

Asymptotically good bosonic Fock state codes

Dor Elimelech^{1,2}, Arda Aydin², and Alexander Barg^{2,1,3}

¹Joint Center for Quantum Information and Computer Science (QuICS), UMD/NIST

²Institute for Systems Research, University of Maryland, College Park, MD 20742

³Department of ECE, University of Maryland, College Park, MD 20742

We study the error-correction properties of multi-mode Fock-state codes under amplitude-damping (AD) noise, focusing on the asymptotic regime in which the total excitation of the code states grows without limit and the number of photon losses induced by the noise scales linearly with it. In this setting, existing code families, which correct only sublinearly many photon losses, do not protect against amplitude-damping (AD) noise with a constant loss parameter. As our main result, we address this gap by constructing asymptotically good Fock-state codes relying on random classical codes in the discrete simplex. Unlike many standard constructions of random quantum codes, our construction introduces randomness through the underlying classical indexing structure. Randomization also enables another desirable feature: bounded per-mode occupancy, which limits the number of photons in any individual mode and thereby increases the coherence lifetime of the code states. Finally, via a relation between Fock-state codes and permutation-invariant codes, our results also yield asymptotically good families of qudit permutation-invariant codes as well as codes in monolithic nuclear state spaces.

Dor Elimelech: dor@umd.edu

Arda Aydin: arda@umd.edu

Alexander Barg: abarg@umd.edu

Contents

1	Introduction	2
2	Preliminaries	5
2.1	Notation and definitions	5
2.2	Exact and approximate quantum codes	6
3	Bosonic Fock state codes and the AD channel	9
3.1	The AD channel	11
3.2	The truncated AD channel and AQEC	13
4	Asymptotically good Fock state codes from ℓ_1 codes	15
4.1	The construction of [2]	16
4.2	Asymptotically good Fock state codes	17
4.3	De-randomized asymptotically good Fock state codes	23
5	Asymptotically good PI and spin codes	24
5.1	PI codes	24
5.2	Spin codes	26
6	Classical ℓ_1 codes	27
6.1	Deterministic ℓ_1 codes	27
6.2	Random ℓ_1 codes	29
7	Conclusion and outlook	35
A	Technicalities and auxiliary results	39
A.1	Channels and state norms	39
A.2	Proof of Theorem 5	41
A.3	Typical vectors in the simplex	44
A.4	Combinatorial and probabilistic results	45
A.5	Additional proofs	46

1 Introduction

Large-scale quantum computing is essential for running advanced quantum algorithms. Maintaining reliability at scale requires error correction codes that satisfy specific criteria imposed by hardware limitations. While various state spaces can be used for that purpose, bosonic systems provide a promising approach by encoding information in the infinite-dimensional Hilbert space of a harmonic oscillator. In this work, we study Fock state codes in the asymptotic regime when the total excitation scales to infinity.

Fock states, a.k.a. number states, are the eigenstates of the number operator $\hat{n}$, forming an infinite basis for the Hamiltonian of the harmonic oscillator. A Fock state $|n\rangle$, $n = 0, 1, 2, \dots$ represents the number of photons occupying a specific mode of the electromagnetic field. Fock state codes are designed to protect quantum information by encoding it in the Fock basis. The encoded information in such systems is subject to decoherence resulting from photon loss. This noise model corresponds to the *amplitude damping (AD)* channel, $\mathcal{N}$, and admits a Kraus

representation

$$\rho \mapsto \sum_{k=0}^{\infty} A_k \rho A_k^\dagger \quad (1)$$

with Kraus operators

$$A_k = \sum_{n=k}^{\infty} \sqrt{\binom{n}{k}} \sqrt{(1-\gamma)^{n-k} \gamma^k} |n-k\rangle\langle n|, \quad k = 0, 1, 2, \dots \quad (2)$$

Here, $\gamma \in (0, 1]$ is a constant denoting the probability of decay, and the integer k represents the number of photons lost. Fock state codes can be generalized to a multi-mode setting by utilizing q harmonic oscillators rather than a single one. The q -mode Fock space is spanned by the product states $|n_0\rangle \otimes |n_1\rangle \otimes \dots \otimes |n_{q-1}\rangle$, $n_i = 0, 1, \dots$; $i = 0, 1, \dots, q-1$. In this paper, we focus on *constant excitation* Fock state codes, in which every state satisfies the fixed excitation condition $\sum_{i=0}^{q-1} n_i = N$.

We study error correction with multi-mode Fock state codes used over the AD channel $\mathcal{N}$. For this noise model, it is typically assumed that the channel acts independently on each mode of the composite system. Error-correcting properties of Fock state codes are often evaluated relying on a quantity termed code fidelity, which serves as a computationally tractable proxy for the worst-case fidelity, $\mathcal{F}$. Given a quantum code Q , the code fidelity is defined as

$$\mathcal{F}_{\mathcal{A}, Q} := \min_{|\psi\rangle \in Q} \sum_{A \in \mathcal{A}} \langle \psi | A^\dagger A | \psi \rangle,$$

where $\mathcal{A}$ is the set of errors that are corrected by the code Q . For the AD channel, we can define $\mathcal{A}$ as the set of operators $\mathcal{A}_{\leq t}$ representing at most t photon losses. For the case of constant excitation codes, prior works rely on an approximation of the code fidelity $\mathcal{F} \geq 1 - O(\gamma^{t+1})$, supporting the claim of error correction for fixed-length codes in channels with a small loss parameter. This approach, initially proposed in [14] and used in a number of follow-up papers [55, 32, 41], has been a common starting point in the code analysis. We argue that the utility of code fidelity is limited for two distinct reasons. First, in the case of constant excitation codes, considered here and in [14, 2], $\mathcal{F}_{\mathcal{A}_{\leq t}}$ is a universal constant that does not depend on the code (see Prop. 3), and second, code fidelity as a performance metric is of little use unless the sequence of codes $(Q_N)_N$ recovers from a proportion of photon losses linear in N . It is easily shown, as we argue below, that codes that do not satisfy this condition cannot perform well on the AD channel unless γ vanishes as N increases.

This leaves the challenge of constructing positive rate Fock state codes that correct a linear proportion of losses, typically referred to as *asymptotically good*, which until this work was left unanswered. As far as code constructions are concerned, for a long while the best known scaling of the number of correctable errors t for codes of total excitation N has been $t \propto \sqrt{N}$, due to a construction of Bergmann and van Loock [9]. A recent paper [2] took another step further by introducing a general construction of multi-mode constant excitation Fock state codes relying on existing classical ℓ_1 -simplex codes, used to produce positive rate codes capable of correcting $o(N/\log N)$ errors. While the codes presented in [2] offer a significant advance in error-correction capability, they suffer from three major pitfalls:

1. The sublinear photon-loss correction implies no protection against the AD channel when N is large.

2. Their construction relies on a result from convex geometry, whose implementation incurs computation cost that scales doubly exponentially with the total number of photons, rendering the construction virtually unfeasible.
3. They lack a desirable feature of Fock states: a limited photon count per mode (balancedness). Deterministic generation of Fock states with higher photon numbers remains a significant experimental bottleneck for realizing scalable Fock state codes [52, 53, 24]. Furthermore, high-photon-number states are also more fragile: the probability of losing a single photon from $|n\rangle$ is approximately n times higher than the $|1\rangle$ state, so the coherence lifetime of a Fock state $|n\rangle$ scales as $1/n$ [54]. This phenomenon limits the maximum number of photons that can be utilized in a single mode for realizing Fock state codes.

In this work, we overcome these shortcomings by (1) replacing the classical deterministic ℓ_1 -simplex codes employed in [2] with randomized constructions and (2) replacing exact error-correction mechanisms with *approximate* ones. Harnessing the resulting randomness in the obtained Fock-state code, we dispense with the convex geometry argument in favor of a simple low-complexity averaging procedure, thereby overcoming the complexity obstacle. In addition, these random ℓ_1 codes have better parameters than the deterministic ℓ_1 codes used in [2], which enables approximate correction of linearly many photon losses and supports strong protection against the AD channel.

This set of results constitutes the main contribution of this work. We next describe the ideas that enabled us to establish these results, highlighting the gains attained by relying on approximate quantum error correction. Although relaxing the exact quantum error correction (QEC) requirement to an approximate one ostensibly reflects a performance loss, this is in fact not the case. Indeed, even exact QEC of t -amplitude errors ultimately translates only to approximate QEC for the AD channel that is not restricted to a fixed number of photon losses. To make this perspective rigorous, we formally define a truncated AD channel $\mathcal{N}_{\leq t}$ and relate the worst-case entanglement fidelity of the code used on it to its performance on the non-truncated AD channel. Specifically, we prove that a code is an approximate quantum error-correcting (AQEC) code for the AD channel if and only if it is an (exact or approximate) QEC code for the truncated AD channel $\mathcal{N}_{\leq t}$. Our formalization enables us to extend the definition of asymptotically good families to the approximate correction case (see Def. 6). Coupled with the equivalence of exact and approximate error correction on the AD channel $\mathcal{N}$ (Theorem 5), this leads to a conclusion that a code sequence is asymptotically good in either approximate or exact sense on the $\mathcal{N}_{\leq t}$ channel if and only if it is approximate asymptotically good for the AD channel (see Corollary 6).

To prove that our codes are indeed asymptotically good, we use the framework of approximate error correction developed by Bény and Oreshkov [8], who derived necessary and sufficient conditions for a code to be ϵ -error-correcting for a noisy channel. The averaging procedure over the random classical ℓ_1 codewords together with concentration arguments ensures that with probability exponentially close to 1, the resulting quantum code satisfies the AQEC conditions for linearly many photon losses while maintaining a positive rate. We also note that, although our randomized construction only succeeds with high probability, once a realization is given, one can efficiently verify whether the resulting code is good by performing a number of checks proportional to $(\text{number of Kraus operators})^2 \times (\text{dimension of the code})^2$.

More broadly, this fits the general theme that random quantum code constructions are natural candidates for approximate quantum error correction. However, unlike existing works in which randomness is introduced at the level of the physical Hilbert space, such as Haar-random codes or constructions based on randomly selected many-body eigenstates [26, 28, 33, 12], our construction introduces randomness at the level of the underlying classical indexing structure from which the quantum codewords are generated. This is advantageous because the resulting quantum code inherits useful properties of random classical codes. In particular, relying on randomization of classical ℓ_1 codes, we obtain Fock state codes that satisfy the balancedness property in a strong form, with at most logarithmic per-mode excitation.

The paper is structured as follows. We start with a formalization of the truncated AD channel $\mathcal{N}_{\leq t}$ and quantify its relation to the “full” AD channel $\mathcal{N}$. In this part (Section 3.2), we also establish a relation between exact and approximate error correction and define asymptotically good Fock state code families. Section 4 is devoted to the construction of the described Fock state code families. Supporting results about classical ℓ_1 codes are collected in Section 6. Finally, in Section 5 we recall that one of the main results of [2] establishes a relation between Fock state codes and qudit permutation invariant (PI) codes [46, 39, 3] as well as families of spin codes [21]. Leveraging our results for Fock state codes, we prove the existence of asymptotically good codes from these code families, advancing the best previously known scaling of $O(N/\log N)$ for the number of correctable errors [2], where N is the code length for PI codes and total spin for spin codes, respectively.

2 Preliminaries

2.1 Notation and definitions

We begin by introducing the notation used throughout the paper. Underlined letters $\underline{n} = (n_0, \dots, n_{q-1})$ refer to q -tuples of nonnegative integers. Random variables are denoted by sans-serif symbols (e.g., X, Y). We use $\mathbb{Z}_0$ and $\mathbb{N}$ to denote the set of non-negative and positive integers, respectively. For $K \in \mathbb{N}$ we define $[K] \triangleq \{0, 1, \dots, K-1\}$. For an integer $N, q \in \mathbb{N}$, consider the set of all ordered integer partitions (a discrete simplex), given by

$$\mathcal{S}_{q,N} \triangleq \left\{ \underline{n} \in \mathbb{Z}_0^q : \sum_{i=0}^{q-1} n_i = N \right\}.$$

The number of points in the simplex is given by the balls-and-bins calculation,

$$|\mathcal{S}_{q,N}| = \binom{N+q-1}{q-1}.$$

By abuse of notation, we will use the notation $\mathcal{S}_{\alpha,N}$ also whenever $\alpha > 1$ is not an integer, with the convention that $\mathcal{S}_{\alpha,N} \triangleq \mathcal{S}_{\lfloor \alpha \rfloor, N}$. For $N \in \mathbb{N}$ and $\underline{n} \in \mathbb{Z}_0^q$, the multinomial coefficient $\binom{N}{\underline{n}}$ is defined in a standard way:

$$\binom{N}{\underline{n}} \triangleq \begin{cases} \frac{N!}{\prod_{i=0}^{q-1} n_i!} & \underline{n} \in \mathcal{S}_{q,N} \\ 0 & \text{otherwise.} \end{cases}$$

We will often use the binary entropy function $h_2 : [0, 1] \rightarrow [0, 1]$, defined as

$$h_2(x) = -x \log_2 x - (1 - x) \log_2 (1 - x),$$

to approximate the growth rate of binomial coefficients.

Given a Hilbert space $\mathcal{H}$, denote by $L(\mathcal{H})$ and $D(\mathcal{H})$ the space of bounded linear operators and density operators on it, respectively. We consider the Hilbert space $\ell_2(\mathbb{N})$ of all sequences $\mathbf{a} = (a_n)_{n=0}^\infty$ with finite ℓ_2 norm: $\|\mathbf{a}\|_2 = \sqrt{\sum_n |a_n|^2} < \infty$. We denote the standard orthonormal basis of $\ell_2(\mathbb{N})$ by $\{|0\rangle, |1\rangle, \dots\}$, where $|n\rangle$ denotes the sequence $\mathbf{a}^{(n)}$ given by $a_i^{(n)} = \delta_n(i)$, with δ_n denoting the Kronecker delta at n .

We use standard asymptotic notation: for functions $f, g : \mathbb{N} \rightarrow \mathbb{R}_+$, we say that $f = O(g)$ if there exists a constant $c > 0$ such that $f(n) \leq cg(n)$ for all $n \in \mathbb{N}$. We say that $f = o(g)$ if $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$.

2.2 Exact and approximate quantum codes

A quantum channel is given by a completely positive trace-preserving (CPTP) map $\mathcal{N} : L(\mathcal{H}) \rightarrow L(\mathcal{H}')$, where $\mathcal{H}$ and $\mathcal{H}'$ are two Hilbert spaces. Such CPTP maps can be equivalently represented by a set of Kraus error operators $\{A_k\}_k$ satisfying the completeness condition $\sum_k A_k^\dagger A_k = I$, such that the action of $\mathcal{N}$ on a state ρ has the form $\mathcal{N}(\rho) = \sum_k A_k \rho A_k^\dagger$. A code Q in $\mathcal{H}$ is said to correct an error set $\mathcal{E}$ if there exists a decoding operation (i.e., a CPTP map) $\mathcal{D} : L(\mathcal{H}') \rightarrow L(\mathcal{H})$ such that for any channel $\mathcal{N}$ with $\{A_k\}_k \subset \mathcal{E}$ we have

$$\mathcal{D} \circ \mathcal{N}|_{L(Q)} = I_{L(Q)},$$

meaning that any input state $|\psi_{\text{in}}\rangle \in Q$ can be perfectly recovered from its noisy version. The well-known Knill-Laflamme (KL) conditions imply that a code Q with an orthonormal basis $|c_0\rangle, \dots, |c_{k-1}\rangle$ corrects $\mathcal{E}$ if there are constants $\{\lambda_{k,l}\}_{k,l}$, such that for any $A_k, A_l \in \mathcal{E}$ and i, j we have

$$\langle c_i | A_k^\dagger A_l | c_j \rangle = \lambda_{k,l} \cdot \delta_{i,j}. \quad (3)$$

Under approximate quantum error-correction (AQEC), we allow the recovered state to be slightly different, but still close, to the original state. Many frameworks for AQEC were studied in the literature [35, 47, 7, 33, 8], suggesting various criteria and tools to gauge the performance of an AQEC code. We follow the formalism introduced by Bény and Oreshkov [8], in which the performance of an AQEC code is measured using the worst-case entanglement fidelity, defined as

$$\mathcal{F}_e(\mathcal{N}, \mathcal{M}) \triangleq \inf_{\rho \in D(\mathcal{H})} \mathcal{F}(\mathcal{N} \otimes I_{L(\mathcal{H})}(|\psi_\rho\rangle\langle\psi_\rho|), \mathcal{M} \otimes I_{L(\mathcal{H})}(|\psi_\rho\rangle\langle\psi_\rho|)), \quad (4)$$

where $\mathcal{N}, \mathcal{M} : L(\mathcal{H}) \rightarrow L(\mathcal{H}')$ are two quantum channels, $|\psi_\rho\rangle$ is any purification of ρ , and $\mathcal{F}(\rho, \tau) = \|\sqrt{\rho}\sqrt{\tau}\|_1$ is the fidelity of a pair of states (the trace norm $\|\cdot\|_1$ is introduced below in Definition 8). While entanglement fidelity provides a natural measure of approximation between channels, it is often technically convenient to work with an equivalent figure of merit. Below, we quantify approximation using the Bures distance, defined as

$$d(\mathcal{N}, \mathcal{M}) := \sqrt{1 - \mathcal{F}_e(\mathcal{N}, \mathcal{M})}.$$

Definition 1 (Approximate quantum error correction). A code $Q \subseteq \mathcal{H}$ is called an ε -approximate error-correcting code (AQECC) for a channel $\mathcal{N}$ if there exists a decoding operation $\mathcal{D}$ such that

$$d(\mathcal{D} \circ \mathcal{N}|_{L(Q)}, I_{L(Q)}) = \sqrt{1 - \mathcal{F}_e(\mathcal{D} \circ \mathcal{N}|_{L(Q)}, I_{L(Q)})} \leq \varepsilon.$$

Extending this definition, consider a sequence of channels $(\mathcal{N})_N$ on Hilbert spaces $(\mathcal{H}_N)_N$ such that there exists a sequence of codes $Q_N \subset \mathcal{H}_N$ with the property that for each N , Q_N is an ε_N -AQECC for $\mathcal{N}_N$. If $\varepsilon_N \rightarrow 0$ as $N \rightarrow \infty$, we say that $\mathcal{Q} = (Q_N)_N$ is an approximate error-correcting sequence of codes for $(\mathcal{N}_N)_N$.

We observe that a code that is an ε -AQECC for a channel $\mathcal{N}$ becomes an exact code if $\varepsilon = 0$.

Remark 1. In contrast to exact quantum error correction, the approximate case does not satisfy the linearity property with respect to the set of correctable errors. For exact quantum codes, the KL conditions imply that if a code corrects a set of error operators, then it also corrects their linear span, making the error-set model rich and meaningful. In the approximate setting, this equivalence fails: approximate correctability of a given set of error operators does not extend in a controlled way to linear combinations or rescalings, and it is therefore unclear which quantum channels can be meaningfully composed from a fixed approximately correctable error set. As a consequence, in the approximate case, we switch perspective from the error set to the channel view, specifying a noise model, and determining whether it can be approximately inverted. This approach generally results in a systematic formulation of AQEC. One exception is presented in a recent work [33], which develops an approximate error-set model; however, it relies crucially on the strong properties of Haar random codes and highly regular unitary error families and does not readily extend beyond this special setting.

In [8], Bény and Oreshkov derived necessary and sufficient conditions for approximate quantum error correction on a given channel, formulated in terms of its Kraus operators:

Theorem 1 (AQEC conditions [8]). *Let $\mathcal{N}$ be a channel with Kraus operators $\mathcal{E} = \{A_k\}_{k=0}^{M-1}$. A code space $Q \subset \mathcal{H}$ is an ε -AQECC if and only if there exists a density matrix $\lambda = \{\lambda_{k,l}\}_{k,l}$ such that $d(\Lambda, \Lambda + \mathcal{B}) \leq \varepsilon$, where $\Lambda, \mathcal{B} : L(\mathcal{H}) \rightarrow L(\mathbb{C}^M)$ are the quantum channels given by*

$$\Lambda(\rho) = \sum_{k,l=0}^{M-1} \lambda_{k,l} \text{tr}(\rho) |k\rangle \langle l|, \quad \mathcal{B}(\rho) = \sum_{k,l=0}^{M-1} \text{tr}(B_{k,l}\rho) |k\rangle \langle l|, \quad (5)$$

where

$$B_{k,l} \triangleq P A_k^\dagger A_l P - \lambda_{k,l} P,$$

P is the projection on Q , and $\{|k\rangle\}_{k=0}^{M-1}$ is an orthonormal basis of $\mathbb{C}^M$.

In the following proposition, we modify the assumptions of Theorem 1 to derive an approximate analog of the KL conditions in (3) that forms a sufficient condition for AQEC and is tailored to the specific structure of the codes that we study in Section 4. We remark that analyses in a similar general spirit, though for different settings, have earlier appeared in the literature, for instance in [12, 6] for Pauli error channels.

Proposition 2 (Approximate KL conditions). *Let $\mathcal{N}$ be a quantum channel with Kraus operators $\mathcal{E} = \{A_k\}_{k=0}^{M-1}$, and let $Q \subseteq H$ be a quantum code with an orthonormal basis $\{|c_i\rangle\}_{i=0}^{K-1}$ that satisfies*

$$\langle c_i | A_k^\dagger A_l | c_j \rangle = 0 \quad (6)$$

if $i \neq j$ or $\ell \neq k$. Then Q is an ε -AQECC with respect to $\mathcal{E}$ if there exist non-negative numbers $\{\lambda_l\}_{l=0}^{M-1}$ such that $\sum_{l=0}^{M-1} \lambda_l = 1$ and

$$\max_{i,l} \left| \langle c_i | A_l^\dagger A_l | c_i \rangle - \lambda_l \right| \leq \frac{\varepsilon^2}{KM}.$$

Proof. Consider the set $\{\lambda_{k,l}\}_{k,l}$ such that $\lambda_{k,l} = \delta_{kl} \lambda_l$, where λ_l are as in the above statement, and let Λ and $\mathcal{B}$ be the operators defined in (5). Note that the matrix defined by $\lambda_{k,l}$ is a density matrix since it is diagonal with positive entries and unit trace by the assumption. By Theorem 1, it is sufficient to show that the Bures distance $d(\Lambda, \Lambda + \mathcal{B}) < \varepsilon$. We will estimate this distance using the diamond norm (see Definition 8) and then use (42) in Lemma 20. It is thus sufficient to show that $\|\Lambda + \mathcal{B} - \Lambda\|_\diamond = \|\mathcal{B}\|_\diamond < \varepsilon^2$, where $\mathcal{B} : L(\mathcal{H}) \rightarrow L(\mathbb{C}^M)$ is now given by

$$\mathcal{B}(\rho) = \sum_{k,l=0}^{M-1} \text{tr}(B_{k,l}\rho) \cdot |k\rangle \langle l| = \sum_{l=0}^{M-1} \text{tr}(B_{l,l}\rho) \cdot |l\rangle \langle l|,$$

where the last equality follows since $B_{k,l} = 0$ for $k \neq l$ by the assumption (6).

