

Decoherence and Wave-Function Deformation of D_4 Non-Abelian Topological Order

Pablo Sala^{1,2,*}, Jason Alicea^{1,2}, and Ruben Verresen^{3,4,5,†}

¹*Department of Physics and Institute for Quantum Information and Matter,
California Institute of Technology, Pasadena, California 91125, USA*

²*Walter Burke Institute for Theoretical Physics, California Institute of Technology,
Pasadena, California 91125, USA*

³*Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA*

⁴*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*

⁵*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*



(Received 16 October 2024; revised 25 March 2025; accepted 8 May 2025; published 1 July 2025)

The effect of decoherence on topological order (TO) has been most deeply understood for the toric code, the paragon of Abelian TOs. We show that certain non-Abelian TOs can be analyzed and understood to a similar degree, despite being significantly richer. We consider both wave-function deformations and quantum channels acting on D_4 TO, which has recently been realized on a quantum processor. By identifying the corresponding local statistical mechanical spin or rotor model with D_4 symmetry, we find a remarkable stability against proliferating non-Abelian anyons. This is shown by leveraging a reformulation in terms of the tractable $O(2)$ loop model in the pure state case and n coupled $O(2)$ loop models for Rényi- n quantities in the decoherence case—corresponding to worldlines of the proliferating anyon with quantum dimension 2. In particular, we find that the purity ($n = 2$) remains deep in the D_4 TO for any decoherence strength, while the $n \rightarrow \infty$ limit becomes critical upon maximally decohering a particular anyon type, similar to our wave-function deformation result. The information-theoretic threshold ($n \rightarrow 1$) appears to be controlled by a disordered version of these statistical mechanical models, akin to the toric code case although significantly more robust. We furthermore use Monte Carlo simulations to explore the phase diagrams when multiple anyon types proliferate at the same time, leading to a continued stability of the D_4 TO in addition to critical phases with emergent $U(1)$ symmetry. Instead of loop models, these are now described by net models corresponding to different anyon types coupled together according to fusion rules. This opens up the exploration of statistical mechanical models for decohered non-Abelian TO, which can inform optimal decoders and which in an ungauged formulation provides examples of non-Abelian strong-to-weak symmetry breaking.

DOI: [10.1103/PhysRevX.15.031002](https://doi.org/10.1103/PhysRevX.15.031002)

Subject Areas: Condensed Matter Physics,
Quantum Physics,
Statistical Physics

I. INTRODUCTION

Exploring and defining mixed-state quantum phases of matter is receiving ever-growing interest [1–40] at the intersection of condensed matter and quantum information. The motivation descends in part from the rapid development of highly controlled quantum platforms, whose sources of decoherence are more accurately

characterized via effective (e.g., Pauli) noise models rather than by more traditional thermal ensembles [41]. This endeavor is also of intrinsic theoretical interest, as pursuing stable quantum coherent phenomena under imperfect conditions requires generalizing the conventional pure “ground-state” paradigm [42,43] to open quantum systems.

One of the most striking emergent quantum phenomena is that of topological order (TO), for which a deep understanding has been achieved under ideal conditions [42,43]. Characterized by emergent anyons [44–46] and ground-state degeneracy [47], these systems can be exploited to robustly store and manipulate quantum information [48–52]. The simplest TO corresponds to a two-dimensional \mathbb{Z}_2 deconfined gauge theory [53,54], which is microscopically realized by the celebrated toric

*Contact author: psala@caltech.edu

†Contact author: verresen@uchicago.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

code model on a square lattice [48]. Correspondingly, much literature [5–10,15–25,55], inspired by the seminal work of Dennis *et al.* [56], has focused on extending this TO to mixed states, appearing as a result of local decohering processes. While this quantum memory breaks down at any finite temperature [1,3,57–60], it has been found—originally in the context of quantum error correction—that quantum information encoded on the ground-state subspace is stable if decoherence does not exceed a certain strength [56]. It turns out that this error-correction threshold p_c (corresponding to an ideal decoder) is set by the finite-temperature phase transition of the 2D random bond Ising model (RBIM), where the disorder is sourced by different error configurations [61]. More recently, Refs. [5–7] followed a different, more agnostic approach to specific error-correction schemes and proposed to directly characterize the decohered density matrix ρ by evaluating various information-theoretic quantities. However, since such quantities are generally difficult to compute, the authors proposed to look at their n -Rényi generalizations associated with different moments $\text{tr}(\rho^n)$ to detect the nontrivial effects of decoherence. Even though these critical thresholds $p_c^{(n)}$ do, in general, depend on the Rényi index n , such generalizations simplify the calculations and connect to simpler statistical mechanical (“stat-mech”). The “intrinsic” (or “information-theoretic”) error threshold p_c can be obtained by taking an $n \rightarrow 1$ replica limit, although finite $p_c^{(n)}$ ’s already inform about singularities in the spectrum of ρ .

However, the rich landscape of TOs consists mostly of non-Abelian states [62–65], whose anyons have internal structure—forming the hardware of a topological quantum computer [48,51]. Previous works considering the effect of decoherence on non-Abelian TO have either focused on error-correcting schemes for certain classes of TOs (without identifying stat-mech models) [66–72] or explored what type of novel intrinsically mixed phases can, in principle, emerge for maximal decoherence rate [16–18]. However, it is *a priori* not clear whether one can hope to understand the effects of decoherence of any non-Abelian TO to a degree that is akin to our understanding of the structurally much simpler Abelian case; it is similarly unclear whether one should expect the physics to be significantly different and more interesting. Here, we answer both these open questions in the *affirmative*. In particular, we study D_4 TO, which has recently been experimentally realized in a trapped-ion processor [73]. The fact that this was the first non-Abelian TO to be realized using a quantum processor is due to it, in some sense, being the simplest non-Abelian TO—thereby making it an ideal candidate to play a paradigmatic role akin to the toric code for Abelian TO. Indeed, although it has as many as 22 anyons [74], it is rather minimal, since all non-Abelian anyons fuse to Abelian ones. This is

known as acyclic or nilpotent property and was crucial to the theory proposal [76] of its realization. This is related to D_4 TO being achingly close to an Abelian TO: It can be thought as a “twisted” Abelian gauge theory [78–81]. Although the (doubled) Ising anyon TO shares a similar property, D_4 TO has an emergent (nonfermionic) gauge group symmetry, similar to the toric code.

With this motivation, the present work is a detailed study of decohering D_4 TO. We consider both *pure wave-function deformations*—realized as imaginary time evolutions for a time β acting on the D_4 topologically ordered ground state—as well as *local quantum channels* acting on an initial pure density matrix with some error rate p . In the following, we refer to both kinds of scenarios as errors. Pure wave-function deformations have been previously considered in the literature (see, e.g., Refs. [82,83] for deformations of the toric code wave function and Refs. [84–88] when considering non-Abelian topological order), exploiting both numerical and analytical tractability, in particular, using tensor network formalism. However, our motivations to consider such wave-function deformations as a prelude to understanding decohered mixed states are the following. First, this scenario provides a conceptually simple way to perturb the pure state while retaining a corresponding local parent Hamiltonian for the modified wave function and involves similar algebraic calculations. Second, the combination of such deformations on top of zero-correlation length wave functions together with local quantum channels allows us to controllably induce a finite-correlation length in the system, which will turn out to provide useful insights. And, finally, as we find, the deformed wave functions turn out to be relevant for understanding the behavior of the stat-mech model appearing in $\text{tr}(\rho^n)$ in the $n \rightarrow \infty$ limit.

A key finding of this work is that while Abelian errors lead, as for the toric code case, to Ising-like stat-mech models, we find that the proliferation of non-Abelian anyons in D_4 TO leads to a D_4 -symmetric spin model. This larger symmetry gives rise to a field theory description at the critical threshold which showcases an emergent $U(1)$ symmetry and, hence, vastly stabilizes the system against both types of errors. We discuss this formulation hand in hand with an exact high-temperature expansion of this model, which is given by an $O(2)$ loop model on a honeycomb lattice [89–91] [indeed, the usual Ising model is an $O(1)$ loop model]. It is no coincidence that $d = 2$ is the quantum dimension of the non-Abelian anyon proliferated by the underlying error.

First, when subjecting D_4 TO to the proliferation of Abelian anyons via Pauli noise, we recover the same phenomenology as found for the decohered toric code ground state [6]. Specifically, D_4 TO breaks down beyond a certain error strength. However, when deforming the wave

function by creating non-Abelian anyons pairs, we find that the system is robust to any finite deformation. This stability relates to the fact that the norm of the deformed wave function, signaling ground-state phase transitions, maps to an $O(2)$ loop model (see the review in Ref. [91]). Here, loop configurations correspond to the worldlines of non-Abelian anyons with quantum dimension $d = 2$, contributing with a topological tension on top of a local string tension given by the error strength. This is then able to give rise to a critical state, which in this case corresponds to a Berezinskii-Kosterlitz-Thouless (BKT) critical point and that can be pushed to an extended gapless phase.

Similarly, when subjecting the system to decoherence, we find that $\text{tr}(\rho^n)$ leads to n coupled $O(2)$ loop models. For example, when computing the purity ($\text{tr}(\rho^2)$), both copies are strongly coupled leading to an $O(4)$ loop model. This translates into the robustness of D_4 TO even at the maximum error rate (under Pauli noise). While the phase diagram of higher Rényis with $n > 2$ is more subtle, we can also argue that in $n \rightarrow \infty$ is very stable and is likely to attain a threshold $p_c^{(\infty)}$ only at the maximum error rate. On more rigorous footing, we have shown that, unlike what has been found for the toric code [6], the thresholds $p_c^{(n)}$ for different Rényi index n are *not a monotonic increasing function* of n . However, the intrinsic threshold p_c (as discussed in Ref. [56]) is controlled by the “replica limit” $n \rightarrow 1$ or, rather, by a transition in information-theoretical quantities that involve the Von Neumann entanglement of ρ . We are able to identify the local stat-mech model determining this threshold (analogous to the RBIM for the toric code) for the decohered density matrix at the maximal error rate, where we can perform a full diagonalization. In particular, we find that the robustness of this non-Abelian quantum memory can be characterized by the free energy cost of inserting a symmetry defect line on a disordered D_4 rotor model.

Finally, we combine both “Abelian” and “non-Abelian” errors and characterize the resulting phase diagram for both pure wave-function deformations and for the decohered mixed state as signaled by a singular behavior in the purity $\text{tr}(\rho^2)$. We find a rich phase diagram that displays several short-range correlated phases, corresponding to various types of TOs, as well as an extended critical phase in the former case which is now stabilized by the presence of Abelian anyons (see Sec. V). Such characterization is possible due to the alternative and equivalent formulations of the resulting loop models using explicitly local variables, which permits an efficient implementation of classical Monte Carlo simulations and the use of local order parameters.

We organize the remainder of the paper as follows. In Sec. II, we start by reviewing a microscopic model realizing D_4 TO and its connection to a

symmetry-enriched $\mathbb{Z}_2 \times \mathbb{Z}_2$ toric code by via a (un)gauging unitary transformation. We then consider the effect of local phase Z errors in Sec. III, which proliferate pairs of Abelian charges. First, we characterize the resulting phase diagram of the corresponding deformed wave function (Sec. III A) and then the fate of the decohered density matrix as measured by the purity (Sec. III B). Section IV then follows a similar structure, first discussing X pure wave-function deformations in Sec. IV A, which proliferates pairs of non-Abelian fluxes. We show that the worldlines of these non-Abelian anyons with quantum dimension $d = 2$ lead to $O(2)$ loop models when computing the norm of the deformed wave function, which is, in turn, reformulated in terms of local Ising-like interactions and as a four-state clock model. In Sec. IV C, we then follow a similar discussion for the purity. However, this section includes three additional results: First, in Sec. IV D, we compute higher moments $\text{tr}(\rho^n)$ of the decohered density matrix where the non-Abelian nature of the proliferating anyon establishes a clear difference with respect to the toric code case. Second, in Sec. IV E, we consider the limit $n \rightarrow \infty$ and prove that the phase diagram as characterized by $\text{tr}(\rho^\infty)$ matches that of the deformed wave function. In Sec. V, we then combine both Z Abelian and X non-Abelian errors on different sublattices and study the phase diagram of both the resulting deformed wave function and decohered density matrix. We also include the statistical models characterizing the effect of proliferating more than one type of non-Abelian anyons in Sec. V C. Finally, we employ the ungauging maps to relate the decoherence transition of D_4 TO to the phenomenon of strong-to-weak spontaneous D_4 symmetry breaking in Sec. VI. We conclude in Sec. VII, discussing the main conclusions of this work as well as various open questions.

In a companion work [92], we explore and confirm some of these results for a broad range of TOs, including the quantum double construction and the Kitaev honeycomb model. The latter provides an example of a nonfixed point wave function with the noise model creating non-Abelian anyons which neither are self-bosons nor have an integer quantum dimension. Moreover, this companion work shows the dependence of the results on the underlying microscopic lattice.

II. REVIEW OF D_4 TOPOLOGICAL ORDER

We consider a microscopic model with D_4 non-Abelian TO following Ref. [78]. It is a solvable spin-1/2 model which is in the same phase of matter as the quantum double $\mathcal{D}(D_4)$ [93–95], which can be thought of as the deconfined phase of an (emergent) D_4 gauge theory. Since the dihedral group $D_4 \cong \mathbb{Z}_4 \rtimes \mathbb{Z}_2$ is the symmetry group of the square and, hence, non-Abelian, this deconfined phase has non-Abelian anyon excitations, as we discuss. However, the model introduced in Ref. [78] can also be thought of as a

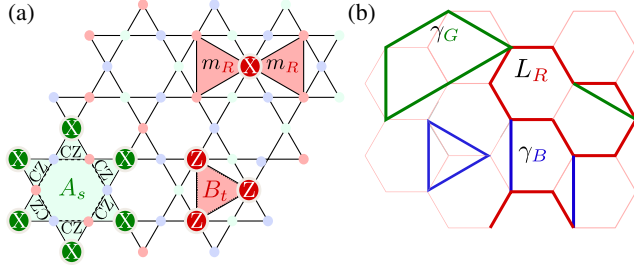


FIG. 1. Microscopic realization of D_4 topological order on the kagome lattice. (a) Local Hamiltonian terms A_s and B_t (specifically, we show A_s and B_t) in Eq. (1). Pairs of non-Abelian (Abelian) m_R (e_B) fluxes (charges) are created by the action of local X_r (Z_b) Paulis acting on the red (blue) sublattice \mathcal{R}_R (\mathcal{R}_B). (b) The worldlines of non-Abelian m_R anyons form closed loops (L_R , red) on the honeycomb lattice (whose vertices lie at the center of triangle terms B_t). On the other hand, Abelian charges e_G and e_B either form closed loops on a triangular lattice (whose vertices lie at the center of star terms A_s^G and A_s^B) or they can end on the red loop, due to the nontrivial fusion $m_R \times m_R = 1 + e_B + e_G + e_B \times e_G$. This configuration space defines a *net model*.

$\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ gauge theory with a subtle kind of “twist” which is responsible for its non-Abelian nature [79–81,96]. In Sec. II B, we use an alternative perspective, where this non-Abelian phase of matter is obtained by gauging the $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry of a symmetry-enriched toric code [97–100].

Our Hilbert space consists of spin-1/2 (qubits) on the vertices of a kagome lattice, and for convenience we work with periodic boundary conditions. Since this lattice is tripartite, we distinguish among red (\mathcal{R}_R), green (\mathcal{R}_G), and blue (\mathcal{R}_B) sublattices; see Fig. 1(a). Equivalently, one can think of qubits lying on the edges of three interleaved honeycomb lattices colored R , G , and B [the red honeycomb lattice corresponding to \mathcal{R}_R is shown in Fig. 1(b)].

A. Microscopic Hamiltonian and anyon content

In this work, we denote by X , Y , and Z the three Pauli matrices corresponding to σ^x , σ^y , and σ^z , respectively. We consider a spin-1/2 Hamiltonian realizing D_4 topological order on the sites of the kagome lattice [78] (following the notation of Ref. [73]):

$$H = - \sum_{s \in \{\star\}} A_s - \sum_{t \in \{\triangleright, \triangleleft\}} B_t \quad (1)$$

with the 12-body operators $A_s = \prod_{i_{\text{in}}=1}^6 \text{CZ}_{i_{\text{in}}, i_{\text{in}}+1} \times \prod_{i_{\text{out}}=1}^6 X_{i_{\text{out}}}$ and the three-body triangle operators $B_t = \prod_{j \in \{\triangleright, \triangleleft\}} Z_j$. For the former, $\text{CZ}_{i,j} = \frac{1}{2}(1 + Z_i + Z_j - Z_i Z_j) = e^{i\pi[(1-Z_i)/2][(1-Z_j)/2]}$ denotes the controlled- Z gate which assigns the phase -1 if and only if the two spins sitting at sites i, j are in the $|\downarrow\rangle$ state. The first product

runs over the i_{in} sites lying within the internal hexagon of a star A , while the second product runs over the exterior i_{out} sites. See Fig. 1(a) for the graphical definition of A_s and B_t .

Analogously to the 2D toric code, A_s ’s can be viewed as plaquette operators defined on the intertwined colored honeycomb lattices, while the B_t ’s correspond to star operators imposing Gauss’s law on each honeycomb-lattice vertex. Each of these operators have an associated color inherited from the three-colorable lattice. Although $[A_s, B_t] = 0$ and $[B_t, B_{t'}] = 0$, the A_s ’s fail to commute with each other (due to the CZ rings lying within each plaquette term). The latter property is important, because otherwise the Hamiltonian would reduce to three decoupled copies of the 2D toric code. Fortunately, the ground-state manifold corresponds to $B_t = +1$ in every triangle. In this sector, $[A_s, A_{s'}]|_{B_t=+1} = 0$. Consequently, $A_s = B_t = +1$ holds in the ground-state subspace.

This topological order hosts a total of 22 anyons (for a detailed description, see the Appendixes of Refs. [73,101]). Of these anyons, eight are Abelian (i.e., have quantum dimension $d = 1$), and the remaining are non-Abelian with quantum dimension $d = 2$. All Abelian anyons, modulo the trivial particle, correspond to self-bosonic electric charges [102] that we label $e_R, e_G, e_B, e_{RG}, e_{RB}, e_{BG}$, and e_{RGB} [103]. The first three charges—which generate the others via fusion (i.e., for example, $e_{RG} = e_R \times e_G$)—arise from violation of a kagome star operator A_s . Namely, a kagome star with $A_s = -1$ corresponds to a charge e_R, e_G , or e_B depending on the color of the flipped operator. These anyons can be created in pairs by the action of local Pauli Z operators. For example, a pair of e_B ’s are created at the blue stars s_i and s_f by the action of the Wilson operator $\mathcal{Z}_{s_i}^{s_f} \equiv \prod_{b \in \gamma_{xy}} Z_b$, where γ_{xy} is any open string with support only on the blue sublattice \mathcal{R}_B and with x and y vertices lying in s_i and s_f , respectively. Hence, for closed contractible loops with $s_i = s_f$, $\mathcal{Z}_{s_i}^{s_i}$ acts as the identity on the D_4 ground state, since such an operator creates and then fuses Abelian anyons. Logical operators $\mathcal{Z}_{R,G,B}$ —which can act nontrivially in the ground-state manifold—correspond to the product of Z operators along closed *non-contractible* (both horizontal and vertical) loops with $s_i = s_f$ and with support only on the corresponding sublattice $\mathcal{R}_{R,G,B}$.

The remaining 14 non-Abelian anyons have either bosonic, fermionic, or semionic self-statistics depending on their topological spin. The fundamental fluxes m_R, m_G , and m_B —which together with the electric charges above generate all anyons via fusion—can also be created in pairs and correspond to violations of the triangle operator $B_t = -1$. For example, a pair of fluxes on nearby triangles can be generated by a single Pauli X operator acting on the common vertex [see the two coinciding red triangles in Fig. 1(a)]. Creation of a pair of distant fluxes m_R at red

triangles t_i and t_f , however, requires a string of X 's supplemented by a unitary circuit whose depth is linear (in the distance between the two triangles) [73]:

$$\mathcal{X}_{t_i}^{t_f} = \prod_{r \in \gamma_{x,y}} X_r \prod_{b < g \in \gamma_{x,y}} CZ_{bg}, \quad (2)$$

where $\gamma_{x,y}$ is an open string connecting sites x and y lying on triangles t_i and t_f , respectively, and where $b < g \in \gamma_{x,y}$ means we go over all blue and green sites which are passed upon traversing the red string and consider all pairs where blue is on the left of green [104]. The additional CZ's along the path appear due to the non-Abelian nature of the m_R 's: Unlike for Abelian anyons, fusion of two non-Abelian fluxes leads to more than one possible outcome, e.g., $m_R \times m_R = 1 + e_G + e_B + e_G e_B$. Consequently, with only a product of X 's, non-Abelian anyons created by adjacent X operators would fuse nontrivially—creating a superposition of allowed anyon fusion products along the entire length of the X string. (This point is very important in later sections.) The CZ's “clean up” these excitations such that the operator $\mathcal{X}_{t_i}^{t_f}$ creates a pair of fluxes in the trivial fusion channel, *without* excess anyons in the intervening region. See additional details in Ref. [73]. A graphical example of this operator is shown in Fig. 14. We note that Eq. (2) also affects the A_s operators with t_i, t_f lying at their center, such that A_s takes an indefinite value with $\langle A_s \rangle = 0$, which is related to the higher-dimensional internal space associated with the non-Abelian anyons. Similarly to the logical operators $\mathcal{Z}_{R,G,B}$ discussed above, a second set of logical operators $\mathcal{X}_{R,G,B}$ analogously correspond to closed noncontractible loop configurations with $t_i = t_f$.

Hereafter, we take $|D_4\rangle$ to correspond to the ground state where all \mathcal{Z} logical operators of the theory act as the identity.

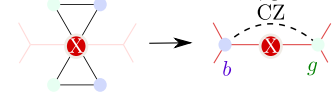
B. Ungauging and disentangling maps

Here, we highlight a particular change of variables that prove useful later on. It involves dualizing some of the sublattices, which can be interpreted as *ungauging*, thereby mapping it to a simpler theory.

