mathbb{E} [\widetilde{M}^{-1} \mathcal{U} |\Pi\rangle\rangle] = \text{Dec} |\rho\rangle\rangle, \quad (54)$$

for sufficiently large sample size the empirical mean state converges to the true mean state and the procedure recovers the expectation value of the logical observable as if it was measured directly after implementing a suitable decoder, that is, the estimator

$$\tilde{o} = \langle\langle \widetilde{M}^{-1}(O) | \mathcal{U} |\Pi\rangle\rangle \quad (55)$$

inherits the mean  $\langle \tilde{o} \rangle = \langle\langle O | \text{Dec} |\rho\rangle\rangle$ .

It is key to the protocol, however, that not only the mean is correctly recovered, but that also the variance is small: on this logical level the estimation essentially does not differ from the typical qubit scenario as already detailed in the construction provided by Huang, Kueng and Preskill [9] and their performance guarantees apply, in that a small number of samples suffices to reproduce the expectation value of many observables with high confidence.

**Theorem 1** (HKP [9]). *A collection of  $NK$  samples  $\{(\mathcal{U}_i, \Pi_i)\}_{i=1}^{NK}$  produced via the above protocol from a CV state  $\rho$  suffice to estimate logical expectation values on observables  $O_i$ ,  $i = 1, \dots, M$  via median of means prediction up to  $\epsilon$  additive error provided that*

$$K = 2 \log(2M/\delta), \quad N = \frac{34}{\epsilon^2} \max_i \left\| O_i - \frac{\text{Tr}[O_i]}{2^n} I \right\|_{\text{shadow}}^2 \quad (56)$$

with probability at least  $1 - \delta$ .

We refer to ref. [9] for the proof and the definition of the shadow norm, for which ref. [9] has also provided the upper bound

$$\left\| O_i - \frac{\text{Tr}[O_i]}{2^n} I \right\|_{\text{shadow}}^2 \leq 3 \text{Tr}[O_i^2]. \quad (57)$$

Median of means prediction is carried out by dividing the set of  $NK$  samples into  $K$  batches of  $N$  samples each, for each of which the arithmetic mean is computed and taking the median over the batches

$$\tilde{o}^{\text{est}} = \text{median} \left\{ N^{-1} \sum_i \tilde{o}_{jN+i} \right\}_{j=1, \dots, K}. \quad (58)$$

In particular, this also implies that the reconstructed state (52) yields a good representation of the decoded state  $\text{Dec} |\rho\rangle\rangle$  in that it reproduces the expectation value of many low-rank observables with only small additive error. By paying an extra cost in sample overhead, we can use the shadow to obtain a full representation of the (finite dimensional) state.

**Theorem 2** (Full representation of the state). *Let  $\mathcal{L} \subset \mathbb{R}^{2n}$  denote the lattice corresponding to a scaled GKP code on  $n$  modes that encodes  $d^n$  logical dimension and let  $\mu$  denote a measure over elements  $\text{Aut}^S(\mathcal{L}^\perp)$  forming a logical Clifford 2-design and let  $\mathcal{M} = \int dz |\Pi(z)\rangle\rangle \langle\langle \Pi(z)|$  denote a physical projective POVM. Let  $|\rho\rangle\rangle$  be an arbitrary state vector on the  $n$ -mode Hilbert space. The state vector*

$$|S\rangle\rangle = N^{-1} \sum_{i=1}^N \mathcal{U}_{S,i} |\Pi(z_i)\rangle\rangle \quad (59)$$

*produced by sampling  $N$  Gaussian unitary operations via the measure  $\mu$  and measurement outcomes from the Born distribution  $z_i \sim \langle\langle \Pi(z) | \mathcal{U}_{S,i}^\dagger | \rho \rangle\rangle$  approximates the logical value of the state in Hilbert-Schmidt distance*

$$d_{\text{HS}}(\text{Dec} |\rho\rangle\rangle, \text{Dec} |S\rangle\rangle) \leq \delta_{\text{HS}}^2 \quad (60)$$

*with probability at least  $1 - \delta$  for*

$$N \geq \frac{2d^{2n}}{\alpha^2 \delta_{\text{HS}}^2} \left[ \ln \left( \frac{2}{\delta} \right) + 2n \ln(d) \right], \quad (61)$$

*where  $\alpha$  is determined by the commutation of the twirled POVM with the decoder*

$$\text{Dec} \mathcal{M}^\tau = \alpha \text{Dec} + \beta |\Pi_{\mathcal{L}}\rangle\rangle. \quad (62)$$

*In particular, we also have*

$$\|\text{Dec} |\rho\rangle\rangle - \text{Dec} |S\rangle\rangle\|_1 \leq d^{\frac{n}{2}} \frac{\delta_{\text{HS}}}{2}. \quad (63)$$

See appendix C for the proof of this statement. Although the sample overhead derived here contains the extra dimensional factor  $O(d^{2n})$  and is therefore rather large compared to the usual situations considered in ref. [9], this statement allows us to apply the shadow tomography toolbox to derive simple representations of bosonic states with the right information content on selected (GKP) subspaces. The extensive overhead, contrasting the result of ref. [9], here essentially stems from “abusing” the shadow tomography routine to perform full (logical) state tomography, requiring the well-estimation of a complete number of observables scaling with the system size. While this overhead may seem daunting, estimating expectation values of observables is not our primary goal. Instead, we emphasize that the shadow tomography protocol allows for a decomposition of the input state into a mixture of pointer states that reproduces the logical information of the encoded state. The sample complexity bounds imply an upper bound on the number of components in this

mixture. This bound is generically loose, but the fact that it can be stated and derived using the tools used for shadow tomography we believe is nontrivial. While decompositions may also be formalized using the language of measurement frames [13], the presented toolbox allows to quantify the strength of the decomposition in a flexible way in the chosen POVM. We exemplify its utility using the heterodyne- and photon-click POVM that resolve the presence and absence of photons.

## 4.2 Gaussian decomposition of the hexagonal GKP code from heterodyne measurements

An interesting use case of our protocol with practical applications is obtained by considering heterodyne measurements paired with the hexagonal GKP code. Heterodyne measurements are such that the measurement pointer states are Gaussian states. Furthermore, the measurement channel is rotationally symmetric, which will turn out to skip the need to apply non-trivial Gaussian operations in the implementation of our protocol when paired with the hexagonal GKP code. After developing the basic ingredients to this protocol, we will later discuss how it can be applied to simulate the forward-propagation of general quantum states through CV circuits.

In our convention for displacement operators, we define generalized coherent state vectors as  $|\alpha\rangle = D(\alpha)|0\rangle$ , where  $|0\rangle$  denotes a  $n$ -mode vacuum state where we have

$$\begin{aligned} |\alpha\rangle\langle\alpha| &= \int_{\mathbb{R}^{2n}} d\beta \operatorname{Tr} \left[ D^\dagger(\beta) |\alpha\rangle\langle\alpha| \right] D(\beta) \\ &= \int_{\mathbb{R}^{2n}} d\beta e^{-\frac{\pi}{2}\beta^T\beta - i2\pi\alpha^T J\beta} D(\beta), \end{aligned} \quad (64)$$

such that a resolution of the identity is given by

$$\int_{\mathbb{R}^{2n}} d\alpha |\alpha\rangle\langle\alpha| = I. \quad (65)$$

In this sense, the coherent states constitute an overcomplete resolution of the identity. We also have

$$\langle\beta|\alpha\rangle = e^{-\frac{\pi}{2}(\|\alpha-\beta\|^2 + i2\pi\alpha^T J\beta)}. \quad (66)$$

Generalized heterodyne measurements are interferometric quantum optical measurements that are known to in effect implement a projective POVM described by the quantum channel

$$\begin{aligned} \mathcal{C} &= \int_{\mathbb{R}^{2n}} d\alpha |\alpha\rangle\langle\alpha| \otimes |\alpha\rangle\langle\alpha|^* \\ &= \int_{\mathbb{R}^{2n}} d\beta e^{-\pi\beta^T\beta} D(\beta) \otimes D(\beta)^*, \end{aligned} \quad (67)$$

equivalent to a Gaussian displacement channel where displacements of amplitude  $\beta$  occur with probability density  $e^{-\pi\beta^T\beta}$ . The above equation is obtained by expanding each coherent state  $|\alpha\rangle\langle\alpha|$  in displacement operators and executing the integral over the variable  $\alpha$ . In this form it becomes clear why this POVM is special. 1. it is already diagonal in displacement operators, such that a displacement twirl has no effect and 2. the chi-function only depends on the Euclidean length  $\|\beta\|$  of the corresponding displacement amplitudes, such that the channel is also invariant under Gaussian unitary twirls realized via orthogonal symplectic transformations of  $\beta$ .

For the single mode hexagonal GKP code we have found that a Clifford 2-design is given by the set of rotations  $R_{\frac{2\pi}{3}}$ , which are such orthogonal symplectic transformations.

Hence this is an example where the twirls have no effect. We can define a sequence of Voronoi-“shells”

$$\mathcal{V}^0 = \mathcal{V}(\mathcal{L}^\perp), \quad (68)$$

$$\mathcal{V}^k = \mathcal{V}((2k+1)\mathcal{L}^\perp) \setminus \mathcal{V}^{k-1}, \quad k \in \mathbb{N}_0, \quad (69)$$

$$\bigcup_{k=0}^{\infty} \mathcal{V}^k = \mathbb{R}^{2n}, \quad (70)$$

where each shell with  $m = 0 \pmod d$  contains logically trivial displacement amplitudes. In particular  $\forall \mathbf{x} \in \mathcal{V}(\mathcal{L}^\perp)$  our decoder maps

$$\text{Dec}\mathcal{D}(\mathbf{x}) = \text{Dec}. \quad (71)$$

For  $d = 2$ , we can compute

$$\text{Dec}\mathcal{C} = \text{Dec} \sum_{k=0}^{\infty} \int_{\mathcal{V}^k} d\boldsymbol{\alpha} e^{-\pi \boldsymbol{\alpha}^T \boldsymbol{\alpha}} \quad (72)$$

$$= (p_0 - p_1)\text{Dec} + 2p_1\Pi_{\mathcal{L}},$$

$$p_i = \sum_{k=0}^{\infty} \int_{\mathcal{V}^{2k+i}} d\boldsymbol{\alpha} e^{-\pi \boldsymbol{\alpha}^T \boldsymbol{\alpha}} \quad (73)$$

where we have used that each displacement in an even shell is removed by the decoder and each displacement in an odd shell is equally likely attributed to a logical Pauli  $X-$ ,  $Y-$ , or  $Z-$  displacement. Here we have that the associated logical depolarizing channel  $\widetilde{M} = (p_0 - p_1)\text{Dec} + 2p_1\Pi_{\mathcal{L}}$  is invertible as long as  $p_0 \neq p_1$ .

Since we have  $p_0 + p_1 = 1$ , this can be verified by showing that  $p_0 \neq 1/2$ . We have  $\mathcal{V}^0 \supseteq \mathcal{B}^2(\rho(\mathcal{L}^\perp))$  and for  $k \geq 1$

$$\mathcal{V}^k \subseteq \mathcal{B}^2((2k+1)\mu(\mathcal{L}^\perp)) \setminus \mathcal{B}^2((2k-1)\rho(\mathcal{L}^\perp)), \quad (74)$$

where  $\rho(L) = \lambda_1(L)/2$  denotes the packing radius of the lattice  $L$  and  $\mu(L)$  denotes the covering radius, which for the hexagonal lattice is given by  $\mu(A_2) = \lambda_1(A_2)/\sqrt{3}$  where  $\lambda_1(A_2) = \sqrt{2/\sqrt{3}}$ . Using  $\lambda_1(\mathcal{L}^\perp) = \lambda_1(A_2)/\sqrt{2}$  we can thus bound

$$\begin{aligned} p_0 &= \sum_{k=0}^{\infty} \int_{\mathcal{V}^{2k}} e^{-\pi \|\boldsymbol{\alpha}\|^2} \leq 2\pi \int_0^{\mu(\mathcal{L}^\perp)} r e^{-\pi r^2} \\ &\quad + 2\pi \sum_{k=1}^{\infty} \int_{(4k-1)\lambda_1(\mathcal{L}^\perp)/\sqrt{3}}^{(4k+1)\lambda_1(\mathcal{L}^\perp)/2} dr r e^{-\pi r^2} \\ &= \left[1 - e^{-\frac{\pi}{3\sqrt{3}}}\right] \\ &\quad + \sum_{k=1}^{\infty} \left(e^{-\frac{\pi}{3\sqrt{3}}(4k-1)^2\lambda_1^2} - e^{-\frac{\pi}{4\sqrt{3}}(4k+1)^2\lambda_1^2}\right) \\ &\approx 0.455 \dots \end{aligned} \quad (75)$$

This yields a bound  $|\alpha| = |p_0 - p_1| = |2p_0 - 1| \geq 0.09$ . We find that while the effective channel  $\widetilde{M}$  is in fact invertible, the logical information retained in the depolarized state is rather minuscule as indicated by the value  $|\alpha| = |p_0 - p_1| = |2p_0 - 1| \geq 0.09$ .

### 4.3 Decoding

We briefly comment on the classical post-processing required to execute this protocol where, w.l.o.g, we focus on the logical Pauli-Z measurement outcomes. Given a sample corresponding to a pure Gaussian state  $\rho_G$  for a GKP code with lattice  $\mathcal{L}$ , there exist a Gaussian unitary transformation implementing a symplectic transformation  $S$  such that  $S\mathcal{L} = \bigoplus_{i=1}^n \mathcal{L}_i$ , where each  $\mathcal{L}_i$  describes a local square single mode GKP code encoding  $d_i$  dimensions. We denote the Gaussian state resulting from this transformation by  $\rho_{G'}$ . The decoding map then effectively implements

$$\overline{\rho_Z} = \bigotimes_{i=1}^n \text{dec}_i(\rho_{G'}), \quad (76)$$

where with  $k_i = 1, \dots, d_i - 1$

$$\text{dec}_i(\rho_{G'}) = \sum_{k_i=0}^{d_i-1} \int_{S_{k_i,i}} \text{Tr}[\rho_{G'} |x_i\rangle \langle x_i|_q] dx_i |\overline{k_i}\rangle \langle \overline{k_i}| \quad (77)$$

represents the “binned homodyne detection” and

$$S_{k_i,i} = \bigcup_{s \in \mathbb{Z}} \left[ -\frac{1}{2} \sqrt{\frac{2\pi}{d_i}} + \sqrt{\frac{2\pi}{d_i}} k_i + \sqrt{2\pi d_i} s, \frac{1}{2} \sqrt{\frac{2\pi}{d_i}} + \sqrt{\frac{2\pi}{d_i}} k_i + \sqrt{2\pi d_i} s \right]. \quad (78)$$

As  $\rho_{G'}$  is Gaussian, this decoding step can be approximated using suitably large truncation  $|s| \leq \overline{s}$ , which can be estimated to be on the scale of  $\text{sq}(S) = \sqrt{\lambda_{\max}(S^T S)}$ . Note that as

$$\bigcup_{i=0}^{d_i-1} S_{k_i,i} = \mathbb{R} \quad \forall i = 1, \dots, n, \quad (79)$$

The single mode decoder in eq. (77) should be approximated by examining the relative magnitude over different  $k_i$ 's of the integral sequence. With each mode being decoded separately, the overall cost in postprocessing can be estimated to scale as  $O(n \text{sq}(S))$ .

We do not examine this question further in this work, and remark that the information theoretic decomposition into Gaussian states, and their subsequence use in simulation are not hindered by a potentially costly decoding step.

