Boosting the Gottesman-Kitaev-Preskill quantum error correction with non-Markovian feedback

Matteo Puviani,[1,*] Sangkha Borah,[1,2] Remmy Zen,[1] Jan Olle,[1] and Florian Marquardt[1,2,†]

[1] *Max Planck Institute for the Science of Light, 91058 Erlangen, Germany*
[2] *Department of Physics, Friedrich-Alexander Universität Erlangen-Nürnberg, 91058 Erlangen, Germany*

Bosonic codes allow the encoding of a logical qubit in a single component device, utilizing the infinitely large Hilbert space of a harmonic oscillator. In particular, the Gottesman-Kitaev-Preskill code has recently been demonstrated to be correctable well beyond the break-even point of the best passive encoding in the same system. Current approaches to quantum error correction (QEC) for this system are based on protocols that use feedback, but the response is based only on the latest measurement outcome. In our work, we use the recently proposed Feedback-GRAPE (Gradient Ascent Pulse Engineering with Feedback) method to train a recurrent neural network that provides a QEC scheme based on memory, responding in a non-Markovian way to the full history of previous measurement outcomes, optimizing all subsequent unitary operations. This approach significantly outperforms current strategies and paves the way for more powerful measurement-based QEC protocols.

Introduction. The key components of quantum computers are quantum bits or qubits, which can be implemented in a variety of hardware platforms, such as superconducting circuits [1–3], trapped ions [4–6], Rydberg atoms [7], and photonic systems [8]. Irrespective of the platform, quantum decoherence that arises from relaxation and dephasing of the quantum components represents the greatest limitation for building quantum computers that are able to run large scale quantum algorithms. The main objective of the current research is the implementation of reliable *quantum error correction* (QEC). Reliable QEC protocols can correct errors on logical qubits, extending the lifetime of the logical information stored in them far beyond the coherence time of individual physical qubits used for encoding a logical qubit. This can be useful for the purposes of quantum memories, reliable quantum operations and quantum computing in general.

The most widespread class of approaches to QEC is represented by the the so-called *stabilizer codes*, where n physical qubits are used to encode $k < n$ logical qubits [10–13]. This redundancy, along with the concept of quantum entanglement, can be used to correct a limited number of independent physical errors [14]. An alternative to such encoding is to use systems with continuous variables, namely bosonic systems, called *bosonic encodings* [15, 16]. Unlike the qubit-based encoding, the redundancy is achieved within a single physical component by exploiting the infinitely large Hilbert space of bosonic systems (i.e. quantum oscillators, such as photons in an optical cavity) to be able to perform QEC. Bosonic codes are more hardware-efficient as there is often only one or very few dominant decay channels to be corrected for, and they offer exceedingly long coherence times in the microwave range. However, they need nonlinear elements for processing, they require a protocol to prepare states and the overall strategies are still less developed than for multi-qubit encoding [17].

A variety of bosonic encodings exist, such as the Pegg-Barnett [18], cat (and kitten) [19–24], *Gottesman-Kitaev-Preskill* (GKP) [2, 17, 25–27] and binomial [28, 29] codes, but recently the original GKP code has received strong renewed attention. Using reinforcement learning (RL)-based controls, Sivak et al. [3] have demonstrated GKP error correction to achieve beyond the so-called break-even point of the best passive encoding in the same system, extending the lifetime by a factor 2. This represents by far the longest qubit lifetime extension reached by any platform [2, 29–34]. However, even though current approaches to QEC for this system are based on protocols that use feedback, this is only limited to a few operations, and the response depends only on the latest measurement outcome.

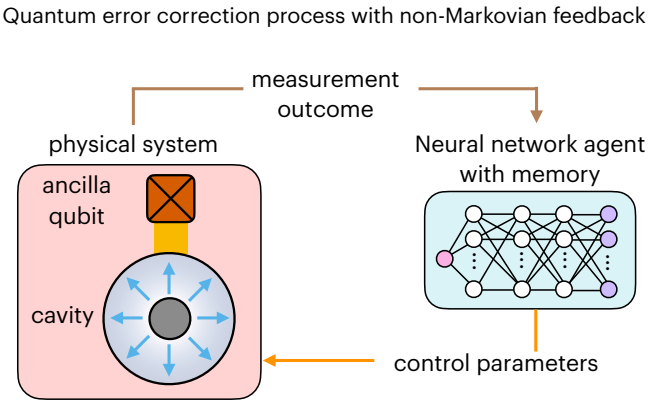

Figure 1. Scheme of the proposed quantum error correction process via non-Markovian feedback. The physical system is represented by a cavity coupled to a control qubit. The result of each binary quantum measurement is used as input of a neural network with memory, which responds to all the sequence of measurements providing the optimized parameters to control the system.

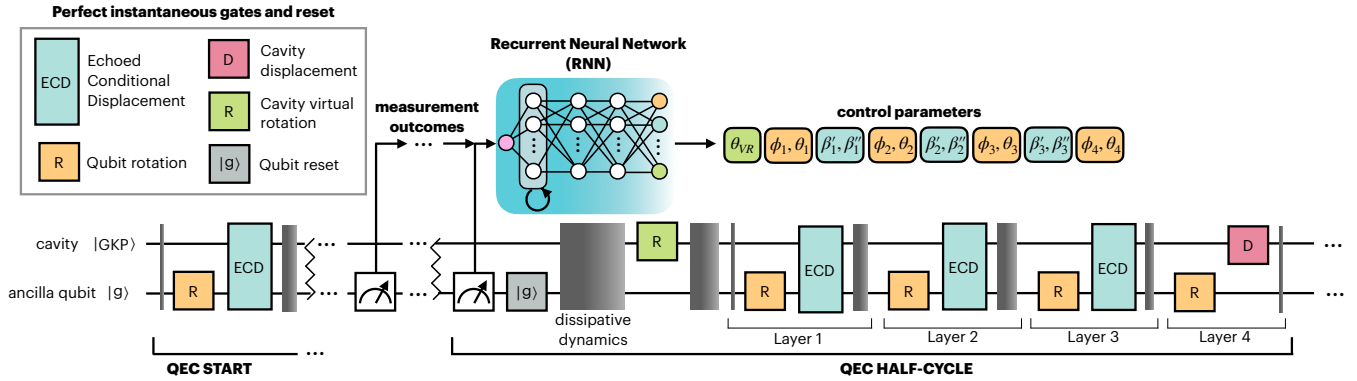

Figure 2. Quantum error correction feedback protocol. The scheme represents the sBs QEC protocol with the NMF approach implemented with a RNN. On each cycle, the latest outcome of the measurement performed on the qubit is used as input for the RNN, which keeps memory of all the previous sequence of inputs and suggests the parameters used in the subsequent gates. The grey areas in the QEC circuit represent the duration of the dissipative dynamics [9].

Recently, Porotti et al. [35] have developed a method for using gradient-ascent pulse engineering (GRAPE) with feedback, known as *Feedback-GRAPE*, that allows for a feedback-based optimal control with a model-based approach. This has been demonstrated to outperform state-of-the-art model-free RL methods in tasks like state purification, state preparation and stabilization, and bosonic code preparation.

In this work, we propose an advanced GKP QEC scheme making use of a recurrent neural network (RNN), which is able to provide an elaborate QEC strategy based on the memory of the entire sequence of measurement outcomes. In particular, we use *Feedback-GRAPE* to train the RNN, which suggests optimized parameters of the GKP QEC circuit based on real-time feedback from measurements of the quantum system (Fig. 1). Thanks to the memory of the neural network, the real-time strategy adopted is non-Markovian and is able to boost the theoretical performance of standard strategies via non-Markovian feedback (NMF), increasing the logical qubit's lifetime by a factor higher than 2.

Theoretical background. In order to be able to counteract the effect of noise in quantum devices, and thus to correct for errors which would lead to the loss of stored quantum information, different QEC protocols have been developed for the GKP code in the past decades [17]. We can conveniently distinguish two classes of QEC strategies: namely, autonomous [26] and feedback-based [2, 3]. While the former does not include any measurement, the latter relies on a feedback loop based on projective measurements of an ancillary qubit to perform operations dependent on the measurement outcomes. Although both approaches have advantages and disadvantages, feedback-based protocols have shown to have a better performance in experiments than autonomous QEC schemes [2, 3], naturally allowing for further extensions and developments. In our work we focus entirely on the up-to-date best optimized measurement-based protocol: namely, the so-called small-BIG-small (sBs) protocol [3, 26, 36]. In particular, we adopt the implementation used by Sivak et al. [3], since it is experimentally feasible and it has been demonstrated to be more versatile for further optimization. This protocol is based on four elementary gates which allow for a universal control of the cavity-qubit system: namely, the *echoed conditional displacement* $E\hat{C}D_{qc}$ (with complex parameter β), acting as an entangling gate on both the ancillary qubit and the cavity; the *qubit rotation* $\hat{R}_q$ (with real parameters ϕ, θ); the cavity *displacement* $\hat{D}_c$ (with real parameter α); and the cavity *virtual rotation* $\hat{VR}_c$ (with real parameter θ_{VR}) [9]. Quantum error correction is performed by applying an even number of half-cycles, each of which includes the same fixed sequence of gates (see Fig. 2). Initially, the ancilla qubit is prepared in its ground state $|g\rangle$, and then 4 layers of the QEC half-cycle consisting of qubit rotation and cavity displacement are applied. At this time, a measurement on the ancillary qubit is performed, projecting its state along the z axis, returning g (e) if projected onto its ground (excited) state $|g\rangle$ ($|e\rangle$). In our work we neglect leakage errors which could get the qubit to be in higher excited states. After the measurement, the ancilla qubit is reset to its ground state just before the next stage of control sequence takes place.

The QEC task presented can be described as a *quantum observable Markov decision process* (a physical application of *hidden Markov models* and *partially observable Markov decision processes*) [37–40], since it is not possible to (experimentally) access the full quantum state during the process: the only available information is the outcome of the projective measurements of the ancillary qubit entangled with the oscillator. However, each observation alone is insufficient to determine the best parameters' set for the applied gates in the following QEC cy-

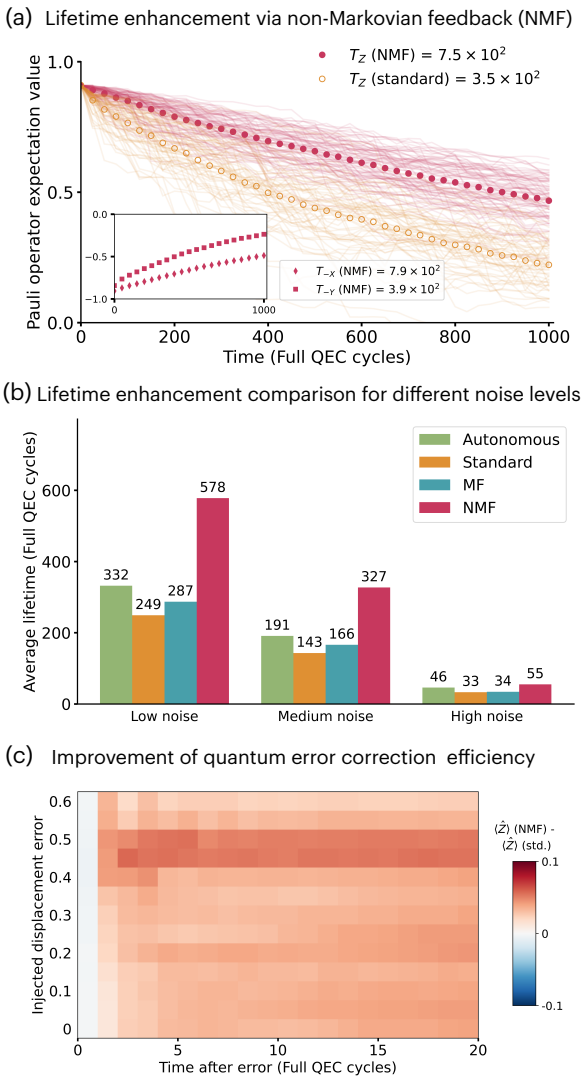

Figure 3. Logical qubit's lifetime extension. (a) Comparison of the logical state $|+Z_L\rangle$ time evolution (dots: average over multiple trajectories, continuous faded lines: sampled trajectories averaged over the batch-size) for the standard sBs protocol and the NMF approach advocated in this work. The inset shows the results obtained with NMF for $|-Y_L\rangle$ and $|-X_L\rangle$. (b) Lifetime of average channel fidelity for four implementations of the sBs protocol: autonomous, standard, MF and NMF. The results are obtained for fixed time intervals of each cycle (Fig. 2) and different noise levels (characterized by decoherence times T_1, T_2, T_s) [9]. (c) QEC after injection of a displacement error: difference between the NMF strategy and the standard sBs protocol, after averaging over 1032 trajectories.

cle. Therefore, we adopt a RNN to integrate information over time, effectively building an internal representation (or belief state) that approximates the full state of the system. In fact, this capability of RNNs has been often exploited in different complex tasks like natural language processing, time series prediction, and many others with applications to scientific problems [41–45]. As a result of the solution adopted, the QEC strategy proposed becomes dependent also on all the past observations: in this sense the strategy is non-Markovian.