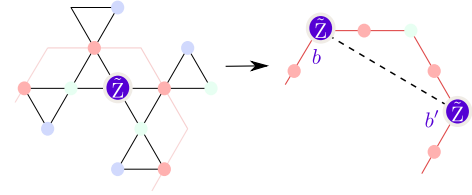
In Sec. IV, we study the proliferation of fundamental non-Abelian fluxes. Since all three sublattices are manifestly on the same footing in the above D_4 TO model, without loss of generality we choose to proliferate fluxes m_R associated to the red sublattice. For this reason, the following change of variables treats the red sublattice on a different footing than blue and green. In particular, we interpret the red sites of the kagome lattice as living on the *links* of a honeycomb lattice. We then dualize the green and blue sublattices of the kagome lattice in the sector $B_t^B = B_t^G = 1$, such that they map to spin-1/2's living on the *vertices* of the aforementioned honeycomb lattice

(with blue and green corresponding to the two sublattices of the honeycomb).

In conclusion, we map the kagome lattice to a “heavy-hexagonal” lattice, with spin-1/2's on the vertices and bonds of the honeycomb lattice. Operators (within the $B_t^B = B_t^G = 1$ sector) map to the new Hilbert space as follows: While $Z^R \rightarrow Z^R$ is unchanged, we have



$$(3)$$



$$(4)$$

and similarly upon exchanging the blue and green sublattices. We have not shown how A_s and A_s map, since it is somewhat more complicated [105] and is not explicitly used in the main text; however, when acting on states where $B_t^R = 1$, they simply map to single-site X^B and X^G , respectively. We use a tilde \sim to refer to the ungauged degrees of freedom on the green and blue sublattices. Here and in the following section, we make a slight abuse of notation and use $\mathcal{R}_{G,B}$ to denote both green and blue sublattices on the kagome lattice, as well as sublattices of the honeycomb lattice, respectively.

One can straightforwardly see that the above mapping indeed preserves the Pauli algebra. The precise structure of the mapping was chosen such that it maps the D_4 TO to a much simpler state:

$$|D_4\rangle_{\text{kagome}} \rightarrow |\text{TC}\rangle_{\mathcal{R}_R} \otimes \underbrace{|\tilde{+}\rangle_g^{\otimes |\mathcal{R}_G|} \otimes |\tilde{+}\rangle_b^{\otimes |\mathcal{R}_B|}}_{=|\tilde{+}\rangle^{\otimes N_{\text{sites}}}}. \quad (5)$$

Indeed, it can be checked that A_s^R maps to the usual toric code stabilizer $\prod_{j \in \square} X_j$. This can be interpreted as mapping the D_4 TO to a toric code which is symmetry enriched with a $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry corresponding to spin-flip symmetries of the blue and green sublattices [99,100], where we have chosen to put the symmetry enrichment in the operator algebra rather than the state [106]. The result of mapping the (red) Wilson operator $\mathcal{X}_{t_i}^{t_f}$ in Eq. (2) is

$$\mathcal{X}_{t_i}^{t_f} \rightarrow \left(\prod_{r \in \gamma_{x,y}} X_r^R \right) \times CZ_{xy}, \quad (6)$$

which is just the usual toric code anyon string, dressed with a CZ connecting its two end points. We provide a visual derivation of this in Fig. 14 in Appendix B.

In the next two sections, we consider Pauli errors of a single type, leaving the combination of several types of errors for later sections. For these, the previous mapping is of great use to simplify the calculations and map the system to local stat-mech models. We first consider errors that generate Abelian charges and, hence, are parallel to previous studies of the toric code ground state under decoherence [5–10,16,19,21–25,55,56]. We then consider the proliferation of non-Abelian anyons. For each case, we examine pure wave-function deformation and then treat decohered mixed states.

III. Z PHASE ERROR: ABELIAN ANYON PROLIFERATION

In this section, we first study the effects of proliferating the *Abelian* anyons e_R , e_G , and e_B of D_4 topological order. This, in large part, reproduces the phenomenology known for the toric code, but it allows us to introduce useful concepts and provides a point of comparison for when we consider the effect of non-Abelian anyons in Sec. IV. Moreover, in Sec. V, we study the interplay of proliferating both Abelian and non-Abelian anyons at the same time.

A. Pure wave-function deformation

We start by considering pure wave-function deformations of the form

$$|\psi(\beta_B^z, \beta_G^z, \beta_R^z)\rangle = \prod_{c \in \{B, G, R\}} e^{\beta_c^z/2 \sum_{j \in \mathcal{R}_c} Z_j} |D_4\rangle. \quad (7)$$

Here, we allow for different deformation strengths β_c^z in each sublattice. The exponential represents a nonunitary operator that restructures the weight on the configurations present in $|D_4\rangle$. As we saw in Sec. II, Z operators generate pairs of Abelian charges e_c , where their color $c = R, G, B$ depends on the sublattice on which the Z 's act. Since this “Abelian” local deformation does not couple different sublattices, it turns out that the resulting phase transitions are decoupled for each of the three colors. Hence, we start by focusing on a single color— B for concreteness.

To characterize the phase diagram of $|\psi(\beta_B^z)\rangle$ (where $\beta_G^z = \beta_R^z = 0$), it is sufficient to look at the wave-function overlap

$$\mathcal{Z}_{|\psi\rangle}(\beta_B^z) = \langle \psi(\beta_B^z) | \psi(\beta_B^z) \rangle. \quad (8)$$

If we expand out $|\psi(\beta_B^z)\rangle$ into a basis which is diagonal in the nonunitary perturbation (in this case, $e^{(\beta_B^z/2) \sum_{j \in \mathcal{R}_B} Z_j}$), we can directly interpret $\mathcal{Z}_{|\psi\rangle}(\beta_B^z)$ as a two-dimensional classical partition function. All its classical correlation functions directly capture observables which are diagonal

in that basis. In fact, we prove in Appendix A that *all* local observables (even off-diagonal ones) of the quantum state are captured by correlation functions in the classical model (due to our perturbation involving only local interactions [107]). We have, thus, reduced this two-dimensional quantum problem to studying a two-dimensional classical problem.

To calculate this partition function, we use the fact that $e^{i\vec{a}\cdot\vec{\sigma}} = \cosh(\alpha) + \sinh(\alpha)\vec{n} \cdot \vec{\sigma}$, where $\vec{\sigma} = (X, Y, Z)$ and \vec{n} is any unit vector:

$$\begin{aligned} \mathcal{Z}_{|\psi\rangle}(\beta_B^z) &= \langle D_4 | e^{\beta_B^z \sum_{j \in \mathcal{R}_B} Z_j} | D_4 \rangle \\ &\propto \langle D_4 | \prod_{j \in \mathcal{R}_B} [1 + \tanh(\beta_B^z) Z_j] | D_4 \rangle, \end{aligned} \quad (9)$$

where we drop the analytic prefactor $\cosh(\beta_B^z)^{|\mathcal{R}_B|}$. Expanding out the product yields

$$\mathcal{Z}_{|\psi\rangle}(\beta_B^z) = \sum_{n_B=0}^{|\mathcal{R}_B|} \tanh(\beta_B^z)^{n_B} \sum_{\{r_j\}_{j=1}^{n_B}} \langle D_4 | \prod_{j=1}^{n_B} Z_{r_j} | D_4 \rangle \quad (10)$$

up to an overall prefactor that in the following we also ignore. Here, n_B counts the number of Pauli Z 's that act on the D_4 ground state at vertices $\{r_j\}$. To evaluate $\langle D_4 | \prod_{j=1}^{n_B} Z_{r_j} | D_4 \rangle$, recall that Z_{r_j} creates pairs of Abelian anyons e_B at the center of the kagome star operators A_s^B . If $\prod_{j=1}^{n_B} Z_{r_j}$ acting on the ket $|D_4\rangle$ includes at least one open string that generates unfused e_B anyons, then the overlap with $\langle D_4 |$ vanishes. Conversely, if $\prod_{j=1}^{n_B} Z_{r_j}$ forms a set of closed loops L_B —either contractible or noncontractible—then it acts on $|D_4\rangle$ as the identity as discussed in the previous section. (Contractible closed loops act trivially because they create and then fuse Abelian anyons in a way that smoothly connects to doing nothing; noncontractible loops are trivial because we consider the ground state where Z -type logical operators act trivially on $|D_4\rangle$.) Closed-loop configurations L_B correspondingly yield $\langle D_4 | \prod_{j=1}^{n_B} Z_{r_j} | D_4 \rangle = 1$.

In summary, we find that (up to an inconsequential prefactor) Eq. (10) reduces to the partition function of an $O(1)$ loop model on the triangular lattice:

$$\mathcal{Z}_{N=1}(t_B) = \sum_{L_B} t_B^{|L_B|}, \quad (11)$$

defined on a triangular lattice with tension $t_B \equiv \tanh(\beta_B^z)$. An example of a closed loop is shown in Fig. 2. Figure 2(a) shows a closed-loop configuration L_B on the original kagome lattice, while Fig. 2(b) shows the same loop configuration in the effective triangular lattice. In the exponent, $|L_B|$ denotes the total loop length in a given closed-loop configuration L_B . The derivation illustrates that

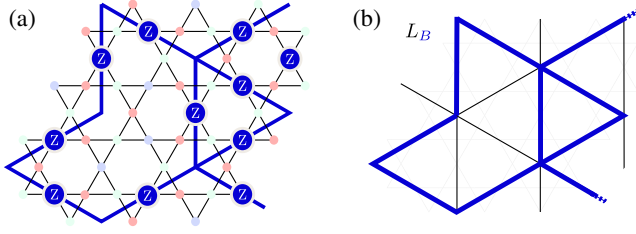


FIG. 2. Abelian anyons moving on a triangular lattice. Abelian charges e_B move on the blue triangular lattice corresponding to the center of the star operators A_s^B . (a) shows a closed-loop configuration formed by the action of local Pauli Z acting on \mathcal{R}_B . These form a closed (intersecting) loop configuration on the triangular lattice (b).

we can think of these closed loops as worldlines of the Abelian anyons. As the name suggests, the $O(1)$ loop model is a special case of the $O(N)$ loop model, where each loop component [108] is weighted by an additional topological factor N [91]. We argue that it is not a coincidence that here we obtain $N = 1$ for an Abelian anyon. As we see in Sec. IV, for non-Abelian anyons with nontrivial quantum dimension d , a loop model with $N = d$ can arise. We explore this for a broad range of non-Abelian TOs in a companion paper [92].

Let us explore the phase diagram of Eq. (11). When the perturbation is small and positive ($0 \leq \beta_B^z \ll 1$), we see that the string tension $t_B \ll 1$ penalizes loops. This small-loop phase corresponds to the initial D_4 TO. As we increase the perturbation, we eventually cross a critical tension t_c , beyond which loops proliferate at all scales. Physically, this corresponds to a condensate of e_B . To infer the resulting phase of matter, let us consider the $\beta_B^z \rightarrow \infty$ limit and examine the local terms A_s and B_t in the Hamiltonian [Eq. (1)] for each of the three colored sublattices. In this limit, all qubits in the blue sublattice \mathcal{R}_B are projected into the $|\uparrow\rangle$ state, and, hence, all B_t^B 's become trivial. Physically, this property reflects the confinement of m_B fluxes by condensing e_B charges. Of the remaining four types of A_s and B_t “stabilizers,” the red B_t and green B_t triangle operators are left untouched, while in the limit $\beta_B^z \rightarrow \infty$ all controlled- Z gates in A_s and A_s become identity [109]. We, therefore, obtain two decoupled sets of (commuting) stabilizers A_s, B_t and A_s, B_t for the red and green sublattices, respectively, corresponding to *two decoupled copies of the toric code*. In particular, the resulting phase of matter [110] is, thus, Abelian.

It turns out the critical value of the string tension is $t_c = 2 - \sqrt{3} \approx 0.268$, i.e., $\beta_c = \ln(3)/4$, and at this point the system is described by the Ising universality class. To see this, it is useful to note that the $O(1)$ loop model Eq. (11) can be rewritten as an *exact* high-temperature expansion of the Ising model on the triangular lattice with inverse temperature β_B^z , for which the critical temperature is known via the Kramers-Wannier duality and star-triangle

relation [112]. In fact, we can directly derive the Ising model by using the *ungauging* transformation in Eq. (4):

$$\mathcal{Z}_{|\psi\rangle}(\beta_B^z) = \langle D_4 | e^{\beta_B^z \sum_{j \in \mathcal{R}_B} Z_j} | D_4 \rangle \quad (12)$$

$$= \frac{1}{2^{\mathcal{N}_B}} \sum_{\vec{\sigma}} e^{\beta_B^z \sum_{\langle b, b' \rangle_{\Delta}} \tilde{\sigma}_b \tilde{\sigma}_{b'}}. \quad (13)$$

The phase where e_B anyons condense corresponds to the ordered phase. Indeed, the two-point correlator $\langle \tilde{Z}_b \tilde{Z}_{b'} \rangle$ on the blue sublattice which diagnoses long-range magnetic order maps to the Wilson operator $\langle \mathcal{Z}_{s_i}^f \rangle$ under the (un)gauging map.

While one benefit of the loop model representation (11) is that it makes manifest the physics of the model (i.e., proliferation of anyon worldlines), one advantage of the Ising model representation (13) is that the partition function remains manifestly positive even if $\beta_B^z < 0$. This introduces *frustration* to the triangular lattice Ising model. This can also be seen in the loop model picture: Different worldlines can now destructively interfere, such that anyon condensation is more difficult to achieve [113]. The consequence is that the D_4 phase is robust for any *finite* negative perturbation strength. In the limit $\beta_B^z \rightarrow -\infty$, the system becomes critical, described by a conformal field theory (CFT) with central charge $c = 1$ [115, 116]. Indeed, in this limit, it is well known that the ground-state manifold of the triangular lattice Ising model can be mapped onto a dimer model on the honeycomb lattice, since each triangle has exactly one frustrated bond [117–121]. Since the honeycomb lattice is bipartite, this gives rise to a $U(1)$ gauge theory with algebraic correlations. Interestingly, similarly to the ferromagnetic case, it can be shown that the remaining (green and red) sublattices define two decoupled toric codes, coexisting with the gapless degrees of freedom residing on the blue sublattice.

Lastly, we note that very similar results follow when considering Abelian deformations on two different sublattices—say, blue and green. For instance, in the strongly ferromagnetic ($t_B, t_G \gg 1$) case, the system becomes a single copy of the toric code on the honeycomb lattice with stabilizers given by A_s and B_t after setting $Z_b = Z_g = +1$.

B. Decohered mixed state

Next, we consider deformations implemented by a composition of local Z quantum channels (i.e., completely positive trace preserving maps) $\mathcal{E}_j(\rho_0) = (1 - p_B)\rho_0 + p_B Z_j \rho_0 Z_j$, with p_B the error rate and $\rho_0 = |D_4\rangle\langle D_4|$ the pure-state density matrix for D_4 topological order. Allowing for such errors on all sites j on the blue sublattice \mathcal{R}_B , the density matrix evolves according to

$$\rho_0 \rightarrow \rho = \mathcal{E}(\rho_0) = \prod_j \mathcal{E}_j(\rho_0). \quad (14)$$

Unlike for ground states, the meaning of a mixed-state phase—even an operational definition—is an active area of research. From the perspective of quantum error correction, one cares about the amount of quantum information preserved under the action of a quantum channel (as measured by the coherent information) [6,7,56]. Alternatively, one might characterize the decohered density matrix ρ according to the existence, or lack thereof, of an unraveling in terms of short-range entangled states [9,122]. These two characterizations correspond to the more “intrinsic” properties of ρ . Yet, such quantities are generally difficult to calculate; see, e.g., Ref. [25]. Recently, various works [5,7,13,122,123] have considered looking for singularities of the moments $\text{tr}(\rho^n)$ of the density matrix [of which the purity $\text{tr}(\rho^2)$ is typically the simplest moment], since these moments determine the full spectrum of ρ [124]. This approach provides valuable insight, although it may not suffice for defining a mixed-state phase [36]. (In the context of strong-to-weak spontaneous symmetry breaking, Ref. [36] showed that mixed-state phases can instead be characterized by the behavior of the quantum fidelity.) In fact, both the universality of the transition as well as its location, in general, depend on n . This property stands to reason: For example, when computing $\text{tr}(\rho^n)$ for a thermal density matrix, the inverse temperature is amplified by a factor of n . Following the above references, in this subsection, we characterize the effect of Z errors on D_4 topological order via the purity $\text{tr}(\rho^2)$, delaying a more detailed discussion of the “intrinsic” threshold to later sections where we also focus on the more interesting case of proliferating non-Abelian anyons.

Although it is not essential and an alternative derivation is discussed in later sections, a similar approach to the deformed wave function (Sec. III A) can be applied to a decohered density matrix ρ . To this end, we vectorize ρ by mapping it to a vector $|\rho\rangle\rangle \in \mathcal{H} \otimes \mathcal{H}$ in a doubled Hilbert space (a detailed explanation can be, e.g., found in Appendix B.2 of Ref. [122]). The quantum channel then becomes a non-negative (and nonunitary) operator acting in the doubled Hilbert space. Upon vectorizing the pure-state density matrix [122] as $|\rho_0\rangle\rangle = |D_4\rangle \otimes |D_4\rangle$, the vectorized decohered density matrix reads

$$|\rho(p_B)\rangle\rangle = |\mathcal{E}(\rho_0)\rangle\rangle = \prod_{j \in \mathcal{R}_B} (1 - p_B + p_B Z_j \otimes Z_j) |\rho_0\rangle\rangle. \quad (15)$$

Equivalently, we can write

$$|\rho(p_B)\rangle\rangle \propto e^{\mu_B \sum_{j \in \mathcal{R}_B} Z_j \otimes Z_j} |\rho_0\rangle\rangle \quad (16)$$

with $\tanh(\mu_B) = [p_B/(1 - p_B)]$, which resembles the form of the deformed pure state $|\psi(\beta_B^z)\rangle\rangle$.

The norm of $|\rho(p_B)\rangle\rangle$ corresponds to the purity $\text{tr}(\rho^2)$ and can be written as

$$\begin{aligned} \text{tr}(\rho^2) &= \langle\langle \rho(p_B) | \rho(p_B) \rangle\rangle \\ &\propto \langle\langle \rho_0 | \prod_{j \in \mathcal{R}_B} (1 + r_B Z_j \otimes Z_j) | \rho_0 \rangle\rangle. \end{aligned} \quad (17)$$

Here, $r_B \equiv [2p_B(1 - p_B)]/[(1 - p_B)^2 + p_B^2]$, which is invariant under $p_B \rightarrow 1 - p_B$. Adopting the same steps as used to derive Eq. (11), we again find the partition function of an $O(1)$ loop model on a triangular lattice:

$$\text{tr}(\rho^2)^{\text{decohere}} \propto_{e_B} \sum_{L_B} r_B^{|L_B|}, \quad (18)$$

with tension given by $r_B \in [0, 1]$. Notice that in this case μ_B (r_B) is always non-negative, and, hence, there is no analog of the antiferromagnetic case discussed in the previous version. Similar to the deformed wave function, this decohered mixed state showcases a finite-temperature phase transition separating a disordered from a ferromagnetically ordered phase at a finite $p_B < 1/2$. This transition indicates that, beyond a finite error threshold $p_c^{(n=2)} = \frac{1}{2}(1 - 3^{-1/4}) \approx 0.12$, the D_4 topological order is lost as measured by the purity.

For completeness, we note that, following the usual computations demonstrated for the square lattice [9,25], one can show that the spectrum of ρ is given by the RBIM [125–132] on the honeycomb lattice with inverse temperature $\beta = \frac{1}{2} \ln[(1 - p)/p]$ or, equivalently, $p = 1/(1 + e^{2\beta})$. More precisely, different eigenvalues are labeled by different disorder realizations of the RBIM, each of which can undergo a critical point. In fact, the largest eigenvalue of this ensemble corresponds to the *clean* Ising model, which on the honeycomb lattice has a critical $\beta_c = \ln(2 + \sqrt{3})/2$. This informs us that $\lim_{n \rightarrow \infty} \text{tr}(\rho^n)$ has a transition at $p_c^{(n=\infty)} = 1/(3 + \sqrt{3}) \approx 0.21$. If one instead takes the whole ensemble into consideration, it turns out the weighting is such that the disorder probability p coincides with the decoherence rate p . This means that the Von Neumann entropy then corresponds to the quenched average free energy of the 2D RBIM along the Nishimori line [127]. On the honeycomb lattice, this is known to occur at $p_c^{(n=1)} \approx 0.068$ [132]. Note that the critical values grow monotonically with the above Rényi indices ($n = 1, 2, \infty$), which has also been observed in earlier work [6]. However, we argue this does not need to hold for non-Abelian cases.

IV. X FLIP ERROR: NON-ABELIAN ANYON PROLIFERATION

We now consider deforming the D_4 topological order by applying local X Paulis. We first consider a single sublattice (e.g., \mathcal{R}_R), since, together with the previous section, this analysis will give us the main insights to obtain the corresponding stat-mech models for general (commuting) deformations. As in the Abelian case, we first consider pure wave-function deformation (Sec. IV A) as a warm-up before advancing to the mixed-state setup (Sec. IV C).

A. Pure wave-function deformation

As we saw in Sec. II, X Paulis creates pairs of non-Abelian m_R charges with quantum dimension $d = 2$. The red sublattice \mathcal{R}_R itself forms a (super)kagome lattice. It is useful to interpret this as the medial lattice of a honeycomb lattice, such that the red qubits live on its bonds [see Fig. 1(b)]. Then, the non-Abelian m_R 's lie on the vertices of this honeycomb lattice. We consider a nonunitarily deformed unnormalized wave function

$$|\psi(\beta_R^x)\rangle = e^{\beta_R^x/2} \sum_{j \in \mathcal{R}_R} X_j |D_4\rangle. \quad (19)$$

Following precisely the same steps as for Z errors in the previous section, the norm can be expressed as

$$\begin{aligned} \mathcal{Z}_{|\psi\rangle}(\beta_R^x) &= \langle D_4 | e^{\beta_R^x \sum_{j \in \mathcal{R}_R} X_j} | D_4 \rangle \\ &\propto \langle D_4 | \prod_{j \in \mathcal{R}_R} [1 + \tanh(\beta_R^x) X_j] | D_4 \rangle \\ &= \sum_{n_R=0}^{|\mathcal{R}_R|} \tanh(\beta_R^x)^{n_R} \sum_{\{r_j\}_{j=1}^{n_R}} \langle D_4 | \prod_{j=1}^{n_R} X_{r_j} | D_4 \rangle. \end{aligned} \quad (20)$$

To proceed, we need to evaluate $\langle D_4 | \prod_{j=1}^{n_R} X_{r_j} | D_4 \rangle$ with $\{r_j\}$ a collection of n_R vertices. Let us first intuitively understand which configurations yield a nonzero contribution. Similar to the case of Z errors, the expectation value vanishes unless $\prod_{j=1}^{n_R} X_{r_j}$ forms a closed-loop configuration L_R . Indeed, any open strings yield unfused non-Abelian anyons in the D_4 ket, leading to orthogonality with the D_4 bra. We thus have

$$\mathcal{Z}_{|\psi\rangle}(\beta_R^x) \propto \sum_{L_R} \tanh(\beta_R^x)^{|L_R|} f(L_R), \quad (21)$$

with L_R any contractible [133] closed-loop configuration, and $f(L_R) \equiv \langle D_4 | \prod_{j \in L_R} X_j | D_4 \rangle$.

