# Limitations of local update recovery in stabilizer-GKP codes: a quantum optimal transport approach

## Robert König

Munich Center for Quantum Science and Technology & Department of Mathematics, School of Computation, Information and Technology, Technical University of Munich, 85748 Garching, Germany

## Cambyse Rouzé \*

Inria, Télécom Paris - LTCI, Institut Polytechnique de Paris, 91120 Palaiseau, France Zentrum Mathematik, Technische Universität München, 85748 Garching, Germany

Local update recovery seeks to maintain quantum information by applying local correction maps alternating with and compensating for the action of noise. Motivated by recent constructions based on quantum LDPC codes in the finite-dimensional setting, we establish an analytic upper bound on the fault-tolerance threshold for concatenated GKP-stabilizer codes with local update recovery. Our bound applies to noise channels that are tensor products of one-mode beamsplitters with arbitrary environment states, capturing, in particular, photon loss occurring independently in each mode. It shows that for loss rates above a threshold given explicitly as a function of the locality of the recovery maps, encoded information is lost at an exponential rate. This extends an early result by Razborov from discrete to continuous variable (CV) quantum systems.

To prove our result, we study a metric on bosonic states akin to the Wasserstein distance between two CV density functions, which we call the bosonic Wasserstein distance. It can be thought of as a CV extension of a quantum Wasserstein distance of order 1 recently introduced by De Palma et al. in the context of qudit systems, in the sense that it captures the notion of locality in a CV setting. We establish several basic properties, including a relation to the trace distance and diameter bounds for states with finite average photon number. We then study its contraction properties under quantum channels, including tensorization, locality and strict contraction under beamsplitter-type noise channels. Due to the simplicity of its formulation, and the established wide applicability of its finite-dimensional counterpart, we believe that the bosonic Wasserstein distance will become a versatile tool in the study of CV quantum systems.

### I. Introduction

Universal quantum computers are envisioned to provide a powerful computational resource. Whether such devices will be available in the future depends on our ability to find experimentally achievable methods to protect the quantum information from the unavoidable presence of noise while simultaneously maintaining the ability to operate on it. The celebrated threshold theorem proves that, at least in theory, such fault-tolerance constructions exist [1], as long as the individual building blocks are sufficiently reliable: If e.g., the noise rate of individual qubits and physical gates is below a constant value called the fault-tolerance threshold, then a size-S quantum circuit can be emulated with error at most  $\varepsilon$  by a noisy circuit of size polynomial in S and  $\log 1/\varepsilon$ . Quantitative lower bounds of this threshold value (obtained by studying concrete fault-tolerance proposals for a given noise model) are of central importance as they provide clear experimental benchmarks. Conversely, upper bounds on the fault-tolerance threshold establish fundamental limitations on quantum computing and may guide the search for better fault-tolerance constructions.

### A. Fault-tolerant quantum memories

Fault-tolerant quantum memories: To establish an upper bound on the fault-tolerance threshold, it suffices to consider the task of fault-tolerantly implementing the concatenation of T identity gates: In other words, this is the problem of realizing a noise-resilient memory that preserves quantum information for a time T. To preserve k logical qubits by using n physical ones, a general error correction strategy relies on a quantum code  $\mathcal{C} \subset (\mathbb{C}^2)^{\otimes n}$ encoding k logical qubits into n physical ones: Starting with an encoded version of the initial state, a recovery map  $\mathcal{R}_t$  is applied to the n qubits at each time step t = 1, ..., T. The choice of  $(\mathcal{C}, \mathcal{R} = \{\mathcal{R}_t\}_{t=1}^T)$  is intended to ensure that the information is preserved even if the execution of the recovery maps is interleaved with (sufficiently weak) noise. For concreteness, assume that each physical qubit undergoes depolarizing noise  $\mathcal{N}_p(\rho) = (1-p)\rho + pI/2$ , with depolarizing parameter  $p \in (0,1)$ . Then we are interested in the (worst-case) fidelity with which the composed channel

$$\mathcal{R}^{(T)} := \bigcap_{t=T}^{1} (\mathcal{R}_t \circ \mathcal{N}_p^{\otimes n}) \tag{1}$$

preserves states supported on C, or more precisely, the trade-off relation between T, this fidelity, and the num-

bers (n,k) of physical and logical qubits, respectively. Clearly, this trade-off depends on the set of allowed recovery maps. We note that to make the notion of preserving states supported on  $\mathcal{C}$  meaningful, one should generally let the final recovery map  $\mathcal{R}_T$  be an arbitrary quantum channel even when  $\{\mathcal{R}_j\}_{j=1}^{T-1}$  are restricted. This is similar to the definition of entanglement fidelity and ensures that we are quantifying the fidelity of the in-principle recoverable information after time T.

As an example, if C is an [n, k, d]-stabilizer code with stabilizer generators  $\{S_j\}_{j=1}^{n-k}$ , one could consider  $\mathcal{R}_t = \mathcal{R}$  for all t with a recovery map  $\mathcal{R}$  of the form

$$\mathcal{R}(\rho) = \sum_{s \in \{0,1\}^{n-k}} C(s)\Pi(s)\rho\Pi(s)C(s)^{\dagger}, \qquad (2)$$

where  $\Pi(s) = \prod_{j=1}^{n-k} \frac{1}{2}(I+(-1)^{s_j}S_j)$  projects onto the syndrome-s subspace and C(s) is a Pauli-correction classically computed from s, which maps the state back to the code space. Then Eq. (1) corresponds to a scenario where each correction and recovery operation is ideally realized. In the case of probabilistic Pauli noise, the logical overall error model after time T is then obtained by taking the T-fold convolution of the effective logical noise associated with a single step and recovery combination  $\mathcal{R} \circ \mathcal{N}_p^{\otimes n}$ .

## B. Local update recovery

A more interesting case – the one considered here – is where locality is imposed on the operators  $\{\mathcal{R}_t\}_{t=1}^{L-1}$ . Specifically, let us say that the recovery scheme  $(C, \mathcal{R} =$  $\{\mathcal{R}_t\}_{t=1}^T$ ) is  $\ell$ -local if for each  $t=1,\ldots,T-1$ , the map  $\mathcal{R}_t = \prod_j \mathcal{R}_{A_j^{(t)}}$  is a product of quantum channels  $\mathcal{R}_{A_{:}^{(t)}}$  each acting on pairwise disjoint subsets  $A_{j}^{(t)}$ of at most  $|A_j^{(t)}| \leq \ell$  qubits  $A_j^{(t)} \subset \{1, \dots, n\}$ . Clearly, an O(1)-local recovery scheme is attractive from an experimental point of view, but it should also be noted that this kind of recovery is highly restricted. For example, in the case of a stabilizer codes, it is not generally sufficient to have local (i.e., low-weight) stabilizer generators: even if syndrome extraction is achieved by measuring constant-size subsets of qubits, typical decoders pool and globally process corresponding measurement results to determine a suitable (global) correction. In contrast, an O(1)-local recovery scheme requires that syndrome extraction, computation and application of an associated correction is achieved by a collection of independent processes acting on disjoint subsets of qubits. We refer to such an O(1)-local scheme as a local update recovery.

Remarkably, despite the highly restrictive locality requirement, O(1)-local recovery schemes exhibiting a threshold property exist. Indeed, recent work focusing on

the break-through discovery of quantum low density parity check (LDPC) codes [2] establishes a threshold property for such a scheme under a more general noise model than independent depolarizing noise on each qubit [3]. In fact, the result of [2] applies even to the case where the implementation of each recovery map  $\mathcal{R}_t$  is noisy. The construction is based on the fact that expander codes, i.e., codes whose Tanner graph is an expander graph, can be decoded by a local process, a fundamental insight first observed by Sipser and Spielman [4]. While the original decoder iteratively goes through the Tanner graph, extracting local syndrome information and flipping bits to decrease the syndrome weight, this process can be parallelized. These ideas were introduced to the quantum context in Refs. [5], where the so-called small-set-flip decoder was introduced and applied in a fault-tolerance construction. Remarkably, the new results of [2] show that for certain quantum LDPC codes, this locally defined error correction process works even if recovery maps are interspersed with noise maps as in (1). It should be noted that this analysis is different from prior works involving local updates such as Toom's rule [6, 7] or certain cellular automata [8–11]: The latter studies consider decoding/recovery by iterative application of local maps (possibly assisted by local memories), but do not analyze the effect of additional errors within iterations.

### C. Bounding local update recovery for qubits

While Ref. [2] gives a scheme for local update recovery for noise rates  $p < p_0$  below some threshold  $p_0$ , here we ask for upper bounds on the threshold. In fact, such an upper bound is known [12, 13]: For any  $\ell$ -local recovery scheme  $(\mathcal{C}, \mathcal{R} = \{\mathcal{R}_t\}_{t=1}^T)$ , if the depolarizing probability is above  $1 - 1/\ell$ , then at any depth  $T \geq C \log(n)$  for a sufficiently large constant C > 0 and for any two input states  $\rho, \sigma$ :

$$\|\mathcal{R}^{(T)}(\rho) - \mathcal{R}^{(T)}(\sigma)\|_{1} = \mathcal{O}(1/n). \tag{3}$$

In words, the circuit loses distinguishability (and hence encoded information) at an exponential rate in T for any depolarizing probability

$$p > 1 - \frac{1}{\ell} =: p_0(\ell) \ .$$
 (4)

We refer to Section III A below for more details on the bound (3).

#### D. Continuous variable error correction

While the previous discussion concerned discrete variable (DV) systems, many systems of practical interest

involve continuous variables (CV), such as harmonic oscillators. The last twenty years have seen the development of various CV codes, e.g. encodings into Fock states [14–16], coherent states [17, 18], or position and momentum eigenstates [19–21]. Among these, the so-called Gottesman-Kitaev-Preskill (GKP) codes [21] hold a prominent position.

By design, a GKP code corrects small random shifts in position and momentum. In contrast, the pure loss channel is the most common incoherent error process in optical and microwave cavities [22]. Despite this apparent discrepancy, recent numerical studies showed that GKP codes significantly outperform all other codes for most values of the loss rate [23]. In [23], these numerical results were also backed by an analytical upper bound on the channel infidelity between the noise composed with the GKP recovery channel and the ideal identity channel. While a CV quantum error-correcting code with a single bosonic mode and a single ancilla qubit like the GKP code can suppress relevant errors such as photon losses in a hardware-efficient manner, it should also be noted that logical error rates cannot be suppressed to an arbitrarily small value with this minimal architecture. To further suppress the residual errors, these bosonic codes should, for example, be concatenated with a qubit errorcorrecting code [10, 24]. Recent proposals include concatenation with a repetition code [25], the surface code [26–28] or cluster states in the setting of measurementbased quantum computing [27, 29, 30]. In those works, the code capacity thresholds were obtained by assuming that only qubits can fail, i.e., gates, state preparations, and measurements are assumed perfect. A more realistic approach was recently taken for the surface code in [31]. However, these works concentrate on upper bounds on the threshold noise.

The goal of the present paper is to provide an analytic lower bound on the accuracy threshold for GKP-stabiliser concatenated codes with local update recovery, realistic input state preparation and measurements and in presence of photon-loss, akin to the threshold (4).

## II. Bosonic local update recovery

## A. The setup

Consider a local recovery scheme  $(C, \mathcal{R} = \{\mathcal{R}_t\}_{t=1}^T)$  using n physical qubits, based on an [n, k]-stabilizer code C and recovery maps  $\mathcal{R}_t$  given by Clifford circuits as that of [21]. Using code concatenation, such a scheme can immediately be lifted to a scheme operating on n bosonic modes. Here we consider the "standard" Gottesman-Kitaev-Preskill (GKP) code encoding a single qubit into a single harmonic oscillator. Concatenating the stabi-

lizer  ${\mathcal C}$  with the GKP code then provides a stabilizer-GKPcode  $C_{osc}$  encoding k logical qubits into n oscillators. Examples include the surface- or toric-GKP codes [32, 33]. A natural recovery (error correction) map for this concatenated code proceeds by applying a bosonic recovery map  $\mathcal{R}_{\mathsf{GKP}}$  to each mode, and subsequently applying  $\mathcal{R}$ to the collection of n modes. Here  $\overline{\mathcal{R}}$  is a GKP-encoded version of the original recovery map  $\mathcal{R}$  for the stabilizer code  $\mathcal{C}$ . We write  $\mathcal{R}^{\mathsf{osc}} := \overline{\mathcal{R}} \circ \mathcal{R}^{\otimes n}_{\mathsf{GKP}}$  for this composed map. Applying this transformation to each recovery map  $\mathcal{R}_t$  yields an  $\ell$ -local recovery scheme ( $\mathcal{C}_{osc}, \mathcal{R}_{osc} =$  $\{\mathcal{R}_t^{\text{osc}}\}_{t=1}^T$ , whose operations act on disjoint subsets of  $\ell$  bosonic modes. To make this construction physical and concrete, we use approximate GKP-states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$ with finite average energy, and a Steane-type error correction circuit for  $\mathcal{R}_{\mathsf{GKP}}$  relying on unsharp (finite-variance) position- and momentum measurements (see Figure 1 in the supplementary material).

Here we study to which degree bosonic local update recovery schemes constructed in this way can withstand photon loss. That is, we consider loss of a photon modeled by a one-mode beamsplitter of transmissivity  $\lambda \in (0,1)$  with arbitrary environment state  $\sigma_E$ :

$$\mathcal{N}_{\lambda}(\rho) = \operatorname{tr}_{E} \left[ U_{\lambda}(\rho \otimes \sigma_{E}) U_{\lambda}^{\dagger} \right], \tag{5}$$

with the beamsplitter unitary  $U_{\lambda} \equiv \exp((a^{\dagger}b - b^{\dagger}a)\arccos(\sqrt{\lambda}))$ . Although this is not required for our results to hold, we will fix the state  $\sigma_E$  to be the same each time the noise channel  $\mathcal{N}_{\lambda}$  is applied, for sake of simplicity. We are interested in the map

$$\mathcal{R}_{\mathsf{osc}}^{(T)} := \bigcap_{t=T}^{1} (\mathcal{R}_{t}^{\mathsf{osc}} \circ \mathcal{N}_{\lambda}^{\otimes n}) , \qquad (6)$$

where each mode undergoes photon loss independently of the other modes, and this noise process is alternated with recovery maps obtained by concatenation.

### B. Bounding local update recovery for concatenated GKP-stabilizer codes

Our main result is an upper bound for the bosonic setup akin to Eq. (3). It states the following:

**Theorem 1** (Limits to local update recovery). Let  $(C_{\text{osc}}, \mathcal{R}_{\text{osc}} = \{\mathcal{R}_t^{\text{osc}}\}_{t=1}^T)$  be a bosonic recovery scheme on n modes derived by code concatenation from a (stabilizer-type) recovery scheme  $(C, \mathcal{R} = \{\mathcal{R}_t\}_{t=1}^T)$ . Then there exist constants  $\kappa$  depending on the noise parameters  $\lambda, \sigma_E$ , and C depending only on  $\{\mathcal{R}_t\}_{t=1}^T$  such that

$$\left\| \mathcal{R}_{\text{osc}}^{(T)}(\rho) - \mathcal{R}_{\text{osc}}^{(T)}(\sigma) \right\|_{1} \leq 4\kappa \left( C \sqrt{\lambda} \right)^{T} \sqrt{2n(n+E)} \,, \quad (7)$$

for any two n-mode quantum states  $\rho, \sigma$  with  $\operatorname{tr}[\rho N], \operatorname{tr}[\sigma N] \leq E$ , and any photon loss strength (transmissivity)  $\lambda \in (0,1)$ . Furthermore, if  $(\mathcal{C}, \mathcal{R} = \{\mathcal{R}_t\}_{t=1}^T)$  is an  $\ell$ -local recovery scheme, then  $C = O(\ell)$ .

Theorem 1 implies, in particular, that the distinguishability between the outputs  $\mathcal{R}_{\mathsf{osc}}^{(T)}(\rho)$  and  $\mathcal{R}_{\mathsf{osc}}^{(T)}(\sigma)$  decays exponentially in the number T of time-steps whenever

$$\lambda < \frac{1}{C^2} \tag{8}$$

thus establishing an upper bound on the fault-tolerance threshold of the construction against photon loss errors.

### III. Quantum optimal transport toolbox

### A. Qubit threshold by quantum optimal transport

Before we introduce our main new tool in the proof of Theorem 1, let us provide a modern (albeit with slightly worse constants) treatment of the result of Razborov [13] via the use of recently introduced quantum optimal transport metrics [34]. The Wasserstein-1 distance between any two n-qubit quantum states  $\rho$ ,  $\sigma$  is defined as

$$W_1(\rho, \sigma) := \sup_{\|X\|_L \le 1} \left| \operatorname{tr} \left[ X(\rho - \sigma) \right] \right|. \tag{9}$$

The optimization in (9) is over observables X of Lipschitz constant

$$||X||_{L} := 2 \max_{i \in [n]} \min_{X^{(i)}} ||X - X^{(i)} \otimes I_{i}||_{\infty}$$
 (10)

less than or equal to 1, where the minimization above is over observables  $X^{(i)}$  supported on  $[n]\backslash\{i\}$ . As opposed to the trace distance, where the optimization is over observables of operator norm  $\|X\|_{\infty} \leq 1$ , the Wasserstein distance is able to capture local differences between the states  $\rho$  and  $\sigma$ . In particular, whenever  $\rho$  and  $\sigma$  share the same marginals over a set of k qubits,

$$\frac{1}{2} \|\rho - \sigma\|_1 \le W_1(\rho, \sigma) \le \frac{3(n-k)}{4} \|\rho - \sigma\|_1.$$
 (11)

This results in the following control on the Wasserstein distance at the outputs of any  $\ell$ -local recovery map  $\mathcal R$  and its inputs:

$$W_1(\mathcal{R}(\rho), \mathcal{R}(\sigma)) \le \frac{3\ell}{2} W_1(\rho, \sigma).$$
 (12)

Moreover, the Wasserstein distance contracts after each layer of independent depolarizing noise:

$$W_1(\mathcal{N}_p^{\otimes n}(\rho), \mathcal{N}_p^{\otimes n}(\sigma)) \le (1-p)W_1(\rho, \sigma). \tag{13}$$

Using the bounds (11), (12) and (13) iteratively, we have

$$\begin{split} \left\| \mathcal{R}^{(T)}(\rho - \sigma) \right\|_1 &\leq 2 \, W_1(\mathcal{R}^{(T)}(\rho), \mathcal{R}^{(T)}(\sigma)) \\ &\leq 2 \left( \frac{3\ell(1-p)}{2} \right)^T W_1(\rho, \sigma) \\ &\leq \frac{3n}{2} \left( \frac{3\ell(1-p)}{2} \right)^T \|\rho - \sigma\|_1 \\ &\leq 3n \left( \frac{3\ell(1-p)}{2} \right)^T. \end{split}$$

Therefore, if  $p > 1 - \frac{2}{3\ell}$ , the above bound implies (3).

### B. A new bosonic Wasserstein distance

Inspired by the argument laid out in Section III A, we aim at introducing a distance on n-mode quantum states of a bosonic system  $\Lambda$  that satisfies similar bounds as (11), (12) and (13), namely: (i) a lower bound in terms of the trace distance as well as a uniform upper bound for a large enough class of physically relevant states; (ii) an input-output control for states undergoing a layer of ℓ-local GKP-stabilizer error correction; and (iii) a strict contraction under the action of the noise channel  $\mathcal{N}_{\lambda}^{\otimes n}$ . A natural candidate in the classical, continuous-variable framework is the distance dual to the Lipschitz constant  $\|\nabla f\| := \max_{j \in [n]} \|\partial_{x_j} f\|_{\infty}$  on the class of bounded, differentiable functions  $f: \mathbb{R}^n \to \mathbb{R}$ . By a standard replacement of partial derivatives  $\partial_{x_i}$  by commutators  $[P_i,(\cdot)],[Q_i,(\cdot)]$  with the position and momentum operators  $P_j, Q_j, j \in [n]$ , we arrive at the following definition.

**Definition 2.** The bosonic Lipschitz constant of an observable X on  $\Lambda$  is defined as

$$\|\nabla X\| := \sup |\langle \psi | [R_j, X] \otimes I_R | \varphi \rangle| \equiv \max_j \|\nabla_j X\|,$$

where the supremum is over all reference systems R, all quadrature operators  $R_j \in \{P_j, Q_j\}$ ,  $j \in \Lambda$ , and any two pure states  $|\psi\rangle, |\varphi\rangle$  of the joint system  $\Lambda R$  with finite total photon number on  $\Lambda$ . When this is finite, it coincides with  $\max_{R_j} ||[R_j, X]||_{\infty}$ . We then define the bosonic Wasserstein distance between two n-mode quantum states  $\rho, \sigma$  as

$$W_{\mathrm{B}}(\rho, \sigma) := \sup_{\|\nabla X\| \le 1} \left| \operatorname{tr} \left[ X(\rho - \sigma) \right] \right|. \tag{14}$$

where the supremum is over all bounded observables X.

In the supplementary material, we show that the distance  $W_{\rm B}$  satisfies bounds analogous to (11), (12) and (13) in the qubit setting. To show a bosonic version of (11), it would be enough to have that, for any observable X,  $\|\nabla X\|_{\infty} \leq C\|X\|_{\infty}$  for some constant C > 0. Indeed, this would imply that  $\|\rho - \sigma\|_1 \leq C W_{\rm B}(\rho_1, \rho_2)$ .