We implement a realistic simulation of the collective entangled dynamics of the cavity and the ancilla qubit, with fixed time dynamics (Fig. 2) and full cycle duration τ_{cycle}. This allows us to compute with high fidelity the probability that a sequence of measurement outcomes is obtained, which is used to train the neural network resulting in a model-aware agent. We call m_i (with $i \in \mathrm{N}$) the binary measurement outcome on the ancilla qubit entangled with the cavity in the half-cycle i, so that $m_i \in \{g, e\}$. The full sequence of measurement outcomes is contained in the vector $\mathbf{m} = (m_0, m_1, \ldots, m_N)$, and $P(\mathbf{m})$ is the probability of the full trajectory to occur. In order to extend the logical qubit's lifetime, our goal is to maximize a return function $\mathcal{R}$ that we choose to be the fidelity of the final density matrix with respect to the initial state: namely, $\mathcal{R} = \mathcal{F}(\rho_{Z_L}, \rho(T))$, with $\mathcal{F}(\sigma, \rho) = [\mathrm{Tr}\{(\sqrt{\hat{\sigma}}\hat{\rho}\sqrt{\hat{\sigma}})^{1/2}\}]^2$. Since we deal with probabilistic sequence of outcomes, we have to maximize the weighted-average cumulative return $\langle \mathcal{R}(\mathbf{m}) \rangle_m = \sum_{\mathbf{m}} P(\mathbf{m})\, \mathcal{R}(\mathbf{m})$ over all the possible trajectories, according to their corresponding probability given. This is performed using automatic differentiation of the expression

$$\frac{\partial \langle \mathcal{R}(\mathbf{m}) \rangle_{\mathbf{m}}}{\partial \theta} = \left\langle \frac{\partial \mathcal{R}(\mathbf{m})}{\partial \theta} + \mathcal{R}(\mathbf{m}) \frac{\partial \ln P_\theta(\mathbf{m})}{\partial \theta} \right\rangle_{\mathbf{m}}, \quad (1)$$

in order to compute gradients through the multi-step dynamics. While the first term on the *rhs* of Eq. (1) is straightforward, the second one is essential to include the dependence of the probability of each trajectory on the parameters vector θ of the RNN, as described in Ref. [35]. The solution provides us with the best weights and biases for the RNN to be used to suggest the optimal QEC strategy.

Results. In our numerical simulations we include the dissipative dynamics via the solution of the Lindblad master equation to simulate an approximation of realistic experimental dynamics, using experimental parameters from Ref. [3], including the qubit measurement and reset time as well as the delay for the application of each gate [9]. The training is performed on a fixed initial logical state ($|Z_L\rangle$) with trajectories of 10 full QEC cycles, in the presence of high noise level [9]. The evaluation of the strategy has been performed on 1000 full QEC cycles, thus extending the time to a much longer range than the one used for the training of the neural network agent itself, and on different noise levels.

In Fig. 3(a), we show the results of QEC with low noise level (with values from the experiment in Ref. [3]) averaged over more than 500 total trajectories for a logical qubit $|+Z_L\rangle$, both with the standard sBs protocol [3] (see [9]) and with our non-Markovian-feedback approach. The lifetime more than doubles with the use of NMF:

in fact it increases from $T_Z(\text{std.})/\tau_{cycle} = 3.5 \times 10^2$ to $T_Z(\text{NFM})/\tau_{cycle} = 7.5 \times 10^2$. Moreover, the learned optimal strategy also generalizes to any initial logical qubit state without re-training, maximizing each corresponding Pauli operator expectation value and extending the qubit's lifetime. Indeed, a comparable result is obtained (inset of Fig. 3(a)) using the same neural network with the initial logical state $|-X_L\rangle$, while for the $|-Y_L\rangle$ we get a lower value of $T_{-Y}(\text{NMF})/\tau_{cycle} = 3.9 \times 10^2$, which is an expected feature of the square GKP code [2, 3]. Moreover, we show that a non-Markovian-feedback strategy based on a RNN is necessary to boost the performance of QEC. In fact, an instantaneous Markovian feedback (MF) implemented with a feed-forward neural network, which is capable of providing distinct parameters' values according to the latest measurement output only, performs comparably with the standard approach with no feedback (Fig. 3(b)). These results hold even for very different noise levels: from the low noise in Ref. [3] to the higher noise of the experiment in Ref. [2], as well as in the presence of imperfect gates [9].

In addition, we notice that our NMF approach considerably outperforms other approaches, including the autonomous QEC scheme which has a shorter protocol duration [9]. Apparently, measurement based protocols are not intrinsically outperforming the autonomous one under all circumstances, but they become strongly advantageous when memory is exploited, as in our scheme. We also compare the standard and the NMF sBs protocols in their ability to correct deliberately injected errors (Fig. 3(c)): by injecting displacement errors to the initial $|+Z_L\rangle$ state before QEC, we compare the ability of the two methods to recover the initial logical information, namely the expectation value of the logical Pauli operator Z_L. The plot in Fig. 3(c) shows that the performance difference of our approach and the standard sBs is always positive, confirming that our approach allows to reach higher values in a shorter time compared to the standard QEC strategy.

We now analyze the strategy discovered by the RNN with non-Markovian feedback, showing how memory is exploited to outperform QEC. In Fig. 4 we present one case of a trajectory for the initial logical state Z_L with noisy dynamics and fixed post-selected measurement outcomes that we prescribe (sequence of outcomes: 10 g, 10 e and 10 g). This allows us to track the values of the 15 parameters optimized with the NMF approach in an easily interpretable situation of subsequent g and e measurement outcomes. For most of the parameters we can notice that the initial values are relatively close to the standard one (represented by dashed lines) [3], and then they evolve over time with an alternating exponential-like behavior according to the previous measurement outcomes.

We also find that our optimal strategy has the same probability of the qubit to be found in the ground state

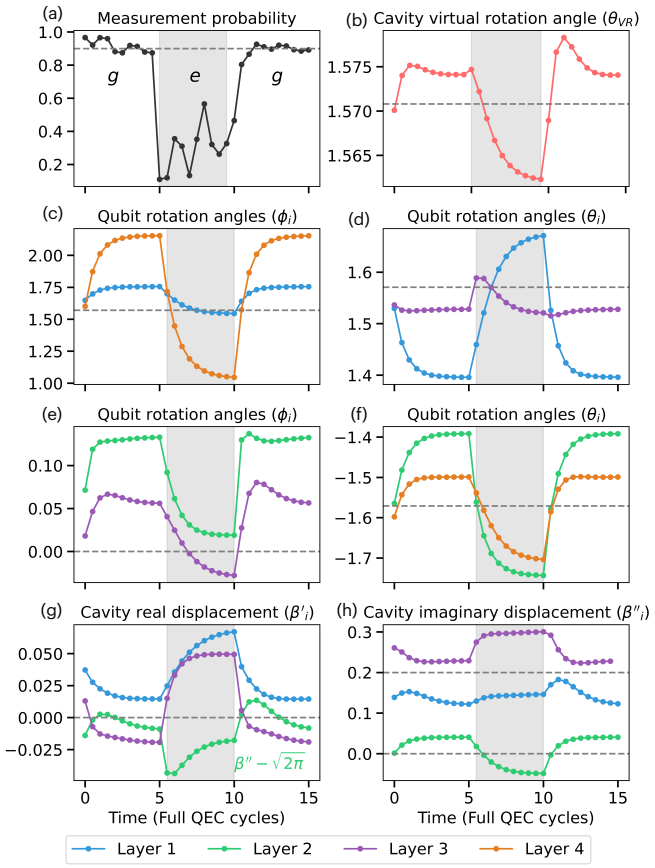

Figure 4. Optimized non-Markovian parameters. (a) Measurement probabilities for a given QEC trajectory, with fixed measurement outcomes. The value g (e) corresponds to an outcome indicating the ancilla qubit state in $|g\rangle$ ($|e\rangle$). The dashed line indicates the probability 0.9. (b) Virtual rotation angle applied at the end of each QEC cycle. The vertical shaded area represents the cycle's timestep at which the measurement $m = e$ is obtained and the virtual rotation is applied. (c)-(h) Parameters entering the 4 layers of the QEC cycle in Fig. 2 suggested by the RNN. The vertical shaded areas represent the cycles in which the RNN is given the e measurement as input. The dashed horizontal lines are a guide for the eyes indicating the corresponding constant values of the standard sBs protocol.

g after the measurement, $p(g) \approx 0.9$, as in the standard approach (see Fig. 4(a)). This observation is important because as explained in Ref. [3] this is experimentally desirable when implementing the sBs measurement-based protocol. Other strategies that we found, instead, occasionally exhibit a longer lifetime in some conditions, but they are less generalizable to different noise levels and show a low probability $p(g)$ averaged over time, thus becoming less auspicable for an experimental implementation. We present more systematic results for other agents, the Markovian-feedback strategy, as well as for biased-noise gates and different dynamics in the Supplemental Material [9].

Conclusion. In this work, we have proposed a complex measurement-based QEC scheme for GKP codes based on non-Markovian feedback. In particular, we have used a recurrent neural network implemented in the QEC circuit in order to optimize all the QEC operations based on the full history of previous measurement outcomes. We have also shown that our QEC scheme remarkably outperforms current autonomous (without feedback) and measurement-based (with Markovian feedback) QEC strategies in different conditions of noise level and dynamics. To reach this result we have used the recently proposed Feedback-GRAPE [35] (Gradient Ascent Pulse Engineering with Feedback) method to train the RNN exploiting a model-based approach.
More generally, our work demonstrates the power of optimizing feedback protocols in the shape of variational quantum circuits [46–48] with in-sequence discrete measurements. We additionally envision the possibility to extend this non-Markovian approach to any measurement-based QEC strategy for the other bosonic codes, with the model-based approach enabled by the Feedback-GRAPE algorithm.

Acknowledgments. Fruitful discussions with Vittorio Peano are thankfully acknowledged. The research is part of the Munich Quantum Valley, which is supported by the Bavarian state government with funds from the Hightech Agenda Bayern Plus.