By Lemma 22, we can write $\|\mathcal{B}\|_\diamond = \|I_{\mathcal{H}} \otimes \mathcal{B}(|\psi\rangle \langle \psi|)\|_1$ for some $|\psi\rangle \in C^{\otimes 2}$. By the Schmidt decomposition, we can write $|\psi\rangle$ as $|\psi\rangle = \sum_{i=0}^{K-1} p_i |i\rangle \otimes |i'\rangle$, where $\sum_i p_i^2 = 1$ and $\{|i\rangle\}_i, \{|i'\rangle\}_i$ are orthonormal bases of (two copies of) Q . Let

$$\|B\|_{\text{op}} = \sup_{\|X\|_1 \leq 1} \|B(X)\|_1$$

be the operator norm of $\mathcal{B}$ with respect to the trace norm.

With these preparations, we compute

$$\begin{aligned} \|\mathcal{B}\|_\diamond &= \|I_{L(C)} \otimes \mathcal{B}(|\psi\rangle \langle \psi|)\|_1 = \left\| \sum_{i,j=0}^{K-1} p_i p_j |i\rangle \langle j| \otimes \mathcal{B}(|i'\rangle \langle j'|) \right\|_1 \\ &\leq \sum_{i,j=0}^{K-1} p_i p_j \| |i\rangle \langle j| \|_1 \cdot \|\mathcal{B}(|i'\rangle \langle j'|)\|_1 \end{aligned} \quad (7)$$

$$\leq \|\mathcal{B}\|_{\text{op}} \cdot \sum_{i,j=0}^{K-1} p_i p_j = \|\mathcal{B}\|_{\text{op}} \cdot \left(\sum_i p_i \right)^2 \quad (8)$$

$$\leq \|\mathcal{B}\|_{\text{op}} \cdot \left(\sqrt{\sum_{i=0}^{K-1} 1^2 \cdot \sum_i p_i^2} \right)^2 \quad (9)$$

$$\begin{aligned} &= K \cdot \sup_{\|X\|_1 \leq 1} \|B(X)\|_1 = K \cdot \sup_{\|X\|_1 \leq 1} \left\| \sum_{l=0}^{M-1} \text{tr}(B_{l,l}X) \cdot |l\rangle \langle l| \right\|_1 \\ &= K \cdot \sup_{\|X\|_1 \leq 1} \sum_{l=0}^{M-1} |\text{tr}(B_{l,l}X)| \leq K \cdot \sum_{l=0}^{M-1} \sup_{\|X\|_1 \leq 1} |\text{tr}(B_{l,l}X)|, \end{aligned} \quad (10)$$

where (7) follows by the triangle inequality and multiplicativity of the trace norm with respect to the tensor product, (8) uses the fact that $\| |i\rangle \langle j| \|_1 = 1$ for orthonormal vectors $|i\rangle, |j\rangle$ as well as the definition of $\|\cdot\|_{\text{op}}$, and (9) follows by the Cauchy–Schwarz inequality.

By the duality principle, Lemma 21, the supremum in (10) can be bounded as

$$\sup_{\|X\|_1 \leq 1} |\text{tr}(B_{l,l}X)| \leq \|B_{l,l}\|_{\infty}, \quad (11)$$

where $\|\cdot\|_{\infty}$ is the usual spectral norm given by the maximal singular value. Recalling that $\{|c_i\rangle\}_{i=0}^{K-1}$ is an orthonormal basis, we obtain

$$\begin{aligned} B_{l,l} &= \sum_{i,j=0}^{K-1} |c_i\rangle \langle c_i| A_l^\dagger A_l |c_j\rangle \langle c_j| - \lambda_{l,l} \cdot \sum_{i=0}^{K-1} |c_i\rangle \langle c_i| \\ &= \sum_{i=0}^{K-1} (\langle c_i| A_l^\dagger A_l |c_i\rangle - \lambda_{l,l}) \cdot |c_i\rangle \langle c_i|, \end{aligned}$$

where the second equality follows by (6). Thus, the set of eigenvalues of $B_{l,l}$ has the form $\{\langle c_i| A_l^\dagger A_l |c_i\rangle - \lambda_{l,l}\}_k$. In particular,

$$\|B_{l,l}\|_{\infty} \leq \max_i |\langle c_i| A_l^\dagger A_l |c_i\rangle - \lambda_{l,l}|.$$

Denote the right-hand side of this inequality by $\varepsilon_{\max}$. Combining this with (10) and (11), we obtain

$$\|\mathcal{B}\|_{\diamond} \leq KM\varepsilon_{\max}.$$

This completes the proof. $\square$

3 Bosonic Fock state codes and the AD channel

We consider a bosonic quantum system comprising q modes, each associated with the separable Hilbert space $\mathcal{H}_{\text{Fock}} = \ell^2(\mathbb{N})$, spanned by the number states $\{|n\rangle\}_{n \in \mathbb{Z}_0}$. The joint Hilbert space of the q -mode system is

$$\mathcal{H}_q \triangleq \bigotimes_{i \in [q]} \mathcal{H}_{\text{Fock}} = \text{Span} \{ |\underline{n}\rangle = |n_0\rangle \otimes \cdots \otimes |n_{q-1}\rangle : \underline{n} \in \mathbb{Z}_0^q \}.$$

The total excitation of a number state $|\underline{n}\rangle \in \mathcal{H}_q$ is given by $N = \sum_{i=0}^{q-1} n_i$. Note that q -mode number states $|\underline{n}\rangle$ with excitation N are in one-to-one correspondence with vertices of the simplex $\mathcal{S}_{q,N}$, a fact that was used in the constructions of [2].

We focus on q -mode bosonic Fock state codes with constant excitation. As we show in Section 3.1, the restriction to constant total excitation both renders the encoding space finite-dimensional and imposes algebraic structure on the action of physically relevant noise operators, which can be exploited in the analysis of quantum error correction. Specifically, a code $Q \subset \mathcal{H}_q$ of dimension K and constant excitation N is a linear subspace of the space

$$\mathcal{H}_{q,N} \triangleq \text{Span} \{ |\underline{n}\rangle : \underline{n} \in \mathcal{S}_{q,N} \} \subset \mathcal{H}_q,$$

spanned by the basis states

$$|c_i\rangle = \sum_{\underline{n} \in \mathcal{S}_{q,N}} \alpha_{\underline{n}}^{(i)} \cdot |\underline{n}\rangle, \quad i = 0, \dots, K-1.$$

The *rate* of a K -dimensional quantum code Q is defined as

$$R(Q) \triangleq \frac{\log_2 K}{\log_2 \dim(\mathcal{H}_{q,N})} = \frac{\log_2 K}{\log_2 \binom{N+q-1}{q-1}}.$$

where in the second equality we assumed constant excitation.

Although states in $\mathcal{H}_{q,N}$ have constant total excitation, this condition alone does not preclude the possibility that a large fraction of photons is concentrated in a single mode. To rule out such highly unbalanced patterns, we will limit ourselves to codewords with bounded per-mode occupancy, assuming that every mode cannot host more than a fixed maximum number of photons. This is a simple technical constraint that keeps the code space inside a low-occupation sector and avoids pathological concentration on a few modes. Similar local bounded-occupancy assumptions appear naturally in bosonic encodings used for optical loss tolerance [45], where logical information is deliberately kept within the vacuum/single-photon sector of each mode. This assumption was earlier used in [20], which constructed families of single-error-correcting codes in this regime.

Definition 2. Let N, q , and $B \leq N$ be positive integers. A state in the space $\mathcal{H}_{q,N}^{(B)}$,

$$\mathcal{H}_{q,N}^{(B)} \triangleq \text{Span} \{ |\underline{n}\rangle : \underline{n} \in \mathcal{S}_{q,N}, \|\underline{n}\|_\infty \leq B \},$$

is said to have B -local excitation. A Fock state code $Q \subseteq \mathcal{H}_{q,N}^{(B)}$ is said to have B -local excitation.

While this definition implies a deterministic per-mode excitation bound for each codeword, it is also natural to consider relaxed formulations in which local excitation is controlled only in a probabilistic sense with respect to measurement outcomes on a given codeword. Rather than requiring a uniform bound on the maximal photon number per mode, one may instead require that the probability of observing excitation above a prescribed threshold B under photon-number measurement is small for any input codeword. Such a relaxation allows states whose amplitudes are not strictly confined to a bounded-occupancy sector, while still suppressing most of the time highly concentrated excitation patterns. From this perspective, approximate local boundedness may provide comparable protection against pathological high-excitation configurations, while allowing a larger collection of admissible codes. Accordingly, we give the following definition.

Definition 3. Let N, q , and $B \leq N$ be positive integers, and let Π_B be the orthogonal projection on $\mathcal{H}_{q,N}^{(B)}$. A Fock state code $Q \subseteq \mathcal{H}_{q,N}$ is said to have an ε -approximate B -local excitation if for any input codeword, the outcome of the measurement $\{M_0 = \Pi_B, M_1 = I - \Pi_B\}$ gives 0 with probability at least $1 - \varepsilon$, i.e.,

$$\inf_{|\psi_{\text{in}}\rangle \in Q} \langle \psi_{\text{in}} | \Pi_B | \psi_{\text{in}} \rangle \geq 1 - \varepsilon.$$

Let $B = (B_N)_N$ be a sequence of positive numbers. A sequence of Fock state codes $(Q_N)_N$, $Q_N \subseteq \mathcal{H}_{q,N}$, is said to have approximate asymptotically B -local excitation if

$$\inf_{|\psi_{\text{in}}\rangle \in Q_N} \langle \psi_{\text{in}} | \Pi_{B_N} | \psi_{\text{in}} \rangle \geq 1 - o(1).$$

3.1 The AD channel

In this section, we re-examine the formulation of error-correcting conditions for the AD channel accepted in the literature.

The dominant noise mechanism in bosonic platforms is AD, which models the loss of energy carriers, such as photons or phonons, from the system. The noise strength is characterized by a loss parameter $\gamma \in [0, 1]$, typically small, which is an intrinsic property of a single bosonic mode and quantifies its coupling to a loss channel. In multi-mode systems, we assume that γ remains a fixed constant for each mode and does not scale with the total number of modes. The AD channel on a q -mode bosonic space, $\mathcal{N} : \mathcal{H}_q \rightarrow \mathcal{H}_q$, is defined by the set of Kraus operators $\{A_{\underline{r}}\}_{\underline{r} \in \mathbb{Z}_0^q}$:

$$A_{\underline{r}} = A_{r_0} \otimes \cdots \otimes A_{r_{q-1}},$$

where A_{r_i} is the single-mode amplitude error corresponding to r_i photon losses, see (2). The action of A_{r_i} is described as follows:

$$|n\rangle \mapsto \sqrt{\binom{n}{r_i}} (1-\gamma)^{\frac{n-r_i}{2}} \gamma^{\frac{r_i}{2}} \cdot |n-r_i\rangle \mathbb{1}_{\{n \geq r_i\}}.$$

The error operator $A_{\underline{r}}$ with $\underline{r} \in \mathcal{S}_{q,r}$, $r \leq t$, is called an r -photon loss error.

Definition 4 (Code distance). A Fock state code Q is said to have distance $d(Q) = t + 1$ if it can correct any combination of up to t photon-loss events across all modes. In other words, Q is a quantum code that corrects the set of errors $\mathcal{E}_{\leq t} = \{A_{\underline{r}} : \underline{r} \in \bigcup_{r \leq t} \mathcal{S}_{q,r}\}$.

Definition 5 (Asymptotically good Fock state codes). A sequence of codes $\mathcal{Q} = (Q_N)_N$, $N \geq 1$, $Q_N \subseteq \mathcal{H}_{q_N, N}$ is called *asymptotically good* if the distance $d(Q_N)$ grows linearly with N and the rate stays positive as N increases:

$$\min \left(R(\mathcal{Q}), \liminf_{N \rightarrow \infty} \frac{d(Q_N)}{N} \right) > 0, \quad (12)$$

where $R(\mathcal{Q}) := \liminf_{N \rightarrow \infty} R(Q_N)$.

Early work on bosonic Fock state codes under AD noise focused on codes that correct up to a total of t photon-loss events. Such codes were introduced in [14, 32] and subsequently studied in numerous follow-up works, e.g., [20, 41]. A key motivation of this view has been provided in [14], which defined the “code fidelity” for constant-excitation Fock state codes under t AD errors:

$$\mathcal{F}_{\leq t}(Q) \triangleq \inf_{|\psi_{\text{in}}\rangle \in Q} \sum_{\substack{\underline{r} \in \mathcal{S}_{q,r} \\ r \leq t}} \langle \psi_{\text{in}} | A_{\underline{r}}^\dagger A_{\underline{r}} | \psi_{\text{in}} \rangle.$$

As shown in [14], once the excitation of the code is fixed, this quantity does not depend on a particular choice of the code that is capable of correcting t AD errors. This observation led the authors of [14] to suggest the code-fidelity as a universal measure of how well a code correcting t AD errors performs in the presence of the (non-truncated) AD channel (see Sec. 3.2 for a discussion of the truncation). Relying on the asymptotic behavior of $\mathcal{F}_{\leq t}$, it was suggested in

[14] that exact correction of t photon losses might suffice to achieve good approximate error correction against the full AD channel (with a potentially arbitrary number of losses) when γ is small.

A closer look at the code fidelity derivation leads us to a different interpretation of this quantity:

$\mathcal{F}_{\leq t}$ is a universal constant describing the entire constant-excitation subspace $\mathcal{S}_{q,N}$, and it does not depend on a specific code.

Thus, the code fidelity is the probability that at most t photons are lost under the AD process for any input state with excitation N , and it is independent of the choice of code within that subspace. We also show that a more delicate asymptotic analysis of this quantity is required to determine how well a t -error-correcting code actually performs when N increases and γ and t change as functions of N .

Proposition 3. *For any quantum state $|\psi\rangle \in \mathcal{H}_{q,N}$ we have*

$$\sum_{r \leq t} \sum_{\underline{r} \in \mathcal{S}_{q,r}} \langle \psi | A_{\underline{r}}^\dagger A_{\underline{r}} | \psi \rangle = \sum_{r=0}^t \binom{N}{r} \gamma^r (1-\gamma)^{N-r}.$$

The right-hand side of this relation is the probability that a binomial random variable is at most t . Below, we use the notation

$$p_{N,t} := \mathbb{P}[\text{Binomial}(N, \gamma) \leq t].$$

Although $p_{N,t}$ depends on γ , we omit it from the notation since it is fixed throughout all arguments.

Proof. Let $|\psi\rangle = \sum_{\underline{n} \in \mathcal{S}_{q,N}} \alpha_{\underline{n}} \cdot |\underline{n}\rangle$. As shown in [2, Eq. (11)], for $\underline{r} \in \mathcal{S}_{q,r}$ we have

$$\langle \psi | A_{\underline{r}}^\dagger A_{\underline{r}} | \psi \rangle = (1-\gamma)^{N-r} \gamma^r \binom{N}{r} \binom{r}{\underline{r}} \cdot \sum_{\underline{n} \in \mathcal{S}_{q,N}} |\alpha_{\underline{n}}|^2 \frac{\binom{N-r}{\underline{n}-\underline{r}}}{\binom{N}{\underline{n}}}.$$

Summing over all $\underline{r} \in \bigcup_{r \leq t} \mathcal{S}_{q,r}$ we obtain

$$\begin{aligned} \sum_{\substack{\underline{r} \in \mathcal{S}_{q,r} \\ r \leq t}} \langle \psi | A_{\underline{r}}^\dagger A_{\underline{r}} | \psi \rangle &= \sum_{r=0}^t \sum_{\underline{r} \in \mathcal{S}_{q,r}} \sum_{\underline{n} \in \mathcal{S}_{q,N}} |\alpha_{\underline{n}}|^2 (1-\gamma)^{N-r} \gamma^r \binom{N}{r} \binom{r}{\underline{r}} \frac{\binom{N-r}{\underline{n}-\underline{r}}}{\binom{N}{\underline{n}}} \\ &= \sum_{r=0}^t (1-\gamma)^{N-r} \gamma^r \sum_{\underline{n} \in \mathcal{S}_{q,N}} |\alpha_{\underline{n}}|^2 \sum_{\underline{r} \in \mathcal{S}_{q,r}} \prod_{i=0}^{q-1} \binom{n_i}{r_i} \\ &= \sum_{r=0}^t (1-\gamma)^{N-r} \gamma^r \binom{N}{r} \sum_{\underline{n} \in \mathcal{S}_{q,N}} |\alpha_{\underline{n}}|^2 \\ &= \sum_{r=0}^t (1-\gamma)^{N-r} \gamma^r \binom{N}{r}, \end{aligned}$$

where the second equality follows by a direct calculation ([2, Lemma A.1]) and the last equality follows since $\sum_{\underline{n} \in \mathcal{S}_{q,N}} |\alpha_{\underline{n}}|^2 = 1$ as $|\psi\rangle$ is a quantum state. $\square$

Relying on this proposition, we argue that the term “code fidelity” is not applicable because $\mathcal{F}_{\leq t}$ does not depend on the code and is simply a binomial probability, $p_{N,t}$, $t \leq N$. In [14, 55, 32, 41], $p_{N,t}$ was approximated as

$$p_{N,t} = 1 - \binom{N}{t+1} \gamma^{t+1} + O(\gamma^{t+2}).$$

This approximation, appropriate for smaller codes, becomes untenable for larger N and t . For instance, the single-error-correcting code with $N = 4$ photons in [14] has fidelity $\approx 1 - 6\gamma^2$, which is a good estimate for small γ . At the same time, the code correcting $t = 4$ errors with $N = 50$ photons presented in the same paper yields a fidelity bound $\approx 1 - 2 \cdot 10^6 \gamma^5$ which becomes vacuous for most values of $\gamma \in (0, 1)$.

Turning to asymptotics, we note that in the physically relevant scenario, γ is a small constant that does not vanish with N , and therefore the above approximation is vacuous unless t exceeds the expected number of photon losses $N\gamma$. Specifically, by the weak law of large numbers, $p_{N,t}$ changes from almost 0 to almost 1 when t increases within the interval $[(1-\varepsilon)\gamma N, (1+\varepsilon)\gamma N]$ for any $\varepsilon > 0$. This implies that codes that cannot support correcting a linear number of AD errors will perform poorly on the AD channel with loss parameter γ since the probability that the input state will be subjected to a correctable error pattern vanishes with the increase of N .

In the next section, we give an operational interpretation of $p_{N,t}$ by relating it to the *worst-case entanglement fidelity*, and show that in the asymptotic regime of $N \rightarrow \infty$, $t > (1+\varepsilon)\gamma N$ a code (approximate or exact) that corrects t AD errors, supports approximate correction of the AD noise with respect to the worst-case entanglement fidelity criterion.

3.2 The truncated AD channel and AQEC

As discussed above, given an AD channel, it is not clear what we mean by saying that a code “corrects t errors”. In this section, we provide a formal answer to this question.

Consider the space $\mathcal{H}_{q,N}$ of constant excitation q -mode Fock states and the operation $\mathcal{N}_{\leq t}$ on $L(\mathcal{H}_{q,N})$ defined by the normalized truncated AD channel:

$$\mathcal{N}_{\leq t}(\rho) = \frac{1}{p_{N,t}} \sum_{\substack{r \in \mathcal{S}_{q,r} \\ r \leq t}} A_r \rho A_r^\dagger. \quad (13)$$

Proposition 3 implies that when restricted to $\mathcal{H}_{q,N}$, $\mathcal{N}_{\leq t}$, it is trace preserving, and hence a CPTP map. This is formulated in the following proposition, whose formal proof appears in Appendix A.5.

Proposition 4. *The map $\mathcal{N}_{\leq t}$ is a quantum channel on $\mathcal{H}_{q,N}$.*

It has been observed since the early days of AQEC that the notion of distance, which plays a central role for exact quantum codes, does not admit a meaningful analog in the approximate setting. In particular, the authors of [15] showed that the equivalence between correcting arbitrary errors and erasures, on which the operational interpretation of distance relies, breaks down for approximate quantum error-correcting codes, and concluded that there is no appreciable notion of distance for AQEC. Proposition 4 nevertheless allows us to formulate

a notion of asymptotically good approximate Fock state codes by quantifying code performance directly through approximate correction guarantees and bypassing the concept of distance. We formalize this approach in the following definition.

Definition 6 (Asymptotically good approximate Fock state codes). A sequence of AQECCs $\mathcal{Q} = (Q_N)_N$, $Q_N \subseteq \mathcal{H}_{q_N, N}$, is called asymptotically good if there exists $\delta > \gamma > 0$ such that $Q_N, N \geq 1$ is an ε_N -AQECC for the channel $\mathcal{N}_{\leq \lfloor \delta N \rfloor}$, where $\varepsilon_N \downarrow 0$, and the sequence $\mathcal{Q}$ has a positive asymptotic rate, $R(\mathcal{Q}) > 0$.

We are now ready to give the main statement in this part, showing that the performance of a code correcting t errors on an AD channel is governed by $p_{N,t}$. We remark that while we formulate the statement for t -AQECCs, it also includes exact t -error correcting codes, which can be viewed as approximate ones with $\varepsilon = 0$.

Theorem 5. Assume that $Q \subseteq \mathcal{H}_{q,N}$ is an ε -AQECC for the truncated channel $\mathcal{N}_{\leq t}$, then Q is an ε' -AQECC for the AD channel $\mathcal{N}$, where

$$\varepsilon' = \sqrt{1 - (1 - \varepsilon^2)p_{N,t}}.$$

Conversely, if $Q \subseteq \mathcal{H}_{q,N}$ is an ε -AQECC for the AD channel $\mathcal{N}$, then it is also ε' -AQECC for the truncated channel $\mathcal{N}_{\leq t}$ with

$$\varepsilon' = \frac{\varepsilon}{\sqrt{p_{N,t}}}.$$

The proof is long and technical, so it is presented in the appendix. We note that this theorem underscores the operational significance of approximate and exact asymptotically good Fock state codes, as made explicit in the following corollary.