In analogy with the classical, continuous variable setting, this condition should, however, not hold since one can easily construct bounded, non-differentiable functions. However, when convoluted with a smooth density g, any function f acquires the regularity of g, so that  $\|\nabla(f*g)\|_{\infty} \leq C\|f\|_{\infty}$ . Since the noise channel  $\mathcal{N}_{\lambda}^{\otimes n}$  can be thought of as a quantum convolution [35], one can expect that a similar smoothing occurs in the bosonic framework. This intuition can be rigorously formalized and leads to the following bound:

**Proposition 3** (see Proposition 9). For any two n-mode states  $\rho$ ,  $\sigma$  and  $\lambda \in (0,1)$ , we have

$$\|\mathcal{N}_{\lambda}^{\otimes n}(\rho_1 - \rho_2)\|_1 \le \kappa W_{\mathcal{B}}(\rho_1, \rho_2), \tag{15}$$

where 
$$\kappa \coloneqq \sqrt{\frac{\lambda}{1-\lambda}} \, \max \left\{ \left\| [Q, \sigma_E] \right\|_1, \left\| [P, \sigma_E] \right\|_1 \right\}.$$

In turn, the bosonic Wasserstein distance can be controlled in terms of the average photon number of the states  $\rho, \sigma$ :

**Proposition 4** (see Proposition 12). For any two n-mode states  $\rho$ ,  $\sigma$  of average total photon number bounded by  $E \geq 0$ ,

$$W_{\rm B}(\rho,\sigma) \le 4\sqrt{2\,n\,(n+E)}\,.\tag{16}$$

The proof of the bound (16) relies on a semigroup interpolation result detailed in Appendix B 2. Next, we can use the canonical commutation relations in order to find the following bosonic analogue of (13):

**Proposition 5** (see Proposition 17). For any  $\lambda \in [0, 1]$  and any two n-mode state  $\rho, \sigma$ 

$$W_{\rm B}(\mathcal{N}_{\lambda}^{\otimes n}(\rho), \mathcal{N}_{\lambda}^{\otimes n}(\sigma)) \le \sqrt{\lambda} W_{\rm B}(\rho, \sigma).$$
 (17)

Finally, we would like to find bosonic analogues of (12). First, we show that the 1-mode GKP error correction scheme satisfies an input-output control bound akin to (12). We refer the interested reader to Appendix A 3 for a short review on GKP error correction.

**Proposition 6** (see Proposition 18). The 1-mode GKP quantum error correction scheme modeled by the channel  $\mathcal{R}_{\mathsf{GKP}}$  introduced in Appendix A 3 satisfies

$$W_{\rm B}(\mathcal{R}_{\mathsf{GKP}}^{\otimes n}(\rho), \mathcal{R}_{\mathsf{GKP}}^{\otimes n}(\sigma)) \le 2W_{\rm B}(\rho, \sigma)$$
 (18)

for all n-mode states  $\rho, \sigma$ .

At this stage, combining (18) together with (17), one can show that the Wasserstein distance decreases exponentially with the circuit depth whenever  $\lambda \leq 1/4$ . This is perhaps not so surprising since the quantum capacity of a quantum loss channel is known to reach 0 at  $\lambda = 1/2$  [36]. However, it is interesting to note that

the convergence in Wasserstein or trace distance that we establish here implies the stronger convergence of any (energy-constrained) capacity of the noisy circuit  $\mathcal{C}$ , and in particular of its energy constrained classical capacity [37]. We also mention that an upper bound on the channel infidelity between the noise composed with the GKP recovery channel and the ideal identity channel as a function of  $\lambda$  was found in [23, Equation (7.24)].

Finally, we consider recovery maps obtained from code Specifically, we consider an [n, k, d]stabilizer code  $\mathcal C$  - where n denotes the number of physical qubits, k the number of logical qubits and d the code distance- with stabilizer generators  $\{S_j\}_{j=1}^{n-k}$  and an associated recovery map  $\mathcal{R}$  of the form given in (2). We are interested in local recovery operations, and formalize this as follows: We assume that there is a Clifford circuit U(composed of one- and two-qubits) on the system and additional r = n - k auxiliary qubits  $A_1 \dots A_{n-k}$  such that syndrome information can be extracted by applying U and subsequently measuring each individual qubit. We capture locality as follows: first, we assume that the backward lightcone of every ancilla qubit  $A_i$  under the action of U has size at most  $\ell_{\rm meas}$ . Furthermore, we assume that the correction operation can be computed by applying local functions to s: There is a partition  $[n] = \bigcup_{j=1}^r \mathcal{F}_j$  of [n] into disjoint subsets  $\mathcal{F}_1, \dots, \mathcal{F}_r$  and functions  $C_i: \{0,1\}^{\mathcal{F}_j} \to \{I,X,Y,Z\}^{\otimes |\mathcal{F}_j'|}$  such that

$$C(s) = \prod_{j=1}^{r} C_j(s_{\mathcal{F}_j})_{\mathcal{F}'_j}$$
 for all  $s \in \{0, 1\}^{n-k}$ , (19)

where  $s_{\mathcal{F}_j}$  denotes the restriction of s to  $\mathcal{F}_j$ , i.e., the corresponding substring of syndrome bits. That is, for each  $j=1,\ldots,r$ , each operator  $C_j(s_{\mathcal{F}_j})$  has support contained in a set  $\mathcal{F}'_j$ . For what follows, we further assume that the sets  $\{\mathcal{F}'_j\}_j$  are pairwise disjoint. Then, we suppose that there are constants  $\ell_{\text{corr}}$ ,  $\ell'_{\text{corr}}$  such that  $|\mathcal{F}_j| \leq \ell_{\text{corr}}$ , resp.  $|\mathcal{F}'_j| \leq \ell'_{\text{corr}}$  for all  $j=1,\ldots,r$ .

**Proposition 7** (see Proposition 21). The recovery map  $\overline{\mathcal{R}}$  satisfies

$$W_{\rm B}(\overline{\mathcal{R}}(\rho), \overline{\mathcal{R}}(\sigma)) \le \Gamma W_{\rm B}(\rho, \sigma)$$
 (20)

for all n-mode states  $\rho, \sigma$ , where

$$\Gamma \coloneqq \ell_{\rm meas} + \frac{\ell_{\rm meas} \, 2^{1 + \ell_{\rm meas}(1 + \ell_{\rm corr})} \ell_{\rm meas} \cdot \ell_{\rm corr}'}{\sqrt{\alpha_{\rm min} \pi}}$$

where  $\alpha_{\min}$  denotes the minimal variance of the unsharp position and momentum measurements done in the scheme.

Finally, by combining (16), (20), (18), (17) and (15), we have proved that, for any two states  $\rho, \sigma$  of average

total photon number bounded by  $E \geq 0$ ,

$$\|\mathcal{R}_{\text{osc}}^{(T)}(\rho - \sigma)\|_{1} \le 4\kappa \left(2\Gamma \sqrt{\lambda}\right)^{T} \sqrt{2n(n+E)}$$

which is what we claimed in Theorem 1 for  $C = 2\Gamma$ .

## Acknowledgments

R.K. gratefully acknowledges support by the European Research Council under grant agreement no. 101001976 (project EQUIPTNT).

C.R. acknowledges the support of the Munich Center for Quantum Science and Technology, as well as that of the Humboldt Foundation.

- \* Cambyse Rouzé cambyse.rouze@tum.de
- D. Aharonov and M. Ben-Or, "Fault-tolerant quantum computation with constant error," in *Proceedings of the* twenty-ninth annual ACM symposium on Theory of computing - STOC '97, ACM Press, 1997.
- [2] P. Panteleev and G. Kalachev, "Quantum ldpc codes with almost linear minimum distance," *IEEE Transactions on Information Theory*, vol. 68, no. 1, pp. 213–229, 2021.
- [3] S. Gu, E. Tang, L. Caha, S. H. Choe, Z. He, and A. Ku-bica, "Single-shot decoding of good quantum ldpc codes," arXiv preprint arXiv:2306.12470, 2023.
- [4] M. Sipser and D. A. Spielman, "Expander codes," *IEEE transactions on Information Theory*, vol. 42, no. 6, pp. 1710–1722, 1996.
- [5] A. Leverrier, J.-P. Tillich, and G. Zémor, "Quantum expander codes," in 2015 IEEE 56th Annual Symposium on Foundations of Computer Science, pp. 810–824, IEEE, 2015
- [6] F. Pastawski, L. Clemente, and J. I. Cirac, "Quantum memories based on engineered dissipation," *Physical Re*view A, vol. 83, no. 1, p. 012304, 2011.
- [7] F. Pastawski, Quantum memory: design and applications. PhD thesis, Ludwig Maximilian Universität München München, 2012.
- [8] N. P. Breuckmann, K. Duivenvoorden, D. Michels, and B. M. Terhal, "Local decoders for the 2d and 4d toric code," arXiv preprint arXiv:1609.00510, 2016.
- [9] J. W. Harrington, Analysis of quantum error-correcting codes: symplectic lattice codes and toric codes. California Institute of Technology, 2004.
- [10] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, "Topological quantum memory," *Journal of Mathematical Physics*, vol. 43, pp. 4452–4505, Sept. 2002.
- [11] M. Herold, E. T. Campbell, J. Eisert, and M. J. Kastoryano, "Cellular-automaton decoders for topological quantum memories," npj Quantum information, vol. 1, no. 1, pp. 1–8, 2015.
- [12] J. Kempe, O. Regev, F. Uunger, and R. de Wolf, "Upper bounds on the noise threshold for fault-tolerant quantum computing," *Quantum Info. Comput.*, vol. 10, p. 361–376, May 2010.

- [13] A. A. Razborov, "An upper bound on the threshold quantum decoherence rate," Quantum Info. Comput., vol. 4, p. 222–228, may 2004.
- [14] I. L. Chuang, D. W. Leung, and Y. Yamamoto, "Bosonic quantum codes for amplitude damping," *Physical Review A*, vol. 56, no. 2, p. 1114, 1997.
- [15] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *Nature*, vol. 409, pp. 46–52, Jan. 2001.
- [16] M. H. Michael, M. Silveri, R. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. Girvin, "New class of quantum error-correcting codes for a bosonic mode," *Physical Review X*, vol. 6, July 2016.
- [17] P. T. Cochrane, G. J. Milburn, and W. J. Munro, "Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping," *Phys. Rev. A*, vol. 59, pp. 2631–2634, Apr 1999.
- [18] Z. Leghtas, G. Kirchmair, B. Vlastakis, R. J. Schoelkopf, M. H. Devoret, and M. Mirrahimi, "Hardware-efficient autonomous quantum memory protection," *Phys. Rev. Lett.*, vol. 111, p. 120501, Sep 2013.
- [19] S. Lloyd and J.-J. E. Slotine, "Analog quantum error correction," Phys. Rev. Lett., vol. 80, pp. 4088–4091, May 1998.
- [20] S. L. Braunstein, "Error correction for continuous quantum variables," in *Quantum Information with Continuous Variables*, pp. 19–29, Springer Netherlands, 1998.
- [21] D. Gottesman, A. Kitaev, and J. Preskill, "Encoding a qubit in an oscillator," *Phys. Rev. A*, vol. 64, p. 012310, Jun 2001.
- [22] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, "New class of quantum error-correcting codes for a bosonic mode," *Phys. Rev. X*, vol. 6, p. 031006, Jul 2016.
- [23] 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, vol. 97, p. 032346, Mar 2018.
- [24] S. B. Bravyi and A. Y. Kitaev, "Quantum codes on a lattice with boundary," arXiv preprint quant-ph/9811052, 1998.
- [25] K. Fukui, A. Tomita, and A. Okamoto, "Analog quantum error correction with encoding a qubit into an oscillator," *Phys. Rev. Lett.*, vol. 119, p. 180507, Nov 2017.
- [26] Y. Wang, "Quantum error correction with the gkp code and concatenation with stabilizer codes," arXiv preprint arXiv:1908.00147, 2019.
- [27] K. Fukui, A. Tomita, A. Okamoto, and K. Fujii, "High-threshold fault-tolerant quantum computation with analog quantum error correction," *Phys. Rev. X*, vol. 8, p. 021054, May 2018.
- [28] C. Vuillot, H. Asasi, Y. Wang, L. P. Pryadko, and B. M. Terhal, "Quantum error correction with the toric gottesman-kitaev-preskill code," *Phys. Rev. A*, vol. 99, p. 032344, Mar 2019.
- [29] N. C. Menicucci, "Fault-tolerant measurement-based quantum computing with continuous-variable cluster states," *Phys. Rev. Lett.*, vol. 112, p. 120504, Mar 2014.
- [30] K. Fukui, "High-threshold fault-tolerant quantum computation with the gkp qubit and realistically noisy devices," arXiv preprint arXiv:1906.09767, 2019.

- [31] K. Noh and C. Chamberland, "Fault-tolerant bosonic quantum error correction with the surface–gottesmankitaev-preskill code," *Phys. Rev. A*, vol. 101, p. 012316, Jan 2020.
- [32] K. Noh and C. Chamberland, "Fault-tolerant bosonic quantum error correction with the surface–gottesman-kitaev-preskill code," *Physical Review A*, vol. 101, no. 1, p. 012316, 2020.
- [33] C. Vuillot, H. Asasi, Y. Wang, L. P. Pryadko, and B. M. Terhal, "Quantum error correction with the toric gottesman-kitaev-preskill code," *Physical Review A*, vol. 99, no. 3, p. 032344, 2019.
- [34] G. D. Palma, M. Marvian, D. Trevisan, and S. Lloyd, "The quantum wasserstein distance of order 1," *IEEE Transactions on Information Theory*, vol. 67, pp. 6627–6643, Oct. 2021.
- [35] R. König and G. Smith, "The entropy power inequality for quantum systems," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1536–1548, 2014.
- [36] M. M. Wolf, D. Pérez-García, and G. Giedke, "Quantum capacities of bosonic channels," *Phys. Rev. Lett.*, vol. 98, p. 130501, Mar 2007.
- [37] M. E. Shirokov, "On the energy-constrained diamond norm and its application in quantum information theory," *Problems of Information Transmission*, vol. 54, pp. 20– 33, Jan. 2018.
- [38] A. Holevo, Quantum Systems, Channels, Information: A Mathematical Introduction. De Gruyter Studies in Mathematical Physics, De Gruyter, 2013.
- [39] A. L. Grimsmo and S. Puri, "Quantum error correction with the gottesman-kitaev-preskill code," PRX Quan-

- tum, vol. 2, p. 020101, Jun 2021.
- [40] A. M. Steane, "Active stabilization, quantum computation, and quantum state synthesis," *Phys. Rev. Lett.*, vol. 78, pp. 2252–2255, Mar 1997.
- [41] E. Knill, "Quantum computing with realistically noisy devices," *Nature*, vol. 434, pp. 39–44, Mar. 2005.
- [42] A. S. Holevo, "Structure of a general quantum Gaussian observable," Proceedings of the Steklov Institute of Mathematics, vol. 313, no. 1, pp. 70–77, 2021.
- [43] L. Gao and C. Rouzé, "Ricci curvature of quantum channels on non-commutative transportation metric spaces," 2021
- [44] L. G. Brown and H. Kosaki, "Jensen's inequality in semifinite von Neumann algebras," *Journal of Operator The*ory, pp. 3–19, 1990.
- [45] M. M. Wolf, "Quantum channels & operations: Guided tour." http://www-m5.ma.tum.de/foswiki/pub/M5/ Allgemeines/MichaelWolf/QChannelLecture.pdf, 2012. Lecture notes based on a course given at the Niels-Bohr Institute.
- [46] E. A. Carlen and J. Maas, "Gradient flow and entropy inequalities for quantum markov semigroups with detailed balance," *Journal of Functional Analysis*, vol. 273, no. 5, pp. 1810–1869, 2017.
- [47] M. E. Shirokov, "On extension of quantum channels and operations to the space of relatively bounded operators," *Lobachevskii Journal of Mathematics*, vol. 41, pp. 714– 727, Apr. 2020.

### Supplementary material

## A. Notations and definitions

### 1. Operators and norms

Given a separable Hilbert space  $\mathcal{H}$ , we denote by  $\mathcal{B}(\mathcal{H})$  the space of bounded linear operators on  $\mathcal{H}$ , and by  $\mathcal{T}_p(\mathcal{H})$ , the *Schatten p-class*, which is the Banach subspace of  $\mathcal{B}(\mathcal{H})$  formed by all bounded linear operators whose Schatten p-norm, defined as  $||X||_p = (\operatorname{tr}|X|^p)^{1/p}$ , is finite. Henceforth, we refer to  $\mathcal{T}_1(\mathcal{H})$  as the set of trace class operators. The set of quantum states (or density matrices), that is positive semi-definite operators  $\rho \in \mathcal{T}_1(\mathcal{H})$  of unit trace, is denoted by  $\mathcal{D}(\mathcal{H})$ . The Schatten 1-norm,  $||\cdot||_1$ , is the trace norm, and the corresponding induced distance (e.g. between quantum states) is the trace distance. Note that the Schatten 2-norm,  $||\cdot||_2$ , coincides with the *Hilbert–Schmidt norm*. For a pair of positive semi-definite operators, X, Y with domains  $\operatorname{dom}(X), \operatorname{dom}(Y) \subseteq \mathcal{H}$ , X > Y if and only if

For a pair of positive semi-definite operators, X, Y with domains  $\operatorname{dom}(X), \operatorname{dom}(Y) \subseteq \mathcal{H}, X \geq Y$  if and only if  $\operatorname{dom}(X^{1/2}) \subseteq \operatorname{dom}(Y^{1/2})$  and  $\|X^{1/2}|\psi\rangle\|^2 \geq \|Y^{1/2}|\psi\rangle\|^2$  for all  $|\psi\rangle \in \operatorname{dom}(X^{1/2})$ . If  $\rho$  is a quantum state with spectral decomposition  $\rho = \sum_i p_i |\varphi_i\rangle \langle \varphi_i|$ , and X is a positive semi-definite operator, the *expected value* of X on  $\rho$  is defined as

$$\operatorname{tr}[\rho X] := \sum_{i: p_i > 0} p_i \left\| X^{1/2} | \varphi_i \rangle \right\|^2 \in \mathbb{R}_+ \cup \{+\infty\};$$
(A1)

here we use the convention that  $\operatorname{tr}[\rho X] = +\infty$  if the above series diverges or if there exists an index i for which  $p_i > 0$  and  $|\varphi_i\rangle \notin \operatorname{dom}\left(X^{1/2}\right)$ . This definition can be extended to a generic densely defined self-adjoint operator X on  $\mathcal{H}$ , by considering its decomposition  $X = X_+ - X_-$  into positive and negative parts, with  $X_\pm$  being positive semi-definite operators with mutually orthogonal supports. The operator X is said to have a finite expected value on  $\rho$  if  $(i) |\varphi_i\rangle \in \operatorname{dom}\left(X_+^{1/2}\right) \cap \operatorname{dom}\left(X_-^{1/2}\right)$  for all i for which  $p_i > 0$ , and (ii) the two series  $\sum_i p_i \|X_\pm^{1/2}|\varphi_i\rangle\|^2$  both converge. In this case, the following quantity is called the expected value of X on  $\rho$ :

$$\operatorname{tr}[\rho X] := \sum_{i: p_i > 0} p_i \left\| X_+^{1/2} \left| \varphi_i \right\rangle \right\|^2 + \sum_{i: p_i < 0} p_i \left\| X_-^{1/2} \left| \varphi_i \right\rangle \right\|^2 \tag{A2}$$

Obviously, for a pair of operators X, Y satisfying  $X \ge Y$ , we have that  $\operatorname{tr}[\rho X] \ge \operatorname{tr}[\rho Y]$ .

Here we adopt standard notations from quantum information theory: given a multipartite quantum system with associated Hilbert space  $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ , and a state  $\rho \in \mathcal{D}(\mathcal{H})$ , we denote by  $\rho_A := \operatorname{tr}_{\mathcal{H}_B}(\rho)$  the marginal state on system A. Similarly, an observable, i.e. a self-adjoint operator X on  $\mathcal{H}_A$  is identified with the observable  $X_A \otimes I_B$  on the system AB, that is over the joint Hilbert space  $\mathcal{H}_A \otimes \mathcal{H}_B$ , where  $I_B$  denotes the identity operator on  $\mathcal{H}_B$ . A quantum channel with input system A and output system B is any completely positive, trace-preserving (CPTP) linear map  $\mathcal{N}: \mathcal{T}_1(\mathcal{H}_A) \to \mathcal{T}_1(\mathcal{H}_B)$ , where  $\mathcal{H}_A, \mathcal{H}_B$  are the Hilbert spaces corresponding to A, B, respectively. We denote identity superoperator over a system A by  $\operatorname{id}_A$ .

