Continuous-variable designs and design-based shadow tomography from random lattices

Jonathan Conrad,[1,*] Joseph T. Iosue,[2,3] Ansgar G. Burchards,[4] and Victor V. Albert[3]

[1] *Institute of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne CH-1015, Switzerland*
[2] *Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA*
[3] *Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, Maryland 20742, USA*
[4] *Dahlem Center for Complex Quantum Systems, Physics Department, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany*

(Dated: August 22, 2025)

We investigate state designs for continuous-variable quantum systems using the aid of lattice-like quantum states. These are code states of Gottesman-Kitaev-Preskill (GKP) codes. We show that the set of all GKP states forms a rigged continuous-variable state 2-design for an n-mode system. We use these lattice state designs to construct a continuous variable shadow tomography protocol, derive sample complexity bounds for both global and local GKP shadows under reasonable physical assumptions, and provide the physical gadgets needed to implement this protocol.

arXiv:2412.17909v3 [quant-ph] 7 Jul 2025

I. INTRODUCTION & SUMMARY

Designs [1] simplify averages over a space X by providing an approximate version, or "skeleton", of the space which can be used instead of X to simplify calculations. Designs are ubiquitous in numerical integration [2] and coding theory [3–5], and designs over the space of quantum states and the unitary group are valuable tools in quantum information processing and even black-hole physics (e.g., see Refs. [6–13]).

Designs exist on finite-dimensional spaces, but recent work [14] heralded an extension of designs to infinite-dimensional spaces of normalizable functions. The prominence of such spaces in quantum information processing, signal processing, machine learning, and quantum field theory warrants further study of such designs.

In quantum mechanics, such function spaces parameterize quantum state spaces of continuous-variable (CV) systems, e.g., photonic or phononic degrees of freedom. As modern technological developments allow for precise control of CV degrees of freedom in engineered quantum systems — and as physics is continuous at its fundamental level — the development of quantum computers embedded in natively CV systems is increasingly recognized as a promising strategy [15–17]. Design-based protocols for CV systems are a well-behaved and quantifiable way to extract information which may further aid in their classification and simulation, analogous to their qubit-based counterparts.

A design of strength t — a *t-design* — is a subset of points $\mathcal{E} \subseteq X$ such that a uniform average over the design of every degree$\leq t$ polynomial equals its average over X. Reference [14] introduced a family of appropriately defined (and necessarily non-Gaussian [18]) CV 2-designs for a space of functions on the real line, $\mathbb{R}^{d=1}$, leaving open the question of existence of natural designs for real-space functions of higher dimension d.

We answer this question in the affirmative using quantum lattice states — infinite quantum superpositions of coherent states at all points in a symplectic lattice in the CV phase space. We show that the set of such states for all displaced versions of all possible $2n$-dimensional symplectic lattices form a family of CV 2-designs *for all* n.

Quantum lattice states form codewords of the Gottesman-Kitaev-Preskill code [19] (see also Refs. [15, 20]) and the Zak basis [21, 22], and are relevant to quantum optics [23, 24], post-quantum cryptography [25, 26], signal processing [27], quantum metrology [28], and condensed-matter physics [21, 22]. Henceforth, we colloquially refer to quantum lattice states as GKP states.

Infinite families of finite-dimensional 2-designs feature prominently in a popular qubit-based tomography protocol called shadow tomography [7, 8, 29], and our results readily yield similarly powerful CV variants of both the "global" and "local" variants of this protocol. This result shows that GKP quantum error correction — realized in several experiments [20, 30–35] — is a powerful tomographic resource for CV systems, on par with Clifford circuits and Pauli measurements for qubit systems.

II. GKP STATES FORM RIGGED 2-DESIGNS

A *rigged t-design* is an ensemble of pure states $\mathcal{E}$ together with a measure $\mu_{\mathcal{E}}$ such that its constituents resolve permutation-symmetric subspaces [14], satisfying

$$\int_{\mathcal{E}} d\mu_{\mathcal{E}}(X)\,|X\rangle\langle X|^{\otimes t'} = a_{t'}\Pi_{t'}\ \forall t' \leq t\,, \tag{1}$$

with non-negative *design coefficients* $a_{t'}$, and with a_1 set to one. The adjective "*rigged*" indicates that our ensemble contains non-normalizable vectors, which are required to satisfy the above equation.

Here, $\Pi_{t'}$ is a projector onto the permutation-symmetric subspace of the tensor-product space of t factors, with each factor consisting of n modes.

We denote the ensemble of all n-mode GKP (stabilizer) states for all lattices by

$$\mathcal{X}_n = \{|\Lambda;\boldsymbol{\alpha}\rangle = D(\boldsymbol{\alpha})\,|\Lambda\rangle\,,\ \Lambda \in Y_n,\ \boldsymbol{\alpha} \in \mathcal{P}(\Lambda)\}\,. \tag{2}$$

* jonathan.conrad@epfl.ch

Here, $|\Lambda\rangle$ is an equal superposition of coherent states at all points in the symplectically self-dual lattice $\Lambda = \Lambda^\perp \subset \mathbb{R}^{2n}$. This state is displaced by all possible vectors $\boldsymbol{\alpha}$ in the lattice's fundamental domain $\mathcal{P}(\Lambda)$ — the unit cell centered at the origin — via displacement operators $D(\boldsymbol{\alpha})$. Each lattice is defined by a symplectic matrix, and the space $Y_n = \mathrm{Sp}_{2n}(\mathbb{Z}) \setminus \mathrm{Sp}_{2n}(\mathbb{R})$ of all lattices is parameterized by all such matrices, up to basis transformations done by integer-valued symplectic matrices. This space happens to be compact and comes with its own Haar measure, $\mu(\Lambda)$ [26, 36].

The core contribution of this manuscript is the following.

Theorem 1 (GKP states form a rigged 2-design)**.** *The ensemble $\mathcal{X}_n$ forms a rigged 2-design with respect to the Haar measure over the space of symplectic lattices Y_n and the uniform measure over the fundamental domain $\mathcal{P}(\Lambda)$.*

We prove this Theorem in three different ways. The key step is afforded by the following formula, which converts certain integrals over lattices to integrals over phase space.

Lemma 1 (Symplectic mean value formula [36][37, theorem 2])**.** *Let $f : \mathbb{R}^{2n} \to \mathbb{R}$ be a (Riemann) integrable function such that the following quantities converge absolutely, and let $\Lambda \in Y_n$ denote a symplectic lattice. For*

$$F(\Lambda) = \sum_{\boldsymbol{\lambda} \in \Lambda - \{0\}} f(\boldsymbol{\lambda}) \ , \tag{3}$$

it holds that

$$\int_{Y_n} d\mu(\Lambda)\, F(\Lambda) = \int_{\mathbb{R}^{2n}} d\boldsymbol{x}\, f(\boldsymbol{x}) \ . \tag{4}$$

We now summarize our proofs. From the decomposition of the CV Hilbert space into orthogonal sectors via displaced GKP states $D(\boldsymbol{\alpha})|\Lambda\rangle$ for each lattice Λ, we obtain a complete covering of phase space. This means that each sub-ensemble $(\Lambda, \boldsymbol{\alpha})$ with Λ fixed forms a 1-design. For the 2-design condition, known formulas about GKP states [38] reduce the left-hand side of Eq. (1) to a lattice integral over $\Lambda \in Y_n$. The mean-value formula allows us to substitute that integral with one over the entire phase space, and solving that integral completes the proof.

By replacing each GKP state in $\mathcal{X}_n$ with a regularized version of the state, this ensemble can further be turned into a set of *physical* (read: square-integrable and finite-energy) quantum states that conforms to the notion of *approximate* rigged designs introduced in Ref. [14]. Without this modification, however, the rigged design is still useful as it allows for the definition of a POVM useful for state tomography.

III. DESIGN-BASED SHADOW TOMOGRAPHY

Our design construction can be used to implement CV classical-shadow tomography protocols using access to GKP states and an appropriately chosen Gaussian circuit. We formulate two protocols — a global and a local GKP shadow protocol — which are CV cousins of global [7] (a.k.a. Clifford) and local [7, 8] (a.k.a. Pauli) qubit shadow protocols, respectively. We first place our protocols in the context of their qubit counterparts, and then proceed with a summary of our technical results.

Any ensemble $\mathcal{E}$ satisfying the 2-design condition, either conventional or rigged, yields the following map,

$$\begin{aligned} \mathcal{M}_{\mathcal{E}}(\rho) &= \int_{\mathcal{E}} d\mu(X)\ \langle X|\rho|X\rangle\, |X\rangle\langle X| \\ &= \frac{a_2}{2}(I + \rho) \ . \end{aligned} \tag{5}$$

This map shows that a POVM over the set of *pointer "states"* $|X\rangle$ with probability densities $\langle X|\rho|X\rangle$ contains complete information about ρ. Plugging the map into the expectation value of an observable O with repsect to ρ expresses this value in terms of a statistical average,

$$\langle O\rangle = \frac{2}{a_2} \mathop{\mathbb{E}}_{X|\rho} \langle X|O|X\rangle - \mathrm{Tr}[O] \equiv \mathop{\mathbb{E}}_{X|\rho} o(X) \ , \tag{6}$$

over estimators $o(X)$ stemming from pointers $|X\rangle$. The subscript $X|\rho$ indicates that the average is relative to the distribution over the pointers X inherited from the state ρ.

The key to classical shadows is choosing an ensemble that yields a favorable scaling of the number of samples needed for a good estimate. This sample complexity can be obtained by examining the estimator variance $\mathrm{Var}\, o(X)$.

Global shadows In the case of global qubit shadows, the ensemble is the set of all n-qubit stabilizer states, $\mathcal{E} = \mathcal{S}_n$, and the POVM reduces to a depolarizing channel with $a_2 = 2/(2^n + 1)$. Stabilizer states form a quantum state 3-design with $a_3 = 6/[(2^n + 1)(2^n + 2)]$, and the variance can be bounded in terms of the trace of O^2 and a ratio of its design coefficients (see supplemental material [39]),

$$\mathop{\mathrm{Var}}_{X|\rho} o(X) \leq \frac{2a_3}{a_2^2} \mathrm{Tr}[O^2] \ . \tag{7}$$

The design-coefficient ratio approaches 3 for large n, implying that the sample complexity depends only on the trace term.

The GKP analogue of this protocol — *global GKP shadows* — uses the set of all GKP states, $\mathcal{E} = \mathcal{X}_n$. While we do not know if GKP states form a 3-design, we show that the ensemble possesses just enough structure to yield a useful characterization of the variance. Using favorable properties of GKP states and lattice moment formulas, we obtain in Thm. 2 a similar overall scaling with $\mathrm{Tr}[O^2]$ and an additional dependence on the L^1 norm of the characteristic function of O.

Local shadows The pointers of the local qubit shadow protocol are n-fold tensor products of single-qubit stabilizer states, $\mathcal{E} = \mathcal{S}_1^{\otimes n}$. For a traceless k-qubit tensor-product observable, the variance is the kth power of the above global shadow variance for the case of a single qubit — upper bounded by $3^k\, \mathrm{Tr}[O^2]$.

Analogously, our *local GKP shadow* protocol utilizes n-fold tensor products of single-mode GKP states, $\mathcal{E} = \mathcal{X}_1^{\otimes n}$. As with its qubit counterpart, each subsystem is measured independently, and no entangled measurements are necessary.

We once again cannot use the 3-design condition to bound the variance, and instead assume a conjecture (14) about finite-energy bosonic states. Equipped with this conjecture, we obtain a similar dependence of the variance on the observable as in the qubit case (see Thm. 3).

A. Global GKP shadows

We employ medians of means (MoM) estimation [8] to implement the classical post-processing of the protocol, where the $N = KB$ samples are divided into K batches of size B, and the estimator MoM returns the median of the arithmetic means over each batch. We establish a general upper bound using a recently derived second-moment bound for functions over the space of symplectic lattices [40], which is parametrized by a bounded sequence $\{f_n\}_{n=1}^{\infty}$, defined as

$$f_1 = 1, \quad 10 \geq f_n = \frac{4\zeta(n)^2}{\zeta(2n)}, \, n > 1, \tag{8}$$

where ζ is the Riemann-zeta function. Combining this bound with the standard approach established in Ref. [8], we show the following theorem.

Theorem 2 (GKP shadow tomography)**.** *Let $\epsilon, \delta > 0$, and let $O_i = O_i^\dagger, i = 1, 2, \cdots, M$ be a set of M observables on an n-mode CV quantum system with integrable characteristic functions c_{O_i}. Let*

$$\begin{aligned} \tilde{V}_{O_i} &= \left(\|c_{O_i}\|_1 + |\mathrm{Tr}\left[O_i\right]|\right)^2 + f_n \, \mathrm{Tr}\left[O_i^2\right] \\ \tilde{V}_O &= \max_i \tilde{V}_{O_i}, \end{aligned} \tag{9}$$

and $N = KB$, with

$$B = 34\tilde{V}_O/\epsilon^2 \quad \textit{and} \quad K = 2\log\left(2M/\delta\right). \tag{10}$$

Then, N samples from the state relative to the POVMs defined by the ensemble of GKP states $\mathcal{X}_n$ suffice to approximate the expectation values of each observable O_i with probability

$$\Pr\left(\max_{i=1,\ldots,M} |\mathrm{MoM}\left[\tilde{o}_i\right] - \mathrm{Tr}\left[O_i\rho\right]| \geq \epsilon\right) \leq \delta. \tag{11}$$

The trace term $\mathrm{Tr}\left[O^2\right]$ in Eq. (D35) stems from the L^2 norm of an observable's characteristic function, $\|c_O\|_2^2 = \mathrm{Tr}\left[O^2\right] \leq \|O\|^2 \mathrm{rank}\left(O\right)$, and is also present in the qubit shadow bound. However, rather than capturing the *rank* of the observable as in the qubit case, the CV domain (where the rank need not be defined) renders this into a measure of how quickly the operator O *decays* in phase space.

The L^1 norm of the characteristic function in Eq. (D35) — not present in the qubit bound — takes a dominant role in the derived bound. For qubits, the l^1 norm of the Pauli coefficient vector of a state (the analogue of the characteristic function) serves as a quantifier for the "non-stabilizerness" of the state [41]. In the CV case, the L^1 norm of a characteristic function of a convex combination of Gaussian pure states $\rho_i = G\left(V_i, \overline{\boldsymbol{x_i}}\right)$ yields a value $\|c_\rho\|_1 \leq 2^n$ [42], such that any state whose L^1 norm exceeds this value must necessarily contain non-Gaussian components. If the 3-design property of GKP states could be shown, more stringent bounds on the estimator variance and sample complexity would follow that do not depend on the observables' characteristic functions [14].

B. Local GKP shadows under physicality conjecture

In order to bound the local GKP shadow sample complexity, we impose two additional "physicality" assumptions on the input state ρ. We assume that the state's average photon number is upper bounded by a constant, $\mathrm{Tr}[\rho\hat{N}] \leq \overline{N}$. and that the state decays exponentially in phase space [43]. Using these assumptions, we can bound the pointer expectation values $\langle X|\rho|X\rangle$ using an effective squeezing parameter [28],

$$\Delta_{\boldsymbol{\lambda}}^2(\rho) = -\frac{2}{\pi}\log\left(|c_\rho(\boldsymbol{\lambda})|\right). \tag{12}$$

We conjecture that this squeezing parameter is lower bounded by the inverse average photon number of the system as

$$\tilde{\Delta}^2(\rho) \coloneqq \min_{M \in \mathrm{Sp}_{2n}(\mathbb{R})} \min_i \Delta_{\boldsymbol{\xi}_i}^2(\rho) \tag{13}$$

$$\geq 4/\overline{N} \qquad \text{(conjecture)}, \tag{14}$$

only depending on $\overline{N}$, and where $M = (\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_{2n})^T$. This conjecture generalizes a previous observation [28] using information-theoretic bounds and corresponds to the minimal squeezing parameter achievable if all the photons of an n-mode system with average photon number $\overline{N}$ are concentrated in an individual mode. For large $\overline{N} \gg 1$, this yields the bound

$$\langle X|\rho|X\rangle \leq \left(\overline{N}/\pi\right)^{2n}. \tag{15}$$

We use this bound to show that the relevant variance parameter $\tilde{V}_{O_i}$ in Thm. 2 can be replaced by

$$\tilde{V}_{O_i} = \left(\overline{N}/\pi\right)^{2n}\left(\mathrm{Tr}\left[O_i^2\right] + \mathrm{Tr}\left[O_i\right]^2\right), \tag{16}$$

which, up to the physicality factor $\left(\overline{N}/\pi\right)^{2n}$, is equivalent to that obtained from rigged CV 3-designs in Ref. [14] and takes the comparable functional form as that of global qubit shadows [8].

Following a similar calculation that leads to the variance scaling in Eq. (16), we find that the performance of the local shadow tomography protocol derived here is not limited by the *locality* of the observable, but rather by its *in-separability*.

Theorem 3 (CV shadow tomography, local and physical)**.** *Let $\epsilon, \delta > 0$, and let $O_i = O_i^\dagger, i = 1, 2, \cdots, M$ be a set of M observables on a n-mode CV quantum system with integrable characteristic functions c_{O_i}. Assume the conjecture in Eq. (14) and that the input state has average photon number*

Method	Sample complexity ($\beta \ll 1$)
Global qubit shadows [8]	$\|O'\|_{\text{shadow}} \sim \text{Tr}\left[O'^2\right] = O\left(2^n\right)$
Global GKP shadows	$O\left(\left(\overline{N}/\pi\beta\right)^{2n}\right)$
Local qubit shadows [8]	$\|O'\|_{\text{shadow}} \sim 4^n \|O'\|_\infty = O\left(4^n\right)$
Local GKP shadows	$O\left(\left(\overline{N}/\pi\beta\right)^{2n}\right)$

TABLE I. Sample complexity scaling of qubit- and GKP based shadow tomography protocols with $\beta \ll 1$ for the thermal state observables $O = e^{-\beta\hat{N}}$ and $O' = e^{-\beta\sum_{i=1}^n \sigma_{z,i}}$. Here, $\overline{N}$ is the average energy of the underlying bosonic state.

at most $\overline{N} \gg 1$ with exponentially decaying characteristic function. Let

$$\tilde{V}_{O_i}^{\text{loc}} = \left(\overline{N}/\pi\right)^{2n} \sum_{\boldsymbol{k}\subseteq\boldsymbol{n}} \text{Tr}_{\boldsymbol{n}-\boldsymbol{k}}\left[\text{Tr}_{\boldsymbol{k}}\left[O_i\right]^2\right] \tag{17}$$

$$\tilde{V}_{O}^{\text{loc}} = \max_i \tilde{V}_{O_i}^{\text{loc}}, \tag{18}$$

and $N = BK$, with

$$B = 34\tilde{V}_O^{\text{loc}}/\epsilon^2 \quad \textit{and} \quad K = 2\log\left(2M/\delta\right). \tag{19}$$

Then, N samples from the states relative to the POVMs defined by the ensemble of GKP states $\mathcal{X}_1^n = \bigotimes_{i=k}^n D\left(\boldsymbol{\alpha}_k\right)|\Lambda_k\rangle$, where $|\Lambda_k\rangle$ is the unique GKP state corresponding to $\Lambda_k = \Lambda_k^\perp \subset \mathbb{R}^2$ and where $\boldsymbol{\alpha}_k \in \mathcal{P}\left(\Lambda_k\right)$, suffice to approximate the expectation values of each observable O_i with

$$\Pr\left(\max_{i=1,\ldots,M} \left|\text{MoM}\left[\tilde{o}_i\right] - \text{Tr}\left[O_i\rho\right]\right| \geq \epsilon\right) \leq \delta. \tag{20}$$

In words, the native estimator variance is given by the *total purity* of the observable, measured over every bi-partition $\boldsymbol{k} \subseteq \boldsymbol{n}$ given by a subset of mode labels $\boldsymbol{n}$ and its complement $\boldsymbol{n}-\boldsymbol{k}$. As long as the total purity of the observable across any bi-partition remains small, the local shadow protocol will be efficient. This is a notable difference from the sample complexity bound on local qubit shadows, which depends heavily on the tensor-product locality of the observable [8].

We test how our protocols fare on estimating an observable proportional to the thermal state of the total-energy Hamiltonian, $H = \overline{N}$, defined on both qubits and modes. We obtain similar exponential scaling with qubit/mode number n for all four protocol combinations — local and global, qubit and GKP — summarized in Table I. In the supplemental material [39], we analyze Hamiltonians with a pertubative polynomial part $\hat{H} = \hat{N} + \epsilon P\left(\hat{\boldsymbol{x}}\right)$, with $P\left(\hat{\boldsymbol{x}}\right)$ a polynomial expression in the quadrature operators, and show how the corresponding estimator values $|\langle X|O|X\rangle|$ are bounded by functions related to so-called weighted theta series [44]. The general techniques outlined there show how these and other nontrivial observables can be examined with our protocol.

IV. IMPLEMENTING SHADOW TOMOGRAPHY

Given an input state ρ, our global shadow protocol proceeds as follows, with the local version substituting $\mathcal{X}_n \to \mathcal{X}_1^{\otimes n}$. To obtain samples from the effective state $\mathcal{M}_{\mathcal{X}_n}\left(\rho\right)$ (5),

1. sample a lattice $\Lambda \in Y_n$ specified by a generator matrix $M = \left(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{2n}\right)^T$ [38];
2. measure each displacement operator $\left\{D\left(\boldsymbol{\xi}_i\right)\right\}_{i=1}^{2n}$ by performing GKP syndrome extraction.

In this strategy, for each sampled lattice Λ, the displacement operator measurement returns an outcome determined by a vector $\boldsymbol{\alpha} \in \mathcal{P}\left(\Lambda\right)$ with probability

$$P_\Lambda\left(\boldsymbol{\alpha}\right) = \left\langle\Lambda\right|D^\dagger\left(\boldsymbol{\alpha}\right)\rho D\left(\boldsymbol{\alpha}\right)\left|\Lambda\right\rangle, \tag{21}$$

which is precisely the expression $\langle X|\rho|X\rangle$ that determines the distribution over pointers $X = \left(\Lambda, \boldsymbol{\alpha}\right)$ that we are seeking. To understand how this expression arises, note that by symplectic self-duality of $\Lambda \in Y_n$, the POVM element $D\left(\boldsymbol{\alpha}\right)|\Lambda\rangle\langle\Lambda|D^\dagger\left(\boldsymbol{\alpha}\right)$ describes the projector onto the $\boldsymbol{\alpha}$-"error sector" of the GKP code specified by Λ. This sector is labeled by syndrome $MJ\boldsymbol{\alpha}$ for $M = \left(\boldsymbol{\xi}_1, ..\boldsymbol{\xi}_{2n}\right)^T$. Sampling over uniformly random $\Lambda \in Y_n$, we find that the shadow can be obtained via standard syndrome extraction for the GKP code. This strategy is analogous to the strategy employed in Ref. [8] and can be implemented with standard qubit-assisted or GKP-state-assisted techniques to implement GKP stabilizer measurements. We depict the corresponding circuits for these in Fig. 3 and provide a detailed analysis in the supplemental material [39].

The qubit-assisted implementation is an experimentally widely used protocol, based on dominant-bit phase estimation [20, 30, 31, 33, 45], and we show how the probabilities $P_\Lambda\left(\boldsymbol{\alpha}\right)$ can be directly estimated from Pauli expectation values on the qubit. This strategy leads to an effective compression of the shadow tomography protocol, and we show how statistical techniques can be used to bound the sample overhead to estimate these probabilities.

Qubit-assisted protocols only yield binary outcomes in each execution of the circuit. An alternative strategy with higher information gain is to execute a continuous-valued phase estimation routine using an auxiliary GKP mode. Here we introduce new circuits that use a single GKP auxiliary mode to measure multimode GKP stabilizers, akin to the displacement sensor routine developed in Ref. [28]. This strategy is, to our knowledge, novel, and we show how the Fourier coefficients of the measurement statistics encode the relevant displacement expectation values. We expect the circuits presented here and their analysis to be of general use also outside of GKP-based shadow tomography.

Finally, we note that, given access to a resource of GKP states $|X\rangle \in \mathcal{X}_n$, one can also directly estimate the probabilities $\langle X|\rho|X\rangle$ by performing a SWAP test [46–48] between the resource state $|X\rangle$ and the input ρ. This further strenghtens the argument that GKP states form powerful resources for quantum tomography.

V. DISCUSSION & OUTLOOK

We show that quantum lattice states (code words of all GKP codes [19]) form a continuous-variable 2-design (in the sense

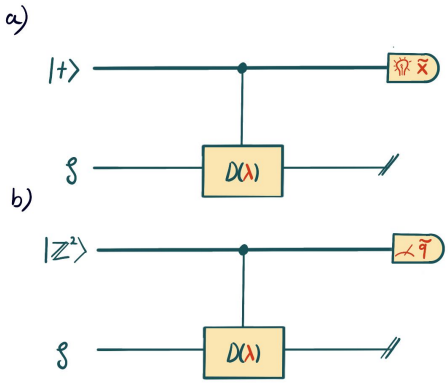


FIG. 1. Qubit-assisted and GKP-state-assisted stabilizer measurement circuits that implement the shadow tomography protocol. a) An implementation using a qubit-controlled displacement, where the qubit expectation value, $\langle \tilde{X} \rangle = P_+ - P_- = \mathrm{Re}\left\{\mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right)\rho\right]\right\}$, encodes the sampling probabilities. b) An auxiliary square-lattice GKP state $|\mathbb{Z}^2\rangle$ together with a mode-controlled displacement operation (a Gaussian unitary gate) is used to sample the probability distribution $P\left(\tilde{p}\right)$ of momenta on the auxiliary mode, whose Fourier series coefficients encode all values $\mathrm{Tr}\left[D\left(\ell\boldsymbol{\lambda}\right)\rho\right]$, $\ell \in \mathbb{Z}$.

of Ref. [14]). This puts them on a similar footing with stabilizer states of qudits of dimension $q > 2$, which also form a 2-design. We use this result to introduce global *and* local GKP-based versions of qubit shadow tomography, present recipes on how to implement them, and derive bounds to understand their performance.