Similarly to the Abelian case, we again find a loop model (albeit now on the honeycomb lattice), and $\tanh(\beta_R^x)$ contributes to the string tension [134]. However, unlike the Abelian case, we now have the remaining expectation value $\langle D_4 | \prod_{j \in L_R} X_j | D_4 \rangle$. In the Abelian case, this would

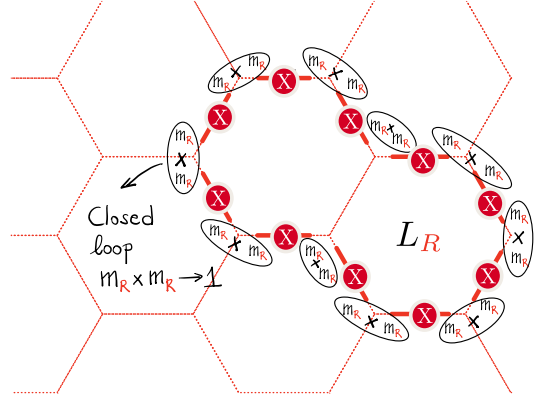


FIG. 3. Closed-loop condition. Every local X_r acting on the red sublattice creates a pair of m_R fluxes. Finite contributions in the partition function (20) correspond to fusing pairs of nearby fluxes $m_R \times m_R$ into the identity channel $m_R \times m_R \rightarrow 1$. For anyons belonging to different pairs, this fusion process occurs with probability $1/d_{m_R}^2$, giving rise to a finite tension in the loop model. Every closed string, moreover, contributes with a topological factor given by the quantum dimension to the number of connected components of the loop configuration $2^{C_{L_R}}$.

simply be unity, since it is the expectation value of a (product of) ground-state stabilizer(s). This is not the case now: Only the expectation value of the true non-Abelian anyon string operator would give a unity value in the ground state, but that is a complicated linear-depth circuit [see Eq. (2)]. Instead, we can think of $f(L_R)$ as the overlap between the ground state $\langle D_4 |$ and a state $\prod_{j \in L_R} X_j | D_4 \rangle$ where we have first created $|L_R|/2$ pairs of adjacent m_R fluxes which are then pairwise fused as sketched in Fig. 3. However, this fusion can contain nontrivial anyons:

$$m_R \times m_R = 1 + e_B + e_G + e_B \times e_G. \quad (22)$$

Only the “1” term gives a nonzero overlap with the ground state. We, thus, expect $f(L_R) < 1$.

We momentarily give a physical way of deriving $f(L_R)$ using the fusion-based interpretation, but let us first show the algebraic way. We exploit the ungauging and disentangling maps introduced in Sec. II B, which tells us:

$$f(L_R) = \langle \tilde{\dagger}_g^{\otimes |\mathcal{R}_G|} | \langle \tilde{\dagger}_b^{\otimes |\mathcal{R}_B|} | \prod_{(g,b) \in L_R} \widetilde{\text{CZ}}_{gb} | \tilde{\dagger}_g^{\otimes |\mathcal{R}_G|} | \tilde{\dagger}_b^{\otimes |\mathcal{R}_B|} \rangle \quad (23)$$

$$= \prod_{\ell_R \in L_R} \frac{1}{2^{|\ell_R|}} \text{tr} \left(\prod_{n=1}^{|\ell_R|} \widetilde{\text{CZ}}_{n,n+1} \right), \quad (24)$$

where we have decomposed $L_R = \bigoplus \ell_R$ into its connected components and where the trace is over a circuit of CZ's on a ring with periodic boundary conditions. We note that this expression was also obtained in a study of

symmetry-protected topological phases [135]. Focusing on a single component and by writing Eq. (24) as tensor network (see Appendix C 1), one can straightforwardly show that

$$f(L_R) = \frac{1}{\sqrt{2}^{|L_R|}} \text{tr}(H^{|L_R|}) = \frac{\text{tr}(\mathbb{1})}{\sqrt{2}^{|L_R|}} = \frac{2}{\sqrt{2}^{|L_R|}}, \quad (25)$$

where $H = [(X + Z)/\sqrt{2}]$ is the Hadamard matrix and we use that $|L_R|$ is automatically even on a honeycomb lattice.

Putting everything together, we obtain that the wavefunction norm is described by the following honeycomb loop model:

$$\mathcal{Z}_{|\psi\rangle}(\beta_R^x) \propto \sum_{L_R} t_R^{|L_R|} 2^{C_{L_R}}, \quad (26)$$

with a string tension $t_R = \tanh(\beta_R^x)/\sqrt{2}$ and where C_{L_R} counts the number of loops (or, equivalently, connected components) in the closed-loop configuration L_R . Consider, for example, the loop configuration in Fig. 3 with only one loop, and, hence, $C_{L_R} = 1$.

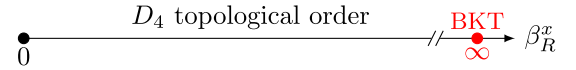
A more physical way of deriving the above expression for $f(L_R)$ is as follows. As we discussed, the (normalized) wave function $|L_R\rangle \equiv \prod_{j \in L_R} X_j |D_4\rangle$ includes all possible outcomes resulting from the fusion of two nearby fluxes m_R , one of which leads to a term proportional to the ground state $|D_4\rangle$. To obtain its numerical prefactor, we can use the fact that different anyon pairs are created separately, and then the fusion outcomes of two (causally disconnected) anyons belonging to two different pairs occur with a probability weighted by the quantum dimension of the fusion product [136,137], i.e., $p(m_R \times m_R \rightarrow a) = d_a/d_{m_R}^2$. Moreover, for every connected component of the closed-loop configuration L_R , an additional d_{m_R} appears. This is a consequence of the fact that, once $|L_R| - 1$ such pairs have fused into the vacuum, then the remaining pair is necessarily in the trivial channel, since all pairs were created from the vacuum (indeed, now the remaining anyons are no longer causally disconnected). Hence, we find

$$|L_R\rangle = \frac{2^{C_{L_R}}}{\sqrt{2}^{|L_R|}} |D_4\rangle + \dots, \quad (27)$$

where \dots is orthogonal to $|D_4\rangle$. Clearly, $f(L_R) = \langle D_4 | L_R \rangle$ exactly picks up this prefactor, arriving at Eq. (25). The reason why here we find $\sqrt{d_{m_R}^{|L_R|}}$ rather than $d_{m_R}^{|L_R|}$ is because for this microscopic model a pair of nearest m_R anyons can fuse only into either the trivial or an Abelian charge e_c whose color $c = G, B$ depends on the lattice site. Indeed, more generally, the length-dependent factor can depend on microscopic details, but the topological piece of $d_{m_R}^{C_{L_R}}$ which depends on the number of components should be universal.

Equation (26) is the partition function for an $O(2)$ loop model with tension t_R defined on the \mathcal{R}_R honeycomb lattice [89,90,138]. Since all loop configurations have even length $|L_R|$, the result does not depend on the sign of β_R^x . It is useful to contrast the $O(2)$ loop model in Eq. (26) to the $O(1)$ loop model we obtained in Eq. (11) when proliferating Abelian e_B charges via Z deformation. Unlike for the latter, proliferating non-Abelian fluxes with quantum dimension $d = 2$ gives rise to an additional topological factor $2^{C_{L_R}}$ that enhances the probability of having a larger number of disconnected loops than in the Abelian case. In fact, this observation is key to understand the robustness of the $|D_4\rangle$ ground state to the proliferation of m_R fluxes.

In general, $O(N)$ loop models with *loop weight* $N \in [-2, 2]$ showcase two different phases separated by a critical point at the critical tension $t_c(N) = (2 + \sqrt{2 - N})^{-1/2}$ [90]: a dilute (or small loop) phase for $t < t_c(N)$ and a dense phase for $t > t_c(N)$. For $N = 2$, the critical point at $t_c(2) = 1/\sqrt{2}$ is described by a BKT transition which extends into an extended gapless phase described by a Luttinger liquid with central charge $c = 1$ (see also Ref. [91]). Since $t_R \leq 1/\sqrt{2}$ for all $\beta_R^x \in [0, \infty)$, we find that $|D_4\rangle$ is *robust to arbitrary large deformations of the type given in Eq. (19)* and turns into a critical state in the projective limit $\beta_R^x = \infty$ (where each red qubit is projected into $X = 1$). In summary, its phase diagram is



B. From loop to local spin models

As in the Abelian case, it is useful to characterize the system by a local statistical-mechanics model, as it can reveal hidden symmetries of the system and also provides insight into the decohered case. As before, such a model can be derived by applying the ungauging maps directly at the level of the wave function. In Sec. II B, we saw how the D_4 state maps to a honeycomb toric code on the red lattice. Using the constraint $A_s = \prod_{j \in \partial s} X_j = +1$ on each plaquette, we can, moreover, dualize this toric code to a trivial paramagnet on the vertices of the honeycomb lattice. This procedure leads to [compare to Eq. (3)]

$$\begin{aligned} |\text{TC}\rangle_{\mathcal{R}_R} &\rightarrow \bigotimes_v |+\rangle_v, \\ X_\ell^R \widetilde{CZ}_{b,g} &\rightarrow Z_b^R Z_g^R \widetilde{CZ}_{b,g}. \end{aligned} \quad (28)$$

Hence, through these mappings, the original kagome lattice is mapped to a bilayer honeycomb lattice with red $\sigma = \pm 1$ Ising variables on the top layer and σ_b and σ_g on the bottom one with $g \in G$ and $b \in B$ denoting the two different sublattices [see Fig. 4(a)].

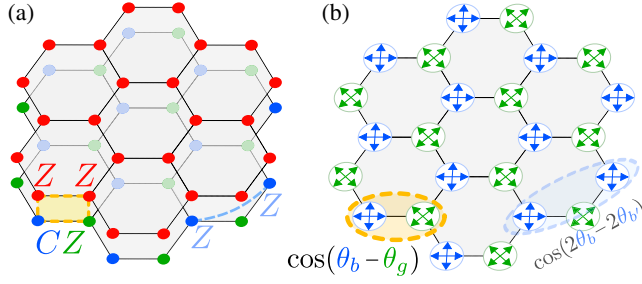


FIG. 4. Explicitly local stat-mech models for the deformed wave function. Formulation of the explicitly local stat-mech models in Secs. IV B and VA to characterize the phase diagram of the deformed D_4 wave function with non-Abelian m_R anyons. (a) corresponds to the formulation in terms of Ising variables on two honeycomb layers, while (b) is given in terms of a single-layer four-state clock model.

In summary, we can write the partition function obtained from the self-overlap of the deformed $|\mathcal{D}_4\rangle$ wave function (19) as an *Ising-like* partition function:

$$\mathcal{Z}_{|\psi\rangle}(\beta_R^x) = \sum_{\{\sigma, \tilde{\sigma}\}} e^{-\beta_R^x H_{|\psi\rangle}(\sigma, \tilde{\sigma})}, \quad (29)$$

with the bilayer honeycomb Hamiltonian

$$H_{|\psi\rangle} = - \sum_{\langle i,j \rangle_{\circ}} \sigma_i \sigma_j \tilde{C}Z_{ij}, \quad (30)$$

where $\tilde{C}Z_{ij} = \frac{1}{2}(1 + \tilde{\sigma}_i + \tilde{\sigma}_j - \tilde{\sigma}_i \tilde{\sigma}_j) \in \{-1, 1\}$; the interaction is shown in Fig. 4(a). It can be shown that an exact high-temperature expansion gives a closed-loop model on the honeycomb model [139]—similar to the Ising case—but the dressing on the bottom layer $\tilde{C}Z_{ij}$ leads to an additional loop factor given by Eq. (24), which, thus, gives an $O(2)$ loop model in agreement with Eq. (26).

Crucially, the spin Hamiltonian (30) has an internal D_4 symmetry. To make this symmetry manifest, it is useful to repackage this honeycomb *bilayer* of Ising spins as a *single* honeycomb lattice with *two-component* spins [140]:

$$H_{|\psi\rangle} = -\sqrt{2} \sum_{\langle b,g \rangle} \mathbf{n}_b \cdot \mathbf{n}_g = -\sqrt{2} \sum_{\langle b,g \rangle} \cos(\theta_b - \theta_g), \quad (31)$$

where

$$\mathbf{n}_b = \frac{1}{2} \begin{pmatrix} \sigma_b + \sigma_b \tilde{\sigma}_b \\ \sigma_b - \sigma_b \tilde{\sigma}_b \end{pmatrix} \quad \text{and} \quad \mathbf{n}_g = \frac{1}{\sqrt{2}} \begin{pmatrix} \sigma_g \\ \sigma_g \tilde{\sigma}_g \end{pmatrix} \quad (32)$$

has been used in the first equality. Moreover, since these vectors are normalized (i.e., $|\mathbf{n}_b| = |\mathbf{n}_g| = 1$), we can interpret them as taking values on the unit circle, although their discrete values depend on whether we are on the blue or green sublattice:

$$\mathbf{n}_b \in \begin{pmatrix} \circlearrowleft \\ \circlearrowright \end{pmatrix} \quad \text{and} \quad \mathbf{n}_g \in \begin{pmatrix} \circlearrowleft \\ \circlearrowright \end{pmatrix}. \quad (33)$$

Hence, we use the angle variables

$$\mathbf{n}_b = \begin{pmatrix} \cos(\theta_b) \\ \sin(\theta_b) \end{pmatrix} \quad \text{and} \quad \mathbf{n}_g = \begin{pmatrix} \cos(\theta_g) \\ \sin(\theta_g) \end{pmatrix} \quad (34)$$

in the second equality of Eq. (31). These representations make clear that $H_{|\psi\rangle}$ is symmetric under an internal $D_4 \cong \mathbb{Z}_4 \rtimes \mathbb{Z}_2 \subset O(2)$ symmetry. More precisely, $\mathbf{n}_b \cdot \mathbf{n}_g$ is invariant if each \mathbf{n}_j transforms under the symmetry group of the square, $D_4 \cong \langle R, S | R^4 = S^2 = 1, SRS = R^{-1} \rangle$, where R rotates the vector counterclockwise by 90° and S mirrors top and bottom:



On the two-component vector, R is represented by $-i\sigma^y$ and S by σ^z , defining a two-dimensional irrep of D_4 . From Eq. (32), we infer that the D_4 symmetry acts as follows in the original variables of Eq. (30):

$$S: \begin{cases} \tilde{\sigma}_g \rightarrow -\tilde{\sigma}_g, \\ \sigma_b \rightarrow \sigma_b \tilde{\sigma}_b, \end{cases}, \quad RS: \begin{cases} \tilde{\sigma}_b \rightarrow -\tilde{\sigma}_b, \\ \sigma_g \rightarrow \sigma_g \tilde{\sigma}_g, \end{cases} \quad (35)$$

where $R = (RS) \times S$.

In addition to this internal D_4 symmetry, the model has a spatial \mathbb{Z}_2 symmetry, denoted M , that swaps the blue and green sublattices in Eq. (30). This operation acts as a Hadamard matrix on the two-component vector, which geometrically corresponds to a mirror across an axis at 22.5° . Interestingly, MS acts as an effective 45° rotation. Hence, the internal and external symmetries combine into a $D_8 \cong \mathbb{Z}_8 \rtimes \mathbb{Z}_2$ symmetry. This enlarged symmetry is important for understanding the stability of a critical phase upon perturbing this model.

Lastly, we note that Eq. (32) also makes apparent that we can write the model as a two-body Ising spin model by performing a change of basis $\sigma_i \rightarrow \sigma_i$, and $\tilde{\sigma}_i \rightarrow \sigma_i \tilde{\sigma}_i$ on every site. We then obtain

$$H_{|\psi\rangle} = -\frac{1}{2} \sum_{\langle i,j \rangle_{\circ}} (\sigma_i \sigma_j + \sigma_i \tilde{\sigma}_j + \tilde{\sigma}_i \sigma_j - \tilde{\sigma}_i \tilde{\sigma}_j). \quad (36)$$

We have found that this representation is especially useful for Monte Carlo simulations of (perturbations of this) model (see Appendix G), as we explore later. On the other hand, the rotor representation (31) proves rather useful for more analytic arguments.

C. Decohered mixed state from purity

The above detailed study of the wave-function-deformed case is useful for understanding the effect of decohering D_4 TO with non-Abelian anyons, which we turn to now. The composition of local quantum channels of the form $\mathcal{E}_j^X(\rho_0) = (1 - p_R)\rho_0 + p_R X_j \rho_0 X_j$ acting on the red sublattice \mathcal{R}_R leads to the proliferation of incoherent m_R fluxes. Similarly to the Abelian case (Sec. III), we first study the purity $\text{tr}(\rho^2)$ of the decohered density matrix $\rho = \prod_j \mathcal{E}_j^X(\rho_0)$. For this, we again turn to the vectorized representation of ρ :

$$|\rho\rangle\rangle = \prod_{j \in \mathcal{R}_R} (1 - p_R + p_R X_j \otimes X_j) |\rho_0\rangle\rangle \\ \propto e^{\mu_R \sum_{j \in \mathcal{R}_R} X_j \otimes X_j} |\rho_0\rangle\rangle \quad (37)$$

with $\tanh(\mu_R) = p_R/(1 - p_R)$ and $|\rho_0\rangle\rangle = |D_4\rangle \otimes |D_4\rangle$. Since $\text{tr}(\rho^2) = \langle\langle \rho | \rho \rangle\rangle$, we can repeat the derivation in the previous subsection, which again leads to a honeycomb loop model, but now $f(L_R)^2$ appears, rather than just $f(L_R)$ as given in Eq. (25). As a result, we now obtain an $O(4)$ honeycomb loop model rather than an $O(2)$ loop model:

$$\text{tr}(\rho^2) \stackrel{\text{decohere}}{\propto_{m_R}} \sum_{L_R} \left(\frac{r_R}{2}\right)^{|L_R|} 4^{C_{L_R}}, \quad (38)$$

with $r_R = [2p_R(1 - p_R)]/[p_R^2 + (1 - p_R)^2] \in [0, 1]$. The fact that the loop weight N is given by the square of the quantum dimension of m_R , i.e., $N = d_a^2$, instead of by $N = d_a$, is understood from the fact that such a local error creates pairs of fluxes $m_R \bar{m}_R$ combining the bra and ket subspaces together, similar to the toric code case [5]. This dependence is clarified when we consider higher moments $\text{tr}(\rho^n)$ of the decohered density matrix.

Unlike the $O(2)$ loop model, the $O(N)$ honeycomb loop model does not proliferate to a large loop phase [142] for $N > 2$. Hence, as far as the purity is concerned, decohering D_4 topological order with m_R anyons does *not* lead to a transition, even if $p = \frac{1}{2}$. It is important to notice that the decohered density matrix $\rho_{1/2} = \mathcal{E}^X(\rho_0)$ when taking the error rate $p_R = \frac{1}{2}$ is a fixed point of the quantum channel \mathcal{E}^X for any p_R , i.e., $\mathcal{E}^X(\rho_{1/2}) = \rho_{1/2}$ for any p_R [144]. Hence, a single application of the channel with $p_R = \frac{1}{2}$ is sufficient to reach its fixed point. This fact in combination with the lack of singularity in $\text{tr}(\rho^2)$ indicates that D_4 topological order is infinitely robust under an X -decoherence channel acting on one of the sublattices, *as far as the purity is concerned*. However, as we mentioned in the Introduction, the critical threshold $p_c^{(n)}$ depends on the moment $\text{tr}(\rho^n)$ of the density matrix that is considered, with the “intrinsic” (i.e., quantum-information-theoretic) threshold p_c appearing when evaluating the Von Neumann

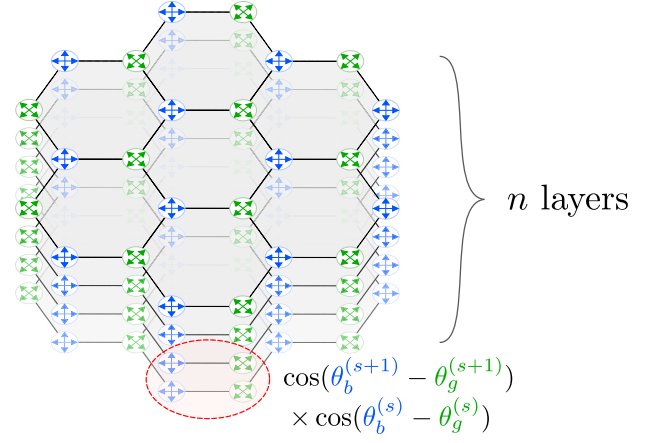


FIG. 5. Stat-mech model for $\text{tr}(\rho^n)$. The n th moment of the density matrix obtained by decohering D_4 TO with incoherent non-Abelian m_R anyons leads to an n -fold stack of honeycomb rotor models, each layer with a discrete D_4 symmetry acting on the four-state rotors. The Hamiltonian of the stat-mech model has a nearest-layer interaction (49), as shown in the figure. Alternatively, instead of four-state rotors, one can also express each layer as an Ising bilayer as in Eq. (50). These interactions resemble those encountered in the wave-function deformation in Secs. IV A and IV B. The case of purity ($n = 2$) is studied in detail in Sec. IV C.

entropy (formally, $n \rightarrow 1$). In the following sections, we find indications of robustness also for other values of n , although we find a transition for $n = \infty$.