### 4.4 Decomposing GKP states from photon click detectors

Another practically relevant measurement channel is provided by photon click detectors [43]. Photon click detectors have dichotomic outcomes and distinguish the presence from the absence of photons. The relevant POVM is given by

$$\mathcal{C} = \Pi_0 \otimes \Pi_0^* + \Pi_1 \otimes \Pi_1^*, \quad (80)$$

$$\Pi_0 = |0\rangle \langle 0|, \quad (81)$$

$$\Pi_1 = I - |0\rangle \langle 0|, \quad (82)$$

with chi-function

$$c(\boldsymbol{\alpha}, \boldsymbol{\beta}) = e^{-\frac{\pi}{2} \|\boldsymbol{\alpha}\|^2 - \frac{\pi}{2} \|\boldsymbol{\beta}\|^2} + \left( \delta(\boldsymbol{\alpha}) - e^{-\frac{\pi}{2} \|\boldsymbol{\alpha}\|^2} \right) \left( \delta(\boldsymbol{\beta}) - e^{-\frac{\pi}{2} \|\boldsymbol{\beta}\|^2} \right). \quad (83)$$

As before, this channel is rotationally invariant, such that the only non-trivial component of the twirl is given by its diagonalizing action on the chi function. Note that the wave function collapse induced by projective measurement of a generating set of GKP stabilizers is equivalent a full stabilizer displacement twirl (with  $m \rightarrow \infty$ ). Ignoring stabilizer coherences, the channel thus takes the form

$$\mathcal{C}^\tau(\rho) = \int d^2\alpha f(\alpha) D(\alpha) \rho D^\dagger(\alpha), \quad (84)$$

$$f(\alpha) = \sum_{\xi \in \mathcal{L}} c(\alpha, \alpha + \xi) e^{-i\pi\alpha^T J \xi}. \quad (85)$$

Commuting through a decoder, the effective depolarizing channel obtains the coefficients

$$\alpha = 1 - 2\vartheta(0 | iG_{A_2}) + I_1, \quad (86)$$

$$\alpha + \beta = 1 - 2\vartheta(0 | iG_{A_2}) + I_2, \quad (87)$$

where we have defined the Riemann theta function as

$$\vartheta(\mathbf{z} | F) = \sum_{\mathbf{m} \in \mathbb{Z}^{2n}} e^{i2\pi(\frac{1}{2}\mathbf{m}^T F \mathbf{m} + \mathbf{m}^T \mathbf{z})} \quad (88)$$

and

$$I_1 = \sqrt{2} \int_{\mathcal{V}(A_2)} d\alpha' e^{-\frac{\pi}{2}\|\alpha'\|^2} \vartheta(M_{A_2}(iI + J)\alpha' | iG_{A_2}) \\ \times \vartheta(M_{A_2}(iI - J)\alpha' | iG_{A_2}), \quad (89)$$

$$I_2 = 2\sqrt{2} \int_{\mathcal{V}(A_2)} d\alpha' e^{-2\pi\|\alpha'\|^2} \vartheta(M_{A_2}(iI + J)\alpha' | iG_{A_2}) \\ \times \vartheta(M_{A_2}(iI - J)\alpha' | iG_{A_2}), \quad (90)$$

where  $G_{A_2} = M_{A_2} M_{A_2}^T$  is the Euclidean Gram matrix for the symplectic basis of the  $A_2$  lattice. By numerical integration and approximating the Riemann theta function with one evaluated using a truncated sum  $\mathbf{m} \in \{-t, \dots, t\}^{2n}$  with  $t$  chosen large enough for convergence of the output values, we evaluate  $\vartheta(0 | iG_{A_2}) = 1.1596$ ,  $I_1 = 1.493$ ,  $I_2 = 1.64$  to obtain the estimates  $\alpha = 0.32$  and  $\beta = 0.15$ . Similar to the previous section, this shows that a representation of the logical content of a single mode GKP code relative to the hexagonal GKP code can be obtained from a finite number of samples from random displaced photon click detectors, of which the number is bounded by theorem 2. As a final remark we mention that potential decoding of the pointer states can proceed analogously as discussed in the previous section.

In this section, we have adapted the (HKP classical) shadow tomography protocol introduced in ref. [9] to operate on the logical degrees of freedom of quantum information encoded via a GKP code in a continuous variable system and considered two of the best-behaved POVMs given by heterodyne detection and a photon-click detector and derived the constant  $\alpha$  to show the invertibility of the effective depolarizing channel. While the full analysis is more complex for more general POVMs, such as photon counting, our approach is sufficiently general to be generalized to such situations.

#### 4.5 Applications to classical simulation

In this section, we briefly speculate about applications of the logical shadow tomography schemes developed here. The main development of the preceding sections has been that

one can view the GKP logical shadow protocol as a physical black box protocol that converts a given physical state via application of appropriate random Gaussian unitaries into a convex combination of Gaussian states when using heterodyne detection as the underlying POVM. Since Gaussian states are efficiently described by mean  $\bar{\mathbf{x}} \in \mathbb{R}^{2n}$  and variance  $V = S^T S \in \mathbb{R}^{2n \times 2n}$ , each part of the decomposition offers an efficient classical description and transforms in a simple manner under application of Gaussian channels [19]. Similar decompositions have been utilized, e.g., in ref. [44] to develop classical simulation methods for CV states by first decomposing them into a linear combination over Gaussian states, analogous to simulation via stabilizer state decompositions found in the qubit-based quantum computing literature, see, e.g., ref. [45] and references therein. The practical drawback of the previously presented methods, however, is that an analytic description of the initial state must be known *a priori*, on which basis the decomposition then proceeds and no bound has been derived on the number of parts necessary to obtain a good approximation of the state. By focusing on reproducing underlying logical properties of the input state and using statistical methods known from classical shadow tomography, the tools developed here yield an experimental method that converts a given physical state into a convex combination of Gaussian states, corresponding to the samples obtained in the protocol and our bounds yield upper bounds on the number of components (samples) necessary to achieve a good approximation to the input state on the logical level. In the development of practical quantum computers based on GKP encoded logic, one may hence apply this protocol to convert experimentally realizable (multi-mode) GKP states into a classical description of Gaussian state components, which can then further be used to assess their performance in algorithms or under noise. While the overhead is not expected to scale favorably in the number of non-Gaussian channels, one may also imagine a recursive version of this procedure to simulate non-Gaussian evolutions while using the logical shadow protocol to repeatedly convert the mid-simulation states into a convex combination of Gaussian states. The protocol outlined here is only one of many possible uses of our scheme and highlights its potential for practical applications.

#### 4.6 Random Wigner tomography

The final measurement channel we consider to exemplify our protocol is photon parity measurements. This measurement channel is practically widely implemented and relevant for its application to Wigner tomography [46] as we will further discuss below. It will turn out that logical GKP shadows can be understood as a randomized Wigner tomography scheme, where the Wigner function of the input state is only evaluated at random points chosen from a distribution close to the symplectic dual lattice  $\mathcal{L}^\perp$  of a GKP code. This yields a protocol which estimates the logical shadow for any state close to a GKP code state but is not limited to such states. For states not close to GKP code states, it merely reproduces a logical shadow of a state *if it was interpreted* as a code state. Later, in sec. 5, we consider this protocol relative to random choices of GKP codes and quantify how such protocol yields a shadow for the full physical input state.

The parity operator  $\hat{\pi} = e^{-i\pi\hat{N}}$ , with  $\hat{N} = \sum_{i=1}^n \hat{n}_i$  has an interesting decomposition into displacement operators with constant characteristic function

$$\begin{aligned} \text{Tr} \left[ D^\dagger(\beta) \hat{\pi} \right] &= \int_{\mathbb{R}^{2n}} d\alpha \left\langle \alpha | D^\dagger(\beta) | -\alpha \right\rangle \\ &= \int_{\mathbb{R}^{2n}} d\alpha e^{-2\pi\alpha^T \alpha} = 2^{-n}. \end{aligned} \quad (91)$$

The fact that this characteristic function is constant has the effect that expectation values



of displaced parity measurements yield

$$\begin{aligned}\langle D(\mathbf{x}) \hat{\pi} D^\dagger(\mathbf{x}) \rangle &= 2^{-n} \int_{\mathbb{R}^{2n}} d\boldsymbol{\beta} e^{-i2\pi \mathbf{x}^T J \boldsymbol{\beta}} \text{Tr}[D(\boldsymbol{\beta}) \rho] \\ &= 2^{-n} W_\rho(\mathbf{x}).\end{aligned}\quad (92)$$

Up to the rescaling  $2^{-n}$  this is precisely the Wigner function we have encountered in eq. (4). This basic insight has been used consistently in quantum optics for the purposes of tomographic recovery [47]. The chi-function associated to the measurement channel

$$c(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{2} [\delta(\boldsymbol{\alpha}) \delta(\boldsymbol{\beta}) + 2^{-2n}] \quad (93)$$

is invariant under symplectic transformations  $S \in \text{Sp}_{2n}(\mathbb{R})$ :  $c(S\boldsymbol{\alpha}, S\boldsymbol{\beta}) = c(\boldsymbol{\alpha}, \boldsymbol{\beta})$ : the only non-trivial component a logical Clifford twirl can have is provided by the displacement twirl over displacements in  $\mathcal{L}^\perp$ . Due to the connection of this POVM to Wigner tomography, each displacement sampled in the displacement twirl  $\mu_{\mathcal{L}^\perp}(\boldsymbol{\gamma})$  corresponds to choosing to estimate the Wigner function of the input state  $\rho$  at the random point  $\boldsymbol{\gamma} \in \mathcal{L}^\perp$ . Operationally, the protocol works by 1. sampling a displacement vector from the distribution specified in the twirl and 2. measuring the photon parity a constant number of times to obtain the expectation value of the displaced photon parity operator or, equivalently, the probabilities to obtain a  $\pm 1$  outcome in the parity measurements. The expectation value of that displaced parity operator is exactly the value of the Wigner function at the sampled displacement vector and can be obtained with fixed target precision at a constant number of samples for each displacement vector  $\boldsymbol{\gamma} \in \mathcal{L}^\perp$  sampled in the twirl. For clarity of presentation we suppress this constant in our analysis. Similar as before, a displacement twirled photon parity measurements – let us call them *Wigner shadows* – can also be understood to behave like a logical depolarizing channel. In contrast to the cases we have considered before, the characteristic function in eq. (93) contains a constant that would yield a diverging contribution when summing over all stabilizer-equivalent coherences in the channel, which forces us to make a more refined model of the displacement twirl. We consider the measure  $d\mu(\boldsymbol{\gamma}) = p_{\sigma^2}(\boldsymbol{\gamma}; \mathcal{L}^\perp) d\boldsymbol{\gamma}$ , with probability density

$$p_{\sigma^2}(\boldsymbol{\gamma}; \mathcal{L}^\perp) = N_{\mathcal{L}^\perp}(\sigma)^{-1} \sum_{\boldsymbol{\xi}^\perp \in \mathcal{L}^\perp} e^{-\frac{\sigma^2}{2} \|\boldsymbol{\xi}^\perp\|^2 - \frac{1}{2\sigma^2} \|\boldsymbol{\gamma} - \boldsymbol{\xi}^\perp\|^2}, \quad (94)$$

$$N_{\mathcal{L}^\perp}(\sigma) = (2\pi\sigma^2)^n \Theta_{\mathcal{L}^\perp}\left(i\frac{\sigma^2}{4\pi}\right), \quad (95)$$

where we have introduced the lattice theta constant [23, 48]  $\Theta_L(z) = \sum_{\mathbf{x} \in L} e^{i2\pi z \|\mathbf{x}\|^2}$  to express the normalization factor.

The characteristic function of this distribution is given by

$$\begin{aligned}\nu_{\sigma^2}(\boldsymbol{\Delta}) &= \int_{\mathbb{R}^{2n}} d\mu(\boldsymbol{\gamma}) e^{-i2\pi \boldsymbol{\gamma}^T J \boldsymbol{\Delta}} \\ &= N_{\mathcal{L}^\perp}(\sigma) \sum_{\boldsymbol{\xi}^\perp \in \mathcal{L}^\perp} e^{-\frac{\sigma^2}{2} \|\boldsymbol{\xi}^\perp\|^2 - i2\pi (\boldsymbol{\xi}^\perp)^T J \boldsymbol{\Delta}} \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma} e^{-\frac{1}{2\sigma^2} \boldsymbol{\gamma}^T \boldsymbol{\gamma} - i2\pi \boldsymbol{\gamma}^T J \boldsymbol{\Delta}},\end{aligned}\quad (96)$$

and we can evaluate this line using the Poisson summation formula (see ref. [23] for a formulation tailored to symplectic inner products and sums over lattices) and Gaussian

integration to obtain

$$\nu_{\sigma^2}(\Delta) = c^{-1} \sum_{\xi \in \mathcal{L}} e^{-2\pi^2 \sigma^2 \|\Delta\|^2} e^{-\frac{2\pi^2}{\sigma^2} \|\Delta - \xi\|^2}, \quad (97)$$

$$c = \det(\mathcal{L}^\perp) \sigma^{2n} \Theta_{\mathcal{L}^\perp}(i\sigma^2/4\pi). \quad (98)$$

This distribution is in particular normalizable, with  $\int d\Delta \nu(\Delta) = 1$ . This allows us to compute the logical fidelity of the depolarizing channel

$$\begin{aligned} 1 - p &\geq \int_{\mathcal{V}(\mathcal{L}^\perp)} d\alpha \int_{\mathbb{R}^{2n}} d\Delta c(\alpha, \alpha - \Delta) \nu_{\sigma^2}(\Delta) \\ &= \frac{1}{2} \left[ 1 + \frac{\det(\mathcal{L}^\perp)}{2^{2n}} \right]. \end{aligned} \quad (99)$$

If the GKP code is chosen so to encode  $k$  qudits each of dimension  $d$  in  $n$  modes, this computation estimates that the logical information accessible from the samples can be understood to have undergone a logical depolarizing channel with depolarizing probability

$$p \leq \frac{1}{2} \left[ 1 - 2^{-n(2 + \frac{k}{n} \log_2(d))} \right]. \quad (100)$$

For a fixed number  $n$  of modes this estimate is reassuring: it tells us that the parameters can be chosen so that the effective logical error probability is bounded away from  $\frac{1}{2}$ . Asymptotically for large  $n$ , however, we are only promised an exponentially small amount of logical information content to survive the procedure.

#### 4.6.1 A quasi-logical estimator

The pointer states output by photon parity measurements are generically relatively unwieldy states and we expect it to be difficult to use them in a similar manner as e.g. the heterodyne pointers discussed earlier. Here it is more interesting to use samples from the Wigner function of an arbitrary state in the following way. Let  $G(\mathbf{x}) = W_G(\mathbf{x})$  be the Wigner function of an arbitrary <sup>3</sup> trace-class observable on a system of  $n$  quantum harmonic oscillators and let  $\mathcal{L} \subseteq \mathcal{L}^\perp$  describe a GKP code with even symplectic Gram matrix  $A$ . We define the estimator

$$\tilde{G}(\mathbf{x}) := W_\rho(\mathbf{x}) G(\mathbf{x}) \quad (101)$$

such that samples from the Wigner function of an input state according to the distribution  $p_\sigma(\gamma; \mathcal{L}^\perp)$  can be combined with the observable Wigner function  $G(\mathbf{x})$  at the same points to produce the expectation value

$$\begin{aligned} \overline{G}_{\mathcal{L}^\perp}^\sigma &= \int d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) W_\rho(\mathbf{x}) G(\mathbf{x}) \\ &\xrightarrow{\sigma \rightarrow 0} \sum_{\xi^\perp \in \mathcal{L}^\perp} W_\rho(\xi^\perp) G(\xi^\perp) =: \overline{G}_{\mathcal{L}^\perp}. \end{aligned} \quad (102)$$

If either the Wigner function of the input state or the observable were solely supported on the lattice  $\mathcal{L}^\perp$ , we see that in this limit  $\sigma \rightarrow 0$  the expectation value

$$\text{Tr}[\rho G] = \int_{\mathbb{R}^{2n}} d\mathbf{x} W_\rho(\mathbf{x}) G(\mathbf{x}) \quad (103)$$

---

<sup>3</sup>Hence, this is not necessarily finite dimensional.

would be exactly reproduced. We discuss an interpretation of eq. (102) in the next subsection and will later see how the expression can be of use. First, we examine how well the lattice sum is approximated for small  $0 < \sigma \ll 1$ . In order to bound the effective convergence of  $p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \rightarrow \text{III}_{\mathcal{L}^\perp}(\mathbf{x}) = \sum_{\boldsymbol{\xi}^\perp \in \mathcal{L}^\perp} \delta(\mathbf{x} - \boldsymbol{\xi}^\perp)$  under the integral, we need to make soft assumptions on the state and observable.