Our key technical tools are two formulas from lattice theory — a mean-value formula and a second-moment bound, both applied over the space of symplectic lattices. Ensembles of random lattices and more general quantum states that mimic lattice structure are relevant objects in classical (post-quantum) cryptography [25, 26]. We believe that the presented applications of these powerful tools are only the beginning of their use in quantum information processing.

For example, it would be interesting to build a better understanding of average-case hardness of the learning with errors (LWE) problem [25], and how exactly average case hardness of the underlying problem relates to the ability of the corresponding lattice ensemble to realize the mean value formula. Any such investigation may also shed light on which discrete subsets of the space of lattices constitute practical sample sets. Provided that a lattice-based cryptosystem was used in Ref. [26] to construct a suitable distribution of lattices, we can speculate whether average-case hardness of typical lattice problems over such distributions could be such a sufficient requirement.

For another example, bounds similar to the utilized second-moment formula have previously been derived for higher moments of euclidean lattices [49]. While they have yet to be derived for the symplectic case, it would be interesting to understand any information theoretic implications.

Our global GKP shadow sample complexity bound features the L^1 norm, a measure of non-Gaussianity of the state [50]. This is conceptually interesting, as the shadow reflects an ensemble of "classical snapshots" of a quantum state. A similar measure for observables on qubit systems is provided by the stabilizer entropy [51], and it would be interesting if the L^1 norm of the characteristic function of observables could be endowed with similar operational meaning as their qubit analogues [52].

The shadow bounds, derived under physicality assumptions and imposed locality, are similar to the bounds derived for qubit-based protocols in terms of their dependency on the trace of the squared observable. In stark contrast to such bounds, however, our GKP shadow bound does not directly depend on an observable's *locality*, but merely on its native variance across bi-partitions, permitting the observable to maintain spatially very non-local support as long as it is of sufficiently fast phase-space decay. In these derivations, we have conjectured a lower bound on the effective squeezing measure for arbitrary states scaling with the inverse average photon number. We note that, through their operational interpretation as variances in a displacement-estimation problem [28], the proof of our conjecture would establish lower bounds for the tomography of displacements, including entangled strategies. It may be valuable to base a resource theory of non-Gaussian quantum states directly off such measures.

It is interesting to note that the sample complexity scaling derived for GKP shadows, albeit assuming a conjecture, exhibits the same scaling in the properties of the observable as the expected scaling for a shadow tomography protocol derived for rigged 3-designs in Ref. [14]. While we have been unable to answer the question whether our ensemble of GKP states forms a 3-design, this is an encouraging factor. Even if the GKP ensemble were to fail to provide a 3-design property, it would be interesting to further investigate the extent to which this failure occurs [53].

ACKNOWLEDGMENTS

J. C. and A. B. thank the University of Maryland and its members for their hospitality, where the essential steps presented in this work were developed and J. C. further thanks Oma Eichler for providing the environment where the manuscript was initially developed. We thank R. Alexander, L. Bittel, A. J. Brady, J. Eisert, S. T. Flammia, M. J. Gullans, J. Haferkamp, L. Leone, A. A. Mele, J. P. Seifert, and N. Walk for helpful discussions as well as F. A. Mele for pointing us to an error in an earlier version of this manuscript. J. T. I. thanks the Joint Quantum Institute at UMD for support through a JQI fellowship. We acknowledge support from NSF grants OMA-2120757 (QLCI) and CIF-2330909, the BMBF (RealistiQ, MUNIQC-Atoms, PhoQuant, QPIC-1, and QSolid), the DFG (CRC 183, project B04, on entangled states of matter), the Munich Quantum Valley (K-8), the ERC (DebuQC), Quantum Berlin as well as the ERU on quantum devices.

[1] A.k.a. averaging sets, quadratures, or cubatures.

[2] A. H. Stroud, *Approximate calculation of multiple integrals* (Prentice Hall, 1971).

[3] P. Delsarte, J. M. Goethals, and J. J. Seidel, Spherical codes and designs, Geometriae Dedicata **6**, 363–388 (1977).

[4] J. Conway and N. Sloane, *Sphere packings, lattices and groups*, Vol. 290 (Springer, New York, NY, 1988).

[5] C. J. Colbourn and J. H. Dinitz, *Handbook of Combinatorial Designs, Second Edition (Discrete Mathematics and Its Applications)* (Chapman & Hall/CRC, 2006).

[6] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, Symmetric informationally complete quantum measurements, Journal of Mathematical Physics **45**, 2171–2180 (2004).

[7] H.-Y. Huang and R. Kueng, Predicting features of quantum systems from very few measurements (2019), arXiv:1908.08909 [quant-ph].

[8] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nature Physics **16**, 1050–1057 (2020).

[9] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, The randomized measurement toolbox, Nature Reviews Physics **5**, 9–24 (2022).

[10] J. Haferkamp, P. Faist, N. B. T. Kothakonda, J. Eisert, and N. Yunger Halpern, Linear growth of quantum circuit complexity, Nature Physics **18**, 528–532 (2022).

[11] D. Gross, K. Audenaert, and J. Eisert, Evenly distributed unitaries: On the structure of unitary designs, Journal of Mathematical Physics **48**, 10.1063/1.2716992 (2007).

[12] J. Haferkamp, *Randomness and complexity in random complex quantum systems*, Ph.D. thesis, Freie Universität Berlin (2022).

[13] S. P. Jain, J. T. Iosue, A. Barg, and V. V. Albert, Quantum spherical codes, Nature Physics **20**, 1300–1305 (2024).

[14] J. T. Iosue, K. Sharma, M. J. Gullans, and V. V. Albert, Continuous-variable quantum state designs: Theory and applications, Physical Review X **14**, 10.1103/physrevx.14.011013 (2024).

[15] V. V. Albert, K. Noh, K. Duivenvoorden, D. J. Young, R. T. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. M. Girvin, B. M. Terhal, and L. Jiang, Performance and structure of single-mode bosonic codes, Phys. Rev. A **97**, 032346 (2018).

[16] B. M. Terhal, J. Conrad, and C. Vuillot, Towards scalable bosonic quantum error correction, Quant. Sc. Tech. **5**, 043001 (2020).

[17] 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).

[18] R. Blume-Kohout and P. S. Turner, The curious nonexistence of gaussian 2-designs, Communications in Mathematical Physics **326**, 755–771 (2014).

[19] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).

[20] 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 (2023).

[21] J. Zak, Finite translations in solid-state physics, Phys. Rev. Lett. **19**, 1385 (1967).

[22] M. Rymarz, S. Bosco, A. Ciani, and D. P. DiVincenzo, Hardware-encoding grid states in a nonreciprocal superconducting circuit, Phys. Rev. X **11**, 011032 (2021).

[23] A. Perelomov, *Generalized Coherent States and Their Applications* (Springer Berlin Heidelberg, 1986).

[24] Y. Aharonov, H. Pendleton, and A. Petersen, Modular variables in quantum theory, International Journal of Theoretical Physics **2**, 213–230 (1969).

[25] O. Regev, On lattices, learning with errors, random linear codes, and cryptography, J. ACM **56**, 10.1145/1568318.1568324 (2009).

[26] J. Conrad, J. Eisert, and J.-P. Seifert, Good Gottesman-Kitaev-Preskill codes from the NTRU cryptosystem, Quantum **8**, 1398 (2024).

[27] H. G. Feichtinger and T. Strohmer, *Gabor Analysis and Algorithms* (Birkhäuser Boston, 1998).

[28] K. Duivenvoorden, B. M. Terhal, and D. Weigand, Single-mode displacement sensor, Phys. Rev. A **95**, 012305 (2017).

[29] A. Acharya, S. Saha, and A. M. Sengupta, Informationally complete povm-based shadow tomography (2021), arXiv:2105.05992 [quant-ph].

[30] C. Flühmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. P. Home, Encoding a qubit in a trapped-ion mechanical oscillator, Nature **566**, 513–517 (2019).

[31] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, and et al., Quantum error correction of a qubit encoded in grid states of an oscillator, Nature **584**, 368 (2020).

[32] B. De Neeve, T. L. Nguyen, T. Behrle, and J. P. Home, Error correction of a logical grid state qubit by dissipative pumping, Nature Physics **18**, 296 (2022).

[33] D. Lachance-Quirion, M. A. Lemonde, J. O. Simoneau, L. St-Jean, P. Lemieux, S. Turcotte, W. Wright, A. Lacroix, J. Fréchette-Viens, R. Shillito, F. Hopfmueller, M. Tremblay, N. E. Frattini, J. Camirand Lemyre, and P. St-Jean, Autonomous quantum error correction of gottesman-kitaev-preskill states, Phys. Rev. Lett. **132**, 150607 (2024).

[34] C. H. Valahu, M. P. Stafford, Z. Huang, V. G. Matsos, M. J. Millican, T. Chalermpusitarak, N. C. Menicucci, J. Combes, B. Q. Baragiola, and T. R. Tan, Quantum-enhanced multi-parameter sensing in a single mode (2024), arXiv:2412.04865 [quant-ph].

[35] B. L. Brock, S. Singh, A. Eickbusch, V. V. Sivak, A. Z. Ding, L. Frunzio, S. M. Girvin, and M. H. Devoret, Quantum error correction of qudits beyond break-even (2024), arXiv:2409.15065 [quant-ph].

[36] 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 (1994).

[37] M. Moskowitz and R. Sacksteder, An extension of a theorem by Hlawka, Mathematika **56**, 203–216 (2010).

[38] J. Conrad, J. Eisert, and F. Arzani, Gottesman-Kitaev-Preskill codes: A lattice perspective, Quantum **6**, 648 (2022).

[39] J. Conrad, T. Iosue, J, A. G. Burchards, and V. V. Albert, Supplemental material: Continuous-variable designs and design-based shadow tomography from random lattices (2024).

[40] D. Kelmer and S. Yu, The second moment of the siegel transform in the space of symplectic lattices, International Mathematics Research Notices **2021**, 5825–5859 (2019).

[41] E. T. Campbell, Catalysis and activation of magic states in fault-tolerant architectures, Physical Review A **83**, 10.1103/phys-

reva.83.032317 (2011).
[42] C. Weedbrook, S. Pirandola, R. Garcia-Patron, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621 (2012).
[43] This is not a very stringent assumption, given that physical states are purported to decay super-polynomially [54], and that all Gaussian states decay exponentially.
[44] N. Elkies, Theta functions and weighted theta functions of euclidean lattices, with some applications, "https://people.math.harvard.edu/~elkies/aws09.pdf" (2009).
[45] B. M. Terhal and D. Weigand, Encoding a qubit into a cavity mode in circuit qed using phase estimation, Physical Review A **93**, 10.1103/physreva.93.012315 (2016).
[46] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Quantum fingerprinting, Physical Review Letters **87**, 10.1103/physrevlett.87.167902 (2001).
[47] G. J. Milburn, Quantum optical fredkin gate, Phys. Rev. Lett. **62**, 2124 (1989).
[48] A. Z. Ding, B. L. Brock, A. Eickbusch, A. Koottandavida, N. E. Frattini, R. G. Cortinas, V. R. Joshi, S. J. de Graaf, B. J. Chapman, S. Ganjam, L. Frunzio, R. J. Schoelkopf, and M. H. Devoret, Quantum control of an oscillator with a kerr-cat qubit (2024), arXiv:2407.10940 [quant-ph].
[49] C. A. Rogers, Mean values over the space of lattices, Acta Mathematica **94**, 249–287 (1955).
[50] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. Ralph, J. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621–669 (2012).
[51] D. Chen and H. Zhu, Nonstabilizerness enhances thrifty shadow estimation (2024), arXiv:2410.23977 [quant-ph].
[52] L. Leone and L. Bittel, Stabilizer entropies are monotones for magic-state resource theory, Physical Review A **110**, 10.1103/physreva.110.l040403 (2024).
[53] H. Zhu, R. Kueng, M. Grassl, and D. Gross, The clifford group fails gracefully to be a unitary 4-design (2016), arXiv:1609.08172 [quant-ph].
[54] V. V. Albert, Bosonic coding: introduction and use cases (2022), arXiv:2211.05714 [quant-ph].
[55] J. Conrad, J. Eisert, and S. T. Flammia, Chasing shadows with gottesman-kitaev-preskill codes (2024), arXiv:2411.00235 [quant-ph].
[56] K. E. Cahill and R. J. Glauber, Ordered expansions in boson amplitude operators, Phys. Rev. **177**, 1857 (1969).
[57] K. Noh and C. Chamberland, Fault-tolerant bosonic quantum error correction with the surface–Gottesman-Kitaev-Preskill code, Phys. Rev. A **101**, 012316 (2020).
[58] F. A. Mele, A. A. Mele, L. Bittel, J. Eisert, V. Giovannetti, L. Lami, L. Leone, and S. F. E. Oliviero, Learning quantum states of continuous variable systems (2024), arXiv:2405.01431 [quant-ph].
[59] C. Gerry and P. Knight, *Introductory Quantum Optics* (Cambridge University Press, 2004).
[60] L. Bittel, F. A. Mele, A. A. Mele, S. Tirone, and L. Lami, Optimal estimates of trace distance between bosonic gaussian states and applications to learning (2024), arXiv:2411.02368 [quant-ph].
[61] W. Rossmann, *Lie Groups*, Oxford Graduate Texts in Mathematics (Oxford University Press, London, England, 2006).
[62] S. Becker, N. Datta, L. Lami, and C. Rouzé, Convergence rates for the quantum central limit theorem, Communications in Mathematical Physics **383**, 223–279 (2021).
[63] B. Royer, S. Singh, and S. M. Girvin, Encoding qubits in multimode grid states, PRX Quantum **3**, 010335 (2022).
[64] A. G. Burchards, S. T. Flammia, and J. Conrad, Fiber bundle fault tolerance of GKP codes (2024), arXiv:2410.07332 [quant-ph].
[65] A. Serafini, *Quantum Continuous Variables* (CRC Press, 2017).
[66] C. L. Siegel, A mean value theorem in geometry of numbers, Annals of Mathematics **46**, 340 (1945).
[67] P. Garret, "https://www-users.cse.umn.edu/~garrett/m/v/volumes.pdf" (2014).
[68] In the literature the mean value formula is typically stated for compact functions. Note that any (Riemann) integrable function can be partitioned into a sum of functions over compactly restricted domains, so that this statement naturally generalizes. This observation is also implicitly present in Ref. [49].
[69] N. C. Menicucci, Fault-tolerant measurement-based quantum computing with continuous-variable cluster states, Phys. Rev. Lett. **112**, 120504 (2014).
[70] I. Tzitrin, J. E. Bourassa, N. C. Menicucci, and K. K. Sabapathy, Progress towards practical qubit computation using approximate Gottesman-Kitaev-Preskill codes, Phys. Rev. A **101**, 032315 (2020).
[71] T. Matsuura, H. Yamasaki, and M. Koashi, Equivalence of approximate gottesman-kitaev-preskill codes, Physical Review A **102**, 10.1103/physreva.102.032408 (2020).
[72] S. Konno, W. Asavanant, F. Hanamura, H. Nagayoshi, K. Fukui, A. Sakaguchi, R. Ide, F. China, M. Yabuno, S. Miki, H. Terai, K. Takase, M. Endo, P. Marek, R. Filip, P. van Loock, and A. Furusawa, Logical states for fault-tolerant quantum computation with propagating light, Science **383**, 289–293 (2024).
[73] K. Noh and C. Chamberland, Fault-tolerant bosonic quantum error correction with the surface–gottesman-kitaev-preskill code, Physical Review A **101**, 10.1103/physreva.101.012316 (2020).
[74] Note that although $f(\boldsymbol{\lambda}) = e^{-i2\pi\boldsymbol{\lambda}^T J\boldsymbol{x}}$ is strictly speaking not an integrable function, the mean value formula may be applied to a regularized form $\tilde{f}_\beta(\boldsymbol{\lambda}) = (2\pi/\beta)^{-1/2} e^{-\frac{\beta}{2}\|\boldsymbol{\lambda}\|^2 - i2\pi\boldsymbol{\lambda}^T J\boldsymbol{x}}$ and the result is obtained in the limit $\beta \to 0$.
[75] L. R. Welch, Lower bounds on the maximum cross correlation of signals, IEEE Transactions on Information Theory **20**, 397 (1974).
[76] S. Datta, S. Howard, and D. Cochran, Geometry of the Welch bounds (2009), arXiv:0909.0206 [cs.IT].
[77] A. Klappenecker and M. Rötteler, Constructions of mutually unbiased bases, in *Finite Fields and Applications*, edited by G. L. Mullen, A. Poli, and H. Stichtenoth (Springer Berlin Heidelberg, Berlin, Heidelberg, 2004) pp. 137–144.
[78] A. Klappenecker and M. Rotteler, Mutually unbiased bases are complex projective 2-designs, in *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.* (2005) pp. 1740–1744.
[79] A. J. Scott, Optimizing quantum process tomography with unitary 2-designs, Journal of Physics A: Mathematical and Theoretical **41**, 10.1088/1751-8113/41/5/055308 (2008), arXiv:0711.1017 [quant-ph].
[80] A. Belovs and J. Smotrovs, A criterion for attaining the Welch bounds with applications for mutually unbiased bases (2008), arXiv:0802.0855 [quant-ph].
[81] D. A. Roberts and B. Yoshida, Chaos and complexity by design, Journal of High Energy Physics **2017**, 121 (2017), arXiv:1610.04903 [quant-ph].
[82] Or $f = (d^n + 1)^{-1}$ for multi-qudit systems.
[83] S. Das, Central limit theorems for lattice point counting on tessellated domains (2024), arXiv:2406.05429 [math.NT].

[84] Y.-C. Chen, A short note on the median-of-means estimator, https://faculty.washington.edu/yenchic/short_note/note_MoM.pdf (2020).
[85] M. Lerasle, Lecture notes: Selected topics on robust statistical learning theory (2019), arXiv:1908.10761 [stat.ML].
[86] S. Gandhari, V. V. Albert, T. Gerrits, J. M. Taylor, and M. J. Gullans, Precision bounds on continuous-variable state tomography using classical shadows (2023), arXiv:2211.05149 [quant-ph].
[87] J. Wright, *How to learn a quantum state*, Ph.D. thesis, Carnegie Mellon University (2016).
[88] S. Becker, N. Datta, L. Lami, and C. Rouze, Classical shadow tomography for continuous variables quantum systems, IEEE Transactions on Information Theory **70**, 3427–3452 (2024).
[89] B. M. Terhal, Quantum error correction for quantum memories, Reviews of Modern Physics **87**, 307 (2015).
[90] K. Noh, C. Chamberland, and F. G. S. L. Brandão, Low-overhead fault-tolerant quantum error correction with the surface-GKP code, PRX Quantum **3**, 010315 (2022).
[91] D. J. Weigand and B. M. Terhal, Generating grid states from schrödinger-cat states without postselection, Physical Review A **97**, 10.1103/physreva.97.022341 (2018).
[92] D. J. Weigand and B. M. Terhal, Realizing modular quadrature measurements via a tunable photon-pressure coupling in circuit qed, Physical Review A **101**, 10.1103/physreva.101.053840 (2020).

APPENDIX TABLE OF CONTENTS

A. The toolbox 10
1. Continuous-variable physics 10
2. Lattices & GKP states 12
3. The mean value formula 12
4. Continuous variable designs 12
5. Warm up: a CV 1-design from random GKP states 13

B. State 2-designs from displaced lattices 14
1. A rigged 2-design from displaced GKP states 15
2. Regularized 2-designs from finitely squeezed GKP states 16
3. Continuous variable frame potentials 17
The meaning of the frame potential 20
4. Diagonal coherent state basis 21

C. Design-based CV shadow tomography with GKP states 22
1. Shadow tomography from 2-designs 22
2. Local shadows vs. global shadows 23
3. Global and local design-based shadow tomography for 3-designs 24
a. Global shadows 24
b. Local shadows 25

D. Continuous variable shadow tomography using GKP states 25
1. Bounding the estimator variance 25
2. Physical state constraints 27
3. Variance bounds with physicality assumptions 29
4. Continuous variable medians of means estimation 31
5. Example: Thermal state observables 32

E. Sample complexity: comparison to related protocols 33
1. Qubit-based classical shadows 33
2. Photon number cutoffs 34
a. Full state tomography 34
3. Lattice based CV shadows 35

F. Implementation of design-based GKP shadow tomography 35
1. Circuit analysis and batched implementation 36
a. Estimating the shadow using qubit-controlled displacements 36
b. Estimating the shadow using auxiliary GKP states 37
c. Typicality 38

G. Application: variational preparation of thermal states many-body states 39

OUTLINE OF APPENDIX

The following material is structured as follows. In section A, we review the technical tools necessary to follow the mathematical developments in this manuscript. In this section we cover the relevant tools from physics, briefly review GKP coding- and lattice theory and explain the structure of CV designs as defined in Ref. [14]. We close this section by constructing a "warm up" example of a CV 1-design based on random lattice states. Section B presents the main result. We show that uniformly randomly displaced GKP states constructed from a random symplectic lattice yield a 2-design as per the definition provided in Ref. [14] and we discuss how random ensembles of lattices can be obtained in practice. Sections C and D apply the core result by developing a CV shadow tomography protocol based on the design property. We investigate the example of estimating a thermal state observable as well as general bounds on the sample complexity without and with additional physical assumptions. In section F we discuss various recipes to physically implement the protocol. While we do not show whether our ensembles also attain the 3-design property, we explore the ramifications of 3-design based classical shadow tomography in sec. C 3. Finally, we outline an application of our protocol in the variational preparation of a bosonic thermal state in section G.

Appendix A: The toolbox

1. Continuous-variable physics

An n-mode continuous-variable system is naturally described using a $2n$-dimensional phase space $\mathbb{R}^{2n}$ which supports the bosonic quadratures $\hat{\boldsymbol{x}} = (\hat{q}_1, \dots \hat{q}_n, \hat{p}_1, \dots, \hat{p}_n)$, which we assumed to be normalized such that $[\hat{x}_i, \hat{x}_j] = iJ_{ij}$ with $J = -J^T = -J^{-1}$ the symplectic form. An important class of operators is given by the displacement operators

$$D(\boldsymbol{\xi}) = e^{-i\sqrt{2\pi}\boldsymbol{\xi}^T J \hat{\boldsymbol{x}}}, \tag{A1}$$

which satisfy

$$D(\boldsymbol{\xi}) D(\boldsymbol{\eta}) = e^{-i2\pi\boldsymbol{\xi}^T J \boldsymbol{\eta}} D(\boldsymbol{\eta}) D(\boldsymbol{\xi}) = e^{-i\pi\boldsymbol{\xi}^T J \boldsymbol{\eta}} D(\boldsymbol{\xi} + \boldsymbol{\eta}) \tag{A2}$$

and form a basis, orthonormal in the Dirac sense, for the space of trace-class operators on the CV system. That is, every trace-class operator O can be expressed as

$$O = \int_{\mathbb{R}^{2n}} d\boldsymbol{\xi}\, c_O(\boldsymbol{\xi})\, D(\boldsymbol{\xi}), \tag{A3}$$

where $c_O(\boldsymbol{\xi}) = \mathrm{Tr}[D^\dagger(\boldsymbol{\xi}) O]$ is called the characteristic function of the operator O.

A relevant set of operators for the CV system is given by the Fock operators $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$ and the annihilation operators $\hat{a}_i = (\hat{q}_i + i\hat{p}_i)/\sqrt{2}$. The Fock operators count the *photon number* of the individual modes, and the total photon number is given by $\hat{N} = \sum_{i=1}^n \hat{n}_i$.

The *total photon parity operator* is given by $\hat{\pi} = e^{i\pi\hat{N}}$ and has a constant characteristic function $c_\pi(\boldsymbol{\xi}) = 2^{-n}$ (see e.g. Ref. [55]) such that the symplectic Fourier transform of the characteristic function of an operator $c_\rho(\boldsymbol{\xi})$ can be obtained via the expectation value of displaced parity operators (a.k.a. phase-space point operators),

$$\Omega(\boldsymbol{\xi}) = D(\boldsymbol{\xi}) \hat{\pi} D^\dagger(\boldsymbol{\xi}), \tag{A4}$$

which satisfy

$$\mathrm{Tr}[\Omega(\boldsymbol{\xi}) \rho] = 2^{-n} \int_{\mathbb{R}^{2n}} d\boldsymbol{\eta}\, e^{-i2\pi\boldsymbol{\xi}^T J \boldsymbol{\eta}} \mathrm{Tr}[D(\boldsymbol{\eta}) \rho] = 2^{-n} W_\rho(\boldsymbol{\xi}). \tag{A5}$$

The function $W_\rho(\boldsymbol{x})$ is the *Wigner function* of the state (or general operator) ρ. For quantum states, it represents a quasi-probability distribution over phase space and the formulas above imply that traces of operators and their Wigner functions can be obtained as

$$\mathrm{Tr}[A] = \int_{\mathbb{R}^{2n}} d\boldsymbol{\xi}\, W_A(\boldsymbol{\xi}), \quad \mathrm{Tr}[AB] = \int_{\mathbb{R}^{2n}} d\boldsymbol{\xi}\, W_A(\boldsymbol{\xi}) W_B(\boldsymbol{\xi}). \tag{A6}$$

The Wigner function of a state $\rho = \rho^\dagger$ always takes real values, and we have the bound $|W_A| \leq 2^n \|A\|$. Throughout this manuscript we will make an important distinction between *formal* and *physical* operators, where we say an operator is physical if the corresponding Wigner function is a Schwartz function – i.e., a function that admits finite expectation values of any power of Fock operators [54]. In terms of the characteristic function of a state this requirement may be phrased as follows. As the Fourier transform is an automorphism of Schwartz space, the requirement of the Wigner function of any physical state to be Schwartz implies this attribute also to the characteristic function, which in particular means that $c_\rho(\boldsymbol{\xi})$ decays superpolynomially as $\|\xi\| \to \infty$ and we can restrict $c_\rho(\boldsymbol{\xi})$ to have compact essential support below a radius R_ρ. We call R_ρ the *bandwidth* of a state (or operator, respectively). As it quantifies the length of the longest relevant displacement appearing in the decomposition of Eq. (A3). As $D(\boldsymbol{x}) D(\boldsymbol{\xi}) D^\dagger(\boldsymbol{x}) = e^{-i2\pi\boldsymbol{x}^T J \boldsymbol{\xi}}$, this length-scale also determines the length $\|\boldsymbol{x}\| \sim \|\boldsymbol{\xi}\|^{-1}$ of the *shortest* displacement $D(\boldsymbol{x})$ under which the corresponding state-component is displacement-invariant. One may hence think of the displacement vectors appearing in the decomposition of Eq. (A3) as generalized *frequencies* of the corresponding operator.