* matteo.puviani@mpl.mpg.de
† florian.marquardt@mpl.mpg.de

[1] M. H. Devoret, A. Wallraff, and J. M. Martinis, Superconducting qubits: A short review (2004), arXiv:cond-mat/0411174 [cond-mat.mes-hall].
[2] 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**, 368 (2020).
[3] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Real-time quantum error correction beyond break-even, Nature **616**, 50 (2023).
[4] D. Kienzler, H.-Y. Lo, V. Negnevitsky, C. Flühmann, M. Marinelli, and J. P. Home, Quantum harmonic oscillator state control in a squeezed fock basis, Phys. Rev. Lett. **119**, 033602 (2017).
[5] C. Flühmann, V. Negnevitsky, M. Marinelli, and J. P. Home, Sequential modular position and momentum measurements of a trapped ion mechanical oscillator, Phys. Rev. X **8**, 021001 (2018).
[6] C. Flühmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. P. Home, Encoding a qubit in a trapped-ion mechanical oscillator, Nature **566**, 513 (2019).
[7] X. Wu, X. Liang, Y. Tian, F. Yang, C. Chen, Y.-C. Liu, M. K. Tey, and L. You, A concise review of rydberg atom based quantum computation and quantum simulation, Chinese Physics B **30**, 020305 (2021).
[8] S. Konno, W. Asavanant, F. Hanamura, H. Nagayoshi, K. Fukui, A. Sakaguchi, R. Ide, F. China, M. Yabuno, S. Miki, H. Terai, K. Takase, M. Endo, P. Marek, R. Filip, P. van Loock, and A. Furusawa, Propagating gottesman-kitaev-preskill states encoded in an optical oscillator (2023), arXiv:2309.02306 [quant-ph].
[9] Supplementary Material for Outperforming Gottesman-Kitaev-Preskill quantum error correction with feedback-memory.
[10] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, Phys. Rev. A **52**, R2493 (1995).
[11] A. R. Calderbank and P. W. Shor, Good quantum error-correcting codes exist, Phys. Rev. A **54**, 1098 (1996).
[12] A. M. Steane, Error correcting codes in quantum theory, Phys. Rev. Lett. **77**, 793 (1996).
[13] R. Raussendorf and J. Harrington, Fault-tolerant quantum computation with high threshold in two dimensions, Phys. Rev. Lett. **98**, 190504 (2007).
[14] D. Gottesman, Stabilizer codes and quantum error correction (1997), arXiv:quant-ph/9705052 [quant-ph].
[15] W. Cai, Y. Ma, W. Wang, C.-L. Zou, and L. Sun, Bosonic quantum error correction codes in superconducting quantum circuits, Fundamental Research **1**, 50 (2021).
[16] Q. Xu, P. Zeng, D. Xu, and L. Jiang, Fault-tolerant operation of bosonic qubits with discrete-variable ancillae (2023), arXiv:2310.20578 [quant-ph].
[17] A. L. Grimsmo and S. Puri, Quantum error correction with the gottesman-kitaev-preskill code, PRX Quantum **2**, 020101 (2021).
[18] S. M. Barnett and D. T. Pegg, Phase in quantum optics, Journal of Physics A: Mathematical and General **19**, 3849 (1986).
[19] P. T. Cochrane, G. J. Milburn, and W. J. Munro, Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping, Phys. Rev. A **59**, 2631 (1999).
[20] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret, Dynamically protected cat-qubits: a new paradigm for universal quantum computation, New Journal of Physics **16**, 045014 (2014).
[21] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Extending the lifetime of a quantum bit with error correction in superconducting circuits, Nature **536**, 441–445 (2016).
[22] R. Lescanne, M. Villiers, T. Peronnin, A. Sarlette, M. Delbecq, B. Huard, T. Kontos, M. Mirrahimi, and Z. Leghtas, Exponential suppression of bit-flips in a qubit encoded in an oscillator, Nature Physics **16**, 509–513 (2020).
[23] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar, and M. H. Devoret, Stabilization and operation of a kerr-cat qubit, Nature **584**, 205–209 (2020).
[24] Q. Xu, G. Zheng, Y.-X. Wang, P. Zoller, A. A. Clerk, and L. Jiang, Autonomous quantum error correction and

fault-tolerant quantum computation with squeezed cat qubits, Npj Quantum Inf. **9** (2023).
[25] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).
[26] B. Royer, S. Singh, and S. M. Girvin, Stabilization of finite-energy gottesman-kitaev-preskill states, Phys. Rev. Lett. **125**, 260509 (2020).
[27] A. J. Brady, A. Eickbusch, S. Singh, J. Wu, and Q. Zhuang, Advances in bosonic quantum error correction with gottesman-kitaev-preskill codes: Theory, engineering and applications (2023), arXiv:2308.02913 [quant-ph].
[28] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New class of quantum error-correcting codes for a bosonic mode, Phys. Rev. X **6**, 031006 (2016).
[29] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. P. Song, C.-L. Zou, S. M. Girvin, L.-M. Duan, and L. Sun, Quantum error correction and universal gate set operation on a binomial bosonic logical qubit, Nature Physics **15**, 503 (2019).
[30] J. M. Gertler, B. Baker, J. Li, S. Shirol, J. Koch, and C. Wang, Protecting a bosonic qubit with autonomous quantum error correction, Nature **590**, 243–248 (2021).
[31] S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann, G. J. Norris, C. K. Andersen, M. Müller, A. Blais, C. Eichler, and A. Wallraff, Realizing repeated quantum error correction in a distance-three surface code, Nature **605**, 669–674 (2022).
[32] Y. Zhao *et al.*, Realization of an error-correcting surface code with superconducting qubits, Phys. Rev. Lett. **129**, 030501 (2022).
[33] N. Sundaresan, T. J. Yoder, Y. Kim, M. Li, E. H. Chen, G. Harper, T. Thorbeck, A. W. Cross, A. D. Córcoles, and M. Takita, Demonstrating multi-round subsystem quantum error correction using matching and maximum likelihood decoders, Nature Communications **14**, 10.1038/s41467-023-38247-5 (2023).
[34] R. Acharya *et al.*, Suppressing quantum errors by scaling a surface code logical qubit, Nature **614**, 676–681 (2023).
[35] R. Porotti, V. Peano, and F. Marquardt, Gradient-ascent pulse engineering with feedback, PRX Quantum **4**, 030305 (2023).
[36] B. de Neeve, T.-L. Nguyen, T. Behrle, and J. P. Home, Error correction of a logical grid state qubit by dissipative pumping, Nature Physics **18**, 296 (2022).
[37] L. Rabiner, A tutorial on hidden markov models and selected applications in speech recognition, Proceedings of the IEEE **77**, 257 (1989).
[38] J. Barry, D. T. Barry, and S. Aaronson, Quantum partially observable markov decision processes, Phys. Rev. A **90**, 032311 (2014).
[39] V. V. Sivak, A. Eickbusch, H. Liu, B. Royer, I. Tsioutsios, and M. H. Devoret, Model-free quantum control with reinforcement learning, Phys. Rev. X **12**, 011059 (2022).
[40] V. Markov, V. Rastunkov, A. Deshmukh, D. Fry, and C. Stefanski, Implementation and learning of quantum hidden markov models (2023), arXiv:2212.03796 [quant-ph].
[41] R. Rico-Martínez, K. Krischer, I. G. Kevrekidis, M. C. Kube, and J. L. Hudson, Discrete- vs. continuous-time nonlinear signal processing of cu electrodissolution data, Chemical Engineering Communications **118**, 25–48 (1992).
[42] N. Gicquel, J. Anderson, and I. Kevrekidis, Noninvertibility and resonance in discrete-time neural networks for time-series processing, Physics Letters A **238**, 8 (1998).
[43] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, A novel connectionist system for unconstrained handwriting recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence **31**, 855 (2009).
[44] X. SHI, Z. Chen, H. Wang, D.-Y. Yeung, W.-k. Wong, and W.-c. WOO, Convolutional lstm network: A machine learning approach for precipitation nowcasting, in *Advances in Neural Information Processing Systems*, Vol. 28, edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (Curran Associates, Inc., 2015).
[45] S.-T. Tsai, E.-J. Kuo, and P. Tiwary, Learning molecular dynamics with simple language model built upon long short-term memory neural network, Nature Communications **11**, 10.1038/s41467-020-18959-8 (2020).
[46] P. K. Barkoutsos, G. Nannicini, A. Robert, I. Tavernelli, and S. Woerner, Improving variational quantum optimization using cvar, Quantum **4**, 256 (2020).
[47] A. Choquette, A. Di Paolo, P. K. Barkoutsos, D. Sénéchal, I. Tavernelli, and A. Blais, Quantum-optimal-control-inspired ansatz for variational quantum algorithms, Phys. Rev. Res. **3**, 023092 (2021).
[48] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).

Supplemental Material for "Boosting the Gottesman-Kitaev-Preskill quantum error correction with non-Markovian feedback"

Matteo Puviani,[1,*] Sangkha Borah,[1,2] Remmy Zen,[1] Jan Olle,[1] and Florian Marquardt[1,2,†]

[1] *Max Planck Institute for the Science of Light, 91058 Erlangen, Germany*

[2] *Department of Physics, Friedrich-Alexander Universität Erlangen-Nürnberg, 91058 Erlangen, Germany*

THE GKP CODE

In contrast to multi-qubit encoding, where the logical qubit information is stored into many physical qubits allowing to implement fault-tolerant operations and to correct for quantum decoherence errors, bosonic codes exploit the theoretically infinite Hilbert space of Fock states of a bosonic system. There are different bosonic codes which we can consider, including the Pegg-Barnett [1], cat (kitten) [2], Gottesman-Kitaev-Preskill (GKP) [3–7] and binomial [8, 9] code. In this work, we consider a single-mode GKP state encoding a single logical qubit.

In the superconducting circuit platform, these codes can be physically realized using an optical superconducting cavity, coupled to an ancilla transmon qubit and a readout resonator. The ancillary qubit is entangled with the cavity state and is used to control the bosonic code itself. In the case of measurement-based quantum error correction (QEC) protocols, projective measurements are performed on the qubit in order to extract weakly information from the cavity, without destroying the encoded logical state.

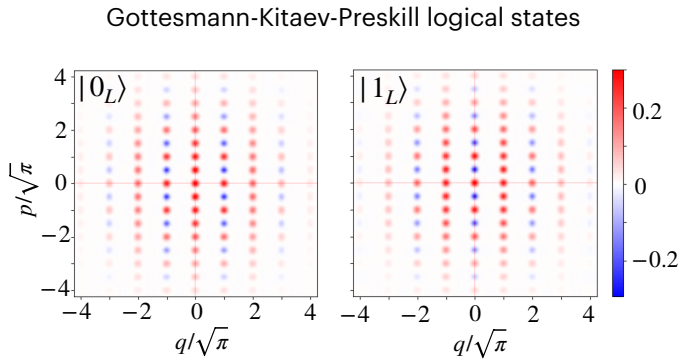

Figure S1. Wigner functions of the GKP physical code for the logical qubits 0_L and 1_L, respectively. The two logical grid states differ only by a displacement of $\sqrt{\pi}$ along q.

Pauli operators and stabilizers

Here we describe the theoretical framework to define and characterize the GKP codes, focusing on the case of a cavity with one single oscillator mode. We start defining the cavity displacement operator

$$\hat{D}(\zeta) = e^{\zeta \hat{a}^\dagger - \zeta^* \hat{a}}, \tag{1}$$

where $\hat{a}(\hat{a}^\dagger)$ is the bosonic annihilation (creation) operator.

There are two stabilizers, $\hat{S}_X = \hat{X}_L^2 = \hat{D}(2\alpha)$ and $\hat{S}_Z = \hat{Z}_L^2 = \hat{D}(2\beta)$, which define the entire GKP codespace spanned by the two logical states $|0_L\rangle$ and $|1_L\rangle$, i.e.: $\mathcal{C} = \langle |0_L\rangle, |1_L\rangle \rangle$. Every state in $\mathcal{C}$ can be written as a linear combination of this two logical states (which form a basis) and is a simultaneous eigenstate of both the stabilizers with eigenvalue $+1$, as the stabilizers also commute with each other. This physically means that every logical qubit is characterized by the periodicity 2α and 2β in the $\{q, p\}$ phase-space.