Corollary 6. A sequence of codes $\mathcal{Q} = (Q_N)_N$ is asymptotically good (in the approximate or exact sense) if and only if it is an AQECC code sequence for the AD channel with a sufficiently small (constant) loss parameter $\gamma > 0$.

Proof. Assume that $\mathcal{Q}$ is an AQECC sequence for the AD channel $\mathcal{N}$ with $\gamma_0 > 0$ such that $Q_N, N \geq 1$ is an ε_N -AQECC for $\mathcal{N}$. Suppose further that the limiting rate of $\mathcal{C}$, (12), is positive. Fix a number $\delta \in (0, 1)$ such that $\gamma_0 < \delta$. The weak law of large numbers implies that

$$p_{N, \lfloor \delta N \rfloor} = 1 - o(1),$$

and by Theorem 5, Q_N is ε' -AQECC for $\mathcal{N}_{\leq \lfloor \delta N \rfloor}$ and $\varepsilon'_N = \varepsilon_N / \sqrt{p_{N, \lfloor \delta N \rfloor}} = o(1)$, as desired.

Conversely, let $\mathcal{Q} = (Q_N)_N$ be an asymptotically good approximate sequence with respect to $\mathcal{N}_{\lfloor \delta N \rfloor}$, with $\delta > \gamma_0 > 0$ and let $\varepsilon_N \downarrow 0$. Similarly to the previous case, we have $p_{N, \lfloor \delta N \rfloor} = 1 - o(1)$, and therefore Q_N is AQECC for $\mathcal{N}$ with

$$\varepsilon'_N = \sqrt{1 - (1 - \varepsilon_N^2)p_{N,t}} = \sqrt{1 - (1 - o(1))(1 - o(1))} = o(1). \quad \square$$

Remark 2 (Average-case error criterion). Theorem 5 shows that the worst-case AQEC performance against the full AD channel can be controlled in terms of the performance against the truncated channel $\mathcal{N}_{\leq t}$ and the tail probability $1 - p_{N,t}$ of losing more than t photons.

The Bény–Oreshkov framework is then used to characterize the optimal worst-case recovery error for $\mathcal{N}_{\leq t}$. Together, these two ingredients yield the bounds on the performance of the code against the full, non-truncated AD channel derived in this subsection.

A parallel analysis can be carried out under the weaker, average-case criterion obtained by replacing the infimum over ρ in (4) with the fixed choice $\rho = P_Q/K$ (where P_Q is the projection on Q and $K = \dim(Q)$), or equivalently by evaluating the entanglement fidelity on a maximally entangled state. This is the usual channel-fidelity criterion. The same decomposition of the AD channel into its truncated and tail components yields an analogous relation between the full and truncated channels under the average-case criterion. In this setting, controlling the tail contribution is simpler, since there is no optimization over the input state. The optimal performance on the truncated channel can then be estimated using explicit near-optimal recovery maps, such as Tyson’s quadratic recovery [51] or the transpose/Petz recovery [5]. These results concern fixed-input entanglement fidelity and, for $\rho = P_Q/K$, specialize to channel fidelity; they therefore provide an average-case counterpart to, rather than an equivalent replacement for, the worst-case Bény–Oreshkov criterion employed here. We choose the stronger worst-case criterion in this work, which guarantees uniform performance over all encoded input states. It is nevertheless possible that relaxing to the average-case criterion could yield code families with improved parameters, in particular a better asymptotic rate–loss tradeoff.

4 Asymptotically good Fock state codes from ℓ_1 codes

We now turn to the construction of bosonic Fock state quantum codes derived from the classical ℓ_1 codes we introduced in Section 6. Our approach follows the general framework of [2], which is based on partitioning a classical code into disjoint subsets and associating a quantum basis codeword to each partition element via an appropriately chosen linear combination of the corresponding Fock states. We briefly review this construction and observe that the resulting quantum codes always satisfy the assumptions of Proposition 2; in particular, they admit reliable identification of the relevant error set.

Using this framework, we obtain three distinct families of asymptotically good Fock state codes derived from the three different ℓ_1 code constructions presented in Section 6. The Gilbert–Varshamov (GV) type ℓ_1 codes lead to Fock state codes achieving exact quantum error correction with the most favorable asymptotic rate–distance tradeoff among the constructions considered. However, this approach relies on computationally involved procedures. One of them is the need to construct a good ℓ_1 code sequence given in Proposition 15, and the other requires finding a partition of the code into blocks (groups of codewords) such that the resulting Fock state codes satisfy the error correction condition. The complexity of these tasks is at least exponential in the number of modes, rendering the construction infeasible for physically relevant system sizes.

To overcome this limitation, we turn to random ℓ_1 codes, which yield families of asymptotically good Fock state codes in the approximate error-correction regime. This construction avoids both computational barriers mentioned above and is therefore significantly simpler from an implementation standpoint. From an operational perspective, this relaxation incurs no essential loss: as established in Corollary 6, both the exact and approximate code families for the truncated amplitude-damping channel imply approximate quantum error correction for the full

(and physically more relevant) amplitude-damping channel. Consequently, the choice between these constructions reflects a tradeoff between implementation complexity and asymptotic distance–rate performance, rather than a fundamental degradation in error-correction capability.

4.1 The construction of [2]

We start with the definition of classical ℓ_1 codes on the simplex, which form the combinatorial backbone of our Fock state code constructions. Define a metric on $\mathcal{S}_{q,N}$, called the ℓ_1 distance:

$$d(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\|_1 = \frac{1}{2} \sum_{i=0}^{q-1} |x_i - y_i|.$$

A classical code $C \in \mathcal{S}_{q,N}$ is said to have ℓ_1 distance at least t , $d(C) \geq t$, if every pair of distinct vectors $\underline{x}, \underline{y} \in C$ satisfies $d(\underline{x}, \underline{y}) \geq t$. The study of ℓ_1 codes has a long history in classical coding theory, motivated by applications ranging from constrained storage systems and flash memories [4] to DNA storage [58, 59] and other communication problems [29, 19]. Constructions of ℓ_1 codes often are based on tools from additive number theory, such as the Bose–Chowla theorem and Sidon sets. Here we focus on random coding and geometric GV-type arguments to show the existence of ℓ_1 codes with favorable rate–distance tradeoffs beyond what is achievable by algebraic constructions.

A general approach to the construction of Fock state codes from classical ℓ_1 codes was suggested in [2].

Construction 1. Let $C_N \subseteq \mathcal{S}_{q,N}$ be a code with ℓ_1 distance $t + 1$. Given a partition of C_N into disjoint subsets $I_0, I_1, \dots, I_K$, consider the Fock state code Q_N with the basis $\{|c_0\rangle, \dots, |c_{K-1}\rangle\} \subset \mathcal{H}_{q,N}$ where

$$|c_i\rangle = \sum_{\underline{n} \in I_i} \alpha_{\underline{n}}^{(i)} |\underline{n}\rangle, \quad (14)$$

and $\{\alpha_{\underline{n}}^{(i)}\}_{\underline{n} \in I_i}$ is a set of coefficients such that $\sum_{\underline{n} \in I_i} |\alpha_{\underline{n}}^{(i)}|^2 = 1$.

We start by observing that Q_N always satisfies the KL orthogonality condition. The only remaining constraint to be addressed for constructing exact or approximate quantum error-correcting codes is the non-deformation condition. Although this fact is established as part of the proof of Theorem 7 in [2], we briefly outline the main idea and use it to show that such a code always admits error identification, in an even more general setting.

Proposition 7. *For $i \neq j$ and $\underline{r} \neq \underline{r}'$ such that $\|\underline{r}\|_1, \|\underline{r}'\|_1 \leq t$, the basis codewords defined in (14) satisfy*

$$\langle c_i | A_{\underline{r}}^\dagger A_{\underline{r}'} | c_j \rangle = 0. \quad (15)$$

Proof. First, note that $A_{\underline{r}}$ maps $|c_i\rangle$ to $\mathcal{H}_{q,N-r}$, and since $\mathcal{H}_{q,N-r}$ and $\mathcal{H}_{q,N-r'}$ are orthogonal to one another, we may assume that $\underline{r}, \underline{r}' \in \mathcal{S}_{q,r}$ for some $r < N$. A simple calculation (see [2,

pp. 8]) shows that

$$\langle c_i | A_{\underline{r}}^\dagger A_{\underline{r}'} | c_j \rangle = \sum_{\substack{\underline{n} \in I_i \\ \underline{n}' \in I_j}} \gamma^r (1 - \gamma)^{N-r} \binom{N}{r} \sqrt{\frac{\binom{r}{\underline{r}} \binom{r'}{\underline{r}'} \binom{N-r}{\underline{n}-\underline{r}} \binom{N-r}{\underline{n}'-\underline{r}'}}{\binom{N}{\underline{n}} \binom{N}{\underline{n}'}}} (\alpha_{\underline{n}}^{(i)})^* \alpha_{\underline{n}'}^{(j)} \cdot \langle \underline{n} - \underline{r} | \underline{n}' - \underline{r}' \rangle.$$

Suppose that $\langle \underline{n} - \underline{r} | \underline{n}' - \underline{r}' \rangle \neq 0$ for some $\underline{n}, \underline{n}', \underline{r}, \underline{r}'$, which then should satisfy $\underline{n} - \underline{r} = \underline{n}' - \underline{r}'$. First, suppose that $i \neq j$, then $I_i \cap I_j = \emptyset$ by assumption. This implies that $\underline{n} \neq \underline{n}'$ and therefore, $d(\underline{n}, \underline{n}') \geq t$ by the assumption on the distance of C_N . On the other hand, $\|\underline{r} - \underline{r}'\|_1 \leq 2r$, implying that $d(\underline{n} - \underline{r}, \underline{n}' - \underline{r}') \geq 1$, so the equality cannot hold. Now let $i = j$, in which case $\underline{n}$ and $\underline{n}'$ can be distinct or coincide. If they are distinct, then the equality cannot be true by the same argument as above. If $\underline{n} = \underline{n}'$, then $\underline{n} - \underline{r} \neq \underline{n}' - \underline{r}'$ by our assumption of $\underline{r} \neq \underline{r}'$. We conclude that $\langle \underline{n} - \underline{r} | \underline{n}' - \underline{r}' \rangle = 0$ for each term in the sum, and so (15) holds. $\square$

Below, we argue that the codes Q_N can identify up to t errors when used on the AD channel. To formalize this claim, we recall that with probability $p_{N,t}$, the action of the channel is described by one of the operators A_r and that with probability $1 - p_{N,t}$ the input state $|\psi_{\text{in}}\rangle$ is subjected to the loss of $> t$ photons. An error resulting in a loss of at most t photons can be identified if there exists a projective measurement $\{\Pi_{\underline{r}} : \underline{r} \in \mathcal{S}_{q,r}, r \leq t\} \cup \{\Pi_{>t}\}$ such that once $|\psi_{\text{in}}\rangle$ is submitted to the channel, the outcome $\underline{r}$ is obtained if the channel acts as $A_{\underline{r}}$, while the outcome $> t$ is obtained with probability $1 - p_{N,t}$. Furthermore, this measurement remains valid even if $I_0, \dots, I_K$ are not disjoint, meaning that the code Q_N is any subspace spanned by Fock states corresponding to classical codewords from C_N . That is formulated in Corollary 8, whose proof appears in Appendix A.5.

Corollary 8. *The code $Q_N = \langle |c_i\rangle, i \in [K] \rangle$ can identify all errors caused by at most t photon losses.*

Remark 3. Quantum error identification, considered here, stops short of correcting the errors, but adds functionality compared to other noise-resistant primitives, such as, for instance, error detection, which plays an essential role in fault tolerance. For instance, in post-selected fault-tolerant architectures and in magic-state distillation protocols, error-detecting codes are used to reject faulty computational branches or ancilla states rather than to actively correct them, enabling arbitrarily low logical error rates through filtering and purification instead of recovery [27, 13]. Error identification goes further by providing classical information about the specific error instance. This additional information enables adaptive feedback and real-time correction strategies, such as photon-loss tracking in bosonic codes [37], and underlies decoding procedures in topological codes where syndrome patterns are mapped to specific error configurations before correction [17].

4.2 Asymptotically good Fock state codes

In this section, we construct two families of asymptotically good approximate Fock state codes using the random ℓ_1 codes from Section 3 within the framework of Construction 1. While orthogonality follows from the ℓ_1 distance guarantees, randomness is used to ensure the non-deformation KL conditions, which are shown to hold with high probability under a

suitable averaging over partitions. We show that the obtained Fock state codes inherit the local boundedness properties from the corresponding random classical codes, which guarantees the approximately bounded local excitation, cf. Definition 3.

We will construct Fock state codes from classical ℓ_1 codes as follows. Let $C \subset \mathcal{S}_{\alpha N, N}$ be a code of size $L = TK$, where K is the target dimension of the quantum code¹. Let $I_0, \dots, I_{K-1}$ be an arbitrary partition of $C = \{\underline{n}^{(0)}, \dots, \underline{n}^{(L-1)}\}$ into equal-size parts. Consider the Fock state code Q_N defined in (14), where the coefficients of the basis states are given by

$$\alpha_{\underline{n}}^{(j)} = \mathbb{1}_{I_j}(\underline{n}) \frac{1}{\sqrt{T}}, \quad j \in [K], \quad (16)$$

where $\mathbb{1}_{I_j}$ is the indicator of the set I_j .

Proposition 9. Fix $\alpha > 0$, $\gamma \in (0, 1]$, and $0 < \delta < 1$. Let $C_N = \{\underline{n}^{(i)}\}_{i=0}^L$ be an ℓ_1 code, whose codewords are drawn independently from $\mathcal{S}_{\alpha N, N}$ according to some distribution $\mathbf{D}_N$ on the simplex. Suppose that $\frac{\ln N}{\ln L} \rightarrow 0$ as $N \rightarrow \infty$. Suppose further that the sequence of Fock state codes $(Q_N)_N$ constructed as in (16) is used on the truncated AD channel $\mathcal{N}_{\leq \lfloor \delta N \rfloor}$ defined in (13) by the Kraus operators $\tilde{A}_{\underline{r}} = A_{\underline{r}}/p_{N, \lfloor \delta N \rfloor}$, $\underline{r} \in \mathcal{S}_{\alpha N, r}$, $r \leq \lfloor \delta N \rfloor$. Let M be the number of Kraus operators of this channel².

There is a choice of non-negative constants $\lambda_{\underline{r}}$, $\underline{r} \in \mathcal{S}_{\alpha N, r}$, $r \leq \lfloor \delta N \rfloor$ that sum to 1 (they may depend on N) and an $o(1)$ function, such that with probability approaching 1 as N increases, the sequence $(C_N)_N$ satisfies

$$\max \left\{ \left| \langle c_i | \tilde{A}_{\underline{r}}^\dagger \tilde{A}_{\underline{r}} | c_i \rangle - \lambda_{\underline{r}} \right| : i \in [K], \underline{r} \in \mathcal{S}_{\alpha N, r}, r \leq \lfloor \delta N \rfloor \right\} = o\left(\frac{1}{KM}\right) \quad (17)$$

as long as

$$K^3 M^2 \leq L^{1-\varepsilon}, \quad (18)$$

for some $\varepsilon > 0$.

Proof. Denote $t = \lfloor \delta N \rfloor$. Our goal is to find non-negative numbers $\{\lambda_{\underline{r}}\}_{\underline{r} \in \mathcal{S}_{\alpha N, r}, r \leq t}$ that sum to 1 and $\xi_N > 0$ such that

$$\mathbb{P} \left[\max_{\underline{r} \in \cup_{r \leq t} \mathcal{S}_{\alpha N, r}} \max_{j=0, \dots, K-1} \left| \langle c_j | \tilde{A}_{\underline{r}}^\dagger \tilde{A}_{\underline{r}} | c_j \rangle - \lambda_{\underline{r}} \right| > \xi_N \right] \xrightarrow{N \rightarrow \infty} 0,$$

where $\xi_N = o\left(\frac{1}{KM}\right)$. To this end, for every $\underline{n} \in \mathcal{S}_{\alpha N, N}$ and $\underline{r} \in \mathcal{S}_{\alpha N, r}$ define:

$$Y_{\underline{r}}(\underline{n}) := \frac{1}{p_{N, t}} (1 - \gamma)^{N-r} \gamma^r \cdot \prod_{i=0}^{q-1} \binom{n_i}{r_i} \leq \frac{1}{p_{N, t}} (1 - \gamma)^{N-r} \gamma^r \binom{N}{r}.$$

Recall that $p_{N, t}$ is the probability that a random variable $Z \sim \text{Bin}(N, \gamma)$ is bounded by t . Since $r \leq t$, we find that

$$Y_{\underline{r}}(\underline{n}) \leq \mathbb{P}[Z = r | Z \leq t] < 1. \quad (19)$$

¹if $K \nmid L$, we can discard $L \bmod K$ points of the code to satisfy the divisibility condition.

²We again suppress the dependence of the parameters K, L , and M on N .

for all $\underline{n} \in \mathcal{S}_{\alpha N, N}$ and $\underline{r} \in \mathcal{S}_{\alpha N, r}$.

As in Prop. 3 above, we use [2, Eq.(11)] to claim that for $\underline{r} \in \mathcal{S}_{N, r}$, we have

$$\begin{aligned}
\langle c_j | \tilde{A}_{\underline{r}}^\dagger \tilde{A}_{\underline{r}} | c_j \rangle &= \frac{1}{p_{N,t}} (1 - \gamma)^{N-r} \gamma^r \binom{N}{r} \binom{r}{\underline{r}} \sum_{\underline{n} \in I_j} |\alpha_{\underline{n}}^{(j)}|^2 \frac{\binom{N-r}{\underline{n}-\underline{r}}}{\binom{N}{\underline{n}}} \\
&= \frac{1}{p_{N,t}} (1 - \gamma)^{N-r} \gamma^r \frac{N!}{r!(N-r)!} \frac{r!}{\prod_{i=0}^{q-1} r_i} \sum_{\underline{n} \in I_j} \frac{1}{T} \frac{(N-r)!}{N!} \prod_{i=0}^{q-1} \frac{\mathbf{n}_i!}{(\mathbf{n}_i - r_i)!} \\
&= \frac{1}{p_{N,t}} (1 - \gamma)^{N-r} \gamma^r \sum_{\underline{n} \in I_j} \frac{1}{T} \prod_{i=0}^{q-1} \binom{\mathbf{n}_i}{r_i} \\
&= \frac{1}{T} \sum_{\underline{n} \in I_j} Y_{\underline{r}}(\underline{n}).
\end{aligned}$$

Define

$$\lambda_{\underline{r}} = \frac{1}{p_{N,t}} (1 - \gamma)^{N-r} \gamma^r \mathbb{E}_{\underline{X} \sim \mathbf{D}_N} \left[\prod_{i=0}^{q-1} \binom{\mathbf{X}_i}{r_i} \right] \geq 0.$$

By Hoeffding's inequality (50) and (19), for any $\xi > 0$ we have

$$\mathbb{P} \left[\left| \frac{1}{T} \sum_{\underline{n} \in I_j} Y_{\underline{r}}(\underline{n}) - \lambda_{\underline{r}} \right| > \xi \right] \leq 2 \exp(-2T\xi^2). \quad (20)$$

Setting $\xi_N = \frac{1}{L^{\varepsilon/4} K M}$ and recalling that $L = TK$, we obtain

$$\begin{aligned}
\mathbb{P} \left[\left| \langle c_j | \tilde{A}_{\underline{r}}^\dagger \tilde{A}_{\underline{r}} | c_j \rangle - \lambda_{\underline{r}} \right| > \xi_N \right] &= \mathbb{P} \left[\left| \frac{1}{T} \sum_{\underline{n} \in I_j} Y_{\underline{r}}(\underline{n}) - \lambda_{\underline{r}} \right| > \frac{1}{L^{\varepsilon/4} K M} \right] \\
&\leq 2 \exp \left(-\frac{2L}{L^{\varepsilon/2} K^3 M^2} \right) \\
&\leq \exp(-L^{\varepsilon/2 - o(1)}), \quad (21)
\end{aligned}$$

where the last equality follows from the assumption in (18). With this, we finish the proof as follows:

$$\begin{aligned}
&\mathbb{P} \left[\max_{\underline{r} \in \cup_{r \leq t} \mathcal{S}_{N,r}} \max_{j=0, \dots, K-1} \left| \langle c_j | \tilde{A}_{\underline{r}}^\dagger \tilde{A}_{\underline{r}} | c_j \rangle - \lambda_{\underline{r}} \right| > \xi_N \right] \\
&= \mathbb{P} \left[\bigcup_{\underline{r} \in \cup_{r \leq t} \mathcal{S}_{N,r}} \bigcup_{j=0, \dots, K-1} \left\{ \left| \langle c_j | \tilde{A}_{\underline{r}}^\dagger \tilde{A}_{\underline{r}} | c_j \rangle - \lambda_{\underline{r}} \right| > \xi_N \right\} \right] \\
&\leq \sum_{\underline{r} \in \cup_{r \leq t} \mathcal{S}_{N,r}} \sum_{j=0}^{K-1} \cdot \mathbb{P} \left[\left| \langle c_j | \tilde{A}_{\underline{r}}^\dagger \tilde{A}_{\underline{r}} | c_j \rangle - \lambda_{\underline{r}} \right| > \xi_N \right] \\
&\leq \left| \bigcup_{r \leq t} \mathcal{S}_{N,t} \right| K \cdot e^{-L^{\varepsilon/2 + o(1)}} \\
&\leq 2^{4N} \cdot L \cdot e^{-L^{\varepsilon/2 + o(1)}} \\
&\leq e^{-L^{\varepsilon/2 - o(1)}},
\end{aligned}$$

which goes to 0 as N increases. On the last line, we used the limiting condition $N/L^\epsilon \rightarrow 0$, implied by our assumption of $\frac{\ln N}{\ln L} \rightarrow 0$.