Furthermore it holds that, for $\beta \in (0, \infty)$ [56, 57],

$$e^{-\beta\hat{N}} = \left(1 - e^{-\beta}\right)^{-n} \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}\, e^{-\frac{\pi\|\boldsymbol{\gamma}\|^2}{\Delta^2}} D(\boldsymbol{\gamma}), \tag{A7}$$

where $\Delta^2 = 2\tanh(\beta/2) = \beta + O(\beta^3) \in (0, 2)$. We say that a state is *physical*, if

$$N[\rho](\beta) = \mathrm{Tr}\left[\rho e^{-\beta\hat{N}}\right] = \left(1 - e^{-\beta}\right)^{-n} \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}\, e^{-\frac{\pi\|\boldsymbol{\gamma}\|^2}{\Delta^2}} c_\rho(\boldsymbol{\gamma}). \tag{A8}$$

is an *analytical* function of β at $\beta = \Delta^2 = 0$. In particular this means that each moment

$$\lim_{\beta\to 0}\left(-\frac{\partial}{\partial\beta}\right)^k N\left[\rho\right](\beta) = \mathrm{Tr}\left[\rho\hat{N}^k\right] \tag{A9}$$

is finite [54, 58]. A direct route to understand this characterization is to recognize the characteristic function as the *moment generating function* of a state, i.e. as

$$c_\rho(\boldsymbol{\gamma}) = \mathrm{Tr}\left[D^\dagger(\boldsymbol{\gamma})\,\rho\right], \tag{A10}$$

the moments

$$\mathrm{Tr}\left[P(\hat{\boldsymbol{x}})\big|_W\rho\right] = \lim_{\boldsymbol{\gamma}\to 0} P\left(-\frac{iJ\boldsymbol{\nabla}}{\sqrt{2\pi}}\right) c_\rho(\boldsymbol{\gamma}), \tag{A11}$$

for any multivariate [59] polynomial $P:\mathbb{R}^{2n}\to\mathbb{R}$ are encoded in the fluctuating behaviour of c_ρ around $\boldsymbol{\gamma}=0$. The index "W" indicates that we assume the polynomial evaluated in the quadrature operators to be Weyl-odered [56, 59], i.e. symmetrized amongst all permutations of ordering the quadrature operators $(\hat{x}_i, \hat{x}_{i+n})$, $i = 1,\ldots,n$.

A more general expression an be obtained by computing the partial derivatives of the displacement operators (see also ref. [60] for an alternative form). Using the derivative of an exponential map [61]

$$\frac{d}{dt}e^{A(t)} = \int_0^1 dy\, e^{yA(t)}A'(t)e^{(1-y)A(t)} \tag{A12}$$

we obtain the expressions

$$\boldsymbol{\nabla}D(\boldsymbol{\gamma}) = -i\sqrt{2\pi}J\left(\hat{\boldsymbol{x}} - \sqrt{\frac{\pi}{2}}\boldsymbol{\gamma}\right)D(\boldsymbol{\gamma}) \tag{A13}$$

$$\Delta D(\boldsymbol{\gamma}) = Q(\boldsymbol{\gamma})\,D(\boldsymbol{\gamma}), \qquad \text{with} \tag{A14}$$

$$Q(\boldsymbol{\gamma}) = -2\pi\left(\hat{\boldsymbol{x}} - \sqrt{\frac{\pi}{2}}\boldsymbol{\gamma}\right)^T\left(\hat{\boldsymbol{x}} - \sqrt{\frac{\pi}{2}}\boldsymbol{\gamma}\right) = -2\pi D\left(\frac{\boldsymbol{\gamma}}{2}\right)\hat{\boldsymbol{x}}^T\hat{\boldsymbol{x}}D^\dagger\left(\frac{\boldsymbol{\gamma}}{2}\right). \tag{A15}$$

We recognize that

$$Q(0) = -4\pi\left(\hat{N} + n/2\right), \tag{A16}$$

as well as $\Delta Q(\boldsymbol{\gamma}) = -(2\pi)^2 n$. Now using the product rule for Laplaciancs

$$\Delta(fg) = (\Delta f)\,g + 2(\boldsymbol{\nabla}f)^T(\boldsymbol{\nabla}g) + f(\Delta g) \tag{A17}$$

it can be shown via induction that for every $k, n \in \mathbb{N}$ there exists a sequence of constants $c_m^{k,n}$ such that

$$\Delta^k D(\boldsymbol{\gamma}) = \sum_{m=0}^{k} c_m^{k,n} Q^m(\boldsymbol{\gamma})\,D(\boldsymbol{\gamma}). \tag{A18}$$

In particular, this also implies that

$$\Delta^k c_\rho(0) = \sum_{m=0}^{k} c_m^{k,n}(-4\pi)^m\,\mathrm{Tr}\left[\left(\hat{N} + \frac{n}{2}\right)^m\rho\right]. \tag{A19}$$

That is, the fluctuations of the characteristic function $c_\rho(\boldsymbol{\gamma})$ around $\boldsymbol{\gamma}=0$ encode the moments of the photon number operator (see also ref. [62]).

2. Lattices & GKP states

Gottesman-Kitaev-Preskill (GKP) codes [19, 38] are stabilizer codes whose stabilizer group is generated by a full-rank set of displacement operators. The properties in Eq. (A2) imply that the stabilizer group becomes isomorphic to a (weakly) symplectically self-dual lattice $\mathcal{L} \subseteq \mathcal{L}^{\perp} \subset \mathbb{R}^{2n}$, i.e. the integer span of a full rank matrix $\mathcal{L} = \mathbb{Z}^{2n} M$ with integral symplectic Gram matrix $A = MJM^T$, whose points label the displacements occuring in the stabilizer group. Here duality is implied by the symplectic inner product represented by matrix J. The subspace symmetric under this stabilizer group has dimension $d = \sqrt{|\mathcal{L}/\mathcal{L}^{\perp}|} = \sqrt{|\det(\mathcal{L})/\det(\mathcal{L}^{\perp})|} = |\det(\mathcal{L})|$ so that in particular symplectic (symplectic self-dual) lattices $\Lambda = \Lambda^{\perp}$ fix a one-dimensional subspace of the CV Hilbert space spanned by a formal state described by the (formal) group projector [38]

$$\Pi_{\Lambda} = |\Lambda\rangle\langle\Lambda| = \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} D(\boldsymbol{\lambda}). \tag{A20}$$

Note, that throughout this manuscript we explicitly include the phases $e^{i\Phi(\boldsymbol{\lambda})}$, which result from the fact that – in our slight abuse of notation – the states $|\Lambda\rangle$ are in fact not fully determined by the lattice Λ as a geometric object, but their specification typically involves a preferred choice of basis [38, 63, 64]. Note, however, that in all subsequent discussions these phases will turn out to become irrelevant.

The phase $e^{i\Phi(\boldsymbol{\lambda})}$ takes values ± 1 depending on the decomposition of $\boldsymbol{\lambda}$ into integer linear combinations of the minimal basis chosen to generate the corresponding stabilizer group [38, 64]. Let $|\boldsymbol{\lambda}\rangle = D(\boldsymbol{\lambda})|0\rangle$ be the coherent state at $\boldsymbol{\lambda} \in \mathbb{R}^{2n}$. Coherent states are such that

$$\langle\boldsymbol{\beta}|\boldsymbol{\alpha}\rangle = e^{-\frac{\pi}{2}\left(\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|^2 + i2\boldsymbol{\alpha}^T J\boldsymbol{\beta}\right)} \tag{A21}$$

and form an overcomplete basis [65]

$$\int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha}\, |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| = I. \tag{A22}$$

3. The mean value formula

The space of symplectic lattices $Y_n = \mathrm{Sp}_{2n}(\mathbb{Z}) \setminus \mathrm{Sp}_{2n}(\mathbb{R})$– parametrized by symplectic matrices $\mathrm{Sp}_{2n}(\mathbb{R})$ *up to basis transformations* $\mathrm{Sp}_{2n}(\mathbb{Z})$ acting from the left admits a natural normalized Haar measure. This insight had been extended from the theory of classical lattices $X_n = \mathrm{SL}_{2n}(\mathbb{Z}) \setminus \mathrm{SL}_{2n}(\mathbb{Z})$, originally examined by Siegel [66], to the case of symplectic lattices [26, 36, 67]. Over this measure, an elegant formula holds for the mean value of a lattice sum when averaged over the space of respective lattices [68].

Lemma 1 (Symplectic mean value formula [36] [37, theorem 2])**.** *Let $f : \mathbb{R}^{2n} \to \mathbb{R}$ be a (Riemann) integrable function following quantities converge absolutely and let $\Lambda \in Y_n$ denote a symplectic lattice. For*

$$F(\Lambda) = \sum_{\boldsymbol{\lambda}\in\Lambda-\{0\}} f(\boldsymbol{\lambda}) \tag{A23}$$

it holds that

$$\int_{Y_n} d\mu(\Lambda)\, F(\Lambda) = \int_{\mathbb{R}^{2n}} d\boldsymbol{x}\, f(\boldsymbol{x}) \tag{A24}$$

relative to the Haar measure over the space of symplectic lattices $Y_n = \mathrm{Sp}_{2n}(\mathbb{Z}) \setminus \mathrm{Sp}_{2n}(\mathbb{R})$.

In what follows, this formula will be the main ingredient for our construction of CV state designs via lattice-states.

4. Continuous variable designs

We briefly recall the relevant definition of continuous-variable state designs for n-mode quantum systems as stated in Ref. [14]. This requires a slightly extended definition of states. Rather than requiring that states are representable as square integrable wave functions in $\mathcal{H} = L^2(\mathbb{R}^n)$, we also allow for *improper* states which can be represented within the space of tempered distributions $S'(\mathbb{R}^n)$, i.e. the continuous dual of the space of Schwartz functions $S(\mathbb{R}^n)$ under the standard inner product. This sequence of spaces forms the well-known Gelfand triple $S(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \subset S'(\mathbb{R}^n)$, with the first two spaces in the triple referred to as a

rigged Hilbert space. The point of this relaxation is, as we will see in the course of the manuscript, that whenever we encounter such improper states, they can still be made sense of as a measurement process that acts on a physical state.

Formally, a continuous-variable, or rigged, t-design is a measure space $(\mathcal{X}_n, \Sigma, \mu)$, where $\mathcal{X}_n \subset S'(\mathbb{R}^n)$ is a subset of tempered distributions, the Σ-algebra defines a means to take unions and complements of elements in $\mathcal{X}_n$, and where $\mu_{\mathcal{X}_n}$ is a measure on $\mathcal{X}_n$.

Definition 1 (Rigged designs [14]). *Let $\mathcal{X}_n \subset S'(\mathbb{R}^n)$. The measure space $(\mathcal{X}_n, \Sigma, \mu_{\mathcal{X}_n})$ is a rigged t-design if there are finite constants $a_{t'} > 0$, such that*

$$\int_{\mathcal{X}_n} d\mu_{\mathcal{X}_n}(X) |X\rangle\langle X|^{\otimes t'} = a_{t'} \Pi_{t'} \ \forall t' \leq t. \tag{A25}$$

In this definition the objects Π_t appearing on the right are the projectors onto the symmetric subspaces of the Hilbert space $\mathcal{H}^{\otimes t}$. The relaxation of $L^2(\mathbb{R}^n)$ to $S'(\mathbb{R}^n)$ allows $(t \geq 2)$-designs to exist. Despite elements of $S'(\mathbb{R}^n)$ not being normalizable states, this relaxation often does not affect the physicality of the design itself. The reason is because vectors corresponding to *physical* states are those with finite energy and finite position and momentum moments, and hence belong to $S(\mathbb{R}^n)$. Inner products of tempered distributions (elements of $S'(\mathbb{R}^n)$) and Schwartz functions (elements of $S(\mathbb{R}^n)$) are always well-defined. It follows that even though a rigged t-design must consist of tempered distributions, measuring from the POVM that it defines is perfectly well-defined, as expectation values of physical observables over tempered distributions are well-defined. This ultimately allows shadow tomography with rigged t-designs to be a valid protocol [14].

However, we may also encounter situations where we want the elements of our designs to correspond to normalizable states. To this end, Ref. [14] defined regularizers $R_\beta : S'(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ for $\beta > 0$, where β is some parameter such that $R_0 = I$, as well as the regularized projectors $\Pi_t^{(R_\beta)} = R_\beta^{\otimes t} \Pi_t R_\beta^{\otimes t}$

Definition 2 (Regularized rigged designs [14]). *An R_β-regularized (rigged) t-design is a measure space $(\mathcal{Y}, \Sigma_Y, \mu_{\mathcal{Y}})$ with $\mathcal{Y} \subset L^2(\mathbb{R}^n)$, such that*

$$\int_{\mathcal{Y}} d\mu_{\mathcal{Y}}(Y) |Y\rangle\langle Y|^{\otimes t'} = a_{t'} \Pi_{t'}^{(R_\beta)} / \operatorname{Tr}\left[\Pi_{t'}^{(R_\beta)}\right] \ \forall t' \leq t. \tag{A26}$$

5. Warm up: a CV 1-design from random GKP states

In this first application we use the mean value formula to show that that GKP states provided by random lattices $\Lambda \in Y_n$ form a CV 1-design. Note that, due to its over-completeness, the set of coherent states $|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|$ also possesses this property. In contrast to the coherent states, labeled by $\boldsymbol{\alpha} \in \mathbb{R}^{2n}$, however, the parametrizing set Y_n is instead compact. The main goal of this example is to provide a first demonstration of how to use the mean value formula. Let us begin by defining the lattice state

$$\rho_\Lambda = \sum_{\boldsymbol{\lambda} \in \Lambda} |\boldsymbol{\lambda}\rangle\langle\boldsymbol{\lambda}| = \sum_{\lambda \in \Lambda} D(\boldsymbol{\lambda}) |0\rangle\langle 0| D^\dagger(\boldsymbol{\lambda}), \tag{A27}$$

which distinguishes itself from $|\Lambda\rangle\langle\Lambda|$ in that its Wigner function is described by a fully classical probability distribution. The state ρ_Λ can be obtained by performing a heterodyne POVM on Λ as

$$\begin{aligned}\int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha} \ |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\Lambda\rangle\langle\Lambda|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| &= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha} \sum_{\boldsymbol{\lambda},\boldsymbol{\lambda}' \in \Lambda} |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\boldsymbol{\lambda}\rangle\langle\boldsymbol{\lambda}'|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| \\ &= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha} \sum_{\boldsymbol{\lambda},\boldsymbol{\lambda}' \in \Lambda} e^{-\frac{\pi}{2}\left(\|\boldsymbol{\alpha}-\boldsymbol{\lambda}\|^2 + \|\boldsymbol{\alpha}-\boldsymbol{\lambda}'\|^2\right) - i\pi\left(\boldsymbol{\lambda}-\boldsymbol{\lambda}'\right)^T J \boldsymbol{\alpha}} |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| .\end{aligned} \tag{A28}$$

In the limit of large shortest lattice vector $\lambda_1(\Lambda) \to \infty$, the sum is essentially supported on parts where $\boldsymbol{\lambda} = \boldsymbol{\lambda}'$, such that the state becomes

$$\widetilde{\rho_\Lambda} = \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha} \sum_{\boldsymbol{\lambda} \in \Lambda} e^{-\pi\|\boldsymbol{\alpha}-\boldsymbol{\lambda}\|^2} |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| , \tag{A29}$$

which approximates ρ_Λ up to an increased uncertainty around each lattice point.

Using the mean value formula, we obtain the following lemma.

Lemma 2. *The set of formal states*

$$\rho_\Lambda^\times = \rho_\Lambda - |0\rangle\langle 0| = \sum_{\boldsymbol{\lambda}\in\Lambda-\{0\}} |\boldsymbol{\lambda}\rangle\langle\boldsymbol{\lambda}| \tag{A30}$$

for all $\Lambda \in Y_n$ forms a CV 1-design. That is, we have that

$$\int_{Y_n} d\mu\left(\Lambda\right) \rho_\Lambda^\times = I. \tag{A31}$$

Proof. We use the fact that coherent sates resolve the identity, and apply the mean-value formula

$$\begin{aligned}
\int_{Y_n} d\mu\left(\Lambda\right) \rho_\Lambda^\times &= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha} d\boldsymbol{\beta} \left[\int_{Y_n} d\mu\left(\Lambda\right) \sum_{\boldsymbol{\lambda}\in\Lambda-\{0\}} \langle\boldsymbol{\alpha}|\boldsymbol{\lambda}\rangle \langle\boldsymbol{\lambda}|\boldsymbol{\beta}\rangle\right] |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\beta}| \\
&= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha} d\boldsymbol{\beta} \left[\int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda} \langle\boldsymbol{\alpha}|\boldsymbol{\lambda}\rangle \langle\boldsymbol{\lambda}|\boldsymbol{\beta}\rangle\right] |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\beta}| \\
&= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha} d\boldsymbol{\beta} \langle\boldsymbol{\alpha}|\boldsymbol{\beta}\rangle |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\beta}| \\
&= I.
\end{aligned} \tag{A32}$$

where we have used the fact that $\langle\boldsymbol{\alpha}|\boldsymbol{\lambda}\rangle \langle\boldsymbol{\lambda}|\boldsymbol{\beta}\rangle$ is an integrable function of $\boldsymbol{\lambda}$. □

The states $\rho^\times$ are GKP states that have decohered due to a heterodyne measurement, making for a poor "classical" approximation to the corresponding GKP states. It is hence reasonable to expect the original family of GKP states to possess an even richer set of properties.

Appendix B: State 2-designs from displaced lattices

Displacement operators can be used to expand not just trace-class operators, but also some unitary operators. Applying this expansion to the parity operator, whose characteristic function is 2^{-n}, and plugging in the mean-value formula yields the following useful fact.

Lemma 3. *It holds that*

$$\int_{Y_n} d\mu\left(\Lambda\right) \sum_{\boldsymbol{\lambda}\in\Lambda-\{0\}} D\left(\boldsymbol{\lambda}\right) = 2^n e^{i\pi\hat{N}}. \tag{B1}$$

Proof. We have that for every trace class operator O that preserves Schwartz space, $\mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right) O\right]$ is an integrable function of $\boldsymbol{\lambda} \in \mathbb{R}^{2n}$. Hence, we can apply the mean value formula to obtain for every O

$$\int_{Y_n} d\mu\left(\Lambda\right) \sum_{\boldsymbol{\lambda}\in\Lambda-\{0\}} \mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right) O\right] = \int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda}\, \mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right) O\right] = 2^n\, \mathrm{Tr}\left[e^{i\pi\hat{N}} O\right]. \tag{B2}$$

Since displacement operators $D\left(\boldsymbol{\alpha}\right)$ form a complete operator basis, the above equality implies that more generally

$$\int_{Y_n} d\mu\left(\Lambda\right) \sum_{\boldsymbol{\lambda}\in\Lambda-\{0\}} D\left(\boldsymbol{\lambda}\right) = 2^n e^{i\pi\hat{N}}. \tag{B3}$$

Note that the above integrals and sums are well-defined in their action on the space of physical states, such that all physical expectation values and inner products involving these operator sums and -integrals are well-defined and consistent with the above equality. □

1. A rigged 2-design from displaced GKP states

Let $\mathcal{P}(\Lambda)$ denote a fundamental domain for the lattice Λ. In the following, we prove a second moment formula.

Theorem 1 (GKP states form a rigged 2-design)**.** *The ensemble $\mathcal{X}_n$, w.r.t. the Haar measure $\mu(\Lambda)$ over the space of lattices Y_n and the uniform measure over the fundamental domain $\mathcal{P}(\Lambda)$, forms a rigged 2-design. In particular, it holds that*

$$\int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, D(\boldsymbol{\alpha})\, \Pi_\Lambda D^\dagger(\boldsymbol{\alpha}) = I. \tag{B4}$$

and

$$\int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, D^{\otimes 2}(\boldsymbol{\alpha})\, \Pi_\Lambda^{\otimes 2} D^{\dagger\otimes 2}(\boldsymbol{\alpha}) = I\otimes I + \widehat{E}, \tag{B5}$$

where

$$\widehat{E} = e^{-i\frac{\pi}{2}\sum_j (\hat{a}_j - \hat{b}_j)^\dagger (\hat{a}_j - \hat{b}_j)} \tag{B6}$$

is the bosonic exchange, or SWAP, operator between n pairs of modes $\{\hat{a}_j, \hat{b}_j\}$.

Proof. We first show that the given set of states yields a rigged 1-design. This can in fact be shown by only the properties of the displacements in $\mathcal{P}(\Lambda)$. We have

$$\int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, D(\boldsymbol{\alpha})\, \Pi_\Lambda D^\dagger(\boldsymbol{\alpha}) = \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} D(\boldsymbol{\lambda}) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, e^{-i2\pi \boldsymbol{\alpha}^T J \boldsymbol{\lambda}}. \tag{B7}$$

To evaluate this integral, first notice that

$$\delta(\boldsymbol{x}) = \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha}\, e^{-i2\pi\boldsymbol{\alpha}^T J \boldsymbol{x}} = \sum_{\boldsymbol{\lambda}\in\Lambda} \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, e^{-i2\pi(\boldsymbol{\alpha}+\boldsymbol{\lambda})^T J \boldsymbol{x}}. \tag{B8}$$

Using Poisson summation, and assuming $\Lambda = \Lambda^\perp$, it holds that this expression equals

$$\begin{aligned}\sum_{\boldsymbol{\lambda}\in\Lambda} \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} \int_{\mathbb{R}^{2n}} d\boldsymbol{y}\, e^{-i2\pi(\boldsymbol{\alpha}+\boldsymbol{y})^T J\boldsymbol{x}} e^{i2\pi \boldsymbol{y}^T J\boldsymbol{\lambda}} &= \sum_{\boldsymbol{\lambda}\in\Lambda} \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} \int_{\mathbb{R}^{2n}} d\boldsymbol{y}\, e^{-i2\pi\boldsymbol{\alpha}^T J\boldsymbol{x}} e^{-i2\pi \boldsymbol{y}^T J(\boldsymbol{x}-\boldsymbol{\lambda})} \\ &= \sum_{\boldsymbol{\lambda}\in\Lambda} \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} e^{-i2\pi\boldsymbol{\alpha}^T J\boldsymbol{\lambda}} \delta(\boldsymbol{x}-\boldsymbol{\lambda}),\end{aligned} \tag{B9}$$

such that it has to hold that

$$\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} e^{-i2\pi\boldsymbol{\alpha}^T J\boldsymbol{\lambda}} = \delta_{\mathbf{0},\boldsymbol{\lambda}},\ \boldsymbol{\lambda}\in\Lambda. \tag{B10}$$

Inserting this expression in Eq. (B7) yields the 1-design property

$$\int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, D(\boldsymbol{\alpha})\, \Pi_\Lambda D^\dagger(\boldsymbol{\alpha}) = I. \tag{B11}$$

We proceed to show the design property also for the second moment. To this end compute

$$\widehat{M}_\Lambda \coloneqq \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, D^{\otimes 2}(\boldsymbol{\alpha})\, \Pi_\Lambda^{\otimes 2} D^{\dagger\otimes 2}(\boldsymbol{\alpha}) = \sum_{\boldsymbol{\lambda},\boldsymbol{\lambda}'\in\Lambda} e^{i\left[\Phi(\boldsymbol{\lambda})+\Phi(\boldsymbol{\lambda}')\right]} D(\boldsymbol{\alpha}) D(\boldsymbol{\lambda}) D^\dagger(\boldsymbol{\alpha}) \otimes D(\boldsymbol{\alpha}) D(\boldsymbol{\lambda}') D^\dagger(\boldsymbol{\alpha}) \tag{B12}$$

$$= \sum_{\boldsymbol{\lambda},\boldsymbol{\lambda}'\in\Lambda} e^{i\left[\Phi(\boldsymbol{\lambda})+\Phi(\boldsymbol{\lambda}')\right]} \left\{ \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, e^{-i2\pi\boldsymbol{\alpha}^T J(\boldsymbol{\lambda}+\boldsymbol{\lambda}')} \right\} D(\boldsymbol{\lambda}) \otimes D(\boldsymbol{\lambda}'). \tag{B13}$$

To evaluate the integral in the curly brackets, we again use the expression derived in Eq. (B10) to obtain

$$\widehat{M}_\Lambda = \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, D^{\otimes 2}(\boldsymbol{\alpha})\, \Pi_\Lambda^{\otimes 2} D^{\dagger\otimes 2}(\boldsymbol{\alpha}) = \sum_{\boldsymbol{\lambda}\in\Lambda} D(\boldsymbol{\lambda}) \otimes D(-\boldsymbol{\lambda}), \tag{B14}$$

where we have also used that $\Phi(\boldsymbol{\lambda}) = \Phi(-\boldsymbol{\lambda})$ is a symmetric function and an integer multiple of π [38] such that the total phase in each part of the sum has now become trivial.

