THE UNIVERSITY OF CHICAGO

QUANTUM ERROR CORRECTION: FROM MEMORY TO LOGIC

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE PRITZKER SCHOOL OF MOLECULAR ENGINEERING
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
GUO ZHENG

CHICAGO, ILLINOIS
DECEMBER 2026

© 2026 by Guo Zheng

ABSTRACT

Quantum computers possess computational power distinct from classical devices, using phase coherence and entanglement. This coherence is fragile and susceptible to environmental noise. Protecting quantum information is a central goal of quantum error correction and fault tolerance, which are essential for scaling quantum devices to utility-scale computation. Such protection, however, often requires a space-time overhead formidable for current experimental platforms. A key question is the minimum space-time cost of protecting a quantum information-processing task and how to attain it. This thesis investigates these questions from quantum memory to logical gates.

For quantum memory, we consider preserving information while it idles under noise. We first develop the near-optimal fidelity, an optimization-free general framework that bounds the optimal performance of a given encoding and noise channel and provides an achievable recovery benchmark. We apply it to Gottesman-Kitaev-Preskill (GKP) codes, which encode information in quantum harmonic oscillators through lattice geometry. We show that finite-size GKP memories can substantially outperform existing recovery schemes. In the asymptotic system size and energy limit, GKP families achieve the capacities of pure-loss and quantum-limited amplification channels, attaining the minimum encoding redundancy permitted by these channels.

For logical gates, we study quantum low-density parity-check codes. We introduce high-rate surgery, which performs parallel, addressable logical Pauli-product measurements with low space-time overhead. Our analytical constructions and complementary numerical benchmarks indicate that logical processing can retain nearly constant space-time overhead relative to quantum memory. These results provide efficient building blocks for utility-scale quantum devices.

To my parents

TABLE OF CONTENTS

LIST OF FIGURES vii

LIST OF TABLES xii

ACKNOWLEDGMENTS xiii

1 Introduction 1

2 Near-optimal performance of quantum error correction codes 6
- 2.1 Background 8
- 2.2 Near-optimal fidelity 13
- 2.3 Infidelity proxy 15
- 2.4 Examples 18
- 2.5 Discussion 23

3 Achieving coherent information rates through Gottesman-Kitaev-Preskill codes 24
- 3.1 GKP codes and lattice theory 25
- 3.2 Gaussian-dilatable noise channels 30
- 3.3 Performance of GKP codes 35
- 3.4 Achievable rates 43
- 3.5 Lattice geometry beyond sphere packing 51
- 3.6 Discussion 55

4 High-rate surgery for quantum low-density parity-check codes 57
- 4.1 Code surgery framework 60
- 4.2 Algebraic construction for structured codes 67
- 4.3 Randomized construction for generic codes 70
- 4.4 Methods 74
- 4.5 Discussion 81

5 Conclusion and outlook 83

A Appendices for Chapter 2 85
- A.1 Convex optimization for recovery 85
- A.2 Transpose channel 86
- A.3 Derivation of the near-optimal fidelity expression 93
- A.4 Derivation of the perturbative form of the near-optimal fidelity and its properties 95
- A.5 Definitions and notations for examples of applications 99
- A.6 Thermodynamic code and derivations 101
- A.7 Infidelity proxy for a general complementary reference 108

B Appendices for Chapter 3 . . . 111
B.1 Basics on information theory . . . 111
B.2 Basics on lattice theory . . . 111
B.3 Ideal GKP code and stabilizer phases . . . 112
B.4 Noise channels and their dilations . . . 113
B.5 Finite-energy GKP codes . . . 114
B.6 Infidelity proxy of GKP codes . . . 115
B.7 Closed-form performance under thermal loss . . . 127
B.8 Vacuum specializations of the general expansion . . . 131
B.9 Energy monotonicity of the full infidelity proxy . . . 132
B.10 Achievable rates . . . 136
B.11 Lattice geometry beyond sphere packing . . . 148
B.12 Vacuum recovery protocols and simulation details . . . 152
B.13 Approximate implementation of the near-optimal recovery. . . . 153
B.14 Numerical results . . . 154
B.15 Mathematical identities . . . 157

C Appendices for Chapter 4 . . . 164
C.1 Notation and background . . . 164
C.2 Expansion properties . . . 167
C.3 General framework for code surgeries . . . 168
C.4 High-rate surgery through structured ancilla . . . 179
C.5 Algorithm design for randomized ancilla construction and simulation details 193
C.6 Measurement of mixed-type logicals . . . 200
C.7 Constructions of spatially coupled codes . . . 202

REFERENCES . . . 203

LIST OF FIGURES

2.1 Encoding, noise, and recovery 8
2.2 QEC matrix and complementary output 12
2.3 Optimal infidelity for qubit codes [[4, 1, 2]], ((5, 6, 2)), and [[9, 1, 3]] and the GKP code with $\bar{n}_C = 7$. The noise channel is amplitude damping noise, i.e. loss for oscillators. The shaded regions represent the optimal infidelity intervals bounded by the two-sided bound given by the near-optimal infidelity in Eq. (2.16). 19
2.4 The near-optimal infidelity of the thermodynamic code with distance d (see Appendix A for the definition of the codewords) and $l = 1$, i.e. single erasure error. The stars represent the exact approach where the QEC matrix is numerically computed, and the circles represent the perturbative expression given in Eq. (2.26). 20
2.5 The entanglement infidelity of square lattice GKP qubit code under loss with transmissivity $\eta = 0.9$ (10% loss). All shaded regions represent the two-sided bound on the optimal fidelity. The red circles represent the exact form of the near-optimal infidelity, where the QEC matrix is computed analytically. The blue dashed line is the perturbative form, evaluated through the closed-form analytical expression derived in Chapter 3. The inset figure shows the performance of GKP codes (square lattice), cat codes (S=4), and binomial codes (S=1) (see Appendix A). The squares (dashed lines) represent the optimal (near-optimal) fidelity. 22

3.1 Square and hexagonal GKP lattices 28
3.2 Gaussian-unitary dilation and complementary output 30
3.3 Achievable GKP rates and finite-energy performance under thermal loss. Rates are in bits per channel use. (a) At $\eta = 0.9$, the attained rate (thick pale line) coincides with the infinite-energy coherent information of a thermal state (dashed), against the pre-amplification GKP construction [1] and the minimum of three established upper bounds [2]. (b) Continuous rates versus the integer-scaling staircase. (c) At $\eta = 0.95$, points show near-optimal infidelity $1 - \tilde{F}^{\mathrm{opt}}$ for square qubit codes at 5–25 mean photons. Dashed curves are the separate D_0-reference infidelity proxy $\tilde{\epsilon}_{D_0}$, from the exact Gram-orthonormalized finite-energy lattice-pair integral, extended to 3000 photons. Dotted lines are its analytic infinite-energy limits, not near-optimal-fidelity asymptotes. 36

3.4 The performance of GKP codes under (a, b) loss and (c, d) amplification. (a, c) The infidelities of near-optimal recovery are compared against that of existing GKP decoders, such as AD [1, 3], the conventional decoder [4], and the sBs circuit [5]. In particular, the sBs circuit is repeated until the system reaches a stabilized state. The GKP codes are based on square lattice. The noise channels are set to be (a) $\eta = 0.9$ (10% loss) for loss and (c) $G = 1.1$ for amplification. (b, d) The performance of various lattices with parameters given in Tab. 3.3 is presented. The encoding dimension is fixed to $d_L = 2$ with $\bar{n}_C/N = 5$ average photon number per mode. 42
3.5 Gaussian-channel landscape and rate scope 47
3.6 (a) The infidelities of GKP code are based on a few lattices with varying numbers of modes with a constant encoding rate such that $d_L = 3^N$. (b) Coherent information of GKP codes is based on various lattices, which are achievable rates. The coherent information is computed through the Werner states as given in Eq. (3.49). (c) Lower bounds on the achievable rates of GKP codes based on scaled symplectically self-dual lattices with an increasing number of modes, N. The infidelity is upper bounded through Eq. (3.51). (d) A comparison of the rates achieved through the amplification decoder, the near-optimal decoder, and the single-mode hexagonal GKP code concatenated with random stabilizer codes for asymptotic system sizes. All subplots are focusing on infinite-energy GKP codes. 48
3.7 Environmental overlap nodes change the preferred GKP lattice geometry. The environment is $|m\rangle$, $m = 0, 1, 2$ [6]. (a) Exact kernels κ_m; triangles mark Laguerre zeros. (b) Full infinite-energy proxy sums for rectangular qubit dual lattices $\Lambda_s^{\perp} = \mathrm{diag}(s, s^{-1})\mathbb{Z}^2/\sqrt{2}$ at $\eta = 0.65$, divided by the square-lattice proxy at $s = 1$ for the same environment. Curves show the normalized infidelity proxy, not the actual infidelity. Dots mark numerical rectangular minima on $1 \le s \le 4$. (c)–(e) Fidelity-optimized dual lattices in β coordinates, obtained by minimizing the infinite-energy infidelity proxy with $s \ge 1$. Dashed rings mark Laguerre zeros, shading marks one cell, and the cross marks the excluded origin. The cell aspect ratio is s^2. 54

4.1 Illustration of high-rate versus low-rate surgery schemes. (a) Comparison of low-rate (sequential [7, 8] and parallel [9, 10]) and high-rate surgery scheme (this work). The horizontal axis represent time axis. Blue circles denote ancilla systems that couple to multiple target logical operators, whose supports are shown by the green, orange, and pink circles, possibly with overlaps. (b) The space (vertical axis) and time (horizontal) costs of surgery schemes on Gross code [11], $[[144, 12, 12]]$, for measuring nine randomly chosen logicals, compared against memory. Space costs are measured in unit of physical qubits. Time costs are in units of logical cycles. 58

4.2 Tanner graph illustration of a general surgery scheme. Checks (resp. qubits) are represented by squares (resp. circles) [7]. The shown figure presents the ancilla system merged with the data system to measure X-logical operators, according to Eq. (4.4). At the end, the ancilla qubits are transversally measured in Z-basis. The example illustrates a transversal-type Γ_1 such that ancilla X checks and the measured operators' support (indicated by the box with a color gradient) are of similar sizes. 62

4.3 Numerical analysis of the space-time costs of surgery schemes. The spacetime costs are normalized by memory spacetime cost. The logical basis and logicals to measure are randomly sampled. The merge code degrees are limited to the data code degree plus two. Each data point is obtained through 6 sets of samples. (a) Cost for measuring varying number of logicals on Gross code. (Inset) the maximum space cost (with only data qubits). (b) Cost for measuring logicals in a set of high-rate spatially-coupled codes $\{[[n_i, k_i, d_i]]\}$ with increasing k_i and d_i (the code parameters are labeled on top). For each code instance, a random sample of $\lfloor 90\% \times k_i \rfloor$ number of logicals are measured. 65

4.4 Illustration of high-rate surgery on a HGP data code using another HGP ancilla code. The logicals to measure on the data code are supported on different columns and are indicated the bars in varying colors. The blue bars in the ancilla code represent a set of ancilla checks, whose stabilizer values are combined to infer the corresponding logical values in the data code. According to Eq. (4.11), we can distance, column-wise connectivites $\{\gamma_1^i\}$, resulting in a large pattern of logical measurements on the data code. 69

4.5 Memory logical error rate of the $[[144, 12, 12]]$ Gross code and its merge codes during surgery. Simulations are performed under standard circuit-level noise, excluding idling errors. The merge codes shown are constructed through the randomized, high-rate surgery scheme for measuring 3, 7, and 11 of its logicals under a randomly chosen logical basis. The logical error rates reported are averaged per logical qubit over 3 code cycles. BP-LSD decoder [12] with the same decoder parameter is used to decode all cases. 72

A.1 Comparison of near-optimal and optimal infidelity for qubit codes [[4, 1, 2]], ((5, 6, 2)), and [[7, 1, 3]] under amplitude damping noise. The black solid line (gray dashed line) represents the upper (lower) bound on the optimal infidelity given by the near-optimal infidelity 90

B.1 Environmental overlap nodes reshape the preferred lattice geometry. The environment is the Fock state $|m\rangle$ [6]. (a) The overlap kernel in Eq. (3.54). (b) Infidelity proxy at the infinite-energy limit, obtained by summing this kernel over $\Lambda_s^\perp := \mathrm{diag}(s, s^{-1})\mathbb{Z}^2/\sqrt{2}$ at $\eta = 0.65$. Its stabilizer is symplectically integral and has $d_L = 2$. Dots mark numerical local minima on $1/4 \le s \le 4$. The dotted line marks the square member $s = 1$, which minimizes the vacuum curve. (c)–(f) Fidelity-optimized dual lattices in β coordinates for $m = 0, 1, 2, 3$, respectively, obtained by minimizing the infinite-energy infidelity proxy with $s \ge 1$. The displayed s_*^2 is the optimal cell aspect ratio. The stretch s determines the kernel radii in (a) through $v = |\beta_{jk}|^2 = (\pi\chi/2)(s^2 j^2 + k^2/s^2)$, with $\chi = \eta/(1-\eta)$ and $(j, k) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$; the nearest axial shell has $v = \pi\chi/(2s^2)$. Dashed circles mark the Laguerre zeros, also indicated by color-matched dashed vertical lines in (a). Shading marks one cell; the origin is excluded from the sum. 151

B.2 Convergence of the first-order energy expansion to the full finite-energy proxy. For square qubit GKP codes under thermal loss at $\eta = 0.95$, dashed curves show the Gram-orthonormalized D_0-reference proxy, dash-dotted curves its first-order expansion, and dotted lines its infinite-energy limit. Points show near-optimal infidelity $1 - \tilde{F}^{\mathrm{opt}}$ only over the original 5–25-photon range. Colors distinguish the bath occupations $\bar{n} = 0.1, 0.3, 0.5$. The horizontal coordinate is the exact mean code energy at the same code energy parameter n_Δ used in each expression. 155

B.3 Comparison of finite-energy recovery infidelity with its marginal-reference proxy. The codes are square qubit GKP codes with Gram orthonormalization and exact mean code photons on the horizontal axis. The environment is a thermal state of mean occupation $\bar{n} = 0$ or 0.3. For (a) thermal loss at $\eta = 0.85$ and (b) thermal amplification at $G = 1.4$, circles show the actual near-optimal infidelity $1 - \tilde{F}^{\mathrm{opt}}$ and squares show $\tilde{\epsilon}_{D_{\mathrm{marg}}}$. For simplicity, the plot labels use D for D_{marg}. (c) The ratio $\tilde{\epsilon}_{D_{\mathrm{marg}}}/(1 - \tilde{F}^{\mathrm{opt}})$ for the same loss channel and codes as in (a), at $\eta = 0.85$. At the marginal reference the proxy is the leading local perturbative term (Ref. [13], Appendix D), so finite-energy corrections affect the ratio. Lemma A.5 bounds this ratio from below by 1/4 whenever the near-optimal infidelity is nonzero. 156

B.4 Reference choice changes the finite-energy infidelity proxy. Square qubit GKP codes undergo loss at $\eta = 0.9$ in all three panels, with (a) a vacuum environment, (b) a thermal state of mean occupation $\bar{n} = 0.3$, and (c) the two-photon Fock state $|2\rangle$. Circles show actual near-optimal infidelity; squares, crosses, and triangles show the proxies for D_0, D_{marg}, and D_{p}, respectively, using Eqs. (2.15) and (2.19). The proxy curves are labeled by their reference alone, with D_P denoting the product reference D_{p}. The D_0 and D_{marg} curves nearly coincide on this scale. The horizontal coordinate is the mean energy of the Gram-isometric code. 157

C.1 Illustrative figure for when the data code is a HGP code. Here the data code is formed by the hypergraph product of two $[7, 4, 3]$ Hamming codes on the right. The ancilla code, on the left, is formed by transpose checks of $[6, 3, 3]$ and $[5, 2, 3]$ on the vertical and horizontal sides respectively. The $[6, 3, 3]$ code is chosen by puncturing the Hamming code. 189
C.2 A demonstration of the random ancilla construction algorithm. (a) Initialization of ancilla system. Green shaded region represent the data code. (b) Layer addition. New layers are added at the end of the layered ancilla system. (c) Intralayer qubit addition. Qubits and Z-checks are added to boost expansion. (d) Interlayer qubit addition. Qubits and Z-checks are added, where Z-checks are connected to qubits in consecutive layers. 194

LIST OF TABLES

3.1 Gaussian-unitary dilations of three noise channels. The additive-noise dilation uses two environmental modes. 32

3.2 Complementary references for the local expansion, thermal-loss proxy, and coding bound. 34

3.3 Examples of symplectic integral lattices parameters [14, 15]. The kissing number is the number of closest lattice points in the dual lattice. The determinant of the lattice is connected to the GKP code's logical dimension through $|\det M_\Lambda| = d_L$, where M_Λ is the lattice generator matrix. . . . 40

4.1 Summary of asymptotical spacetime overheads of various surgery schemes, compared against memory, for a $[[n, k, d]]$ qLDPC code. Space overhead is normalized by unencoded circuit cost, and time overhead is normalized by the length of a logical cycle, $\Theta(d)$. For simplicity, assume $\Theta(k)$ logicals are measured, and the memory is constant encoding rate. $\Theta(1)^*$ assumes Theorem 4.6. $\tilde{\Theta}(d)^\dagger$ assumes minimum-weight basis is available. 60

4.2 Summary of spatially-coupled code parameters. The check degrees shown are the qubit degrees and check weights respectively. Distances are estimated through GAP [16] and should be taken as upper bounds. . . 72

B.1 Heisenberg-picture input-output relations and interaction coefficients. . 114

ACKNOWLEDGMENTS

My deepest thanks go to my advisor, Liang Jiang. Whatever I brought to him, in research or in life, Liang thought about it from my side first, and he has been unfailingly supportive of whatever I wanted to pursue next. From him I learned what it takes to be an outstanding researcher, and how to choose a research problem well: one that matters to the community, is hard enough to be worth the years it takes, and that I can contribute to with my own perspective. Above all, I learned to be ambitious. The clearest example is the work in Chapters 2 and 3 of this thesis: without his persistence and curiosity it would have ended long before it reached where it stands today. I am grateful, too, for the freedom to follow whatever I was excited about and the unhurried time to chase stray thoughts wherever they led. Much of how I grew as a researcher came out of that. He is someone I look up to, and hope to live up to, in whatever work I go on to do.

Much of the work presented in this thesis would not have been possible without my research collaborators, from whom I learned a lot about research, life, and beyond: Dolev Bluvstein, J. Pablo Bonilla Ataides, Aashish A. Clerk, Connor T. Hann, Wenhao He, Bikun Li, Simon Lieu, Mikhail D. Lukin, Kyungjoo Noh, Emma L. Rosenfeld, Mahadevan Subramanian, Yu-Xin Wang, Zhaoyou Wang, Yat Wong, Han Zheng, Hengyun Zhou, and Peter Zoller.

I would also like to thank all members of the Jiang group, which forms a unique research community that is open, collaborative, and supportive. Since there are so many diverse and fascinating directions of active research in the group, it is always fun and illuminating to learn about what others are up to. In particular, I would like to thank Senrui Chen, Ming Yuan, and Pei Zeng for the many meals we shared and for introducing me to the academic world when I was first starting out. Moreover, I owe a lot to many of the group's alumni, whose earlier work on bosonic codes inspired much of my own. I have only come this far by standing on the shoulders of others.

Through the ups and downs of my PhD, my friends have been a steady light that guided my way. I thank Gideon Lee for countless rambling conversations over the years, about life, academia, and much else besides. I am also grateful to Qian Xu, with whom I have collaborated since the very start of my PhD, and with whom I have shared the most academic trips over the years; I greatly enjoyed the journey we made together. I thank, too, the friends I met at the Ratner gym—basketball has always been the thing that most reliably quiets my mind, pulling me out of work and into the present. I also thank my roommates of the last two years, Zehan Mi, Mingxing (Jeffrey) Yao, Ruiming Lin, and Hong Qiao, who made those years a great deal brighter. I will miss our board game nights. To the friends who filled the hours that had nothing to do with physics—the Zoom calls and the long conversations about everything. I carry all of it.

Last, but certainly not least, I thank my parents, Shirong Bao and Bin Zheng. None of this would have been possible without them. They have been there whenever I needed them, heard every unfiltered thought I have had along the way, and stayed with me through the hardest stretches of these years. Going home has always been where I recharge, and I have always left again stronger for it. They are the anchor for every journey I take, and whatever I have managed to do here, I owe to them.

CHAPTER 1

INTRODUCTION

Quantum computers hold the promise of solving important problems far more efficiently than the best known classical methods. Shor's algorithm for prime factorization and discrete logarithms [17] is one example, among many others [18–20]. These examples motivate the effort to build quantum computers that can preserve and process information reliably over the long computations needed for useful applications.

Quantum information is also fragile. A classical digital device can stabilize two distinguishable bit values, whereas a quantum device must preserve phase coherence and entanglement [21]. Uncontrolled coupling to the environment introduces errors during storage, control, and measurement. The problem of protecting quantum information can be divided into memory and logical gates. Memory is the simpler case in which the intended operation is to leave the information unchanged; logical gates transform it while it remains exposed to noise. In both cases, errors accumulate with time. Repeated error correction must remove the entropy introduced by noise without revealing or destroying the encoded quantum state [22]. Fault-tolerant schemes must also prevent faults in the correction and gate operations themselves from spreading into uncorrectable errors [23, 24]. Quantum error correction (QEC) and fault tolerance therefore lie at the heart of both long-lived quantum memory and reliable quantum computation.

The major bottleneck is the space-time overhead of this protection. Encoding a logical qubit uses additional physical systems, and extracting error information repeatedly takes additional operations and time. The surface code offers strong protection using local checks and is a leading approach to fault tolerance [22, 25, 26]. Its resource cost can nevertheless be demanding for platforms whose qubit numbers, connectivity, and control resources are limited. Surface-code estimates for cryptographically relevant factoring have required millions of physical qubits, with newer circuit and memory designs reducing the estimate to below

a million under specified superconducting-hardware assumptions [27, 28]. Other studies based on atomic platforms estimate that around 10,000 reconfigurable qubits could suffice for cryptographically relevant Shor computations [29]. These cutting-edge resource estimates already rely on aggressive hardware assumptions. Further suppressing the space-time overhead remains a challenging task given current experimental progress.

A central goal of QEC and fault tolerance is to explore this design space and minimize the space-time cost of a task while maintaining the required protection against noise. This requires understanding the limits of an encoding and finding operations that approach those limits. In this thesis, we study these questions from memory to logic. For memory, we ask how well a given code can preserve information through a noisy channel and how much redundancy reliable protection requires. For logical gates, we ask whether codes that store information efficiently can also process it efficiently. The two problems share a concern with overhead, but the resources being optimized differ: the memory results bound recovery performance and encoding rate, while the logical-operation results address the additional physical systems and time needed to act on encoded information.

Memory. We first develop a general framework, the near-optimal fidelity, to quantify the performance of a code under a given noise channel [30]. It can be understood as a generalization of the celebrated Knill–Laflamme conditions [31], and it is straightforward to compute from the same QEC matrix used to examine those conditions. It gives an achievable recovery performance together with a narrow bound on the best possible infidelity, without optimizing over recovery maps. Infidelity measures the failure probability for preserving the encoded quantum information. This framework reduces the numerical cost of benchmarking and provides analytical expressions that reveal how performance depends on code and noise parameters. The detailed conditions for exact correction and their quantitative extension are developed in Chapter 2.

With this tool in hand, we focus on encoding information in quantum harmonic oscillators

using the lattice-based Gottesman–Kitaev–Preskill (GKP) code [4]. Oscillator encodings are motivated by the large Hilbert space available in superconducting cavities [32–34], trapped-ion motion [35, 36], and optical modes [37, 38]. GKP codes have attracted sustained theoretical and experimental interest, with demonstrated grid-state preparation and correction and improving recovery protocols [32, 33, 35, 36]. Recent experiments also demonstrate GKP qudit memories beyond break-even [39], a universal logical gate set in trapped-ion motion [40], and integrated photonic GKP-state generation [41]. They are especially appealing because important logical operations can be implemented using Gaussian operations, and their lattice structure connects error correction to a well-developed geometry [4, 42].

Understanding the limits of GKP protection has been a long-standing goal. Harrington and Preskill established early achievable rates for GKP codes under Gaussian displacement noise [43]. Later studies [1, 3] found that, with optimized recovery, finite-size GKP codes perform remarkably well under excitation loss compared with other bosonic codes. Loss is particularly relevant to microwave cavities and optical communication. The numerical results suggested that GKP codes could offer stronger protection than their original displacement-noise motivation implied. Yet the known analytical rate guarantees remained suboptimal [1]. It was unclear whether the gap reflected a limitation of the recovery or of the GKP encoding itself.

Our analysis resolves this long-standing question for pure loss and quantum-limited amplification [13]. At finite energy, near-optimal recovery can outperform existing GKP recovery schemes by orders of magnitude in experimentally relevant regimes. By connecting GKP performance to lattice theory, we show that GKP families asymptotically achieve the channel capacities, so their encoding redundancy approaches the minimum allowed by these channels. We further extend the achievable-rate results to Gaussian-dilatable channels, which represent a much broader class of quantum channels [44].

Logical gates. Scalable fault-tolerant quantum computation requires not only codes with low memory overhead but also logical operations with comparably low space-time overhead. Quantum low-density parity-check (qLDPC) codes have become promising candidates for efficient memory because they can encode many logical qubits while retaining sparse parity checks [45, 46]. Circuit-level studies show substantial memory-overhead advantages over surface-code approaches [11, 47], and recent experiments demonstrate the implementation of low-overhead qLDPC codes on superconducting processors [48]. These developments make it natural to ask whether qLDPC codes retain their advantage when the stored information is used for computation.

A high encoding rate alone does not ensure a high information-processing rate. If logical qubits can only be accessed one at a time, then a densely encoded memory can still require a long sequence of operations to execute a logical circuit. The resulting space-time cost can undermine the benefit of storing many logical qubits in one block. Flexible access is difficult because qLDPC logical operators often extend across many physical qubits and have overlapping supports. Acting on one logical subsystem while protecting the others requires a fault-tolerant operation that respects this shared physical structure.

We develop high-rate surgery to address this bottleneck [49]. The key idea is to use a shared, high-rate ancilla system to perform many addressable logical Pauli-product measurements in parallel. These measurements are building blocks for logical gates, state preparation, and gate teleportation. The theory describes their logical action and gives conditions for fault tolerance without requiring special code symmetries. For codes with suitable structure, including hypergraph-product codes, analytical ancilla constructions preserve protection while achieving amortized constant overhead for a number of logical operations proportional to the number of encoded qubits. For more general codes, a randomized construction supplies complementary numerical evidence that flexible access can be obtained with modest additional connectivity.

We benchmark the randomized construction on the Gross code and new spatially coupled

codes, using randomly sampled logical bases and measurements [49]. Hundreds of logical measurements can be performed simultaneously in the studied instances, with total space-time overhead within a factor of two of an idling memory's syndrome-extraction cycle. The analytical constructions and numerical results support logical processing with nearly constant overhead relative to memory.

Reading guide. Chapter 2 develops the near-optimal fidelity and its complementary-reference extension [30, 44], introducing the QEC background where it is needed. Chapter 3 studies GKP protection, beginning with lattice theory and noise channels, then examining finite- and infinite-energy performance, achievable rates, and environmental lattice geometry [13, 44]. Chapter 4 presents the theory, analytical constructions, and numerical benchmarks of high-rate surgery [49]. Chapter 5 summarizes the results and discusses open directions. Appendices A–C contain supporting proofs, derivations, and simulation details for the three research chapters.

CHAPTER 2

NEAR-OPTIMAL PERFORMANCE OF QUANTUM ERROR CORRECTION CODES

Quantum error correction (QEC) has central importance in scaling up quantum devices. The seminal work [31] by Knill and Laflamme (KL) outlines the necessary and sufficient conditions for exact QEC. The KL conditions are celebrated for their conciseness and computational efficiency. Practically, they have guided the discoveries and analysis of many state-of-the-art qubit codes [50–52] and oscillator codes [4, 53–56].

The KL conditions deal with *exact* error correction, in that they tell us whether or not any given error is exactly correctable by a given code. However, there are two issues with this: first, the set of correctable errors may not correspond exactly to the errors that occur in real devices – one often needs to approximate practical noise sources by considering a truncated set of Kraus operators and/or through techniques such as Pauli Twirling [57, 58]. Furthermore, codes that *exact* correct leading order errors do not necessarily outperform codes that "approximately" correct errors at all orders [3, 52]. Therefore, it is critical to develop a concise and efficiently computable metric (similar to the KL conditions) that quantifies the capabilities of general codes. Such an extension uncovers the fundamental limit set by the encoding and noise, which is a critical benchmark for code designs.

One such widely accepted benchmark is the entanglement fidelity [59–62], also called the channel or process fidelity (Section 2.1). Under this metric, the optimal recovery can be found through convex optimizations, which motivated works in optimization-based QEC [1, 3, 59, 60, 63–69]. The optimal recovery fidelity can also serve as a guide for encoding designs [3, 56, 61, 65]. However, these methods have two drawbacks. Firstly, although convex optimization algorithms are decently optimized, they remain computationally expensive compared to optimization-free methods. Consequently, past numerical works only work with small Hilbert space sizes, such as systems with fewer than five qubits [59, 70, 71] or oscillators

containing at most ten average excitations [1, 3, 72]. Secondly, these optimization techniques are inherently numerical. While powerful, they do not yield analytical forms for many of the quantities we are interested in, eg. the parametric dependence of the code performance.

Near-optimal recoveries [69, 73–78] have the potential to circumvent these limitations. These channels have constructive forms and solve a relaxed optimization problem: their performances provide two-sided bounds on the optimal fidelity. These channels have led to attempts to generalize the KL conditions [69, 74, 79–81], further leading to the development of codes like the thermodynamic code [82–84]. However, the generalized conditions still involve optimizations and/or the Bures metric, which are generally challenging to analyze. The common solutions were to derive bounds on the scaling of the near-optimal fidelity assuming large system sizes, which no longer provides a two-sided bound on the optimal performance. Moreover, the parametric dependence on other system parameters remains unknown.

In this chapter, we derive a concise and optimization-free performance metric, the near-optimal fidelity. The metric is achievable by the transpose channel [73], also known as the Petz channel [75, 76, 85], and provides a narrow two-sided bound on the optimal fidelity. Crucially, the only input to our expression is the QEC matrix, which is exactly what is required to verify the KL conditions. Therefore, our result is a quantitative generalization of the KL conditions for arbitrary codes and noise processes. We also develop a perturbative expression, providing intuition on how are codes' performances connected to the structures of their error subspaces. More importantly, the perturbative form well approximates the near-optimal fidelity and can be computed analytically. Taking multiple qubit codes as examples, we numerically validate the proximity of the closed-form near-optimal expression and the optimal fidelity obtained from convex optimizations. Furthermore, we analytically compute the near-optimal fidelity for the thermodynamic code in the thermodynamic limit, for which only a scaling with system size was known in past works. After examining a few representative oscillator codes, we provide rigorous insights on why certain codes' performances under excitation

loss improve monotonically with increased energy, such as the Gottesman-Kitaev-Preskill (GKP) code [4]. While past numerical simulations of the GKP code were limited to a few average excitations, we extend our result to hundreds of excitations. We also obtain GKP's performance analytically, parameterized by system parameters and loss rates. With the analytical expression, we find the GKP's performance admits an asymptotic limit at infinite energy.

2.1 Background

2.1.1 Encodings in finite and infinite dimensions

Quantum information is protected by encoding it with redundancy in the available physical levels. Systems such as superconducting transmons, spins, and atoms can be described by finite-level models. Restricting each system to two levels gives a physical qubit; encodings built from these qubits are called *qubit codes*. Optical modes and superconducting cavities are naturally modeled as infinite-dimensional harmonic oscillators. Codes that use their oscillator levels to encode a finite logical system are called *bosonic codes*.

An encoding embeds d_L logical basis states into codewords $\{|\mu_L\rangle\}$ through an isometry $\hat{V} : \mathbb{C}^{d_L} \to \mathcal{H}$, with $\hat{V}|\mu\rangle = |\mu_L\rangle$. Its channel is $\mathcal{E}(\rho) := \hat{V}\rho\hat{V}^\dagger$, and the codespace projector is $\hat{P}_L := \hat{V}\hat{V}^\dagger$. After noise acts on the encoded state, a recovery channel $\mathcal{R}$ maps the physical system back to the logical space, as illustrated in Fig. 2.1.

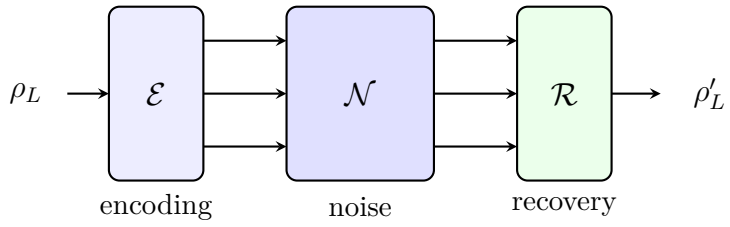


Figure 2.1: Encoding, noise, and recovery. Multiple physical systems pass through a single channel $\mathcal{N}$, so the diagram also accommodates correlated errors. The composite logical channel is $\mathcal{Q} = \mathcal{R} \circ \mathcal{N} \circ \mathcal{E}$.

For qubit codes, we most often work with stabilizer codes [24]. An $[[n, k, d]]$ qubit stabilizer code encodes k logical qubits into n physical qubits, has $d_L = 2^k$, and is the common $+1$

eigenspace of an abelian Pauli subgroup $\mathcal{S}$ that does not contain $-I$. Its projector is

$$\hat{P}_L = \frac{1}{|\mathcal{S}|} \sum_{g \in \mathcal{S}} g, \qquad |\mathcal{S}| = 2^{n-k}. \tag{2.1}$$

Calderbank–Shor–Steane (CSS) codes have separate X-type and Z-type stabilizer generators [86, 87]. Their binary check matrices satisfy $H_X H_Z^T = 0$ over $\mathbb{F}_2$, expressing mutual commutation. Chapter 4 develops the additional algebraic structure needed for logical operations.

Bosonic codes encode a finite logical system in one or more oscillator modes. Examples include cat, squeezed-cat, binomial, and GKP codes [4, 54–56]. Multimode constructions with additional algebraic structure include quantum spherical codes [88] and tiger codes [89].

The mean codeword energy is an important parameter for bosonic codes because it affects their performance. Higher photon numbers can also amplify errors from cavity dephasing and inherited Kerr nonlinearities [90]. Comparing codes at the same mean energy therefore helps account for their physical resource cost. With $\hat{a}_j$ the annihilation operator of mode j and $\hat{n}_{\text{tot}} := \sum_j \hat{a}_j^\dagger \hat{a}_j$, the mean code energy per codeword, measured in photons, is

$$\bar{n}_C := \frac{1}{d_L} \operatorname{Tr}\Big(\hat{P}_L \hat{n}_{\text{tot}}\Big). \tag{2.2}$$

For one mode, $\hat{n}_{\text{tot}} = \hat{n} = \hat{a}^\dagger \hat{a}$.

2.1.2 Noise models

A noise channel $\mathcal{N}$ is a completely positive trace-preserving map with Kraus representation

$$\mathcal{N}(\rho) = \sum_a \hat{K}_a \rho \hat{K}_a^\dagger, \qquad \sum_a \hat{K}_a^\dagger \hat{K}_a = I. \tag{2.3}$$

Unless stated otherwise, noise is independent and identically distributed (i.i.d.) among the physical systems: $\mathcal{N} = \mathcal{N}_1^{\otimes n}$. The diagram in Fig. 2.1 also accommodates general joint noise channels.

For qubit codes, two common models are amplitude damping and depolarizing noise [21, 52]. Amplitude damping with decay probability $0 \leq \kappa \leq 1$ has Kraus operators

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\kappa} \end{pmatrix}, \qquad K_1 = \sqrt{\kappa}\,|0\rangle\langle 1| = \begin{pmatrix} 0 & \sqrt{\kappa} \\ 0 & 0 \end{pmatrix}. \tag{2.4}$$

For depolarizing noise with total Pauli-error probability p,

$$K_I = \sqrt{1-p}\,I, \qquad K_X = \sqrt{p/3}\,X, \quad K_Y = \sqrt{p/3}\,Y, \quad K_Z = \sqrt{p/3}\,Z. \tag{2.5}$$

Depolarizing and other stochastic Pauli models are common benchmarks in fault-tolerant QEC. They are motivated by twirling techniques: Pauli twirling removes off-diagonal terms of the process matrix in the Pauli basis, producing a stochastic Pauli channel; Clifford twirling additionally averages its nonidentity Pauli probabilities and gives a depolarizing channel [91]. Randomized compiling implements noise tailoring into stochastic Pauli errors while preserving the logical computation [57, 58]. Amplitude damping retains non-Pauli structure relevant to physical relaxation.

For bosonic codes, common models include random displacements, pure loss, and dephasing [3, 4]. In one mode, random displacement noise is

$$\mathcal{N}_{\text{disp}}(\rho) = \int_{\mathbb{C}} d^2\alpha\, p(\alpha)\hat{D}(\alpha)\rho\hat{D}(\alpha)^\dagger, \quad \hat{D}(\alpha) := e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}, \tag{2.6}$$

where p is a normalized probability density. Gaussian random displacement noise has $p(\alpha) = e^{-|\alpha|^2/\sigma^2}/(\pi\sigma^2)$. Pure loss with transmissivity η has Kraus operators

$$\hat{E}_\ell = \sqrt{\frac{(1-\eta)^\ell}{\ell!}}\,\eta^{\hat{n}/2}\hat{a}^\ell, \qquad \ell \geq 0. \tag{2.7}$$

Dephasing is a random phase rotation,

$$\mathcal{N}_{\text{deph}}(\rho) = \int_{-\pi}^{\pi} d\theta \, p_\theta(\theta) e^{-i\theta\hat{n}} \rho e^{i\theta\hat{n}}. \tag{2.8}$$

Chapter 3 introduces the broader class of Gaussian-dilatable channels, including random displacement noise and loss and amplification with general environmental states.

2.1.3 Quantum error correction matrix

The encoding and noise channel together determine the overlaps between error subspaces. These overlaps are collected in the quantum error correction (QEC) matrix.

Definition 2.1 (Quantum error correction matrix)**.** *For an isometric encoding $\hat{V}$ and noise Kraus operators $\{\hat{K}_a\}$, the QEC matrix is*

$$M_{\mu a,\nu b} := \langle \mu_L | \, \hat{K}_a^\dagger \hat{K}_b \, | \nu_L \rangle = \langle \mu | \, \hat{V}^\dagger \hat{K}_a^\dagger \hat{K}_b \hat{V} \, | \nu \rangle \, . \tag{2.9}$$

Here μ, ν label logical basis states and a, b label Kraus operators. The matrix acts on the logical space L and Kraus-index space C.

Remark 2.2 (Connection to the complementary Choi state; [92])**.** *The QEC matrix can also be interpreted through the complementary channel. Fix the complementary-channel convention by $[\mathcal{N}^c_{\hat{V}}(|\mu\rangle \langle\nu|)]_{ab} = \langle \nu | \, \hat{V}^\dagger \hat{K}_b^\dagger \hat{K}_a \hat{V} \, | \mu \rangle$. The normalized Choi state of this encoded complementary channel is*

$$J^c := \frac{1}{d_L} \sum_{\mu,\nu} |\mu\rangle \langle\nu| \otimes \mathcal{N}^c_{\hat{V}}(|\mu\rangle \langle\nu|) = \frac{M^T}{d_L}. \tag{2.10}$$

Generally, the Choi state can be entangled between the logical reference L' and complementary system C, as illustrated in Fig. 2.2.

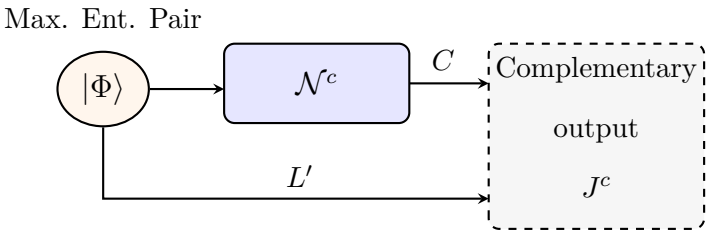


Figure 2.2: A complementary channel acting on one half of the maximally entangled pair $|\Phi\rangle$ produces a generally correlated output J^c on $L' \otimes C$. With the fixed basis transpose convention, $M/d_L = (J^c)^T$. Its logical marginal is I_L/d_L, a maximally mixed state. Under exact correction, this output factorizes between L' and C, so the complement learns no logical information.

For a trace-preserving channel, tracing out the Kraus index gives

$$(\mathrm{Tr}_C\, M)_{\mu\nu} = \sum_a M_{\mu a,\nu a} = \delta_{\mu\nu}, \qquad \mathrm{Tr}\, M = d_L. \tag{2.11}$$

We write $(\mathrm{Tr}_L\, B)_{ab} := \sum_\mu B_{\mu a,\mu b}$ for the logical partial trace and $\|B\|_F^2 := \mathrm{Tr}\Big(B^\dagger B\Big)$ for the Frobenius norm.

Lemma 2.3 (Knill–Laflamme conditions; [31])**.** *An isometric code exactly corrects a trace-preserving noise channel if and only if*

$$M = \frac{I_L}{d_L} \otimes \mathrm{Tr}_L\, M. \tag{2.12}$$

Thus exact correction requires every Kraus-pair overlap to be proportional to the logical identity.

The performance of an encoding against a noise channel is determined by the best allowed recovery. We adopt *entanglement fidelity* as the performance metric, evaluated throughout this chapter at the maximally mixed logical input. For a logical channel $\mathcal{Q}$,

$$F(\mathcal{Q}) := \langle\Phi|\,(\mathcal{Q}\otimes\mathcal{I}_{L'})(|\Phi\rangle\langle\Phi|)\,|\Phi\rangle\,, \quad |\Phi\rangle := \frac{1}{\sqrt{d_L}}\sum_{\mu=1}^{d_L} |\mu\rangle_L\,|\mu\rangle_{L'}\,. \tag{2.13}$$

The auxiliary system L' purifies the maximally mixed logical input. This quantity is also

called channel or process fidelity in the source papers [59, 60, 62]. The optimal fidelity is

$$F^{\text{opt}} := \sup_{\mathcal{R}} F(\mathcal{R} \circ \mathcal{N} \circ \mathcal{E}), \tag{2.14}$$

with the supremum over completely positive trace-preserving recoveries. Entanglement fidelity is directly related to average input-output fidelity [93, 94]. It is also linear in the recovery's Choi matrix, so optimizing it is a semidefinite program (SDP), as reviewed in Appendix A.

2.2 Near-optimal fidelity

Here we propose the near-optimal fidelity as a quantitative metric that can be evaluated without optimization.

Theorem 2.4 (The near-optimal fidelity)**.** *For a d_L-dimensional encoding, $\mathcal{E}$, and a noise channel, $\mathcal{N}$, the near-optimal fidelity is*

$$\tilde{F}^{opt} = \frac{1}{d_L^2} \left\| \text{Tr}_L \sqrt{M} \right\|_F^2, \tag{2.15}$$

where M is the QEC matrix, $(\text{Tr}_L B)_{l,k} = \sum_\mu B_{\mu l,\mu k}$ denotes the partial trace over the code space indices, and $||\cdot||_F$ is the Frobenius norm. The near-optimal fidelity gives a two-sided bound on the optimal fidelity as

$$\frac{1}{2}\left(1 - \tilde{F}^{opt}\right) \leq 1 - F^{opt} \leq 1 - \tilde{F}^{opt}. \tag{2.16}$$

Worth noticing, the gap between the two-sided bound is proportional to the optimal infidelity. Therefore, when the code performs well against the noise channel, the near-optimal fidelity is a close approximation to the optimal fidelity. Equally importantly, it is remarkable that Eq. (2.15) is only dependent on the QEC matrix. Therefore, our metric requires exactly the same resources as the KL conditions, but it provides a quantitative performance metric beyond a binary Yes-or-No output. Moreover, our result can further extended to general

channel reversals [95, 96], subsystem codes [97], and mixed-state codes [98]. Our result implies that the QEC matrix contains richer information about the code and noise structure beyond the KL conditions. For example, one can greatly reduce the complexity of optimization-based methods via adopting the error subspaces as an efficient basis to describe the action of the noise channel [99, 100] (see Appendix A).

The near-optimal fidelity is achievable by the transpose channel [73, 74] by construction, which is exactly where our expression inherits the two-sided bound. While there are other channels possessing similar near-optimal properties [69, 77], the transpose channel has the most concise fidelity expression. The derivation of Eq. (2.15) is based on the observation that the QEC matrix is the Gram matrix of the error subspaces. When the QEC matrix is invertible, the error subspaces can be orthonormalized by the Gram matrix. In such an orthonormal basis, the transpose channel is equivalent to a measure-and-recover operation, which leads to the expression of Eq. (2.15). The derivation can be generalized to scenarios of degenerate QEC matrix (see Appendix A).

A reasonable question is when the transpose channel attains the supremum over recoveries exactly, making the right-hand inequality in Eq. (2.16) an equality.

Lemma 2.5 (Optimality of the transpose channel; [92])**.** *For a finite-dimensional isometric encoding and trace-preserving noise, the transpose channel is optimal for entanglement fidelity if and only if*

$$[M, I_L \otimes \mathrm{Tr}_L \sqrt{M}] = 0. \tag{2.17}$$

The KL condition in Lemma 2.3 satisfies this commutation condition and gives fidelity one. The same condition can hold for a nonproduct QEC matrix, yielding an optimal recovery with fidelity below one. It distinguishes optimal recovery from perfect correction.

Computationally, our approach has a drastically reduced cost compared to optimization-based methods. In many cases, the QEC matrix can be analytically computed. Otherwise, it is possible to efficiently compute the matrix depending on the code and noise of interest.

In such cases, suppose we are considering N_K number of noise Kraus operators, the cost of evaluating $\tilde{F}^{\text{opt}}$ is $\mathcal{O}\left((d_L N_K)^3\right)$. As a reference, the SDP for optimal recovery costs $\tilde{\mathcal{O}}\left((d_L N)^{5.246}\right)$ [101][1], where N is the physical Hilbert space dimension. For example, if we consider qubit codes, N scales exponentially with the number of qubits, n. However, it is sufficient numerically to truncate the number of noise Kraus operators to a polynomial scaling, $N_K \propto n^r$. In many cases, r only depends on the target precision and physical error rates.

Remark 2.6 (Iterative recovery optimization)**.** *Iterative methods also improve on the cost of generic SDP solvers [59, 67]. With the full recovery Choi matrix of dimension $d_L N$, their dense updates cost $O(n_{\text{iter}}(d_L N)^3)$ for n_{iter} iterations [92]. Restricting the output space to the span of the error states reduces the matrix size. By comparison, the QEC matrix has dimension $d_L N_K$, and Eq.* (2.15) *requires a single matrix square root and a partial trace, with cost $O((d_L N_K)^3)$, without convergence iterations.*

2.3 Infidelity proxy

The exact near-optimal fidelity reduces numerical cost, but the matrix square root can make analytical expressions cumbersome. We therefore compare the QEC matrix with a product operator on the logical and complementary spaces. This comparison gives both a perturbative expression for the near-optimal infidelity and a nonperturbative upper bound.

Definition 2.7 (Complementary reference and admissibility)**.** *A positive trace-class operator D on the Kraus-index space C is called a* complementary reference*. It is* admissible *for a QEC matrix M when*

$$\operatorname{supp} M \subseteq \mathbb{C}^{d_L} \otimes \operatorname{supp} D, \qquad \tilde{\epsilon}_D(M) < \infty, \tag{2.18}$$

where $\tilde{\epsilon}_D(M)$ is the infidelity proxy defined in Eq. (2.19)*.*

[1] $\tilde{\mathcal{O}}(g(n))$ is shortened for $\mathcal{O}(g(n)\log^k g(n))$

For N_K Kraus operators, M has dimension $d_L N_K$, whereas D acts only on the N_K-dimensional complementary space. This space represents the discarded output and, for a mixed environmental input, its purifier. A normalized D represents a guessed complementary state independent of the logical input, up to the transpose convention of Remark 2.2. An unnormalized reference can also serve as a comparator. The complementary reference differs from the logical input/reference state used to define a general Petz recovery in Ref. [92].

Definition 2.8 (Infidelity proxy)**.** *Let D be a positive trace-class complementary reference. Diagonalize it on its support as $D_U := UDU^\dagger = \sum_a d_a |a\rangle\langle a|$, with $d_a > 0$, and set $M_U := (I_L \otimes U)M(I_L \otimes U^\dagger)$. The* infidelity proxy *is*

$$\tilde{\epsilon}_D(M) := \frac{1}{d_L} \sum_{\mu,\nu,a,b} \frac{\left|(M_U - I_L \otimes D_U)_{\mu a,\nu b}\right|^2}{(\sqrt{d_a} + \sqrt{d_b})^2}. \tag{2.19}$$

The residual $M_U - I_L \otimes D_U$ measures departure from a guessed complementary output that carries no logical information. The proxy weights this departure by the reference spectrum.

The square root uses the same spectral projectors as D_U: $\sqrt{D_U} = \sum_a \sqrt{d_a} |a\rangle\langle a|$. Thus Eq. (2.19) is a weighted spectral sum of squared residual overlaps. Each term is divided by $(\sqrt{d_a} + \sqrt{d_b})^2$, so deviations involving small reference eigenvalues contribute more strongly. We can also understand the proxy as the squared Frobenius norm of the solution to a Sylvester equation. With $C_D := I_L \otimes \sqrt{D_U}$, the equation and proxy are

$$C_D X_D + X_D C_D = M_U - I_L \otimes D_U \quad \text{satisfies} \quad \tilde{\epsilon}_D(M) = \frac{\|X_D\|_F^2}{d_L}. \tag{2.20}$$

This is the linearized square-root comparison used in Appendix A.7.

For normalized references, the comparison states form the class

$$\mathfrak{P}_{L'C} := \left\{ \frac{I_{L'}}{d_L} \otimes \sigma_C : \sigma_C \geq 0,\ \operatorname{Tr} \sigma_C = 1 \right\}. \tag{2.21}$$

These are decoupled product outputs, equivalently normalized Choi states of constant-output

channels $\rho \mapsto \mathrm{Tr}(\rho)\sigma_C$. Their role in exact and approximate QEC is established in the complementary-channel formulation [69]. For normalized D, the proxy compares J^c with the member $I_{L'}/d_L \otimes D^T$ of this class. Minimizing over D would compare with the whole class. We have the freedom to choose a reference that makes the calculation tractable.

The *marginal reference* is

$$D_{\mathrm{marg}} := \frac{\mathrm{Tr}_L M}{d_L}. \tag{2.22}$$

It is the transposed complementary output of the maximally mixed encoded input and has unit trace. The KL condition becomes $M = I_L \otimes D_{\mathrm{marg}}$: exact correction is possible when the complementary channel reveals no logical information.

Corollary 2.9 (Marginal-reference perturbative expansion)**.** *With the marginal reference defined in Eq.* (2.22)*, the proxy gives the first-order approximation to the near-optimal infidelity,*

$$1 - \tilde{F}^{opt} = \tilde{\epsilon}_{D_{\mathrm{marg}}}(M) + \mathcal{O}\Big(\tilde{\epsilon}_{D_{\mathrm{marg}}}(M)^{3/2}\Big)\,. \tag{2.23}$$

This corollary follows from the Daleckii–Krein expansion of the matrix square root [102, 103]; its proof is in Appendix A. It weights overlaps between error subspaces by their complementary probabilities. Past performance estimators [3, 65, 104–107] used the uncorrectable part of the QEC matrix or heuristic weights; the proxy supplies the weights from the square-root expansion.

For the marginal reference, the proxy is the leading perturbative infidelity with coefficient one. Its quadratic dependence on the QEC-matrix residual shows how small overlaps suppress the error; quadratic error suppression was also found in Ref. [78]. If the proxy vanishes, the KL condition holds exactly. Computing or diagonalizing D_{marg} may be inconvenient. A reference such as its diagonal part in a fixed Kraus basis then gives a different weighted comparison, whose general guarantee is the following upper bound.

Lemma 2.10 (Complementary-reference bound)**.** *For every isometric code, trace-preserving channel, and admissible complementary reference D, the near-optimal infidelity obeys*

$$1 - \tilde{F}^{opt} \leq 4\,\tilde{\epsilon}_D(M). \tag{2.24}$$

The factor four is independent of dimension and reference spectrum, so this nonperturbative bound applies along the coding limit, where the logical dimension and complementary spectrum can change with the number of channel uses. A small proxy certifies high fidelity even when D is not the marginal reference. The proof is in Appendix A.7. The local expansion gives a more accurate estimate near exact correctability, while the upper bound gives a recovery guarantee for any admissible reference.

2.4 Examples

The two-sided bound in Eq. (2.16) holds for arbitrary isometric codes, including qubit, bosonic, and non-stabilizer codes. It depends on the encoding and noise through the QEC matrix and requires no stabilizer structure. In Fig. 2.3, we numerically validate the two-sided bound presented in Eq. (2.16) for qubit codes under amplitude damping noise (see Appendix A), which is a practical but non-Pauli noise channel. We adopt the convention that $[[n, k, d]]$ represents encoding k logical qubits in n physical qubits with distance d, while for $((n, k, d))$, k represents the logical dimension instead. Even for well-studied codes like stabilizer codes, their optimal decoders under non-Pauli noise are in general unknown. The shown qubit codes include the classic $[[9, 1, 3]]$ stabilizer code [51] and approximate $[[4, 1, 2]]$ code [52]. Another example is the $((5, 6, 2))$ code [108], which is a qudit non-stabilizer code.

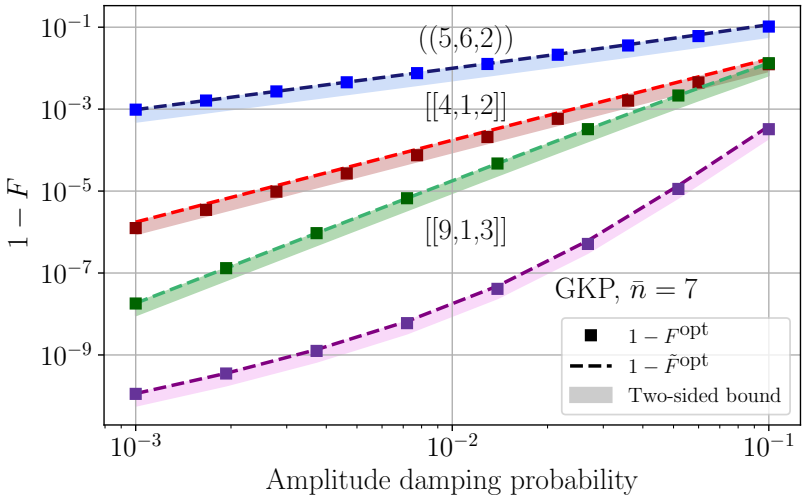


Figure 2.3: Optimal infidelity for qubit codes $[[4,1,2]]$, $((5,6,2))$, and $[[9,1,3]]$ and the GKP code with $\bar{n}_C = 7$. The noise channel is amplitude damping noise, i.e. loss for oscillators. The shaded regions represent the optimal infidelity intervals bounded by the two-sided bound given by the near-optimal infidelity in Eq. (2.16).

The contributions of our expression lie not only in numerics but also in analytical aspects. We take the thermodynamic code [82] (see Appendix A) as an example. The code has sparked interest because of its close connections with the Eigenstate Thermalization Hypothesis [109], as well as being an instance of a covariant code [83].

For N physical qubits, let $|\mathbf{s}\rangle$ denote a joint Pauli-Z eigenstate with $s_j \in \{-1,+1\}$. The normalized state of magnetization m is

$$\left|h_m^N\right\rangle := \binom{N}{(N+m)/2}^{-1/2} \sum_{\mathbf{s}:\sum_j s_j = m} |\mathbf{s}\rangle\,, \qquad |0_L\rangle = \left|h_{-d/2}^N\right\rangle, \quad |1_L\rangle = \left|h_{d/2}^N\right\rangle. \tag{2.25}$$

The two codewords have magnetization separation d, with N and d chosen so that both magnetization sectors exist. These centered codewords are used in the erasure calculation below; Appendix A gives the derivation.

When considering a constant number of erasure errors and thermodynamic limit, the optimal infidelity was proven to scale as $1 - F^{\mathrm{opt}} = \Theta(\frac{1}{N^2})$ [83], where N is the number of qubits. Beyond a scaling argument, our expression makes it possible to derive the marginal-

reference infidelity proxy analytically (see Appendix A),

$$\tilde{\epsilon}_{D_{\text{marg}}}(M) = \frac{l}{16}\frac{d^2}{N^2} + \mathcal{O}\left(\frac{1}{N^3}\right). \tag{2.26}$$

for l erasure errors. The derivation of Eq. (2.26) with our formalism is straightforward and can be easily extended to consider more erasure errors (see Appendix A). The comparison of the perturbative and the exact forms of the near-optimal infidelity is shown in Fig. 2.4, where it is clear they closely follow each other. For the exact form, the QEC matrix is numerically computed, and the simulation stops at 14 qubits because of the increased computational cost. To compare, if we optimize the recovery using SDP, it is only possible for fewer than 5 qubits under the same time constraint.

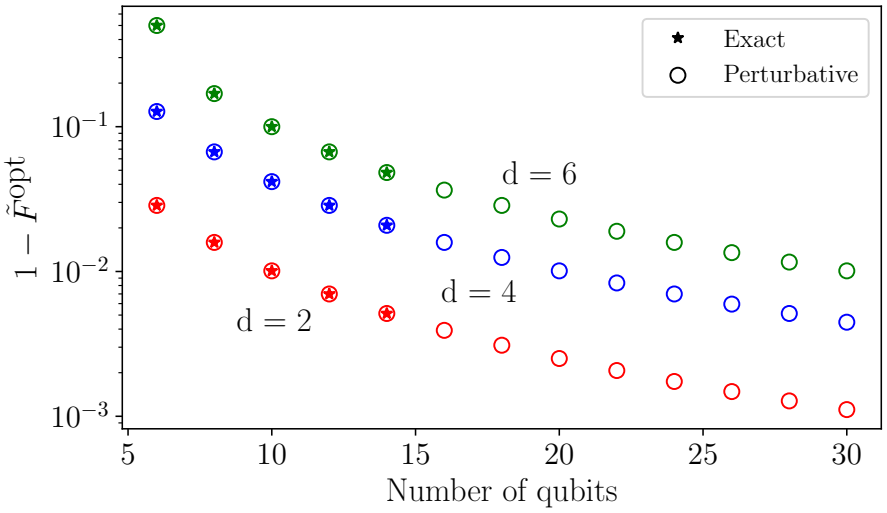


Figure 2.4: The near-optimal infidelity of the thermodynamic code with distance d (see Appendix A for the definition of the codewords) and $l = 1$, i.e. single erasure error. The stars represent the exact approach where the QEC matrix is numerically computed, and the circles represent the perturbative expression given in Eq. (2.26).

In this section we also briefly touch on bosonic codes, including GKP, cat, and binomial codes. We illustrate two uses of the near-optimal fidelity: its perturbative form gives closed-form performance estimates in regimes that numerical recovery optimization could not reach, and its exact matrix expression reduces the power of Hilbert-space dimension entering the runtime. These gains extend calculations to larger oscillator truncations. Chapter 3 studies GKP protection itself, its lattice dependence, and its communication rates.

Bosonic codes, or oscillator codes, are codes that encode a d_L-dimensional logical space in the infinite-dimensional Hilbert space of oscillator(s). For a bosonic code, the average excitation number per codeword, $\bar{n}_C$ of Eq. (2.2), affects code properties and is also a parameter of experimental interest. In the inset plot of Fig. 2.5, we show the optimal performance of a few popular bosonic codes under excitation loss (see Appendix A), including the binomial code [55], the cat code [54], and the GKP code [4]. An emerging feature is that for the binomial and cat codes, their performance do not improve monotonically with energy with fixed S. Here, S represents the fock basis spacing, which can be understood as the distance against loss. On the contrary, the GKP code's performance improves monotonically for the range of $\bar{n}_C$ shown. Similar numerical observations were made in previous works [3, 55], where heuristic explanations were given based on the QEC matrix. Our expression Eq. (2.23) supports their arguments with rigor: like exact qubit codes, the cat code and binomial code exactly correct no more than S excitation loss. However, they are completely unprotected against more than S loss, which is more likely to occur at larger $\bar{n}_C$. Therefore, the infidelity is not suppressed with energy.

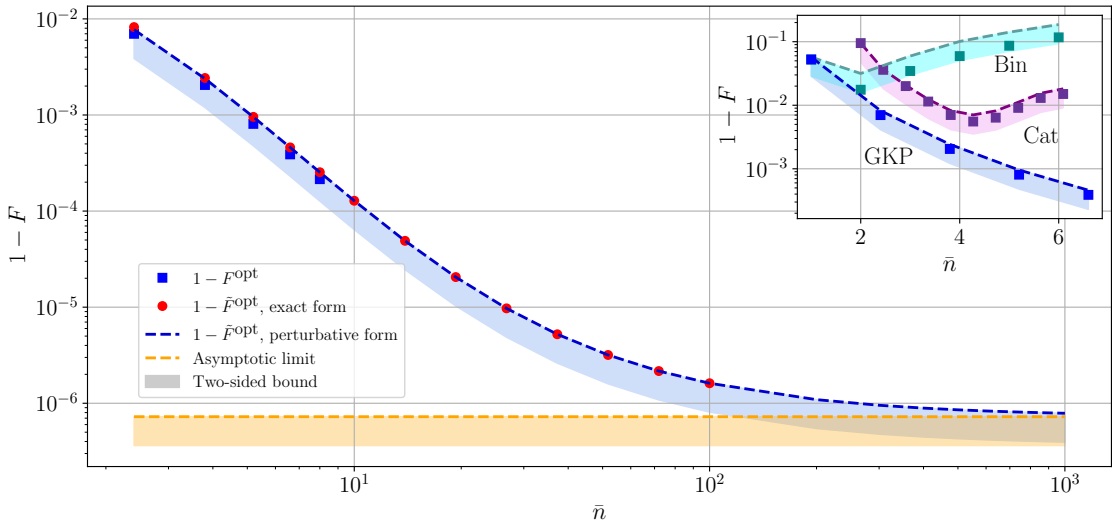


Figure 2.5: The entanglement infidelity of square lattice GKP qubit code under loss with transmissivity $\eta = 0.9$ (10% loss). All shaded regions represent the two-sided bound on the optimal fidelity. The red circles represent the exact form of the near-optimal infidelity, where the QEC matrix is computed analytically. The blue dashed line is the perturbative form, evaluated through the closed-form analytical expression derived in Chapter 3. The inset figure shows the performance of GKP codes (square lattice), cat codes (S=4), and binomial codes (S=1) (see Appendix A). The squares (dashed lines) represent the optimal (near-optimal) fidelity.

The critical difference of GKP codes is that while it generally cannot even correct a single loss, the uncorrectable elements for all orders of loss are suppressed by a factor of $e^{-\frac{1-\eta}{(1-\eta)+\frac{1}{\bar{n}_C}}}$ [3]. One open question was whether the suppression of the infidelity holds asymptotically: optimizing for the optimal fidelity was only possible for $\bar{n}_C \leq 10$ [3] because the Hilbert space size quickly becomes unmanageable. The near-optimal fidelity solves the problem. Firstly, with the QEC matrix being analytically computable, the only complexity cost for the exact form in Eq. (2.15) is to compute the matrix square root. Thus, it is much cheaper than SDP optimizations, and the numerical results can reach $\bar{n}_C \sim 10^2$. Secondly, we can express the near-optimal performance analytically with the perturbative form (Chapter 3), which reveals the near-optimal fidelity for arbitrarily large $\bar{n}_C$. In Fig. 2.5, we demonstrate the results for a GKP square code under loss with transmissivity $\eta = 0.9$, i.e. a loss probability $1 - \eta = 0.1$. In the nearest-shell perturbative approximation, the

large-energy near-optimal infidelity is estimated by [13, 30]

$$1 - \tilde{F}^{\text{opt}} \approx \tilde{\epsilon}_{D_{\text{marg}}}(M) \approx e^{-\frac{\pi}{2}\frac{\eta}{1-\eta}} \qquad (\bar{n}_C \to \infty), \tag{2.27}$$

which is approximately 7.2×10^{-7} for $\eta = 0.9$. As a comparison, the best decoder with known performance at infinite energy is the amplification decoder (AD) [1, 3]. For the same loss, AD gives a logical error rate of $1 - F^{\text{AD}} \approx e^{-\frac{\pi}{8}\frac{\eta}{1-\eta}} = 2.9 \times 10^{-2}$.

2.5 Discussion

We have derived the near-optimal fidelity as a quantitative metric for arbitrary codes and noise channels [30]. Since it depends only on the QEC matrix, it uses the same input as the KL conditions and bounds the best recovery performance beyond exact correction. Its matrix expression reduces numerical cost, while its perturbative form gives analytical descriptions of code performance.

The complementary-reference formulation extends this approach. The marginal reference yields the perturbative estimate near exact correctability; another admissible reference can simplify an analytical calculation. The global bound relates every such proxy to recovery infidelity with a dimension-independent coefficient, allowing the logical dimension and complementary spectrum to change along an asymptotic code family.

These tools support numerical benchmarking, encoding optimization, and recovery design for realistic noise. They also separate the information-theoretic limit of an encoding from the performance of a particular recovery circuit. Developing recoveries that approach the benchmark is a further step toward realizing the code's potential. The reference bound applies to arbitrary isometric codes, so it can also support rate analyses for code families with tractable QEC matrices.

CHAPTER 3

ACHIEVING COHERENT INFORMATION RATES THROUGH GOTTESMAN-KITAEV-PRESKILL CODES

An ultimate question in the theory of communication is to determine the maximum information transmission rate, also known as capacity, of communication models and channels [21, 76]. Extending Shannon's theory to quantum channels [110], the one-way quantum capacity is the supremum of QEC coding rates with asymptotically perfect decoding and no auxiliary classical communication [111–114]. Achievable rates of specific codes provide lower bounds to this capacity. Computing these rates directly requires both an encoding and a recovery whose performance can be controlled over asymptotically many channel uses. Coherent information instead requires only an input state [112, 113, 115]. For pure loss and quantum-limited amplification, thermal input states give coherent information equal to the known capacities [116–119]. For thermal Gaussian channels, they give a standard achievable-rate benchmark [2]. These existence arguments do not by themselves provide code structure suited to logical operations, decoding, and implementation.

Quantum noise can be described by coupling a signal to an environment and discarding the environmental output. Gaussian channels admit Gaussian-unitary dilations with Gaussian environments [120]; allowing non-Gaussian environmental states gives the broader class of Gaussian-dilatable channels [121, 122]. Structured approaches to these channels include closest-lattice-point decoding [123] and quantum low-density lattice codes [124].

The Gottesman-Kitaev-Preskill (GKP) code [4] is a promising bosonic code that has attracted active theoretical and experimental effort in superconducting circuits [32–34], ion traps [35, 36], and optical platforms [37, 38, 41]. Its popularity is largely due to the Gaussian nature of its logical operations [4] and the recent advances in its decoding [5, 125]. Another key motivation is that while the GKP code was originally designed against displacement noise [4], there has been numerical evidence [1, 3, 72] that demonstrates the superiority of the

optimally decoded single-mode GKP code under loss compared to other single-mode bosonic codes. Compared to displacement noise, loss is more experimentally motivated and is the dominant noise process in various hardware platforms such as microwave cavities [55] and optical fibers [118]. The outstanding performance of GKP code under loss is rather surprising. While there have been some qualitative arguments [3, 126], these numerical studies did not provide a rigorous and quantitative explanation that connects GKP code's performance with its parameters, such as its dual lattice and energy. In addition, the optimal performance of GKP code under loss is obtained through numerical convex optimization [1, 3], leading to restrictions on the accessible system size. They left open whether GKP code can maintain its performance at larger average photon numbers.