To complete the proof, it remains to show that indeed $\sum_{\underline{r}} \lambda_{\underline{r}} = 1$. Indeed,

$$\begin{aligned} \sum_{r \leq t} \sum_{\underline{r} \in \mathcal{S}_{\alpha N, r}} \lambda_{\underline{r}} &= \sum_{r \leq t} \sum_{\underline{r} \in \mathcal{S}_{\alpha N, r}} \frac{1}{p_{N, t}} (1 - \gamma)^{N-r} \gamma^r \mathbb{E}_{\underline{X} \sim \mathbf{D}_N} \left[\prod_{i=0}^{q-1} \binom{X_i}{r_i} \right] \\ &= \sum_{r \leq t} \frac{1}{p_{N, t}} (1 - \gamma)^{N-r} \gamma^r \mathbb{E}_{\underline{X} \sim \mathbf{D}_N} \left[\sum_{\underline{r} \in \mathcal{S}_{\alpha N, r}} \prod_{i=0}^{q-1} \binom{X_i}{r_i} \right] \\ &= \sum_{r \leq t} \frac{1}{p_{N, t}} (1 - \gamma)^{N-r} \gamma^r \mathbb{E}_{\underline{X} \sim \mathbf{D}_N} \left[\binom{\sum_{i=0}^{q-1} X_i}{r} \right] \\ &= \sum_{r \leq t} \frac{1}{p_{N, t}} (1 - \gamma)^{N-r} \gamma^r \binom{N}{r} = 1, \end{aligned}$$

where the equality at the last line follows since under the distribution $\mathbf{D}_N$, with probability 1, we have $\underline{X} \in \mathcal{S}_{\alpha N, N}$. $\square$

We are now ready to state the main result of this section, which gives the parameters of asymptotically good AQECCs that are obtained from our construction.

Theorem 10. Fix $\alpha > 0$, $0 < \gamma, \delta < 1$ and $\epsilon > 0$. Let $C_N = \{\underline{n}^{(i)}\}_{i=0}^L$ be an ℓ_1 code, whose codewords are drawn independently from $\mathcal{S}_{\alpha N, N}$ according to some distribution $\mathbf{D}_N$ on the simplex. Suppose that the sequence of K_N -dimensional Fock state codes $(Q_N)_N$ constructed as in (16) is used on the truncated AD channel $\mathcal{N}_{\leq \lfloor \alpha N \rfloor}$ defined in (13). If

$$\begin{aligned} 0 &< \limsup_{N \rightarrow \infty} \frac{\log_2 L}{N} = \bar{R}(\delta) - o(1) \\ \frac{\log_2 K_N}{N} &< \frac{1}{3} \bar{R}(\delta) - \frac{2}{3} (1 + \delta) h_2 \left(\frac{1}{1 + \delta} \right) - \epsilon, \end{aligned} \quad (22)$$

then with probability $1 - o(1)$ the codes $(Q_N)_N$ are AQECCs for the channel sequence $(\mathcal{N}_{\leq \lfloor \alpha N \rfloor})_N$ and they also have approximate asymptotically B -local excitation (cf. Def. 3). Here

$$(\bar{R}(\delta), B) = \begin{cases} (R_U(\delta), (1 + \epsilon) \log_{1+\alpha} N) & \text{if } \mathbf{D}_N = \text{Unif}(\mathcal{S}_{\alpha N, N}) \\ (R_M(\delta), (1 + \epsilon) \frac{\ln N}{\ln \ln N}) & \text{if } \mathbf{D}_N = \text{Multinomial}(\lfloor \alpha N \rfloor, N). \end{cases}$$

For instance, for $\alpha = 5$, the rate of the quantum codes constructed in this theorem is positive for $\delta \leq .15$ for the uniform ensemble and $\delta \leq 0.05$ for the multinomial ensemble.

Proof. Throughout this proof, we write K instead of K_N . Both cases of the claim follow the same argument, so we only verify the case of the uniform distribution. The number of Kraus operators of the channel $\mathcal{N}_{\leq \lfloor \delta N \rfloor}$ is given by

$$M = \sum_{i=0}^t |\mathcal{S}_{N, r}| = \sum_{i=0}^t \binom{r + N - 1}{N - 1} \leq 2^{N(1+\delta)(h_2(\frac{1}{1+\delta}) + o(1))}, \quad (23)$$

Coupled with (22), we conclude that

$$K^3 \leq \frac{L^{1-\varepsilon}}{M^2}, \quad \frac{\ln N}{\ln L} = o(1). \quad (24)$$

Thus, by Propositions 16, 9, and 7, with probability $1 - o(1)$, the Fock state code sequence $(Q_N)_N$ and the underlying random ℓ_1 code sequence $(C_N)_N$ have the following properties:

1. The ℓ_1 distance of C_N is at least $\lfloor \delta N \rfloor$, and therefore the corresponding Fock state code Q_N satisfies (15).
2. There are constants $\{\lambda_{\underline{L}}\}$ such that the approximate non-deformation condition (17) holds.

Using Proposition 2, we conclude with probability $1 - o(1)$, the sequence $(Q_N)_N$ is ε_N -AQECC for the channels $\mathcal{N}_{\leq \lfloor \delta N \rfloor}$ for some $\varepsilon_N \downarrow 0$.

It remains to prove the claim about the bounded local excitation. Let $|\psi_{\text{in}}\rangle \in Q_N$ be a codeword and let Π be the orthogonal projection on $\mathcal{H}_{\lfloor \alpha N \rfloor, N}^{(B_N)}$, where $B_N = (1 + \varepsilon) \log_{1+\alpha}(N)$. Let us write $|\psi_{\text{in}}\rangle = \sum_{i=0}^{K-1} \beta_i |c_i\rangle$, where $(|c_i\rangle)$ is the basis defined in (14), (16). A simple calculation shows that

$$\langle \psi_{\text{in}} | \Pi | \psi_{\text{in}} \rangle = \sum_{i,j=0}^{K-1} \beta_i^* \beta_j \langle c_i | \Pi | c_j \rangle = \sum_{i,j=0}^{K-1} \frac{K}{L} \cdot \beta_i^* \beta_j \sum_{\underline{n} \in I_i} \sum_{\underline{n}' \in I_j} \langle \underline{n} | \Pi | \underline{n}' \rangle. \quad (25)$$

To estimate this expression, observe that for any $\underline{n}, \underline{n}' \in \mathcal{S}_{\lfloor \alpha N \rfloor, N}$, we have $\langle \underline{n} | \Pi | \underline{n}' \rangle = \delta_{\underline{n}, \underline{n}'} \mathbb{1}_{\{\|\underline{n}\|_{\infty} \leq B_N\}}$, and therefore,

$$\begin{aligned} \langle \psi_{\text{in}} | \Pi | \psi_{\text{in}} \rangle &= \sum_{i,j=0}^{K-1} \frac{K}{L} \beta_i^* \beta_j \sum_{\underline{n} \in I_i} \sum_{\underline{n}' \in I_j} \delta_{\underline{n}, \underline{n}'} \mathbb{1}_{\{\|\underline{n}\|_{\infty} \leq B_N\}} \\ &= \sum_{i=0}^{K-1} \frac{K}{L} |\beta_i|^2 |\{\underline{n} \in I_i : \|\underline{n}\|_{\infty} \leq B_N\}| \\ &= \sum_{i=0}^{K-1} |\beta_i|^2 \left(1 - \frac{K}{L} \cdot |\{\underline{n} \in I_i : \|\underline{n}\|_{\infty} > B_N\}| \right) \end{aligned} \quad (26)$$

$$\geq 1 - \max_{i \in [K]} \frac{K}{L} \cdot |\{\underline{n} \in I_i : \|\underline{n}\|_{\infty} > B_N\}|, \quad (27)$$

where we used the fact that $I_i \cap I_j = \emptyset$ and that $\sum_i |\beta_i|^2 = 1$. Recall that I_i is a partition block formed of $T = L/K$ i.i.d. random vectors distributed uniformly on the simplex. Let Y_i be the random variable that counts the fraction of ‘bad’ words in I_i :

$$Y_i = \frac{1}{T} \cdot |\{\underline{n} \in I_i : \|\underline{n}\|_{\infty} > B_N\}|.$$

By (27), to show that $\min_{|\psi_{\text{in}}\rangle} \langle \psi_{\text{in}} | \Pi | \psi_{\text{in}} \rangle = 1 - o(1)$ with high probability, it is sufficient to show that

$$\mathbb{P} \left[\max_{i \in [K]} Y_i \leq g(N) \right] \xrightarrow{N \rightarrow \infty} 1, \quad (28)$$

where $g(N) \rightarrow 0$. To that end, let p_N be the probability that a uniform vector on the simplex $\underline{X} \sim \text{Unif}(\mathcal{S}_{\alpha N, N})$ satisfies $\|\underline{X}\|_\infty > B_N = (1 + \varepsilon) \log_{1+\alpha}(N)$. By (34) and the definition of the random ℓ_1 codeword $\underline{n} \in I_i$

$$p_N := \mathbb{E}[\mathbb{1}_{\{\|\underline{n}\|_\infty > B_N\}}] = \mathbb{P}[\|\underline{X}\|_\infty > B_N] = O(N^{\frac{\varepsilon^2 - \varepsilon}{2}}).$$

For some $0 < \xi < \frac{1}{2}$, let $g(N) = p_N + T^{-\xi} = o(1)$. Using Hoeffding's inequality (50), we have that

$$\begin{aligned} \mathbb{P}[Y_i > g(N)] &= \mathbb{P}\left[\frac{1}{T} \cdot \sum_{\underline{n} \in I_i} \left(\mathbb{1}_{\{\|\underline{n}\|_\infty > B_N\}} - \mathbb{E}[\mathbb{1}_{\{\|\underline{n}\|_\infty < B_N\}}]\right) > T^{-\xi}\right] \\ &\leq 2 \exp\left(-\frac{2T^{2-2\xi}}{T}\right) = 2e^{-2T^{1-2\xi}}. \end{aligned}$$

We are now ready to make the final step:

$$\begin{aligned} \mathbb{P}[\max_{i \in [K]} Y_i > g(N)] &\leq \sum_{i \in [K]} \mathbb{P}[Y_i > g(N)] \leq 2Ke^{-2T^{1-2\xi}} \\ &= 2 \exp\left(-\left(\frac{L}{K}\right)^{1-2\xi} + \log K\right). \end{aligned}$$

Rewriting (24), we obtain $K < \frac{L^{1/3-\varepsilon}}{M^{2/3}} < L^{1/3}$. Using this on the previous line,

$$\mathbb{P}[\max_{i \in [K]} Y_i > g(N)] \leq 2 \exp\left(-L^{\left(\frac{2}{3}(1-2\xi)\right)} + \frac{1}{3} \log L\right).$$

Since the right-hand side vanishes as $N \rightarrow \infty$, we have proved (28) and therefore, the theorem. $\square$

Remark 4. [Discussion] The codes defined in (14), (16), which we employ in this section, rely on an arbitrary partition of the classical code into equal-sized blocks, followed by averaging of the associated quantum states within each subset, yielding approximate non-deformation KL conditions. In [2] we used partitions arising from Tverberg's theorem to prove validity of the KL conditions (cf. Theorem 11). Here, condition (18) replaces the use of Tverberg-type arguments, supporting exact non-deformation conditions.

There is also another relaxation of the convex-geometric arguments, known as the no-dimensional (approximate) Tverberg theorem [1], which asserts that any set of L points contained in the Euclidean ball of diameter D can be partitioned into K subsets whose convex hulls all intersect a common ball of radius $O(\sqrt{K/L})D$. Interestingly, this non-constructive geometric argument yields exactly the same scaling relation between K, L , and M as (18). It is also known that there exists a procedure that finds such partitions as well as the common ball in time that is polynomial with respect to the number of points and the ambient dimension.

However, in our setting, this geometric approach has two important drawbacks. First, when instantiated for Fock state codes, the ambient Euclidean dimension corresponds to the number of modes, which grows exponentially with the number of physical qudits. As a result, even algorithms that are polynomial in the dimension become exponential when the

complexity is measured by the relevant physical parameters. Second, the resulting convex combinations are not guaranteed to preserve the required local boundedness (per-mode excitation) constraints, which are essential for the implementation of the code. For these reasons, even though both approaches apply under the same set of assumptions on the parameters, we adopt the uniform partition and averaging construction.

4.3 De-randomized asymptotically good Fock state codes

In this section, we present a deterministic construction of asymptotically good Fock-state codes, obtained by derandomizing the uniform random ℓ_1 simplex code via a Gilbert–Varshamov-type argument (see Section 6.1) and then reverting to the formalism of [2]. Although this approach entails high complexity of construction similar to [2], we include it to show that, at least theoretically, a more favorable error-correction/distance tradeoff is possible. As discussed above, the improvement of parameters over the randomized approach obtained here stems from the fact that Tverberg partitions result in less stringent constraints on the code dimension than those arising from the more general concentration arguments used in the random construction.

Theorem 11. ([2, Theorem 7]) *Let $C_N \subseteq \mathcal{S}_{q,N}$ be an ℓ_1 code with distance at least $t + 1$. If*

$$|C_N| \geq (K - 1) \binom{t + q - 1}{t - 1} + 1.$$

then there exist a partition $C_N = \sqcup_{i=0}^{K-1} I_i$ and a choice of coefficients $\{\alpha_n^{(i)}\}$ such that the corresponding Fock state code from Construction 1 has dimension K and distance $t + 1$.

We use Theorem 11 and the GV-type bounds presented in Section 6 to construct exact asymptotically good Fock state codes with the additional property of logarithmic per-mode excitation, inherited from the structure of the classical codes that underlie the construction.

Theorem 12 (Exact asymptotically good Fock state codes). *For any $0 < \delta < 1$, $\varepsilon > 0$ and $\alpha > 0$ there exists a sequence of Fock state codes $(Q_N)_N$, $Q_N \subseteq \mathcal{H}_{\alpha N, N}$ with distance at least $\lfloor \delta N \rfloor$, local excitation $B = (1 + \varepsilon) \log_{1+\alpha} N$ and dimension K_N as long as*

$$\frac{1}{N} \log_2 K_N \leq R_{\text{GV}}(\delta, \alpha) - (\alpha + \delta) h_2 \left(\frac{\alpha}{\alpha + \delta} \right) + o(1), \quad (29)$$

for some fixed $o(1)$ function.

Proof. Consider the sequence of GV-type codes C_N constructed in Proposition 15. By Theorem 11, there exists a partition $I_0, \dots, I_{K-1}$ of C_N and a choice of coefficients $\{\alpha_n^{(i)}\}$ such that the corresponding Fock state code obtained from Construction 1 has dimension K and distance $\lfloor \delta N \rfloor + 1$ given that

$$K_N < |C_N| \binom{\lfloor \delta N \rfloor + \lfloor \alpha N \rfloor - 1}{\lfloor \alpha N \rfloor - 1}^{-1}.$$

For $N \rightarrow \infty$,

$$\begin{aligned} \log_2 K_N &< \log_2 |C_N| - \log_2 \binom{\lfloor \delta N \rfloor + \lfloor \alpha N \rfloor - 1}{\lfloor \alpha N \rfloor - 1} \\ &= N(R_{\text{GV}}(\delta, \alpha) - (\alpha + \delta)h_2\left(\frac{\alpha}{\alpha + \delta}\right) + o(1)), \end{aligned}$$

where the last equality follows from Proposition 15 and (51). Finally, the fact that Q_N has local excitation $B = (1 + \varepsilon) \log_{1+\alpha} N$ follows immediately since for any classical codeword $\underline{n}$, it is guaranteed by Proposition 15 that $\|\underline{n}\|_\infty \leq B$. $\square$

This theorem implies the existence of codes of positive rate for any $\delta > 0$ such that the right-hand side of (29) is positive. To argue that this region is non-vacuous, note that it is positive for any $\alpha > 0$ and $\delta = 0$, and so there exists a neighborhood of $\delta = 0$ for which there exist infinite sequences of codes of distance δN and strictly positive rate.

5 Asymptotically good PI and spin codes

5.1 PI codes

Our construction of asymptotically good Fock state codes implies existence of another asymptotically good family of quantum codes called *permutation-invariant codes*. PI codes, which lie in the symmetric subspaces of the Hilbert space, were introduced in the work of Ruskai [46, 44] with the motivation of constructing quantum codes that protect against particle exchange errors. These errors occur due to the nature of identical particles, and better isolation from the noisy environment does not reduce the probability of decoherence. Due to the symmetric structure of permutation-invariant codes, they can tolerate exchange (permutation) errors.

Subsequently, it was shown that PI codewords form ground states of the ferromagnetic Heisenberg model and thus can be realized in a Heisenberg ferromagnet [39]. Another appealing feature of PI codes is their capability of correcting quantum deletion errors [40, 48], which is not the case for conventional stabilizer codes. More recently, it was discovered that PI codes support implementation of transversal gates from finite groups beyond the Clifford group [31, 30, 43].

Definition 7. A permutation invariant code Q^{PI} is an N -qudit quantum code stabilized by the symmetric group: for all $|\psi\rangle \in Q^{\text{PI}}$ and $g \in S_N$, $g|\psi\rangle = |\psi\rangle$.

PI codes are often described using Dicke states. To define them, we recall that the composition of a string $\underline{x}$ of length N over a q -ary alphabet is a q -tuple

$$C(\underline{x}) = (n_0(\underline{x}), n_1(\underline{x}), \dots, n_{q-1}(\underline{x})),$$

where

$$n_i(\underline{x}) = \sum_{j=1}^N \delta_{x_j, i}$$

is the number of occurrences of the character i in the string $\underline{x}$. Note that $\sum_x C(\underline{x}) = N$. A qudit *Dicke state* is the linear combination of all qudit states with the same composition, i.e.,

$$|D_{\underline{n}}\rangle = \frac{1}{\sqrt{\binom{N}{\underline{n}}}} \sum_{C(\underline{x})=\underline{n}} |\underline{x}\rangle.$$

For example, the Dicke state corresponding to the composition string $(1, 1, 1)$ is

$$|D_{(1,1,1)}\rangle = \frac{1}{\sqrt{6}}(|0, 1, 2\rangle + |0, 2, 1\rangle + |1, 0, 2\rangle + |1, 2, 0\rangle + |2, 0, 1\rangle + |2, 1, 0\rangle).$$

By construction, such states are permutation invariant: for any $\underline{n} \in \mathcal{S}_{q,N}$, we have $g|D_{\underline{n}}\rangle = |D_{\underline{n}}\rangle$ for all $g \in S_N$. Therefore, the code space of a permutation-invariant code is formed of linear combinations of Dicke states with complex coefficients. A K -dimensional qudit PI code $\mathcal{Q}^{\text{PI}}$ of length N is defined as the subspace spanned by the following basis states:

$$|c_i\rangle = \sum_{\underline{n} \in \mathcal{S}_{q,N}} \alpha_{\underline{n}}^{(i)} \cdot |D_{\underline{n}}\rangle, \quad i = 0, \dots, K-1.$$

The first family of qubit PI codes correcting an arbitrary number of errors was found in [39]. It can be defined using integer parameters g, n, u , and so is called the *gnu codes*. The shortest gnu code correcting t Pauli errors has length at least $(2t+1)^2$. PI codes encoding multiple qudits were presented in [42]. These works introduced an explicit construction of K -dimensional qudit PI codes of length $(K-1)(2t+1)^2$ correcting t Pauli errors for any alphabet size $q \geq 2$. A more efficient construction of qubit PI codes generalizing gnu codes was found in [3]. These qubit codes encode two logical qubits, correct t errors, and require $(2t+1)^2 - 2t$ physical qubits. Recently, a general framework for constructing high-dimensional qudit PI codes using classical ℓ_1 codes was suggested in [2]. Relying on it, the authors of [2] constructed a family of K -dimensional qudit PI codes of length N , distance $d = t+1$, and alphabet size $q = N$, where $N = (K-1)t(t+1)$. They also constructed a family of PI codes of growing length N with $q = N$, distance close to $N/\log N$, and dimension 2^N . In the same paper, the authors established a connection between Fock state codes and PI codes. This result can be summarized as follows.

Proposition 13. [2] *Let $C_N \in \mathcal{S}_{q,N}$ be a classical ℓ_1 code with distance $t+1$. Let Q be a K -dimensional Fock state code with total excitation N constructed from C_N as in (14). Define the mapping*

$$|\underline{n}\rangle_b \xleftrightarrow[h^{-1}]{h} |D_{\underline{n}}\rangle.$$

Applying this mapping to a Fock state code Q yields a PI code with alphabet size q , length N , and distance $d = t+1$ defined by the basis

$$|c_i\rangle = \sum_{\underline{n} \in I_i} \alpha_{\underline{n}}^{(i)} \cdot |D_{\underline{n}}\rangle, \quad i \in \{0, 1, \dots, K-1\}. \quad (30)$$

Proof sketch. We start with the code in (14), which is a q -mode Fock state code with total excitation N . The mapping h sends a q -mode Fock state with total excitation N to a Dicke state with alphabet size q and total length N . Since the Fock state code Q in (14) is constructed using the framework in [2], the coefficients $\{\alpha_{\underline{n}}^{(i)}\}$ associated to the partition $I_0, I_1, \dots, I_{K-1}$ of C_N must be a solution for the set of equations (C1)-(C4) of that paper. As shown in Theorem 1 on p.7 of [2], these equations guarantee that the PI code in (30) has distance $d = t+1$. $\square$

Proposition 13 combined with Theorem 12 gives a construction of asymptotically good PI codes.

Theorem 14 (Asymptotically good PI codes). *For any $0 < \delta < 1$, $\varepsilon > 0$, and $\alpha > 0$ there exists a sequence of PI codes $(Q_N^{\text{PI}})_N$, $Q_N^{\text{PI}} \subseteq \mathcal{H}_{\alpha N, N}$ of increasing length N , distance at least $\lfloor \delta N \rfloor$, and dimension K_N as long as*

$$\frac{1}{N} \log_2 K_N \leq R_{\text{GV}}(\delta, \alpha) - (\alpha + \delta) h_2\left(\frac{\alpha}{\alpha + \delta}\right) + o(1),$$

for some fixed $o(1)$ function.

We note that by the previous results, the codewords of the codes $(Q_N^{\text{PI}})_N$ are balanced in the sense that every symbol of the alphabet is used at most $O(\log N)$ times.