**Lemma 1** (Random lattice point sampling). *Let  $\rho, G$  be operators such that the product of their Wigner functions  $\tilde{G}(\mathbf{x}) = W_\rho(\mathbf{x}) G(\mathbf{x})$  is well defined,  $\text{Tr}[G^2] < \infty$  is finite and further assume that  $W_\rho(\mathbf{x}), G(\mathbf{x})$  are Lipschitz-continuous, with  $\|\nabla W_\rho(\mathbf{x})\| \leq l_\rho$  and  $\|\nabla G(\mathbf{x})\| \leq l_G$ . Set  $\tilde{l} = l_\rho + l_G$ . Let  $\sigma \ll \lambda_1(\mathcal{L}^\perp)$  be a small parameter. It holds that*

$$\begin{aligned} \overline{G}_{\mathcal{L}^\perp}^\sigma &= \int_{\mathbb{R}^{2n}} d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}(\mathbf{x}) \\ &= \sum_{\boldsymbol{\xi}^\perp \in \mathcal{L}^\perp} \tilde{G}(\boldsymbol{\xi}^\perp) + \epsilon(\sigma) \\ &= \overline{G}_{\mathcal{L}^\perp} + \epsilon(\sigma) \end{aligned} \quad (104)$$

with

$$|\epsilon(\sigma)| \leq \sqrt{2}\sigma \frac{\sqrt{2}\Gamma(n + \frac{1}{2})}{\Gamma(n)} \tilde{l}. \quad (105)$$

Furthermore, it holds that

$$\begin{aligned} \overline{G}_{\mathcal{L}^\perp}^{\sigma,2} &= \int d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}^2(\mathbf{x}) \\ &\leq \left(\frac{2}{\pi\sigma^2}\right)^n \text{Tr}[G^2]. \end{aligned} \quad (106)$$

Note that for large  $n$ , we have that  $\Gamma(n + \frac{1}{2})/\Gamma(n) = \sqrt{n} + O(n^{-1/2})$  while  $\lambda_1(\mathcal{L}^\perp) \propto \sqrt{n}$  is the maximally achievable shortest vector scaling for any family of lattices  $\mathcal{L}^\perp = \mathcal{L}_n^\perp$ . Picking  $\sigma = o(1/\sqrt{n})$  essentially allows to reduce the error  $\epsilon(\sigma)$  error to a negligible amount. Again here we observe a non-trivial scaling with the system size, contrasting the bounds obtained in ref. [9]. Note that the above statement concerns the variance of the estimator, which in ref. [9] is upper bounded by defining the *shadow norm*. The shadow norm is generically upper bounded using the rank of the observable, which, for finite dimensional systems, scales at most exponentially in the system size. Here the behavior is qualitatively different as  $\text{Tr}[G^2]$  captures the variance of the observable in phase space, for a potentially infinite-rank observable. The phase space integral that evaluates the trace may be split into separate integrals over each of the  $n$  subspaces that represent the individual modes, for which the values are each suppressed by the factor  $2/\pi\sigma^2$ .

#### 4.6.2 Interpreting $\overline{G}_{\mathcal{L}^\perp}$ & structure of general code states

We wish to find an operational interpretation of the estimator  $\overline{G}_{\mathcal{L}^\perp}$  in the limit  $\sigma \rightarrow 0$ , which can be understood as the expectation value  $\overline{G}_{\mathcal{L}^\perp} = \text{Tr}[\rho|_{\mathcal{L}^\perp} G]$  of a state  $\rho|_{\mathcal{L}^\perp}$  whose Wigner function has been projected to be exclusively supported on the lattice  $\mathcal{L}^\perp$ . A naive expectation would be that the code space projection  $\rho \mapsto \Pi_{\mathcal{L}} \rho \Pi_{\mathcal{L}}$  (or performing optimal quantum error correction) could perform such a restriction. In appendix A we show that this is indeed not the case and that the Wigner function of a GKP code state is also supported on points in  $\mathcal{L}^\perp/2$  outside of  $\mathcal{L}^\perp$ . And even if those points were not present, the

projection  $\rho \mapsto \Pi_{\mathcal{L}} \rho \Pi_{\mathcal{L}}$  does not clearly correspond to a restriction of the Wigner function of  $\rho$  onto the desired support.

Here we defined the code space projector of a GKP code associated to lattice  $\mathcal{L} \subset \mathbb{R}^{2n}$ , which is generally given by <sup>4</sup>

$$\Pi_{\mathcal{L},\phi} = \sum_{\xi \in \mathcal{L}} e^{i\phi(\xi)} D(\xi). \quad (107)$$

The function  $\phi$  specifies the phase associated with a particular choice of *pivot basis* [23, 25] or *gauge* [26]. We omit indicating  $\phi$  since it will not be relevant for the present discussion and simply write  $\Pi_{\mathcal{L}}$ . In particular, for scaled GKP codes with even  $d$  it holds that  $\phi = 0$ , which we now also assume to be the case. Earlier we have already encountered the Weyl- or phase-space-point operators  $\Omega(\mathbf{x}) = D(\mathbf{x}) \hat{\pi} D^\dagger(\mathbf{x})$ . These operators implement a degree  $\pi$  rotation of a state around the point  $\mathbf{x}$ , such that  $\Omega^2(\mathbf{x}) = I$ , are complete and orthogonal as

$$\int_{\mathbb{R}^{2n}} d\mathbf{x} \Omega(\mathbf{x}) = 2^{-n} I, \quad \text{Tr}[\Omega(\mathbf{x}) \Omega(\mathbf{y})] = 2^{-2n} \delta(\mathbf{x} - \mathbf{y}). \quad (108)$$

Similar to their expansion in displacement operators, this allows to express any Hermitian trace class operator as

$$O = \int_{\mathbb{R}^{2n}} d\mathbf{x} 2^{2n} \text{Tr}[\Omega(\mathbf{x}) O] \Omega(\mathbf{x}) = \int_{\mathbb{R}^{2n}} d\mathbf{x} 2^n W_O(\mathbf{x}) \Omega(\mathbf{x}), \quad (109)$$

where we used that  $\text{Tr}[\Omega(\mathbf{x}) O] = 2^{-n} W_O(\mathbf{x})$ . They further satisfy

$$D(\mathbf{y}) \Omega(\mathbf{x}) D(\mathbf{y}) = e^{-i4\pi \mathbf{y}^T J \mathbf{x}} \Omega(\mathbf{x}), \quad (110)$$

such that we observe that the operation

$$\Gamma(O) = \int d\mu^\Gamma(\mathbf{y}) D(\mathbf{y}) O D(\mathbf{y}) \quad (111)$$

can be understood as a map

$$W_O(\mathbf{x}) \mapsto W_O(\mathbf{x}) \left[ \int d\mu^\Gamma(\mathbf{y}) e^{-i4\pi \mathbf{y}^T J \mathbf{x}} \right]. \quad (112)$$

This form of modifying the Wigner function of the observable is similar to the modification of the characteristic function under a displacement twirl encountered in sec. 3. A striking difference is an additional factor of 2 in the exponent, which has the effect that a projective factor in eq. (112) only uniformly projects onto  $\mathcal{L}^\perp$  if we choose

$$d\mu^\Gamma(\mathbf{y}) = d\mathbf{y} \text{III}_{\mathcal{L}/2}(\mathbf{x}). \quad (113)$$

with the Dirac comb defined for lattice  $L$  as

$$\text{III}_L(\mathbf{x}) = \sum_{\mathbf{v} \in L} \delta(\mathbf{x} - \mathbf{v}). \quad (114)$$

Combining this, the operation in eq. (111) still does not generally correspond to a displacement twirl obtained from by randomly displacing an input state by vectors in  $\mathcal{L}/2$ . This, however, will be the case if we further specify the following.

---

<sup>4</sup>Note that, formally,  $\Pi_{\mathcal{L}}$  is not a physically well-defined object and needs to be replaced by appropriate physically regularized versions. It describes the correct support of the code space and satisfies  $\text{supp}(\Pi_{\mathcal{L}}^2) \subseteq \text{supp}(\Pi_{\mathcal{L}})$  in the phase space description of operators but, due to the infinite size of the lattice, is not normalized to  $\Pi_{\mathcal{L}}^2 = \Pi_{\mathcal{L}}$ . In a more careful treatment one would replace  $\Pi_{\mathcal{L}}$  by a regularized version, e.g. by introducing a probability measure over  $\mathcal{L}$ , and would consider the present treatment as limiting case. To ease the presentation, and since  $\Pi_{\mathcal{L}}$  behaves like a projector on physical states in phase space, we ignore this artifact.

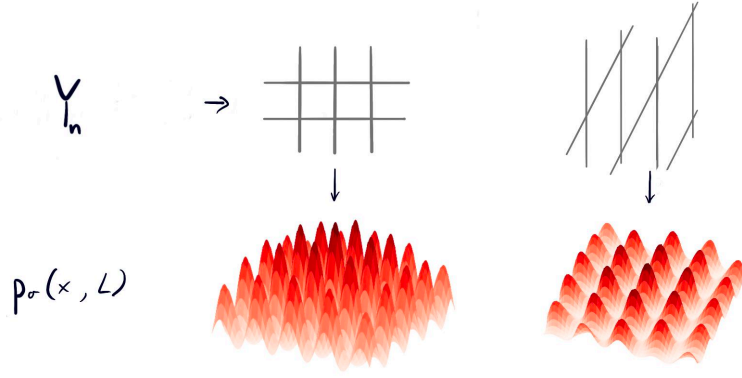


Figure 5: Phase-space distribution  $p_\sigma(\cdot, L)$  for the GKP Wigner shadow protocol, which are determined by a randomly chosen lattice  $L$ .

1.  $\mathcal{L}$  is such that  $\forall \xi, \xi' \in \mathcal{L} : \xi^T J \xi' \in 4\mathbb{Z}$ . This is such that two displacements by  $z, z' \in \mathcal{L}/2$  combine without additional phase  $D(z)D(z') = D(z + z')$ .
2. The output state is evaluated on the code space of  $\mathcal{L}$ , e.g., by performing syndrome measurements and post-selecting on zero syndrome on the input state or by projecting  $G \mapsto \Pi_{\mathcal{L}} G \Pi_{\mathcal{L}}$  onto the logical observable associated to  $G$ . Together, this has the consequence that we can replace the rightmost displacement via  $y \in \mathcal{L}/2$  in eq. (111) by  $D(y) \Pi_{\mathcal{L}} = D(y)^\dagger \Pi_{\mathcal{L}}$ , so that the map is understood as random displacement channel.

Restricted to the code space of  $\mathcal{L}$ , the infinite twirl specified in eq. (113) can be replaced by a uniform measure over the finite set  $\frac{1}{2}\mathcal{L}/\mathcal{L}$ . Observe that for a scaled GKP code with  $\mathcal{L} = d\mathcal{L}^\perp$  we have in the case  $d = 2$  precisely that a twirl over  $\frac{1}{2}\mathcal{L}/\mathcal{L}$  is logically maximally depolarizing. In consequence, the state is projected onto a contribution proportional to  $\Pi_{\mathcal{L}}$  and loses all information except for the proportionality constant. We hence focus on scaled GKP codes with  $d \in 4\mathbb{Z}$ , which also automatically satisfied point 1. above. In general, the code space projected state is specified by  $|\mathcal{L}^\perp/\mathcal{L}| = d^{2n}$  coefficients. The additional twirl over displacements in  $\frac{1}{2}\mathcal{L}/\mathcal{L}$  will only preserve characteristic function contributions in  $(\frac{1}{2}\mathcal{L})^\perp = 2\mathcal{L}^\perp \subset \mathcal{L}^\perp$ , which amounts to  $|2\mathcal{L}^\perp/\mathcal{L}| = |\mathcal{L}^\perp/2\mathcal{L}^\perp|^{-1} |\mathcal{L}^\perp/\mathcal{L}| = (d/2)^{2n}$  coefficients specifying *logical* information of the GKP code associated to  $\mathcal{L}/2$ . We summarize: for scaled GKP codes with  $d \in 4\mathbb{N}_{>1}$ , the limit  $\sigma \rightarrow 0$  of the estimator  $\overline{G}_{\mathcal{L}^\perp}^\sigma$  in eq. (102) encodes the logical expectation values of  $G$  relative to  $\rho$  for a GKP code specified by lattice  $\mathcal{L}/2$ .

## 5 Chasing shadows

In this section, we investigate how the estimator derived in the previous section together with a random choice of lattice allows to construct a shadow tomography protocol for the Hilbert space of a  $2n$ -dimensional continuous variable system. Crucially, by averaging over lattices, the reference to any particular GKP code is removed. The following presents a randomized scheme to select phase space points on which the Wigner function of an arbitrary input state are evaluated so to estimate any observable without restriction to any particular GKP code space (see fig. 5 for an illustration of such randomly selected distributions). We continue to focus on scaled GKP codes obtained from rescaling a symplectically self-dual lattices  $L = L^\perp \subseteq \mathbb{R}^{2n}$ . Such a lattice is spanned by the rows of a

symplectic generator matrix  $M \in \text{Sp}_{2n}(\mathbb{R})$  and the space of all possible symplectic lattices in dimension  $2n$  is simply parametrized by the group of symplectic matrices *up to basis transformations*  $Y_n = \text{Sp}_{2n}(\mathbb{Z}) \backslash \text{Sp}_{2n}(\mathbb{R})$ , where here the “ $\backslash$ ” operator is meant to denote a left-modulo operation since basis transformations are implemented by left-multiplication on a given generator matrix.

### 5.1 Moments of symplectic lattices

Following work by Siegel and Rogers, it has been realized by Buser and Sarnak in ref. [49] and by Kelmer and Yu in ref. [50] that the space of symplectic lattices possesses a Haar measure  $\mu(L)$ , relative to which functions of the lattice can be integrated. Define the Siegel transform of a sufficiently fast decaying function  $f : \mathbb{R}^{2n} \rightarrow \mathbb{R}$  by

$$F_f(L) = \sum_{\mathbf{v} \in L_{\text{pr}}} f(\mathbf{v}), \quad (115)$$

where  $L_{\text{pr}}$  denotes the set of primitive vectors of  $L$ , that is the set of minimal vectors that cannot be obtained by another via integer multiplication. This (infinite) set is such that every point in the lattice  $L = \mathbb{N} \otimes L_{\text{pr}}$  is uniquely reproduced a positive integer multiple of a point in  $L_{\text{pr}}$ . We also define

$$\tilde{F}_f(L) = \sum_{\mathbf{v} \in L - \{0\}} f(\mathbf{v}) = \sum_{k \in \mathbb{N}} \sum_{\mathbf{v} \in L_{\text{pr}}} f(k\mathbf{v}). \quad (116)$$

The symplectic version of Siegel’s mean value theorem, as derived in ref. [49], can thus be formulated as

$$\int_{Y_n} d\mu(L) F_f(L) = \frac{1}{\zeta(2n)} \int_{\mathbb{R}^{2n}} d\mathbf{x} f(\mathbf{x}), \quad (117)$$

with  $\zeta(z) = \sum_{k \in \mathbb{N}} k^{-z}$  being the Riemann zeta function. Applying this formula once allows to derive its perhaps more standard variant

$$\begin{aligned} \int_{Y_n} d\mu(L) \tilde{F}_f(L) &= \frac{1}{\zeta(2n)} \sum_{k \in \mathbb{N}} \int_{\mathbb{R}^{2n}} d\mathbf{x} f(k\mathbf{x}) \\ &= \int_{\mathbb{R}^{2n}} d\mathbf{x} f(\mathbf{x}). \end{aligned} \quad (118)$$

Analogous to the inner product over  $L^2(\mathbb{R}^{2n})$ ,

$$\langle f, g \rangle = \int_{\mathbb{R}^{2n}} d\mathbf{x} f(\mathbf{x})^* g(\mathbf{x}), \quad (119)$$

one can define an inner product over  $Y_n$  as

$$\langle F, G \rangle_{Y_n} = \int_{Y_n} d\mu(L) F(L)^* G(L). \quad (120)$$

In this notation, the mean value formula can be simply expressed as

$$\langle 1, F_f \rangle_{Y_n} = \frac{1}{\zeta(2n)} \langle 1, f \rangle. \quad (121)$$

In ref. [50], Kelmer and Yu have derived a formula for second moments over  $Y_n$ , which for two even and compactly supported functions  $f, g$  can be compactly written as

$$\langle F_f, F_g \rangle_{Y_n} = \frac{\langle f, 1 \rangle \langle 1, g \rangle}{\zeta(2n)^2} + \frac{2}{\zeta(2n)} (\langle f, g \rangle + \langle \iota(f), g \rangle), \quad (122)$$

where  $\iota$  is an isometry such that  $\|\iota(f)\| = \|f\| = \sqrt{\langle f, f \rangle}$ . Note that, since  $L_{\text{pr}}$  and  $L$  are even sets, any non-even function  $f$  on the LHS of this formula can simply be replaced by their even projection  $\frac{f(\mathbf{x}) + f(-\mathbf{x})}{2}$  (similar for  $g$ ): the assumption of evenness of the functions is without loss of generality. Using this expression, we show the following statement in the appendix.