We study GKP protection under a general noise-channel form, Gaussian-dilatable channels, and their complementary outputs. Through the QEC matrix, the near-optimal fidelity of Chapter 2 turns the channel and code structure into recovery benchmarks. We derive the QEC matrix for beam-splitter and two-mode-squeezer interactions with general environments, then apply it to finite-energy GKP performance under pure loss and quantum-limited amplification, where existing decoders can be compared directly [13]. The rate analysis [44] combines the complementary-reference bound with lattice averaging and extends to Gaussian-dilatable channels with finite-energy environments. We introduce the notion of an environmental kernel to illustrate how lattice design changes beyond vacuum and thermal noise.

3.1 GKP codes and lattice theory

We use the following phase-space conventions for oscillator modes. For N modes, let $\hat{x} := (\hat{q}_1, \hat{p}_1, \ldots, \hat{q}_N, \hat{p}_N)^T$, $\Omega := I_N \otimes \omega$, and $\omega := \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Thus $[\hat{x}_j, \hat{x}_k] = i\Omega_{jk}$, $\hat{a}_j := (\hat{q}_j + i\hat{p}_j)/\sqrt{2}$, $\hat{n}_j := \hat{a}_j^\dagger \hat{a}_j$, and the vacuum covariance is $I_{2N}/2$. The real displacement operator is

$$\hat{T}(u) := e^{-iu^T \Omega \hat{x}}. \tag{3.1}$$

The complex displacement operator is $\hat{D}(\alpha) := \exp\Big(\sum_j \alpha_j \hat{a}_j^\dagger - \alpha_j^* \hat{a}_j\Big)$. For $u \in \mathbb{R}^{2N}$, the map $C := I_N \otimes (1 \;\; i)/\sqrt{2}$ gives $\alpha = Cu \in \mathbb{C}^N$, with $\alpha_j = (u_{2j-1} + iu_{2j})/\sqrt{2}$, and relates the two operators by $\hat{T}(u) = \hat{D}(Cu)$. Real displacements obey the Weyl relation

$$\hat{T}(u)\hat{T}(v) = e^{-iu^T \Omega v/2} \hat{T}(u+v). \tag{3.2}$$

Equivalently, $\hat{D}(\alpha)\hat{D}(\beta) = e^{(\alpha\cdot\beta^* - \alpha^*\cdot\beta)/2}\hat{D}(\alpha+\beta)$. To distinguish dimensionless lattice coordinates from physical displacements, for $x \in \mathbb{R}^{2N}$ we write

$$\xi(x) := \sqrt{2\pi}\, Cx, \tag{3.3}$$

with components $\xi_j(x) = \sqrt{\pi}\,(x_{2j-1} + ix_{2j})$, so that the displacement associated with x is $\hat{T}(\sqrt{2\pi}\,x) = \hat{D}(\xi(x))$ and $\|\xi(x)\|^2 = \pi\|x\|^2$. The Frobenius (Hilbert–Schmidt) norm is $\|A\|_F^2 := \mathrm{Tr}\Big(A^\dagger A\Big)$; the trace norm retains subscript 1; an unsubscripted operator norm is the spectral norm, and an unsubscripted vector norm is Euclidean. Logarithms are natural unless a base is displayed. Entropy in nats and bits is related by

$$S(\rho) = (\ln 2) S_{\mathrm{bits}}(\rho). \tag{3.4}$$

A full-rank lattice in $2N$ dimensions is specified by independent vectors v_i, written as rows of a generator matrix:

$$M_\Lambda := \begin{pmatrix} v_1^T \\ \vdots \\ v_{2N}^T \end{pmatrix}, \qquad \Lambda := \{M_\Lambda^T a : a \in \mathbb{Z}^{2N}\}, \qquad A_\Lambda := M_\Lambda \Omega M_\Lambda^T. \tag{3.5}$$

Its symplectic dual is

$$\Lambda^\perp := \{x : x^T \Omega y \in \mathbb{Z} \text{ for all } y \in \Lambda\}. \tag{3.6}$$

Symplectic integrality means that A_Λ has integer entries, equivalently $\Lambda \subseteq \Lambda^\perp$. The com-

muting GKP stabilizer generators are $\hat{S}_i := \hat{T}(\sqrt{2\pi}v_i)$. Their joint $+1$ eigenspace defines the ideal code [4]. Products carry the Weyl phase fixed by the generators:

$$\prod_{i=1}^{2N} \hat{S}_i^{a_i} = \exp\left[-i\pi \sum_{i<j} a_i a_j (A_\Lambda)_{ij}\right] \hat{T}(\sqrt{2\pi}\, M_\Lambda^T a). \tag{3.7}$$

The logical Weyl matrices used below include these phases.

An integer change of lattice basis puts A_Λ into the canonical form $\mathrm{diag}(d_1, \ldots, d_N) \otimes \omega$, with positive integers d_j. In that basis set $M_{\mathrm{sq}} := \mathrm{diag}(\sqrt{d_1}, \ldots, \sqrt{d_N}) \otimes I_2$. Writing $M_\Lambda = M_{\mathrm{sq}} S^T$ gives $S^T \Omega S = \Omega$. If $\hat{U}_S^\dagger \hat{x} \hat{U}_S = S\hat{x}$, then $\hat{U}_S \hat{T}(u) \hat{U}_S^\dagger = \hat{T}(Su)$. The Gaussian unitary $\hat{U}_S$ maps the square-code space to the space for Λ, up to a displacement setting the chosen stabilizer eigenvalues. This displacement accounts for phases introduced by an integer change of generators and does not change the logical dimension. Each square-code factor has d_j logical states. Thus the logical dimension and number of logical displacement classes are

$$d_L = \prod_j d_j = |\det M_\Lambda| = \mathrm{covol}\,\Lambda, \qquad |\Lambda^\perp/\Lambda| = d_L^2. \tag{3.8}$$

The covolume is the volume of a fundamental lattice cell. Displacements in $\Lambda^\perp$ preserve stabilizer eigenvalues and act logically; their cosets modulo Λ specify the logical Weyl operators. The dual covolume is $\mathrm{covol}\,\Lambda^\perp = (\mathrm{covol}\,\Lambda)^{-1}$. A self-dual lattice has covolume one. See Appendix B.3 for the Gaussian transformation and explicit ideal codewords.

The dual-lattice length entering the performance sums is

$$|\Lambda^\perp|_{\min} := \min_{x \in \Lambda^\perp \setminus \{0\}} \|x\|. \tag{3.9}$$

For random displacement noise, nontrivial logical displacements determine decoding distance [42, 123]; the sum used here includes all nonzero dual vectors, including stabilizer vectors. The same dual lattice controls the loss and amplification expressions below.

In Fig. 3.1, we provide phase-space visualizations of square and hexagonal GKP codes.

The plotted quantities are Wigner functions. The green and blue arrows represent logical displacements associated with dual-lattice points. The hexagonal lattice, the tightest packing lattice in two dimensions, has longer logical displacement vectors than the square lattice at the same logical dimension.

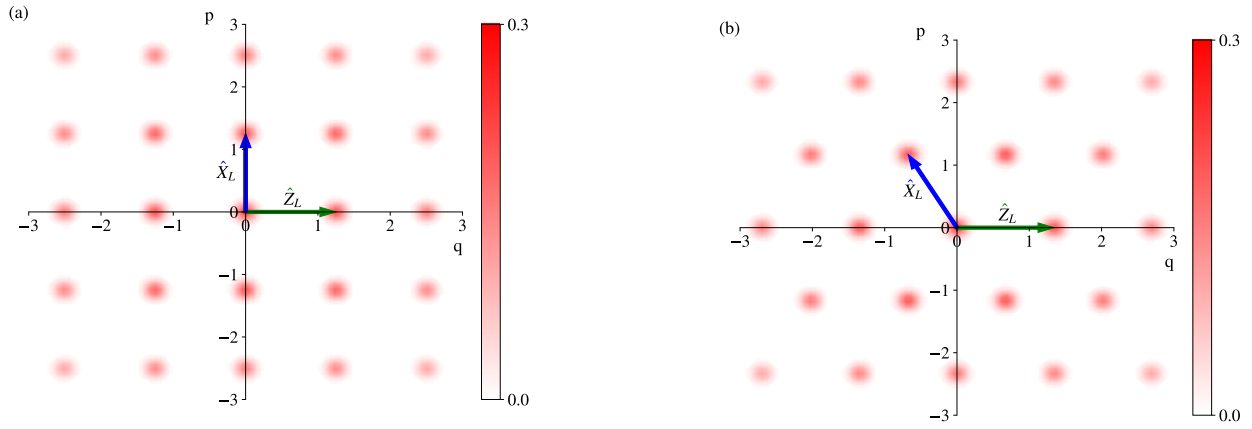


Figure 3.1: The Wigner functions of maximally mixed GKP states based on (a) square and (b) hexagonal lattices. The maximally mixed states are given by $\rho = \frac{1}{d_L}\sum_{\mu=0}^{d_L-1} |\mu\rangle\langle\mu|$, where we set $d_L = 2$ here. The green and blue arrows represent displacements that correspond to logical $\hat{Z}_L$ and $\hat{X}_L$ operators respectively.

Finite-energy codes. The ideal codewords $|\mu\rangle_0$ are not normalizable. Applying a Gaussian energy envelope gives the individually normalized codewords

$$|\mu\rangle_\Delta := N_{\Delta,\mu} e^{-\Delta^2 \hat{n}_{\text{tot}}} |\mu\rangle_0 , \qquad \hat{n}_{\text{tot}} := \sum_{j=1}^{N} \hat{a}_j^\dagger \hat{a}_j, \quad \Delta > 0. \tag{3.10}$$

Individual normalization does not impose orthogonality between different codewords. For the overlap expansion, define the raw operator $\hat{C}_\Delta |\mu\rangle := \mathcal{N}_\Delta e^{-\Delta^2 \hat{n}_{\text{tot}}} |\mu\rangle_0$. Its common positive factor sets the zero-displacement term of $\hat{G}_\Delta := \hat{C}_\Delta^\dagger \hat{C}_\Delta$ to I_L.

For each fixed $x \in \Lambda^\perp$, define the logical Weyl matrix by

$$(\hat{W}_x)_{\mu\nu} := \lim_{\Delta\to 0} {}_\Delta\langle\mu| \hat{D}(\xi(x)) |\nu\rangle_\Delta . \tag{3.11}$$

These matrices obey $\hat{W}_0 = I_L$, $\hat{W}_{-x} = \hat{W}_x^\dagger$, and $\|\hat{W}_x\|_F^2 = d_L$. With $n_\Delta := (e^{2\Delta^2} - 1)^{-1}$, the

raw displaced overlap is

$$\begin{aligned}\hat{C}_\Delta^\dagger \hat{D}(\alpha)\hat{C}_\Delta = \sum_{x\in\Lambda^\perp} \hat{W}_x e^{-(n_\Delta+1/2)(\|\xi(x)\|^2+\|\alpha\|^2)} \\ \times e^{2\sqrt{n_\Delta(n_\Delta+1)}\,\mathrm{Re}(\xi(x)^\dagger\alpha)}.\end{aligned} \tag{3.12}$$

Setting $\alpha = 0$ gives $\hat{G}_\Delta \to I_L$ at a fixed lattice as $\Delta \to 0$. The physical finite-energy encoding is the Gram-orthonormalized isometry

$$\hat{V}_\Delta := \hat{C}_\Delta \hat{G}_\Delta^{-1/2}, \qquad \hat{V}_\Delta^\dagger \hat{V}_\Delta = I_L. \tag{3.13}$$

The Gram matrix is positive for every $\Delta > 0$. See Appendix B.5 for the displaced-overlap source convention and the proof of Gram orthonormalization.

Mean codeword energy. For a fixed one-mode lattice, the codeword-averaged mean photon number of this isometry satisfies

$$\bar{n}_C := \frac{1}{d_L}\sum_\mu \langle\mu|\, \hat{V}_\Delta^\dagger \hat{n} \hat{V}_\Delta\, |\mu\rangle = n_\Delta + \mathcal{O}(n_\Delta^2 e^{-cn_\Delta}), \qquad c > 0. \tag{3.14}$$

Here c depends on the lattice. Differentiating the displaced overlap and taking the Gram-orthonormalized logical trace gives this exponentially suppressed correction. For multimode codes, $\bar{n}_C := \mathrm{Tr}\big(\hat{P}_L \hat{n}_{\mathrm{tot}}\big)/d_L$ is the total mean photon number, and $\bar{n}_C/N$ is the mean per mode.

The correction is exponentially suppressed as energy grows and can already be negligible for codes with a few photons. The commonly cited expression $\bar{n}_C \approx 1/(2\Delta^2) - 1/2$ [3, 4] is the small-Δ Taylor expansion of n_Δ in Eq. (3.14).

3.2 Gaussian-dilatable noise channels

Definition 3.1 (Gaussian-dilatable channel; [121, 122])**.** *A channel is Gaussian-dilatable if it admits*

$$\mathcal{N}(\rho) := \mathrm{Tr}_F[\hat{U}(\rho \otimes \rho_E)\hat{U}^\dagger], \tag{3.15}$$

where $\hat{U} : aE \to bF$ is a Gaussian unitary and ρ_E is a state on finitely many environmental modes.

Here a, b label the signal input and retained output, and E, F label the environmental input and discarded output. For a mixed bath, the complementary output comprises F and a purifier R of ρ_E. Figure 3.2 illustrates this dilation. The environment need not be Gaussian.

$$S_{\mathrm{bits}}(\rho_E) := -\,\mathrm{Tr}(\rho_E \log_2 \rho_E). \tag{3.16}$$

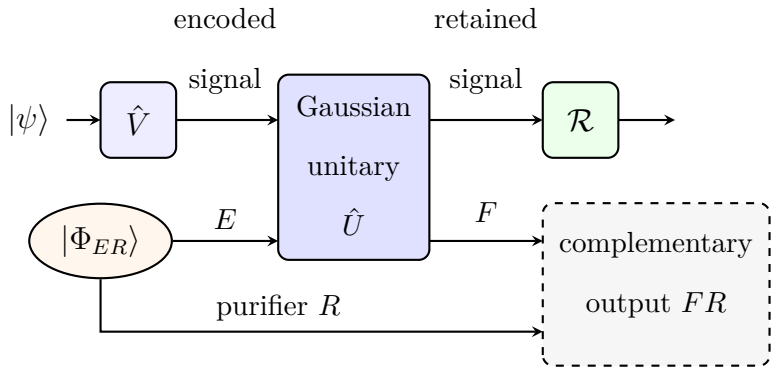


Figure 3.2: A Gaussian unitary $\hat{U}$ couples the signal to an environment E. The encoded signal is retained for recovery, while the discarded output F and the environmental purifier R form the complementary output.

For attenuation and amplification, one signal mode $\hat{a}$ interacts with one environment mode $\hat{e}$. The input factors are ordered as $\mathcal{H}_a \otimes \mathcal{H}_E$ and the output factors as $\mathcal{H}_b \otimes \mathcal{H}_f$. For a mixed bath, the purified dilation gives

$$\rho_{bfR} := (\hat{U} \otimes I_R)(\rho \otimes |\psi_{ER}\rangle\langle\psi_{ER}|)(\hat{U}^\dagger \otimes I_R), \qquad \mathcal{N}^{\mathrm{c}}(\rho) := \mathrm{Tr}_b\, \rho_{bfR}, \tag{3.17}$$

with $\rho_E = \mathrm{Tr}_R\,|\psi_{ER}\rangle\,\langle\psi_{ER}|$. A block of N independent uses has bath $\rho_E^{\otimes N}$.

The two interactions are

$$\begin{aligned} \hat{U}_{\mathrm{BS}}(\theta) &:= e^{\theta(\hat{a}^\dagger\otimes\hat{e}-\hat{a}\otimes\hat{e}^\dagger)}, \qquad & \cos\theta &= \sqrt{\eta}, \\ \hat{U}_{\mathrm{TMS}}(r) &:= e^{r(\hat{a}^\dagger\otimes\hat{e}^\dagger-\hat{a}\otimes\hat{e})}, \qquad & \cosh r &= \sqrt{G}. \end{aligned} \tag{3.18}$$

For $0 < \eta < 1$ and $G > 1$, the retained Heisenberg-picture mode obeys

$$\hat{U}^\dagger(\hat{a}_b \otimes I_f)\hat{U} = \sqrt{c_2}\,\hat{a} \otimes I_E + \sqrt{c_1}\,I_a \otimes \hat{e}, \tag{3.19}$$

with $\hat{e}$ replaced by $\hat{e}^\dagger$ for two-mode squeezing, and

$$(c_1, c_2) := \begin{cases} (1-\eta, \eta), & \text{beam splitter}, \\ (G-1, G), & \text{two-mode squeezing}. \end{cases} \tag{3.20}$$

The complementary-output displacement has the unified transformation

$$\begin{aligned} \hat{U}^\dagger(I_b \otimes \hat{D}_f(\alpha))\hat{U} &= \hat{D}_a(g(\alpha)) \otimes \hat{D}_E(\sqrt{c_2}\alpha), \\ g(\alpha) &:= \begin{cases} -\sqrt{c_1}\alpha, & \text{beam splitter}, \\ -\sqrt{c_1}\alpha^*, & \text{two-mode squeezing}. \end{cases} \end{aligned} \tag{3.21}$$

Define $\chi := c_2/c_1$, the squared bath-to-signal displacement-amplitude ratio, and

$$\beta(x) := \begin{cases} -\sqrt{\chi}\,\xi(x), & \text{beam splitter}, \\ -\sqrt{\chi}\,\xi(x)^*, & \text{two-mode squeezing}. \end{cases} \tag{3.22}$$

The Jacobian is $d^{2N}\beta = \pi^N \chi^N d^{2N}x$.

Pure loss and quantum-limited amplification are special cases where the environment is

vacuum. A thermal bath instead gives thermal loss or amplification, with

$$\rho_{\rm th}(\nu) := \sum_{m=0}^{\infty} \frac{\nu^m}{(1+\nu)^{m+1}} \left|m\right\rangle \left\langle m\right|, \qquad \nu \geq 0. \tag{3.23}$$

Table 3.1 summarizes these channels and the Gaussian random displacement channel of Eq. (2.6).

Table 3.1: Gaussian-unitary dilations of three noise channels. The additive-noise dilation uses two environmental modes.

Channel	Interaction unitary	Environmental state
Thermal loss	$\hat{U}_{\rm BS}(\theta)$, $\cos^2\theta = \eta$	$\rho_{\rm th}(\bar{n})$
Thermal amplification	$\hat{U}_{\rm TMS}(r)$, $\cosh^2 r = G$	$\rho_{\rm th}(\bar{n})$
Gaussian random displacement	$\hat{U}_{\rm add}(\sigma)$	Two-mode vacuum states

A two-mode vacuum dilation of additive noise is given by the two QND couplings $\hat{U}_{\rm add}(\sigma) := e^{-i\sqrt{2}\sigma\hat{q}\hat{p}_1} e^{i\sqrt{2}\sigma\hat{p}\hat{p}_2}$. Here $\hat{p}_1, \hat{p}_2$ act on distinct environmental modes. This gives $\hat{q}_b = \hat{q} - \sqrt{2}\sigma\hat{p}_2$ and $\hat{p}_b = \hat{p} - \sqrt{2}\sigma\hat{p}_1$, with independent Gaussian shifts of variance σ^2. Additive Gaussian noise is the canonical class B_2 and admits such a dilation with additional modes [118, 120]; it also arises as a thermal-loss limit with $\eta \to 1$ and $(1-\eta)\bar{n}$ fixed. A non-Gaussian bath generally gives a non-Gaussian channel under the same interaction.

For later vacuum specializations, the Kraus operators are

$$\hat{E}_\ell = \sqrt{\frac{(1-\eta)^\ell}{\ell!}} \eta^{\hat{n}/2} \hat{a}^\ell, \qquad \hat{A}_\ell = \sqrt{\frac{(G-1)^\ell}{G\ell!}} G^{-\hat{n}/2} (\hat{a}^\dagger)^\ell. \tag{3.24}$$

We use $\gamma := 1 - \eta$ where it simplifies a vacuum expression. For a thermal bath with probabilities $p_r = \bar{n}^r/(1+\bar{n})^{r+1}$, one can similarly write

$$\hat{K}_{\ell r} = \sqrt{p_r}\, {}_f\!\left\langle \ell \right| \hat{U} \left| r \right\rangle_E, \qquad \sum_{\ell,r} \hat{K}_{\ell r}^\dagger \hat{K}_{\ell r} = I_a. \tag{3.25}$$

The two indices resolve the discarded-output outcome and the bath eigenstate, equivalently its purification label. Choosing $\hat{U}_{\mathrm{BS}}$ or $\hat{U}_{\mathrm{TMS}}$ gives thermal loss or amplification.

3.2.1 Complementary channel view of the QEC matrix

We now express the QEC matrix of Eq. (2.9), and hence the near-optimal fidelity $\tilde{F}^{\mathrm{opt}}$ of Theorem 2.4, for a code sent through the dilated channels of Section 3.2. The QEC matrix describes information leaked to the environment and is proportional to the transposed Choi state of the encoded complementary channel. For a mixed environment the complementary output includes both the discarded output F and the environment's purifier R.

The key to expressing the QEC matrix is to purify the environmental state and use the dilated input-output relation in Eq. (3.19). Write its spectral decomposition as

$$\rho_E = \sum_r \lambda_r \left|\phi_r\right\rangle \left\langle\phi_r\right| . \tag{3.26}$$

Resolving the Kraus indices into discarded-output outcomes and purification labels gives Eq. (3.27).

$$\begin{aligned} &M_{\mu\ell r,\nu k s} \\ &= \sqrt{\lambda_r \lambda_s} \int \frac{d^{2N}\alpha}{\pi^N} \left\langle k\right| \hat{D}(-\alpha) \left|\ell\right\rangle \, \left\langle\phi_r\right| \hat{D}(\sqrt{c_2}\alpha) \left|\phi_s\right\rangle \, \times \left\langle\mu\right| \hat{V}^\dagger \hat{D}(g(\alpha)) \hat{V} \left|\nu\right\rangle , \end{aligned} \tag{3.27}$$

Here μ, ν label logical codewords, ℓ, k are discarded-output Fock outcomes, and r, s label the environmental eigenvectors and their purification register. The states $|\phi_r\rangle$ are the environmental eigenvectors in Eq. (3.26), not themselves purified states. A purification of ρ_E gives $\sum_r \sqrt{\lambda_r} \left|\phi_r\right\rangle_E \otimes \left|r\right\rangle_R$. For N independent uses, these labels denote the corresponding product eigenvectors and eigenvalues of $\rho_E^{\otimes N}$. Here $g(\alpha) := -\sqrt{c_1}\alpha$ for a beam-splitter interaction and $g(\alpha) := -\sqrt{c_1}\alpha^*$ for two-mode squeezing. The three factors are overlaps in the discarded output, environment, and code, respectively. Equation (3.27) applies to any isometric code for these dilations and is useful when $\hat{V}^\dagger \hat{D}(g(\alpha)) \hat{V}$ has a tractable form.

The reference operators in Table 3.2 serve different purposes in the infidelity proxy of Section 2.3. Define the discarded-mode thermal occupation by

$$\tilde{n}_\Delta := c_1 n_\Delta + (c_1 + c_2 - 1)/2. \tag{3.28}$$

The reference D_0 uses a thermal input with occupation equal to the code energy parameter n_Δ. For a mixed environment it need not equal D_{marg}. In vacuum it is thermal with mean $(1-\eta)n_\Delta$ for loss or $(G-1)(n_\Delta+1)$ for amplification. In D_{p}, each thermal factor acts on one discarded-output register; $\text{diag}(\lambda_r)$ acts on the environmental purification register and depends only on the bath spectrum. For N independent uses, λ_r denotes the spectrum of $\rho_E^{\otimes N}$. The global bound in Lemma 2.10 supplies a recovery guarantee for any admissible reference. Weighted-norm estimates show that the higher-order QEC terms vanish as $\Delta \to 0$ with the product reference. See Appendix B.6.1 for the expansion and Appendix B.6.2 for these estimates.

Table 3.2: Complementary references for the local expansion, thermal-loss proxy, and coding bound.

Reference	Definition	Purpose
D_{marg}	$\text{Tr}_L M/d_L$, the complementary marginal for the maximally mixed encoded input.	Unit local coefficient in the perturbative expansion of Chapter 2; numerical marginal-reference proxies.
D_0	$[\mathcal{N}^c(\rho_{\text{th}}(n_\Delta))]^T$ per use; take the tensor product over N uses.	Zero-displacement term of the raw QEC matrix; closed-form thermal-loss proxy.
D_{p}	$\rho_{\text{th}}(\tilde{n}_\Delta)^{\otimes N} \otimes \text{diag}(\lambda_r)$.	Product reference with full support on the retained complementary space; coding-limit bound.

3.3 Performance of GKP codes

In this section we use the infidelity proxy to approximate GKP performance. We start with finite-energy expressions and then take the infinite-energy limit, where only the zeroth-order environmental kernel remains. Pure loss and amplification provide explicit vacuum benchmarks and comparisons with existing decoders.

3.3.1 Finite-energy case

Inserting the displaced-codeword overlap into the Weyl integral gives an ordered expansion of the QEC matrix in powers of $1/n_\Delta$, which is small in the large-photon-number regime. Each dual-lattice contribution contains an overlap with a displaced copy of the lattice and a correction from the Gaussian envelope. The ordered expansion controls the finite-energy perturbative expressions below. With the product reference, the higher-order terms vanish in the infinite-energy limit, leaving only the zeroth-order environmental kernel for the rate calculation. See Appendix B for the full QEC-matrix expression and convergence proof.

For a kernel f, define its lattice sum

$$S_f(\Lambda) := \sum_{x\in\Lambda\setminus\{0\}} f(x). \tag{3.29}$$

Given the codeword overlaps, the thermal-loss Gaussian integrals can be evaluated analytically. The performance sums below are over the dual lattice $\Lambda^\perp$, rather than the stabilizer lattice Λ. For example, at a fixed one-mode lattice and thermal-loss parameters, the QEC matrix M_Δ of the normalized encoding has the proxy

$$\tilde{\epsilon}_{D_0}(M_\Delta) = S_{\kappa_{\bar{n}}}(\Lambda^\perp) + \mathcal{O}(n_\Delta^{-1}). \tag{3.30}$$

Here $\kappa_{\bar{n}}(x)$ is the radial thermal-loss kernel, with $v := \chi\|\xi(x)\|^2$,

$$\kappa_{\bar{n}}(x) := e^{-(2\bar{n}+1)v} \sum_{j\in\mathbb{Z}} \frac{I_{|j|}\left(2v\sqrt{\bar{n}(1+\bar{n})}\right)}{4\cosh^2\left(\frac{j}{4}\ln t_E\right)}. \tag{3.31}$$

Here $t_E := \bar{n}/(1+\bar{n}) > 0$ and I_j is the modified Bessel function. In vacuum, $\kappa_0(x) = e^{-v}/4$. Thermal occupation changes this leading kernel, whereas finite-energy corrections are suppressed by $1/n_\Delta$. Summable Gaussian bounds control the remainder after lattice summation. The high-energy expression isolates the lattice contributions and their dependence on the channel parameters. The dashed curves in Fig. 3.3(c) retain the full finite-energy dependence, including higher-order corrections omitted in Eq. (3.30) and exact Gram orthonormalization. They agree closely with the near-optimal infidelity points. Notably, evaluating this proxy through lattice sums and scalar quadrature is substantially cheaper at high energy than our direct evaluation of Eq. (2.15), which represents the codewords and QEC matrix in a growing Fock basis.

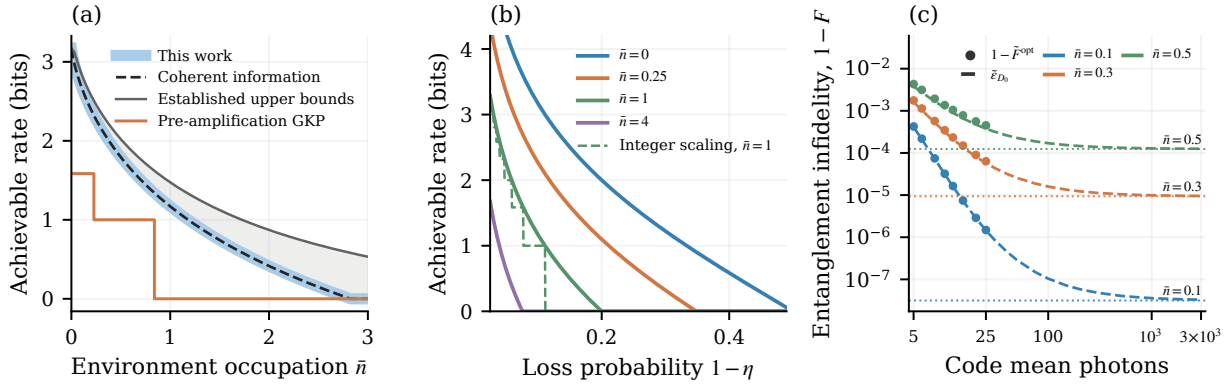


Figure 3.3: Achievable GKP rates and finite-energy performance under thermal loss. Rates are in bits per channel use. (a) At $\eta = 0.9$, the attained rate (thick pale line) coincides with the infinite-energy coherent information of a thermal state (dashed), against the pre-amplification GKP construction [1] and the minimum of three established upper bounds [2]. (b) Continuous rates versus the integer-scaling staircase. (c) At $\eta = 0.95$, points show near-optimal infidelity $1 - \tilde{F}^{\mathrm{opt}}$ for square qubit codes at 5–25 mean photons. Dashed curves are the separate D_0-reference infidelity proxy $\tilde{\epsilon}_{D_0}$, from the exact Gram-orthonormalized finite-energy lattice-pair integral, extended to 3000 photons. Dotted lines are its analytic infinite-energy limits, not near-optimal-fidelity asymptotes.

Monotonicity with added energy. Previous numerical studies observed decreasing GKP infidelity with increasing energy under pure loss in the regimes studied [3, 13, 30]. By contrast, improvement with added energy is not guaranteed for cat or binomial codes [3]. Our proxy gives an analytical account of the high-energy trend. For a fixed one-mode lattice under vacuum loss or quantum-limited amplification, the high-energy derivative with reference D_0 is

$$\frac{d\tilde{\epsilon}_{D_0}(M_\Delta)}{dn_\Delta} = -\frac{8c_2 - 3}{32c_1^2 n_\Delta^2} S_h(\Lambda^\perp) + \mathcal{O}(n_\Delta^{-3}), \tag{3.32}$$

where $h(x) := \|\xi(x)\|^2 e^{-\chi\left\|\xi(x)\right\|^2}$ and $S_h(\Lambda^\perp) > 0$. This proves decrease of the proxy at large energy for vacuum loss with $\eta > 3/8$ and all quantum-limited amplification channels. For each fixed lattice and $\eta > 3/8$, eventual decrease also holds for thermal loss with sufficiently small fixed environmental occupation. This is monotonicity of the proxy and its certified bound; it does not imply universal strict monotonicity of exact recovery fidelity. See Appendix B.9 for the proof and parameter restrictions.

3.3.2 Infinite-energy limit

As $\Delta \to 0$, the proxy reduces to environmental displacement overlaps sampled at dual-lattice points. Using the spectrum of the block environment $\rho_E^{\otimes N}$ as in Eq. (3.26), the QEC matrix M_Δ and product reference D_{p} in Table 3.2 give

$$\lim_{\Delta\to 0} \tilde{\epsilon}_{D_{\mathrm{p}}}(M_\Delta) = S_{f_{\rho_E}}(\Lambda^\perp). \tag{3.33}$$

Here $\beta(x) := -\sqrt{\chi}\,\xi(x)$ for a beam splitter, with $\xi(x)$ complex conjugated for two-mode squeezing. This is a lattice sum of the environmental kernel function, with the block length N implicit in f_{ρ_E},

$$f_{\rho_E}(x) := \sum_{r,s} \frac{\lambda_r \lambda_s |\langle\phi_r|\hat{D}(\beta(x))|\phi_s\rangle|^2}{(\sqrt{\lambda_r} + \sqrt{\lambda_s})^2}. \tag{3.34}$$

The spectrum enters through the prefactor weights, and the eigenvectors through their displacement overlaps.

Evaluating the spectral kernel in Eq. (3.34) for a thermal state gives $\kappa_{\bar{n}}$. The D_0 limit gives the dotted lines in Fig. 3.3(c). In vacuum, $D_{\mathrm{p}} = D_0$ and $f_{\rho_E}(x) = \kappa_0(x)$, recovering the pure-loss and amplification results of Section 3.3.3, Eq. (3.40). The thermal-loss kernel $\kappa_{\bar{n}}(x)$ is radial and positive. Non-Gaussian environments can instead produce directional structure or zeros in $f_{\rho_E}(x)$. Section 3.5 examines how such zeros change the preferred lattice geometry.

3.3.3 Pure loss and amplification

Setting $\rho_E = |0\rangle\langle 0|$ specializes the general expansion to pure loss and quantum-limited amplification. The vacuum displacement overlap is Gaussian. The corresponding Kraus-product representation is

$$\hat{K}_l^\dagger \hat{K}_k = \int \frac{d^2\alpha}{\pi} e^{-\tau|\alpha|^2/2} \langle l|\, \hat{D}(\alpha^*)\, |k\rangle\ \hat{D}\Big(\alpha\sqrt{|1-\tau|}\Big)\,, \tag{3.35}$$

where $\tau = \eta$ for loss and $\tau = G$ for amplification. This is the vacuum corollary of the general Weyl representation, with the displacement conventions reconciled in Appendix B.8. The same general expansion yields the following finite-energy formula and the decoder comparisons of Ref. [13].

In the following, we focus on single-mode GKP code under loss to highlight the key takeaways. See Appendix B.8 for the full multimode loss and amplification expressions. For a GKP code with d_L logical dimensions and lattice Λ undergoing loss with transmissivity η, writing $\gamma := 1-\eta$ for the loss probability in this expression and in Eq. (3.39), its QEC matrix expression can be written as

$$\begin{aligned} M_{\mu l,\nu k} = (1+\delta)\frac{t^{(l+k)/2}}{\gamma n_\Delta + 1} \sum_{n_1,n_2\in\mathbb{Z}} & e^{i\pi n_2(n_1+(\mu+\nu)/d_L)} \\ & \times e^{-\frac{\pi}{2}\frac{1-\gamma}{\gamma+1/n_\Delta}|\mathbf{L}|^2} \langle l|\, \hat{D}\left(\sqrt{\frac{\pi(n_\Delta+1)}{\gamma n_\Delta+1}}(L_1 - iL_2)\right)|k\rangle\,. \end{aligned} \tag{3.36}$$

where $\delta = O\left(e^{-\pi\left(n_\Delta+\frac{1}{2}\right)\left|\Lambda^\perp\right|_{\min}^2}\right)$, $t := \frac{\gamma n_\Delta}{\gamma n_\Delta+1}$, and $n_\Delta := \frac{1}{e^{2\Delta^2}-1}$. Here, we have defined the symplectic dual lattice to be $\Lambda^\perp$, and $\left|\Lambda^\perp\right|_{\min}$ is defined in Eq. (3.9). δ exponentially decreases

with the input energy and is the direct consequence of the GKP codewords being non-orthogonal for finite Δ. Suppose $M_{\Lambda^\perp}$ generates $\Lambda^\perp$, the vector $\mathbf{L} := M_{\Lambda^\perp}^T \left(d_L n_1 + \mu - \nu, n_2\right)^T$ enumerates lattice points over a sublattice of the dual lattice such that

$$\Lambda^\perp = \bigcup_{\mu-\nu\in\mathbb{Z}_{d_L}} \{\mathbf{L} | n_1, n_2 \in \mathbb{Z}\} \,. \tag{3.37}$$

The appearance of the symplectic dual lattice is natural since we are working with displacement operators, and the displacements that exactly connect infinite-energy codewords are the ones that correspond to dual lattice points.

With a different motivation, Ref. [3] derived the QEC matrix expression for single-mode square lattice GKP qubit code. While the general form is similar, their derivation is restricted to square lattices and assumed $\Delta \ll 1$ to simplify the derivations along the way. In contrast, the only approximation we made is in the prefactor of $(1+\delta)$. Moreover, our derivation is valid for arbitrary multimode lattices.

For these vacuum perturbative benchmarks, write

$$\tilde{\epsilon} := \tilde{\epsilon}_{D_{\mathrm{diag}}}(M), \qquad D_{\mathrm{diag}} := \mathrm{diag}(\mathrm{Tr}_L\, M/d_L) \tag{3.38}$$

in the displayed Kraus basis. This provides a different reference from D_0 and D_{p}. In vacuum, it is a good approximation to the marginal reference D_{marg}. In the infinite-energy limit for pure loss and amplification, it converges to $D_0 = D_{\mathrm{p}}$, yielding the same limiting proxy.

With the QEC matrix expression, we arrive at the near-optimal performance of the GKP code through the perturbative infidelity. The full infidelity expression contains summations over all dual sublattices given in Eq. (3.37). Nevertheless, the leading order contributions are given by the closest lattice points of the dual lattice. Retaining only those leading terms gives [13]

$$\tilde{\epsilon} \approx \frac{m}{2} e^{-2\pi(n_\Delta + 1/2)|\Lambda^\perp|^2_{\min}} \left[8 \sum_{j=0}^{\infty} \frac{t^j I_{2j}(z)}{(1+t^j)^2} - I_0(z) - 1 \right]. \tag{3.39}$$

Table 3.3: Examples of symplectic integral lattices parameters [14, 15]. The kissing number is the number of closest lattice points in the dual lattice. The determinant of the lattice is connected to the GKP code's logical dimension through $|\det M_\Lambda| = d_L$, where M_Λ is the lattice generator matrix.

Lattice	Modes N	$\lvert\Lambda^\perp\rvert^2_{\min}$	Kissing number m	$\lvert\det M_\Lambda\rvert$
Square	1	1	4	1
Hexagonal	1	$2/\sqrt{3}$	6	1
Tesseract	2	$1/\sqrt{2}$	8	2
D_4	2	$\sqrt{2}$	24	1
D_6	3	1	12	2
E_8	4	2	240	1
Leech	12	4	196560	1

where m denotes the kissing number (the number of closest lattice points with distance $\left|\Lambda^\perp\right|_{\min}$), I denotes modified Bessel functions, and $z := \pi(n_\Delta + 1)\sqrt{t}\left|\Lambda^\perp\right|^2_{\min}$. Some examples of single-mode GKP codes that are of interest include the square and hexagonal lattice GKP codes, whose relevant parameters are provided in Tab. 3.3.

From Eq. (3.34), the general environmental kernel reduces in vacuum to $f_0(x) = e^{-\pi\chi\|x\|^2}/4$. In the infinite-energy limit, D_{p} and D_{diag} converge to each other and give

$$\lim_{\Delta\to 0} \tilde{\epsilon}_{D_{\mathrm{p}}}(M_\Delta) = \frac{1}{4} \sum_{x\in\Lambda^\perp\setminus\{0\}} e^{-\pi\chi\|x\|^2}. \tag{3.40}$$

The earlier vacuum calculation [13] uses a diagonal-reference perturbative proxy and obtains the same Gaussian lattice sum as its single-mode high-energy expression and multimode upper estimate. The global reference bound supplies the corresponding recovery guarantee.

As an example, if we consider the leading contributions from the closest lattice points, the single-mode square lattice GKP code's performance is given by $\tilde{\epsilon} \approx e^{-\frac{\pi}{d_L}\frac{\eta}{1-\eta}}$. Such scaling can be compared with the performance of GKP code under amplification decoder (AD), which gives an infinite-energy error probability of $e^{-\frac{\pi}{4d_L}\frac{\eta}{1-\eta}}$ [3] for square lattice. Our infidelity expression for near-optimal recovery shows a significant improvement over that of AD – increasing the exponent by a factor of 4. This enhancement is not only practically important

for high-performance bosonic QEC, but also crucial for our later proof that the GKP code achieves quantum channel capacity.

When the GKP code is considered in the context of displacement noise, the connection between the dual lattice and the GKP code's performance is apparent. However, to the best of our knowledge, this is the first time GKP's performance under loss (or amplification) has been rigorously connected to its dual lattice. The dual lattice spacings and the relevant phase have their origins in the QEC matrix expression since we decomposed the noise channel Kraus operators as displacement operators.

In Fig. 3.4, we explicitly compare the performance of GKP decoders and lattices under loss and amplification. In subfigures (a, b) and (c, d), the noise channels are loss channels with transmissivity $\eta = 0.9$ (loss probability 0.1) and amplification channels with gain $G = 1.1$, respectively. In (a, c), we focus on the square lattice GKP code encoding $d_L = 2$ logical dimensions. We present a comparison between the analytical and perturbative expression given in Eq. (3.39) (green lines) and the *exact* values of the near-optimal infidelity, $1 - \tilde{F}^{\text{opt}}$ (blue triangles). The exact values are computed through Eq. (2.15). Clearly, the perturbative expression approximates the exact values very well. Moreover, the asymptotic limits are shown in orange dashed lines, which performs surprisingly well: even with 10% loss, GKP code can still achieve infidelities of order 10^{-6}.

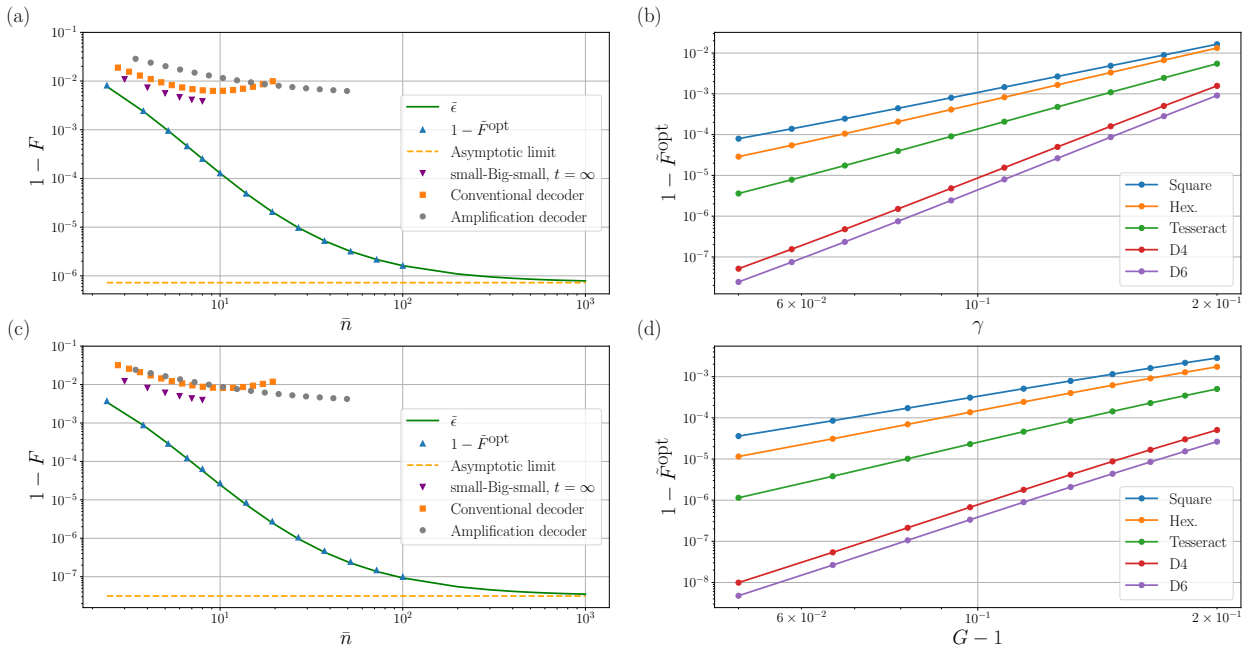


Figure 3.4: The performance of GKP codes under (a, b) loss and (c, d) amplification. (a, c) The infidelities of near-optimal recovery are compared against that of existing GKP decoders, such as AD [1, 3], the conventional decoder [4], and the sBs circuit [5]. In particular, the sBs circuit is repeated until the system reaches a stabilized state. The GKP codes are based on square lattice. The noise channels are set to be (a) $\eta = 0.9$ (10% loss) for loss and (c) $G = 1.1$ for amplification. (b, d) The performance of various lattices with parameters given in Tab. 3.3 is presented. The encoding dimension is fixed to $d_L = 2$ with $\bar{n}_C/N = 5$ average photon number per mode.

While the near-optimal fidelities reveal the limit of GKP code, we examine the performance of a few existing GKP code recoveries. For example, proposed along with the GKP code [4], the conventional decoder is designed against displacement errors and recovers the information through modular quadrature measurement and adaptive displacement recovery. As shown in Fig. 3.4 (a, c), when it is applied against loss or amplification noise, its performance does not always improve with increasing photon numbers. Small-Big-small (sBs) [5] is arguably the state-of-the-art recovery against loss, and it is widely adopted in superconducting circuit [32, 33, 127] and ion trap experiments [35, 36] for the correction and preparation of GKP states. A round of sBs consists of a short circuit comprised of conditional displacements and ancilla qubit rotations. As an optimistic benchmark, we show its stabilized fidelity after multiple rounds of sBs. Last but not least, AD is the decoder that established the state-of-the-art rate of GKP code under pure and thermal loss [1]. At the cost of injecting

additional noise, it converts loss or amplification channels into Gaussian random displacement channels, whose optimal decoder at infinite energy is known to be the conventional decoder. See Appendix B.12 for the simulation details and exact definition of the recoveries.

Overall, the comparisons between recoveries lead to the conclusion that in experimentally accessible regimes, e.g. $\bar{n}_C \sim 10$ and $\eta = 0.9$, GKP code's performance can be improved by order(s)-of-magnitude solely through refining its recovery. The improvement can be even more significant for smaller loss. Recall that the near-optimal fidelity is achievable by the transpose channel. Thus, one promising route is to approximate the transpose channel or other near-optimal recovery channels through techniques such as optimal control. We provide insights on the form of the truncated transpose channel in Appendix B.13 and leave the construction of experimentally feasible circuits for future work.

In Fig. 3.4 (b, d), we show the near-optimal infidelities that correspond to various lattices types. The average photon number per mode is fixed to $\bar{n}_C/N = 5$, and all codes encode $d_L = 2$ dimensional logical space. The lattices and their respective number of modes, N, shortest dual lattice vector sizes, and the kissing numbers are given in Tab. 3.3. In particular, it is apparent that as we increase the number of modes that encode the same amount of information, we effectively have more redundancy that we can use to protect the information, which is quantified by the distance of the dual lattice $\left|\Lambda^{\perp}\right|_{\min}$. Some lattices utilize space more effectively than others, which is related to the well-known sphere packing problem. For example, in two- and four-dimensional space, the tightest packing is achieved by the hexagonal and the $D4$ lattice respectively, rendering them more potential in correcting errors.

3.4 Achievable rates

Harrington and Preskill established GKP-achievable rates under Gaussian random displacement noise [43]. Using the near-optimal-fidelity method of Ref. [30], our loss and amplification analysis [13] established capacity attainment at the noise strengths allowed by integer scaling of symplectically self-dual lattices. Both constructions have an integer floor in their rate guar-

antees. Their lattice arguments use the Buser–Sarnak bound over the family of symplectically self-dual lattices for rotationally invariant kernels [128].

We subsequently removed the floor for Gaussian random displacement and pure-loss channels by concatenating square GKP qudits with discrete-variable codes [129]. The displacement-noise construction uses quantum polar codes with analog syndrome information; the pure-loss proof uses random outer stabilizer codes, finite-field code averaging, and Gaussian theta-function estimates. The outer code gives $d_L = d^K$ for an $[[N, K]]_d$ code, allowing the rate $(K/N)\log_2 d$ to vary continuously in the large-block limit.

Here we address the broader class of Gaussian-dilatable channels and remove the floor with a different lattice construction [44]. We combine the symplectic Siegel first-moment identity [130], valid for nonradial kernels, with prime-index sublattices. This accommodates the displacement-overlap kernels of general mixed environments.

3.4.1 Gaussian-dilatable channels

We consider one-way transmission through N independent uses of a channel $\mathcal{N}$, with collective encoding and recovery [76]. An achievable rate combines the amount of logical information with its transmission fidelity:

$$R := \lim_{N\to\infty} \frac{1}{N} \log_2 d_L, \qquad F \to 1. \tag{3.41}$$

The quantum capacity C_Q is the supremum over encoding-recovery families. Coherent information supplies a lower bound, and its regularization gives capacity [112, 113, 115]:

$$C_Q = \lim_{N\to\infty} \frac{1}{N} \sup_{\rho} I_c(\rho, \mathcal{N}^{\otimes N}), \tag{3.42}$$

where

$$I_c(\rho, \mathcal{N}) := S_{\text{bits}}(\mathcal{N}(\rho)) - S_{\text{bits}}(\mathcal{N}^c(\rho)). \tag{3.43}$$

Entropies and rates in these expressions are in bits. For pure loss and quantum-limited amplification,

$$C_Q = \left[\log_2 \frac{\tau}{|1-\tau|}\right]_+, \quad \tau = \eta \text{ or } G, \tag{3.44}$$

which sets the benchmark for the vacuum special cases [118].

The general kernel gives the following achievable-rate result [44].

Theorem 3.2 (GKP rates for general environments)**.** *For finite $S_{\text{bits}}(\rho_E)$, the supremum of GKP-achievable rates satisfies*

$$R_{\text{GKP}} \;\geq\; \left[\log_2 \chi - S_{\text{bits}}(\rho_E)\right]_+, \tag{3.45}$$

Every positive rate below the right-hand side is achieved by a family with entanglement fidelity tending to one.

The environment need not be Gaussian or have finite mean photon number. Achievability is energy-unconstrained in the limit $\Delta \to 0$.

For a thermal environment of occupation $\bar{n}$, $S_{\text{bits}} = g_2(\bar{n})$ with $g_2(n) := (n+1)\log_2(n+1) - n\log_2 n$, and Eq. (3.45) becomes the positive part of the infinite-energy coherent information of a thermal state for both loss and amplification [Fig. 3.3(a,b)]. This lower bound was also known for non-Gaussian environments of finite mean energy [131]. Our result establishes GKP-family attainment under the weaker finite-entropy assumption.

The proof uses the product reference in Table 3.2 and the global bound from Section 2.3. The limiting lattice sum is averaged over an ensemble of symplectic lattices, while typical-subspace and finite-Fock approximations control a general environmental state. See Appendix B.10 for the construction and fidelity transfer.

Equation (3.33) separates channel properties from code geometry. The coupling and environmental state determine the kernel $f_{\rho_E}(x)$, while the GKP lattice determines the points at which it is sampled. Lattice averaging replaces this discrete sum by an integral through Eq. (3.46). Combining the symplectic mean-value identity with prime-index sublattice

counting gives a lattice average that holds for nonradial kernels. For a uniformly random index-m sublattice Λ of a Haar-random symplectic self-dual lattice, with m prime, every nonnegative integrable kernel f obeys

$$\mathbb{E}_\Lambda S_f(\Lambda^\perp) = m \int_{\mathbb{R}^{2N}} f(x)\, d^{2N}x. \tag{3.46}$$

This ensemble has logical dimension $d_L = m$ and removes the integer-floor restriction of uniformly scaled lattices.

Equation (3.33) implicitly assumes a finite environmental spectrum and finite Fock support. Rigorously, typicality bounds the retained rank by $2^{N[S_{\mathrm{bits}}(\rho_E)+\delta_b]}$. Finite-Fock approximation regularizes the retained eigenvectors, with vanishing trace-distance error. Moreover, lattice averaging and Weyl orthogonality remove dependence on the environmental eigenbasis. For every fixed $\delta_b > 0$ and sufficiently large N, the regularized block channel obeys

$$\mathbb{E}_\Lambda\left[\lim_{\Delta\to 0} \tilde{\epsilon}_{D_{\mathrm{p}}}(M_\Delta)\right] \le \frac{m}{4}\, 2^{-N[\log_2 \chi - S_{\mathrm{bits}}(\rho_E)-\delta_b]}. \tag{3.47}$$

This leads to Theorem 3.2 when the logical dimension m is suitably chosen for the target encoding rate. See Appendix B.10 for the full derivation and fidelity transfer to the original channel.

Recall that single-mode Gaussian channels fall into six canonical classes A_1, A_2, B_1, B_2, C, and D, up to Gaussian unitaries [118]. These unitaries can be absorbed into the encoder and decoder. We further extend this result to Gaussian-dilatable channels with finite-energy environments through the following corollary.

Phase-insensitive, phase-preserving Gaussian channels are also a subclass of Gaussian-dilatable channels [118] and admit the decomposition [132]

$$\mathcal{N} = \mathcal{A}_G \circ \mathcal{L}_\eta. \tag{3.48}$$

Here $\mathcal{L}_\eta$ is pure loss ($0 < \eta \leq 1$) and $\mathcal{A}_G$ is quantum-limited amplification ($G \geq 1$). Figure 3.5 uses this decomposition to compare the scope of the rate result with prior GKP results.

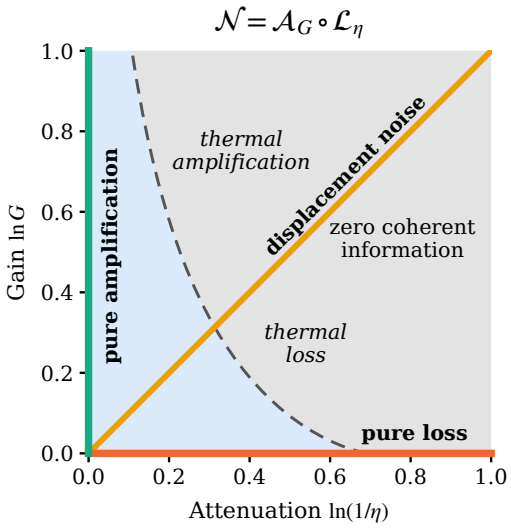


Figure 3.5: Phase-insensitive, phase-preserving Gaussian channels factor as $\mathcal{A}_G \circ \mathcal{L}_\eta$. Pure loss, pure amplification, thermal loss and amplification, and additive displacement noise occur within this landscape. Colored lines mark earlier GKP results [13, 43, 129]; the shaded region identifies the scope of the present coherent-information guarantee. The gray region has zero infinite-energy thermal-state coherent information; the dashed curve marks the boundary of that benchmark.

Corollary 3.3. *Consider such a channel $\mathcal{N}$ with one input and one output mode. Suppose a thermal state achieves infinite-energy coherent information $I_c^\infty(\mathcal{N}) := \lim_{n\to\infty} I_c(\rho_{\text{th}}(n), \mathcal{N})$. Then, for every $0 < R < [I_c^\infty(\mathcal{N})]_+$, there exists a GKP family achieving rate R.*

See Appendix B.10.4 for the canonical-channel reductions and full proof.

3.4.2 Pure loss constructions from few-mode lattices

The finite-energy code benchmarks of Section 3.3 focus on recovery fidelity, while the following vacuum examples also account for the logical information transmitted. Another equally important metric is the amount of logical information transmitted. In fact, the near-optimal fidelity derived has close connections with the logical dimension, d_L, as shown in Eq. (3.39). In Fig. 3.4(b,d), the comparison is made between various lattices transmitting the same amount of information, $d_L = 2$. However, this is not necessarily a fair comparison since some use more resources, e.g. number of modes, than others. In other words, the series of

lattices we examine have a vanishing rate. In this light, we show a similar performance plot in Fig. 3.6(a). The difference is that here we are showing the upper bound for the infinite-energy performance, as given in Eq. (3.40), and the logical dimensions are set to $d_L = 3^N$ with N being the number of modes used for each lattice type. Recall that the quantum capacity of loss channel is $\log_2 \frac{\eta}{1-\eta}$ number of qubits per mode. This essentially means that at this logical dimension rate, there exists an asymptotic code family that can transmit information arbitrarily well as we increase the number of modes if $1 - \eta < \frac{1}{d_L^{1/N}+1} = 0.25$.

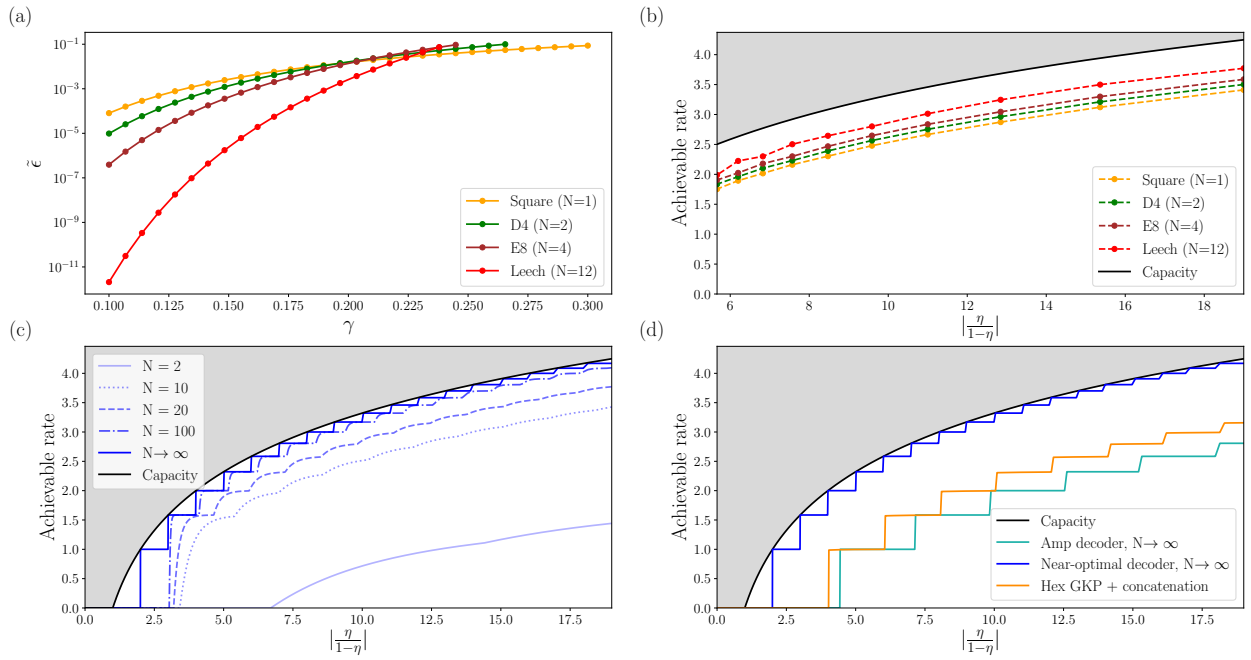


Figure 3.6: (a) The infidelities of GKP code are based on a few lattices with varying numbers of modes with a constant encoding rate such that $d_L = 3^N$. (b) Coherent information of GKP codes is based on various lattices, which are achievable rates. The coherent information is computed through the Werner states as given in Eq. (3.49). (c) Lower bounds on the achievable rates of GKP codes based on scaled symplectically self-dual lattices with an increasing number of modes, N. The infidelity is upper bounded through Eq. (3.51). (d) A comparison of the rates achieved through the amplification decoder, the near-optimal decoder, and the single-mode hexagonal GKP code concatenated with random stabilizer codes for asymptotic system sizes. All subplots are focusing on infinite-energy GKP codes.

The main purpose of Fig. 3.6(a) is to show qualitative support on the existence of a threshold for GKP code, i.e. below certain noise threshold, the error is arbitarily suppressed with increasing system size. The (pseudo)thresholds of these lattices are not expected to

be strictly 0.25 for the following reasons. Firstly, while the threshold is defined in the asymptotic limit and corresponds to certain good lattice families, the lattices shown are low-dimensional ones and are not guaranteed to be in the same optimal lattice family. Moreover, the estimated pseudothreshold could be inaccurate: the plotted metric is $\tilde{\epsilon}$, which is a perturbative approximation of the near-optimal fidelity. At finite error, the near-optimal fidelity can have non-negligible difference in its crossover points compared to the optimal one. Additionally, the approximation error, as mentioned below Eq. 3.38, has a dependence on the logical dimension in the finite-error regime, leading to an increased approximation error for high-dimensional lattices, such as the Leech lattice, especially at around the threshold.

The achievable rate serves as a single benchmark that incorporates both the fidelity and the logical dimension. Other than following the original definition of the achievable rate, we can lower bound the achievable rates through coherent information. In particular, with a code of fidelity F encoding d_L dimensional logical space, it can be converted into a Werner state with coherent information [1]

$$I(F, d_L) = \log_2 d_L + F \log_2 F + (1 - F) \log_2 \frac{1 - F}{d_L^2 - 1}. \tag{3.49}$$

whose regularized version, $I\left(F, d_L\right)/N$, is an achievable rate. In Fig. 3.6(b), the achievable rates based on each base lattice are shown. For each base lattice, we compute their achievable rates with varying logical dimensions and plot the supremum. The logical dimensions are varied through scaling the base lattices. However, not all logical dimensions can be achieved through this approach. Assuming the base lattice has a logical dimension of d_0, we are only able to achieve logical dimensions of the form [4, 15, 43]

$$d_L = \lambda^N d_0 \tag{3.50}$$

where λ is any positive integer and N is the number of modes. See Appendix B for more details. Therefore, it is possible for the envelope of the achievable rates to be not smooth.

Moreover, the rates are computed with $\tilde{\epsilon}$ as an approximation of $\tilde{F}^{\mathrm{opt}}$. As shown in Eq. (3.38), such an approximation only holds when the code performance is expected to be promising. Recall that the rates are optimized over all possible logical dimensions, we set the numerical constraint that we only optimize over logical dimensions such that $\tilde{\epsilon} \leq 10^{-2}$ to ensure the rates we compute are reliable.

For comparison, the earlier integer-scaling construction uses $\sqrt{\lambda}\Lambda_0$, where Λ_0 is symplectically self-dual and λ is a positive integer. It encodes $d_L = \lambda^N$ and, for pure loss, gives the benchmark

$$\lim_{n_\Delta \to \infty} \tilde{\epsilon} \leq \frac{1}{4}\left(\lambda \frac{1-\eta}{\eta}\right)^N. \tag{3.51}$$

Figure 3.6(c) shows rates inferred from this estimate at finite block lengths. The large-block staircase in panel (d) reflects the restriction to integer λ. The prime-index ensemble above removes this restriction and attains every rate below the vacuum-channel capacities. The few-mode curves remain useful benchmarks for finite blocks; their packing geometry need not belong to an asymptotically capacity-achieving family.

In the past, there have been multiple works on the achievable rates of GKP code. In particular, Ref. [1] established the state-of-the-art rate against loss channel before this work. The key to their approach is the amplification decoder (AD), which converts the (thermal) pure loss channel into a Gaussian random displacement channel through an amplification channel. However, the amplification channel inevitably introduces noise from the environment port of the amplification channel, leading to the suboptimality of AD. The rates of GKP code under AD are shown in Fig. 3.6(d), which has a finite gap of $\log_2 e \approx 1.4$ compared to the capacity. Whether the gap was due to the deficiency in the recovery was left as an open question. Our result demonstrates that the GKP encoding with a (near-)optimal recovery is sufficient to close such a gap.

An alternative method to obtain rates from bosonic codes is to concatenate them with qudit codes. For example, on the lower level, we can select infinite-energy single-mode hexagonal GKP code. If we perform twirling, the noise channel can be understood as a

generalized Pauli channel or a depolarizing channel, depending on the specific twirl we perform. The only constraint is that the summed error probability should equal $1 - F$. To obtain a lower bound on the rates, we assume the channel is twirled such that we are dealing with a qudit depolarizing channel, i.e. the error probability vector is $\mathbf{p} = \left(F, \frac{1-F}{d_L^2-1}, \ldots, \frac{1-F}{d_L^2-1}\right)$. The achievable rate of concatenating it with random stabilizer codes is given by the hashing bound (see Appendix B for more details) and is shown in Fig. 3.6(d). Note that the hashing bound is strictly below the achievable rate of the genuine multimode lattices. As pointed out in Ref. [43], while concatenated single mode GKP lattices can also be understood as lattices living in $2N$ dimensions, they do not achieve as good a packing as the sphere-packing lattices like the scaled self-dual lattices.

One interesting conclusion from our approach based on scaled self-dual lattice is that there exists a GKP code that simultaneously achieves the capacity of loss and amplification given that

$$\frac{\eta}{1-\eta} = \frac{G}{G-1} \tag{3.52}$$

with tailored decoders. This observation has its root in the almost identical form of loss and amplification Kraus operator combinations in Eq. (3.35). See Appendix B.8 for the derivation.

3.5 Lattice geometry beyond sphere packing

Summing the environmental displacement-overlap kernel over nonzero dual-lattice vectors gives the infinite-energy infidelity proxy. A smaller sum gives a smaller certified upper bound on infidelity. For vacuum loss and quantum-limited amplification, the kernel is a radially decreasing Gaussian. Maximizing the shortest dual-lattice distance is then a natural design principle, especially when the nearest shell dominates, much as packing motivates GKP design against Gaussian displacement noise [1, 13]. Exact optimization still involves the full lattice sum. A non-Gaussian environment can produce nodes and nonmonotone radial

weights, so increasing the shortest distance alone need not improve the proxy.

Thermal-loss lattice optimality. For fixed $0 < \eta < 1$ and finite $\bar{n} \geq 0$, the hexagonal lattice uniquely minimizes the infinite-energy D_0 proxy among one-mode GKP lattices of fixed covolume, up to orthogonal transformations. For $\bar{n} > 0$, the thermal kernel has the positive Gaussian-mixture representation

$$\begin{aligned} \kappa_{\bar{n}}(x) &= \int_{\mathbb{R}} \varrho_F(k) e^{-a_{\bar{n}}(k)v}\, dk, \qquad v := \chi\|\xi(x)\|^2, \\ a_{\bar{n}}(k) &:= 2\bar{n} + 1 - 2\sqrt{\bar{n}(1+\bar{n})}\cos(k \ln t_E). \end{aligned} \tag{3.53}$$

Here $t_E = \bar{n}/(1+\bar{n})$ and $\varrho_F(k) := 2k/\sinh(2\pi k)$, with $\varrho_F(0) = 1/\pi$. This weight is positive and has integral $1/4$, while

$$0 < (\sqrt{1+\bar{n}} - \sqrt{\bar{n}})^2 \leq a_{\bar{n}}(k) \leq (\sqrt{1+\bar{n}} + \sqrt{\bar{n}})^2.$$

The kernel is therefore completely monotone in the squared displacement v. Montgomery's theorem [133] minimizes each planar Gaussian lattice sum by the hexagonal lattice at fixed covolume. Integration with positive weight preserves this minimum and its uniqueness. In vacuum, $\kappa_0(x) = e^{-v}/4$ and the same theorem applies directly. See Appendix B.11 for the full proof. In higher dimensions, Gaussian-sum optimization also favors the universally optimal E_8 and Leech lattices at fixed density [134].

Non-Gaussian environmental nodes. For the Fock-state environment $\rho_E = |m\rangle\langle m|$ [6], the identity $\langle m|\hat{D}(\beta)|m\rangle = e^{-|\beta|^2/2} L_m(|\beta|^2)$ gives

$$\kappa_m(x) = \tfrac{1}{4} e^{-v} L_m(v)^2, \qquad v := |\beta(x)|^2 = \chi\|\xi(x)\|^2. \tag{3.54}$$

Here L_m is the Laguerre polynomial of degree m. Its zeros are zeros of the overlap kernel. For $m = 1$, the derivative with respect to v is $e^{-v}(1-v)(v-3)/4$: the weight decreases on

$(0,1)$, increases on $(1,3)$, and decreases thereafter. The full lattice sum depends on shell multiplicities and their positions relative to these nodes.

Figure 3.7 compares the full proxy sums for rectangular qubit lattices $\Lambda_s^{\perp} = \mathrm{diag}(s, s^{-1})\mathbb{Z}^2/\sqrt{2}$ at $\eta = 0.65$. For $m = 1$, a rectangular member has a smaller proxy than the square and hexagonal lattices. Its shortest shell lies near a Laguerre zero. In environmental displacement coordinates the points are

$$\beta_{jk} = \sqrt{\frac{\pi\chi}{2}}(sj + ik/s), \qquad (j,k) \in \mathbb{Z}^2. \tag{3.55}$$

The cell aspect ratio s^2 is distinct from the kernel radius v. For $s \geq 1$, the nearest axial shell has $v = \pi\chi/(2s^2)$. At the $m = 1$ minimum it lies at $v \simeq 1.115$, close to the node at 1; contributions from the remaining shells shift the minimum away from the node.