Similarly to the previous section for the pure wave-function deformation, one can rewrite the resulting loop model in terms of an explicitly local Ising-like Hamiltonian:

$$\text{tr}(\rho^2) \stackrel{\text{decohere}}{\propto_{m_R}} \sum e^{-2\mu_R H_\rho}. \quad (39)$$

In this case, due to the doubling of degrees of freedom, we find a *tetralayer* (rather than bilayer) honeycomb model, governed by a Hamiltonian of the form

$$H_\rho = - \sum_{\langle i,j \rangle_\circ} \sigma_i^1 \sigma_j^1 \widetilde{C} Z_{ij}^{(2)} \sigma_i^3 \sigma_j^3 \widetilde{C} Z_{ij}^{(4)}, \quad (40)$$

with $\tanh(\mu_R) = [p_R/(1 - p_R)]$. Here, σ^1 and $\tilde{\sigma}_i^2, \tilde{\sigma}_i^2$ live on the first and second layer, respectively [with the second set living on the two different sublattices as in Eq. (30)], while σ^3 and $\tilde{\sigma}_i^4, \tilde{\sigma}_i^4$ live on layers 3 and 4 (see Fig. 5 expressed in terms of rotor variables). However, one can effectively reduce the number of layers from four to three by defining a new variable $\sigma_i^1 \rightarrow \sigma_i = \sigma_i^1 \sigma_i^3$. The Hamiltonian then simplifies to a *trilayer* honeycomb model with interactions

$$H_\rho = - \sum_{(i,j)_\circ} \sigma_i \sigma_j \widetilde{CZ}_{ij}^{(2)} \widetilde{CZ}_{ij}^{(3)}. \quad (41)$$

This has at least two independent D_4 symmetries involving the pair of layers (1, 2) and (1, 3).

We can use this statistical mechanical model to evaluate specific information-theoretic diagnostics—or, rather, simpler-to-compute Rényi versions—of the stability of this topological order, as Ref. [6] did for the toric code. In particular, let us consider the Rényi-2 quantum relative entropy measuring the distinguishability of two states ρ (namely, the decohered D_4 TO state) and ρ_{xy} (to be specified later on). This quantity is defined as

$$D^{(2)}(\rho || \rho_{xy}) \equiv - \log \left(\frac{\text{tr}(\rho \rho_{xy})}{\text{tr}(\rho^2)} \right) \quad (42)$$

and diverges when ρ and σ are orthogonal (hence, distinguishable), while it saturates to a finite value when the two corrupted states cannot be distinguished. Since anyonic excitations are a defining characteristic of TO, checking whether they are still well defined (i.e., orthogonal to the vacuum) for the decohered D_4 TO is relevant. Hence, we take ρ_{xy} to be the decohered state after applying the quantum channel \mathcal{E}^X to the initial state $\mathcal{X}_x^y |D_4\rangle$, hosting two non-Abelian anyons m_R located on triangles x and y ; see Eq. (2). Using the same ungauging maps previously employed in this section (in particular, Appendix B), the Rényi-2 quantum relative entropy becomes the thermal correlation function

$$D^{(2)}(\rho || \rho_{xy}) = - \ln \left(\langle \sigma_x \sigma_y \widetilde{CZ}_{xy}^{(2)} \widetilde{CZ}_{xy}^{(3)} \rangle \right) \quad (43)$$

evaluated on the stat-mech model of Eq. (41). Since for all error rates the resulting $O(4)$ loop model lies within the small loop phase, $\langle \sigma_x \sigma_y \widetilde{CZ}_{xy}^{(2)} \widetilde{CZ}_{xy}^{(3)} \rangle$ decays exponentially with the distance and, hence, $D^{(2)}(\rho || \rho_{xy}) \sim |y - x|$ diverges with the distance between the two m_R anyons.

The previous discussion, in particular, the content of Eq. (38), appears to be at odds with the fact that fusion of m_R anyons can also lead to Abelian charges $e_{G,B}$. In particular, one might wonder why additional contributions coming from such anyons are absent from the partition function $\mathcal{Z}_\rho(p_R)$. More plainly put, why do we obtain clean loop models where $m_R \times m_R$ always fuses trivially? This property relates to the particular choice of quantum channel which proliferates only m_R (which is discussed more in later sections), but it also reflects the fixed-point nature of the initial D_4 wave function we are considering. In Sec. V, we study the deformation and decoherence proliferating both Abelian and non-Abelian anyons. Let us state a key result here, in the context of X error on top of the deformed wave function we already introduced in Sec. III A for perturbing with Abelian anyons. First, the relevant stat-

mech model corresponds to a *net* (rather than loop) model, with closed red loops L_R together with blue γ_B and green γ_G closed loops or strings ending on L_R loop configurations [see Fig. 1(b)]. For the pure wave-function deformation, we again derive Eq. (21), however now with

$$f(L_R; t_G, t_B) = f(L_R) \times \mathcal{Z}_{L_R}(t_G, t_B), \quad (44)$$

where $f(L_R)$ was given in Eq. (25). Here, $t_{G,B} = \tanh(\beta_{G,B}^z)$, and $\mathcal{Z}_{L_R}(t_G, t_B)$ is given by

$$\mathcal{Z}_{L_R}(t_G, t_B) = \sum_{\gamma_G, \gamma_B} \sigma_{L_R}(\gamma_B, \gamma_G) t_G^{|\gamma_G|} t_B^{|\gamma_B|}. \quad (45)$$

The symbol $\sigma_{L_R}(\gamma_B, \gamma_G) = \pm 1$ is a sign assignment which is consistent with anyon braiding properties and is specified in Appendix C 2. Hence, Abelian charges $e_{G,B}$ are generically generated when detuning from the fixed-point wave function. Notice that, when either t_G or t_B vanish, $\mathcal{Z}_{L_R}(t_G, t_B)$ becomes positive, and, hence, it corresponds to a classical partition function for every L_R .

Computing the purity $\text{tr}(\rho^2)$ involves a similar calculation where the main difference is replacing the topological factor in Eq. (44), by its square: $f(L_R; t_G, t_B) \rightarrow f(L_R; t_G, t_B)^2$. Hence, the purity of the decohered density matrix reads

$$\text{tr}(\rho^2) \stackrel{\text{decohere}}{\propto}_{m_R} \sum_{L_R} \left(\frac{r_R}{2} \right)^{|L_R|} 4^{C_{L_R}} \mathcal{Z}_{L_R}^2(t_G, t_B). \quad (46)$$

This expression simplifies when both $t_G, t_B \ll 1$. First-order contributions in these tensions are given by the shortest configurations γ_G and γ_B . These correspond either to $|\gamma_G| = |\gamma_B| = 0$ or to a string with the shortest possible length of either blue or green errors (namely, with $|\gamma_G|, |\gamma_B| = 1$) whose end points terminate on a closed-loop configuration L_R . An example of such configuration is shown in Fig. 1(b). For a given L_R , the partition function $\mathcal{Z}_{L_R}(t_G, t_B)$ then becomes $\mathcal{Z}_{L_R}(t_G, t_B) \approx 1 + (|L_R|/2)(t_G + t_B) \approx [1 + (t_G + t_B)/2]^{|L_R|}$, which can be interpreted as dressing the loop L_R . Hence, we find that the purity

$$\text{tr}(\rho^2) \stackrel{\text{decohere}}{\propto}_{m_R} \sum_{L_R} \left(\frac{r_R(2 + t_G + t_B)}{4} \right)^{|L_R|} 4^{C_{L_R}} \quad (47)$$

is again given by an $O(4)$ honeycomb loop model where the tension is tuned by r_R as well as by the tensions t_G and t_B . This proves useful later for detuning away from accidentally fine-tuned points.

D. Higher moments $\text{tr}(\rho^n)$

So far, we characterized the decohered mixed state ρ in terms of quantities that can be computed from its second

moments [i.e., evaluating the expectation value of the density matrix $\rho^2/\text{tr}(\rho^2)$]. We can similarly derive loop model descriptions of the higher moments $\text{tr}(\rho^n)$. This is most convenient to do in terms of the so-called error picture [6,56], where we observe that we can write ρ as a classical mixture of corrupted states $\prod_j X_{r_j}|D_4\rangle$. Then, $\text{tr}(\rho^n)$ is naturally expressed in terms of overlaps of such corrupted wave functions, similarly to what we saw in Secs. IV A and IV C. Since we have already obtained a general expression for this wave-function overlap in Eq. (25), we straightforwardly obtain stat-mech models for these higher moments. In particular, since Eq. (25) took the form of the weight of an O(2) loop model, we naturally write $\text{tr}(\rho^n)$ as $\sim n$ coupled O(2) loop models. Indeed, the O(4) loop model we obtained for $n = 2$ can be thought of as two tightly bound O(2) loop models. Similarly, $\text{tr}(\rho^3)$ admits a particularly simple expression as a coupled O(2) loop model:

$$\text{tr}(\rho^3) \propto \sum_{L^{(1)}, L^{(2)}} \prod_{L \in \{L^{(1)}, L^{(2)}, L^{(1)} \oplus L^{(2)}\}} \tilde{t}^{|L|/2} C_L, \quad (48)$$

with $\tilde{t}^2 = [(p - p^2)/(2 - 6p + 6p^2)]$ and where the sums are over contractible closed-loop configurations $L^{(1)}$ and $L^{(2)}$ on the honeycomb lattice. This can be thought of as three O(2) loop models, with a strong coupling that enforces the symmetric difference of all three loops to vanish. The coupled loop models for higher n are summarized in Appendix D, although they take a more complicated form. In a companion work, we show how such coupled O(N) loop models arise more generally for other types of topological order, with the loop weight N corresponding to the quantum dimension of the proliferating anyon [92].

However, to the best of our knowledge, the physics of these stat-mech loop models has not yet been explored in the literature, in contrast to the O(2) and O(4) loop models discussed above. For this reason, we focus on the local spin model representation of these stat-mech models. In particular, similar to the previous section, we can show that $\text{tr}(\rho^n)$ can be written in terms of n coupled “ZZCZ” models:

$$\text{tr}(\rho^n) \propto \sum_{\{\sigma_j^{(s)}, \tilde{\sigma}_j^{(s)}\}_{s=1}^n} \exp \left(\beta \sum_{\langle i,j \rangle} \sum_{s=1}^n h_{i,j}^{(s)} h_{i,j}^{(s+1)} \right), \quad (49)$$

where $\tanh(\beta) = [p_R/(1 - p_R)] \in [0, 1]$ and each replica is a Ising honeycomb bilayer with

$$h_{i,j}^{(s)} = \sigma_i^{(s)} \sigma_j^{(s)} \widetilde{\text{CZ}}_{ij}^{(s)} \quad \text{and} \quad h_{i,j}^{(n+1)} = h_{i,j}^{(1)}. \quad (50)$$

This is direct generalization of the $n = 2$ case in Eq. (39). This is derived in Appendix D using the error picture.

There, we also show an equivalent rewriting, which reduces to the known result for the toric code case [6] if we were to set $\widetilde{\text{CZ}}_{ij} \rightarrow 1$. We note that $h_{i,j}^{(s)}$ is the same interaction we encountered in the section on wave-function deformation [Secs. IV A and IV B, in particular, Eq. (30)]. One can, thus, also express it in terms of rotors as in Eq. (31), to make the internal D_4 symmetry manifest. In other words, $\text{tr}(\rho^n)$ can be regarded as a local coupling of n layers of rotors, each with D_4 symmetry. We visually represent the model in Fig. 5.

We have already seen that for $n = 2$ (i.e., purity), the local stat-mech model fails to order, even at zero temperature. This tells us that the purity does not detect any transition out of D_4 TO. If we increase n and thereby add more layers to our stat-mech model (49), can this increase the tendency to order? To answer this question, it is instructive to consider the zero-temperature limit $\beta \rightarrow \infty$, more precisely, setting $p_R = \frac{1}{2}$, where the model has the highest chance of ordering. Note that, in this limit, Eq. (49) has nonzero (and equal) weight only for configurations where $h_{i,j}^{(s)} h_{i,j}^{(s+1)} = 1$. In other words, for the maximal decoherence rate $p_R = \frac{1}{2}$, the resulting stat-mech model is the infinite-temperature ensemble in the constrained configuration space where $h_{i,j}^{(s)} = h_{i,j}^{(s')} \equiv \eta_{ij} \in \{\pm 1\}$. Hence,

$$\text{tr}(\rho^n)_{p_R \rightarrow 1/2} \propto \sum_{\{\eta_{ij}\}} \left(\sum_{\{\sigma_j, \tilde{\sigma}_j\}} \prod_{\langle i,j \rangle} [1 + \eta_{ij} \sigma_i \sigma_j \widetilde{\text{CZ}}_{ij}] \right)^n. \quad (51)$$

Considerable insight can be gained from Eq. (51). For instance, it indicates the fate of the $n \rightarrow \infty$ limit, which is where the spin model (49) becomes a full-fledged 3D stat-mech model. Taking this limit in Eq. (51) means only that the $\{\eta_{ij}\}$ configuration leading to the largest weight will survive. This turns out to be $\eta_{ij} = 1$, as we prove in the next subsection [145]. Hence,

$$\lim_{n \rightarrow \infty} \sqrt[n]{\text{tr}(\rho^n)}_{p_R \rightarrow 1/2} \propto \sum_{\{\sigma_i, \tilde{\sigma}_j\}} \prod_{\langle i,j \rangle} [1 + \sigma_i \sigma_j \widetilde{\text{CZ}}_{ij}]. \quad (52)$$

This is *exactly* the zero-temperature limit of the stat-mech model we encountered when studying the pure wave-function deformation in Secs. IV A and IV B. There, we found that this model is perched at BKT criticality, with algebraic correlations. This shows that, unlike for $n = 2$, there exists a critical threshold $p_c^{(\infty)}$ for the $n \rightarrow \infty$ replica limit whose value is $p_c^{(\infty)} \leq \frac{1}{2}$ (in fact, in the next subsection, we argue $p_c^{(\infty)} = \frac{1}{2}$) [146]. In particular, this shows that $p_c^{(n)}$ is not an increasing function of n , unlike for the decohered toric code case [6]. This distinct behavior can be linked to the topological weights in the loop model

representation: For $n = 2$ the strong coupling of the two layers led to an $O(4)$ loop model which is unable to proliferate, but for $n \rightarrow \infty$ the effects of the coupling are suppressed, revealing the critical point of the $O(2)$ loop model, of which the local spin model in Eq. (52) is an equivalent representation.

The result in Eq. (51) even allows us to infer properties of the “true” threshold. Indeed, the form is amenable to taking the replica limit $n \rightarrow 1$. From this, one can infer that the Von Neumann entropy (at $p_R = \frac{1}{2}$) is dictated by a random-bond version of the ZZCZ model, which we discuss more rigorously and in detail in the next subsection. This suggests that the fate of D_4 as a quantum memory subjected to decoherence is tied up with whether or not this disordered spin model orders—and we have already established that even the clean model reaches quasi-long-range order only at zero temperature. However, we can avoid having to take the replica limit altogether: We now show how one can fully diagonalize ρ for $p_R = \frac{1}{2}$, which means we have access not only to the full spectrum, but also to all eigenstates, which, in turn, gives insight to how these results inform about $p_R < \frac{1}{2}$.

E. Solving maximal decoherence $p_R = 1/2$

We have seen that, unlike decohering with Abelian anyons, TO can be remarkably stable to decohering with non-Abelian anyons. For instance, the purity sees no transition at all upon maximally decohering m_R in D_4 . This raises the question: Can the TO be stable even at maximal decoherence to the proliferation of a given (set of) anyons? Here, we show how one can shed light on this question by exactly diagonalizing the decohered density matrix $\rho_{1/2}$ at maximum error rate $p_R = \frac{1}{2}$ in terms of a disordered spin model.

To diagonalize $\rho_{1/2}$, we notice that the local quantum channel \mathcal{E}_r^X (for $p_R = \frac{1}{2}$) acting on a site r of \mathcal{R}_R can be written as a random projector channel:

$$\mathcal{E}_r^X(\cdot) = \sum_{\eta_r = \pm 1} P_{\eta_r}(\cdot) P_{\eta_r}, \quad (53)$$

with $P_{\eta_r} = \frac{1}{2}(\mathbb{1} + \eta_r X_r)$. Hence, acting on $|D_4\rangle$ and on all sites of the red sublattice leads to the maximally decohered density matrix

$$\rho_{1/2} = \frac{1}{4^{|\mathcal{R}_R|}} \sum_{\eta} |\eta\rangle \langle \eta| \quad (54)$$

with $\eta = \{\eta_r = \pm 1\}$ and where the (unnormalized) states $|\eta\rangle$ associated to a given configuration η are given by

$$|\eta\rangle = \prod_{r \in \mathcal{R}_R} (1 + \eta_r X_r) |D_4\rangle. \quad (55)$$

The orthogonality of the projectors P_{s_r} directly leads to the orthogonality condition $\langle \eta | \eta' \rangle = \langle \eta | \eta \rangle \prod_{r \in \mathcal{R}_R} \delta_{\eta_r, \eta'_r}$; hence, the states $|\eta\rangle$ correspond to eigenvectors of $\rho_{1/2}$. In other words, Eq. (54) constitutes a diagonalized density matrix. The corresponding (non-negative) eigenvalues are then given by $P(\eta) = \langle \eta | \eta \rangle / 4^{|\mathcal{R}_R|}$, which more explicitly reads

$$P(\eta) = \frac{1}{4^{|\mathcal{R}_R|}} \sum_{L_R} \left(\prod_{e \in L_R} \frac{\eta_e}{\sqrt{2}} \right) 2^{C_{L_R}}, \quad (56)$$

where we have explicitly used that $f(L_R) = 2^{C_{L_R}} / \sqrt{2}^{|L_R|}$ [see Eq. (25)]. Therefore, the eigenvalues $P(\eta)$ correspond to the partition function of a random $O(2)$ loop model with a signed-disorder tension given by $\eta_e / \sqrt{2}$ on a bond e of the honeycomb lattice.

Let us now consider the stat-mech model appearing in the limit $\lim_{n \rightarrow \infty} \text{tr}(\rho^n)^{1/n}$ which equals to $\max_{\eta} P(\eta)$ (up to a system-size-dependent constant factor related to its degeneracy). Since $f(L_R) \geq 0$, one finds the largest eigenvalue to be

$$\max_{\eta} P(\eta) = \frac{1}{4^{|\mathcal{R}_R|}} \sum_{L_R} \frac{2^{C_{L_R}}}{\sqrt{2}^{|L_R|}}, \quad (57)$$

which is attained for all eigenstates for which $\prod_{e \in \square} \eta_r = +1$ for all plaquettes \square . This provides a derivation of Eq. (52), which is written as the local stat-mech model formulation of the $O(2)$ loop model at the BKT critical point. As discussed in the previous subsection, this shows that, unlike the case for purity ($n = 2$), there is a $p_c^{(\infty)} \leq \frac{1}{2}$. In fact, we now argue $p_c^{(\infty)} = \frac{1}{2}$.

The previous calculation can be repeated when considering the deformed wave function $|\psi(\beta_G^z, \beta_B^z)\rangle$ instead of $|D_4\rangle$ (see Sec. IV C, where we discuss the case of purity). In this case, the eigenvalues become

$$P(\eta) = \frac{1}{4^{|\mathcal{R}_R|}} \sum_{L_R} \left(\prod_{e \in L_R} \frac{\eta_e}{\sqrt{2}} \right) 2^{C_{L_R}} \frac{\mathcal{Z}_{L_R}(t_G, t_B)}{\mathcal{Z}_{\emptyset}(t_G, t_B)}, \quad (58)$$

where $\mathcal{Z}_{\emptyset}(t_G, t_B) \geq 0$ is the norm of $|\psi(\beta_G^z, \beta_B^z)\rangle$, which agrees with Eq. (45) for the trivial loop configuration $L_R = \emptyset$. Moreover, whenever $\mathcal{Z}_{L_R}(t_G, t_B) \geq 0$, one finds the largest eigenvalue is given by

$$\max_{\eta} P(\eta) = \frac{1}{4^{|\mathcal{R}_R|}} \sum_{L_R} \frac{2^{C_{L_R}}}{\sqrt{2}^{|L_R|}} \frac{\mathcal{Z}_{L_R}(t_G, t_B)}{\mathcal{Z}_{\emptyset}(t_G, t_B)}. \quad (59)$$

The condition $\mathcal{Z}_{L_R}(t_G, t_B) \geq 0$ for all L_R is clearly satisfied when either t_G or t_B vanish [since in this case $\sigma_{L_R}(\gamma_G, \gamma_B)$ in Eq. (45) is positive] or when both of them are positive but

sufficiently small as shown at the end of Sec. IV C. In this case, we then find that

$$\lim_{n \rightarrow \infty} \text{tr}(\rho^n)^{1/n} \approx \sum_{L_R} 2^{C_{L_R}} \left(\frac{t_R}{\sqrt{2}} \right)^{|L_R|} \quad (60)$$

corresponds to the partition function of the O(2) loop model on the honeycomb lattice with (tunable) tension $t_R \approx 1 + [(t_B + t_G)/2]$. Therefore, for $t_B, t_G > 0$, the system lies within the extended gapless phase; attains a critical point for $t_B = t_G = 0$; and finally lies within the shortly correlated phase for negative t_B, t_G . This means that any infinitesimal negative value of t_B or t_G is sufficient to prevent there being a transition for $n = \infty$, similar to the purity case ($n = 2$). If we presume $p_c^{(\infty)}$ is a continuous function of t_B, t_G , this suggests that, for $t_B = 0 = t_G$, we have $p_c^{(\infty)} = \frac{1}{2}$. However, this is based on the aforementioned perturbative expansion, the validity of which would be interesting to explore in future work.

Knowledge of the full spectrum of $\rho_{1/2}$ allows us to compute various information-theoretic quantities which relate to the underlying random O(2) loop model stat-mech model. To do so, it is desirable to express the eigenvalues $P(\eta)$ in Eq. (58) as the partition function of a local stat-mech model with *non-negative Boltzmann weights*. In particular, (up to η -independent overall factors) Eq. (58) can be rewritten as

$$P(\eta) \propto \sum_{\{\sigma, \tilde{\sigma}\}} \prod_{\langle i, j \rangle} (1 + \eta_{ij} \sigma_i \sigma_j \widetilde{CZ}_{ij}) \times e^{\beta_B^z \sum_{\langle b, b' \rangle} \tilde{\sigma}_b \tilde{\sigma}_{b'} + \beta_G^z \sum_{\langle g, g' \rangle} \tilde{\sigma}_g \tilde{\sigma}_{g'}}, \quad (61)$$

defined on a bilayer honeycomb lattice as in Fig. 4(a) [recall that $t_G = \tanh(\beta_G^z)$ and $t_B = \tanh(\beta_B^z)$]. Note that this agrees with Eq. (51) for $\beta_B^z = \beta_G^z = 0$. The Von Neumann entropy of $\rho_{1/2}$ then corresponds to the quenched disorder average of the free energy of the partition function in Eq. (61):

$$S_{\text{VN}}(\rho_{1/2}) = - \sum_{\eta} P(\eta) \ln[P(\eta)]. \quad (62)$$

The fact that the probability of a given disorder configuration η is proportional to the partition function $P(\eta)$ itself is an example of a Nishimori condition [127].