## 2. Phase-space formalism

In this paper, given  $m \in \mathbb{N}$ , we are concerned with the Hilbert space  $\mathcal{H}_m := L^2(\mathbb{R}^m)$  of a so-called m-mode oscillator, which is the space of square-integrable functions on  $\mathbb{R}^m$ . We often denote by  $\Lambda = \{1, \ldots, m\}$  the (arbitrarily ordered) set of m-modes. We denote by  $Q_j$  and  $P_j$  the canonical position and momentum operators on the  $j^{\text{th}}$  mode. The  $j^{\text{th}}$  creation and annihilation operators  $a_j = (Q_j + iP_j)/\sqrt{2}$  and  $a_j^{\dagger} = (Q_j - iP_j)/\sqrt{2}$  satisfy the well-known canonical commutation relations (CCR):

$$[a_j, a_k] = 0, [a_j, a_k^{\dagger}] = \delta_{jk} I, (A3)$$

where I denotes the identity operator on  $\mathcal{H}_m$ . The Hilbert space  $\mathcal{H}_m$  can be understood as a projective representation of the symplectic group [38]. We denote by  $\mathbb{M}_{2m}(\mathbb{R})$  the set of  $2m \times 2m$  real matrices, and by  $\operatorname{Sp}_{2m}(\mathbb{R})$ , the set of symplectic matrices in  $\mathbb{M}_{2m}(\mathbb{R})$ , i.e. matrices  $S \in \mathbb{M}_{2m}(\mathbb{R})$  satisfying the condition  $S\Omega_m S^T = \Omega_m$ , where  $\Omega_m$  denotes the  $2m \times 2m$  commutation matrix:

$$\Omega_m := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{\oplus m} . \tag{A4}$$

Any symplectic matrix S has determinant equal to one and is invertible with  $S^{-1} \in \operatorname{Sp}_{2m}(\mathbb{R})$ . Hence,  $\operatorname{Sp}_{2m}(\mathbb{R})$  is a subgroup of the special linear group  $\operatorname{SL}_{2m}(\mathbb{R})$ . In terms of the vector  $R := (Q_1, P_1, \ldots, Q_m, P_m)$ , the above relations take the compact form  $[R_j, R_k] = -i(\Omega_m)_{jk}$ , where  $\Omega_m$  denotes the  $2m \times 2m$  standard symplectic form defined in (A4). We will often omit the subscript m if the number of modes is fixed. The total photon number is defined by

$$N := \sum_{j=1}^{m} a_j^{\dagger} a_j = \sum_{j=1}^{m} \frac{Q_j^2 + P_j^2}{2} - \frac{m}{2}.$$
 (A5)

The operator N is diagonal in the multi-mode Fock basis  $\{|\mathbf{k}\rangle\}_{k\in\mathbb{N}^m}$ , with

$$N | \mathbf{k} \rangle = \left( \sum_{i=1}^{m} k_i \right) | \mathbf{k} \rangle, \qquad \mathbf{k} \equiv (k_1, \dots, k_m).$$
 (A6)

We define the displacement operator  $\mathfrak{D}(z)$  associated with a complex vector  $z \in \mathbb{C}^m$  as

$$\mathfrak{D}(z) = \exp\left[\sum_{j} (z_j a_j^{\dagger} - \overline{z}_j a_j)\right]. \tag{A7}$$

Thus,  $\mathfrak{D}(z)$  is a unitary operator and satisfies  $\mathfrak{D}(z)^{\dagger} = \mathfrak{D}(-z)$  and

$$\mathfrak{D}(z)\mathfrak{D}(w) = \mathfrak{D}(z+w) e^{\frac{1}{2}(z^{\mathsf{T}}\overline{w}-z^{\dagger}w)},\tag{A8}$$

valid for all  $z, w \in \mathbb{C}^m$ . We will also consider the real representation of the displacement operators: given  $x = \in \mathbb{R}^{2m}$ ,

$$V_x := e^{iR^{\mathsf{T}}x} \equiv \mathfrak{D}\left(\frac{-x_2 + ix_1}{\sqrt{2}}, \dots, \frac{-x_{2m} + ix_{2m-1}}{\sqrt{2}}\right).$$

A quantum state on  $\mathcal{H}_m$  is fully determined by its *characteristic function*  $\chi_{\rho}: \mathbb{C}^m \to \mathbb{C}$ , given by

$$\chi_{\rho}(z) := \operatorname{tr}[\rho \,\mathfrak{D}(z)] \,. \tag{A9}$$

In what follows, we are going to use coherent states: given  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{C}^m$ ,

$$|\alpha\rangle \equiv |\alpha_1, \dots, \alpha_m\rangle := \mathfrak{D}(\alpha)|0\rangle.$$
 (A10)

The inner product between two single mode coherent states satisfies

$$\langle \beta | \alpha \rangle = e^{-\frac{|\alpha|^2 + |\beta|^2}{2} + \alpha \overline{\beta}}.$$

In this article, we mainly consider a class of noises which includes the realistic description of the loss of a photon in each mode independently of the other, namely a tensor product of one-mode beamsplitters of transmissivity  $\lambda \in (0,1)$  with arbitrary environment state  $\sigma_E$ :

$$\mathcal{N}_{\lambda}(\rho) = \operatorname{tr}_{E} \left[ U_{\lambda}(\rho \otimes \sigma_{E}) U_{\lambda}^{\dagger} \right], \tag{A11}$$

where the beamsplitter unitary  $U_{\lambda} \equiv \exp\left((a^{\dagger}b - b^{\dagger}a)\arccos(\sqrt{\lambda})\right)$  satisfies the relations

$$U_{\lambda}^{\dagger} a U_{\lambda} = \sqrt{\lambda} a + \sqrt{1 - \lambda} b;$$

$$U_{\lambda}^{\dagger} b U_{\lambda} = -\sqrt{1 - \lambda} a + \sqrt{\lambda} b.$$
(A12)

where b stands for the annihilation operator associated to the environment mode. Note that we do not assume the environment state to be Gaussian. Although this is not required for our results to hold, we will fix the state  $\sigma_E$  to be the same each time the noise  $\mathcal{N}_{\lambda}$  is applied, for sake of simplicity. For any subset  $A \subseteq \Lambda$ , we denote by  $\mathcal{N}_A := \mathcal{N}_{\lambda}^{\otimes |A|}$  the |A|-fold product of the channel  $\mathcal{N}_{\lambda}$  defined in Equation (A11) acting on A.

#### 3. Gottesman-Kitaev-Preskill codes

The Weyl commutation relations (A8) imply the commutation of  $\mathfrak{D}(\alpha)$  and  $\mathfrak{D}(\beta)$  for any two complex numbers  $\alpha, \beta$  up to a phase. In particular, if

$$\beta \overline{\alpha} - \overline{\beta} \alpha = i\pi \,, \tag{A13}$$

the two operators anticommute, while if  $\beta \overline{\alpha} - \alpha \overline{\beta} = 2i\pi$ , they commute. This leads to the choice of logical Pauli operators  $\overline{X} = \mathfrak{D}(\alpha)$  and  $\overline{Z} = \mathfrak{D}(\beta)$  where  $\alpha$  and  $\beta$  are any two complex numbers that satisfy (A13), which ensures that  $\overline{X} \, \overline{Z} = -\overline{Z} \, \overline{X}$ . We also define the logical Pauli operator  $\overline{Y} := \mathfrak{D}(\alpha + \beta) = i \overline{X} \, \overline{Z}$ . In order for the operators  $\overline{X}$ ,  $\overline{Y}$  and  $\overline{Z}$  to behave like true Pauli operators, we then define the GKP logical codespace to be the simultaneous +1 eigenspace of the two operators

$$S_X := \overline{X}^2 = \mathfrak{D}(2\alpha), \qquad S_Z := \overline{Z}^2 = \mathfrak{D}(2\beta).$$

This can be done since these operators commute. Therefore, the set  $\{S_X^k, S_Z^l\}$  for  $k, l \in \mathbb{Z}$  can be interpreted as the stabilizer group of the GKP code. The latter can be defined by choosing  $\pm 1$  eigenstates of  $\overline{Z}$ . Unfortunately, the codewords are non-normalizable, which means that there is no physical process that can prepare a state lying exactly in the GKP codespace. In practice, one is required to resort to approximations of the latter. There exist various ways of defining approximate GKP states, and we refer to [39] for a comprehensive review. For our purposes, we will not need an explicit expression for the codestates, and we simply refer to them as  $|\widetilde{0}\rangle$  and  $|\widetilde{1}\rangle$ . The only characteristic of the state that we will need is that they have finite average energy with respect to the photon number operator  $N := a^{\dagger}a$ ,

$$\langle \widetilde{0}|N|\widetilde{0}\rangle \leq E_0, \quad \langle \widetilde{1}|N|\widetilde{1}\rangle \leq E_0.$$

Next, we discuss the ideal procedure for quantum error correction with the GKP code. To this end, we first define two generalized quadratures  $\hat{Q} := i(\overline{\beta}a - \beta a^{\dagger})/\sqrt{\pi}$  and  $\hat{P} := -i(\overline{\alpha}a - \alpha a^{\dagger})/\sqrt{\pi}$ , such that  $[\hat{Q}, \hat{P}] = iI$ . Here and for sake of simplicity, we stick to the so-called square code for which  $\alpha = \sqrt{\pi/2}$  and  $\beta = i\sqrt{\pi/2}$ , and will often identify  $Q = \hat{Q}$  and  $P = \hat{P}$ . Next, there are two canonical ways to perform GKP error correction using GKP-encoded ancillae. These are continuous variables versions of Steane [40] and Knill [41]. Here, we focus on Steane's method, see Figure 1. There, the entanglement gate is a 2-mode CNOT gate, whose dual action on quadrature operators takes the form

$$(\mathcal{U}_{AA'}^{\mathrm{CNOT}})^{\dagger} : Q_A \to Q_A, P_A \to P_A - P_{A'},$$
  
 $Q_{A'} \to Q_A + Q_{A'}, P_{A'} \to P_{A'},$  (A14)

where the control register is A.

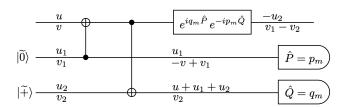


FIG. 1. Illustration of Steane's error correction circuit. The labels  $\{u,v\}$  next to a rail indicate a general displacement error  $e^{-iu\hat{P}}e^{iv\hat{Q}}$ , and the diagram indicates how the incoming errors propagate through the circuit. The two measurements are of the  $\hat{P}$  and  $\hat{Q}$  quadratures of the GKP code, respectively. For the correction shifts, we use the measurement outcomes modulo lattice spacing  $\sqrt{\pi}$ . Here, the approximate GKP state  $|\widetilde{+}\rangle$  corresponds to the logical  $|+\rangle$  state.

In this paper, we also consider the concatenation of the GKP code with qubit stabilizer codes such as the toric code. For this, we need destructive logical measurements in any Pauli basis:

$$\overline{X} = e^{-i\sqrt{\pi}\hat{P}}, \qquad \overline{Z} = e^{i\sqrt{\pi}\hat{Q}}, \qquad \overline{Y} = e^{i\sqrt{\pi}(\hat{Q} - \hat{P})}. \tag{A15}$$

Then, Pauli measurements are performed by measuring one of three respective quadratures

 $\mathcal{M}_X$ : measure  $-\hat{P}$   $\mathcal{M}_Y$ : measure  $\hat{Q} - \hat{P}$   $\mathcal{M}_Z$ : measure  $\hat{Q}$ ,

and rounding to the nearest multiple of  $\sqrt{\pi}$ . If the result is an even multiple, report a +1 outcome, and otherwise report a -1 outcome. The error correction schemes described so far include homodyne-type measurements of the quadratures  $\hat{Q}$  and  $\hat{P}$ . However, the latter correspond to the infinite-squeezing limit of more realistic Gaussian POVM measurements as introduced in [42]. For simplicity, we consider the unsharp (noisy) position and momentum measurements with variances  $\alpha_q$ ,  $\alpha_p > 0$  and  $q, p \in \mathbb{R}$ 

$$m_{\hat{Q}}(q) = \frac{1}{2\pi\sqrt{\alpha_q}} V_{(0,-q)^t} e^{-\frac{1}{2\alpha_q} \hat{Q}^2} V_{(0,-q)^t}^{\dagger} = \frac{1}{2\pi\sqrt{\alpha_q}} e^{-\frac{1}{2\alpha_q} (\hat{Q} - q)^2}, \tag{A16}$$

$$m_{\hat{P}}(p) = \frac{1}{2\pi\sqrt{\alpha_p}} V_{(p,0)^t} e^{-\frac{1}{2\alpha_p} \hat{P}^2} V_{(p,0)^t}^{\dagger} = \frac{1}{2\pi\sqrt{\alpha_p}} e^{-\frac{1}{2\alpha_p} (\hat{P} - p)^2}, \tag{A17}$$

and denote  $\alpha_{\min} := \min\{\alpha_p, \alpha_q\}$ . The above operator valued densities are associated to POVMs  $A \mapsto M(A) = \int_A m(x) dx$  for  $m \in \{m_{\hat{P}}, m_{\hat{Q}}\}$ .

For our qubit stabilizer code, we consider a general situation where each of the m modes corresponds to a vertex of a given graph G = (V, E). Then, we examine a simple scheme of nondestructive measurement involving the introduction of one bosonic ancilla system per stabilizer, logical CNOT gates entangling modes associated with a stabilizer together with an ancilla, and logical Pauli measurements on the latter (see e.g. Figure 2 for the case of the 2D Toric code). This scheme was first considered in [10] for the qubit setting. Then, upon the measurement of all the stabilizers, the corresponding outcomes  $s = \{s_k\}_{k=1}^K$  form a syndrome which is then used to construct an error correction unitary in the form of a displacement operator D(f(s)) for some classical function f of the syndromes. We denote by  $\Phi_f^{\text{Stab}}$  the overall quantum error correction channel acting on the m modes.

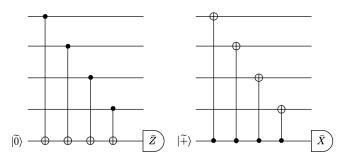


FIG. 2. Circuits for measurement of plaquette  $\bar{Z}^{\otimes 4}$  and star  $\bar{X}^{\otimes 4}$  stabilizers for the concatenated 2D Toric code.

## B. Bosonic Wasserstein distance

Our main technical contribution is the introduction and use of a non-commutative extension of the classical Lipschitz-type constant  $\|\nabla f\| := \max_{j \in [m]} \|\partial_j f\|_{\infty}$  of a real, continuously differentiable function f of m variables, and the dual notion of a Wasserstein-type distance between two distributions over  $\mathbb{R}^m$ .

**Definition 8.** The bosonic Lipschitz constant of an operator  $X \in \mathcal{B}(\mathcal{H}_m)$  is defined as

$$\|\nabla X\|^2 \coloneqq \max_{j \in \Lambda} \max_{R_j \in \{Q_j, P_j\}} \sup_{R} \sup_{|\psi\rangle, |\varphi\rangle} |\langle \psi|[R_j, X] \otimes I_R |\varphi\rangle|^2 \equiv \max_{j} \|\nabla_j X\|,$$

where the supremum is over all pure states  $|\psi\rangle, |\varphi\rangle \in \text{dom}(\sqrt{N} \otimes I_R)$  for an arbitrarily large reference system R. By duality, we then define the bosonic Wasserstein distance between two m-mode quantum states  $\rho, \sigma \in \mathcal{D}(L_2(\mathbb{R}^m))$  as

$$W_{\mathrm{B}}(\rho,\sigma) := \sup_{\|\nabla X\| \le 1} \left| \operatorname{tr} \left[ X(\rho - \sigma) \right] \right|. \tag{B1}$$

where the supremum is over all bounded, self-adjoint operators X.

In the next sections, we provide some basic properties of the bosonic Wasserstein distance  $W_{\rm B}$ .

### 1. Relation to the trace distance

First, we seek for an upper bound on the trace distance in terms of  $W_B$ . By duality of both metrics, this amounts to finding an upper bound on the Lipschitz constant  $\|\nabla X\|$  of any bounded operator X in terms of its operator norm  $\|X\|_{\infty}$ . However, a bound of that sort does not exist (as classically, one can easily think of bounded observables which are not smooth). In the classical setting, the problem can be handled by first smoothing the function f, e.g. by convolving it with a Gaussian density g. In that case, one proves that there exists a finite constant C > 0 such that  $\|\nabla (f * g)\| \le C\|f\|_{\infty}$ . Our first technical result consists in a bosonic analogue of this bound (see also [43, Proposition 6.4] for a single-mode variant).

**Proposition 9.** For any two states  $\rho_1, \rho_2 \in \mathcal{D}(L_2(\mathbb{R}^m))$  and  $\lambda \in (0,1)$ , we have

$$\|\mathcal{N}_{\lambda}^{\otimes m}(\rho_{1} - \rho_{2})\|_{1} \leq \sqrt{\frac{\lambda}{1 - \lambda}} \max \{\|[Q, \sigma_{E}]\|_{1}, \|[P, \sigma_{E}]\|_{1}\} W_{B}(\rho_{1}, \rho_{2}).$$
 (B2)

In order to prove Proposition 9, we need the following technical lemma.

**Lemma 10.** Given a two-mode bosonic system AE with associated annihilation operators a, b, any reference system C, any  $X \in \mathcal{B}(L^2(\mathbb{R}))$  and any  $|\varphi\rangle, |\psi\rangle \in \text{dom}(\sqrt{N} \otimes I_C)$ ,

$$\langle \varphi | [a, \mathcal{N}_{\lambda}^{\dagger}(X)] | \psi \rangle = \sqrt{\lambda} \, \langle \varphi | \mathcal{N}_{\lambda}^{\dagger}([a, X]) | \psi \rangle = -\sqrt{\frac{\lambda}{1 - \lambda}} \, \langle \varphi | \operatorname{tr}_{E} \left( (I \otimes [b, \sigma]) U_{1 - \lambda}(I \otimes X_{E}) U_{1 - \lambda}^{\dagger} \right) | \psi \rangle \tag{B3}$$

$$\langle \varphi | [a^{\dagger}, \mathcal{N}_{\lambda}^{\dagger}(X)] | \psi \rangle = \sqrt{\lambda} \langle \varphi | \mathcal{N}_{\lambda}^{\dagger}([a^{\dagger}, X]) | \psi \rangle = -\sqrt{\frac{\lambda}{1 - \lambda}} \langle \varphi | \operatorname{tr}_{E} \left( (I \otimes [b^{\dagger}, \sigma]) U_{1 - \lambda} (I \otimes X_{E}) U_{1 - \lambda}^{\dagger} \right) | \psi \rangle$$
 (B4)

where  $\mathcal{N}_{\lambda}^{\dagger}$  is assumed to act on the system A above, and we omitted the operator  $I_C$  for ease of notations.