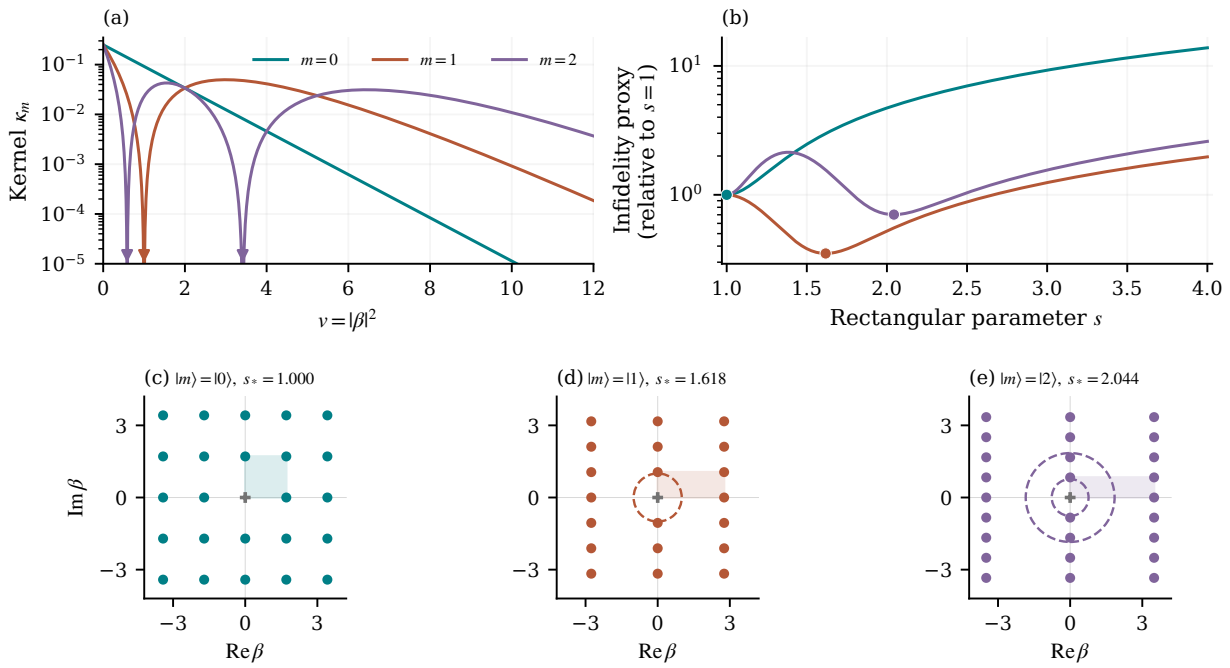


Figure 3.7: Environmental overlap nodes change the preferred GKP lattice geometry. The environment is $|m\rangle$, $m = 0, 1, 2$ [6]. (a) Exact kernels κ_m; triangles mark Laguerre zeros. (b) Full infinite-energy proxy sums for rectangular qubit dual lattices $\Lambda_s^{\perp} = \mathrm{diag}(s, s^{-1})\mathbb{Z}^2/\sqrt{2}$ at $\eta = 0.65$, divided by the square-lattice proxy at $s = 1$ for the same environment. Curves show the normalized infidelity proxy, not the actual infidelity. Dots mark numerical rectangular minima on $1 \le s \le 4$. (c)–(e) Fidelity-optimized dual lattices in β coordinates, obtained by minimizing the infinite-energy infidelity proxy with $s \ge 1$. Dashed rings mark Laguerre zeros, shading marks one cell, and the cross marks the excluded origin. The cell aspect ratio is s^2.

Voss *et al.* [135] explain zero-ring suppression of complementary logical information for hexagonal GKP codes with pure Fock environments in the characteristic-function picture. The spectral overlap kernel in Eq. (3.34) includes mixed environments and their purifiers. For a mixed environment, the discarded mode alone is not the full complementary output: a zero of $\mathrm{Tr}\left[\rho_E \hat{D}(\beta)\right]$ need not cancel the off-diagonal spectral overlaps entering the kernel. For a pure environment the kernel reduces to $\frac{1}{4}|\mathrm{Tr}\left[\rho_E \hat{D}(\beta)\right]|^2$, recovering the zero-ring mechanism. Minimizing its lattice sum optimizes the certified infidelity bound.

Environmental structure can also support higher rates than the entropy-only guarantee of Theorem 3.2. Wang and Jiang [136] construct lattice-matched encodings and pure GKP environmental states that permit perfect transmission of a d-dimensional logical system in

the ideal limit under their matching conditions. For example, their construction transmits a qubit at $\eta = 1/3$, and finite-energy approximations give positive coherent information below $\eta = 1/2$. Our beam-splitter lower bound $[\log_2\{\eta/(1-\eta)\} - S_{\text{bits}}(\rho_E)]_+$ vanishes there even for pure environments; it is an achievable-rate guarantee, not the maximum GKP rate.

A squeezed-vacuum environment $\hat{S}(r)\,|0\rangle$, with $\hat{S}(r) := \exp\Big[r(\hat{e}^2 - \hat{e}^{\dagger 2})/2\Big]$, instead gives the anisotropic kernel

$$f_{\text{sq}}(\beta) = \tfrac{1}{4}\exp\Big[-e^{2r}(\operatorname{Re}\beta)^2 - e^{-2r}(\operatorname{Im}\beta)^2\Big]. \tag{3.56}$$

Vacuum, one-photon, and squeezed-vacuum environments are all pure, with $S_{\text{bits}} = 0$ and $\int_{\mathbb{C}} f(\beta)\, d^2\beta/\pi = 1/4$, yet their kernels have different shapes. In lattice coordinates this normalization includes the χ^{-1} Jacobian. Their rate lower bounds agree, while their preferred finite-dimensional lattices can differ.

3.6 Discussion

We have connected finite-energy GKP performance under loss and amplification to communication rates through the QEC matrix and complementary-channel overlaps. The vacuum expressions apply to multimode codes with arbitrary lattice shapes. At experimentally accessible parameters, near-optimal recovery gives a benchmark that can outperform existing decoders by orders of magnitude. The dual lattice controls the leading errors, and refining recovery alone can substantially improve performance without changing the encoding [13].

The near-optimal-fidelity method of Chapter 2 supplies both numerical benchmarks and analytical rate guarantees. The global complementary-reference bound makes this approach valid in the coding limit, where logical dimensions and reference spectra change with block length. Combined with symplectic averaging and prime-index sublattices, it removes the integer floor of the earlier vacuum construction and extends GKP rate guarantees to finite-entropy environments coupled by attenuation or amplification. For Gaussian-dilatable channels with finite-energy environments, the extension attains the infinite-energy thermal-

input coherent-information benchmark. Thermal loss and amplification, which remained open in the vacuum study, are included in this result.

Environmental kernels also guide lattice design. Thermal-loss kernels preserve planar hexagonal optimality through a positive Gaussian-mixture representation. Non-Gaussian overlap nodes and directional weights can favor different shell distributions; the entropy-only rate guarantee does not exhaust these environmental advantages. The dilated QEC expression and reference bound apply to arbitrary isometric encodings and may help analyze other multimode families, including quantum spherical codes [88].

Developing practical recoveries that approach these benchmarks remains a challenge. Approximating the transpose channel through optimal control is one possible route; see Appendix B.13 for its truncated form. Quantifying achievable rates under finite-energy constraints is another direction: the present coding theorem uses finite-energy codes at every block length but allows their energy to grow without a fixed bound.

Ancilla error propagation is also a main bottleneck for practical GKP error correction. In the small–Big–small (sBs) stabilization protocol, ancilla decay during a large controlled displacement can produce an uncorrectable shift of the oscillator [5]. Routes toward fault tolerance include flag qubits that detect dangerous faults [137], path-independent gates that make the intended operation insensitive to correctable ancilla errors [138], and biased-noise ancillas whose dominant errors do not propagate into the encoded mode. Kerr-cat ancillas offer one such approach [139]. Fluxonium is another candidate: our cross-resonance proposal provides oscillator conditional-displacement gates with a long-lived, noise-biased fluxonium ancilla [90].

CHAPTER 4

HIGH-RATE SURGERY FOR QUANTUM LOW-DENSITY PARITY-CHECK CODES

Quantum error correction (QEC) is a cornerstone of large-scale, fault-tolerant quantum computation [22, 24, 25, 51]. Much of the literature has focused on reducing the space-time overhead of quantum memories—that is, the cost of storing quantum information reliably over time. However, beyond protecting idle qubits, practical quantum computers should also support logical operations with comparably low overhead [26].

Quantum low-density parity-check (qLDPC) codes [45] are quantum error-correcting codes with sparse parity-check matrices. Though nonlocal, they are ideal for low-overhead quantum memories and form the basis of many recent experiments and proposals [11, 47, 48, 140–142]. Nevertheless, it remains open whether they retain this overhead advantage for processing quantum information. Well-known approaches to fault-tolerant logical operations include transversal (and related) gates [45] and automorphism gates [143–146], but these are often tailored to specific code families and lack full addressability. Originating in lattice surgery for topological codes [8, 26, 147–149], code surgery [7, 9, 10, 150–157] is one of the few schemes available for arbitrary qLDPC codes. It enables universal computation via Pauli-product measurements (PPMs) [26, 158] combined with magic-state injection [159].

Inspired by techniques from quantum weight reduction [160–162], there has been tremendous progress in reducing the overhead of code surgery. Nevertheless, most work has focused on measuring a single logical operator. Such an assumption fundamentally limits the scalability of surgery schemes for dense computation. From a space-time perspective, these protocols adopt ancilla systems that are inherently low rate, which limits the information-processing density—how many logical operations are performed per physical operation. Asymptotically, this can negate the benefits of high-rate quantum memory.

In this work, we propose a framework for high-rate surgery that significantly lowers

overhead; see Fig. 4.1(a) for a comparative illustration. The key idea is to employ high-rate ancilla patches for information processing. We formalize the logical action of such surgeries in a homological framework [151, 156, 163, 164] and prove sufficient conditions for fault tolerance. We further give explicit constructions of high-rate ancillas that are distance preserving. When the memory possesses suitable structure, our ancilla construction achieves amortized constant overhead for up to $\mathcal{O}(k)$ logical operations. Hypergraph-product (HGP) codes [165] are examples of such structured codes.

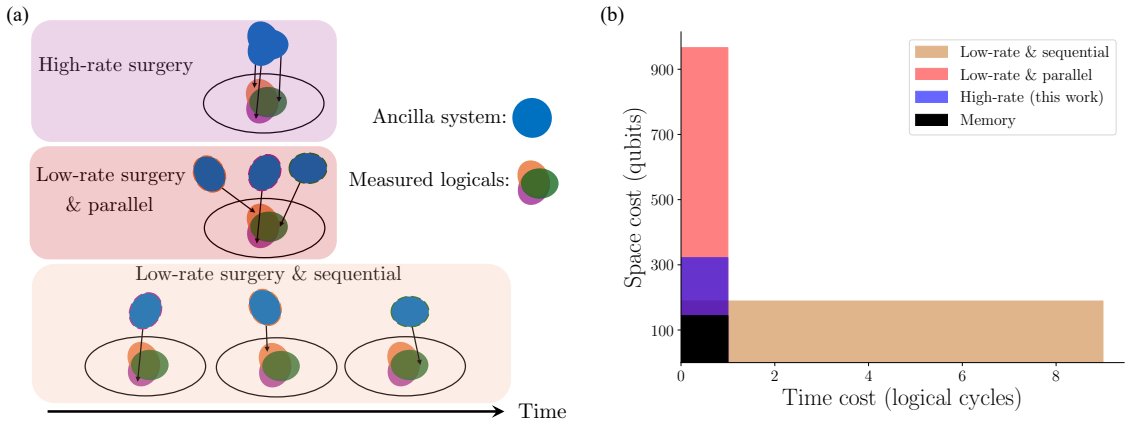


Figure 4.1: Illustration of high-rate versus low-rate surgery schemes. (a) Comparison of low-rate (sequential [7, 8] and parallel [9, 10]) and high-rate surgery scheme (this work). The horizontal axis represent time axis. Blue circles denote ancilla systems that couple to multiple target logical operators, whose supports are shown by the green, orange, and pink circles, possibly with overlaps. (b) The space (vertical axis) and time (horizontal) costs of surgery schemes on Gross code [11], $[[144, 12, 12]]$, for measuring nine randomly chosen logicals, compared against memory. Space costs are measured in unit of physical qubits. Time costs are in units of logical cycles.

In the absence of code structure, we demonstrate the power of high-rate surgery with a randomized ancilla-construction algorithm that is high rate and introduces only minimal increases in connectivity. We benchmark its performance on Gross codes [11, 140] and spatially coupled (SC) codes [166], using a randomly sampled logical basis and measurements. For the latter, we present new code families with encoding rates up to 25% and block lengths up to 1125 qubits. On these practical instances, high-rate surgery achieves nearly constant space-time overhead across a range of code distances and sizes. The constant factor is modest—within a factor of two of the memory baseline (the space-time cost of

one syndrome-extraction cycle on an idling system). Figure 4.1(b) compares space-time costs across schemes. In contrast to low-rate approaches, high-rate surgery has spatial and temporal footprints comparable to memory. Our results suggest that logical operations can be made nearly as cheap as memory, marking a step toward constant-overhead, fault-tolerant quantum computation.

The space overhead of a quantum memory with code parameters $[[n, k, d]]$ is captured by its *encoding rate*, $r := k/n$. High-rate qLDPC codes cap this overhead by achieving a *constant encoding rate*, $r = \Theta(1)$, whereas in low-rate codes $r \to 0$ as $n \to \infty$.

When computing on encoded logical qubits, one must also account for space–time trade-offs. Thus, the total cost of computation is best quantified by the *space–time overhead*. For a circuit C with width $W(C)$ and depth $D(C)$, define its space–time cost as $S(C) := W(C) \times D(C)$. An encoded circuit $\overline{C}$ implementing the same logical operations has a corresponding cost $S(\overline{C})$. The *space–time overhead*, $S(\overline{C})/S(C)$, quantifies the additional resources—extra physical qubits for encoding redundancy and additional circuit depth for fault-tolerant logical operations—needed to perform the computation reliably in the presence of noise.

In typical fault-tolerant protocols, each elementary logical operation consists of a *logical cycle*, comprising multiple rounds of syndrome extraction to reliably infer syndromes and remove entropy from the system. Although complementary works explore ways to shorten the logical cycle [146, 167–171], we assume that for a distance-d code each logical cycle takes time $t_L = \Theta(d)$ unless stated otherwise. We then define the *normalized space–time overhead* as

$$\alpha := \frac{1}{t_L} \frac{S(\overline{C})}{S(C)}, \tag{4.1}$$

which measures overhead in units of (physical qubits) $\times$ (logical cycles).

As a baseline, an idling memory operation—one code block undergoing a single logical cycle—has normalized overhead equal to the inverse of the encoding rate, $\alpha = \Theta(1/r)$. For constant-rate qLDPC codes, this baseline overhead is $\Theta(1)$. Our goal is to design logical

Table 4.1: Summary of asymptotical spacetime overheads of various surgery schemes, compared against memory, for a $[[n, k, d]]$ qLDPC code. Space overhead is normalized by unencoded circuit cost, and time overhead is normalized by the length of a logical cycle, $\Theta(d)$. For simplicity, assume $\Theta(k)$ logicals are measured, and the memory is constant encoding rate. $\Theta(1)^*$ assumes Theorem 4.6. $\tilde{\Theta}(d)^\dagger$ assumes minimum-weight basis is available.

Operation	Scheme	Space overhead	Time overhead
Logical measurement	Low-rate surgery, sequential	$\Theta(1)$	$\Theta(k)$
Logical measurement	Low-rate surgery, parallel	$\tilde{\Theta}(d)^\dagger$	$\Theta(1)$
Logical measurement	High-rate surgery	$\Theta(1)^*$	$\Theta(1)$
Memory	Syndrome extraction	$\Theta(1)$	$\Theta(1)$

operations whose overhead approaches this memory limit.

4.1 Code surgery framework

In what follows, we consider logical circuits built from logical Pauli-product measurements (PPMs), which, together with magic-state injection, realize universal Pauli-based computation [26, 159]. For concreteness, we focus on X-type PPMs for a Calderbank–Shor–Steane (CSS) code [86, 87]; measurements of Z-type logicals follow symmetrically. Simultaneous measurement of mixed-type Paulis is discussed in Methods.

CSS codes are quantum codes whose parity-check matrices admit a generating basis split into X- and Z-type checks, H_X and H_Z, respectively. Their commutation relations can be expressed with a 2-chain complex $\mathcal{C}$:

$$\mathcal{C}: \; C_2 \xrightarrow{\partial_2^C} C_1 \xrightarrow{\partial_1^C} C_0 \;, \tag{4.2}$$

with boundary maps $\partial_2^C = H_X^T$ and $\partial_1^C = H_Z$ satisfying $\partial_1^C \partial_2^C = 0 \bmod 2$. The spaces C_l, for $l = 2, 1, 0$, are $\mathbb{F}_2$-vector spaces whose bases are associated with X-checks, qubits, and Z-checks, respectively.

To perform a surgery logical operation on a data code $\mathcal{C}$, we attach an ancilla system

$$\mathcal{A}: \quad A_1 \xrightarrow{\partial_1^A} A_0 \xrightarrow{\partial_0^A} A_{-1} \ , \tag{4.3}$$

where A_l with $l = 1, 0, -1$ are associated with the X checks, qubits, and Z checks of $\mathcal{A}$, respectively. We attach $\mathcal{A}$ to $\mathcal{C}$ and aim to extract certain logical information of $\mathcal{C}$ through measurements of the X checks A_1 on $\mathcal{A}$. The "attaching" procedure, and the resulting joint data-ancilla code, are described by the following complex [156]:

Definition 4.1 (Code surgery)**.** *For a data code $\mathcal{C}$ and an ancilla system $\mathcal{A}$, code surgery is uniquely defined by the complex*

$$\begin{array}{ccccc} A_1 & \xrightarrow{\partial_1^A} & A_0 & \xrightarrow{\partial_0^A} & A_{-1} \\ & \searrow^{\Gamma_1} & & \searrow^{\Gamma_0} & \\ C_2 & \xrightarrow{\partial_2^C} & C_1 & \xrightarrow{\partial_1^C} & C_0 \end{array} \tag{4.4}$$

where the diagram commutes, i.e., $\Gamma_0 \partial_1^A = \partial_1^C \Gamma_1$. Here Γ_1 (respectively Γ_0) describes how the X checks A_1 of $\mathcal{A}$ (respectively the Z checks C_0 of $\mathcal{C}$) are connected to the qubits C_1 of $\mathcal{C}$ (respectively A_0 of $\mathcal{A}$).

See also Figure 4.2 for a Tanner-graph illustration of the merged code. After one logical cycle on the merged code in Eq. (4.4), during which the logical information of $\mathcal{C}$ is reliably extracted, we split the two codes and return to the codespace of $\mathcal{C}$ by transversally measuring out the ancilla $\mathcal{A}$. While we omit the details of the merge-and-split procedure here, it is crucial to initialize and measure $\mathcal{A}$ in the correct bases to ensure fault tolerance. The required procedure has been established in prior works [7, 8, 147, 172], and a complete description is provided in Methods. We also refer interested readers to Ref. [152] for a review of the history and development of surgery techniques.

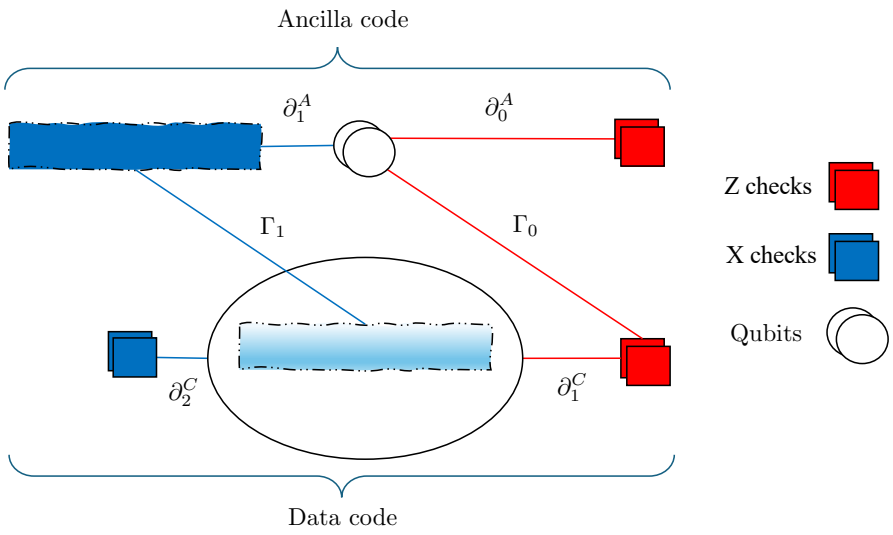


Figure 4.2: Tanner graph illustration of a general surgery scheme. Checks (resp. qubits) are represented by squares (resp. circles) [7]. The shown figure presents the ancilla system merged with the data system to measure X-logical operators, according to Eq. (4.4). At the end, the ancilla qubits are transversally measured in Z-basis. The example illustrates a transversal-type Γ_1 such that ancilla X checks and the measured operators' support (indicated by the box with a color gradient) are of similar sizes.

The essential idea of the surgery procedure is weight reduction: global logical X operators in the data code are decomposed into products of local stabilizers from the X checks A_1 of the merged code, whose values are reliably inferred during a logical cycle. This merge-and-split weight-reduction process is also known in the condensed-matter community as gauging and ungauging symmetries [173, 174]. More concretely, the logical information extracted by surgery is given by the product of stabilizers in A_1 with no support on the ancilla qubits A_0 [156]:

Lemma 4.2 (Information extracted through code surgery)**.** *For a surgery procedure in Def. 4.1, the space of all measured operators of* $\mathcal{C}$ *is*

$$M = \Gamma_1 \left(\ker \partial_1^A \right) \subseteq S_X^C \oplus L_X^C, \tag{4.5}$$

where S_X^C *and* L_X^C *denote the spaces of the* X *checks and the* X *logical operators of* $\mathcal{C}$*, respectively.*

As implied by Lemma 4.2, the number of bits of logical information extracted is $\dim M$, which is closely tied to the kernel dimension of ∂_1^A. Formally, the information-extraction efficiency is measured by the information-extraction rate:

Definition 4.3 (Information extraction rate (IER))**.** *Consider a surgery procedure with space of measured operators M in Lemma 4.2 using an ancilla system $\mathcal{A}$. The information-extraction rate is*

$$r_M := \dim(M)/|\mathcal{A}|, \tag{4.6}$$

where $|\mathcal{A}| = \sum_{i=-1}^{1} \dim A_i$ is the total size of the ancilla system (including qubits and checks).

We now show that, just as the encoding rate r of a code governs the space overhead of a quantum memory, the IER r_M controls the space–time overhead of a surgery-based quantum processor. Consider a shallow yet dense circuit executing $t = \Theta(k)$ PPMs on k qubits, where each qubit participates in $\mathcal{O}(1)$ PPMs. This implies each PPM involves only $\mathcal{O}(1)$ qubits. Such a pattern corresponds to parallel computation on all $\Theta(k)$ logical qubits, for example injecting magic states on all logical qubits in parallel, which in an unencoded setting can be executed in depth $\mathcal{O}(1)$. The spacetime overhead associated with surgery measurements using an ancilla system $\mathcal{A}$ with fixed IER r_M is [1]

$$\alpha = \underbrace{\frac{|\mathcal{A}|}{k}}_{\text{space}} \times \underbrace{\frac{t}{r_M\,|\mathcal{A}|}}_{\text{time}} = \Theta\left(\frac{1}{r_M}\right), \tag{4.7}$$

where the space factor $\frac{|\mathcal{A}|}{k}$ compares the total ancilla size to the space cost of the unencoded circuit, and the time factor reflects measuring t PPMs in $\frac{t}{r_M|\mathcal{A}|}$ logical cycles, each cycle completing $r_M|\mathcal{A}|$ PPMs using $\mathcal{A}$ once.

The above analysis admits space–time trade-offs. In a *sequential* surgery scheme with $r_M|\mathcal{A}| \ll t$, the space cost can be small because one can take $|\mathcal{A}| = \mathcal{O}(k)$, but the time cost is large due to sequential execution (see the yellow box in Fig. 4.1(a)). In a *parallel* surgery

[1] The spacetime overhead of the full scheme is $\alpha = \frac{|\mathcal{C}|+|\mathcal{A}|}{k+t} \times \frac{t}{r_M|\mathcal{A}|} = \mathcal{O}(1/r_M)$.

scheme with $r_M|\mathcal{A}| = \Theta(t)$, the PPMs complete in parallel with small time cost, yet the space cost can become large with $|\mathcal{A}| = \Theta(t)/r_M$ when r_M is low (see the pink box in Fig. 4.1(a)). In all cases, the overall space–time overhead is characterized by the IER r_M.

Most prior surgery schemes are low rate: they operate under the restriction that the ancilla system consists of one or more disjoint subsystems each with $\dim\ker\partial_1^A = 1$ and $|\mathcal{A}| = \tilde{\Omega}(d)$, which yields IER $r_M = \tilde{\mathcal{O}}(1/d)$. Consequently, they incur large overhead in either space or time, depending on whether the implementation is parallel or sequential (see Fig. 4.1(b)).

Therefore, in this work we introduce *high-rate* surgery schemes. In particular, we design ancilla systems that extract $\dim(M) = \Theta(k)$ bits of information collectively, using a joint ancilla of bounded size $|\mathcal{A}| = \mathcal{O}(n)$ (see the purple box in Fig. 4.1), which for constant-rate codes implies a constant IER $r_M = \Theta(1)$ and thus low space-time overhead by the analysis above (see Fig. 4.1(b)). We note that in the numerical analysis of the space-time overhead in Fig. 4.1(b) and Fig. 4.3, we also include the memory overhead in Eq. (4.7), which significantly increases the overall cost of sequential schemes but only marginally impact the parallel schemes.

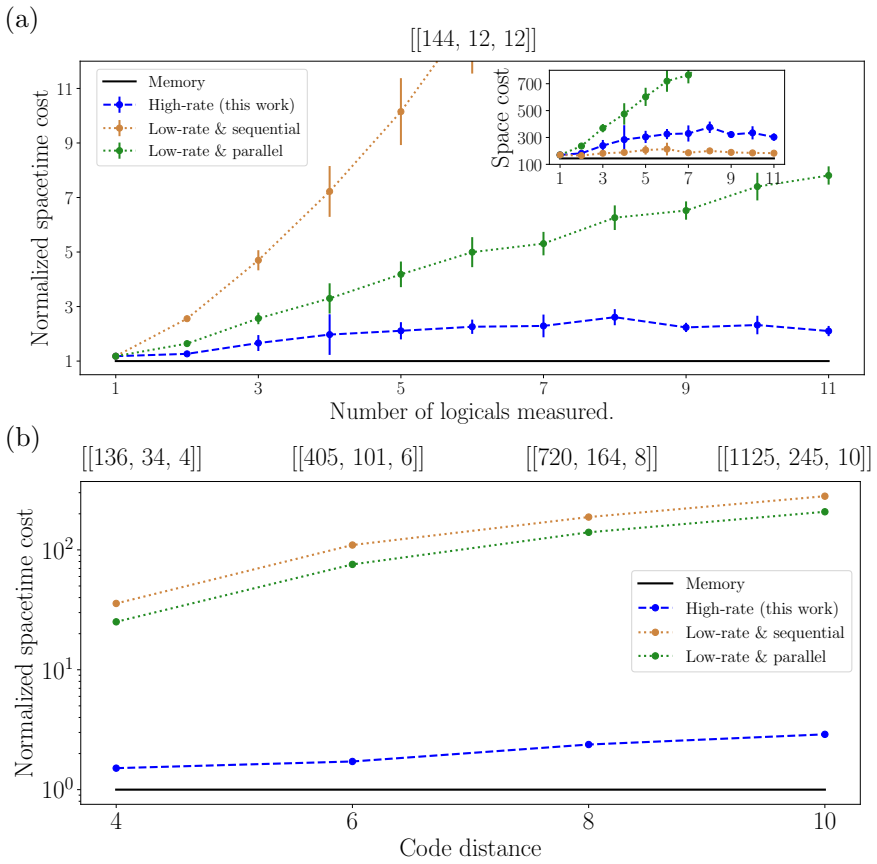


Figure 4.3: Numerical analysis of the space-time costs of surgery schemes. The spacetime costs are normalized by memory spacetime cost. The logical basis and logicals to measure are randomly sampled. The merge code degrees are limited to the data code degree plus two. Each data point is obtained through 6 sets of samples. (a) Cost for measuring varying number of logicals on Gross code. (Inset) the maximum space cost (with only data qubits). (b) Cost for measuring logicals in a set of high-rate spatially-coupled codes $\{[[n_i, k_i, d_i]]\}$ with increasing k_i and d_i (the code parameters are labeled on top). For each code instance, a random sample of $\lfloor 90\% \times k_i \rfloor$ number of logicals are measured.

A code-surgery procedure is fault-tolerant if (i) the boundary and chain maps in Eq. (4.4) are LDPC and (ii) the phenomenological distance of the procedure is $\Omega(d)$. The first condition guarantees low-degree circuit connectivity, limiting error propagation under circuit-level noise; the second ensures sufficient error protection throughout surgery. Together these conditions imply the existence of thresholds for surgery-based computation [175].

Past work shows that the phenomenological distance is preserved if and only if the merged-code distance is preserved [7, 150]. Although the merged code in Eq. (4.4) is a stabilizer code with commuting checks, we treat it as a subsystem code by identifying the

new logical operators in the ancilla—logical Z and X associated with $H_0(\mathcal{A})$ and $H^0(\mathcal{A})$—as gauge operators, since these degrees of freedom do not carry logical information. We refer to the dressed distance of this subsystem code as the merged-code distance d_{merge} for brevity.

The low-rate surgery schemes satisfy these two fault-tolerance conditions by forcing $\mathcal{A}$ to be a graph and leveraging graph-theoretic tools [176]. However, high-rate schemes inevitably require $\mathcal{A}$ to be a hypergraph, which creates new challenges for fault tolerance. For example, random-expander arguments no longer apply, and the gauge logicals of $\mathcal{A}$ cannot be promoted to stabilizers. We therefore introduce the following lemma as a sufficient condition for preserving the merged-code distance, using the "soundness" of a linear map [167]. We say a boundary ∂ has soundness ρ up to t if $\rho \leq \frac{|\partial x|}{d(x,\ker \partial)}$ for all x with $|\partial x| \leq t$, where the reduced distance is $d(x, V) := \min_{u \in V} |x + u|$. see Appendix C.2 for a more detailed introduction to expansion-related properties.

Lemma 4.4 (Sufficient condition for preserving merged-code distance)**.** *If ∂_1^A has soundness ρ up to $t \geq d$, and the ancilla code $\mathcal{A}$ has X-distance $d_A^X \geq d$, then the merged-code distance satisfies*

$$d_{\text{merge}} \geq \min\left(1,\ \frac{\rho}{|\Gamma_1|}\right) d, \tag{4.8}$$

where $|\Gamma_1|$ denotes the maximum degree of Γ_1.

We remark that this condition is sufficient but not necessary [157], especially when additional structure is present in the data code or the chain maps. A more detailed discussion appears in Appendix C.3, and we note where this condition can be relaxed when appropriate.

Next, we present high-rate, fault-tolerant surgery schemes, using algebraic constructions for structured codes and random constructions for unstructured codes.

4.2 Algebraic construction for structured codes

A core challenge in designing high-rate surgery schemes is to design ancilla codes and chain maps that preserve the merged-code distance, since the soundness condition in Lemma 4.4 can be hard to satisfy. To begin, we provide an ancilla construction that provably preserves distance; we later analyze which logical measurements it can perform. The ancilla is a hypergraph-product (HGP) code, denoted $\mathcal{A} = \mathrm{HGP}(\mathcal{B}, \mathcal{D})$, formed as a tensor product of two base classical codes $\mathcal{B}\colon\ B_1 \xrightarrow{\partial_B} B_0$ and $\mathcal{D}\colon\ D_1 \xrightarrow{\partial_D} D_0$, where ∂_B and ∂_D are the classical check matrices. Let $\mathcal{B}$ and $\mathcal{D}$ have parameters $[n_B, k_B, d_B]$ and $[n_D, k_D, d_D]$, respectively. See Methods for a detailed definition of the HGP code. Under this construction, the transpose of its X-check, ∂_1^A, is known to satisfy a soundness property [167] that is weaker than the required linear soundness in Lemma 4.4. We close this gap by taking one base code, e.g., $\mathcal{D}$, to be an expander code.

Theorem 4.5 (HGP ancilla with one expander base code is distance-preserving)**.** *Consider a surgery procedure equipped with a LDPC commuting diagram and a data code, $\mathcal{C}$, with distance d. Let the ancilla code $\mathcal{A} = HGP(\mathcal{B}, \mathcal{D})$. If $\min(d_B, d_D) \geq d$ and $\mathcal{D}$ is an expanding code up to weight $t = \Theta(d)$, the merge code distance is preserved as $d_{merge} = \Theta(d)$.*

With the fault tolerance guarantee, we now analyze HGP ancilla's IER. First, its kernel is

$$\ker \partial_1^A = \ker \partial_B \otimes \ker \partial_D, \tag{4.9}$$

which has $\dim \ker \partial_1^A = k_B k_D$. In addition, the ancilla size $|\mathcal{A}| = \Theta(n_B n_D)$. Assuming the existence of the commuting diagram, we arrive at

$$r_M = \frac{\dim \Gamma_1(\ker_1^A)}{|\mathcal{A}|} \sim \Theta\left(\frac{k_B k_D}{n_B n_D}\right). \tag{4.10}$$

Thus, using high-rate classical codes $\mathcal{B}$ and $\mathcal{D}$ leads to high IER of the surgery scheme, which is formalized as a theorem below.

Theorem 4.6 (High-rate surgery with HGP ancilla achieves constant overhead)**.** *For qLDPC code $\mathcal{C}$, a surgery scheme with HGP ancilla $\mathcal{A} = \mathrm{HGP}(\mathcal{B}, \mathcal{D})$ achieves constant IER if (1) the base codes $\mathcal{B}$ and $\mathcal{D}$ have constant encoding rate and (2) there exist a commuting diagram with chain maps Γ_1 and Γ_0, where Γ_1 does not significantly reduce the kernel dimension ∂_1^A, i.e. $\dim \Gamma_1(\partial_1^A) = \Theta(\dim \partial_1^A)$. The scheme is distance-preserving if it additionally satisfies Theorem 4.5.*

Our ancilla construction can be interpreted as a high-rate "HGP extractor", as opposed to the low-rate graph-based extractors [152]. Here, the philosophy of an extractor is that given a fixed ancilla with guaranteed fault-tolerance properties, one design its connectivity, Γ_1 and Γ_0, depending on the structures in target quantum codes and logical measurements.

To characterize the logicals that can be measured in parallel requires specification of the data code and a logical basis. While we defer a detailed discussion of the general symmetries required to Appendix C.4.2, the discussion is straightforward when we consider the HGP code as the data code, which admits a canonical logical basis. For example, consider data code $\mathcal{C} = \mathrm{HGP}(\mathcal{B}', \mathcal{F})$. Due to the product structure of $\mathcal{C}$, any X-logical in the canonical basis is supported in one of the columns and taking the form of a codeword of $\mathcal{B}'$. As a result, the challenge of finding a quantum code homomorphism in Theorem 4.6 reduces to finding classical code homomorphisms between the classical base codes $\mathcal{B}$ and $\mathcal{B}'$,

$$\begin{array}{ccc} B_1 & \xrightarrow{\partial_B} & B_0 \\ \downarrow{\scriptstyle \gamma_1^i} & & \downarrow{\scriptstyle \gamma_0^i} \\ B'_1 & \xrightarrow{\partial_{B'}} & B'_0 \end{array}, \tag{4.11}$$

where γ_1^i connects the i-th column of $\mathcal{A}$ to a distinct column of $\mathcal{C}$, and we allow these column-wise connections $\{\gamma_1^i\}$ to be different. See Figure 4.4 for an illustration.

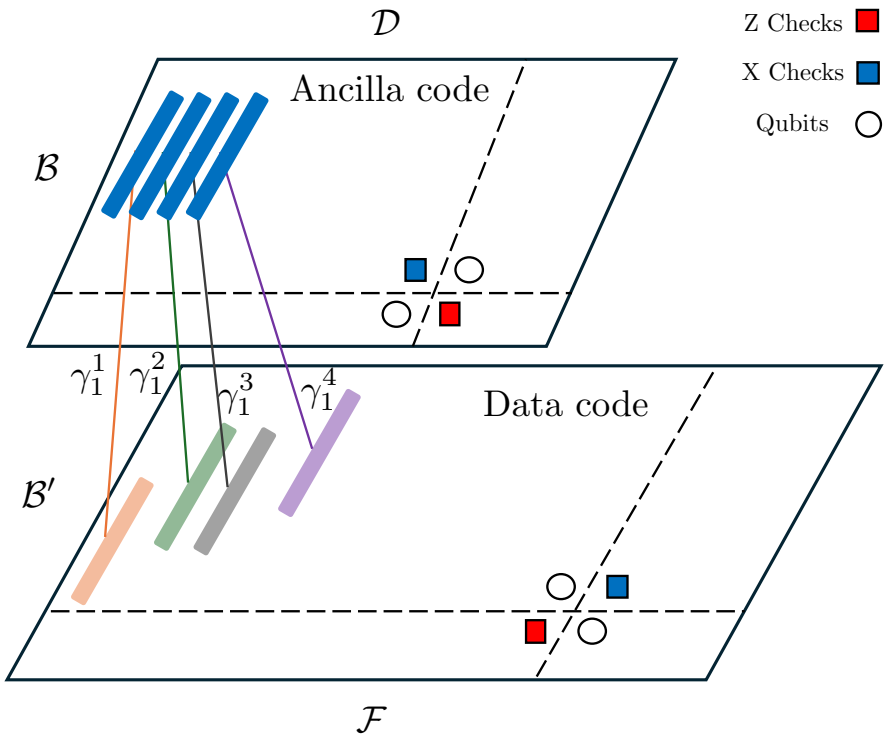


Figure 4.4: Illustration of high-rate surgery on a HGP data code using another HGP ancilla code. The logicals to measure on the data code are supported on different columns and are indicated the bars in varying colors. The blue bars in the ancilla code represent a set of ancilla checks, whose stabilizer values are combined to infer the corresponding logical values in the data code. According to Eq. (4.11), we can distance, column-wise connectivites $\{\gamma_1^i\}$, resulting in a large pattern of logical measurements on the data code.

Several known approaches to construct classical code homomorphisms include puncturing and augmentation [163] and code automorphisms [177, 178]. However, prior works utilize symmetries mostly through a direct lift from classical to quantum product codes. As a result, the available induced logical operations are limited by the strict product structure. For example, when $\mathcal{B}$ has an automorphism group $\mathrm{Aut}\,(\mathcal{B})$, directly lifting it to quantum-code automorphism [177, 178] induces an automorphism group of order $\mathcal{O}\left(|\mathrm{Aut}\,(\mathcal{B})|k_D!\right)$. In surgery, we can instead have the column-wise addressing in Eq. (4.11) by e.g., using distance automorphisms of $\mathcal{B}$ as $\{\gamma_1^i\}$, resulting in an exponential pattern of logical measurements $\mathcal{O}\left(|\mathrm{Aut}\,(\mathcal{B})|^{k_D}\right)$ and high logical addressibility. Such an addressibility is beyond the capabilities of the above-mentioned automorphism gates and the recently developed homomorphic measurements [163] for HGP codes.

We emphasize that algebraic scheme in Theorem 4.6 is general and applicable to other codes. It is particularly straightforward to extend to codes with a generalized product structure, such as LP code [179], fiber bundle code [180], and balanced product codes [181]. We leave the exploration of HGP ancilla tailored to these codes and their logical basis to future work.

4.3 Randomized construction for generic codes

In the last section we presented an algebraic ancilla construction with sufficient soundness, enabling application to structured quantum codes with a provable distance guarantee. In this section we go further: without any assumption on code structure, we can still perform high-rate surgeries in practice using a randomized construction for the ancilla system $\mathcal{A}$. The heart of the approach is a randomized algorithm with optimized heuristics, motivated by prior randomized constructions of low-rate, graph-based ancillas [7, 150]. At a high level, the construction proceeds as follows:

1. **Initialization:** Construct ancilla X-checks A_1 that transversally connect to the union of all generating logicals. Add low-degree qubits A_0 and low-degree Z-checks A_{-1} so as to measure the target logical space defined in Lemma 4.2.

2. **Boost expansion:** At each step, sample qubits connected to the X-checks under the constraint that the induced logical measurements are unchanged. This typically boosts soundness.

3. **Reduce check degrees:** Add either Z- or X-checks depending on the current sparsity of the parity checks.

4. **Iterate:** Repeat steps 2 and 3 until the distance is sufficient.

We remark that previous randomized graph algorithms do not extend straightforwardly to high-rate surgery. Key reasons include the lack of hypergraph analogs of decongestion and

cellulation [176]. We design an algorithm that boosts the dressed distance with convergence guarantees: it adaptively alternates between boosting expansion and reducing average degree. Another crucial ingredient is a local greedy selection rule to evaluate and choose randomly sampled qubits so as to maximize the expansion gain. For a full discussion, see Appendix C.5.

In practice, finite increases in connectivity matter. Surgery-based protocols typically incur a minimum increase of two in the check degree. In our simulations, we constrain the algorithm to the same benchmark. As a result, we restrict to transversally connected chain maps Γ_0 and Γ_1.

We next apply this randomized construction to the $[[144, 12, 12]]$ Gross code [11] and to new instances of spatially coupled (SC) codes [166]. Both codes possess strong symmetries that could be leveraged to build structured ancillas. Nevertheless, to demonstrate generality, we do not exploit any symmetry: we measure a randomly sampled set of logical operators with a randomly generated logical basis.

As shown in Fig. 4.3(a), using the Gross code as an example, low-rate surgery schemes—implemented sequentially [140] with a graph-based ancilla of rate $r_M = \tilde{\mathcal{O}}(1/d)$ over t logical cycles, or in parallel with t disjoint ancillas [9]—exhibit rapidly increasing space-time overhead as the number of logical measurements t grows. The total overhead can exceed $\approx 10\times$ the memory cost when t approaches the number of logical qubits $k = 12$. Moreover, in both sequential and parallel implementations, either the spatial or the temporal overhead becomes prohibitive. In contrast, the high-rate scheme with a randomized hypergraph ancilla performs all t measurements within a single logical cycle, using an ancilla no larger than the memory code (i.e., $|\mathcal{A}| \leq |\mathcal{C}|$). Consequently, the resulting space-time overhead remains within a factor of two of the memory cost, even for large t.

Table 4.2: Summary of spatially-coupled code parameters. The check degrees shown are the qubit degrees and check weights respectively. Distances are estimated through GAP [16] and should be taken as upper bounds.

Parameters	Rate	Degrees
$[[136, 34, 4]]$	25%	$(5, 8)$
$[[405, 101, 6]]$	24%	$(6, 9)$
$[[720, 164, 8]]$	22%	$(4, 6)$
$[[1125, 245, 10]]$	21%	$(6, 9)$

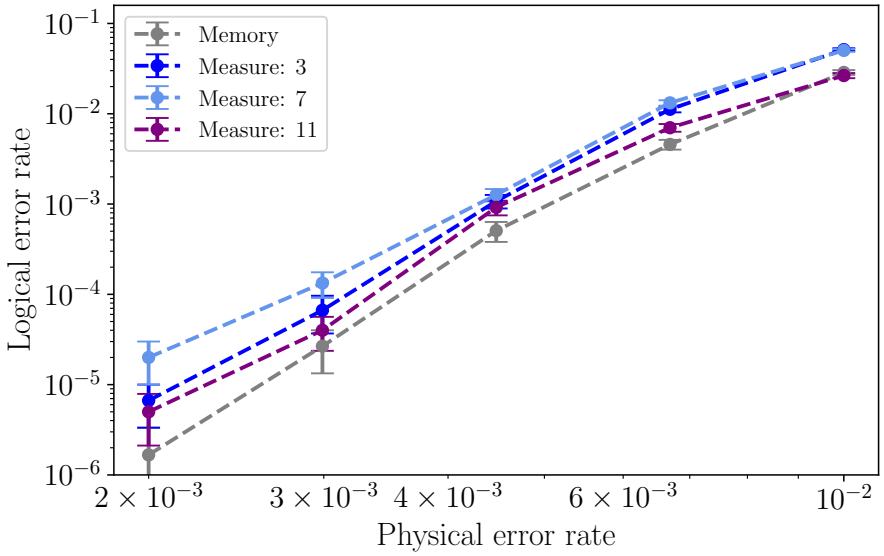


Figure 4.5: Memory logical error rate of the $[[144, 12, 12]]$ Gross code and its merge codes during surgery. Simulations are performed under standard circuit-level noise, excluding idling errors. The merge codes shown are constructed through the randomized, high-rate surgery scheme for measuring 3, 7, and 11 of its logicals under a randomly chosen logical basis. The logical error rates reported are averaged per logical qubit over 3 code cycles. BP-LSD decoder [12] with the same decoder parameter is used to decode all cases.

Next, to demonstrate constant overhead using high-rate surgeries with constant IER r_M, we consider a set of spatially coupled codes with similar encoding rates (between 20% and 25%) and increasing distances (from 4 to 10). See Tab. 4.2 for a summary of their code parameters and Appendix C.7 for their construction. As shown in Fig. 4.3(b), for a $[[n, k, d]]$ SC code the space-time cost relative to the memory overhead for measuring a randomly sampled set of $\lfloor 90\% \times k \rfloor$ logical operators increases very slowly and remains within a factor of 2 as d increases. In contrast, low-rate schemes have overhead that grows quickly with d,

consistent with their decreasing IER.

We highlight that for the largest SC code with 245 logical qubits, our scheme performs ≈ 200 randomly sampled logical measurements in parallel with roughly twice the memory overhead. Here, the memory overhead is already extremely low due to the high encoding rate. This demonstrates the potential of these codes, combined with our technique, for executing complex, addressable, large-scale logical computations with low overhead. Although these SC codes have extremely high rate, their logical operations and code structures remain largely unexplored. These results therefore highlight the power of our randomized construction: addressable, low-cost logical operations are generated automatically with no prior assumptions on the code.

We conjecture that the slight increase in overhead observed in Fig. 4.3(b) reflects limitations of our current numerical algorithms, which could be further improved to approach genuinely constant scaling; we leave this optimization to future work.

Finally, we simulate the logical error rate for the Gross code to show that the high-rate surgery scheme does not significantly impact data-code performance. In particular, we simulate the memory performance of the merged code and compare it with that of the data code, decoding both with the BP-LSD decoder [12] for consistency. Similar to prior work [140], we observe an increase in logical error rate, but the magnitude is acceptable. Moreover, the merged code often has significantly increased Z-distance, leading to orders-of-magnitude bias between Z and X errors. This presents opportunities for further optimization, such as a randomized distance-balancing algorithm.

4.4 Methods

4.4.1 Canonical basis

Here we give an overview of the canonical basis for completeness. We define a basis for a classical code codeword, $[n, k]$, if it possess the following form

$$G = \begin{pmatrix} I_k & Q \end{pmatrix} \tag{4.12}$$

$$H = \begin{pmatrix} Q^T & I_{n-k} \end{pmatrix} \tag{4.13}$$

up to column permutations. Here I_k is a $k \times k$ dimensional matrix. We say the bit support of the I_k submatrix are the information bits, whose union is the information set. Worth noting that the information bits are not unique. Similarly we can introduce a canonical basis form [24] for quantum codes

$$G_X = \begin{pmatrix} I_k & Q_X & 0 \end{pmatrix} \tag{4.14}$$

$$G_Z = \begin{pmatrix} I_k & 0 & Q_Z \end{pmatrix} \tag{4.15}$$

and we can similarly define the information (qu)bits. In general, the canonical basis is not a minimum weight basis.

4.4.2 Classical code homomorphisms

Consider a classical homomorphism as the following commuting diagram,

$$\begin{array}{ccc} B_1 & \xrightarrow{\partial^B} & B_0 \\ \downarrow{\scriptstyle \gamma_1} & & \downarrow{\scriptstyle \gamma_0} \\ B'_1 & \xrightarrow{\partial^{B'}} & B'_0 \end{array}, \tag{4.16}$$

The commuting diagram implies $\gamma_1 \partial^{B'} = \partial^B \gamma_1$, which induces a logical action, w, given by

$$G_B \gamma_1^T = w G_{B'} \tag{4.17}$$

w depends on the logical basis chosen. Here we given two specific examples.

1. Code modification: the modification technique here is referred to as logical checks in our work. It is similar to the puncture and augmentation techniques in classical coding theory. Consider the choice of $\mathcal{B}$ as

$$\partial^B = \begin{pmatrix} \partial^{B'} \\ H_L \end{pmatrix} \tag{4.18}$$

and chain maps

$$\gamma_1 = I_1, \gamma_0 = \begin{pmatrix} I_0 & 0 \end{pmatrix} \tag{4.19}$$

where $I_{0,1}$ and 0 should be interpreted as identity and zero block matrices with matching dimensions. The kernel space follows as $\ker \partial^B = \ker \partial^{B'} \cap \ker H_L$. Suppose $\mathcal{B}$ and $\mathcal{B}'$ are the ancilla and data codes respectively, H_L reduce the ancilla code kernel space to a subspace of the data code. Thus, it is effectively a logical check that filters unwanted logicals. For clarity, consider an exaple where weadopt the canonical basis as $G_{B'} = \begin{pmatrix} I_k & Q \end{pmatrix}$ and assume the logical checks only have support on the information set as $H_L = \begin{pmatrix} H_I & 0 \end{pmatrix}$. The kernel space follows as

$$G_B = \ker H_I \begin{pmatrix} I_k & Q \end{pmatrix}. \tag{4.20}$$

Thus, we tailor the kernel space by designing the kernel space of H_I. In the context of surgery, H_I correspond to the unencoded circuit's sparsity. When we assume it is a shallow circuit, H_I is necessarily LDPC. As a simple example, consider

$$G_{B'} = \begin{pmatrix} 1 & 0 & 0 & a_1^T \\ 0 & 1 & 0 & a_2^T \\ 0 & 0 & 1 & a_3^T \end{pmatrix}, H_I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \tag{4.21}$$

where one can understand the first and second lines of H_I realizes puncturing the first logical and coupling the second and third logicals respectively. As a result, we realize

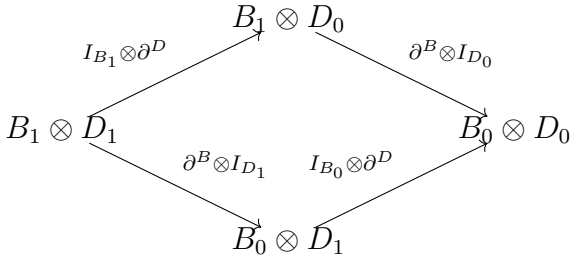
$$G_B = \begin{pmatrix} 0 & 1 & 1 & a_2^T + a_3^T \end{pmatrix} \tag{4.22}$$

2. Code automorphisms: automorphism refer to the special case where $\mathcal{B}' = \mathcal{B}$ and is often a result of the code's symmetry. The automorphism group, $\mathrm{Aut}(\mathcal{B}')$, consists of γ_1 such that there exists γ_0 where the diagram commutes. The automorphism group can be generated by permutations and the induced logical actions are affine, $w \in \mathrm{GL}_{k_B}(\mathbb{F}_2)$. The restriction to affine actions makes it not universal for classical codes.

We highlight that the commuting diagram can be superposed when $\mathcal{B}, \mathcal{B}'$ are the same through $\gamma_1 = \gamma_1^a + \gamma_1^b, \gamma_0 = \gamma_0^a + \gamma_0^b$, which is different from group compositions. This can, e.g., generate logical actions beyond affine ones with superposed automorphism gadgets. In addition, any code homomophism $\mathcal{B} \to \mathcal{B}'$ can be composed with any automorphisms $\mathcal{B}' \to \mathcal{B}'$. For example, one can combine logical checks with translational automorphisms, as we have demonstrated in the main text.

4.4.3 Hypergraph product code (HGP)

For completeness, we provide the definition of $\mathrm{HGP}(\mathcal{B}, \mathcal{D})$. As length-2 chain complex, they can be written as



(4.23)

Through Kunneth theorem, it has parameters

$$
\begin{aligned}
n &= n_{B_1} n_{D_0} + n_{B_0} n_{D_1} && (4.24)\\
k &= k_B^T k_D + k_B k_D^T && (4.25)\\
d_X &= \min\left(d_B, d_D\right) && (4.26)\\
d_Z &= \min\left(d_B^T, d_D^T\right). && (4.27)
\end{aligned}
$$

Note that our definition is different from some definitions in literature by a code transpose.

4.4.4 Surgery procedure

Consider a general surgery scheme with the following diagram

$$
\begin{array}{ccccc}
A_1 & \xrightarrow{\partial_1^A} & A_0 & \xrightarrow{\partial_0^A} & A_{-1} \\
 & \searrow^{\Gamma_1} & & \searrow^{\Gamma_0} & \\
C_2 & \xrightarrow{\partial_2^C} & C_1 & \xrightarrow{\partial_1^C} & C_0
\end{array}
\tag{4.28}
$$

One begins with a data code $\mathcal{C}$, and we follow the convention of X-Q-Z vector spaces in the main text. The conventional surgery procedure consists of the follow steps:

1. Initialize qubits in A_0 in $|0\rangle^{\otimes |A_0|}$.

2. Measure X-checks in A_1, C_2 and Z-checks in C_0, A_{-1}. Repeat for d rounds.

3. Transversally measure qubits in A_0 in Z-basis. Apply Pauli frame correction to track the Z-stabilizer values in C_0.

The logical information is measured and can be inferred through step 2. Since the qubits are initialized in the Z-basis, X-stabilizers in A_1 are not random if and only if they do not have support on A_0. For the details on the Pauli correction see Ref. [7].

4.4.5 Generalizations from graph to hypergraph subroutines

When the ancilla system is formed through disjoint patches of $\dim\ker\partial_1^A = 1$, each patch can be understood as a graph. Naturally, the generalization to high-rate ancilla translates to generalization from graphs to hypergraphs. Such a translation has two main difficulties: Z logicals of the ancilla system, G_Z^A, might not admit a low-weight generating basis, and the construction of LDPC commuting diagram such that the correct logicals are measured, $\Gamma_1\left(\ker\partial_1^A\right) = \mathrm{Span}\left\{L_M\right\}$, where L_M is a generating basis for the logicals to measure. In the following we address each of these problems.

In hypergraphs, many of the graph subroutines are no longer valid. The most notable ones being cellulation and decongestion, both of which are key components in weight-reducing G_Z^A, which is the graph's cycle basis. One observation is that when data code has Z-metachecks, they often serve as a reverse-directioned surgery that effectively measures out (part of) G_Z^A. While this does not provide useful information, it reduces the number of gauge logicals in the merge code. A similar observation was initially made in Ref. [150] for HGP codes. In rare occasions where some large G_Z^A remain unmeasured, we accept them as gauge logicals instead of attempting to cellulate and decongest. Our choice is motivated by heuristics: while cellulation and decongestion algorithms can be naively generalized to hypergraphs, they no longer have a guarantee on the resulting code's degree. In fact, we find it challenging to suppress both X and Z check degrees for decently large codes, especially when we want to limit the increase in code degree to 2. In cases where the merge code is a subsystem code, we observe a modest increase in the overhead.

Another challenge is posed by measuring the right logicals, i.e. the measured operator space $\mathrm{Span}\left\{\Gamma_1\left(\ker\partial_1^A\right)\right\}$ can contain unwanted logicals. For low-rate surgery this is initially dealt with by assuming the logical to be irreducible [8, 150]. In later works, the ancilla system can be constructed by randomly adding edges to the graph, which effectively eliminates all other kernel dimensions except the target logical. However, these techniques and assumptions are only valid for low-rate surgery. There is no universal guarantee that for arbitrary logical

basis, one can measure exactly an arbitrary logical set through a LDPC choice of Γ_1 and $\ker \partial_1^A$.

For simplicitly, we consider $\Gamma_{0,1}$ to be of transversal-type, but our conclusion is valid for any connections that are LDPC. Following the discussion on logical checks, a reasonable assumption is that the logicals to measure admits a LDPC logical check for the unencoded circuit to be of depth-$\mathcal{O}(1)$. Depending on the choice of logical basis and construction of commuting diagram, we discuss case-by-case as following:

- The basis is canonical, and the classical code homomorphism is constructed by puncturing. Here puncturing refers to taking the largest union necessary. In general codes this means $B_1' = B_1$. In specific codes such as HGP code B_1' can only be the size of a column because of the product structure. In this case we are guaranteed to have both a LDPC homomorphism that is measures correctly and a stabilizer code with no gauge logicals.

- The basis is canonical, but the classical code homomorphism is not punctured. For example, the homomorphism can be obtained through taking the union. In this case we still have LDPC guarantee on measuring the correct logicals, but no longer the stabilizer code guarantee.

- The basis is not canonical. We don't have a guarantee in this scenario. However, in practice we see that for decent-sized codes and sampled joint logicals to measure, one finds logical checks that are low in degree in most cases. This is only violated in rare instances. We want to point out that in simulations we never encounter large-degree logical checks if a random basis is chosen. This can be understood intuitively since these basis have relatively large weights compared to d, which means that they have more qubits per logical that one can place logical checks on such that the overall weight is reduced.

4.4.6 Randomized algorithm for ancilla construction

In the last section we discussed from a high-level why a randomized algorithm can still retain degree for high-rate surgery. Here, we give a more detailed procedure description. For the full protocol, see Appendix C.5. For a general quantum code and a randomized set of logical operators L_M, we perform the following steps

1. Take the union of the support of L_M, i.e. choose a set of qubits $Q \subseteq C_1$ such that for every logical operator $l_i \in L_M$ we have $l_i \subseteq Q$. Moreover, for every qubit in Q, it is in the support of at least one logical operator in L_M.

2. Construct Z-checks in B_1 that are transversally connected to Q. Construct qubits in B_0 that are transversally connected to $\text{Supp}\left\{\left(H_Z|_Q\right)^T\right\}$, i.e. we first restrict the Z-checks to Q then delete all checks that have empty rows.

3. Optimize for low-weighted logical checks, which corresponds to adding qubits in B_0 that are connected to B_1. The optimization is done through a distance estimation algorithm in GAP [16]. The check after this step is denoted as ∂''_B

4. Verify if the Z-logicals of $\mathcal{A}$ are in the stabilizer group. For those that are not, implement if one can find a low-weight basis.

5. Sample low-weight (dressed) logicals of the merged code that are below the target distance, d. Sample valid ancilla qubits that do not modify $\ker \partial_1^A$. A qubit is valid if its connections are low-weight. Assign a score to each qubit based on how many low-weight X-logicals they anticommute with. Implement the one with the largest score.

6. If there is no valid qubits with non-zero score, extend the ancilla by one layer by adding a full copy of the checks in B_1 and connect them to the previous layer through transversal qubits.

7. Repeat the above procedure until target distance achieved.

We present a more detailed analysis of the algorithm and its convergence in Appendix C.5.

4.4.7 Measurement of mixed-type Paulis

For simplicity, we have restricted to measurement of logicals of single Pauli type in a CSS code. Now, we comment on how it generalizes to general logical measurements. Suppose we consider CSS codes and measurement of mixed types of logicals. One can always write the logical measurement set as a commuting set of logical Paulis. However, it is not generally guaranteed that all composite X-logicals and Z-logical commute with each other. This poses a challenge if we construct the X and Z ancilla system separately since we would be measuring operators that anticommute with the stabilizer group, which goes beyond our framework. A remedy is to employ the canonical logical basis [9, 24] and generalize the framework in Lemma 4.2 to mixed-type Pauli connectivities. Another remedy is to consider twist-free surgery and Y resource states, similar to Refs. [9, 10]. See Appendix C for an extended discussion.

4.5 Discussion

In this paper, we provide a new framework and concrete constructions for high-rate code surgeries. By formalizing the logical action in a homological framework and giving a sufficient condition for distance preservation, we establish a general recipe for constructing ancilla patches that exploit the dense logical encoding of quantum LDPC codes. The framework motivates explicit ancilla constructions that achieve provably constant overhead when suitable structure is present in the data code. We further present randomized algorithms for finite-size codes that are agnostic to code structure. Remarkably, the computational cost is typically within a factor of two of the memory cost. Taken together, these results demonstrate the potential of high-rate surgery by showing its low space-time cost both asymptotically and for practical code instances.

Several promising research directions follow. As noted, surgery schemes are essentially a weight-reduction method in which one or more logical operators are the targets. Analogous

to low-rate surgery, prior approaches to weight-reducing parity checks [160–162] apply the same procedure iteratively to all checks above a weight limit. A natural extension of our work is a parallel weight reduction of stabilizers, which we expect to yield resource savings similar to those observed here.

Another direction is to extend the framework to single-shot quantum error correction and related work [146, 167–171], potentially further reducing temporal overhead. Exploring optimized ancilla designs based on other good code families—including bivariate-bicycle codes and quantum expander codes—may lower constant factors and relax the need for product-structure symmetries. Finally, our randomized algorithm is a proof-of-principle demonstration; improved heuristics or combinatorial optimization should further reduce resource costs, especially for large codes.

CHAPTER 5

CONCLUSION AND OUTLOOK

This thesis investigated how quantum information can be protected efficiently, from memory to logical computation. We developed a general way to benchmark the protection offered by an encoding under noise, used it to establish the performance and encoding limits of oscillator-based GKP memories, and constructed logical operations that retain the memory efficiency of high-rate qLDPC codes. Together, these results clarify where existing protection schemes leave room for improvement and how code design and recovery or gate operations can reduce the cost of preserving and processing quantum information.

I expect the near-optimal fidelity to be useful more broadly for qubit codes under noise models that are poorly captured by standard Pauli-error descriptions. A first goal is to identify codes and channels for which noise-adapted recovery offers an advantage, including codes with non-Pauli Clifford stabilizers and other encodings whose decoding is not yet well understood. Recent constructions of non-Abelian qLDPC codes provide examples of such Clifford-stabilizer settings [182, 183]. Pauli twirling and Pauli-frame updates are convenient tools, but replacing a channel by its Pauli approximation can discard structure that a physical recovery could exploit. The bosonic results in Chapter 3 motivate testing whether similar gains are possible for qubit codes under their actual noise channels.

Scaling these calculations remains a challenge. The near-optimal fidelity avoids an optimization over recovery maps, but the matrices required to evaluate it can still grow rapidly with code size. Symmetries of the code and noise may reduce that cost, and analytical calculations may make selected families accessible. Related progress has been made using coherent information: small-code calculations give accurate estimates of error thresholds [184], and exact analytical expressions for decohered toric codes connect the information-theoretic threshold to a statistical-mechanical transition [185]. Whether comparable reductions can be obtained for the near-optimal fidelity is an open direction. For unconventional codes that lack

a good decoder, the Petz recovery provides an explicit, provably near-optimal benchmark even when its implementation is costly. Understanding which parts of that recovery are responsible for its performance could guide the construction of practical decoders and recovery circuits.

For qLDPC computation, an important missing piece is efficient generation and use of magic states. High-rate surgery supplies flexible logical Pauli measurements, but the cost of a universal computation also depends on preparing non-Clifford resources, delivering them to the required logical qubits, and consuming them at the required rate. Recent work on gauging logical Clifford operators provides routes to magic-state preparation in qLDPC codes [182, 183]. Magic-state cultivation offers another approach to reducing preparation cost [186]. The more conventional route uses distillation-based magic-state factories [159, 187]. These results leave substantial design space in the choice of code, preparation protocol, and interface to the computational memory.

More generally, I expect further overhead reductions to require closer attention to quantum algorithms. General schemes that address arbitrary logical gates with few assumptions on the code are useful, but obtaining additional savings in this setting is increasingly difficult. Assumptions about code structure can restrict a scheme's applicability; they can also provide the operations that an algorithm needs at much lower cost. Chapter 4 illustrates how matched code structure and surgery constructions enable efficient parallel operations. Related work on homological and lifted product codes develops this connection further [163, 188]. Algorithm–code co-design asks which logical operations dominate a computation and how to choose the encoding, logical basis, and fault-tolerant scheme around them. A concrete algorithm supplies a useful measure of success: the physical systems and time required to reach a specified output accuracy.

Co-design may also include the physical device. Available connectivity, native gates, measurement speed, and the noise channel can change which code and logical scheme are economical. A resource estimate that incorporates these capabilities can test whether an abstract reduction in overhead survives implementation.

APPENDIX A

APPENDICES FOR CHAPTER 2

A.1 Convex optimization for recovery

The problem setup for channel reversal is that we want to revert a general channel $\mathcal{A}: \mathcal{B}\left(\mathcal{H}_{d_L}\right) \to \mathcal{B}\left(\mathcal{H}_n\right)$ with a physical recovery, $\mathcal{R}: \mathcal{B}\left(\mathcal{H}_n\right) \to \mathcal{B}\left(\mathcal{H}_{d_L}\right)$. Then, we would want to minimize the distance between the composed channel, $\mathcal{R} \circ \mathcal{A}$, and the identity channel. One of the distance metrics is the channel fidelity. For a quantum channel $\mathcal{Q}: \mathcal{B}\left(\mathcal{H}_{d_L}\right) \to \mathcal{B}\left(\mathcal{H}_{d_L}\right)$ with an operator sum representation, $\mathcal{Q}(\rho) = \sum_i Q_i \rho Q_i^\dagger$, the entanglement fidelity of Eq. (2.13) can also be written as

$$F(\mathcal{Q}) = \frac{1}{d_L^2} \sum_i |\operatorname{Tr} \hat{Q}_i|^2 = \frac{1}{d_L^2} \operatorname{Tr} X_{\mathcal{Q}}. \tag{A.1}$$

Here $X_{\mathcal{Q}}$ is the superoperator representation [21, 189]. If $\mathcal{Q} = \mathcal{R} \circ \mathcal{A}$, a nice property of the channel fidelity is that it can be expressed as a linear function of the Choi matrix form of $\mathcal{R}$:

$$F(\mathcal{R} \circ \mathcal{A}) = \frac{1}{d_L^2} \operatorname{Tr}\left(X_{\mathcal{R}} X_{\mathcal{A}}\right) = \frac{1}{d_L^2} \operatorname{Tr}\left(C_{\mathcal{R}} \tilde{C}_{\mathcal{A}}\right) \tag{A.2}$$

where the Choi matrix $C_{\mathcal{M}}$ of any channel $\mathcal{M}$ is defined to have an index form of $(C_{\mathcal{M}})_{[\mu\rho],[\nu\sigma]} = \langle\rho| \mathcal{M}\left(|\mu_L\rangle \langle\nu_L|\right) |\sigma\rangle$. We also define the linear map $\left(\tilde{C}_{\mathcal{M}}\right)_{[\mu\rho],[\nu\sigma]} = (C_{\mathcal{M}})_{[\sigma\nu],[\rho\mu]}$ to make the linear dependence of F on $C_{\mathcal{R}}$ more apparent. Moreover, the constraints on the recovery channel being physical, i.e. a completely positive trace-preserving map, can also be expressed through linear functions of $C_{\mathcal{R}}$. As a result, we can formulate the channel reversal problem as a convex optimization problem. The optimal recovery $\mathcal{R}^{\text{opt}}$ and its performance can be found through

$$\max_{C_{\mathcal{R}}} \quad \operatorname{Tr}\left\{C_{\mathcal{R}} \tilde{C}_{\mathcal{A}}\right\}, \tag{A.3}$$

$$\text{s.t.} \quad C_{\mathcal{R}} \succeq 0, \operatorname{Tr}_{\mathcal{H}_n} C_{\mathcal{R}} = I_{d_L} \tag{A.4}$$

which is a semidefinite program (SDP). Here, the partial trace notation $\mathrm{Tr}_{\mathcal{H}_n}$ denotes tracing out the indices of the input system, and I_d is a $d_L \times d_L$ identity matrix. The first condition ensures the recovery to be a completely positive map, while the second represents the constraint of trace-preserving maps.