The logical Pauli operators are instead defined as $\hat{X}_L = \hat{D}(\alpha)$ and $\hat{Z}_L = \hat{D}(\beta)$, such that their eigenstates are the logical states $\pm X$ ($|\pm X_L\rangle$) and $\pm Z$ ($|\pm Z_L\rangle$), respectively. The operator $\hat{Z}_L$ is used to define the basis logical states: $|+Z_L\rangle = |0_L\rangle$ and $|-Z_L\rangle = |1_L\rangle$. Physically, the eigenstates of the logical Pauli operators with opposite eigenvalues

are transformed into each other with the application of a displacement $\hat{D}(\alpha)$ or $\hat{D}\beta$, respectively. For this reason, the Wigner function of a logical state in this encoding is characterized by dots forming a periodic lattice, which is why the GKP code is also referred to as the *grid code* [Fig. S1].

Logical qubit states

The values of the complex code parameters α and β, which enter the definition of the Pauli operators and stabilizers, are not unique, but they are constrained by a relation according to anti-/commutation relation of the logical operators:

$$\alpha^*\beta - \alpha\beta^* = i\pi \ , \tag{2}$$

or

$$\alpha^*\beta - \alpha\beta^* = 2i\pi \ , \tag{3}$$

respectively. We choose here the condition for which the logical operators anti-commute. We can now define the logical qubits encoding $|0_L\rangle = |+Z_L\rangle$ and $|1_L\rangle = |-Z_L\rangle$ as [6]

$$|0_L\rangle = \sum_j |2j\sqrt{\pi}\rangle_Q \ , \tag{4}$$

$$|1_L\rangle = \sum_j |(2j+1)\sqrt{\pi}\rangle_Q \ , \tag{5}$$

with the eigenstates of the position operator $\hat{Q} = i(\beta^*\hat{a} - \beta\hat{a}^\dagger)/\sqrt{\pi}$, or equivalently

$$|0_L\rangle = \mathcal{N}\sum_{k,l} e^{-i\pi kl}|2k\alpha + l\beta\rangle \ , \tag{6}$$

$$|1_L\rangle = \mathcal{N}\sum_{k,l} e^{-i\pi(kl+l/2)}|(2k+1)\alpha + l\beta\rangle \ , \tag{7}$$

using a linear combination of the coherent states of the cavity oscillator, $\mathcal{N}$ being a normalization coefficient.
Since the anti-commutation relation is the only constrain over the two parameters of the code, it is possible to define GKP codes that have different symmetries in the phase space according to the values of α and β, such as the square, rectangular and hexagonal code. For the square code, characterized by a square lattice periodicity of the Wigner function, we have $\alpha = \sqrt{\pi/2}$ and $\beta = i\alpha$. Therefore α is the unitary lattice displacement along the $\hat{q}$ direction, while β is the corresponding one along $\hat{p}$. The rectangular code is generalized as $\alpha = l\sqrt{\pi/2}$, $\beta = i\sqrt{\pi/2}/l$, where l is a free scaling parameter, while the hexagonal code has $\alpha = \sqrt{\pi/\sqrt{3}}$ and $\beta = e^{2i\pi/3}\sqrt{\pi/\sqrt{3}}$.
In this work, we have focused entirely on the square GKP code.

Finite-energy GKP code

The ideal GKP logical qubit encoding as described in the previous section, however, is not physical since it requires an infinite combination of coherent states and therefore an infinite amount of photons and energy to be implemented. Thus, a real finite-energy GKP code can be described by truncating the number of coherent states up to a maximum value M: physically this corresponds to considering a finite-size grid in phase space instead of an infinite periodic lattice, as represented in Fig. S1. Mathematically, this can be described by a convolution of a gaussian envelope with the ideal logical state in Eqs. (6) and (7):

$$|\tilde{\mu}_L\rangle \propto \ e^{-\Delta^2\hat{a}^\dagger\hat{a}}|\mu_L\rangle \ , \tag{8}$$

with $\mu = 0, 1$ for the corresponding logical qubit states. The ideal limit is obtained for $\Delta \to 0$.
As we can notice, the gaussian envelope introduces a cut-off both in real and phase space, as well as in the Hilbert space of Fock states, providing an upper limit to the actual number of photons involved. Therefore, it is possible to

accurately simulate the dynamics of a finite-energy GKP code imposing a large but finite number of Fock states to be considered. In our work we adopted an upper limit of the Fock space of $N_{max} = 100$.
However, the main consequence of this is that the GKP code is no longer invariant under discrete translations, but only approximately and within a few translations: therefore, one additional challenge of QEC is to keep the grid centered in phase space around the $(q, p) = (0, 0)$ point.
It is also possible to introduce *finite-energy* or *approximate* Pauli operators and stabilizers, in order to normalize them to the GKP finite energy [6]: however, we did not use such operators, but rather the standard definition provided in the previous section.

GKP QUANTUM ERROR CORRECTION

Dynamics, noise and errors of the GKP code

In order to be able to counteract the effect of photons' dissipation in the quantum cavity by acting and applying controls on it, the cavity itself is coupled to a two-level system that acts as control qubit. However, the qubit that is adopted as ancilla is also affected by quantum noise, namely relaxation, dephasing and leakage events. The former two are responsible for the decoherence of the qubit, while the latter is the most harmful error and is due to the excitation to higher levels in an actual multi-level system. A protocol must be developed to correct for these additional qubit errors that would affect the cavity after entanglement, as these errors (especially the leakage error) can be the most detrimental to GKP coding. A readout resonator completes the hardware setup that enables the measurement of the ancilla qubit: a strong measurement of the qubit in this configuration is reflected into a weak measurement of the cavity so that the logical quantum information contained therein is not destroyed. In our treatment we will not consider the qubit leakage and the associated errors, as well as SPAM errors: this would have equal impact on all the measurement-based protocols and would not significantly affect the results from a quantitative point of view. In fact, in an experimental setup such as the one in Ref. [10], the probability of a leakage error is $p_{leak.} < 0.01$, and it can be reset in the same way as the excited qubit state.

We consider the ground state of the ancillary qubit $|g\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, so that the Pauli operators $\hat{\sigma}_- = (\hat{\sigma}_x - i\hat{\sigma}_y)/2$ and $\hat{\sigma}_+ = (\hat{\sigma}_x + i\hat{\sigma}_y)/2$ act as creation and annihilation ladder operators, respectively. We define the dissipator of a damping operator $\hat{A}$ applied to the density matrix $\hat{\rho}$ as

$$\mathcal{D}[\hat{A}]\hat{\rho} = \hat{A}\hat{\rho}\hat{A}^\dagger - \left(\hat{A}^\dagger\hat{A}\hat{\rho} + \hat{\rho}\hat{A}^\dagger\hat{A}\right)/2 \ , \tag{9}$$

in order to be able to write the time evolution of the density matrix with the Lindblad master equation for the cavity-ancilla system in the following form:

$$\dot{\hat{\rho}} = -i\left[\hat{H}, \hat{\rho}\right] + \mathcal{D}[\hat{a}/T_s]\hat{\rho} + \mathcal{D}[\hat{\sigma}_+/T_1]\hat{\rho} + \mathcal{D}[\hat{\sigma}_z/2T_\phi]\hat{\rho} \ , \tag{10}$$

where T_s is the cavity relaxation time due to dissipation, T_1 is the qubit time for relaxation, while $T_\phi = \dfrac{1}{1/T_1 + 1/T_2}$, T_2 being the qubit dephasing time. The Hamiltonian of the system is $\hat{H}$, and it includes a coupling and a Kerr term:

$$\hat{H} = \frac{1}{2}\chi\hat{a}^\dagger\hat{a}\hat{\sigma}_z + \frac{1}{2}K\left(\hat{a}^\dagger\hat{a}\right)^2 \ , \tag{11}$$

where χ is the first order dispersive shift term and K is the Kerr constant. Since the dynamics generated by the Hamiltonian is used to model the pulses to give rise to the gates of the QEC circuit [5], in our work we neglected its contribution to the dynamics in Eq.(9) because we adopted perfect instantaneous parameterized gates followed by idling dissipative dynamics, without modelling the electromagnetic pulses.

For the purpose of quantum error correction, we can add projective quantum measurements on the ancillary qubit, which act as weak (syndrome) measurement on the cavity state providing information on whether an error has occurred [10]. Since the quantum measurements have an intrinsically statistical nature, we stochastically sample the measurement's outcome with a Monte-Carlo algorithm. The density matrix of the full quantum system (cavity and ancilla) after a measurement is

$$\hat{\rho}(t_+) = \sum_{m=\pm 1} r_m \ \hat{P}_m \ \hat{\rho}(t_-) \ \hat{P}_m^\dagger \ , \tag{12}$$

with the sampling algorithm

$$r_m = \begin{cases} 1, & \text{if } p_m > s \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

where $s \in [0,1]$ is sampled from a uniform distribution, $p_m = \mathrm{Tr}\{\hat{P}_m \hat{\rho} \hat{P}_m^\dagger\}$ is the probability of measurement $m \in \{g, e\}$, and $\hat{P}_m = \hat{\mathbb{1}}_c \otimes |m\rangle_q\langle m|$ is the projection operator onto the $|m\rangle$ qubit state. Here we introduced the subscripts q and c to indicate the qubit and the cavity, respectively.
The Lindblad master equation (9) including the discrete stochastic measurements (12) becomes overall probabilistic. Moreover, we notice that both of them define the evolution of the density matrix immediately after the measurement $\hat{\rho}(t_+)$ in a Markovian way, as they only depend on the density matrix $\hat{\rho}(t_-)$ at the step immediately before.

Physical gates

To implement the small-BIG-small QEC protocol as represented in Fig. 2 in the main text following Ref. [10], we adopted a set of universal physical gates acting on the cavity [the cavity virtual rotation $\hat{VR}_c(\theta_{VR})$ and the cavity displacement $\hat{D}_c(\alpha)$], on the ancillary qubit [the the qubit rotation $\hat{R}_q(\phi,\theta)$], and on both [the echoed conditional displacement $\hat{ECD}_{qc}(\beta)$]. For these gates we adopt the definition given in Ref. [11], with the echoed conditional displacement defined as

$$\hat{ECD}_{qc}(\beta) = \hat{D}(\beta/2) \otimes \hat{\sigma}_- + \hat{D}(-\beta/2) \otimes \hat{\sigma}_+ \ , \tag{14}$$

where $\beta \in \mathbb{C}$. The operators $\hat{\sigma}_\pm$ are acting on the ancillary qubit, while the displacement $\hat{D}$ acting on the cavity is defined as

$$\hat{D}_c(\beta) = \exp\{\beta \hat{a}^\dagger - \beta^* \hat{a}\} \ . \tag{15}$$

The ancillary qubit rotation is given by

$$\hat{R}_q(\phi,\theta) = \exp\Big\{ -i\frac{\theta}{2}(\hat{\sigma}_x \cos\phi + \hat{\sigma}_y \sin\phi)\Big\}. \tag{16}$$

The cavity virtual rotation that allows to rotate the q and p axis in phase space is given by

$$\hat{VR}_c(\theta_{VR}) = \exp\{i\theta_{VR}\hat{a}^\dagger \hat{a}\} \ . \tag{17}$$

This is used in order to keep the same fixed parameters in the QEC circuit without swapping between real and complex parameters' values.

small-BIG-small QEC protocol

In order to perform QEC we adopted the so-called small-BIG-small (sBs) protocol with measurement-based feedback as described in Ref. [10]. In order to simulate the realistic dynamics we modelled the circuit with perfect instantaneous gates followed by idling dissipative dynamics solving the Lindblad master equation (9), neglecting the Hamiltonian's dynamics. As a consequence, the feedback coming from the measurement, which is used in the experiments to reset the ancilla and to determine the correct angle for the cavity virtual rotation gate, is not neede here in our modeling. For the time dynamics during the QEC circuit, we considered a realistic model based on experiment in Ref. [10]: namely, we assumed a total duration $\tau_{\text{cycle}} = 10$ μs, so that the half-cycle shown in Fig. S2 will have a duration of 5 μs. The absolute and relative times of each step are shown in Table S1.