**Lemma 2** (Scalar product bound). *Let  $f, g$  be two even compactly supported functions, it holds that*

$$\left| \langle \tilde{F}_f, \tilde{F}_g \rangle_{Y_n} \right| \leq |\langle f, 1 \rangle \langle 1, g \rangle| + \frac{4\zeta(n)^2}{\zeta(2n)} \|f\| \|g\|. \quad (123)$$

## 5.2 Estimating CV observables with random lattices

The mean value theorem and Lemma 2 are the key tools to analyse the behaviour of the estimator  $\tilde{G}$  when in addition to the points  $\mathbf{x} \sim p_\sigma(\mathbf{x}, L)$  being randomly sampled from a lattice-Gaussian-like distribution also the lattice  $L = L^\perp$  is chosen according to a uniformly random distribution  $\mu(L)$  over the space of lattices  $L = L^\perp$ . Under this procedure, the estimator takes the expectation value

$$\begin{aligned} \int_{Y_n} d\mu(L) \int_{\mathbb{R}^{2n}} d\mathbf{x} p_\sigma(\mathbf{x}; L) \tilde{G}(\mathbf{x}) &= \int_{Y_n} d\mu(L) \bar{G}_L \\ &= \tilde{G}(0) + \int_{\mathbb{R}^{2n}} d\mathbf{x} \tilde{G}(\mathbf{x}), \\ &= \tilde{G}(0) + \text{Tr}[\rho G] \\ &=: \bar{G} \end{aligned} \quad (124)$$

which is attained up to a small error  $\epsilon(\sigma)$  as in Lemma 1. Note, that  $\tilde{G}(0) = \text{Tr}[\hat{\pi}\rho] \text{Tr}[\hat{\pi}G]$  is simply the total photon parity of the state and observable together and can be offset by an independent estimation of the photon parity of the unknown input state.

We consider the following two scenarios. In the first, we simply sample a point  $\mathbf{x}_{i,k}$  according to the distribution  $p_\sigma(\mathbf{x}, L_k)$  for each randomly chosen lattice  $L_k$ . In this case a second moment bound is obtained as follows.

**Lemma 3** (Second moment: random lattice sampling). *Let  $\tilde{G}$  be an estimator as in Lemma 1 for an observable on a  $n$ -mode continuous variable quantum system, symplectic lattice  $L = L^\perp \subset \mathbb{R}^{2n}$  and  $d \in \mathbb{N}$  a natural number. Assume that the observable  $G$  is such that*

$$\|G\|_2^2 := \int d\mathbf{x} |G(\mathbf{x})|^2 < \infty \quad (125)$$

*is finite. It holds that*

$$\int_{Y_n} d\mu(L) \int d\mathbf{x} p_\sigma(\mathbf{x}; d^{-1/2}L) \tilde{G}(\mathbf{x})^2 \leq \left(\frac{d}{2\pi}\right)^n \{1 + c_{\sigma,d}^n\} \|G\|_2^2 := V_1(G, n, d, \sigma), \quad (126)$$

*with  $c_{\sigma,d} = \frac{2\pi d}{\sigma^2 + \sigma^{-2}} \in [0, d\pi]$*

We have  $c_{\sigma,d} \leq 1$  for  $\sigma^2 \leq \pi d \pm \sqrt{\pi^2 d^2 - 1}$ . The second scenario that we consider is such that the “inner” average is computed with high precision so that the only contributions to the variance arise from the choice of random lattices. Using similar techniques and Lemma 2, in appendix G we show that estimation of  $\bar{G}_L^\sigma$  in this setting has bounded variance as follows.



**Lemma 4** (Second moment: random lattice sampling for exact inner means). *Let  $\tilde{G}$  be an estimator as in Lemma 1 for an observable on a  $n$ -mode continuous variable quantum system with  $n > 1$ ,  $L = L^\perp \subset \mathbb{R}^{2n}$  denote a symplectic lattice and let  $d \in \mathbb{N}$  be a natural number. Assume that the observable  $G$  is such that*

$$\|G\|_1 := \int d\mathbf{x} |G(\mathbf{x})| < \infty \quad (127)$$

*is finite. Let*

$$\overline{G}_{\mathcal{L}^\perp}^\sigma = \int d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}(\mathbf{x}) \quad (128)$$

*be the estimator from Lemma 1 with  $\mathcal{L}^\perp = d^{-1/2}L$  and let  $\sigma > 0$ . It holds that*

$$\left| \int_{Y_n} d\mu(L) (\overline{G}_L^\sigma)^2 \right| \leq \left( \frac{d}{4\pi} \right)^{2n} \left[ 1 + \left( 2 + 2^{2-n} \frac{\zeta(n)^2}{\zeta(2n)} \right) c_{\sigma,d}^n + c_{\sigma,d}^{2n} \right] \|G\|_1^2 \quad (129)$$

*with  $c_{\sigma,d} = \frac{2\pi d}{\sigma^2 + \sigma^{-2}} \in [0, d\pi]$ .*

This second moment bound naturally implies a variance bound for the expectation value  $\overline{G}$  in the scenario where the inner average is taken with high precision and the only fluctuations around the mean essentially arise from the choice of random lattice. We compare this behavior with the scaling in lem. 3 for small  $\sigma < \pi d - \sqrt{\pi^2 d^2 - 1}$  with  $n$ . First note that the factor  $\zeta(n)^2/\zeta(2n) \xrightarrow{n \rightarrow \infty} 1$  converges rapidly to a constant. We can hence approximate to low order in  $c_{\sigma,d}$

$$(4\pi)^{-2n} \left[ 1 + \left( 2 + 2^{2-n} \frac{\zeta(n)^2}{\zeta(2n)} \right) c_{\sigma,d}^n + c_{\sigma,d}^{2n} \right] \|G\|_1^2 \approx \frac{1 + 2c_{\sigma,d}^n}{(4\pi)^{2n}} \|G\|_1^2. \quad (130)$$

The factor of  $(4\pi)^{-2n}$  suggests an exponentially faster decay of the prefactor as compared to the bound in lem. 3. On the other hand, for large  $\sigma$  such that  $c_{\sigma,d} > 1$ , the dominant contribution of eq. (130) scales as

$$\left( \frac{c_{\sigma,d}}{4\pi} \right)^{2n} \|G\|_1^2 \leq \left( \frac{d}{4} \right)^{2n} \|G\|_1^2, \quad (131)$$

where we finally used that it generally holds that  $c_{\sigma,d} \leq \pi d$ . This suggests that the estimator variance of the second scenario generally remains well-behaved and favorable relative to the previous setting for  $d \leq 4$ , but  $\sigma$  needs to be tuned sufficiently small in all other cases. A further relevant difference is in the observable norm appearing in these bounds. While the 2-norm, appearing in lem. 3 captures the phase-space variance of the observable  $G$ , the bound in lem. 4 depends on its one-norm, which quantifies the *Wigner-negativity* of the observable  $G$  when its Wigner function is normalized [51].

### 5.3 Protocols

We outline two shadow estimation protocols for arbitrary states and appropriate observables, also described in figs. 6 and 7. Both protocols build on sampling points  $\mathbf{x} \xleftarrow{p_\sigma(\cdot, \mathcal{L}^\perp)} \mathbb{R}^{2n}$  according to the probability density  $p_\sigma(\cdot, \mathcal{L}^\perp)$ , relative to a random lattice  $L$ , corresponding to a scaled GKP code with  $\mathcal{L} = \sqrt{d}L$  and  $\mathcal{L}^\perp = L/\sqrt{d}$ , on which the Wigner function of the state is to be evaluated, e.g., by performing displaced photon parity measurements. To achieve a high precision in this step, an extra constant overhead will be required, which we however suppress in our notation. This is the *POVM twirl* in this shadow tomography setting and the Wigner function values evaluated over these point samples form the classical shadow of this protocol. In a post-processing step, they can be combined with selected observables (i.e, their Wigner function) to estimate its expectation.

**Protocol 1.**

1. Sample  $N = CKB$  points  $\mathbf{x}_i \xleftarrow{p_\sigma(\mathbf{x}, \mathcal{L}_k^\perp)} \mathbb{R}^{2n} \supset \mathcal{L}_k^\perp = d^{-1/2} L_k \xleftarrow{\mu(L)} Y_n$ , where

$$K = 2 \log(2M/\delta)$$

$$B = \frac{1 + c_{\sigma,d}^n}{(2\pi)^n} \max_m \|G_m\|_2^2 / \tilde{\epsilon}^2$$

and  $C$  is a constant.

2. Evaluate the Wigner function of an unknown input states at these points  $\{\mathbf{x}_i\}_{i=1}^N$  and combine the output with the Wigner function of the observable  $G(\mathbf{x}_i)$  at these points.
3. For each point, return the estimate  $\tilde{G}(\mathbf{x}_i) = W_\rho(\mathbf{x}_i) G(\mathbf{x}_i)$ .

Figure 6: Random CV shadow protocol.

### 5.3.1 Protocol 1

Protocol 1 proceeds by sampling a lattice from the uniform measure  $\mu(L)$  over  $Y_n$ , the space of all symplectic lattices. For each sampled lattice, the probability density  $p_\sigma(\mathbf{x}, \mathcal{L}^\perp)$  is used to sample a point  $\mathbf{x} \in \mathbb{R}^{2n}$ , which, by construction, is close (with variance  $\sigma^2$ ) to the lattice  $\mathcal{L}^\perp$ . The experiment evaluates the Wigner function at this point and stores its value  $(\mathbf{x}, W_\rho(\mathbf{x}))$  together with the sampled point as one data point of the classical shadow. For an appropriate observable  $G$ , later estimation then proceeds by using median of means estimation on the sample set  $\{\tilde{G}(\mathbf{x}_i) = W_\rho(\mathbf{x}_i) G(\mathbf{x}_i)\}_{i=1}^N$ . The estimator variance here is captured by lemma 3, which provides an upper bound on its second moment (and hence also its variance). In this protocol the random lattice simply serves as an auxiliary object to sample points in a well-distributed manner from  $\mathbb{R}^{2n}$ .

### 5.3.2 Protocol 2

Protocol 2 proceeds similar to protocol 1, but rather than sampling only one point  $\mathbf{x} \xleftarrow{p_\sigma(\cdot, \mathcal{L}^\perp)} \mathbb{R}^{2n}$  for every sampled lattice  $L = \sqrt{d} \mathcal{L}^\perp$ , we try to obtain a good estimate for

$$\bar{G}_{\mathcal{L}^\perp}^\sigma = \int_{\mathbb{R}^{2n}} d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}(\mathbf{x}) \quad (132)$$

for each sampled lattice. According to lemma 1, this estimation step is governed by the variance upper bound (again using the second moment as proxy)

$$V_1(G; n, \sigma) = \left(\frac{d}{2\pi}\right)^n \{1 + c_{\sigma,d}^n\} \|G\|_2^2, \quad c_{\sigma,d} = \frac{2\pi d}{\sigma^2 + \sigma^{-2}} \in [0, \pi d]. \quad (133)$$

For each individual sampled lattice,  $O\left((2/\pi\sigma^2)^n \text{Tr}[G^2]\right)$  samples suffice in suppressing the additive error on  $\bar{G}_{\mathcal{L}^\perp}^\sigma$  to constant average variance, independent of the chosen lattice. The total estimation error from a sequential protocol, 1. sampling a lattice  $L \xleftarrow{\mu} Y_n$  and 2. sampling a point  $\mathbf{x} \xleftarrow{p_\sigma(\cdot, \mathcal{L}^\perp)} \mathbb{R}^{2n}$  is additive.

Together with the bound in lemma 4 we hence obtain the following.

**Corollary 1** (Variance upper bound of protocol 2). *Let  $\tilde{G}$  be an estimator as in Lemma 1 for an observable on a  $n$ -mode continuous variable quantum system where  $n > 1$  and a GKP code with  $\mathcal{L} = \sqrt{d}L$  given by symplectic lattice  $L = L^\perp \subset \mathbb{R}^{2n}$ . Assume that the corresponding observable  $G$  is such that*

$$\|G\|_1 := \int_{\mathbb{R}^{2n}} d\mathbf{x} |G(\mathbf{x})| < \infty \quad (134)$$

*is finite. Let*

$$\overline{G}_{\mathcal{L}^\perp}^\sigma = \int_{\mathbb{R}^{2n}} d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}(\mathbf{x}) \quad (135)$$

*be the estimator from Lemma 1 and let  $\sigma > 0$ . It holds that the estimator of protocol 2 has variance upper bounded by*

$$\text{Var}(\overline{G}_2) \leq \left(\frac{2}{\pi\sigma^2}\right)^n \text{Tr}[G^2] + \left(\frac{d}{4\pi}\right)^n \left[1 + \left(2 + 2^{2-n} \frac{\zeta(n)^2}{\zeta(2n)}\right) c_{\sigma,d}^n + c_{\sigma,d}^{2n}\right] \|G\|_1^2 := V_2(G; n, d, \sigma). \quad (136)$$

It is relevant to keep in mind that protocols 1 and 2 estimate the expectation value of  $\text{Tr}[\rho G]$  only up to the photon parity offset  $\tilde{G}(0)$ . Correcting for this offset, assuming it can be pre-estimated with high confidence, however has no effect on the variances of the estimators. We can combine the variance estimates with the shadow tomography protocol of Huang et al. [9], which, adapted to our setting, proceeds by a median of means strategy to quantify how the data obtained from our protocols can be used to estimate a large number of observables. We state the conclusions here and provide the relevant proofs in the appendix.

**Theorem 3** (Random lattice CV shadows). *Let  $\tilde{\epsilon}, \delta, \sigma > 0$  be small parameters,  $d \in \mathbb{N}$  an integer, let  $G_m$ ,  $m = 1, \dots, M$  be operators with finite  $\|G_m\|_2^2$  and set  $K, B$  as described in protocol 1.  $N = CKB$ , samples from the distribution of phase space points  $\{\mathbf{x}_i\}_{i=1}^N$  sampled according to protocol 1 approximate the expectation values*

$$\overline{G}_m = \text{Tr}[\rho G_m], m = 1, \dots, M \quad (137)$$

*of an arbitrary state on an  $n$ -mode continuous variable quantum system up to photon parity offsets  $\tilde{G}_m(0)$  and error  $\tilde{\epsilon}$  with probability at least  $1 - \delta$ .*

**Theorem 4** (Random lattice CV shadows with exact inner means). *Let  $\tilde{\epsilon}, \delta, \sigma > 0$  be small parameters,  $d \in \mathbb{N}$  an integer, let  $G_m$ ,  $m = 1, \dots, M$  be operators with finite  $\|G_m\|_2^2$  and set  $K, B$  as described in protocol 2.  $N = CC'KB N_P$ , samples from the distribution of phase space points  $\{\mathbf{x}_i\}_{i=1}^N$  sampled according to protocol 2 approximate the expectation values*

$$\overline{G}_m = \text{Tr}[\rho G_m], m = 1, \dots, M \quad (138)$$

*of an arbitrary state on an  $n$ -mode continuous variable quantum system up to photon parity offsets  $\tilde{G}_m(0)$  and error  $\tilde{\epsilon}$  with probability at least  $1 - \delta$ .*

These protocols describe particularly structured ways to sample phase space points  $\mathbf{x} \in \mathbb{R}^{2n}$  to determine Wigner shadows. The significance of these protocols lies in the properties of the Haar measure over the space of symplectic lattices, which allow variances of the resulting estimation procedure to be well-quantified with minimal assumptions on the state and observables. A further feature is that the parameter  $\sigma$  can serve as an

**Protocol 2.**

1. Sample  $N_L = CKB$  lattices  $\mathcal{L}_k = \sqrt{d}L_k$ , where  $L_k \xleftarrow{\mu(L)} Y_n$  is sampled uniformly and, where

$$K = 2 \log(2M/\delta),$$

$$B = (4\pi)^{-2n} \left[ 1 + \left( 2 + 2^{2-n} \frac{\zeta(n)^2}{\zeta(2n)} \right) c_{\sigma,d}^n + c_{\sigma,d}^{2n} \right] \max_m \|G_m\|_1^2 / \epsilon^2,$$

with  $C$  a constant. For each lattice sample

$$N_P = C' \left( 2/\pi\sigma^2 \right)^n \max_m \text{Tr} [G_m^2] \quad (139)$$

points  $\mathbf{x}_{kN_P+i} \xleftarrow{p_\sigma(\mathbf{x}, \mathcal{L}_k^\perp)} \mathbb{R}^{2n}$ ,  $k = 0, \dots, N_L - 1$ ,  $i = 0, \dots, N_P - 1$ , where  $C'$  is another constant.