*Proof.* We only prove the first line (B3), since (B4) follows analogously with a replaced by  $a^{\dagger}$  and b by  $b^{\dagger}$ . The adjoint map  $\mathcal{N}_{\lambda}^{\dagger}$  is given by

$$\mathcal{N}_{\lambda}^{\dagger}(X) = \operatorname{tr}_{E}\left((I \otimes \sigma)U_{\lambda}^{\dagger}(X \otimes I)U_{\lambda}\right) , \tag{B5}$$

and the first identity follows from this expression and Eq. (A12) by a straightforward computation (see [43, Proof of Proposition 6.2]). The proof of the second identity uses the identity  $U_{\lambda}^{\dagger} = U_{1-\lambda}e^{-\frac{\pi}{2}(a^{\dagger}b-b^{\dagger}a)} = U_{1-\lambda}\mathsf{SWAP}(I\otimes Z)$  where SWAP interchanges the two modes and Z is a Gaussian unitary acting as  $ZbZ^{\dagger} = -b$ ,  $Zb^{\dagger}Z^{\dagger} = -b^{\dagger}$  on the second system. This implies that  $U_{\lambda}^{\dagger}(X\otimes I)U_{\lambda} = U_{1-\lambda}(I\otimes X)U_{1-\lambda}^{\dagger}$  and thus

$$[a, \mathcal{N}_{\lambda}^{\dagger}(X)] = \left[ a, \operatorname{tr}_{E} \left( (I \otimes \sigma_{E}) U_{\lambda}^{\dagger}(X \otimes I_{E}) U_{\lambda} \right) \right]$$

$$= \left[ a, \operatorname{tr}_{E} \left( (I \otimes \sigma_{E}) U_{1-\lambda} (I \otimes X) U_{1-\lambda}^{\dagger} \right) \right]$$

$$= \operatorname{tr}_{E} \left( (I \otimes \sigma) \left[ a \otimes I, U_{1-\lambda} (I \otimes X) U_{1-\lambda}^{\dagger} \right] \right) , \tag{B6}$$

where these identities hold when evaluated on states  $|\varphi\rangle, |\psi\rangle \in \text{dom}(\sqrt{N} \otimes I_C)$ . Observe that

$$\left[a\otimes I, U_{1-\lambda}(I\otimes X)U_{1-\lambda}^{\dagger}\right] = \left[a\otimes I, U_{1-\lambda}\right](I\otimes X)U_{1-\lambda}^{\dagger} + U_{1-\lambda}(I\otimes X)\left[a\otimes I, U_{1-\lambda}^{\dagger}\right] + U_{1-\lambda}[a\otimes I, I\otimes X]U_{1-\lambda}^{\dagger}$$

where the third term on the right vanishes. Using the commutation relations (A12) then gives

$$\begin{bmatrix} a \otimes I, U_{1-\lambda}(I \otimes X)U_{1-\lambda}^{\dagger} \end{bmatrix} = (1 - \sqrt{1 - \lambda}) \left[ (a \otimes I), U_{1-\lambda}(I \otimes X)U_{1-\lambda}^{\dagger} \right] + \sqrt{\lambda} \left[ (I \otimes b), U_{1-\lambda}(I \otimes X)U_{1-\lambda}^{\dagger} \right] 
= (1 - \sqrt{1 - \lambda}) \left[ a \otimes I, U_{\lambda}^{\dagger}(X \otimes I)U_{\lambda} \right] + \sqrt{\lambda} \left[ (I \otimes b), U_{1-\lambda}(I \otimes X)U_{1-\lambda}^{\dagger} \right] .$$
(B7)

Inserting Eq. (B7) into Eq. (B6) gives

$$[a, \mathcal{N}_{\lambda}^{\dagger}(X)] = (1 - \sqrt{1 - \lambda})[a, \mathcal{N}_{\lambda}^{\dagger}(X)] + \sqrt{\lambda} \operatorname{tr}_{E} \left( (I \otimes \sigma)[I \otimes b, U_{1 - \lambda}(I \otimes X)U_{1 - \lambda}^{\dagger}] \right)$$
$$= (1 - \sqrt{1 - \lambda})[a, \mathcal{N}_{\lambda}^{\dagger}(X)] - \sqrt{\lambda} \operatorname{tr}_{E} \left( (I \otimes [b, \sigma])U_{1 - \lambda}(I \otimes X)U_{1 - \lambda}^{\dagger}] \right)$$

as claimed. Here we used the tracial property in [44, Theorem 17]. The second identity in Eq. (B3) follows from this.  $\Box$ 

Proof of Proposition 9. Eq. (B3) immediately extends to several modes, giving for every  $j \in \Lambda$  the identities

$$[a_{j}, \mathcal{N}_{\lambda}^{\dagger \otimes m}(X)] = \sqrt{\lambda} \mathcal{N}_{\lambda}^{\dagger \otimes m}([a_{j}, X]) [a_{j}^{\dagger}, \mathcal{N}_{\lambda}^{\dagger \otimes m}(X)] = \sqrt{\lambda} \mathcal{N}_{\lambda}^{\dagger \otimes m}([a_{j}^{\dagger}, X])$$
(B8)

because e.g.,  $[a_j, \mathcal{N}^{\otimes m}(X)] = \bigotimes_{k \neq j} \mathcal{N}_k ([a_j, \mathcal{N}_j(X)]) = \sqrt{\lambda} \left(\bigotimes_{k \neq j} \mathcal{N}_k\right) \circ \mathcal{N}_j ([a_j, X]) = \sqrt{\lambda} \mathcal{N}^{\otimes m} ([a_j, X])$  where  $\mathcal{N}_k$  denotes the channel  $\mathcal{N} = \mathcal{N}_{\lambda}$  acting on the j-th factor. We can similarly show that, for every  $j \in \Lambda$ ,

$$[a_{j}, \mathcal{N}_{\lambda}^{\dagger \otimes m}(X)] = -\sqrt{\frac{\lambda}{1-\lambda}} \, \mathcal{N}_{\Lambda \setminus \{j\}}^{\dagger} \left( \operatorname{tr}_{E_{j}} \left[ (I \otimes [b_{j}, \sigma_{E_{j}}]) U_{1-\lambda}(I \otimes X) U_{1-\lambda}^{\dagger} \right] \right) ,$$

$$[a_{j}^{\dagger}, \mathcal{N}_{\lambda}^{\dagger \otimes m}(X)] = -\sqrt{\frac{\lambda}{1-\lambda}} \, \mathcal{N}_{\Lambda \setminus \{j\}}^{\dagger} \left( \operatorname{tr}_{E_{j}} \left[ (I \otimes [b_{j}^{\dagger}, \sigma_{E_{j}}]) U_{1-\lambda}(I \otimes X) U_{1-\lambda}^{\dagger} \right] \right) .$$
(B9)

where  $b_j$  denotes the annihilation operator on the environment system  $E_j$ . Indeed, we have  $[a_j, \mathcal{N}_{\lambda}^{\dagger \otimes m}(X)] = \mathcal{N}_{\Lambda \setminus \{j\}}^{\dagger} ([a_j, \mathcal{N}_j(X)])$  and similarly for  $a_j^{\dagger}$ , hence Eq. (B9) is an immediate adaptation of the proof of (10). Hence, for any two pure states  $|\psi\rangle$ ,  $|\varphi\rangle \in \text{dom}(\sqrt{N} \otimes I_R)$ ,

$$\begin{aligned} |\langle \psi | [Q_j, \mathcal{N}_{\lambda}^{\dagger \otimes m}(X)] \otimes I_R | \varphi \rangle| &\leq \sqrt{\frac{\lambda}{1 - \lambda}} \sup_{\|y\|_1 \leq 1} \operatorname{tr} \left[ (y \otimes [Q_j', \sigma_{E_j}]) \, U_{1 - \lambda}(I \otimes X) U_{1 - \lambda}^{\dagger} \right] \\ &\leq \sqrt{\frac{\lambda}{1 - \lambda}} \left\| [Q_j', \sigma_{E_j}] \right\|_1 \|X\|_{\infty} \,, \end{aligned}$$

and similarly for  $|\langle \psi | [P_j, \mathcal{N}_{\lambda}^{\dagger \otimes m}(X)] \otimes I_R | \varphi \rangle|$ . The result follows by duality.

Proposition 9 can be used to prove that the bosonic Wasserstein distance is a metric on quantum states:

Corollary 11.  $W_B: \mathcal{D}(L_2(\mathbb{R}^m)) \times \mathcal{D}(L_2(\mathbb{R}^m)) \to \mathbb{R}_+$  is a metric on the set of quantum states over  $L_2(\mathbb{R}^m)$ .

Proof. The triangle inequality and positivity are clear from the definition. Moreover  $W_{\rm B}(\rho,\rho)=0$  for all  $\rho\in\mathcal{D}(L_2(\mathbb{R}^m))$ . Finally, Given two states  $\rho_1,\rho_2\in\mathcal{D}(L_2(\mathbb{R}^m))$ , Proposition 9 shows that  $W_{\rm B}(\rho_1,\rho_2)=0$  implies  $\mathcal{N}_{\lambda}^{\otimes m}(\rho_1)=\mathcal{N}_{\lambda}^{\otimes m}(\rho_2)$  for any  $\lambda\in(0,1)$ . Therefore, that  $\rho_1=\rho_2$  follows from the strong continuity of  $\lambda\mapsto\mathcal{N}_{\lambda}^{\otimes m}$  with respect to the trace distance.

## 2. Diameter bound under moment constraints

Next, we prove an upper bound on the Wasserstein distance between two quantum states under a constraint on their moments.

**Proposition 12.** For any two states  $\rho, \sigma \in \mathcal{D}(\mathcal{H}_m)$  such that  $\operatorname{tr}[\rho N], \operatorname{tr}[\sigma N] \leq E < \infty$ ,

$$W_{\rm B}(\rho,\sigma) \le 4\sqrt{2\,m\,(m+E)}$$
.

Before proving Proposition 12, we first state and prove a simple technical lemma which we will use multiple times in the section.

**Lemma 13.** Let  $\Phi: \mathcal{T}_1(\mathcal{H}_m) \to \mathcal{T}_1(\mathcal{H}_{m'})$  be a moment-limited quantum channel, i.e. for any  $\rho \in \mathcal{D}(\mathcal{H}_m)$  with  $\operatorname{tr}[\rho N_m] < \infty$ ,  $\operatorname{tr}[\Phi(\rho)N_{m'}] < \infty$ . Then, for any reference system R and any two pure states  $|\varphi\rangle, |\psi\rangle \in \operatorname{dom}(\sqrt{N_m} \otimes I_R)$ , the operator  $A := (\Phi \otimes id_R)(|\psi\rangle\langle\varphi|)$  is trace-class,  $||A||_1 \leq 1$ , and  $\operatorname{ran}(A), \operatorname{ran}(A^{\dagger}) \subseteq \operatorname{dom}(\sqrt{N_{m'}} \otimes I_R)$ . In other words, A admits the following representation

$$A = \sum_{\ell} \alpha_{\ell} |e_{\ell}\rangle\langle f_{\ell}|,$$

for some sequences  $\alpha \equiv \{\alpha_\ell\}$  of positive numbers with  $\|\alpha\|_{\ell_1} \leq 1$ , and of orthonormal vectors  $\{|e_\ell\rangle\}, \{|f_\ell\rangle\}$  in the domain of  $\sqrt{N_{m'}} \otimes I_R$ .

*Proof.* Let  $\{|e_i\rangle\}$  be an orthonormal basis in  $\mathcal{H}_R$  and let  $|\chi\rangle\in\mathcal{H}_{m'}\otimes\mathcal{H}_R$ . We have

$$\begin{split} \|N_{m'}^{s} \otimes I_{R}(\Phi \otimes \mathrm{id}_{R})(|\psi\rangle\langle\varphi|)|\chi\rangle\|^{2} &= \sum_{k \in \mathbb{N}^{m'}} \sum_{i} E_{k} \left| \langle \chi | (\Phi \otimes \mathrm{id}_{R})(|\varphi\rangle\langle\psi|)(|k\rangle \otimes |e_{i}\rangle\rangle) \right|^{2} \\ &= \sum_{k \in \mathbb{N}^{m'}} \sum_{i} E_{k} \left| \langle \psi | (\Phi^{\dagger} \otimes \mathrm{id}_{R})(|k\rangle \otimes |e_{i}\rangle\langle\chi|) |\varphi\rangle|^{2} \\ &\leq \sum_{k \in \mathbb{N}^{m'}} \sum_{i} E_{k} \left| \langle \psi | (\Phi^{\dagger} \otimes \mathrm{id}_{R})(|k\rangle \otimes |e_{i}\rangle\langle\chi|) (\Phi^{\dagger} \otimes \mathrm{id}_{R})(|\chi\rangle\langle k| \otimes \langle e_{i}|) |\psi\rangle \\ &\leq \sum_{k \in \mathbb{N}^{m'}} \sum_{i} E_{k} \left| \langle \psi | (\Phi^{\dagger} \otimes \mathrm{id}_{R})(|k\rangle\langle k| \otimes |e_{i}\rangle\langle e_{i}|) |\psi\rangle \\ &= \mathrm{tr}[(\Phi \otimes \mathrm{id}_{R})(|\psi\rangle\langle\psi|) (N_{m'} \otimes I_{R})] \end{split}$$

where  $N_{m'}|k\rangle = E_k|k\rangle$ , and where in the last inequality we have used the operator Schwarz inequality [45, Theorem 5.3]. Now, since  $\Phi$  is moment-limited and  $|\psi\rangle \in \text{dom}(\sqrt{N_m} \otimes I_R)$ , we have proved that  $||\sqrt{N_{m'}} \otimes I_R(\Phi \otimes id_R)(|\psi\rangle\langle\varphi|)|\chi\rangle|| < \infty$ , and therefore  $\text{ran}(A) \subseteq \text{dom}(\sqrt{N_{m'}} \otimes I_R)$ . The same holds for  $A^{\dagger}$  after exchanging the vectors  $|\psi\rangle$  and  $|\varphi\rangle$ . Finally, since  $\Phi$  is a channel, we have  $||\alpha||_{\ell_1} = ||(\Phi \otimes id_R)(|\psi\rangle\langle\varphi|)||_1 \le ||\psi\rangle\langle\varphi||_1 = 1$ .

Proof of Proposition 12. First, we denote the thermal Gaussian state at inverse temperature  $\beta > 0$  as

$$\sigma_{\beta} \coloneqq \frac{e^{-\beta N_{\Lambda}}}{\operatorname{tr}[e^{-\beta N_{\Lambda}}]}.$$

The state  $\sigma_{\beta}$  is the unique invariant state of the so-called Bose Ornstein-Uhlenbeck semigroup  $(e^{t\mathcal{L}_{\beta}})_{t\geq 0}$  of generator in the Heisenberg picture defined on a suitable subspace as [46]

$$\mathcal{L}_{\beta}^{\dagger}(X) \coloneqq \frac{1}{2} \sum_{j \in \Lambda} \left( e^{\frac{\beta}{2}} \left( a_j^{\dagger}[X, a_j] + [a_j^{\dagger}, X] a_j \right) + e^{-\frac{\beta}{2}} \left( a_j[X, a_j^{\dagger}] + [a_j, X] a_j^{\dagger} \right) \right).$$

Next, we show an upper bound on the Wasserstein distance between an arbitrary state  $\rho$  satisfying the constraints of the Proposition and  $\sigma_{\beta}$ , see Eq. (B22) below. To this end, we proceed as follows. Given a bounded, self-adjoint operator X on  $\mathcal{H}_m$  satisfying the constraints in the definition of  $W_B$ , we write  $X_t := e^{t\mathcal{L}_{\beta}^{\dagger}}(X)$  and  $\rho_t := e^{t\mathcal{L}_{\beta}}(\rho)$  such that

$$\operatorname{tr}[X(\sigma_{\beta} - \rho)] = \operatorname{tr}\left(X \lim_{t \to \infty} (e^{t\mathcal{L}_{\beta}}(\rho) - \rho)\right)$$

$$= \operatorname{tr}\left(X \int_{0}^{\infty} \mathcal{L}_{\beta}(\rho_{t}) dt\right)$$

$$= \frac{1}{2} \sum_{j \in \Lambda} \int_{0}^{\infty} e^{\frac{\beta}{2}} \left(\operatorname{tr}\left[[X_{t}, a_{j}]\rho a_{j}^{\dagger}\right] + \operatorname{tr}\left[a_{j}\rho[a_{j}^{\dagger}, X_{t}]\right]\right) + e^{-\frac{\beta}{2}} \left(\operatorname{tr}\left[[X_{t}, a_{j}^{\dagger}]\rho a_{j}\right] + \operatorname{tr}\left[a_{j}^{\dagger}\rho[a_{j}, X_{t}]\right]\right) dt.$$
(B10)

The maps  $e^{t\mathcal{L}_{\beta}}$  turn out to be equal to the channels  $\mathcal{N}_{\lambda}^{\otimes m}$  for  $\lambda \equiv e^{-2\sinh(\beta/2)t}$  and environment states  $\sigma_{E_j} = \sigma_{\beta}$ . Therefore, by the relations (B8) and the definition of  $X_t$  we have

$$\operatorname{tr}\left([X_t, a_j] \rho a_j^{\dagger}\right) = e^{-\sinh(\beta/2)t} \operatorname{tr}\left(e^{t\mathcal{L}_{\beta}^{\dagger}}([X, a_j]) \rho a_j^{\dagger}\right) = e^{-\sinh(\beta/2)t} \operatorname{tr}\left([X, a_j] e^{t\mathcal{L}_{\beta}}(\rho a_j^{\dagger})\right) \tag{B11}$$

and similarly

$$\operatorname{tr}\left(a_{j}\rho[a_{j}^{\dagger},X_{t}]\right) = e^{-\sinh(\beta/2)t}\operatorname{tr}\left(e^{t\mathcal{L}_{\beta}}(a_{j}\rho)[a_{j}^{\dagger},X]\right)$$

$$\operatorname{tr}\left([X_{t},a_{j}^{\dagger}]\rho a_{j}\right) = e^{-\sinh(\beta/2)t}\operatorname{tr}\left([X,a_{j}^{\dagger}]e^{t\mathcal{L}_{\beta}}(\rho a_{j})\right) .$$

$$\operatorname{tr}\left(a_{j}^{\dagger}\rho[a_{j},X_{t}]\right) = e^{-\sinh(\beta/2)t}\operatorname{tr}\left(e^{t\mathcal{L}_{\beta}}(a_{j}^{\dagger}\rho)[a_{j},X]\right)$$
(B12)

Inserting (B11) and (B12) into (B10) we obtain

$$\operatorname{tr}[X(\sigma_{\beta} - \rho)] = \frac{1}{2} \int_{0}^{\infty} e^{-\sinh(\beta/2)t} \sum_{j \in \Lambda} e^{\frac{\beta}{2}} \left( \operatorname{tr}\left( [X, a_{j}] e^{t\mathcal{L}_{\beta}}(\rho a_{j}^{\dagger}) \right) + \operatorname{tr}\left( e^{t\mathcal{L}_{\beta}}(a_{j}\rho)[a_{j}^{\dagger}, X] \right) \right) + e^{-\frac{\beta}{2}} \left( \operatorname{tr}\left( [X, a_{j}^{\dagger}] e^{t\mathcal{L}_{\beta}}(\rho a_{j}) \right) + \operatorname{tr}\left( e^{t\mathcal{L}_{\beta}}(a_{j}^{\dagger}\rho)[a_{j}, X] \right) \right) dt .$$
(B13)

Next, we control each of the traces above separately. For this, we consider the spectral decomposition  $\rho := \sum_{\ell} \lambda_{\ell} |\psi_{\ell}\rangle \langle \psi_{\ell}|$ . We first consider the operator  $e^{t\mathcal{L}_{\beta}}(|\psi_{\ell}\rangle \langle \varphi_{\ell}|)$  for some arbitrary eigenvector  $|\psi_{\ell}\rangle$  of  $\rho$  such that  $a_{j}|\psi_{\ell}\rangle \neq 0$  and  $|\varphi_{\ell}\rangle := a_{j}|\psi_{\ell}\rangle / ||a_{j}|\psi_{\ell}\rangle||$ . Then, by assumption on  $\rho$ ,  $|\psi_{\ell}\rangle$ ,  $|\varphi_{\ell}\rangle \in \text{dom}(\sqrt{N})$ . Moreover, since  $e^{t\mathcal{L}_{\beta}}$  is Gaussian for all  $t \geq 0$ , we have that for any state  $\rho'$  such that  $\text{tr}[\rho'N] < \infty$ ,  $\text{tr}[e^{t\mathcal{L}_{\beta}}(\rho')N] < \infty$ . By Lemma 13, the operator  $e^{t\mathcal{L}_{\beta}}(|\psi_{\ell}\rangle \langle \varphi_{\ell}|)$  admits a decomposition onto rank-one operators of the form

$$e^{t\mathcal{L}_{\beta}}(|\psi_{\ell}\rangle\langle\varphi_{\ell}|) = \sum_{i} \alpha_{i}^{(\ell)} |e_{i}^{(\ell)}\rangle\langle f_{i}^{(\ell)}|, \quad \text{with} \quad \|\alpha^{(\ell)}\|_{\ell_{1}} \leq 1, \quad \{|e_{i}^{(\ell)}\rangle, |f_{i}^{(\ell)}\rangle\} \in \text{dom}(\sqrt{N}).$$

Moreover, we have

$$\operatorname{tr}\left(e^{t\mathcal{L}_{\beta}^{\dagger}}([X,a_{j}])\rho a_{j}^{\dagger}\right) = \operatorname{tr}\left([X,a_{j}]e^{t\mathcal{L}_{\beta}}(\rho a_{j}^{\dagger})\right)$$
(B14)

$$= \sum_{\ell} \lambda_{\ell} \|a_j |\psi_{\ell}\rangle \| \cdot \operatorname{tr} \left( [X, a_j] e^{t\mathcal{L}_{\beta}} (|\psi_{\ell}\rangle \langle \varphi_{\ell}|) \right)$$
 (B15)

$$= \sum_{\ell,i} \lambda_{\ell} \alpha_i^{(\ell)} \|a_j|\psi_{\ell}\rangle \| |\langle f_i^{(\ell)}|[X,a_j]|e_i^{(\ell)}\rangle |$$
 (B16)

$$\leq \sqrt{2} \|\nabla X\| \sum_{\ell} \lambda_{\ell} \|a_{j}|\psi_{\ell}\rangle\| \tag{B17}$$

$$\leq \sqrt{2} \|\nabla X\| \operatorname{tr}(\rho N_j)^{\frac{1}{2}}, \tag{B18}$$

where the last inequality follows by Jensen's inequality. Therefore, for all  $t \geq 0$ ,

$$\left|\operatorname{tr}\left[e^{t\mathcal{L}_{\beta}^{\dagger}}([X, a_{i}])\rho a_{i}^{\dagger}\right]\right| \leq \sqrt{2} \left\|\nabla X\right\| \operatorname{tr}(\rho N_{i})^{\frac{1}{2}}.$$
(B19)

Similarly, we find that

$$\left|\operatorname{tr}[a_{j}\rho e^{t\mathcal{L}_{\beta}^{\dagger}}([a_{j}^{\dagger},X])]\right| \leq \sqrt{2} \|\nabla X\| \operatorname{tr}(\rho N_{j})^{\frac{1}{2}}. \tag{B20}$$

and

$$\left|\operatorname{tr}\left[e^{t\mathcal{L}_{\beta}^{\dagger}}([X,a_{j}^{\dagger}])\rho a_{j}\right]\right|, \left|\operatorname{tr}\left[a_{j}^{\dagger}\rho\,e^{t\mathcal{L}_{\beta}^{\dagger}}([a_{j},X])\right]\right| \leq \sqrt{2}\,\|\nabla X\|\,\operatorname{tr}(\rho(I+N_{j}))^{\frac{1}{2}}\,. \tag{B21}$$

Therefore, inserting (B19), (B20) and (B21) into Eq. (B13) and integrating over t gives

$$\begin{aligned} \left| \operatorname{tr}[X(\sigma_{\beta} - \rho)] \right| &\leq \frac{\sqrt{2} \|\nabla X\|}{\sinh(\beta/2)} \sum_{j \in \Lambda} \left( e^{\frac{\beta}{2}} \operatorname{tr}(\rho N_{j})^{\frac{1}{2}} + e^{-\frac{\beta}{2}} \operatorname{tr}(\rho (I + N_{j}))^{\frac{1}{2}} \right) \\ &\leq \frac{2\sqrt{2} \|\nabla X\| \cosh(\beta/2)}{\sinh(\beta/2)} \sum_{j \in \Lambda} \operatorname{tr}(\rho (I + N_{j}))^{\frac{1}{2}} \\ &\leq \frac{2\sqrt{2} |\Lambda|^{\frac{1}{2}} \|\nabla X\| \cosh(\beta/2)}{\sinh(\beta/2)} \left( \sum_{j \in \Lambda} \operatorname{tr}(\rho (I + N_{j})) \right)^{\frac{1}{2}} \\ &= \frac{2\sqrt{2} |\Lambda|^{\frac{1}{2}} \|\nabla X\| \cosh(\beta/2)}{\sinh(\beta/2)} \left( |\Lambda| + \operatorname{tr}(\rho N) \right)^{\frac{1}{2}}. \end{aligned}$$

From duality, i.e., expression (B1), it follows immediately that

$$W_{\mathcal{B}}(\rho, \sigma_{\beta}) \le \frac{2\sqrt{2} |\Lambda|^{\frac{1}{2}} \cosh(\beta/2)}{\sinh(\beta/2)} \left( |\Lambda| + \operatorname{tr}(\rho N) \right)^{\frac{1}{2}}. \tag{B22}$$

Applying the same reasoning to  $\sigma$  instead of  $\rho$  and using the triangle inequality implies

$$W_{\rm B}(\rho,\sigma) \leq \frac{4\sqrt{2} |\Lambda|^{\frac{1}{2}} \cosh(\beta/2)}{\sinh(\beta/2)} \left(|\Lambda| + \operatorname{tr}(\rho N)\right)^{\frac{1}{2}}.$$

The result then follows optimizing over  $\beta$ , i.e., taking the limit  $\beta \to \infty$ .