The initial condition of the QEC protocol comprises a cavity with a prepared GKP code, and an ancillary transmon qubit set in its ground state, $|g\rangle$. The sBs standard protocol consists of 4 layers of qubit ancilla rotation and echoed conditional displacement, which entangles the cavity and the qubit states (Fig. S2). In the fourth layer the entangling gate is replaced by a simple displacement gate acting on the cavity. Then, a projective measurement (POVM) takes place on the ancillary qubit, which projects its state onto either the ground ($|g\rangle$) or the excited ($|e\rangle$) state. After that, the qubit is reset to its ground state and a virtual rotation is applied on the cavity, in order to switch between the position and momentum space variables. We notice here that we did not include the Hamiltonian dynamics in the time evolution, and the reset is performed numerically and not with gates or pulses. As a consequence, the feedback that is usually applied in experiments for the qubit reset and to compensate the Hamiltonian dynamics in the cavity virtual rotation does not apply to our numerical simulations.

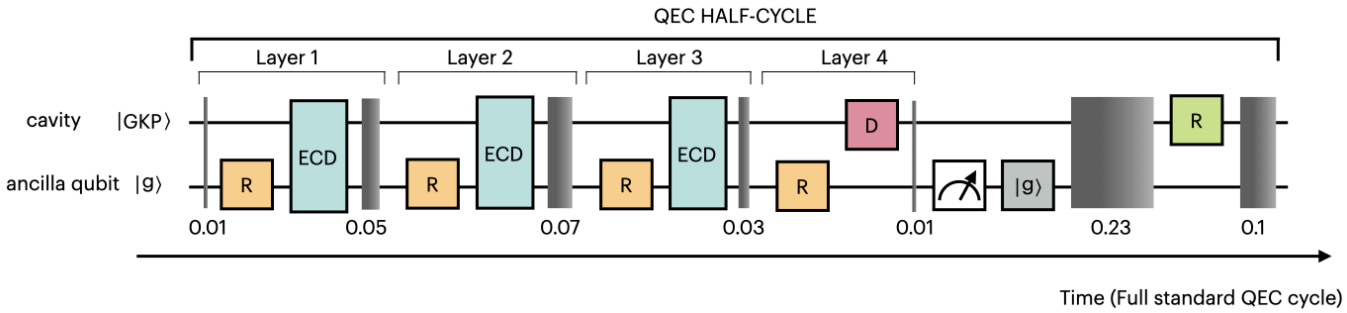

Figure S2. Scheme showing the half-cycle measurement-based sBs QEC protocol as adopted in the main paper and shown in Fig. 2, with the dynamics represented by the shaded grey areas, with the time indicated in units of the full QEC cycle.

Dynamics	Duration time (μs)	Duration time (τ_{cycle})
Entering cycle	0.1 μs	0.01
Layer 1	0.5 μs	0.05
Layer 2	0.7 μs	0.07
Layer 3	0.3 μs	0.03
Layer 4	0.1 μs	0.01
Measurement and reset	2.3 μs	0.23
Virtual rotation and idling	1 μs	0.1

Table S1. Time duration of the dissipative dynamics in half QEC cycle shown in Fig. S2.

FEEDBACK-GRAPE OPTIMIZATION WITH NON-MARKOVIAN FEEDBACK

Feedback-GRAPE

Feedback-GRAPE is an algorithm recently developed by Porotti et. al [12] for Gradient Ascent Pulse Engineering (GRAPE) with feedback. It works for a quantum system with feedback that is subject to repeated strong projective measurements at times $t_1, t_2, ..., t_N$ with the corresponding measurement results m_j, $j \in [1, N]$. The task of Feedback-GRAPE is to find the accurate feedback control functions $F_\theta^j(\mathbf{m}_j)$ for all possible measurement results $\mathbf{m}_j = \{m_1, m_2, ..., m_j\}$, as a feedback strategy parameterized with the trainable parameters θ that are optimized by gradient ascent. It is assumed that F_θ^j is differentiable with respect to θ, which can be provided by a neural network or alternatively a lookup table. In order to optimize them, the task would be to to maximize a cumulative total reward R (this is called return in the nomenclature of reinforcement learning). In a state preparation task, for example, this could be the final fidelity with respect to a target state $\hat{\sigma}$, $\mathcal{R}(\mathbf{m}) = \left(\mathrm{tr}\sqrt{\sqrt{\hat{\sigma}}\hat{\rho}(T|\mathbf{m})\sqrt{\hat{\sigma}}}\right)^2$, where $\mathbf{m}$ represents a particular measurement sequence, which would finally be averaged over all possible sequences of measurement results to yield $\bar{\mathcal{R}} = \langle \mathcal{R}(\mathbf{m}) \rangle_{\mathbf{m}}$. For strong projective qubit measurements, as used in the present work, however, the probabilities P for the different measurement results themselves depend on all controls F_θ^j that were applied in the previous time intervals, which must be taken into account when analyzing the gradients with respect to θ. Thus

$$\langle \mathcal{R}(\mathbf{m}) \rangle_{\mathbf{m}} = \sum_{\mathbf{m}} P(\mathbf{m}) \mathcal{R}(\mathbf{m}) \,, \tag{18}$$

where $P(\mathbf{m})$ is the probability for measurement outcomes $\mathbf{m}$. Thus, when evaluating $\partial \langle R(\mathbf{m}) \rangle_{\mathbf{m}} / \partial\theta$, we will get two contributions,

$$\frac{\partial [\mathcal{R}(\mathbf{m}) P(\mathbf{m})]}{\partial \theta} = P(\mathbf{m}) \frac{\partial \mathcal{R}(\mathbf{m})}{\partial \theta} + \mathcal{R}(\mathbf{m}) \frac{\partial P(m)}{\partial \theta} \,. \tag{19}$$

To enable stochastic sampling of the second term, we rewrite it using

$$\frac{\partial P(\mathbf{m})}{\partial \theta} = P(\mathbf{m}) \frac{\partial \ln P(\mathbf{m})}{\partial \theta} \,, \tag{20}$$

which leads to

$$\frac{\partial \langle \mathcal{R}(\mathbf{m}) \rangle_{\mathbf{m}}}{\partial \theta} = \left\langle \frac{\partial \mathcal{R}(\mathbf{m})}{\partial \theta} \right\rangle_{\mathbf{m}} + \left\langle \mathcal{R}(\mathbf{m}) \frac{\partial \ln P_\theta(\mathbf{m})}{\partial \theta} \right\rangle_{\mathbf{m}}, \tag{21}$$

where the $P_\theta(\mathbf{m})$, with the explicit parameter dependence θ represents the probability of getting the full sequence of measurement outcomes $\mathbf{m} = (m_1, m_2, \ldots)$. The log-likelihood term can be conveniently computed as the cumulative sum over each trajectory. The gradients can then be taken for a batch of trajectories.

Numerical implementation

The feedback-GRAPE algorithm relies on a gradient-based optimization: therefore we used TensorFlow to implement automatic differentiation through the whole multi-step dynamics, allowing for an efficient setup which can run in parallel on a GPU to speed up the calculations. Moreover, in order to numerically simulate the dissipative dynamics of the system, we developed a custom TensorFlow time-evolution simulator implementing the Runge-Kutta-4 algorithm, which can be run in parallel also on a GPU.

Even though the implementation of a look-up table containing the optimal parameters for a given measurement outcomes sequence would be optimal and easier to implement, this would however be limiting as the number of parameters of the look-up table increase exponentially with the number of QEC cycles, and therefore with time, as shown in [12]. Therefore, when evaluating trajectories with many cycles the effective memory of the look-up table and its performance would be limited. For this reason, we adopted a recurrent neural network (RNN, whose hyperparameters are shown in Table S2) which can be used to suggest parameters even at time scales much longer than the ones used for training. In fact, the RNN responds in a non-Markovian way to the measurement outcomes, building an internal representation (or *belief state*) of the state of the quantum system. However, the update of the RNN's hidden state only depends on the current state and the current measurement outcome, making this update itself Markovian. For comparison we also used a feed-forward neural network (FNN), whose hyperparameters are provided in Table S3. In this latter case, however, the response is not subject to memory and therefore Markovian, as it will be explained in the results' section.

Parameter	Value
RNN cell	GRU
Neurons	[10, 256, 256, 15]
Input shape	[batch_size,1,1]
RNN cell activation	tanh
Dense layers activation	tanh
Initializer	Random Uniform
Initial bias	0.01

Table S2. Hyperparameters of the RNN used, with batch_size = 6. The core of the network is the memory provided by the *Gated Recurrent Unit* (GRU) in the first layer of the neural network.

Parameter	Value
Neurons	[256, 256, 15]
Input shape	[batch_size,1,1]
Dense layers activation	tanh
Initializer	Random Uniform
Initial bias	0.01

Table S3. Hyperparameters of the FNN used, with batch_size = 8.

The training of the RNN to suggest the gates parameters from scratch was not successful: therefore we provided as initial values the optimal parameters θ_i used in Ref. [10] (Table S4), allowing the neural network to provide at each step corrections $\delta\theta_i$, applying to the gates the final values $\theta_i + \delta\theta_i$. The gates' parameters which are suggested are in total 15 for each cycle: the rotation angles ϕ and θ of the qubit rotations of each layer, the real and imaginary part of the ECD, β' and β'', respectively, in layers 1-3, and the virtual rotation angle θ_{VR}. The range allowed for the corrections $\delta\theta_i$ is $[-2, 2]$, except for the virtual rotation for which $\delta\theta_{VR} \in [-1, 1]$. We decided not to change the parameter of the displacement gate in layer 4 of the QEC cycle, as its value α contains crucial information about the geometry and the symmetry of the grid code, which could be affected if modifications were allowed.

Gate	Parameter	L1	L2	L3	L4
$\hat{R}_q$	ϕ	$\pi/2$	0	0	$\pi/2$
$\hat{R}_q$	θ	$\pi/2$	$-\pi/2$	$\pi/2$	$-\pi/2$
$E\hat{C}D_{qc}$	β	$0.2i$	$\sqrt{2\pi}$	$0.2i$	-

Table S4. Parameters' values optimized for a single half-cycle derived from [10]: qubit rotation ($\hat{R}_q(\phi,\theta)$), echoed conditional displacement ($E\hat{C}D_{qc}(\beta)$), cavity virtual rotation ($\hat{VR}_c(\theta_{VR})$).