Following Ref. [18], we now express each pair of modes in terms of a center-of-mass mode and a relative difference mode. There exists a Gaussian unitary (a beam splitter transformation) U inducing a symplectic transformation $S \in \mathrm{Sp}_{4n}(\mathbb{R})$ such that

$$S \begin{pmatrix} \boldsymbol{\lambda} \\ -\boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \sqrt{2}\boldsymbol{\lambda} \\ \mathbf{0} \end{pmatrix}, \tag{B15}$$

and in particular

$$U\widehat{M}_\Lambda U^\dagger = \sum_{\boldsymbol{\lambda}\in\Lambda} D\left(\sqrt{2}\boldsymbol{\lambda}\right) \otimes I\,. \tag{B16}$$

We can now average this expression over the space of all lattices in Y_n, again using the fact that displacement operators form a complete operator basis and are operators with compact characteristic function so that the mean value formula becomes applicable. Note that the following sums and integrals converge in terms of well-defined operators on Schwartz space. This yields

$$U\widehat{M}U^\dagger := \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}\in\Lambda} D\left(\sqrt{2}\boldsymbol{\lambda}\right) \otimes I \tag{B17}$$

$$= I\otimes I + \int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda} D\left(\sqrt{2}\boldsymbol{\lambda}\right) \otimes I \tag{B18}$$

$$= I\otimes I + 2^{-n}\left[\int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda} D(\boldsymbol{\lambda})\right] \otimes I \tag{B19}$$

$$= I\otimes I + e^{i\pi\hat{N}} \otimes I. \tag{B20}$$

Finally, let us introduce generalized annihilation operators $\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}$ on the two tensor factors. Using the fact that $U^\dagger \hat{a}_i U = (\hat{a}_i - \hat{b}_i)/\sqrt{2}$ yields

$$U\left(e^{i\pi\hat{N}} \otimes I\right)U^\dagger = e^{-i\frac{\pi}{2}\sum_j (\hat{a}_j - \hat{b}_j)^\dagger (\hat{a}_j - \hat{b}_j)} =: \widehat{E}, \tag{B21}$$

with $\widehat{E}$ the bosonic swap operator. □

2. Regularized 2-designs from finitely squeezed GKP states

There is a fundamental problem with the states $|\Lambda\rangle$ defined above: they are unphysical (i.e., are not in Schwartz space). This is most easily recognized by computing their Wigner function $W_{\Pi_\Lambda}(\boldsymbol{x}) = 2^n \mathrm{Tr}\left[\Omega(\boldsymbol{x})\,\Pi_\Lambda\right]$, which can be verified to constitute a non-compact function on phase space and fails to be normalizable as $\mathrm{Tr}\left[\Pi_\Lambda\right]$ formally diverges. The average photon number $\langle\hat{N}\rangle$ of such a state diverges as well.

The typical way to deal with this problem within analyses of GKP error correction is to replace $|\Lambda\rangle$ by an approximate state [19, 69]

$$|\Lambda_\beta\rangle = N_\beta e^{-\beta\hat{N}} |\Lambda\rangle\ , \tag{B22}$$

where the envelope operator $R_\beta = e^{-\beta\hat{N}}$ implements an isotropic smooth cutoff to the state and converts the unphysical, formal, state into one that is producible in experiments [16, 19, 20, 30, 31, 70–72]. Here, N_β is the relevant normalization constant (see Ref. [16] for an explicit expression).

The finite squeezing error admits a decompositon into displacement operators given by [56, 73]

$$e^{-\beta\hat{N}} = \left(1 - e^{-\beta}\right)^{-n} \int_{\mathbb{R}^{2n}} d\boldsymbol{x}\, e^{-\frac{\pi\|\boldsymbol{x}\|^2}{\Delta^2}} D(\boldsymbol{x}), \tag{B23}$$

where $\Delta^2 = 2\tanh(\beta/2) = \beta + O(\beta^2)$ is typically regarded as the square of the *squeezing parameter* as it corresponds to the variance of the width of each Gaussian peak of $|\Lambda_\beta\rangle$ in phase space. Furthermore, Δ^{-2} quantifies the variance of an enveloping Gaussian on the state that renders it normalizable (see Fig. 4 for an illustration of the structure of such states).

A simple strategy to obtain a physically valid set of states for Thm. 1 is to replace every state $|X\rangle$ in the ensemble by states of the form $N_\beta R_\beta |X\rangle$ with $N_\beta = \langle X|R_{2\beta}|X\rangle$ the normalization factor. By linearity of the theorem above, we immediately obtain the following implication.

Corollary 1. *For lattices $\Lambda \in Y_n$ and $\boldsymbol{\alpha} \in \mathcal{P}(\Lambda)$ vectors in the fundamental domains of each lattice Λ, let $\beta > 0$ be a small parameter. The set of states*

$$R_\beta \left(D(\boldsymbol{\alpha}) \Pi_\Lambda D^\dagger(\boldsymbol{\alpha}) \right) R_\beta^\dagger. \tag{B24}$$

forms a regularized 2-design.

Proof. By linearity of the integral, we have analogous to the previous proof

$$\int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, R_\beta^{\otimes 2} D^{\otimes 2}(\boldsymbol{\alpha}) \Pi_\Lambda^{\otimes 2} D^{\dagger\otimes 2}(\boldsymbol{\alpha}) R_\beta^{\dagger\otimes 2} = R_\beta^{\otimes 2} \left(I \otimes I + \widehat{E} \right) R_\beta^{\dagger\otimes 2}, \tag{B25}$$

with $\widehat{E}$ as in Thm. 1. □

3. Continuous variable frame potentials

Building on Ref. [14], we examine the ensemble of GKP states defined in the main text through the lens of continuous variable *frame potentials*. Frame potentials are an important tool in the analysis of evenly distributed states and we discuss their interpretation further below. To maintain well-defined expressions throughout, we need to work with appropriately regularized quantities, regularized through a Hermitian regularization operator R. The end of this section contains a general discussion on the frame potential and its interpretation.

Let Π_t be the projector onto the symmetric subspace of $L^2(\mathbb{R}^n)^{\otimes t}$, and for a positive definite Hermitian trace-class operator R, define $\Pi_t^{(R)} = R^{\otimes t} \Pi_t R^{\otimes t}$. For any measure space $\mathcal{X}_n$ over $S'(\mathbb{R}^n)$, define

$$\begin{aligned} W_t^{(R)}(\mathcal{X}_n) &\coloneqq \int_{\mathcal{X}_n} d\mu_{\mathcal{X}_n}(X)\, d\mu_{\mathcal{X}_n}(Y)\, |\langle X|R|Y\rangle|^{2t} \\ X_t^{(R)}(\mathcal{X}_n) &\coloneqq \int_{\mathcal{X}_n} d\mu_{\mathcal{X}_n}(X) \left\langle X|R^2|X\right\rangle^t \\ \mathcal{F}_t^{(R)}(\mathcal{X}_n, a_t) &\coloneqq W_t^{(R)}(\mathcal{X}_n) - 2a_t X_t^{(R)}(\mathcal{X}_n) + a_t^2 \operatorname{Tr}\left[\Pi_t^{(R)}\right]. \end{aligned} \tag{B26}$$

The following proposition is a modification of [14, Prop. 6].

Proposition 1 ([14])**.** *For any ensemble $\mathcal{X}_n$ and any positive real number a_t, $\mathcal{F}_t^{(R)}(\mathcal{X}_n, a_t) \geq 0$. Furthermore, there exists positive real numbers $a_{t'}$ such that $\mathcal{F}_{t'}^{(R)}(\mathcal{X}_n, a_{t'}) = 0$ for each $t' = 1, 2, \ldots, t$ if and only if $\mathcal{X}_n$ is a rigged t-design.*

There are a few easy facts that we need:

$$\begin{aligned} \operatorname{Tr}\left[\Pi_1^{(R)}\right] &= \operatorname{Tr}\left[R^2\right] \\ \operatorname{Tr}\left[\Pi_2^{(R)}\right] &= \frac{1}{2}\left(\operatorname{Tr}\left[R^2\right]^2 + \operatorname{Tr}\left[R^4\right]\right) \\ \operatorname{Tr}\left[\Pi_3^{(R)}\right] &= \frac{1}{6}\left(\operatorname{Tr}\left[R^2\right]^3 + 3\left(\operatorname{Tr}\left[R^4\right]\right)\left(\operatorname{Tr}\left[R^2\right]\right) + 2\operatorname{Tr}\left[R^6\right]\right). \end{aligned} \tag{B27}$$

Finally, for the operator R, let $c_R(\boldsymbol{\alpha}) \coloneqq \operatorname{Tr}\left[D^\dagger(\boldsymbol{\alpha}) R\right]$ denote its characteristic function. We have

$$\begin{aligned} c_{R^2}(0) = \operatorname{Tr}\left[R^2\right] &= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha}\, |c_R(\alpha)|^2 \\ c_{R^4}(0) = \operatorname{Tr}\left[R^4\right] &= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha}_1 d\boldsymbol{\alpha}_2 d\boldsymbol{\alpha}_3\, c_R(\boldsymbol{\alpha}_1)\, c_R(\boldsymbol{\alpha}_2)\, c_R(\boldsymbol{\alpha}_3)\, c_R^*(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3)\, e^{-i\pi\left(\boldsymbol{\alpha}_1^T J \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_1^T J \boldsymbol{\alpha}_3 + \boldsymbol{\alpha}_2^T J \boldsymbol{\alpha}_3\right)}. \end{aligned} \tag{B28}$$

In a very long calculation we will now proceed to show the following theorem, using the proposition above. This calculation serves as independent verification that the ensemble $\mathcal{X}_n$ of GKP states discussed in the main text indeed yields a rigged 2-design.

Theorem 2. *The ensemble of GKP states $\mathcal{X}_n \subset \mathcal{S}'(\mathbb{R}^n)$ defined in the main text forms a rigged 2-design with $a_1 = 1$ and $a_2 = 2$.*

Proof. Using the frame potential, we can check if our ensemble $\mathcal{X}_n$ defined in the main text is a $t = 1, 2$-design. Specifically, we need to express $X_t^{(R)}(\mathcal{G})$ and $W_t^{(R)}(\mathcal{G})$ for each t in terms of $\mathrm{Tr}\left[R^2\right]$ and $\mathrm{Tr}\left[R^4\right]$.

Recall that elements of $\mathcal{X}_n$ are of the form $D(\boldsymbol{\alpha})|\Lambda\rangle$, where $|\Lambda\rangle\langle\Lambda| = \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} D(\boldsymbol{\lambda})$. In general, our computations will include expressions of the form

$$\left\langle\Lambda|D^\dagger(\boldsymbol{\alpha})\, RD(\boldsymbol{\alpha}')\,|\Lambda'\right\rangle = \sum_{\boldsymbol{\lambda}\in\Lambda}\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}\, c_R(\boldsymbol{\gamma})\, e^{i\Phi(\boldsymbol{\lambda})}\,\mathrm{Tr}\left[D(\boldsymbol{\lambda})\, D^\dagger(\boldsymbol{\alpha})\, D(\boldsymbol{\gamma})\, D(\boldsymbol{\alpha}')\right] \tag{B29}$$

$$= \sum_{\boldsymbol{\lambda}\in\Lambda}\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}\, c_R(\boldsymbol{\gamma})\, e^{i\Phi(\boldsymbol{\lambda})} e^{i\pi\boldsymbol{\lambda}^T J\boldsymbol{\alpha} - i\pi\boldsymbol{\gamma}^T J\boldsymbol{\alpha}'} \underbrace{\mathrm{Tr}\left[D(\boldsymbol{\lambda}-\boldsymbol{\alpha})\, D(\boldsymbol{\gamma}+\boldsymbol{\alpha}')\right]}_{\delta(\boldsymbol{\gamma}-\boldsymbol{\alpha}+\boldsymbol{\alpha}'+\boldsymbol{\lambda})} \tag{B30}$$

$$= \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} c_R(\boldsymbol{\alpha}-\boldsymbol{\alpha}'-\boldsymbol{\lambda})\, e^{i\pi\boldsymbol{\lambda}^T J\boldsymbol{\alpha} - i\pi(\boldsymbol{\alpha}-\boldsymbol{\alpha}'-\boldsymbol{\lambda})^T J\boldsymbol{\alpha}'}. \tag{B31}$$

If $\boldsymbol{\alpha} = \boldsymbol{\alpha}'$ this is

$$\left\langle\Lambda|D^\dagger(\boldsymbol{\alpha})\, RD(\boldsymbol{\alpha})\,|\Lambda\right\rangle = \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} c_R(-\boldsymbol{\lambda})\, e^{i2\pi\boldsymbol{\lambda}^T J\boldsymbol{\alpha}}. \tag{B32}$$

We begin with $t = 1$. We compute

$$X_1^{(R)}(\mathcal{X}_n) = \int_{Y_n} d\mu(\Lambda)\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\left\langle\Lambda|D^\dagger(\boldsymbol{\alpha})\, R^2 D(\boldsymbol{\alpha})\,|\Lambda\right\rangle \tag{B33}$$

$$= \int_{Y_n} d\mu(\Lambda)\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}\sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} e^{-i2\pi\boldsymbol{\alpha}^T J\boldsymbol{\lambda}} c_{R^2}(\boldsymbol{\gamma})\,\mathrm{Tr}\left[D(\boldsymbol{\lambda})\, D(\boldsymbol{\gamma})\right] \tag{B34}$$

$$= c_{R^2}(0) = \mathrm{Tr}\left[R^2\right]. \tag{B35}$$

Similarly, we have

$$W_1^{(R)}(\mathcal{X}_n) = \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda')\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\int_{\mathcal{P}(\Lambda)'} d\boldsymbol{\alpha}'\,\left\langle\Lambda|D^\dagger(\boldsymbol{\alpha})\, RD(\boldsymbol{\alpha}')\,|\Lambda'\right\rangle\left\langle\Lambda'|D^\dagger(\boldsymbol{\alpha}')\, RD(\boldsymbol{\alpha})\,|\Lambda\right\rangle \tag{B36}$$

$$= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda')\sum_{\boldsymbol{\lambda}\in\Lambda}\sum_{\boldsymbol{\lambda}'\in\Lambda'}\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\int_{\mathcal{P}(\Lambda)'} d\boldsymbol{\alpha}'\, e^{i\Phi(\boldsymbol{\lambda})} e^{i\Phi(\boldsymbol{\lambda}')} e^{-i2\pi\boldsymbol{\alpha}^T J\boldsymbol{\lambda} - i2\pi\boldsymbol{\alpha}'^T J\boldsymbol{\lambda}'}\,\mathrm{Tr}\left[D(\boldsymbol{\lambda})\, RD(\boldsymbol{\lambda}')\, R\right] \tag{B37}$$

$$= \mathrm{Tr}\left[R^2\right], \tag{B38}$$

where we have again used Eq. (B10). Altogether, this yields

$$\mathcal{F}_1^{(R)}(\mathcal{X}_n, a_1) = \mathrm{Tr}\left[R^2\right] - 2a_1\,\mathrm{Tr}\left[R^2\right] + a_1^2\,\mathrm{Tr}\left[R^2\right] \tag{B39}$$

We see that, with $a_1 = 1$, we obtain $\mathcal{F}_1^{(R)}(\mathcal{X}_n, a_1 = 1) = 0$.

Now, on to $t = 2$. Here we have

$$X_2^{(R)}(\mathcal{X}_n) = \int_{Y_n} d\mu(\Lambda)\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\left\langle\Lambda|D^\dagger(\boldsymbol{\alpha})\, R^2 D(\boldsymbol{\alpha})\,|\Lambda\right\rangle^2 \tag{B40}$$

$$= \int_{Y_n} d\mu(\Lambda)\sum_{\boldsymbol{\lambda},\boldsymbol{\lambda}'\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})-i\Phi(\boldsymbol{\lambda}')} c_{R^2}(-\boldsymbol{\lambda})\, c_{R^2}(\boldsymbol{\lambda}')\underbrace{\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, e^{i2\pi(\boldsymbol{\lambda}-\boldsymbol{\lambda}')^T J\boldsymbol{\alpha}}}_{\delta_{\boldsymbol{\lambda},\boldsymbol{\lambda}'}} \tag{B41}$$

$$= \int_{Y_n} d\mu(\Lambda)\sum_{\boldsymbol{\lambda}\in\Lambda}|c_{R^2}(\boldsymbol{\lambda})|^2 \tag{B42}$$

$$= |c_{R^2}(0)|^2 + \int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda}\,|c_{R^2}(\boldsymbol{\lambda})|^2 \tag{B43}$$

$$= \mathrm{Tr}\left[R^2\right]^2 + \mathrm{Tr}\left[R^4\right]. \tag{B44}$$

We also need to compute

$$
\begin{aligned}
W_2^{(R)}(\mathcal{X}_n) &= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} \int_{\mathcal{P}(\Lambda)'} d\boldsymbol{\alpha}' \left|\langle\Lambda| D^\dagger(\boldsymbol{\alpha})\, RD(\boldsymbol{\alpha}')\, |\Lambda'\rangle\right|^4 \\
&= \ldots \langle\Lambda| D^\dagger(\boldsymbol{\alpha})\, RD(\boldsymbol{\alpha}')\, |\Lambda'\rangle \langle\Lambda'| D^\dagger(\boldsymbol{\alpha}')\, RD(\boldsymbol{\alpha})\, |\Lambda\rangle \langle\Lambda| D^\dagger(\boldsymbol{\alpha})\, RD(\boldsymbol{\alpha}')\, |\Lambda'\rangle \langle\Lambda'| D^\dagger(\boldsymbol{\alpha}')\, RD(\boldsymbol{\alpha})\, |\Lambda\rangle \\
&= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 d\boldsymbol{\gamma}_4\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R(\boldsymbol{\gamma}_4) \sum_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2\in\Lambda;\boldsymbol{\lambda}_1',\boldsymbol{\lambda}_2'\in\Lambda'} \\
&\qquad \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} \int_{\mathcal{P}(\Lambda)'} d\boldsymbol{\alpha}' e^{-i2\pi\boldsymbol{\alpha}'^T J\left(\boldsymbol{\lambda}_1'+\boldsymbol{\lambda}_2'\right)-i2\pi\boldsymbol{\alpha}^T J(\boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2)} \qquad \text{(B45)} \\
&\qquad \mathrm{Tr}\left[D(\boldsymbol{\gamma}_1)\, D(\boldsymbol{\lambda}_1')\, D(\boldsymbol{\gamma}_2)\, D(\boldsymbol{\lambda}_1)\, D(\boldsymbol{\gamma}_3)\, D(\boldsymbol{\lambda}_2')\, D(\boldsymbol{\gamma}_4)\, D(\boldsymbol{\lambda}_2)\right] \\
&= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 d\boldsymbol{\gamma}_4\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R(\boldsymbol{\gamma}_4) \sum_{\boldsymbol{\lambda}\in\Lambda;\boldsymbol{\lambda}'\in\Lambda'} \\
&\qquad \mathrm{Tr}\left[D(\boldsymbol{\gamma}_1)\, D(\boldsymbol{\lambda}')\, D(\boldsymbol{\gamma}_2)\, D(\boldsymbol{\lambda})\, D(\boldsymbol{\gamma}_3)\, D^\dagger(\boldsymbol{\lambda}')\, D(\boldsymbol{\gamma}_4)\, D^\dagger(\boldsymbol{\lambda})\right] \qquad \text{(B46)} \\
&= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 d\boldsymbol{\gamma}_4\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R(\boldsymbol{\gamma}_4) \sum_{\boldsymbol{\lambda}\in\Lambda;\boldsymbol{\lambda}'\in\Lambda'} \\
&\qquad \underbrace{\mathrm{Tr}\left[D(\boldsymbol{\gamma}_1)\, D(\boldsymbol{\gamma}_2)\, D(\boldsymbol{\gamma}_3)\, D(\boldsymbol{\gamma}_4)\right]}_{e^{-i\pi\boldsymbol{\gamma}_1^T J\boldsymbol{\gamma}_2} e^{-i\pi\boldsymbol{\gamma}_3^T J\boldsymbol{\gamma}_4} \delta(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3+\boldsymbol{\gamma}_4)} e^{-i2\pi\boldsymbol{\lambda}^T J\left(\boldsymbol{\gamma}_3+\boldsymbol{\gamma}_4-\boldsymbol{\lambda}'\right)} e^{-i2\pi\boldsymbol{\lambda}'^T J(\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3)} \qquad \text{(B47)} \\
&= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R^*(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3) \sum_{\boldsymbol{\lambda}\in\Lambda;\boldsymbol{\lambda}'\in\Lambda'} \\
&\qquad e^{-i\pi\boldsymbol{\gamma}_1^T J\boldsymbol{\gamma}_2} e^{-i\pi\boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1-\boldsymbol{\gamma}_2)} e^{i2\pi\boldsymbol{\lambda}^T J\left(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\lambda}'\right)} e^{-i2\pi\boldsymbol{\lambda}'^T J(\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3)} \qquad \text{(B48)}
\end{aligned}
$$

We have [74],

$$
\int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-i2\pi\boldsymbol{\lambda}^T J\boldsymbol{x}} = 1 + \delta(\boldsymbol{x}), \qquad \text{(B49)}
$$

such that we can further simplify

$$
\begin{aligned}
\ldots &= \int_{Y_n} d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R^*(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3) \\
&\qquad \sum_{\boldsymbol{\lambda}'\in\Lambda'} e^{-i\pi\boldsymbol{\gamma}_1^T J\boldsymbol{\gamma}_2} e^{-i\pi\boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1-\boldsymbol{\gamma}_2)} \left\{1+\delta(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\lambda}')\right\} e^{-i2\pi\boldsymbol{\lambda}'^T J(\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3)} \qquad \text{(B50)} \\
&= \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R^*(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3) \\
&\qquad e^{-i\pi\boldsymbol{\gamma}_1^T J\boldsymbol{\gamma}_2} e^{-i\pi\boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1-\boldsymbol{\gamma}_2)} \left\{1+\delta(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2)\right\} \\
&+ \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R^*(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3) \\
&\quad e^{-i\pi\boldsymbol{\gamma}_1^T J\boldsymbol{\gamma}_2} e^{-i\pi\boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1-\boldsymbol{\gamma}_2)} \int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda}' \left\{1+\delta(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\lambda}')\right\} e^{-i2\pi\boldsymbol{\lambda}'^T J(\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3)} \qquad \text{(B51)} \\
&= \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R^*(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3) \\
&\qquad e^{-i\pi\boldsymbol{\gamma}_1^T J\boldsymbol{\gamma}_2} e^{-i\pi\boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1-\boldsymbol{\gamma}_2)} \left\{1+\delta(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2)\right\} \\
&+ \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R^*(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3) \\
&\quad e^{-i\pi\boldsymbol{\gamma}_1^T J\boldsymbol{\gamma}_2} e^{i\pi\boldsymbol{\gamma}_3^T J(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2)} \left\{\delta(\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3) + e^{i2\pi(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2)^T J(\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3)}\right\} \qquad \text{(B52)}
\end{aligned}
$$

This final sum has 4 parts, the first equals $\mathrm{Tr}\left[R^4\right]$. The second is

$$\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_3 \left|c_R\left(\boldsymbol{\gamma}_1\right)\right|^2 \left|c_R\left(\boldsymbol{\gamma}_3\right)\right|^2 = \mathrm{Tr}\left[R^2\right]^2, \tag{B53}$$

same as the third

$$\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 \left|c_R\left(\boldsymbol{\gamma}_1\right)\right|^2 \left|c_R\left(\boldsymbol{\gamma}_2\right)\right|^2 = \mathrm{Tr}\left[R^2\right]^2. \tag{B54}$$

Finally, the last term is computed as

$$\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3\, c_R\left(\boldsymbol{\gamma}_1\right) c_R\left(\boldsymbol{\gamma}_2\right) c_R\left(\boldsymbol{\gamma}_3\right) c_R^*\left(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3\right) e^{-i\pi \boldsymbol{\gamma}_1^T J \boldsymbol{\gamma}_2} e^{i\pi \boldsymbol{\gamma}_3^T J\left(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2\right)} e^{i2\pi\left(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2\right)^T J\left(\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3\right)} \tag{B55}$$

$$= \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_4\, c_R\left(\boldsymbol{\gamma}_1\right) c_R\left(\boldsymbol{\gamma}_2\right) c_R^*\left(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_4\right) c_R\left(\boldsymbol{\gamma}_4\right) e^{-i\pi \boldsymbol{\gamma}_1^T J \boldsymbol{\gamma}_2} e^{-i\pi \boldsymbol{\gamma}_4^T\left(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2\right)} = \mathrm{Tr}\left[R^4\right]. \tag{B56}$$

Thus in total, we have

$$W_2^{(R)}(\mathcal{X}_n) = 2\,\mathrm{Tr}\left[R^4\right] + 2\,\mathrm{Tr}\left[R^2\right]^2 \tag{B57}$$

and

$$\mathcal{F}_2^{(R)}(\mathcal{X}_n, a_2) = 2\,\mathrm{Tr}\left[R^4\right] + 2\,\mathrm{Tr}\left[R^2\right]^2 - 2a_2\left\{\mathrm{Tr}\left[R^2\right]^2 + \mathrm{Tr}\left[R^4\right]\right\} + \frac{a_t^2}{2}\left\{\mathrm{Tr}\left[R^2\right]^2 + \mathrm{Tr}\left[R^4\right]\right\}. \tag{B58}$$

From this equation, we can observe that we have, with $a_2 = 2$, $\mathcal{F}_2^{(R)}(\mathcal{X}_n, a_2 = 2) = 0$. Under the premise of prop. 1 we hence have verified that $\mathcal{X}_n$ is a rigged 2-design.

□

The meaning of the frame potential

We now give a basic interpretation of the frame potential in *finite-dimensional* systems. The frame potential has its roots in Welch bounds [75, 76]. There are many great expositions of the frame potential, its relation to designs, mutually unbiased bases, and scrambling and chaos [77–81]. Here, we give a slightly different argument for its applicability.