2. Evaluate the Wigner function of an unknown input states at these points  $\{\mathbf{x}_i\}_{i=1}^{N_L N_P}$ .
3. To estimate an observable, combine the output with the Wigner function of the observable  $G(\mathbf{x}_i)$  at these points.
4. For each  $k = 0, \dots, N_L - 1$ , return the estimate

$$\tilde{G}_{\mathcal{L}^\perp, k}^\sigma = N_P^{-1} \sum_{i=0}^{N_P-1} W_\rho(\mathbf{x}_{kN_P+i}) G(\mathbf{x}_{kN_P+i}). \quad (140)$$

Figure 7: Random CV shadow protocol.

interesting dial on this protocol. While, in the limit  $\sigma \rightarrow \infty$ , the distribution  $p_\sigma$  simply becomes increasingly flat, which may also be a choice in a naive attempt to sample points from phase space under physical assumptions on state and/or observable. The opposite limit  $\sigma \rightarrow 0$  is instructive in that the distribution  $p_\sigma(\cdot, \mathcal{L}^\perp)$  tends to a narrow Gaussian distribution around the lattice points of  $\mathcal{L}^\perp$ . Such distributions have significance in cryptography [52], and more importantly, resemble the structure of GKP states [3]. This motivates to investigate how such resource state, or their induced ability to perform GKP error correction [3, 53] can serve as a tomographic resource.

In sec. 4.6.2 we have discussed the interpretation of the estimator

$$\overline{G}_{\mathcal{L}^\perp}^\sigma = \int_{\mathbb{R}^{2n}} d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}(\mathbf{x}) \xrightarrow{\sigma \rightarrow 0} \sum_{\boldsymbol{\xi}^\perp \in \mathcal{L}^\perp} W_\rho(\boldsymbol{\xi}^\perp) G(\boldsymbol{\xi}^\perp), \quad (141)$$

in the  $\sigma \rightarrow 0$  limit as the logical expectation value of the observable  $G$  under a GKP code specified by lattice  $\mathcal{L}/2$ . To obtain this perspective we assumed  $\mathcal{L}$  to correspond to a scaled GKP code  $\mathcal{L} = d\mathcal{L}^\perp$  with scaling parameter  $d \in 4\mathbb{N}_{>1}$ .

Comparing to lemma. 1, in this setting we observe that idealized GKP error correction allows to estimate the inner means  $\overline{G}_{\mathcal{L}^\perp}$  of protocol 2 exactly, such that the ability to perform code space projections (or implementing optimal GKP quantum error correction) can be understood as a resource equivalent to additional sample complexity of asymptotic size  $\lim_{\sigma \rightarrow 0} O(\exp((2/\pi\sigma^2)^n))$ , which is otherwise necessary to obtain an estimate for  $\overline{G}_L^\sigma$  with perfect confidence. While we do not further explore this here, it would be interesting to investigate whether a physical protocol can be designed that estimates  $G_L^\sigma$  for  $\sigma > 0$ .

## 6 Further related work

It is a valid question to ask whether involved procedures as designed in this work really are necessary to design a shadow protocol for observables on a CV system. Many alternative strategies seem equally valid, in particular through alternative means to regularize the Hilbert space to an effectively finite-dimensional system. This is the approach that has been taken by Iosue et al. and Ghandari et al. in refs. [16, 17]. A related technical ingredient to draw random unitaries on a CV Hilbert space is also presented in ref. [54]. The interesting observation to be taken away from this section is that it is in fact possible to interpolate between *logical* shadow tomography of a discrete quantum system embedded in a CV Hilbert space and *physical* shadow tomography applied to the full CV Hilbert space. The bounds derived in this section explain exactly the cost of this interpolation. This property appears to be special to GKP codes and is implied by the fact that the group algebra of displacement operators defining the GKP stabilizer group together by their real powers form a complete operator basis for the CV Hilbert space and the existence of a Haar measure over the possible codes. These are properties not shared by other bosonic quantum error correcting codes that we know of where such an interpolation does not seem to be possible. Note that in ref. [55] an information theoretic bound for learning CV quantum states has been derived, which we recall.

**Theorem 5** (Obstructions against general quantum state tomography [55]). *Let  $\rho$  be an unknown quantum state on  $n$  bosonic modes satisfying an energy constraint  $\text{Tr}(\hat{N}\rho) \leq n\overline{N}$  for some absolute constant  $\overline{N}$ . Then the number of copies of  $\rho$  required to perform quantum state tomography with precision  $\varepsilon$  in trace distance has to scale at least as  $(\Theta(\overline{N}/\varepsilon))^{2n}$ .*

From the sample overhead scaling in Theorem 3, we clearly see that our protocol appears consistent with this bound, scaling both exponential in the system size as well as

the lattice approximation parameter  $\sigma^{-1}$ . We can model an average energy constraint on an input state by assuming it to be in the support of a Gaussian regularizer  $\hat{R} = e^{-\sigma \hat{N}}$ , which constrains the typical support of any input state to be within a phase-space radius of

$$r = \Omega\left(\sqrt{n\sigma^{-1}}\right). \quad (142)$$

In a rough estimate, we relate the typical support radius of such a state in phase space  $r^2 \propto n\bar{N}$  to its average photon number, which conversely leads us to the estimate that, given a state with average photon number constrained as above, sampling the Wigner function from a probability distribution with Gaussian envelope with variance  $\sigma^{-1} \propto \bar{N}$  suffices to cover the essential support of the state. With this estimate the theorem above thus expects a scaling of the sample overhead with  $\sigma^{-2n} \propto N_P$ , stemming (for small  $\sigma$ ) mainly from the “inner” estimation procedure. In contrast, when the inner estimator  $\overline{G}_L^\sigma$  is obtained with high confidence, the average over different lattices only incurs an overhead growing polynomially in  $c_{\sigma,d} \approx 2\pi d\sigma^2$ , while remaining exponential in  $n$  for sufficiently large choice of parameter  $d$ .

In ref. [56] it was moreover shown that the space of GKP codes yields a rigged CV 2-design [16] and that this property also allows to derive CV shadow tomography protocols. The key concept of interest presented here is that our scheme allows to interpolate between encoded logical information relative to a random code and physical tomographic data of the CV system and is used to construct an explicit randomized Wigner tomography protocol.

## 7 Discussion and outlook

In this article, we have developed the toolbox of shadow tomography for a continuous variable system relative to GKP codes, which themselves possess an intrinsic infinite structure and so that the infinitude of the CV phase space can be matched. Our analysis and bounds on logical shadow tomography for GKP codes is particularly useful to design and test logical properties of experimental realizations of GKP code states such as pursued for *GKP-measurement based quantum computations* (MBQC) [57, 58] or realizations via superconducting architectures, where a photon parity measurement is inexpensive [21, 59]. The twirling-centric perspective on shadow tomography highlighted in this manuscript complements perspectives obtained using the language of generalized POVMs and frame theory developed in refs. [13–15, 56]. Our treatment exploits that this perspective neatly captures how encoded logical information presents itself in a physical POVM.

The logical channel twirl of a bosonic POVM introduced here is a powerful technique in its own right, as the space of bosonic POVMs is plentiful and one often does not have native access to a POVM tailored to the observable one desires to measure. As we have discussed, one may, however, tailor a given POVM to one’s needs by suitable channel twirling to generate classical snapshots that capture relevant information of a given quantum state. It would be interesting to investigate how well properties of quantum states can be learned for POVMs strongly different from the measurement of interest by combining our twirling technique with a classical learning strategy and explore where the boundaries of such schemes lie. We have restricted our attention to simple, idealized, models of CV POVMs and note that applying our methods to more accurate models of physical POVMs for tomography would be an interesting and potentially useful line of future work. Finally, we note that the projective nature of the twirling allows to adapt our base protocol to noisy measurement channels by following the strategy in ref. [11]. In ref. [11] the authors extract the effective logical channel parameters in a calibration experiment using known states

prior to applying the shadow tomography protocol with suitably adapted postprocessing. This technique is readily applicable to our setup as well.

The probabilistic state decomposition derived from heterodyne measurements and displacement twirling relative to a GKP code also warrants further investigation. This protocol can be understood as a “black-box” scheme that produces relevant Gaussian samples from an arbitrary state in the input such that logical expectation values are aligned. Since Gaussian states are computationally easy to tract through Gaussian evolutions, this tool may find application in the simulation of realistically producible states with a priori unknown decomposition into Gaussian states as they are evolved under Gaussian evolutions which is relevant in the design of large scale experiments with the GKP code [44]. It would be interesting to investigate *lower bounds* for this kind of decomposition: how many Gaussian states need to be mixed in order to emulate the logic of an encoded (GKP) state? This question, here motivated by our protocol and the proposed applications to simulation, motivate the general exploration of effective logical properties of easily accessible physical states which need not be restricted to the representation of encoded quantum states. It would further be interesting to understand what kind of logical dynamics can be effectively generated – in expectation – via Gaussian quantum channels.

Finally, we view the extension of the logical GKP shadow to a full-fledged CV shadow discussed in sec. 5 as the most interesting contribution presented here. Here we proposed a random Wigner tomography scheme, where the statistical tools used in shadow tomography [10] translate into a genuine continuous variable setting. We showed how logical expectation values of observables relative to a random GKP code can be used to reconstruct the full CV expectation value of said observable. In this part the application of tools from the theory of random lattices were the key element. It is interesting that random coding techniques are useful for (continuous variable) shadow tomography, which is an observation similarly made in ref. [56].

A natural question is whether the techniques investigated here also apply to other quantum error correcting codes. At its core, the relevant ingredients to our base protocol are 1. a physical POVM that is complete on the code subspace, 2. a unitary realization of the logical Clifford group, 3. a measure over the set of representatives of the logical Clifford group. In the presented construction we used a measure that limits to a uniform distribution over *all* representatives by including displacements over the full (infinite) stabilizer group. This choice ensured that the twirled POVM indeed results into a logical depolarizing channel without logical coherent terms. While this aligns with the construction presented in ref. [9], this choice is not strictly necessary: any reduced set of Clifford representatives is expected to yield an invertible twirled POVM.

We hence expect our protocol to also apply to other bosonic codes such as the cat- [60] or binomial codes [61]. We however do not expect the extrapolation to a physical shadow tomography protocol based on random codes to generalize as easily to other bosonic codes. For this protocol, our construction relies in particular on the mean value formula for symplectic lattices. It is conceivable that a similar effect can be obtained using quantum spherical codes [62] by leveraging their intrinsic relation to spherical designs, which would be an interesting item of future research. Another route would be to further investigate the extent to which the presented ideas are applicable to discrete variable systems, e.g., whether logical shadow tomography relative to random qubit-QECCs also allows to estimate properties of the full physical system. Here, it would be interesting to investigate the extractable information from elements of random stabilizer or subsystem codes on a multi-qubit quantum system through the lens of shadow tomography, where we speculate that techniques similar to those presented here may be applicable. The toolbox developed



here allows for a wide range of generalizations and possible applications, for which we hope this work stimulates curiosity.

## Acknowledgments

We are grateful to V. Albert, C. Bertoni, A. Ciani, J. Haferkamp, J. Iosue, R. Kueng, J. Magdalena de la Fuente, Y. Teng and N. Walk for many inspiring and helpful discussions. JC thanks in particular C. Bertoni for sharing his toolbox of ideas on statistical bounds and V. Albert and J. Iosue for discussions on the relation between the techniques discussed in sec. 5 and CV state designs. We further thank the anonymous referees for detailed and constructive feedback on this manuscript. JC and JE gratefully acknowledge support from the BMFTR (RealistiQ, MUNIQC-Atoms, PhoQuant, QPIC-1, QSolid, PasQuops), the DFG (CRC 183, SPP 2541, BoLaCo), the Munich Quantum Valley (K-8), the ERC (DebuQC), the Cluster of Excellence ML4Q, Quantum Berlin as well as the Einstein Research Unit on quantum devices.

## References

- [1] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, Bonilla A. J. P., N. Maskara, I. Cong, X. Gao, Pedro Sales R., T. Karolyshyn, G. Semeghini, M. J. Gullans, M. Greiner, V. Vuletić, and M. D. Lukin. “Logical quantum processor based on reconfigurable atom arrays”. *Nature* **626**, 58–65 (2023).
- [2] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret. “Real-time quantum error correction beyond break-even”. *Nature* **616**, 50–55 (2023).
- [3] D. Gottesman, A. Kitaev, and J. Preskill. “Encoding a qubit in an oscillator”. *Phys. Rev. A* **64**, 012310 (2001).
- [4] Shunya Konno, Warit Asavanant, Fumiya Hanamura, Hironari Nagayoshi, Kosuke Fukui, Atsushi Sakaguchi, Ryuhoh Ide, Fumihiro China, Masahiro Yabuno, Shigehito Miki, Hirotaka Terai, Kan Takase, Mamoru Endo, Petr Marek, Radim Filip, Peter van Loock, and Akira Furusawa. “Propagating Gottesman-Kitaev-Preskill states encoded in an optical oscillator” (2023). [arXiv:2309.02306](https://arxiv.org/abs/2309.02306).
- [5] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles. “Variational quantum algorithms”. *Nature Rev. Phys.* **3**, 625–644 (2021).
- [6] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, et al. “Quantum supremacy using a programmable superconducting processor”. *Nature* **574**, 505–510 (2019).
- [7] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi. “Quantum certification and benchmarking”. *Nature Rev. Phys.* **2**, 382–390 (2020).
- [8] S. Aaronson. “Shadow tomography of quantum states” (2018). [arXiv:1711.01053](https://arxiv.org/abs/1711.01053).
- [9] H.-Y. Huang, R. Kueng, and J. Preskill. “Predicting many properties of a quantum system from very few measurements”. *Nature Phys.* **16**, 1050–1057 (2020).
- [10] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller. “The randomized measurement toolbox”. *Nature Rev. Phys.* **5**, 9–24 (2023).

- [11] S. Chen, W. Yu, P. Zeng, and S. T. Flammia. “Robust shadow estimation”. *PRX Quantum* **2**, 030348 (2021).
- [12] Dax Enshan Koh and Sabee Grewal. “Classical Shadows With Noise”. *Quantum* **6**, 776 (2022).
- [13] L. Innocenti, S. Lorenzo, I. Palmisano, F. Albarelli, A. Ferraro, M. Paternostro, and G. M. Palma. “Shadow tomography on general measurement frames”. *PRX Quantum* **4**, 040328 (2023).
- [14] A. Acharya, S. Saha, and A. M. Sengupta. “Shadow tomography based on informationally complete positive operator-valued measure”. *Phys. Rev. A* **104**, 052418 (2021).
- [15] H. C. Nguyen, J. L. Bönsel, J. Steinberg, and O. Gühne. “Optimizing shadow tomography with generalized measurements”. *Phys. Rev. Lett.* **129**, 220502 (2022).
- [16] J. T. Iosue, K. Sharma, M. J. Gullans, and V. V. Albert. “Continuous-variable quantum state designs: theory and applications” (2022). [arXiv:2211.05127](https://arxiv.org/abs/2211.05127).
- [17] S. Gandhari, V. V. Albert, T. Gerrits, J. M. Taylor, and M. J. Gullans. “Continuous-variable shadow tomography” (2022). [arXiv:2211.05149](https://arxiv.org/abs/2211.05149).
- [18] Takaya Matsuura, Hayata Yamasaki, and Masato Koashi. “Equivalence of approximate Gottesman-Kitaev-Preskill codes”. *Phys. Rev. A* **102**, 032408 (2020).
- [19] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T.C. Ralph, J.H. Shapiro, and S. Lloyd. “Gaussian quantum information”. *Rev. Mod. Phys.* **84**, 621–669 (2012).
- [20] Christopher C. Gerry and Peter L. Knight. “Introductory quantum optics”. Cambridge University Press. (2023). 2 edition. url: <https://doi.org/10.1017/CB09780511791239>.
- [21] B. M. Terhal, J. Conrad, and C. Vuillot. “Towards scalable bosonic quantum error correction”. *Quant. Sc. Tech.* **5**, 043001 (2020).
- [22] J. Eisert and M. B. Plenio. “Introduction to the basics of entanglement theory in continuous-variable systems”. *Int. J. Quant. Inf.* **1**, 479 (2003).
- [23] J. Conrad, J. Eisert, and F. Arzani. “Gottesman-Kitaev-Preskill codes: A lattice perspective”. *Quantum* **6**, 648 (2022).
- [24] J. Conrad, A. G. Burchards, and S. T. Flammia. “Lattices, gates, and curves: GKP codes as a Rosetta stone” (2024). [arXiv:2407.03270](https://arxiv.org/abs/2407.03270).
- [25] Ansgar G. Burchards, Steven T. Flammia, and Jonathan Conrad. “Fiber bundle fault tolerance of GKP codes” (2024). [arXiv:2410.07332](https://arxiv.org/abs/2410.07332).
- [26] B. Royer, S. Singh, and S. M. Girvin. “Encoding qubits in multimode grid states”. *PRX Quantum* **3**, 010335 (2022).
- [27] A. L. Grimsmo, J. Combes, and B. Q. Baragiola. “Quantum computing with rotation-symmetric bosonic codes”. *Phys. Rev. X* **10**, 011058 (2020).
- [28] J. Conrad. “Twirling and Hamiltonian engineering via dynamical decoupling for Gottesman-Kitaev-Preskill quantum computing”. *Phys. Rev. A* **103**, 022404 (2021).
- [29] Paolo Zanardi. “Symmetrizing evolutions”. *Phys. Lett. A* **258**, 77–82 (1999).
- [30] R. F. Werner. “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model”. *Phys. Rev. A* **40**, 4277–4281 (1989).
- [31] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters. “Purification of noisy entanglement and faithful teleportation via noisy channels”. *Phys. Rev. Lett.* **76**, 722–725 (1996).
- [32] C. Dankert, R. Cleve, J. Emerson, and E. Livine. “Exact and approximate unitary 2-designs and their application to fidelity estimation”. *Phys. Rev. A* **80**, 012304 (2009).
- [33] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal. “Quantum data hiding”. *IEEE Trans. Inf. Th.* **48**, 580–598 (2002).