5.2 Spin codes

Both PI codes and multimode Fock state codes encode quantum information into a composite many-body quantum system. Another physical platform to realize quantum codes is nuclear state spaces of molecules, atoms, or ions. Spin codes, introduced in [21], aim to protect quantum information by encoding it into a single large spin. Spin- J systems are generally described using $(2J + 1)$ -dimensional irreducible representations of $SU(2)$. Therefore, spin codes were initially defined and studied using $SU(2)$ geometry. The physical Hilbert space of these platforms has the basis $\{|J, m\rangle \mid m = -J, -J + 1, \dots, J - 1, J\}$, where each basis state $|J, m\rangle$ is an eigenstate of the angular momentum operator J_z with eigenvalue m . It is also possible to interpret the same physical Hilbert space using irreducible representations of $SU(q)$ of the same dimension. The basis states of this Hilbert space can also be defined using vectors in the discrete simplex $\mathcal{S}_{q, N}$. In other words, the D -dimensional irreducible representation of $SU(q)$ is defined using the basis $\{|\underline{n}\rangle_s, \underline{n} \in \mathcal{S}_{q, N}\}$, where $D = \binom{q+N-1}{q-1}$. Hence, a spin- J system that houses an irreducible representation of $SU(2)$ also houses an irreducible representation of $SU(q)$ for any $q, N > 0$, where $2J + 1 = \binom{q+N-1}{q-1}$.

A K -dimensional $SU(q)$ spin code with total spin N has the following basis states:

$$|c_i\rangle = \sum_{\underline{n} \in \mathcal{S}_{q, N}} \alpha_{\underline{n}}^{(i)} |\underline{n}\rangle_s, \quad i \in [K] \quad (31)$$

The noise process for spin codes is described by small random rotations. Let $\mathcal{J}_q = \{J_i : i = 1, 2, \dots, q^2 - 1\}$ be the generators of the Lie algebra $\mathfrak{su}(q)$ in the fundamental irreducible representation. When $q = 2$, the generators $\mathcal{J}_2 = \{J_x, J_y, J_z\}$ are the angular momentum operators. Then we say that a spin code has a distance $d = t + 1$, if it detects errors from the set

$$\mathcal{E}_t^{(s)} = \{J_{i_1} J_{i_2} \dots J_{i_t} : J_{i_k} \in \mathcal{J}_q\}$$

Relying on the results of [2], it is possible to show that Proposition 13 can be modified to yield a connection between spin codes and Fock state codes. In conclusion, the construction of asymptotically good Fock state codes yields a family of spin codes with increasing J that detect errors from the set $\mathcal{E}_t^s$ with t growing linearly with N .

6 Classical ℓ_1 codes

In this section, we state and prove results about the asymptotic behavior of classical ℓ_1 codes used in our main constructions.

6.1 Deterministic ℓ_1 codes

GV bounds for codes in $\mathcal{S}_{q,N}$ were previously considered in the literature. A standard GV argument implies that, in a finite metric space X , there is a code of size $\geq |X|/\bar{B}(d-1)$, where the denominator is the *average volume* of the ball of radius $d-1$ in X [22, 50]. Coupled with a recent work [19], which provides the asymptotic growth rate of $\bar{B}(\cdot) \subset \mathcal{S}_{q,N}$, this gives an existence bound for sequences of ℓ_1 codes. However, this result lacks certain features that we need for the construction. Specifically, we are interested in codes whose codewords have bounded support or the ℓ_∞ norm, neither of which is implied by the above argument. At the same time, choosing a point in $\mathcal{S}_{q,N}$ at random will yield such points with high probability, so we can limit our sample space to a subset of “typical” vectors and apply greedy search or randomized constructions only on this subset. When used in the construction of quantum codes that we employ below, this results in quantum codewords with bounded excitation per mode, yielding a well-balanced Fock state code.

Limiting ourselves to the subset of typical vectors, we consider approaches to constructing ℓ_1 codes. The first one relies on a Gilbert greedy procedure on the subset of typical vectors, resulting in a nonconstructive existence proof of asymptotically good codes. We also propose a simple randomized construction that yields similar code families relying on a low-complexity procedure.

Given a classical code $C \subset \mathcal{S}_{q,N}$, we define its rate as

$$R(C) = \frac{\log_2 |C|}{\log_2 |\mathcal{S}_{q,N}|}. \quad (32)$$

For a sequence of codes $(C_N)_N$, $N \rightarrow \infty$, the asymptotic rate and relative distance are defined as $\liminf_{N \rightarrow \infty} R(C_N)$ and $\liminf_{N \rightarrow \infty} d(C_N)/N$, respectively. Note that the diameter of the metric space $\mathcal{S}_{q,N}$ is N , so the relative distance is at most 1.

For $\underline{x} \in \mathcal{S}_{q,N}$ let $B(\underline{x}, r) := \{\underline{y} \in \mathcal{S}_{q,N} : d(\underline{x}, \underline{y}) \leq r\}$ denote the metric ball centered at $\underline{x}$. Below we use typical properties of vectors in $\mathcal{S}_{q,N}$ with respect to the support, $\text{supp}(\underline{x}) := \{i \in [q] : x_i \neq 0\}$ and l_∞ norm, $\|\underline{x}\|_\infty := \max_{i \in [q]} |x_i|$, collected in Lemmas 23 to 25 in the Appendix. We show that with high probability, the support of a random vector is about $q/(1+\alpha)$, and, perhaps more unexpectedly, its coordinates are bounded above by $O(\log N)$.

Proposition 15 (GV-construction of balanced ℓ_1 codes). *Given $\alpha > 0, \varepsilon > 0$ and $0 < \delta < 1$, consider the function*

$$R_{\text{GV}}(\delta) = (1 + \alpha)h_2\left(\frac{1}{1 + \alpha}\right) - \left(\frac{\alpha}{1 + \alpha} + \delta\right)h_2\left(\frac{\alpha}{\alpha + (1 + \alpha)\delta}\right) - (\alpha + \delta)h_2\left(\frac{\alpha}{\alpha + \delta}\right).$$

There exists a sequence of ℓ_1 codes $(C_N)_N$, where $C_N \subseteq \mathcal{S}_{\alpha N, N}$, such that

1. $\log_2 |C_N| \geq N(R_{\text{GV}}(\delta) + o(1))$;
2. C_N has distance at least $\lfloor \delta N \rfloor$;
3. $\max_{\underline{n} \in C_N} \left| \frac{1}{N} \text{supp}(\underline{n}) - \frac{\alpha}{1 + \alpha} \right| = o(1)$;
4. $\max_{\underline{n} \in C_N} \|\underline{n}\|_\infty < (1 + \varepsilon) \log_{1 + \alpha} N$.

Proof. By abuse of notation, below we assume that αN is an integer. Let us fix some $0 < \xi < 1$ and let $\mathcal{S}'_{\alpha N, N}$ be the set of all $\underline{n} \in \mathcal{S}_{\alpha N, N}$ that satisfy

$$\left| \frac{1}{N} |\text{supp}(\underline{n})| - \frac{\alpha}{1 + \alpha} \right| \leq N^{-\frac{1}{2} + \xi}, \quad \|\underline{n}\|_\infty \leq (1 + \varepsilon) \log_{1 + \alpha} N. \quad (33)$$

1. First, we observe that by limiting ourselves to $\mathcal{S}'_{\alpha N, N}$ we remove only a negligible part of the space, in other words that $\frac{|\mathcal{S}'_{\alpha N, N}|}{|\mathcal{S}_{\alpha N, N}|} = 1 - o(1)$ as N increases. Indeed, let $\underline{X}$ be uniform on $\mathcal{S}_{\alpha N, N}$ and consider the event $\mathcal{E}$ defined in (33) with $\underline{X}$ replacing $\underline{n}$. We will show that $\mathbb{P}(\mathcal{E}) = 1 - o(1)$.

Let us start with $\text{supp}(\underline{X})$. By Lemma 24, for sufficiently large N , for some decaying sequence $\gamma_N \rightarrow 0$, we have

$$\begin{aligned} \mathbb{P}\left[\left|\frac{1}{N} |\text{supp}(\underline{X})| - \frac{\alpha}{1 + \alpha}\right| > N^{-\frac{1}{2} + \xi/2}\right] \\ \leq \mathbb{P}\left[\left|\frac{1}{\alpha N} |\text{supp}(\underline{X})| - \mathbb{E}\left[\frac{1}{\alpha N} |\text{supp}(\underline{X})|\right]\right| > N^{-\frac{1}{2} + \xi/2} \cdot (1 - \gamma_N)\right] \\ \leq \frac{N(N - 1)(\alpha N - 1)N^{1 - \xi}}{\alpha N(N(1 + \alpha) - 1)^2(N(1 + \alpha) - 2)(1 - \gamma_N)^2} = O(N^{-\xi}). \end{aligned}$$

Now let us address the second part of the event $\mathcal{E}$. By Lemma 23, for any integer $B > 0$

$$\begin{aligned} \mathbb{P}[\|\underline{X}\|_\infty \geq B] &\leq \alpha N \frac{\binom{N(1 + \alpha) - B - 1}{\alpha N - 1}}{\binom{N(1 + \alpha) - 1}{\alpha N - 1}} = \alpha N \cdot \prod_{i=0}^{B-1} \frac{N - i}{N(1 + \alpha) - i - 1} \\ &\leq \alpha N \left(\frac{N}{N(1 + \alpha) - B} \right)^B. \end{aligned}$$

Let us take $B = (1 + \varepsilon) \log_{1 + \alpha} N$ and let $\gamma_N = B/N$. For a sufficiently large N , we have

$$\frac{N}{N(1 + \alpha) - B} = \frac{1}{1 + \alpha + \gamma_N} < (1 + \alpha)^{-(1 - \frac{\varepsilon}{2})},$$

and therefore

$$\mathbb{P}[\|\underline{X}\|_\infty \geq B] \leq \alpha N(1 + \alpha)^{-(1-\frac{\epsilon}{2})B} = \alpha N^{(\epsilon^2 - \epsilon)/2} \xrightarrow{N \rightarrow \infty} 0. \quad (34)$$

We have shown that both parts of $\mathcal{E}$ occur with probability $1 - o(1)$, but then so does their intersection, the event $\mathcal{E}$ itself.

2. Let us turn to code construction. We start with an arbitrary $\underline{n}_1 \in \mathcal{S}'_{\alpha N, N}$ and then choose an arbitrary vector $\underline{n}_2$ from $\mathcal{S}'_{\alpha N, N} \setminus \mathcal{B}(\underline{n}_1, \lfloor \delta N \rfloor)$. Arguing inductively, in the i th step we pick an arbitrary codeword $\underline{n}_i \in \mathcal{S}'_{\alpha N, N} \setminus \bigcup_{j=1}^{i-1} \mathcal{B}(\underline{n}_j, \lfloor \delta N \rfloor)$ as long as this set is not empty. Thus, to complete the proof, it is sufficient to prove that, as N increases,

$$\frac{|\mathcal{S}'_{\alpha N, N}|}{\max_{\underline{n} \in \mathcal{S}'_{\alpha N, N}} |\mathcal{B}(\underline{n}, \lfloor \delta N \rfloor)|} \geq 2^{N(R_{\text{GV}}(\delta) + o(1))}.$$

By Part 1 above and (51),

$$\begin{aligned} |\mathcal{S}'_{\alpha N, N}| &= |\mathcal{S}_{\alpha N, N}|(1 + o(1)) = \binom{N(1 + \alpha) - 1}{\alpha N - 1} (1 + o(1)) \\ &= 2^{N(1 + \alpha)(h_2(\frac{1}{1 + \alpha}) + o(1))}. \end{aligned} \quad (35)$$

Next, let us fix $\underline{n} \in \mathcal{S}'_{\alpha N, N}$ and denote $s_{\underline{n}} := |\text{supp}(\underline{n})| \leq \frac{\alpha N}{1 + \alpha}(1 + \gamma_N)$, where $\gamma_N \downarrow 0$ (not necessarily the same as above), where the support size is given in (33). Then, using Lemma 25,

$$\begin{aligned} |\mathcal{B}(\underline{n}, \lfloor \delta N \rfloor)| &\leq \sum_{j=0}^{\lfloor \delta N \rfloor} \binom{j + s_{\underline{n}} - 1}{s_{\underline{n}} - 1} \binom{j + \alpha N - 1}{\alpha N - 1} \\ &\leq \delta N \max_{0 \leq j \leq \lfloor \delta N \rfloor} \binom{j + s_{\underline{n}} - 1}{s_{\underline{n}} - 1} \binom{j + \alpha N - 1}{\alpha N - 1} \\ &= \delta N \binom{\lfloor \delta N \rfloor + s_{\underline{n}} - 1}{s_{\underline{n}} - 1} \binom{\lfloor \delta N \rfloor + \alpha N - 1}{\alpha N - 1}. \end{aligned}$$

Letting $N \rightarrow \infty$ and again using (51), we obtain

$$\frac{1}{N} \log_2 |\mathcal{B}(\underline{n}, \lfloor \delta N \rfloor)| \leq \left(\frac{\alpha}{1 + \alpha} + \delta \right) h_2\left(\frac{\alpha}{\alpha + (1 + \alpha)\delta} \right) + (\alpha + \delta) h_2\left(\frac{\alpha}{\alpha + \delta} \right) + o(1)$$

Since this is true for any $\underline{n}$, the proof follows by combining this inequality and (35). $\square$

6.2 Random ℓ_1 codes

In this section, we consider randomized procedures of constructing ℓ_1 codes. Although still asymptotically good, their rate–distance trade-offs are inferior to those obtained in the previous section. At the same time, these random ensembles offer a distinct structural advantage. As we will show in the next section, their inherent randomness enables a simple and low-complexity procedure for constructing asymptotically good bosonic Fock state codes, bypassing a computationally involved step required in an earlier construction, [2]. See Remark 4 for additional details.

We introduce two randomized ℓ_1 code constructions. The first one samples codewords independently and uniformly from the simplex $\mathcal{S}_{q,N}$. Using arguments similar in spirit to the GV-type analysis, we show that with high probability the resulting code achieves linear distance and positive rate while satisfying an $O(\log N)$ bound on the coordinate values. We then show that replacing the uniform simplex distribution with a multinomial distribution allows us to further reduce the maximal coordinate value to $O(\log N / \log \log N)$, at the cost of a weaker achievable rate. We emphasize that this flexibility in choosing the per-codeword sampling distribution is quite general, and other distributions can be used within the same framework. In particular, identifying a distribution that simultaneously achieves linear distance and rate while ensuring $O(1)$ boundedness remains an intriguing and important open question.

Proposition 16 (Uniform ℓ_1 codes). *Let $R_U : [0, 1] \rightarrow \mathbb{R}$ be a function given by*

$$R_U(\delta) = \frac{(1+\alpha)}{2} h_2\left(\frac{1}{1+\alpha}\right) - (\alpha + \delta) h_2\left(\frac{\alpha}{\alpha + \delta}\right).$$

Let $C_N^U = \{\underline{n}^{(i)}\}_{i=1}^{L_N}$ be a code formed of L_N uniform random vectors chosen independently from the simplex $\mathcal{S}_{\alpha N, N}$. For $N \rightarrow \infty$, if

$$\log_2 L_N \leq N(R_U(\delta) + o(1)), \quad (36)$$

then $\lim_{N \rightarrow \infty} \mathbb{P}[d(C_N^U) \geq \lfloor \delta N \rfloor] = 1$.

Proof. Let $t = \lfloor \delta N \rfloor$ and for $i, j \in [L], i \neq j$, consider a random variable $Y_{ij} = \mathbb{1}(d(\underline{n}^{(i)}, \underline{n}^{(j)}) \leq t)$. We have

$$\mathbb{P}\left[\sum_{i \neq j} Y_{i,j} = 0\right] = 1 - \mathbb{P}\left[\bigcup_{i \neq j} \{Y_{i,j} = 1\}\right] \geq 1 - \sum_{i \neq j} \mathbb{P}[Y_{ij} = 1]. \quad (37)$$

Using independence, we compute

$$\begin{aligned} \mathbb{P}[Y_{ij} = 1] &= \sum_{\underline{n} \in \mathcal{S}_{\alpha N, N}} \mathbb{P}[Y_{ij} = 1 \mid \underline{n}^{(i)} = \underline{n}] \mathbb{P}[\underline{n}^{(i)} = \underline{n}] \\ &= \sum_{\underline{n}} \mathbb{P}[\underline{n}^{(j)} \in B(\underline{n}, t) \mid \underline{n}^{(i)} = \underline{n}] \mathbb{P}[\underline{n}^{(i)} = \underline{n}] \\ &= \sum_{\underline{n}} \mathbb{P}[\underline{n}^{(j)} \in B(\underline{n}, t)] \mathbb{P}[\underline{n}^{(i)} = \underline{n}] \\ &= \sum_{\underline{n}} \frac{|B(\underline{n}, t)|}{|\mathcal{S}_{\alpha N, N}|} \mathbb{P}[\underline{n}^{(i)} = \underline{n}] \\ &\leq \sum_{\underline{n} \in \mathcal{S}_{\alpha N, N}} t \frac{\binom{\lfloor \delta N \rfloor + \alpha N - 1}{\alpha N - 1}}{\binom{N(1+\alpha) - 1}{\alpha N - 1}} \cdot \mathbb{P}[\underline{n}^{(i)} = \underline{n}] \\ &= 2^{-N(2R_U(\delta) + o(1))} \end{aligned}$$

where the inequality follows from Lemma 25, and the last line uses (51). Combining this estimate with (37) and the assumption in (36), we find that

$$\mathbb{P}\left[\sum_{i \neq j} Y_{i,j} = 0\right] \geq 1 - \sum_{i \neq j} \mathbb{P}[Y_{i,j} = 1] \geq 1 - \binom{L}{2} \cdot 2^{-N(2R_U(\delta) + o(1))}$$

$$\geq 1 - 2^{2(\log_2 L - N(R_{\mathcal{U}}(\delta) + o(1)))} = 1 - o(1).$$

We have shown that the distance of the code with probability approaching one is at least $\lfloor \delta N \rfloor$. $\square$

Remark 5. Note that uniform sampling from $\mathcal{S}_{q,N}$ is easily done by the balls-and-bins argument. To start, sample numbers $b_1, b_2, \dots, b_{q-1}$ uniformly without replacement from the set $\{1, \dots, N + q - 1\}$. W.l.o.g. assume that $b_1 < b_2 < \dots < b_{q-1}$. Then the vector $(x_1, \dots, x_q)$, where $x_1 = b_1 - 1$, $x_i = b_i - b_{i-1} - 1$, $i = 2, \dots, q-1$, and $x_q = (N + q - 1) - b_{q-1}$, is a uniform random vector on $\mathcal{S}_{q,N}$.

To further improve the guaranteed per-mode occupancy, we replace the uniform simplex distribution with a multinomial distribution. Specifically, each codeword is generated by independently placing N balls into q bins, where each ball is assigned to a bin uniformly at random, and the resulting occupancy vector is taken as the codeword.

Construction 2. Let $Y_1, \dots, Y_N$ be i.i.d. random variables, distributed uniformly on the set $[q]$. Define a random vector $\underline{X} \in \mathcal{S}_{q,N}$ by

$$X_i = \sum_{j=1}^N \mathbb{1}_{\{Y_j=i\}}, \quad i = 0, \dots, q-1 \quad (38)$$

We refer to its distribution as $\text{Multinomial}(q, N)$. A code in $\mathcal{S}_{q,N}$ of size L is constructed by sampling L independent vectors $\underline{X}_i$ according to this procedure.

This construction relies on independent sampling and is therefore simpler to implement than uniform sampling from the simplex, while still producing codewords with well-behaved typical properties. We start by showing that the resulting codes with high probability are asymptotically good. At the end of this section (Lemma 19) we also prove that the coordinates of a random vector $\underline{X} \sim \text{Multinomial}(q, N)$ are uniformly bounded above by a sublogarithmic function of N .

Proposition 17 (Multinomial ℓ_1 codes). *Fix $\delta \in (0, 1)$ and $\alpha > 0$, and define the function $R_M : [0, 1] \rightarrow \mathbb{R}$ by*

$$R_M(\delta) = \frac{(\Delta_\alpha - \delta)^2}{8 \ln 2},$$

where

$$\Delta_\alpha := \frac{1}{\pi} \int_0^\pi e^{-\frac{2(1-\cos(\theta))}{\alpha}} (1 - \cos(\theta)) d\theta.$$

Let $C_N^M = \{\underline{n}^{(i)}\}_{i=1}^{L_N}$ be a code formed of L_N i.i.d. random vectors sampled out of $\mathcal{S}_{\alpha N, N}$ according to $\text{Multinomial}(\lfloor \alpha N \rfloor, N)$. There exists an $o(1)$ function such that for $N \rightarrow \infty$, if,

$$\log_2 L_N \leq N(R_M(\delta) + o(1)) \quad (39)$$

then $\lim_{N \rightarrow \infty} \mathbb{P} \left[d(C_N^M) \geq \lfloor \delta N \rfloor \right] = 1$.

For the proof of Proposition 17, we will need the expected distance between two random vectors as above. It is given in the following lemma.

Lemma 18. Let $\underline{X}$ and $\underline{Z}$ be independent $\text{Multinomial}(\lfloor \alpha N \rfloor, N)$ random vectors. Then,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} [d(\underline{X}, \underline{Z})] = \Delta_\alpha.$$

Proof. Although $\underline{X}$ and $\underline{Z}$ depends on N , below we will suppress N from the notation. As before, by abuse of notation, we assume that αN is an integer. Observe that

$$\frac{1}{N} \mathbb{E} [d(\underline{X}, \underline{Z})] = \frac{1}{2N} \sum_{i=0}^{\alpha N - 1} \mathbb{E} |X_i - Z_i| = \frac{\lfloor \alpha N \rfloor}{2N} \cdot \mathbb{E} |X_1 - Z_1|, \quad (40)$$

where the last equality follows since the pairs (X_i, Z_i) are identically distributed and therefore all have the same expectation. From the definition of the multinomial distribution, X_i and Z_i are independent $\text{Binomial}(N, p_N)$, with $p_N = \frac{1}{\alpha N}$. As is well known ([16, Sec. VI.5]), such binomial random variables converge in distribution to a $\text{Poi}(\frac{1}{\alpha})$ random variable, and thus their distribution functions F_{X_1}, F_{Z_1} converge to F_α , the cumulative distribution function of $\text{Poi}(\frac{1}{\alpha})$. By the independence of X_1 and Z_1 , the joint distribution satisfies

$$F_{X_1, Z_1}(x, z) = F_{X_1}(x) F_{Z_1}(z) \rightarrow F_\alpha(x) F_\alpha(z),$$

and therefore as $N \rightarrow \infty$, the pairs (X_1, Z_1) converge in distribution to (X, Z) , distributed as a pair of independent Poisson random variables $\text{Poi}(\frac{1}{\alpha})$. By the continuous mapping theorem (see Lemma 26), we now can claim that as N increases, the random variables $|X_1 - Z_1|$ converge in distribution to $|X - Z|$. Next, note that this sequence of random variables is uniformly integrable. Indeed, since X_1 and Z_1 are non-negative, we have:

$$\mathbb{E} |X_1 - Z_1| \leq \mathbb{E} [X_1 + Z_1] = 2 \frac{N}{\lfloor \alpha N \rfloor} < \frac{2}{\alpha} + \varepsilon,$$

for sufficiently large N , as desired. By Lemma 27, this suffices to claim that $\mathbb{E} |X_1 - Z_1| \rightarrow \mathbb{E} |X - Z|$ as $N \rightarrow \infty$.