Proposition 12 can be easily refined in the case where the states  $\rho$  and  $\sigma$  have coinciding marginals on a subset of the modes.

**Proposition 14.** Let  $A \subset [m]$  and  $\rho, \sigma \in \mathcal{D}(\mathcal{H}_m)$  be such that  $\rho_A = \sigma_A$  and  $\operatorname{tr}[\rho N_{A^c}], \operatorname{tr}[\sigma N_{A^c}] \leq E < \infty$  Then

$$W_{\rm B}(\rho,\sigma) \leq 4\sqrt{2|A^c|(|A^c|+E)}$$
.

*Proof.* The proof is similar to tat of Proposition 12. Instead, we consider the decomposition  $W_B(\rho, \sigma) \leq W_B(\rho, \rho_A \otimes \sigma_\beta) + W_B(\sigma_A \otimes \sigma_\beta, \sigma)$ , where  $\sigma_\beta$  is the invariant state of the quantum Ornstein-Uhlenbeck semigroup acting on the subsystem  $A^c$ .

### C. Bosonic Wasserstein contraction coefficients

The proof of the main result of this work relies on the analysis of the Wasserstein distance between two arbitrary states passing through layers of the noisy error correction protocol. For this, we need to relate the Wasserstein distance at the output of a gate included in the latter in terms of the Wasserstein distance at its input. This notion is well captured by the concept of a contraction coefficient:

**Definition 15.** The bosonic Lipschitz contraction coefficient of a quantum channel  $\Phi: \mathcal{T}_1(\mathcal{H}_m) \to \mathcal{T}_1(\mathcal{H}_m)$  is defined as

$$\|\Phi^{\dagger}\|_{\nabla \to \nabla} \coloneqq \sup_{\|\nabla X\| \le 1} \|\nabla \Phi^{\dagger}(X)\|,$$

where the optimization is over all bounded, self-adjoint operators X. Similarly, the bosonic Wasserstein contraction coefficient of  $\Phi$  is defined as

$$\|\Phi\|_{W_{\mathrm{B}} \to W_{\mathrm{B}}} \coloneqq \sup_{\rho \neq \sigma \in \mathcal{D}(\mathcal{H}_m)} \frac{W_{\mathrm{B}}(\Phi(\rho), \Phi(\sigma))}{W_{\mathrm{B}}(\rho, \sigma)} \,.$$

By duality of  $W_{\rm B}$  and  $\|\nabla(.)\|$ , we clearly have that

$$\|\Phi\|_{W_{\mathcal{R}} \to W_{\mathcal{R}}} \le \|\Phi^{\dagger}\|_{\nabla \to \nabla} \,. \tag{C1}$$

Indeed, for any X with  $\|\nabla X\| \leq 1$ , we have that  $Y := \frac{1}{\|\Phi^{\dagger}\|_{\Delta \to \Delta}} \Phi^{\dagger}(X)$  satisfies  $\|\nabla Y\| \leq 1$ , hence

$$|\operatorname{tr}\left(\Phi^{\dagger}(X)(\rho_{1}-\rho_{2})\right)| = \|\Phi^{\dagger}\|_{\nabla \to \nabla} |\operatorname{tr}(Y(\rho_{1}-\rho_{2}))| \leq \|\Phi^{\dagger}\|_{\nabla \to \nabla} W_{B}(\rho_{1},\rho_{2})$$

and the claim (C1) follows by taking the supremum over X.

In the next proposition, we show that the bosonic Lipschitz contraction coefficient tensorizes:

**Proposition 16.** Given  $\Phi = \bigotimes_{j=1}^m \Phi_j$  a tensor product of moment-limited quantum channels  $\Phi_j : \mathcal{T}_1(\mathcal{H}_1) \to \mathcal{T}_1(\mathcal{H}_1)$ . Then

$$\|\Phi^{\dagger}\|_{\nabla \to \nabla} \leq \max_{j \in [m]} \|\Phi_{j}^{\dagger} \otimes id_{[m] \setminus \{j\}}\|_{\nabla_{j} \to \nabla_{j}},$$

 $where \ \|\Phi_j^{\dagger} \otimes id_{[m] \backslash \{j\}}\|_{\nabla_j \to \nabla_j} := \sup_{\|\nabla_j X\| \leq 1} \ \|\nabla_j \Phi_j^{\dagger} \otimes id_{[m] \backslash \{j\}}(X)\|.$ 

Proof. Let  $j \in [m]$ , R be an arbitrary reference system and  $|\varphi\rangle, |\psi\rangle \in \text{dom}(\sqrt{N} \otimes I_R)$ . Since  $\Phi_k$  is moment-limited for each  $k \in [m]$ , we have by Lemma 13 that there exist two sequences of orthogonal, normalized vectors  $\{|e_i\rangle\}_{i\in I}$  and  $\{|f_i\rangle\}_{i\in I}$  in the domain of  $\sqrt{N} \otimes I_R$ , as well as a sequence  $\alpha = \{\alpha_i\}_{i\in I}$  of positive numbers in  $\ell_1$  with  $\|\alpha\|_{\ell_1} \leq 1$  such that  $(\Phi_{\Lambda \setminus \Lambda_j} \otimes \text{id}_{\Lambda_j R})(|\psi\rangle\langle\varphi|) = \sum_i \alpha_i |e_i\rangle\langle f_i|$ , where we denote  $\Lambda_j := \{j\}$ . Therefore, for any self-adjoint bounded X and  $R_i \in \{P_i, Q_i\}$ ,

$$|\langle \varphi | \left[ R_j, \Phi^{\dagger}(X) \right] | \otimes I_R | \psi \rangle| = \left| \sum_i \alpha_i \langle f_i | \left[ R_j, \Phi_j^{\dagger}(X) \right] \otimes I_R | e_i \rangle \right| \stackrel{(1)}{\leq} \sup_{R', |\psi'\rangle, |\varphi'\rangle} |\langle \varphi' | \left[ R_j, \Phi_j^{\dagger}(X) \right] \otimes I_{R'} | \psi' \rangle | \leq \| \nabla_j \Phi_j^{\dagger}(X) \| ,$$

where the inequality (1) follows from choosing R' = RI,  $|\psi'\rangle = \sum_i \sqrt{\alpha_i} |e_i\rangle \otimes |i\rangle$  and  $|\varphi'\rangle = \sum_i \sqrt{\alpha_i} |f_i\rangle \otimes |i\rangle$ , for some orthonormal basis  $\{|i\rangle\}_{i\in I}$ . We have therefore that

$$\|\nabla \Phi^{\dagger}(X)\| \leq \max_{j \in [m]} \|\nabla_j \Phi_j^{\dagger}(X)\| \leq \max_{j \in [m]} \|\Phi_j^{\dagger} \otimes \operatorname{id}_{[m] \setminus \{j\}}\|_{\nabla_j \to \nabla_j} \|\nabla X\|$$

and the result follows.  $\Box$ 

In the next sections, we provide estimates for the bosonic Wasserstein contraction coefficients for the channels appearing in the approximate GKP error correction scheme described in Appendix A 3.

## 1. Contraction under the noise channel $\mathcal{N}_{\lambda}$

We start by showing that the bosonic Wasserstein distance contracts under the

**Proposition 17.** For any  $\lambda \in [0,1]$ ,

$$\|\mathcal{N}_{\lambda}^{\otimes m}\|_{W_{\mathcal{B}} \to W_{\mathcal{B}}} \le \sqrt{\lambda} \,. \tag{C2}$$

*Proof.* By (C1), it is enough to show that  $\|\nabla \mathcal{N}_{\Lambda}^{\dagger}(X)\| \leq \sqrt{\lambda} \|\nabla X\|$  for any bounded, Lipschitz observable X. For any reference system R, any  $j \in \Lambda$  and all  $|\psi\rangle, |\varphi\rangle \in \text{dom}(\sqrt{N} \otimes I_R)$ , we have

$$\begin{split} \langle \varphi | [Q_j, (\mathcal{N}_{\Lambda})^{\dagger}(X)] \otimes I_R | \psi \rangle &= \operatorname{tr}[(\mathcal{N}_{\Lambda})^{\dagger}(X) \otimes I_R | \psi \rangle \langle \varphi | Q_j \otimes I_R] - \operatorname{tr}[Q_j \otimes I_R | \psi \rangle \langle \varphi | (\mathcal{N}_{\Lambda})^{\dagger}(X) \otimes I_R] \\ &= \operatorname{tr}[X \otimes I_R(\mathcal{N}_{\Lambda} \otimes \operatorname{id}_R)(|\psi\rangle \langle \varphi | Q_j \otimes I_R)] - \operatorname{tr}[(\mathcal{N}_{\Lambda} \otimes \operatorname{id}_R)(Q_j \otimes I_R | \psi \rangle \langle \varphi |) X \otimes I_R] \\ &= \sqrt{\lambda} \operatorname{tr}\left[ (\mathcal{N}_{\Lambda} \otimes \operatorname{id}_R)(|\psi\rangle \langle \varphi |) [Q_j, X] \otimes I_R \right], \end{split}$$

where the last identity follows from the commutation relations (A12). The same holds for  $\langle \varphi | [P_j, (\mathcal{N}_\Lambda)^\dagger(X)] \otimes I_R | \psi \rangle$ . Now, since  $\mathcal{N}_\Lambda$  is moment-limited (see [47]), we have by Lemma 13 that there exist two sequences of orthogonal, normalized vectors  $\{|e_i\rangle\}$  and  $\{|f_i\rangle\}$  in the domain of  $\sqrt{N} \otimes I_R$ , as well as a sequence  $\alpha = \{\alpha_i\}$  of positive numbers in  $\ell_1$  with  $\|\alpha\|_{\ell_1} \leq 1$  such that  $(\mathcal{N}_\Lambda \otimes \mathrm{id}_R)(|\psi\rangle\langle\varphi|) = \sum_i \alpha_i |e_i\rangle\langle f_i|$ . Therefore, we have

$$\|\nabla \mathcal{N}_{\Lambda}^{\dagger}(X)\| \leq \sqrt{\lambda} \max_{j} \max_{R_{j} \in \{Q_{j}, P_{j}\}} \sup_{R} \sup_{|\psi\rangle, |\varphi\rangle} \sum_{i} \alpha_{i} \left| \langle f_{i} | [R_{j}, X] \otimes I_{R} | e_{i} \rangle \right| \leq \sqrt{\lambda} \|\nabla X\|. \tag{C3}$$

## 2. Boundedness under GKP error correction

In this section, we control the growth of the Wasserstein distance due to the error correction schemes for the GKP code. We start by Steane's quantum error correction procedure which is illustrated in Figure 1: when acting on a mode  $j \in \Lambda$ , it corresponds to the map

$$\Phi_{j}^{\mathrm{St}}(\rho_{\Lambda}) = \int \mathrm{tr}_{B_{j}C_{j}} \left[ \left( I_{\Lambda} \otimes m_{\hat{P}}(p)_{B_{j}} \otimes m_{\hat{Q}}(q)_{C_{j}} \right) \mathcal{V}_{\Lambda_{j}}^{(q,p)} \left( \mathcal{U}_{\Lambda_{j}C_{j}}^{\mathrm{CNOT}} \circ \mathcal{U}_{B_{j}\Lambda_{j}}^{\mathrm{CNOT}}(\rho_{\Lambda} \otimes \sigma_{B_{j}C_{j}}) \right) \right] dq \, dp$$

where we write  $\Lambda_j := \{j\}$ ,  $\mathcal{V}^{(q,p)} := e^{iqP} e^{-ipQ} (.) e^{ipQ} e^{-iqP}$  and  $\sigma_{B_j C_j} := |\widetilde{0}\rangle \langle \widetilde{0}|_{B_j} \otimes |\widetilde{+}\rangle \langle \widetilde{+}|_{C_j}$ . Using the abbreviations  $\mathcal{U}_{\Lambda_j B_j C_j} := \mathcal{U}^{\text{CNOT}}_{\Lambda_j C_j} \circ \mathcal{U}^{\text{CNOT}}_{B_j \Lambda_j}$  and  $m(q,p) := m_{\widehat{P}}(p) \otimes m_{\widehat{Q}}(q)$ , we have

$$\Phi_j^{\text{St}}(\rho_{\Lambda}) = \int \operatorname{tr}_{B_j C_j} \left[ \left( I_{\Lambda} \otimes m(q, p)_{B_j C_j} \right) \mathcal{V}_{\Lambda_j}^{(q, p)} \mathcal{U}_{\Lambda_j B_j C_j} (\rho_{\Lambda} \otimes \sigma_{B_j C_j}) \right) \right] dq \, dp \; . \tag{C4}$$

17

**Proposition 18.** The quantum channel  $\Phi^{St}$  satisfies

$$\|(\Phi^{\operatorname{St}})^{\otimes m}\|_{W_{\operatorname{B}}\to W_{\operatorname{B}}} \le 2.$$

To prove Proposition 18, we argue once again by duality. We need the adjoint map  $\Phi_i^{\text{St}}$ 

$$(\Phi_{j}^{\mathrm{St}})^{\dagger}(X) = \int \operatorname{tr}_{B_{j}C_{j}} \left( (I_{\Lambda} \otimes \sigma_{B_{j}C_{j}}) \mathcal{U}_{\Lambda_{j}B_{j}C_{j}}^{\dagger} ((\mathcal{V}_{\Lambda_{j}}^{(q,p)})^{\dagger}(X) \otimes m(q,p)_{B_{j}C_{j}}) \right) dq \, dp$$

$$= \operatorname{tr}_{B_{j}C_{j}} \left( (I_{\Lambda} \otimes \sigma_{B_{j}C_{j}}) \mathcal{U}_{\Lambda_{j}B_{j}C_{j}}^{\dagger} \Phi_{j}^{\dagger}(X) \right)$$
(C5)

where we introduced the completely positive unital map

$$\Phi_j^{\dagger}: X \mapsto \int (\mathcal{V}_{\Lambda_j}^{(q,p)})^{\dagger}(X) \otimes m(q,p)_{B_jC_j} dq dp.$$

By Proposition 16, it is enough to consider the maps  $\Phi_j^{\text{St}\dagger}$  and commutations with  $P_j$  and  $R_j$ . Inserting the expression (C5) for  $\Phi^{\text{St}\dagger}$  and using the abbreviation

$$Y_{\Lambda B_j C_j} := \mathcal{U}_{\Lambda_j B_j C_j}^{\dagger} \Phi_j^{\dagger}(X)$$
,

this takes the form

$$\langle \varphi | \left( \left[ P_{j}, \Phi_{j}^{\operatorname{St}\dagger}(X) \right] \otimes I_{R} \right) | \psi \rangle = \operatorname{tr} \left( \left[ P_{j}, \operatorname{tr}_{B_{j}C_{j}} \left( \left( I_{\Lambda} \otimes \sigma_{B_{j}C_{j}} \right) Y_{\Lambda B_{j}C_{j}} \right) \right] \otimes I_{R} | \psi \rangle \langle \varphi | \right)$$

$$= \operatorname{tr} \left( \left( \left[ P_{j} \otimes I_{B_{j}C_{j}}, Y_{\Lambda B_{j}C_{j}} \right] \otimes I_{R} \right) \left( \sigma_{B_{j}C_{j}} \otimes | \psi \rangle \langle \varphi | \right) \right)$$

where we used that  $\operatorname{tr}_{B_jC_j}\left((I_{\Lambda}\otimes\sigma_{B_jC_j})Y_{\Lambda B_jC_j}\right)=\operatorname{tr}_{B_jC_j}\left(Y_{\Lambda B_jC_j}(I_{\Lambda}\otimes\sigma_{B_jC_j})\right)$  in the second identity. That is, we have

$$\langle \varphi | \left( \left[ P_{j}, \Phi_{j}^{\text{St} \dagger}(X) \right] \otimes I_{R} \right) | \psi \rangle = \operatorname{tr} \left( \left( \left[ P_{j} \otimes I_{B_{j}C_{j}}, \mathcal{U}_{\Lambda_{j}B_{j}C_{j}}^{\dagger} \Phi_{j}^{\dagger}(X) \right] \otimes I_{R} \right) (\sigma_{B_{j}C_{j}} \otimes | \psi \rangle \langle \varphi | \right) \right)$$

$$= \operatorname{tr} \left( \left( \left[ \mathcal{U}_{\Lambda_{j}B_{j}C_{j}}(P_{j} \otimes I_{B_{j}C_{j}}), \Phi_{j}^{\dagger}(X) \right] \otimes I_{R} \right) \mathcal{U}_{\Lambda_{j}B_{j}C_{j}}(\sigma_{B_{j}C_{j}} \otimes | \psi \rangle \langle \varphi | \right) \right)$$

$$= \operatorname{tr} \left( \left( \left[ P_{j} \otimes I_{B_{j}C_{j}} - P_{C_{j}} \otimes I_{\Lambda B_{j}}, \Phi_{j}^{\dagger}(X) \right] \otimes I_{R} \right) \mathcal{U}_{\Lambda_{j}B_{j}C_{j}}(\sigma_{B_{j}C_{j}} \otimes (| \psi \rangle \langle \varphi | ) \right)$$
(C6)

where we applied the identity

$$\operatorname{tr}\left([A, \mathcal{U}^{\dagger}(B)]C\right) = \operatorname{tr}\left([\mathcal{U}(A), B)]\mathcal{U}(C)\right) \tag{C7}$$

for conjugation  $\mathcal{U}(\cdot) = U \cdot U^{\dagger}$  by a unitary U in the second step, and used the definition of  $\mathcal{U}_{\Lambda_j B_j C_j}$  (and the corresponding action of CNOT defined in Equation (A14)) in the last step.