Ultimately, we want to characterize some ensemble $\mathcal{X}_n$ of states via a number $F(\mathcal{X}_n) \in \mathbb{R}$. Because states are elements of projective space, this function F must be a function of projectors rather than states. Additionally, there is no ordering on $\mathcal{X}_n$, meaning $F(\mathcal{X}_n)$ should be independent of reordering or relabeling the elements of $\mathcal{X}_n$. We are thus led to

$$F_t^{(k)}(\mathcal{X}_n) = \mathrm{Tr}\,\mathbb{E}_{\psi_1,\ldots,\psi_k\in\mathcal{X}_n}\left(|\psi_1\rangle\langle\psi_1|\right)^{\otimes t}\ldots\left(|\psi_k\rangle\langle\psi_k|\right)^{\otimes t} = \mathbb{E}_{\psi_1,\ldots,\psi_k\in\mathcal{X}_n}\langle\psi_1|\psi_2\rangle^t\langle\psi_2|\psi_3\rangle^t\ldots\langle\psi_k|\psi_1\rangle^t, \tag{B59}$$

where we can tune t to probe different moments of $\mathcal{X}_n$.

We then notice that $F_t^{(1)}(\mathcal{X}_n) = 1$ for any $\mathcal{X}_n$ (because in finite-dimensions we can always assume the states are unit-normalized), and so is trivial. The first nontrivial function that can characterize the t^{th} moment of $\mathcal{X}_n$ is thus $F_t^{(2)}(\mathcal{X}_n)$. $F_t^{(2)}(\mathcal{X}_n)$ is in fact the standard frame potential for finite-dimensional systems, and as such we will notate $F_t(\mathcal{X}_n) \equiv F_t^{(2)}(\mathcal{X}_n)$.

If we are interested in measuring how distributed an ensemble of states is (*i.e.* how much it "covers" the full space of states and how spread it is over the space), we want the average magnitude of the overlap between the states in the ensemble to be small. Therefore, the $F_t(\mathcal{X}_n)$ being small is a measure of how evenly spread $\mathcal{X}_n$ is. This then gives some intuition as to why the frame potential shows up as a measure of scrambling and chaos.

When thinking about t-designs, we need ensembles $\mathcal{X}_n$ that maximally cover the whole space of states in order for them to mimic the whole space itself. Therefore, $\mathcal{X}_n$ is a t-design if and only if $F_t(\mathcal{X}_n)$ is minimized (the proof of this is analogous to the proof of Prop. 1). When $\mathcal{X}_n$ does mimic the t^{th} moment of the whole space of states, then clearly $F_t^{(k)}(\mathcal{X}_n)$ is determined for all k. Thus, we have found that even though $F_t^{(k)}(\mathcal{X}_n)$ in principle contains more information about the t^{th} moment of $\mathcal{X}_n$ than just $F_t(\mathcal{X}_n) = F_t^{(2)}(\mathcal{X}_n)$, $F_t(\mathcal{X}_n)$ alone already characterizes how far $\mathcal{X}_n$ is from being sufficiently spread over the whole space of states. In particular, via some continuity analysis, $F_t^{(k)}(\mathcal{X}_n)$ can be bounded in terms of how far $F_t(\mathcal{X}_n)$ is from its minimum, and this bound gets tighter as $F_t(\mathcal{X}_n)$ approaches its minimum.

In the CV case, $W_t^{(R)}(\mathcal{X}_n)$ is the analogue of $F_t(\mathcal{X}_n)$. The function $X_t^{(R)}(\mathcal{X}_n)$ is an artifact of the fact that the states are not normalized. In the case of regularized rigged designs, $X_t^{(R)}(\mathcal{X}_n)$ can in fact be gotten rid of [14, Prop. 6], making the analogy between the finite- and infinte-dimensional frame potentials more direct.

4. Diagonal coherent state basis

In this appendix, we provide another proof that the displaced GKP ensemble forms a rigged 2-design by working in the coherent state basis (a.k.a. the P-representation). We then perform the analogous 3-design calculation, but are unfortunately unable to complete it.

The key point is that the diagonal elements of an operator in the coherent state basis fully determine the operator [56, 59]. Hence, in order to check if an ensemble $\mathcal{X}_n$ is a rigged t-design, we only need to check that

$$\int_{\mathcal{X}_n} dX \prod_{i=1}^{t} |\langle \boldsymbol{\alpha}_i | X \rangle|^2 = a_t \, \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t | \Pi_t | \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t \rangle \tag{B60}$$

for some $a_t > 0$. For $\mathcal{X}_n$ the ensemble of GKP codes defined in the main text, this amounts to showing that

$$g(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t) = a_t \, \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t | \, \Pi_t \, | \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t \rangle \, , \tag{B61}$$

where

$$\begin{aligned}
g(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t) &\coloneqq \int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\beta} \prod_{i=1}^{t} \mathrm{Tr} \left[D(\boldsymbol{\beta}) \, |\Lambda\rangle \, \langle \Lambda| \, D(-\boldsymbol{\beta}) D(\boldsymbol{\alpha}_i) \, |0\rangle \, \langle 0| \, D(-\boldsymbol{\alpha}_i) \right] \\
&= \int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\beta} \prod_{i=1}^{t} \sum_{\boldsymbol{\lambda}_i \in \Lambda} e^{i\Phi(\boldsymbol{\lambda}_i)} \, \mathrm{Tr} \left[D(\boldsymbol{\beta}) D(\boldsymbol{\lambda}_i) D(-\boldsymbol{\beta}) D(\boldsymbol{\alpha}_i) \, |0\rangle \, \langle 0| \, D(-\boldsymbol{\alpha}_i) \right] \\
&= \int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\beta} \prod_{i=1}^{t} \sum_{\boldsymbol{\lambda}_i \in \Lambda} e^{i\Phi(\boldsymbol{\lambda}_i)} \, \langle 0| \, D(-\boldsymbol{\alpha}_i) D(\boldsymbol{\beta}) D(\boldsymbol{\lambda}_i) D(-\boldsymbol{\beta}) D(\boldsymbol{\alpha}_i) \, |0\rangle \\
&= \int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\beta} \prod_{i=1}^{t} \sum_{\boldsymbol{\lambda}_i \in \Lambda} e^{i\Phi(\boldsymbol{\lambda}_i)} e^{2\pi i \boldsymbol{\lambda}_i^T J (\boldsymbol{\beta} - \boldsymbol{\alpha}_i)} \, \langle 0| \, D(\boldsymbol{\lambda}_i) \, |0\rangle \\
&\overset{\text{Eq. (B10)}}{=} \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_t \in \Lambda \\ \text{s.t. } \boldsymbol{\lambda}_1 + \cdots + \boldsymbol{\lambda}_t = 0}} e^{i \sum_{j=1}^{t} \Phi(\boldsymbol{\lambda}_j)} e^{2\pi i \sum_{j=1}^{t} \boldsymbol{\alpha}_j^T J \boldsymbol{\lambda}_j} \prod_{j=1}^{t} \langle 0 | \boldsymbol{\lambda}_j \rangle \\
&= \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_t \in \Lambda \\ \text{s.t. } \boldsymbol{\lambda}_1 + \cdots + \boldsymbol{\lambda}_t = 0}} e^{i \sum_{j=1}^{t} \Phi(\boldsymbol{\lambda}_j)} e^{2\pi i \sum_{j=1}^{t} \boldsymbol{\alpha}_j^T J \boldsymbol{\lambda}_j} e^{-\frac{1}{2} \sum_{j=1}^{t} ||\boldsymbol{\lambda}_j||^2} .
\end{aligned} \tag{B62}$$

The RHS of Eq. (B60) yields

$$\begin{aligned}
\Pi_1 = I \qquad &\implies \langle \boldsymbol{\alpha} | \, \Pi_1 \, | \boldsymbol{\alpha} \rangle = 1 \\
\Pi_2 = \frac{1}{2}(I \otimes I + \mathrm{SWAP}) \qquad &\implies \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 | \, \Pi_2 \, | \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \rangle = \frac{1}{2} + \frac{1}{2} |\langle \boldsymbol{\alpha}_1 | \boldsymbol{\alpha}_2 \rangle|^2 .
\end{aligned} \tag{B63}$$

Furthermore, using $\Pi_3 = \frac{1}{6} \sum_{\sigma \in S_3} W_\sigma$, where W_σ permutes the tensor factors, we also have

$$\langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3 | \, \Pi_3 \, | \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3 \rangle = \frac{1}{6} \left[1 + |\langle \boldsymbol{\alpha}_1 | \boldsymbol{\alpha}_2 \rangle|^2 + |\langle \boldsymbol{\alpha}_1 | \boldsymbol{\alpha}_3 \rangle|^2 + |\langle \boldsymbol{\alpha}_2 | \boldsymbol{\alpha}_3 \rangle|^2 + 2 \, \mathrm{Re} \left(\langle \boldsymbol{\alpha}_1 | \boldsymbol{\alpha}_2 \rangle \, \langle \boldsymbol{\alpha}_2 | \boldsymbol{\alpha}_3 \rangle \, \langle \boldsymbol{\alpha}_3 | \boldsymbol{\alpha}_1 \rangle \right) \right] . \tag{B64}$$

For $t = 1$, we have

$$\begin{aligned}
g(\boldsymbol{\alpha}) &= \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda} \in \Lambda \\ \text{s.t. } \boldsymbol{\lambda} = 0}} e^{i\Phi(\boldsymbol{\lambda})} e^{2\pi i \boldsymbol{\alpha}^T J \boldsymbol{\lambda}} e^{-||\boldsymbol{\lambda}||^2/2} \\
&= 1 \\
&= \langle \boldsymbol{\alpha} | \, \Pi_1 \, | \boldsymbol{\alpha} \rangle \, ,
\end{aligned} \tag{B65}$$

therefore showing that the displaced GKP ensemble forms a 1-design with $a_1 = 1$.

For $t = 2$, we have

$$\begin{aligned} g(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) &= \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \Lambda \\ \text{s.t. } \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 = 0}} e^{i \sum_{j=1}^{2} \Phi(\boldsymbol{\lambda}_j)} e^{2\pi i \sum_{j=1}^{2} \boldsymbol{\alpha}_j^T J \boldsymbol{\lambda}_j} e^{-\frac{1}{2} \sum_{j=1}^{2} ||\boldsymbol{\lambda}_j||^2} \\ &= \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda} \in \Lambda} e^{2\pi i \boldsymbol{\lambda}^T J (\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_1)} e^{-||\boldsymbol{\lambda}||^2} \\ &= 1 + \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda} \in \Lambda \setminus \{0\}} e^{2\pi i \boldsymbol{\lambda}^T J (\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_1)} e^{-||\boldsymbol{\lambda}||^2} \\ &\overset{\text{lem. 1}}{=} 1 + \int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda}\, e^{2\pi i \boldsymbol{\lambda}^T J (\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_1)} e^{-||\boldsymbol{\lambda}||^2} \\ &= 1 + e^{-||\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_1||^2} \\ &= 1 + |\langle \boldsymbol{\alpha}_1 | \boldsymbol{\alpha}_2 \rangle|^2 \\ &= 2 \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 | \, \Pi_2 \, | \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \rangle \, . \end{aligned} \tag{B66}$$

This therefore shows that the displaced GKP ensemble forms a rigged 2-design with $a_2 = 2$.

The analogous calculation for $t = 3$ requires a second moment analogue of Lem. 1 beyond the extension provided in Ref. [40]. For $t = 3$, we have

$$\begin{aligned} g(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3) &= \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3 \in \Lambda \\ \text{s.t. } \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \boldsymbol{\lambda}_t = 0}} e^{i \sum_{j=1}^{t} \Phi(\boldsymbol{\lambda}_j)} e^{2\pi i \sum_{j=1}^{t} \boldsymbol{\alpha}_j^T J \boldsymbol{\lambda}_j} e^{-\frac{1}{2} \sum_{j=1}^{3} ||\boldsymbol{\lambda}_j||^2} \\ &= \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \Lambda} e^{2\pi i \left(\boldsymbol{\alpha}_1^T J \boldsymbol{\lambda}_1 + \boldsymbol{\alpha}_2^T J \boldsymbol{\lambda}_2 - \boldsymbol{\alpha}_3^T J (\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2) \right)} e^{-||\boldsymbol{\lambda}_1||^2/2 - ||\boldsymbol{\lambda}_2||^2/2 - ||\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2||^2/2} \\ &= \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \Lambda} e^{2\pi i \left(\boldsymbol{\lambda}_1^T J (\boldsymbol{\alpha}_3 - \boldsymbol{\alpha}_1) + \boldsymbol{\lambda}_2^T J (\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_1) \right)} e^{-||\boldsymbol{\lambda}_1||^2/2 - ||\boldsymbol{\lambda}_2||^2/2 - ||\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2||^2/2} . \end{aligned} \tag{B67}$$

To complete this calculation, we need an expression for the second moment integral over Y_n. Once this is known, we would need to compare the result to Eq. (B64).

Appendix C: Design-based CV shadow tomography with GKP states

It was shown in Ref. [14] that a rigged 2-design immediately gives rise to a classical shadow tomography protocol for states on the CV Hilbert space. A distinctive feature of the classical shadow tomography protocol is that it constructs a set of POVM elements, whose measurement statistics suffice to reconstruct expectation values of many observables with specific properties while utilizing only a relatively small number of samples [8].

1. Shadow tomography from 2-designs

We recall the CV classical shadow procedure proposed in Ref. [14]. Let $(\mathcal{X}_n, \mu_{\mathcal{X}_n})$ form a set of (improper) states $|X\rangle\langle X|$, $|X\rangle \in \mathcal{X}_n$ together with a normalized measure $\mu_{\mathcal{X}_n} : \mathcal{X}_n \to \mathbb{R}$, such that it forms a rigged 2-design. Then measuring a state ρ with respect to $|X\rangle\langle X|$ yields the POVM

$$\begin{aligned} \mathcal{M}_{\mathcal{X}_n}(\rho) &= \int_{\mathcal{X}_n} d\mu_{\mathcal{X}_n}(X) \, \langle X|\rho|X\rangle \, |X\rangle\langle X| \\ &= \int_{\mathcal{X}_n} d\mu_{\mathcal{X}_n}(X) \, \mathrm{Tr}_1 \left[(\rho \otimes I) \, |X\rangle\langle X|^{\otimes 2} \right] \\ &= \mathrm{Tr}_1 \left[(\rho \otimes I) \left\{ I \otimes I + \widehat{E} \right\} \right] \\ &= I + \rho, \end{aligned} \tag{C1}$$

where we inserted the 2-design property in the third line. Note, that in this expression it is not actually necessary to think about the $|X\rangle\langle X|$ as physical states, instead they can be regarded as the classical pointer of the associated measurement outcome that

was obtained with probability distribution $\mathrm{Tr}\left[|X\rangle\langle X|\,\rho\right] = \langle X|\rho|X\rangle$. Inverting this expression as a function of ρ, we find that expectation values of observables O are obtained as

$$\begin{aligned}\langle O\rangle &= \mathrm{Tr}\left[\left(\int_{\mathcal{X}_n} d\mu_{\mathcal{X}_n}(X)\,\langle X|\rho|X\rangle\,|X\rangle\langle X| - I\right)O\right] \\ &= \int_{\mathcal{X}_n} d\mu_{\mathcal{X}_n}(X)\,\langle X|\rho|X\rangle\,\langle X|O|X\rangle - \mathrm{Tr}\left[O\right].\end{aligned} \tag{C2}$$

In this recipe, for each sampled set of measurements X_i, $i = 1, 2, \cdots, N$, the ensemble of reconstructed "states"

$$\{\tilde{\rho}_i = |X\rangle\langle X|_i - I\}_{i=1}^{N} \tag{C3}$$

is regarded as the classical shadow of the input quantum state. Again, here we notice that it is in fact not relevant whether one can regard the expression $|X\rangle\langle X|$ as physical states. The only elements necessary for the classical shadow tomography protocol to function are that they form rigged 2-designs, that a quantum POVM exists with measurement probability distribution $P(X|\rho) = \langle X|\rho|X\rangle$ exists for input states ρ, and that the expression $\langle X|O|X\rangle$ is sufficiently well behaved to be evaluated under the above integral.

2. Local shadows vs. global shadows

In Clifford-based classical shadow tomography for discrete variable systems [8] the ensemble of states used to build the shadow protocol is typically chosen to be a random set of stabilizer states obtained from Clifford-twirling a Pauli measurement channel. In this construction, the Clifford-twirled measurement channel is projected onto a depolarizing channel such that the outcome statistics over the stabilizer-state-like pointer stats reflects that of the depolarized input state. This depolarizing channel shortens the length of the Bloch vector of the input state by a factor of $f = (2^n + 1)^{-1}$ [82] and to reproduce the expectation values of the input state faithfully, in classical post-processing of the output stabilizer states, one needs to reamplify by this factor. While this re-amplification process is necessary to obtain the correct expectation values, it also amplifies the variance of the estimator to scale exponentially in the system (or subsystem) size n. It is for this reason that a distinction is typically being made between *local* and *global* classical shadows and, for practical purposes, the tomography protocol is typically restricted to be only applied locally. Interestingly, when the POVM used is a Pauli-Z measurement on each of the qubits, it is known that this depolarizing channel with probability $p = 1/3$ that is projected onto is *entanglement breaking* despite remaining classically invertible as a matrix. Via the classical-post processing protocol of classical shadow tomography, this then allows to estimate the expectation values of non-local observables of entangled states by passing through their depolarized, unentangled shadows [8]. The price to pay for this trick, however, is an exponential overhead in the the degree of locality k of the observable, which enters the variance upper bounds with a factor of f^{-k}, where $f = 1/3$ in the case of qubit systems.

From Eq. (C1) we recognize that the tensor power of the random measurement channel over a local ensemble $\mathcal{X}_1$ randomly deletes the supported mode or preserves it. Similar to the strategy in Ref. [8], it is possible to define a local shadow protocol by reconstructing the shadow state

$$\bigotimes_{i=1}^{n}\left(|X_i\rangle\langle X_i| - I\right) \tag{C4}$$

upon sampling outcome X_i on each mode $i = 1, ..., n$. As, together with the inversion, each local channel $\mathcal{M}_{\mathcal{X}_1}^{-1}|X\rangle\langle X|\cdot|X\rangle\langle X|$ averages to the identity channel, we see that again the full expectation value of the state can be reconstructed.

An interesting observation that can be made in Eq. (C1) is that, in our GKP-state based shadow protocol, the corresponding factor is $f = 1$. As the tomographic data obtained from measuring relative to the ensemble of GKP states $\mathcal{X}_n$ does not correspond to a depolarized version of the input state but reproduces it exactly on average up to the constant identity offset. Note that the output states of the channel in Eq. (C1) are not physical, such that the POVM is generally not a trace preserving channel. We have shown that the interpretation of the classical shadow as an actual quantum state in our protocol is not necessary to reproduce meaningful expectation values for the given observables, which ultimately circumvents the appearance of a factor $f < 1$ as discussed above. The implication is interesting: Estimating non-local observables using local ($n = 1$) shadows as per our protocol is a-priori not accompanied by a variance-, and sample overhead -blow up in the locality of the observable originating from the need to reamplify the result. Yet, in the bounds we will be able to provide, we still observe a non-trivial scaling of the variance in the number of modes n. We investigate this further as follows.

3. Global and local design-based shadow tomography for 3-designs

In the previous sections, we have leveraged the rigged 2-design property of our ensemble $\mathcal{X}_n$ together with knowledge about the specific structure of the ensemble and/or additional physical assumptions in order to bound the variance scaling. While we have not been able to show whether a stronger 3-design property also holds for our ensemble, we explore what this property would imply. The following discussion applies to design-based classical shadow tomography protocols for either finite or inifinite systems (via the definition of rigged designs).

Throughout this section, we will always assume that O is traceless in order to simplify calculations. Assume that we have a measure space $(\mathcal{X}, \mu)$ that is a 3-design. In other words, for all $t \leq 3$, $\int_{\mathcal{X}} X^{\otimes t} d\mu(X) = a_t \Pi_t$, where each $X \in \mathcal{X}$ is a positive operator, $a_1 = 1$, and $a_t > 0$. A rigged 3-design in infinite-dimensions and a quantum state 3-design in finite dimensions satisfies exactly this criteria. In particular, in d dimensions for d finite, if we choose $a_t = d/\operatorname{Tr}(\Pi_t)$ (ie $a_2 = 2/(d+1)$ and $a_3 = 6/(d+1)(d+2)$) and require each X to be rank 1, then we get the standard definition of a quantum state 3-design. In infinite-dimensions, if we allow a_2, a_3 to be arbitrary constants and require each X to be rank 1, then we get the definition of a rigged 3-design.

a. *Global shadows*

Define

$$M(\rho) = \int_{\mathcal{X}} \operatorname{Tr}[\rho X] X \, \mathrm{d}\mu(X) = \frac{a_2}{2}(I + \rho). \tag{C5}$$

Suppose that we can sample from the probability distribution $dP(X, \rho) = \operatorname{Tr}(\rho X) \, \mathrm{d}\mu(X)$. Given a sample X from this distribution, we have the estimator $o(X) = \frac{2}{a_2} \operatorname{Tr}[XO]$. The expectation value of our estimator is

$$\begin{aligned}
\mathbb{E}_{X \sim dP}[o(X)] &= \int_{\mathcal{X}} o(X) \, dP(X) \\
&= \frac{2}{a_2} \int_{\mathcal{X}} \operatorname{Tr}[OX] \operatorname{Tr}[\rho X] \, d\mu(X) \\
&= 2 \operatorname{Tr}[\Pi_2(\rho \otimes O)] \\
&= \operatorname{Tr}[\rho O]
\end{aligned} \tag{C6}$$

The variance of our estimator is

$$\begin{aligned}
\operatorname{Var}_{X \sim dP}[o(X)] &= \int_{\mathcal{X}} o(X)^2 \, dP(X) \; - \operatorname{Tr}[\rho O]^2 \\
&= \frac{4}{a_2^2} \int_{\mathcal{X}} \operatorname{Tr}[XO]^2 \operatorname{Tr}[\rho X] \, d\mu(X) \; - \operatorname{Tr}[\rho O]^2 \\
&= \frac{4a_3}{a_2^2} \operatorname{Tr}[\Pi_3(O \otimes O \otimes \rho)] - \operatorname{Tr}[\rho O]^2 \\
&\leq \frac{2a_3}{a_2^2} \operatorname{Tr}[O^2],
\end{aligned} \tag{C7}$$

where the last step is executed using Eq. (E9) in Ref. [14]. We see that the efficiency of a global shadow tomography protocol with a 3-design depends on $2a_3/a_2^2$. In finite dimension d, when using quantum state 3-designs, the scaling dependencies of this fraction on d cancel and it follows that the variance of the corresponding estimator is $\sim \operatorname{Tr}[O^2]$. In infinite dimensions, a rigged 3-design satisfies $a_2, a_3 \sim 1$, and thus again we obtain a variance scaling of $\sim \operatorname{Tr}[O^2]$.

We highlight that the factor a_2 alone is not important, but rather the ratio $2a_3/a_2^2$. In general, even in the displaced GKP case (where we are unsure if the ensemble forms a rigged 3-design), the fact that $a_2 \sim 1$ (as opposed to $a_2 \sim 2^{-n}$ as is typically the case in qubit protocols) is not the key difference dictating the variance scaling of the corresponding shadow tomography protocol. Ultimately, what matters is the ratio $2a_3/a_2^2$. For finite d and utilizing quantum state 3-designs, we have

$$a_2 = \frac{2}{d+1} \quad \text{and} \quad a_3 = \frac{6}{(d+1)(d+2)}, \tag{C8}$$

yielding $\frac{2a_3}{a_2^2} \xrightarrow{d=2^n \to \infty} 3$, which is precisely the scaling observed in Ref. [8].

b. Local shadows

In this section, we will assume that our observable O is of form $O = O_1 \otimes \cdots \otimes O_n$, and we assume that $\mathrm{Tr}[O] = 0$ (for example, in finite dimensions, O can be a Pauli observable). We expand our state as $\rho = \sum_\alpha c_\alpha \rho_1^{(\alpha)} \otimes \cdots \otimes \rho_n^{(\alpha)}$, with $\mathrm{Tr}[\rho_i^{(\alpha)}] = 1$ for each i and α. We sample from the ensemble $\mathcal{X} = \mathcal{X}_1^{\otimes n}$ with measure $dP(X_1 \otimes \cdots \otimes X_n)$ defined by

$$dP(X_1 \otimes \cdots \otimes X_n) = \mathrm{Tr}[\rho(X_1 \otimes \cdots \otimes X_n)]\ d\mu(X_1) \dots d\mu(X_n). \tag{C9}$$

Our estimator is

$$o(X_1 \otimes \cdots \otimes X_n) = \left(\frac{2}{a_2}\right)^n \mathrm{Tr}[X_1 O_1] \dots \mathrm{Tr}[X_n O_n]. \tag{C10}$$

The expectation value of our estimator is then

$$\begin{aligned}
\mathbb{E}_{X_1 \otimes \cdots \otimes X_n \sim dP}[o(X_1 \otimes \cdots \otimes X_n)] &= \int_{\mathcal{X}_1^{\otimes n}} o(X_1 \otimes \cdots \otimes X_n)\ dP(X_1 \otimes \cdots \otimes X_n) \\
&= \left(\frac{2}{a_2}\right)^n \sum_\alpha c_\alpha \prod_{i=1}^n \int_{\mathcal{X}_1} \mathrm{Tr}[X_i O_i] \,\mathrm{Tr}[X_i \rho_i^{(\alpha)}]\ d\mu(X_i) \\
&= \sum_\alpha c_\alpha \prod_{i=1}^n \mathrm{Tr}[O_i \rho_i^{(\alpha)}] \\
&= \mathrm{Tr}[\rho O]
\end{aligned} \tag{C11}$$

The variance of our estimator is

$$\begin{aligned}
\mathrm{Var}_{X_1 \otimes \cdots \otimes X_n \sim dP}[o(X_1 \otimes \cdots \otimes X_n)] &\le \int_{\mathcal{X}_1^{\otimes n}} o(X_1 \otimes \cdots \otimes X_n)^2\ dP(X_1 \otimes \cdots \otimes X_n) \\
&= \sum_\alpha c_\alpha \prod_{i=1}^n \frac{4}{a_2^2} \int_{\mathcal{X}_1} \mathrm{Tr}[X_i O_i]^2 \,\mathrm{Tr}[X_i \rho_i^{(\alpha)}]\ d\mu(X_i) \\
&= \sum_\alpha c_\alpha \prod_{i=1}^n \frac{2a_3}{a_2^2} \,\mathrm{Tr}[O_i^2] \\
&= \left(\frac{2a_3}{a_2^2}\right)^n \prod_{i=1}^n \mathrm{Tr}[O_i^2] \\
&= \left(\frac{2a_3}{a_2^2}\right)^n \mathrm{Tr}[O^2]
\end{aligned} \tag{C12}$$

Again we see that the relevant factor is $\frac{2a_3}{a_2^2}$. For fixed $d = 2$, using Eq. (C8), this value is $2a_3/a_2^2 = 9/4 < 3$.