It now remains to compute the expectation of $|X - Z|$, which is the absolute value of $S \triangleq X - Z$ —the difference between two independent $\text{Poi}(\frac{1}{\alpha})$. The distribution of S is known as Skellam($\frac{1}{\alpha}, \frac{1}{\alpha}$) distribution [49], and its PDF is given by

$$P_S(k) = e^{-\frac{2}{\alpha}} I_k \left(\frac{2}{\alpha} \right), \quad k \in \mathbb{Z},$$

where for $k \in \mathbb{Z}$, I_k is the k^{th} order modified Bessel function of the first kind,

$$I_k(x) := \frac{1}{\pi} \int_0^\pi e^{x \cos \theta} \cos(k\theta) d\theta.$$

Note that the above distribution is symmetric around zero, and therefore

$$\begin{aligned} \mathbb{E}[|S|] &= \sum_{k \in \mathbb{Z}} |k| \cdot P_S(k) = 2 \cdot \sum_{k=1}^{\infty} k \cdot P_S(k) = e^{-\frac{2}{\alpha}} \sum_{k=1}^{\infty} k I_k \left(\frac{2}{\alpha} \right) \\ &\stackrel{(a)}{=} \frac{e^{-\frac{2}{\alpha}}}{\alpha} \lim_{m \rightarrow \infty} \sum_{k=1}^m \left(I_{k-1} \left(\frac{2}{\alpha} \right) - I_{k+1} \left(\frac{2}{\alpha} \right) \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{e^{-\frac{2}{\alpha}}}{\alpha} \lim_{m \rightarrow \infty} \sum_{k=1}^m \left(I_0\left(\frac{2}{\alpha}\right) + I_1\left(\frac{2}{\alpha}\right) - I_m\left(\frac{2}{\alpha}\right) - I_{m+1}\left(\frac{2}{\alpha}\right) \right) \\
&\stackrel{(b)}{=} \frac{2e^{-\frac{2}{\alpha}}}{\alpha} \left(I_0\left(\frac{2}{\alpha}\right) + I_1\left(\frac{2}{\alpha}\right) \right),
\end{aligned}$$

where (a) uses the following recurrence relation for the Bessel functions [38, Sec. 10.29]:

$$\frac{2k}{x} I_k(x) = I_{k-1}(x) - I_{k+1}(x), \quad k \in \mathbb{Z}$$

and (b) follows since $I_m(1/\alpha) \xrightarrow{m \rightarrow \infty} 0$ as $P_S(K)$ sums to 1. Combining the above with (40) we conclude:

$$\begin{aligned}
\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[d(\mathbf{X}, \mathbf{Z})] &= \lim_{N \rightarrow \infty} \frac{\lfloor \alpha N \rfloor}{2N} \cdot \mathbb{E}[\mathbf{X}_1 - \mathbf{Z}_1] = \frac{\alpha}{2} \cdot \lim_{N \rightarrow \infty} \mathbb{E}[\mathbf{X}_1 - \mathbf{Z}_1] \\
&= \frac{\alpha}{2} \cdot \mathbb{E}[|\mathbf{X} - \mathbf{Z}|] = e^{-\frac{2}{\alpha}} \left(I_0\left(\frac{2}{\alpha}\right) + I_1\left(\frac{2}{\alpha}\right) \right). \quad \square
\end{aligned}$$

We can now prove Proposition 17:

Proof of Proposition 17. The proof uses the notation and closely follows the logic of the proof of Proposition 16 with the only change in the evaluation of $\mathbb{P}[\mathbf{Y}_{i,j} = 1] = \mathbb{P}[d(\mathbf{n}^{(i)}, \mathbf{n}^{(j)}) < t]$, which we now bound using McDiarmid's inequality. Indeed, let $\{Z_{i,l} : i = 1, \dots, N, l = 1, \dots, L\}$ be the $\text{Unif}([q])$ i.i.d. random variables such that $\mathbf{n}^{(i)}$ is defined by $(Z_{1,i}, \dots, Z_{N,i}) \triangleq \mathbf{Z}_i$ via (38). Consider the function $f : [q]^{2N} \rightarrow \mathbb{R}_+$ defined by

$$f(Z_{1,i}, \dots, Z_{N,i}, Z_{1,j}, \dots, Z_{N,j}) = \frac{1}{N} d(\mathbf{n}^{(i)}, \mathbf{n}^{(j)}).$$

We note that f satisfies the bounded differences property (49) with $c_1 = c_2 = \dots = c_{2N} = \frac{2}{N}$. Indeed, changing the bin of one ball will increase one coordinate of the corresponding simplex vector by 1 and decrease another coordinate by 1. Setting $t = \lfloor \delta N \rfloor$ and $q = \lfloor \alpha N \rfloor$, we have that

$$\begin{aligned}
\mathbb{P}[d(\mathbf{n}^{(i)}, \mathbf{n}^{(j)}) < \lfloor \delta N \rfloor] &\leq \mathbb{P}[|d(\mathbf{n}^{(i)}, \mathbf{n}^{(j)}) - \mathbb{E} d(\mathbf{n}^{(i)}, \mathbf{n}^{(j)})| > \mathbb{E} d(\mathbf{n}^{(i)}, \mathbf{n}^{(j)}) - \lfloor \delta N \rfloor] \\
&\stackrel{(a)}{=} \mathbb{P}\left[\left|\frac{1}{N} d(\mathbf{n}^{(i)}, \mathbf{n}^{(j)}) - \mathbb{E}\left[\frac{1}{N} d(\mathbf{n}^{(i)}, \mathbf{n}^{(j)})\right]\right| > R_\alpha - \delta + o(1)\right] \\
&= \mathbb{P}\left[\left|f(\mathbf{Z}_i, \mathbf{Z}_j) - \mathbb{E}[f(\mathbf{Z}_i, \mathbf{Z}_j)]\right| > R_\alpha - \delta + o(1)\right] \\
&\stackrel{(b)}{\leq} 2 \exp\left(-\frac{2(\Delta_\alpha - \delta + o(1))^2}{\sum_{i=1}^{2N} \frac{4}{N^2}}\right) \\
&= \exp\left(-\frac{N}{4} \left((\Delta_\alpha - \delta)^2 + o(1)\right)\right),
\end{aligned}$$

where (a) follows since $\frac{1}{N} \mathbb{E} d(\mathbf{n}^{(i)}, \mathbf{n}^{(j)}) = \Delta_\alpha + o(1)$ by Lemma 18, and (b) follows from MacDiarmid's inequality (see Lemma 28). We now conclude following the steps of Proposition 16:

$$\mathbb{P}\left[d(\mathbf{C}_N^M) \geq \lfloor \delta N \rfloor\right] \geq 1 - \sum_{i \neq j} \mathbb{P}[\mathbf{Y}_{i,j} = 1] \leq 1 - \binom{L}{2} \cdot \exp\left(-\frac{N}{4} \left((R_\alpha - \delta)^2 + o(1)\right)\right)$$

$$> 1 - 2^{2 \log_2 L - N^{\frac{(R\alpha - \delta)^2 + o(1)}{4 \ln 2}}} = 1 - o(1),$$

where the last equality from the assumption (39). This completes the proof. $\square$

In Figure 1, we plot the curves $R_{GV}(\delta)$ and $R_U(\delta)$ for $\alpha = 1$. The bound $R_M(\delta)$ is positive for $\delta \leq 0.292$; however, the values of the rate are much smaller than the other two bounds, so we are not showing it in the same plot (the loss occurs because we rely on a concentration inequality). We also remark that the *code rate* as defined in (32) above is obtained by dividing the values R_{GV} etc. by $\frac{1}{N} \log_2 |S_{\alpha N, N}| \sim (1 + \alpha)h_2(1/(1 + \alpha))$, which does not change the conclusion regarding asymptotic goodness of the code sequences. We also observe that [19, Thm. 10] gives a better asymptotic bound for codes in the entire simplex, not accounting for the typical subset restriction.

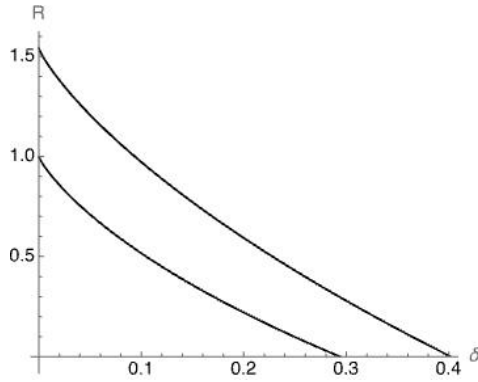


Figure 1: Bounds $R_{GV}(\delta)$ (the upper curve) and $R_U(\delta)$ for $\alpha = 1$.

We conclude this section by showing that vectors drawn from the multinomial distribution typically have sub-logarithmic ∞ -norm, mimicking the properties of the uniform distribution on the simplex. In fact, it turns out that the ∞ -norm of a multinomial random vector decays as $(\ln N)/\ln \ln N$, which is even faster than the decay rate of $\ln N$ for a uniform vector.

Lemma 19. *Let $\mathbf{X} \sim \text{Multinomial}(\lfloor \alpha N \rfloor, N)$ be a random vector on the simplex. Then, for any fixed $\varepsilon > 0$*

$$\mathbb{P} \left[\|\mathbf{X}\|_{\infty} \geq \frac{(1 + \varepsilon) \ln N}{\ln \ln N} \right] = o(1).$$

Proof. As shown in the proof of Lemma 18, each coordinate $\mathbf{X}_i$ is distributed as $\text{Binomial}(N, p_N)$, with $p_N = \frac{1}{\lfloor \alpha N \rfloor}$, which converge in distribution to a $\text{Poi}(\frac{1}{\alpha})$ distribution. Thus, by the union bound

$$\begin{aligned} \mathbb{P} \left[\|\mathbf{X}\|_{\infty} \geq \frac{(1 + \varepsilon) \ln N}{\ln \ln N} \right] &= \mathbb{P} \left[\bigcup_{i=0}^{\lfloor \alpha N \rfloor - 1} \left\{ \mathbf{X}_i \geq \frac{(1 + \varepsilon) \ln N}{\ln \ln N} \right\} \right] \\ &\leq \sum_{i=0}^{\lfloor \alpha N \rfloor - 1} \mathbb{P} \left[\mathbf{X}_i \geq \frac{(1 + \varepsilon) \ln N}{\ln \ln N} \right] \\ &\leq \alpha N \cdot \left(\mathbb{P} \left[\text{Poi} \left(\frac{1}{\alpha} \right) \geq \frac{(1 + \varepsilon) \ln N}{\ln \ln N} \right] \cdot (1 + o(1)) \right), \end{aligned}$$

where the last inequality follows from convergence in distribution. It is therefore remains to show that

$$\mathbb{P} \left[\text{Poi} \left(\frac{1}{\alpha} \right) \geq \frac{(1 + \varepsilon) \ln N}{\ln \ln N} \right] = o \left(\frac{1}{N} \right) \quad (41)$$

Denote $\lambda = 1/\alpha$, and observe that for any $k \in \mathbb{N}$ by Stirling approximation

$$\begin{aligned} \mathbb{P} \left[\text{Poi} \left(\frac{1}{\alpha} \right) \geq k \right] &= \sum_{n=k}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} = \lambda \frac{\lambda^k}{k!} \left(1 + o \left(\frac{1}{k} \right) \right) \\ &= e^{-\lambda} \exp(-\lambda + k \ln \lambda - k \ln k + o(k)) \left(1 + o \left(\frac{1}{k} \right) \right) \\ &= \exp(-k \ln k (1 + o(1))). \end{aligned}$$

Setting $k = (1 + \varepsilon) \ln N / \ln \ln N$ we obtain

$$\begin{aligned} N \cdot \mathbb{P} \left[\text{Poi} \left(\frac{1}{\alpha} \right) \geq \frac{(1 + \varepsilon) \ln N}{\ln \ln N} \right] &\leq \exp(-(1 + \varepsilon) \cdot (\ln N) (1 - o(1)) + \ln N) \\ &= N^{-\varepsilon + o(1)} = o(1). \end{aligned} \quad \square$$

7 Conclusion and outlook

In this work, we studied quantum error correction for amplitude damping noise from two complementary perspectives. First, we refined the understanding of quantum codes' performance when used over the physically relevant, non-truncated AD channel. Second, we constructed new families of codes whose performance on this channel can be rigorously analyzed and shown to improve upon the known results. These two aspects, analysis of the physical noise model and explicit code constructions, are closely intertwined and together provide a more complete picture of error correction in this setting.

One conceptual message that emerges from our study concerns the relationship between the abstract framework of quantum coding theory and the performance of codes on concrete physical channels. The Knill–Laflamme conditions provide a powerful and general formalism for reasoning about correctable error sets, and they remain a cornerstone of quantum error-correction theory. At the same time, our results highlight the importance of carefully analyzing how a code behaves under the actual physical noise model of interest. When this connection is not examined directly, one may obtain codes that satisfy formal QECC conditions yet provide limited protection against the underlying physical channel. In this sense, our analysis can be viewed as an amplitude-damping/Fock-state analogue of the well-known approximation principle whereby general independent noise processes are related to restricted error models (see, e.g., [18, Theorem 1.3.4]).

Another observation arising from this work concerns the role of randomness in approximate quantum error correction. Much of the existing literature introduces randomness directly at the level of the *physical Hilbert space*. A canonical example is the framework of Haar-random codes, where the encoding subspace itself is chosen as a random subspace of the physical system [26, 28, 33]. A related perspective appears in the construction of codes from many-body eigenstates of translation-invariant spin chains [12], where the code space is formed by

randomly choosing physical eigenstates of a quantum Hamiltonian within a suitable energy window.

In contrast, the randomness in our construction appears at a different structural level. Rather than randomizing the physical code subspace, we introduce randomness in the *logical indexing geometry* from which the codewords are generated. Concretely, the codewords are indexed by points in an underlying classical metric space (the discrete simplex equipped with the ℓ_1 metric), and randomness is applied to this geometric structure. The performance guarantees in our analysis arise from concentration phenomena on the metric space. While our construction focuses on the ℓ_1 geometry of the simplex, the mechanism driving the result is not tied to this specific setting. Similar concentration effects are known to arise in a wide variety of high-dimensional metric spaces, suggesting that introducing randomness at the level of the geometric or combinatorial structure underlying a quantum code may provide a flexible and broadly applicable design principle. Investigating how this perspective can be leveraged in other coding geometries and for different physical noise models remains an intriguing direction for future work.

We conclude with a remark concerning spin codes. The errors in spin codes are typically modeled as small random rotations that transform spin states according to the Lindbladian master equation. We note that in the asymptotic regime, one can address the question similar to the one treated in this work for the AD channel, namely, relating the continuous quantum channel generated by the Lindbladian to a truncated error set defined by a subset of generators of the Lie algebra $\mathfrak{su}(q)$. We hope to address this problem in future research.

Acknowledgements

The research of the first author was supported by a Rothschild Fellowship granted by the Yad HaNadiv Foundation. The research of the last two authors was partially supported by NSF grants CIF-2330909 and CIF-2526035.

AI usage: AI tools were used to assist with proofreading the final paper.

Author contributions: All authors contributed to the mathematical development and writing of the manuscript.

References

- [1] K. Adiprasito, I. Bárány, N. H. Mustafa, and T. Terpai. Theorems of Carathéodory, Helly, and Tverberg without dimension. *Discrete & Computational Geometry*, 64(2):233–258, 2020. doi:[10.1007/s00454-020-00172-5](https://doi.org/10.1007/s00454-020-00172-5).
- [2] A. Aydin, V. V. Albert, and A. Barg. Quantum error correction beyond $SU(2)$: Spin, bosonic, and permutation-invariant codes from convex geometry. *PRX Quantum*, 7:010341, 2026. doi:[10.1103/kx3b-4nrp](https://doi.org/10.1103/kx3b-4nrp).
- [3] A. Aydin, M. A. Alekseyev, and A. Barg. A family of permutationally invariant quantum codes. *Quantum*, 8:1321, 2024. arXiv:[2310.05358](https://arxiv.org/abs/2310.05358), doi:[10.22331/q-2024-04-30-1321](https://doi.org/10.22331/q-2024-04-30-1321).
- [4] A. Barg and A. Mazumdar. Codes in permutations and error correction for rank modulation. *IEEE Trans. Inform. Theory*, 56(7):3158–3165, 2010. doi:[10.1109/TIT.2010.2048455](https://doi.org/10.1109/TIT.2010.2048455).
- [5] H. Barnum and E. Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *Journal of Mathematical Physics*, 43(5):2097–2106, 2002. doi:[10.1063/1.1459754](https://doi.org/10.1063/1.1459754).

- [6] F. Battistel. *General quantum error correction for MERA codes*. PhD thesis, Technische Universität and Ludwig-Maximilians-Universität München, Germany, 2017.
- [7] C. Bény. Conditions for the approximate correction of algebras. In *Workshop on Quantum Computation, Communication, and Cryptography*, pages 66–75. Springer, 2009. doi:[10.1007/978-3-642-10698-9_8](https://doi.org/10.1007/978-3-642-10698-9_8).
- [8] C. Bény and O. Oreshkov. General conditions for approximate quantum error correction and near-optimal recovery channels. *Physical Review Letters*, 104(12):120501, 2010. doi:[10.1103/PhysRevLett.104.120501](https://doi.org/10.1103/PhysRevLett.104.120501).
- [9] M. Bergmann and P. van Loock. Quantum error correction against photon loss using NOON states. *Phys. Rev. A*, 94:012311, Jul 2016. doi:[10.1103/PhysRevA.94.012311](https://doi.org/10.1103/PhysRevA.94.012311).
- [10] P. Billingsley. *Convergence of Probability Measures*. J. Wiley & Sons, Hoboken, NJ, 2nd edition, 1999. doi:[10.1002/9780470316962](https://doi.org/10.1002/9780470316962)
- [11] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013. doi:[10.1093/acprof:oso/9780199535255.001.0001](https://doi.org/10.1093/acprof:oso/9780199535255.001.0001).
- [12] F. G. Brandão, E. Crosson, M. B. Şahinoğlu, and J. Bowen. Quantum error correcting codes in eigenstates of translation-invariant spin chains. *Physical Review Letters*, 123(11):110502, 2019. doi:[10.1103/PhysRevLett.123.110502](https://doi.org/10.1103/PhysRevLett.123.110502).
- [13] S. Bravyi and A. Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Physical Review A—Atomic, Molecular, and Optical Physics*, 71(2):022316, 2005. doi:[10.1103/PhysRevA.71.022316](https://doi.org/10.1103/PhysRevA.71.022316).
- [14] I. L. Chuang, D. W. Leung, and Y. Yamamoto. Bosonic quantum codes for amplitude damping. *Physical Review A*, 56(2):1114, 1997. doi:[10.1103/PhysRevA.56.1114](https://doi.org/10.1103/PhysRevA.56.1114).
- [15] C. Crépeau, D. Gottesman, and A. Smith. Approximate quantum error-correcting codes and secret sharing schemes. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 285–301. Springer, 2005. doi:[10.1007/11426639_17](https://doi.org/10.1007/11426639_17).
- [16] W. Feller. *An Introduction to Probability Theory and its Applications*. Wiley, Hoboken, NJ, 3rd edition, 1968.
- [17] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A—Atomic, Molecular, and Optical Physics*, 86(3):032324, 2012. doi:[10.1103/PhysRevA.86.032324](https://doi.org/10.1103/PhysRevA.86.032324).
- [18] D. Gottesman. Surviving as a quantum computer in a classical world. *Textbook manuscript*, 2024.
- [19] K. Goyal, D. T. Dao, M. Kovačević, and H. M. Kiah. Gilbert–Varshamov bound for codes in l_1 metric using multivariate analytic combinatorics. *IEEE Trans. Inform. Theory*, 71(1):244–262, 2025. doi:[10.1109/TIT.2024.3483303](https://doi.org/10.1109/TIT.2024.3483303).
- [20] M. Grassl, L. Kong, Z. Wei, Z.-Q. Yin, and B. Zeng. Quantum error-correcting codes for qudit amplitude damping. *IEEE Transactions on Information Theory*, 64(6):4674–4685, 2018. doi:[10.1109/TIT.2018.2790423](https://doi.org/10.1109/TIT.2018.2790423).
- [21] J. A. Gross. Designing codes around interactions: The case of a spin. *Phys. Rev. Lett.*, 127:010504, 2021. doi:[10.1103/PhysRevLett.127.010504](https://doi.org/10.1103/PhysRevLett.127.010504).
- [22] J. Gu and T. Fuja. A generalized Gilbert–Varshamov bound derived via analysis of a code-search algorithm. *IEEE Trans. Inform. Theory*, 39(3):1089–1093, 1993. doi:[10.1109/18.256522](https://doi.org/10.1109/18.256522).
- [23] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi:[10.1080/01621459.1963.10500830](https://doi.org/10.1080/01621459.1963.10500830).
- [24] M. Hofheinz, E. M. Weig, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O’Connell, H. Wang, J. M. Martinis, and A. N. Cleland. Generation of Fock states in a superconducting quantum circuit. *Nature*, 454(7202):310–314, 2008. doi:[10.1038/nature07136](https://doi.org/10.1038/nature07136).