Proof of Proposition 18. By (C1) and Proposition 16, it is enough to show that  $\|\Phi_j^{\operatorname{St}^{\dagger}} \otimes \operatorname{id}_{[m]\setminus\{j\}}\|_{\nabla_j \to \nabla_j} \leq 2$  for each  $j \in [m]$ . For sake of clarity, we will omit the identity  $\operatorname{id}_{[m]\setminus\{j\}}$  in the following. Let X be a self-adjoint, bounded operator on  $\mathcal{H}_m$  satisfying the constraints in the definition of  $W_B$  and fix a reference system R. For any two states  $|\psi\rangle, |\varphi\rangle \in \operatorname{dom}(\sqrt{N} \otimes I_R)$ , we denote  $(U_{\Lambda_j B_j C_j} \otimes I_R)\sigma_{B_j C_j} \otimes |\psi\rangle\langle\varphi|(U_{\Lambda_j B_j C_j} \otimes I_R)^{\dagger} = |\psi'\rangle\langle\varphi'|$ , where  $|\psi'\rangle$  and  $|\varphi'\rangle$  are in the domain of  $\sqrt{N_{B_j C_j} + N_{\Lambda}} \otimes I_R$  [47]. Therefore, by (C6), we obtain

$$\langle \varphi | [P_j, \Phi_j^{\text{St}\dagger}(X)] \otimes I_R | \psi \rangle = \Delta_j - \Delta_j'$$
 (C8)

where

$$\Delta_j = \langle \varphi' | ([P_j \otimes I_{B_i C_i}, \Phi_i^{\dagger}(X)] \otimes I_R) | \psi' \rangle \tag{C9}$$

$$\Delta_j' = \langle \varphi' | ([P_{C_j} \otimes I_{\Lambda B_j}, \Phi_j^{\dagger}(X)] \otimes I_R) | \psi' \rangle.$$
 (C10)

We first analyze the expression  $\Delta_i$ . Observe that, again by (C7) we have that

$$\langle \varphi' | ([P_j \otimes I_{B_j C_j}, \Phi_j^{\dagger}(X)] \otimes I_R) | \psi' \rangle = \int \langle \varphi' | \left( [P_j \otimes I_{B_j C_j}, (\mathcal{V}_{\Lambda_j}^{(q,p)})^{\dagger}(X) \otimes m(q, p)_{B_j C_j}] \otimes I_R \right) | \psi' \rangle dq \, dp$$

$$= \int \langle \varphi' | (\mathcal{V}_{\Lambda_j}^{(q,p)})^{\dagger} \left( [\mathcal{V}_{\Lambda_j}^{(q,p)}(P_j) \otimes I_{B_j C_j}, X \otimes m(q, p)_{B_j C_j}] \otimes I_R \right) | \psi' \rangle dq \, dp \, .$$

Since  $\mathcal{V}_{\Lambda_j}^{(q,p)}(P_j) = P_j + pI_{\Lambda}$  for all (q,p), it follows that

$$\Delta_{j} = \langle \varphi' | ([P_{j} \otimes I_{B_{j}C_{j}}, \Phi_{j}^{\dagger}(X)] \otimes I_{R}) | \psi' \rangle = \int \langle \varphi' | (\mathcal{V}_{\Lambda_{j}}^{(q,p)})^{\dagger} ([P_{j} \otimes I_{B_{j}C_{j}}, X \otimes m(q, p)_{B_{j}C_{j}}] \otimes I_{R}) | \psi' \rangle dq \, dp$$

$$= \int \langle \varphi' | ((\mathcal{V}_{\Lambda_{j}}^{(q,p)})^{\dagger} ([P_{j}, X]) \otimes m(q, p)_{B_{j}C_{j}} \otimes I_{R}) | \psi' \rangle dq \, dp$$

$$= \langle \varphi' | (\Phi_{j}^{\dagger}([P_{j}, X]) \otimes I_{R}) | \psi' \rangle .$$

Now, by assumption, the channel  $\Phi_j$  is moment-limited by Lemma 20, so we can use Lemma 13 to argue that, for there exist a family  $\{\alpha_v\}_v$  of positive numbers in  $\ell_1$  with  $\|v\mapsto\alpha_v\|_{\ell_1}\leq 1$ , and two families  $\{|e_v\rangle\}_v$  and  $\{|f_v\rangle\}_v$  of orthonormal vectors in  $\mathrm{dom}(\sqrt{N_\Lambda}\otimes I_R)$  such that

$$\Delta_j = \sum_{v} \alpha_v \langle f_v | ([P_j, X] \otimes I_R) | e_v \rangle.$$
 (C11)

We next analyze the expression  $\Delta'_{i}$ . We have

$$\left[P_{C_{j}} \otimes I_{\Lambda B_{j}}, \Phi_{j}^{\dagger}(X)\right] \otimes I_{R} = \int \left(\left[P_{C_{j}} \otimes I_{\Lambda B_{j}}, (\mathcal{V}_{\Lambda_{j}}^{(q,p)})^{\dagger}(X) \otimes m(q,p)_{B_{j}C_{j}}\right] \otimes I_{R}\right) dq \, dp \\
= \int \left((\mathcal{V}_{\Lambda_{j}}^{(q,p)})^{\dagger}(X) \otimes \left[P_{C_{j}} \otimes I_{B_{j}}, m(q,p)_{B_{j}C_{j}}\right] \otimes I_{R}\right) dq \, dp \\
= \int \left(e^{ipQ_{j}} e^{-iqP_{j}} X e^{iqP_{j}} e^{-ipQ_{j}} \otimes m_{\hat{P}}(p)_{B_{j}} \otimes \left[P_{C_{j}}, m_{\hat{Q}}(q)_{C_{j}}\right] \otimes I_{R}\right) dq \, dp \\
= \int \left(\Psi_{j}^{\dagger} (e^{-iqP_{j}} X e^{iqP_{j}})_{\Lambda B_{j}} \otimes \left[P, m_{\hat{Q}}(q)\right]_{C_{j}} \otimes I_{R}\right) dq \, , \tag{C12}$$

where we introduced the completely positive unital map

$$\Psi_j^{\dagger}: X \mapsto \int e^{ipQ_j} X e^{-ipQ_j} \otimes m_{\hat{P}}(p)_{B_j} dp.$$

Once again,  $\Psi_j$  is moment-limited by Lemma 20, so we can use Lemma 13 and Eq. (C12) to argue that there exist a family  $\{\beta_v\}_v$  of positive numbers and two families  $\{|g_v\rangle\}_v$  and  $\{|h_v\rangle\}_v$  of orthogonal vectors in dom $(\sqrt{N_{\Lambda}+N_{C_j}}\otimes I_R)$  such that

$$\Delta'_{j} = \sum_{v} \beta_{v} \int_{\mathbb{R}} \langle h_{v} | \left( e^{-iqP_{j}} X e^{iqP_{j}} \otimes [P, m_{\hat{Q}}(q)]_{C_{j}} \otimes I_{R} \right) |g_{v}\rangle dq$$

with  $||v \mapsto \beta_v||_{\ell_1} \le 1$ . Lemma 19 below gives an alternative expression for the integral, which allows us to rewrite this as

$$\Delta'_{j} = -\sum_{v} \beta_{v} \int_{\mathbb{R}} \langle h_{v} | \left( e^{-iqP_{j}} [P_{j}, X] e^{iqP_{j}} \otimes m_{\hat{Q}}(q)_{C_{j}} \otimes I_{R} \right) |g_{v}\rangle dq$$

$$= -\langle \varphi' | \left( \Phi_{j}^{\dagger} ([P_{j}, X]) \otimes I_{R} \right) |\psi'\rangle$$

$$= -\Delta_{i}$$

where we re-summed over v in the second identity above. Therefore, we have arrived at the following identity

$$\langle \varphi | \left[ P_j, \Phi_j^{\text{St}\dagger}(X) \right] \otimes I_R | \psi \rangle = 2\Delta_j = 2\sum_v \alpha_v \langle f_v | [P_j, X] \otimes I_R | e_v \rangle.$$
 (C13)

One can also show a similar expression for  $|\langle \varphi | \left( [Q_j, \Phi_j^{\text{St}\dagger}(X)] \otimes I_R \right) | \psi \rangle|$ : following the steps leading to Equation (C6), we find that

$$\langle \varphi | \left( \left[ P_j, (\Phi^{\operatorname{St}\dagger})^{\otimes m}(X) \right] \otimes I_R \right) | \psi \rangle = \Gamma_j + \Gamma_j',$$
 (C14)

where

$$\Gamma_{j} = \langle \varphi' | \left( \left[ Q_{j} \otimes I_{B_{j}C_{j}}, \Phi_{j}^{\dagger}(X) \right] \otimes I_{R} \right) | \psi' \rangle \tag{C15}$$

$$\Gamma'_{j} = \langle \varphi' | \left( \left[ Q_{B_{j}} \otimes I_{\Lambda C_{j}}, \Phi_{j}^{\dagger}(X) \right] \otimes I_{R} \right) | \psi' \rangle. \tag{C16}$$

Then, a similar analysis to that performed for  $\Delta_j$ , using that  $\mathcal{V}_{\Lambda_j}^{(q,p)}(Q_j) = Q_j + qI_{\Lambda}$ , leads to

$$\Gamma_j = \langle \varphi' | \left( \Phi_i^{\dagger}([Q_j, X]) \otimes I_R \right) | \psi' \rangle.$$

Similarly, the analysis of  $\Gamma'_i$  follows that of  $\Delta'_i$ . We have

$$\left[Q_{B_{j}} \otimes I_{\Lambda C_{j}}, \Phi_{j}^{\dagger}(X)\right] \otimes I_{R} = \int \left(\left[Q_{B_{j}} \otimes I_{\Lambda C_{j}}, (\mathcal{V}_{\Lambda_{j}}^{(q,p)})^{\dagger}(X) \otimes m(q,p)_{B_{j}C_{j}}\right] \otimes I_{R}\right) dq \, dp \\
= \int \left((\mathcal{V}_{\Lambda_{j}}^{(q,p)})^{\dagger}(X) \otimes \left[Q_{B_{j}} \otimes I_{C_{j}}, m(q,p)_{B_{j}C_{j}}\right] \otimes I_{R}\right) dq \, dp \\
= \int \left(e^{ipQ_{j}} e^{-iqP_{j}} X e^{iqP_{j}} e^{-ipQ_{j}} \otimes m_{\hat{Q}}(q)_{C_{j}} \otimes \left[Q_{B_{j}}, m_{\hat{P}}(p)_{B_{j}}\right] \otimes I_{R}\right) dq \, dp \\
= \int \left(\Upsilon_{j}^{\dagger} (e^{ipQ_{j}} X e^{-ipQ_{j}})_{\Lambda C_{j}} \otimes \left[Q, m_{\hat{P}}(p)\right]_{B_{j}} \otimes I_{R}\right) dp , \tag{C17}$$

where we introduced the completely positive unital map

$$\Upsilon_j^{\dagger}: X \mapsto \int e^{-iqP_j} X e^{iqP_j} \otimes m_{\hat{Q}}(q)_{C_j} dq.$$

Once again,  $\Upsilon_j$  is moment-limited by Lemma 20, so we can use Lemma 13 and Eq. (C12) to argue that there exist a family  $\{\gamma_v\}_v$  of positive numbers and two families  $\{|p_v\rangle\}_v$  and  $\{|q_v\rangle\}_v$  of orthogonal vectors in dom $(\sqrt{N_{\Lambda}+N_{B_j}}\otimes I_R)$  such that

$$\langle \varphi' | \left( \left[ Q_{B_j} \otimes I_{\Lambda C_j}, \Phi_j^{\dagger}(X) \right] \otimes I_R \right) | \psi' \rangle = \sum_{v} \gamma_v \int_{\mathbb{R}} \langle q_v | \left( e^{ipQ_j} X e^{-ipQ_j} \otimes [Q, m_{\hat{P}}(p)]_{B_j} \otimes I_R \right) | p_v \rangle dp$$

with  $||v \mapsto \gamma_v||_{\ell_1} \le 1$ . Lemma 19 below gives an alternative expression for the integral, which allows us to rewrite this as

$$\langle \varphi' | \left( \left[ Q_{B_j} \otimes I_{\Lambda C_j}, \Phi_j^{\dagger}(X) \right] \otimes I_R \right) | \psi' \rangle = -\sum_{v} \gamma_v \int_{\mathbb{R}} \langle q_v | \left( e^{ipQ_j} [Q_j, X] e^{-ipQ_j} \otimes m_{\hat{P}}(p)_{B_j} \otimes I_R \right) | p_v \rangle dq$$

$$= -\langle \varphi' | \left( \Phi_j^{\dagger} ([Q_j, X]) \otimes I_R \right) | \psi' \rangle,$$

where we re-summed over v in the second identity above. By inserting this expression into (C16), we obtain

$$\Gamma'_{j} = -\langle \varphi' | \left( \Phi_{j}^{\dagger} ([Q_{j}, X]) \otimes I_{R} \right) | \psi' \rangle = -\Gamma_{j} . \tag{C18}$$

Therefore,

$$\langle \varphi | \left[ Q_j, \Phi_j^{\text{St}\dagger}(X) \right] \otimes I_R | \psi \rangle = 0.$$
 (C19)

By combining (C13) and (C19), we have

$$\begin{split} \|\nabla_{j}\Phi_{j}^{\operatorname{St}\dagger}(X)\| &= \max_{R_{j} \in \{Q_{j},P_{j}\}} \sup_{R} \sup_{|\psi\rangle,|\varphi\rangle} \left| \langle \varphi | [R_{j},\Phi_{j}^{\operatorname{St}\dagger}(X)] \otimes I_{R} | \psi \rangle \right| \\ &= 2 \sup_{R} \sup_{|\psi\rangle,|\varphi\rangle} \left| \Delta_{j} \right| \\ &= 2 \sup_{|\psi\rangle,|\varphi\rangle} \left| \sum_{v} \alpha_{v} \left\langle f_{v} | ([P_{j},X] \otimes I_{R}) | e_{v} \right\rangle \right| \\ &\stackrel{(1)}{\leq} 2 \sup_{R'} \sup_{|\psi'\rangle,|\varphi'\rangle} \left| \left\langle \varphi' | [P_{j},X] \otimes I_{R'} | \psi' \right\rangle \right| \\ &\leq 2 \|\nabla_{j}X\| \,, \end{split}$$

where (1) follows from choosing R' = RV,  $|\varphi'\rangle = \sum_v \alpha_v |f_v\rangle \otimes |v\rangle_V$  and  $|\psi'\rangle = \sum_v \alpha_v |e_v\rangle \otimes |v\rangle_V$ .

**Lemma 19.** Let  $|e\rangle, |f\rangle \in \text{dom}(\sqrt{N_{\Lambda} + N_{C_j}} \otimes I_R)$ . Then

$$\int_{\mathbb{R}} \langle f | \left( e^{-iqP_j} X e^{iqP_j} \otimes [P, m_{\hat{Q}}(q)]_{C_j} \otimes I_R \right) | e \rangle dq = -\int_{\mathbb{R}} \langle f | \left( e^{-iqP_j} [P_j, X] e^{iqP_j} \otimes m_{\hat{Q}}(q)_{C_j} \otimes I_R \right) | e \rangle dq .$$

Similarly, for any  $|e\rangle, |f\rangle \in \text{dom}(\sqrt{N_{\Lambda} + N_{B_i}} \otimes I_R)$ :

$$\int_{\mathbb{R}} \langle f | \left( e^{ipQ_j} X e^{-ipQ_j} \otimes [Q, m_{\hat{P}}(p)]_{B_j} \otimes I_R \right) | e \rangle \, dp = -\int_{\mathbb{R}} \langle f | \left( e^{ipQ_j} [Q_j, X] e^{-ipQ_j} \otimes m_{\hat{P}}(p)_{B_j} \otimes I_R \right) | e \rangle \, dp$$

*Proof.* Let us define

$$G(q) := e^{-iqP_j} X e^{iqP_j} \otimes m_{\hat{O}}(q)_{C_i} \otimes I_R \quad \text{and} \quad g(q) := \langle f|G(q)|e \rangle .$$
 (C20)

We compute the derivative  $g'(q) = \langle f|G'(q)|e\rangle$ . Because  $m_{\hat{Q}}(q) = \frac{1}{2\pi\sqrt{\alpha_q}}e^{-iqP}e^{-\frac{1}{2\alpha_q}Q^2}e^{iqP}$  we have

$$m_{\hat{O}}'(q) = -[iP, e^{-iqP} m_{\hat{O}}(0) e^{iqP}] = -i[P, m_{\hat{O}}(q)]$$

and thus

$$G'(q) = e^{-iqP_j} [-iP_j, X] e^{iqP_j} \otimes m_{\hat{Q}}(q)_{C_j} \otimes I_R - e^{-iqP_j} X e^{iqP_j} \otimes [iP, m_{\hat{Q}}(q)] \otimes I_R .$$
 (C21)

Eq. (C27) implies that the claim of the Lemma is equivalent to the statement that  $\int_{\mathbb{R}} g'(q)dq = 0$ . By the fundamental theorem of calculus, it thus suffices to show that

$$\lim_{|q| \to \infty} g(q) = 0. \tag{C22}$$

In order to show (C28), we introduce the following representations of the bipartite states  $|f\rangle$  and  $|e\rangle$ :

$$|f\rangle = \sum_{a} \sqrt{\lambda_a} |\psi_a\rangle_{\Lambda R} \otimes |a\rangle_{C_j}$$
 and  $|e\rangle = \sum_{a} \sqrt{\mu_a} |\varphi_a\rangle_{\Lambda R} \otimes |a\rangle_{C_j}$ 

where  $\{|a\rangle\}$  is the eigenbasis of  $N_{C_j}$ . Since both states belong to the domain of  $I_{\Lambda R} \otimes \sqrt{N_{C_j}}$ , we have that

$$\langle f|I_{\Lambda R}\otimes N_{C_j}|f\rangle = \sum_a \lambda_a \, a < +\infty \qquad \Rightarrow \qquad \lambda_a = \mathop{o}_{a\to\infty}(a^{-2})\,,$$
 (C23)

and similarly  $\mu_a = \underset{a \to \infty}{o}(a^{-2})$ . Next, we consider the decomposition  $g(q) = \sum_{aa'} \sqrt{\lambda_a \mu_{a'}} g_{aa'}(q)$  with overlaps

$$g_{aa'}: q \mapsto \langle \psi_a | \left( e^{-iqP_j} X e^{iqP_j} \right) \, \otimes I_R | \varphi_{a'} \rangle \, \left\langle a | m_{\hat{Q}}(q)_{C_j} | a' \right\rangle.$$

We analyze the second overlap in the definition of  $g_{aa'}$ : denoting by  $\Pi_Q$  the projection-valued measure of the position operator Q, we have

$$m_{\hat{Q}}(q) = \frac{1}{2\pi\sqrt{\alpha_q}} e^{-\frac{1}{2\alpha_q}(Q-q)^2} = \frac{1}{2\pi\sqrt{\alpha_q}} \int_{\mathbb{R}} e^{-\frac{1}{2\alpha_q}(y-q)^2} \Pi_Q(dy) \,.$$

Therefore, for any  $q \in \mathbb{R}$ ,

$$|q| \langle a|m_{\hat{Q}}(q)|a\rangle = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1}{\sqrt{\alpha_{q}}} |q| e^{-\frac{1}{2\alpha_{q}}(y-q)^{2}} \langle a|\Pi_{Q}(dy)|a\rangle$$

$$\stackrel{(1)}{\leq} \frac{\sqrt{\alpha_{q}}}{\pi} \sup_{x \in \mathbb{R}} |x| e^{-x^{2}} + \frac{1}{2\pi\sqrt{\alpha_{q}}} \int_{\mathbb{R}} |y| \langle a|\Pi_{Q}(dy)|a\rangle$$

$$= \frac{\sqrt{\alpha_{q}} e^{-\frac{1}{4}}}{2\pi} + \frac{1}{4\pi^{2}\alpha_{q}} \langle a||Q||a\rangle$$

$$\stackrel{(2)}{\leq} \frac{\sqrt{\alpha_{q}} e^{-\frac{1}{4}}}{2\pi} + \frac{1}{2\pi\sqrt{\alpha_{q}}} \sqrt{\langle a|Q^{2}|a\rangle}$$

$$\stackrel{(2)}{\leq} \frac{\sqrt{\alpha_{q}} e^{-\frac{1}{4}}}{2\pi} + \frac{1}{2\pi\sqrt{\alpha_{q}}} \sqrt{2a+1} .$$

Above, in (1) we simply used the triangle inequality  $|q| \le |q-y| + |y|$  together with the fact that  $\langle a|\Pi_Q(dy)|a\rangle$  defines a probability measure; and in (2) we used that  $Q^2 \le Q^2 + P^2 = 2N + 1$ . Hence,

$$|q| |\langle a| m_{\hat{Q}}(q) |a'\rangle| \leq |q| \left( \langle a| m_{\hat{Q}}(q) |a\rangle \langle a' | m_{\hat{Q}}(q) |a'\rangle \right)^{\frac{1}{2}}$$

$$\leq |q| \left( \frac{\sqrt{\alpha_q} e^{-\frac{1}{4}}}{2\pi} + \frac{1}{2\pi\sqrt{\alpha_q}} \sqrt{2a+1} \right)^{\frac{1}{2}} \left( \frac{\sqrt{\alpha_q} e^{-\frac{1}{4}}}{2\pi} + \frac{1}{2\pi\sqrt{\alpha_q}} \sqrt{2a'+1} \right)^{\frac{1}{2}}. \tag{C25}$$

Therefore, for any  $q \in \mathbb{R} \setminus \{0\}$ , we have

$$\begin{split} |g(q)| & \leq \sum_{aa'} \sqrt{\lambda_a \mu_{a'}} \, |g_{aa'}(q)| \\ & \leq \sum_{aa'} \sqrt{\lambda_a \mu_{a'}} \, \|X\|_{\infty} \, |\langle a| m_{\hat{Q}}(q)_{C_j} |a'\rangle| \\ & \leq \frac{\|X\|_{\infty}}{|q|} \, \sum_{a} \left( \frac{\lambda_a \sqrt{\alpha_q} \, e^{-\frac{1}{4}}}{2\pi} + \frac{\lambda_a}{2\pi \sqrt{\alpha_q}} \, \sqrt{2a+1} \right)^{\frac{1}{2}} \, \sum_{a'} \left( \frac{\mu_{a'} \sqrt{\alpha_q} \, e^{-\frac{1}{4}}}{2\pi} + \frac{\mu_{a'}}{2\pi \sqrt{\alpha_q}} \, \sqrt{2a'+1} \right)^{\frac{1}{2}} \, . \end{split}$$

Since  $\lambda_a, \mu_a = o(a^{-2})$ , the series above converge and therefore  $g(q) \to 0$  as  $|q| \to \infty$ , as required. The second identity claimed proceeds similarly: we define

$$H(p) := e^{ipQ_j} X e^{-ipQ_j} \otimes m_{\hat{P}}(p)_{B_j} \otimes I_R$$
 and  $h(p) := \langle f | h(p) | e \rangle$ . (C26)