If the design was such that $2a_3 \le a_2^2$, this design-based shadow tomography protocol is expected attain an asymptotic variance of 0 for *any* observable. This suggests that in general a parameter tradeoff $2a_3 \ge a_2^2$ must hold and the optimal ensemble is the one that minimizes this inequality. Observe that here we have used the *local shadow tomography* protocol to derive a statement about the global shadow as well as general parameter tradeoffs between the factors a_t of any 3-design.

Appendix D: Continuous variable shadow tomography using GKP states

1. Bounding the estimator variance

When estimating the expectation value of a set of observables $O_i,\ i = 1, 2, \cdots, M$, using estimators given by

$$\tilde{o}_i = \langle X|\rho|X\rangle \langle X|O_i|X\rangle - \mathrm{Tr}\left[O_i\right] \tag{D1}$$

as in Eq. (C2), we need to bound its variance as it dictates the necessary sample overhead to obtain a good estimate. The variance $\mathrm{Var}\left[\tilde{o}\right] = \left\langle \tilde{o}^2 \right\rangle - \left\langle \tilde{o} \right\rangle^2 \le \left\langle \tilde{o}^2 \right\rangle$ is naturally upper bounded by the second moment, so we focus on bounding that.

To bound the estimator second moments without additional assumptions, we employ existing bound for second moments of functions over symplectic lattices derived in Ref. [40]. Let $f \in L^p\left(\mathbb{R}^{2n}\right)$, the p-norm is given by

$$\|f\|_p = \left(\int_{\mathbb{R}^{2n}} d\boldsymbol{x}\, f(x)^p\right)^{\frac{1}{p}}. \tag{D2}$$

We also make use of Riemann zeta function

$$\zeta(z) = \sum_{k\in\mathbb{N}} k^{-z}. \tag{D3}$$

The first statement is a second moment bound for functions of symplectic lattices, which we adapt from Refs. [55, 83].

Lemma 4 (symplectic second moments [55, 83])**.** *Let $f \in L^1\left(\mathbb{R}^{2n}\right) \cap L^2\left(\mathbb{R}^{2n}\right)$ be a non-negative function and $\Lambda \in Y_n$, with*

$$F_f(\Lambda) = \sum_{\boldsymbol{x}\in\Lambda-\{0\}} f(\boldsymbol{x}). \tag{D4}$$

For $n = 1$, we have that

$$\int_{Y_1} d\mu(\Lambda)\, |F_f(\Lambda)|^2 = \|f\|_1^2 + \|f\|_2^2, \tag{D5}$$

and for $n > 2$ it holds that

$$\int_{Y_n} d\mu(\Lambda)\, |F_f(\Lambda)|^2 \leq \|f\|_1^2 + \frac{4\zeta(n)^2}{\zeta(2n)}\|f\|_2^2. \tag{D6}$$

Note that the restriction to non-negativity of f is purely cosmetic can be easily removed. In the following application of this lemma f will be non-negative by default.

Proof. In the case $n = 1$ this statement follows from the fact that $\mathrm{Sp}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})$, such that in particular the spaces $X_2 = \mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ and Y_1 are identical. The statement is then provided by Ref. [83, Proposition 2.2]. For $n > 1$ this bound was proven in Ref. [55, Lemma 2] using the results of Ref. [40]. Note that the typically cited requirement of compact support for f can generally be relaxed to any f that decays sufficiently fast such that the series converge and integrals and sums may be exchanged or alternatively, that $f \in L^2\left(\mathbb{R}^{2n}\right)$ is bounded and non-negative (see Ref. [40, p.3]) □

Note that while the second statement in this lemma also principally also holds true for $n = 1$, as $\lim_{n\to 1}\zeta(n) = \infty$, the bound becomes trivial. As the next step we show that

Lemma 5 (Second moment bound)**.** *Let $|X\rangle \in \mathcal{X}_n$ for an n-mode CV quantum system as defined in Eq. (2) and let $O = O^\dagger$ be a trace class observable with characteristic function c_O. Relative to the uniform measure over $\mathcal{X}_n$, it holds that*

$$\int_{\mathcal{X}_n} d\mu(X)\, \langle X|\rho|X\rangle\, \langle X|O|X\rangle^2 \leq |\mathrm{Tr}[O]|^2 + 2|\mathrm{Tr}[O]|\|c_O\|_1 + \|c_O\|_1^2 + f_n\, \mathrm{Tr}\left[O^2\right] =: \widehat{V}_O^n, \tag{D7}$$

where $f_1 = 1$ and for $n > 1$,

$$f_n = \frac{4\zeta(n)^2}{\zeta(2n)}. \tag{D8}$$

Proof. Denote $|X\rangle = |\Lambda, \boldsymbol{\alpha}\rangle = D(\boldsymbol{\alpha})|\Lambda\rangle$. We have

$$\begin{aligned} M &= \int_{Y_n} d\mu(\Lambda) \int_{P(\Lambda)} d\boldsymbol{\alpha}\, \langle\Lambda, \boldsymbol{\alpha}|\rho|\Lambda, \boldsymbol{\alpha}\rangle\, \langle\Lambda, \boldsymbol{\alpha}|O|\Lambda, \boldsymbol{\alpha}\rangle^2 \\ &= \int_{Y_n} d\mu(\Lambda) \int_{P(\Lambda)} \sum_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_3\in\Lambda} e^{i\Phi(\boldsymbol{\lambda}_1)+i\Phi(\boldsymbol{\lambda}_2)+i\Phi(\boldsymbol{\lambda}_3)+i2\pi\boldsymbol{\alpha}^T J(\boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2+\boldsymbol{\lambda}_3)} c_\rho(\boldsymbol{\lambda}_1)\, c_O(\boldsymbol{\lambda}_2)\, c_O(\boldsymbol{\lambda}_3) \\ &= \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_3\in\Lambda} e^{i\Phi(\boldsymbol{\lambda}_2+\boldsymbol{\lambda}_3)+i\Phi(\boldsymbol{\lambda}_2)+i\Phi(\boldsymbol{\lambda}_3)} c_\rho^*(\boldsymbol{\lambda}_2+\boldsymbol{\lambda}_3)\, c_O(\boldsymbol{\lambda_2})\, c_O(\boldsymbol{\lambda_3}). \end{aligned} \tag{D9}$$

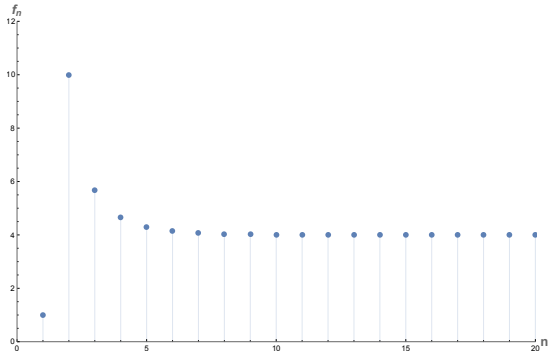


FIG. 2. The sequence $\{f_n\}_{n=1}^{\infty}$ that parametrizes the variance upper bound for the GKP shadow tomography estimators. Note that the sequence converges to $f_n \to 4$ for $n \to \infty$.

Since $|c_\rho(\boldsymbol{\gamma})| = |\mathrm{Tr}\left[D^\dagger(\boldsymbol{\gamma})\,\rho\right]| \leq 1$ by the unitarity of $D^\dagger(\boldsymbol{\gamma})$, we can bound

$$\begin{aligned}
|M| &\leq \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda_2},\boldsymbol{\lambda_3}\in\Lambda} |c_O(\boldsymbol{\lambda_2})||c_O(\boldsymbol{\lambda_3})| \\
&= \int_{Y_n} d\mu(\Lambda) \left\{|c_O(0)| + F_{|c_O|}(\Lambda)\right\}^2 \\
&= \int_{Y_n} d\mu(\Lambda)\, |c_O(0)|^2 + 2|c_O(0)|F_{|c_O|}(\Lambda) + F_{|c_O|}(\Lambda)^2
\end{aligned} \tag{D10}$$

Using the mean value formula, lemma 1, together with lemma 4, we obtain for $n = 1$

$$|M| \leq |c_O(0)|^2 + 2|c_O(0)|\|c_O\|_1 + \|c_O\|_1^2 + \|c_O\|_2^2, \tag{D11}$$

and for $n > 1$,

$$|M| \leq |c_O(0)|^2 + 2|c_O(0)|\|c_O\|_1 + \|c_O\|_1^2 + \frac{4\zeta(n)^2}{\zeta(2n)}\|c_O\|_2^2. \tag{D12}$$

Finally, we use that $c_O(0) = \mathrm{Tr}[O]$ and $\|c_O\|_2^2 = \mathrm{Tr}\left[O^2\right]$. □

We illustrate the sequence f_n in Fig. 2. We observe that the sequence is bounded from above by $f_n \leq f_2 = 10$ and, as n grows, we have

$$\lim_{n\to\infty} f_n = 4. \tag{D13}$$

The boundedness of the factor can also be verified by the fact that for $n \geq 2$, $\zeta(n) \leq \zeta(2n)$, such that $\zeta(n)^2/\zeta(2n) \leq \zeta(n) \leq \zeta(2)$. The actual bound observed indeed is a little bit stronger, and stems from that observation that for $n \geq 2$, f_n is monotonically decreasing such that the maximum is attained at f_2.

2. Physical state constraints

The previous statements have been derived entirely without external assumptions on the input state or strong constraints on the locality structure of the ensemble $\mathcal{X}$. In the following sections, we investigate how the incorporation of such constraints manifests in the sample complexity of the shadow tomography protocol. We define the generalized effective squeezing parameter via [28]

$$\Delta_{\boldsymbol{\lambda}}^2(\rho) = -\frac{2}{\pi}\log\left(|c_\rho(\boldsymbol{\lambda})|\right). \tag{D14}$$

This parameter captures *how translationally* invariant a state is relative to the lattice vector $\boldsymbol{\lambda} \in \Lambda$. Equivalently, it parametrizes how strongly the state is squeezed along the phase space hyperplane $\boldsymbol{\lambda}^{\perp} \subset \mathbb{R}^{2n}$. For $n = 1$, it holds that approximatively $\Delta^2 \geq \frac{4}{\mathrm{Tr}[\rho \hat{n}]}$ [28].

We make the following assumption on physical states.

Assumption 1 (structure of physical state)**.** *Let $M = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{2n})^T$ be a basis for the lattice Λ. Then it holds that for any physical state ρ and $\boldsymbol{n} \in \mathbb{Z}^{2n}$*

$$|c_\rho \left(\sum_{i=1}^{2n} n_i \boldsymbol{\xi}_i \right)| \leq \prod_{i=1}^{2n} |c_\rho \left(\boldsymbol{\xi}_i \right)|^{|n_i|}. \tag{D15}$$

Note that this assumption is well motivated, as the characteristic function of any physical state is Schwartz, i.e. it decays super-polynomially in phase space. As a random lattice will typically have a shortest vector length $\lambda_1 (\Lambda) \sim \sqrt{n}$, we expect all vectors $\boldsymbol{\xi}_i$ in the selected basis to follow at least this length scale.

Using this assumption, we have

Proposition 2 (physicality)**.** *For any physical state ρ, and any $|X\rangle \in \mathcal{X}_n$ it holds that*

$$\langle X|\rho|X\rangle \leq \left(2 \left(1 - e^{-\frac{\pi}{2} \tilde{\Delta}^2 (\rho)} \right)^{-1} - 1 \right)^{2n}, \tag{D16}$$

where

$$\tilde{\Delta} (\rho) \coloneqq \min_{M \in \mathrm{Sp}_{2n}(\mathbb{R})} \min_i \Delta_{\boldsymbol{\xi}_i} (\rho) \tag{D17}$$

with $M = (\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_{2n})^T$.

Proof. Let $M = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{2n})^T$ be a minimal basis for Λ. It holds that, using assumption 1,

$$\begin{aligned}
|\langle X|\rho|X\rangle| &\leq \sum_{\boldsymbol{\lambda} \in \Lambda} |c_\rho (\boldsymbol{\lambda})| = \sum_{\boldsymbol{\lambda} \in \Lambda} e^{-\frac{\pi}{2} \Delta^2_{\boldsymbol{\lambda}} (\rho)} \\
&\leq \prod_{i=1}^{2n} \left(|c_\rho (0)| + 2 \sum_{n_i \in \mathbb{N}} |c_\rho (\boldsymbol{\xi}_i)|^{n_i} \right) \\
&= \prod_{i=1}^{2n} \left(1 + 2 \sum_{n_i \in \mathbb{N}} e^{-\frac{\pi}{2} n_i \Delta^2_{\boldsymbol{\xi}_i} (\rho)} \right) \\
&= \prod_{i=1}^{2n} \left(2 \left(1 - e^{-\frac{\pi}{2} \Delta^2_{\boldsymbol{\xi}_i} (\rho)} \right)^{-1} - 1 \right).
\end{aligned} \tag{D18}$$

Let $\tilde{\Delta}_M (\rho) = \min_i \Delta_{\boldsymbol{\xi}_i} (\rho)$, where $M = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{2n})^T$.

We can thus bound

$$|\langle X|\rho|X\rangle| \leq \left(2 \left(1 - e^{-\frac{\pi}{2} \tilde{\Delta}^2_M (\rho)} \right)^{-1} - 1 \right)^{2n}. \tag{D19}$$

Finally, minimizing over the possible lattice choices and all possible choices of bases yields the result. □

Following the arguments in Refs. [28, 45], we further estimate that a general lower bound for the effective squeezing is set by the average photon number of the state, $\tilde{\Delta} (\rho) \geq \frac{4}{\mathrm{Tr}[\rho \hat{N}]}$. For any fixed lattice Λ, let $M = S^T = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{2n})^T$ denote a choice of shortest basis. It is reasonable to expect that any state that delivered a small value of effective squeezing $\tilde{\Delta}_M$ for the lattice basis M should also have a high degree of approximative symmetry under the corresponding group of GKP stabilizer operators. The state $|\Lambda\rangle$ may be obtained as $|\Lambda\rangle = U_S |\mathbb{Z}^{2n}\rangle = U_S |\mathbb{Z}^2\rangle^{\otimes n}$ for U_S a Gaussian unitary transform [38]. To obtain the approximate version, we replace each individual state $|\mathbb{Z}^2\rangle$ with an approximate state $|\widetilde{\mathbb{Z}^2}\rangle \approx e^{-\Delta^2 \hat{n}} |\mathbb{Z}^2\rangle / \sqrt{\langle \mathbb{Z}^2 | e^{-2\Delta^2 \hat{n}} | \mathbb{Z}^2 \rangle}$

with effective squeezing parameter $\Delta^2 \geq 4/\overline{n}$, where $\overline{n} = \left\langle \widetilde{\mathbb{Z}^2} | \hat{n} | \widetilde{\mathbb{Z}^2} \right\rangle$ is the average photon number of the individual state [28, 57]. Let $I = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{2n})^T$ be a basis that stabilizes the state $|\mathbb{Z}^2\rangle$ and $S^T = IS^T$. We have, by the properties of the trace,

$$\mathrm{Tr}\left[D\left(S\boldsymbol{e}_i\right) U_S \left|\widetilde{\mathbb{Z}^{2n}}\right\rangle\!\left\langle\widetilde{\mathbb{Z}^{2n}}\right| U_S^\dagger\right] = \mathrm{Tr}\left[D\left(\boldsymbol{e}_i\right) \left|\widetilde{\mathbb{Z}^{2n}}\right\rangle\!\left\langle\widetilde{\mathbb{Z}^{2n}}\right|\right]. \tag{D20}$$

In particular, the finite squeezing parameters satisfy

$$\Delta^2_{S^T}\left(U_S \left|\widetilde{\mathbb{Z}^{2n}}\right\rangle\!\left\langle\widetilde{\mathbb{Z}^{2n}}\right| U_S^\dagger\right) = \Delta^2_I\left(\left|\widetilde{\mathbb{Z}^{2n}}\right\rangle\!\left\langle\widetilde{\mathbb{Z}^{2n}}\right|\right) \tag{D21}$$

$$\geq 4/\overline{N}, \tag{D22}$$

where the last inequality stems from the fact that, optimizing over all possible states, a minimal value for $\Delta^2_I = \Delta^2$ could be obtained by considering states of the form $|0\rangle^{\otimes(n-1)} \otimes |\widetilde{\mathbb{Z}^2}\rangle$, where $|\widetilde{\mathbb{Z}^2}\rangle$ attaims the minimal Δ^2 the total average photon number allows for. We hence have motivated the following conjecture.

Conjecture 1 (Lower bound for the effective squeezing parameter)**.** *Given a physical state ρ on an n-mode quantum system with average photon number $\overline{N} = \mathrm{Tr}\left[\rho\hat{N}\right]$, it holds with $M = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{2n})^T$, that*

$$\tilde{\Delta}^2(\rho) \coloneqq \min_{M \in \mathrm{Sp}_{2n}(\mathbb{R})} \min_i \Delta^2_{\boldsymbol{\xi}_i}(\rho) \geq \frac{4}{\overline{N}}. \tag{D23}$$

Using conjecture 1, for any state with average photon number at most $\overline{N}$ the factor under the exponent can be bounded by

$$\begin{aligned}\gamma &= 2\left(1 - e^{-\frac{\pi}{2}\tilde{\Delta}^2}\right)^{-1} - 1 \\ &\leq 2\left(1 - e^{-2\pi/\overline{N}}\right)^{-1} - 1 \\ &\approx \overline{N}/\pi\end{aligned} \tag{D24}$$

in the limit $\overline{N} \gg 1$.

We note that, following a strategy proposed in Ref. [28], a possible proof to this the conjecture may be obtained by viewing ρ as a "sensor state" that is employed to estimate a displacement $D(\boldsymbol{x})$ by executing measurements on copies of $\rho_{\boldsymbol{x}} = D(\boldsymbol{x})\rho D^\dagger(\boldsymbol{x})$. We expect that a Cramer-Rao bound for this estimation problem would imply the result.

3. Variance bounds with physicality assumptions

We have also attempted to derive a similar bound for the local ensemble $\mathcal{X}_1^{\otimes n}$, but have not been able to distill a sufficiently nice and interpretable expression. Instead, to derive a corresponding bound for the ensemble $\mathcal{X}_1^{\otimes n}$, we invoke the physicality bound from assumption 1 and proposition 2.

Lemma 6 (Second moment of global ensemble, physical)**.** *Let $\mathcal{X}_n$ be the ensemble of GKP states and let ρ be any physical state, for which assumption 1 holds with average photon number at most $\overline{N}$. We have that for any hermitian trace class observable O,*

$$\int_{\mathcal{X}_1^{\otimes n}} d\mu(X)\ \langle X|\rho|X\rangle \langle X|O|X\rangle^2 \leq \left(\overline{N}/\pi\right)^{2n}\left(\mathrm{Tr}\left[O\right]^2 + \mathrm{Tr}\left[O^2\right]\right) =: \left(\overline{N}/\pi\right)^{2n}\tilde{V}_O. \tag{D25}$$

Proof. We have that for any hermitian trace class observable O

$$\begin{aligned}\int_{\mathcal{X}_1^{\otimes n}} d\mu(X)\ \langle X|\rho|X\rangle \langle X|O|X\rangle^2 &\leq \left(\overline{N}/\pi\right)^{2n}\int_{\mathcal{X}_n} d\mu(X)\ \mathrm{Tr}\left[\left(O\otimes O\right)|X\rangle\!\langle X|^{\otimes 2}\right] \\ &= \left(\overline{N}/\pi\right)^{2n}\left(\mathrm{Tr}\left[O\right]^2 + \mathrm{Tr}\left[O^2\right]\right) \\ &=: \left(\overline{N}/\pi\right)^{2n}\tilde{V}_O^n,\end{aligned} \tag{D26}$$

where we have used bound in prop. 2 the 2-design property, Thm. 1 and the bound in Eq. D24. □

Lemma 7 (Second moment of local ensemble, physical)**.** *Let $\mathcal{X}_1^{\otimes n}$ be the n-fold ensemble of single-mode GKP states and let ρ be any physical state, for which assumption 1 holds with average photon number at most $\overline{N}$. We have that for any hermitian trace class observable O,*

$$\int_{\mathcal{X}_1^{\otimes n}} d\mu\left(X\right)\,\langle X|\rho|X\rangle\,\langle X|O|X\rangle^2 \leq \left(\overline{N}/\pi\right)^{2n}\sum_{\boldsymbol{k}\subseteq\boldsymbol{n}}\mathrm{Tr}_{\boldsymbol{n}-\boldsymbol{k}}\left[\mathrm{Tr}_{\boldsymbol{k}}\left[O\right]^2\right] =: \left(\overline{N}/\pi\right)^{2n}\tilde{V}_O^{\mathrm{loc}}, \tag{D27}$$

where $\sum_{\boldsymbol{k}\subseteq\boldsymbol{n}}$ denotes the sum over all elements $\boldsymbol{k}$ of the power set of $\boldsymbol{n}=\{1,\ldots,n\}$ and the outer trace ranges over the complement of $\boldsymbol{k}$.

Proof. Let $\Lambda_\oplus^n=\Lambda_1\oplus\ldots\oplus\Lambda_n$ denote a direct sum composite lattice and let $\boldsymbol{\alpha}_\oplus^n=\boldsymbol{\alpha}_1\oplus\ldots\oplus\boldsymbol{\alpha}_n$ denote a composite vector, which is such that $\boldsymbol{\alpha}_\oplus^n\in\mathcal{P}\left(\Lambda_\oplus^n\right)$. Note that $\Lambda_\oplus^n$ is symplectic if its components Λ_i are. Each local shadow sample is represented by a composite state

$$\bigotimes_{i=1}^n|X_i\rangle=|\Lambda_\oplus^n;\boldsymbol{\alpha}_\oplus^n\rangle \tag{D28}$$

such that we have using prop. 2

$$\begin{aligned}
M&=\int_{\mathcal{X}_1^{\otimes n}} d\mu\left(X\right)\,\langle X|\rho|X\rangle\,\langle X|O|X\rangle^2\\
&\leq\left(\overline{N}/\pi\right)^{2n}\int_{\mathcal{X}_1^{\otimes n}}\langle X|O|X\rangle^2\\
&=\left(\overline{N}/\pi\right)^{2n}\int_{Y_1^{\oplus n}} d\mu\left(\Lambda_\oplus^n\right)\sum_{\boldsymbol{\lambda}_{\oplus,1}^n,\boldsymbol{\lambda}_{\oplus,2}^n\in\Lambda_\oplus^n} e^{i\Phi\left(\boldsymbol{\lambda}_{\oplus,1}^n+\boldsymbol{\lambda}_{\oplus,2}^n\right)}\\
&\qquad\int_{\mathcal{P}\left(\Lambda_\oplus^n\right)} d\boldsymbol{\alpha}_\oplus^n\, e^{-i2\pi{\boldsymbol{\alpha}_\oplus^n}^T J\left(\boldsymbol{\lambda}_{\oplus,1}^n+\boldsymbol{\lambda}_{\oplus,2}^n\right)}c_O\left(\boldsymbol{\lambda}_{\oplus,1}^n\right)c_O\left(\boldsymbol{\lambda}_{\oplus,2}^n\right)\\
&=\left(\overline{N}/\pi\right)^{2n}\int_{Y_1^{\oplus n}} d\mu\left(\Lambda_\oplus^n\right)\sum_{\boldsymbol{\lambda}_\oplus^n\in\Lambda_\oplus^n}|c_O\left(\boldsymbol{\lambda}_\oplus^n\right)|^2.
\end{aligned} \tag{D29}$$

We split the lattice into partitions $\boldsymbol{k}\subseteq\boldsymbol{n}=\{1,\ldots,n\}$, where the multi-index $\boldsymbol{k}$ indicates the indices where the lattice vector has non-zero components,

$$\Lambda_\oplus^n=\bigcup_{\boldsymbol{k}\subseteq\boldsymbol{n}}\left(\Lambda-\{0\}\right)^{\boldsymbol{k}}\cup\{0\}^{\boldsymbol{n}-\boldsymbol{k}}. \tag{D30}$$

Using the fact that

$$c_O\left(\boldsymbol{\lambda}^{\boldsymbol{k}}\oplus\{0\}^{\boldsymbol{n}-\boldsymbol{k}}\right)=c_{\mathrm{Tr}_{\boldsymbol{n}-\boldsymbol{k}}[O]}\left(\boldsymbol{\lambda}\right), \tag{D31}$$

we rewrite the sum above as

$$\begin{aligned}
M&\leq\left(\overline{N}/\pi\right)^{2n}\sum_{\boldsymbol{k}\subseteq\boldsymbol{n}}\int_{Y_1^{\oplus n}} d\mu\left(\Lambda_\oplus^n\right)\sum_{\boldsymbol{\lambda}^{\boldsymbol{k}}\in\left(\Lambda-\{0\}\right)^{\boldsymbol{k}}}|c_{\mathrm{Tr}_{\boldsymbol{n}-\boldsymbol{k}}[O]}\left(\boldsymbol{\lambda}^{\boldsymbol{k}}\right)|^2\\
&=\left(\overline{N}/\pi\right)^{2n}\sum_{\boldsymbol{k}\subseteq\boldsymbol{n}}\mathrm{Tr}_{\boldsymbol{k}}\left[\mathrm{Tr}_{\boldsymbol{n}-\boldsymbol{k}}\left[O\right]^2\right].
\end{aligned} \tag{D32}$$

Where, in the last step, we have performed the integral over the independent components in $\boldsymbol{k}$ and used that for $O=O^\dagger$,

$$\int_{\mathbb{R}^{2k}} d\boldsymbol{x}\,|c_O\left(\boldsymbol{x}\right)|^2=\mathrm{Tr}\left[O^2\right]. \tag{D33}$$

□

4. Continuous variable medians of means estimation

We employ median of means (MoM) estimation to set up an algorithm for the estimation of the expectation values of M observables $\{O_i\}$ relative to the statistics obtained from measuring an arbitrary (physical) input state relative to a GKP state, such that the measurement probability is given by $\langle X|\rho|X\rangle, |X\rangle \in \mathcal{X}$ as outlined before.