Equation (62) suggests that determining whether the D_4 topological quantum memory persists to the maximal-decoherence limit is equivalent to asking whether this disordered ZZCZ model remains in the paramagnetic phase [147]. Indeed, since we have full access to the *eigenstates* of $\rho_{1/2}$, one can confirm that the fidelity between the decohered copies of two (initially) distinct logical states [148] of the D_4 can be related to a thermodynamic quantity

of the disordered spin model. Writing $\rho = \mathcal{E}^X(|D_4\rangle\langle D_4|)$ and $\sigma = \mathcal{E}^X(\mathcal{X}|D_4\rangle\langle D_4|\mathcal{X})$ with \mathcal{X} the logical operator as defined below Eq. (2), acting on a noncontractible loop (e.g., the horizontal direction) around the torus, the quantum fidelity between these two quantities reads

$$F(\rho, \sigma) \equiv \text{tr} \left(\sqrt{\sqrt{\rho} \sigma \sqrt{\rho}} \right) \propto \frac{1}{2} \sum_{\{\eta\}} P(\eta) \frac{|1 - e^{-\Delta F_C}|}{1 + e^{-\Delta F_C}}, \quad (63)$$

where ΔF_C is the free energy difference of inserting a line of flipped (antiferromagnetic) bonds along a noncontractible loop \mathcal{C} perpendicular to \mathcal{X} , i.e., with a symmetry defect line [149]. See details of the derivation in Appendix E. In the paramagnetic phase, this symmetry defect line is invisible such that $F(\rho, \sigma) \approx 0$ in the thermodynamic limit, whereas in the ordered phase the defect would frustrate the order $\Delta F_C \approx |\mathcal{C}|$, making the fidelity nonzero and, hence, signaling the breakdown of the (non-Abelian) quantum memory. Notice that the same condition, namely, the free energy cost of inserting a domain wall, is usually employed (although in the opposite direction with the ferromagnetic phase corresponding to a good quantum memory) to diagnose the breakdown of an Abelian quantum memory [55,56,150].

This then raises the question: What is the fate of the disordered spin model? We already discussed that in the clean case, i.e., for $\eta_{gb} = +1$ on every link, the system showcases quasi-long-range order only for $\beta_R^x = \infty$. It is then suggestive that if this “ferromagnetic” model fails to develop long-range order, the disordered one will follow the same fate. Hence, we expect that disorder will hinder even more the system from ordering, leading to lower ordering temperatures and, in turn, to the absence of a transition all together. This intuition appears to be supported by the observation that

$$P(\eta) \approx \sum_{L_R} 2^{C_{L_R}} \left(\prod_{e \in L_R} \frac{\eta_e t_R}{\sqrt{2}} \right), \quad (64)$$

appearing in the regime $t_G, t_B \ll 1$ for any disorder configuration η , resembles disordered XY models previously studied in the literature [151,152]. Notice that the clean versions of the XY model and that of the O(2) loop model in Eq. (64) share the same critical behavior. It has been found that, for the former, various forms of disorder cause the critical temperature separating a quasi-long-range order from a paramagnetic phase to decrease with increasing disorder strength. Hence, it is suggestive that a similar outcome holds for Eq. (64), in which case either $p_c^{(1)} = \frac{1}{2}$ or the D_4 quantum memory would be stable even in the maximal-decoherence limit (which seems especially likely if we allow ourselves to tune to negative t_G or t_B in the initial wave function, as discussed above). Nonetheless, we notice that, unlike the XY model, the system in Eq. (64) appears to

be highly frustrated, so we cannot exclude that disorder increases the critical temperature by partially alleviating the frustration (which would, in turn, imply a finite $p_c^{(1)}$). We leave a more careful analysis for future work.

V. COMBINED ABELIAN AND NON-ABELIAN ERRORS

We now combine both Abelian Z and non-Abelian errors X acting on different sublattices on the initial wave function $|D_4\rangle$ [153]. As the previous section has made explicit, the evaluation of the expectation values of the corresponding errors on the $|D_4\rangle$ TO ground state gives rise to a nontrivial coupling among different errors. In particular, we consider X acting on \mathcal{R}_R and Z acting on \mathcal{R}_B and \mathcal{R}_G with different tunable strengths.

A. Pure wave-function deformation

Let us consider the unnormalized wave function

$$|\psi(\beta_R^x, \beta_G^z = 0, \beta_B^z)\rangle = e^{(\beta_R^x/2) \sum_{j \in \mathcal{R}_R} X_j} e^{(\beta_B^z/2) \sum_{j \in \mathcal{R}_B} Z_j} |D_4\rangle. \quad (65)$$

One finds that its norm reads (again neglecting overall analytic prefactors)

$$\begin{aligned} \mathcal{Z}_{|\psi\rangle}(\beta_R^x, t_B) &= \sum_{L_R} \sum_{\gamma_B} \tanh(\beta_R^x)^{|L_R|} \\ &\times t_B^{|\gamma_B|} \langle D_4 | \prod_{j \in L_R} X_j \prod_{i \in \gamma_B} Z_i | D_4 \rangle, \end{aligned} \quad (66)$$

with $t_B = \tanh(\beta_B^z)$ and where now the errors are coupled through the last factor. As before, only closed red loop configurations L_R on the honeycomb lattice lead to a non-vanishing contribution (since \mathcal{R}_R and \mathcal{R}_B do not overlap). Deferring detailed calculations to Appendix C 2, we now explain the constraint on the possible configurations of γ_B (since it is no longer simply closed loops, we use γ_B rather than L_B to emphasize this). First, recall that, while the O(2) loop model describing the proliferation of non-Abelian fluxes is defined on the \mathcal{R}_R honeycomb lattice, the analogous formulation for Abelian e_B charges is given in terms of O(1) loop model on a triangular lattice (see Fig. 2). This corresponds to the dual triangular lattice of the \mathcal{R}_B honeycomb. The case where L_R and γ_B do not intersect reduces to the discussion in the previous section, with γ_B required to be a closed-loop configuration. However, since an L_R loop contains e_B Abelian anyons coming from the fusion channel $m_R \times m_R = 1 + e_B + e_G + e_B e_G$, a γ_B configuration can end on a L_R loop. Putting it all together, we find

$$\mathcal{Z}_{|\psi\rangle}(\beta_R^x, t_B) = \sum_{L_R} \sum_{\gamma_B} t_B^{|\gamma_B|} \left(\frac{\tanh(\beta_R^x)}{\sqrt{2}} \right)^{|L_R|} 2^{C_{L_R}}, \quad (67)$$

i.e., coupled O(2) and O(1) loop models where \sum_{γ_B} corresponds to the sum over all configurations in the triangular blue sublattice that either form closed loops or strings whose end points terminate on a L_R loop [see Fig. 1(b)].

As a next step, we add Abelian Z errors on the remaining \mathcal{R}_G sublattice. A similar calculation to the one above leads to the classical partition function

$$\begin{aligned} \mathcal{Z}_{|\psi\rangle}(\beta_R^x, t_G, t_B) &= \sum_{L_R} \sum_{\gamma_B} \sum_{\gamma_G} \sigma_{L_R}(\gamma_B, \gamma_G) \\ &\times t_B^{|\gamma_B|} t_G^{|\gamma_G|} \left(\frac{\tanh(\beta_R^x)}{\sqrt{2}} \right)^{|L_R|} 2^{C_{L_R}}, \end{aligned} \quad (68)$$

where now both γ_B and γ_G are loop configurations that can be either closed or end up on a L_R loop configuration on the blue and green triangular sublattices, respectively. This corresponds to two O(1) loop models (one per excited Abelian anyon type) coupled to a O(2) loop model describing the proliferation of non-Abelian charges. Moreover, unlike for vanishing β_G^z , there is an additional sign contribution $\sigma_{L_R}(\gamma_B, \gamma_G) = \pm 1$. Since in the following this sign turns out not to be necessary, we defer a more detailed discussion to Appendix C 2.

Similarly to Eq. (29), we can alternatively derive an Ising-like Hamiltonian which now reads (including the inverse temperature dependence)

$$\begin{aligned} H_{|\psi\rangle} &= -\beta_R^x \sum_{\langle i,j \rangle_\circ} \sigma_i \sigma_j \widetilde{CZ}_{i,j} - \beta_B^z \sum_{\langle\langle b,b' \rangle\rangle_\circ} \tilde{\sigma}_b \tilde{\sigma}_{b'} \\ &- \beta_G^z \sum_{\langle\langle g,g' \rangle\rangle_\circ} \tilde{\sigma}_g \tilde{\sigma}_{g'}, \end{aligned} \quad (69)$$

with $b, b'; g, g'$ sites on the blue B and green G sublattices of the bottom layer, respectively, as shown in Fig. 4(a). Notice that D_4 symmetry is enlarged to a D_8 symmetry group when $\beta_G^z = \beta_B^z$, due to the additional M mirror symmetry defined below Eq. (35).

1. Order parameters

The D_4 symmetry group has eight different proper nontrivial subgroups which correspond to either \mathbb{Z}_2 (appearing 5 times), \mathbb{Z}_4 (appearing once), or $\mathbb{Z}_2 \times \mathbb{Z}_2$ (appearing twice). These label the different possible thermal phases that can appear as a result of spontaneous symmetry breaking in the corresponding stat-mech model. The advantage of the explicitly local formulation in terms of Ising variables as in Eq. (69), relative to that in Eq. (68),

is to permit an efficient characterization of the system using Monte Carlo methods. In particular, one can numerically investigate the phase diagram via the three order parameters discussed in the following.

First, and as discussed in Sec. III A, condensation of e_B charges is measured by long-range order of the two-point correlation $\mathcal{W}_{e_B}(b, b') \equiv \langle \tilde{\sigma}_b \tilde{\sigma}_{b'} \rangle$ in the limit $|x - y| \rightarrow \infty$. It is useful to repackage this into a single number which is nonzero only in the condensed phase:

$$\mathcal{W}_{e_B} \equiv \frac{\sum_{b, b'} \mathcal{W}_{e_B}(b, b')}{\sum_{b, b'} 1} = \frac{1}{|\mathcal{R}_B|^2} \left\langle \sum_{b, b' \in \mathcal{R}_B} \tilde{\sigma}_b \tilde{\sigma}_{b'} \right\rangle \quad (70)$$

with a similar definition of \mathcal{W}_{e_G} for the green sublattice. As usual, $\mathcal{W}_{e_B} \neq 0$ implies condensation of the order parameter $\langle \sum_b \tilde{\sigma}_b \rangle$. Using Eq. (34), we can also rewrite \mathcal{W}_{e_B} as follows:

$$\mathcal{W}_{e_B} \propto \left\langle \sum_{b, b' \in \mathcal{R}_B} \cos(2\theta_b - 2\theta_{b'}) \right\rangle. \quad (71)$$

If in the thermodynamic limit $\mathcal{W}_{e_B} \neq 0$ ($\mathcal{W}_{e_G} \neq 0$), the resulting state can be invariant only under the $\mathbb{Z}_2 \times \mathbb{Z}_2$ group generated by S, R^2 (SR, R^2) defined in Eq. (35). If both \mathcal{W}_{e_B} and \mathcal{W}_{e_G} are nonzero, only the \mathbb{Z}_2 subgroup generated by R^2 remains as a potential symmetry.

Analogous to the Abelian case, an order parameter can be defined for the “condensation” of non-Abelian anyons m_R . Apart from the application of a string of X on \mathcal{R}_R , one requires a linear-depth circuit to fix the $m_R \times m_R$ fusion channel into the identity one as described in Eq. (2). Under the ungauging map specified in Sec. II B, the Wilson operator maps to the (local) thermal correlation $\mathcal{W}_{m_R}(g, b) \equiv \langle \sigma_g \sigma_b \widetilde{CZ}_{g,b} \rangle$ (see the derivation at the end of Appendix B). Similarly to the Abelian case, we can repackage this into a single symmetric order parameter that is efficiently computable via Monte Carlo methods:

$$\mathcal{W}_{m_R} \equiv \frac{1}{|\mathcal{R}_G||\mathcal{R}_B|} \left\langle \sum_{g \in \mathcal{R}_G} \sum_{b \in \mathcal{R}_B} \sigma_g \sigma_b \widetilde{CZ}_{g,b} \right\rangle. \quad (72)$$

In terms of the four-state clock formulation discussed in Sec. IV B [in particular, around Eq. (31)], this correlation becomes

$$\begin{aligned} \mathcal{W}_{m_R} &\propto \left\langle \sum_{g \in \mathcal{R}_G} \sum_{b \in \mathcal{R}_B} \mathbf{n}_g \cdot \mathbf{n}_b \right\rangle \\ &\propto \left\langle \sum_{g \in \mathcal{R}_G} \sum_{b \in \mathcal{R}_B} \cos(\theta_g - \theta_b) \right\rangle, \end{aligned} \quad (73)$$

TABLE I. Order parameters and remaining symmetry group for a deformed pure wave function. Phase diagram as characterized by the order parameters \mathcal{W}_{e_G} , \mathcal{W}_{e_B} , and \mathcal{W}_{m_R} . These signal spontaneous symmetry breaking of the D_4 symmetry down to a subgroup specified in the first column. The resulting quantum many-body phase is specified in the rightmost column. When $\beta_G^z = \beta_B^z$, an additional mirror sublattice symmetry M exists.

Remaining symmetry	\mathcal{W}_{e_G}	\mathcal{W}_{e_B}	\mathcal{W}_{m_R}	Phase
$\mathbb{Z}_2 \times \mathbb{Z}_2 \cong \langle SR, R^2 \rangle$	$\neq 0$	0	0	$ \text{TC}\rangle \text{TC}\rangle$
$\mathbb{Z}_2 \times \mathbb{Z}_2 \cong \langle S, R^2 \rangle$	0	$\neq 0$	0	$ \text{TC}\rangle \text{TC}\rangle$
$\mathbb{Z}_2 \cong \langle SR \rangle$	$\neq 0$	0	$\neq 0$	$ \text{TC}\rangle$
$\mathbb{Z}_2 \cong \langle S \rangle$	0	$\neq 0$	$\neq 0$	$ \text{TC}\rangle$
$\mathbb{Z}_2 \cong \langle R^2 \rangle$	$\neq 0$	$\neq 0$	0	$ \text{TC}\rangle$
\emptyset	$\neq 0$	$\neq 0$	$\neq 0$	Trivial

a form that we exploit to understand the remaining symmetry after different symmetry-breaking patterns. Notice that the variable $\theta_g - \theta_b$ is invariant under the symmetry transformation R , i.e., $\pi/2$ rotations. In the following, we find that \mathcal{W}_{m_R} takes a finite value only when, in addition, either \mathcal{W}_{e_G} or \mathcal{W}_{e_B} is also nonvanishing. As detailed in Table I, for a finite value of \mathcal{W}_{m_R} the D_4 symmetry is either broken down to a \mathbb{Z}_2 subgroup generated by either SR or S or fully broken. Other symmetry-breaking patterns can potentially also arise and would be interesting to explore but do not appear to be relevant for the present study.

2. Phase diagram

In the following, we combine the knowledge of the phase diagram of decoupled $O(1)$ and $O(2)$ loop models, together with Monte Carlo simulations, to characterize the phase diagram of the deformed wave function in Eq. (65). Recall that, at the end of Sec. IV E, we found that this phase diagram corresponds to that of $\text{tr}(\rho^n)$ for the decohered density matrix in the limit $n \rightarrow \infty$. Sketches of the resulting phase diagrams for $\beta_G^z = 0$ and for $\beta_G^z = \beta_B^z$, respectively, appear in Figs. 6(a) and 6(b). Numerical results obtained via Monte Carlo simulations are additionally shown in Figs. 7 and 8, respectively. These were obtained using a bilayer honeycomb lattice with linear sizes $\mathcal{N}_x = 80$ and $\mathcal{N}_y = 40$ and periodic boundary conditions in both directions. Final data were then obtained averaging over $\mathcal{N}_x \times \mathcal{N}_y \times 7 \times 10^4$ realizations. Additional details regarding Monte Carlo simulations are contained in Appendix G.

Let us first consider small deformations. We note that, by combining the ungauging and disentangling mappings in Sec. II B with Witten’s conjugation method [154, 155], one can find a local Hamiltonian for which $|\psi(\beta_R^x, \beta_G^z, \beta_B^z)\rangle$ is a ground state. In particular, small deformations correspond to small perturbations at the Hamiltonian

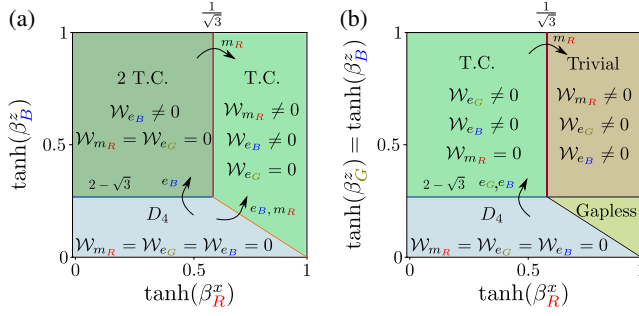


FIG. 6. Schematic phase diagram for pure wave-function deformations. Both panels include the value of the relevant order parameters indicating condensation of anyons as well as the resulting phases (see Ref. [102]) labeled as $D_4 \rightarrow D_4$ topological order, $2TC \rightarrow$ topological order equivalent to two copies of the toric code phase, $TC \rightarrow$ toric code phase. (a) shows the phase diagram as a function of β_R^x and β_B^z for $\beta_G^z = 0$. (b) displays the phase diagram as a function of β_R^x and $\beta_G^z = \beta_B^z$. This shows an additional “gapless” phase diagnosed by the power-law decay of two-point correlation functions as shown in Fig. 9. Transitions among phases can be understood as condensation of anyons indicated by an arrow. The underlying stat-mech model in the Ising-like representation is displayed in Eq. (69).

level. Since $|D_4\rangle$ is the ground state of a gapped Hamiltonian, $|\psi(\beta_R^x, \beta_G^z, \beta_B^z)\rangle$ remains within the D_4 TO phase for sufficiently small β_R^x , β_G^z , and β_B^z . Within this phase, none of the order parameters introduced in the previous section picks up a finite expectation value—consistent with the presence of a fully D_4 symmetric phase. This phase corresponds to the lower left corners in Figs. 6(a) and 6(b).

Now, by tuning β_R^x , β_G^z , and β_B^z to larger values, a ground-state phase transition associated to the condensation of different types of anyons can occur. To simplify matters, let us first consider the scenario with $\beta_G = 0$. Weights appearing in the loop model partition function $\mathcal{Z}_{|\psi\rangle}(\beta_R^x, \beta_B^z)$ then become positive, leading to the simplified expression in Eq. (67). The cases where either β_R^x or β_B^z are set to zero, respectively, correspond to $O(1)$ and $O(2)$ loop models that we already encountered in Secs. III A and IV A. The former admits a finite-temperature phase transition at $\tanh(\beta_B^z) = 2 - \sqrt{3}$, beyond which the D_4 TO becomes an Abelian TO equivalent to two copies of the TC, as already discussed in Sec. III A, including the subtlety mentioned in Ref. [102]. We label this phase by the state $|2TC\rangle$ [upper left corner in Fig. 6(a)]. This phase transition can be diagnosed by the quantity \mathcal{W}_{e_B} acquiring a finite value for sufficiently large β_B^z . This corresponds to the spontaneous symmetry breaking of D_4 down to the $\mathbb{Z}_2 \times \mathbb{Z}_2$ subgroup generated by $\langle R^2, S \rangle$. This characterization is consistent with our numerical results in Fig. 7(a), where indeed \mathcal{W}_{e_B} acquires a finite value at $\tanh(\beta_B^z) > 2 - \sqrt{3}$ (dashed black horizontal line). Moreover, Figs. 7(b) and 7(c) show that in this phase $\mathcal{W}_{m_R} = \mathcal{W}_{e_B} = 0$. In Fig. 8, we also characterize the phase diagram when tuning $\beta_G^z = \beta_B^z$. While the loop model for β_B^z and β_G^z finite has negative Boltzmann weights, its Ising-like formulation in Eqs. (30) and (36) does not, allowing one to address them using Monte Carlo. Numerical results are shown in Fig. 8(a), where \mathcal{W}_{e_B} and \mathcal{W}_{e_G} take the same numerical values due to the M mirror symmetry. Only the \mathbb{Z}_2 symmetry group generated by R^2 then remains, and the system reduces to the TC phase $|TC\rangle$.

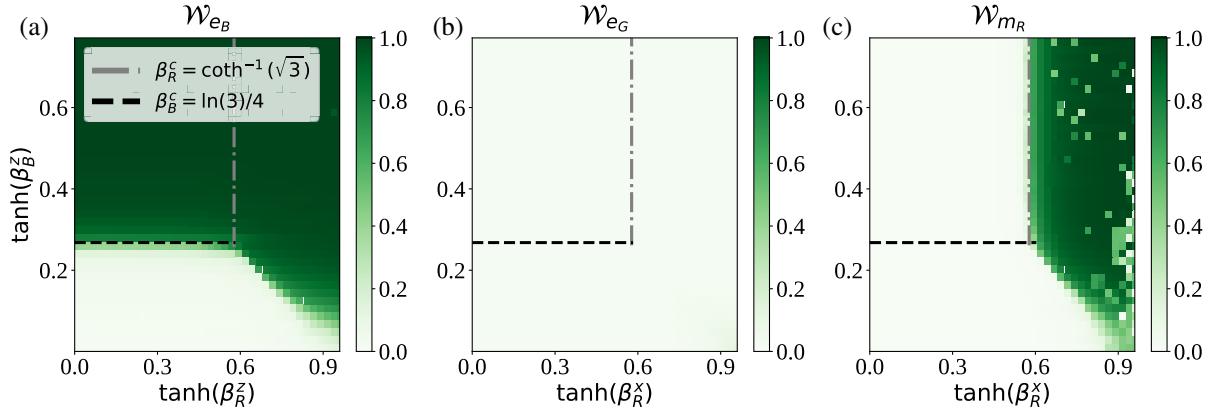


FIG. 7. Numerical phase diagram for pure wave-function deformations with $\beta_G^z = 0$ using Monte Carlo. Numerical Monte Carlo simulations with linear sizes $\mathcal{N}_x = 80$ and $\mathcal{N}_y = 40$ and periodic boundary conditions. The dashed black horizontal line indicates the value of $\beta_B^z = \ln(3)/4$ at which an Ising transition takes place for $\beta_R^x = 0$. Similarly, the dot-dashed gray vertical line marks the Ising transition occurring in the limit $\beta_B^z = \infty$ attained at $\beta_R^x = \coth^{-1}(\sqrt{3})$. (a) and (b) show the behavior of the order parameters \mathcal{W}_{e_B} and \mathcal{W}_{e_G} defined in Eq. (70) signaling the condensation of Abelian e_B and e_G charges, respectively. (c) shows instead the behavior of \mathcal{W}_{m_R} signaling the condensation of non-Abelian fluxes m_R . Results are averaged over 70000 metropolis steps. Additional details of the simulations are provided in Appendix G.