In the context of quantum error correction (QEC), the target channel to be reverted is composed of the noise channel and the encoding channel, $\mathcal{A} = \mathcal{N} \circ \mathcal{E}$. The scenario considered by the original Knill-Laflamme conditions is the special case of pure state encodings, where $\mathcal{E}$ is an isometry. The optimization result specifies the optimal recovery and the optimal channel fidelity, $F^{\mathrm{opt}} = F\left(\mathcal{R}^{\mathrm{opt}} \circ \mathcal{N} \circ \mathcal{E}\right)$. Similarly, we can utilize the permutation invariant property of the trace and similarly optimize the encoding. Since we optimize for the Choi matrix of the recovery and/or encoding, the computational complexity is a high-order polynomial or even exponential function of the system size.

Worth noticing, the Fock state basis may not be the most efficient basis for representing the Choi matrix and, more generally, the code after noise. We can learn from the transpose channel form and devise better forms to drastically reduce the complexity required. Some recent works on antidistinguishability [100] and quantum state discrimination [99] adopted similar ideas to improve their SDP forms, and the improvement can be also extended to QEC.

A.2 Transpose channel

The transpose channel (TC) [73, 74] is a member of the near-optimal channel family. It has close connections with pretty good measurements [190–192] and has applications in quantum information theory [193], quantum thermodynamics [194–196], and many others. There is also active effort in devising implementation schemes. There have been efforts concerning the systematic implementation of general CPTP maps in both bosonic [197] and qubit systems [198]. Fault-tolerant error detection has been demonstrated in a bosonic encoding [199], and path-independent gates make the intended ancilla-mediated operation insensitive to specified correctable ancilla errors [138]. Focusing on the transpose channel,

Ref. [85] proposed an algorithm to implement a Petz channel based on quantum singular value transformation, block encoding, and a set of unitary gates. Moreover, Ref. [200] proposed and compared several implementation schemes of the transpose channel on practical devices.

For a given logical code space projector $\hat{P}_L$ and noise channel $\mathcal{N} \sim \left\{\hat{N}_i\right\}$, the transpose channel admits a Kraus operator representation of $\mathcal{R}^{\mathrm{TC}} \sim \left\{\hat{R}_i^{\mathrm{TC}}\right\}$, where

$$\hat{R}_i^{\mathrm{TC}} := \hat{P}_L \hat{N}_i^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} \tag{A.5}$$

for $i = 0, \ldots, L-1$. Here, we define $\hat{P}_L$ to be the projector onto the logical code space. The inverse of $\mathcal{N}\left(\hat{P}\right)$ is defined on its support, which is spanned by the error subspaces.

The transpose channel has a critical property: its fidelity gives a two-sided bound of the optimal recovery fidelity. While there exist other recovery channels with similar two-sided bounds, such as the quadratic recovery [77], the near-optimal expression we obtained from the transpose channel is the most concise and informative by far. For completeness, we provide a proof for the near-optimality.

Lemma A.1 (Near-optimality of transpose channel; [73])**.** *For any QEC code with encoding channel $\mathcal{E}$ and noise channel $\mathcal{N}$, the optimal channel fidelity $F^{\mathrm{opt}} = F\left(\mathcal{R}^{\mathrm{opt}} \circ \mathcal{N} \circ \mathcal{E}\right)$ is bounded by the channel fidelity of the transpose channel $F\left(\mathcal{R}^{\mathrm{TC}} \circ \mathcal{N} \circ \mathcal{E}\right)$ through*

$$\frac{1}{2}\left(1 - F(\mathcal{R}^{TC} \circ \mathcal{N} \circ \mathcal{E})\right) \leq 1 - F^{opt} \leq 1 - F(\mathcal{R}^{TC} \circ \mathcal{N} \circ \mathcal{E}). \tag{A.6}$$

Proof. The second inequality holds with definition of optimal recovery. To prove the first inequality, notice that the recovery should have its domain on $\mathrm{supp}(\mathcal{N}\hat{\rho})$ and its range on $\mathrm{supp}(\hat{\rho})$. The optimal recovery channel $\mathcal{R}^{\mathrm{opt}} \sim \{R_i^{\mathrm{opt}}\}$ can be represented as

$$\hat{R}_i^{\mathrm{opt}} = \hat{P}_L \hat{B}_i^\dagger \mathcal{N}(\hat{P}_L)^{-1/2}. \tag{A.7}$$

where we can choose $\hat{B}_i^\dagger = \hat{P}_L \hat{R}_i^{\text{opt}} \mathcal{N}(\hat{P}_L)^{1/2}$. The channel fidelity can be written as

$$F(\mathcal{R}^{\text{opt}} \circ \mathcal{N} \circ \mathcal{E}) = \frac{1}{d_L^2} \sum_{i,j} \left| \text{Tr}\left(\hat{R}_i^{\text{opt}} \hat{N}_j\right) \right|^2 \tag{A.8}$$

$$= \frac{1}{d_L^2} \sum_{i,j} \left| \text{Tr}\left(\hat{P}_L \hat{B}_i^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} \hat{N}_j\right) \right|^2 \tag{A.9}$$

$$= \frac{1}{d_L^2} \sum_{i,j} \left| \text{Tr}\left(\hat{P}_L \hat{B}_i^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} \hat{N}_j \hat{P}_L\right) \right|^2 \tag{A.10}$$

$$\tag{A.11}$$

By appropriate choice of Kraus operators $\hat{B}_i$ and $\hat{N}_j$, the inner sum can be diagonalized such that the trace vanishes for $i \neq j$, which is equivalent to performing singular value

decomposition. Then, we have

$$
\begin{aligned}
&F(\mathcal{R}^{\mathrm{opt}} \circ \mathcal{N} \circ \mathcal{E}) && \text{(A.12)}\\
&= \frac{1}{d_L^2} \sum_i \left| \mathrm{Tr}\left(\hat{P}_L \hat{B}_i^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} \hat{N}_i \hat{P}_L \right) \right|^2 && \text{(A.13)}\\
&= \frac{1}{d_L^2} \sum_i \left| \mathrm{Tr}\left((\hat{P}_L \hat{B}_i^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}}) (\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}} \hat{N}_i \hat{P}_L) \right) \right|^2 && \text{(A.14)}\\
&\leq \frac{1}{d_L^2} \sum_{i,j} \left| \mathrm{Tr}\left((\hat{P}_L \hat{B}_i^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}}) (\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}} \hat{B}_i \hat{P}_L) \right) \right| \\
&\quad \times \left| \mathrm{Tr}\left((\hat{P}_L \hat{N}_j^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}}) (\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}} \hat{N}_j \hat{P}_L) \right) \right| && \text{(A.15)}\\
&\leq \frac{1}{d_L^2} \left(\sum_i \left| \mathrm{Tr}\left((\hat{P}_L \hat{B}_i^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}}) (\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}} \hat{B}_i \hat{P}_L) \right) \right|^2 \right)^{1/2} \\
&\quad \times \left(\sum_j \left| \mathrm{Tr}\left((\hat{P}_L \hat{N}_j^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}}) (\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{4}} \hat{N}_j \hat{P}_L) \right) \right|^2 \right)^{1/2} && \text{(A.16)}\\
&\leq \frac{1}{d_L^2} \left(\sum_{ik} \left| \mathrm{Tr}\left(\hat{P}_L \hat{B}_i^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} \hat{B}_k \hat{P}_L \right) \right|^2 \right)^{1/2} \\
&\quad \times \left(\sum_{jl} \left| \mathrm{Tr}\left(\hat{P}_L \hat{N}_j^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} \hat{N}_l \hat{P}_L \right) \right|^2 \right)^{1/2} && \text{(A.17)}
\end{aligned}
$$

where the second term inside square root can be simplified as

$$
\frac{1}{d_L^2} \sum_{jl} \left| \mathrm{Tr}\left(\hat{P}_L \hat{N}_j^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} \hat{N}_l \hat{P}_L \right) \right|^2 = F\left(\mathcal{R}^{\mathrm{TC}} \circ \mathcal{N} \circ \mathcal{E}\right). \quad \text{(A.18)}
$$

The first term can be upper bounded as

$$
\frac{1}{d_L^2} \sum_{ik} \left| \mathrm{Tr}\left(\hat{P}_L \hat{B}_i^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} \hat{B}_k \hat{P}_L \right) \right|^2 = \frac{1}{d_L^2} \sum_{ik} \left| \mathrm{Tr}\left(\hat{R}_i^{\mathrm{opt}} \hat{B}_k \right) \right|^2 = F\left(\mathcal{R}^{\mathrm{opt}} \circ \mathcal{B}\right) \leq 1. \quad \text{(A.19)}
$$

The proof of the upper bound is from the trace-preserving constraint on the channel $\mathcal{R}^{\mathrm{opt}}$, such

that $\mathcal{B} \sim \left\{\hat{B}_i\right\}$ acts as a CPTP map on $\hat{P}_L$. See Ref. [73] for a more detailed proof. Substituting Eq. (A.18) and Eq. (A.19) into Eq. (A.17), we have $\left(F^{\text{opt}}\right)^2 \leq F\left(\mathcal{R}^{\text{TC}} \circ \mathcal{N} \circ \mathcal{E}\right) \leq F^{\text{opt}}$. As a straightforward corollary, we have

$$\frac{1}{2}\left(1 - F(\mathcal{R}^{\text{TC}} \circ \mathcal{N} \circ \mathcal{E})\right) \leq 1 - F^{\text{opt}} \leq 1 - F(\mathcal{R}^{\text{TC}} \circ \mathcal{N} \circ \mathcal{E}). \tag{A.20}$$

and conclude the proof. □

Since the goal is to find and benchmark good codes, we are mostly only interested in codes that perform well against the noise channels of interest. In these scenarios, the bound given by the TC channel is quite a tight bound on the optimal recovery. Therefore, the transpose channel provides a constructive channel that solves the slightly relaxed convex optimization problem. To numerically demonstrate the two-sided bound, we can look at some example qubit codes in Fig. A.1. It is clear that while the two-sided bounds are narrow, they nicely bound the optimal code performances.

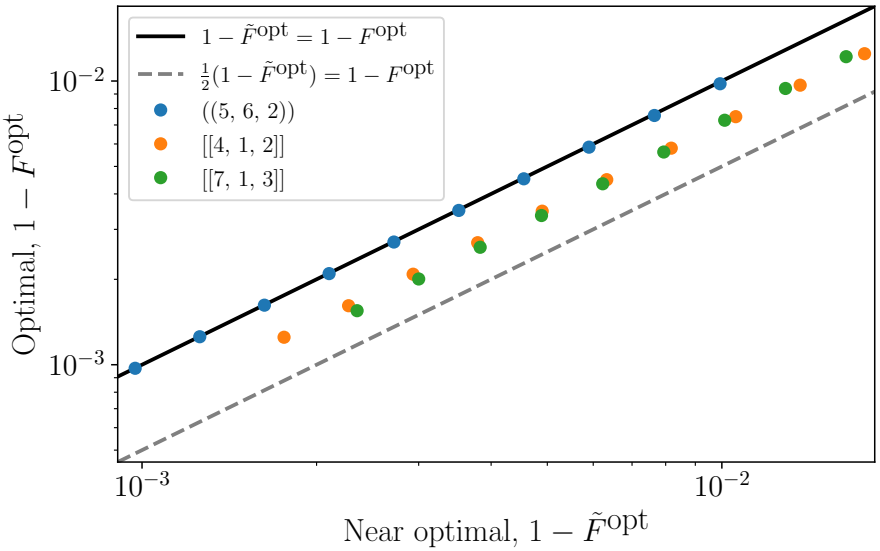


Figure A.1: Comparison of near-optimal and optimal infidelity for qubit codes [[4, 1, 2]], ((5, 6, 2)), and [[7, 1, 3]] under amplitude damping noise. The black solid line (gray dashed line) represents the upper (lower) bound on the optimal infidelity given by the near-optimal infidelity

As a high-level understanding, the TC channel is a generalized version of the QEC recovery

procedure from KL conditions: $\mathcal{A}\left(\hat{P}\right)^{-\frac{1}{2}}$ performs measurements that project the state into carefully chosen subspaces, and $\hat{P}\hat{N}_i^\dagger$ rotates them back into the corresponding logical space. More rigorously, we can derive the following lemma.

Lemma A.2. *Given noise channel* $\mathcal{N} \sim \left\{\hat{N}_i\right\}$ *and logical codewords* $|\mu_L\rangle$*, the form of the transpose channel Kraus operators is equivalent to*

$$\hat{R}_l^{TC} = \sum_{\mu,\nu,k} \left(M^{-\frac{1}{2}}\right)_{[\mu l],[\nu k]} |\mu_L\rangle \langle\nu_L| \hat{N}_k^\dagger \tag{A.21}$$

where the inverse represents matrix pseudoinverse and the QEC matrix $M_{[\mu l],[\nu k]} := \left\langle \mu_L \left| \hat{N}_l^\dagger \hat{N}_k \right| \nu_L \right\rangle$.

Proof. The error subspace basis is defined by the set $\left\{\hat{N}_l |\mu_L\rangle\right\}$. It is worth noting that this basis set is neither necessarily linear independent nor necessarily complete on the Hilbert space. However, it provides us the sufficient basis to express $\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}}$ and consequently $\hat{R}_l^{\mathrm{TC}}$. To start with, we can write a projection operator on this subspace

$$\hat{P}_N = \sum_{l,k,\mu,\nu} \left(M^{-1}\right)_{[\mu l],[\nu k]} \hat{N}_l|\mu\rangle \langle\nu| \hat{N}_k^\dagger, \tag{A.22}$$

where M^{-1} is the pseudoinverse of M. The projector expression can be verified through checking its effect in the error subspace $\hat{P}_N \hat{N}_k|\nu_L\rangle = \hat{N}_k|\nu_L\rangle$.

Notice that $\mathcal{N}\left(\hat{P}_L\right) = \sum_{l,\mu} \hat{N}_l |\mu_L\rangle \langle\mu_L| \hat{N}_l^\dagger$. We can then derive representation of $\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}}$ in the error subspace basis as

$$\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} = \sum_{l,k,\mu,\nu} \left(M^{-\frac{3}{2}}\right)_{[\mu l],[\nu k]} \hat{N}_l|\mu_L\rangle \langle\nu_L| \hat{N}_k^\dagger. \tag{A.23}$$

We can either derive step-by-step or straightforwardly verify from the definition of $\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}}$:

$$\mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} \mathcal{N}\left(\hat{P}_L\right) \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} = \hat{P}_N. \tag{A.24}$$

With the expression presented in Eq. (A.23), the transpose channel Kraus can be written in the form of

$$\begin{aligned}
\hat{R}_l^{\mathrm{TC}} &:= \hat{P}_L \hat{N}_l^\dagger \mathcal{N}\left(\hat{P}_L\right)^{-\frac{1}{2}} && \text{(A.25)}\\
&= \sum_\mu |\mu_L\rangle\langle\mu_L| \hat{N}_l^\dagger \sum_{p,k,\eta,\nu} \left(M^{-\frac{3}{2}}\right)_{[\eta p],[\nu k]} \hat{N}_p |\eta_L\rangle\langle\nu_L| \hat{N}_k^\dagger && \text{(A.26)}\\
&= \sum_\mu \sum_{p,k,\eta,\nu} \left(M^{-\frac{3}{2}}\right)_{[\eta p],[\nu k]} \langle\mu_L| \hat{N}_l^\dagger \hat{N}_p |\eta_L\rangle |\mu_L\rangle\langle\nu_L| \hat{N}_k^\dagger && \text{(A.27)}\\
&= \sum_\mu \sum_{p,k,\eta,\nu} \left(M^{-\frac{3}{2}}\right)_{[\eta p],[\nu k]} M_{[\mu l],[\eta p]} |\mu_L\rangle\langle\nu_L| \hat{N}_k^\dagger && \text{(A.28)}\\
&= \sum_{\mu,\nu,k} \left(M^{-\frac{1}{2}}\right)_{[\mu l],[\nu k]} |\mu_L\rangle\langle\nu_L| \hat{N}_k^\dagger. && \text{(A.29)}
\end{aligned}$$

□

Although we assume isometry encoding for simplicity, it is straightforward to extend our result to the context of general channel reversal problems. For example, for mixed-state encodings, we can combine the encoding channel with the noise channel and revert the composed channel. For other codes like subsystem codes, the transpose channel is also valid following, for example, Ref. [80].

While Eq. (A.21) is useful for our proof of the near-optimal fidelity expression, it is not so intuitive. We can arrive at a more intuitive expression by defining a set of new basis

$$|\psi_i\rangle := \sum_{[\nu k]} \left(M^{-\frac{1}{2}}\right)^*_{i,[\nu k]} \hat{N}_k |\nu_L\rangle, \tag{A.30}$$

which is not necessarily a set of orthonormal basis. In fact, since M is a Hermitian matrix, we can check that $\langle\phi_j|\phi_i\rangle = M^{-\frac{1}{2}}_{j,[\mu l]} \left(M^{-\frac{1}{2}}\right)^*_{i,[\nu k]} \langle\mu_L| \hat{N}_l^\dagger \hat{N}_k |\nu_L\rangle = \left(M^{-1/2} M M^{-1/2}\right)_{j,i}$, which only guarantees orthonormality of the basis if M is full rank. The transpose channel Kraus operators can be conveniently expressed as

$$\hat{R}_l^{\mathrm{TC}} = \sum_\mu |\mu_L\rangle\langle\psi_{[\mu l]}|. \tag{A.31}$$

This form of the transpose channel Kraus operators can be interpreted as following: for each logical codeword $|\mu_L\rangle$ and noise Kraus operator with label l, the physical operation of the transpose channel is to measure the projector $|\psi_{[\mu l]}\rangle\langle\psi_{[\mu l]}|$ and perform a rotation back to the logical codeword $|\mu_L\rangle$ accordingly.

A.3 Derivation of the near-optimal fidelity expression

With Lemma. A.2, we arrive at our main result, Theorem 2.4 in the main text.

Theorem. *For a d_L-dimensional encoding, $\mathcal{E}$, and noise channel $\mathcal{N}$, the near-optimal fidelity is defined as*

$$\tilde{F}^{opt} = \frac{1}{d_L^2} \left\| \mathrm{Tr}_L \sqrt{M} \right\|_F^2, \tag{A.32}$$

where M is the QEC matrix, $(\mathrm{Tr}_L B)_{l,k} = \sum_\mu B_{[\mu l],[\mu k]}$ denotes the partial trace over the code space indices, and $||\cdot||_F$ is the Frobenius norm. The near-optimal fidelity gives a two-sided bound on the optimal fidelity as

$$\frac{1}{2}\left(1 - \tilde{F}^{opt}\right) \leq 1 - F^{opt} \leq 1 - \tilde{F}^{opt}. \tag{A.33}$$

Proof. The near-optimal fidelity is defined as the fidelity achieved by the transpose channel

recovery, $\tilde{F}^{\text{opt}} := F\left(\mathcal{R}^{\text{TC}} \circ \mathcal{N} \circ \mathcal{E}\right)$. Substituting Lemma A.2 into Eq. (A.1), we have

$$\begin{aligned}
\tilde{F}^{\text{opt}} &= \frac{1}{d_L^2} \sum_{i,j} \left| \text{Tr}\left(\hat{R}_i^{\text{TC}} \hat{N}_j\right) \right|^2 && \text{(A.34)} \\
&= \frac{1}{d_L^2} \sum_{i,j} \left| \text{Tr}\left(\sum_{\mu,\nu,k} \left(M^{-\frac{1}{2}}\right)_{[\mu i],[\nu k]} |\mu_L\rangle \langle \nu_L| \hat{N}_k^\dagger \hat{N}_j \right) \right|^2 && \text{(A.35)} \\
&= \frac{1}{d_L^2} \sum_{i,j} \left| \text{Tr}\left(\sum_{\mu,\nu,k} \left(M^{-\frac{1}{2}}\right)_{[\mu i],[\nu k]} \langle \nu_L| \hat{N}_k^\dagger \hat{N}_j |\mu_L\rangle \right) \right|^2 && \text{(A.36)} \\
&= \frac{1}{d_L^2} \sum_{i,j} \left| \text{Tr}\left(\sum_{\mu,\nu,k} \left(M^{-\frac{1}{2}}\right)_{[\mu i],[\nu k]} M_{[\nu k],[\mu j]} \right) \right|^2 && \text{(A.37)} \\
&= \frac{1}{d_L^2} \sum_{i,j} \left| \sum_{\mu} \left(M^{\frac{1}{2}}\right)_{[\mu i],[\mu j]} \right|^2 && \text{(A.38)} \\
&= \frac{1}{d_L^2} \left\| \text{Tr}_L \sqrt{M} \right\|_F^2 && \text{(A.39)}
\end{aligned}$$

where we have used the cyclicity of trace and the definition of the Frobenius norm $||B||_F := \sum_{i,j} \left|B_{i,j}\right|^2$

□

In some cases, we might also be dealing with non-orthonormal codewords. As an example, finite-energy GKP codes generally have non-orthonormal codewords. When we want to obtain a valid near-optimal fidelity for the orthonormalized codewords, it is convenient to have the following corollary.

Corollary A.3. *For any non-orthonormal codeword with overlap matrix*

$$m_{\mu,\nu} := \langle \mu_L | \nu_L \rangle \,, \tag{A.40}$$

the near-optimal fidelity in Theorem 2.4 has the form of

$$\tilde{F}^{opt} = \frac{1}{d_L^2} \left\| \text{Tr}_L \sqrt{(m^{-1} \otimes I_l) M} \right\|_F^2 , \tag{A.41}$$

where I_l is a $l \times l$ identity matrix, and l is the number of Kraus operators of the noise channel $\mathcal{N}$.

Proof. Assume that with the, not necessarily orthonormal, original codewords $|\mu_L\rangle$, we obtain the QEC matrix M. Moreover, we can always find a matrix A, such that $AA^\dagger = m^{-1} \otimes I_l$. Consequently, $A^\dagger(m \otimes I_l)A = I$ and orthonormalizes the codewords. The QEC matrix obtained from the orthonormalized codewords, M', is related to the original QEC matrix via $M' = A^\dagger M A$. Therefore, based on Theorem 2.4, we have

$$\tilde{F}^{\text{opt}} = \frac{1}{d_L^2} \left\| \text{Tr}_L \sqrt{M'} \right\|_F^2 . \tag{A.42}$$

Because of the tensor product structure of $m \otimes I_l$, we can write $A = B \otimes I_l$ and still utilize the cyclicity of the partial trace:

$$\text{Tr}_L \sqrt{M'} = \text{Tr}_L A\sqrt{M'}A^{-1} = \text{Tr}_L \sqrt{AM'A^{-1}} = \text{Tr}_L \sqrt{AA^\dagger M} = \text{Tr}_L \sqrt{(m^{-1} \otimes I_l)M}, \tag{A.43}$$

where the second equality is a consequence of the identity $\left(A\sqrt{M'}A^{-1}\right)^2 = AM'A^{-1}$. Substituting Eq. (A.43) into Eq. (A.42), we conclude the proof.

□

A.4 Derivation of the perturbative form of the near-optimal fidelity and its properties

Here we attempt to prove Corollary. 2.9, which is a perturbative expansion of Theorem 2.4. Such a perturbative expansion can lead to fruitful analytical results, as we have demonstrated for the thermodynamic code and the GKP code in the main text.

Corollary. *With a suitable choice of unitary gauge, there exists a decomposition of the QEC matrix M such that $D = \frac{1}{d_L} \text{Tr}_L M$ is diagonal. With the residual matrix $\Delta M := M - I_L \otimes D$,*

the near-optimal infidelity has a perturbative expansion through

$$1-\tilde{F}^{opt}=\frac{1}{d_L}\|f(D)\odot\Delta M\|_F^2+\mathcal{O}\left(\frac{1}{d_L}\|f(D)\odot\Delta M\|_F^3\right) \tag{A.44}$$

where we define $f(D)_{[\mu l],[\nu k]}=\frac{1}{\sqrt{D_{ll}}+\sqrt{D_{kk}}}$ *and the Hadamard product to be* $(A\odot B)_{ij}=A_{ij}B_{ij}$.

Proof. Given the decomposition $M=I_L\otimes D+\Delta M$, where matrix D is diagonal and $D\succ 0$ (since we can truncate the M matrix and discard the channel subspaces which vanish, $D_{ii}=0$). If we treat ΔM as a perturbation, we can perturbatively expand the matrix square root such that

$$\begin{aligned}\sqrt{M} &= I_L\otimes\sqrt{D}+f(D)\odot\Delta M-f(D)\odot\left(f(D)\odot\Delta M\right)^2+\delta, &\text{(A.45)}\\ ||\delta||_F &\leq \mathcal{O}(\|f(D)\odot\Delta M\|_F^3), &\text{(A.46)}\end{aligned}$$

which is a second-order expansion from the Daleckii-Krein theorem [102, 103]. Notice that $\mathrm{Tr}D=1$ and $\mathrm{Tr}_L\Delta M=0$, then we have

$$\mathrm{Tr}_L\sqrt{M}=d_L\sqrt{D}-f(D)\odot\mathrm{Tr}_L\left(\left(f(D)\odot\Delta M\right)^2\right)+\mathrm{Tr}_L\delta. \tag{A.47}$$

Therefore, an expansion of the Frobenius norm lead to the final expression

$$\begin{aligned}1-\tilde{F}^{\mathrm{opt}} &= 1-\frac{1}{d_L^2}\left\|\mathrm{Tr}_L\sqrt{M}\right\|_F^2 &\text{(A.48)}\\ &= 1-\frac{1}{d_L^2}\mathrm{Tr}(\mathrm{Tr}_L\sqrt{M})^2 &\text{(A.49)}\\ &= 1-\mathrm{Tr}D+\mathrm{Tr}\left(\frac{1}{d_L}\mathrm{Tr}_L\left(f(D)\odot\Delta M\right)^2\right)+\mathcal{O}\left(\frac{1}{d_L}\|f(D)\odot\Delta M\|_F^3\right) &\text{(A.50)}\\ &= \frac{1}{d_L}\|f(D)\odot\Delta M\|_F^2+\mathcal{O}\left(\frac{1}{d_L}\|f(D)\odot\Delta M\|_F^3\right). &\text{(A.51)}\end{aligned}$$

□

As mentioned in the main text, we can choose to determine D not through performing a

unitary rotation but instead truncating the off-diagonal terms in $\frac{1}{d_L}\operatorname{Tr}_L M$. Such a choice can be more convenient when the analytical expression for the unitary rotation is unknown. In this scenario, we end up with $\operatorname{Tr}\Delta M \neq 0$ and $\operatorname{Tr} D \neq 1$. Physically, we are distributing some of the correctable parts into the uncorrectable off-diagonal part of the QEC matrix, thus overestimating the near-optimal channel infidelity. The perturbative form is modified to be

$$1-\tilde{F}^{\text{opt}} = \frac{1}{d_L}\|f(D)\odot\Delta M\|_F^2 - \frac{1}{d_L^2}\operatorname{Tr}\left(\operatorname{Tr}_L f(D)\odot\Delta M\right)^2 + \mathcal{O}\left(\frac{1}{d_L}\|f(D)\odot\Delta M\|_F^3\right). \tag{A.52}$$

Therefore, with such a choice of D, the resulting infidelity is a good approximation as long as $\left\|f(D)\odot\left(\frac{1}{d_L}\operatorname{Tr}_L M - D\right)\right\|_F^2 \ll \frac{1}{d_L}\|f(D)\odot\Delta M\|_F^2$. Moreover, if higher precision is required, it is certainly possible to have higher orders of perturbative expressions through the definition of Fretchet derivatives or through solving the iterative relations order-by-order.

For a more intuitive understanding of the perturbative form, we can rewrite it as

$$\begin{aligned} 1-F^{\text{opt}} = \frac{1}{d_L}\Bigg(&\sum_{[\mu l]\neq[\nu k]}\frac{p_{[\mu l]}p_{[\nu k]}}{\left(\sqrt{\bar{p}_l}+\sqrt{\bar{p}_k}\right)^2}\operatorname{Tr}\left\{\hat{P}_{[\mu l]}\hat{P}_{[\nu k]}\right\} \\ &+\sum_{[\mu l]}\frac{\left(p_{[\mu l]}-\bar{p}_l\right)^2}{4\bar{p}_l}\Bigg) \\ &+\mathcal{O}\left(\frac{1}{d_L}\|f(D)\odot\Delta M\|_F^3\right). \end{aligned} \tag{A.53}$$

Here, we define the error subspace projectors as $\hat{P}_{[\mu l]} := \frac{\hat{N}_l|\mu_L\rangle\langle\mu_L|\hat{N}_l^\dagger}{p_{[\mu l]}}$, where $p_{[\mu l]} := \langle\mu_L|\hat{N}_l^\dagger\hat{N}_l|\mu_L\rangle$ and represents the probability of measuring $|\mu,l\rangle$ if we initialize in $|\mu_L\rangle$. The probability of measuring noise Kraus operator l is $\bar{p}_l = \frac{1}{d_L}\sum_{\mu=0}^{d_L-1}p_{[\mu l]}$. In the expression, the second term contains the errors from different logical codewords reacting differently to the same noise Kraus operator, which leaks information to the environment. The first

term includes logical flip errors (reverting to a different logical state) and dephasing errors (reverting to the same logical state but with an incorrect phase).

Moreover, the perturbative form possesses a special property when it vanishes:

Corollary A.4. *As* $\frac{1}{d_L}\|f(D)\odot\Delta M\|_F^2\to 0$*, the near-optimal fidelity*

$$\tilde{F}^{opt}\to \mathrm{Tr}\{D\}. \tag{A.54}$$

Proof. Here, we define $\Delta := \frac{1}{d_L}\left|\left\|\mathrm{Tr}_L\sqrt{M}\right\|_F - \left\|\mathrm{Tr}_L\sqrt{I_\mu\otimes D}\right\|_F\right|$, and we show that Δ vanishes when the perturbative expression vanishes, i.e. $\frac{1}{d_L}\|f(D)\odot\Delta M\|_F^2\to 0$. First, notice some matrix identities

$$\begin{aligned}
\Delta &\leq \frac{1}{d_L}\left\|\mathrm{Tr}_L\sqrt{M}-\mathrm{Tr}_L\sqrt{I_L\otimes D}\right\|_F && \text{(A.55)}\\
&\leq \frac{1}{\sqrt{d_L}}\left\|\sqrt{M}-\sqrt{I_L\otimes D}\right\|_F && \text{(A.56)}\\
&\leq \left\|\sqrt{\frac{M}{d_L}}-\sqrt{\frac{I_L\otimes D}{d_L}}\right\|_F && \text{(A.57)}
\end{aligned}$$

The first inequality applies Von Neumann's trace inequality to obtain that for any two Hermitian matrices $A,B\succeq 0$, $\|A-B\|_F\geq\|A\|_F-\|B\|_F$. In the second inequality, we proceeded by decomposing any matrix $X=\sum_i Y_i\otimes\sigma_i$, where σ_i are the Gell-mann matrices, and show that $\|X\|_F\geq\frac{1}{\sqrt{d_L}}\|\mathrm{Tr}_L X\|_F$, where d is the logical dimension. Notice that from the definition of Fretchet derivatives and the Daleckii-Krein theorem,

$$\sqrt{M}=\sqrt{I_L\otimes D}+f(D)\odot\Delta M+o(\Delta M). \tag{A.58}$$

Since $D \succ 0$ and $\mathrm{Tr}\{D\} \leq 1$, $\left\| f(D) \odot \frac{\Delta M}{\sqrt{d_L}} \right\|_F \to 0$ implies that $\left\| \frac{\Delta M}{\sqrt{d_L}} \right\|_F \to 0$. Thus,

$$\begin{aligned} \lim_{\frac{\Delta M}{\sqrt{d_L}} \to 0} \Delta &\leq \lim_{\frac{\Delta M}{\sqrt{d_L}} \to 0} \left\| f(\frac{D}{d_L}) \odot \frac{\Delta M}{d_L} + o(\frac{\Delta M}{d_L}) \right\|_F && \text{(A.59)} \\ &= \left\| f(D) \odot \frac{\Delta M}{\sqrt{d_L}} \right\|_F && \text{(A.60)} \\ &= 0 && \text{(A.61)} \end{aligned}$$

where the last equality holds by assumption. Therefore, in the limit of $\frac{1}{d_L} \| f(D) \odot \Delta M \|_F^2 \to 0$,

$$\tilde{F}^{\mathrm{opt}} = \frac{1}{d_L^2} \left\| \mathrm{Tr}_L \sqrt{M} \right\|_F^2 = \frac{1}{d_L^2} \left\| \mathrm{Tr}_L \sqrt{I_L \otimes D} \right\|_F^2 = \mathrm{Tr}\, D. \tag{A.62}$$

□

A special case is when D contains all the correctable elements, i.e. $\mathrm{Tr}D = 1$, where both the near-optimal and the optimal fidelity will be unity.

A.5 Definitions and notations for examples of applications

The amplitude-damping, Pauli, and pure-loss noise conventions are given in Section 2.1.2; we use those definitions in the examples below.

A.5.1 Code definitions

For the standard qubit codes, we refer interested readers to the original references [51, 52, 87, 108]. We focus on the definitions for the bosonic codes, which largely follow the definitions in Ref. [3].

Cat codes are characterized by the coherent state amplitude, α, and the loss spacing, S, which is also known as the number of legs. The codewords are defined as

$$\begin{aligned} |0_L\rangle &\propto \hat{\Pi}_0 \, |\alpha\rangle \\ |1_L\rangle &\propto \hat{\Pi}_{S+1} \, |\alpha\rangle \, , \end{aligned} \tag{A.63}$$

where we omit the normalization factors for simplicity. Here, $\hat{\Pi}_0 = \sum_{n=0}^{\infty} |2n(S+1)\rangle\langle 2n(S+1)|$ and $\hat{\Pi}_{S+1} = \sum_{n=0}^{\infty} |(2n+1)(S+1)\rangle\langle (2n+1)(S+1)|$ are projectors onto Fock states that are 0 and $S+1$ mod $2(S+1)$ respectively. In the special case of $S=0$, they are the even and odd parity space projectors. Notice that since the codewords' support in the Fock basis has a spacing of $S+1$, the error subspace is orthogonal and exactly correctable as long as we consider less than $S+1$ excitation loss. Nevertheless, it is unnecessary to have a complete separation in terms of their Fock state support: two states supported in the same generalized parity space can still be orthogonal. For example, the recently discovered squeezed cat code. [56] explores such a property. The squeezed cat code has its codewords living in the even and odd parity spaces, i.e. $S=0$ in the convention of cat codes. However, in the limit of infinite squeezing, the code can still correct single excitation loss because of the orthogonality mentioned.

Binomial codes are characterized by the loss spacing, S, and the dephasing spacing, N. The codewords are defined as

$$\begin{aligned} |0_L\rangle &= \frac{1}{\sqrt{2^{N+1}}} \sum_{m=0}^{N+1} \sqrt{\binom{N+1}{m}} |(S+1)m\rangle \\ |1_L\rangle &= \frac{1}{\sqrt{2^{N+1}}} \sum_{m=0}^{N+1} (-1)^m \sqrt{\binom{N+1}{m}} |(S+1)m\rangle \,. \end{aligned} \tag{A.64}$$

It is clear that if we switch the basis to $|\pm_L\rangle$, the binomial code is protected by the Fock state subspace spacing very similar to the cat code. Therefore, it possesses the same property that it can exactly correct less than $S+1$ excitation loss.

The ideal GKP, which has infinite energy is defined on top of a symplective lattice. In the main text, we gave the square lattice as an example. In the position basis, we have that

$$\begin{aligned} |0_L\rangle &\propto \sum_{n\in\mathbb{Z}} \left|2n\sqrt{\pi}\right\rangle_x \\ |1_L\rangle &\propto \sum_{n\in\mathbb{Z}} \left|(2n+1)\sqrt{\pi}\right\rangle_x . \end{aligned} \tag{A.65}$$

The finite-energy envelope and Gram orthonormalization are defined in Section 3.1. For GKP codewords with a common eigenvalue under the two-fold rotation $e^{i\pi\hat{n}}$, the Gaussian envelope preserves their photon-number parity. A single-photon loss flips this parity, so the two-fold symmetry allows exact single-photon error detection by a parity measurement: $\hat{P}_L\hat{a}\hat{P}_L = 0$ [3, 31]. The unshifted square-lattice qubit GKP code above has this property. For general lattice phases and logical dimensions, the codewords need not share a common parity.

In the main text, we considered the performance comparison of cat, binomial, and GKP codes with increasing photon number. In particular, for cat and binomial codes, we increase their photon number by fixing the parameter S and increasing the coherent state amplitude, α, and the dephasing spacing, N, respectively. For the GKP code, we fix the underlying lattice geometry and decrease Δ.

A.6 Thermodynamic code and derivations

A.6.1 Code definition

Thermodynamic codes were first proposed in Ref. [82], where it was observed that the Eigenstate Thermalization Hypothesis automatically yielded approximate error correction codes in the bulk of the spectrum. We focus here on the simplest thermodynamic code, which can be seen as either an ETH code with respect to the one local Hamiltonian $\sum_i (I - \sigma_z^i)$, or a ground state encoding in the ferromagnetic Heisenberg model. The codewords on N qubits lie in magnetization sectors $m = -N, -N+2, \ldots, N$ and are

$$|h_m^N\rangle = \binom{N}{N/2 + m/2}^{-1/2} \sum_{\mathbf{s}:\sum_j s_j = m} |\mathbf{s}\rangle_N, \tag{A.66}$$

where $s_j = \pm 1$.

We focus on a logical qubit with $k = 2$. To define a logical qubit, we pick some

$d < N, m_0 < N - d$ and take

$$\begin{aligned} |0_L\rangle &= |h^N_{m_0}\rangle, \\ |1_L\rangle &= |h^N_{m_0+d}\rangle, \end{aligned} \tag{A.67}$$

where d is the distance of the code. In the main-text example, we use the centered choice $m_0 = -d/2$, as in the erasure expansion below.

A.6.2 One erasure

We first work out the case of a single erasure. The error channel can be written with Kraus operators

$$\begin{aligned} K_0 &= \sqrt{1-p}I, \\ K_1 &= \sqrt{p}|\Omega\rangle\langle -1|, \\ K_2 &= \sqrt{p}|\Omega\rangle\langle 1|, \end{aligned} \tag{A.68}$$

where p is the probability of erasure, and $|\Omega\rangle$ is some external state, orthonormal to $|1\rangle, |-1\rangle$, representing an erased qubit. Since the codewords are permutation invariant, without loss of generality, we can set the error-prone qubit to be the first qubit. Then the above Kraus operators are only supported on the first qubit, with identity on the rest of the system. We want to calculate the relevant QEC matrix.

First, note that if we simply pick $d > 2$ we will have $M_{l,\mu,l',\nu} \propto \delta_{\mu,\nu}\delta_{l,l'}$ such that

$$M = \sum_{l,\mu} c_{\mu,l}|l\rangle\langle l| \otimes |\mu\rangle\langle\mu|. \tag{A.69}$$

Then, for this problem, Eq. (A.32) simplifies to

$$1 - \tilde{F}^{\text{opt}} = 1 - \frac{1}{4}\sum_l \left(\sum_\mu \sqrt{c_{\mu,l}}\right)^2. \tag{A.70}$$

We simply have to calculate $c_{\mu,l}$, which are defined by

$$\langle\mu_L|K_l^\dagger K_l|\mu_L\rangle. \tag{A.71}$$

For $l = 0$ this is simply $1 - p$. Suppose μ_L has magnetization m. Then K_1 increases the magnetization by 1, so that this simply counts the number of states on $N - 1$ qubits with magnetization $m + 1$. Similarly, K_2 decreases the magnetization by 1. Then we have

$$\begin{aligned}
\langle h_m^N | K_1^\dagger K_1 | h_m^N \rangle &= p \frac{\binom{N-1}{(N+m)/2}}{\binom{N}{(N+m)/2}} = \frac{p}{2}(1 - m/N), \\
\langle h_m^N | K_2^\dagger K_2 | h_m^N \rangle &= p \frac{\binom{N-1}{(N+m-2)/2}}{\binom{N}{(N+m)/2}} = \frac{p}{2}(1 + m/N)
\end{aligned} \tag{A.72}$$

Specializing to the chosen codewords,

$$\begin{aligned}
\sum_l \left(\sum_\mu \sqrt{c_{\mu,l}} \right)^2 &= \sum_{l=0}^{2} \left(\sqrt{\langle h_{m_0}^N | K_l^\dagger K_l | h_{m_0}^N \rangle} + \sqrt{\langle h_{m_0+d}^N | K_l^\dagger K_l | h_{m_0+d}^N \rangle} \right)^2 \\
&= 4(1-p) + \frac{p}{2} \left(\sqrt{1 - \frac{m_0}{N}} + \sqrt{1 - \frac{m_0 + d}{N}} \right)^2 \\
&\quad + \frac{p}{2} \left(\sqrt{1 + \frac{m_0}{N}} + \sqrt{1 + \frac{m_0 + d}{N}} \right)^2 \\
&= 4(1-p) + 2p + p\sqrt{1 - \frac{m_0}{N}} \sqrt{1 - \frac{m_0 + d}{N}} \\
&\quad + p\sqrt{1 + \frac{m_0}{N}} \sqrt{1 + \frac{m_0 + d}{N}}
\end{aligned} \tag{A.73}$$

The most straightforward way to pick two states in the middle of the spectrum is to take $m_0 = -d/2$. We can treat $x \equiv d/N$ as a small parameter in the thermodynamic limit. The exact expression is

$$1 - \tilde{F}^{\text{opt}} = \frac{1}{2} p (1 - \sqrt{1 - x^2/4}). \tag{A.74}$$

For small x, we get

$$1 - \tilde{F}^{\text{opt}} = \frac{1}{16} p x^2 = \frac{1}{16} p \frac{d^2}{N^2} \tag{A.75}$$

which has the expected $1/N^2$ scaling [83]. At first glance, a strange property of this scaling is that infidelity increases with distance. We can understand this as follows: once $d > 2$, the

QEC matrix becomes diagonal. Then, infidelity is given by how close the diagonal elements are to each other. The smaller d is, the closer the codewords are to each other, such that the diagonal elements become more similar. Another way to think about this is that now d controls how far from the exact middle of the spectrum we are.

A.6.3 Two erasures

We now fix two qubits i, j where erasures occur. Again, without loss of generality, we can simply think of these as the first two qubits. If we pick $d > 4$, the QEC matrix will be diagonal along the codewords – however, it will not be diagonal along the Kraus operators. We can index Kraus operators by (α, β), and write

$$M = \sum_{\mu} \sum_{\alpha,\beta,\alpha',\beta'} c_{\mu,(\alpha,\beta),(\alpha',\beta')} |\alpha,\beta\rangle\langle\alpha',\beta'| \otimes |\mu\rangle\langle\mu| \tag{A.76}$$

This is certainly diagonal for the case where there is only an error on one qubit, as we know from studying one erasure, so

$$\begin{aligned} M = &\sum_{\mu} c_{\mu,(0,0),(0,0)} |0,0\rangle\langle 0,0| \otimes |\mu\rangle\langle\mu| \\ &+ \sum_{\mu} \sum_{\alpha\neq 0} c_{\mu,(\alpha,0),(\alpha,0)} \left(|\alpha,0\rangle\langle\alpha,0| + |0,\alpha\rangle\langle 0,\alpha|\right) \otimes |\mu\rangle\langle\mu| \\ &+ \sum_{\mu} \sum_{\alpha\neq 0,\beta\neq 0,\alpha'\neq 0,\beta'\neq 0} c_{\mu,(\alpha,\beta),(\alpha',\beta')} |\alpha,\beta\rangle\langle\alpha',\beta'| \otimes |\mu\rangle\langle\mu|, \end{aligned} \tag{A.77}$$

where the first row corresponds to no error, the second row to one error, and the last row to two errors. The one error elements can be read off using the results of the previous section: $c_{\mu,(0,0),(0,0)} = (1-p)^2, c_{\mu,(1,0),(1,0)} = \frac{1}{2}p(1-p)(1-m_0/N), c_{\mu,(2,0),(2,0)} = \frac{1}{2}p(1-p)(1+m_0/N)$.

Now, let's consider the two error parts of the QEC matrix (i.e. both α, β non-zero). First, note that if $\alpha = \beta$, then we must have $\alpha' = \beta' = \alpha = \beta$ for the matrix element to be nonzero. The reason for this is that we are either taking away two excitations, or adding two excitations, so in order to have a non-zero inner product, we must do the same to the other

side. With that, we have

$$M_{2\text{err}}^{(\mu)} = \begin{pmatrix} c_{\mu,(1,1),(1,1)} & 0 \\ 0 & c_{\mu,(2,2),(2,2)} \end{pmatrix} \oplus \begin{pmatrix} c_{\mu,(1,2),(1,2)} & c_{\mu,(1,2),(2,1)} \\ c_{\mu,(2,1),(1,2)} & c_{\mu,(2,1),(2,1)} \end{pmatrix} \quad , \tag{A.78}$$

so thankfully, we only have the one 2×2 matrix to diagonalize. The diagonal part has elements,

$$\begin{aligned} c_{\mu,(1,1),(1,1)} &= \langle h_m^N | K_1^{\otimes 2\dagger} K_1^{\otimes 2} | h_m^N \rangle \\ &= p^2 \binom{N}{(N+m)/2}^{-1} \binom{N-2}{(N+m)/2} \\ &= p^2 \left(\frac{(N-m)^2}{4N(N-1)} - \frac{N-m}{2N(N-1)} \right), \\ c_{\mu,(2,2),(2,2)} &= \langle h_m^N | K_2^{\otimes 2\dagger} K_2^{\otimes 2} | h_m^N \rangle \\ &= p^2 \binom{N}{(N+m)/2}^{-1} \binom{N-2}{(N+m-4)/2} \\ &= p^2 \left(\frac{(N+m)^2}{4N(N-1)} - \frac{N+m}{2N(N-1)} \right). \end{aligned} \tag{A.79}$$

Now for the non-diagonal part. It is straightforward to see that all the elements are the same, which we will denote

$$M_{12} = \begin{pmatrix} c_{\mu,(1,2),(1,2)} & c_{\mu,(1,2),(2,1)} \\ c_{\mu,(2,1),(1,2)} & c_{\mu,(2,1),(2,1)} \end{pmatrix} \equiv c_{\mu,\times} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \to \sqrt{M_{12}} = \sqrt{\frac{c_{\mu,\times}}{2}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{A.80}$$

Since errors of the form $(1,2)$ do not change the magnetization, we can calculate $c_{\mu,\times}$ as follows:

$$\begin{aligned} c_{\mu,\times} &= \langle h_m^N | (K_1^\dagger \otimes K_2^\dagger)(K_1 \otimes K_2) | h_m^N \rangle \\ &= p^2 \binom{N}{(N+m)/2}^{-1} \binom{N-2}{(N+m-2)/2} \\ &= p^2 \frac{N^2 - m^2}{4N(N-1)} \end{aligned} \tag{A.81}$$

With that we can now calculate the infidelity. Denoting the diagonal and non diagonal

parts of M by M_1, M_2 so that $M = M_1 \oplus M_2$ and

$$\begin{aligned} M_1 =& \sum_\mu c_{\mu,(0,0),(0,0)} |0,0\rangle\langle 0,0| \otimes |\mu\rangle\langle\mu| \\ &+ \sum_\mu \sum_{\alpha\neq 0} c_{\mu,(\alpha,0),(\alpha,0)} \left(|\alpha,0\rangle\langle\alpha,0| + |0,\alpha\rangle\langle 0,\alpha|\right) \otimes |\mu\rangle\langle\mu| \\ &+ \sum_\mu \sum_{\alpha=\beta\neq 0} c_{\mu,(\alpha,\alpha),(\alpha,\alpha)} |\alpha,\alpha\rangle\langle\alpha,\alpha| \otimes |\mu\rangle\langle\mu|, \\ M_2 =& \oplus_\mu \, c_{\mu,\times} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \end{aligned} \tag{A.82}$$

one can plug and chug to obtain,

$$\begin{aligned} (\mathrm{Tr}_\mu \sqrt{M})^2 =& (\mathrm{Tr}_\mu \sqrt{M_1})^2 + (\mathrm{Tr}_\mu \sqrt{M_2})^2 \\ \|(\mathrm{Tr}_\mu \sqrt{M_1})\|^2 =& \left(\sum_\mu \sqrt{c_{\mu,(0,0),(0,0)}} \right)^2 \\ &+ 2 \sum_{\alpha=1,2} \left(\sum_\mu \sqrt{c_{\mu,(\alpha,0),(\alpha,0)}} \right)^2 + \sum_{\alpha=1,2} \left(\sum_\mu \sqrt{c_{\mu,(\alpha,\alpha),(\alpha,\alpha)}} \right)^2 \\ \|\mathrm{Tr}_\mu \sqrt{M_2}\|^2 =& 2 \left(\sum_\mu \sqrt{c_{\mu,\times}} \right)^2 \end{aligned} \tag{A.83}$$

Finally, the fidelity will be

$$\tilde{F}^{\mathrm{opt}} = (1-p)^2 + \frac{1}{2} \sum_{\alpha=1,2} \left(\sum_\mu \sqrt{c_{\mu,(\alpha,0),(\alpha,0)}} \right)^2 + \frac{1}{4} \sum_{\alpha=1,2} \left(\sum_\mu \sqrt{c_{\mu,(\alpha,\alpha),(\alpha,\alpha)}} \right)^2 + \frac{1}{2} \left(\sum_\mu \sqrt{c_{\mu,\times}} \right)^2 . \tag{A.84}$$

We can now set $m_0 = -d/2, m_1 = d/2$ as usual, and obtain exact expressions. The full expression however, is long and tedious, so we take the limit $N \gg 1, d/N \ll 1$ in turn, and obtain

$$1 - \tilde{F}^{\mathrm{opt}} = \left(\frac{1}{2} p - \frac{3}{8} p^2 \right) \frac{d^2}{N^2} + O\left(\frac{d^2}{N^3} \right). \tag{A.85}$$

Note that in the 1 erasure case, we could have taken $m \sim O(N)$ and still obtained the d^2/N^2 scaling. However, here we cannot – if we take $m_0 = (N-d)/2$ for instance, it turns out there

will remain uncorrectable errors even in the thermodynamic limit.

A.6.4 Constant number of erasures under sufficiently large distance

Now we consider the case where there are l erasures for an instance of a thermodynamic code with distance $d > l$. Again, WLOG we place these errors on the first l qubits. The matrix elements we have to calculate all have the form

$$\begin{aligned} c_\mu^{(l_1)} &= p^l(1-p)^{N-l}\langle h_{m_\mu}^N| \left(K_1^{\otimes l_1}\otimes K_2^{\otimes l-l_1}\otimes K_0^{\otimes N-l}\right)^\dagger \left(K_1^{\otimes l_1}\otimes K_2^{\otimes l-l_1}\otimes K_0^{\otimes N-l}\right)|h_{m_\mu}^N\rangle \\ &= p^l(1-p)^{N-l}\binom{N}{(N+m_\mu)/2}^{-1}\binom{N-l}{((N+m_\mu)/2-(l-l_1)}, \end{aligned} \tag{A.86}$$

where the second factor counts the number of states on $N-l$ qubits with magnetization $m_\mu - l + 2l_1$. Here, m_μ is the magnetization of the codeword $|\mu_L\rangle$. We will set $p = 1$, since we are interested in what happens if these l qubits were deterministically erased. We will assume that $m_\mu \ll N$, in which case we can expand the above in powers of m_μ/N as follows:

$$\begin{aligned} c_\mu^{(l_1)} &\simeq \frac{1}{N^l}\left(\frac{N+m_\mu}{2}\right)^{l-l_1}\left(\frac{N-m_\mu}{2}\right)^{l_1} \\ &= \frac{1}{2^l}\left(1+\frac{m_\mu}{N}\right)^{l-l_1}\left(1-\frac{m_\mu}{N}\right)^{l_1} \\ &\simeq \frac{1}{2^l}\left(1+(l-2l_1)\frac{m_\mu}{N}+\frac{1}{2}(l^2+4l_1^2-4ll_1-l)\frac{m_\mu^2}{N^2}\right)+O\left(\frac{1}{N^3}\right). \end{aligned} \tag{A.87}$$

where we have obtained the first approximate equality by only keeping track of terms with N, m_μ.

Now, any permutation of l_1 K_1's and $l-l_1$ K_2's on the first l qubits will give the same overlap (and any other configuration of Kraus operators will give 0 overlap), so the above gives a block of the QEC matrix of size $\binom{l}{l_1}$ by $\binom{l}{l_1}$, with every matrix element being the same number above. Let us denote a matrix of ones of size $d\times d$ by O_1^d. The full QEC matrix then takes the form

$$M = \bigoplus_{\mu=0,1}\bigoplus_{l_1=0}^{l} c_\mu^{(l_1)} O_1^{\binom{l}{l_1}}, \tag{A.88}$$

such that

$$\begin{aligned}\widetilde{F}^{\text{opt}} = \frac{1}{4}\|\text{Tr}_\mu\sqrt{M}\|_F^2 &= \frac{1}{4}\sum_{l_1=0}^{l}\text{Tr}\left(\sum_\mu \sqrt{c_\mu^{(l_1)}O_1^{\binom{l}{l_1}}}\right)^2 \\ &= \frac{1}{4}\sum_{l_1=0}^{l}\binom{l}{l_1}\left(\sum_\mu\sqrt{c_\mu^{(l_1)}}\right)^2.\end{aligned} \tag{A.89}$$

The exact expression can now be easily evaluated numerically, or we can continue with our expansion in terms of d/N as before. Setting $m_\mu = \pm d/2, x = d/N$, and using the approximate form in Eq. (A.86), we have

$$\left(\sum_\mu\sqrt{c_\mu^{(l_1)}}\right)^2 \simeq \frac{4}{2^l}\left(1+\frac{1}{16}\left(-2l+l^2-4ll_1+4l_1^2\right)x^2\right) \tag{A.90}$$

To evaluate the prefactor of x^2, we note that $\sum_{l_1=0}^{l}\binom{l}{l_1} = 2^l, \sum_{l_1=0}^{l}\binom{l}{l_1}l_1 = 2^{l-1}l, \sum_{l_1=0}^{l}\binom{l}{l_1}l_1^2 = 2^{l-2}l(l+1)$. This leads us to the simple final expression,

$$1-\widetilde{F}^{\text{opt}} = \frac{l}{16}x^2 + O\left(\frac{1}{N^3}\right) = \frac{l}{16}\frac{d^2}{N^2} + O\left(\frac{1}{N^3}\right). \tag{A.91}$$

Note that the above equation is only valid when we keep l constant and take N to be large, but will not hold if we allow l to scale with N. Eq. (A.91) matches with the single and two erasure error expressions by setting $l = 1, 2$ and unit erasure probability $p = 1$.

A.7 Infidelity proxy for a general complementary reference

This section supports Section 2.3. The transpose (Petz) recovery attains the near-optimal entanglement fidelity $\tilde{F}^{\text{opt}}$ of Theorem 2.4 [73]. Lemma 2.10 bounds its infidelity by a weighted deviation of the QEC matrix from the Knill–Laflamme form; this infidelity proxy is the quantity used in the rate proofs of Appendix B. We use the QEC matrix $M_{\mu a,\nu b}$ of Eq. (2.9) for an isometric encoding $\hat{V}$ and Kraus operators $\{\hat{K}_a\}$, with $(\text{Tr}_L B)_{ab} := \sum_{\mu=1}^{d_L} B_{\mu a,\mu b}$, and write $F^{\text{opt}} := \sup_{\mathcal{R}} F(\mathcal{R}\circ\mathcal{N}\circ\mathcal{E})$ for the optimal entanglement fidelity, so that $\frac{1}{2}(1-\tilde{F}^{\text{opt}}) \leq 1-F^{\text{opt}} \leq 1-\tilde{F}^{\text{opt}}$.

Use the complementary reference, admissibility condition, and infidelity proxy of Definitions 2.7 and 2.8. The marginal reference and its Choi-state interpretation are given in Eq. (2.22) and Remark 2.2; Corollary 2.9 gives the local perturbative expansion. We prove the global bound below.

The following restates Lemma 2.10.

Lemma A.5 (Complementary-reference bound)**.** *Let M be the QEC matrix of an isometric encoding and a trace-preserving channel. For every admissible reference D from Definition 2.8, the near-optimal* infidelity *satisfies*

$$1 - \tilde{F}^{opt} \leq 4\,\tilde{\epsilon}_D(M). \tag{A.92}$$

Proof. Apply the square-root comparison of Lemma B.30 (Appendix B.15) to the QEC matrix and its complementary reference. An orthogonal projection then relates that distance to the near-optimal infidelity.

Using Lemma B.30, the QEC matrix and its admissible reference obey

$$\frac{1}{d_L}\left\|\sqrt{M} - I_L \otimes \sqrt{D}\right\|_F^2 \leq 4\tilde{\epsilon}_D(M). \tag{A.93}$$

Let $\Pi(Z) := (I_L/d_L) \otimes \mathrm{Tr}_L\, Z$ be the Hilbert–Schmidt orthogonal projection onto operators proportional to the logical identity. Equation (2.15) gives $\left\|\Pi\sqrt{M}\right\|_F^2 = d_L \tilde{F}^{\mathrm{opt}}$, while $\mathrm{Tr}\, M = d_L$. Since $I_L \otimes \sqrt{D}$ lies in the range of Π, orthogonality and the bound above give

$$\begin{aligned} 1 - \tilde{F}^{\mathrm{opt}} &= \frac{1}{d_L}\left\|(I - \Pi)\sqrt{M}\right\|_F^2 \\ &\leq \frac{1}{d_L}\left\|\sqrt{M} - I_L \otimes \sqrt{D}\right\|_F^2 \leq 4\tilde{\epsilon}_D(M). \end{aligned} \tag{A.94}$$

□

The coefficient four is independent of dimension and reference spectrum, so the bound applies along the coding limit. The marginal reference gives the tighter unit coefficient in the

perturbative limit of Eq. (2.23).

APPENDIX B

APPENDICES FOR CHAPTER 3

This appendix follows the general derivation used in Chapter 3: lattice and information-theory foundations, channel dilations, finite-energy GKP overlaps, the ordered QEC expansion, and its limiting environmental kernel. Thermal and vacuum performance formulas specialize that calculation. The coding theorem then combines the general-reference bound with environment approximation and lattice averaging. Numerical protocols and mathematical identities complete the derivation.

B.1 Basics on information theory

Section 3.4 defines coherent information and capacity and relates entanglement fidelity to an achievable rate. Here we give the generalized-Pauli hashing expression used in the calculations. For a generalized Pauli channel in the Weyl operator basis, the hashing bound provides an achievable rate for random stabilizer codes and can be written as

$$D_{\mathcal{N}} = (1 - H_{d_L}(\mathbf{p})) \log_2 d_L \tag{B.1}$$

where d_L is the number of levels of the qudits, $\mathbf{p}$ is the probability vector, H_n is the entropy with base n, i.e. $H_n(\mathbf{x}) := -\sum_i x_i \log_n x_i$. $\log_2 d_L$ converts the unit from qudit to qubit. Worth noting, there are also other more general definitions of the Hashing bound (see, e.g., Refs. [72] and references therein) for general channels, but it converges to the expression above for qudit Pauli channels.

B.2 Basics on lattice theory

We use the generator, symplectic Gram matrix, dual lattice, and canonical-basis conventions of Section 3.1. The integer canonical-basis reduction is established in Ref. [123]. The following verifies the square-lattice transformation used there.

Lemma B.1 (Square lattice with a symplectic transformation)**.** *The canonical generator matrix of any symplectic integral lattice, M_Λ, can be connected with a square lattice through*

$$M_\Lambda = L_{sq} S^T \tag{B.2}$$

where $L_{sq} = D_{sq} \otimes I_2 = \mathrm{diag}\left(\sqrt{d_1}, \cdots, \sqrt{d_N}\right) \otimes I_2$, and $S^T = L_{sq}^{-1} M_\Lambda$ is a symplectic matrix.

Lemma B.1 is straightforward, and S can be easily verified to be symplectic.

Lemma B.2 (Existence of good self-dual lattice; [128])**.** *For each nonnegative integrable, rotationally invariant function f, there exists a $2N$-dimensional symplectic self-dual lattice Λ such that*

$$\sum_{\mathbf{x} \in \Lambda \notin \{\mathbf{0}\}} f(\mathbf{x}) \leq \int f(\mathbf{x}) d^{2N}\mathbf{x} \tag{B.3}$$

The chosen lattice may depend on f.

B.3 Ideal GKP code and stabilizer phases

The stabilizers and logical dimension are defined in Section 3.1. We retain the Gaussian transformation proof and explicit ideal codewords needed for the overlap calculation.

Corollary B.3 (Gaussian transformation to a square lattice)**.** *Suppose a GKP code's underlying symplectic integral lattice, Λ, has a canonical symplectic Gram matrix of the form $A_\Lambda = \mathrm{diag}\left(d_1, \cdots, d_N\right) \otimes \omega$. There exists a Gaussian transformation, $\hat{U}_S$, that transforms code into a GKP code that corresponds to a square lattice, Λ_{sq}, with a generator of the form $L_{sq} = D_{sq} \otimes I_2 = \mathrm{diag}\left(\sqrt{d_1}, \cdots, \sqrt{d_N}\right) \otimes I_2$.*

Proof. With a Gaussian unitary, $\hat{U}_S$, the quadrature operators transform as

$$\hat{U}_S^\dagger \hat{x} \hat{U}_S = S\hat{x} \tag{B.4}$$

where S is a symplectic matrix. Equivalently, one can show that the GKP stabilizer group

with a generator matrix M_Λ is transformed into another GKP code with $M'_\Lambda = M_\Lambda S^T$. Therefore, from Lemma B.1, it is straightforward to take $S^T = M_\Lambda^{-1} L_{\text{sq}}$, leading to $M'_\Lambda = L_{\text{sq}} = \text{diag}\left(\sqrt{d_1}, \cdots, \sqrt{d_N}\right) \otimes I_2$. □

Therefore, we can understand each GKP code as being Gaussian transformed from a square lattice GKP code, where the ith mode encodes d_i logical dimensions. For an infinite energy GKP code, the codewords can then be denoted as $|\boldsymbol{\mu}\rangle_0$ where $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix}$ and $\mu_i \in \mathbb{Z}_{d_i}$. The codewords follow as

$$|\boldsymbol{\mu}\rangle_0 := \hat{U}_S^\dagger |\boldsymbol{\mu}\rangle_0^{\text{sq}} = \hat{U}_S^\dagger \bigotimes_{i=1}^{N} |\mu_i\rangle_0^{\text{sq}} \tag{B.5}$$

where the square-lattice codeword of mode i is

$$|\mu_i\rangle_0^{\text{sq}} := \sum_{n\in\mathbb{Z}} \left| \sqrt{2\pi/d_i}(d_i n + \mu_i) \right\rangle_x \tag{B.6}$$

with the codewords defined in the conventional Z basis.

B.4 Noise channels and their dilations

Section 3.2 defines Gaussian-dilatable channels, their purification, and the attenuation and amplification interactions. We use that setup here and keep the complementary-output operator relations needed in the derivation.

For attenuation and amplification, the signal mode $\hat{a}$ interacts with one environment mode $\hat{e}$ in state ρ_E. We order the input factors as $\mathcal{H}_a \otimes \mathcal{H}_E$ and the output factors as $\mathcal{H}_b \otimes \mathcal{H}_f$, so $\hat{U} : \mathcal{H}_a \otimes \mathcal{H}_E \to \mathcal{H}_b \otimes \mathcal{H}_f$. The identity on a factor X is I_X. For a mixed bath, the complementary output contains both f and its environmental purifier R, as specified in Section 3.2. The coding proof approximates an independent bath $\rho_E^{\otimes N}$ by correlated finite-block states.

B.4.1 Attenuation and amplification

Take the beam-splitter or two-mode-squeezing unitary of Section 3.2. Let $\hat{a}_b$ and $\hat{a}_f$ be the local annihilation operators on the two output factors. The corresponding Heisenberg-picture mode operators are

$$\hat{b} := \hat{U}^\dagger(\hat{a}_b \otimes I_f)\hat{U}, \qquad \hat{f} := \hat{U}^\dagger(I_b \otimes \hat{a}_f)\hat{U}. \tag{B.7}$$

The operators $\hat{b}$ and $\hat{f}$ act on the joint input space, while b, f label output subsystems. Table B.1 gives their input-output relations and the coefficients c_1, c_2.

Table B.1: Heisenberg-picture input-output relations and interaction coefficients.

Interaction	$\hat{b}$	$\hat{f}$	c_1	c_2	$\chi = c_2/c_1$
Beam splitter	$\sqrt{\eta}\,\hat{a} \otimes I_E + \sqrt{1-\eta}\, I_a \otimes \hat{e}$	$-\sqrt{1-\eta}\,\hat{a} \otimes I_E + \sqrt{\eta}\, I_a \otimes \hat{e}$	$1-\eta$	η	$\eta/(1-\eta)$
Two-mode squeezing	$\sqrt{G}\,\hat{a} \otimes I_E + \sqrt{G-1}\, I_a \otimes \hat{e}^\dagger$	$\sqrt{G-1}\,\hat{a}^\dagger \otimes I_E + \sqrt{G}\, I_a \otimes \hat{e}$	$G-1$	G	$G/(G-1)$

The complementary-displacement transformation and the environmental coordinate $\beta(x)$ are given in Eqs. (3.21) and (3.22). They include the complex conjugation for amplification and the Jacobian $d^{2N}\beta = \pi^N \chi^N d^{2N}x$.

B.4.2 Vacuum, thermal, and non-Gaussian baths

A vacuum bath gives pure loss or quantum-limited amplification, with the Kraus operators $\hat{E}_\ell$ and $\hat{A}_\ell$ of Eq. (3.24) and the loss probability $\gamma := 1 - \eta$ (Section 2.1.2 and Appendix B.4). With the thermal state $\rho_{\mathrm{th}}(\nu)$ of Eq. (3.23), the bath $\rho_E = \rho_{\mathrm{th}}(\bar{n})$ gives thermal attenuation or amplification. A non-Gaussian bath generally gives a non-Gaussian channel under the same interaction. Additive Gaussian displacement noise and other Gaussian canonical forms are treated in Sec. B.10.4.

B.5 Finite-energy GKP codes

Use the individually normalized codewords, raw operator $\hat{C}_\Delta$, logical Weyl matrices $\hat{W}_x$, and energy parameter n_Δ defined in Section 3.1. The individual normalization is $N_{\Delta,\mu} = \mathcal{N}_\Delta / \sqrt{(\hat{G}_\Delta)_{\mu\mu}}$; it does not impose orthogonality. Equation (3.12) gives the displaced overlap.

Its scalar prefactor is absorbed into $\hat{C}_\Delta$, following Appendix E of Ref. [13]. Using an isometry in that expression requires Gram correction on both sides.

Proposition B.4 (Gram orthonormalization)**.** *At a fixed full-rank lattice,*

$$\hat{G}_\Delta := \hat{C}_\Delta^\dagger \hat{C}_\Delta = I_L + \sum_{x \in \Lambda^\perp \setminus \{0\}} \hat{W}_x e^{-(n_\Delta + 1/2)\left\|\xi(x)\right\|^2} \longrightarrow I_L \quad (\Delta \to 0).$$

For every $\Delta > 0$, $\hat{G}_\Delta > 0$ *and* $\hat{V}_\Delta := \hat{C}_\Delta \hat{G}_\Delta^{-1/2}$ *is a finite-energy isometry.*

Proof. Set $\alpha = 0$ in Eq. (3.12). Bounded $\hat{W}_x$ and polynomial lattice-point counts make the nonzero-vector sum tend to zero as $\Delta \to 0$. The ideal codewords are linearly independent tempered distributions. Lemma B.31 shows that the envelope operator preserves their linear independence for every $\Delta > 0$. Thus $\hat{G}_\Delta > 0$ for every $\Delta > 0$ and $\hat{V}_\Delta^\dagger \hat{V}_\Delta = I_L$. The oscillator heat kernel maps the ideal tempered combs to Schwartz functions for every $\Delta > 0$, so their energy is finite; Gram orthonormalization preserves this property.

□

The mean-energy expression is Eq. (3.14). It follows by differentiating the displaced overlap at $\alpha = 0$ and taking the Gram-orthonormalized logical trace; the nonzero-vector terms supply the exponentially suppressed correction.