- [34] D. Gross, K. Audenaert, and J. Eisert. “Evenly distributed unitaries: On the structure of unitary designs”. *J. Math. Phys.* **48**, 052104 (2007).
- [35] S. Flammia. “[scirate.com/arxiv/quant-ph/0611002](https://scirate.com/arxiv/quant-ph/0611002)”. see comment section.
- [36] E. Onorati, J. Kitzinger, J. Helsen, M. Ioannou, A. H. Werner, I. Roth, and J. Eisert. “Noise-mitigated randomized measurements and self-calibrating shadow estimation” (2024). [arXiv:2403.04751](https://arxiv.org/abs/2403.04751).
- [37] Ying Li and Simon C. Benjamin. “Efficient variational quantum simulator incorporating active error minimization”. *Phys. Rev. X* **7**, 021050 (2017).
- [38] K. Temme, S. Bravyi, and J. M. Gambetta. “Error mitigation for short-depth quantum circuits”. *Phys. Rev. Lett.* **119**, 180509 (2017).
- [39] S. Bravyi and A. Kitaev. “Universal quantum computation with ideal Clifford gates and noisy ancillas”. *Phys. Rev. A* **71**, 022316 (2005).
- [40] P. P. Varjú. “Random walks in compact groups”. *Doc. Math.* **18**, 1137 (2012).
- [41] P. Diaconis and M. Shahshahani. “On the eigenvalues of random matrices”. *J. Appl. Prob.* **31**, 49–62 (1994). url: <http://www.jstor.org/stable/3214948>.
- [42] Hong-Ye Hu, Ryan LaRose, Yi-Zhuang You, Eleanor Rieffel, and Zhihui Wang. “Logical shadow tomography: Efficient estimation of error-mitigated observables” (2022). [arXiv:2203.07263](https://arxiv.org/abs/2203.07263).
- [43] S. Cova, M. Ghioni, A. Lotito, I. Rech, and F. Zappa. “Evolution and prospects for single-photon avalanche diodes and quenching circuits”. *J. Mod. Opt.* **51**, 1267–1288 (2004).
- [44] J. Eli Bourassa, Nicolás Quesada, Ilan Tzitrin, Antal Száva, Theodor Isacsson, Josh Izaac, Krishna Kumar Sabapathy, Guillaume Dauphinais, and Ish Dhand. “Fast simulation of bosonic qubits via Gaussian functions in phase space”. *PRX Quantum* **2**, 040315 (2021).
- [45] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard. “Simulation of quantum circuits by low-rank stabilizer decompositions”. *Quantum* **3**, 181 (2019).
- [46] Kevin He, Ming Yuan, Yat Wong, Srivatsan Chakram, Alireza Seif, Liang Jiang, and David I. Schuster. “Efficient multimode wigner tomography”. *Nature Comm.* **15** (2024).
- [47] Konrad Banaszek and Krzysztof Wódkiewicz. “Direct probing of quantum phase space by photon counting”. *Phys. Rev. Lett.* **76**, 4344–4347 (1996).
- [48] J. Conway and N. Sloane. “Sphere packings, lattices and groups”. *Volume 290*. Springer, New York, NY. (1988).
- [49] P. Sarnak and P. Buser. “On the period matrix of a Riemann surface of large genus (with an Appendix by J. H. Conway and N. J. A. Sloane)”. *Inv. Math.* **117**, 27–56 (1994). url: <http://eudml.org/doc/144207>.
- [50] D. Kelmer and S. Yu. “The second moment of the Siegel transform in the space of symplectic lattices”. *Int. Math. Res. Not.* **2021**, 5825–5859 (2019).
- [51] R.L. Hudson. “When is the wigner quasi-probability density non-negative?”. *Rep. Math. Phys.* **6**, 249–252 (1974).
- [52] D. J. Bernstein, J. Buchmann, and Dahmen E. “Post-quantum cryptography”. *Springer Berlin Heidelberg*. Berlin, Heidelberg (2009).
- [53] S. Glancy and E. Knill. “Error analysis for encoding a qubit in an oscillator”. *Phys. Rev. A* **73** (2006).
- [54] C. Zhong, C. Oh, and L. Jiang. “Information transmission with continuous variable quantum erasure channels”. *Quantum* **7**, 939 (2023).
- [55] F. Anna Mele, A. Anna Mele, L. Bittel, J. Eisert, V. Giovannetti, L. Lami, L. Leone, and S. F. E. Oliviero. “Learning quantum states of continuous variable systems”. *Nature Phys.* **21**, in press (2025).

- [56] Jonathan Conrad, Joseph T. Iosue, Ansgar G. Burchards, and Victor V. Albert. “Continuous-variable designs and design-based shadow tomography from random lattices”. *Phys. Rev. Lett.* **135**, 060802 (2025).
- [57] J. E. Bourassa, R. N. Alexander, M. Vasmer, A. Patil, I. Tzitrin, T. Matsuura, D. Su, B. Q. Baragiola, S. Guha, G. Dauphinais, K. K. Sabapathy, N. C. Menicucci, and I. Dhand. “Blueprint for a scalable photonic fault-tolerant quantum computer”. *Quantum* **5**, 392 (2021).
- [58] Nicolas C. Menicucci. “Fault-tolerant measurement-based quantum computing with continuous-variable cluster states”. *Phys. Rev. Lett.* **112**, 120504 (2014).
- [59] Dany Lachance-Quirion, Marc-Antoine Lemonde, Jean Olivier Simoneau, Lucas St-Jean, Pascal Lemieux, Sara Turcotte, Wyatt Wright, Amélie Lacroix, Joëlle Fréchette-Viens, Ross Shillito, Florian Hopfmueller, Maxime Tremblay, Nicholas E. Frattini, Julien Camirand Lemyre, and Philippe St-Jean. “Autonomous quantum error correction of gottesman–kitaev–preskill states” (2023). [arXiv:2310.11400](https://arxiv.org/abs/2310.11400).
- [60] Zaki Leghtas, Gerhard Kirchmair, Brian Vlastakis, Robert J. Schoelkopf, Michel H. Devoret, and Mazyar Mirrahimi. “Hardware-efficient autonomous quantum memory protection”. *Phys. Rev. Lett.* **111**, 120501 (2013).
- [61] Marios H. Michael, Matti Silveri, R. T. Brierley, Victor V. Albert, Juha Salmilehto, Liang Jiang, and S. M. Girvin. “New class of quantum error-correcting codes for a bosonic mode”. *Phys. Rev. X* **6**, 031006 (2016).
- [62] Shubham P. Jain, Joseph T. Iosue, Alexander Barg, and Victor V. Albert. “Quantum spherical codes”. *Nature Physics* **20**, 1300–1305 (2024).
- [63] Scott Aaronson and Daniel Gottesman. “Improved simulation of stabilizer circuits”. *Phys. Rev. A* **70**, 052328 (2004).
- [64] Ketan N. Patel, Igor L. Markov, and John P. Hayes. “Optimal synthesis of linear reversible circuits”. *Quant. Inf. Comp.* **8**, 282–294 (2008).
- [65] Froilán M. Dopico and Charles R. Johnson. “Parametrization of the matrix symplectic group and applications”. *SIAM J. Matrix Ana. Appl.* **31**, 650–673 (2009).

## A Wigner function for general GKP states

In this section we derive the Wigner function for a general idealized GKP state. Let  $\rho$  be an arbitrary state and consider the projection of the state onto code space  $\bar{\rho} = \Pi_{\mathcal{L}}\rho\Pi_{\mathcal{L}}$ , which satisfies

$$\forall \xi \in \mathcal{L} : \bar{\rho} = D(\xi) \bar{\rho} = \bar{\rho} D^\dagger(\xi), \quad (143)$$

in particular this state is an operator in the commutant of the stabilizer group given by displacements in  $\mathcal{L}$  (see sec. 3). Let  $\rho = \int_{\mathbb{R}^{2n}} d\mathbf{x} \rho(\mathbf{x}) D(\mathbf{x})$ . We have in the limit of a uniform distribution  $P_{\mathcal{L}}(\xi) \rightarrow$

$$\bar{\rho} = \Pi_{\mathcal{L}}\rho\Pi_{\mathcal{L}} = \Pi_{\mathcal{L}} \sum_{\xi \in \mathcal{L}} P_{\mathcal{L}}(\xi) D(\xi) \rho D^\dagger(\xi), \quad (144)$$

where  $P_{\mathcal{L}}$  is a priori any probability measure supported on  $\mathcal{L}$  and the limiting constant is arbitrarily chosen. Using Poisson summation, we obtain

$$\bar{\rho} = |\det(\mathcal{L})| \Pi_{\mathcal{L}} \sum_{\xi^{\perp} \in \mathcal{L}^{\perp}} \rho(\xi^{\perp}) D(\xi^{\perp}) \quad (145)$$

$$= |\det(\mathcal{L})| \Pi_{\mathcal{L}} \sum_{[\xi^{\perp}] \in \mathcal{L}^{\perp}/\mathcal{L}} \bar{\rho}([\xi^{\perp}]) D([\xi^{\perp}]) \quad (146)$$

$$= |\det(\mathcal{L})| \sum_{\xi^{\perp} \in \mathcal{L}^{\perp}} \underbrace{\left[ \sum_{\xi \in \mathcal{L}} \rho(\xi^{\perp} - \xi) e^{-i\pi \xi^T J \xi^{\perp}} \right]}_{c_{\xi^{\perp}}} D(\xi^{\perp}), \quad (147)$$

where the coefficients are such that  $c_{\xi^{\perp} + \xi} = c_{\xi^{\perp}} e^{-i\pi \xi^T J \xi^{\perp}}$ . In eq. (146) we have defined  $[\cdot] : \mathcal{L} \rightarrow \mathcal{L}^{\perp}/\mathcal{L}$ , which reduces a point in  $\mathcal{L}$  to an element of a fixed choice of representatives of  $\mathcal{L}^{\perp}/\mathcal{L}$  as well as the logical Bloch coefficients

$$\bar{\rho}([\xi^{\perp}]) = \sum_{\xi \in \mathcal{L}} \rho([\xi^{\perp}] + \xi) e^{i\pi \xi^T J [\xi^{\perp}]} = \bar{\rho}([\xi^{\perp}] + \xi') \quad \forall \xi' \in \mathcal{L}. \quad (148)$$

The Wigner function associated to this state is

$$\begin{aligned} W_{\Pi_{\mathcal{L}} \rho \Pi_{\mathcal{L}}}(\mathbf{x}) &= |\det(\mathcal{L})| \sum_{\xi^{\perp} \in \mathcal{L}^{\perp}} c_{\xi^{\perp}} e^{-i2\pi \mathbf{x}^T J \xi^{\perp}} \\ &= |\det(\mathcal{L})| \sum_{[\xi^{\perp}] \in \mathcal{L}^{\perp}/\mathcal{L}} \sum_{\xi \in \mathcal{L}} c_{[\xi^{\perp}] + \xi} e^{-i2\pi \mathbf{x}^T J ([\xi^{\perp}] + \xi)} \\ &= |\det(\mathcal{L})| \sum_{[\xi^{\perp}] \in \mathcal{L}^{\perp}/\mathcal{L}} c_{[\xi^{\perp}]} e^{-i2\pi \mathbf{x}^T J [\xi^{\perp}]} \sum_{\xi \in \mathcal{L}} e^{-i2\pi (\mathbf{x} - [\xi^{\perp}]/2)^T J \xi} \\ &= \sum_{[\xi^{\perp}] \in \mathcal{L}^{\perp}/\mathcal{L}} c_{[\xi^{\perp}]} e^{-i2\pi \mathbf{x}^T J [\xi^{\perp}]} \text{III}_{\mathcal{L}^{\perp}}(\mathbf{x} - [\xi^{\perp}]/2). \end{aligned} \quad (149)$$

where again we have use Poisson summation, the Dirac comb  $\text{III}_{\mathcal{L}^{\perp}}$  defined in the main text. Eq. (149) shows that, different from the characteristic function, the Wigner function can be supported on points outside of  $\mathcal{L}^{\perp}$ . The only general restriction is that it no supports outside of  $\mathcal{L}^{\perp}/2$ .

## B Generating the symplectic group

In this section, we discuss schemes for generating all elements in the symplectic group [19, 22]. Block matrices

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathbb{Z}_d^{2n \times 2n} \quad (150)$$

are symplectic if  $S^T J S = J$ , which requires  $A^T C = C^T A$ ,  $B^T D = D^T B$  as well as  $A^T D - C^T B = I$ . In particular, we have that for  $B = C = 0$  it is symplectic if  $D = A^{-T}$  such that  $S = A \oplus A^{-T}$ . If  $A = D = I$  and  $C = 0$  ( $B = 0$ ) it becomes necessary that  $B = B^T$  is symmetric ( $C$  is symmetric). Matrices of these constrained types feature a particularly

simple structure and follow the simple multiplication rules

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_1^{-T} \end{pmatrix} \begin{pmatrix} A_2 & 0 \\ 0 & A_2^{-T} \end{pmatrix} = \begin{pmatrix} A_1 A_2 & 0 \\ 0 & (A_1 A_2)^{-T} \end{pmatrix}, \quad (151)$$

$$\begin{pmatrix} I & B_1 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & B_2 \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & B_1 + B_2 \\ 0 & I \end{pmatrix}, \quad (152)$$

$$\begin{pmatrix} I & 0 \\ C_1 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ C_2 & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ C_1 + C_2 & I \end{pmatrix}. \quad (153)$$

In this section, we show specifically – building on previous work on qubits [63, 64] – how for prime dimension  $q$ , symplectic matrices in  $\text{Sp}_{2n}(\mathbb{Z}_d)$  can be synthesized from an elementary gate set  $S = \{J_i, P_i, C_{i \rightarrow j}\}$  of such constrained block matrices consisting of the following matrices in block form, where  $\pi_i = \mathbf{e}_i \mathbf{e}_i^T$  and  $e_{i,j} = \mathbf{e}_i \mathbf{e}_j^T$ :

- The quantum Fourier transform on qudit  $i$

$$J_i = \begin{pmatrix} I - \pi_i & \pi_i \\ -\pi_i & I - \pi_i \end{pmatrix}, \quad i \in [1, n] \quad (154)$$

with  $J_i^2 = -I$ , mapping  $X_i \mapsto Z_i^{-1}$ ,  $Z_i \mapsto X_i$ ,

- the phase gate

$$P_i = \begin{pmatrix} I & 0 \\ \pi_i & I \end{pmatrix}, \quad i \in [1, n], \quad (155)$$

mapping  $X_i \mapsto X_i Z_i$  and

- the CNOT gate

$$C_{i \rightarrow j} = \begin{pmatrix} I + e_{j,i} & 0 \\ 0 & I - e_{i,j} \end{pmatrix}, \quad i \neq j \in [1, n] \quad (156)$$

that maps  $X_i \mapsto X_i X_j$ .

- The CNOT gate is of block diagonal form, and it can be shown by performing the matrix multiplication that the upper triangular elementary block matrix

$$B_{i,j} = \begin{pmatrix} I & e_{i,j} + e_{i,j} \\ 0 & I \end{pmatrix} = J_j^{-1} C_{j \rightarrow i} J_j \quad (157)$$

mapping  $Z_i \mapsto X_j Z_i$  and  $Z_j \mapsto X_i Z_j$  can be obtained by conjugating the CNOT with a Hadamard type gate. This generating set has  $|S| = 2n + n(n-1)$  elements, where the contribution  $n(n-1)$  comes from the fact that we assume all-to-all connectivity for the CNOTs in use. This set can be reduced down to a set of  $3n-1$  generators with CNOTs only between a linear number of pairs analogous to the Lickorish generators for the Dehn-twists mentioned in the main text, which however would come at the cost of needing to mediate CNOTs not included in the set via a  $O(n)$  number of those that are.