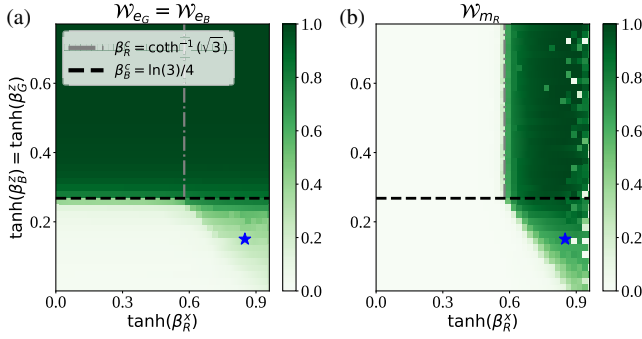


FIG. 8. Numerical phase diagram for pure wave-function deformations with $\beta_G^z = \beta_B^z$ using Monte Carlo. Numerical Monte Carlo simulations with linear sizes $\mathcal{N}_x = 80$ and $\mathcal{N}_y = 40$ and periodic boundary conditions. Because of the choice $\beta_G^z = \beta_B^z$, the order parameters \mathcal{W}_{e_B} and \mathcal{W}_{e_G} shown in (a) numerically agree throughout the phase diagram. (b) shows instead the behavior of \mathcal{W}_{m_R} signaling the condensation of non-Abelian fluxes m_R . The lightly shaded region in both panels corresponds to a gapless phase. For any point in this region, the three order parameters decay with system size. Moreover, in Fig. 9, we show the power-law behavior of two-point correlations evaluated for the parameters indicated by \star . Additional details of numerical simulations are the same as in Fig. 7.

Our numerical results also show that this critical point [corresponding to one (two decoupled) Ising critical points when $\beta_G^z = 0$ ($\beta_G^z = \beta_B^z$)] extends into a line when tuning β_R^x to finite values [see the dashed black horizontal line in Figs. 7(a) and 8(a)]. But how large can β_R^x become before a different transition occurs? Let us consider the regime with $\beta_B^z \rightarrow \infty$, i.e., $\tanh(\beta_B^z) \rightarrow 1$ (upper horizontal boundaries in Fig. 6). To understand this limit, we set again $\beta_G^z = 0$ and exactly rewrite the partition function in Eq. (67) as

$$\mathcal{Z}_{|\psi\rangle}(\beta_R^x, \beta_B^z) = \sum_{L_R} \left(\frac{\tanh(\beta_R^x)}{\sqrt{2}} \right)^{|L_R|} \prod_{\ell_R \in L_R} \mathcal{Z}_{\ell_R}^{\text{Ising}}(\tilde{\beta}_B^z), \quad (74)$$

where $\mathcal{Z}_{\ell_R}^{\text{Ising}}(\tilde{\beta}_B^z)$ is the partition function of the 2D Ising model on the honeycomb lattice within the interior of a connected component ℓ_R of the loop configuration $L_R = \bigoplus \ell_R$, at inverse temperature $e^{-2\tilde{\beta}_B^z} = \tanh(\beta_B^z)$. Here, the factor $2^{C_{L_R}}$ is reabsorbed into the Ising partition function $\mathcal{Z}_{\ell_R}^{\text{Ising}}(\tilde{\beta}_B^z)$ in a given component ℓ_R , using the exact relation $\mathcal{Z}_{\ell_R}^{\text{Ising}}(\tilde{\beta}_B^z) = 2\mathcal{Z}_{\ell_R}^{\text{O}(1)}(\beta_B^z)$ [156]. To perform this mapping, one can imagine there are two blue Ising spins on every vertex of the honeycomb lattice, which are decoupled if a red loop runs through and which are infinitely coupled otherwise. In this formulation, we can see that in the limit $\beta_B^z \rightarrow \infty$ (i.e., $\tilde{\beta}_B^z \rightarrow 0$),

$\mathcal{Z}_{\ell_R}^{\text{Ising}}(\tilde{\beta}_B^z) = 2^{N(\ell_R)}$, where $N(\ell_R)$ corresponds to the number of spins in a component ℓ_R . Hence,

$$\begin{aligned} \mathcal{Z}_{|\psi\rangle}(\beta_R^x, \beta_B^z \rightarrow \infty) &= \sum_{L_R} \left(\frac{\tanh(\beta_R^x)}{\sqrt{2}} \right)^{|L_R|} \sqrt{2}^{|L_R|} \\ &= \sum_{L_R} \tanh(\beta_R^x)^{|L_R|}, \end{aligned} \quad (75)$$

namely, we recover an $O(1)$ loop model on the honeycomb lattice which has an Ising critical point at $\tanh(\beta_R^x) = 1/\sqrt{3}$ [112]. Physically, we can interpret this result as showing that condensing e_B anyons dresses up the non-Abelian m_R anyons—downgrading them to Abelian, i.e., $d_{m_R} = 2 \rightarrow 1$. In fact, this is consistent with the condensation picture: Upon condensing e_B , the non-Abelian flux m_R (and similarly for m_G) splits up into quantum-dimension-one $m_R^{(1)}$ anyons, $m_R = m_R^{(1)} \oplus m_R^{(1)} e_G$. These correspond to the Abelian anyons of the $|2, \text{TC}\rangle$ phase with stabilizers A_s^{TC}, B_t and A_s^{TC}, B_t , where A_s^{TC} is defined via $A_s^{\text{TC}} = \prod_{j \in \square} X_j$, and analogously for A_s^{TC} . This is analogous to splitting of the boson $\sigma\bar{\sigma}$ into charge and flux when condensing $\psi\bar{\psi}$ in an Ising \times Ising TO [157].

Suppose now that the system lies deep within the $|2\text{TC}\rangle$ phase (namely, $\beta_B^z \rightarrow \infty$ and $\beta_R^x \ll 1$) and that we increase the value of β_R^x . We just found that a phase transition can occur at a finite β_R^x , where the $\mathbb{Z}_2 \times \mathbb{Z}_2$ can break down to either the trivial or a \mathbb{Z}_2 subgroup. The latter can be generated by either S or SR . As specified in Table I, the former path ($\mathbb{Z}_2 \times \mathbb{Z}_2 \rightarrow \langle S \rangle$) is signaled by a nonzero value of the quantity \mathcal{W}_{m_R} . Figure 7(c) shows that \mathcal{W}_{m_R} acquires a finite value roughly beyond $\tanh(\beta_R^x) = 1/\sqrt{3}$, as also happens in the limit $\beta_B^z \rightarrow \infty$. Nonetheless, we can follow a similar discussion as in Sec. III A and understand the fate of $|2\text{TC}\rangle$ by looking at its stabilizers in the limit $\beta_R^x \rightarrow \infty$. In this limit, m_R anyons—now corresponding to Abelian excitations of B_t —condense, leading to the confinement of e_R charges, with which they braid non-trivially. Algebraically, this makes the stabilizers A_s^{TC} to become trivial $A_s^{\text{TC}} \rightarrow \mathbb{1}$, and, hence, a single TC copy is left. This is stabilized by the commuting operators A_s^{TC} and B_t .

Instead, when taking $\beta_G^z = \beta_B^z$, and assuming that the system lies deeply within the $|\text{TC}\rangle$ phase [upper left phase in Fig. 8(b)], the system is a common eigenstate of the stabilizers A_s^{TC} and B_t . In the limit $\beta_R^x \rightarrow \infty$, one then predicts a condensation transition of the (now Abelian) fluxes m_R leading to a topologically trivial phase. This is captured by \mathcal{W}_{m_R} acquiring a finite value as shown in Fig. 8(b).

Finally, let us explore the regime with $\beta_R^x \rightarrow \infty$ but finite β_B^z and β_G^z . Following our discussion at the end of Sec. IV E, we find that in the limit $\beta_R^x \rightarrow \infty$ [vertical right

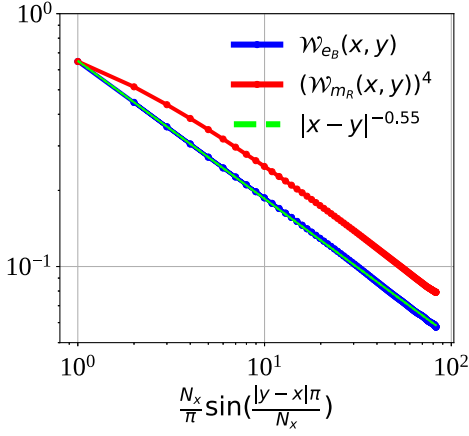


FIG. 9. Power-law decay of correlations for $\beta_G^z = \beta_B^z$. Numerical evaluation of correlation functions $\mathcal{W}_{e_B}(b, b') \equiv \langle \tilde{\sigma}_b \tilde{\sigma}_{b'} \rangle$ and $\mathcal{W}_{m_R}(b, g) \equiv \langle \sigma_b \sigma_g \tilde{C}_{b,g} \rangle$ within the extended gapless phase for parameters $\tanh(\beta_B^z) = 0.15$ and $\tanh(\beta_B^z) = 0.85$ indicated with \star in Fig. 8(a). For the former x and y belong to the blue sublattice, while for the latter $x \in G$ and $y \in B$. The average $\langle \cdot \rangle$ corresponds to the thermal expectation value evaluated with respect the classical Hamiltonian in Eq. (69). Numerical data correspond to system size $N_x = 520, N_y = 260$ with periodic boundary conditions, 7×10^4 equilibration time steps, and 12×10^4 Metropolis steps to obtain the plotted data.

boundary in Figs. 6(a) and 6(b)], the system is described by the stat-mech model $\sum_{L_R} (2^{C_{L_R}} / \sqrt{2}^{|L_R|}) \mathcal{Z}_{L_R}(t_G, t_B)$, which corresponds an $O(2)$ loop model for $t_G = t_B = 0$ at the BKT critical point separating an extended gapless phase from one with exponentially decaying correlations [91]. As we already found, in the limit of small (and positive) t_G and t_B , one can then explore its extended gapless phase. In the next section, we argue that, while only a critical line separating the D_4 TO from the TC phase remains when setting $t_G = 0$ [see Fig. 6(a)], an extended gapless phase appears for small but finite $t_G = t_B$ [see lower right triangular region in Fig. 6(b)], whose stability is related to the enhanced $D_4 \rightarrow D_8$ symmetry for this choice of parameters. See the next subsection. In fact, a closer look at the numerical results in Fig. 8 shows that, for a point [blue star \star in Fig. 8(a)] within the lightly shaded triangular region below the dashed black horizontal critical line, all order parameters \mathcal{W}_{m_R} , \mathcal{W}_{e_G} , and \mathcal{W}_{e_B} scale down with increasing system size—consistent with a gapless, symmetric phase with power-law-decaying correlations (see Fig. 9) emerging in those regions.

3. Field theory description

In this section, we aim to provide a consistent field theory interpretation of the results we discussed above. For this is useful to rewrite the local Hamiltonian in Eq. (69) using four-state clock variables in Eq. (32). Recalling that

$n_b = [\cos(\theta_b), \sin(\theta_b)]^T$ and $n_g = [\cos(\theta_g), \sin(\theta_g)]^T$, we find that the Hamiltonian (69) takes the form

$$H = -\sqrt{2}\beta_R^x \sum_{\langle b,g \rangle_\circ} \cos(\theta_b - \theta_g) - \beta_B^z \sum_{\langle\langle b,b' \rangle\rangle_\circ} \cos(2\theta_b - 2\theta_{b'}) - \beta_G^z \sum_{\langle\langle g,g' \rangle\rangle_\circ} \cos(2\theta_g - 2\theta_{g'}). \quad (76)$$

Here, we have used that $\tilde{\sigma}_j \tilde{\sigma}_{j'} = \cos(2\theta_j - 2\theta_{j'})$ when j and j' lie on the same sublattice. Recall that, for all values of the parameters, the model has a $D_4 = \langle R, S \rangle$ symmetry.

Let us first set $\beta_G^z = \beta_B^z = 0$. In this case, the system is described by an $O(2)$ loop model which at the critical point—attained at $\beta_R^x = \infty$ —corresponds to a compact boson at the BKT point. With this continuum formulation in mind, we define a 2π -periodic, real-valued function $\theta(x, y)$ living on lattice sites now labeled (x, y) , constrained such that $\theta(x, y) = \theta_g$ if $(x, y) \in G$ and $\theta(x, y) = \theta_b$ if $(x, y) \in B$. On $\theta(x, y)$, symmetry transformations act via $R: \theta \rightarrow \theta + \pi/2$, while $S: \theta \rightarrow -\theta$. Moreover, the mirror symmetry M now acts as $M: \theta(x, y) \rightarrow \pi/4 - \theta(-x, y)$, namely, $\theta_g \rightarrow \pi/4 - \theta_b$ and vice versa. To relax the four-state constraint in Eq. (32), we can include a softening potential with minima at the physical values of $\theta(x, y)$; in general, this potential takes the form $+g_B \cos(4\theta_b) - g_G \cos(4\theta_g) + \dots$, where the ellipsis denotes higher harmonics. Importantly, when M symmetry is enforced, as is the case here, since $\beta_G^z = \beta_B^z$, we must have $g_B = g_G$. Taking now the continuum limit by assuming that slowly varying configurations dominate [158], the leading symmetry-allowed harmonic has the form $-g \cos(8\theta)$ due to cancellation of the $\cos(4\theta)$ contributions by M symmetry. One can view the surviving $-g \cos(8\theta)$ piece as implementing a common field constraint on the two sublattices.

We then arrive at the Hamiltonian density

$$\mathcal{H} = \frac{K}{2\pi} \cos(\nabla\theta) - g \cos(8\theta). \quad (77)$$

One can recognize \mathcal{H} as the XY model [159,160], whose low-temperature physics can be recovered [provided the $\cos(8\theta)$ is irrelevant] by the $(2+0)$ -dimensional quadratic theory $\mathcal{H} = [K/(2\pi)](\nabla\theta)^2$ or as the $(1+1)$ -dimensional quantum field theory

$$\mathcal{H} = \frac{1}{2\pi} \left(K(\partial_x \theta)^2 + \frac{1}{4K} (\partial_x \varphi)^2 \right) - g \cos(8\theta) - \lambda \cos(\varphi), \quad (78)$$

namely, a Luttinger liquid with Luttinger parameter K in the presence of additional potential terms. Here, φ is a dual 2π -periodic field satisfying $[\varphi(x), \partial_y \theta(y)] = i2\pi \delta(x-y)$, symmetric under R and transforming as $\varphi \rightarrow -\varphi$ under S [161]. Hence, $\cos(\varphi)$ is a symmetry-allowed local term

with the smallest scaling dimension. The scaling dimension of $\cos(m\theta)$ is given by $m^2/(4K)$ and that of $\cos(n\varphi)$ by n^2K . The resulting compact boson field theory formulation when additional cosine terms are irrelevant exactly matches the fact that the $O(2)$ loop model can be rewritten as a restricted solid-on-solid height-field model [discretized version of $\int(\nabla\varphi)^2$], whose critical point is described by a BKT transition [163,164]. It turns out that, when $\beta_G^z = \beta_B^z = 0$, the $O(2)$ loop model is parked [165] at $K = 2$ [166]. For this value of the Luttinger parameter, the $\cos(\varphi)$ term is marginally irrelevant for one choice of the sign of λ and marginally relevant for the other and is then responsible for gapping out the system into the “short loop phase.” In this work, λ is tuned by β_R^x such that $\lambda \rightarrow 0$ when $\beta_R^x \rightarrow \infty$ (i.e., only one sign of λ can be explored), and the short loop phase corresponds to D_4 TO.

So far, we provided a field theory description along the horizontal axis in Fig. 6 and close to the critical point $\beta_R^x = \infty$. Let us now include weak $\beta_G^z, \beta_B^z \neq 0$. One important effect of this perturbation is to modify the value of K , and, hence, we expect additional transitions. First, when β_G^z and β_B^z differ, the M symmetry is broken—allowing a $\cos(4\theta)$ perturbation which is marginal for $K = 2$, coexisting with the $\cos(\varphi)$ term. When the former becomes relevant (i.e., for $K > 2$), the BKT point will be gapped out due to the pinning of θ by $\cos(4\theta)$. This scenario takes place when setting $\beta_G^z = 0$ while keeping β_B^z finite, leading to the toric code phase in Fig. 6(a), where θ_b is pinned, and, hence, \mathcal{W}_{e_b} takes a finite value. On the other hand, for $K < 2$, $\cos(\varphi)$ becomes relevant, driving a transition into the gapped D_4 TO phase. Finally, $\beta_G^z = \beta_B^z \neq 0$ [for which the $\cos(4\theta)$ term is again forbidden by symmetry] can lead to an extended gapless phase [see lower right triangle labeled “gapless” in Fig. 6(b)] if $2 < K < 8$, since then both $\cos(\varphi)$ and $\cos(8\theta)$ become irrelevant.

We validate the consistency of this hypothesis by numerically evaluating the correlations $\mathcal{W}_{e_b}(x, y)$ and $\mathcal{W}_{m_R}(x, y)$ on a point belonging to this phase with parameters $\tanh(\beta_R^x) = 0.85$ and $\tanh(\beta_G^z) = \tanh(\beta_B^z) = 0.15$. Using the previous mapping and assuming the existence of this gapless phase, one finds that

$$\mathcal{W}_{m_R}(x, y) = \langle \cos[\theta(x) - \theta(y)] \rangle \sim |x - y|^{-1/2K} \quad (79)$$

and, similarly,

$$\mathcal{W}_{e_b}(x, y) = \langle \cos\{2[\theta(x) - \theta(y)]\} \rangle \sim |x - y|^{-2/K}. \quad (80)$$

Hence, both correlations are expected to decay as a power law with the distance, where the power-law exponent of $\mathcal{W}_{e_b}(x, y)$ is 4 times larger than that of $\mathcal{W}_{m_R}(x, y)$ regardless the value of K . Figure 9 shows that (i) both correlations indeed decay as a power law with $|x - y|$ and

(ii) the data approximately satisfy the relation $\mathcal{W}_{e_b}(x, y) \sim [\mathcal{W}_{m_R}(x, y)]^4$ at long distances. Moreover, assuming Eq. (80), we extract that the Luttinger parameter is given by $K \approx 4 > 2$.

Finally, if β_G^z and β_B^z are strong, one sees from the lattice Hamiltonian that we should be able to condense, e.g., $\cos(2\theta)$ without condensing $\cos(\theta)$. This can be accomplished by, e.g., introducing new “Ising” variables σ that have the same symmetry charges as $\cos(2\theta)$. We provide a consistent field theory formulation in Appendix F.

B. Local quantum channels

We now describe the fate of the decohered density matrix as measured by the purity $\text{tr}(\rho^2)$, when subjecting D_4 TO to the composition of the following commuting local quantum channels:

$$\begin{aligned} \mathcal{E}_r^X(\rho_0) &= (1 - p_R)\rho_0 + p_R X_r \rho_0 X_r, \\ \mathcal{E}_b^Z(\rho_0) &= (1 - p_B)\rho_0 + p_B Z_b \rho_0 Z_b, \\ \mathcal{E}_g^Z(\rho_0) &= (1 - p_G)\rho_0 + p_G Z_g \rho_0 Z_g, \end{aligned} \quad (81)$$

with tunable error rates $p_R, p_G, p_B \in [0, 1/2]$.

1. Phase diagram from $\text{tr}(\rho^\infty)$

As discussed at the end of Sec. IV E, the limit $\lim_{n \rightarrow \infty} \text{tr}(\rho^n)^{1/n}$ at $p_R = \frac{1}{2}$ is given by the norm of the deformed wave function $|\psi(\beta_R^x, \beta_G^z, \beta_B^z)\rangle$ (up to possibly overall factors) when $\beta_R^x = \infty$, as long as the weight $\mathcal{Z}_{L_R}(t_G, t_B) \geq 0$ in Eq. (45) is non-negative. First, we showed that this is always true when $\beta_G^z = 0$, and, hence, the phase diagram of $\text{tr}(\rho^\infty)$ is given by the right vertical boundary of the phase diagram in Fig. 6(a). As explicitly shown in Fig. 10(a), this showcases an unstable BKT critical point at $\beta_B^z = 0$ beyond which D_4 TO is completely lost.

On the other hand, we also argued that, even when β_G^z and β_B^z are finite, $\mathcal{Z}_{L_R}(t_G, t_B)$ remains non-negative. We use this fact to conjecture that the phase diagram of $\lim_{n \rightarrow \infty} \text{tr}(\rho^n)^{1/n}$ as a function of β_G^z and β_B^z corresponds to the right vertical boundary in Fig. 6(b). Unlike the previous case, this contains an extended gapless phase described by a compact boson for values of β_G^z and β_B^z below a threshold. The corresponding phase diagram is shown in Fig. 10(b). Recall that as argued in Sec. IV E the critical threshold for the $n \rightarrow \infty$ limit, $p_c^{(\infty)}$, lower bounds the intrinsic threshold

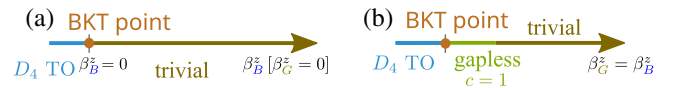


FIG. 10. Phase diagram of $\text{tr}(\rho^\infty)$ for initial state $|\psi(\beta_G^z, \beta_B^z, \beta_R^z = 0)\rangle$ in Eq. (7) at $p_R = \frac{1}{2}$. The phase diagrams include the D_4 topological order phase, a gapless phase, and a “trivial” phase [102].