B.6 Infidelity proxy of GKP codes

Section B.6.1 derives an exact ordered expansion of the finite-energy QEC matrix. Section B.6.2 proves that its higher-order terms vanish and evaluates the remaining infidelity proxy as a convergent lattice sum.

B.6.1 Ordered QEC expansion

The finite-energy QEC matrix admits an exact expansion in inverse powers of $(1 + \tilde{n}_\Delta)$, with $\tilde{n}_\Delta$ growing linearly with the code energy parameter n_Δ. Expanding the Kraus products in displacement operators and inserting the GKP overlap separates the environmental

displacement overlap from higher-order corrections. Section B.6.2 controls these corrections in the weighted norm used for recovery.

Throughout Sec. B.6, fix the number of modes N, a symplectically integral lattice Λ, and the environmental state

$$\sigma := \sum_{r=1}^{K} \lambda_r \left|\phi_r\right\rangle \left\langle\phi_r\right|, \qquad \lambda_r > 0, \quad \sum_r \lambda_r = 1, \tag{B.8}$$

whose orthonormal eigenvectors lie in a finite N-mode Fock subspace. The physical channel has a one-mode environment ρ_E. For the block-channel approximation in Sec. B.10.1, we also allow σ to have correlations across the N environmental modes. Purifying by the eigenvector label, the Kraus operators are

$$K_{\ell r} := \sqrt{\lambda_r}\,(I_b \otimes {}_f\langle\ell|)\hat{U}(I_a \otimes |\phi_r\rangle_E) : \mathcal{H}_a \longrightarrow \mathcal{H}_b. \tag{B.9}$$

For N modes, $\hat{U}$ denotes the product of the N interaction unitaries, with factors regrouped into signal and environment blocks.

We use the interaction coefficients c_1, c_2 from Table B.1. The environmental state has finite Fock support, so choose $m_* < \infty$ with $\boldsymbol{m} \in \mathcal{I} := \{0, \ldots, m_*\}^N$ for every retained Fock component. For a multi-index $\boldsymbol{i}$, its 1-norm is $|\boldsymbol{i}|_1 := \sum_j i_j$. Define

$$\tilde{n}_\Delta := c_1 n_\Delta + \frac{c_1 + c_2 - 1}{2}. \tag{B.10}$$

The interaction coefficients give the exact identities

$$\tilde{n}_\Delta = \begin{cases} (1-\eta)n_\Delta, & \text{beam splitter,} \\ (G-1)(n_\Delta + 1), & \text{two-mode squeezing.} \end{cases} \tag{B.11}$$

Thus $\tilde{n}_\Delta$ is the code energy parameter scaled by the interaction strength, with the vacuum contribution $+1$ for amplification. For a vacuum environment it is the mean occupation of

the discarded output when the channel input is $\rho_{\mathrm{th}}(n_\Delta)$. Define $t := \tilde{n}_\Delta/(1+\tilde{n}_\Delta)$. At fixed interaction, $\Delta \to 0$ gives $\tilde{n}_\Delta \to \infty$ and $t \to 1$. We write $\hat{f}_j$ for the local annihilation operator on output mode j, and $\hat{\boldsymbol{f}}^{\boldsymbol{i}} := \prod_j \hat{f}_j^{i_j}$.

For the raw encoding $\hat{C}_\Delta$ and environmental state σ, write M_Δ for the QEC matrix with entries $\langle \mu_L | \hat{C}_\Delta^\dagger K_{\ell r}^\dagger K_{ks} \hat{C}_\Delta | \nu_L \rangle$. The encoding isometry $\hat{V}_\Delta = \hat{C}_\Delta \hat{G}_\Delta^{-1/2}$ incorporates the finite-energy overlaps; its QEC matrix is obtained by the same logical congruence, as in Remark B.9.

Lemma B.5 (Ordered expansion of the QEC matrix)**.** *Under the preceding setup, M_Δ has the exact decomposition*

$$M_\Delta = \sum_{x \in \Lambda^\perp} \hat{W}_x \otimes \sum_{r,s=1}^{K} M_x^{rs} \otimes |r\rangle \langle s| \,, \tag{B.12}$$

where the tensor factors are the logical space, the N output environmental modes, and the K-dimensional purification-label space. The last factor labels the environment eigenvectors. Define $\hat{R}_x := \bigotimes_{j=1}^N \hat{R}_{x,j}$ by

$$\hat{R}_{x,j} := (1-t) \exp\left[\frac{\sqrt{c_1 n_\Delta (n_\Delta+1)}}{(1+\tilde{n}_\Delta)\sqrt{\chi}} \beta_j(x)^* \hat{f}_j^\dagger \right] t^{\hat{f}_j^\dagger \hat{f}_j} \exp\left[-\frac{\sqrt{c_1 n_\Delta (n_\Delta+1)}}{(1+\tilde{n}_\Delta)\sqrt{\chi}} \beta_j(x) \hat{f}_j \right]. \tag{B.13}$$

Using Lemma B.32, the ordered operator has a same-sided displacement representation. Each complementary block is

$$\begin{aligned} M_x^{rs} = \sqrt{\lambda_r \lambda_s}\, e^{\left(\frac{c_1 n_\Delta (n_\Delta+1)}{1+\tilde{n}_\Delta} - n_\Delta - \frac{1}{2}\right) \|\beta(x)/\sqrt{\chi}\|^2} \Big[& d_{rs,\Delta}(x) \hat{R}_x \\ & + \sum_{\substack{\boldsymbol{i},\boldsymbol{j} \in \mathcal{I} \\ |\boldsymbol{i}|_1 + |\boldsymbol{j}|_1 > 0}} \left(\frac{\sqrt{c_2}}{1+\tilde{n}_\Delta} \right)^{|\boldsymbol{i}|_1 + |\boldsymbol{j}|_1} d_{rs\boldsymbol{ij},\Delta}(x) (\hat{\boldsymbol{f}}^\dagger)^{\boldsymbol{i}} \hat{R}_x \hat{\boldsymbol{f}}^{\boldsymbol{j}} \Big]. \end{aligned} \tag{B.14}$$

The coefficients $d_{rs\boldsymbol{ij},\Delta}(x)$ can be expressed through complex Hermite polynomials. Every higher-order term carries the explicit inverse power $(1+\tilde{n}_\Delta)^{-|\boldsymbol{i}|_1 - |\boldsymbol{j}|_1}$; its weighted-norm suppression is established in Lemma B.7. The zeroth-order coefficient is

$$
\begin{aligned}
d_{rs,\Delta}(x) = \exp\!\left[\frac{c_1^2 n_\Delta(n_\Delta+1)}{2(1+\tilde{n}_\Delta)((1+\tilde{n}_\Delta)-c_2)}\|\beta(x)\|^2\right] \\
\times \langle\phi_r|\,(1-c_2/(1+\tilde{n}_\Delta))^{\hat{n}_{\rm tot}/2} \\
\times \hat{D}\left(\frac{c_1\sqrt{n_\Delta(n_\Delta+1)}}{\sqrt{(1+\tilde{n}_\Delta)((1+\tilde{n}_\Delta)-c_2)}}\beta(x)\right)(1-c_2/(1+\tilde{n}_\Delta))^{\hat{n}_{\rm tot}/2}\,|\phi_s\rangle\,.
\end{aligned}
\tag{B.15}
$$

For each fixed x, its large-energy expansion is

$$
\begin{aligned}
&d_{rs,\Delta}(x) = d_{rs,\infty}(x) + \mathcal{O}(\tilde{n}_\Delta^{-1}), \\
&d_{rs,\infty}(x) := e^{\left\|\beta(x)\right\|^2/2}\,\langle\phi_r|\,\hat{D}(\beta(x))\,|\phi_s\rangle\,.
\end{aligned}
\tag{B.16}
$$

Order counts explicit inverse powers of $(1+\tilde{n}_\Delta)$, with finite-energy dependence retained in the coefficients and $\hat{R}_x$. The weighted-norm estimate is proved in Sec. B.6.2.

Proof. We first express each lattice contribution by a Gaussian generating function, then extract the finite expansion, and finally identify the zeroth-order coefficient for arbitrary environment eigenvectors.

Displacement representation and generating function. Expanding the Kraus products in the displacement-operator basis gives

$$
K_{\ell r}^\dagger K_{ks} = \sqrt{\lambda_r\lambda_s}\int\frac{d^{2N}\alpha}{\pi^N}\,\langle k|\,\hat{D}_f(-\alpha)\,|\ell\rangle\,\langle\phi_r|\,\hat{D}_E(\sqrt{c_2}\alpha)\,|\phi_s\rangle\,\hat{D}_a(g(\alpha)).
\tag{B.17}
$$

Here g is defined in Eq. (3.21). Inserting Eq. (3.12) gives a dual-lattice sum. Its scalar coefficient in one mode for environment Fock states $|m\rangle\,,|n\rangle$ is

$$
\begin{aligned}
\mathcal{B}_x(\ell,m;k,n) := \int\frac{d^2\alpha}{\pi}\,\langle k|\,\hat{D}(-\alpha)\,|\ell\rangle\,\langle m|\,\hat{D}(\sqrt{c_2}\alpha)\,|n\rangle \\
\times\exp\Big[-(n_\Delta+\tfrac{1}{2})(c_1|\alpha|^2+|\beta(x)/\sqrt{\chi}|^2) \\
+\frac{\sqrt{c_1 n_\Delta(n_\Delta+1)}}{\sqrt{\chi}}(\alpha\beta(x)^*+\alpha^*\beta(x))\Big]\,.
\end{aligned}
\tag{B.18}
$$

This excludes $\hat{W}_x$ and $\sqrt{\lambda_r\lambda_s}$. Writing $\phi_{r,\boldsymbol{m}} := \langle \boldsymbol{m}|\phi_r\rangle$, let $\mathcal{B}_{x,j}$ be the same formula with $\beta(x)$ replaced by its component $\beta_j(x)$. The multimode complementary block is

$$(M_x^{rs})_{\boldsymbol{\ell},\boldsymbol{k}} := \sqrt{\lambda_r\lambda_s} \sum_{\boldsymbol{m},\boldsymbol{n}\in\mathcal{I}} \phi^*_{r,\boldsymbol{m}}\phi_{s,\boldsymbol{n}} \prod_{j=1}^{N} \mathcal{B}_{x,j}(\ell_j, m_j; k_j, n_j). \tag{B.19}$$

This proves Eq. (B.12). To evaluate the coefficients, introduce independent formal variables v, y, z, w and the exponential generating function

$$\mathcal{G}_x(v, y, z, w) := \sum_{\ell,m,k,n\geq 0} \frac{\mathcal{B}_x(\ell, m; k, n)}{\sqrt{\ell!m!k!n!}}\, v^\ell y^n z^k w^m.$$

The formal variables y, w label the environment ket and bra. The Fock generating functions of the two displacement matrix elements are $\exp\Big[vz - |\alpha|^2/2 + \alpha^* v - \alpha z\Big]$ and $\exp\Big[yw - c_2|\alpha|^2/2 + \sqrt{c_2}(\alpha w - \alpha^* y)\Big]$. Gaussian integration gives

$$\begin{aligned}\mathcal{G}_x = (1+\tilde{n}_\Delta)^{-1} \exp&\left[\left(\frac{c_1 n_\Delta(n_\Delta+1)}{1+\tilde{n}_\Delta} - n_\Delta - \frac{1}{2}\right)\frac{|\beta(x)|^2}{\chi}\right] \\ &\times \exp\left[tvz + (1 - c_2/(1+\tilde{n}_\Delta))yw + \frac{\sqrt{c_2}}{1+\tilde{n}_\Delta}(vw + yz)\right] \\ &\times \exp\left[\frac{\sqrt{c_1 n_\Delta(n_\Delta+1)}}{(1+\tilde{n}_\Delta)\sqrt{\chi}}\Big(\beta(x)^*(v - \sqrt{c_2}y) - \beta(x)(z - \sqrt{c_2}w)\Big)\right].\end{aligned} \tag{B.20}$$

Coefficient extraction. Temporarily take a one-mode Fock environment, $\phi_r = |r\rangle$ and $\phi_s = |s\rangle$. The factor $(1+\tilde{n}_\Delta)^{-1}\exp\left[tvz + \frac{\sqrt{c_1 n_\Delta(n_\Delta+1)}}{(1+\tilde{n}_\Delta)\sqrt{\chi}}(\beta(x)^* v - \beta(x)z)\right]$ generates the matrix entries of $\hat{R}_x$. The remaining coupling factor expands as

$$e^{\frac{\sqrt{c_2}}{1+\tilde{n}_\Delta}(vw+yz)} = \sum_{i,j\geq 0} \frac{\left(\frac{\sqrt{c_2}}{1+\tilde{n}_\Delta}\right)^{i+j} v^i w^i y^j z^j}{i!j!}.$$

To extract $y^s w^r$, a term with $w^i y^j$ must be multiplied by $w^{r-i}y^{s-j}$ from the other factor. That factor has only nonnegative powers, so $i > r$ or $j > s$ gives zero. The coefficient extraction is therefore finite.

Define the scaled complex Hermite coefficients by

$$h_{ab}(x) := [y^a w^b] \exp\left[(1 - c_2/(1+\tilde{n}_\Delta))yw + \frac{c_1\sqrt{n_\Delta(n_\Delta+1)}}{1+\tilde{n}_\Delta}(-\beta(x)^* y + \beta(x) w)\right],$$

where $[y^a w^b]$ denotes coefficient extraction, a, b are nonnegative integer indices, and $h_{ab} := 0$ if either index is negative. Expansion gives

$$\begin{aligned} h_{ab}(x) &= \sum_{h=0}^{\min(a,b)} \left\{ \frac{(1 - c_2/(1+\tilde{n}_\Delta))^h}{h!(a-h)!(b-h)!} \right. \\ &\quad \left. \times \left(-\frac{c_1\sqrt{n_\Delta(n_\Delta+1)}\beta(x)^*}{1+\tilde{n}_\Delta}\right)^{a-h} \left(\frac{c_1\sqrt{n_\Delta(n_\Delta+1)}\beta(x)}{1+\tilde{n}_\Delta}\right)^{b-h} \right\} \\ &= \frac{(1 - c_2/(1+\tilde{n}_\Delta))^{(a+b)/2}}{\sqrt{a!b!}}\, e^{\frac{c_1^2 n_\Delta(n_\Delta+1)|\beta(x)|^2}{2(1+\tilde{n}_\Delta)((1+\tilde{n}_\Delta)-c_2)}} \\ &\quad \times \langle b|\, \hat{D}\left(\frac{c_1\sqrt{n_\Delta(n_\Delta+1)}}{\sqrt{(1+\tilde{n}_\Delta)((1+\tilde{n}_\Delta)-c_2)}}\beta(x)\right)|a\rangle\,. \end{aligned} \tag{B.21}$$

The second equality is the Hermite coefficient identity in Lemma B.33. Multiplication of the outcome generating function by $v^i z^j$ corresponds to $(\hat{f}^\dagger)^i \hat{R}_x \hat{f}^j$. Hence the one-mode coefficients are

$$d_{rsij,\Delta}(x) = \frac{\sqrt{r!s!}}{i!j!} h_{s-j,r-i}(x), \qquad 0 \le i \le r, \quad 0 \le j \le s. \tag{B.22}$$

Let $h_{ab}^{(k)}(x)$ be the one-mode coefficient evaluated at $\beta_k(x)$. Tensoring the identities and contracting the finite Fock expansions gives

$$d_{rs\boldsymbol{ij},\Delta}(x) = \sum_{\boldsymbol{m},\boldsymbol{n}\in\mathcal{I}} \phi^*_{r,\boldsymbol{m}} \phi_{s,\boldsymbol{n}} \prod_{k=1}^{N} \frac{\sqrt{m_k! n_k!}}{i_k! j_k!} h^{(k)}_{n_k - j_k, m_k - i_k}(x). \tag{B.23}$$

Separating $d_{rs,\Delta} := d_{rs\mathbf{00},\Delta}$ gives Eq. (B.14).

Zeroth-order coefficient and its limit. Setting $i = j = 0$ in Eq. (B.21) and contracting the

finite Fock expansions gives Eq. (B.15). At fixed interaction, Eq. (B.10) gives

$$\begin{aligned} \frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2} &\to \frac{1}{c_1}, \\ \frac{c_1 n_\Delta(n_\Delta+1)}{1+\tilde{n}_\Delta} - n_\Delta - \frac{1}{2} &= -\frac{(c_2+1)n_\Delta + (c_1+c_2+1)/2}{2(1+\tilde{n}_\Delta)} \to -\frac{c_2+1}{2c_1}. \end{aligned} \tag{B.24}$$

The Gaussian in each Fock displacement matrix element cancels the prefactor in Eq. (B.15), leaving a finite polynomial. Its coefficients depend smoothly on $(1+\tilde{n}_\Delta)^{-1}$ near zero. The fixed finite Fock support therefore gives the stated $\mathcal{O}(\tilde{n}_\Delta^{-1})$ coefficientwise error.

□

B.6.2 Infinite-energy limit

For the recovery estimate, choose the product complementary reference

$$D_{\mathrm{p}} := \rho_{\mathrm{th}}(\tilde{n}_\Delta)^{\otimes N} \otimes \mathrm{diag}(\lambda_r). \tag{B.25}$$

It has trace one and full support on the complementary space of the retained purification. Its weighted proxy is finite for sufficiently small Δ by Sec. B.6.2. At the zero lattice vector, $d_{rs,\Delta}(0) = \langle\phi_r|(1-c_2/(1+\tilde{n}_\Delta))^{\hat{n}_{\mathrm{tot}}}|\phi_s\rangle \to \delta_{rs}$, so the zero-vector block approaches D_{p}.

With the definition

$$f_\sigma(x) := \sum_{r,s} w_{rs} \Big| \langle\phi_r| \, \hat{D}(\beta(x)) \, |\phi_s\rangle \Big|^2, \qquad w_{rs} := \frac{\lambda_r \lambda_s}{(\sqrt{\lambda_r}+\sqrt{\lambda_s})^2}. \tag{B.26}$$

the infinite-energy proxy has the following form.

Proposition B.6 (Infinite-energy proxy and infidelity bound)**.** *For σ as in Eq.* (B.8)*, the isometry of Proposition B.4, and the reference of Eq.* (B.25)*, the infidelity proxy obeys*

$$\lim_{\Delta\to 0} \tilde{\epsilon}_{D_{\mathrm{p}}}(M_{\hat{V}_\Delta,\sigma}) = \sum_{x\in\Lambda^\perp\setminus\{0\}} f_\sigma(x). \tag{B.27}$$

Together with Lemma A.5, this gives $\limsup_{\Delta\to 0}(1-\tilde{F}^{\mathrm{opt}}) \le 4\sum_{x\neq 0} f_\sigma(x)$.

Higher-order terms

For simplicity, write $q_{\boldsymbol{\ell}} := (1-t)^N t^{|\boldsymbol{\ell}|_1}$ for the eigenvalues of the output thermal factor in D_{p}. Suppressing boldface on the multi-indices, define the weighted norm of a residual matrix B in the diagonal basis of D_{p} by

$$\|B\|_{D_{\mathrm{p}}}^2 := \frac{1}{d_L}\sum_{\mu,\nu,\ell r,ks} \frac{\left|B_{\mu\ell r,\nu ks}\right|^2}{(\sqrt{q_\ell\lambda_r}+\sqrt{q_k\lambda_s})^2}. \tag{B.28}$$

Thus $\tilde{\epsilon}_{D_{\mathrm{p}}}(M) = \left\|M - I_L\otimes D_{\mathrm{p}}\right\|_{D_{\mathrm{p}}}^2$.

Lemma B.7 (Suppression of higher-order terms)**.** *Under the fixed setup of Sec. B.6, let $B_{>0,\Delta}(x)$ be the contribution to M_Δ from the terms in Eq.* (B.14) *with $|\boldsymbol{i}|_1+|\boldsymbol{j}|_1>0$, including their scalar, logical, and purification-label factors. Then*

$$\left\|\sum_x B_{>0,\Delta}(x)\right\|_{D_{\mathrm{p}}} = \mathcal{O}(\tilde{n}_\Delta^{-1/2}). \tag{B.29}$$

Proof. We first remove the geometric weights from the norm, then bound the remaining creation and annihilation operators by geometric moments, and finally sum the resulting Gaussian bound over the lattice. Constants may increase between estimates and depend on the fixed channel, environment support, and block length. An integer $i_0 \ge 0$ bounds the polynomial degree in the displacement.

The arithmetic–geometric mean inequality gives

$$\frac{\lambda_r\lambda_s}{(\sqrt{q_\ell\lambda_r}+\sqrt{q_k\lambda_s})^2} \le \frac{\sqrt{\lambda_r\lambda_s}}{4\sqrt{q_\ell q_k}}. \tag{B.30}$$

Each block is controlled by a Frobenius norm with exterior weights $\rho_{\mathrm{th}}(\tilde{n}_\Delta)^{-1/4}$. In one mode, Lemma B.32 gives

$$
\begin{aligned}
&\rho_{\mathrm{th}}(\tilde{n}_\Delta)^{-1/4}\hat{R}_x\rho_{\mathrm{th}}(\tilde{n}_\Delta)^{-1/4} \\
&= \sqrt{1-t}\exp\left[\frac{c_1 n_\Delta(n_\Delta+1)|\beta(x)|^2}{(1+\tilde{n}_\Delta)^2\chi\sqrt{t}(1+\sqrt{t})}\right] \\
&\quad\times \hat{D}\left(\frac{\sqrt{c_1 n_\Delta(n_\Delta+1)}\beta(x)^*}{(1+\tilde{n}_\Delta)\sqrt{\chi}\,t^{1/4}(1+\sqrt{t})}\right) t^{\hat{n}/2}\hat{D}\left(\frac{\sqrt{c_1 n_\Delta(n_\Delta+1)}\beta(x)^*}{(1+\tilde{n}_\Delta)\sqrt{\chi}\,t^{1/4}(1+\sqrt{t})}\right).
\end{aligned}
\tag{B.31}
$$

Commuting the exterior weights through an order-$(i+j)$ term gives $t^{-(i+j)/4}$. Removing the displacement unitaries leaves binomials in the creation and annihilation operators shifted by the displacement arguments in Eq. (B.31). Using Lemma B.39, including the factor $\sqrt{1-t}$ in Eq. (B.31), therefore bounds the norm before multiplication by $\left(\frac{\sqrt{c_2}}{1+\tilde{n}_\Delta}\right)^{i+j} e^{\left(\frac{c_1 n_\Delta(n_\Delta+1)}{1+\tilde{n}_\Delta}-n_\Delta-\frac{1}{2}\right)|\beta(x)/\sqrt{\chi}|^2}$ by

$$
b_*(1+\tilde{n}_\Delta)^{(i+j)/2}(1+|\beta(x)/\sqrt{\chi}|)^{i_0}\exp\left[\frac{\frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2}|\beta(x)/\sqrt{\chi}|^2}{\sqrt{t}(1+\sqrt{t})}\right]. \tag{B.32}
$$

The explicit coupling is $\left(\frac{\sqrt{c_2}}{1+\tilde{n}_\Delta}\right)^{i+j} = \mathcal{O}((1+\tilde{n}_\Delta)^{-i-j})$, so every positive order is at most $\mathcal{O}((1+\tilde{n}_\Delta)^{-1/2})$. Combining the two exponentials gives

$$
h_\Delta := \left(\frac{c_1 n_\Delta(n_\Delta+1)}{1+\tilde{n}_\Delta}-n_\Delta-\frac{1}{2}\right)+\frac{\frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2}}{\sqrt{t}(1+\sqrt{t})} \longrightarrow -\frac{c_2}{2c_1}<0. \tag{B.33}
$$

This follows from $\frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2}\to 1/c_1$, $t\to 1$, and $\left(\frac{c_1 n_\Delta(n_\Delta+1)}{1+\tilde{n}_\Delta}-n_\Delta-\frac{1}{2}\right)\to -(c_2+1)/(2c_1)$. Thus $h_\Delta \le -b'_*$ for all sufficiently large $(1+\tilde{n}_\Delta)$. Product Frobenius norms factorize across modes, and the finite eigenvector contractions in Eq. (B.23) obey the same bound by the triangle inequality. In particular, there are $b_*>0$, $b'_*>0$, and an integer $i_0\ge 0$, independent of $(1+\tilde{n}_\Delta), x$, such that

$$
\left\|B_{>0,\Delta}(x)\right\|_{D_{\mathrm{p}}} \le b_*(1+\tilde{n}_\Delta)^{-1/2}(1+\left\|\beta(x)/\sqrt{\chi}\right\|)^{i_0}e^{-b'_*\left\|\beta(x)/\sqrt{\chi}\right\|^2}. \tag{B.34}
$$

Finally, a polynomial times a decaying Gaussian is summable on a fixed full-rank lattice,

proving Eq. (B.29). The coefficient errors $\mathcal{O}((1+\tilde{n}_\Delta)^{-1})$ in Lemma B.5, including the zero-vector discrepancy from D_{p}, satisfy the same Gaussian bound and vanish in this norm as well.

□

The zeroth-order limit

For fixed spectral labels r, s, define the weighted contraction of the zeroth-order outcome operators by

$$\mathcal{K}_\Delta(x,x') := e^{\left(\frac{c_1 n_\Delta(n_\Delta+1)}{1+\tilde{n}_\Delta} - n_\Delta - \frac{1}{2}\right)\left(\left\|\beta(x)/\sqrt{\chi}\right\|^2 + \left\|\beta(x')/\sqrt{\chi}\right\|^2\right)} \sum_{\ell,k} \frac{\lambda_r \lambda_s \prod_{j=1}^N (\hat{R}_{x,j})_{\ell_j k_j} \overline{(\hat{R}_{x',j})_{\ell_j k_j}}}{(\sqrt{q_\ell \lambda_r} + \sqrt{q_k \lambda_s})^2}. \tag{B.35}$$

The overbar denotes complex conjugation of the Fock-basis matrix element. Only equal lattice vectors $x = x'$ survive as $\Delta \to 0$.

Proposition B.8 (Zeroth-order pair limits)**.** *At fixed x, x' as $\Delta \to 0$, $\mathcal{K}_\Delta(x,x') \to 0$ if $x \neq x'$, whereas*

$$\mathcal{K}_\Delta(x,x) \left| d_{rs,\infty}(x) \right|^2 \longrightarrow w_{rs} \left| \langle \phi_r | \, \hat{D}(\beta(x)) \, | \phi_s \rangle \right|^2. \tag{B.36}$$

Proof. We sum the outcome occupations using the Hardy–Hille identity, write the remaining spectral weight as a Fourier integral, and take the limit under an integrable bound. Unequal radii are suppressed by a Gaussian, unequal angles by oscillatory cancellation, and equal vectors give the displacement-overlap kernel. Write $\rho_j := |\beta_j(x)/\sqrt{\chi}|$, $\rho'_j := |\beta_j(x')/\sqrt{\chi}|$ and $\theta_j := \arg \beta_j(x) - \arg \beta_j(x')$ when both radii are nonzero.

For $j \geq 0$ the outcome entries of Eq. (B.13) are

$$(\hat{R}_x)_{k+j,k} = (1-t)\sqrt{\frac{k!}{(k+j)!}}\left(\frac{\sqrt{c_1 n_\Delta(n_\Delta+1)}}{1+\tilde{n}_\Delta}\frac{\beta(x)^*}{\sqrt{\chi}}\right)^j t^k \times L_k^{(j)}\left(\frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2}\frac{|\beta(x)/\sqrt{\chi}|^2}{t}\right).$$

the other ordering carrying $(-\frac{\sqrt{c_1 n_\Delta(n_\Delta+1)}}{1+\tilde{n}_\Delta}\beta(x)/\sqrt{\chi})^j$. Using Lemma B.37, dividing the product of entries by $\sqrt{q_{k+j}q_k}$, and summing k, each signed difference $j \in \mathbb{Z}$ contributes

$$e^{-c_1 n_\Delta(n_\Delta+1)(\rho^2+\rho'^2)/[(1+\tilde{n}_\Delta)^2(1-t)]} I_{|j|}(z) e^{-ij\theta},$$

with $z := 2\frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2}\rho\rho'/[\sqrt{t}(1-t)]$ and $\theta := \arg\beta(x) - \arg\beta(x')$.

Use the spectral weight $w_{rs}(\omega)$ defined in Eq. (B.105). The metric factor in Eq. (2.19) splits as

$$\frac{\lambda_r\lambda_s}{(\sqrt{q_\ell\lambda_r}+\sqrt{q_k\lambda_s})^2} = \frac{w_{rs}(\ln(q_\ell/q_k))}{\sqrt{q_\ell q_k}}. \tag{B.37}$$

The Hardy–Hille sum includes $1/\sqrt{q_\ell q_k}$. The remaining weight depends on the total outcome-number difference, $\omega = (|\ell|_1 - |k|_1)\ln t$, and satisfies $w_{rs}(0) = w_{rs}$ of Proposition B.6.

Lemma B.38 gives the Fourier representation of the spectral weight and its contraction with independent Skellam variables.

Combining Eqs. (B.104) and (B.107) with $z_j = 2\frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2}\rho_j\rho'_j/[\sqrt{t}(1-t)]$, the zeroth-order pair contribution, before logical and polynomial factors, is exactly

$$\begin{aligned}\mathcal{K}_\Delta(x,x') &= e^{\varphi_\Delta}\,\mathbb{E}\Big[w_{rs}\Big(\ln t\sum_j J_j\Big)e^{-i\sum_j\theta_j J_j}\Big],\\ \varphi_\Delta &:= \sum_j\left[\left(\frac{c_1 n_\Delta(n_\Delta+1)}{1+\tilde{n}_\Delta} - n_\Delta - \frac{1}{2} - \frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2(1-t)}\right)(\rho_j^2+\rho_j'^2) + z_j\right].\end{aligned} \tag{B.38}$$

The exponent decomposes as

$$\varphi_\Delta = h_\Delta(\left\|\beta(x)/\sqrt{\chi}\right\|^2 + \left\|\beta(x')/\sqrt{\chi}\right\|^2) - \frac{\frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2}}{\sqrt{t}(1-t)} \sum_j (\rho_j - \rho'_j)^2. \tag{B.39}$$

The expectation in Eq. (B.38) is bounded by $\sqrt{\lambda_r \lambda_s}/4$ (Eq. (B.107)). If radii differ, the second term of Eq. (B.39) tends to $-\infty$ because $\frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2}/[\sqrt{t}(1-t)] = \Theta((1+\tilde{n}_\Delta))$, so $\mathcal{K}_\Delta \to 0$. If all radii agree and $\beta(x)/\sqrt{\chi} \neq \beta(x')/\sqrt{\chi}$, some coordinate of nonzero radius has $\theta_j \notin 2\pi\mathbb{Z}$ and $z_j = \Theta((1+\tilde{n}_\Delta)) \to \infty$; in Eq. (B.107) with $\tau_j = \ln t \to 0$ the exponent at each fixed k tends to $-\infty$, and ϱ_F dominates, so the integral vanishes. Coordinates of zero radius have $z_j = 0$ and need no angle.

For $\beta(x)/\sqrt{\chi} = \beta(x')/\sqrt{\chi}$ every $\theta_j = 0$, and $\left|z_j\{\cos(k \ln t) - 1\}\right| \leq z_j k^2 (\ln t)^2/2 = \mathcal{O}((1+\tilde{n}_\Delta)^{-1})$ at each fixed k, using $z_j = \mathcal{O}((1+\tilde{n}_\Delta))$ and $\ln t = \mathcal{O}((1+\tilde{n}_\Delta)^{-1})$. The exponent therefore tends to zero. Dominated convergence then gives $\sqrt{\lambda_r \lambda_s} \int \varrho_F e^{2ik\nu_{rs}} dk = w_{rs}(0)$ by Eq. (B.106). Meanwhile $\varphi_\Delta \to -\chi\left\|\beta(x)/\sqrt{\chi}\right\|^2$ by Eq. (B.24), and $\left|d_{rs,\infty}\right|^2 = e^{\chi\left\|\beta(x)/\sqrt{\chi}\right\|^2} \left|\langle\phi_r| \hat{D}(\beta(x)) |\phi_s\rangle\right|^2$ cancels that Gaussian exactly.

Every pair, including its finite polynomial coefficients, is bounded by a fixed polynomial times $e^{-b'_*(\left\|\beta(x)/\sqrt{\chi}\right\|^2 + \left\|\beta(x')/\sqrt{\chi}\right\|^2)}$, which is summable over the double-lattice indices $(x, x\prime)$, so dominated convergence passes Proposition B.8 through the lattice sum. The diagonal logical coefficient contributes $\mathrm{Tr}\left(\hat{W}_x^\dagger \hat{W}_x\right) = d_L$, cancelling the $1/d_L$ in Eq. (2.19), and subtracting $I_L \otimes D_\mathrm{p}$ removes the zero vector.

□

Remark B.9 (Gram correction)**.** *Exact Gram orthonormalization controls the exponentially small codeword overlaps. The resulting correction vanishes in the same weighted norm. Let* $\hat{H}_\Delta := \hat{G}_\Delta^{-1/2}$, $E_C := M_\Delta - I_L \otimes D_\mathrm{p}$ *and* $E_{\hat{V}} := M_{\hat{V}_\Delta} - I_L \otimes D_\mathrm{p}$. *Exactly,*

$$E_{\hat{V}} = (\hat{H}_\Delta \otimes I) E_C (\hat{H}_\Delta \otimes I) + (\hat{H}_\Delta^2 - I_L) \otimes D_\mathrm{p}. \tag{B.40}$$

The weight in Eq. (2.19) *acts on complementary indices and commutes with logical multiplication, so the congruence changes the weighted norm of the first term by a factor tending*

to one, while the second has squared weighted norm $\mathrm{Tr}\big(D_{\mathrm{p}}\big)\big\|\hat{H}_\Delta^2 - I_L\big\|_F^2/(4d_L) \to 0$ *by Proposition B.4. This proves Proposition B.6 for the physical isometry.*

Table 3.2 defines the marginal, zero-displacement, and product references used below. The product reference is written explicitly in Eq. (B.25) for the weighted-norm proof.

B.7 Closed-form performance under thermal loss

For a fixed one-mode GKP lattice of arbitrary shape and logical dimension, thermal loss has the closed-form infinite-energy proxy given below. We use the reference D_0, the zero-displacement QEC term before Gram orthonormalization (Table 3.2). Write $\mathcal{L} := \{\xi(x) : x \in \Lambda^\perp\}$ for the lattice in physical displacement coordinates and $\hat{W}_\xi := \hat{W}_x$ when $\xi = \xi(x)$. Thus $|\xi(x)|^2 = \pi|x|^2$. Throughout the fixed-code calculations below, M_Δ denotes the QEC matrix of the encoding $\hat{V}_\Delta$ for the channel under discussion.

Proposition B.10 (Fixed-lattice thermal-loss asymptote)**.** *For fixed* $0 < \eta < 1$, $0 \le \bar{n} < \infty$, *and a fixed one-mode GKP lattice,*

$$\lim_{\Delta\to 0} \tilde{\epsilon}_{D_0}(M_\Delta) = \sum_{\xi\in\mathcal{L}\setminus\{0\}} f_{\bar{n}}(\chi|\xi|^2), \qquad \chi = \frac{\eta}{1-\eta}, \tag{B.41}$$

For $t_E = \bar{n}/(1+\bar{n}) > 0$,

$$\begin{aligned} f_{\bar{n}}(v) &:= e^{-v(1+t_E)/(1-t_E)} \sum_{j\in\mathbb{Z}} I_{|j|}\left(\frac{2v\sqrt{t_E}}{1-t_E}\right) \frac{t_E^{j/2}}{(1+t_E^{j/2})^2}, \\ f_0(v) &:= \tfrac{1}{4}e^{-v}. \end{aligned} \tag{B.42}$$

Lemma A.5 consequently gives $\limsup_{\Delta\to 0}(1-\tilde{F}^{\mathrm{opt}}) \le 4\sum_{\xi\neq 0} f_{\bar{n}}(\chi|\xi|^2)$. At finite energy the exact proxy is given by Eqs. (B.49) and (B.92), as used in Sec. B.14.1.

Proof. First diagonalize the Gaussian reference while preserving its mode labels. Next contract each pair of lattice operators by a one-dimensional Fourier integral. Finally take the limit, retaining the equal-displacement terms and controlling the entire lattice sum.

The Gaussian reference and lattice operators. To construct D_0 for thermal loss, prepare the channel input in the thermal state $\rho_{\mathrm{th}}(n_\Delta)$ of Eq. (3.23), and purify the environment of mean occupation $\bar{n}$ as $\sum_r \sqrt{(1-t_E)t_E^r}\,|r\rangle_E \otimes |r\rangle_R$, where t_E is the thermal ratio from Remark B.18 (with $t_E = 0$ for vacuum). Table B.1 then gives

$$\langle \hat{f}^\dagger \hat{f}\rangle = (1-\eta)n_\Delta + \eta\bar{n}, \qquad \langle \hat{r}^\dagger \hat{r}\rangle = \bar{n}, \qquad \langle \hat{f}\otimes\hat{r}\rangle = \sqrt{\eta\bar{n}(1+\bar{n})}. \tag{B.43}$$

Here $\hat{r}$ is the purifier's annihilation operator. The first moments and all other independent anomalous or intermode moments vanish. Using Lemma B.35, the normal-mode occupations of this reference are

$$n_{1,2}^{\mathrm{th}} := \frac{1}{2}\left[\sqrt{\Big(\gamma n_\Delta + (1+\eta)\bar{n} + 1\Big)^2 - 4\eta\bar{n}(1+\bar{n})} - 1 \pm \gamma(n_\Delta - \bar{n})\right]. \tag{B.44}$$

The associated squeezing parameter satisfies $\tanh(2r) = 2\sqrt{\eta\bar{n}(1+\bar{n})}/[\gamma n_\Delta + (1+\eta)\bar{n} + 1]$. Let $t_{1,2} := n_{1,2}^{\mathrm{th}}/(1+n_{1,2}^{\mathrm{th}})$, and

$$\begin{aligned}
u_1 &:= \frac{\sqrt{\gamma n_\Delta(n_\Delta+1)}}{1+\tilde{n}_\Delta}(\cosh r - t_1\sqrt{\eta t_E}\sinh r),\\
u_2 &:= \frac{\sqrt{\gamma n_\Delta(n_\Delta+1)}}{1+\tilde{n}_\Delta}(\sqrt{\eta t_E}\cosh r - t_2\sinh r),\\
k_0 &:= -\tfrac{1}{2} - \frac{\eta n_\Delta}{(1+\tilde{n}_\Delta)} + \frac{\gamma n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2}[\sqrt{\eta t_E}\sinh(2r) - (\eta t_E t_1 + t_2)\sinh^2 r].
\end{aligned} \tag{B.45}$$

one has $D_0 = [(1-t_1){t_1}^{\hat{n}_1}] \otimes [(1-t_2){t_2}^{\hat{n}_2}]$, where $\hat{n}_j := \hat{a}_j^\dagger \hat{a}_j$ are the normal-mode number operators, and the raw lattice operator

$$\hat{O}_\xi = e^{k_0|\xi|^2}\Big(e^{u_1\xi^*\hat{a}_1^\dagger}\otimes e^{u_2\xi\hat{a}_2^\dagger}\Big)D_0\Big(e^{-u_1\xi\hat{a}_1}\otimes e^{-u_2\xi^*\hat{a}_2}\Big). \tag{B.46}$$

This follows from Gaussian integration of the comb characteristic function through the beam splitter, then two-mode squeezing and normal ordering; equivalently, both sides have the

same exponential generating function. For vacuum $t_2 = u_2 = 0$ and the second mode drops out.

Lattice-pair contraction. For vacuum, omit the second mode. Put $\theta := \arg\xi - \arg\xi\prime$ if $|\xi||\xi\prime| > 0$, with $\theta := 0$ otherwise. Define

$$c_a := \frac{u_1^2}{\sqrt{t_1}(1-t_1)}, \quad c_b := \frac{u_2^2}{\sqrt{t_2}(1-t_2)}, \quad z_a := 2c_a|\xi||\xi\prime|, \quad z_b := 2c_b|\xi||\xi\prime|, \tag{B.47}$$

$$h_\Delta := k_0 + \frac{u_1^2}{\sqrt{t_1}(1+\sqrt{t_1})} + \frac{u_2^2}{\sqrt{t_2}(1+\sqrt{t_2})}, \qquad \varphi := h_\Delta(|\xi|^2 + |\xi\prime|^2) - (c_a + c_b)(|\xi| - |\xi\prime|)^2. \tag{B.48}$$

Define $K_\Delta(\xi,\xi') := \sum_{a'b'}(\hat{O}_\xi)^*_{a'b'}(\hat{O}_{\xi'})_{a'b'}/(\sqrt{(D_0)_{a'}} + \sqrt{(D_0)_{b'}})^2$. Then

$$K_\Delta(\xi,\xi') = e^\varphi \int_{\mathbb{R}} \frac{2k\,dk}{\sinh(2\pi k)} \exp\Big[- 2z_a \sin^2 \tfrac{k\ln t_1 - \theta}{2} - 2z_b \sin^2 \tfrac{k\ln t_2 + \theta}{2}\Big]. \tag{B.49}$$

The kernel is real and symmetric; logical lattice phases remain in $\hat{W}_\xi$. Using Lemma B.37 for the normal modes in Eq. (B.46), the trace-one prefactors of D_0 cancel the Hardy–Hille denominators, giving independent Skellam variables J, I with parameters z_a, z_b and weight $1/\{4\cosh^2[(J\ln t_1 + I\ln t_2)/4]\}$ times $e^{-i(J-I)\theta}$. Using Lemma B.38, this contraction has the Fourier representation in Eq. (B.49), with opposite normal-mode angles. Equation (B.49) resums both Bessel-order sums into one Fourier integral.

The fixed-lattice limit. For fixed $\bar{n} > 0$ the preceding parameters obey

$$\begin{gathered} t_1 \to 1, \quad n_\Delta(1-t_1) \to \gamma^{-1}, \quad t_2 \to t_E, \quad u_1^2 \to \gamma^{-1}, \quad u_2^2 \to \chi t_E, \quad k_0 \to -\chi - \tfrac{1}{2}, \\ c_a \sim n_\Delta, \qquad c_b \to \frac{\chi\sqrt{t_E}}{1-t_E}, \qquad h_\Delta \to h_\infty := -\frac{\chi(1-\sqrt{t_E})}{2(1+\sqrt{t_E})} < 0. \end{gathered} \tag{B.50}$$

The first-mode factor in Eq. (B.49) is $\exp\big[-c_a|\xi - e^{ik\ln t_1}\xi'|^2\big]$. For each fixed k it tends to zero if $\xi \neq \xi'$, whereas for $\xi = \xi'$ its exponent tends to zero because $c_a(\ln t_1)^2 = \mathcal{O}(n_\Delta^{-1})$. The

density $\varrho_F(k) = 2k/\sinh(2\pi k)$ is integrable, so dominated convergence removes all unequal displacements. For an equal pair with $v = \chi|\xi|^2$, the remaining integral is

$$e^{-v(1+t_E)/(1-t_E)} \int_{\mathbb{R}} \varrho_F(k) \exp\left[\frac{2v\sqrt{t_E}}{1-t_E} \cos(k \ln t_E)\right] dk. \tag{B.51}$$

Expanding $e^{z\cos\theta} = \sum_{j\in\mathbb{Z}} I_{|j|}(z)e^{ij\theta}$ and applying Eq. (B.106) gives Eq. (B.42). The Bessel coefficients have finite sum e^z, justifying this interchange. In vacuum the second mode is absent and $\int \varrho_F = 1/4$.

It remains to include Gram orthonormalization and the infinite sum. Since $\hat{W}_\xi$ is unitary, the comb formula gives

$$\|\hat{G}_\Delta - I_L\| \leq \sum_{\xi\in\mathcal{L}\setminus\{0\}} e^{-(n_\Delta+1/2)|\xi|^2} = \mathcal{O}(e^{-cn_\Delta})$$

for some $c > 0$. Hence $\hat{H}_\Delta = \hat{G}_\Delta^{-1/2} \to I_L$; the coefficients in Eq. (B.92) tend to $\hat{W}_\xi$ for $\xi \neq 0$, while $B_0 \to 0$. Their norms are uniformly bounded at sufficiently large n_Δ, and $\mathrm{Tr}\big(\hat{W}_\xi^\dagger \hat{W}_\xi\big)/d_L = 1$. Finally $|K_\Delta(\xi,\xi')| \leq e^{h_\Delta(|\xi|^2+|\xi'|^2)}/4$ and $h_\infty < 0$ supply a summable Gaussian dominator on $\mathcal{L}\times\mathcal{L}$. Dominated convergence therefore yields Eq. (B.41).

□

Remark B.11 (Relation to pure loss and amplification)**.** *For vacuum loss, Eq.* (B.41) *reduces to*

$$\frac{1}{4} \sum_{x\in\Lambda^\perp\setminus\{0\}} e^{-\pi\chi|x|^2},$$

the Gaussian dual-lattice expression in Ref. [13]. For quantum-limited amplification the same expression holds with $\chi = G/(G-1)$, as follows either from its vacuum QEC operator in Eq. (B.13) *or the amplification treatment in that reference. The present thermal kernel extends this fixed-lattice proxy expression to $\bar{n} > 0$; the exact Gram correction is retained before taking the limit.*

B.8 Vacuum specializations of the general expansion

The main-text vacuum formulas are obtained by setting $\sigma = |0\rangle\langle 0|$ in the ordered expansion. Only one environmental eigenvector remains, so the purification-label factor is one-dimensional. In the coefficient extraction of Eq. (B.22), $r = s = 0$ forces $i = j = 0$, and $d_{00,\Delta}(x) = 1$. Thus all higher-order environmental terms disappear exactly.

Corollary B.12 (Vacuum QEC matrix)**.** *For vacuum loss or amplification, the raw finite-energy QEC matrix is*

$$M_\Delta^{\mathrm{raw}} = \sum_{x\in\Lambda^\perp} \hat{W}_x \otimes \exp\left[\left(\frac{c_1 n_\Delta(n_\Delta+1)}{1+\tilde{n}_\Delta} - n_\Delta - \frac{1}{2}\right)\|\beta(x)/\sqrt{\chi}\|^2\right]\hat{R}_x, \tag{B.52}$$

where $\hat{R}_x$ is the ordered thermal operator of Eq. (B.13)*. The normalized-isometry QEC matrix is obtained by logical congruence with $\hat{G}_\Delta^{-1/2}$.*

Proof. Set the environmental Fock indices to zero in Lemma B.5. The generating-function factor involving those indices has constant coefficient one, giving the displayed expression. Gram orthonormalization follows from Proposition B.4. □

For one mode, rewriting $\hat{R}_x$ as same-sided displacements using Lemma B.32, and resolving $\hat{W}_x$ in the logical basis, yields Eq. (3.36) after the individual codeword normalizations used in Ref. [13]. That form keeps an exponentially small normalization correction. Exact finite-energy calculations instead apply the full Gram congruence.

The vacuum Kraus-product expression in Eq. (3.35) is likewise obtained from Eq. (B.17) by inserting the vacuum environmental characteristic function. For attenuation or amplification, the coefficients are those of Table B.1. The index transpose and complex conjugation are changed together when translating between its Fock-outcome convention and the convention of Ref. [13].

Corollary B.13 (Vacuum environmental kernel)**.** *For a vacuum environment,* $D_{\mathrm{p}} = D_0$ *and*

$$f_0(x) = \frac{1}{4} e^{-\chi \|\xi(x)\|^2}, \qquad \lim_{\Delta \to 0} \tilde{\epsilon}_{D_{\mathrm{p}}}(M_\Delta) = \frac{1}{4} \sum_{x \in \Lambda^\perp \setminus \{0\}} e^{-\pi \chi \|x\|^2}. \tag{B.53}$$

Proof. For $\rho_E = |0\rangle\langle 0|$, the spectral weight is $1/4$ and $|\langle 0|\hat{D}(\beta)|0\rangle|^2 = e^{-|\beta|^2}$. Substitute this into Proposition B.6 and use $\|\xi(x)\|^2 = \pi \|x\|^2$. □

At finite energy the vacuum specialization of Eqs. (B.49) and (B.92) retains every lattice pair and the Gram correction. To recover the nearest-shell perturbative expression of Eq. (3.39), use the diagonal marginal-reference approximation of Ref. [13], keep the nearest logical displacement pairs, and contract the thermal Fock sums by the Hardy–Hille identity (Lemma B.37). The square and hexagonal cases have four and six nearest vectors, with squared lengths $1/d_L$ and $2/(\sqrt{3}d_L)$, respectively. These give the different exponents and prefactors plotted in Fig. 3.4. The full finite-energy proxy calculation above retains the terms discarded by this nearest-shell approximation.

B.9 Energy monotonicity of the full infidelity proxy

At sufficiently large energy, the sign of the full D_0 proxy's derivative is fixed by its first nonzero inverse-energy coefficient. This includes cross-displacement terms and Gram orthonormalization and also orders the certified infidelity bound $4\tilde{\epsilon}_{D_0}$.

Proposition B.14 (Eventual energy monotonicity)**.** *For vacuum loss and quantum-limited amplification, define*

$$\epsilon_\infty := \frac{1}{4} \sum_{\xi \in \mathcal{L} \setminus \{0\}} e^{-\chi|\xi|^2}, \qquad a_{\mathcal{L}} := \frac{8c_2 - 3}{32 c_1^2} \sum_{\xi \in \mathcal{L} \setminus \{0\}} |\xi|^2 e^{-\chi|\xi|^2}. \tag{B.54}$$

using c_1, c_2, χ *from Table B.1.*

For a fixed one-mode GKP lattice under vacuum loss ($0 < \eta < 1$) *or quantum-limited*

amplification ($G > 1$),

$$\tilde{\epsilon}_{D_0}(M_\Delta) = \epsilon_\infty + \frac{a_\mathcal{L}}{n_\Delta} + \mathcal{O}(n_\Delta^{-2}), \qquad \frac{d\tilde{\epsilon}_{D_0}(M_\Delta)}{dn_\Delta} = -\frac{a_\mathcal{L}}{n_\Delta^2} + \mathcal{O}(n_\Delta^{-3}). \tag{B.55}$$

Thus the proxy eventually decreases for vacuum loss with $\eta > 3/8$ and for every quantum-limited amplifier, and eventually increases for vacuum loss with $\eta < 3/8$. For each fixed lattice and $\eta > 3/8$, eventual decrease also holds for thermal loss with all sufficiently small fixed $\bar{n} > 0$.

Energy thresholds and the thermal-occupation interval may depend on the fixed channel and lattice. At $\eta = 3/8$, $a_\mathcal{L} = 0$, so the derivative's sign requires the next nonzero coefficient.

Proof. We expand the equal-displacement contractions and their derivatives under a summable bound. We then control unequal displacements and Gram corrections. Continuity of the leading coefficient gives the thermal extension.

Equal-displacement expansion. For either vacuum channel, the zeroth-order operator of Eq. (B.13) is exact, $D_0 = (1-t)t^{\hat{n}}$, and the pair kernel takes the form

$$\begin{aligned} K_\Delta(\xi, \xi') &= e^{h_\Delta(|\xi|^2 + |\xi'|^2)} \int_\mathbb{R} \varrho_F(k) e^{-c_a|\xi - e^{ik\ln t}\xi'|^2}\, dk, \\ h_\Delta &= \frac{c_1 n_\Delta(n_\Delta+1)}{1+\tilde{n}_\Delta} - n_\Delta - \frac{1}{2} + \frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2\sqrt{t}(1+\sqrt{t})}, \\ c_a &= \frac{c_1 n_\Delta(n_\Delta+1)}{(1+\tilde{n}_\Delta)^2\sqrt{t}(1-t)}. \end{aligned} \tag{B.56}$$

For amplification one may use the reflected coordinates $\beta(x)/\sqrt{\chi} = -\xi(x)^*$ of Eq. (3.22); this is a fixed isometry and leaves all norms and the argument below unchanged. The parameters of Eq. (B.10) give

$$h_\Delta = -\frac{\chi}{2} + \frac{4c_2 - 1}{8c_1^2 n_\Delta} + \mathcal{O}(n_\Delta^{-2}), \qquad c_a(\ln t)^2 = \frac{1}{c_1^2 n_\Delta} + \mathcal{O}(n_\Delta^{-2}). \tag{B.57}$$

For loss, $t = t_1$ and $c_a = (n_\Delta + 1)\sqrt{t_1}$; for amplification, $(1 + \tilde{n}_\Delta) = (G-1)n_\Delta + G$ and $c_a = n_\Delta\sqrt{t}$. Together with $\int \varrho_F = 1/4$ and $\int k^2 \varrho_F = 1/32$, expansion of the diagonal integral

gives

$$K_\Delta(\xi,\xi) = \frac{1}{4}e^{-\chi|\xi|^2} + \frac{(8c_2-3)|\xi|^2 e^{-\chi|\xi|^2}}{32c_1^2 n_\Delta} + R_\Delta(\xi). \tag{B.58}$$

The remainder is bounded by $b_* q(|\xi|)e^{-c|\xi|^2}n_\Delta^{-2}$, and its n_Δ derivative by $b_* q(|\xi|)e^{-c|\xi|^2}n_\Delta^{-3}$, for a fixed polynomial q and positive constants b_*, c. These bounds follow from the second-order exponential and cosine remainders and their differentiated versions. All Fourier moments are finite, and $h_\Delta \to -\chi/2 < 0$ absorbs polynomial displacement factors into a summable Gaussian. Thus the expansion can be summed and differentiated over the lattice.

Unequal displacements and Gram orthonormalization. Let $\ell > 0$ be the minimum lattice separation. On $|\xi|, |\xi\prime|, |k| \le n_\Delta^{1/8}$, the rotation $e^{ik\ln t}$ moves $\xi\prime$ by $\mathcal{O}(n_\Delta^{-3/4})$, so unequal vectors remain at least $\ell/2$ apart and their contractions are exponentially small since $c_a \sim n_\Delta$. Outside this region, the Gaussian lattice envelope and exponential Fourier tail decay faster than every inverse power of n_Δ. Differentiation adds only polynomial factors, preserving these bounds. The Gram correction and its derivative are exponentially small by the comb formula. Consequently unequal pairs and Gram corrections contribute $\mathcal{O}(n_\Delta^{-m})$ for every fixed $m > 0$, including after differentiation. Summing Eq. (B.58) gives Eq. (B.55), and the sign of $a_{\mathcal{L}}$ is that of $8c_2 - 3$. Differentiating the same Gram-corrected overlap as in Eq. (3.14) gives $d\bar{n}_C/dn_\Delta = 1 + \mathcal{O}(n_\Delta^2 e^{-cn_\Delta}) > 0$ at large energy, so the eventual ordering also holds in actual mean photons.

Thermal loss. At fixed $0 < t_E < 1$ the second normal mode survives. Its first-order parameters, obtained directly from the Gaussian-reference calculation in Sec. B.7, are

Write $h^{(1)} := [4\eta(1+t_E) - 4\eta^2\sqrt{t_E} - (1+\sqrt{t_E})^2]/[8\gamma^2(1+\sqrt{t_E})^2]$, $c_b^{(0)} := \chi\sqrt{t_E}/(1-t_E)$, and $c_b^{(1)} := -\eta(\eta+2)\sqrt{t_E}(1+t_E)/[2\gamma^2(1-t_E)^2]$. Then

$$\begin{aligned} h_\Delta &= h_\infty + \frac{h^{(1)}}{n_\Delta} + \mathcal{O}(n_\Delta^{-2}), & c_b &= c_b^{(0)} + \frac{c_b^{(1)}}{n_\Delta} + \mathcal{O}(n_\Delta^{-2}), \\ \ln t_2 &= \ln t_E - \frac{\eta}{\gamma n_\Delta} + \mathcal{O}(n_\Delta^{-2}), & c_a(\ln t_1)^2 &= \frac{1}{\gamma^2 n_\Delta} + \mathcal{O}(n_\Delta^{-2}). \end{aligned} \tag{B.59}$$

The same estimates as in the vacuum calculation yield Eq. (B.55) with $\epsilon_\infty = \sum_{\xi\neq 0} f_{\bar{n}}(\chi|\xi|^2)$ and the coefficient

$$a_{\mathcal{L}}(\eta,\bar{n}) = \sum_{\xi\in\mathcal{L}\setminus\{0\}} |\xi|^2 e^{2h_\infty|\xi|^2} \int_{\mathbb{R}} \varrho_F(k) e^{-2c_b^{(0)}|\xi|^2[1-\cos(k\ln t_E)]} \times \left[2h^{(1)} - \frac{k^2}{\gamma^2} - 2c_b^{(1)}[1-\cos(k\ln t_E)] + \frac{2\eta c_b^{(0)}}{\gamma} k\sin(k\ln t_E)\right] dk. \tag{B.60}$$

For example, the last term follows by differentiating $1-\cos(k\ln t_2)$, including the change of t_2. The extra second-mode exponential is at most one and its parameter derivatives introduce only polynomial factors, so the lattice, Fourier, and derivative bounds remain valid at every fixed temperature.

As $t_E \to 0$, $c_b^{(0)}, c_b^{(1)} \to 0$ and $h^{(1)} \to (4\eta-1)/(8\gamma^2)$. Although $\ln t_E$ diverges, its sine and cosine are bounded; the terms they multiply vanish, with at most a factor $|k|$. Also h_∞ stays strictly negative uniformly near vacuum. Dominated convergence in the integral and lattice sum therefore gives $a_{\mathcal{L}}(\eta,\bar{n}) \to a_{\mathcal{L}}(\eta,0)$. For $\eta > 3/8$ this limit is positive, proving the claimed small-temperature interval. More generally, positivity or negativity of Eq. (B.60) is the sufficient sign test at any fixed temperature.

□

For thermal loss, a nonzero coefficient in Eq. (B.60) determines the eventual sign. The monotonicity result concerns the proxy and its certified infidelity bound; actual recovery fidelity is evaluated separately below.

The decreasing-proxy interval includes $3/8 < \eta \leq 1/2$, where vacuum loss is antidegradable [201]. This decrease does not establish a zero infidelity floor. Antidegradability forbids asymptotically vanishing actual infidelity for logical dimension $d_L > 1$; the limiting proxy ϵ_∞ is distinct from that actual infidelity floor.

Remark B.15 (Interpretation of proxy monotonicity)**.** *The large-energy proxy is a useful indicator of recovery performance and explains the improving-energy trends reported for GKP*

codes in Ref. [13]. Proposition B.14 orders the proxy and its certified upper bound, not optimal fidelity. Quantitative identification with near-optimal infidelity requires the local marginal-reference hypotheses of Eq. (2.23)*; proxy monotonicity does not imply universal strict monotonicity of optimal fidelity.*

B.10 Achievable rates

Theorem B.16 (Operational coding bound)**.** *Let ρ_E be any one-mode state with $S_{\mathrm{bits}}(\rho_E) < \infty$ and let χ be as in Table B.1. For every rate satisfying*

$$0 < R < \log_2 \chi - S_{\mathrm{bits}}(\rho_E),$$

there are Gram-orthonormalized GKP isometries with finite-energy envelopes $\hat{V}_N : \mathbb{C}^{d_{L,N}} \to \mathcal{H}^{\otimes N}$ and trace-preserving recoveries $\mathcal{R}_N$ with

$$\frac{\log_2 d_{L,N}}{N} \to R, \qquad F_e\Big(\mathcal{R}_N \circ \mathcal{N}_{\rho_E}^{\otimes N} \circ \mathcal{E}_N\Big) \to 1, \tag{B.61}$$

where $\mathcal{E}_N(\rho) := \hat{V}_N \rho \hat{V}_N^\dagger$. Hence the achievable rate is at least $[\log_2 \chi - S_{\mathrm{bits}}(\rho_E)]_+$.

The proof combines an entropy-controlled environment approximation (Sec. B.10.1) with lattice selection at the required logical dimension (Sec. B.10.2). Proposition B.6 then gives a finite-energy code, and fidelity transfer returns it to the original channel (Sec. B.10.3).

B.10.1 Spectral cost and environment approximation

The N-mode environment can be approximated in trace norm by a finite-Fock state whose integrated kernel has entropy exponent S. Here f_σ is the kernel in Eq. (B.26).

Proposition B.17 (Entropy-controlled environment approximation)**.** *Let ρ_E have finite entropy S in nats. For every $\delta > 0$ and $\varepsilon_N > 0$, and all sufficiently large N, there are finite-Fock-support N-mode states σ'_N and retained probabilities $p_N \to 1$ such that*

$$\left\| \sigma'_N - \rho_E^{\otimes N} \right\|_1 \le 2(1 - p_N + \varepsilon_N), \tag{B.62}$$

$$\int_{\mathbb{R}^{2N}} f_{\sigma'_N}(x)\, d^{2N}x \le \frac{1}{4} e^{N(S+\delta-\ln\chi)}. \tag{B.63}$$

Proof. We first integrate the displacement overlaps. We restrict the spectrum to a typical set, then approximate its eigenvectors in a finite Fock subspace while preserving the retained eigenvalues. *Integrating the displacement overlaps.* Using Weyl orthogonality and $d^{2N}\beta = \pi^N \chi^N\, d^{2N}x$ from Eq. (3.22), for any state $\sigma = \sum_r \lambda_r \left|\phi_r\right\rangle \left\langle\phi_r\right|$ we obtain

$$\int f_\sigma(x)\, d^{2N}x = \chi^{-N} \sum_{r,s} \frac{\lambda_r \lambda_s}{(\sqrt{\lambda_r} + \sqrt{\lambda_s})^2} \le \frac{\chi^{-N}}{4} \Big(\sum_r \sqrt{\lambda_r} \Big)^2. \tag{B.64}$$

Tonelli's theorem interchanges the nonnegative sum and integral, including when the common value is infinite. The kernel is integrable exactly when $\sum_{r,s} w_{rs} < \infty$. Lattice averaging enters separately through Eq. (B.67). Rotating the eigenvectors changes the shape of the kernel but leaves its total integral unchanged.

Typical restriction. Write the nonzero eigenvalues of ρ_E as λ_i. For a product label $r = (r_1, \ldots, r_N)$, put $\lambda_r := \prod_j \lambda_{r_j}$ and let $\left|\phi_r\right\rangle$ be the corresponding product eigenvector of $\rho_E^{\otimes N}$. Define

$$T_N := \{r : -\ln \lambda_r \le N(S+\delta)\}, \quad p_N := \sum_{r \in T_N} \lambda_r, \quad K_N := |T_N|, \quad \sigma_N := \sum_{r \in T_N} \frac{\lambda_r}{p_N} \left|\phi_r\right\rangle \left\langle\phi_r\right|. \tag{B.65}$$

Thus $K_N = \operatorname{rank} \sigma_N$ is the dimension of the typical environmental subspace. The spectrum-preserving approximation below also has rank K_N. The law of large numbers for the integrable random variable $-\ln \lambda_i$ under the probabilities λ_i gives $p_N \to 1$, including countably infinite spectra. Each retained eigenvalue is at least $e^{-N(S+\delta)}$, so $K_N \le e^{N(S+\delta)}$. For sufficiently large N, $p_N > 0$, and orthogonality of retained and discarded eigenspaces gives

$$\left\| \rho_E^{\otimes N} - \sigma_N \right\|_1 = 2(1 - p_N). \tag{B.66}$$

The normalized retained spectrum has rank $K_N \le e^{N(S+\delta)}$, which bounds its integrated

kernel.

Finite-Fock approximation. For fixed N, let $\hat{\Phi}$ have the retained eigenvectors as columns, and let $\hat{\Pi}_L$ be the projections onto $\mathrm{span}\{|\boldsymbol{m}\rangle : 0 \le m_j \le L\}$. They converge strongly to I, meaning $\left\|\hat{\Pi}_L\psi - \psi\right\| \to 0$ for each vector ψ. Lemma B.34 gives an isometry $\hat{\Psi}_L$ with columns in this cutoff space and $\left\|\hat{\Psi}_L - \hat{\Phi}\right\| \to 0$. For each fixed block length N, choose the Fock cutoff L_N so that $\left\|\hat{\Psi}_{L_N} - \hat{\Phi}\right\| \le \varepsilon_N$, and put $\sigma'_N := \hat{\Psi}_{L_N}\,\mathrm{diag}(\lambda_r/p_N)\hat{\Psi}^\dagger_{L_N}$. It has finite Fock support and exactly the normalized typical spectrum. The same lemma gives $\|\sigma'_N - \sigma_N\|_1 \le 2\varepsilon_N$. Together with Eq. (B.66), this proves Eq. (B.62).

Finally, writing $\lambda'_r := \lambda_r/p_N$ on T_N, Cauchy–Schwarz gives $(\sum_r \sqrt{\lambda'_r})^2 \le K_N \sum_r \lambda'_r = K_N$. Equation (B.64) proves Eq. (B.63). The approximation preserves the retained spectrum and may introduce correlations, as allowed in Proposition B.6.

□

Remark B.18 (Typicality and the thermal example)**.** *The tolerance δ controls the retained rank exponent $S + \delta$, while ε_N controls the finite-Fock approximation of its eigenvectors.*

For a thermal environment with $\bar{n} > 0$, put $t_E := \bar{n}/(1+\bar{n})$. The product eigenvalues obey $-\ln\lambda_r = -N\ln(1-t_E) + |r|_1 \ln(1/t_E)$, so the typical set in Eq. (B.65) *is simply the total-photon-number cutoff $|r|_1 \le \lfloor N(\bar{n} + \delta_{\mathrm{ph}})\rfloor$, where $\delta_{\mathrm{ph}} := \delta/\ln(1/t_E)$. This subspace has finite Fock support, so $\sigma'_N = \sigma_N$. For vacuum, the retained state is $|0\rangle\langle 0|$.*

B.10.2 Lattice averaging

Lattice averaging converts the integrated kernel into a bound on the dual-lattice sum in Proposition B.6. The symplectic mean-value formula recalled by Kelmer and Yu [130] covers nonradial kernels, including those from arbitrary environmental eigenvectors. An index-m construction then sets the total logical dimension to a prime m, making the achievable rate continuous in the large-block limit. For radial kernels, the Buser–Sarnak route used in earlier GKP rate proofs [13, 43, 128] gives the same bound (Remark B.23).

For a nonnegative kernel f, its lattice sum on Λ is $S_f(\Lambda) := \sum_{x\in\Lambda\setminus\{0\}} f(x)$.

Lemma B.19 (Symplectic Siegel first moment; [130])**.** *With Haar probability on symplectic self-dual lattices $\Lambda_0 \subset \mathbb{R}^{2N}$ and nonnegative integrable f,*

$$\mathbb{E}_{\Lambda_0} S_f(\Lambda_0) = \int_{\mathbb{R}^{2N}} f(x)\, d^{2N}x. \tag{B.67}$$

Definition B.20 (Index-m sublattice)**.** *A full-rank sublattice $\Lambda' \subseteq \Lambda$ has index m when $|\Lambda/\Lambda'| = m$. Equivalently,* $\mathrm{covol}(\Lambda') = m\,\mathrm{covol}(\Lambda)$.

Lemma B.21 (Prime-index sublattice identities)**.** *Let $\Lambda_0 \subset \mathbb{R}^{2N}$ be symplectic self-dual and $\Lambda' \subseteq \Lambda_0$ have prime index m. Then Λ' is symplectically integral, has covolume and GKP logical dimension m, and satisfies $m\Lambda_0 \subseteq \Lambda' \subseteq \Lambda_0, \Lambda_0 \subseteq \Lambda'^{\perp} \subseteq m^{-1}\Lambda_0, \Lambda'^{\perp}/\Lambda_0 \cong \mathbb{F}_m, \mathrm{covol}(\Lambda'^{\perp}) = m^{-1}$. For a uniformly random index-m sublattice and any nonnegative f, with $q := (m-1)/(m^{2N}-1)$,*

$$\mathbb{E}_{\mathrm{sub}}\Big[S_f(\Lambda'^{\perp}) \mid \Lambda_0\Big] = (1-q)S_f(\Lambda_0) + q\, S_{f(\cdot/m)}(\Lambda_0). \tag{B.68}$$

Proof. The quotient Λ_0/Λ' has prime order m, so $m\Lambda_0 \subseteq \Lambda'$. Taking symplectic duals gives the dual inclusions. Since Λ_0 is self-dual, it has covolume one; the index and dual-covolume identities then give the stated covolumes. Integrality follows from $\Lambda' \subseteq \Lambda_0 \subseteq \Lambda'^{\perp}$, and the GKP dimension equals the covolume of a symplectically integral lattice.