Denote sequences generated by a finite product from  $S$  as  $S^k := \{g_1, g_2, \dots, g_k, , g_i \in S\}$ . Similar to previous work on generating  $\text{Sp}_{2n}(\mathbb{Z}_2)$  we show here a result for prime dimension.

**Lemma 5** (Universality for prime dimensions). *Let  $d$  be prime. For the generating set  $S = J_i, P_i, C_{i \rightarrow j}$  defined above, we have*

$$\text{Sp}_{2n}(\mathbb{Z}_d) \subseteq S^k \quad (158)$$

for  $k = O(dn^2)$ .

That is, sequences of  $O(dn^2)$  of gates from  $S$  suffice to generate all elements in  $\text{Sp}_{2n}(d)$ .

*Proof.* It has been shown in ref. [65] that every symplectic matrix  $S \in \text{Sp}_{2n}(\mathbb{Z}_d)$ ,  $d$  prime admits a decomposition into symplectic matrices

$$S = Q \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & A^{-T} \end{pmatrix} \begin{pmatrix} I & B \\ 0 & I \end{pmatrix}, \quad (159)$$

where  $A \in \text{GL}(n, d)$  is invertible and  $C \in \text{GL}_n(d)$  and  $B \in \text{GL}_n(d)$  are symmetric and  $Q$  is a  $O(n)$  length product of the matrices  $J_i$  we have defined above. Reference [65], in fact, has shown this for the field of complex numbers  $\mathbb{C}$  but the proof carries over to any number field, such as  $\mathbb{Z}_d = \mathbb{F}_d$  for  $d$  prime. Using this decomposition, it suffices to check how each individual block matrix can be compiled from the generating set above. Using that

$$J \begin{pmatrix} I & B \\ 0 & I \end{pmatrix} J^T = \begin{pmatrix} I & 0 \\ -B & I \end{pmatrix} \quad (160)$$

together with  $J = \prod_{i=1}^n J_i$  we have that the every upper block triangular matrix can be converted to a lower block triangular one with  $O(n)$  overhead and that every block upper triangular matrix

$$\begin{pmatrix} I & B \\ 0 & I \end{pmatrix} \quad (161)$$

with  $B = B^T$  can be obtained from a  $O(dn^2)$  fold product of matrices of type  $J_i P_i J_i^T$  and  $B_{i,j}$  for their simple multiplication structure. It remains to bound the complexity of compiling the block diagonal part  $A \oplus A^{-T}$ . Note that due to the simple multiplication structure of these matrices this problem is equivalent to bounding the complexity of compiling the blocks  $A$  as generated by elements  $I + e_{j,i}$ . This is bounded using the same argument as in ref. [63], which has employed a result from Patel et al. [64], who have shown that for the underlying field  $\mathbb{Z}_2$ , an achievable lower bound is given by  $O(n^2/\log_2(n))$ . As was already noticed in ref. [64], their technique generalizes for any finite field with order  $d$ , where it yields a bound  $O(n^2/\log_d(n))$ . In total, we hence obtain a bound  $O(dn^2)$  for the length of the product from  $S$  to generate any element in  $\text{Sp}_{2n}(\mathbb{Z}_d)$ .  $\square$

## C Proof of Theorem 2

In this section, we provide a proof of Theorem 2 of the main text.

**Theorem 2** (Full representation of the state). *Let  $\mathcal{L} \subset \mathbb{R}^{2n}$  denote the lattice corresponding to a scaled GKP code on  $n$  modes that encodes  $d^n$  logical dimension and let  $\mu$  denote a measure over elements  $\text{Aut}^S(\mathcal{L}^\perp)$  forming a logical Clifford 2-design and let  $\mathcal{M} = \int dz |\Pi(z)\rangle\langle\Pi(z)|$  denote a physical POVM. Let  $|\rho\rangle\rangle$  be an arbitrary state on the  $n$ -mode Hilbert space. The state*

$$|S\rangle\rangle = N^{-1} \sum_{i=1}^N \mathcal{U}_{S,i} |\Pi(z_i)\rangle\rangle \quad (162)$$



produced by sampling  $N$  Gaussian unitary operations via the measure  $\mu$  and measurement outcomes from the Born distribution  $z_i \sim \langle \Pi(z) | \mathcal{U}_{S,i}^\dagger | \rho \rangle \rangle$  approximates the logical value of the state in Hilbert-Schmidt distance

$$d_{\text{HS}}(\text{Dec} |\rho\rangle, \text{Dec} |S\rangle) \leq \delta_{\text{HS}}^2 \quad (163)$$

with probability at least  $1 - \delta$  for

$$N \geq \frac{2d^{2n}}{\alpha^2 \delta_{\text{HS}}^2} \left[ \ln \left( \frac{2}{\delta} \right) + 2n \ln(d) \right], \quad (164)$$

where  $\alpha$  is determined by the commutation of the twirled POVM with the decoder

$$\text{Dec} \mathcal{M}^\tau = \alpha \text{Dec} + \beta |\Pi_{\mathcal{L}}\rangle. \quad (165)$$

In particular, we also have

$$\|\text{Dec} |\rho\rangle - \text{Dec} |S\rangle\|_1 \leq d^{\frac{n}{2}} \frac{\delta_{\text{HS}}}{2}. \quad (166)$$

*Proof.* The random variable  $x_i = \frac{1}{\alpha N} \langle P_\alpha | \text{Dec} \mathcal{U}_{S,i} | \Pi(z_i) \rangle \rangle$  lies in a range  $[-1/\alpha N, 1/\alpha N]$ , such that Höfdding's inequality implies that for all logical Pauli operators  $P_\alpha$ ,  $\alpha = 1, \dots, d^{2n}$  the probability of the arithmetic mean  $\langle P_\alpha | \text{Dec} | S \rangle \rangle$  to deviate from its expectation value is bounded by

$$P(|\langle P_\alpha | \text{Dec} | S \rangle \rangle - \mathbb{E}[\langle P_\alpha | \text{Dec} | S \rangle \rangle] \geq \epsilon) \leq 2^{-N\alpha^2 \epsilon^2/2}, \quad (167)$$

such that, requiring the rhs to be upper bounded by  $\delta/d^{2n}$ ,

$$N = \frac{2}{\alpha^2 \epsilon^2} \left[ \ln \left( \frac{2}{\delta} \right) + 2n \ln(d) \right] \quad (168)$$

asserts that with probability  $1 - \delta$ , the arithmetic mean recovers the true expectation value up to  $\epsilon$  accuracy. In particular, this also bounds the logical Hilbert-Schmidt distance

$$d_{\text{HS}}(\text{Dec} |\rho\rangle, \text{Dec} |S\rangle) \leq d^{2n} \epsilon^2 \quad (169)$$

with probability  $1 - \delta$ . Setting  $\delta_{\text{HS}}^2 = d^{2n} \epsilon^2$  then yields the result. Finally, eq. (166) is recovered using the Cauchy-Schwarz inequality to trade off the trace distance with the Hilbert-Schmidt distance.  $\square$

The preceding proof features non-trivial sample complexity scaling in the number of modes  $n$ . The origin of this scaling is mainly the choice of norm we make as we aim at a full logical state-reconstruction. Firstly, in eq. (168) we required that the error probability is bounded by  $\delta/d^{2n}$ . This guarantees an anti-concentration error bounded by  $\delta$  for each of the  $d^{2n}$  logical Pauli observables. The logarithmic scaling with the number of observables is consistent with the scaling obtained in ref. [9]. Similarly, bounding the Pauli expectation value for each of the  $d^{2n}$  logical Pauli operators by a value of  $\epsilon$  implies only a  $d^{2n} \epsilon$  bound on the Hilbert-Schmidt distance. Each of these contributions essentially arise from our “abuse” of the shadow tomography protocol to perform full tomography on the logical state, which requires well approximation of an exponentially large complete set of operators. Finally, we obtain a trade-off from Hilbert-Schmidt distance to trace distance.

## D Proof of Lemma 1

In this section, we prove Lemma 1.

**Lemma 1** (Random lattice point sampling). *Let  $\rho, G$  be operators such that the product of their Wigner functions  $\tilde{G}(\mathbf{x}) = W_\rho(\mathbf{x}) G(\mathbf{x})$  is well defined,  $\text{Tr}[G^2] < \infty$  is finite and further assume that  $W_\rho(\mathbf{x}), G(\mathbf{x})$  are Lipschitz-continuous, with  $\|\nabla W_\rho(\mathbf{x})\| \leq l_\rho$  and  $\|\nabla G(\mathbf{x})\| \leq l_G$ . Set  $\tilde{l} = l_\rho + l_G$ . Let  $\sigma \ll \lambda_1(\mathcal{L}^\perp)$  be a small parameter. It holds that*

$$\begin{aligned}\overline{G}_{\mathcal{L}^\perp}^\sigma &= \int d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}(\mathbf{x}) \\ &= \sum_{\boldsymbol{\xi}^\perp \in \mathcal{L}^\perp} \tilde{G}(\boldsymbol{\xi}^\perp) + \epsilon(\sigma) \\ &= \overline{G}_{\mathcal{L}^\perp} + \epsilon(\sigma)\end{aligned}\tag{170}$$

with

$$|\epsilon(\sigma)| \leq \sqrt{2}\sigma \frac{\sqrt{2}\Gamma(n + \frac{1}{2})}{\Gamma(n)} \tilde{l}.\tag{171}$$

and furthermore

$$\begin{aligned}\overline{G}_{\mathcal{L}^\perp}^{\sigma,2} &= \int d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}^2(\mathbf{x}) \\ &\leq \left(\frac{2}{\pi\sigma^2}\right)^n (\sigma) \text{Tr}[G^2].\end{aligned}\tag{172}$$

*Proof.* We have

$$\overline{G}_{\mathcal{L}^\perp}^\sigma = N_{\mathcal{L}^\perp}^{-1} \sum_{\boldsymbol{\xi}^\perp \in \mathcal{L}^\perp} e^{-\frac{\sigma^2}{2}\|\boldsymbol{\xi}^\perp\|^2} \int d\mathbf{x} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}-\boldsymbol{\xi}^\perp\|^2} \tilde{G}(\mathbf{x}).\tag{173}$$

In each summand, we have to first order

$$\int d\mathbf{x} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}-\boldsymbol{\xi}^\perp\|^2} \tilde{G}(\mathbf{x}) = \int d\mathbf{x} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}-\boldsymbol{\xi}^\perp\|^2} \tilde{G}(\boldsymbol{\xi}^\perp) + \int d\mathbf{x} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}-\boldsymbol{\xi}^\perp\|^2} (\mathbf{x} - \boldsymbol{\xi}^\perp)^T \nabla \tilde{G}(\boldsymbol{\xi}^\perp).\tag{174}$$

The first term in this expression simply evaluates to

$$\int d\mathbf{x} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}-\boldsymbol{\xi}^\perp\|^2} \tilde{G}(\boldsymbol{\xi}^\perp) = (2\pi\sigma^2)^n \tilde{G}(\boldsymbol{\xi}^\perp)\tag{175}$$

and contributes the term  $\overline{G}_{\mathcal{L}^\perp}$  to the final expression, where the normalization factor  $N_{\mathcal{L}^\perp}$  perfectly cancels out.

Using Lipschitz continuity and the boundedness of the Wigner functions, it holds that  $\|\nabla \tilde{G}(\mathbf{x})\| \leq \tilde{l}$ . Together with the Cauchy-Schwartz inequality, we can bound the second term

$$\begin{aligned}\left| \int d\mathbf{x} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}-\boldsymbol{\xi}^\perp\|^2} (\mathbf{x} - \boldsymbol{\xi}^\perp)^T \nabla \tilde{G}(\boldsymbol{\xi}^\perp) \right| &\leq \tilde{l} \int d\mathbf{x} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}-\boldsymbol{\xi}^\perp\|^2} \|\mathbf{x} - \boldsymbol{\xi}^\perp\| \\ &= \sqrt{2}\sigma \left(2\pi\sigma^2\right)^n \frac{\Gamma(n + \frac{1}{2})}{\Gamma(n)} \tilde{l}.\end{aligned}\tag{176}$$

Combining this with the normalization and sum yields an expression for the error

$$|\epsilon(\sigma)| \leq \sqrt{2}\sigma \frac{\sqrt{2}\Gamma\left(n + \frac{1}{2}\right)}{\Gamma(n)} \tilde{l}. \quad (177)$$

Finally, we use  $|W_\rho(\mathbf{x})| \leq 2^n$  and  $p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \leq (2\pi\sigma^2)^{-n}$  to obtain

$$\overline{G}_{\mathcal{L}^\perp}^{\sigma,2} = \int d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}^2(\mathbf{x}) \leq \left(\frac{2}{\pi\sigma^2}\right)^n \int d\mathbf{x} G^2(\mathbf{x}) = \left(\frac{2}{\pi\sigma^2}\right)^n \text{Tr}[G^2]. \quad (178)$$

□

## E Proof of Lemma 2

Here, we prove the scalar product bound in Lemma 2. The main ingredient to the following derivation is a scalar product formula for functions on the Hilbert space associated to the space of lattices,  $Y_n = \text{Sp}_{2n}(\mathbb{Z}) \setminus \text{Sp}_{2n}(\mathbb{R})$ , provided in ref. [50].

**Lemma 2** (Scalar product bound). *Let  $f, g$  be two even compactly supported functions, it holds that*

$$\left| \langle \tilde{F}_f, \tilde{F}_g \rangle_{Y_n} \right| \leq |\langle f, 1 \rangle \langle 1, g \rangle| + \frac{4\zeta(n)^2}{\zeta(2n)} \|f\| \|g\|. \quad (179)$$

*Proof.* Define  $f_k(\mathbf{x}) = f(k\mathbf{x})$ ,  $g_{k'}(\mathbf{x}) = g(k'\mathbf{x})$ . We have that

$$\begin{aligned} \left| \langle \tilde{F}_f, \tilde{F}_g \rangle_{Y_n} \right| &= \left| \sum_{k,k' \in \mathbb{N}} \langle F_{f_k}, F_{g_{k'}} \rangle \right| \\ &= \left| \sum_{k,k' \in \mathbb{N}} \frac{\langle f_k, 1 \rangle \langle 1, g_{k'} \rangle}{\zeta(2n)^2} + \frac{2}{\zeta(2n)} (\langle f_k, g_{k'} \rangle + \langle \iota(f_k), g_{k'} \rangle) \right| \\ &= \left| \langle f, 1 \rangle \langle 1, g \rangle + \frac{2}{\zeta(2n)} \sum_{k,k' \in \mathbb{N}} (\langle f_k, g_{k'} \rangle + \langle \iota(f_k), g_{k'} \rangle) \right| \\ &\leq \langle f, 1 \rangle \langle 1, g \rangle + \frac{4}{\zeta(2n)} \sum_{k,k' \in \mathbb{N}} \|f_k\| \|g_{k'}\|, \end{aligned} \quad (180)$$

where we repeatedly use the triangle inequality, the fact that  $\langle f_k, 1 \rangle = k^{-2n} \langle f, 1 \rangle$ ,  $\langle 1, g_{k'} \rangle = k'^{-2n} \langle 1, g \rangle$  and the last line is derived using the Cauchy-Schwartz inequality  $\langle \iota(f_k), g_{k'} \rangle \leq \|\iota(f_k)\| \|g_{k'}\| = \|f_k\| \|g_{k'}\|$ . Finally, note that it holds that

$$\sum_k \|f_k\| = \sum_k \sqrt{\int d\mathbf{x} |f(k\mathbf{x})|^2} = \sum_k k^{-n} \sqrt{\int d\mathbf{x} |f(\mathbf{x})|^2} = \zeta(n) \|f\|, \quad (181)$$

which implies the final result. □

## F Proof of Lemma 3

**Lemma 3** (Second moment: random lattice sampling). *Let  $\tilde{G}$  be an estimator as in Lemma 1 for an observable on a  $n$ -mode continuous variable quantum system, symplectic lattice  $L = L^\perp \subset \mathbb{R}^{2n}$  and  $d \in \mathbb{N}$  a natural number. Assume that the observable  $G$  is such that*

$$\|G\|_2^2 := \int d\mathbf{x} |G(\mathbf{x})|^2 < \infty \quad (182)$$

*is finite. It holds that*

$$\int_{Y_n} d\mu(L) \int d\mathbf{x} p_\sigma(\mathbf{x}; d^{-1/2}L) \tilde{G}(\mathbf{x})^2 \leq \left(\frac{d}{2\pi}\right)^n \{1 + c_{\sigma,d}^n\} \|G\|_2^2, \quad (183)$$