We compute the derivative  $h'(p) = \langle f|H'(p)|e\rangle$ . Because  $m_{\hat{P}}(p) = \frac{1}{2\pi\sqrt{\alpha_p}}e^{ipQ}e^{-\frac{1}{2\alpha_p}P^2}e^{-ipQ}$  we have

$$m_{\hat{P}}'(p) = [iQ, e^{ipQ} m_{\hat{P}}(0) e^{-ipQ}] = i[Q, m_{\hat{P}}(p)]$$

and thus

$$H'(p) = e^{ipQ_j}[iQ_j, X]e^{-ipQ_j} \otimes m_{\hat{P}}(p)_{B_j} \otimes I_R + e^{ipQ_j}Xe^{-ipQ_j} \otimes [iQ, m_{\hat{P}}(p)] \otimes I_R .$$
 (C27)

Eq. (C27) implies that the claim of the Lemma is equivalent to the statement that  $\int_{\mathbb{R}} h'(q)dq = 0$ . Again, it suffices to show that

$$\lim_{|p| \to \infty} h(p) = 0 . \tag{C28}$$

For this, we consider the decomposition  $h(p) = \sum_{aa'} \sqrt{\lambda_a \mu_{a'}} h_{aa'}(p)$  with overlaps

$$h_{aa'}: p \mapsto \langle \psi_a | \left( e^{ipQ_j} X e^{-ipQ_j} \right) \otimes I_R | \varphi_{a'} \rangle \langle a | m_{\hat{P}}(p)_{B_j} | a' \rangle.$$

We analyze the second overlap in the definition of  $h_{aa'}$ : as for  $m_{\hat{O}}(q)$  above, we have

$$|p| |\langle a|m_{\hat{P}}(p)|a'\rangle| \le |p| \left(\frac{\sqrt{\alpha_p} e^{-\frac{1}{4}}}{2\pi} + \frac{1}{2\pi\sqrt{\alpha_p}} \sqrt{2a+1}\right)^{\frac{1}{2}} \left(\frac{\sqrt{\alpha_p} e^{-\frac{1}{4}}}{2\pi} + \frac{1}{2\pi\sqrt{\alpha_p}} \sqrt{2a'+1}\right)^{\frac{1}{2}}. \tag{C29}$$

Therefore, for any  $p \in \mathbb{R} \setminus \{0\}$ , we have

$$|h(p)| \leq \frac{\|X\|_{\infty}}{|p|} \sum_{a} \left( \frac{\lambda_a \sqrt{\alpha_p} e^{-\frac{1}{4}}}{2\pi} + \frac{\lambda_a}{2\pi \sqrt{\alpha_p}} \sqrt{2a+1} \right)^{\frac{1}{2}} \sum_{a'} \left( \frac{\mu_{a'} \sqrt{\alpha_p} e^{-\frac{1}{4}}}{2\pi} + \frac{\mu_{a'}}{2\pi \sqrt{\alpha_p}} \sqrt{2a'+1} \right)^{\frac{1}{2}}.$$

Since  $\lambda_a, \mu_a = o(a^{-2})$ , the series above converge and therefore  $h(p) \to 0$  as  $|p| \to \infty$ , as required.

**Lemma 20.** For all  $j \in [m]$ , the maps  $\Phi_j$ ,  $\Psi_j$ ,  $\Upsilon_j$  and  $\Phi_j^{\text{St}}$  are moment-limited, i.e. for all  $\rho \in \mathcal{D}(\mathcal{H}_m)$ ,

$$\operatorname{tr}[\rho N] < \infty \quad \Rightarrow \quad \operatorname{tr}[(\Phi_j \otimes id_{j^c})(\rho)N], \ \operatorname{tr}[(\Psi_j \otimes id_{j^c})(\rho)N], \ \operatorname{tr}[(\Upsilon_j \otimes id_{j^c})(\rho)N], \ \operatorname{tr}[(\Phi_j^{\operatorname{St}} \otimes id_{j^c})(\rho)N] < \infty.$$

*Proof.* The maps  $\Psi_j, \Phi_j$  and  $\Upsilon_j$  are Gaussian and therefore moment-limited (see [47]). As for  $\Phi_j^{\text{St}}$ , it is the composition of the channel  $\Phi_\sigma: \rho \mapsto \rho \otimes \sigma_{B_jC_j}$ , where the approximate GKP state  $\sigma_{B_jC_j}$  has finite energy, with a Gaussian channel.  $\Phi_j^{\text{St}}$  is therefore also moment-limited.

### 3. Boundedness under concatenated stabilizer decoding

In this section, we consider recovery maps obtained from code concatenation. Specifically, we consider an [m,k,d]-stabilizer code  $\mathcal{C}$  - where m denotes the number of physical qubits, k the number of logical qubits and d the code distance- with stabilizer generators  $\{S_j\}_{j=1}^{m-k}$  and an associated recovery map  $\mathcal{R}$  of the form

$$\mathcal{R}(\rho) = \sum_{s \in \{0,1\}^{m-k}} C(s)\Pi(s)\rho\Pi(s)C(s)^{\dagger} , \qquad (C30)$$

where for  $s \in \{0,1\}^{m-k}$ , the operator  $\Pi(s) = \prod_{j=1}^{m-k} \frac{1}{2}(I + (-1)^{s_j}S_j)$  projects onto the syndrome-s subspace, and C(s) is a Pauli-correction classically computed from s, which maps the state back to the code space.

We are interested in local recovery operations, and formalize this as follows: We assume that there is a Clifford circuit U (composed of one- and two-qubits) on the system and additional m-k auxiliary qubits  $A_1 \ldots A_{m-k}$  such that syndrome information can be extracted by applying U and subsequently measuring each individual qubits: We have

$$\Pi(s)\rho\Pi(s) = \operatorname{tr}_{A^{m-k}} \left( \bigotimes_{j=1}^{m-k} |s_j\rangle \langle s_j|_{A_j} \mathcal{U}(\rho \otimes |0\rangle \langle 0|^{\otimes m-k}) \right) \qquad \text{for every} \qquad s \in \{0,1\}^{m-k} \ . \tag{C31}$$

We capture locality as follows: We assume that the backward lightcone of every ancilla qubit has size

$$|\mathcal{L}_{\mathcal{U}}^{\leftarrow}(A_j)| \le \ell_{\text{meas}} \quad \text{for} \quad j = 1, \dots, m - k .$$
 (C32)

Furthermore, we assume that the correction operation can be computed by applying local functions to s: There is a partition  $[m] = \bigcup_{j=1}^r \mathcal{F}_j$  of [m] into disjoint subsets  $\mathcal{F}_1, \ldots, \mathcal{F}_r$  and functions  $C_j : \{0,1\}^{\mathcal{F}_j} \to \{I,X,Y,Z\}^{\otimes |\mathcal{F}_j'|}$  such that

$$C(s) = \prod_{j=1}^{r} C_j(s_{\mathcal{F}_j})_{\mathcal{F}'_j}$$
 for all  $s \in \{0, 1\}^{m-k}$ , (C33)

where  $s_{\mathcal{F}_j}$  denotes the restriction of s to  $\mathcal{F}_j$ , i.e., the corresponding substring of syndrome bits. That is, for each  $j = 1, \ldots, r$ , each operator  $C_j(s_{\mathcal{F}_j})$  has support contained in a set  $\mathcal{F}'_j$ . For what follows, we further assume that the sets  $\{\mathcal{F}'_j\}_j$  are pairwise disjoint.

To formalize locality, assume that there is a constant  $\ell_{\rm corr}$  such that

$$|\mathcal{F}_j| \le \ell_{\text{corr}}$$
 for all  $j = 1, \dots, r$ . (C34)

and a constant  $\ell'_{\rm corr}$  such that

$$|\mathcal{F}'_i| \le \ell'_{\text{corr}}$$
 for all  $j = 1, \dots, r$ . (C35)

In the following, let  $\overline{\mathcal{R}}$  denote a GKP-encoded version of the original recovery map  $\mathcal{R}$ . Let us specify this map in more detail. It has the form

$$\overline{\mathcal{R}}(\rho) = \sum_{s} V_{f(s)} \operatorname{tr}_{A^{(m-k)}} \left[ \bigotimes_{j=1}^{m-k} \overline{P}(s_j)_{A_j} \overline{\mathcal{U}} \left( \rho \otimes \bigotimes_{j=1}^{m-k} \sigma_{A_j} \right) \right] V_{f(s)}^{\dagger} . \tag{C36}$$

In this expression, the bosonic ancilla states  $\sigma_{A_j}$  are assumed to have finite moments – typically, they are approximate GKP states. The unitary  $\overline{\mathcal{U}}$  is a circuit composed of Gaussian one- and two-mode unitaries, e.g., a product of CNOT gates if the stabilizer code  $\mathcal{C}$  is of CSS-type. It acts on m system modes augmented by m-k auxiliary modes. The assumption (C32) amounts to the following: denoting the systems's quadrature operators by  $(R_1,\ldots,R_{2m})=(Q_1,P_1,\ldots,Q_m,P_m)$ , and those of the auxiliary modes by  $\{R_{2m+j}\}_{j=1}^{2(m-k)}$ , we have the condition

$$\mathcal{U}(R_{2m+j}) = \sum_{i=1}^{2m+2(m-k)} u_{j,i} R_i, \quad \text{with} \quad |\{i \in \{1, \dots, 2m+2(m-k)\} \mid u_{j,i} \neq 0\}| \le \ell_{\text{meas}}. \quad (C37)$$

for every j = 1, ..., 2(m - k).

The projection  $\overline{P}(s_k)_{A_k}$  corresponds to a measurement of logical Paulis on the ancilla  $A_k$ . In practice, this can be done by an approximate homodyne measurement in the form of Equations (A16) and (A17) modulo a fixed real parameter. Since for approximate GKP states, the measurement generally gives real outcomes, we choose  $s_k \in \{0,1\}$  by rounding the outcome. More precisely, each POVM element  $\overline{P}(s_i)$  can be written as an integral over the operator density  $m_{\hat{Q}}$  or  $m_{\hat{P}}$ , for instance

$$\overline{P}(0) = \int_{\Gamma} m_{\hat{P}}(z) \, dz \,, \qquad \overline{P}(1) = I - \overline{P}(0) = \int_{\Gamma^c} m_{\hat{P}}(z) \, dz \,,$$

where  $\Gamma$  denotes the domain of the real line such that whenever the parameter  $z \in \Gamma$  is measured, it is interpreted as syndrome 0.

Finally, the correction is done via a displacement  $V_{f(s)}$  of parameter  $f(s) \in \mathbb{R}^{2m}$ : This is the GKP-encoded counterpart of applying the Pauli correction C(s). Recalling that logical  $\overline{X}$ ,  $\overline{Y}$ - and  $\overline{Z}$ -operators in the GKP code are given by the displacements in Equation (A15), namely

$$\overline{X} = e^{-i\sqrt{\pi}\hat{P}}, \qquad \overline{Z} = e^{i\sqrt{\pi}\hat{Q}}, \qquad \overline{Y} = e^{i\sqrt{\pi}(\hat{Q}-\hat{P})},$$
 (C38)

we conclude that

$$||f(s)||_{\infty} \le \sqrt{\pi}$$
 for every  $s \in \{0,1\}^{m-k}$ . (C39)

It follows from Eq. (C33) that the function  $f:\{0,1\}^{m-k}\to\mathbb{R}^{2m}$  satisfies the following: We have

$$f(s) = \left(\bigoplus_{u=1}^{r} f_u(s_{\mathcal{F}_u})_{\mathcal{F}'_u}\right) \oplus \left(0^{2(m-\sum_{u=1}^{r} |\mathcal{F}'_u|)}\right)_{[m]\setminus\bigcup_{u=1}^{r} \mathcal{F}'_u} \quad \text{for every} \quad s \in \{0,1\}^{m-k} . \tag{C40}$$

for some functions  $f_u: \{0,1\}^{\mathcal{F}_u} \to \{0,1\}^{\mathcal{F}'_u}$ . In particular,  $V_{f(s)} = \prod_{u=1}^r V_{f_u(s_{\mathcal{F}_u})_{\mathcal{F}'_u}}$  is a product of (possibly multimode) displacements on disjoint subsets  $\{\mathcal{F}'_u\}_u$  of modes, each depending on at most  $\ell_{\text{corr}}$  syndrome bits. This completes the specification of the recovery map  $\overline{\mathcal{R}}$  in (C36). Our main result for this recovery map is the following:

**Proposition 21.** The recovery map  $\overline{\mathcal{R}}$  satisfies

$$\|\overline{\mathcal{R}}\|_{W_{\mathrm{B}} \to W_{\mathrm{B}}} \leq \left(\ell_{\mathrm{meas}} + \frac{\ell_{\mathrm{meas}} \, 2^{1 + \ell_{\mathrm{meas}} (1 + \ell_{\mathrm{corr}})} \ell_{\mathrm{meas}} \cdot \ell_{\mathrm{corr}}'}{\sqrt{\alpha_{\mathrm{min}} \pi}}\right).$$

Before discussing the proof of Proposition 21, let us briefly summarize some of the relevant properties of the function f derived from Eqs. (C39) and (C40). For any two syndromes  $s, s' \in \{0, 1\}^{m-k}$  differing on  $\delta$  fixed ancillas, f(s) - f(s') only depends on at most  $\delta \cdot \ell_{\text{corr}}$  ancilla bits, and has support on at most  $\delta \cdot \ell'_{\text{corr}}$  modes. More precisely, defining

$$D := \{ j \in [m-k] \mid s_j \neq s_j' \} \tag{C41}$$

as the set of indices where the syndromes s and s' differ, there are  $v \in \{0,1\}^{|D|}$   $w \in \{0,1\}^{m-k-|D|}$  such that

$$\begin{aligned}
s_D &= v & s_{D^c} &= w \\
s'_D &= \overline{v} & s'_{D^c} &= w
\end{aligned}$$
(C42)

where  $D^c = [m-k] \setminus D$  and  $\overline{v}_j = 1 - v_j$  for each  $j \in [|D|]$ . It follows immediately that for  $u \in [r]$ , we have

$$s_{\mathcal{F}_u} \neq s'_{\mathcal{F}_u}$$
 if and only if  $D \cap \mathcal{F}_u \neq \emptyset$ . (C43)

Omitting "padding" zeros, we thus have (using Eq. (C40))

$$f(s) - f(s') = \bigoplus_{u \mid D \cap \mathcal{F}_u \neq \emptyset} \left( f_u(s_{\mathcal{F}_u}) - f_u(s'_{\mathcal{F}_u}) \right)_{\mathcal{F}'_u} = \bigoplus_{u \mid D \cap \mathcal{F}_u \neq \emptyset} g_u \left( v, s_{\mathcal{F}_u \setminus D} \right)_{\mathcal{F}'_u}$$
(C44)

with  $g_u(v,t) := f_u(v_{D \cap \mathcal{F}_u} \oplus t_{\mathcal{F}_u \setminus D}) - f_u(\overline{v}_{D \cap \mathcal{F}_u} \oplus t_{\mathcal{F}_u \setminus D})$ . In this expression, the only dependence on  $w = s_{D^c}$  is through the collection of substrings  $\{s_{\mathcal{F}_u \setminus D}\}_{u \in D}$ , hence the total number of bits of w the expression f(s) - f(s') depends on is bounded by

$$|\{u|D \cap \mathcal{F}_u \neq \emptyset\}| \cdot \max_{u|D \cap \mathcal{F}_u \neq \emptyset} |\mathcal{F}_u \setminus D| \le |D| \cdot \max_{u \in D} |\mathcal{F}_u \setminus D| \le |D| \cdot \ell_{\text{corr}}.$$
(C45)

In summary, there is a subset  $M \subset D^c$  of size  $|M| \leq |D| \cdot \ell_{\text{corr}}$  and a function g such that

$$f(s) - f(s') = g(v, w_M)$$
 for all  $v, w$  where  $s = s(v, w), s' = s'(v, w)$  is given by (C44). (C46)

Observe that (C44) also implies that the number of non-zero components of  $g(v, w_M)$  is bounded by

$$|\{j \in [2m] \mid g_j(v, w_M) \neq 0\}| \leq 2|\{u|D \cap \mathcal{F}_u \neq \emptyset\}| \cdot \max_{\substack{u \mid D \cap \mathcal{F}_u \neq \emptyset}} |\mathcal{F}_u'| \leq 2|D| \cdot \ell_{\text{corr}}', \tag{C47}$$

that is, f(s) - f(s') has support on at most  $|D| \cdot \ell'_{\text{corr}}$  modes, as claimed.

Proof of Proposition 21. Once again, we argue by duality: we first fix a reference system R, choose a vertex  $j \in [m]$  and consider a self-adjoint, bounded operator X as well as  $|\psi\rangle, |\varphi\rangle \in \text{dom}(\sqrt{N} \otimes I_R)$ . Then with a slight abuse of notation, we denote  $\overline{P}(s) := \bigotimes_{j=1}^{m-k} \overline{P}(s_j)$  for a given syndrome  $s = (s_1, \dots, s_{m-k}), X(s) := V_{f(s)}^{\dagger} X V_{f(s)}, \sigma_{A^{(m-k)}} = \bigotimes_{j=1}^{m-k} \sigma_{Aj},$  and compute

$$\overline{\mathcal{R}}^{\dagger}(X) = \sum_{s \in \{0,1\}^{m-k}} \operatorname{tr}_{A^{(m-k)}} \left[ \left( I \otimes \sigma_{A^{(m-k)}} \right) \overline{\mathcal{U}}^{\dagger} \left( X(s) \otimes \overline{P}(s) \right) \right].$$

Next, we consider a quadrature  $R_{2j+1} = P_j$  corresponding to a vertex  $j \in [m]$  and partition each syndrome  $s = (v^{(j)}, w^{(j)})$  onto a part  $v^{(j)}$  of measurement outcomes  $s_k$  corresponding to projections  $\overline{P}(s_k)$  of support overlapping with the light-cone of vertex j with respect to the unitary map  $\overline{\mathcal{U}}$ , and a part  $w^{(j)}$  of measurement outcomes corresponding to projections whose support does not overlap with that light-cone. That is, we set

$$v^{(j)} = \{s_r\}_{r: j \in \mathcal{L}_{\mathcal{U}}^{\leftarrow}(A_r)} \quad \text{and} \quad w^{(j)} = \{s_r\}_{r: j \notin \mathcal{L}_{\mathcal{U}}^{\leftarrow}(A_r)}$$
 (C48)

Let us reorder (relabel) the syndrome bits in such a way that the first  $\delta_j$  bits constitute  $v^{(j)}$  and the remaining  $m-k-\delta_j$  bits constitute  $w^{(j)}$ . By assumption, we have  $\delta_j \leq \ell_{\text{meas}}$ . Then we have

$$\sum_{s \in \{0,1\}^{m-k}} X(s) \otimes \overline{P}(s) = \sum_{(v,w) \in \{0,1\}^{\delta_j} \times \{0,1\}^{m-k-\delta_j}} X(v,w) \otimes \overline{P}(v,w) 
= \sum_{w \in \{0,1\}^{m-k-\delta_j}} X(0^{\delta_j},w) \otimes P(0^{\delta_j},w) + \sum_{\substack{v \in \{0,1\}^{\delta_j} \setminus \{0^{\delta_j}\}\\w \in \{0,1\}^{m-k-\delta_j}}} X(v,w) \otimes \overline{P}(v,w) 
= \sum_{w \in \{0,1\}^{m-k-\delta_j}} X(0^{\delta_j},w) \otimes \overline{P}(w) + \sum_{\substack{v \in \{0,1\}^{\delta_j} \setminus \{0^{\delta_j}\}\\w \in \{0,1\}^{m-k-\delta_j}}} \Delta_v X(w) \otimes \overline{P}(v,w)$$
(C49)

where we introduced the operators

$$\overline{P}(w) \coloneqq \sum_{v \in \{0,1\}^{\delta_j}} \overline{P}(v, w)$$
$$\Delta_v X(w) \coloneqq X(v, w) - X(0^{\delta_j}, w)$$

for  $w \in \{0,1\}^{m-k-\delta_j}$ . By definition,  $\overline{P}(w)$  acts non-trivially only on auxiliary qubits  $A_i$  with the property that  $j \notin \mathcal{L}_{\mathcal{U}}^{\leftarrow}(A_i)$ . Note that we have

$$\overline{\mathcal{U}}(P_j) = \sum_{i=1}^{2m} u_{2j+1,i} R_i + \sum_{i \in \{1,\dots,m-k\}: j \in \mathcal{L}_{\mathcal{U}}^{\leftarrow}(A_i)}^{m-k} \left( u_{2j+1,2m+2i-1} R_{2m+2i-1} + u_{2j+1,2m+2i} R_{2m+2i} \right)$$
(C50)

since all terms i with j not belonging to the backward lightcone of  $A_i$  do not contribute, i.e.,

$$u_{2j+1,2m+2i-1} = u_{2j+1,2m+2i} = 0$$
 for all  $j \notin \mathcal{L}_{\mathcal{U}}^{\leftarrow}(A_i)$ . (C51)