The quotient $\Lambda'^{\perp}/\Lambda_0$ has order m and is a line in $(m^{-1}\Lambda_0)/\Lambda_0 \cong \mathbb{F}_m^{2N}$. Taking the dual gives a bijection between index-m sublattices and these lines. There are $(m^{2N}-1)/(m-1)$ lines, and each nonzero residue belongs to exactly one, giving inclusion probability q. Vectors in Λ_0 are always included. Summing over the residue classes and using $S_f(m^{-1}\Lambda_0) = S_{f(\cdot/m)}(\Lambda_0)$ proves Eq. (B.68).

□

Uniform scaling $\sqrt{\lambda}\Lambda_0$ requires integer λ for symplectic integrality and gives $d_L = \lambda^N$. By contrast, index-m sublattices give $d_L = m$. Choosing a prime $2^{NR} \le m_N \le 2^{NR+1}$ yields $R \le N^{-1}\log_2 m_N \le R + 1/N$, removing the integer-floor restriction in the large-block limit.

The next proposition bounds the dual-lattice sum at this logical dimension.

Subramanian *et al.* [129] remove the floor by concatenating square GKP qudits with an outer $[[N, K]]_d$ stabilizer code, giving $d_L = d^K$ and rate $(K/N)\log_2 d$. Their pure-loss proof uses finite-field code averaging and Gaussian theta-function estimates; the lattice bound has a remainder controlled as $N, d \to \infty$. Their displacement-noise construction instead uses quantum polar codes with analog syndrome information. Here finite-field line counting is combined with a Haar average over self-dual base lattices, giving Eq. (B.69) exactly for every nonnegative integrable kernel. Both approaches tune the total block dimension; our distinction is this ensemble and its applicability to general environmental kernels. Their Gaussian theta-function estimates do not directly extend to the nonradial displacement-overlap weights of arbitrary mixed environments.

Proposition B.22 (Lattice selection at prime logical dimension)**.** *For a nonnegative integrable $f : \mathbb{R}^{2N} \to \mathbb{R}$ and a prime m, choose a Haar-random symplectic self-dual lattice Λ_0 and a uniformly random index-m sublattice $\Lambda' \subseteq \Lambda_0$. Then*

$$\mathbb{E}\, S_f(\Lambda'^{\perp}) = m \int_{\mathbb{R}^{2N}} f(x)\, d^{2N}x. \tag{B.69}$$

Every Λ' encodes logical dimension m, and some choice satisfies $S_f(\Lambda'^{\perp}) \le m \int f$.

Proof. Using Lemmas B.21 and B.19, and $\int f(x/m)\, d^{2N}x = m^{2N} \int f$, we obtain

$$\mathbb{E}S_f(\Lambda'^{\perp}) = [1 - q + qm^{2N}] \int f = m \int f, \tag{B.70}$$

since $1 + q(m^{2N} - 1) = m$. Nonnegativity permits exchanging sums and averages. Some lattice sum is no larger than its mean. The logical dimension follows from Lemma B.21.

□

Remark B.23 (A radial-only alternative)**.** *For the thermal cutoff in Remark B.18, the retained state is a function of total photon number. Let P_k project onto total photon number k and let λ'_k*

be the normalized eigenvalue on that shell. Its kernel is $f_T(\beta) := \sum_{k,h} w(\lambda'_k, \lambda'_h) \left\| P_k \hat{D}(\beta) P_h \right\|_F^2$, *summed over retained shells. Passive mode rotations commute with each* P_k *and send* $\hat{D}(\beta)$ *to* $\hat{D}(U\beta)$. *Unitary invariance of the Frobenius norm therefore gives* $f_T(U\beta) = f_T(\beta)$. *Since* $U(N)$ *acts transitively on spheres in* $\mathbb{C}^N$, *the kernel remains radial after truncation. Vacuum has the radial kernel* $e^{-\|\beta\|^2}/4$.

The same radiality holds for any retained state invariant under passive $U(N)$ *rotations. In contrast, independent phase invariance can leave an anisotropic block kernel, such as* $e^{-\|\beta\|^2} \prod_j (1 - |\beta_j|^2)^2$ *for a product of one-photon states with* $N > 1$.

For radial f, *the rotationally invariant Minkowski–Hlawka/Buser–Sarnak statement [43, 128] applied to the radial test function* $(1-q)f(\cdot) + qf(\cdot/m)$ *supplies a self-dual* Λ_0 *for which the conditional average in Eq.* (B.68) *is at most* $[1 - q + qm^{2N}] \int f = m \int f$. *Choosing a sublattice no larger than that average gives the same conclusion.*

B.10.3 Proof of the coding theorem

Proof of Theorem B.16. We approximate the environment, select a lattice, and choose its finite-energy envelope. The global bound supplies a recovery for the approximating channel. Trace-norm contractivity then transfers its fidelity to the original channel. Fix $0 < R < \log_2 \chi - S_{\text{bits}}(\rho_E)$ and choose $\delta > 0$ with $a_* := \ln \chi - S - \delta - R \ln 2 > 0$.

Environment approximation. Apply Proposition B.17 with fixed δ and $\varepsilon_N \to 0$ to obtain σ'_N satisfying Eqs. (B.62) and (B.63).

Lattice selection. Bertrand's postulate guarantees a prime between x and $2x$ for $x \geq 2$. Choose $2^{NR} \leq m_N \leq 2^{NR+1}$; then $N^{-1} \log_2 m_N \to R$. Apply Proposition B.22 to $f_{\sigma'_N}$ and use Eq. (B.63) to obtain a symplectically integral Λ_N with $d_{L,N} = m_N$ and

$$s_N := \sum_{x \in \Lambda_N^\perp \setminus \{0\}} f_{\sigma'_N}(x) \leq \tfrac{1}{2} e^{-N a_*}. \tag{B.71}$$

Finite energy. With N, σ'_N, its support, and Λ_N now fixed, Proposition B.6 permits choosing $0 < \Delta_N \leq 1/N$ small enough that the physical isometry $\hat{V}_N$ satisfies $\tilde{\epsilon}_{D_{\text{p}}^{(N)}}(M_{\hat{V}_N, \sigma'_N}) \leq$

$s_N + e^{-N}$. This choice may depend on the block length, environment approximation, and lattice. Each code has finite energy, and $\Delta_N \to 0$.

Recovery. The trace-one reference $D_{\mathrm{p}}^{(N)}$ has full complementary support and a finite proxy by the finite-energy choice, so it is admissible. Lemma A.5 gives a Petz recovery $\mathcal{R}_N$ with $1 - F_e^{\mathrm{approx}} \leq 4(s_N + e^{-N})$.

Fidelity transfer. Apply the same $\hat{V}_N, \mathcal{R}_N$ to the true channel. The common dilation and trace-norm contractivity bound the output-state trace distance by $\left\|\rho_E^{\otimes N} - \sigma'_N\right\|_1$. Entanglement fidelity is the probability of the maximally entangled projector, so Eq. (B.62) gives

$$|F_e^{\mathrm{true}} - F_e^{\mathrm{approx}}| \leq \tfrac{1}{2}\left\|\rho_E^{\otimes N} - \sigma'_N\right\|_1 \leq (1 - p_N) + \varepsilon_N. \tag{B.72}$$

Combining this with the recovery bound gives

$$1 - F_e^{\mathrm{true}} \leq 4(s_N + e^{-N}) + (1 - p_N) + \varepsilon_N \longrightarrow 0. \tag{B.73}$$

Since $\log_2 d_{L,N}/N \to R$ and R was arbitrary below the threshold, the achievable rate is at least $[\log_2 \chi - S_{\mathrm{bits}}]_+$, the case of a nonpositive threshold being covered by a one-dimensional logical space. □

B.10.4 Gaussian-channel corollaries

A thermal bath has entropy $S_{\mathrm{bits}}(\rho_E) = g_2(\bar{n})$, where $g_2(n) := (n+1)\log_2(n+1) - n\log_2 n$ with $g_2(0) = 0$. Theorem B.16 gives the thermal-channel rates below. The canonical classification in Ref. [201] reduces a single-mode Gaussian channel, up to input and output Gaussian unitaries, to attenuator/amplifier, additive-noise, or singular canonical forms. Receiver postprocessing and these unitaries give the remaining cases.

Corollary B.24 (Thermal loss and amplification)**.** $R_{\mathrm{loss}} \geq [\log_2 \frac{\eta}{1-\eta} - g_2(\bar{n})]_+$ *and* $R_{\mathrm{amp}} \geq [\log_2 \frac{G}{G-1} - g_2(\bar{n})]_+$.

Lim *et al.* [131], Sec. IV, also use a thermal state at the channel input, bounding its coherent information for attenuators and amplifiers with arbitrary finite-mean-energy environments. In the infinite input-energy limit their bounds give $[\log_2 \chi - S_{\text{bits}}(\rho_E)]_+$. Theorem B.16 attains this lower bound with structured GKP families under the finite-entropy assumption.

The entropy-only rate guarantee follows from three averaging steps. Lattice averaging replaces the relative placement and orientation of dual-lattice samples and environmental overlap patterns by the phase-space integral. Weyl orthogonality removes the environmental eigenbasis from that integral. Typicality then restricts the environmental alternatives to $2^{N[S_{\text{bits}}(\rho_E)+o(1)]}$ states. Their exponential count supplies the entropy cost, while the code family can still depend on the environment.

These rates equal the infinite-energy coherent-information lower bounds obtained from thermal states [116]. For thermal loss, this benchmark need not be the tightest lower bound on quantum capacity. At $\eta = 0.8, \bar{n} = 1$, it vanishes, whereas an explicit non-Gaussian state has certified positive coherent information and therefore gives a strictly stronger capacity lower bound [202]. At $\bar{n} = 0$, the expressions coincide with the known pure-channel capacities [117].

Corollary B.25 (Additive noise)**.** *For $\mathcal{B}_2(v) : V \mapsto V + vI$ with $v > 0$, $R_{\text{GKP}} \geq [-\log_2(ev)]_+$.*

Proof. Appending pure loss η at the receiver gives $V \mapsto \eta V + [\eta v + (1-\eta)/2]I$, a thermal attenuator with $\bar{n} = \eta v/(1-\eta)$. Since $g_2(n) = \log_2(en) + o(1)$, Corollary B.24 tends to $[-\log_2(ev)]_+$ as $\eta \to 1$. For any strictly smaller target rate fix $\eta < 1$ first, then apply Theorem B.16 and absorb the appended attenuation into the recovery; receiver postprocessing cannot increase the achievable rate. □

This recovers the established displacement-noise rate [43].

Theorem B.26 (Canonical Gaussian classes)**.** *GKP families attain the standard coherent-information lower bound obtained from thermal states in the infinite-energy limit for single-mode Gaussian channels.*

Proof. Input and output Gaussian unitaries are absorbed into the encoding and recovery and preserve finite code energy. The transformed envelope may be anisotropic. Classes C (loss and amplification) are Corollary B.24, and B_2 is Corollary B.25. For the singular class $B_1 : V \mapsto V + \mathrm{diag}(0, w)$, conjugating by $\mathrm{diag}(s, s^{-1})$ gives noise $\mathrm{diag}(0, w/s^2)$; adding position noise w/s^2 at the receiver yields $\mathcal{B}_2(w/s^2)$, and Corollary B.25 is unbounded as $s \to \infty$, so fixing s before the coding limit gives every finite rate. Classes A_1, A_2 and D are entanglement breaking [201] with zero capacity. For class D, $\tau < 0$ gives $\chi = |\tau|/(1+|\tau|) < 1$, so the bound returns zero. The identity channel has unbounded capacity. □

Comparison with pre-amplification GKP rates. The pre-amplification GKP construction of Ref. [1] is dominated by Corollary B.24. For $0 < \eta < 1$ and $\bar{n} \geq 0$,

$$\left[\log_2 \frac{\eta}{1-\eta} - g_2(\bar{n})\right]_+ \geq \log_2 \max\left\{1, \left\lfloor \frac{1}{e(1-\eta)(1+\bar{n})} \right\rfloor\right\}. \tag{B.74}$$

The inequality is strict whenever $e(1-\eta)(1+\bar{n}) < 1$, already against the unfloored expression $\log_2[1/(e(1-\eta)(1+\bar{n}))]$. In particular, it is strict wherever the pre-amplification rate is positive; elsewhere both guarantees can vanish. The construction also removes the integer floor. For comparison, amplification after the channel gives the distinct benchmark $[\log_2\{\eta/[e(1-\eta)(1+\bar{n})]\}]_+$.

B.10.5 Extension to Gaussian-dilatable channels

In the previous subsection, we focused on Gaussian channels. The coding result extends more generally to Gaussian-dilatable channels as defined in Definition 3.1. Each independent use has one signal input a, one retained output b, and K environmental modes in a fixed, possibly non-Gaussian and internally correlated state ρ_E, independent of the input.

Definition B.27 (Gaussian unitary and input–output relations)**.** *A Gaussian unitary $\hat{U} : aE \to bF$ acts affinely on the interleaved quadratures, with real displacement d and symplectic*

matrix S,

$$\hat{U}^\dagger \begin{pmatrix} \hat{x}_b \\ \hat{x}_F \end{pmatrix} \hat{U} = S \begin{pmatrix} \hat{x}_a \\ \hat{x}_E \end{pmatrix} + d, \qquad S := \begin{pmatrix} X & B \\ C_F & D_F \end{pmatrix}, \qquad S(\omega \oplus \Omega_E)S^T = \omega \oplus \Omega_E. \quad \text{(B.75)}$$

Here F contains the K discarded modes and $\Omega_E := I_K \otimes \omega$. Removing the output displacement gives the retained relation

$$\hat{x}_b = X\hat{x}_a + B\hat{x}_E, \qquad B\Omega_E B^T = (1-\tau)\omega, \qquad \tau := \det X, \quad \text{(B.76)}$$

with output operators in the Heisenberg picture and $B \in \mathbb{R}^{2\times 2K}$.

The induced channel need not preserve Gaussian states when ρ_E is non-Gaussian.

We assume finite total mean energy in the environmental state, $\operatorname{Tr}\rho_E \sum_{j=1}^{K} \hat{e}_j^\dagger \hat{e}_j < \infty$. Finite mean energy is an additional restriction on the dilation, not part of the Gaussian-dilatable definition [122]. This is stronger than the assumption of Theorem B.16, since finite mean energy implies finite entropy but not conversely. Let $|\Phi_{ER}\rangle$ purify $\rho_E = \operatorname{Tr}_R |\Phi_{ER}\rangle\langle\Phi_{ER}|$. For an input state ρ, set $\omega_{bFR}(\rho) := (\hat{U} \otimes I_R)(\rho \otimes |\Phi_{ER}\rangle\langle\Phi_{ER}|)(\hat{U}^\dagger \otimes I_R)$. The coherent information is $I_c(\rho, \mathcal{N}) := S_{\text{bits}}(\omega_b) - S_{\text{bits}}(\omega_{FR})$, with $\omega_b := \operatorname{Tr}_{FR}\omega_{bFR}$ and $\omega_{FR} := \operatorname{Tr}_b \omega_{bFR}$. The thermal-state benchmark is

$$I_c^\infty(\mathcal{N}) := \lim_{n\to\infty} I_c(\rho_{\text{th}}(n), \mathcal{N}). \quad \text{(B.77)}$$

Corollary B.28 (Gaussian-dilatable channels)**.** *For a Gaussian-dilatable channel $\mathcal{N}$ with one input and one output mode, admitting a dilation with a finite-mode environment of finite total mean energy, GKP families attain every positive rate below $[I_c^\infty(\mathcal{N})]_+$.*

For $\tau > 0$, this attains the thermal-state rate benchmark, including the displacement case $\tau = 1$. For $\tau \le 0$, its positive part vanishes. At $\tau = 0$ the channel is entanglement breaking and has zero quantum capacity.

Proof. For $\tau \neq 1$, the interaction reduces to one effective environmental mode. Theorem B.16 applies when $\tau > 0$; $\tau = 0$ is handled by measurement followed by preparation. For $\tau = 1$, receiver attenuation gives the displacement-noise rate and its thermal-state limit. All limits below also hold for a fixed squeezed thermal family with covariance $(n + 1/2)H$, $H > 0$, $\det H = 1$.

Effective environmental mode ($\tau \neq 1$). Let $Z := \mathrm{diag}(1, -1)$ and define

$$L_E := \begin{cases} B/\sqrt{1-\tau}, & \tau < 1, \\ ZB/\sqrt{\tau - 1}, & \tau > 1, \end{cases} \qquad L_E \Omega_E L_E^T = \omega. \tag{B.78}$$

Extend these rows to a symplectic matrix S_E, implemented by $\hat{U}_E^\dagger \hat{x}_E \hat{U}_E = S_E \hat{x}_E$, and set

$$\rho_{E,\mathrm{eff}} := \mathrm{Tr}_{E_2 \cdots E_K}(\hat{U}_E \rho_E \hat{U}_E^\dagger). \tag{B.79}$$

It has finite mean energy. Signal preprocessing by $\sqrt{\tau} X^{-1}$ for $\tau > 0$ gives exactly the attenuator $\eta = \tau$ or amplifier $G = \tau$ of Table B.1. For $\tau = -\kappa < 0$, preprocessing by $\sqrt{\kappa} X^{-1} Z$ gives $\hat{x}_b = \sqrt{\kappa} Z \hat{x}_a + \sqrt{1+\kappa}\hat{x}_{E_1}$. Characteristic-function factorization shows equality of channel maps. Environmental spectators are absorbed into a purification of the effective mode.

For $\tau \neq 0, 1$, a consequence of Theorem 6.1 of Ref. [203] is

$$I_c^\infty(\mathcal{N}) = \log_2 \frac{|\tau|}{|1-\tau|} - S_{\mathrm{bits}}(\rho_{E,\mathrm{eff}}). \tag{B.80}$$

The cited conditional entropy-power inequality and Gaussian maximal entropy give $S_{\mathrm{bits}}(b) = \log_2(e|\tau|n) + o(1)$ and $S_{\mathrm{bits}}(f|R_{\mathrm{eff}}) = \log_2(e|1-\tau|n) + o(1)$, where $R_{\mathrm{eff}} := E_2 \cdots E_K R$ purifies the effective mode. The equivalent one-mode dilation has complementary output $f R_{\mathrm{eff}}$, with $S_{\mathrm{bits}}(R_{\mathrm{eff}}) = S_{\mathrm{bits}}(\rho_{E,\mathrm{eff}})$. For positive τ, the rate is already Theorem B.16. Lim *et al.*, Sec. IV, Eqs. (22) and (26) of Ref. [131], obtained the same asymptotic lower bound for

general environments; those bounds alone do not state the exact limit. The Gaussian special cases were evaluated in Sec. V of Ref. [116]. For $K = 1$, $S_{\text{bits}}(\rho_{E,\text{eff}}) = S_{\text{bits}}(\rho_E)$. For $K > 1$, the reduced effective entropy, not the full-bath entropy, sets the rate. A pure entangled bath can have positive effective-mode entropy even though its total entropy vanishes.

For $\tau = 0$, the effective environmental rows are still a canonical pair. If $X = 0$ the channel is replacement. Otherwise $X = vw^T$ and

$$\mathcal{N}(\rho) = \int p_\rho(dz)\, \hat{T}(vz)\rho_{E,\text{eff}}\hat{T}(vz)^\dagger, \tag{B.81}$$

where p_ρ is the spectral law of $w^T\hat{x}_a$. This is measurement followed by preparation. For a thermal state at the input and its purifying system A', $S_{\text{bits}}(b) \le \frac{1}{2}\log_2(n+1) + \mathcal{O}(1)$ and $I(A' : b) \le I(z : b) = S_{\text{bits}}(b) - S_{\text{bits}}(\rho_{E,\text{eff}})$. Thus $I_c = I(A' : b) - g_2(n) \to -\infty$. Replacement has $I_c = -g_2(n)$ directly. For $\tau < 0$, Eq. (B.80) is negative; this does not assert that every non-Gaussian phase-conjugating channel is entanglement breaking.

Displacement noise ($\tau = 1$). The components of $B\hat{x}_E$ commute. Preprocessing by X^{-1} gives a random displacement with joint law μ. Appending receiver loss $\eta < 1$ gives an attenuator with environmental state

$$\mathcal{N}_\mu(\rho) := \int \mu(du)\, \hat{T}(u)\rho\hat{T}(u)^\dagger, \tag{B.82}$$

$$\rho_{E,\eta} := \int \mu(du)\, \hat{T}\left(\sqrt{\frac{\eta}{1-\eta}}u\right) |0\rangle\langle 0|\, \hat{T}\left(\sqrt{\frac{\eta}{1-\eta}}u\right)^\dagger. \tag{B.83}$$

Fixing η before the coding limit and absorbing loss into recovery, Theorem B.16 gives

$$R_{\text{GKP}}(\mathcal{N}_\mu) \ge \sup_{0<\eta<1} \left[\log_2 \frac{\eta}{1-\eta} - S_{\text{bits}}(\rho_{E,\eta})\right]_+. \tag{B.84}$$

For $\operatorname{rank} B = 2$, $\mu(du) = p(u)d^2u$ has finite second moments. Its differential entropy $h_2(p) := -\int p(u)\log_2 p(u)d^2u$ is finite. Indeed, measurement data processing against a full-rank product thermal reference gives finite relative entropy $D_{\text{bits}}(p\|q)$ to a nonsingular

Gaussian q; finite Gaussian cross entropy then gives $h_2(p) = -\int p(u)\log_2 q(u)d^2u - D_{\text{bits}}(p\|q)$. Theorem 5 of Ref. [204], its Lemma 3, and Equation (48) of Ref. [205] give

$$S_{\text{bits}}(\rho_{E,\eta}) = \log_2 \frac{\eta}{1-\eta} + h_2(p) - \log_2(2\pi) + o(1) \quad (\eta \to 1), \tag{B.85}$$

with the constant fixed by the Weyl measure $d^2u/(2\pi)$. For a thermal state at the input, the same Theorem 5 with its purification as memory gives $S_{\text{bits}}(\omega_{FR}) \geq g_2(n) + h_2(p) - \log_2(2\pi)$, while $S_{\text{bits}}(\omega_b) = g_2(n) + o(1)$. Data processing and Eq. (B.80) give

$$\liminf_{n\to\infty} I_c(\rho_{\text{th}}(n), \mathcal{N}_p) \geq \log_2 \frac{\eta}{1-\eta} - S_{\text{bits}}(\rho_{E,\eta}). \tag{B.86}$$

Taking $\eta \to 1$ after $n \to \infty$ proves

$$I_c^{\infty}(\mathcal{N}_p) = \log_2(2\pi) - h_2(p), \qquad R_{\text{GKP}}(\mathcal{N}_p) \geq [\log_2(2\pi) - h_2(p)]_+. \tag{B.87}$$

For $\operatorname{rank} B \leq 1$, symplecticity gives $\operatorname{rank} C_F \leq 1$. Gaussian maximal entropy gives $S_{\text{bits}}(\rho_{E,\eta}) \leq \frac{1}{2}\log_2[\eta/(1-\eta)] + \mathcal{O}(1)$ and $S_{\text{bits}}(F) \leq \frac{1}{2}\log_2(n+1) + \mathcal{O}(1)$. The first bound permits every finite GKP rate; the second, together with $S_{\text{bits}}(b) \geq g_2(n)$ and finite $S_{\text{bits}}(R)$, gives $I_c^{\infty} = +\infty$. □

B.11 Lattice geometry beyond sphere packing

We prove the thermal-loss geometry result of Section 3.5 and give additional numerical geometry examples below.

Proposition B.29 (Planar thermal-loss lattice optimality)**.** *For fixed $0 < \eta < 1$ and $0 \leq \bar{n} < \infty$, the hexagonal lattice uniquely minimizes the infinite-energy D_0 infidelity proxy among one-mode GKP lattices of fixed covolume, up to orthogonal transformations.*

Proof. We express the thermal kernel as a positive mixture of Gaussians and apply the planar

theta-function minimum to each Gaussian. For $\bar{n} > 0$, Eq. (B.51) gives

$$\begin{aligned} f_{\bar{n}}(v) &= \int_{\mathbb{R}} \varrho_F(k) e^{-a_{\bar{n}}(k)v}\, dk, \\ a_{\bar{n}}(k) &:= 2\bar{n} + 1 - 2\sqrt{\bar{n}(1+\bar{n})}\cos(k \ln t_E), \qquad t_E = \frac{\bar{n}}{1+\bar{n}}. \end{aligned} \tag{B.88}$$

Here $\varrho_F(k) > 0$ and $\int \varrho_F = 1/4$ by Eq. (B.106), while

$$0 < \left(\sqrt{1+\bar{n}} - \sqrt{\bar{n}}\right)^2 \le a_{\bar{n}}(k) \le \left(\sqrt{1+\bar{n}} + \sqrt{\bar{n}}\right)^2.$$

In particular, the kernel is completely monotone in the squared displacement v. For every integer $q \ge 0$,

$$(-1)^q f_{\bar{n}}^{(q)}(v) = \int_{\mathbb{R}} \varrho_F(k) a_{\bar{n}}(k)^q e^{-a_{\bar{n}}(k)v}\, dk > 0, \qquad v \ge 0. \tag{B.89}$$

The finite upper bound on $a_{\bar{n}}$ justifies differentiation under the integral. For a planar displacement lattice $\mathcal{L}$, Tonelli's theorem and Eq. (B.41) give

$$\sum_{\xi \in \mathcal{L}\setminus\{0\}} f_{\bar{n}}(\chi|\xi|^2) = \int_{\mathbb{R}} \varrho_F(k) \sum_{\xi \in \mathcal{L}\setminus\{0\}} e^{-a_{\bar{n}}(k)\chi|\xi|^2}\, dk.$$

The positive lower bound on $a_{\bar{n}}$ also makes this sum finite. At each k, Montgomery's theorem [133] uniquely minimizes the inner Gaussian sum by the hexagonal lattice at the prescribed covolume. The omitted origin contributes the same constant for every lattice. Integration with positive weight preserves this minimum and its uniqueness. In vacuum, $f_0(v) = e^{-v}/4$ and the same theorem applies directly. □

The Fock-state kernel, Laguerre-node mechanism, and comparisons with environment-assisted transmission are discussed in Section 3.5. The following figure supplements those examples with a rectangular-lattice search.

Figure B.1 compares full proxy sums over rectangular lattices at $\eta = 0.65$. Figure B.1(c)–

(f) shows the minimizing members found in the rectangular search, with environmental displacements

$$\beta_{jk} = \sqrt{\frac{\pi\chi}{2}}\,(sj + ik/s), \qquad (j,k) \in \mathbb{Z}^2.$$

The cell aspect ratio s^2 is distinct from the kernel argument $v = |\beta|^2$ in panel (a). At a lattice point, $|\beta_{jk}|^2 = (\pi\chi/2)(s^2j^2 + k^2/s^2)$; the nearest axial shell for $s \geq 1$ has $v = \pi\chi/(2s^2)$. Thus changing s^2 changes which kernel radii the lattice samples. A zero $v_{m,\ell}$ of L_m becomes a circle $|\beta| = \sqrt{v_{m,\ell}}$; points near such a circle contribute little to the proxy. For example, at $m = 1$ the minimizing member has nearest axial shell $|\beta|^2 \simeq 1.115$, close to the node at 1. The full sum weighs every shell, shifting the minimum away from the node.

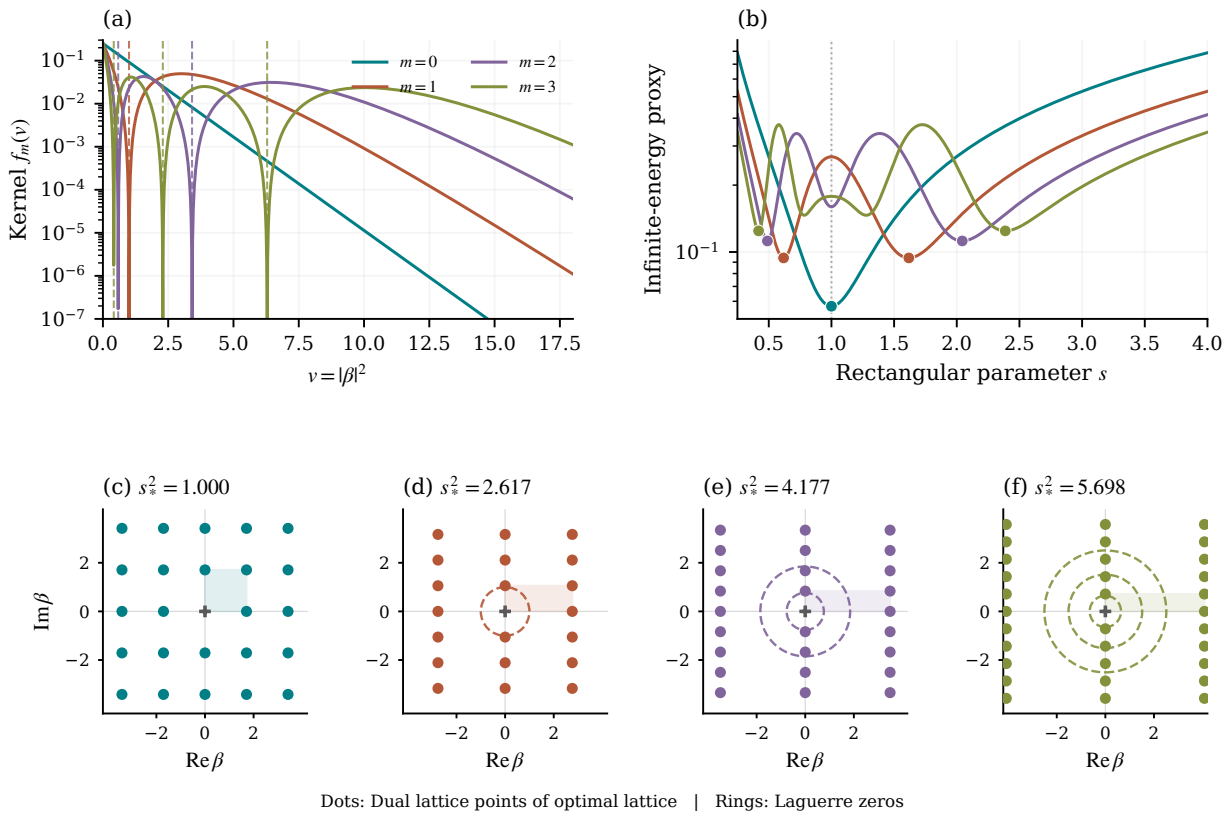


Figure B.1: Environmental overlap nodes reshape the preferred lattice geometry. The environment is the Fock state $|m\rangle$ [6]. (a) The overlap kernel in Eq. (3.54). (b) Infidelity proxy at the infinite-energy limit, obtained by summing this kernel over $\Lambda_s^{\perp} := \mathrm{diag}(s, s^{-1})\mathbb{Z}^2/\sqrt{2}$ at $\eta = 0.65$. Its stabilizer is symplectically integral and has $d_L = 2$. Dots mark numerical local minima on $1/4 \le s \le 4$. The dotted line marks the square member $s = 1$, which minimizes the vacuum curve. (c)–(f) Fidelity-optimized dual lattices in β coordinates for $m = 0, 1, 2, 3$, respectively, obtained by minimizing the infinite-energy infidelity proxy with $s \ge 1$. The displayed s_*^2 is the optimal cell aspect ratio. The stretch s determines the kernel radii in (a) through $v = |\beta_{jk}|^2 = (\pi\chi/2)(s^2 j^2 + k^2/s^2)$, with $\chi = \eta/(1-\eta)$ and $(j, k) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$; the nearest axial shell has $v = \pi\chi/(2s^2)$. Dashed circles mark the Laguerre zeros, also indicated by color-matched dashed vertical lines in (a). Shading marks one cell; the origin is excluded from the sum.

The kernels $f_0(v) = e^{-v}/4$ and $f_1(v) = e^{-v}(1-v)^2/4$ correspond respectively to a vacuum environment $|0\rangle$ and a one-photon environment $|1\rangle$. A squeezed-vacuum environment $\hat{S}(r)\,|0\rangle$, with $\hat{S}(r) := \exp\!\left[r(\hat{e}^2 - \hat{e}^{\dagger 2})/2\right]$, instead has

$$f_{\mathrm{sq}}(\beta) = \tfrac{1}{4}\exp\!\left[-e^{2r}(\mathrm{Re}\,\beta)^2 - e^{-2r}(\mathrm{Im}\,\beta)^2\right].$$

All three environments are pure, hence have $S_{\mathrm{bits}} = 0$, and their kernels obey

$\int_{\mathbb{C}} f(\beta)\, d^2\beta/\pi = 1/4$. Their shapes nevertheless differ. The normalization here uses environmental displacement β; integration in lattice coordinates includes the χ^{-1} Jacobian of Eq. (B.64).

B.12 Vacuum recovery protocols and simulation details

In this section, we give a brief review on recovery protocols that are compared against in Section 3.3.3 and the simulation details. To start with, the conventional decoder is the decoder proposed by the original work [4]. Consider a single-mode square lattice GKP code, the decoder measures the quadrature operators modular the lattice spacings and performs a corresponding displacement. It can be implemented through, e.g., a supply of ancillary GKP states along with SUM gates [4, 206], phase estimations [126], or teleportation-based error correction [207]. However, while such a strategy is optimal for infinite-energy GKP code under displacement channel, it is far from optimal when we consider realistic noises since it does not take into account, for example, the contraction features of loss.

A closely related decoder is the amplification decoder (AD) [1, 3]. The key idea is that through channel composition, we can convert loss (amplification) channel into Gaussian random displacement noise channel through amplifying (contracting) the logical states,

$$\mathcal{N}_L(\eta) \circ \mathcal{N}_A(G) = \mathcal{N}_{B_2}(\sigma^2) \tag{B.90}$$

with $\eta G = 1$ and $\sigma^2 = 1 - \eta$. Here, $\mathcal{N}_L$, $\mathcal{N}_A$, and $\mathcal{N}_{B_2}$ represent pure loss, pure amplification, and Gaussian random displacements, respectively. After the noise channel is converted into Gaussian noise, it is natural to then append it with the conventional decoding. Nevertheless, the channel conversion process injects additional uncertainty into the system since it involves the participation of an input ancillary mode in a vacuum, which is eventually traced out. Therefore, it is also suboptimal. The simulations of both the conventional decoder and AD are performed through the Zak basis representation of the GKP code [208], which is computationally efficient.

The small-Big-small (sBs) is another recovery that is practical and has attracted experimental interest. The core advantage of it is its components are all experimentally accessible and are free from measurements, which could be the bottleneck for both operating speed and quality for many experimental platforms. More specifically, one round of sBs recovery circuit that stabilizes the position quadrature consists of a series of unitary operations on the oscillator and an ancillary qubit,

$$C\hat{D}\left(\delta/2\right)\hat{R}^{\dagger}(\pi/2)C\hat{D}\left(-il\right)\hat{R}(\pi/2)C\hat{D}\left(\delta/2\right) \tag{B.91}$$

where the conditional displacement is defined as $C\hat{D}\left(\alpha\right) := |0\rangle\langle 0|\otimes\hat{D}(\alpha/2)+|1\rangle\langle 1|\otimes\hat{D}(-\alpha/2)$ and $\hat{R}(\pi/2) := \exp\left\{-i\hat{\sigma}_x\pi/4\right\}$. The displacement lengths $\delta \approx \Delta^2 l$ and $l = \sqrt{2\pi}$. The ancilla qubit is initialized in $|+\rangle$ and reset at the end. The sBs circuit is developed through the trotterization of the nullifier of GKP's finite-energy stabilizers. Since sBs does not stabilize the finite-energy GKP codespace with the Gaussian envelope, our simulation first discovers the stabilized codewords by initializing the oscillator in GKP codewords and applying the sBs circuits till convergence. Then, the stabilized codewords undergo the noise channel of interest and are later recovered through repeated applications of the stabilization circuits. The performance is evaluated through comparing the fidelities between the states before the noise channel and after the recovery.

B.13 Approximate implementation of the near-optimal recovery.

While our core message is the analytical performance of the GKP code, there are also implications of practical interest. One of them is the possibility to design more optimal recoveries inspired by the near-optimal recoveries such as the transpose channel.

Suppose one focuses on the two leading orders of the TC Kraus operators, i.e. $i = 0, 1$ in Eq. (A.21), it is straightforward to see that it roughly corresponds to a projection onto the no-error and one-error subspace and a rotation back to the codespace. The unshifted square-lattice qubit GKP code possesses a two-fold rotation symmetry [209] such that both

of its codewords live in the even parity subspace. For codewords with a common parity, this two-fold symmetry allows exact single-photon error detection, as discussed in Appendix A. Therefore, take square lattice qubit GKP code as an example, one can approximate the error subspace projection through a parity measurement. Then, one can apply a unitary $\hat{U}_{0\to L}$ or $\hat{U}_{1\to L}$ depending on the parity being measured to be even or odd, respectively. Here, the unitary $\hat{U}_{i\to L}$ rotates the oscillator from the i-th error subspace back to the logical codespace. The exact implementation of such a unitary can be discovered through techniques such as optimal control. While we have been focused on the explicit implementation of the first two Kraus operators, it is worth noting that it is in principle possible to extend such analysis to higher orders or through systematic techniques that implement general CPTP maps [197]

B.14 Numerical results

Near-optimal (Petz) fidelity is evaluated using Eq. (2.15), and the infidelity proxy using Eq. (2.19) with its specified complementary reference. We plot $1-\tilde{F}^{\mathrm{opt}}$, the proxy, and their ratio to compare recovery performance with the proxy and its large-energy limits.

B.14.1 Finite-energy evaluation of the infidelity proxy

Equation (B.92) evaluates the full finite-energy proxy plotted in Fig. 2(c) of the Letter. It combines the thermal-loss pair kernel (B.49) with exact Gram orthonormalization. Quantum-limited amplification uses Eq. (B.56); finite-Fock environments use the QEC expansion (B.12). The thermal-loss lattice sum converges absolutely when $h_\Delta < 0$, as holds at sufficiently large energy by Eq. (B.50).

Write $M_{\hat{C}_\Delta} = \sum_\xi \hat{W}_\xi \otimes \hat{O}_\xi$ with $\hat{W}_0 = I_L$, and use the Gram matrix $\hat{G}_\Delta = \sum_\xi \hat{W}_\xi e^{-(n_\Delta+1/2)|\xi|^2}$ and $\hat{H}_\Delta = \hat{G}_\Delta^{-1/2}$ from Remark B.9. Put $B_0 := \hat{H}_\Delta^2 - I_L$ and $B_\xi := \hat{H}_\Delta \hat{W}_\xi \hat{H}_\Delta$ for $\xi \neq 0$. Logical congruence and $\hat{O}_0 = D_0$ give

$$\tilde{\epsilon}_{D_0}(M_\Delta) = \frac{1}{d_L} \sum_{\xi,\xi' \in \mathcal{L}} \mathrm{Tr}\Big(B_\xi^\dagger B_{\xi'}\Big) K_\Delta(\xi,\xi'). \tag{B.92}$$

The numerical curves retain every Gram term and lattice pair in Eq. (B.92). The leading large-

energy approximation is Eq. (B.41), with first-order correction $\epsilon_\infty + a_{\mathcal{L}}/n_\Delta$ from Eqs. (B.55) and (B.60). These are proxy approximations at a fixed channel and lattice; the controlled remainder is $\mathcal{O}(n_\Delta^{-2})$.

Asymptotic limit. In Fig. B.2, dashed curves evaluate the exact finite-energy D_0 proxy, dotted lines show ϵ_∞, and dash-dotted curves show $\epsilon_\infty + a_{\mathcal{L}}/n_\Delta$. The first-order expansion approaches the exact proxy as energy grows.

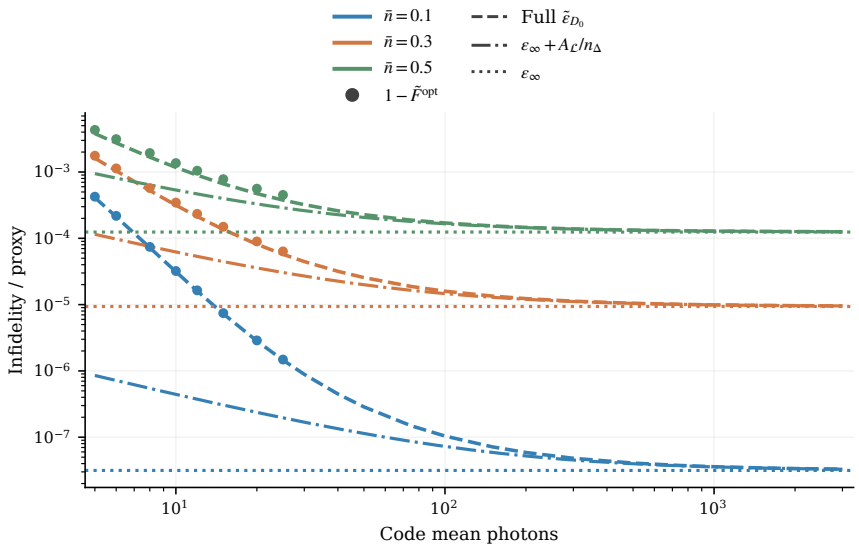


Figure B.2: Convergence of the first-order energy expansion to the full finite-energy proxy. For square qubit GKP codes under thermal loss at $\eta = 0.95$, dashed curves show the Gram-orthonormalized D_0-reference proxy, dash-dotted curves its first-order expansion, and dotted lines its infinite-energy limit. Points show near-optimal infidelity $1 - \tilde{F}^{\mathrm{opt}}$ only over the original 5–25-photon range. Colors distinguish the bath occupations $\bar{n} = 0.1, 0.3, 0.5$. The horizontal coordinate is the exact mean code energy at the same code energy parameter n_Δ used in each expression.

B.14.2 Plotted results and provenance

Figure B.3 uses the saved thermal-channel calculations at $\eta = 0.85$ and $G = 1.4$, with environment occupations 0 and 0.3. Both the Petz infidelity and the marginal-reference proxy use the same Gram-isometric codes and the same physical channel. The ratio in panel (c) is formed directly from the quantities in panel (a). Lemma A.5 gives $\tilde{\epsilon}_{D_{\mathrm{marg}}}/(1 - \tilde{F}^{\mathrm{opt}}) \geq 1/4$ whenever $1 - \tilde{F}^{\mathrm{opt}} > 0$.

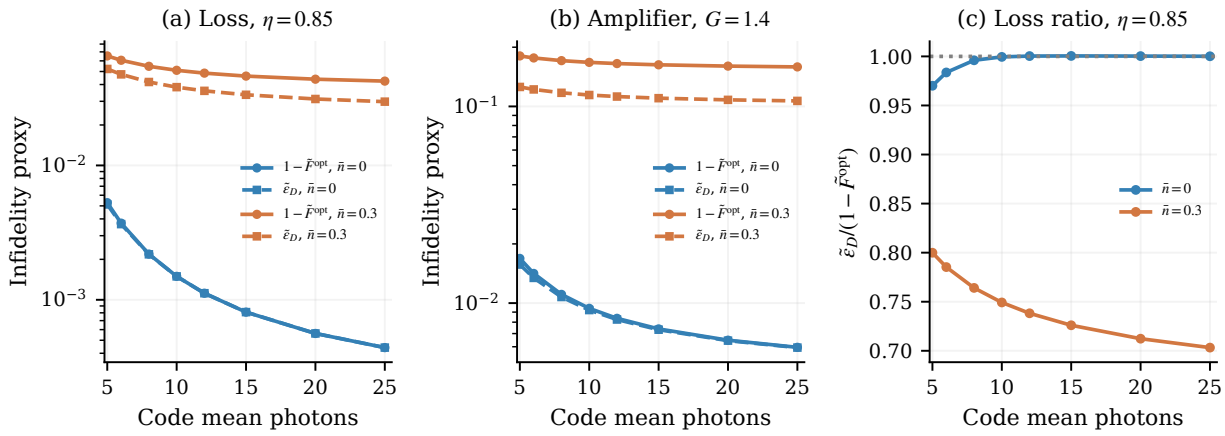


Figure B.3: Comparison of finite-energy recovery infidelity with its marginal-reference proxy. The codes are square qubit GKP codes with Gram orthonormalization and exact mean code photons on the horizontal axis. The environment is a thermal state of mean occupation $\bar{n} = 0$ or 0.3. For (a) thermal loss at $\eta = 0.85$ and (b) thermal amplification at $G = 1.4$, circles show the actual near-optimal infidelity $1 - \tilde{F}^{\mathrm{opt}}$ and squares show $\tilde{\epsilon}_{D_{\mathrm{marg}}}$. For simplicity, the plot labels use D for D_{marg}. (c) The ratio $\tilde{\epsilon}_{D_{\mathrm{marg}}}/(1 - \tilde{F}^{\mathrm{opt}})$ for the same loss channel and codes as in (a), at $\eta = 0.85$. At the marginal reference the proxy is the leading local perturbative term (Ref. [13], Appendix D), so finite-energy corrections affect the ratio. Lemma A.5 bounds this ratio from below by 1/4 whenever the near-optimal infidelity is nonzero.

Figure B.4 evaluates the three references in the same discarded-output and purification-label basis. The zero-displacement reference is computed from the exact finite-Fock QEC polynomial, and the product reference from Eq. (B.25). Input, comb, outcome, bath, and inverse-reference cutoffs were refined. These are cutoff-refined floating-point evaluations.

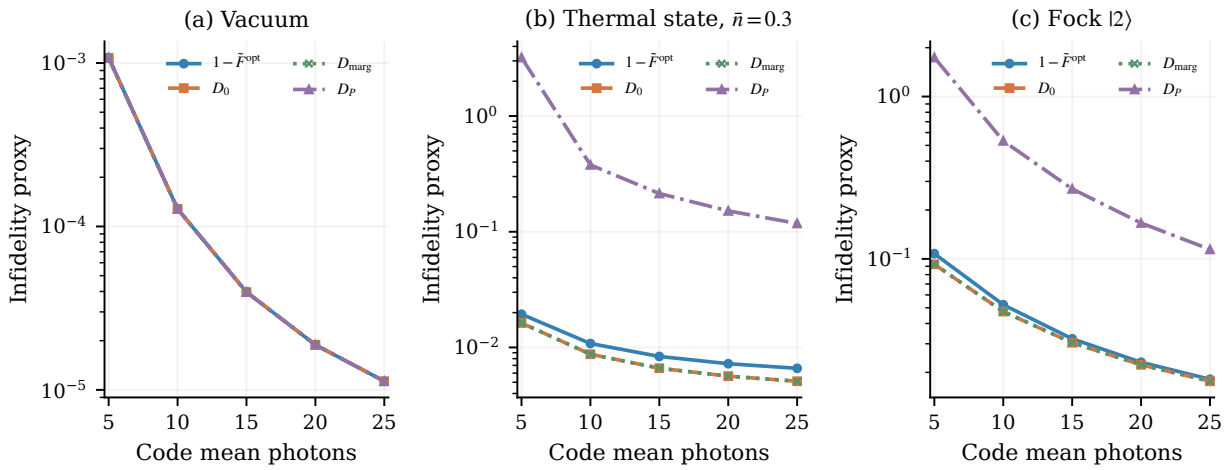


Figure B.4: Reference choice changes the finite-energy infidelity proxy. Square qubit GKP codes undergo loss at $\eta = 0.9$ in all three panels, with (a) a vacuum environment, (b) a thermal state of mean occupation $\bar{n} = 0.3$, and (c) the two-photon Fock state $|2\rangle$. Circles show actual near-optimal infidelity; squares, crosses, and triangles show the proxies for D_0, D_{marg}, and D_{p}, respectively, using Eqs. (2.15) and (2.19). The proxy curves are labeled by their reference alone, with D_P denoting the product reference D_{p}. The D_0 and D_{marg} curves nearly coincide on this scale. The horizontal coordinate is the mean energy of the Gram-isometric code.

Figure 2(c) of the Letter instead uses thermal loss at $\eta = 0.95$ and $\bar{n} = 0.1, 0.3, 0.5$. Direct Petz evaluation covers 5–25 mean photons. The dashed curves evaluate Eq. (B.92), with two lattice radii (8 and 10) and two quadrature orders (384 and 768). Their horizontal coordinate uses Eq. (3.14). The dotted horizontal lines use the infinite-energy limit (B.41). The D_0 proxy extends to 3000 photons.

B.15 Mathematical identities

This section collects the operator inequalities and integral identities used in the QEC and coding proofs.

Lemma B.30 (Square-root comparison)**.** *Let A, B be positive trace-class operators with* $\operatorname{supp} A \subseteq \operatorname{supp} B$. *In an eigenbasis $B = \sum_m b_m |m\rangle\langle m|$ on its support, set $X_{mn} := (A - B)_{mn}/(\sqrt{b_m} + \sqrt{b_n})$. If X is Hilbert–Schmidt, then $\left\|\sqrt{A} - \sqrt{B}\right\|_F^2 \leq 4\|X\|_F^2$.*

Proof. The finite-matrix Sylvester equation gives the bound. Spectral compression and trace-norm continuity of square roots extend it to trace-class operators. *Finite matrices.*

Let $A \geq 0$, $B > 0$ be finite matrices, $C := \sqrt{B}$, $Y := \sqrt{A} - C$, and $\mathcal{T}_C(Z) := CZ + ZC$, with inverse $\mathcal{T}_C^{-1}(Z) = \int_0^\infty e^{-tC} Z e^{-tC} dt$; the integral converges because $C > 0$, and $\mathcal{T}_C^{-1}$ is positivity preserving and Hilbert–Schmidt self-adjoint. The defining Sylvester equation for X is

$$\mathcal{T}_C(X) = CX + XC = A - B, \qquad X := \mathcal{T}_C^{-1}(A - B). \tag{B.93}$$

Expanding $A = (C + Y)^2 = B + CY + YC + Y^2$ in that order gives

$$X = Y + \mathcal{T}_C^{-1}(Y^2). \tag{B.94}$$

Since $\sqrt{A} \geq 0$ we have $Y \geq -C$, and $\mathcal{T}_C(I/2) = C$ gives $\mathcal{T}_C^{-1}(Y) \geq -I/2$. Pairing Eq. (B.94) with Y and using self-adjointness,

$$\operatorname{Tr}(YX) = \operatorname{Tr}\left(Y^2\right) + \operatorname{Tr}\left[\mathcal{T}_C^{-1}(Y)\, Y^2\right] \geq \tfrac{1}{2}\|Y\|_F^2, \tag{B.95}$$

because $\operatorname{Tr}(HP) \geq 0$ for $H := \mathcal{T}_C^{-1}(Y) + I/2 \geq 0$ and $P := Y^2 \geq 0$. Cauchy–Schwarz then gives $\|Y\|_F \leq 2\|X\|_F$, i.e.

$$\left\|\sqrt{A} - \sqrt{B}\right\|_F^2 \leq 4\|X\|_F^2. \tag{B.96}$$

Trace-class operators. Now let A, B be positive trace-class operators, with $\operatorname{supp} A \subseteq \operatorname{supp} B$. On $\operatorname{supp} B$, write $B = \sum_m b_m \left|m\right\rangle \left\langle m\right|$, $b_m > 0$, and set

$$P_j := \mathbf{1}_{[1/j,\infty)}(B) = \sum_{b_m \geq 1/j} \left|m\right\rangle \left\langle m\right|, \qquad A_j := P_j A P_j, \qquad B_j := P_j B P_j. \tag{B.97}$$

The projections P_j increase strongly to the identity on $\operatorname{supp} B$ and have finite rank, since $\operatorname{rank} P_j \leq j \operatorname{Tr} B$. Apply the finite-matrix bound on $\operatorname{span}\{|m\rangle : b_m \geq 1/j\}$, where $B_j > 0$, and extend the resulting operators by zero.

Assume the weighted square sum is finite and define the Hilbert–Schmidt matrix X by $X_{mn} := (A-B)_{mn}/(\sqrt{b_m}+\sqrt{b_n})$. The solution for the compressed problem is $X_j = P_j X P_j$, and

$$\left\|X_j\right\|_F^2 = \sum_{\substack{b_m \geq 1/j \\ b_n \geq 1/j}} |X_{mn}|^2 \uparrow \sum_{m,n} |X_{mn}|^2 = \|X\|_F^2. \tag{B.98}$$

Also $\left\|A_j - A\right\|_1 \to 0$ and $\left\|B_j - B\right\|_1 \to 0$.

The square-root continuity estimate

$$\left\|\sqrt{A} - \sqrt{B}\right\|_F^2 \leq \|A-B\|_1 \tag{B.99}$$

holds for positive trace-class operators. Indeed, with $Y = \sqrt{A} - \sqrt{B}$, trace-norm duality and cyclicity give $\|A-B\|_1 \geq \operatorname{Tr}[\operatorname{sign}(Y)(A-B)] = \operatorname{Tr}\left[|Y|(\sqrt{A}+\sqrt{B})\right] \geq \operatorname{Tr} Y^2$. For the last inequality, in an eigenbasis of Y each diagonal entry of $\sqrt{A}+\sqrt{B}$ is at least the absolute value of the corresponding entry of Y, because both square roots are positive. These traces are finite by Hilbert–Schmidt Hölder. Applying Eq. (B.99) to (A_j, A) and (B_j, B) gives convergence of their square roots in Frobenius norm. Together with Eq. (B.98), this passes Eq. (B.96) to the trace-class limit.

□

Lemma B.31 (Oscillator envelope on tempered distributions)**.** *For $s > 0$, $e^{-s\hat{n}_{\mathrm{tot}}}$ is injective on tempered distributions and maps them into Schwartz functions.*

Proof. The Hermite coefficients of a tempered distribution grow at most polynomially. The envelope multiplies the coefficient with multi-index $\boldsymbol{m}$ by the nonzero factor $e^{-s|\boldsymbol{m}|_1}$. The resulting coefficients decay faster than every power, characterizing Schwartz functions. If the enveloped distribution vanishes, all its Hermite coefficients vanish, hence the original distribution is zero.

□

Lemma B.32 (Normal ordering of a displaced geometric operator)**.** *Let $\hat{f}$ be a bosonic annihilation operator and $\hat{n} := \hat{f}^\dagger \hat{f}$. For $0 < r < 1$ and $z \in \mathbb{C}$,*

$$e^{z\hat{f}^\dagger} r^{\hat{n}} e^{-z^* \hat{f}} = e^{|z|^2/(1+r)} \hat{D}\left(\frac{z}{1+r}\right) r^{\hat{n}} \hat{D}\left(\frac{z}{1+r}\right). \tag{B.100}$$

Proof. Normal ordering the displacement operators and using $\hat{f} r^{\hat{n}} = r\, r^{\hat{n}} \hat{f}$ and $r^{\hat{n}} \hat{f}^\dagger = r \hat{f}^\dagger r^{\hat{n}}$ gives

$$\hat{D}(h) r^{\hat{n}} \hat{D}(h) = e^{-(1+r)|h|^2} e^{(1+r)h\hat{f}^\dagger} r^{\hat{n}} e^{-(1+r)h^* \hat{f}}.$$

Set $h = z/(1+r)$.

□

Lemma B.33 (Hermite coefficient identity)**.** *For $b > 0$, $z \in \mathbb{C}$ and nonnegative integers a, h,*

$$[y^a w^h] e^{byw - z^* y + zw} = \frac{b^{(a+h)/2}}{\sqrt{a!h!}} e^{|z|^2/(2b)} \langle h | \hat{D}(z/\sqrt{b}) | a \rangle. \tag{B.101}$$

Proof. Expansion of the left side gives $\sum_{j=0}^{\min(a,h)} b^j (-z^*)^{a-j} z^{h-j} / [j!(a-j)!(h-j)!]$. The normal-ordered form $\hat{D}(\alpha) = e^{-|\alpha|^2/2} e^{\alpha \hat{f}^\dagger} e^{-\alpha^* \hat{f}}$ gives the same sum on the right after setting $\alpha = z/\sqrt{b}$.

□

Lemma B.34 (Finite-projection approximation)**.** *Let $\hat{\Phi} : \mathbb{C}^k \to \mathcal{H}$ be an isometry and let $\hat{\Pi}_L \to I$ strongly be finite-rank orthogonal projections. For sufficiently large L,*

$$G_L := \hat{\Phi}^\dagger \hat{\Pi}_L \hat{\Phi} \to I_k, \qquad \hat{\Psi}_L := \hat{\Pi}_L \hat{\Phi} G_L^{-1/2} \to \hat{\Phi}, \qquad \hat{\Psi}_L^\dagger \hat{\Psi}_L = I_k. \tag{B.102}$$

The convergence is in operator norm. For every state d on $\mathbb{C}^k$, $\left\| \hat{\Psi}_L d \hat{\Psi}_L^\dagger - \hat{\Phi} d \hat{\Phi}^\dagger \right\|_1 \leq 2 \left\| \hat{\Psi}_L - \hat{\Phi} \right\|$.

Proof. Strong convergence is uniform on the unit sphere of the finite-dimensional range of $\hat{\Phi}$. Thus $G_L \to I_k$ in operator norm, its inverse square root exists for large L and tends to I_k,

and the displayed assertions follow. Split the state difference as $(\hat{\Psi}_L - \hat{\Phi})d\hat{\Psi}_L^\dagger + \hat{\Phi}d(\hat{\Psi}_L^\dagger - \hat{\Phi}^\dagger)$. Trace-norm submultiplicativity and the unit norms of the isometries give the claimed bound.

□

Lemma B.35 (Two-mode Gaussian normal modes)**.** *Let d be a zero-mean two-mode Gaussian state with finite occupations, $\langle \hat{f}_j^\dagger \hat{f}_j \rangle = n_j$, $\langle \hat{f}_1 \hat{f}_2 \rangle = c \geq 0$, and all other independent anomalous and intermode moments zero. Under the two-mode squeezing transformation with* $\tanh(2r) = 2c/(n_1 + n_2 + 1)$*, its thermal occupations are*

$$n_\pm := \frac{1}{2}\left[\sqrt{(n_1 + n_2 + 1)^2 - 4c^2} - 1 \pm (n_1 - n_2)\right]. \tag{B.103}$$

The labels $+$ and $-$ refer to the first and second modes, respectively.

Proof. Set $\hat{a}_1 = \cosh r\, \hat{f}_1 - \sinh r\, \hat{f}_2^\dagger$ and $\hat{a}_2 = \cosh r\, \hat{f}_2 - \sinh r\, \hat{f}_1^\dagger$. Their correlation is $c\cosh(2r) - (n_1 + n_2 + 1)\sinh(2r)/2 = 0$. Their occupation difference is $n_1 - n_2$ and their sum plus one is $\sqrt{(n_1 + n_2 + 1)^2 - 4c^2}$, giving $n_\pm$. The vanishing cross moments and Gaussianity give the product state $\rho_{\mathrm{th}}(n_+) \otimes \rho_{\mathrm{th}}(n_-)$ in the new modes.

□

Lemma B.36 (All-vector symplectic first moment; [130])**.** *The primitive-vector mean in Eq. (0.6) of Ref. [130] implies Eq.* (B.67) *for every nonnegative integrable f.*

Proof. Every nonzero vector is uniquely a positive integer multiple of a primitive vector. Tonelli's theorem and the primitive-vector constant $1/\zeta(2N)$ give $\mathbb{E}S_f = \zeta(2N)^{-1}\sum_{j\geq 1}\int f(jx)\, d^{2N}x = \zeta(2N)^{-1}\sum_{j\geq 1} j^{-2N}\int f = \int f$. Monotone convergence extends the compactly supported formula to nonnegative integrable functions.

□

Lemma B.37 (Hardy–Hille identity)**.** *For $j \geq 0$, $x, y \geq 0$, and $0 < t < 1$,*

$$\sum_{k\geq 0} \frac{k!}{(k+j)!} t^k L_k^{(j)}(x) L_k^{(j)}(y) = \frac{(xyt)^{-j/2}}{1-t} e^{-t(x+y)/(1-t)} I_j\left(\frac{2\sqrt{xyt}}{1-t}\right). \tag{B.104}$$

Zero arguments are interpreted by continuity.

Lemma B.38 (Fourier representation of spectral weights)**.** *For $\lambda_r, \lambda_s > 0$ and $\nu_{rs} := \frac{1}{2}\ln(\lambda_r/\lambda_s)$,*

$$w_{rs}(\omega) := \frac{\lambda_r\lambda_s}{(\sqrt{\lambda_r}e^{\omega/4} + \sqrt{\lambda_s}e^{-\omega/4})^2} = \frac{\sqrt{\lambda_r\lambda_s}}{4\cosh^2[(\omega + 2\nu_{rs})/4]} \;\le\; \frac{\sqrt{\lambda_r\lambda_s}}{4}. \tag{B.105}$$

Let $\varrho_F(k) := 2k/\sinh(2\pi k)$, extended by $\varrho_F(0) = 1/\pi$; it is positive, even, integrable with $\int \varrho_F = 1/4$, and

$$\frac{1}{4\cosh^2(s/4)} = \int_{\mathbb{R}} \varrho_F(k) e^{iks}\, dk. \tag{B.106}$$

Let J_j be independent symmetric Skellam variables, $\Pr\Big(J_j = m\Big) = e^{-z_j} I_{|m|}(z_j)$ with $z_j \ge 0$. Then for real τ_j, θ_j,

$$\mathbb{E}\Big[w_{rs}\Big(\sum_j \tau_j J_j\Big) e^{-i\sum_j \theta_j J_j}\Big] = \sqrt{\lambda_r\lambda_s}\int_{\mathbb{R}} \varrho_F(k)\, e^{2ik\nu_{rs}} \exp\Big[\sum_j z_j\{\cos\Big(k\tau_j - \theta_j\Big) - 1\}\Big] dk, \tag{B.107}$$

whose modulus is at most $\sqrt{\lambda_r\lambda_s}/4$.

Proof. With $v = e^{s/2}$, the forward Fourier transform is

$$\int_{\mathbb{R}} \frac{e^{-iks}}{4\cosh^2(s/4)}\, ds = 2\int_0^\infty \frac{v^{-2ik}}{(1+v)^2}\, dv = 2\Gamma(1-2ik)\Gamma(1+2ik) = \frac{4\pi k}{\sinh(2\pi k)}.$$

The middle equality is the beta integral and the last equality uses the gamma reflection formula. Fourier inversion gives Eq. (B.106); evaluating it at $s = 0$ gives $\int \varrho_F = 1/4$. Insert it into Eq. (B.105); integrability of ϱ_F and unit modulus of the phases permit exchanging expectation and integration. Independence gives the product of Skellam characteristic

functions $\mathbb{E}e^{i\vartheta J_j} = \exp\Big[z_j(\cos\vartheta - 1)\Big]$, obtained from the difference of two independent Poisson variables of mean $z_j/2$. Each has modulus at most one, giving the bound.

□

Lemma B.39 (Geometric moments)**.** *For $0 < t < 1$ and nonnegative integers v, w, there is a constant $b_{vw} > 0$, independent of t, such that*

$$\left\|(\hat{f}^\dagger)^v t^{\hat{n}/2} \hat{f}^w\right\|_F^2 = \sum_{m\geq 0} t^m \frac{(m+v)!(m+w)!}{(m!)^2} \leq b_{vw}(1-t)^{-(v+w+1)}. \tag{B.108}$$

Proof. The factorial ratio is a degree-$(v+w)$ polynomial in m, so its geometric sum has the displayed order, with b_{vw} depending only on v, w. □

APPENDIX C

APPENDICES FOR CHAPTER 4

C.1 Notation and background

We start by introducing some notations. Throughout this work, $[n]$ is used to denote the set of integers $\{1, \ldots, n\}$. Without specification, we assume vectors are column vectors. For any vector $v \in \mathbb{F}_2^n$, its support is defined as $\operatorname{supp}(v) = \{i | v_i = 1\}$, where v_i denotes its entry at index i. The supports of a matrix, A, are defined correspondly as

$$\operatorname{rowsupp}(A) = \bigcup_i \operatorname{supp}\left(A_{:,i}\right) \tag{C.1}$$

$$\operatorname{colsupp}(A) = \bigcup_i \operatorname{supp}\left(A_{i,:}\right) \tag{C.2}$$

i.e. column support represent all the columns that any rows of A has nontrivial support. We follow the convention that $\operatorname{Span}\{A\}$ denotes the vector space spanned by row vectors of A. With a slight abuse of notation, we write for a given vector space V and linear map γ that $\gamma(V) := \{\gamma v | v \in V\}$.

A chain complex, $\mathcal{C}$, is a sequence of vector spaces $\{C_i\}$ with linear maps known as boundary operators, $\partial_i : C_i \to C_{i-1}$. A chain complex composed of $n+1$ vector spaces is called a length-n chain complex. The boundary maps satisfy the boundary condition $\forall i \in [n]$,

$$\partial_{i-1}\partial_i = 0. \tag{C.3}$$

The ith homology group is defined as $H_i(\mathcal{C}) := \ker \partial_i / \operatorname{im} \partial_{i+1}$. One can define the cochain complex and cohomology groups analogously. A chain complex is exact at C_i if $\dim H_i(\mathcal{C}) = 0$ or equivalently $\operatorname{im} \partial_{i+1} = \ker \partial_i$.

Since we mostly work in $\mathbb{F}_2$, the notation $|\cdot|$ is used to denote the Hamming weight of a vector. When acted on a matrix, it denote the maximum Hamming weight of all row and

column vectors, i.e. the maximum matrix degree.

A classical code is defined through a parity check matrix, H, and can be written as a length-1 chain complex as

$$C_1 \xrightarrow{\partial} C_0 \tag{C.4}$$

with $\partial = H$. The logical space is $\ker H$.

Section 4.1 gives the CSS chain-complex representation, with $\partial_2 = H_X^T$ and $\partial_1 = H_Z$. We use that convention below. The stabilizer groups of a code is defined as $S_X := \text{Span}\{H_X\}$ and $S_Z := \text{Span}\{H_Z\}$. The X-logicals are in $L_X = H_1(\mathcal{C})$. The logical dimension of the code is $k := \dim H_1(\mathcal{C})$, with the X-distance being

$$d_X = \min_{v \in H_1(\mathcal{C})} |v| \tag{C.5}$$

and d_Z is defined analogously for logical Z operators in the first co-homology group. These definitions correspond to the (co)systoic distances of a chain complex. The code distance is $d = \min(d_X, d_Z)$. When $k = 0$, we take $d \to \infty$ by convention.

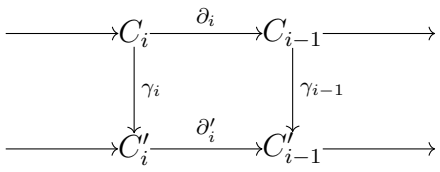
A quantum CSS subsystem code [210] is defined through a non-abelian gauge group, and the stabilizer group is the center of the gauge group. We can write a generating set for the stabilizer group similarly as H_X and H_Z, and the gauge logicals are generated by G_X and G_Z. Together, they generate the full gauge group. Without loss of generality, we let

$$G_{X,Z} H_{Z,X}^T = \mathbf{0} \tag{C.6}$$

$$G_X G_Z^T = I_g \tag{C.7}$$

where I_g has dimension equal to the number of gauge logicals. The dressed logicals, $L_{X,Z}$ are defined as the centralizer of $S_{X,Z}$ that are not in the gauge group. The dressed distance is the minimum weight of all dressed logicals.

One recurring component in our work is commuting diagrams, which describes relations between chain complexes. Consider the following diagram



(C.8)

and we say it is commuting if and only if

$$\gamma_{i-1}\partial_i = \partial'_i\gamma_i. \tag{C.9}$$

While being a simple result, we can conclude the condition for the existence of a commuting diagram in the following lemma.