Medians of means (MoM) estimation provides a beneficial strategy to implement this task, which has the upshot that the necessary sample complexity admits scaling linearly with the variance of the estimator and *logarithmically* in the number of observables one is interested in estimating. Given $N = KB$ samples of a random variable $\tilde{o}$, MoM estimation proceeds by dividing this set into K batches each of size B and outputs the median value of the arithmetic mean of $\tilde{o}$ on each batch. This strategy biases the expected value away from low-probability events at the tail of the distribution and allows to prove exponential concentration of the resulting estimate around the true mean (see Ref. [84] for a pedagogical introduction).

MoM proceeds as follows. Divide the $N = KB$ samples of $\tilde{o}_i$ into K batches $\mathcal{B}_i^{(j)}$, $j = 1, 2, \cdots, K$ each of size B, and

$$\mathrm{MoM}\left[\tilde{o}_i\right] = \mathrm{median}\left\{\overline{o}_i^{(1)}, \ldots, \overline{o}_i^{(K)}\right\}, \quad \overline{o}_i^{(j)} = \frac{1}{B}\sum_{\tilde{o}\in\mathcal{B}_i^{(j)}} \tilde{o} \tag{D34}$$

is the medians of means estimator. See Refs. [8, 84, 85] for good explanations of this estimator.

Theorem 3 (GKP shadow tomography)**.** *Let $\epsilon, \delta > 0$, and let $O_i = O_i^\dagger, i = 1, 2, \cdots, M$ be a set of M observables on an n-mode CV quantum system with integrable characteristic functions c_{O_i}. Let*

$$\begin{aligned} \tilde{V}_{O_i} &= \left(\|c_{O_i}\|_1 + |\mathrm{Tr}\left[O_i\right]|\right)^2 + f_n \,\mathrm{Tr}\left[O_i^2\right] \\ \tilde{V}_O &= \max_i \tilde{V}_{O_i}, \end{aligned} \tag{D35}$$

and $N = KB$, with

$$B = 34\tilde{V}_O/\epsilon^2 \quad \textit{and} \quad K = 2\log\left(2M/\delta\right). \tag{D36}$$

Then, N samples from the state relative to the POVMs defined by the ensemble of GKP states $\mathcal{X}_n$ suffice to approximate the expectation values of each observable O_i with probability

$$\Pr\left(\max_{i=1,\ldots,M} |\mathrm{MoM}\left[\tilde{o}_i\right] - \mathrm{Tr}\left[O_i\rho\right]| \geq \epsilon\right) \leq \delta. \tag{D37}$$

Proof. The proof follows the standard proof using the variance bound $\tilde{V}_O$ from lem. 5 for each observable, as discussed in Ref. [8]. □

On the technical level, the key ingredient is that for medians of means, the error probability of the estimator is provided by counting the number of means that deviated from the median. This counting function can be expressed with a simple bounded random variable, such that Hoeffding's inequality establishes its anticoncentration. Naturally, if the estimator is bounded by default, one may directly apply Hoeffding's inequality to obtain a tight anticoncentration bound using a simple arithmetic mean estimator. This was the strategy employed in Ref. [14] and is particularly useful when $\langle X|O|X\rangle$ can be bounded. A special case where this is possible is outlined in appendix G such that one can employ Popoviciu's inequality on the variance of a bounded estimator $|\sigma| \leq M$ to simply replace $\tilde{V}_O = M^2$ in the statement above.

Note that, using the physicality assumption we may replace the variance estimates in the theorem to obtain the following.

Theorem 4 (CV shadow tomography, physical)**.** *Let $\epsilon, \delta > 0$, and let $O_i = O_i^\dagger$, $i = 1, 2, \cdots, M$ be a set of M observables on a n-mode CV quantum system with integrable characteristic functions c_{O_i}. Let ρ be any physical state, for which assumption 1 holds with average photon number at most $\overline{N}$ and let*

$$\tilde{V}_{O_i} = \left(\overline{N}/\pi\right)^{2n}\left(\mathrm{Tr}\left[O_i\right]^2 + \mathrm{Tr}\left[O_i^2\right]\right) \tag{D38}$$

$$\tilde{V}_O = \max_i \tilde{V}_{O_i}, \tag{D39}$$

and $N = cBK$, with

$$B = 34\tilde{V}_O/\epsilon^2 \quad \textit{and} \quad K = 2\log\left(2M/\delta\right), \tag{D40}$$

and c a numerical constant. Then, N samples from the state relative to the POVMs defined by the ensemble of GKP states $\mathcal{X}_n \ni |X\rangle = D(\boldsymbol{\alpha})|\Lambda\rangle$, where $|\Lambda\rangle$ is the unique GKP state corresponding to $\Lambda = \Lambda^\perp \subset \mathbb{R}^{2n}$, and where $\boldsymbol{\alpha} \in \mathcal{P}(\Lambda)$, suffice to approximate the expectation values of each observable O_i with high probability,

$$\Pr\left(\max_{i=1,\dots,M} |\mathrm{MoM}[\tilde{o}_i] - \mathrm{Tr}[O_i\rho]| \geq \epsilon\right) \leq \delta. \tag{D41}$$

Proof. The proof follows the standard proof using the variance bound $\tilde{V}_O$ from lem. 6 for each observable, as discussed in Ref. [8] together with the estimate from Eq. (D24). □

To fully bound the performance of the local GKP shadow protocol, we employ the result derived in lemma 7.

Note that the partition-based structure of the bound is consistent with the fact that, written in vectorized notation, $\mathcal{M}_{\mathcal{X}_1}^{\otimes n} = (I\otimes I^* + |I\rangle\!\rangle\langle\!\langle I|)^{\otimes n}$ behaves like a *random* deletion channel. For $n = 2$, the expression in lemma 7 yields (up to the factor f^{2n})

$$\sum_{\boldsymbol{k}\subseteq\boldsymbol{n}} \mathrm{Tr}_{\boldsymbol{n}-\boldsymbol{k}}\left[\mathrm{Tr}_{\boldsymbol{k}}[O]^2\right] = \mathrm{Tr}\left[O^2\right] + \mathrm{Tr}[O]^2 + \mathrm{Tr}_2\left[\mathrm{Tr}_1[O]^2\right] + \mathrm{Tr}_1\left[\mathrm{Tr}_2[O]^2\right]. \tag{D42}$$

In comparison to the global shadow protocol, we see that we obtain additional terms from all the non-trivial bi-partitions of the n-mode system. It is notable that, in contrast to e.g. the qubit-based shadow protocol in Ref. [8], the performance of the local shadow tomography protocol derived here is not limited by the *locality* of the observable but rather, in the opposite, by its *in-separability*. As long as the total purity of the observable across any bi-partition remains small, the local shadow protocol will be efficient.

Equipped with the above derivation, we arrive at following theorem.

Theorem 5 (CV shadow tomography, local and physical)**.** *Let $\epsilon, \delta > 0$, and let $O_i = O_i^\dagger, i = 1, 2, \cdots, M$ be a set of M observables on a n-mode CV quantum system with integrable characteristic functions c_{O_i}. Assume conjecture 1 and that the input state has average photon number at most $\overline{N} \gg 1$ with exponentially decaying characteristic function according to assumption 1. Let*

$$\tilde{V}_{O_i}^{\mathrm{loc}} = \left(\overline{N}/\pi\right)^{2n} \sum_{\boldsymbol{k}\subseteq\boldsymbol{n}} \mathrm{Tr}_{\boldsymbol{n}-\boldsymbol{k}}\left[\mathrm{Tr}_{\boldsymbol{k}}[O_i]^2\right] \tag{D43}$$

$$\tilde{V}_{O}^{\mathrm{loc}} = \max_i \tilde{V}_{O_i}^{\mathrm{loc}}, \tag{D44}$$

and $N = BK$, with

$$B = 34\tilde{V}_O^{\mathrm{loc}}/\epsilon^2 \quad \textit{and} \quad K = 2\log(2M/\delta). \tag{D45}$$

Then, N samples from the states relative to the POVMs defined by the ensemble of GKP states $\mathcal{X}_1^n = \bigotimes_{i=k}^n D(\boldsymbol{\alpha}_k)|\Lambda_k\rangle$, where $|\Lambda_k\rangle$ is the unique GKP state corresponding to $\Lambda_k = \Lambda_k^\perp \subset \mathbb{R}^2$ and where $\boldsymbol{\alpha}_k \in \mathcal{P}(\Lambda_k)$, suffice to approximate the expectation values of each observable O_i with

$$\Pr\left(\max_{i=1,\dots,M} |\mathrm{MoM}[\tilde{o}_i] - \mathrm{Tr}[O_i\rho]| \geq \epsilon\right) \leq \delta. \tag{D46}$$

Proof. The proof follows the standard proof using the variance bound $\tilde{V}_O^{\mathrm{loc}}$ from lem. 7 for each observable, as discussed in Ref. [8] as well as the estimate in Eq. (D24). □

5. Example: Thermal state observables

We investigate the derived sample complexities for a simple observable, the n-mode (unnormalized) thermal state $O = e^{-\beta\hat{N}}$. To this end, we compute the variances $\widehat{V}_O^n, \tilde{V}_O, \tilde{V}_O^{\mathrm{loc}}$ derived in lemmas 5, 6 and 7, respectively.

We begin with the variance per lemma 5. We have

$$\mathrm{Tr}\left[e^{-\beta\hat{N}}\right] = \left(1-e^{-\beta}\right)^{-n}, \quad \mathrm{Tr}\left[e^{-2\beta\hat{N}}\right] = \left(1-e^{-2\beta}\right)^{-n}. \tag{D47}$$

Using [56, 57]

$$O = e^{-\beta \hat{N}} = \left(1 - e^{-\beta}\right)^{-n} \int_{\mathbb{R}^{2n}} d\boldsymbol{x}\, e^{-\frac{\|\boldsymbol{x}\|^2}{\Delta^2}} D\left(\boldsymbol{x}\right), \tag{D48}$$

where $\Delta^2 = 2\tanh\left(\beta/2\right)$, we can also obtain the norm

$$\begin{aligned}
\|c_O\|_1 &= \left(1 - e^{-\beta}\right)^n \int_{\mathbb{R}^{2n}} d\boldsymbol{x}\, e^{-\frac{\pi\|\boldsymbol{x}\|^2}{\Delta^2}} \\
&= \left(1 - e^{-\beta}\right)^n \left|\frac{2\pi}{(2\pi/\Delta^2)}\right|^n \\
&= \left(1 - e^{-\beta}\right)^n \left(2\frac{1 + e^{-\beta}}{1 - e^{-\beta}}\right)^n \\
&= 2^n \left(1 + e^{-\beta}\right)^n \\
&= (4 - 2\beta)^n + O\left(\beta^2\right)
\end{aligned} \tag{D49}$$

Thus, the bound in lemma 5 becomes

$$\begin{aligned}
\widehat{V}_O^n &= |\mathrm{Tr}\left[O\right]|^2 + 2|\mathrm{Tr}\left[O\right]|\|c_O\|_1 + \|c_O\|_1^2 + f_n\,\mathrm{Tr}\left[O^2\right] \\
&= 2^{n+1} + 2^n\left(1 + e^{-2\beta}\right)^n + 2^{2n}\left(1 + e^{-\beta}\right)^{2n} + f_n\left(1 - e^{-2\beta}\right)^{-n} \\
&= O\left(\left(2\left(1 + e^{-\beta}\right)\right)^{2n}\right).
\end{aligned} \tag{D50}$$

For the remaining variances, a quick computation shows that

$$\begin{aligned}
\tilde{V}_O &= \mathrm{Tr}\left[O^2\right] + \mathrm{Tr}\left[O\right]^2 \\
&= \left(1 - e^{-2\beta}\right)^{-n} + \left(1 - e^{-\beta}\right)^{-2n} \\
&= \left(1 + e^{-\beta}\right)^{-n}\left(1 - e^{-\beta}\right)^{-n} + \left(1 - e^{-\beta}\right)^{-2n}
\end{aligned} \tag{D51}$$

and

$$\begin{aligned}
\tilde{V}_O^{\mathrm{loc}} &= \sum_{\boldsymbol{k}\subseteq\boldsymbol{n}} \mathrm{Tr}_{\boldsymbol{n}-\boldsymbol{k}}\left[\mathrm{Tr}_{\boldsymbol{k}}\left[O\right]^2\right] \\
&= \sum_{k=0}^{n}\binom{n}{k}\left(1 - e^{-2\beta}\right)^{n-k}\left(1 - e^{-\beta}\right)^{-2k} \\
&= \left(1 - e^{-\beta}\right)^{-2n}\left(2e^{-3\beta} - 2e^{-\beta} - e^{-4\beta} + 2\right)^n.
\end{aligned} \tag{D52}$$

In the high-temperature regime ($\beta \to 0$), the local shadow and global shadow follow the same variance scaling with n. Restoring the prefactor from the energy assumption, we hence obtain

$$f^{2n}\tilde{V}_O, f^{2n}\tilde{V}_O^{\mathrm{loc}} \sim \left(\frac{\overline{N}}{\pi\beta}\right)^{2n}. \tag{D53}$$

In the reverse limit, $\beta \to \infty$, we also have $\tilde{V}_O \to 2$, $\tilde{V}_O^{\mathrm{loc}} \to 1$.

It is notable that, under the energy constraint, the global- and local shadow attain the same effective scaling. They exhibit different scalings as $n \to \infty$ depending on whether $\overline{N} > \pi\beta$ or $\overline{N} < \pi\beta$. I particular, if the average photon number of the input system can be assumed to be very low $\overline{N} < \pi\beta$ (remember that we are aready assuming a high temperature limit $\beta \ll 1$), the given variance together with the sample complexity of the shadow tomography protocol may scale benignly in the system size.

Appendix E: Sample complexity: comparison to related protocols

1. Qubit-based classical shadows

We would like to make a comparison of the thermal-observable calculation in Sec. D 5 to related qubit-based protocols. To this end consider, analogous to the above, a canonical choice of thermal state

$$O' = e^{-\beta\sum_{i=1}^n \hat{\sigma}_z}, \tag{E1}$$

where each $\hat{\sigma}_z$ is the Pauli-Z operator on qubit "i". We evaluate the scaling bound estimated in Ref. [8],

$$\mathrm{Tr}\left[O'^2\right] = \left[e^{2\beta} - e^{-2\beta}\right]^n = \left[2\cosh\left(2\beta\right)\right]^n = \left(2 + 4\beta^2 + O\left(\beta^4\right)\right)^n. \tag{E2}$$

In the limit of small $\beta \ll 1$, this bound scales as 2^n with n. Note that this bound differs from the bound derived from lemma 5, 4^n (for $\beta \to 0$) only by an exponential factor, although the Hilbert space employed in our setting is of infinite size, rather than 2^n.

2. Photon number cutoffs

Building on the intuition gained from the previous section, we further discuss the sample complexity of our GKP-design based shadow tomography scheme. We first investigate the sample complexity under the assumption that the input state is well represented by a state with a total photon number cutoff $N_{\max}$, which is a scenario also investigated in Refs. [14, 86].

Naturally, given a total photon number cutoff on the input state $N_{\max}$, this implies a trivial upper bound on the average photon number on the input $\overline{N} \leq N_{\max}$.

We can also use this derivation to compare to the multimode homodyne shadow protocol discussed in Ref. [86], where each mode is cutoff at maximum photon number $N_{1,\max}$, such that the total photon number admits a cutoff of $N_{\max} = nN_{1,\max}$. Thus, our variance bound and sample complexity scaling of our shadow tomography protocol scale as $O\left(\left(nN_{1,\max}\right)^{2n}\right)$ under a very pessimistic estimate. While this scaling appears more beneficial in the local photon number $N_{1,\max}$, the scaling with the mode number n appears worse. We expect this to be mainly an effect of the conservative estimate provided in conjecture 1.

Homodyne tomography can be bounded using the same type of shadow analysis using a combination of Corr. IV.1.2 and Eq. (15) from Ref. [86]. The former yields a shadow sample complexity bound of the same type as our main theorem, while the latter fills in a key variable in said bound. Keeping track of $N_{1,\max}$ and n yields a homodyne sample complexity of order $O(nC_1^{2n}N_{1,\max}^{20n/3}\log 2N_{1,\max})$ (with C_1 a constant), roughly a quartic deterioration over our protocol in $N_{1,\max}$. However, the shadow bounds of Ref. [86] are not tight, and it remains unclear how well they quantify the performance of homodyne tomography.

a. *Full state tomography*

Using these findings, we further investigate the application of our shadow tomography protocol to full state tomography. To this end, note that, in Ref. [58], the following lower bound on energy constrained CV state tomography has been proven.

Theorem 6 (Optimal CV tomography [58])**.** *Let ρ be an unknown quantum state on n bosonic modes satisfying an energy constraint $\mathrm{Tr}\left[\hat{N}\rho\right] \leq \overline{N}$ for some absolute constant $\overline{N}$. Then the number of copies of ρ required to perform quantum state tomography with precision ε in trace distance has to scale at least as $(\Theta(\overline{N}/n\varepsilon))^{2n}$.*

This theorem is proven by showing that an energy constrained quantum state $\rho : \mathrm{Tr}\left[\hat{N}\rho\right] \leq \overline{N}$ with average photon number $\overline{N}$ is approximated by a rank $r = O\left(\left(e\overline{N}/n\varepsilon\right)^n\right)$ state projected onto the subspace cutoff at total photon number $N_{\max} = \lceil \overline{N}/\varepsilon^2 \rceil$ up to an error ε in trace distance. The corresponding effective subspace has dimension $D = O\left(\left(e\overline{N}/n\varepsilon^2\right)^n\right)$. This approximation of the state by one restricted to the subspace of the CV Hilbert space then allows to leverage the optimal sampling bound of $O\left(Dr\right)$ needed to learn a quantum state [87].

We denote by ρ_D the input state ρ projected onto the D-dimensional subspace, which with $D = O\left(\left(e\overline{N}/\nu^2\varepsilon^2\right)^n\right)$ is $\nu\varepsilon$-close in trace distance to the average energy constrained input state ρ. To reconstruct this state using our shadow tomography protocol, we compute the allowed estimation error on each element a complete basis of D^2 generalized (normalized) Pauli operators $\{P_i\} : \mathrm{Tr}\left[P_i^\dagger P_j\right] = \delta_{ij}$. Assume each Pauli operator is estimated to within additive error ϵ_S. The trace distance between our estimate $\hat{\rho}$ and the finite target state then is

$$\|\hat{\rho} - \rho_D\| \leq \epsilon_S \sum_{i=0}^{D^2-1} \|P_i\|_1 = \epsilon_S D^{\frac{5}{2}}, \tag{E3}$$

which is bounded from above by $(1-\nu)\,\varepsilon$ for

$$\epsilon_S \leq (1-\nu)\,\varepsilon D^{-\frac{5}{2}}. \tag{E4}$$

In total, we hence obtain with Thm. 4 an estimation with error

$$\|\hat{\rho}-\rho_D+\rho_D-\rho\|_1 \leq (1-\nu)\,\varepsilon+\nu\varepsilon=\varepsilon, \tag{E5}$$

with a sample complexity scaling as

$$O\left(\left(\frac{\overline{N}}{\pi}\right)^{2n}\frac{\tilde{V}_O}{\epsilon_S^2}\log\left(\frac{D^2}{\delta}\right)\right)=O\left(\left(\frac{\overline{N}}{\pi}\right)^{2n}\frac{1}{\epsilon_S^2}\log\left(\frac{D}{\delta^{1/2}}\right)\right). \tag{E6}$$

Inserting $\epsilon_S \leq \varepsilon D^{-5/2}$ and $D=O\left(\left(e\overline{N}/\nu^2\varepsilon^2\right)^n\right)$ now yields the sample complexity estimate of

$$O\left(\left(\frac{\overline{N}}{\pi}\right)^{2n}\frac{D^5}{\varepsilon^2}\log\left(\frac{D}{\delta^{1/2}}\right)\right)=O\left(\left(\frac{e^5\overline{N}^7}{\nu^{10}\varepsilon^{10}\pi^2}\right)^n n\log\left(\frac{e\overline{N}}{\nu^2\varepsilon^2\delta^{1/2}}\right)\right) \tag{E7}$$

In Refs. [58, 88] it has been shown that, in addition to the average photon number constraint as in Thm. 6, a restriction on the higher moments of the total photon number of the input state $\mathrm{Tr}\left[\hat{N}^k\rho\right]$ allows for an improved sample complexity in its dependence on ε.

Nevertheless, this scaling is significantly worse as compared to the optimal bound in Thm. 6. This can be attributed to an inefficient choice of observable basis to implement the state tomography protocol as well as the fact that classical shadow tomography is not a particularly optimal way of performing this task.

As full state tomography requires the execution of D^2 known POVMs, it is to be expected that a randomized approach, such as classical shadow tomography, necessarily underperforms. Note that even for qubit systems, the naive shadow-tomography based full-tomography routine discussed here similarly underperforms. To obtain a classical shadow that is indistinguishable from the original state up to trace distance ε, using the results of Ref. [8] and $D=2^n$ together with Eq. (E3) also here we estimate a sample complexity scaling as

$$O\left(\frac{2^{5n}}{\varepsilon^2}\log\left(\frac{4^n}{\delta}\right)\right), \tag{E8}$$

which is not quite as bad as the scaling in Eq. (E6) since the system is natively finite and doesn't require to truncate the tails of the input state, but still remains in a similar ballpark.

3. Lattice based CV shadows

In Ref. [55], the first author of the present manuscript discusses a random Wigner tomography scheme that finds an interpretation of shadow tomography in the style of Ref. [8] via a suitable logical channel twirl of a physical POVM. While also randomizing over lattices, however, the protocol utilizes *classical* probability distributions that mimic the structure of random GKP states. No design condition is used or required in the protocol. By contrast, in showing that GKP states form a 2-design, the present work shows that access to true GKP states serves as a powerful resource for state tomography. This yields a protocol that is more closely aligned with the original design-based shadow tomography scheme.

Appendix F: Implementation of design-based GKP shadow tomography

Given an input state ρ, the original shadow strategy to obtain samples from the effective state $\mathcal{M}_{\mathcal{X}_n}(\rho)$ proceeds as follows.

1. Sample a lattice $\Lambda \in Y_n$ specified by a generator matrix $M=(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_{2n})^T$ [38].
2. Measure each displacement operator $\{D(\boldsymbol{\xi}_i)\}_{i=1}^{2n}$ by performing GKP syndrome extraction.

In this strategy, for each sampled lattice Λ, the displacement operator measurement returns an outcome determined by a vector $\boldsymbol{\alpha} \in \mathcal{P}(\Lambda)$ with probability

$$P_\Lambda(\boldsymbol{\alpha})=\left\langle \Lambda|D^\dagger(\boldsymbol{\alpha})\,\rho D(\boldsymbol{\alpha})\,|\Lambda\right\rangle, \tag{F1}$$

which is precisely the expression $\langle X|\rho|X\rangle$ that determines the distribution over pointers $X=(\Lambda,\boldsymbol{\alpha})$ that we are seeking. This strategy is analogous to the strategy employed in [8] and can be implemented with standard techniques to implement GKP stabilizer measurements, for which the circuits are depicted in Fig. 3.



FIG. 3. Qubit-assisted and GKP-state-assisted stabilizer measurement circuits that implement the shadow tomography protocol. a) An implementation using a qubit-controlled displacement, where the qubit expectation value, $\langle \tilde{X} \rangle = P_+ - P_- = \mathrm{Re}\left\{\mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right)\rho\right]\right\}$, encodes the sampling probabilities. b) An auxiliary square-lattice GKP state $|\mathbb{Z}^2\rangle$ together with a mode-controlled displacement operation (a Gaussian unitary gate) is used to sample the probability distribution $P\left(\tilde{p}\right)$ of momenta on the auxiliary mode, whose Fourier series coefficients encode all values $\mathrm{Re}\left\{\mathrm{Tr}\left[D\left(\ell\boldsymbol{\lambda}\right)\rho\right]\right\}$, $\ell \in \mathbb{Z}$.

On the other extreme, a full batching of all samples to directly obtain the desired statistics $\langle X|\rho|X\rangle$ could be obtained using a SWAP test, provided access to GKP states $|X\rangle \in \mathcal{X}_n$ [46–48]. While this strategy highlights the power of access to GKP states $|X\rangle \in \mathcal{X}_n$, such states are unphysical and this is not a practically feasible protocol.

In this section we seek to improve the former strategy while also explaining the mechanics of the circuits in Fig. 3. In particular, we show how repeated executions of the circuits allow to directly estimate the values $P\left(\boldsymbol{\alpha}\right)$ from Eq. (F1), effectively already processing the sample statistics to probabilities for X for the experiment outlined above.

1. Circuit analysis and batched implementation

In the following it suffices to show how, provided a state ρ, we can estimate $P_\Lambda\left(\mathbf{0}\right) = \langle\Lambda|\rho|\Lambda\rangle$, and the outlined strategy can immediately be extended to estimate $P_\Lambda\left(\boldsymbol{\alpha}\right)$ by inserting appropriate phases. This yields an alternative procedure to estimate the values $\langle X|\rho|X\rangle$ for the set of states constructed in Thm. 1, which is expected to simplify the implementation of our protocol in practice.