- [25] S. Khatri and M. M. Wilde. Principles of Quantum Communication Theory: A Modern Approach. arXiv preprint [arXiv:2011.04672v2](https://arxiv.org/abs/2011.04672v2), 2024. book draft.
- [26] R. Klesse. Approximate quantum error correction, random codes, and quantum channel capacity. *Physical Review A*, 75(6):062315, 2007. doi:[10.1103/PhysRevA.75.062315](https://doi.org/10.1103/PhysRevA.75.062315).
- [27] E. Knill. Quantum computing with realistically noisy devices. *Nature*, 434(7029):39–44, 2005. doi:[10.1038/nature03350](https://doi.org/10.1038/nature03350).
- [28] L. Kong and Z.-W. Liu. Near-optimal covariant quantum error-correcting codes from random unitaries with symmetries. *PRX Quantum*, 3(2):020314, 2022. doi:[10.1103/PRXQuantum.3.020314](https://doi.org/10.1103/PRXQuantum.3.020314).
- [29] M. Kovačević and V. Y. F. Tan. Codes in the space of multisets—Coding for permutation channels with impairments. *IEEE Trans. Inform. Theory*, 64(7):5156–5169, 2018. doi:[10.1109/TIT.2017.2789292](https://doi.org/10.1109/TIT.2017.2789292).
- [30] E. Kubischta and I. Teixeira. Family of quantum codes with exotic transversal gates. *Phys. Rev. Lett.*, 131(24):240601, 2023. doi:[10.1103/PhysRevLett.131.240601](https://doi.org/10.1103/PhysRevLett.131.240601).
- [31] E. Kubischta and I. Teixeira. Quantum codes from twisted unitary t -groups. *Phys. Rev. Lett.*, 133:030602, Jul 2024. doi:[10.1103/PhysRevLett.133.030602](https://doi.org/10.1103/PhysRevLett.133.030602).
- [32] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto. Approximate quantum error correction can lead to better codes. *Physical Review A*, 56(4):2567, 1997. doi:[10.1103/PhysRevA.56.2567](https://doi.org/10.1103/PhysRevA.56.2567).
- [33] F. Ma, X. Tan, and J. Wright. Haar random codes attain the quantum hamming bound, approximately. *arXiv preprint arXiv:2510.07158*, 2025.
- [34] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. Elsevier, 1977.
- [35] P. Mandayam and H. K. Ng. Towards a unified framework for approximate quantum error correction. *Physical Review A—Atomic, Molecular, and Optical Physics*, 86(1):012335, 2012. doi:[10.1103/PhysRevA.86.012335](https://doi.org/10.1103/PhysRevA.86.012335).
- [36] C. McDiarmid. On the method of bounded differences. In *Surveys in combinatorics, 1989 (Norwich, 1989)*, volume 141 of *London Math. Soc. Lecture Note Ser.*, pages 148–188. Cambridge Univ. Press, Cambridge, 1989. doi:[10.1017/CB09781107359949.008](https://doi.org/10.1017/CB09781107359949.008).
- [37] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, et al. Extending the lifetime of a quantum bit with error correction in superconducting circuits. *Nature*, 536(7617):441–445, 2016. doi:[10.1038/nature17665](https://doi.org/10.1038/nature17665).
- [38] F. W. Olver et al., editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010. online at <https://dlmf.nist.gov/10.29>.
- [39] Y. Ouyang. Permutation-invariant quantum codes. *Phys. Rev. A*, 90:062317, 2014. doi:[10.1103/PhysRevA.90.062317](https://doi.org/10.1103/PhysRevA.90.062317).
- [40] Y. Ouyang. Permutation-invariant quantum coding for quantum deletion channels. In *2021 IEEE International Symposium on Information Theory (ISIT)*, pages 1499–1503, 2021. doi:[10.1109/ISIT45174.2021.9518078](https://doi.org/10.1109/ISIT45174.2021.9518078).
- [41] Y. Ouyang and R. Chao. Permutation-invariant constant-excitation quantum codes for amplitude damping. *IEEE Trans. Inform. Theory*, 66(5):2921–2933, 2020. doi:[10.1109/TIT.2019.2956142](https://doi.org/10.1109/TIT.2019.2956142).
- [42] Y. Ouyang and J. Fitzsimons. Permutation-invariant codes encoding more than one qubit. *Phys. Rev. A*, 93:042340, Apr 2016. doi:[10.1103/PhysRevA.93.042340](https://doi.org/10.1103/PhysRevA.93.042340).
- [43] Y. Ouyang, Y. Jing, and G. K. Brennen. Measurement-free code-switching for low overhead quantum computation using permutation invariant codes, 2025. [arXiv:2411.13142](https://arxiv.org/abs/2411.13142).
- [44] H. Pollatsek and M. B. Ruskai. Permutationally invariant codes for quantum error correction. *Linear Alg. Appl.*, 392:255–288, 2004. doi:[10.1016/j.laa.2004.06.014](https://doi.org/10.1016/j.laa.2004.06.014).

- [45] T. C. Ralph, A. Hayes, and A. Gilchrist. Loss-tolerant optical qubits. *Physical Review Letters*, 95(10):100501, 2005. [doi:10.1103/PhysRevLett.95.100501](#).
- [46] M. B. Ruskai. Pauli exchange and quantum error correction, 2000. [arXiv quant-ph/0006008](#).
- [47] B. Schumacher and M. D. Westmoreland. Approximate quantum error correction. *Quantum Information Processing*, 1(1):5–12, 2002. [doi:10.1023/A:1019653202562](#).
- [48] T. Shibayama and M. Hagiwara. Permutation-invariant quantum codes for deletion errors. In *2021 IEEE International Symposium on Information Theory (ISIT)*, pages 1493–1498. IEEE, 2021. [doi:10.1109/ISIT45174.2021.9517870](#).
- [49] J. G. Skellam. The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 109(3):296–296, 1946. [doi:10.1111/j.2397-2335.1946.tb04670.x](#).
- [50] L. Tolhuizen. The generalized Gilbert-Varshamov bound is implied by Turán’s theorem. *IEEE Trans. Inform. Theory*, 43(5):1605–1606, 1997. [doi:10.1109/18.623158](#).
- [51] J. Tyson. Two-sided bounds on minimum-error quantum measurement, on the reversibility of quantum dynamics, and on the maximum overlap problem using directional iterates. *Journal of Mathematical Physics*, 51(9):092204, 2010. [doi:10.1063/1.3463451](#).
- [52] M. Uria, P. Solano, and C. Hermann-Avigliano. Deterministic generation of large Fock states. *Phys. Rev. Lett.*, 125:093603, Aug 2020. [doi:10.1103/PhysRevLett.125.093603](#).
- [53] E. Waks, E. Diamanti, and Y. Yamamoto. Generation of photon number states. *New Journal of Physics*, 8(1):4, jan 2006. [doi:10.1088/1367-2630/8/1/004](#).
- [54] H. Wang, M. Hofheinz, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O’Connell, D. Sank, J. Wenner, A. N. Cleland, and J. M. Martinis. Measurement of the decay of Fock states in a superconducting quantum circuit. *Phys. Rev. Lett.*, 101:240401, Dec 2008. [doi:10.1103/PhysRevLett.101.240401](#).
- [55] W. Wasilewski and K. Banaszek. Protecting an optical qubit against photon loss. *Physical Review A—Atomic, Molecular, and Optical Physics*, 75(4):042316, 2007. [doi:10.1103/PhysRevA.75.042316](#).
- [56] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. [doi:10.1017/9781316848142](#).
- [57] M. M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2nd edition, 2017. [doi:10.1017/9781316809976](#).
- [58] Y. Yehezkeally and M. Schwartz. Reconstruction codes for DNA sequences with uniform tandem-duplication errors. *IEEE Transactions on Information Theory*, 66(5):2658–2668, 2019. [doi:10.1109/TIT.2019.2940256](#).
- [59] Y. Yehezkeally and M. Schwartz. Uncertainty of reconstruction with list-decoding from uniform-tandem-duplication noise. *IEEE Transactions on Information Theory*, 67(7):4276–4287, 2021. [doi:10.1109/TIT.2021.3070466](#).

A Technicalities and auxiliary results

In this section, we give formal definitions and technical results that we shall use throughout the paper.

A.1 Channels and state norms

Definition 8 (Trace norm, Diamond norm, Completely Bounded norm). Let $\mathcal{H}$ and $\mathcal{H}'$ be finite-dimensional Hilbert spaces. Let $X : \mathcal{H} \rightarrow \mathcal{H}'$ and $\mathcal{N} : L(\mathcal{H}) \rightarrow L(\mathcal{H}')$ linear maps.

1. The trace norm on $L(\mathcal{H})$ is defined by

$$\|X\|_1 \triangleq \text{tr}(|X|), \quad |X| \triangleq \sqrt{X^\dagger X},$$

and its just the sum of absolute values of singular values of X .

2. The completely bounded norm of $\mathcal{N}$ is defined as

$$\|\mathcal{N}\|_{\text{cb}} = \sup_{\substack{X \in L(\mathcal{H}^{\otimes 2}) \\ \|X\|_1 \leq 1}} \|(I_{L(\mathcal{H})} \otimes \mathcal{N})(X)\|_1.$$

3. The diamond norm of $\mathcal{N}$ is defined as

$$\|\mathcal{N}\|_\diamond = \sup_{\rho \in D(\mathcal{H}^{\otimes 2})} \|(I_{L(\mathcal{H})} \otimes \mathcal{N})(\rho)\|_1.$$

Note that it is immediate from the definition of the diamond and completely bounded norms that $\|\mathcal{N}\|_\diamond \leq \|\mathcal{N}\|_{\text{cb}}$. An extension of Fuchs–van de Graaf inequalities (see [25, Section 6.2]) implies that for two quantum channels $\mathcal{N}, \mathcal{M}$ we have

$$1 - \sqrt{\mathcal{F}_e(\mathcal{N}, \mathcal{M})} \leq \frac{1}{2} \|\mathcal{N} - \mathcal{M}\|_\diamond \leq \sqrt{1 - \mathcal{F}_e(\mathcal{N}, \mathcal{M})}.$$

Translating to Bures distance, the above implies the following:

Lemma 20. *For any two quantum channels $\mathcal{N}, \mathcal{M}$ we have*

$$\frac{1}{2} \|\mathcal{N} - \mathcal{M}\|_\diamond \leq d(\mathcal{N}, \mathcal{M}) \leq \sqrt{\|\mathcal{N} - \mathcal{M}\|_\diamond}. \quad (42)$$

We shall also require the Duality principle and a well-known inequality between norms (see [56, Section 1.1, pp 32]):

Lemma 21 (The duality principle). *For any $X, Y \in L(H_A)$ we have*

$$\text{tr}(Y^\dagger X) \leq \|X\|_1 \cdot \|Y\|_\infty$$

Lemma 22. *Let $\mathcal{B}$ be the channel from Theorem 1. There exists $|\psi\rangle \in Q^{\otimes 2}$ such that*

$$\|\mathcal{B}\|_\diamond = \|\mathcal{B}\|_{\text{cb}} = \|I_{L(\mathcal{H})} \otimes \mathcal{B}(|\psi\rangle\langle\psi|)\|_1.$$

Proof. We begin by noticing a simple fact: if $\mathcal{N} : L(\mathcal{H}) \rightarrow L(\mathcal{H}')$ is a Hermitian-preserving map (which means if ρ is Hermitian, then so is $\mathcal{N}(\rho)$). Then, for some state $|\psi\rangle \in \mathcal{H}^{\otimes 2}$,

$$\|\mathcal{N}\|_\diamond = \|\mathcal{N}\|_{\text{cb}} = \|I_{L(\mathcal{H})} \otimes \mathcal{N}(|\psi\rangle\langle\psi|)\|_1.$$

To see that, we recall that Hermitian-preserving operators, the completely bounded norm is maximized by a one-dimensional projection (see [56, Theorem 3.51]); in other words,

$$\|\mathcal{N}\|_{\text{cb}} = \sup_{\substack{|\psi\rangle \in \mathcal{H}^{\otimes 2} \\ \langle\psi|\psi\rangle=1}} \|I_{L(\mathcal{H})} \otimes \mathcal{N}(|\psi\rangle\langle\psi|)\|_1 \leq \sup_{\rho \in D(\mathcal{H})} \|I_{L(\mathcal{H})} \otimes \mathcal{N}(\rho)\|_1 = \|\mathcal{N}\|_\diamond, \quad (43)$$

and since the opposite inequality is trivially satisfied, equality holds. Note that, as it was shown in [8], $\mathcal{B}$ is the difference between two quantum channels and therefore is Hermitian preserving. Thus, it only remains to show that we can find such a $|\psi\rangle$, satisfying (43), which also belongs in $C^{\otimes 2}$. Indeed, note that for any $\rho = \sum_i \eta_i \otimes \tau_i \in L(\mathcal{H}^{\otimes 2})$ we can write

$$I_{L(\mathcal{H})} \otimes \mathcal{B}(\rho) = \sum_{k,l} \sum_i \text{tr}(\tau_i(PA_k^\dagger A_l P - \lambda_{k,l} P)) \cdot \eta_i \otimes |k\rangle\langle l|$$

$$\begin{aligned}
&= \sum_{k,l} \sum_i \text{tr}((P\tau_i P)(PA_k^\dagger A_l P - \lambda_{k,l} P)) \cdot \eta_i \otimes |k\rangle \langle l| \\
&= (I_{L(\mathcal{H})} \otimes \mathcal{B}) \circ (I_{L(\mathcal{H})} \otimes \mathcal{P})(\rho),
\end{aligned}$$

where $\mathcal{P}(\tau) = P\tau P$. Note that $(I_{L(\mathcal{H})} \otimes \mathcal{P})$ is a completely positive trace non-increasing map, whose image is $L(\mathcal{H} \otimes Q)$. Thus, for any $\rho \in D(\mathcal{H}^{\otimes 2})$ outside of the kernel, we have that the operator

$$\rho' = \frac{1}{\text{tr}(I_{L(\mathcal{H})} \otimes \mathcal{P}(\rho))} \cdot I_{L(\mathcal{H})} \otimes \mathcal{P}(\rho)$$

is a density operator in $D(\mathcal{H} \otimes Q)$, and

$$\|I_{L(\mathcal{H})} \otimes \mathcal{B}(\rho)\|_1 = \text{tr}(I_{L(\mathcal{H})} \otimes \mathcal{P}(\rho)) \cdot \|I_{L(\mathcal{H})} \otimes \mathcal{B}(\rho')\|_1 \leq \|I_{L(\mathcal{H})} \otimes \mathcal{B}(\rho')\|_1.$$

Thus,

$$\begin{aligned}
\|\mathcal{B}\|_\diamond &= \sup_{\rho \in D(\mathcal{H}^{\otimes 2})} \|I_{L(\mathcal{H})} \otimes \mathcal{B}(\rho)\|_1 \\
&\leq \sup_{\rho' \in D(\mathcal{H} \otimes Q)} \|I_{L(\mathcal{H})} \otimes \mathcal{B}(\rho')\|_1 \\
&= \sup_{\tau \in D(Q^{\otimes 2})} \|I_{L(Q)} \otimes \mathcal{B}(\tau)\|_1 \\
&= \sup_{|\psi_Q\rangle \in Q} \|I_{L(Q)} \otimes \mathcal{B}(|\psi_Q\rangle \langle \psi_Q|)\|_1,
\end{aligned} \tag{44}$$

where (44) follows since this expression is exactly the diamond norm of $\mathcal{B}$ when considered as a channel on $L(Q)$, and the diamond norm is equivalently defined by an additional optimization on all auxiliary reference systems of any finite dimension (see [57, Theorem 9.1.1]):

$$\|N\|_\diamond = \sup_{n \in \mathbb{N}} \sup_{\rho_n \in D(\mathbb{C}^n \otimes \mathcal{H})} \|I_{L(\mathbb{C}^n)} \otimes N(\rho_n)\|_1.$$

Finally, (b) follows from (43). This completes the proof. $\square$

A.2 Proof of Theorem 5

The idea of the proof is quite simple. Let $\mathcal{D}_{\leq t}$ be a recovery operation for Q under the truncated noise channel $\mathcal{N}_{\leq t}$ such that $\mathcal{F}_e(\mathcal{D}_{\leq t} \circ \mathcal{N}_{\leq t}|_Q, I_{L(Q)}) \geq 1 - \varepsilon^2$, which exists by the assumption. Consider the decoding operation $\mathcal{D}$ given by the following procedure:

- Given a state $\mathcal{N}(\rho)$, perform a measurement to determine whether ρ has lost no more than t photons.
- If that is the case, use the decoder $\mathcal{D}_{\leq t}$ on the resulting state.
- Otherwise, decode $\mathcal{N}(\rho)$ to some arbitrary state τ .

Let us proceed to implement this idea.

Defining the decoder. Consider a quantum channel

$$\begin{aligned}
\mathcal{M} : \mathcal{H}_{q,N} &\rightarrow \text{Span} \left(\bigcup_{r \leq N} \mathcal{H}_{q,r} \right) \otimes \mathbb{C}^2 \\
\rho &\mapsto P_{\leq t} \rho P_{\leq t} \otimes |0\rangle \langle 0| + \text{tr}((I - P_{\leq t}) \rho) \cdot \tau \otimes |1\rangle \langle 1|,
\end{aligned}$$

where $|0\rangle$ and $|1\rangle$ are ancillas representing *success* and *fail* flag states, $P_{\leq t}$ is the projection on $\text{Span}\left(\bigcup_{i=N-t}^N \mathcal{H}_{q,i}\right)$, and $\tau = |\psi_\tau\rangle\langle\psi_\tau|$ is an arbitrary pure state in Q . It is easy to check that $\mathcal{M}$ is completely positive, with Kraus operators

$$E_0 = P_{\leq t} \otimes |0\rangle, \quad E_{\underline{n}} = |\psi_\tau\rangle\langle\underline{n}|(I - P_{\leq t}) \otimes |1\rangle, \quad \underline{n} \in \bigcup_{r \leq N} \mathcal{S}_{q,r}.$$

$\mathcal{M}$ is also trace preserving:

$$\begin{aligned} \text{tr}(\mathcal{M}(\rho)) &= \text{tr}(P_{\leq t} \rho P_{\leq t} \otimes |0\rangle\langle 0|) + \text{tr}((I - P_{\leq t}) \rho) \cdot \text{tr}(\tau \otimes |1\rangle\langle 1|) \\ &= \text{tr}(P_{\leq t} \rho) + \text{tr}((I - P_{\leq t}) \rho) = \text{tr}(\rho). \end{aligned}$$

Let P_0 be the projection on the success space $\text{Span}\left(\bigcup_{r \leq N} \mathcal{H}_{q,r}\right) \otimes \text{Span}(|0\rangle)$ and $P_1 = I - P_0$. Define $\tilde{\mathcal{D}} : \text{Span}\left(\bigcup_{r \leq N} \mathcal{H}_{q,r}\right) \otimes \mathbb{C}^2 \rightarrow \mathcal{H}_{q,N}$ given by

$$\tilde{\mathcal{D}}(\rho) = \mathcal{D}_{\leq t}(\text{tr}_{\mathbb{C}^2}(P_0 \rho P_0)) + \text{tr}_{\mathbb{C}^2}(P_1 \rho P_1),$$

where $\text{tr}_{\mathbb{C}^2}(\cdot)$ denotes the operation of tracing out the success/failure qubit, given by partial trace (which is CPTP). Note that the maps $\rho \rightarrow P_i \rho P_i, i = 0, 1$ are trivially CP, and thus each of the maps $\rho \mapsto \text{tr}_{\mathbb{C}^2}(P_i \rho P_i)$ and $\rho \mapsto \mathcal{D}_{\leq t}(\text{tr}_{\mathbb{C}^2}(P_0 \rho P_0))$ is also CP as a composition of CP maps (since $\mathcal{D}_{\leq t}$ is a quantum channel, it is CPTP). This shows that $\tilde{\mathcal{D}}$ is CP. Furthermore, since $\mathcal{D}$ is CPTP we have

$$\text{tr}(\tilde{\mathcal{D}}(\rho)) = \text{tr}(\mathcal{D}_{\leq t}(\text{tr}_{\mathbb{C}^2}(P_0 \rho P_0))) + \text{tr}(\text{tr}_{\mathbb{C}^2}(P_1 \rho P_1)) = \text{tr}(P_0 \rho P_0) + \text{tr}(P_1 \rho P_1) = \text{tr}(\rho),$$

To conclude, we have shown that $\tilde{\mathcal{D}}$ is CPTP. We now define the decoding operation to be $\mathcal{D} = \tilde{\mathcal{D}} \circ \mathcal{M}$ and note that it is CPTP as a composition of two CPTP operators.