Lemma C.1 (Condition for commuting diagrams)**.** *Given boundary operators ∂_i and ∂'_i, a commuting diagram exists if and only if either of the following holds,*

$$\gamma_i\left(\ker\partial_i\right) \subseteq \ker\partial'_i \tag{C.10}$$

$$\gamma_{i-1}\left(\text{im }\partial_i\right) \subseteq \text{im }\partial'_i. \tag{C.11}$$

Proof. We focus on the first condition, and the proof for the second proceeds analagously. We prove $\gamma_i\left(\ker\partial_i\right) \subseteq \ker\partial'_i$ is a necessary and sufficient condition for the existence of a γ_{i-1} that satisfies Eq. (C.9). The necessity is straightforward as for all $x \in \ker\partial_i$ we have

$$\partial'_i\gamma_i x = \gamma_{i-1}\partial_i x = 0, \tag{C.12}$$

so $\gamma_i x \in \ker\partial'_i$. The converse is shown constructively. Consider $v_{i-1} \in \text{im }\partial_i$ such that $v_{i-1} = \partial_i v_i$. The map is defined as $\gamma_{i-1}\left(v_{i-1}\right) := \partial'_i\gamma_1 v_i$. For $u_{i-1} \in C_{i-1} \setminus \text{im }\partial_i$, γ_{i-1} can be assigned arbitrarily. Such a γ_{i-1} is unique from the assumption, and it satisfies the commutation relations. □

With a commuting diagram, the maps γ_i form a chain map.

C.2 Expansion properties

In this work, the expansion properties of the parity check matrices are a key to guarantee its distance properties. There are many notions of expansions, and the one most relevant to our purpose is soundness defined as below. We start with a slightly generalized notion of reduced distance that will be useful later.

Definition C.2 (Reduced distance)**.** *The reduced distance of a vector $x \in U$ with respect to a subspace $V \subseteq U$ and a linear map $\gamma : U \to U'$ is defined as*

$$d_\gamma(x, V) := \min_{v \in V} |\gamma(x + v)| \tag{C.13}$$

where $|\cdot|$ denotes Hamming weight, and γ should be taken as the identity map when omitted.

A simple inequality follows as

$$d_\gamma(x, V) \leq |\gamma| \cdot d(x, V) \tag{C.14}$$

where $|\gamma|$ represents the weight of the matrix representation of γ.

Definition C.3 (Soundness; [167])**.** *A map, $H : V \to V'$, is (t, f)-sound for some integer t and a function $f : \mathbb{Z} \to \mathbb{R}$ if for all $x \in V$ and $|Hx| < t$, we have*

$$d(x, \ker H) \leq f(|Hx|) \tag{C.15}$$

In our interest, H is often the parity check matrix of a classical or quantum code. In this context, soundness have been shown to be sufficient conditions for single-shot error correction if t grows with n and f is a monotonically increasing function independent of n [167]. Here n denotes the system size.

The definition of soundness is later relaxed to confinement [211], where the weight

constraint is put on the error weight $|x| < t$ instead of syndrome weight. Confinement successfully explained single-shot properties of codes like quantum expander code [212], which are confined but not sound. Soundness is also closely related to (co-)boundary expansion notions of a chain complex. The major difference is in whether a limit is put on $|Hx|$ and if one restricts to only considering linear functions for f. While there is a deep literature on expansions properties, we focus on the relevant definitions only.

In general, for surgery schemes to fully preserve distance, it requires the strong condition of soundness, which is when f is a linear function. Moreover, while the above definition of soundness targets at a code family, we extend it to the linear soundness for a single code or linear map.

Definition C.4 (Linear soundness)**.** *A map $H : V \to V'$ has soundness ρ up to t if for all x with $|Hx| \leq t$, we have*

$$\rho \leq \frac{|Hx|}{d\left(x, \ker H\right)}. \tag{C.16}$$

C.3 General framework for code surgeries

In this section and the following few, we derive main theorems and lemmas for surgeries focusing on CSS data code and measurements of X-type logicals. While such scenarios are already of interest for many practical codes and circuits, we remark on how our results generalize to more general measurements in Appendix C.6.

From here on, we fix the key notation of data (memory) code and ancilla code as $\mathcal{C}$ and $\mathcal{A}$ for clarity. Other curly math symbols are used for classical or quantum codes depending on the context. We assume the data code has code parameters $[[n_C, k_C, d_C = d]]$.

C.3.1 General surgery procedure

Follow the initialization, merge, and split procedure in Section 4.4.4, using the merged-code diagram of Eq. (4.4). Its chain-map condition is $\Gamma_0 \partial_1^A = \partial_1^C \Gamma_1$. Initializing and measuring

ancilla qubits in the Z basis preserves the data Z-check detectors throughout the procedure. The following distance result supports that construction.

Lemma C.5 (Phenomenological distance is guaranteed by merge code distance; [7])**.** *Following the described surgery protocol, a sufficient merge code distance, $d_{merge} \geq d$, is sufficient and necessary condition for*

$$d_{ph} \geq d \tag{C.17}$$

where d_{ph} is the phenomenological distance.

We refer interested readers to Ref. [7] for a complete proof, but we remark that this is nontrivial since in general the merge code distance is only a necessary condition for the merged code distance. To guarantee the phenomenological distance, one needs to consider the subsystem spacetime code distance [172]. The proof of the dressed distance comes from the careful preservation of detectors following the designed procedures outlined above. With Lemma C.5, we implicitly assume the above procedure is followed in later discussions and focus on the merged code distance.

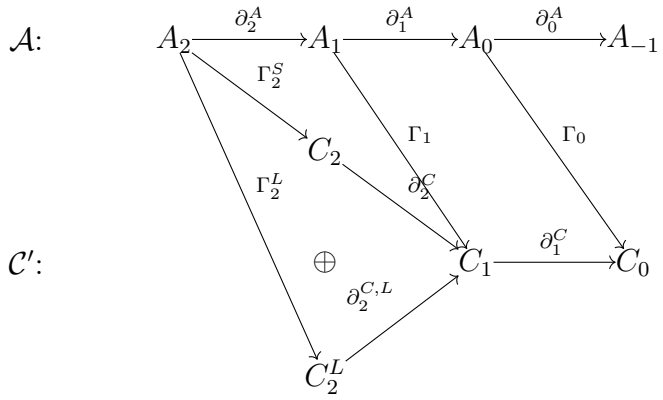
Remark C.6. *To guarantee fault-tolerance, the merge code should be LDPC as well [7]. Thus, the additional boundaries and chain maps, including $\partial_{1,2}^{A,C}$ and $\Gamma_{0,1}$, should be sparse. This is in general not guaranteed and where the challenge comes from. In Appendix C.4, we present constructive approaches to construct these sparse connections. In other sections, we assume this desiderata is satisfied.*

C.3.2 Measurements through metachecks

While the ancilla chain complex never exists by itself during surgery, its property is indicative and sufficient for characterizing the effect of the surgery. To start with, we add an additional vector space, A_2, to $\mathcal{A}$ known as the metacheck [167] such that $\mathcal{A} : A_2 \to A_1 \to A_0 \to A_{-1}$. The X-metachecks characterizes the deterministic outcomes in combination of X-checks, i.e.

the redundancies in the X-checks. While there is a degree of freedom to choose A_2 as long as the boundary condition is satisfied, we will choose it such that $\mathcal{A}$ is exact at A_1 by default.

The criticality of metachecks in A_2 rises from that we follow the surgery protocol where A_0 are initialized in Z-basis. As a result, by the stabilizer update rule, any X-stabilizer that has support in A_0 has random outcome at the initial merge step, providing no information on data code by itself. im ∂_2^A exactly represents the combination of X-checks that have no support in A_0. To understand what information one is measuring, we can draw the following commuting diagram,



(C.18)

where $\partial_2^{C,L} = \left(L_X^C\right)^T$. We have added C_2^L to represent the space of X-logicals so that the metacheck mapping relation is clear. The addition of C_2^L turn $\mathcal{C}'$ exact at C_1. The existence of $\Gamma_2^{S,L}$ are then guaranteed by the exactness of both chain complex and Lemma C.1.

It is now straightforward to understand why such metachecks enable the measurement of large-weighted X-logicals in C_2^L: instead of using the metachecks for correcting measurement errors in single-shot error correction [167], they are used to infer the values of C_2^L through Γ_2^L and measured parity checks of A_1, C_2. More concreteley, the measured operators are characterized through the following lemma, which is a generalization of Lemma 1 in Ref. [156]. This gives the measured-operator space stated in Lemma 4.2: metachecks select combinations of merged X checks with no ancilla support, leaving exactly the image $\Gamma_1[\ker \partial_1^A]$ on the data code.

Recall that the central idea of surgery schemes is to weight-reduce weight-$\Omega(d)$ logicals into weight-$\mathcal{O}(1)$ stabilizers. Under the reasonable assumptions ∂_2^C, Γ_1 in Eq. (C.18) are LDPC and that elements of $\operatorname{im} \partial_2^{C,L}$ have weight $\Omega(d)$, a necessary requirement is $\left|\partial_2^A\right| = \Omega(d)$, i.e. metachecks of $\mathcal{A}$ should be global. This rules out the potential of adopting codes whose metachecks are generated by local metachecks, such as the gauge color code [213, 214].

The efficiency of a surgery scheme is determined by the spacetime cost scaling per logical measurement. Most of the existing works are low-rate in this regards as their ancilla are composed of one or more disjoint ancilla systems with $\dim\left(\ker \partial_1^A\right) = 1$. This is a direct consequence of the kernel space of a connected graph being one-dimensional. A more formal definition of the efficiency through information extraction rate is given below. Use the information extraction rate of Definition 4.3. Under a $\Theta(d)$-weight logical basis, a low-rate surgery has $r_M = \mathcal{O}(1/d)$; a constant-rate surgery has $r_M = \Theta(1)$. The information extraction rate and the space overhead is directly connected by the number of logical measurements. A constant-rate surgery scheme leads to constant spacetime overhead, $\alpha = \Theta(1)$, regardless of number of measurements. While in the definition we assumed a minimum-weight logical basis for simplicity, it can be naturally generalized by consider low-rate schemes to be $r_M = \mathcal{O}(1/w)$ with $w := \min|l|$ for $l \in M_L$.

C.3.3 Merge code distance

In addition to the logical action, it is crucial to ensure the protocol is distance-preserved, which is determined by the merge code distance. In a surgery scheme that measures X-logicals following Eq. (4.4), the Z-distance is preserved regardless. On a high-level, the reason is that all X-checks of the data code, C_2 remained intact. Therefore, any dressed Z-logical's support on C_1 must be a Z-logical of the data code, which has sufficient distance by definition. If a Z-logical do not have support on C_1, it cannot be a dressed logical.

Lemma C.7 (Z-distance preservation; [7, 8, 150, 156])**.** *For a surgery scheme, the merge*

code Z-distance is always preserved,

$$d^Z_{merge} \geq d. \tag{C.19}$$

The X-distance is more challenging to prove. Before diving into the technical, it is important to interpret the logical space of $\mathcal{A}$. These logicals, if exist, are initialized deterministically in $|0_L\rangle^{\otimes k_A}$ and measured in Z-basis. Since they do not encode information, they are effectively gauge logicalsof the merge code. It is important to note that one generally cannot use these logicals to correct errors since they are of weight $\Omega(d_A)$. These gauge logicals can reduce the code distance by forming dressed logicals as we defined in Appendix C.1. Three remedies have been taken to treat these gauge logicals.

1. Gauging [7, 156]: the gauge degrees of freedom are promoted into large-weighted stabilizers. These stabilizer are then treated with techniques from quantum weight reduction [160–162].

2. Eliminating gauge logicals by construction: under certain assumptions on, e.g., the data code encoding and chain maps [150], one can provide explicit constructions of the ancilla system such that $k_A = 0$. In Methods, we outlined a more general approach where we do not place assumptions on the data code but only assume the logical basis takes on a canonical form. This approach also falls in this category.

3. Subsystem code [8, 10]: the merge code is accepted as a subsystem code where one focuses on boosting its dressed distance. A typical technique is to boost the initial ancilla system by thickening.

The caveat of the last approach is that the gauge logicals grow as one thickens, so the required number of thickened layers cannot be bounded by the expansion of the initial ancilla system as one does when there are no gauge logicals [150]. To measure a single logical, the first approach is asymptotically optimal up to polylogarithm factors, and it has shown

decent performance in practical regimes as well [140, 156]. However, it heavily relies on graph subroutines such as cellulation and decongestion [176], which are not straightforwardly extendable to hypergraphs. The second approach is tightly connected to the structure of the data code and logical basis chosen. As an example, when $\mathcal{C}$ has Z-metachecks, it is possible that the logicals of $\mathcal{A}$ are in turn measured through these Z-metachecks. This is straightforward to see if one consider the cohomology picture of Eq. (4.4), i.e. reverting all arrows. In this dual picture, $\mathcal{C}$ is the surgery ancilla system for $\mathcal{C}$ for measuring Z-logicals or Z-stabilizers, following Lemma 4.2. Therefore, some gauge logicals of $\mathcal{A}$ might be naturally elimiated. This is first proposed in Ref. [150] for when $\mathcal{C}$ is a HGP codes. While this is not uncommon to occur, we wish to focus on general data codes in this section, so we do not make similar assumptions. Therefore, in this work we mainly consider the last approach and work with subsystem codes with the same gauge logical number as k_A.

A general sufficient condition for the X-distance is most conveniently expressed with the guaged ancilla code and commuting diagram as following

$$\begin{array}{ccccc} A_1^g & \xrightarrow{\partial_1^{A,g}} & A_0 & \xrightarrow{\partial_0^A} & A_{-1} \\ & \searrow^{\Gamma_1^g} & & \searrow^{\Gamma_0} & \\ C_2 & \xrightarrow{\partial_2^C} & C_1 & \xrightarrow{\partial_1^C} & C_0 \end{array} \tag{C.20}$$

where $A_1^g = A_1 \oplus A_1^L$, and the new boundaries are

$$\partial_1^{A,g} = \begin{pmatrix} \partial_1^A & \left(G_A^X\right)^T \end{pmatrix}, \Gamma_1^g = \begin{pmatrix} \Gamma_1 & \Gamma_g \end{pmatrix}. \tag{C.21}$$

Here, G_A^X is a generating basis for the gauge X-logicals in A_1^L, and Γ_g is an induced chain map. The existence of this commuting diagram is similarly guaranteed by the exactness of chain complex, Lemma C.1, and the assumption that $\mathcal{C}$ has no Z-metachecks. Based on this gauged system, we arrive at a sufficient condition for general X-distance. To arrive at the concise lemmas, we relax the inequalities in multiple steps. We comment at the end of this

section how these overlooked structures can be crucial and utilized. We also remark on how our result connects to prior works, which are based on (relative) Cheeger constants [150, 157].

Theorem C.8 (X-distance)**.** *Consider the gauged boundary $\partial_1^{A,g}$ and the induced commuting diagram in Eq.* (C.20)*. The merge code X-distance satsify*

$$d^X_{merge} \geq \min\left(1, \frac{\rho}{\left|\Gamma_1^g\right|}\right) d. \tag{C.22}$$

if $\partial_1^{A,g}$ has soundness ρ up to $t \geq d$,

Proof. The distance proof is a generalization of past works but the underlying idea is very similar. An observation is that $\forall l_C^X \in L_C^X$, it stills commutes with all Z-checks in the merge code since the merge step does not apply additional Z-checks on C_1. Therefore, a bare logicals basis is given by $L_C^X \setminus M_L$, where M_L are the measured logicals such that $M_L \in S^X_{\text{merge}}$.

Given a bare logicals basis, the weight can be reduced by any combination of stabilizers and gauge logicals. Before merging, the gauge logicals have a generating basis of G_A^X by definition. After merging, the gauge logicals' supports extends to $A_0 \oplus C_1$ with a generating basis of $\begin{pmatrix} \left(G_A^X\right)^T \\ \Gamma_g \end{pmatrix}$. One can straightforwardly see that such a basis is not in the stabilizer group, S^X_{merge}, and it commutes with all Z-checks since Eq. (C.20) forms a commuting diagram. The ancilla X-stabilizers are similarly generated by a basis of $\begin{pmatrix} \partial_1^A \\ \Gamma_1 \end{pmatrix}$. Therefore, dressed logicals in the merged code have the following form. $\forall l^X_{\text{merge}} \in L^X_{\text{merge}}$, $\exists l_C^X \in L_C^X, x_s, x_g$ such that

$$l^X_{\text{merge}} = \begin{pmatrix} l_C^X \\ 0 \end{pmatrix} + \begin{pmatrix} \partial_1^A \\ \Gamma_1 \end{pmatrix} x_s + \begin{pmatrix} \left(G_A^X\right)^T \\ \Gamma_g \end{pmatrix} x_g \tag{C.23}$$

Therefore, the dressed logical weight is

$$\left|l^X_{\text{merge}}\right| = \underbrace{\left|l^X_C + \Gamma_1 x_s + \Gamma_g x_g\right|}_{\text{Supp. on } C_1} + \underbrace{\left|\partial^A_1 x_s + \left(G^X_A\right)^T x_g\right|}_{\text{Supp. on } A_0} \tag{C.24}$$

$$= \left|l^X_C + \Gamma^g_1 x\right| + \left|\partial^{A,g}_1 x\right| \tag{C.25}$$

where we define $x := \begin{pmatrix} x_s \\ x_g \end{pmatrix}$. Here we do not need to explicitly include stabilizers of $\mathcal{C}$ because we accept any logical representation of l^X_C. It is important to observe that this logical weight and the following derivation is equivalent with respect to any kernel vectors $u \in \ker \partial^{A,g}_1$: let $x = x' + u$, we have

$$\left|l^X_{\text{merge}}\right| = \left|l^X_C + \Gamma^g_1 x\right| + \left|\partial^{A,g}_1 x\right| = \left|l^{X,\prime}_C + \Gamma^g_1 x'\right| + \left|\partial^{A,g}_1 x'\right| \tag{C.26}$$

where $l^{X,\prime}_C = l^X_C + \Gamma^g_1 u$. Note that the gauge logicals generators are by default an independent basis, and it is not in the image of ∂^A_1. Thus, $\ker \partial^{A,g}_1 = \left\{ \begin{pmatrix} u_s \\ 0 \end{pmatrix} | u_s \in \ker \partial^A_1 \right\}$. Through Lemma 4.2, we know $\Gamma_1\left(\ker \partial^A_1\right) \subseteq S^C_X \oplus L^C_X$. Thus, $l^{X,\prime}_C \in L^C_X$. Since $\left|l^X_C\right|, \left|l^{X,\prime}_C\right| \geq d$ by definition of distance, to prove $\left|l^X_{\text{merge}}\right| \geq d$ it is sufficient to discuss when $\left|\Gamma^g_1 x\right| = d_{\Gamma^g_1}\left(x, \ker \partial^{A,g}_1\right)$. Then, we have

$$\left|l^X_{\text{merge}}\right| = \left|l^X_C + \Gamma^g_1 x\right| + \left|\partial^{A,g}_1 x\right| \tag{C.27}$$

$$\geq \left|\left|l^X_C\right| - \left|\Gamma^g_1 x\right|\right| + \left|\partial^{A,g}_1 x\right| \tag{C.28}$$

$$= \left|\left|l^X_C\right| - d_{\Gamma^g_1}\left(x, \ker \partial^{A,g}_1\right)\right| + \left|\partial^{A,g}_1 x\right| \tag{C.29}$$

One have already, $\left|l^X_{\text{merge}}\right| \geq \left|\partial^{A,g}_1 x\right|$, so we only require below the soundness condition to hold for $\left|\partial^{A,g}_1 x\right| < d$. Now we discuss case by case.

1. When $\left|l_C^X\right| < d_{\Gamma_1^g}\left(x, \ker \partial_1^{A,g}\right)$, we have

$$\left|l_{\text{merge}}^X\right| \;>\; \left|\partial_1^{A,g} x\right| \geq \rho \cdot d\left(x, \ker \partial_1^{A,g}\right) \geq \frac{\rho}{\left|\Gamma_1^g\right|} d \tag{C.30}$$

where the inequalities follow from the definition of soundness and Eq. (C.14)

2. When $\left|l_C^X\right| \geq d_{\Gamma_1^g}\left(x, \ker \partial_1^{A,g}\right)$, we have

$$\begin{aligned} \left|l_{\text{merge}}^X\right| &\geq \left|l_C^X\right| - d_{\Gamma_1^g}\left(x, \ker \partial_1^{A,g}\right) + \left|\partial_1^{A,g} x\right| && \text{(C.31)} \\ &\geq d - \left(\left|\Gamma_1^g\right| - \rho\right) d\left(x, \ker \partial_1^{A,g}\right) && \text{(C.32)} \end{aligned}$$

Note that to expand the absolute value, we have imposed the initial assumption and that $d - \left|\Gamma_1^g\right| d\left(x, \ker \partial_1^{A,g}\right) \geq 0$, leading to $d\left(x, \ker \partial_1^{A,g}\right) \leq \frac{d}{\max\left(\rho, \left|\Gamma_1^g\right|\right)}$. Thus, we arrive at

$$\left|l_{\text{merge}}^X\right| \geq \begin{cases} \frac{\rho}{\left|\Gamma_1^g\right|} d, & \text{if } \left|\Gamma_1^g\right| \geq \rho \\ \frac{2\rho - \left|\Gamma_1^g\right|}{\rho} d, & \text{if } \left|\Gamma_1^g\right| < \rho \end{cases} \tag{C.33}$$

The second case is guaranteed to be larger than d.

Therefore, above concludes that we can guarantee

$$d_{\text{merge}}^X \geq \min\left(1, \frac{\rho}{\left|\Gamma_1^g\right|}\right) d. \tag{C.34}$$

□

While the above considers codes with linear soundness ρ, it is straightforward to generalize the result to good but nonlinear soundness as in Definition C.3.

Corollary C.9 (X-distance under general soundness)**.** *Assume the boundary* $\partial_1^{A,g}$ *is* (t, f)-

sound for $t \geq d$, the merge code X-distance satsify

$$d^X_{merge} \geq \min\left(d, f^{-1}\left(\frac{d}{\left|\Gamma^g_1\right|}\right)\right) \tag{C.35}$$

where f is assumed to be monotoically increasing.

Proof. The proof proceeds largely the same as in Theorem C.8. The major difference is that we replace the syndrome weight constraints with

$$\left|\partial^{A,g}_1\right| \geq f^{-1}\left(d\left(x, \ker \partial^{A,g}_1\right)\right) \geq f^{-1}\left(\frac{1}{\left|\Gamma^g_1\right|} d\left(x, \ker \partial^{A,g}_1\right)\right). \tag{C.36}$$

The invertibility of f is guaranteed by it being monotonically increasing. The corollary then follows the same case-by-case discussion. □

The above lemma puts a constraint on the ratio between soundness and weight of $\Gamma^g_1 = \begin{pmatrix}\Gamma_1 & \Gamma_g\end{pmatrix}$. However, this is not very useful in general since while $\rho = \mathcal{O}(1)$ by the constraint of LDPC ∂^A_1, there is no constraint on $\left|\Gamma_g\right|$. The following corollary replaces the constraint on $\left|\Gamma_g\right|$ by an assumption on the X-distance of $\mathcal{A}$.

Corollary C.10 (X-distance with sufficient ancilla X-distance)**.** *Assume the boundary ∂^A_1 has soundness ρ up to $t \geq d$, the merge code X-distance satsify*

$$d^X_{merge} \geq \min\left(1, \frac{\rho}{|\Gamma_1|}\right) d. \tag{C.37}$$

if the ancilla code, $\mathcal{A}$, has sufficient X-distance, $d^X_A \geq d$.

Proof. The proof extends from Theorem C.8. The key difference is that $\forall x \in A^g_1$, we have

$$\partial^{A,g}_1 x = \begin{pmatrix}\partial^A_1 & \left(G^X_A\right)^T\end{pmatrix}\begin{pmatrix}x_{A_1} \\ x_g\end{pmatrix} = \partial^A_1 x_{A_1} + \left(G^X_A\right)^T x_g \tag{C.38}$$

where we have assumed $x = \begin{pmatrix} x_{A_1} \\ x_g \end{pmatrix}$ with $x_{A_1} \in A_1, x_g \in A_1^L$. Consider the case where $\partial_1^{A,g} x$ is equivalent to a gauge logical up to stabilizer group, i.e. $x_g \neq 0$. The gauge X-logicals are the X-logicals of the ancilla code $\mathcal{A}$ viewed independently. Therefore, under the assumption $d_A^X \geq d$, we have by definition

$$\left|\partial_1^{A,g} x\right| \geq d. \tag{C.39}$$

Therefore, we no longer need to consider dressed logicals, i.e. one only needs to prove the minimum weight restricted to $x = \begin{pmatrix} x_{A_1} \\ 0 \end{pmatrix}$. This effectively leads us back to the scenario where the code $\mathcal{A}$ has no gauge logicals, so the distance can be guaranteed by the expansion of the added X-checks. The corollary then follows analagously. □

Notably, under the assumption that $|\Gamma_1|$ is LDPC, Corollary C.10 is built upon conditions solely on the independent ancilla code. From a high-level, our condition presents a general desiderata sufficient for distance-preservation: the ancilla code should have sufficient X-distance, and the transpose of its X-check should have constant classical soundness. The condition on ancilla distance is not very restrictive since it is only single-sided. This enable the opportunity to consider codes with guaranteed distance in one side only, which is natural in code constructions such as lifted product codes [179, 188], fiber bundle codes [180], and others [215, 216] before distance balancing. Moreover, the search for expanding classical and quantum codes is an active direction of research [46, 217–220]. We highlight that the sufficient condition derived is distinct from the desiderata for conventional expanding codes and could motivate future research directions.

One should note that Corollary C.10 is a sufficient condition with relaxations. An example is that we have applied the inequality in Eq. (C.14) multiple times. In fact, in prior works and some constructions we present in Appendix C.4, Γ_1 only has support on a limited subset of A_1. In these cases, it is unnecessary to require global expansion and local expansion is

sufficient. When one narrow down to considering low-rate surgery and transversal-type chain maps $\Gamma_1 = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, this condition is equivalent to the relative Cheeger constant condition in Ref. [157].

Another piece of information that could greatly relax our conditions is the structure of the logicals. Here, structure refers to both structure in the logical basis and in the chosen set of logical measurements, M. In many cases, especially when there is translation symmetry or product structures, these information reduces the overhead greatly. Some well-known examples include logical measurements of surface code and toric codes through lattice surgery [147], which require almost minimal ancilla overhead. The ancilla system does not have sufficient global or local expansion, but it is still distance-preserving. In Appendix C.4, we utilize such structures in the data code to demonstrate examples when the requirement on the expansion property in Corollary C.10 can be completely removed.

C.4 High-rate surgery through structured ancilla

In this section, we utilize the structure in the ancilla and/or the data code to (1) reduce overhead and (2) provide more concrete and constructive schemes for high-rate surgery. In Appendix C.4.1, we specify the ancilla construction as a HGP code, which is an example where the ancilla code satisfy the desiderata listed in Corollary C.10 while being high-rate. In Appendix C.4.2, we further specify the memory to be encoded in a HGP code. Such assumption largely alleviates the requirement on expansion, which is often argued in the asymptotic limit.

While our discussions focus on HGP codes, a large portion of them can be applicable to other codes, specifically those that bears a similar product structure, such as LP code [179], fiber bundle code [180], and balanced product codes [181]. The key reason we perform the analysis on HGP codes is that its distance is well understood and possess a known complete logical basis, both of which are lacking in the other codes with better code parameters in general. We present some progress in this direction in Ref. [188] and leave the adaptation of

discussions in this sections to those codes to future work.

C.4.1 Surgery on general data code

We begin with quoting some well-known results from classical coding theory, which constructs expander codes from left-expanding bipartite graphs. We only outline the most relevant result here for simplicity.

Definition C.11 (Expander codes; [212, 221])**.** *A (γ, δ)-expanding code with $\delta < 1/2$ is a classical LDPC code with parameters $[n, k = \Theta(n), d \geq \gamma n]$. For any subset of $S \subseteq F_2^n$ such that $|S| \leq \gamma n$, we have*

$$\Gamma_u(S) \geq (1 - 2\delta)\Delta|S| \tag{C.40}$$

where Δ is the column weight of its biregular parity check matrix. $\Gamma_u(S)$ denotes the number of checks only connected to one qubit in S.

Next, we formally define the product construction. In particular, we first define the product complex and the hypergraph product code [45, 165, 222, 223].

Definition C.12 (Product complex and hypergraph product code)**.** *Consider two chain complex, $\mathcal{B}, \mathcal{D}$, with boundaries $\partial^{B,D}$, the product complex is defined as $\mathcal{P} = Prod(\mathcal{B}, \mathcal{D})$ with*

$$P_k := \bigoplus_{i+j=k} B_i \otimes D_j. \tag{C.41}$$

$\forall b_i \otimes d_j \in P_k$, the boundary maps act as $\partial_k^P \left(b_i \otimes d_j\right) = \partial_i^B b_i \otimes d_j + b_i \otimes \partial_j^D d_j$. A (2-dimensional) hypergraph product code (HGP) is a length-2 complex given by $HGP(\mathcal{B}, \mathcal{D}) := Prod(\mathcal{B}, \mathcal{D})$ with $\mathcal{B}, \mathcal{D}$ being length-1 complexes, i.e. classical codes.

The definition of product complex and HGP can be naturally extended to higher dimensions to homological product codes [222, 223], but we restrict our discussion to HGP and discuss high-dimensional extensions when possible. The HGP diagram, parameters,

and transpose-code convention appear in Section 4.4.3. The general product-complex definition above also applies in higher dimensions; the expansion proof below uses its two-term subcomplex.

Consider a HGP code, $\mathcal{P} = \mathrm{HGP}(\mathcal{B}, \mathcal{D})$, and restrict to the subcomplex, $P_2 \longrightarrow P_1$. This exactly corresponds to a tensor product code in classical codes [224, 225]. We denote a complex with this structure as $\mathrm{Tensor}(\mathcal{B}, \mathcal{D})$. While tensor product codes are well-known examples of robustly testable codes under certain assumptions [226], we derive the following expansion property for our purpose.

Lemma C.13 (Soundness properties of tensor product codes)**.** *Consider a tensor product code, $\mathrm{Tensor}(\mathcal{B}, \mathcal{D})$, its parity check is (t, f)-sound with $f(x) = x^2/4$ and $t = \min(d_B, d_D)$ [167]. Under the assumption that at least one of $\mathcal{B}, \mathcal{D}$ is a (γ, δ)-expanding code, the parity check of $\mathrm{Tensor}(\mathcal{B}, \mathcal{D})$ has linear soundness of*

$$\rho = (1 - 2\delta)\,\Delta \tag{C.42}$$

up to $t = \gamma n_{D_1}$, where n_{D_1} denotes the number of bits in $\mathcal{D}$ and Δ is the weight of ∂^D.

Proof. The first part of the lemma was proved in Ref. [167] to show single-shot properties of two- and high-dimensional hypergraph product codes. We provide a rough sketch of the idea here. The bits of a tensor code can be equivalently considered to be in $F_2^{n_{B_1} \otimes n_{D_1}}$ up to rearrangement. In this representation, the bits form a matrix of size $n_{B_1} \times n_{D_1}$. The checks are a direct sum of row checks and column checks. Consider an error configuration $e \in F_2^{n_{B_1} \otimes n_{D_1}}$, with equivalent row and column vector forms

$$e = \begin{pmatrix} c_1 & c_2 & \dots & c_{n_{B_1}} \end{pmatrix} = \begin{pmatrix} r_1^T \\ r_2^T \\ \vdots \\ r_{n_{A_1}}^T \end{pmatrix}. \tag{C.43}$$

Suppose the boundary operator of Tensor $(\mathcal{B}, \mathcal{D})$ is ∂. The syndrome is given by

$$\partial e = \begin{pmatrix} \partial^B c_1 & \partial^B c_2 & \dots & \partial^B c_{n_{B_1}} \end{pmatrix} \oplus \begin{pmatrix} \partial^D r_1 \\ \partial^D r_2 \\ \vdots \\ \partial^D r_{n_{A_1}} \end{pmatrix} \tag{C.44}$$

A case-by-case discussion follows as below.

1. If none of the columns and rows of e are in $\ker \partial^{B,D}$ respectively, the syndrome size is lower bounded by the sum of its column and row support size, while the error weight is upper bounded by the product of its column and row support size. This leads to

$$|e| \leq |\text{rowsupp}(e)||\text{rowsupp}(e)| \leq \frac{1}{2}\sqrt{|\text{rowsupp}(e)| + |\text{colsupp}(e)|} \leq \frac{1}{2}\sqrt{|\partial e|}, \tag{C.45}$$

2. If certain columns of e are in $\ker \partial^B$ but none of its rows are in $\ker \partial^D$, since the columns are of weight at least d_B, the row syndrome size is lower bounded by d_B. The symmetric case where rows of e are in $\ker \partial^D$ is argued similarly.

3. If there exists a pair of index (i, j) with $i \in [n_{B_1}], j \in [n_{A_1}]$ such that both $r_i \in \ker \partial^A$ and $c_j \in \ker \partial^B$, we can consider an equivalent configuration $e' = e + r_i \otimes c_j$. Considering these two error configurations are equivalent under soundness analysis because it is only concerned with the reduced distance, $d(e, \ker \partial)$ with $\ker \partial = \ker \partial^A \times \ker \partial^B$. After applying one such transformation for each paired index (i, j), e' will reduce to either case (1) or (2).

The conclusion from the above three cases lead to the result presented in the first half of the lemma.

If additional expansion assumption is made on either of the codes, the soundness is boosted to linear scaling. Without loss of generality, consider $\mathcal{D}$ to be a (γ, δ)-expanding

code as defined in Definition C.11. Previous discussion in case 2 and 3 follow identically. Under the assumption of case 1, we note that

$$|\partial e| \geq |\text{rowsupp}(e)| \geq \max_i |r_i|. \tag{C.46}$$

Thus, if we restrict to considering $\forall r_i, |r_i| \leq t = \gamma n_{B_1}$, the error weight is upper bounded as $\partial^B r_i \geq (1-2\delta)\,\Delta |r_i|$, which leads to

$$|\partial e| \geq \left| \begin{pmatrix} \partial^B r_1 \\ \partial^B r_2 \\ \vdots \\ \partial^B r_{n_{A_1}} \end{pmatrix} \right| \geq (1-2\delta)\,\Delta |e|. \tag{C.47}$$

Here, we have used $|e| = \sum_i |r_i|$. This inequality immediately implies linear soundness with

$$\rho = (1-2\delta)\,\Delta \tag{C.48}$$

up to t. □

The above construction might seem similar to quantum expander codes [212], which are known to be confined but not sound. However, the difference lies in that their construction, in our notation, is HGP $(\mathcal{B}, \mathcal{D})$ with $\mathcal{B}, \mathcal{D}^T$ both based on expander graphs. Moreover, the expansion property of interest to them are the X, Z check matrices, while our soundness condition is on the transpose of the X check matrix only.

Theorem C.14 (HGP ancilla with expander code is distance-preserving)**.** *Consider a surgery procedure equipped with a LDPC commuting diagram and a data code, $\mathcal{C}$, with distance d. Let the ancilla code $\mathcal{A} = HGP(\mathcal{B}, \mathcal{D})$, where the classical codes have parameters $[n_\alpha, k_\alpha, d_\alpha \geq d]$ with $\alpha = B, D$. Suppose $\mathcal{D}$ is a (γ, δ)-expanding code with soundness $\rho = (1-2\delta)\,\Delta$ up to γn_{D_1}, where n_{D_1} denotes the number of bits in $\mathcal{D}$ and Δ is the weight of ∂^D. The merge code*

distance is

$$d_{merge} \geq \min\left(d, \frac{(1-2\delta)\,\Delta}{|\Gamma_1|} d, \gamma n_B\right) = \Theta\left(d\right) \tag{C.49}$$

Proof. The proof is a straightforward combination of Lemma C.13 and Corollary C.10. □

By far, we have derived the conditions for distance-preservation with a HGP ancilla. However, it is based on the assumption that there exists a LDPC commuting diagram composed of $\Gamma_{0,1}$ given a data code $\mathcal{C}$ and logical measurements M. It is nontrivial to construct such diagrams. In the following, we provide examples on how to construct these diagrams and what logical measurements can be performed in these cases. With a HGP ancilla code, $\mathcal{A} := \text{Tensor}\left(\mathcal{B}, \mathcal{D}\right)$, the X metachecks are given by the codewords of the tensor product code, i.e.

$$\ker \partial_1^A = \ker \partial^B \otimes \ker \partial^D. \tag{C.50}$$

An example of structured logicals one can measure is when the following condition holds.

Definition C.15 (Logical operators with generalized parallel structure)**.** *Suppose the set of logical operators to measure is M. A partition of M is collection of subsets M_i with $i \in [l]$ satisfying $M_i \subseteq M$ and $\bigcup_i M_i = M$. We say M possess a generalized parallel structure if $\exists \mathcal{B}, M_i, \gamma_1^i$ such that $\forall i \in [l]$*

$$\gamma_1^i\left(\ker \partial^B\right) = \mathit{Span}\left\{M_i\right\} \tag{C.51}$$

Recall that if Eq. (C.51) holds, a commuting diagram follows from Lemma C.1, i.e. $\exists \gamma_0^i$ with

$$\partial_1^C \gamma_1^i = \gamma_0^i \partial^B. \tag{C.52}$$

Given the maps $\gamma_{0,1}^i$, one can construct the commuting diagram as

$$\Gamma_{0,1} = \sum_{i \leq k_D} \gamma_{0,1}^i \otimes \left(e_i^D\right)^T \tag{C.53}$$

where $e_i^D \in D_1$ is a unit vector with its nonzero element on the i-th information bit of code $\mathcal{D}$. One can verify with Lemma 4.2 that this connection indeed measures the full set M provided that the number of subsets is upper bounded as $l \leq k_D$. The above construction utilize the property that each logical of $\mathcal{D}$ in the canonical form have a unique support on its information bits. Adopting the canonical basis here is only for the purpose of extracting logical information and does not pose restrictions on $\gamma_{0,1}^i$ or the data code.

The above structure is generalized compared to strictly parallel structures. Those strictly parallel structures have been explored previously in the context of homomorphic measurements [163] under the name of grid PPMs (GPPMs) for the measurement of a HGP data code with a HGP ancilla code. We give more detailed comparisons in Appendix C.4.2 when we narrow down to HGP data codes.

To analyze the overhead more systematically, we consider the optimal and worst case scenario. Since $\mathcal{D}$ is asymptotically good and can be chosen with parameters $[n_D = \Theta(d), k_D = \Theta(d), d_D = d]$, the spacetime overhead of measuring t logicals is generally given by $\alpha = \mathcal{O}\left(\frac{n_C + n_B n_D}{k_C + t} \frac{t}{k_B l}\right)$. Notice that $n_B \in [d, n_C]$, depending on the choice of logical basis and the specific logical measurements. In the optimal scenario, the logicals can be partitioned such that each subset $|M_i| = \Theta(n_B) = \Theta(d_C)$,

$$\alpha_{\text{opt}} = \mathcal{O}\left(\frac{d_C^2}{n_C}\right) \tag{C.54}$$

where we have assumed $t = \Theta(k_B l) = \mathcal{O}(k_C)$ and $k_C = \Theta(n_C)$ for simplicity. We will see later that this holds for HGP data codes. In the worst scenario, we have $l = 1$ and $k_B = t$. $\mathcal{D}$ can be chosen as a repetition code with sufficient distance, which leads to an overhead similar to the devised sticker approach in Ref. [10], where we generally expect $n_A \sim n_C$ and

an overhead of

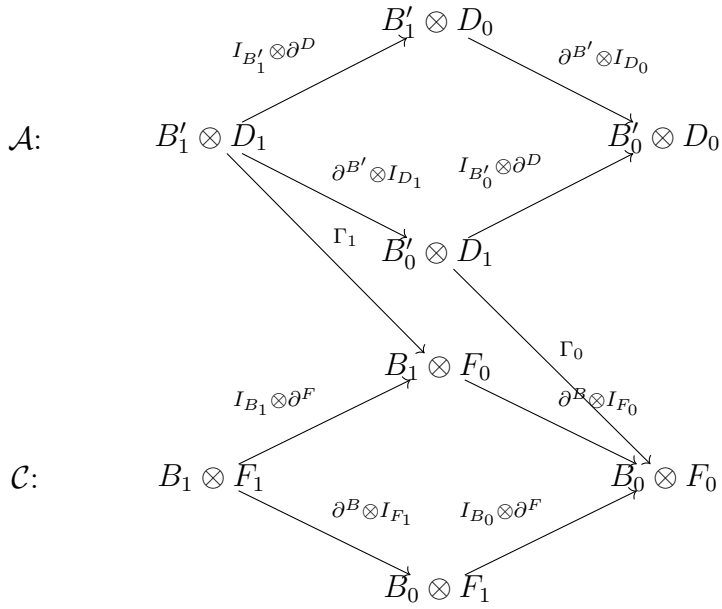
$$\alpha_{\text{worse}} = \mathcal{O}\left(d_C\right). \tag{C.55}$$

The above construction still requires an expanding code to form the HGP ancilla. While this is acceptable as an asymptotic argument, it is not necessarily practical for near-term implementations since the hidden constant factors might be daunting. One example where the expansion requirement can be alleviated is when M only consists of a single element and we construct $\mathcal{B}$ with $k_B = 1$. Here, the ancilla X-check A_1 is connected to the data code at C_1 only through one non-zero entry per row, thus the relative expansion of ∂_1^A is at least 1. Therefore, merge code distance is guaranteed whenever $\mathcal{D}$ has sufficient distance. This is precisely why past schemes that rely on thickening works. However, this argument is restricted to low-rate surgery. In the following, we provide a high-rate surgery example.

C.4.2 Surgery on Hypergraph Product Codes

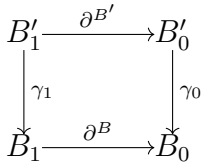
In the last subsection, we refrained from specifying the data code for maximum generality. In this subsection, we consider specifically the HGP code as data code to give explicit constructions. In particular, we consider the case where $\mathcal{C} = \text{HGP}\left(\mathcal{B}, \mathcal{D}\right)$ and the ancilla

code is written as $\mathcal{A} = \mathrm{HGP}\,(\mathcal{B}', \mathcal{F})$. Then, we have the commuting diagram



(C.56)

Here, we construct the ancilla code such that $\mathcal{B}'$ and $\mathcal{B}$ are connected through code homomorphism,



(C.57)

where $\gamma_{0,1}$ are LDPC. This is equivalent to a general requirement of

$$\gamma_0 \partial^{B'} = \partial^B \gamma_1 \tag{C.58}$$

and the logical action is defined through matrix w with

$$G_{B'} \gamma_1^T = w G_B \tag{C.59}$$

which depends on the generator basis $G_{B,B'}$ chosen. For simplicity of discussion, we consider a canonical basis of HGP, where d_B number of X-logicals are solely supported in each column

of the labelled by each information bit of $\mathcal{F}^T$. This is also a minimum weight basis as all logicals' weights are $\Theta(d_B)$ if $\mathcal{B}$. The disjoint column support structure allows us to tailor the connectivities column-by-column. More concretely, we implement the full chain map as

$$\Gamma_{0,1} = \sum_i \gamma_{0,1} \otimes e_i^{F^T} \left(e_i^D\right)^T \tag{C.60}$$

where $e_i^{F^T}$ and $e_i^{F^T}$ denote the unit vector on the i-th information bit on $\mathcal{D}$ and $\mathcal{F}^T$, respectively. Suppose the implementable logical actions are $W = \{w\}$, the generalized parallel structure in Definition C.15 implies that we can simultaneously measure all logical set in each column if they are connected by logical actions in W. In other words, logicals to measure in the i-th column are simultaneously measurable if they are connected to a common logical basis, $G_{B'}$, through tailored $w_i \in W$. The key is to realize the connectivity can vary from column to column.

While our scheme has great flexibility, it remains unknown if it is possible to construct LDPC $\gamma_{0,1}$, or compile in constant-depth, for arbitrary logical codewords in each column through w_i. However, it is sufficient to see that the power is beyond prior works through several examples. In below we give several examples and compare with relevant literatures.

Example C.16 (Code modifications)**.** *Assume we work in the canonical basis of the classical code $\mathcal{B}$ in each column, the code puncturing and augmentation techniques induce classical code homomorphisms. Through these classical code homomorphisms, we can measure M if its logicals form a hypergraph product pattern. Such a pattern correspond exactly to the GPPMs pattern in Ref. [163].*

If one restrict to puncturing and augmentation, our scheme's capability is equivalent to the homomorphic measurement scheme in Ref. [163]. There are some constant factor gains because in their scheme they require an additional ancilla patch for masking to complete the full grid PPM measurements while we do not. However, the asymptotic scaling is the same.

Example C.17 (Code automorphisms)**.** *Consider the automorphism group associated with*

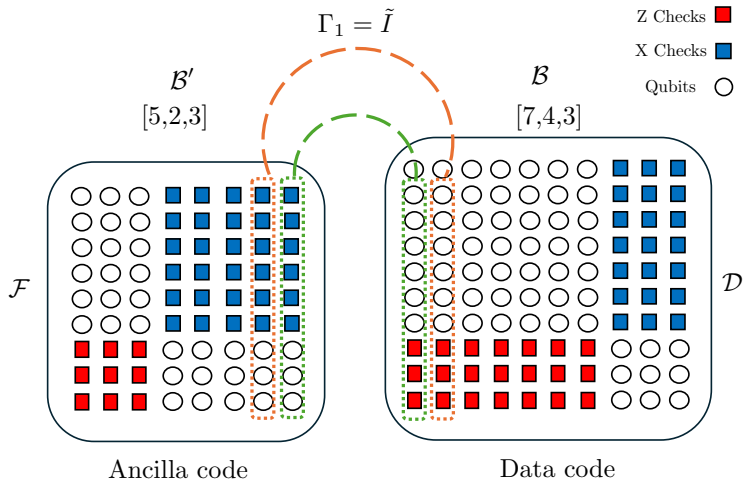


Figure C.1: Illustrative figure for when the data code is a HGP code. Here the data code is formed by the hypergraph product of two $[7, 4, 3]$ Hamming codes on the right. The ancilla code, on the left, is formed by transpose checks of $[6, 3, 3]$ and $[5, 2, 3]$ on the vertical and horizontal sides respectively. The $[6, 3, 3]$ code is chosen by puncturing the Hamming code.

$\mathcal{B}$, which is equivalent to the following conditions

$$G_B\sigma = wG_B \tag{C.61}$$

$$\partial^B\sigma = v\partial^B \tag{C.62}$$

generated by permutations $\sigma \in S_{n_B}$ and $w \in GL_{k_B}(\mathbb{F}_2)$. Our scheme can simulatenously measure logicals in each column if they are connected by automorphisms.

Quantum and classical code automorphisms have been extensively studied [143, 144, 177, 178, 227]. In particular, prior works [177, 178] have considered explicitly how (subsystem) HGP codes inherit the classical code automorphisms, which enables efficient computation when codes with rich symmetry are adopted as base codes. For example, in subsystem HGP (sHGP) the automorphism group is inherited from classical base codes as [177]

$$\mathrm{Aut}(\mathcal{B}) \otimes \mathrm{Aut}\left(\mathcal{D}^T\right) \subseteq \mathrm{Aut}(\mathrm{HGP}(\mathcal{B}, \mathcal{D})) \tag{C.63}$$

The choice of sHGP, at the cost of a lack of threshold, is to preserve the tensor product

structure. In HGP, while the automorphism group size is similarly $|\mathrm{Aut}(\mathcal{B})| \cdot |\mathrm{Aut}(\mathcal{D})|$ [178], not all automorphisms are guaranteed to preserve the tensor product structure. Combining the automorphism gadgets with transversal or homomorphic measurements, the number of distinct measurements one can perform scales as $\mathcal{O}\left(|\mathrm{Aut}(\mathcal{B})| \cdot |\mathrm{Aut}(\mathcal{D})|\right)$. In our scheme, since we can freely apply distinct classical automorphisms for each of the k_D columns, the distinct measurements scale as $\mathcal{O}\left(|\mathrm{Aut}(\mathcal{B})|^{k_D} \cdot |\mathrm{Aut}(\mathcal{D})|\right)$.

While the presented surgery scheme can already utilize the underlying code's automorphism group more flexibly than direct inerhitence, it is unnecessarily restrictive to only consider automorphisms. One limitation of automorphism gadgets [177, 178] is that they are not universal since they are restricted to affine logical actions, $w \in \mathrm{GL}_{k_B}(\mathbb{F}_2)$. However, our scheme is not limited by such rules: one can simply compose the automorphism logical actions to go beyond affine gates. For example, assume $\mathcal{B}$ has $k_B = 3$ and its automorphism group include a logical permutation action $w \in \mathrm{Aut}(\mathcal{B})$ such that

$$w = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{C.64}$$

induced by physical permutation matrices σ. We can construct chain maps by linearity as $G_B\left(\sigma + I_{n_B}\right) = \left(w + I_{k_B}\right) G_B$, which realizes a logical action of

$$w + I_{k_B} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{C.65}$$

This specific logical measurement performs effectively a repeated measurement of the joint logicals, $L_1^X L_2^X$ and leaves L_3^X intact. Through simple compositions as such, we arrive at logical actions no longer in the automorphism group. Moreover, this provides an example where we tailor the number of logicals measued for each column through chain maps with

nontrivial kernels. Notably, if we only rely on automorphism plus classical code modifications, the number of logicals measured per column must stay the same since automorphisms are invertible.

Last but not least, we remark that HGP as data code do not require the ancilla HGP code to have an expander base code, which present it as a practical and explicit near-term code example.

Theorem C.18 (HGP ancilla is distance-preserving for HGP data code)**.** *Consider a surgery procedure equipped with a LDPC commuting diagram exists and the data code is a HGP code* $\mathcal{C} = HGP(\mathcal{B}, \mathcal{F})$ *with distance* d*. Let the ancilla code be* $\mathcal{A} = HGP(\mathcal{B}', \mathcal{D})$*. If* $d_D \geq d$*, we have*

$$d_{merge} \geq \frac{1}{|\Gamma_1|} d \tag{C.66}$$

Proof. The proof relies on the special structure of the HGP code and the way the connectivities are set. Recall that one can define the canonical basis of the classical code $\mathcal{F}^T$ and its information set, $\mathcal{I}_{F^T}$. The canonical basis for the X-logicals of the HGP $\mathcal{C}$ is spanned by logical operators of the form [146, 228]

$$l = l_B^i \otimes e_i^{F^T}, \tag{C.67}$$

where $l_B^i \in \ker \partial^B$, and $e_i^{F^T}$ are unit vectors in $\mathcal{I}_{F^T}$. The considered logical is fully contained in the i-th column.

Consider the logical in canonical basis, we have

$$\text{rowsupp}\,(l) \geq d \tag{C.68}$$

and by the form of Eq. (C.60), the column i in $B_0 \otimes F_1$ is connected to only one column in $B_1' \otimes D_1$. Key is to note that given such connection, suppose we consider a stabilizer

combination $x \in B_1' \otimes D_1$. The logical is now

$$\begin{aligned} \left|l'\right| &\geq |l + \Gamma_1 x| + \left|\left(I_{B_1'} \otimes \partial^D\right) x\right| && \text{(C.69)} \\ &\geq |\text{rowsupp}\,(l) - |\Gamma_1|\text{rowsupp}\,(x)| + \text{rowsupp}\,(x) && \text{(C.70)} \\ &\geq \frac{1}{|\Gamma_1|} d && \text{(C.71)} \end{aligned}$$

where we did not use any expansion property. All that is assumed is that for any row of x not in kernel space of ∂^D, the incurred qubit support is non-zero $\left|\partial^D x|_j\right| \geq 1$.

To prove distance, one would need to consider combinations of any logical with the stabilizer group. One can first notice that considering joint paulis from different columns would not decrease weight since they are fully disjoint. The X-check matrix, restricted to $B_0 \otimes F_1$, is written as

$$H_X|_{B_0 \otimes F_1} = e^j \otimes \partial^{F^T} \tag{C.72}$$

where in the canonical basis

$$\partial^{F^T} = \begin{pmatrix} A & I' \end{pmatrix} \tag{C.73}$$

with I' being an identity matrix of size $n_{F^T} - k_{F^T}$. Therefore, if one add a combination of checks, s, to row j of the logical, it will incur support on the $n_{F^T} - k_{F^T}$ qubits that are the complementary of the information set. The weights of these qubits cannot be reduced by stabilizers in $B_0 \otimes F_1$ since im $\Gamma_1 \cap (\mathcal{I}_{F^T})^c = \varnothing$. □

Consider a HGP code with good base codes such that the data code parameters are $[[n_C, k_C = \Theta\,(n_C)\,, d_C = \Theta\left(\sqrt{n_C}\right)]]$. If we adopt the canonical basis, for each simultaneously measurable logical set defined by Definition C.15, the logicals partitions in each column occupy at most $\mathcal{O}\,(d_C)$ space. Therefore, the spacetime overhead is always in the optimal scenario described in Eq. (C.54).

Corollary C.19 (Spacetime overhead of surgery on HGP data code)**.** *The spacetime overhead of high-rate surgery on HGP data code is*

$$\alpha = \mathcal{O}(1) \tag{C.74}$$

for measurement of $t = \mathcal{O}(k_C)$ *logicals that form a generalized parallel pattern.*

C.5 Algorithm design for randomized ancilla construction and simulation details

In this section, we provide a detailed description of the randomized algorithm for ancilla construction. The inputs are given by the data code parity check matrices, $H_{X,Z}$, and the logicals to be measured, L_X. Here we assume the logicals to be measured are all of X-type. As mentioned in the main text, in simulations, we aim to minimally increase the parity check matrix. In low-rate gauging, it is commonly found a minimum of weight-2 increase in the check degree [140, 150, 156]. Thus, we have chosen to adopt the simplest transversal-type connections in $\Gamma_{0,1}$, which has $\left|\Gamma_{0,1}\right| = 1$, to guarantee such properties.

One key motivation for the numerics is to demonstrate that even without the symmetry defined in Definition C.15, the idea of high-rate surgery brings significant savings in overhead. Therefore, we assume no symmetry in the logicals to measure. With such an assumption, our HGP ancilla construction would construct an HGP ancilla as the product of a $k = t$ classical code and a repetition code. Such a construction is not efficient in practice since for finite size the constant factors are highly nontrivial. Therefore, we consider a more flexible randomized construction, which is similar to the randomized edge addition algorithms in Refs. [7, 156]. Our major innovation is to adapt it to general check matrices, which can be understood as adding hyperedges. Such a generalization loses the asymptotic guarantee of expander graphs, but it is observed to perform well in practice. Moreover, while the way we picked which hyperedges to add still has random components, we developed well-defined heuristics to boost its effectiveness.

See Figure C.2 for an illustration of the algorithm's key steps and see Algorithm 1 for a complete description. The inputs to the algorithm is $\partial^C_{1,2}$ and a logical representation of L_X. We also set a check degree restriction, c, on the merge code degree. From discussion before, c is set degree of data code plus 2 in simulations we presented in this work. On a high-level, the algorithm constructs a layered ancilla system. Broadly speaking, the algorithm can be divided into 4 steps, where one initializes the ancilla system to perform the right measurements (Figure C.2(a)), adds qubits and Z-checks to increase expansion (Figure C.2(c, d)), and adds qubits and X-checks to lower check degrees (Figure C.2(b)). The second step is a generalization of the random add edge construction [140, 150, 156]. The last step is crucial, especially for larger codes, and is performed when one cannot increase expansion while maintaining the degree limit of the parity check.

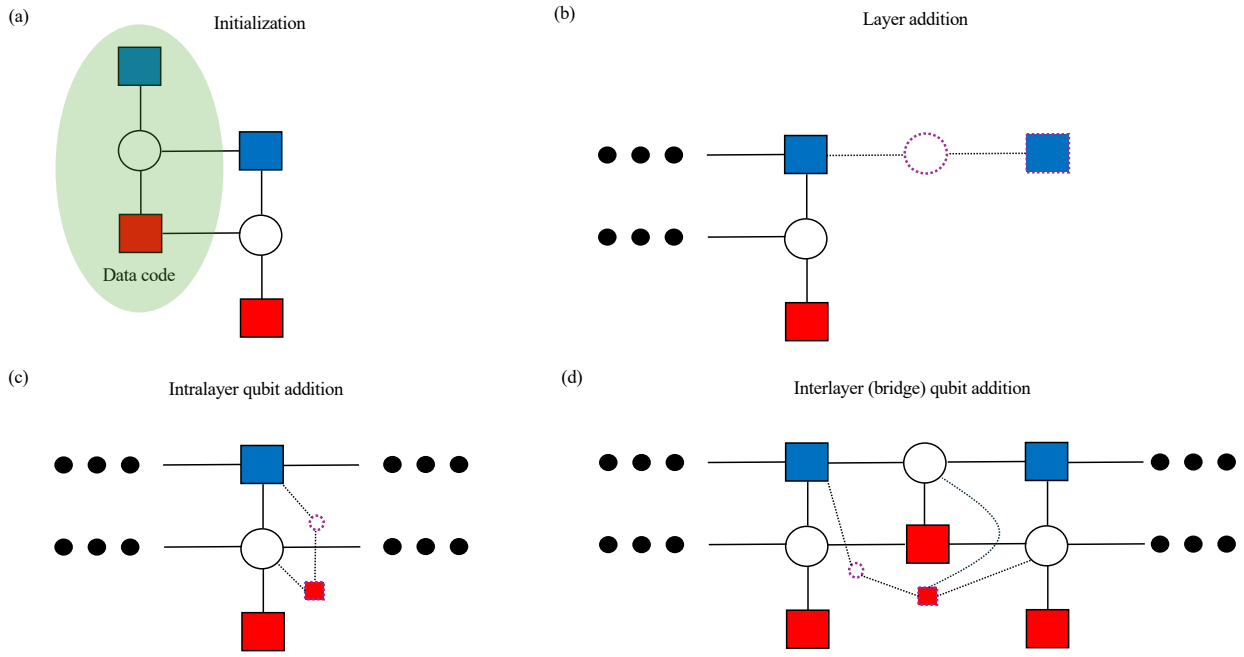


Figure C.2: A demonstration of the random ancilla construction algorithm. (a) Initialization of ancilla system. Green shaded region represent the data code. (b) Layer addition. New layers are added at the end of the layered ancilla system. (c) Intralayer qubit addition. Qubits and Z-checks are added to boost expansion. (d) Interlayer qubit addition. Qubits and Z-checks are added, where Z-checks are connected to qubits in consecutive layers.

Algorithm 1: Randomized construction of distance-preserving ancilla system

Input: Data code parity checks $\partial_2^C, \partial_1^C$, logicals to measure L_X, degree limit c.
Output: Ancilla parity checks $\partial_1^A, \partial_0^A$, chain maps $\Gamma_{0,1}$.

Initialize ancilla parity check as $\partial_1^A \leftarrow \begin{pmatrix} \partial_1^C|_{L_X} \\ \partial^l \end{pmatrix}$, where ∂_1^C is truncated to the union of L_X and low-weight logical checks, ∂^l, are found through a greedy search. Construct generating basis for $\ker \left(\partial_2^A\right)^T$ with weight reduction heuristics. Initialize ∂_1^A by the rows of the generating basis that has degree less than c.

if $\left|\partial_1^A\right| < c$ **then**
    Continue
else
    Return fail.

Initialize chain maps as $\Gamma_{0,1} \leftarrow \begin{pmatrix} I_{0,1} \\ \mathbf{0} \end{pmatrix}$ are of size $\left|C_{0,1}\right| \times \left|A_{0,1}\right|$, respectively.

// $I_{0,1}$ are projections onto the support of L_X and onto the adjacent checks in ∂_1^C respectively.

Initialize a list of layers, with the initial checks as the first layer with index 1. Each layer with index i is parametrized by the layer's check matrices, $\partial_{0,1}^i$.

Initialize empty bridge checks list.

while $d_{merge} < d$ **do**
    Sample logically distinct low-weight X-type logicals, l_X of weight d_{merge}.
    for $i \leftarrow 0$ **to** *max layer index* **do**
        Sample low-weight row vectors through (1) generate low-weight combinations, v, to obtain $w = \left(\partial_1^i\right)^T v$, such that $|v|, |w| \leq c$ or (2) copying checks from layers of index $i-1, i+1$, if exist.
        Compute the score of each low-weight check generated, w. The score is the number of l_X that anticommutes with w.
        if *exists w with score* > 0 **then**
            Choose the w with max score. Modify the layer's check matrix as $\partial_1^i \leftarrow \begin{pmatrix} \partial_1^i \\ w \end{pmatrix}$. If w is generated through check combination, update $\partial_0^i \rightarrow \begin{pmatrix} \partial_0^i & \mathbf{0} \\ v^T & 1 \end{pmatrix}$. Otherwise, update the bridging checks.
        else
            Continue to next loop
    if *no checks have been added* **then**
        Let current last layer index be m. Add a new layer, $m+1$. Initialize the new layer's checks as null matrices. Add bridge checks that connects transversally the X-checks of $m, m+1$ through bridge qubits.
        // See Figure C.2 and discussion.

return $\partial_1^A, \partial_0^A, \Gamma_{0,1}$

Initialization

The ancilla system, $\partial^A_{0,1}$, and the chain maps are initialized. We start by initializing ∂^A_1 as $\partial^C_1|_{L_X}$, where the subscript denotes restricted to the support of L_X, and $\Gamma_{0,1} \leftarrow \begin{pmatrix} I_{0,1} \\ \mathbf{0} \end{pmatrix}$. Here, $I_{0,1}$ are projections onto the support of L_X and all adjacent checks respectively. Such initializations guarantee the measurement of all logicals in L_X. However, to avoid measuring additional logicals, we add an additional check ∂^l such that the full initial boundary is

$$\partial^A_1 = \begin{pmatrix} \partial^C_1|_{L_X} \\ \partial^l \end{pmatrix} \tag{C.75}$$

In general, there is no theoretical guarantee that ∂^l is LDPC for general basis (see Methods for a detailed discussion). However, we find as a heuristic that as long as the physical circuit has circuit degree less than c, we can find a low-weight $\left|\partial^l\right|$. We search for ∂^l with an iterative local greedy algorithm: for each step, we find checks ∂^l through Algorithm 2 by letting the inputs $H = L_X\Gamma_1^{-1}$, $H_c = \mathbf{0}$, and $L^c = \ker\Gamma_1^{-1} \setminus H$. Here, the inverse should be understood as the pseudoinverse, and ∂^l is a random lowest-weight choice from the output, $\{L\}$. This step can be understood as finding low-weight checks to add to a classical code Γ_1^{-1} such that a specified kernel space can be retained. This process is repeated until $\Gamma_1\left(\ker\Gamma_1^{-1}\right)$ is in the span of $\begin{pmatrix} \left(\partial^C_2\right)^T \\ L_X \end{pmatrix}$. Note that this step is the only procedure in the algorithm that can lead to failure in ancilla construction. Nevertheless, the failure is mostly due to a poor choice in logical basis or a high physical circuit degree, so there is little one can do from a algorithmic perspective. Lastly, we search for a low-weight generating basis for $\ker\left(\Gamma_1^{-1}\right)^T$ and include all low-weight checks with degree less than c to be the initial ∂^A_0. The ancilla system at this point constitutes layer with index 1.

Algorithm 2: Search for low-weight subsystem logicals

Input: Parity check H, conjugate parity check H^c, bare conjugate logical L^c, max logical weight w_m, number of trials n_{trials}.

Output: Set of low-weight dressed logicals $\{L\}$, dressed distance d.

```
Initialize sets of dressed logicals, {L}, as empty set.
for i ← 1 to n_trials do
    Randomly shuffle rows of H to get H_i.
    Perform Gaussian elimination on H_i to get a generating basis for its kernel,
     defined as a matrix K.
    Compute M := K (L^c)^T  mod 2, and assume it is a R × C dimensional matrix.
    for r ← 1 to R do
        if Sum(M_{r,:}) > 0 and Sum(K_{r,:}) <= w_m then
            {L} ← {L} + K_{r,:}
Dress distance, d ← min {|L|}.
return {L}, d
```

Increase expansion

The expansion is increased through addition of ancilla qubits and Z checks. This subroutine is a generalization of the previous graph algorithms. However, instead of a fully random choice of hyperedge, we make a local greedy choice: we start with computing the current dressed distance and sample a set of distinct low-weight logicals, $\{L\}$, with Algorithm 2 of the whole merged code. We then search for new qubits to add from layer-to-layer. Consider a layer with index i and the intralayer checks are $\partial^i_{0,1}$, there are two types of qubits one can add.

1. The new qubit is connected to ancilla X-checks as a low-weight combination of the X-checks in the current layer. In practice, we low-weight vectors v and defined $w = \left(\partial^i_1\right)^T v$. If we add such a qubit, the layer checks are modified as

$$\partial^i_1 \to \begin{pmatrix} \partial^i_1 \\ w \end{pmatrix}, \partial^i_0 \;\; \to \;\; \begin{pmatrix} \partial^i_0 & \mathbf{0} \\ v^T & 1 \end{pmatrix} \tag{C.76}$$

   so we filter by requiring the modified check matrices to have degree under c.

2. The new qubit is connected to current layer X-checks identically with the connection of an existing qubit in layer $i-1$ or $i+1$ to that layer's X checks. As we will see later and also in Figure. C.2, the X-checks in each layer are of the same size, so to have identical connections is well-defined. Without loss of generality, consider duplicating a X-check in the $i-1$th layer with X-check w, we modify the check matrices as

$$\partial_1^i \;\rightarrow\; \begin{pmatrix} \partial_1^i \\ w \end{pmatrix} \tag{C.77}$$

$$\begin{pmatrix} & \vdots & \vdots & \vdots & \\ \dots & \partial_0^{i-1} & \mathbf{0} & \mathbf{0} & \dots \\ \dots & I^{i-1} & \delta^{i-1,i} & I^i & \dots \\ \dots & \mathbf{0} & \mathbf{0} & \partial_0^i & \dots \\ & \underbrace{\vdots}_{\text{Layer } i-1} & \underbrace{\vdots}_{\text{Bridge}} & \underbrace{\vdots}_{\text{Layer } i} & \end{pmatrix} \rightarrow \begin{pmatrix} & \vdots & \vdots & \vdots & \vdots & \\ \dots & \partial_0^{i-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots \\ \dots & I^{i-1} & \delta^{i-1,i} & \mathbf{0} & I^i & \dots \\ \dots & e_w^T & w^T & e_w^T & \mathbf{0} & \dots \\ \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \partial_0^i & \dots \\ & \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \tag{C.78}$$

where in the second line we show the modification of relevant intralayer Z-checks and also bridge checks and qubits. $\delta^{i-1,i}$ describes the bridge checks, and its rows are formed by common subsets of columns of $\partial_0^{i-1,i}$. e_w is unit vectors with nonzero entry at the corresponding duplicated qubit's position. One can see clearly from Eq. (C.78) that the added bridge Z-checks will have weight $|w|+2$. Thus, overall we restrict $|w| \leq c-2$ so that all qubits are available for copying.

With either choice of qubits, suppose the Z-checks added is Z_v, we compute a score for each qubit addition as $s_v := \text{Sum}\left(Z_v^T L \mod 2\right)$. Intuitively, such a score represents how many low-weight logicals anticommutes with the candidate Z-check. By the stabilizer measurement update rule [24], a logical operator that commutes with the added Z checks acts within the postmeasurement stabilizer sectors. If its extension acts trivially on the new qubits and remains outside the updated gauge group, it represents a surviving dressed logical with the same weight in the enlarged code [7]. We iterate over existing layers from 1 to the end, and

as long as we find a layer with nonzero s_v, we add the qubit with max s_v and continue to the next iteration. Note that there is no guarantee one can always find low-weight v, w such that the score is nonzero. When one fails to do so, we enter the procedure of adding layers.

Reducing code degrees

To have a mechanism of adding X-checks meaningfully is crucial for the success of our algorithm since notably the steps before are all adding qubits and Z-checks. If no X-checks can be added, the weight of the X-checks would soon exceed the code degree limit set. We choose to add X-checks in a conservative way to guarantee that the algorithm succeeds, i.e. in the worst case we recover the HGP ancilla construction and still return an ancilla construction with sufficient distance. We add X-checks by adding a whole layer of X-checks and qubits, connected 1-to-1. Assume the current maximum number of layers is m, we add a new layer through

$$\begin{pmatrix} & \vdots \\ \dots & \partial_1^m \end{pmatrix} \rightarrow \begin{pmatrix} & \vdots & \vdots \\ \dots & \partial_1^m & I \\ \dots & \mathbf{0} & I \end{pmatrix} \tag{C.79}$$

$$\partial_0^A \rightarrow \begin{pmatrix} \partial_0^A & \mathbf{0} \end{pmatrix} \tag{C.80}$$

This step cannot increase the distance or eliminate any low-weight logicals, but after adding such a layer, one is guaranteed to find low-weight qubits that can increase expansion in the previous procedure.