It thus follows with Eq. (C49), (C49) and the support property of  $\overline{P}(w)$  that

$$[\overline{\mathcal{U}}(P_j), \Lambda] = \sum_{i=1}^{2m} u_{2j+1,i}[R_i, \Lambda] + \sum_{i=1}^{2(m-k)} u_{2j+1,2m+i}[R_{2m+i}, \Lambda']$$
(C52)

where we used the abbreviations

$$\Lambda := \sum_{s \in \{0,1\}^{m-k}} X(s) \otimes \overline{P}(s) \quad \text{and} \quad \Lambda' := \sum_{\substack{v \in \{0,1\}^{\delta_j} \setminus \{0^{\delta_j}\}\\ w \in \{0,1\}^{m-k-\delta_j}}} \Delta_v X(w) \otimes \overline{P}(v,w). \quad (C53)$$

With (C53), we obtain

$$[P_{j}, \overline{\mathcal{R}}^{\dagger}(X)] = \sum_{s \in \{0,1\}^{m-k}} \operatorname{tr}_{A^{(m-k)}} \left[ (I \otimes \sigma_{A^{(m-k)}}) [P_{j}, \overline{\mathcal{U}}^{\dagger}(X(s) \otimes \overline{P}(s))] \right]$$

$$= \sum_{s \in \{0,1\}^{m-k}} \operatorname{tr}_{A^{(m-k)}} \left[ (I \otimes \sigma_{A^{(m-k)}}) \overline{\mathcal{U}}^{\dagger} \left( [\overline{\mathcal{U}}(P_{j}), X(s) \otimes \overline{P}(s)] \right) \right]$$

$$= \Delta_{j} + \Delta'_{j}$$
(C54)

where

$$\Delta_{j} := \sum_{i=1}^{2m} u_{2j+1,i} \sum_{s \in \{0,1\}^{m-k}} \operatorname{tr}_{A^{(m-k)}} \left[ \left( I \otimes \sigma_{A^{(m-k)}} \right) \overline{\mathcal{U}}^{\dagger} \left( \left[ R_{i}, X(s) \right] \otimes \overline{P}(s) \right) \right]$$
(C55)

$$\Delta_{j}' := \sum_{i=1}^{2(m-k)} u_{2j+1,2m+i} \sum_{\substack{v \in \{0,1\}^{\delta_{j}} \setminus \{0^{\delta_{j}}\}\\ w \in \{0,1\}^{m-k-\delta_{j}}}} \operatorname{tr}_{A^{(m-k)}} \left[ \left( I \otimes \sigma_{A^{(m-k)}} \right) \overline{\mathcal{U}}^{\dagger} \left( \Delta_{v} X(w) \otimes \left[ R_{2m+i}, \overline{P}(v,w) \right] \right) \right]. \tag{C56}$$

Next, we bound each of these sums separately. We start with  $\Delta_j$ : first, we observe by the commutation relations between quadratures and displacements that  $[R_i, X(s)] = V_{f(s)}^{\dagger}[R_i, X]V_{f(s)}$ . Therefore, we can rewrite  $\Delta_j$  as

$$\Delta_j = \sum_{i=1}^{2m} u_{2j+1,i} \, \Phi^{\dagger}([R_i, X]) \quad \text{with} \quad \Phi^{\dagger}(Y) := \sum_{s \in \{0,1\}^{m-k}} \operatorname{tr}_{A^{(m-k)}} \left[ (I \otimes \sigma_{A^{(m-k)}}) \, \overline{\mathcal{U}}^{\dagger}(V_{f(s))}^{\dagger} Y V_{f(s)} \otimes \overline{P}(s) \right].$$

Then we use Lemma 13 together with Lemma 24 to argue that since  $|\varphi\rangle, |\psi\rangle \in \text{dom}(\sqrt{N} \otimes I_R)$ , there exists a sequence  $\{\alpha_u\}$  of positive numbers with  $\|u\mapsto\alpha_u\|_{\ell_1}\leq 1$ , and sequences  $\{|e_u\rangle\}$  and  $\{|f_u\rangle\}$  of orthogonal vectors in  $\operatorname{dom}(\sqrt{N} \otimes I_R)$  such that

$$\langle \varphi | (\Delta_j \otimes I_R) | \psi \rangle = \sum_{i=1}^{2m} u_{2j+1,i} \sum_u \alpha_u \langle e_u | [R_i, X] \otimes I_R | f_u \rangle.$$
 (C57)

Therefore,

$$|\langle \varphi | (\Delta_j \otimes I_R) | \psi \rangle| \le \sum_{i=1}^{2m} |u_{2j+1,i}| \|\nabla X\| \le \ell_{\text{meas}} \|\nabla X\|.$$
 (C58)

where in the second inequality we used that the number of indices  $i \in [2(m-k)]$  such that  $u_{2j+1,2m+1} \neq 0$  is upper bounded by  $\ell_{\text{meas}}$ , see Eq. (C37). Next, we consider  $\langle \varphi | (\Delta'_j \otimes I_R) | \psi \rangle$ : recalling that the state  $\sigma_{A^{(m-k)}}$  is a tensor product of pure GKP states which we denote by  $|\widetilde{\psi}_{A^{(m-k)}}\rangle$ , we can rewrite the quantity as

$$\langle \varphi | (\Delta'_{j} \otimes I_{R}) | \psi \rangle = \sum_{i=1}^{2(m-k)} u_{2j+1,2m+i} \sum_{\substack{v \in \{0,1\}^{\delta_{j}} \setminus \{0^{\delta_{j}}\}\\ w \in \{0,1\}^{m-k-\delta_{j}}}} \langle \widetilde{\varphi} | \left( \overline{\mathcal{U}}^{\dagger} \left( \Delta_{v} X(w) \otimes [R_{2m+i}, \overline{P}(v,w)] \right) \otimes I_{R} \right) | \widetilde{\psi} \rangle$$
 (C59)

where we denoted  $|\widetilde{\varphi}\rangle \coloneqq |\varphi\rangle \otimes |\widetilde{\psi}_{A^{(m-k)}}\rangle, \ |\widetilde{\psi}\rangle \coloneqq |\psi\rangle \otimes |\widetilde{\psi}_{A^{(m-k)}}\rangle.$  By definition, we have with the definition  $\mathcal{V}_x(\cdot) = |\psi\rangle \otimes |\widetilde{\psi}_{A^{(m-k)}}\rangle$  $V_x^{\dagger}(\cdot)V_x$  for  $x \in \mathbb{R}^{2m}$  that

$$\Delta_{v}X(w) = \mathcal{V}_{f(v,w)}(X) - \mathcal{V}_{f(0_{j}^{\delta},w)}(X)$$

$$= \mathcal{V}_{f(v,w)}\left(X - \mathcal{V}_{f(v,w)}^{\dagger}(\mathcal{V}_{f(0_{j}^{\delta},w)}(X))\right)$$

$$= \mathcal{V}_{f(v,w)}\left(X - \mathcal{V}_{-\Delta_{v}f(w)}(X)\right) \tag{C60}$$

where we used that  $\mathcal{V}_{\alpha} \circ \mathcal{V}_{\beta} = \mathcal{V}_{\alpha+\beta}$  and  $\mathcal{V}_{\alpha}^{\dagger} = \mathcal{V}_{-\alpha}$  and the abbreviation

$$\Delta_v f(w) := f(v, w) - f(0^{\delta_j}, w) \quad \text{for} \quad v \in \{0, 1\}^{\delta_j} \text{ and } w \in \{0, 1\}^{m-k-\delta_j} .$$
 (C61)

From the local structure of the function f, (i.e., expression (C40) and the fact that  $|\mathcal{F}_i| \leq \ell_{\text{corr}}$ ), it follows that we can decompose the vector w into two subvectors  $w = (w_1, w_2)$  where  $w_1 \in \{0, 1\}^{\delta_j \cdot \ell_{\text{corr}}}$ , such that

$$\Delta_v f(w) = \delta_v f(w_1) \tag{C62}$$

for a function  $\delta_v f: \{0,1\}^{\delta_j \cdot \ell_{\text{corr}}} \to \mathbb{R}^{2m}$ , see Eq. (C46). For fixed  $v \in \{0,1\}^{\delta_j} \setminus \{0^{\delta_j}\}$  and  $w = (w_1, w_2) \in \{0,1\}^{m-k-\delta_j}$  we have by Eq. (C60)

$$\Delta_{v}X(w) \otimes [R_{2m+i}, \overline{P}(v, w)] = (\mathcal{V}_{f(v,w)}(X - \mathcal{V}_{-\Delta_{v}f(w)}(X))) \otimes [R_{2m+i}, \overline{P}(v, w)]$$

$$= (\mathcal{V}_{f(v,w)}(X - \mathcal{V}_{-\delta_{v}f(w_{1})}(X))) \otimes [R_{2m+i}, \overline{P}(v, w_{1})\overline{P}(w_{2})]$$

$$= (\mathcal{V}_{f(v,w)}(\delta_{v}(X)(w_{1}))) \otimes [R_{2m+i}, \overline{P}(v, w_{1})\overline{P}(w_{2})],$$

where

$$\delta_v(X)(w_1) := X - \mathcal{V}_{-\delta_v, f(w_1)}(X). \tag{C63}$$

It follows that

$$\sum_{\substack{w=(w_1,w_2)\\ \in \{0,1\}^{m-k-\delta_j}}} \langle \widetilde{\varphi} | \left( \mathcal{U}^{\dagger} \left( \Delta_v X(w) \otimes [R_{2m+i}, \overline{P}(v,w)] \right) \otimes I_R \right) | \widetilde{\psi} \rangle$$
(C64)

$$= \sum_{w_1 \in \{0,1\}^{\delta_j \cdot \ell_{\mathrm{corr}}}} \langle \widetilde{\varphi} | \left( \Psi_{w_1}^{\dagger} \left( (\delta_v(X)(w_1) \otimes [R_{2m+i}, \overline{P}(v, w_1)] \right) \otimes I_R \right) | \widetilde{\psi} \rangle,$$

where

$$\Psi_{w_1}^{\dagger}(Z) := \sum_{w_2 \in \{0,1\}^{m-k-\delta_j(\ell_{\operatorname{corr}}+1)}} \overline{\mathcal{U}}^{\dagger} \left( \mathcal{V}_{f(v,(w_1,w_2))}(Z) \otimes \overline{P}(w_2) \right) . \tag{C65}$$

The map  $\Psi_{w_1}^{\dagger}$  is completely positive, unital and moment-limited by Lemma 24. Therefore, by Lemma 13 together with Lemma 24 there exists a sequence  $\{\alpha_u^{(w_1)}\}_u$  of positive numbers with  $\|u\mapsto\alpha_u^{(w_1)}\|_{\ell_1}\leq 1$ , and sequences  $\{|e_u^{(w_1)}\rangle\}_u$  and  $\{|f_u^{(w_1)}\rangle\}_u$  of vectors in  $\mathrm{dom}(\sqrt{N}\otimes I_R)$  such that

$$\langle \widetilde{\varphi} | \left( \Psi_{w_1}^{\dagger} \left( (\delta_v(X)(w_1) \otimes [R_{2m+i}, \overline{P}(v, w_1)] \right) \otimes I_R \right) | \widetilde{\psi} \rangle = \sum_u \alpha_u^{(w_1)} \langle e_u^{(w_1)} | (\delta_v X(w_1) \otimes [R_{2m+i}, \overline{P}(v, w_1)] \otimes I_R) | f_u^{(w_1)} \rangle \ .$$

We show in Lemma 22 below that

$$||[R_{2m+i}, \overline{P}(v, w_1)]|| \le \frac{1}{\pi \sqrt{\alpha_q}}$$
 (C66)

when i is even (i.e.,  $R_{2m+i}$  is a momentum-operator). Defining the normalized vectors

$$|\tilde{f}_u^{(w_1)}\rangle := \frac{([R_{2m+i}, \overline{P}(v, w_1)] \otimes I_R)|f_u^{(w_1)}\rangle}{\|([R_{2m+i}, \overline{P}(v, w_1)] \otimes I_R)|f_u^{(w_1)}\rangle\|},$$
(C67)

it thus follows that

$$\left| \langle \widetilde{\varphi} | \left( \Psi_{w_1}^{\dagger} \left( (\delta_v(X)(w_1) \otimes [R_{2m+i}, \overline{P}(v, w_1)] \right) \otimes I_R \right) | \widetilde{\psi} \rangle \right| \leq \frac{1}{\pi \sqrt{\alpha_q}} \sum_{u} \alpha_u^{(w_1)} \left| \langle e_u^{(w_1)} | (\delta_v X(w_1) \otimes I_{A^{(m-k)}R}) | \widetilde{f}_u^{(w_1)} \rangle \right| .$$

Applying (C74) from Lemma 23 to each term in this sum gives

$$\left| \langle \widetilde{\varphi} | \left( \Psi_{w_1}^{\dagger} \left( (\delta_v(X)(w_1) \otimes [R_{2m+i}, \overline{P}(v, w_1)] \right) \otimes I_R \right) | \widetilde{\psi} \rangle \right| \leq \frac{2\ell_{\text{meas}} \cdot \ell'_{\text{corr}}}{\sqrt{\alpha_q \pi}} \| \nabla X \| .$$
 (C68)

Inserting (C68) into (C64) and combining this with (C59) gives

$$\left| \langle \varphi | (\Delta'_{j} \otimes I_{R}) | \psi \rangle \right| \leq \sum_{i=1}^{2(m-k)} |u_{2j+1,2m+i}| \sum_{\substack{v \in \{0,1\}^{\delta_{j}} \setminus \{0^{\delta_{j}}\}\\ w_{i} \in \{0,1\}^{\delta_{j}} \cdot \ell_{\text{corr}}}} \left| \langle \widetilde{\varphi} | \left( \Psi_{w_{1}}^{\dagger} \left( (\delta_{v}(X)(w_{1}) \otimes [R_{2m+i}, \overline{P}(v, w_{1})] \right) \otimes I_{R} \right) | \widetilde{\psi} \rangle \right|$$
(C69)

$$\leq \frac{2^{1+\ell_{\text{meas}}(1+\ell_{\text{corr}})}\ell_{\text{meas}} \cdot \ell'_{\text{corr}}}{\sqrt{\alpha_q \pi}} \|\nabla X\| \cdot \sum_{i=1}^{2(m-k)} |u_{2j+1,2m+i}| .$$
(C70)

Recall that the number of indices  $i \in [2(m-k)]$  such that  $u_{2j+1,2m+1} \neq 0$  is upper bounded by  $\ell_{\text{meas}}$ , see Eq. (C37) hence

$$\left| \langle \varphi | (\Delta_j' \otimes I_R) | \psi \rangle \right| \le \frac{\ell_{\text{meas}} 2^{1 + \ell_{\text{meas}} (1 + \ell_{\text{corr}})} \ell_{\text{meas}} \cdot \ell_{\text{corr}}'}{\sqrt{\alpha_q \pi}} \| \nabla X \| . \tag{C71}$$

Combining (C58) and (C71) with Eq. (C54) gives

$$\left| \langle \varphi | [P_j, \overline{\mathcal{R}}^{\dagger}(X)] \psi \rangle \right| \le \left( \ell_{\text{meas}} + \frac{\ell_{\text{meas}} 2^{1 + \ell_{\text{meas}} (1 + \ell_{\text{corr}})} \ell_{\text{meas}} \cdot \ell'_{\text{corr}}}{\sqrt{\alpha_q \pi}} \right) \| \nabla X \| . \tag{C72}$$

Repeating the steps that led to this bound for  $P_j$ , with  $P_j$  replaced by  $Q_j$ , we then have found that

$$\|\nabla \Phi^{\operatorname{Stab}\,\dagger}(X)\| \leq \left(\ell_{\operatorname{meas}} + \frac{\ell_{\operatorname{meas}} \, 2^{1+\ell_{\operatorname{meas}}(1+\ell_{\operatorname{corr}})} \ell_{\operatorname{meas}} \cdot \ell'_{\operatorname{corr}}}{\sqrt{\alpha_{\min} \pi}}\right) \|\nabla X\| \ .$$

The claim then follows by duality, cf. (C1).

Lemma 22. We have

$$||[P_{m+i}, \overline{P}(v, w_1)]|| \le \frac{1}{\pi \sqrt{\alpha_q}}. \tag{C73}$$

*Proof.*  $R_i = P$  is a local momentum operator, so that it is enough to consider that  $P(s_i) = \int_{\Gamma} m_{\hat{P}}(z) dz$  where  $\Gamma$  denotes the domain of the real line such that whenever the parameter  $z \in \Gamma$  is measured, it is interpreted as syndrome  $s_i$ . Then,

$$\begin{split} |[R_i,P(s_i)]| &= \Big| \int_{\Gamma} [R_i,m_{\hat{P}}(z)] \, dz \Big| \\ &= \frac{1}{2\pi\sqrt{\alpha_q}} \Big| \int_{\Gamma} V_{(0,-z)^t} [P,e^{-\frac{1}{2\alpha_q}Q^2}] V_{(0,z)^t} \, dz \Big| \\ &= \frac{1}{2\pi\alpha_q^{\frac{3}{2}}} \Big| \int_{\Gamma} V_{(0,-z)^t} \, Q e^{-\frac{1}{2\alpha_q}Q^2} \, V_{(0,z)^t} \, dz \Big| \\ &= \frac{1}{2\pi\alpha_q^{\frac{3}{2}}} \Big| \int_{\Gamma} (Q-z) e^{-\frac{1}{2\alpha_q}(Q-z)^2} \, dz \Big| \\ &\leq \frac{1}{2\pi\alpha_q^{\frac{3}{2}}} \int |Q-z| \, e^{-\frac{1}{2\alpha_q}(Q-z)^2} \, dz \\ &= \frac{1}{\pi\sqrt{\alpha_q}} \, I \,, \end{split}$$

and therefore  $||[R_i, P(s_i)]||_{\infty} \leq (\pi \sqrt{\alpha_q})^{-1}$ .

**Lemma 23.** Let  $|e\rangle$  and  $|f\rangle$  be vectors in dom $(\sqrt{N_{\Lambda} + N_{C_j}} \otimes I_R)$ . Then there is a distribution over pairs  $(|e'\rangle, |f'\rangle)$  of vectors in dom $(\sqrt{N_{\Lambda} + N_{C_j}} \otimes I_R)$  such that

$$|\langle e|(\delta_v X(w_1) \otimes I_{A^{(m-k)}R})|f\rangle| \leq 2\sqrt{\pi}\ell_{\text{meas}} \cdot \ell'_{\text{corr}} \cdot \max_{i \in [2m]} \mathbb{E}_{|e'\rangle,|f'\rangle} \left[ |\langle e'|([R_i,X] \otimes I_{A^{(m-k)}R})|f'\rangle| \right].$$

In particular, we have

$$|\langle e|(\delta_v X(w_1) \otimes I_{A^{(m-k)}R})|f\rangle| \le 2\sqrt{\pi} \,\ell_{\text{meas}} \cdot \ell'_{\text{corr}} \cdot \|\nabla X\| \ . \tag{C74}$$

Proof. Recall (cf. Eq. (C63)) that  $\delta_v(X)(w_1) \coloneqq X - \mathcal{D}_{-\delta_v f(w_1)}(X)$  is the difference of two displaced operators. Define  $\xi \coloneqq -\delta_v f(w_1)$ . Recall that by Eq. (C47), the number of non-zero entries of  $\xi$  is upper bounded by  $\ell'_{\text{corr}} \cdot \ell_{\text{meas}}$  and  $\|\xi\|_{\infty} \le 2\sqrt{\pi}$  by (C39). We now consider a linear interpolation between 0 and  $\xi \in \mathbb{R}^{2m}$ : We have

$$\begin{split} |\langle e|\delta_v X(w_1)\otimes I_{A^{(m-k)}R}|f\rangle| &= \Big|\int_0^1 \frac{d}{dx} \left\langle e|(V_{x\xi}^\dagger X \, V_{x\xi}\otimes I_{A^{(m-k)}R})|f\rangle \, dx \Big| \\ &\leq \int_0^1 \sum_{i=1}^{2m} |\xi_i| \cdot \left| \langle e|(V_{x\xi}^\dagger [R_i,X] V_{x\xi}\otimes I_{A^{(m-k)}R})|f\rangle \right| \, dx \Big| \\ &\leq 2\sqrt{\pi} \, \ell_{\text{meas}} \cdot \ell'_{\text{corr}} \cdot \max_{i \in [2m]} \int_0^1 |\langle V_{x\xi}e|([R_i,X] \otimes I_{A^{(m-k)}R})|V_{x\xi}f\rangle| \, dx \, . \end{split}$$

This implies the claim.

Lemma 24. The maps  $\Phi$  and  $\Psi_{w_1}^{\dagger}$  are moment-limited.

*Proof.* This is clear since these maps are linear combinations of compositions of Gaussian channels and partial measurements. The latter are of the form  $\rho_{AB} \mapsto \operatorname{tr}_B(I_A \otimes P_B \rho_{AB})$  for some local projection  $P_B$ , so that

$$\operatorname{tr}(\operatorname{tr}_B(I_A \otimes P_B \rho_{AB}) N_A) = \operatorname{tr}(\rho_{AB} N_A \otimes P_B) \leq \operatorname{tr}(\rho_{AB} N_A) \leq \operatorname{tr}(\rho_{AB} N_{AB}).$$

Therefore, these maps are moment-limited.

30