RNN-agent training

For the agent to be included in the QEC scheme as in Fig. 2 of the main text, we adopted a recurrent neural-network (RNN) with 10 *Gated Recurrent Units* (GRU) cells [13], and 2 hidden layers (256, 256), with an activation function $f = tanh()$ in each layer (Table S2). The training was performed over 1000 epochs, with different batch-sizes (6 and 8 for the sBs optimization) for the Monte-Carlo sampling of the measurements giving rise to different trajectories, and collecting the gradient from each of them and applying it to the same network. The learning rate was set to $lr = 10^{-4}$, the biases were all randomly initialized within the range $[-0.1, 0.1]$.
For every optimization, multiple agents were trained in parallel in different parallel environments: the best performing agent was then evaluated and post-selected according to the best performance, namely the longer lifetime achieved.
For the reward function to maximize we used the fidelity of the final state density matrix $\hat{\rho} = \hat{\rho}(T)$ with respect to the target $\hat{\sigma}$, which is the initial state ($\hat{\sigma} = \hat{\rho}(0)$):

$$\mathcal{F}[\hat{\rho},\hat{\sigma}] = \left[\mathrm{Tr}\left(\sqrt{\sqrt{\hat{\sigma}}\hat{\rho}\sqrt{\hat{\sigma}}}\right)\right]^2 , \tag{22}$$

which for pure target (initial) states is equivalent to

$$\mathcal{F}[\hat{\rho},\hat{\sigma}] = \left[\mathrm{Tr}\left(\sqrt{\hat{\sigma}\hat{\rho}}\right)\right]^2 . \tag{23}$$

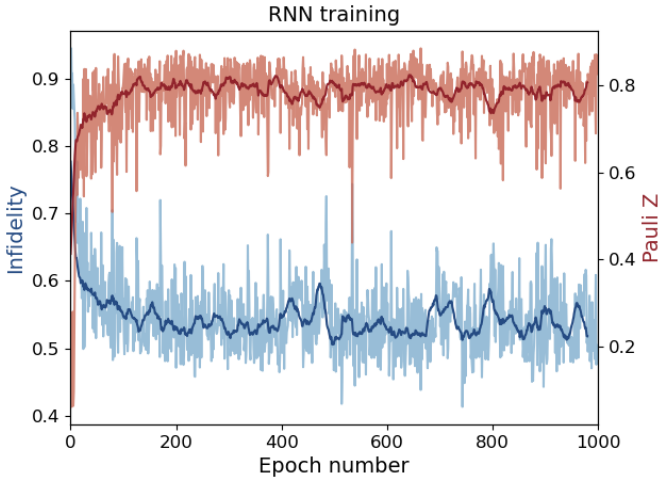

Figure S3. Training of the RNN for the NMF QEC strategy. Plot of the final infidelity ($1-\mathcal{F}$, $\mathcal{F}$ being the fidelity between the final and the initial state) and the Pauli Z expectation value $\langle\hat{Z}\rangle$ as a function of the epoch number during the training with Feedabck-GRAPE. The final fidelity $\mathcal{F}$ has been used as a reward for the agent (see text).

An example of the training of a RNN for the optimization of 15 parameters of the sBs QEC protocol is shown in Fig. S3 for the initial state $|+Z_L\rangle$: the training has been performed over 1000 epochs for trajectories of 10 full sBs QEC cycles, including 2 measurements each. In blue we show the behavior of the infidelity (defined as $1-\mathcal{F}$, $\mathcal{F}$ being the fidelity) with respect to the initial state as a function of the training epoch. The dark straight line represents the running average over 40 data points, while the light blue line corresponds to the raw data. The red lines, instead, represent the running average and the raw data of the logical Pauli Z expectation value, respectively. Since the initial

state is the logical $|Z_L\rangle$, the maximization of the fidelity via QEC after 10 cycles (or, equivalently, the minimization of the infidelity) corresponds to the maximization of the expectation value of the logical Pauli Z, which is the actual goal of QEC to preserve the logical information contained in it. However, the opposite is not true: maximizing the logical corresponding Pauli Z, in fact, does not lead to maximal fidelity to the initial state, and is not generalizable to other logical states, which is possible using the fidelity, as shown later. We can also notice that the minimum fidelity is already reached after around 300 epochs, so that in principle the results could have been comparable with a shorter training. However, we decided to opt for a longer training in order for the agent to learn a proper strategy which also minimizes the probability of a drop of logical information, as shown in the Pauli Z expectation value around epoch 1000, in a sort of over-fitting to the problem.

ADDITIONAL NUMERICAL RESULTS

Autonomous quantum error correction

Recently it has been shown in an experiment that it is possible to perform autonomous quantum error correction on the GKP code with the sBs protocol, resulting in a longer qubit's lifetime with respect to the vanilla measurement-based sBs protocol [14]. We considered the autonomous QEC scheme as represented in Fig. S4, realistically reducing the time dynamics for the reset from 0.23 to 0.08 expressed in units of the full measurement-based QEC cycle, τ_{cycle}. As a consequence, the duration of half-cycle for the autonomous QEC will be 0.35 instead of 0.50 as for the measurement-based QEC half-cycle. This value has been used as a realistic estimate resulting from the absence of the measurement process. As a consequence, the ancillary qubit reset happens in the autonomous QEC $\approx 43\%$ more times than for the measurement-based sBs protocol.

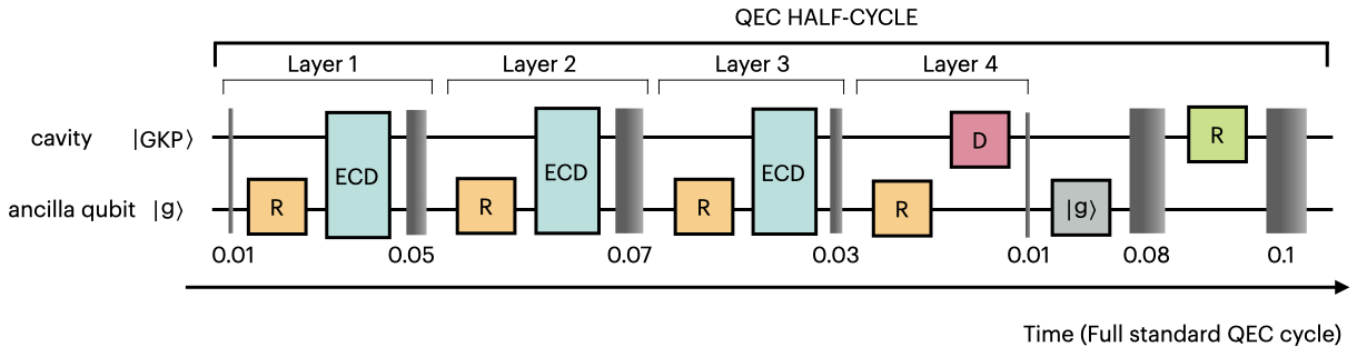

Figure S4. Autonomous QEC scheme. Scheme of the half-cycle sBs autonomous QEC, without measurements. The ancillary qubit is still reset after layer 4. The grey vertical strips represent the dynamics duration referred to the length of a full standard QEC as in Fig. 2: the summation returns 0.35 instead of 0.5 because the dynamics for the qubit reset has been reduced from 0.23 to 0.08 due to the absence of the measurement and readout. This means that a full cycle of the autonomous QEC lasts 0.7 of the measurement-feedback one including measurements.

Average channel fidelity: evaluation and lifetime

The average trajectories of the Pauli operators' expectation values, which are shown e.g. in Fig. 3(a) in the main text, are calculated from 6×85 (for NMF approach) and 8×64 (for the standard protocol) raw trajectories, where the first number is the batch-size of the sampling and the second number is the repetition number for each of them. In Fig. S5 we show these raw trajectories (from which Fig.3(a) is derived) which already contain the average over the batch-size. The standard deviation shown in the plots is calculated on these trajectories as well.

It is also interesting and insightful to have a look at the behavior of the state fidelity during the QEC process. In fact, our neural-network agents are given as a reward, i.e. function to maximize, the final fidelity $\mathcal{F}$ of the density matrix of the state with respect to the initial state $|+Z_L\rangle$, as defined in Eq. (23). However, in order to evaluate the QEC performance, the different agents and their strategies are evaluated on the logical qubits' lifetime, namely calculating the time-evolution and the decay rate of the logical Pauli operators. In Fig. S6 we show the fidelity as a

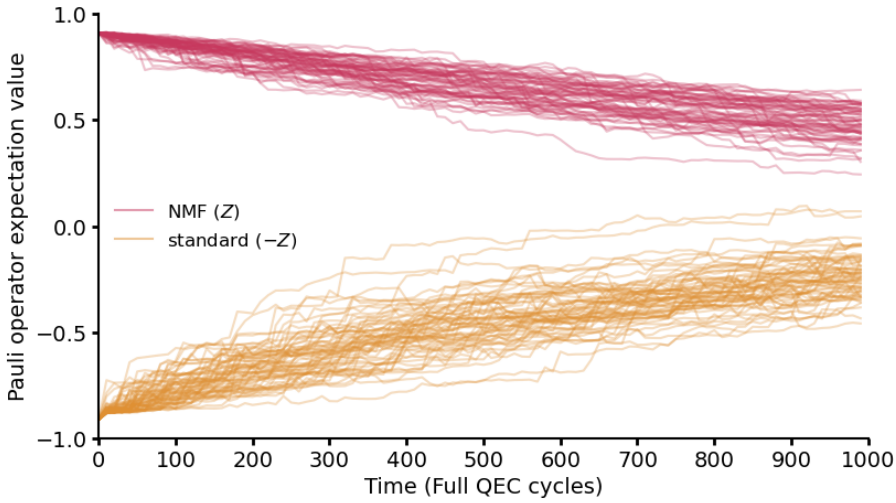

Figure S5. Trajectories' sampling. Trajectories (each of them averaged over a batch of sampled trajectories) of the logical Pauli operator expectation value as a function of time, when applying the standard and the NMF protocol, respectively. For the standard strategy the value $-\langle\hat{Z}\rangle$ is shown for better readability.

function of the QEC cycle number for different logical states for the standard and the best NMF strategies. Similarly to the case of the logical lifetime, even the fidelity remains higher over time with the time-dependent approach with memory compared with the standard sBs protocol.

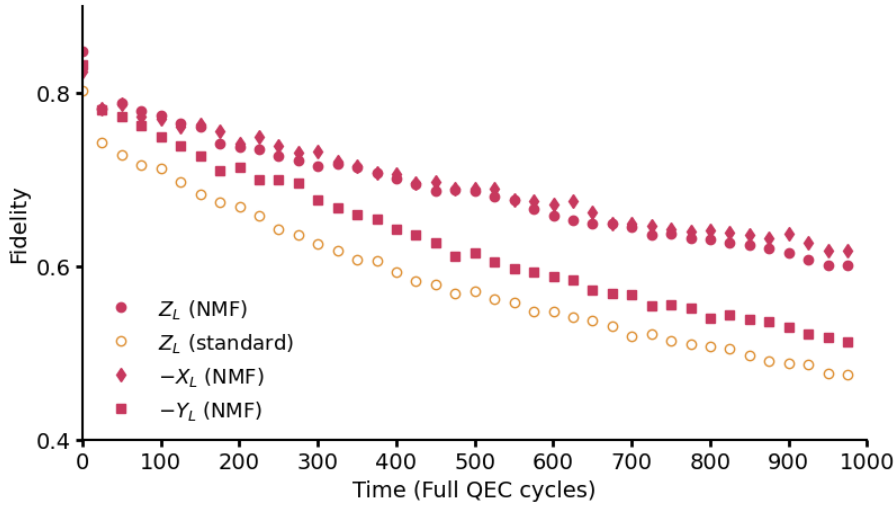

Figure S6. Time evolution of the fidelity of the GKP code with respect to the initial state when performing QEC according to the sBs protocol without (standard) and with non-Markovian feedback. The values displayed refer to different initial logical states, namely Z_L, $-X_L$ and $-Y_L$.

The average channel fidelity shown in Fig. 3(b) in the main text, as well as the ones in Fig. S7, are computed as the average lifetime $T = \frac{3}{\frac{1}{T_{-X}} + \frac{1}{T_{-Y}} + \frac{1}{T_Z}}$, since $T_{-X} = T_X$ and $T_{-Y} = T_Y$ hold. In this way we also verify that $T_X = T_Z$ within the variation due to the sampling of a limited number of trajectories (500+).