*with  $c_{\sigma,d} = \frac{2\pi d}{\sigma^2 + \sigma^{-2}} \in [0, d\pi]$*

*Proof.* Note that from  $|W_\rho(\mathbf{x})| \leq 2^n$  we generally have  $\tilde{G}(\mathbf{x})^2 \leq 2^{2n} G(\mathbf{x})^2$ . We can use the functional equation (Poisson summation) [23]

$$\Theta_{\mathcal{L}^\perp}(z) = \det(\mathcal{L}) \left(\frac{i}{z}\right)^n \Theta_{\mathcal{L}}\left(\frac{-1}{z}\right) \quad (184)$$

to bound, using  $\Theta_{\mathcal{L}}\left(i\frac{4\pi}{\sigma^2}\right) \geq 1$ ,

$$N_{\mathcal{L}^\perp}(\sigma) = \det(\mathcal{L}^\perp) (8\pi)^n \Theta_{\mathcal{L}}\left(i\frac{4\pi}{\sigma^2}\right) \geq \det(\mathcal{L}^\perp) (8\pi)^n. \quad (185)$$

For lattices  $\mathcal{L} = \sqrt{d}L$ , where  $L$  is symplectic and  $\det(\mathcal{L}) = d^n$  this yields the bound  $N_{\mathcal{L}^\perp}(\sigma)^{-1} \leq (d/8\pi)^n$ . We thus estimate

$$\begin{aligned} \int_{Y_n} d\mu(L) \int d\mathbf{x} p_\sigma(\mathbf{x}; d^{-1/2}L) \tilde{G}(\mathbf{x})^2 &\leq \left(\frac{d}{2\pi}\right)^n \int_{Y_n} d\mu(L) \int d\mathbf{x} \left\{e^{-\frac{1}{2\sigma^2}\|\mathbf{x}\|^2} + \tilde{F}_{f_{\mathbf{x}}}\right\} G(\mathbf{x})^2 \\ &= (d/2\pi)^n \int d\mathbf{x} \left\{e^{-\frac{1}{2\sigma^2}\|\mathbf{x}\|^2} + \langle \tilde{F}_{f_{\mathbf{x}}}, 1 \rangle_{Y_n}\right\} G(\mathbf{x})^2, \end{aligned} \quad (186)$$

where we defined  $f_{\mathbf{x}}(\mathbf{v}) = e^{-\frac{\sigma^2}{2d}\|\mathbf{v}\|^2 - \frac{1}{2\sigma^2}\|\mathbf{x} - d^{-1/2}\mathbf{v}\|^2}$ . Using the mean value formula, it holds that

$$\langle \tilde{F}_{f_{\mathbf{x}}}, 1 \rangle_{Y_n} = \int d\mathbf{v} e^{-\frac{\sigma^2}{2d}\|\mathbf{v}\|^2 - \frac{1}{2\sigma^2}\|\mathbf{x} - d^{-1/2}\mathbf{v}\|^2} = \left(\frac{2\pi d}{\sigma^2 + \sigma^{-2}}\right)^n e^{-\frac{\Sigma}{2}\|\mathbf{x}\|^2} =: c_{\sigma,d}^n e^{-\frac{\Sigma}{2}\|\mathbf{x}\|^2}, \quad (187)$$

where we have called the expression in the final bracket  $c_{\sigma,d} = \frac{2\pi d}{\sigma^2 + \sigma^{-2}}$ , which is bounded in the range  $c_{\sigma,d} \in [0, d\pi]$  and have defined  $\Sigma := \sigma^{-2} \left(1 - (1 + \sigma^4)^{-1}\right) = \frac{\sigma^2}{1 + \sigma^4} \approx \sigma^2$ . Note that  $\forall s > 0$ , it holds that

$$\int d\mathbf{x} e^{-\frac{1}{2s^2}\|\mathbf{x}\|^2} G(\mathbf{x})^2 \leq \int d\mathbf{x} G(\mathbf{x})^2 =: \|G\|_2^2. \quad (188)$$

Inserting this inequality yields the final bound

$$\int_{Y_n} d\mu(L) \int d\mathbf{x} p_\sigma(\mathbf{x}; d^{-1/2}L) \tilde{G}(\mathbf{x})^2 \leq \left(\frac{d}{2\pi}\right)^n \{1 + c_{\sigma,d}^n\} \|G\|_2^2. \quad (189)$$

□

## G Second moment bound for exact inner means

In this section, we prove the following lemma 4.

**Lemma 4** (Second moment: random lattice sampling for exact inner means). *Let  $\tilde{G}$  be an estimator as in Lemma 1 for an observable on a  $n$ -mode continuous variable quantum system with  $n > 1$ ,  $L = L^\perp \subset \mathbb{R}^{2n}$  denote a symplectic lattice and let  $d \in \mathbb{N}$  be a natural number. Assume that the observable  $G$  is such that*

$$\|G\|_1 := \int d\mathbf{x} |G(\mathbf{x})| < \infty \quad (190)$$

is finite. Let

$$\overline{G}_{\mathcal{L}^\perp}^\sigma = \int d\mathbf{x} p_\sigma(\mathbf{x}; \mathcal{L}^\perp) \tilde{G}(\mathbf{x}) \quad (191)$$

be the estimator from Lemma 1 with  $\mathcal{L}^\perp = d^{-1/2}L$  and let  $\sigma > 0$ . It holds that

$$\left| \int_{Y_n} d\mu(L) (\overline{G}_L^\sigma)^2 \right| \leq \left( \frac{d}{4\pi} \right)^{2n} \left[ 1 + \left( 2 + 2^{2-n} \frac{\zeta(n)^2}{\zeta(2n)} \right) c_{\sigma,d}^n + c_{\sigma,d}^{2n} \right] \|G\|_1^2 \quad (192)$$

with  $c_{\sigma,d} = \frac{2\pi d}{\sigma^2 + \sigma^{-2}} \in [0, d\pi]$

*Proof.* We start by computing the second moment

$$\int d\mu(L) \left[ \int d\mathbf{x} p_\sigma(\mathbf{x}; d^{-1/2}L) \tilde{G}(\mathbf{x}) \right]^2 = \int d\mu(L) \int d\mathbf{x} d\mathbf{y} p_\sigma(\mathbf{x}; d^{-1/2}L) p_\sigma(\mathbf{y}; d^{-1/2}L) \tilde{G}(\mathbf{x}) \tilde{G}(\mathbf{y}). \quad (193)$$

Swapping the order of integration, we first compute

$$\begin{aligned} \int d\mu(L) p_\sigma(\mathbf{x}; d^{-1/2}L) p_\sigma(\mathbf{y}; d^{-1/2}L) &= \int d\mu(L) N_{d^{-1/2}}(\sigma)^{-2} \left\{ e^{-\frac{1}{2\sigma^2}\|\mathbf{x}\|^2} + \tilde{F}_{f_{\mathbf{x}}} \right\} \left\{ e^{-\frac{1}{2\sigma^2}\|\mathbf{y}\|^2} + \tilde{F}_{f_{\mathbf{y}}} \right\}, \\ &\leq \left( \frac{d}{8\pi} \right)^{2n} \int d\mu(L) \left\{ e^{-\frac{1}{2\sigma^2}\|\mathbf{x}\|^2} + \tilde{F}_{f_{\mathbf{x}}} \right\} \left\{ e^{-\frac{1}{2\sigma^2}\|\mathbf{y}\|^2} + \tilde{F}_{f_{\mathbf{y}}} \right\} \\ &= \left( \frac{d}{8\pi} \right)^{2n} \left\{ e^{-\frac{1}{2\sigma^2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)} + \langle \tilde{F}_{f_{\mathbf{x}}}, 1 \rangle_{Y_n} e^{-\frac{1}{2\sigma^2}\|\mathbf{y}\|^2} + \langle 1, \tilde{F}_{f_{\mathbf{y}}} \rangle_{Y_n} \right. \\ &\quad \left. \times e^{-\frac{1}{2\sigma^2}\|\mathbf{x}\|^2} + \langle \tilde{F}_{f_{\mathbf{x}}}, \tilde{F}_{f_{\mathbf{y}}} \rangle_{Y_n} \right\} \end{aligned} \quad (194)$$

where we denote again  $f_{\mathbf{x}}(\mathbf{v}) = e^{-\frac{\sigma^2}{2d}\|\mathbf{v}\|^2 - \frac{1}{2\sigma^2}\|\mathbf{x} - d^{-1/2}\mathbf{v}\|^2}$  and have used that  $N_L(\sigma)^{-1} \leq (d/8\pi)^n$  as in the previous proof. Here, we have inserted the relations from the proof of Lemma 3. Using Lemma 2, we also obtain the bound

$$\left| \langle \tilde{F}_{f_{\mathbf{x}}}, \tilde{F}_{f_{\mathbf{y}}} \rangle_{Y_n} \right| \leq |\langle f_{\mathbf{x}}, 1 \rangle \langle 1, f_{\mathbf{y}} \rangle| + \frac{4\zeta(n)^2}{\zeta(2n)} \|f_{\mathbf{x}}\| \|f_{\mathbf{y}}\|. \quad (195)$$

We have already computed  $\langle f_{\mathbf{x}}, 1 \rangle = \langle 1, f_{\mathbf{y}} \rangle$ . Similarly, we obtain by completing the square

$$\|f_{\mathbf{x}}\|^2 = \int d\mathbf{v} e^{-\frac{\sigma^2}{d}\|\mathbf{v}\|^2 - \frac{1}{\sigma^2}\|\mathbf{x} - d^{-1/2}\mathbf{v}\|^2} = \left( \frac{c_{\sigma,d}}{2} \right)^n e^{-\Sigma\|\mathbf{x}\|^2}. \quad (196)$$

Taking everything together and using the triangle inequality gives rise to the bound

$$\begin{aligned} \left( \frac{8\pi}{d} \right)^{2n} \left| \int d\mu(L) p_\sigma(\mathbf{x}; d^{-1/2}L) p_\sigma(\mathbf{y}; d^{-1/2}L) \right| &\leq e^{-\frac{1}{2\sigma^2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)} + c_{\sigma,d}^n e^{-\frac{\Sigma}{2}\|\mathbf{x}\|^2} e^{-\frac{1}{2\sigma^2}\|\mathbf{y}\|^2} \\ &\quad + c_{\sigma,d}^n e^{-\frac{\Sigma}{2}\|\mathbf{y}\|^2} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}\|^2} + c_{\sigma,d}^{2n} e^{-\frac{\Sigma}{2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)} \\ &\quad + \frac{4\zeta(n)^2}{\zeta(2n)} \left( \frac{c_{\sigma,d}}{2} \right)^n e^{-\frac{\Sigma}{2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)}. \end{aligned} \quad (197)$$

Again we use

$$\int d\mathbf{x} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}\|^2} \tilde{G}(\mathbf{x}) \leq 2^n \int d\mathbf{x} |G(\mathbf{x})| =: 2^n \|G\|_1. \quad (198)$$

We can now combine all the above elements and obtain

$$\left| \int d\mu(L) \left[ \int d\mathbf{x} p_\sigma(\mathbf{x}; d^{-1/2}L) \tilde{G}(\mathbf{x}) \right]^2 \right| \leq \left( \frac{d}{4\pi} \right)^{2n} \left[ 1 + \left( 2 + 2^{2-n} \frac{\zeta(n)^2}{\zeta(2n)} \right) c_{\sigma,d}^n + c_{\sigma,d}^{2n} \right] \|G\|_1^2. \quad (199)$$

□

## H Proof of Theorem 3 and 4

The following proofs are essentially restatements of the strategy followed in ref. [9] with the appropriately adapted variance upper bounds.

**Theorem 3** (Random lattice CV shadows). *Let  $\tilde{\epsilon}, \delta, \sigma > 0$  be small parameters, let  $G_m, m = 1, \dots, M$  be operators with finite  $\|G_m\|_2^2$  and set  $K, B$  as described in protocol 1.*

*$N = CKB$ , samples from the distribution of phase space points  $\{\mathbf{x}_i\}_{i=1}^N$  sampled according to protocol 1 approximate the expectation values*

$$\overline{G}_m = \text{Tr}[\rho G_m], m = 1, \dots, M \quad (200)$$

*of an arbitrary state on an  $n$ -mode continuous variable quantum system up to individual photon parity offsets  $\tilde{G}_m(0)$  and error  $\tilde{\epsilon}$  with probability at least  $1 - \delta$ .*

*Proof.* We proceed by a medians of means strategy equivalent that in ref. [9]. Block the  $N = KB$  estimates into  $K$  batches each of size  $B$ . As per the result of Lemma 1 and using Chebychevs inequality the arithmetic mean  $\hat{G}_k$  of each of these batches approximates the final mean  $\overline{G}$  up to an error of size  $|\epsilon(\sigma)|$  from Lemma 1 with failure probability

$$P(|\hat{G}_k - \overline{G}| > |\epsilon(\sigma)| + \tilde{\epsilon}) \leq \frac{V_1(G; n, d, \sigma)}{B\tilde{\epsilon}^2}, \quad (201)$$

with  $V_1(G; n, \sigma)$  as in lem. 3. Choosing  $B \geq 2^{2n} \left\{ 1 + \left( \frac{2\pi}{\sigma^2} \right)^n \right\} \|G\|_2^2 / \tilde{\epsilon}^2$  and using Hoeffding's bound, the probability of deviation of the median of these estimates

$$G_{\text{MoM}} := \text{median} \{ \hat{G}_1, \dots, \hat{G}_K \} \quad (202)$$

from the real mean is bounded by

$$P(|G_{\text{MoM}} - \overline{G}| \geq |\epsilon(\sigma)| + \tilde{\epsilon}) \leq 2e^{-K/2} \quad (203)$$

for all  $\tilde{\epsilon} > 0$ . Choosing  $K = 2 \log(2M/\delta)$  suppresses the failure probability uniformly for  $M$  observables. □

**Theorem 4** (Random lattice CV shadows with exact inner means). *Let  $\tilde{\epsilon}, \delta, \sigma > 0$  be small parameters, let  $G_m, m = 1, \dots, M$  be operators with finite  $\|G_m\|_2^2$  and set  $K, B$  as described in protocol 2.  $N = CC'KB N_P$ , samples from the distribution of phase space points  $\{\mathbf{x}_i\}_{i=1}^N$  sampled according to protocol 2 approximate the expectation values*

$$\overline{G}_m = \text{Tr}[\rho G_m] + \tilde{G}_m(0), m = 1, \dots, M \quad (204)$$

*of an arbitrary state on an  $n$ -mode continuous variable quantum system up to a photon parity offset  $\tilde{G}(0)$  and error  $\tilde{\epsilon}$  with probability at least  $1 - \delta$ .*

*Proof.* We proceed by a medians of means strategy equivalent that in ref. [9]. Block the  $N = CKB$  estimates into  $K$  batches each of size  $B$ . As per the result of Lemma 1 and using Chebychevs inequality the arithmetic mean  $\hat{G}_k$  of each of these batches approximates the final mean  $\overline{G}$  up to an error of size  $|\epsilon(\sigma)|$  from Lemma 1 with failure probability

$$P\left(|\hat{G}_k - \overline{G}| > |\epsilon(\sigma)| + \tilde{\epsilon}\right) \leq \frac{V_1(G; n, d, \sigma)}{B\tilde{\epsilon}^2}, \quad (205)$$

with  $V_1(G; n, \sigma)$  as in lem. 3. Choosing  $B \geq 2^{2n} \left\{1 + \left(\frac{2\pi}{\sigma^2}\right)^n\right\} \|G\|_2^2 / \tilde{\epsilon}^2$  and using Hoeffding's bound, the probability of deviation of the median of these estimates

$$G_{\text{MoM}} := \text{median}\{\hat{G}_1, \dots, \hat{G}_K\} \quad (206)$$

from the real mean is bounded by

$$P\left(|G_{\text{MoM}} - \overline{G}| \geq |\epsilon(\sigma)| + \tilde{\epsilon}\right) \leq 2e^{-K/2} \quad (207)$$

for all  $\tilde{\epsilon} > 0$ . Choosing  $K = 2 \log(2M/\delta)$  suppresses the failure probability uniformly for  $M$  observables.  $\square$