By the standard structure of group projectors (see e.g. Ref. [38]) and the uniqueness of the states fixed by symplectic lattices, we have that

$$\langle\Lambda|\rho|\Lambda\rangle = \mathrm{Tr}\left[\Pi_\Lambda\rho\right] = \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})}\,\mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right)\rho\right]. \tag{F2}$$

This is precisely the expression for the probability of a syndrome measurement on the state ρ relative to the GKP state specified by Λ to yield a zero syndrome. As this is a measure zero set within the possible syndromes, we can't rely on measuring syndromes and counting how often a zero event occurs to estimate this number but need to find a better strategy. First, note that the phase $e^{i\Phi(\Lambda)} = \pm 1$ will always be real due to the symplecticity of the lattice [38]. By symmetry of the lattice, this means that only the real parts $\mathrm{Re}\left\{\mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right)\rho\right]\right\}$ of the displacement expectations on the r.h.s. will contribute to the sum. The individual parts of the above sum can be evaluated by finding the values of the characteristic function $c_\rho\left(\boldsymbol{\lambda}\right) = \mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right)\rho\right]$ of the state ρ over the lattice Λ. Although the sum formally necessitates the evaluation of all such values over $\boldsymbol{\lambda} \in \Lambda$, we make the further assumption on the input state ρ to be of *limited bandwidth*, i.e. it has a characteristic function with bounded support, such that there exists some $R_\rho \in \mathbb{R}_+$, such that $c_\rho\left(\boldsymbol{x}\right) = 0\ \forall\, \|\boldsymbol{x}\| \geq R_\rho$.

This assumption effectively limits the maximal resolution of the fine-grained structure of the state but can be physically motivated from the bounded behavior of Wigner functions physical states, see also section A. By choosing a model for the regularization of the state, e.g. using the regularizer $R_\beta = e^{-\beta\hat{N}}$ an estimate for R_ρ can be evaluated, which then provides a means to cutoff the infinite sum to the set $\boldsymbol{\lambda} \in \Lambda \cap B_{2n}\left(R_\rho\right)$, where $B_{2n}\left(R_\epsilon\right)$ is the $2n$-dimensional R_ρ ball (see also a discussion in the previous section and in Ref. [55]).

Given a quantum state ρ, we provide two different settings displayed in Fig. 3 to estimate the values $\langle D\left(\boldsymbol{\lambda}\right)\rangle = \mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right)\rho\right]$.

a. *Estimating the shadow using qubit-controlled displacements*

The first scheme uses an auxiliary qubit initialized in the $|+\rangle = \left(|0\rangle + |1\rangle\right)/\sqrt{2}$ that is coupled to the input n-mode state via a generalized qubit-controlled displacement operation

$$D_{\mathrm{qc}}\left(\boldsymbol{\lambda}\right) = |0\rangle\langle 0|_a \otimes I + |1\rangle\langle 1|_a \otimes D\left(\boldsymbol{\lambda}\right) \tag{F3}$$

and finally measured in the Pauli X basis with probabilities to obtain an outcome ± 1 given by

$$\begin{aligned} P_{\pm} &= \mathrm{Tr}\left(|\pm\rangle\langle\pm| \otimes I\right) D_{\mathrm{qc}}\left(\boldsymbol{\lambda}\right)\left(|+\rangle\langle+|_a \otimes \rho\right) D_{\mathrm{qc}}^{\dagger}\left(\boldsymbol{\lambda}\right) \\ &= \frac{1}{2} \pm \frac{\mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right)\rho\right]}{2}, \end{aligned} \tag{F4}$$

such that we obtain the desired expectation value as $\langle X\rangle_a = P_+ - P_- = \mathrm{Re}\left\{\mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right)\rho\right]\right\}$. This procedure corresponds to the established way to implement stabilizer measurements for qubit-controlled GKP quantum error correction.

Qubit-controlled displacement operators of this form, albeit usually only considered to couple to a single mode, are the typical gates considered for implementations of the GKP code, where qubit-like degrees of freedom are natural to the physical architecture. We refer to Refs. [16, 20, 30, 31, 89] for a more in-depth discussion on the implementation and analysis of this scheme.

While the expectation value of this estimator exactly yields the desired outcome, the information gain in each individual execution of the circuit is rather low at a single bit. Using Hoeffdings inequality on the sampled values for a X measurement on the qubit, whose range is bounded to $\{\pm 1\}$, we obtain for S samples that the sample average $\overline{X}_a^S$ approximates the desired value with accuracy

$$P\left(|\overline{X}_a^S - \mathrm{Re}\left\{\mathrm{Tr}\left[D\left(\boldsymbol{\lambda}\right)\rho\right]\right\}| \geq \epsilon\right) \leq 2e^{-\frac{\epsilon^2}{2}S}, \tag{F5}$$

such that, with high probability $1-\delta$ a target accuracy $\epsilon \ll 1$ is achieved with a number of samples scaling as $S = O\left(\epsilon^{-2}\log\left(1/\delta\right)\right)$.

b. Estimating the shadow using auxiliary GKP states

Alternatively, we may use a GKP encoded auxiliary mode together with a mode-controlled displacement operator (which is a Gaussian unitary operation) to estimate the desired expectation value. This procedure (depicted in panel b in Fig. 3) requires a *fixed* square-lattice GKP resource state and a controlled displacement by primitive vectors of a desired lattice.

We use an auxiliary GKP state initialized in the single mode GKP "sensor state" [28], which is fixed by displacement in the lattice $\Lambda = \mathbb{Z}^2$, and couple it via the Gaussian unitary operation

$$D_c\left(\boldsymbol{\lambda}\right) = e^{-i\hat{q}_a \otimes \boldsymbol{\lambda}^T J \hat{\boldsymbol{x}}} \tag{F6}$$

to the input state. We denote by $|\tilde{p}\rangle\langle\tilde{p}|_p$ the projector onto a fixed momentum $\tilde{p} \in \mathbb{R}$ and by $\rho_{\mathbb{Z}^2} = |\mathbb{Z}^2\rangle\langle\mathbb{Z}^2|$ the (approximative) unique GKP state stabilized with lattice $\mathbb{Z}^2$ with wave function in position- or momentum basis $\psi_{\mathbb{Z}^2}\left(x\right) = {}_q\langle x|\mathbb{Z}^2\rangle = {}_p\langle x|\mathbb{Z}^2\rangle$. After executing the circuit in Fig. 3, a measurement of the auxiliary state in the momentum basis produces the distribution

$$\begin{aligned} P\left(\tilde{p}\right) &= \mathrm{Tr}\left[\left(|\tilde{p}\rangle\langle\tilde{p}| \otimes I\right) D_c\left(\boldsymbol{\lambda}\right)\left(\rho_{\mathbb{Z}^2} \otimes \rho\right) D_c^{\dagger}\left(\boldsymbol{\lambda}\right)\right] \\ &= \int_{\mathbb{R}} dq dq'\, e^{i\tilde{p}\left(q-q'\right)} \langle q'|\rho_{\mathbb{Z}^2}|q\rangle\, \mathrm{Tr}\left[D\left(\left(q-q'\right)\boldsymbol{\lambda}/\sqrt{2\pi}\right)\rho\right] \\ &= \int_{\mathbb{R}} dq\, e^{i\tilde{p}q} \left\langle \mathbb{Z}^2|e^{-iq\hat{p}}|\mathbb{Z}^2\right\rangle \mathrm{Tr}\left[D\left(q\boldsymbol{\lambda}/\sqrt{2\pi}\right)\rho\right]. \end{aligned} \tag{F7}$$

Using the fact that $\left\langle \mathbb{Z}^2|e^{-iq\hat{p}}|\mathbb{Z}^2\right\rangle = \text{Ш}_{\sqrt{2\pi}}\left(q\right) = \sum_{n\in\mathbb{Z}} \delta\left(q - \sqrt{2\pi}n\right)$, we find that

$$P_{\boldsymbol{\lambda}}\left(\tilde{p}\right) = \sum_{n\in\mathbb{Z}} e^{i\sqrt{2\pi}n\tilde{p}}\, \mathrm{Tr}\left[D\left(n\boldsymbol{\lambda}\right)\rho\right]. \tag{F8}$$

The outcome statistics $P\left(\tilde{p}\right)$ is a periodic distribution invariant under $\tilde{p} \mapsto \tilde{p} + \sqrt{2\pi}$, $n \in \mathbb{Z}$ and its Fourier coefficients encode the expectation values $\mathrm{Tr}\left[D\left(n\boldsymbol{\lambda}\right)\rho\right]$ for all $n \in \mathbb{Z}$. Given the GKP state $|\mathbb{Z}^2\rangle$, it hence suffices to estimate the statistics $P_{\boldsymbol{\lambda}}\left(\tilde{p}\right)$ for *primitive* lattice vectors $\boldsymbol{\lambda} \in \Lambda_{\mathrm{pr}}$. The set of primitive lattice vectors Λ_{pr} is the set of smallest lattice vectors that are not proportional to each other. For example, for $\mathbb{Z}^2$, $\mathbb{Z}^2_{\mathrm{pr}}$ contains all vectors $(a,b)^T \in \mathbb{Z}^2$ where the coefficients are coprime $\gcd\left(a,b\right) = 1$. Crucially, every lattice vector $\boldsymbol{\lambda} \in \Lambda$ can be written as $n\boldsymbol{\lambda}_{\mathrm{pr}}$ with a primitive lattice vector $\boldsymbol{\lambda}_{\mathrm{pr}}$ and a unique integer n. Using this circuit and again using the finite bandwith of physical states, we see that the set of vector for which the quantum circuit needs to be evaluated really is $\Lambda_{\mathrm{pr}} \cap B_{2n}\left(R_\rho\right)$.

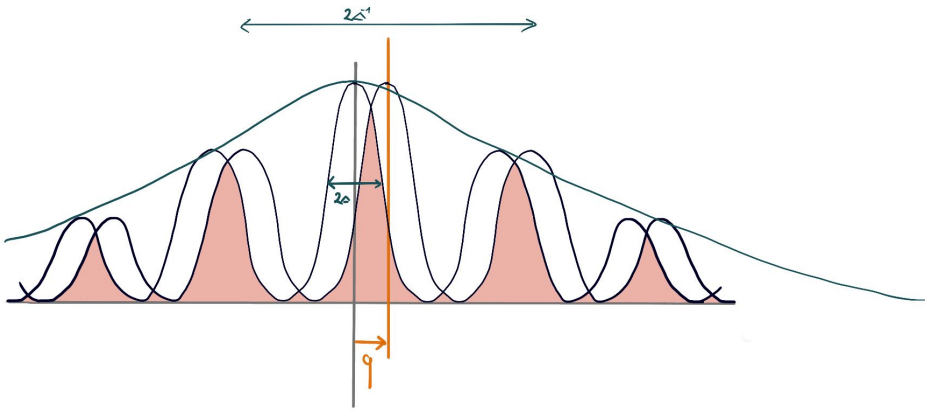


FIG. 4. A sketch of the overlap integral of an approximate GKP state with its own displacement. An approximate GKP state approximates a state with periodic structure. For finite bandwidth states the individual peaks are of finite resolution and the state is globally deformed so that it is normalizable. Here, we illustrate the "standard" approximation implemented by regularizer $R_\beta = e^{-\beta \hat{n}}$, where $\Delta^2 = \tanh(\beta/2)$ quantifies the variance of a train of Gaussian peaks regularized by an envelope with variance Δ^{-2} [19, 90]

In physical implementations, we won't have access to the formal GKP states $|\mathbb{Z}^2\rangle$ but rather states of a regularized form such as $|\mathbb{Z}^2_\beta\rangle = N_\beta e^{-\beta \hat{N}}$. The effect of using these approximate states can be interpreted using Eq. (F7). Substituting $|\mathbb{Z}^2\rangle$ with its approximate version changes the factor $\langle \mathbb{Z}^2 | e^{-iq\hat{p}} | \mathbb{Z}^2 \rangle$ into the position basis overlap integral

$$\langle \mathbb{Z}^2_\beta | e^{-iq\hat{p}} | \mathbb{Z}^2_\beta \rangle = \int_{\mathbb{R}} dq' \, \psi^*_{\mathbb{Z}^2_\beta}(q') \, \psi_{\mathbb{Z}^2_\beta}(q'+q) \, . \tag{F9}$$

This hopping-like integral computes "how periodic" the approximate state is by computing its overlap with its displaced version. We illustrate the form of this integral in Fig. 4. Due to the finite bandwidth of these approximate states, each peak of the distribution now has acquired a finite width $\propto \Delta$, such that values of q close to, but not exactly lying on, integer multiples of $\sqrt{2\pi}$ now can also produce non-zero contributions. Furthermore, due to the presence of an envelope of width $\propto \Delta^{-1}$ on the state (that renders it normalizable), now large values of $q \ll \Delta^{-1}$ can render the corresponding contributions zero, even if $q \in \sqrt{2\pi}\mathbb{Z}$ should be an allowed value by the intended periodicity. While this analysis has been provided with the aim of obtaining displacement expectation values, it showcases how the Fourier spectrum of homodyne measurement outcomes under the circuit discussed above contains the information of an infinite number of such displacement operator expectation values. It is to be expected that this information can be of use in the refined characterization of GKP states, e.g. through generalizatons of the effective squeezing parameter [91, 92], and we leave a refined study to future work.

c. *Typicality*

In the constructions above, we approximate the sum

$$\langle \Lambda | \rho | \Lambda \rangle = \sum_{\boldsymbol{\lambda} \in \Lambda} e^{i\Phi(\boldsymbol{\lambda})} \operatorname{Tr}\left[D(\boldsymbol{\lambda}) \, \rho \right] \tag{F10}$$

by truncation beyond a radius R. Under a randomly chosen lattice $\Lambda \in Y_n$, we wish to estimate how many points there are in our set of displacements to measure. The relevant number of points for our two protocols we need to estimate is hence

$$|\Lambda \cap B_{2n}(R)| \quad \text{and} \quad |\Lambda_{\mathrm{pr}} \cap B_{2n}(R)| . \tag{F11}$$

In fact, one may apply the mean value formula to obtain the average values of these estimates simply by choosing $f(\lambda) = \chi_R(0 \neq \|\boldsymbol{\lambda}\| \leq R)$ in the mean value formula in Eq. (A24). We call $\Lambda^\times = \Lambda - \{0\}$. Applying the mean value formula shows that, that the expected number of (primitive) lattice points of $\Lambda^\times$ in $B_{2n}(R)$ is given by the volume of a $2n$-ball $\langle \Lambda | \rho | \Lambda \rangle$.

$$\left\langle |\Lambda^\times \cap B_{2n}(R)| \right\rangle_{Y_n} = V_{2n}(R) \quad \text{and} \quad \left\langle |\Lambda_{\mathrm{pr}} \cap B_{2n}(R)| \right\rangle_{Y_n} = V_{2n}(R)/\zeta(n) \, , \tag{F12}$$

where $\zeta(n)$ is the Riemann zeta function [40].

The volume of an $2n$-ball is

$$V_{2n}(R) = \frac{\pi^n}{n!} R^{2n} \overset{\text{Stirling}}{\approx} \frac{1}{\sqrt{2\pi n}} \left(2\pi e \frac{R^2}{n}\right)^{n/2}, \tag{F13}$$

which vanishes in the limit $n \to \infty$ for $R = o(\sqrt{n})$ and grows as R^n otherwise, while $\lim_{n\to\infty} \zeta(n) = 1$. One observation is hence this: as long as the bandwidth of the state can be assumed to scale as $R = o(\sqrt{n})$, our protocol to estimate the expectation values $\langle\Lambda|\rho|\Lambda\rangle$ remains efficient on average. Ref. [40] even derives second moment extensions to the mean value formula over Y_n, which yield bounds on the variances of these counters for $n \geq 2$ given by

$$\begin{aligned} \left\langle \left(|\Lambda^\times \cap B_{2n}(R)| - V_{2n}(R)\right)^2 \right\rangle_{Y_n} &\leq 4\zeta(n) V_{2n}(R), \\ \left\langle \left(|\Lambda_{\text{pr}} \cap B_{2n}(R)| - V_{2n}(R)\right)^2 \right\rangle_{Y_n} &\leq 4\frac{V_{2n}(R)}{\zeta(2n)}. \end{aligned} \tag{F14}$$

These bounds tell us that, in the limit of large n and for states of bandwidth bounded by $R = o(\sqrt{n})$, the number of displacement operators whose expectation we need to estimate to obtain a good approximation to $\langle\Lambda|\rho|\Lambda\rangle$ remains small and vanishes asymptotically in n. In this limit the typical lattice will have shortest vector length scaling with $\sqrt{n}$ and only the trivial point $\{0\}$ and a small number of short vectors close by remain within the ball $B_{2n}(R)$ in phase space. To contain non-trivial components for typical lattices, the input state is hence required to have bandwidth limit of $R = O(n)$.

We understand this artifact in analogy to a qubit-based protocol, where the natural analogue of our ensemble $\mathcal{X}_n$ are the qubit stabilizer states. If the input state is *itself* a stabilizer state, most Pauli expectation values will evaluate to 0, except on the exponentially small fraction of Pauli operators that exactly align with the stabilizer group of the input state. Most information can hence be gained if the input state is instead well-distributed among a large fraction of stabilizer states.

Back to the present situation, a meaningful quantifier for how much a state is overlapping with the various GKP pointer states is the phase-space bandwidth R and lattice instances will yield non-trivial shadows for $R = \Omega(\sqrt{n})$. Provided a fixed length-scale, the number of relevant displacement expectation values is estimated by $V_{2n}(R)/\zeta(n) + 2k\sqrt{V_{2n}(R)/\zeta(2n)}$, where $k = O(1)$ determines the scale up to which we wish to consider fluctuations around the typical instance. When n is fixed, the number of expectation values needed scales efficiently as R^{2n}.

Appendix G: Application: variational preparation of thermal states many-body states

In this appendix we sketch the application of our protocol to variationally prepare a many-body thermal state based on a perturbation series approach. Given a description of a thermal state $\rho_\beta \propto e^{-\beta H}$ for an n-mode Hamiltonian H, we show for a GKP state $|\Lambda\rangle$, how to compute the value of $\langle\Lambda|e^{-\beta H}|\Lambda\rangle$ under a perturbative expansion. Having access to these values then one allows to apply the above described protocol to estimate the state overlap $\text{Tr}[\rho_\beta \rho_{\text{model}}]$ between the target state ρ_β and another state ρ_{model}, which in turn can be used to estimate their Hilbert-Schmidt distance $d_{\text{HS}}(\rho_\beta, \rho_{\text{model}})$. By preparing the model state $\rho_{\text{model}} = |\psi_{\text{model}}\rangle\langle\psi_{\text{model}}|$ via a parametrized quantum channel then would allow to approximate the thermal state using well-studied variational techniques. The key element in this chain of argumentation is the ability to classically find the values $\langle\Lambda|e^{-\beta H}|\Lambda\rangle$ needed to implement the classical post-processing in the shadow tomography protocol. Furthermore, as discussed in the previous section, evaluating (or bounding) this value also directly implies a bound on the sample overhead necessary to implement this protocol.

Consider a n-mode bosonic system with a Hamiltonian given by

$$H = \hat{N} + \epsilon P(\hat{\boldsymbol{x}}), \tag{G1}$$

where $P(\hat{\boldsymbol{x}})$ is a finite polynomial in the quadratures $\hat{q}_1, \ldots, \hat{p}_n$. The thermal state of such a system at inverse temperature β is (up to the potentially unknown partition function) proportional to

$$\rho_\beta = e^{-\beta\left(\hat{N} + \epsilon P(\hat{\boldsymbol{x}})\right)}, \tag{G2}$$

and, in the limit of small $0 \leq \epsilon \ll 1$ can be expanded in a perturbative series

$$\rho_\beta = \tilde{P}(\hat{\boldsymbol{x}})\, e^{-\beta\hat{N}}, \tag{G3}$$

where we assume $\tilde{P}(\hat{\boldsymbol{x}})$ to be a polynomial in Weyl-ordered form [56, 59]. The existence of this form of expression can be again found using the formula for exponential derivatives [61] to expand $\rho_\beta(\epsilon) = \rho_\beta(0) + \epsilon \frac{\partial \rho_\beta(\epsilon)}{\partial \epsilon}|_{\epsilon=0}$, which yields

$$\frac{\partial \rho_\beta(\epsilon)}{\partial \epsilon}|_{\epsilon=0} = -\beta \int_0^1 dy\, e^{-y\beta\hat{N}} P(\hat{\boldsymbol{x}})\, e^{y\beta\hat{N}} e^{-\beta\hat{N}} = -\beta \int_0^1 dy\, P\left(e^{-y\beta\hat{N}} \hat{\boldsymbol{x}} e^{y\beta\hat{N}}\right) e^{-\beta\hat{N}}. \tag{G4}$$

It holds for each pair of quadrature and value $r \in \mathbb{R}$ that

$$e^{-r\hat{n}} \begin{pmatrix} \hat{q} \\ \hat{p} \end{pmatrix} e^{r\hat{n}} = \begin{pmatrix} \cosh(r) & i\sinh(r) \\ -i\sinh(r) & \cosh(r) \end{pmatrix} \begin{pmatrix} \hat{q} \\ \hat{p} \end{pmatrix}, \tag{G5}$$

such that we see that the expression in eq. (G4) can be integrated and rewritten in the desired form.

We now leverage the decomposition already provided in Eq. (B23),

$$e^{-\beta\hat{N}} = \left(1 - e^{-\beta}\right)^{-n} \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha}\, e^{-\pi\|\boldsymbol{\alpha}\|^2/2\Delta^2} D(\boldsymbol{\alpha}), \tag{G6}$$

with $\Delta^2 = 2\tanh(\beta/2)$ to write

$$\begin{aligned} \left(1 - e^{-\beta}\right)^n \tilde{P}(\hat{\boldsymbol{x}})\, e^{-\beta\hat{N}} &= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha}\, e^{-\pi\|\boldsymbol{\alpha}\|^2/2\Delta^2} \tilde{P}(\hat{\boldsymbol{x}})\, D(\boldsymbol{\alpha}) \\ &= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha}\, e^{-\pi\|\boldsymbol{\alpha}\|^2/2\Delta^2} \left\{\tilde{P}(\boldsymbol{\nabla}_{\boldsymbol{\alpha}})\, D(\boldsymbol{\alpha})\right\} \\ &= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha} \left\{\tilde{P}\left((1 + i\Delta^{-2})\sqrt{\frac{\pi}{2}} J\boldsymbol{\alpha}\right) e^{-\pi\|\boldsymbol{\alpha}\|^2/2\Delta^2}\right\} D(\boldsymbol{\alpha}), \end{aligned} \tag{G7}$$

where we defined

$$\boldsymbol{\nabla}_{\boldsymbol{\alpha}} \coloneqq \frac{-i}{\sqrt{2\pi}} J\boldsymbol{\nabla} + \sqrt{\frac{\pi}{2}}\boldsymbol{\alpha} \tag{G8}$$

and used integration by parts together with the compactness of the Gaussian factor. This expression determines the characteristic function of the perturbative expansion of the thermal state. Using this expression, we can compute for a GKP state $|\Lambda\rangle$ the expectation value

$$\left\langle \Lambda \middle| \tilde{P}(\hat{\boldsymbol{x}})\, e^{-\beta\hat{N}} \middle| \Lambda \right\rangle = \left(1 - e^{-\beta}\right)^{-n} \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} \tilde{P}\left((1 + i\Delta^{-2})\sqrt{\frac{\pi}{2}} J\boldsymbol{\lambda}\right) e^{-\pi\|\boldsymbol{\lambda}\|^2/2\Delta^2}. \tag{G9}$$

This series naturally converges for any finite polynomial $\tilde{P}$. Let $\overline{P}(\boldsymbol{x})$ be a polynomial function of the coefficients of $\boldsymbol{x} \in \mathbb{R}^{2n}$ such that

$$\overline{P}(\boldsymbol{x}) \geq \left|\tilde{P}\left((1 + i\Delta^{-2})\sqrt{\frac{\pi}{2}} J\boldsymbol{x}\right)\right| \quad \forall \boldsymbol{x} \in \mathbb{R}^{2n}. \tag{G10}$$

The above expression is bounded by

$$\left|\left\langle \Lambda \middle| \tilde{P}(\hat{\boldsymbol{x}})\, e^{-\beta\hat{N}} \middle| \Lambda \right\rangle\right| \leq \left(1 - e^{-\beta}\right)^n \sum_{\boldsymbol{\lambda}\in\Lambda} \overline{P}(\boldsymbol{x})\, e^{-\pi\|\boldsymbol{\lambda}\|^2/2\Delta^2} = \Theta_{\Lambda,\overline{P}}\left(\frac{i}{4\Delta^2}\right), \tag{G11}$$

where Θ are known as a *weighted theta series* in the special case where $\overline{P}$ is a harmonic polymonial [44].

What have we gained? We have shown that the GKP-state expectation value of the perturbative expansion of any (unnormalized) multi-mode thermal state can be evaluated in terms of a Gaussian lattice sum. We have also bounded this sum by a Gaussian sum which is easy to study and where one can devise tractable approximations. Under the specific assumptions that the perturbative expansion is bounded by a harmonic polynomial, this expression turns into a weighted theta series.

Weighted theta series are number theoretically interesting objects, encoding information about the distribution of lengths of lattice vectors together with their distribution over spheres. Furthermore, they turn out to be modular forms [44], which are functions with a high degree of symmetry that are potentially useful for the derivation of stronger bounds and evaluation tactics for these functions. While we leave a further study of the protocol sketched in this appendix to future work, we have demonstrated that the CV shadow tomography protocol devised in this manuscript can find applications in non-trivial algorithmic tasks, such as estimating the expectation value of multi-mode thermal states relative to a given bosonic state. This can be leveraged for the variational preparation of such a thermal state.