One of the subtlety we have skipped is that to guarantee absolutely one can reach HGP ancilla construction as worst case, it is necessary to only consider case 1 in the increase expansion procedure when the current layer has a full kernl,

$$\ker \partial_1^i = \ker \partial_1^1 \tag{C.81}$$

where $\ker \partial_1^1$ denotes the kernel space of the first layer. Such condition is required since we have allowed the algorithm to randomly add hyperedge combinations, which is not present in HGP constructions. If this condition is not present, it is possible for the algorithm to not find a possible nontrivial qubit choice even with added layers. However, we remark that this failure scenario only appeared when we simulated very large code examples, and it contributed the increased overhead per logicals measured. For smaller codes, one can often find much smaller distance-preserving ancilla constructions by not enforcing this condition. In addition, we only considered bridge qubits between adjacent layers, which is not necessary since locality is not a focus in this work. Therefore, we imagine that there are many ways to improve the algorithm by adding flexibility and clever algorithms. We leave such optimizations to future work.

In conclusion, we found that two points are crucial in the algorithm design. The first is to provide the ancilla system procedures to increase the number of X-checks. This is, in our opinion, an overlooked mechanism by prior works[140, 150, 156]. Their approaches worked both because they did not enforce a strict degree constraint and that they are working with small code examples measuring single logicals. However, to scale up while preserving sparsity, one would require a procedure similar to the layer addition procedure we outlined. Secondly, we found that to rank the possible qubit additions by their scores significantly reduces the system overhead. This holds even if we apply our algorithm to measure a single logical.

C.6 Measurement of mixed-type logicals

In the above, we have restricted to measurement of single Pauli type logicals in CSS codes for simplicity. In this section, we comment on how our approach generalizes to general logical measurements. Suppose we consider CSS codes and measurement of mixed types of logicals. One can generally write the logical measurement set as

$$M = \left\{ X_i^L Z_j^L | i \in K_X, j \in K_Z \right\} \tag{C.82}$$

where $K_{X,Z}$ denotes the set of logical X and Z indices that are in M. Here, we neglect the ± 1 factors for simplicity. The discussion is relatively simple when $K_X \cap K_Z = \varnothing$: without lost of generality, one can first construct the ancilla system for measuring all $M_X = \left\{X_i^L | i \in K_X\right\}$ and then for the Z-logicals in K_Z. To properly measure the products instead of measuring them individually, one add logical checks following the procedure described in Methods and in Appendix C.5. Similar ideas for low-rate surgery has been known as bridge qubits [140, 150] or adapters [157].

It is less straightforward when there are overlap in $K_{X,Z}$. The above construction works is that the construction of each X and Z ancilla system measure only operators that commutes with the current stabilizer group, which aligns with the surgery framework we have developed. Nevertheless, when there is overlap, such an assumption no longer holds. One remedy is to use twist-free surgery [9, 10, 229], where the measurement circuit is decomposed into two parts, each measuring a commuting joint Pauli product. The drawback is that one would need a resource state, $|Y\rangle$. This overhead is acceptable in the asymptotic regime since the Y resource state is unchanged after the measurement and can be used repeatedly. One way of preparing these Y resource states would be to perform low-rate surgery one-by-one. However, it is unknown how to prepare these Y resource states in parallel. Thus, more work is needed to assess the overhead associated with such resource states.

In Ref. [9], the canonical logical basis [9, 24] was proposed to circumvent this challenge. In such a basis, for any physical qubit in C_1, it is either X-support or Z-support of any logicals, but not both. This enables the scheme in Ref. [9] to prepare the Y states simulatenously. We refer interested readers to Ref. [9] for a more detailed discussion. While our scheme can directly adopt a similar solution, there is also an alternative approach: from our discussion in Appendix C.3 and Lemma 4.2, we remarked that the key to the logical measurement is through redundancies in the ancilla X-checks. In fact, this idea is not limited to measuring a

single type of logicals: in the symplectic form, consider the chain maps $\Gamma_{0,1}$ as

$$\Gamma_\alpha = \begin{pmatrix} \Gamma_\alpha^X & \Gamma_\alpha^Z \end{pmatrix} \tag{C.83}$$

with $\alpha = 0, 1$. Γ_α is a matrix of size $|A_1| \times 2|C_1|$. Γ_α allows the ancilla A_1 checks to connect to the data qubits, C_1, through either X, Y, Z connections. One the chain maps form a commuting diagram, the logical action is given by $\Gamma_1 \left(\ker \partial_1^A\right)$ as in Lemma 4.2. However, while it is not so different from the perspective of design the kernel, $\ker \partial_1^A$, there is generally no guarantee that all checks in Γ_α will commute with itself, i.e. $\Gamma_\alpha \Omega \Gamma_\alpha^T = 0$, with Ω being the symplectic matrix. For example, the logical set, M, can contain pairs of commuting logicals that have anticommuting supports on even number of qubits. Nevertheless, this will not occur if the canonical basis is adopted: Γ_1 will either not connect to a data qubit or connect to it through a single type of Pauli operator. One should note that the canonical basis is not limited to CSS codes and so is our discussion above. The major change would be that when one construct the ancilla system through, for example, taking union of the conjugate checks (see discussion in Appendix C.5), one should consider all checks that anticommute with all individual parity checks added through the ancilla system.

C.7 Constructions of spatially coupled codes

We follow the standard construction of Ref. [166] and identify the instances listed in Table 4.2 via numerical search. Here, we briefly sketch the construction and refer the reader to Ref. [166] for full details. A $[[n, k, d]]$ spatially coupled hypergraph product (SC-HGP) code is built by starting from an $L \times L$ array of identical base HGP codes — each obtained as the hypergraph product of two identical classical codes with check matrix $H_C \in \mathbb{F}_2^{r_C \times n_C}$ — and introducing couplings between them. Each base HGP code has parameters $[[n_0 = r_C^2 + n_C^2,\ k_0 \geq n_0 - 2r_C n_C,\ d_0]]$, leading to a SC-HGP code with parameters $n = n_0 L^2$ and $k \geq k_0 L^2$. As an example, for the $[[1125, 245, 10]]$, we choose $r_C = 3$, $n_C = 6$, and $L = 5$.

REFERENCES

[1] Kyungjoo Noh, Victor V. Albert, and Liang Jiang. Quantum capacity bounds of gaussian thermal loss channels and achievable rates with gottesman-kitaev-preskill codes. *IEEE Transactions on Information Theory*, 65(4):2563–2582, 2019. doi: 10.1109/TIT.2018.2873764.

[2] Matteo Rosati, Andrea Mari, and Vittorio Giovannetti. Narrow bounds for the quantum capacity of thermal attenuators. *Nat. Commun.*, 9:4339, 2018. doi: 10.1038/s41467-018-06848-0.

[3] Victor V. Albert, Kyungjoo Noh, Kasper Duivenvoorden, Dylan J. Young, R. T. Brierley, Philip Reinhold, Christophe Vuillot, Linshu Li, Chao Shen, S. M. Girvin, Barbara M. Terhal, and Liang Jiang. Performance and structure of single-mode bosonic codes. *Phys. Rev. A*, 97:032346, Mar 2018. doi: 10.1103/PhysRevA.97.032346. URL https://link.aps.org/doi/10.1103/PhysRevA.97.032346.

[4] Daniel Gottesman, Alexei Kitaev, and John Preskill. Encoding a qubit in an oscillator. *Phys. Rev. A*, 64:012310, Jun 2001. doi: 10.1103/PhysRevA.64.012310. URL https://link.aps.org/doi/10.1103/PhysRevA.64.012310.

[5] Baptiste Royer, Shraddha Singh, and S. M. Girvin. Stabilization of finite-energy gottesman-kitaev-preskill states. *Phys. Rev. Lett.*, 125:260509, Dec 2020. doi: 10.1103/PhysRevLett.125.260509. URL https://link.aps.org/doi/10.1103/PhysRevLett.125.260509.

[6] Ludovico Lami, Martin B. Plenio, Vittorio Giovannetti, and Alexander S. Holevo. Bosonic quantum communication across arbitrarily high loss channels. *Phys. Rev. Lett.*, 125:110504, 2020. doi: 10.1103/PhysRevLett.125.110504.

[7] Dominic J. Williamson and Theodore J. Yoder. Low-overhead fault-tolerant quantum computation by gauging logical operators, 2024. URL https://arxiv.org/abs/2410.02213.

[8] Lawrence Z. Cohen, Isaac H. Kim, Stephen D. Bartlett, and Benjamin J. Brown. Low-overhead fault-tolerant quantum computing using long-range connectivity. *Science Advances*, 8(20):eabn1717, 2022. doi: 10.1126/sciadv.abn1717.

[9] Alexander Cowtan, Zhiyang He, Dominic J. Williamson, and Theodore J. Yoder. Parallel logical measurements via quantum code surgery, 2025. URL https://arxiv.org/abs/2503.05003.

[10] Guo Zhang and Ying Li. Time-efficient logical operations on quantum low-density parity-check codes. *Phys. Rev. Lett.*, 134:070602, 2025. doi: 10.1103/PhysRevLett.134.070602.

[11] Sergey Bravyi et al. High-threshold and low-overhead fault-tolerant quantum memory. *Nature*, 627:778–784, 2024. doi: 10.1038/s41586-024-07107-7.

[12] Timo Hillmann, Lucas Berent, Armanda O Quintavalle, Jens Eisert, Robert Wille, and Joschka Roffe. Localized statistics decoding: A parallel decoding algorithm for quantum low-density parity-check codes. *arXiv preprint arXiv:2406.18655*, 2024.

[13] Guo Zheng, Wenhao He, Gideon Lee, Kyungjoo Noh, and Liang Jiang. Performance and achievable rates of the Gottesman-Kitaev-Preskill code for pure-loss and amplification channels. *PRX Quantum*, 6:030314, 2025. doi: 10.1103/gh1c-xyn1.

[14] Florian Pfender and Günter Ziegler. Kissing numbers, sphere packings, and some unexpected proofs. *Notices of the American Mathematical Society*, 06 2004.

[15] Baptiste Royer, Shraddha Singh, and S.M. Girvin. Encoding qubits in multimode grid states. *PRX Quantum*, 3:010335, Mar 2022. doi: 10.1103/PRXQuantum.3.010335. URL https://link.aps.org/doi/10.1103/PRXQuantum.3.010335.

[16] Leonid P. Pryadko, Vadim A. Shabashov, and Valerii K. Kozin. Qdistrnd: A gap package for computing the distance of quantum error-correcting codes. *Journal of Open Source Software*, 7(71):4120, 2022. doi: 10.21105/joss.04120. URL https://doi.org/10.21105/joss.04120.

[17] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 1997. doi: 10.1137/S0097539795293172.

[18] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, pages 212–219, 1996.

[19] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical Review Letters*, 103:150502, 2009. doi: 10.1103/PhysRevLett.103.150502.

[20] Haimeng Zhao, Alexander Zlokapa, Hartmut Neven, Ryan Babbush, John Preskill, Jarrod R. McClean, and Hsin-Yuan Huang. Exponential quantum advantage in processing massive classical data, 2026.

[21] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. doi: 10.1017/CBO9780511976667.

[22] Barbara M. Terhal. Quantum error correction for quantum memories. *Rev. Mod. Phys.*, 87:307–346, 2015. doi: 10.1103/RevModPhys.87.307.

[23] Peter W. Shor. Fault-tolerant quantum computation, 1996.

[24] Daniel Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology, 1997. arXiv:quant-ph/9705052.

[25] Alexei Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, 2003. doi: 10.1016/S0003-4916(02)00018-0.

[26] Daniel Litinski. A game of surface codes: Large-scale quantum computing with lattice surgery. *Quantum*, 3:128, 2019. doi: 10.22331/q-2019-03-05-128.

[27] Craig Gidney and Martin Ekerå. How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. *Quantum*, 5:433, 2021. doi: 10.22331/q-2021-04-15-433.

[28] Craig Gidney. How to factor 2048 bit RSA integers with less than a million noisy qubits, 2025.

[29] Madelyn Cain, Qian Xu, Robbie King, Lewis R. B. Picard, Harry Levine, Manuel Endres, John Preskill, Hsin-Yuan Huang, and Dolev Bluvstein. Shor's algorithm is possible with as few as 10,000 reconfigurable atomic qubits, 2026.

[30] Guo Zheng, Wenhao He, Gideon Lee, and Liang Jiang. Near-optimal performance of quantum error correction codes. *Phys. Rev. Lett.*, 132:250602, Jun 2024. doi: 10.1103/PhysRevLett.132.250602. URL https://link.aps.org/doi/10.1103/PhysRevLett.132.250602.

[31] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Phys. Rev. A*, 55:900–911, Feb 1997. doi: 10.1103/PhysRevA.55.900. URL https://link.aps.org/doi/10.1103/PhysRevA.55.900.

[32] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret. Quantum error correction of a qubit encoded in grid states of an oscillator. *Nature*, 584(7821):368–372, August 2020. ISSN 1476-4687. doi: 10.1038/s41586-020-2603-3. URL http://dx.doi.org/10.1038/s41586-020-2603-3.

[33] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret. Real-time quantum error correction beyond break-even. *Nature*, 616(7955):50–55, 2023. doi: 10.1038/s41586-023-05782-6. URL https://doi.org/10.1038/s41586-023-05782-6.

[34] Dany Lachance-Quirion, Marc-Antoine Lemonde, Jean Olivier Simoneau, Lucas St-Jean, Pascal Lemieux, Sara Turcotte, Wyatt Wright, Amélie Lacroix, Joëlle Fréchette-Viens, Ross Shillito, Florian Hopfmueller, Maxime Tremblay, Nicholas E. Frattini, Julien Camirand Lemyre, and Philippe St-Jean. Autonomous quantum error correction of gottesman-kitaev-preskill states. *Phys. Rev. Lett.*, 132:150607, Apr 2024. doi: 10.1103/PhysRevLett.132.150607. URL https://link.aps.org/doi/10.1103/PhysRevLett.132.150607.

[35] C. Flühmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. P. Home. Encoding a qubit in a trapped-ion mechanical oscillator. *Nature*, 566(7745): 513–517, February 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-0960-6. URL http://dx.doi.org/10.1038/s41586-019-0960-6.

[36] Brennan de Neeve, Thanh-Long Nguyen, Tanja Behrle, and Jonathan P. Home. Error correction of a logical grid state qubit by dissipative pumping. *Nature Physics*, 18(3): 296–300, 2022. doi: 10.1038/s41567-021-01487-7. URL https://doi.org/10.1038/s41567-021-01487-7.

[37] N. Fabre, G. Maltese, F. Appas, S. Felicetti, A. Ketterer, A. Keller, T. Coudreau, F. Baboux, M. I. Amanti, S. Ducci, and P. Milman. Generation of a time-frequency grid state with integrated biphoton frequency combs. *Phys. Rev. A*, 102:012607, Jul 2020. doi: 10.1103/PhysRevA.102.012607. URL https://link.aps.org/doi/10.1103/PhysRevA.102.012607.

[38] J. Eli Bourassa, Rafael N. Alexander, Michael Vasmer, Ashlesha Patil, Ilan Tzitrin, Takaya Matsuura, Daiqin Su, Ben Q. Baragiola, Saikat Guha, Guillaume Dauphinais, Krishna K. Sabapathy, Nicolas C. Menicucci, and Ish Dhand. Blueprint for a Scalable Photonic Fault-Tolerant Quantum Computer. *Quantum*, 5:392, February 2021. ISSN 2521-327X. doi: 10.22331/q-2021-02-04-392. URL https://doi.org/10.22331/q-2021-02-04-392.

[39] Benjamin L. Brock, Shraddha Singh, Alec Eickbusch, Volodymyr V. Sivak, Andy Z. Ding, Luigi Frunzio, Steven M. Girvin, and Michel H. Devoret. Quantum error correction of qudits beyond break-even. *Nature*, 641:612–618, 2025. doi: 10.1038/s41586-025-08899-y. URL https://www.nature.com/articles/s41586-025-08899-y.

[40] V. G. Matsos, C. H. Valahu, M. J. Millican, T. Navickas, X. C. Kolesnikow, M. J. Biercuk, and T. R. Tan. Universal quantum gate set for Gottesman–Kitaev–Preskill logical qubits. *Nature Physics*, 21:1664–1669, 2025. doi: 10.1038/s41567-025-03002-8. URL https://www.nature.com/articles/s41567-025-03002-8.

[41] M. V. Larsen, J. E. Bourassa, S. Kocsis, J. F. Tasker, R. S. Chadwick, C. González-Arciniegas, J. Hastrup, C. E. Lopetegui-González, F. M. Miatto, A. Motamedi, R. Noro, G. Roeland, R. Baby, H. Chen, P. Contu, I. Di Luch, C. Drago, M. Giesbrecht, T. Grainge, I. Krasnokutska, M. Menotti, B. Morrison, C. Puviraj, K. Rezaei Shad, B. Hussain, J. McMahon, J. E. Ortmann, M. J. Collins, C. Ma, D. S. Phillips, M. Seymour, Q. Y. Tang, B. Yang, Z. Vernon, R. N. Alexander, and D. H. Mahler. Integrated photonic source of Gottesman–Kitaev–Preskill qubits. *Nature*, 642:587–591, 2025. doi: 10.1038/s41586-025-09044-5. URL https://www.nature.com/articles/s41586-025-09044-5.

[42] Jonathan Conrad, Jens Eisert, and Francesco Arzani. Gottesman-kitaev-preskill codes: A lattice perspective. *Quantum*, 6:648, February 2022. ISSN 2521-327X. doi: 10.22331/q-2022-02-10-648. URL http://dx.doi.org/10.22331/q-2022-02-10-648.

[43] Jim Harrington and John Preskill. Achievable rates for the gaussian quantum channel. *Phys. Rev. A*, 64:062301, Nov 2001. doi: 10.1103/PhysRevA.64.062301. URL https://link.aps.org/doi/10.1103/PhysRevA.64.062301.

[44] Guo Zheng and Liang Jiang. Achieving coherent information rates for Gaussian-dilatable channels through Gottesman-Kitaev-Preskill codes. Manuscript in preparation; local source snapshot dated October 9, 2026, 2026.

[45] Nikolas P. Breuckmann and Jens Niklas Eberhardt. Quantum low-density parity-check codes. *PRX Quantum*, 2:040101, 2021. doi: 10.1103/PRXQuantum.2.040101.

[46] Pavel Panteleev and Gleb Kalachev. Asymptotically good quantum and locally testable classical ldpc codes. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, page 375–388, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450392648. doi: 10.1145/3519935.3520017. URL https://doi.org/10.1145/3519935.3520017.

[47] Qian Xu, J. Pablo Bonilla Ataides, Christopher A. Pattison, Nithin Raveendran, Dolev Bluvstein, Jonathan Wurtz, Bane Vasić, Mikhail D. Lukin, Liang Jiang, and Hengyun Zhou. Constant-overhead fault-tolerant quantum computation with reconfigurable atom arrays. *Nature Physics*, 20(7):1084–1090, 2024. doi: 10.1038/s41567-024-02479-z. URL https://doi.org/10.1038/s41567-024-02479-z.

[48] Ke Wang, Zhide Lu, Chuanyu Zhang, Gongyu Liu, Jiachen Chen, Yanzhe Wang, Yaozu Wu, Shibo Xu, Xuhao Zhu, Feitong Jin, Yu Gao, Ziqi Tan, Zhengyi Cui, Ning Wang, Yiren Zou, Aosai Zhang, Tingting Li, Fanhao Shen, Jiarun Zhong, Zehang Bao, Zitian Zhu, Yihang Han, Yiyang He, Jiayuan Shen, Han Wang, Jia-Nan Yang, Zixuan Song, Jinfeng Deng, Hang Dong, Zheng-Zhi Sun, Weikang Li, Qi Ye, Si Jiang, Yixuan Ma, Pei-Xin Shen, Pengfei Zhang, Hekang Li, Qiujiang Guo, Zhen Wang, Chao Song, H. Wang, and Dong-Ling Deng. Demonstration of low-overhead quantum error correction codes, 2025. URL https://arxiv.org/abs/2505.09684.

[49] Guo Zheng, Liang Jiang, and Qian Xu. High-rate surgery: Towards constant-overhead logical operations. *arXiv preprint*, 2025.

[50] S. B. Bravyi and A. Yu. Kitaev. Quantum codes on a lattice with boundary, 1998.

[51] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52:R2493–R2496, Oct 1995. doi: 10.1103/PhysRevA.52.R2493. URL https://link.aps.org/doi/10.1103/PhysRevA.52.R2493.

[52] Debbie W. Leung, M. A. Nielsen, Isaac L. Chuang, and Yoshihisa Yamamoto. Approximate quantum error correction can lead to better codes. *Phys. Rev. A*, 56:2567–2573, Oct 1997. doi: 10.1103/PhysRevA.56.2567. URL https://link.aps.org/doi/10.1103/PhysRevA.56.2567.

[53] Zaki Leghtas, Gerhard Kirchmair, Brian Vlastakis, Robert J. Schoelkopf, Michel H. Devoret, and Mazyar Mirrahimi. Hardware-efficient autonomous quantum memory protection. *Phys. Rev. Lett.*, 111:120501, Sep 2013. doi: 10.1103/PhysRevLett.111.120501. URL https://link.aps.org/doi/10.1103/PhysRevLett.111.120501.

[54] Raphaël Lescanne, Marius Villiers, Théau Peronnin, Alain Sarlette, Matthieu Delbecq, Benjamin Huard, Takis Kontos, Mazyar Mirrahimi, and Zaki Leghtas. Exponential suppression of bit-flips in a qubit encoded in an oscillator. *Nature Physics*, 16(5): 509–513, 2020. doi: 10.1038/s41567-020-0824-x. URL https://doi.org/10.1038/s41567-020-0824-x.

[55] Marios H. Michael, Matti Silveri, R. T. Brierley, Victor V. Albert, Juha Salmilehto, Liang Jiang, and S. M. Girvin. New class of quantum error-correcting codes for a bosonic mode. *Phys. Rev. X*, 6:031006, Jul 2016. doi: 10.1103/PhysRevX.6.031006. URL https://link.aps.org/doi/10.1103/PhysRevX.6.031006.

[56] Qian Xu, Guo Zheng, Yu-Xin Wang, Peter Zoller, Aashish A. Clerk, and Liang Jiang. Autonomous quantum error correction and fault-tolerant quantum computation with squeezed cat qubits. *npj Quantum Information*, 9(1):78, 2023. doi: 10.1038/s41534-023-00746-0. URL https://doi.org/10.1038/s41534-023-00746-0.

[57] Akel Hashim, Ravi K. Naik, Alexis Morvan, Jean-Loup Ville, Bradley Mitchell, John Mark Kreikebaum, Marc Davis, Ethan Smith, Costin Iancu, Kevin P. O'Brien, Ian Hincks, Joel J. Wallman, Joseph Emerson, and Irfan Siddiqi. Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor. *Phys. Rev. X*, 11:041039, Nov 2021. doi: 10.1103/PhysRevX.11.041039. URL https://link.aps.org/doi/10.1103/PhysRevX.11.041039.

[58] Joel J. Wallman and Joseph Emerson. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A*, 94:052325, Nov 2016. doi: 10.1103/PhysRevA.94.052325. URL https://link.aps.org/doi/10.1103/PhysRevA.94.052325.

[59] M. Reimpell and R. F. Werner. Iterative optimization of quantum error correcting codes. *Phys. Rev. Lett.*, 94:080501, Mar 2005. doi: 10.1103/PhysRevLett.94.080501. URL https://link.aps.org/doi/10.1103/PhysRevLett.94.080501.

[60] Andrew S. Fletcher, Peter W. Shor, and Moe Z. Win. Optimum quantum error recovery using semidefinite programming. *Phys. Rev. A*, 75:012338, Jan 2007. doi: 10.1103/PhysRevA.75.012338. URL https://link.aps.org/doi/10.1103/PhysRevA.75.012338.

[61] Aurélie Denys and Anthony Leverrier. The $2T$-qutrit, a two-mode bosonic qutrit. *Quantum*, 7:1032, June 2023. ISSN 2521-327X. doi: 10.22331/q-2023-06-05-1032. URL https://doi.org/10.22331/q-2023-06-05-1032.

[62] Benjamin Schumacher. Sending entanglement through noisy quantum channels. *Phys. Rev. A*, 54:2614–2628, Oct 1996. doi: 10.1103/PhysRevA.54.2614. URL https://link.aps.org/doi/10.1103/PhysRevA.54.2614.

[63] Koenraad Audenaert and Bart De Moor. Optimizing completely positive maps using semidefinite programming. *Phys. Rev. A*, 65:030302, Feb 2002. doi: 10.1103/PhysRevA.65.030302. URL https://link.aps.org/doi/10.1103/PhysRevA.65.030302.

[64] Robert L. Kosut and Daniel A. Lidar. Quantum error correction via convex optimization. *Quantum Information Processing*, 8(5):443–459, July 2009. ISSN 1573-1332. doi: 10.1007/s11128-009-0120-2. URL http://dx.doi.org/10.1007/s11128-009-0120-2.

[65] David S. Schlegel, Fabrizio Minganti, and Vincenzo Savona. Quantum error correction using squeezed schrödinger cat states. *Phys. Rev. A*, 106:022431, Aug 2022. doi: 10.1103/PhysRevA.106.022431. URL https://link.aps.org/doi/10.1103/PhysRevA.106.022431.

[66] Andrew S. Fletcher. Channel-adapted quantum error correction. 2007.

[67] M. Reimpell, R. F. Werner, and K. Audenaert. Comment on "optimum quantum error recovery using semidefinite programming". 2006.

[68] Naoki Yamamoto, Shinji Hara, and Koji Tsumura. Suboptimal quantum-error-correcting procedure based on semidefinite programming. *Physical Review A*, 71(2), February 2005. ISSN 1094-1622. doi: 10.1103/physreva.71.022322. URL http://dx.doi.org/10.1103/PhysRevA.71.022322.

[69] Cédric Bény and Ognyan Oreshkov. General conditions for approximate quantum error correction and near-optimal recovery channels. *Phys. Rev. Lett.*, 104:120501, Mar 2010. doi: 10.1103/PhysRevLett.104.120501. URL https://link.aps.org/doi/10.1103/PhysRevLett.104.120501.

[70] Robert L. Kosut, Alireza Shabani, and Daniel A. Lidar. Robust quantum error correction via convex optimization. *Phys. Rev. Lett.*, 100:020502, Jan 2008. doi: 10.1103/PhysRevLett.100.020502. URL https://link.aps.org/doi/10.1103/PhysRevLett.100.020502.

[71] Akshaya Jayashankar, Anjala M. Babu, Hui Khoon Ng, and Prabha Mandayam. Finding good quantum codes using the cartan form. *Phys. Rev. A*, 101:042307, Apr 2020. doi: 10.1103/PhysRevA.101.042307. URL https://link.aps.org/doi/10.1103/PhysRevA.101.042307.

[72] Peter Leviant, Qian Xu, Liang Jiang, and Serge Rosenblum. Quantum capacity and codes for the bosonic loss-dephasing channel. *Quantum*, 6:821, September 2022. ISSN 2521-327X. doi: 10.22331/q-2022-09-29-821. URL https://doi.org/10.22331/q-2022-09-29-821.

[73] H. Barnum and E. Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *Journal of Mathematical Physics*, 43(5):2097–2106, 04 2002. ISSN 0022-2488. doi: 10.1063/1.1459754. URL https://doi.org/10.1063/1.1459754.

[74] Hui Khoon Ng and Prabha Mandayam. Simple approach to approximate quantum error correction based on the transpose channel. *Phys. Rev. A*, 81:062342, Jun 2010. doi: 10.1103/PhysRevA.81.062342. URL https://link.aps.org/doi/10.1103/PhysRevA.81.062342.

[75] Denes Petz. SUFFICIENCY OF CHANNELS OVER VON NEUMANN ALGEBRAS. *Quart. J. Math. Oxford Ser.*, 39(1):97–108, 1988. doi: 10.1093/qmath/39.1.97.

[76] M.M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2017. ISBN 9781316813300. URL https://books.google.com/books?id=gYcHDgAAQBAJ.

[77] Jon Tyson. Two-sided bounds on minimum-error quantum measurement, on the reversibility of quantum dynamics, and on maximum overlap using directional iterates. *Journal of Mathematical Physics*, 51(9):092204, 09 2010. ISSN 0022-2488. doi: 10.1063/1.3463451. URL https://doi.org/10.1063/1.3463451.

[78] Cédric Bény. Perturbative quantum error correction. *Phys. Rev. Lett.*, 107:080501, Aug 2011. doi: 10.1103/PhysRevLett.107.080501. URL https://link.aps.org/doi/10.1103/PhysRevLett.107.080501.

[79] Cédric Bény and Ognyan Oreshkov. Approximate simulation of quantum channels. *Physical Review A*, 84(2), August 2011. ISSN 1094-1622. doi: 10.1103/physreva.84.022333. URL http://dx.doi.org/10.1103/PhysRevA.84.022333.

[80] Prabha Mandayam and Hui Khoon Ng. Towards a unified framework for approximate quantum error correction. *Phys. Rev. A*, 86:012335, Jul 2012. doi: 10.1103/PhysRevA.86.012335. URL https://link.aps.org/doi/10.1103/PhysRevA.86.012335.

[81] Marius Junge, Renato Renner, David Sutter, Mark M. Wilde, and Andreas Winter. Universal recovery maps and approximate sufficiency of quantum relative entropy. *Annales Henri Poincaré*, 19(10):2955–2978, August 2018. ISSN 1424-0661. doi: 10.1007/s00023-018-0716-0. URL http://dx.doi.org/10.1007/s00023-018-0716-0.

[82] Fernando G. S. L. Brandão, Elizabeth Crosson, M. Burak Şahinoğlu, and John Bowen. Quantum error correcting codes in eigenstates of translation-invariant spin chains. *Phys. Rev. Lett.*, 123:110502, Sep 2019. doi: 10.1103/PhysRevLett.123.110502. URL https://link.aps.org/doi/10.1103/PhysRevLett.123.110502.

[83] Philippe Faist, Sepehr Nezami, Victor V. Albert, Grant Salton, Fernando Pastawski, Patrick Hayden, and John Preskill. Continuous symmetries and approximate quantum error correction. *Phys. Rev. X*, 10:041018, Oct 2020. doi: 10.1103/PhysRevX.10.041018. URL https://link.aps.org/doi/10.1103/PhysRevX.10.041018.

[84] Sisi Zhou, Zi-Wen Liu, and Liang Jiang. New perspectives on covariant quantum error correction. *Quantum*, 5:521, August 2021. ISSN 2521-327X. doi: 10.22331/q-2021-08-09-521. URL https://doi.org/10.22331/q-2021-08-09-521.

[85] András Gilyén, Seth Lloyd, Iman Marvian, Yihui Quek, and Mark M. Wilde. Quantum algorithm for petz recovery channels and pretty good measurements. *Phys. Rev. Lett.*, 128:220502, Jun 2022. doi: 10.1103/PhysRevLett.128.220502. URL https://link.aps.org/doi/10.1103/PhysRevLett.128.220502.

[86] A. R. Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54:1098–1105, Aug 1996. doi: 10.1103/PhysRevA.54.1098. URL https://link.aps.org/doi/10.1103/PhysRevA.54.1098.

[87] A. M. Steane. Error correcting codes in quantum theory. *Phys. Rev. Lett.*, 77:793–797, Jul 1996. doi: 10.1103/PhysRevLett.77.793. URL https://link.aps.org/doi/10.1103/PhysRevLett.77.793.

[88] Shubham P. Jain, Joseph T. Iosue, Alexander Barg, and Victor V. Albert. Quantum spherical codes. *Nature Physics*, May 2024. ISSN 1745-2481. doi: 10.1038/s41567-024-02496-y. URL http://dx.doi.org/10.1038/s41567-024-02496-y.

[89] Yijia Xu, Yixu Wang, Christophe Vuillot, and Victor V. Albert. Letting the tiger out of its cage: bosonic coding without concatenation. *Physical Review X*, 15:041025, 2025. doi: 10.1103/ls5r-vj7r.

[90] Guo Zheng, Simon Lieu, Emma L. Rosenfeld, Kyungjoo Noh, and Connor T. Hann. Cross-resonance control of an oscillator with an auxiliary fluxonium qubit. *Physical Review Applied*, 23(2):024067, 2025. doi: 10.1103/PhysRevApplied.23.024067.

[91] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Physical Review A*, 80:012304, 2009. doi: 10.1103/PhysRevA.80.012304.

[92] Bikun Li, Zhaoyou Wang, Guo Zheng, Yat Wong, and Liang Jiang. Optimality condition for the Petz map. *Physical Review Letters*, 134(20):200602, 2025. doi: 10.1103/PhysRevLett.134.200602.

[93] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Phys. Rev. A*, 60:1888–1898, Sep 1999. doi: 10.1103/PhysRevA.60.1888. URL https://link.aps.org/doi/10.1103/PhysRevA.60.1888.

[94] Michael A Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, 303(4):249–252, October 2002. ISSN 0375-9601. doi: 10.1016/s0375-9601(02)01272-0. URL http://dx.doi.org/10.1016/S0375-9601(02)01272-0.

[95] Mario Berta and Marco Tomamichel. The fidelity of recovery is multiplicative. *IEEE Transactions on Information Theory*, 62(4):1758–1763, 2016. doi: 10.1109/TIT.2016.2527683.

[96] Vahid Karimipour, Fabio Benatti, and Roberto Floreanini. Quasi-inversion of qubit channels. *Phys. Rev. A*, 101:032109, Mar 2020. doi: 10.1103/PhysRevA.101.032109. URL https://link.aps.org/doi/10.1103/PhysRevA.101.032109.

[97] David Kribs, Raymond Laflamme, and David Poulin. Unified and generalized approach to quantum error correction. *Phys. Rev. Lett.*, 94:180501, May 2005. doi: 10.1103/PhysRevLett.94.180501. URL https://link.aps.org/doi/10.1103/PhysRevLett.94.180501.

[98] Masato Koashi and Masahito Ueda. Reversing measurement and probabilistic quantum error correction. *Phys. Rev. Lett.*, 82:2598–2601, Mar 1999. doi: 10.1103/PhysRevLett.82.2598. URL https://link.aps.org/doi/10.1103/PhysRevLett.82.2598.

[99] Ankith Mohan, Jamie Sikora, and Sarvagya Upadhyay. A hybrid quantum-classical algorithm for Bayes-optimal quantum state discrimination using the source code. 2023.

[100] Nathaniel Johnston, Vincent Russo, and Jamie Sikora. Tight bounds for antidistinguishability and circulant sets of pure quantum states. 2023.

[101] Haotian Jiang, Tarun Kathuria, Yin Tat Lee, Swati Padmanabhan, and Zhao Song. A faster interior point method for semidefinite programming. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 910–918, 2020. doi: 10.1109/FOCS46700.2020.00089.

[102] J. L. Daletskii and S.G. Krein. Integration and differentiation of functions of hermitian operators and applications to the theory of perturbations. *American Mathematical Society Translations, Series 2*, 47:1–30, 1965. doi: 10.1090/trans2/047/01.

[103] Marcus Carlsson. Perturbation theory for the matrix square root and matrix modulus. 2018.

[104] Jan Olle, Remmy Zen, Matteo Puviani, and Florian Marquardt. Simultaneous discovery of quantum error correction codes and encoders with a noise-aware reinforcement learning agent. 2023.

[105] Yexiong Zeng, Zheng-Yang Zhou, Enrico Rinaldi, Clemens Gneiting, and Franco Nori. Approximate autonomous quantum error correction with reinforcement learning. *Phys. Rev. Lett.*, 131:050601, Jul 2023. doi: 10.1103/PhysRevLett.131.050601. URL https://link.aps.org/doi/10.1103/PhysRevLett.131.050601.

[106] S. B. Korolev, E. N. Bashmakova, and T. Yu. Golubeva. Error correction using squeezed fock states. 2023.

[107] Chenfeng Cao, Chao Zhang, Zipeng Wu, Markus Grassl, and Bei Zeng. Quantum variational learning for quantum error-correcting codes. *Quantum*, 6:828, October 2022. ISSN 2521-327X. doi: 10.22331/q-2022-10-06-828. URL http://dx.doi.org/10.22331/q-2022-10-06-828.

[108] Eric M. Rains, R. H. Hardin, Peter W. Shor, and N. J. A. Sloane. A nonadditive quantum code. *Phys. Rev. Lett.*, 79:953–954, Aug 1997. doi: 10.1103/PhysRevLett.79.953. URL https://link.aps.org/doi/10.1103/PhysRevLett.79.953.

[109] Mark Srednicki. Chaos and quantum thermalization. *Phys. Rev. E*, 50:888–901, Aug 1994. doi: 10.1103/PhysRevE.50.888. URL https://link.aps.org/doi/10.1103/PhysRevE.50.888.

[110] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3–4):379–423, 623–656, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x. Published in two installments; second-installment DOI: https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.

[111] Benjamin Schumacher and M. A. Nielsen. Quantum data processing and error correction. *Phys. Rev. A*, 54:2629–2635, Oct 1996. doi: 10.1103/PhysRevA.54.2629. URL https://link.aps.org/doi/10.1103/PhysRevA.54.2629.

[112] Seth Lloyd. Capacity of the noisy quantum channel. *Phys. Rev. A*, 55:1613–1622, Mar 1997. doi: 10.1103/PhysRevA.55.1613. URL https://link.aps.org/doi/10.1103/PhysRevA.55.1613.

[113] I. Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55, 2005. doi: 10.1109/TIT.2004.839515.

[114] Howard Barnum, M. A. Nielsen, and Benjamin Schumacher. Information transmission through a noisy quantum channel. *Phys. Rev. A*, 57:4153–4175, Jun 1998. doi: 10.1103/PhysRevA.57.4153. URL https://link.aps.org/doi/10.1103/PhysRevA.57.4153.

[115] P. W. Shor. The quantum channel capacity and coherent information., February 2002.

[116] A. S. Holevo and R. F. Werner. Evaluating capacities of bosonic gaussian channels. *Phys. Rev. A*, 63:032312, Feb 2001. doi: 10.1103/PhysRevA.63.032312. URL https://link.aps.org/doi/10.1103/PhysRevA.63.032312.

[117] Michael M. Wolf, David Pérez-García, and Geza Giedke. Quantum capacities of bosonic channels. *Phys. Rev. Lett.*, 98:130501, Mar 2007. doi: 10.1103/PhysRevLett.98.130501. URL https://link.aps.org/doi/10.1103/PhysRevLett.98.130501.

[118] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd. Gaussian quantum information. *Rev. Mod. Phys.*, 84:621–669, May 2012. doi: 10.1103/RevModPhys.84.621. URL https://link.aps.org/doi/10.1103/RevModPhys.84.621.

[119] Filippo Caruso and Vittorio Giovannetti. Degradability of bosonic gaussian channels. *Phys. Rev. A*, 74:062307, Dec 2006. doi: 10.1103/PhysRevA.74.062307. URL https://link.aps.org/doi/10.1103/PhysRevA.74.062307.

[120] F. Caruso, J. Eisert, V. Giovannetti, and A. S. Holevo. Multi-mode bosonic Gaussian channels. *New J. Phys.*, 10:083030, 2008. doi: 10.1088/1367-2630/10/8/083030.

[121] Krishna Kumar Sabapathy and Andreas Winter. Non-Gaussian operations on bosonic modes of light: Photon-added Gaussian channels. *Physical Review A*, 95:062309, 2017. doi: 10.1103/PhysRevA.95.062309.

[122] Ludovico Lami, Krishna Kumar Sabapathy, and Andreas Winter. All phase-space linear bosonic channels are approximately gaussian dilatable. *New Journal of Physics*, 20: 113012, 2018. doi: 10.1088/1367-2630/aae738.

[123] Mao Lin, Christopher Chamberland, and Kyungjoo Noh. Closest lattice point decoding for multimode gottesman-kitaev-preskill codes. *PRX Quantum*, 4:040334, Dec 2023. doi: 10.1103/PRXQuantum.4.040334. URL https://link.aps.org/doi/10.1103/PRXQuantum.4.040334.

[124] Timo Hillmann, Jens Eisert, and Francesco Arzani. Quantum low-density lattice codes, 2026.

[125] Jacob Hastrup, Kimin Park, Jonatan Bohr Brask, Radim Filip, and Ulrik Lund Andersen. Measurement-free preparation of grid states. *npj Quantum Information*, 7 (1):17, 2021. doi: 10.1038/s41534-020-00353-3. URL https://doi.org/10.1038/s41534-020-00353-3.

[126] B. M. Terhal and D. Weigand. Encoding a qubit into a cavity mode in circuit qed using phase estimation. *Phys. Rev. A*, 93:012315, Jan 2016. doi: 10.1103/PhysRevA.93.012315. URL https://link.aps.org/doi/10.1103/PhysRevA.93.012315.

[127] Alec Eickbusch, Volodymyr Sivak, Andy Z. Ding, Salvatore S. Elder, Shantanu R. Jha, Jayameenakshi Venkatraman, Baptiste Royer, S. M. Girvin, Robert J. Schoelkopf, and Michel H. Devoret. Fast universal control of an oscillator with weak dispersive coupling to a qubit. *Nature Physics*, 18(12):1464–1469, October 2022. ISSN 1745-2481. doi: 10.1038/s41567-022-01776-9. URL http://dx.doi.org/10.1038/s41567-022-01776-9.

[128] P. Buser and P. Sarnak. On the period matrix of a riemann surface of large genus (with an appendix by j.h. conway and n.j.a. sloane). *Inventiones mathematicae*, 117(1):27–56, 1994. doi: 10.1007/BF01232233. URL https://doi.org/10.1007/BF01232233.

[129] Mahadevan Subramanian, Guo Zheng, and Liang Jiang. Achievable rates for concatenated square Gottesman-Kitaev-Preskill codes. *PRX Quantum*, 6:040342, 2025. doi: 10.1103/56vj-z7h1.

[130] Dubi Kelmer and Shucheng Yu. Second moment of the Siegel transform in the space of symplectic lattices. *Int. Math. Res. Not.*, 2021:5825–5859, 2021. doi: 10.1093/imrn/rnz027.

[131] Youngrong Lim, Soojoon Lee, Jaewan Kim, and Kabgyun Jeong. Upper bounds on the quantum capacity for general attenuator and amplifier. *Phys. Rev. A*, 99:052326, 2019. doi: 10.1103/PhysRevA.99.052326.

[132] Vittorio Giovannetti, Alexander S. Holevo, and Raúl García-Patrón. A solution of Gaussian optimizer conjecture for quantum channels. *Commun. Math. Phys.*, 334: 1553–1571, 2015. doi: 10.1007/s00220-014-2150-6.

[133] Hugh L. Montgomery. Minimal theta functions. *Glasgow Mathematical Journal*, 30(1): 75–85, 1988. doi: 10.1017/S0017089500007047.

[134] Henry Cohn, Abhinav Kumar, Stephen D. Miller, Danylo Radchenko, and Maryna Viazovska. Universal optimality of the E_8 and Leech lattices and interpolation formulas. *Annals of Mathematics*, 196(3):983–1082, 2022. doi: 10.4007/annals.2022.196.3.3.

[135] Evelyn Voss, Bikun Li, Zhaoyou Wang, and Liang Jiang. Passive environment-assisted quantum communication, 2026.

[136] Zhaoyou Wang and Liang Jiang. Passive environment-assisted quantum communication with GKP states. *Phys. Rev. X*, 15:021003, 2025. doi: 10.1103/PhysRevX.15.021003.

[137] Yunong Shi, Christopher Chamberland, and Andrew W. Cross. Fault-tolerant preparation of approximate GKP states. *New Journal of Physics*, 21:093007, 2019. doi: 10.1088/1367-2630/ab3a62. URL https://arxiv.org/abs/1905.00903.

[138] Wen-Long Ma, Mengzhen Zhang, Yat Wong, Kyungjoo Noh, Serge Rosenblum, Philip Reinhold, Robert J. Schoelkopf, and Liang Jiang. Path-independent quantum gates with noisy ancilla. *Phys. Rev. Lett.*, 125:110503, Sep 2020. doi: 10.1103/PhysRevLett.125.110503. URL https://link.aps.org/doi/10.1103/PhysRevLett.125.110503.

[139] Shruti Puri, Alexander Grimm, Philippe Campagne-Ibarcq, Alec Eickbusch, Kyungjoo Noh, Gabrielle Roberts, Liang Jiang, Mazyar Mirrahimi, Michel H. Devoret, and S. M. Girvin. Stabilized cat in a driven nonlinear cavity: A fault-tolerant error syndrome detector. *Physical Review X*, 9:041009, 2019. doi: 10.1103/PhysRevX.9.041009. URL https://journals.aps.org/prx/abstract/10.1103/PhysRevX.9.041009.

[140] Theodore J. Yoder, Eddie Schoute, Patrick Rall, Emily Pritchett, Jay M. Gambetta, Andrew W. Cross, Malcolm Carroll, and Michael E. Beverland. Tour de gross: A modular quantum computer based on bivariate bicycle codes, 2025. URL https://arxiv.org/abs/2506.03094.

[141] Christian Kraglund Andersen and Eliška Greplová. Small quantum low parity density check codes for near-term experiments, 2025. URL https://arxiv.org/abs/2507.09690.

[142] Laura Pecorari, Sven Jandura, Gavin K. Brennen, and Guido Pupillo. High-rate quantum ldpc codes for long-range-connected neutral atom registers. *Nature Communications*, 16(1):1111, 2025. doi: 10.1038/s41467-025-56255-5. URL https://doi.org/10.1038/s41467-025-56255-5.

[143] Nikolas P. Breuckmann and Simon Burton. Fold-transversal clifford gates for quantum codes. *Quantum*, 8:1372, 2024. doi: 10.22331/q-2024-06-13-1372. arXiv:2202.06647.

[144] Hasan Sayginel, Stergios Koutsioumpas, Mark Webster, Abhishek Rajput, and Dan E. Browne. Fault-tolerant logical clifford gates from code automorphisms, 2025. URL https://arxiv.org/abs/2409.18175.

[145] Jens Niklas Eberhardt and Vincent Steffan. Logical operators and fold-transversal gates of bivariate bicycle codes, 2024. URL https://arxiv.org/abs/2407.03973.

[146] Armanda O. Quintavalle, Paul Webster, and Michael Vasmer. Partitioning qubits in hypergraph product codes to implement logical gates. *Quantum*, 7:1153, October 2023. ISSN 2521-327X. doi: 10.22331/q-2023-10-24-1153. URL https://doi.org/10.22331/q-2023-10-24-1153.

[147] Dominic Horsman, Austin G Fowler, Simon Devitt, and Rodney Van Meter. Surface code quantum computing by lattice surgery. *New Journal of Physics*, 14(12):123011, December 2012. ISSN 1367-2630. doi: 10.1088/1367-2630/14/12/123011. URL http://dx.doi.org/10.1088/1367-2630/14/12/123011.

[148] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86(3), September 2012. ISSN 1094-1622. doi: 10.1103/physreva.86.032324. URL http://dx.doi.org/10.1103/PhysRevA.86.032324.

[149] Andrew J. Landahl and Ciaran Ryan-Anderson. Quantum computing by color-code lattice surgery, 2014. URL https://arxiv.org/abs/1407.5103.

[150] Andrew W. Cross, Zhiyang He, Patrick J. Rall, and Theodore J. Yoder. Improved QLDPC surgery: Logical measurements and bridging codes, 2024. URL https://arxiv.org/abs/2407.18393.

[151] Alexander Cowtan. Homology, hopf algebras and quantum code surgery, 2025. URL https://arxiv.org/abs/2508.01496.

[152] Zhiyang He, Alexander Cowtan, Dominic J. Williamson, and Theodore J. Yoder. Extractors: Qldpc architectures for efficient pauli-based computation, 2025. URL https://arxiv.org/abs/2503.10390.

[153] Alexander Cowtan. Ssip: automated surgery with quantum ldpc codes, 2024. URL https://arxiv.org/abs/2407.09423.

[154] Alexander Cowtan and Simon Burton. Css code surgery as a universal construction. *Quantum*, 8:1344, May 2024. ISSN 2521-327X. doi: 10.22331/q-2024-05-14-1344. URL http://dx.doi.org/10.22331/q-2024-05-14-1344.

[155] Alexander Cowtan. Qudit lattice surgery, 2022. URL https://arxiv.org/abs/2204.13228.

[156] Benjamin Ide, Manoj G. Gowda, Priya J. Nadkarni, and Guillaume Dauphinais. Fault-tolerant logical measurements via homological measurement. *Physical Review X*, 15(2), June 2025. ISSN 2160-3308. doi: 10.1103/physrevx.15.021088. URL http://dx.doi.org/10.1103/PhysRevX.15.021088.

[157] Esha Swaroop, Tomas Jochym-O'Connor, and Theodore J. Yoder. Universal adapters between quantum ldpc codes, 2025. URL https://arxiv.org/abs/2410.03628.

[158] Sergey Bravyi, Graeme Smith, and John A. Smolin. Trading classical and quantum computational resources. *Phys. Rev. X*, 6:021043, Jun 2016. doi: 10.1103/PhysRevX.6.021043. URL https://link.aps.org/doi/10.1103/PhysRevX.6.021043.

[159] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Phys. Rev. A*, 71:022316, Feb 2005. doi: 10.1103/PhysRevA.71.022316. URL https://link.aps.org/doi/10.1103/PhysRevA.71.022316.

[160] M. B. Hastings. Weight reduction for quantum codes, 2016. URL https://arxiv.org/abs/1611.03790.

[161] M. B. Hastings. On quantum weight reduction, 2023. URL https://arxiv.org/abs/2102.10030.

[162] Eric Sabo, Lane G. Gunderman, Benjamin Ide, Michael Vasmer, and Guillaume Dauphinais. Weight-reduced stabilizer codes with lower overhead. *PRX Quantum*, 5(4), October 2024. ISSN 2691-3399. doi: 10.1103/prxquantum.5.040302. URL http://dx.doi.org/10.1103/PRXQuantum.5.040302.

[163] Qian Xu, Hengyun Zhou, Guo Zheng, Dolev Bluvstein, J. Pablo Bonilla Ataides, Mikhail D. Lukin, and Liang Jiang. Fast and parallelizable logical computation with homological product codes. *Physical Review X*, 15(2):021065, 2025. doi: 10.1103/PhysRevX.15.021065.

[164] Shilin Huang, Tomas Jochym-O'Connor, and Theodore J. Yoder. Homomorphic logical measurements. *PRX Quantum*, 4(3):030301, 2023. doi: 10.1103/PRXQuantum.4.030301.

[165] Jean-Pierre Tillich and Gilles Zémor. Quantum LDPC codes with positive rate and minimum distance proportional to the square root of the blocklength. *IEEE Transactions on Information Theory*, 60(2):1193–1202, 2014. doi: 10.1109/TIT.2013.2292061.

[166] Siyi Yang and Robert Calderbank. Spatially-coupled qldpc codes. *Quantum*, 9:1693, 2025. doi: 10.22331/q-2025-04-07-1693. arXiv:2305.00137.

[167] Earl T Campbell. A theory of single-shot error correction for adversarial noise. *Quantum Science and Technology*, 4(2):025006, February 2019. ISSN 2058-9565. doi: 10.1088/2058-9565/aafc8f. URL http://dx.doi.org/10.1088/2058-9565/aafc8f.

[168] Hengyun Zhou, Chen Zhao, Madelyn Cain, Dolev Bluvstein, Nishad Maskara, Casey Duckering, Hong-Ye Hu, Sheng-Tao Wang, Aleksander Kubica, and Mikhail D. Lukin. Low-overhead transversal fault tolerance for universal quantum computation, 2025. URL https://arxiv.org/abs/2406.17653.

[169] Madelyn Cain, Chen Zhao, Hengyun Zhou, Nadine Meister, J. Pablo Bonilla Ataides, Arthur Jaffe, Dolev Bluvstein, and Mikhail D. Lukin. Correlated decoding of logical algorithms with transversal gates. *Phys. Rev. Lett.*, 133:240602, Dec 2024. doi: 10.1103/PhysRevLett.133.240602. URL https://link.aps.org/doi/10.1103/PhysRevLett.133.240602.

[170] Nouédyn Baspin, Lucas Berent, and Lawrence Z. Cohen. Fast surgery for quantum ldpc codes, 2025. URL https://arxiv.org/abs/2510.04521.

[171] Qian Xu, Hengyun Zhou, Dolev Bluvstein, Madelyn Cain, Marcin Kalinowski, John Preskill, Mikhail D. Lukin, and Nishad Maskara. Batched high-rate logical operations for quantum ldpc codes, 2025. URL https://arxiv.org/abs/2510.06159.

[172] Christophe Vuillot, Lingling Lao, Ben Criger, Carmen García Almudéver, Koen Bertels, and Barbara M Terhal. Code deformation and lattice surgery are gauge fixing. *New Journal of Physics*, 21(3):033028, March 2019. ISSN 1367-2630. doi: 10.1088/1367-2630/ab0199. URL http://dx.doi.org/10.1088/1367-2630/ab0199.

[173] Dominic J. Williamson. Fractal symmetries: Ungauging the cubic code. *Phys. Rev. B*, 94:155128, Oct 2016. doi: 10.1103/PhysRevB.94.155128. URL https://link.aps.org/doi/10.1103/PhysRevB.94.155128.

[174] Aleksander Kubica and Beni Yoshida. Ungauging quantum error-correcting codes, 2018. URL https://arxiv.org/abs/1805.01836.

[175] Daniel Gottesman. Fault-Tolerant Quantum Computation with Constant Overhead. *Quantum Information and Computation*, 14(15-16):1338–1372, 2013. ISSN 15337146. doi: 10.48550/arxiv.1310.2984. URL https://arxiv.org/abs/1310.2984v3.

[176] Michael Freedman and Matthew Hastings. Building manifolds from quantum codes. *Geometric and Functional Analysis*, 31, 08 2021. doi: 10.1007/s00039-021-00567-3.

[177] Alexander J. Malcolm, Andrew N. Glaudell, Patricio Fuentes, Daryus Chandra, Alexis Schotte, Colby DeLisle, Rafael Haenel, Amir Ebrahimi, Joschka Roffe, Armanda O. Quintavalle, Stefanie J. Beale, Nicholas R. Lee-Hone, and Stephanie Simmons. Computing efficiently in qldpc codes, 2025. URL https://arxiv.org/abs/2502.07150.

[178] Noah Berthusen, Michael J. Gullans, Yifan Hong, Maryam Mudassar, and Shi Jie Samuel Tan. Automorphism gadgets in homological product codes, 2025. URL https://arxiv.org/abs/2508.04794.

[179] Pavel Panteleev and Gleb Kalachev. Quantum ldpc codes with almost linear minimum distance. *IEEE Transactions on Information Theory*, 68(1):213–229, January 2022. ISSN 1557-9654. doi: 10.1109/tit.2021.3119384. URL http://dx.doi.org/10.1109/TIT.2021.3119384.

[180] Matthew B. Hastings, Jeongwan Haah, and Ryan O'Donnell. Fiber bundle codes: breaking the n 1/2 polylog(n) barrier for quantum ldpc codes. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC '21, page 1276–1288. ACM, June 2021. doi: 10.1145/3406325.3451005. URL http://dx.doi.org/10.1145/3406325.3451005.

[181] Nikolas P. Breuckmann and Jens N. Eberhardt. Balanced product quantum codes. *IEEE Transactions on Information Theory*, 67(10):6653–6674, October 2021. ISSN 1557-9654. doi: 10.1109/tit.2021.3097347. URL http://dx.doi.org/10.1109/TIT.2021.3097347.

[182] Maine Christos, Chiu Fan Bowen Lo, Vedika Khemani, and Rahul Sahay. Non-abelian quantum low-density parity check codes and non-clifford operations from gauging logical gates via measurements, 2026. URL https://arxiv.org/abs/2602.12228.

[183] Guanyu Zhu, Ryohei Kobayashi, and Po-Shen Hsin. Non-abelian qLDPC: TQFT formalism, addressable gauging measurement and application to magic state fountain on 2D product codes, 2026. URL https://arxiv.org/abs/2601.06736.

[184] Luis Colmenarez, Ze-Min Huang, Sebastian Diehl, and Markus Müller. Accurate optimal quantum error correction thresholds from coherent information. *Physical Review Research*, 6:L042014, 2024. doi: 10.1103/PhysRevResearch.6.L042014.

[185] Jong Yeon Lee. Exact calculations of coherent information for toric codes under decoherence: Identifying the fundamental error threshold. *Physical Review Letters*, 134: 250601, 2025. doi: 10.1103/hlfh-86yz.

[186] Craig Gidney, Noah Shutty, and Cody Jones. Magic state cultivation: growing T states as cheap as CNOT gates, 2024. URL https://arxiv.org/abs/2409.17595.

[187] Daniel Litinski. Magic state distillation: Not as costly as you think. *Quantum*, 3:205, 2019. doi: 10.22331/q-2019-12-02-205. URL https://arxiv.org/abs/1905.06903.

[188] Han Zheng, Guo Zheng, Liang Jiang, and Qian Xu. Logical computation with canonical lifted product codes. *arXiv preprint*, 2026.

[189] Carlton M. Caves. Quantum error correction and reversible operations. *Journal of Superconductivity*, 12(6):707–718, 1999. doi: 10.1023/A:1007720606911. URL https://doi.org/10.1023/A:1007720606911.

[190] V. P. Belavkin. Optimum distinction of non-orthogonal quantum signals. *Radiotekhnika i Elektronika*, 20:1177–1185, June 1975.

[191] A. S. Kholevo. On asymptotically optimal hypothesis testing in quantum statistics. *Theory of Probability & Its Applications*, 23(2):411–415, 1979. doi: 10.1137/1123048. URL https://doi.org/10.1137/1123048.

[192] Paul A. Hausladen and William K. Wootters. A 'pretty good' measurement for distinguishing quantum states. *Journal of Modern Optics*, 41:2385–2390, 1994. URL https://api.semanticscholar.org/CorpusID:122223374.

[193] Gavin E. Crooks. Quantum operation time reversal. *Phys. Rev. A*, 77:034101, Mar 2008. doi: 10.1103/PhysRevA.77.034101. URL https://link.aps.org/doi/10.1103/PhysRevA.77.034101.

[194] Clive Cenxin Aw, Francesco Buscemi, and Valerio Scarani. Fluctuation theorems with retrodiction rather than reverse processes. *AVS Quantum Science*, 3(4):045601, 10 2021. ISSN 2639-0213. doi: 10.1116/5.0060893. URL https://doi.org/10.1116/5.0060893.

[195] Hyukjoon Kwon and M. S. Kim. Fluctuation theorems for a quantum channel. *Phys. Rev. X*, 9:031029, Aug 2019. doi: 10.1103/PhysRevX.9.031029. URL https://link.aps.org/doi/10.1103/PhysRevX.9.031029.

[196] Francesco Buscemi and Valerio Scarani. Fluctuation theorems from bayesian retrodiction. *Phys. Rev. E*, 103:052111, May 2021. doi: 10.1103/PhysRevE.103.052111. URL https://link.aps.org/doi/10.1103/PhysRevE.103.052111.

[197] Chao Shen, Kyungjoo Noh, Victor V. Albert, Stefan Krastanov, M. H. Devoret, R. J. Schoelkopf, S. M. Girvin, and Liang Jiang. Quantum channel construction with circuit quantum electrodynamics. *Phys. Rev. B*, 95:134501, Apr 2017. doi: 10.1103/PhysRevB.95.134501. URL https://link.aps.org/doi/10.1103/PhysRevB.95.134501.

[198] Raban Iten, Roger Colbeck, and Matthias Christandl. Quantum circuits for quantum channels. *Phys. Rev. A*, 95:052316, May 2017. doi: 10.1103/PhysRevA.95.052316. URL https://link.aps.org/doi/10.1103/PhysRevA.95.052316.

[199] S. Rosenblum, P. Reinhold, M. Mirrahimi, Liang Jiang, L. Frunzio, and R. J. Schoelkopf. Fault-tolerant detection of a quantum error. *Science*, 361(6399):266–270, July 2018. ISSN 1095-9203. doi: 10.1126/science.aat3996. URL http://dx.doi.org/10.1126/science.aat3996.

[200] Debjyoti Biswas, Gaurav M. Vaidya, and Prabha Mandayam. Noise-adapted recovery circuits for quantum error correction. 2023.

[201] F. Caruso, V. Giovannetti, and A. S. Holevo. One-mode bosonic Gaussian channels: A full weak-degradability classification. *New J. Phys.*, 8:310, 2006. doi: 10.1088/1367-2630/8/12/310.

[202] Francesco Anna Mele, Giuseppe Catalano, Marco Fanizza, Vittorio Giovannetti, and Ludovico Lami. Bosonic quantum communication beyond the thermal threshold, 2026.

[203] Giacomo De Palma and Dario Trevisan. The conditional entropy power inequality for bosonic quantum systems. *Communications in Mathematical Physics*, 360(2):639–662, 2018. doi: 10.1007/s00220-017-3082-8.

[204] Giacomo De Palma and Stefan Huber. The conditional entropy power inequality for quantum additive noise channels. *Journal of Mathematical Physics*, 59:122201, 2018. doi: 10.1063/1.5027495.

[205] Elliott H. Lieb and Robert Seiringer. Stronger subadditivity of entropy. *Physical Review A*, 71:062329, 2005. doi: 10.1103/PhysRevA.71.062329.

[206] Kyungjoo Noh, S. M. Girvin, and Liang Jiang. Encoding an oscillator into many oscillators. *Phys. Rev. Lett.*, 125:080503, Aug 2020. doi: 10.1103/PhysRevLett.125.080503. URL https://link.aps.org/doi/10.1103/PhysRevLett.125.080503.

[207] Kyungjoo Noh, Christopher Chamberland, and Fernando G.S.L. Brandão. Low-overhead fault-tolerant quantum error correction with the surface-gkp code. *PRX Quantum*, 3:010315, Jan 2022. doi: 10.1103/PRXQuantum.3.010315. URL https://link.aps.org/doi/10.1103/PRXQuantum.3.010315.

[208] Mackenzie H. Shaw, Andrew C. Doherty, and Arne L. Grimsmo. Stabilizer subsystem decompositions for single- and multimode gottesman-kitaev-preskill codes. *PRX Quantum*, 5(1), February 2024. ISSN 2691-3399. doi: 10.1103/prxquantum.5.010331. URL http://dx.doi.org/10.1103/PRXQuantum.5.010331.

[209] Arne L. Grimsmo, Joshua Combes, and Ben Q. Baragiola. Quantum computing with rotation-symmetric bosonic codes. *Phys. Rev. X*, 10:011058, Mar 2020. doi: 10.1103/PhysRevX.10.011058. URL https://link.aps.org/doi/10.1103/PhysRevX.10.011058.

[210] David Poulin. Stabilizer formalism for operator quantum error correction. *Phys. Rev. Lett.*, 95:230504, Dec 2005. doi: 10.1103/PhysRevLett.95.230504. URL https://link.aps.org/doi/10.1103/PhysRevLett.95.230504.

[211] Armanda O. Quintavalle, Michael Vasmer, Joschka Roffe, and Earl T. Campbell. Single-shot error correction of three-dimensional homological product codes. *PRX Quantum*, 2:020340, Jun 2021. doi: 10.1103/PRXQuantum.2.020340. URL https://link.aps.org/doi/10.1103/PRXQuantum.2.020340.

[212] Anthony Leverrier, Jean-Pierre Tillich, and Gilles Zemor. Quantum expander codes. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, page 810–824. IEEE, October 2015. doi: 10.1109/focs.2015.55. URL http://dx.doi.org/10.1109/FOCS.2015.55.

[213] Héctor Bombín. Single-shot fault-tolerant quantum error correction. *Physical Review X*, 5(3), September 2015. ISSN 2160-3308. doi: 10.1103/physrevx.5.031043. URL http://dx.doi.org/10.1103/PhysRevX.5.031043.

[214] Benjamin J. Brown, Naomi H. Nickerson, and Dan E. Browne. Fault-tolerant error correction with the gauge color code. *Nature Communications*, 7(1), July 2016. ISSN 2041-1723. doi: 10.1038/ncomms12302. URL http://dx.doi.org/10.1038/ncomms12302.

[215] Shai Evra, Tali Kaufman, and Gilles Zémor. Decodable quantum ldpc codes beyond the square root distance barrier using high dimensional expanders. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 218–227, 2020. doi: 10.1109/FOCS46700.2020.00029.

[216] Tali Kaufman and Ran J. Tessler. New cosystolic expanders from tensors imply explicit quantum ldpc codes with $\Omega(\sqrt{n} \log n)$ distance. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, page 1317–1329, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380539. doi: 10.1145/3406325.3451029. URL https://doi.org/10.1145/3406325.3451029.

[217] Irit Dinur, Shai Evra, Ron Livne, Alexander Lubotzky, and Shahar Mozes. Locally testable codes with constant rate, distance, and locality, 2021. URL https://arxiv.org/abs/2111.04808.

[218] Ting-Chun Lin and Min-Hsiu Hsieh. c3-locally testable codes from lossless expanders. In *2022 IEEE International Symposium on Information Theory (ISIT)*, pages 1175–1180, 2022. doi: 10.1109/ISIT50566.2022.9834679.

[219] Adam Wills, Ting-Chun Lin, and Min-Hsiu Hsieh. Tradeoff constructions for quantum locally testable codes. *IEEE Transactions on Information Theory*, 71(1):426–458, 2025. doi: 10.1109/TIT.2024.3503500.

[220] Andrew Cross, Zhiyang He, Anand Natarajan, Mario Szegedy, and Guanyu Zhu. Quantum Locally Testable Code with Constant Soundness. *Quantum*, 8:1501, October 2024. ISSN 2521-327X. doi: 10.22331/q-2024-10-18-1501. URL https://doi.org/10.22331/q-2024-10-18-1501.

[221] Michael Sipser and Daniel A. Spielman. Expander codes. *IEEE Transactions on Information Theory*, 42(6):1710–1722, November 1996. doi: 10.1109/18.556667.

[222] Sergey Bravyi and Matthew B. Hastings. Homological product codes. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, page 273–282, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450327107. doi: 10.1145/2591796.2591870. URL https://doi.org/10.1145/2591796.2591870.

[223] Weilei Zeng and Leonid P. Pryadko. Higher-dimensional quantum hypergraph-product codes with finite rates. *Phys. Rev. Lett.*, 122:230501, Jun 2019. doi: 10.1103/PhysRevLett.122.230501. URL https://link.aps.org/doi/10.1103/PhysRevLett.122.230501.

[224] Eli Ben-Sasson and Madhu Sudan. Robust locally testable codes and products of codes, 2004. URL https://arxiv.org/abs/cs/0408066.

[225] Gleb Kalachev and Pavel Panteleev. Two-sided robustly testable codes, 2023. URL https://arxiv.org/abs/2206.09973.

[226] Irit Dinur, Madhu Sudan, and Avi Wigderson. Robust local testability of tensor products of ldpc codes. In *Proceedings of the 9th International Conference on Approximation Algorithms for Combinatorial Optimization Problems, and 10th International Conference on Randomization and Computation*, APPROX'06/RANDOM'06, page 304–315, Berlin, Heidelberg, 2006. Springer-Verlag. ISBN 3540380442. doi: 10.1007/11830924_29. URL https://doi.org/10.1007/11830924_29.

[227] Markus Grassl and Martin Roetteler. Leveraging automorphisms of quantum codes for fault-tolerant quantum computation. In *2013 IEEE International Symposium on Information Theory*, pages 534–538, 2013. doi: 10.1109/ISIT.2013.6620283.

[228] Armanda O. Quintavalle and Earl T. Campbell. Reshape: A decoder for hypergraph product codes. *IEEE Transactions on Information Theory*, 68(10):6569–6584, 2022. doi: 10.1109/TIT.2022.3184108.

[229] Christopher Chamberland and Earl T. Campbell. Universal quantum computing with twist-free and temporally encoded lattice surgery. *PRX Quantum*, 3:010331, 2022. doi: 10.1103/PRXQuantum.3.010331.