Action on noisy states. For any state $\rho \in D(\mathcal{H}_{q,N})$ and error operator $A_{\underline{r}} \in \mathcal{S}_{q,r}$,

$$A_{\underline{r}} \rho A_{\underline{r}}^\dagger \in \begin{cases} \mathcal{H}_{q,N-r} & r \leq N \\ \{0\} & r > N, \end{cases}$$

and in particular,

$$P_{\leq t}(A_{\underline{r}} \rho A_{\underline{r}}^\dagger) P_{\leq t} = \begin{cases} A_{\underline{r}} \rho A_{\underline{r}}^\dagger & r \leq t \\ 0 & r > t. \end{cases}$$

Thus,

$$\begin{aligned} \mathcal{M} \circ \mathcal{N}(\rho) &= \sum_{\substack{\underline{r} \in \mathcal{S}_{q,r} \\ r \leq t}} \mathcal{M}(A_{\underline{r}} \rho A_{\underline{r}}^\dagger) + \sum_{\substack{\underline{r} \in \mathcal{S}_{q,r} \\ r > t}} \mathcal{M}(A_{\underline{r}} \rho A_{\underline{r}}^\dagger) \\ &= \sum_{\substack{\underline{r} \in \mathcal{S}_{q,r} \\ r \leq t}} A_{\underline{r}} \rho A_{\underline{r}}^\dagger \otimes |0\rangle\langle 0| + \text{tr}\left(\sum_{\substack{\underline{r} \in \mathcal{S}_{q,r} \\ r > t}} A_{\underline{r}} \rho A_{\underline{r}}^\dagger\right) \cdot \tau \otimes |1\rangle\langle 1| \\ &= p_{N,t} \cdot \mathcal{N}_{\leq t}(\rho) \otimes |0\rangle\langle 0| + (1 - p_{N,t}) \cdot \tau \otimes |1\rangle\langle 1|, \end{aligned}$$

where the last equality follows (13) and the completeness condition $\sum_{\underline{r} \in \mathbb{Z}_0^q} A_{\underline{r}}^\dagger A_{\underline{r}} = I$. We can now derive the action of the decoder on noisy states on the output of $\mathcal{N}$:

$$\begin{aligned} \mathcal{D} \circ \mathcal{N}(\rho) &= \tilde{\mathcal{D}}(\mathcal{M} \circ \mathcal{N}(\rho)) \\ &= \tilde{\mathcal{D}}(p_{N,t} \cdot \mathcal{N}_{\leq t}(\rho) \otimes |0\rangle\langle 0| + (1 - p_{N,t}) \cdot \tau \otimes |1\rangle\langle 1|) \\ &= p_{N,t} \cdot \mathcal{D}_{\leq t} \circ \mathcal{N}_{\leq t}(\rho) + (1 - p_{N,t}) \cdot \tau. \end{aligned} \tag{45}$$

It is easy to verify that the unique linear map that acts on density operators as (45) acts on general linear maps $X \in L(Q)$ as

$$\mathcal{D} \circ \mathcal{N}(X) = p_{N,t} \cdot \mathcal{D}_{\leq t} \circ \mathcal{N}_{\leq t}(X) + (1 - p_{N,t}) \text{tr}(X) \cdot \tau.$$

where $\mathcal{N}(X)$ denotes the action of the AD channel on a general linear map. *Computing the entanglement fidelity.* For a density operator $\rho \in D(Q)$, let ψ be its purification. Let $|\psi\rangle = \sum_{i=0}^{K-1} p_i |i\rangle \otimes |i'\rangle$ be a Schmidt decomposition of $|\psi\rangle$, where $\{|i\rangle\}$ is an orthonormal basis of Q , $p_i \geq 0$ for all i , and $\sum_i p_i^2 = 1$. We set out to compute the entanglement fidelity $\mathcal{F}_e$, (4). First, observe that

$$\begin{aligned} \mathcal{F}(I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N}(|\psi\rangle\langle\psi|), |\psi\rangle\langle\psi|) &= \langle\psi| I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N}(|\psi\rangle\langle\psi|) |\psi\rangle \\ &= \langle\psi| I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N}\left(\sum_{i,j} p_i p_j |i\rangle\langle j| \otimes |i'\rangle\langle j'|\right) |\psi\rangle \\ &= \langle\psi| \left(\sum_{i,j} p_i p_j |i\rangle\langle j| \otimes \left(p_{N,t} \mathcal{D}_{\leq t} \circ \mathcal{N}_{\leq t}(|i'\rangle\langle j'|) + (1 - p_{N,t}) \cdot \tau\right)\right) |\psi\rangle \\ &= \langle\psi| \left(\sum_{i,j} p_i p_j |i\rangle\langle j| \otimes \left(p_{N,t} \mathcal{D}_{\leq t} \circ \mathcal{N}_{\leq t}(|i'\rangle\langle j'|) + (1 - p_{N,t}) \text{tr}(|i'\rangle\langle j'|) \cdot \tau\right)\right) |\psi\rangle \\ &= p_{N,t} \cdot \langle\psi| I_{L(Q)} \otimes \mathcal{D}_{\leq t} \circ \mathcal{N}_{\leq t}(|\psi\rangle\langle\psi|) |\psi\rangle + (1 - p_{N,t}) \cdot \langle\psi| \tilde{\tau} |\psi\rangle, \end{aligned}$$

where

$$\tilde{\tau} = \rho_\psi \otimes \tau, \quad \rho_\psi = \sum_{i=0}^{K-1} p_i^2 \cdot |i\rangle\langle i|.$$

Since $\tilde{\tau}$ is positive semidefinite, we have $\langle\psi| \tilde{\tau} |\psi\rangle \geq 0$, so the second term can be discarded, and we obtain

$$\begin{aligned} \mathcal{F}(I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N}(|\psi\rangle\langle\psi|), |\psi\rangle\langle\psi|) &\geq p_{N,t} \cdot \langle\psi| I_{L(Q)} \otimes \mathcal{D}_{\leq t} \circ \mathcal{N}_{\leq t}(|\psi\rangle\langle\psi|) |\psi\rangle \\ &\geq p_{N,t} \cdot (1 - \varepsilon^2), \end{aligned}$$

where the last inequality follows from the assumption of the theorem (that the entanglement fidelity for $\mathcal{D}_{\leq t}$ is close to 1 under the action of the truncated noise $\mathcal{N}_{\leq t}$. Since this is true for an arbitrary pure state in $C^{\otimes 2}$, it is also true for the infimum in (4), completing the forward part.

Let us now prove the converse. Assume that Q is ε -AQECC for $\mathcal{N}$, with a decoding operation $\mathcal{D}$ such that $\mathcal{F}_e(\mathcal{D} \circ \mathcal{N}|_Q, I_{L(Q)}) \geq 1 - \varepsilon^2$. We will show that

$$\mathcal{F}_e(\mathcal{D} \circ \mathcal{N}_{\leq t}|_Q, I_{L(Q)}) \geq 1 - \frac{\varepsilon^2}{p_{N,t}}.$$

We will also use the “complementary channel”, with more than t errors, defined as

$$\mathcal{N}_{>t}(\rho) = \frac{1}{1 - p_{N,t}} \cdot \sum_{\substack{\underline{r} \in \mathcal{S}_{q,r} \\ r > t}} A_{\underline{r}} \rho A_{\underline{r}}^\dagger. \quad (46)$$

By the same argument as in Proposition 4, we claim that $\mathcal{N}_{>t}$ is a quantum channel on $\mathcal{H}_{q,N}$ (as for any $\rho \in L(Q)$ we have $\text{tr}(\rho) = 1 - p_{N,t}$). We also observe that

$$\mathcal{N}(\rho) = p_{N,t} \cdot \mathcal{N}_{\leq t}(\rho) + (1 - p_{N,t}) \cdot \mathcal{N}_{>t}(\rho).$$

Fix a pure state $|\psi\rangle \in Q^{\otimes 2}$. We will show that

$$\mathcal{F}(I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N}_{\leq t}(|\psi\rangle\langle\psi|), |\psi\rangle\langle\psi|) \geq 1 - \frac{\varepsilon^2}{p_{N,t}}. \quad (47)$$

Note that we may write $I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N} = (I_{L(Q)} \otimes \mathcal{D}) \circ (I_{L(Q)} \otimes \mathcal{N})$. Thus, since fidelity is linear whenever one of the involved states is pure, we have:

$$\begin{aligned}
1 - \varepsilon^2 &\leq \mathcal{F}(I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N}(|\psi\rangle\langle\psi|), |\psi\rangle\langle\psi|) \\
&= \mathcal{F}((I_{L(Q)} \otimes \mathcal{D}) \circ (I_{L(Q)} \otimes \mathcal{N})(|\psi\rangle\langle\psi|), |\psi\rangle\langle\psi|) \\
&= \mathcal{F}((I_{L(Q)} \otimes \mathcal{D}) \circ (I_{L(Q)} \otimes (p_{N,t} \cdot \mathcal{N}_{\leq t} + (1 - p_{N,t})\mathcal{N}_{>t}))(|\psi\rangle\langle\psi|), |\psi\rangle\langle\psi|) \\
&= p_{N,t} \cdot \mathcal{F}((I_{L(Q)} \otimes \mathcal{D}) \circ (I_{L(Q)} \otimes \mathcal{N}_{\leq t})(|\psi\rangle\langle\psi|), |\psi\rangle\langle\psi|) \\
&\quad + (1 - p_{N,t}) \cdot \mathcal{F}((I_{L(Q)} \otimes \mathcal{D}) \circ (I_{L(Q)} \otimes \mathcal{N}_{>t})(|\psi\rangle\langle\psi|), |\psi\rangle\langle\psi|) \\
&\leq p_{N,t} \cdot \mathcal{F}(I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N}_{\leq t}(|\psi\rangle\langle\psi|), |\psi\rangle\langle\psi|) + (1 - p_{N,t}),
\end{aligned}$$

where the last inequality follows since $I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N}_{>t}$ is a quantum channel, and in particular $I_{L(Q)} \otimes \mathcal{D} \circ \mathcal{N}_{\leq t}(|\psi\rangle\langle\psi|)$ is a quantum state (and therefore the fidelity is bounded by 1). Rearranging the above inequality, we obtain (47), which completes the proof.

A.3 Typical vectors in the simplex

For $\underline{x} \in \mathcal{S}_{q,N}$ let

$$\text{supp}(\underline{x}) = \{i \in [q] : x_i \neq 0\}, \quad \|\underline{x}\|_\infty = \max_{i \in [q]} |x_i|.$$

Lemma 23 (Tail bound for ∞ -norm on the simplex). *Let $\underline{X} = (X_0, \dots, X_{q-1})$ be a uniform random vector in $\mathcal{S}_{q,N}$, $q \in \mathbb{N}$, and let $B \in \{0, 1, \dots, N\}$ be an integer. Then*

$$\mathbb{P}[\|\underline{X}\|_\infty \geq B] \leq q \cdot \frac{\binom{N-B+q-1}{q-1}}{\binom{N+q-1}{q-1}}.$$

Proof. Let us fix some $i \in [q]$ and bound the probability $\mathbb{P}[X_i \geq B]$. We start by observing that

$$\mathbb{P}[X_i \geq B] = \frac{|\{\underline{n} \in \mathcal{S}_{q,N} : n_i \geq B\}|}{|\mathcal{S}_{q,N}|} = \frac{\binom{N-B+q-1}{q-1}}{\binom{N+q-1}{q-1}},$$

where the last equality follows since once B balls are placed, we are left to distribute $N - B$ balls among q bins. Next,

$$\mathbb{P}[\max_i X_i \geq B] \leq \sum_{i=1}^q \mathbb{P}[X_i \geq B] = q \cdot \frac{\binom{N-B+q-1}{q-1}}{\binom{N+q-1}{q-1}}. \quad \square$$

Lemma 24 (Tail bound in the support size on the simplex). *Let $\underline{X} = (X_0, \dots, X_{q-1})$ be a uniform random vector in $\mathcal{S}_{q,N}$, $q \in \mathbb{N}$, and let $\varepsilon > 0$. Let $S = \frac{1}{q}|\text{supp}(\underline{X})|$. Then*

$$\mathbb{P}\left[\left|S - \frac{N}{N+q-1}\right| > \varepsilon\right] = \mathbb{P}[|S - \mathbb{E}[S]| > \varepsilon] \leq \frac{N(N-1)(q-1)}{\varepsilon^2 q(N+q-1)^2(N+q-2)}. \quad (48)$$

Proof. Let $Y = 1 - S$, be the normalized number of zero coordinates of $\underline{X}$:

$$Y = \frac{1}{q} \sum_{i=0}^{q-1} Y_i, \quad Y_i := \mathbb{1}_{\{X_i=0\}}.$$

Note that

$$\mathbb{E}[Y_i] = \mathbb{P}[X_i = 0] = \frac{\binom{N+q-2}{q-2}}{\binom{N+q-1}{q-1}} = \frac{q-1}{N+q-1}.$$

and therefore

$$E[S] = 1 - \mathbb{E}[Y] = 1 - \frac{1}{q} \sum_{i=0}^{q-1} E[Y_i] = \frac{N}{N+q-1}.$$

Similarly, we compute the variance:

$$\text{Var}(S) = \text{Var}(1 - Y) = \text{Var}(Y) = \frac{1}{q^2} \left(\sum_{i=1}^q \text{Var}(Y_i) + \sum_{i \neq j} \text{cov}(Y_i, Y_j) \right).$$

Since Y_i is a $\text{Ber}(p)$ random variable with $p = \frac{q-1}{N+q-1}$ we have $\text{Var}(Y_i) = p(1-p)$. For the cross terms, we have:

$$\begin{aligned} \text{cov}(Y_i, Y_j) &= \mathbb{E}[Y_i Y_j] - E[Y_i] E[Y_j] = \mathbb{P}[X_i = X_j = 0] - \left(\frac{q-1}{N+q-1} \right)^2 \\ &= \frac{\binom{N+q-3}{q-3}}{\binom{N+q-1}{q-1}} - \left(\frac{q-1}{N+q-1} \right)^2 = -\frac{(q-1)N}{(N+q-1)^2(N+q-2)} \end{aligned}$$

after simplifications. This yields

$$\begin{aligned} \text{Var}(S) &= \frac{1}{q^2} \left(q \cdot \frac{q-1}{N+q-1} \left(1 - \frac{q-1}{N+q-1} \right) - q(q-1) \cdot \frac{(q-1)N}{(N+q-1)^2(N+q-2)} \right) \\ &= \frac{N(q-1)}{q(N+q-1)} \left(\frac{N-1}{(N+q-1)(N+q-2)} \right). \end{aligned}$$

By Chebyshev's inequality, $\mathbb{P}[|S - \mathbb{E}[S]| > \varepsilon] \leq \frac{\text{Var}(S)}{\varepsilon^2}$, and substitutions yield (48). $\square$

In the next lemma, we compute a bound on the volume of the ball with center of bounded support.

Lemma 25. *Let $\underline{n} \in S_{q,N}$ be a vector with $|\text{supp}(\underline{n})| = m$. Then*

$$|B(\underline{n}, r)| \leq \sum_{j=0}^r \binom{j+m-1}{m-1} \binom{j+q-1}{q-1}.$$

Proof. Let $|\text{supp}(\underline{n})| = m$ and let $\underline{n}'$ be such that $d(\underline{n}, \underline{n}') = j$. Let $A_1 = \{i : \underline{n}_i > \underline{n}'_i\}$ and $A_2 = \{i : \underline{n}_i < \underline{n}'_i\}$. Since $\sum_i \underline{n}_i = \sum_i \underline{n}'_i$, we have that $\sum_{i \in A_1} (\underline{n}_i - \underline{n}'_i) = \sum_{i \in A_2} (\underline{n}'_i - \underline{n}_i) = j$. Thus, we must place j units within $A_1 \subset \text{supp}(\underline{n})$ and the other j units in $A_2 \subset [q] \setminus A_1$. Since we do not have control over the size of A_2 , the best upper bound on the number of points $\underline{n}'$ that we can get is $\binom{j+m-1}{m-1} \binom{j+q-1}{q-1}$. Summing over all j , we obtain the claimed estimate. $\square$

A.4 Combinatorial and probabilistic results

Lemma 26. (The continuous mapping theorem, [10, Theorem 2.7]) *Let $(S_n)_n$ be a sequence of random variables with values in a metric space (M, d) , and assume that $T : M \rightarrow M'$ is a continuous map to another metric space (M', d') . Then if S_n converge to some variable S in distribution, then $T(S_n)$ converge to $T(S)$ in distribution.*

Another useful result concerns the convergence of expectations for sequences that converge in distribution:

Lemma 27. [10, Theorem 3.5] *Let $(S_n)_n$ be a sequence of random variables that converge in distribution to S . Then $\mathbb{E}[S_n] \rightarrow \mathbb{E}[S]$ if $(S_n)_n$ are uniformly integrable, that is,*

$$\sup_n \mathbb{E}[|S_n|] < \infty.$$

A function $f : A_1, \dots, A_n$ and $f : A_1 \times \dots \times A_n \rightarrow \mathbb{R}$ is said to satisfy the *bounded differences property* if changing the value of the i th coordinate x_i changes the value of f by at most c_i . Formally, this means that

$$\max_{x'_i \in A_i} |f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n)| \leq c_i \quad (49)$$

for all $i \in \{1, \dots, n\}$.

Lemma 28. (McDiarmid's inequality, [36, Theorem 3.1]; [11, Sec. 6.1]) *Let $X_1, \dots, X_n$ be independent random variables taking values in $A_1, \dots, A_n$ and $f : A_1 \times \dots \times A_n \rightarrow \mathbb{R}$ be a measurable function with bounded differences $c_1, \dots, c_n$. Then*

$$\mathbb{P}[|f(\mathbf{X}) - \mathbb{E}[f(\mathbf{X})]| > \varepsilon] \leq 2 \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n c_i^2}\right).$$

Lemma 29. (Hoeffding's inequality, [23, Theorem 2]; [11, Lemma 2.2]) *Let $X_1, \dots, X_n$ be independent random variables such that $a_i \leq X_i \leq b_i$ with probability 1. Then*

$$\mathbb{P}\left[\left|\sum_{i=1}^n (X_i - \mathbb{E}[X_i])\right| > \varepsilon\right] \leq 2 \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \quad (50)$$

Finally, we repeatedly use the asymptotic approximation for binomial coefficients: for a fixed α and $N \rightarrow \infty$,

$$\binom{N}{\lfloor \alpha N \rfloor} = 2^{N(h_2(\alpha) + o(1))}, \quad (51)$$

where $h_2(\cdot)$ is the binary entropy function; e.g., [34, Ch. 10, Lemma 7].

A.5 Additional proofs

Proof of Corollary 8. For $\underline{r} \in \mathcal{S}_{q,r}$ let $\Pi_{\underline{r}}$ be the orthogonal projection on the space $A_{\underline{r}}(Q_N) = \{A_{\underline{r}}|\underline{n}\rangle : \underline{n} \in C_N\}$, and let $\Pi_{t>1} = I - \sum_{r < t} \Pi_r$. Let $|\psi\rangle \in Q_N$ be any input state, and let $\underline{r} \in \mathcal{S}_{q,r}$, $r \leq t$ be fixed. Note that the noisy state $\mathcal{N}(|\psi\rangle\langle\psi|)$ satisfies

$$\Pi_{\underline{r}}\mathcal{N}(|\psi\rangle\langle\psi|)\Pi_{\underline{r}} = \Pi_{\underline{r}}\left\{(1 - p_{N,t}) \cdot \mathcal{N}_{>t}(|\psi\rangle\langle\psi|) + \sum_{\substack{\underline{r}' \in \mathcal{S}_{q,r'} \\ r' \leq t}} A_{\underline{r}'}|\psi\rangle\langle\psi|A_{\underline{r}'}^\dagger\right\}\Pi_{\underline{r}}$$

where the channel $\mathcal{N}_{>t}$ is defined in (46). We observe that $\mathcal{N}_{>t}$ maps states in $L(\mathcal{H}_{q,N})$ to states supported on $\mathcal{H}_{q,<N-t} \triangleq \text{Span}\{\bigcup_{0 < i < N-t} \mathcal{H}_{q,i}\}$. We also observe that $A_{\underline{r}}(Q_N) \subseteq \mathcal{H}_{q,N-r} \subset \mathcal{H}_{q,<N-t}^\perp$ and therefore $\Pi_{\underline{r}}\mathcal{N}(|\psi\rangle\langle\psi|)\Pi_{\underline{r}} = 0$. By (15), any error operators $A_{\underline{r}'}$, $\underline{r} \neq \underline{r}' \in \mathcal{S}_{q,r}$, $r \leq t$, maps the code to orthogonal subspaces of $\mathcal{H}_q$. In particular, $\Pi_{\underline{r}}A_{\underline{r}'}|\psi\rangle = 0$. Combining these observations, we obtain:

$$\begin{aligned} \Pi_{\underline{r}}\mathcal{N}(|\psi\rangle\langle\psi|)\Pi_{\underline{r}} &= (1 - p_{N,t}) \cdot \overbrace{\Pi_{\underline{r}}\mathcal{N}_{>t}(|\psi\rangle\langle\psi|)\Pi_{\underline{r}}}^0 \\ &\quad + \sum_{\substack{\underline{r}' \in \mathcal{S}_{q,r} \\ r' \leq t}} \overbrace{\Pi_{\underline{r}}A_{\underline{r}'}\mathcal{N}(|\psi\rangle\langle\psi|)A_{\underline{r}'}^\dagger\Pi_{\underline{r}}}^{\delta_{\underline{r},\underline{r}'} \cdot A_{\underline{r}}|\psi\rangle\langle\psi|A_{\underline{r}}^\dagger} = A_{\underline{r}}|\psi\rangle\langle\psi|A_{\underline{r}}^\dagger. \end{aligned}$$

We also note that the operator $\Pi_{>t}$ is measured with probability:

$$\text{tr}(\Pi_{>t} \mathcal{N}(|\psi\rangle\langle\psi|) \Pi_{>t}) = (1 - p_{N,t}) \cdot \text{tr}(\mathcal{N}_{>t}(|\psi\rangle\langle\psi|)) = 1 - p_{N,t}. \quad (52)$$

We conclude by observing that we do not need the full power of the condition (15), only that $\langle c_i | A_{\underline{r}} A_{\underline{r}'}^\dagger | c_j \rangle = 0$ holds for $\underline{r} \neq \underline{r}'$. Note that (52) holds in general, without any assumption on I_i and I_j . Note that if $\underline{r} \neq \underline{r}'$ the inner products $\langle \underline{n} - \underline{r} | \underline{n}' - \underline{r}' \rangle$ in (52) must be all zero, since $\underline{n} \neq \underline{n}' + \underline{r} - \underline{r}'$ as $\underline{r} - \underline{r}'$ is a nonzero vector of ℓ_1 norm at most $2t$ and $d(\underline{n}, \underline{n}') > t$ by the distance assumption. $\square$

Proof of Proposition 4. We note that $\mathcal{N}_{\leq t}$ is a completely positive map since any map that admits a Kraus representation is such. To show that it is also trace-preserving, let $X \in L(\mathcal{H}_{q,N})$ be any operator. It is always possible to decompose it as $X = T_1 + i \cdot T_2$, where T_1 and T_2 are Hermitian. Since $\mathcal{H}_{q,N}$ is finite-dimensional, these operators can be written as $T_l = \sum_j \lambda_{l,j} |\psi_{l,j}\rangle\langle\psi_{l,j}|$, $l = 1, 2$. By Proposition 3 we have

$$\begin{aligned} \text{tr}(\mathcal{N}_{\leq t}(X)) &= \text{tr}(\mathcal{N}_{\leq t}(T_1 + i \cdot T_2)) \\ &= \frac{1}{p_{N,t}} \cdot \sum_j \left(\lambda_{1,j} \cdot \sum_{\substack{\underline{r} \in \mathcal{S}_{q,r} \\ r \leq t}} \langle \psi_{1,j} | A_{\underline{r}}^\dagger A_{\underline{r}} | \psi_{1,j} \rangle + i \cdot \lambda_{2,j} \cdot \sum_{\substack{\underline{r} \in \mathcal{S}_{q,r} \\ r \leq t}} \langle \psi_{2,j} | A_{\underline{r}}^\dagger A_{\underline{r}} | \psi_{2,j} \rangle \right) \\ &= \frac{1}{p_{N,t}} \cdot \sum_j (\lambda_{1,j} \cdot p_{N,t} + i \cdot \lambda_{2,j} \cdot p_{N,t}) = \text{tr}(T_1) + i \cdot \text{tr}(T_2) = \text{tr}(X). \quad \square \end{aligned}$$