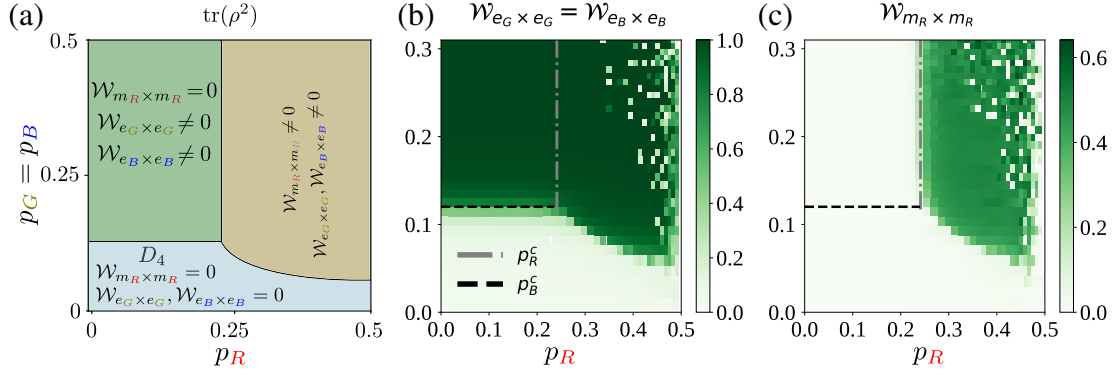


FIG. 11. Phase diagram for decohered mixed state with $p_G = p_B$ via $\text{tr}(\rho^2)$. (a) Schematic phase diagram including the values of the relevant order parameters. D_4 TO is robust to finite error rates (light blue region). This is signaled by the order parameters acquiring a finite value for sufficiently large error rates. (b),(c) Numerical phase diagram with linear system sizes $\mathcal{N}_x = 40$ and $\mathcal{N}_y = 20$ and with periodic boundary conditions. Dashed black horizontal and gray dot-dashed lines indicate the critical values $p_G = p_B = (1 - 3^{-1/4})/2 \approx 0.12$ and $p_R = (1 - \sqrt{2 - \sqrt{3}})/2 \approx 0.24$, respectively. (b) and (c) show numerical results for the order parameters $\mathcal{W}_{e_G \times e_G} = \mathcal{W}_{e_B \times e_B}$ [Eq. (85)] and $\mathcal{W}_{m_R \times m_R}$ [Eq. (86)], respectively.

beyond which the capacity of D_4 TO as a good quantum memory decreases.

2. Phase diagram from purity

Following a similar derivation as in the previous sections, one finds that the purity of the decohered density matrix takes the general form

$$\text{tr}(\rho^2) \stackrel{\text{decohere}}{\propto} \sum_{m_R, e_G, e_B} \sum_{L_R} \sum_{\gamma_B} \sum_{\gamma_G} (r_B^z)^{|\gamma_B|} (r_G^z)^{|\gamma_G|} \left(\frac{r_R^x}{2} \right)^{|L_R|} 4^{C_{L_R}}, \quad (82)$$

i.e., three coupled $O(N)$ loop models with $N = 1, 4$. Notice the absence of the sign function σ_{L_R} when comparing to Eq. (68), the reason being that now every contribution for the pure wave-function deformation is squared. Alternatively, one can write these coupled loop models in an explicitly local way similar to Eq. (41) and given by (including temperature factors)

$$H_\rho = -\beta_R^x \sum_{\langle i,j \rangle_\square} \sigma_i \sigma_j \widetilde{\text{CZ}}_{ij}^{(2)} \widetilde{\text{CZ}}_{ij}^{(3)} - \beta_B^z \sum_{\langle\langle b,b' \rangle\rangle_\square} \tilde{\sigma}_b^{(2)} \tilde{\sigma}_{b'}^{(2)} \tilde{\sigma}_b^{(3)} \tilde{\sigma}_{b'}^{(3)} - \beta_G^z \sum_{\langle\langle g,g' \rangle\rangle_\square} \tilde{\sigma}_g^{(2)} \tilde{\sigma}_{g'}^{(2)} \tilde{\sigma}_g^{(3)} \tilde{\sigma}_{g'}^{(3)}, \quad (83)$$

defined on three honeycomb layers (one more than the stat-mech model associated to the deformed wave function). As discussed in previous sections, this model has an enhanced symmetry group, since H_ρ is invariant under D_4 symmetry transformations involving either layers 1 and 2 or 1 and 3.

While a more detailed characterization of the resulting phase diagram as in Ref. [16] is left for future work, we now briefly discuss the thresholds beyond which D_4 TO is lost in the presence of decoherence assuming $p_G = p_B$. As for the pure wave-function deformation, the previous formulation permits a numerical evaluation of the phase diagram of the decohered density matrix using standard Monte Carlo methods. We directly simulate H_ρ in Eq. (83) on a trilayer honeycomb lattice with a brick-wall configuration with linear system sizes $\mathcal{N}_x = 80$ and $\mathcal{N}_y = 40$. Numerical results together with a sketch of the resulting phase diagram are shown in Fig. 11(a).

From a physics perspective, the main difference between decohering a density matrix and then computing its purity and analyzing a pure wave-function deformation is that anyons are simultaneously created on both the bra and ket in the former case. Indeed, let us, e.g., consider the effect of the quantum channel \mathcal{E}_b^Z in Eq. (81) and recall that the purity corresponds (up to an overall constant) to the norm of the unnormalized state

$$|\rho\rangle\rangle = e^{\mu_B \sum_{b \in \mathcal{R}_B} Z_j \otimes Z_j} |D_4\rangle \otimes |D_4\rangle, \quad (84)$$

with $\tanh(\mu_B) = p_B/(1 - p_B)$. Hence, pairs of fluxes $e_B \times \bar{e}_B$ are simultaneously created on the first and second copies [5]. The corresponding order parameter capturing the condensation of these pairs is given by

$$\mathcal{W}_{e_B \times e_B} \equiv \frac{1}{|\mathcal{R}_B|^2} \left\langle \sum_{b, b' \in \mathcal{R}_B} \tilde{\sigma}_b^{(2)} \tilde{\sigma}_b^{(3)} \tilde{\sigma}_{b'}^{(2)} \tilde{\sigma}_{b'}^{(3)} \right\rangle \quad (85)$$

and analogously defined for $\mathcal{W}_{e_G \times e_G}$. On the other hand, the condensation of $m_R \times \overline{m_R}$ is signaled by

$$\mathcal{W}_{m_R \times m_R} \equiv \frac{1}{|\mathcal{R}_G||\mathcal{R}_B|} \left\langle \sum_{g \in \mathcal{R}_G} \sum_{b \in \mathcal{R}_B} \sigma_g \sigma_b \widetilde{\mathcal{C}Z}_{g,b}^{(2)} \widetilde{\mathcal{C}Z}_{g,b}^{(3)} \right\rangle. \quad (86)$$

Figures 11(b) and 11(c) show the dependence of these order parameters as a function of p_R and $p_G = p_B$, resulting in the sketched diagram appearing in Fig. 11(a). For this numerical analysis, we considered a smaller system size $\mathcal{N}_x = 40, \mathcal{N}_y = 20$, since we found that close to the zero-temperature regime ($p_R \rightarrow 1/2$) the algorithm takes longer to converge. We believe the reason is the use of single site updates while interactions in H_ρ involve local six-spin terms. Similar to the pure wave-function deformation, an Ising phase transition occurs when $p_R = 0$ and for $p_G = p_B = (1 - 3^{-1/4})/2$. In this case, the partition function with Hamiltonian H_ρ in Eq. (83) corresponds to two decoupled 2D ferromagnetic Ising model lying on a triangular lattice. Figure 11(b) shows that, in fact, this point extends into a critical line (dashed black horizontal line) for finite values of p_R . Also, similarly to our analysis of the deformed wave function, we find an Ising critical point in the limit $p_G = p_B \rightarrow \infty$ for $p_R = (1 - \sqrt{2 - \sqrt{3}})/2$, which also extends into a critical line (dot-dashed gray vertical line) as shown in Fig. 11(c). However, unlike for the deformed wave function, attaining a critical point at $\beta_R^x \rightarrow \infty$ for $\beta_G^z = \beta_B^z = 0$, the horizontal axis with vanishing p_G and p_B is described by an O(4) loop model—which shows no critical behavior for any value of p_R . Finally, our numerical results indicate that, even for $p_R = 1/2, D_4$ topological order is robust to finite error rates $p_G = p_B$. This seems natural, since, when $p_G = p_B = 0$, the system is well within the shortly correlated small loop phase of the O(4) loop model, and so, by infinitesimally increasing the coupling to γ_G and γ_B , large red loop configurations L_R (indicative of the dense loop phase) are not expected to be suddenly generated. Hence, a finite threshold is expected.

C. Competing non-Abelian errors

At this point, we have explored scenarios involving either a single non-Abelian error or a combination of Abelian and non-Abelian errors acting on two different sublattices. In this last subsection, we consider the scenario of non-Abelian X errors acting on either two or all three sublattices. Suppose that X errors act on both \mathcal{R}_R and \mathcal{R}_B sublattices, hence proliferating pairs of m_R and m_B fluxes, respectively. Unlike in our previous analysis, two non-Abelian anyons of different color have nontrivial braiding statistics. In fact, even when two pairs of m_R and m_B are generated out of the vacuum, their

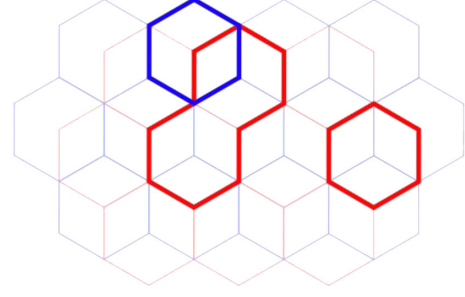


FIG. 12. Competing non-Abelian errors. Resulting coupled O(2) loop models when considering the proliferation of two non-Abelian errors acting on different sublattices. For the configuration shown in the figure, $C_{L_R \cup L_B} = 4$.

braiding toggles the fusion channel of both pairs from 1 to e_G . Hence, when annihilated, a pair of e_G are left behind [73]. This scenario corresponds to either the deformed wave function

$$|\psi(\beta_R^x, \beta_B^x)\rangle = e^{\beta_R^x/2 \sum_{r \in \mathcal{R}_R} X_r} e^{\beta_B^x/2 \sum_{b \in \mathcal{R}_B} X_b} |D_4\rangle \quad (87)$$

or the composition of the local quantum channels

$$\begin{aligned} \mathcal{E}_r^X(\rho_0) &= (1 - p_R)\rho_0 + p_R X_r \rho_0 X_r, \\ \mathcal{E}_b^X(\rho_0) &= (1 - p_B)\rho_0 + p_B X_b \rho_0 X_b, \end{aligned} \quad (88)$$

with $r \in \mathcal{R}_R$ and $b \in \mathcal{R}_B$. We find the classical partition function (see Appendix C 2 b)

$$\mathcal{Z}_{|\psi\rangle}(\beta_R^x, \beta_B^x) = \sum_{L_B, L_R} \frac{\tanh(\beta_R^x)^{|L_R|} \tanh(\beta_B^x)^{|L_B|}}{\sqrt{2}^{|L_R \cup L_B|}} 2^{C_{L_R \cup L_B}} \quad (89)$$

for the deformed pure wave function and

$$\text{tr}(\rho^2) \stackrel{\text{decohere}}{\propto} \sum_{m_R, m_B} \frac{r_R^{|L_R|} r_B^{|L_B|}}{2^{|L_R| + |L_B|}} 4^{C_{L_R \cup L_B}} \quad (90)$$

when computing the purity of the resulting decohered density matrix. Here, we have used the fact that only closed-loop configurations L_R and L_B on the red \mathcal{R}_R and blue \mathcal{R}_B sublattices, respectively, lead to nonvanishing contributions. Moreover, $C_{L_R \cup L_B}$ corresponds to the number of independent loops in the (colorless) union $L_R \cup L_B$ or, equivalently and perhaps more clearly, the number of faces enclosed [see Fig. 12(a) for an example with $C_{L_R \cup L_B} = 4$]. This coupling effectively works as an attractive potential between loops of different color that

increases their probability of crossing. This enhancement is directly related to their nontrivial braiding. If the models were decoupled, we would have instead found $2^{C_{LR}+C_{LB}}$. Hence, we find that the combination of two non-Abelian errors leads to two coupled $O(2)$ loop models. Unfortunately, we are not aware of previous works studying the phase diagram of the resulting coupled loop model, so we leave their study to future work. Finally, including the action of an X error on the remaining green sublattice \mathcal{R}_G follows a similar logic. However, in this case, three (instead of two) different colored loops can intersect in two different ways and results in vanishing contributions. What would be the fate of D_4 TO under the simultaneous action of these errors? Does the braiding have an impact on the robustness of D_4 TO under decoherence?

VI. D_4 STRONG-TO-WEAK SPONTANEOUS SYMMETRY BREAKING

Previous sections have dealt with the loss of coherence arising from coupling to an environment, which can result in a topologically ordered state losing its ability to store quantum information. However, the effect of a quantum channel does not necessarily lead to featureless stationary mixed states stripped of quantum correlations. There are various ways to avoid this fate—e.g., by fine-tuning the environment [167–171] or restricting to strongly symmetric quantum channels, namely, those for which every Kraus operator is symmetric [7,22,26,36,37,172,173]. As an example, the phenomenon of strong-to-weak spontaneous symmetry breaking (sw-SSB) has been shown to trigger long-range order in mixed states as measured by nonlinear observables in the density matrix [7,22,26,36].

Following these works, let us introduce the concept by considering a \mathbb{Z}_2 -symmetric short-range correlated initial state $|\psi_0\rangle$, under the action of a strongly \mathbb{Z}_2 -symmetric local quantum channel. For example, take $|\psi_0\rangle = |+\rangle^{\otimes N}$ in the X basis and suppose that every Kraus operator commutes with the \mathbb{Z}_2 generator $G = \prod_j X_j$. Such a symmetric channel can be given by the composition of local quantum channels

$$\rho_0 \rightarrow \mathcal{E}_{i,j}(\rho_0) = (1-p)\rho_0 + pZ_i Z_j \rho_0 Z_i Z_j \quad (91)$$

acting on every bond (i, j) of a square lattice. Notice that the resulting decohered density matrix ρ is invariant under both left and right multiplication by the symmetry transformation G . When this is the case, ρ is said to be strongly G symmetric.

While a local quantum channel cannot modify the scaling of linear correlations of the form $\text{tr}(\rho O_x O_y)$ as imposed by Lieb-Robinson bounds, the phenomenon of sw-SSB is

signaled by the generation of long-range correlations in nonlinear quantities like the Rényi-2 correlators

$$R_{\mathbb{Z}_2}^{(2)}(x, y) \equiv \frac{\text{tr}(\rho O_x O_y \rho O_x O_y)}{\text{tr}(\rho^2)}, \quad (92)$$

with O_x a charged operator under G , e.g., Z_x . The definition of sw-SSB, up to subtleties that are discussed later on, is as follows [36]: A strongly G -symmetric mixed state ρ has sw-SSB if the Rényi-2 correlator $R_{\mathbb{Z}_2}^{(2)}(x, y)$ of some charged local operator (e.g., Z_x in this case) is nonvanishing in the limit $|x - y| \rightarrow \infty$, while the conventional correlation function $\text{tr}(\rho Z_x Z_y)$ shows no long-range order. When evaluating $R^{(2)}(x, y)$ for $|\psi_0\rangle$, one finds that this corresponds to the thermal correlation $\langle \sigma_x \sigma_y \rangle_{2\beta}$ on the 2D Ising model at inverse temperature 2β with $\tanh(\beta) = p/(1-p)$. Hence, for error rates larger than a critical value $p_c^{(2)}$, this metric reveals that the density matrix ρ displays sw-SSB. It follows that for $p^{(2)} > p_c^{(2)}$ the decohered density matrix is not “symmetrically invertible”; namely, no mixed state $\tilde{\rho}$ exists such that $\rho \otimes \tilde{\rho}$ can be connected to a symmetric product state via a local quantum channel [36], showcasing the nontrivial structure of ρ .

While *a priori* this phenomenon appears to be unrelated to a decoherence transition for TO states, it has been shown that, in fact, these phenomena are intimately related. One can connect the \mathbb{Z}_2 sw-SSB transition discussed above to decohering the TC wave function via an Z quantum channel using Wegner’s duality [7,9,174]. This duality maps the toric code wave function to a trivial paramagnet $|+\rangle^{\otimes N}$ and vice versa. Similarly here, the combination of the gauging and entangling maps introduced in Sec. II B together with Eq. (28) provides a complementary dual picture of subjecting the D_4 topological order to decoherence as a D_4 sw-SSB transition. To reveal this connection, we consider a bilayer system where each of the layers corresponds to a honeycomb lattice with qubits lying on the vertices as in Fig. 4(a). The upper layer hosts red qubits initialized in the product state $|+\rangle \equiv |+\rangle^{\otimes |\mathcal{R}_R|}$, while the lower layer hosts green $|\tilde{+}\rangle \equiv |\tilde{+}\rangle^{\otimes |\mathcal{R}_G|}$ and blue $|\bar{+}\rangle \equiv |\bar{+}\rangle^{\otimes |\mathcal{R}_B|}$ qubits lying on the two \mathcal{R}_G and \mathcal{R}_B sublattices, respectively. The state of the system $|\psi_0\rangle = |+\rangle|\tilde{+}\rangle|\bar{+}\rangle$ is then invariant under the D_4 symmetry transformation generated by R and S as defined below Eq. (35). Following our discussion in the previous sections, one can readily find strongly D_4 symmetric local quantum channels corresponding to the “ungauged” version of the quantum channels in Eqs. (81). These are given by the composition of the commuting local quantum channels

$$\mathcal{E}_{i,j}^X(\rho_0) = (1-p_x)\rho_0 + p_x Z_i Z_j \tilde{C} Z_{ij} \rho_0 Z_i Z_j \tilde{C} Z_{ij}, \quad (93)$$

acting on every link (i, j) of the bilayer system, and

$$\mathcal{E}_{b,b'}^Z(\rho_0) = (1 - p_z)\rho_0 + p_z\tilde{Z}_b\tilde{Z}_{b'}\rho_0\tilde{Z}_b\tilde{Z}_{b'}, \quad (94)$$

$$\mathcal{E}_{g,g'}^Z(\rho_0) = (1 - p_z)\rho_0 + p_z\tilde{Z}_g\tilde{Z}_{g'}\rho_0\tilde{Z}_g\tilde{Z}_{g'}, \quad (95)$$

with $b, b' \in \mathcal{R}_B$ and $g, g' \in \mathcal{R}_G$. Consequently, characterizing the possible spontaneous symmetry-breaking patterns of D_4 is equivalent to the characterization of the possible ways on which the decoherence channels in Eqs. (81) affect D_4 TO.

In this case, one can diagnose sw-SSB by the non-vanishing of the Rényi-2 correlators

$$R_{D_4}^{(2)}(b, g) = \frac{\text{tr}(\rho Z_b Z_g \tilde{C} Z_{bg} \rho Z_b Z_g \tilde{C} Z_{bg})}{\text{tr}(\rho^2)} \quad (96)$$

and

$$R_G^{(2)}(g, g') = \frac{\text{tr}(\rho \tilde{Z}_g \tilde{Z}_{g'} \rho \tilde{Z}_g \tilde{Z}_{g'})}{\text{tr}(\rho^2)}, \quad (97)$$

$$R_B^{(2)}(b, b') = \frac{\text{tr}(\rho \tilde{Z}_b \tilde{Z}_{b'} \rho \tilde{Z}_b \tilde{Z}_{b'})}{\text{tr}(\rho^2)} \quad (98)$$

in the limit $|g - g'|, |b - b'| \rightarrow \infty$. The former $R_{D_4}^{(2)}(b, g)$ corresponds to the thermal expectation value $\langle Z_b Z_g \tilde{C} Z_{bg}^{(2)} \tilde{C} Z_{bg}^{(3)} \rangle$ evaluated on the three-honeycomb-layer stat-mech model introduced in Eq. (39). This quantity exactly matches the expectation value $\mathcal{W}_{m_R \times m_R}$ (when summed over all sites x, y) signaling the condensation of flux pairs $m_R \times m_R$. The latter two in Eq. (97) correspond to the condensation of $e_G \times e_G$ and $e_B \times e_B$ pairs, respectively, as signaled by the respective order parameters $\mathcal{W}_{e_G \times e_G}$ and $\mathcal{W}_{e_B \times e_B}$. From this perspective, these correlators detect different symmetry-breaking patterns with the corresponding phase diagram resembling that in Fig. 11. A detailed analysis of this phase diagram will be included somewhere else.

Before closing this section, we notice that an alternative, and inequivalent, characterization of sw-SSB is via the so-called fidelity correlator [36] (corresponding to a generalization of the quantum fidelity involving the state ρ). Unlike for Rényi-2 correlators, this quantity acquiring a finite value permits one to prove a stability theorem that establishes sw-SSB as a universal mixed-state property [36]. Similar to the relation, through Kramers-Wannier duality, between the toric code ground state under bit-flip noise and the \mathbb{Z}_2 sw-SSB [7,9,36], we notice that the fidelity correlator to diagnose D_4 sw-SSB maps to the quantum fidelity defined in Eq. (63), after replacing the closed noncontractible loop \mathcal{X} by the open string $\mathcal{X}_{t_i}^{f_i}$ and performing the ungauging map in Sec. II B. Hence, in

combination to our findings in Sec. IV E for $p_G = p_B = 0$, it is plausible that the critical point for D_4 sw-SSB is attained at maximum error rate.

VII. CONCLUSIONS AND OUTLOOK

We have characterized the effect of decoherence and wave-function deformation for a microscopic realization of D_4 TO. For the case of proliferating Abelian anyons, we recovered the phenomenology found in previous works for the toric code ground state, where the random bond Ising model governs the (intrinsic) quantum correction threshold p_c [6,56]. Our main focus is on the proliferation of non-Abelian anyons. We found that, even if we decohere *all* possible non-Abelian anyons of a certain type (here, m_R), the D_4 TO is infinitely robust as characterized by the purity. However, a finite threshold exists for the largest moment $\text{tr}(\rho^\infty)$ of the decohered density matrix ρ , and we further argued that $p_c^{(\infty)} = 1/2$ (i.e., even for $n \rightarrow \infty$ we lose only D_4 TO at the maximal decoherence rate). This scenario contrasts to the behavior of the toric code ground state, where the finite threshold $p_c^{(n)}$ monotonically *increases* with the Rényi-index n [6]. We argued that the underlying cause is the presence of a nontrivial loop weight appearing for non-Abelian anyons. Indeed, we showed how loop models provide a natural stat-mech model for these systems, physically corresponding to the worldlines of the “poisoning” anyons, whose quantum dimensions lead to topological loop weights.

Furthermore, by exactly diagonalizing ρ at maximal error rate $p = 1/2$ for the error proliferating non-Abelian anyons (namely, with quantum dimension $d_{m_R} = 2$), we found that its spectrum is given by a random $O(2)$ loop model. We equivalently write this as a random D_4 symmetric four-state clock model, at a tuning parameter where the clean model is critical. This can be further pushed into the extended gapless phase by initially considering a deformed wave function with a finite correlation length. We argued that the presence of disorder likely lowers the ordering temperature compared to the clean model, implying that the non-Abelian quantum memory might be stable to arbitrary decoherence of m_R anyons—although future work will have to clarify the effect of the frustration of the model. As a concrete diagnostic, we showed that the robustness of this non-Abelian memory (in terms of the quantum fidelity between two topologically distinct ground states) is encoded in the free energy cost of inserting a symmetry defect line in this disordered D_4 rotor model.