RNN agents and post-selection

The results shown for the NMF approach have been obtained by training multiple (20) RNN agents, evaluating all of them and post-selecting the one which was offering the highest performance in terms of logical lifetime's extension. The results shown in the main text were obtained with the best agent among the ones trained using the dynamics shown in Fig. 2 and high noise level, with an initial physical GKP state $|Z_L\rangle$ with $\Delta = 0.2$.
We will discuss about the generalizability of these agents in the following sections. However, as it was evident even from the main results, the agent trained with given dynamics and noise still outperforms the standard strategy just as well even for different circuit duration and noise level. Despite this, we also evaluated in the same conditions another agent based on the same RNN model, that we called NMF_2, which was post-selected after training on a simplified dynamics (as depicted in Fig. S8) and high noise level. Its performance is shown in Fig. S7 in comparison with all other strategies shown also in Fig. 3(b) in the main text. As it is visible, it outperforms by a factor ≈ 1.7 even the other RNN with NMF in the low noise level. However, this gain decreases for medium noise and at high noise it even fails to perform even just as well the standard sBs protocol. The reason of this reduced generalizability is found in the different probability distribution of the ancillary measurements' outcomes (Fig. S10), as discussed in the next sections. For this reason, and because of its lower adaptability to a real experimental setup, we decided to discard this strategy from the main results.

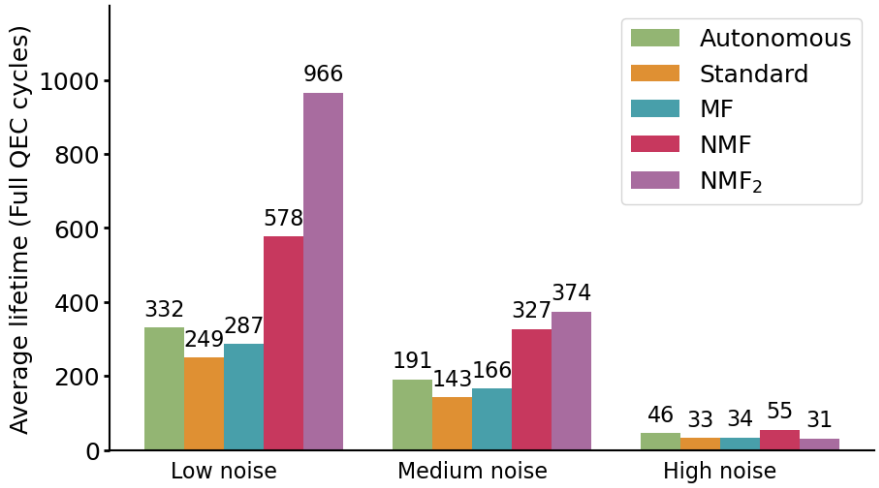

Figure S7. Average channel fidelity of different strategies. Comparison of the average lifetime for different strategies and noise levels as defined in Table S5.

Generalization to different noise levels

In our work we considered three different noise levels, reflecting real values of recent experiments: namely, in our numerical results we simulated the dissipative dynamics of the cavity, given by the cavity relaxation time (T_s), and the qubit relaxation (T_1) and dephasing (T_2). For the high noise level we used the values in Ref. [5], while for the low level we referred to the more recent Ref. [10]. The medium noise level was obtained by simply doubling the times of the high level. These values are represented in Table S5, with the times expressed in units of duration of a full sBs QEC cycle, which similarly to Ref. [10] we set to $\tau_{\text{cycle}} = 10\ \mu$s.

Noise level	T_s/τ_{cycle}	T_1/τ_{cycle}	T_2/τ_{cycle}
Low [10]	61	28	23.8
Medium	49	10	12
High [5]	24.5	5	6

Table S5. Noise levels. Table of the components' lifetimes used for the dynamics, in units of a full standard sBs cycle's duration, provided in Fig. 2, which is $\tau_{\text{cycle}} = 10\ \mu$s. T_s is the cavity relaxation time, T_1 and T_2 are the ancillary qubit's relaxation and dephasing times, respectively.

In particular, as we can notice in Fig. 3(b) of the main text and in Fig. S7, the logical qubit's lifetime decreases as the noise level increases, but the ratio is not linear: in fact, by doubling the components' lifetime from high to medium, the GKP code's lifetime increases by a factor higher than 2. In particular, this value depends on the strategy adopted: by decreasing the noise, the standard strategy increases by a factor ≈ 4, while the NMF approach provides an improvement of almost a factor ≈ 6.

Generalization to different dynamics

One strength of our NMF approach based on recurrent neural network is its versatility and generalizability to different noise and dynamics, which proves the presence of an underlying common feature for the quantum error correction of the GKP code. In fact, the same neural network agent trained for a given dynamics model and noise level, can perform just as well for a different combination of dynamics and noise, even if subtle remarks have to be raised.

As an example, we considered a simplified dynamics scheme considering instantaneous measurement and qubit reset, and equal time duration of each layer of the sBs protocol, as shown in Fig. S8. In particular, we compared the standard protocol, as well as two RNN with non-Markovian feedback, namely NMF and NMF_2 (this latter described in the previous section): the former was trained on the original dynamics of Fig. S2, while the latter on the simple dynamics of Fig. S8 (and both on high noise level). The lifetime of the Z_L logical state is shown in Fig. S9 for the different strategies and for three different noise levels (as described in Table S5). As we can notice, both the NMF strategies outperform the standard protocol, confirming that the memory is a widespread advantage for improving QEC, independent of the circuit dynamics. However, we also do observe that the agent trained with the same circuit dynamics outperforms the standard strategy by almost a factor ≈ 10. The reason can be found in the specific characteristics of this circuit, where the measurement ratio is higher (because the overall cycle duration is lower) and no dissipation is present during measurement and reset. This leads to a better error correctability, which makes not useful to have a high probability to measure the ancillary qubit in its ground state g, unlike in the realistic case. As a consequence, the agent trained on this simplified dynamics, NMF_2, optimized the parameters of the gates in disregard of the measurement outcome, as it is visible in the section on the strategy analysis. In fact, as the noise level increases, the advantage of such an approach becomes less evident: with high noise, the NMF_2 outperforms the standard strategy by a factor ≈ 4. This behavior is similar for the evaluation on a realistic dynamics (see Fig. S7), even though in that case the NMF_2 agent is failing even to perform as good as the standard strategy in the high noise conditions.

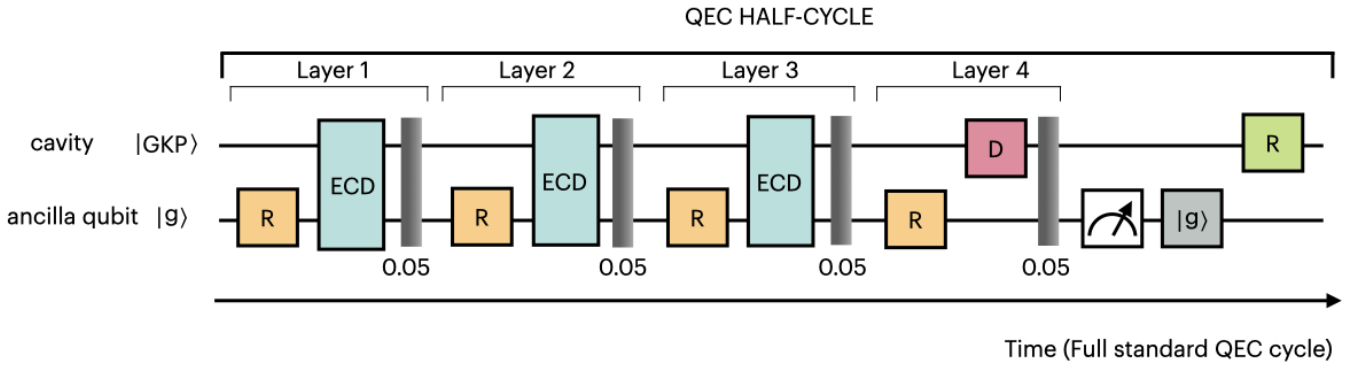

Figure S8. Simplified dynamics QEC scheme. Scheme representing the sBs QEC protocol similarly to Fig. 2, with a simpler dynamics scheme, represented by the grey vertical strips. In this case, we adopted a constant duration of the dynamics of each layer, and instantaneous measurement and reset (i.e.: no time evolution for the measurement and the ancillary qubit reset).

Strategy analysis

We analyze here more in detail the strategies that the different RNNs find when trained in different conditions. In particular, we observe that all the recurrent neural networks show a characteristic time dependence of the output

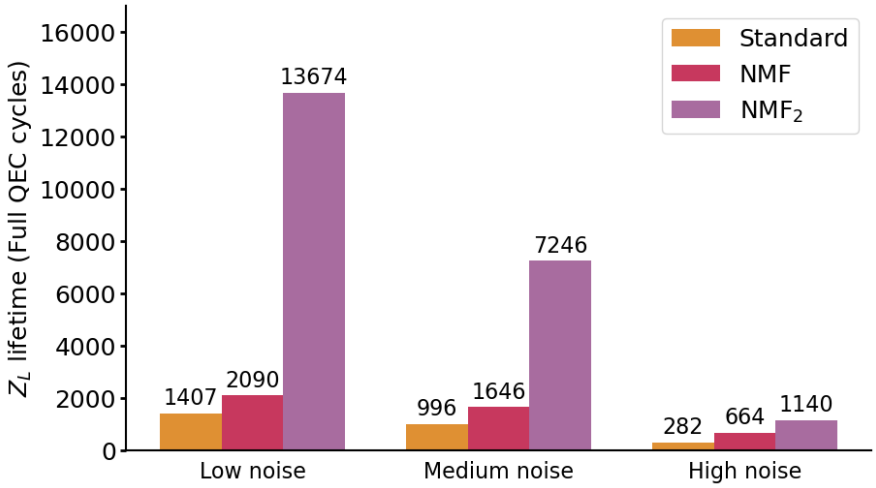

Figure S9. Comparison of the lifetimes of the Z_L state for different sBs protocols with the dynamics simulation in Fig. S8. The standard sBs protocol uses the constant optimal gates' parameters, while NMF and NMF$_2$ use the non-Markovian-feedback approach, and they were trained on different dynamics (see text).

parameters, which behaves as an exponential saturation to a certain value. This is true for both sequences of g as well as e measurements' outcomes, but with different saturation levels. Due to this behavior, we can fit the parameters' values after a g/e qubit's measurement's outcome as

$$\pi_{g/e}(t-\bar{t}) = \pi_{e/g(0)}(t=\bar{t})\ e^{-\gamma_{g/e}(t-\bar{t})} + \pi^{\infty}_{g/e}\left(1-e^{-\gamma_{g/e}(t-\bar{t})}\right)\ , \tag{24}$$

where $\bar{t}$ is the first time-step at which the sequence of identical g/e measurements started, t is the current time-step, $\pi^{\infty}_{g/e}$ is the saturation level for a theoretical infinite number of consecutive g/e measurements' outcomes, $\gamma_{g/e}$ is the parameter decay rate for consecutive g/e measurements. The initial value of one parameter for the very first step is $\pi_0(0)$. Then, if we have consecutive g measurements' outcomes, the neural network adjusts the outputs to $\pi_g(t) = \pi_0(0)\ e^{-\gamma_g t} + \pi^{\infty}_g\,(1-e^{-\gamma_g t})$. Ideally, if we measure only g on the ancillary qubit, we would end up with the values $\pi_g(t\to\infty) = \pi^{\infty}_g$. If we get a sequence of 5 times g we will get the parameter $\pi_g(5) = \pi(0)\ e^{-5\gamma_g} + \pi^{\infty}_g\left(1-e^{-5\gamma_g}\right) \approx \pi^{\infty}_g$. However, if we get a measure e after a sequence of 5 times g, then the next parameter's value will be $\pi_e(1) = \pi_g(5)\ e^{-\gamma_e} + \pi^{\infty}_e\,(1-e^{-\gamma_e})$, where $\pi_g(5)$ was the last value used for the g measurement, as calculated before. All in all, however, we realize that the most important term of this expression which determines whether a strategy is failing in QEC or not, is the saturation value $\pi^{\infty}_{g/e}$, even if the time dependence with an exponential saturation is necessary to outperform. Moreover, it is important to notice that the optimal parameters' choice for a given experimental setup (circuit duration and noise level) is not unique, as there is some degree of freedom due to symmetry in the choice of the rotation angles and conditional displacement's parameters.

Another relevant aspect to be analyzed is the probability measurement's outcome determined by the noise level and, in particular, by the parameters' values chosen. We can then compare the strategy of the NMF shown in Fig. 4 in the main text, and the one NMF$_2$ in Fig. S10: despite having different signs in the exponential saturation behavior, which is part of the degree of freedom mentioned above, the main difference is the amplitude of those values. This is reflected into the measurement's probability as shown in the top left panel of Fig. S10 and Fig. 4 in the main text: while the NMF strategy returns a probability for the g measurement $p(g) = 0.9$ constant as long as g are indeed measured, for the NMF$_2$ this probability drops to $p(g) \approx 0.3$, making more frequent the measurement of the qubit in the $|e\rangle$ state. Although this has no negative consequence here, except the limitation of generalizability to very different noise levels, this can be detrimental in experiments, where it has been shown that increasing the probability to measure e can lead to more errors [10].

When using a feed-forward neural network, instead, the output parameters can only assume three possible values, as there is no memory and therefore only a Markovian response: namely, π_0 as initial value, π_g in case of g measurement outcome, and π_e in case of e measurement (Fig. S11). Moreover, as we can notice, the values π_g and π_e assume symmetrically opposite values with respect to the initial value π_0.

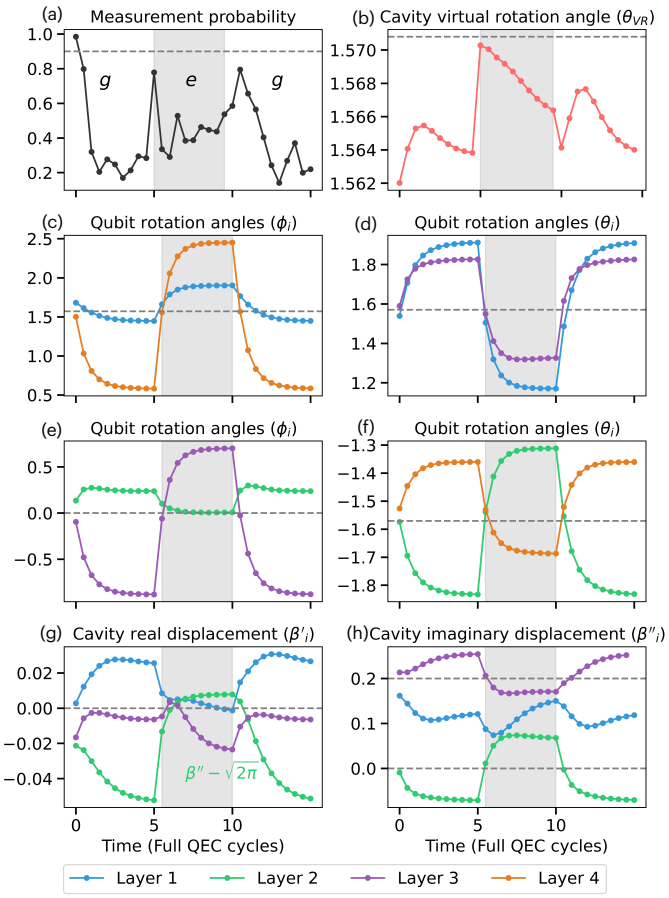

Figure S10. Optimized time-dependent parameters for the RNN of the NMF_2 approach, evaluated with the dynamics in Fig. S2 with low noise rate. The plot corresponds to the one in Fig. 4 for NMF, fixed deterministic measurements' outcome.

Biased-noise gates QEC

We show here that the optimized memory-feedback QEC strategy for the sBs protocol used so far modeling perfect gates, can be generalized also to the case of gates with biased errors constant in time. In particular, we choose the random biased errors for the gates' parameters as shown in Table S6.

Comparing the Pauli expectation value $\langle \hat{Z} \rangle$ as a function of time for the optimized and the standard sBs proto-

Gate	Parameter	Biased error
$\hat{R}_q$	ϕ_1	+ 0.05
$\hat{R}_q$	ϕ_2	- 0.03
$\hat{R}_q$	ϕ_3	- 0.06
$\hat{R}_q$	ϕ_4	+ 0.04
$\hat{R}_q$	θ_1	- 0.03
$\hat{R}_q$	θ_2	+ 0.05
$\hat{R}_q$	θ_3	+ 0.06
$\hat{R}_q$	θ_4	+0.04
$EC\hat{D}_{qc}$	β_1	+ 0.06 - 0.04 i
$EC\hat{D}_{qc}$	β_2	+ 0.04 - 0.02 i
$EC\hat{D}_{qc}$	β_3	+ 0.04 - 0.05 i

Table S6. Biased errors on gates' parameters. List of the biased errors for each parameter of the gates used in the circuit in Fig. 2 for the test of biased noise gates.

col, given the initial state $|+Z_L\rangle$ and applying periodically the QEC cycles, we notice that the lifetime decreases

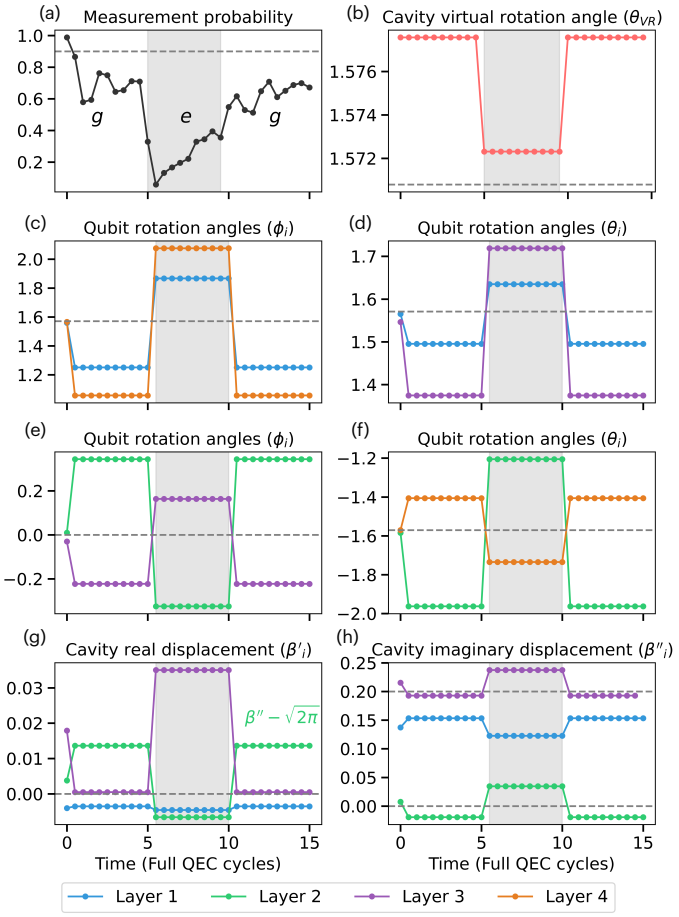

Figure S11. Optimized time-dependent parameters for the feedback agent based on feed-forward neural network without memory (MF), evaluated with the dynamics in Fig. S2 with low noise rate. The plot corresponds to the one in Fig. 4 for NMF, fixed deterministic measurements' outcome.

with respect to the case of perfect gates, as expected (Fig. S12). However, the optimized strategy has a lifetime $T_Z(\text{NMF})/\tau_{cycle} = 3.3 \times 10^2$, while the standard one is $T_Z(\text{standard})/\tau_{cycle} = 1.6 \times 10^2$, i.e. a factor 2 smaller, similarly to the case of perfect gates.

* matteo.puviani@mpl.mpg.de
† florian.marquardt@mpl.mpg.de

[1] S. M. Barnett and D. T. Pegg, Phase in quantum optics, Journal of Physics A: Mathematical and General **19**, 3849 (1986).
[2] P. T. Cochrane, G. J. Milburn, and W. J. Munro, Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping, Phys. Rev. A **59**, 2631 (1999).
[3] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).
[4] B. Royer, S. Singh, and S. M. Girvin, Stabilization of finite-energy gottesman-kitaev-preskill states, Phys. Rev. Lett. **125**, 260509 (2020).
[5] 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**, 368 (2020).
[6] A. L. Grimsmo and S. Puri, Quantum error correction with the gottesman-kitaev-preskill code, PRX Quantum **2**, 020101 (2021).
[7] A. J. Brady, A. Eickbusch, S. Singh, J. Wu, and Q. Zhuang, Advances in bosonic quantum error correction with gottesman-kitaev-preskill codes: Theory, engineering and applications (2023), arXiv:2308.02913 [quant-ph].
[8] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New class of quantum error-correcting codes for a bosonic mode, Phys. Rev. X **6**, 031006 (2016).

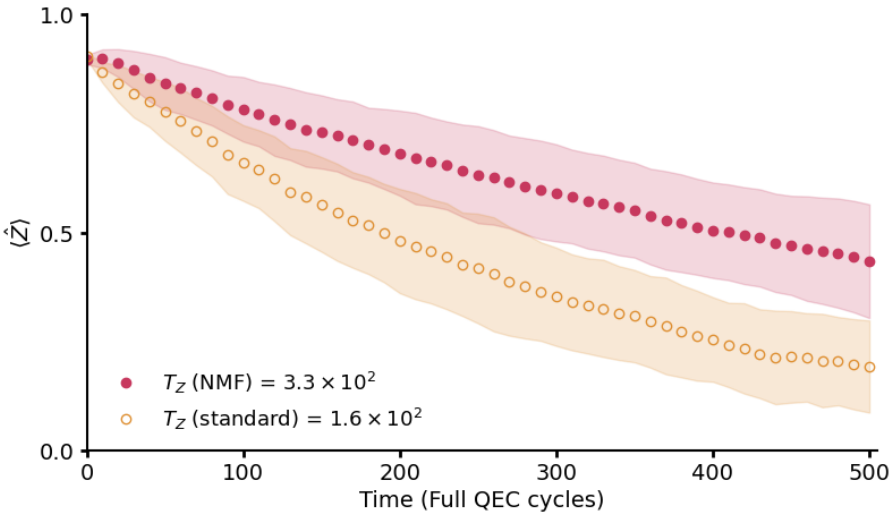

Figure S12. QEC with biased gates. Comparison of the QEC performance between the standard (empty dots, orange) and the NMF (full dots, red) protocol using gates with biased errors as in Table S6. The shaded area represents the standard deviation over all the sampled trajectories.

[9] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. P. Song, C.-L. Zou, S. M. Girvin, L.-M. Duan, and L. Sun, Quantum error correction and universal gate set operation on a binomial bosonic logical qubit, Nature Physics **15**, 503 (2019).

[10] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Real-time quantum error correction beyond break-even, Nature **616**, 50 (2023).

[11] A. Eickbusch, V. Sivak, A. Z. Ding, S. S. Elder, S. R. Jha, J. Venkatraman, B. Royer, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Fast universal control of an oscillator with weak dispersive coupling to a qubit, Nature Physics **18**, 1464 (2022).

[12] R. Porotti, V. Peano, and F. Marquardt, Gradient-ascent pulse engineering with feedback, PRX Quantum **4**, 030305 (2023).

[13] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, Learning phrase representations using rnn encoder-decoder for statistical machine translation (2014), arXiv:1406.1078 [cs.CL].

[14] D. Lachance-Quirion, M.-A. Lemonde, J. O. Simoneau, L. St-Jean, P. Lemieux, S. Turcotte, W. Wright, A. Lacroix, J. Fréchette-Viens, R. Shillito, F. Hopfmueller, M. Tremblay, N. E. Frattini, J. C. Lemyre, and P. St-Jean, Autonomous quantum error correction of gottesman-kitaev-preskill states (2023), arXiv:2310.11400 [quant-ph].