We then considered the action of various types of Pauli errors proliferating both Abelian and non-Abelian anyons, the former appearing as fusion channels of the latter. This gave rise to a rich phase diagram both for pure wave-function deformations (Fig. 6) as well as for the effect of

decoherence as characterized by the purity (Fig. 11). These were obtained by a combination of analytical results pertaining the analysis of the underlying stat-mech models in certain regimes and numerical analysis via Monte Carlo methods. Here, we found that different (condensation) transitions can be diagnosed by local order parameters of the underlying stat-mech model. An important finding for the deformed wave function was the appearance of an extended gapless phase showcasing power-law decaying correlations of the Wilson operators associated with the proliferated anyons. It would be interesting for future work to further explore the physical reason for this critical phase. It is suggestive that it arises due to an inability to condense m_R anyons in isolation, which require a formalism beyond those of the recent Refs. [17,18], which discussed the emergence of nonmodular TOs in mixed states. Moreover, we notice that this gapless phase is reminiscent of that appearing in Ref. [20]. It would be interesting to understand whether the tools discussed in this work can be used to investigate that setup. For the decohered density matrix, we then found that, even when non-Abelian errors are maximally proliferated on top of D_4 TO, the system is still robust to a finite proliferation of Abelian anyons. The characterization of the resulting mixed-state “phases” appearing beyond these thresholds is left as an open question. It would also be worthwhile to analyze the setting where the initial wave function is not at the fixed point of the topologically ordered phase and whether this allows one to tune the effective string tension of the resulting loop model. Finally, we discussed the direct application of these results to the analysis of D_4 strong-to-weak spontaneous symmetry breaking, exploiting a similar relation as the relating \mathbb{Z}_2 sw-SSB to the decoherence transition of the toric code [7,9].

Our work also gives rise to other open questions both regarding the fate of D_4 TO to decoherence, as well as to that of other non-Abelian topological orders. First, it appears that the non-Abelian nature of the phase makes it robust against errors proliferating anyons of nontrivial quantum dimension. In a companion paper [92], we indeed show that decohering quantum doubles with anyons with larger quantum dimension again produces loop models which prevent a straightforward proliferation. This motivates the question: *Are non-Abelian topological orders with only large quantum dimensions more stable?* Second, while here we considered a combination of errors leading to anyons with trivial braiding, it would be interesting to explore whether the case with nontrivial braiding leads to distinct error thresholds. This scenario could be explored by the configuration discussed in Sec. V C, where two or three different non-Abelian anyons with nontrivial braiding are produced. Here, the underlying stat-mech model corresponds to two (or three) coupled $O(2)$ loop models where the braiding leads to a cancellation of certain configurations.

Our work shows that a simple generalization of the Ising model (the ZZCZ model) dictates much of what happens to deforming and decohering D_4 non-Abelian TO. Here, we focused on the “intrinsic” properties of the decohered density matrix as in Ref. [6], agnostic to a particular error-correction scheme. The aforementioned work found that the threshold p_c does indeed match the one obtained from an optimal error-correction protocol (when assuming perfect syndrome measurements). It would be relevant to perform a similar analysis for D_4 TO: Can one indeed correct for arbitrary strong decoherence proliferating (only) m_R anyons? Moreover, how would the presence of faulty measurements affect this conclusion? Does a 3D version of the ZZCZ model then emerge? Moreover, our work shows that detailed studies for non-Abelian decohered models similar to those for the toric code are possible, including the formulation of relevant stat-mech models. With this new insight, it would be exciting to explore whether similar structures can be identified for other non-Abelian TOs.

Finally, while we were able to harness known results about loop models on the honeycomb lattice, our work points the way to various new kinds of loop models [e.g., the coupled $O(2)$ loop models appearing for $\text{tr}(\rho^n)$ in Eq. (49) or the coupled $O(2)$ loop models appearing when including two or three non-Abelian anyons discussed in Sec. V C]. We believe these new stat-mech models can be of interest in their own right, and, in addition, these can, in turn, inform about error thresholds for non-Abelian TO. In particular, what is the fate of the corresponding random models even at the maximum error rate? It would be interesting to perform a numerical study of the fidelity as shown in Eq. (63), an analysis we leave for previous work.

ACKNOWLEDGMENTS

We are grateful to Ehud Altman, Henrik Dreyer, Ruihua Fan, Paul Fendley, Tarun Grover, Wenjie Ji, Yujie Liu, Roger Mong, Lesik Motrunich, Benedikt Placke, John Preskill, Daniel Ranard, Ramanjit Sohal, Nathanan Tantivasadakarn, Robijn Vanhove, and Sagar Vijay for helpful discussions and feedback. We also thank Yizhi You for a previous collaboration on a project about strong-to-weak symmetry breaking. P.S. acknowledges the Les Houches School “Topological order: Anyons and Fractons” and its participants for their insightful lectures and discussions. This work was partially conceived at the Aspen Center for Physics (P.S. and R.V.), which is supported by National Science Foundation (NSF) Grant No. PHY-2210452 and Durand Fund. P.S. and J.A. acknowledge support from the Caltech Institute for Quantum Information and Matter, an NSF Physics Frontiers Center (NSF Grant No. PHY-1733907), and the Walter Burke Institute for Theoretical Physics at Caltech. The U.S. Department of Energy, Office of

Science, National Quantum Information Science Research Centers, Quantum Science Center partially supported the field theory analysis of this work.

APPENDIX A: PURE WAVE-FUNCTION DEFORMATIONS AND GROUND-STATE PHASE TRANSITIONS

Let us consider pure wave-function deformations of the form

$$|\psi(\beta)\rangle = e^{\beta/2 \sum_{j \in \mathcal{R}} \hat{\sigma}_j} |\text{TO}\rangle \quad (\text{A1})$$

with \mathcal{R} a subset of the lattice, $\hat{\sigma}_j$ a Pauli or combination of Pauli operators of different strength, and $\beta > 0$. In the main text, we focus on the norm of $|\psi(\beta)\rangle$ as a function of β , which can be interpreted as the partition function of a classical stat-mech model

$$\mathcal{Z}(\beta) \equiv \langle \psi(\beta) | \psi(\beta) \rangle. \quad (\text{A2})$$

One may wonder whether singularities in $\mathcal{Z}(\beta)$, i.e., finite-temperature phase transitions in the $(2+0)$ -dimensional classical stat-mech model, necessarily correspond to ground-state phase transitions in the corresponding parent Hamiltonians. We claim that these are equivalent as long as the corresponding stat-mech model involves only local interactions as we argue in the following. Let us assume that the partition function can be written as

$$\mathcal{Z}(\beta) = \langle \psi(\beta) | \psi(\beta) \rangle = \sum_{\{\sigma\}} e^{-\beta H(\{\sigma\})} \quad (\text{A3})$$

with $H(\{\sigma\}) = \sum_r h_r$ a local spin Hamiltonian with $h_r = -\frac{1}{2} \sum_r \sigma_r \sum_{j \in \mathcal{N}_r} \sigma_j$ in terms of “Ising”-like variables σ_r . Then, diagonal correlations (i.e., those involving only \hat{Z} Paulis) are given by thermal correlations evaluated on the corresponding stat-mech model:

$$\langle \psi(\beta) | Z_r Z_{r'} | \psi(\beta) \rangle = \langle \sigma_r \sigma_{r'} \rangle_\beta. \quad (\text{A4})$$

Moreover, off-diagonal correlations can also be computed as a thermal average of *local* observables [107]:

$$\begin{aligned} \langle X_r X_{r'} \rangle &= \frac{1}{\mathcal{Z}} \sum_{\{\sigma\}} e^{-\beta H(\{\sigma\})} e^{\beta h_r} e^{\beta h_{r'}} \\ &= \langle e^{\beta h_r} e^{\beta h_{r'}} \rangle_\beta \end{aligned} \quad (\text{A5})$$

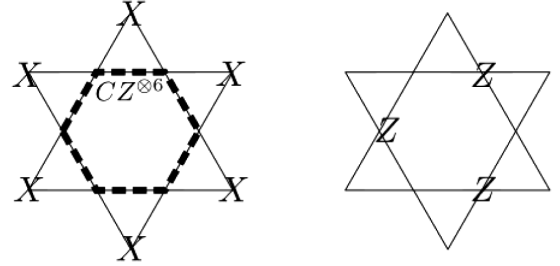
and

$$\begin{aligned} \langle Y_r Y_{r'} \rangle &= \frac{1}{\mathcal{Z}} \sum_{\{\sigma\}} e^{-\beta H(\{\sigma\})} (i\sigma_r e^{\beta h_r}) (i\sigma_{r'} e^{\beta h_{r'}}) \\ &= -\langle \sigma_r e^{\beta h_r} \sigma_{r'} e^{\beta h_{r'}} \rangle_\beta. \end{aligned} \quad (\text{A6})$$

On the one hand, we recall that a ground-state phase transition corresponds to closing the gap, and this affects the scaling of certain correlation functions, e.g., modifying their scaling from exponential, to constant or to power-law decay. On the other, it is clear that phase transitions in $\mathcal{Z}(\beta)$ (as a function of β) are captured by the scaling of diagonal correlations, but as we showed before, the stat-mech model can also be used to evaluate off-diagonal ones.

APPENDIX B: (UN)GAUGING AND (DIS) ENTANGLING MAPPINGS

The ground state of the D_4 topological order model in Eq. (1) is specified by the following stabilizers (including the mirrored version) equaling unity:



It is useful to keep track of the three \mathcal{R}_R , \mathcal{R}_G , and \mathcal{R}_B sublattices of the model (on the kagome lattice) as shown in Fig. 13. A key property we use is that if, e.g., X appears only on red, then blue and green have conserved $B_t = Z^{\otimes 3}$, which we can use to dualize (i.e., gauge the 1-form or ungauged the \mathbb{Z}_2 symmetry).

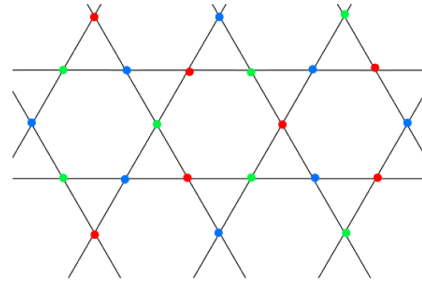
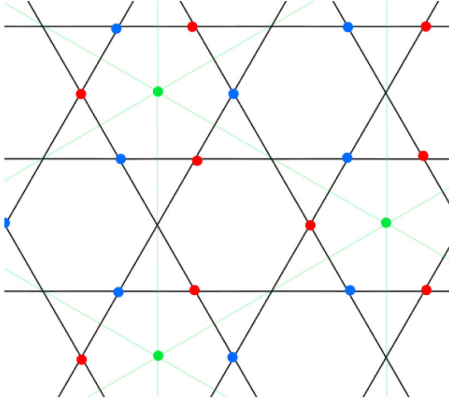


FIG. 13. Kagome lattice. Three color sublattices red, green, and blue.

1. “Ungauging” the green sublattice

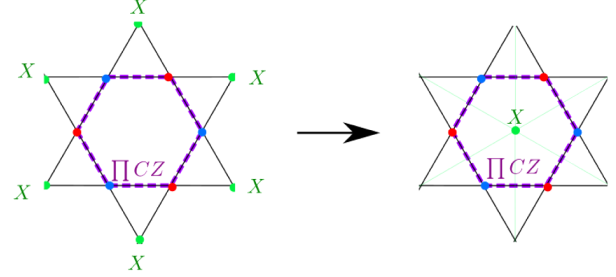
Doing this first for green, we use $Z^{\otimes 3} = 1$ to map the problem to the following lattice:



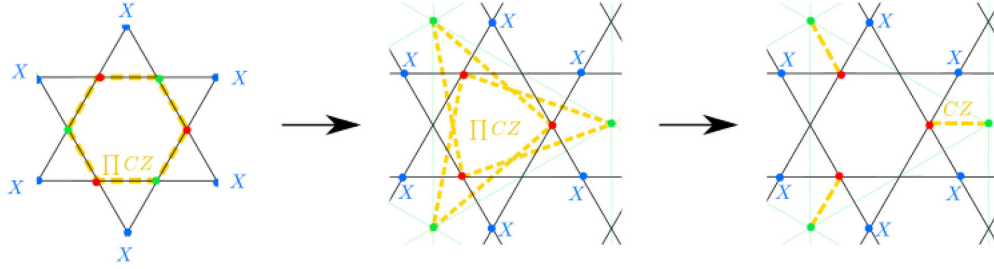
where a green Z on the original lattice maps to ZZ on the new lattice. Now our state has mapped to $|D_4\rangle \rightarrow |R\tilde{G}B\rangle$.

We now work out the structure of $|R\tilde{G}B\rangle$ by specifying its stabilizers on this new lattice.

The stabilizers for this state are as follows: The simplest ones are the $Z^{\otimes 3}$ stabilizers on the red and blue sublattices, which have not changed at all. Second, the 12-body stabilizer whose X 's were living on the green sublattice now map to the following:



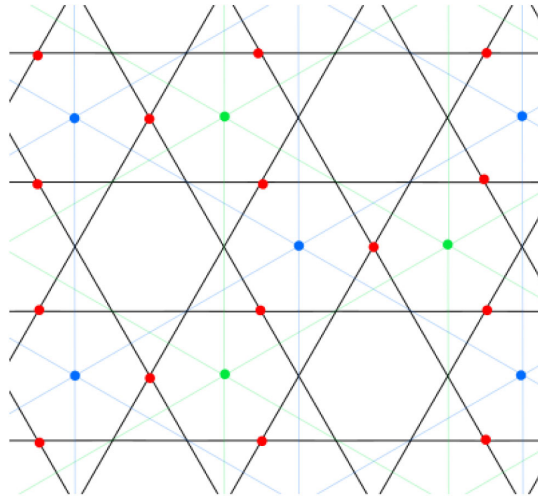
Finally, the other 12-body stabilizers (i.e., where the X 's are on blue or red) map as follows:



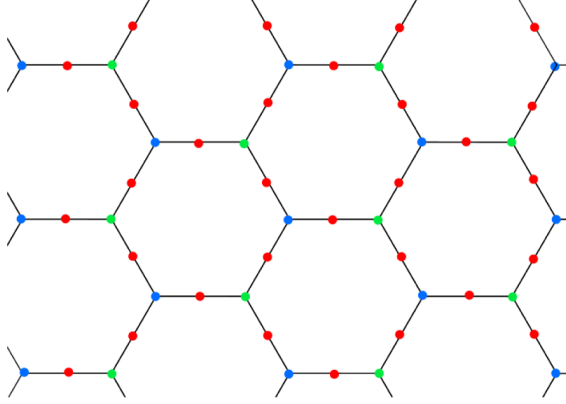
where in the last step we use the $Z^{\otimes 3} = 1$ stabilizer to simplify the CZ 's. A similar stabilizer of course holds upon replacing blue and red.

2. Ungauging the blue sublattice

We now use the $Z^{\otimes 3}$ stabilizer on the blue sublattice to similarly map to the following lattice, where a single blue Z maps to two ZZ 's on the new blue sublattice:

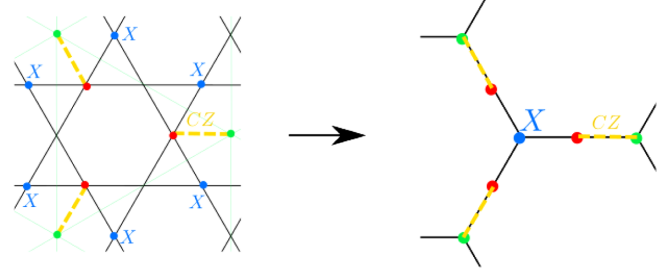


Since this is starting to get a bit messy, note that we can also draw these points as lying on the following Lieb (“heavy hex”) honeycomb lattice:

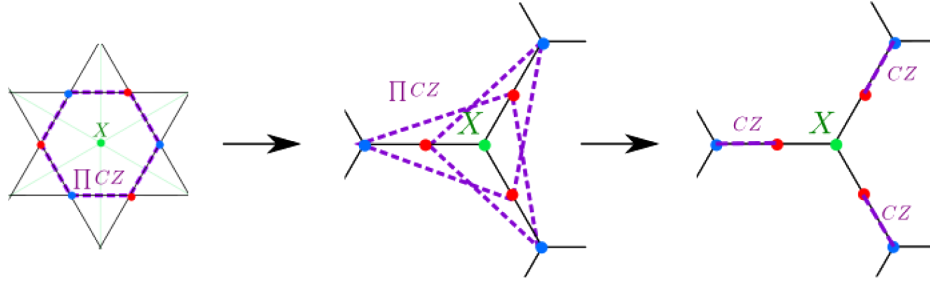


Our wave function on this lattice then becomes $|R\tilde{G}\tilde{B}\rangle$, with stabilizers:

- (1) The red sublattice still has the usual $Z^{\otimes 3}$ stabilizer.
- (2) For every blue vertex, we have

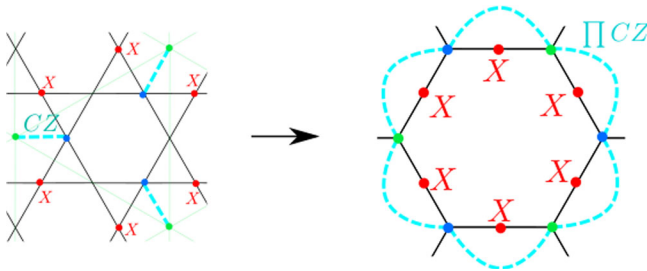


- (3) For every green vertex, we have



where the last step again involves using the $Z^{\otimes 3}$ stabilizer. The resulting stabilizer is equivalent to the blue one (after exchanging blue and green), which confirms that blue and green are playing equivalent roles in this formulation.

- (4) Lastly, there is an X -type stabilizer associated to the red sublattice:



3. (Dis)entangling

Thus far, we got $|R\tilde{G}\tilde{B}\rangle$ on a honeycomb Lieb lattice (i.e., vertices and bonds of the honeycomb lattice), where $|R\tilde{G}\tilde{B}\rangle$ is specified by the above list of stabilizers.

We now further simplify this state by acting with a (dis)entangler U consisting of a three-body CCZ gate on every bond. Note that every bond has three qubits: blue and green vertex qubits and a red bond qubit. We can, thus, write

$$U = \prod_{\langle v, v' \rangle} CCZ_{v, \langle v, v' \rangle, v'}. \quad (\text{B1})$$

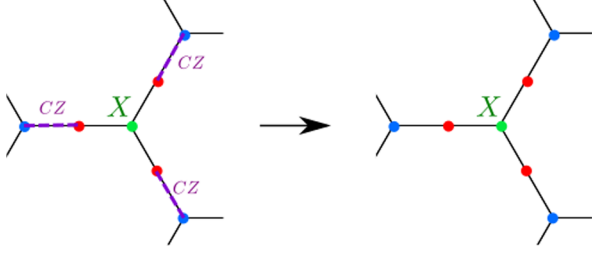
The action of this unitary on our state is quite simple, by just repeatedly using the fact that conjugating X_i (or X_j) by $CCZ_{i,j,k}$ gives $X_i CZ_{j,k}$ (or $X_j CZ_{i,k}$). Thus,

$$U|\psi(\beta)\rangle \propto e^{\beta \sum_{r \in R} X_r CZ_{b,g}} (U|R\tilde{G}\tilde{B}\rangle), \quad (\text{B2})$$

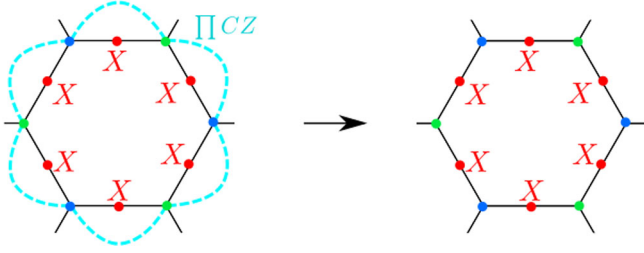
where $X_r CZ_{b,g}$ is to be understood as a three-body operator on a single bond, acting with X on the red bond qubit and acting with CZ on the blue and green vertices of that bond, and where $U|R\tilde{G}\tilde{B}\rangle$ is defined by the following list of stabilizers:

- (1) The red sublattice still has the usual $Z^{\otimes 3}$ stabilizer.

- (2) For every blue and green vertex, we just have the stabilizers $X_b = 1$ and $X_g = 1$. Explicitly, for the green sublattice,



- (3) Lastly, the red X stabilizer also disentangles:



In conclusion, at this point, $U|R\tilde{G}\tilde{B}\rangle$ consists of *trivial product states* on the blue and green sublattices and a usual toric code state on the red sublattice.

4. Dualizing and then ungauging the red sublattice

Since our imaginary time evolution does not commute with the red $Z^{\otimes 3}$ stabilizer of our initial state, we cannot use the latter to ungauging the model, unlike what we did for the green and blue sublattices. However, we can do so if we first perform an $e - m$ duality, i.e., using the X instead of Z stabilizer. In particular, note that the red six-body $X^{\otimes 6}$ stabilizer *does* commute with our imaginary time evolution. This means we can perform a Kramers-Wannier transformation such that the red qubits now live on the sites of the honeycomb lattice. For convenience, we choose the convention that a bond X qubit maps to a ZZ interaction on the new lattice. This maps the toric code on the red sublattice into a $|+\rangle^{\otimes N}$ product state. Note that the effective lattice is a honeycomb lattice with *two* qubits per site. If we denote the Pauli operators as X, Z and \tilde{X}, \tilde{Z} , then we can write our resulting state as follows:

$$|\psi(\beta)\rangle \propto e^{\beta \sum_{\langle v, v' \rangle} \tilde{Z}_v \tilde{Z}_{v'} CZ_{v, v'}} |+\rangle^{\otimes N} \otimes |\tilde{+}\rangle^{\otimes N}. \quad (\text{B3})$$

We can think of this as a *honeycomb bilayer*, where the coupling consists of $\tilde{Z} \tilde{Z}$ on the bond of one layer coupled to CZ on the bond of the other layer.

The result of applying these mappings to the Wilson operator $X_{t_i}^{t_f}$ in Eq. (2) is provided in Fig. 14.

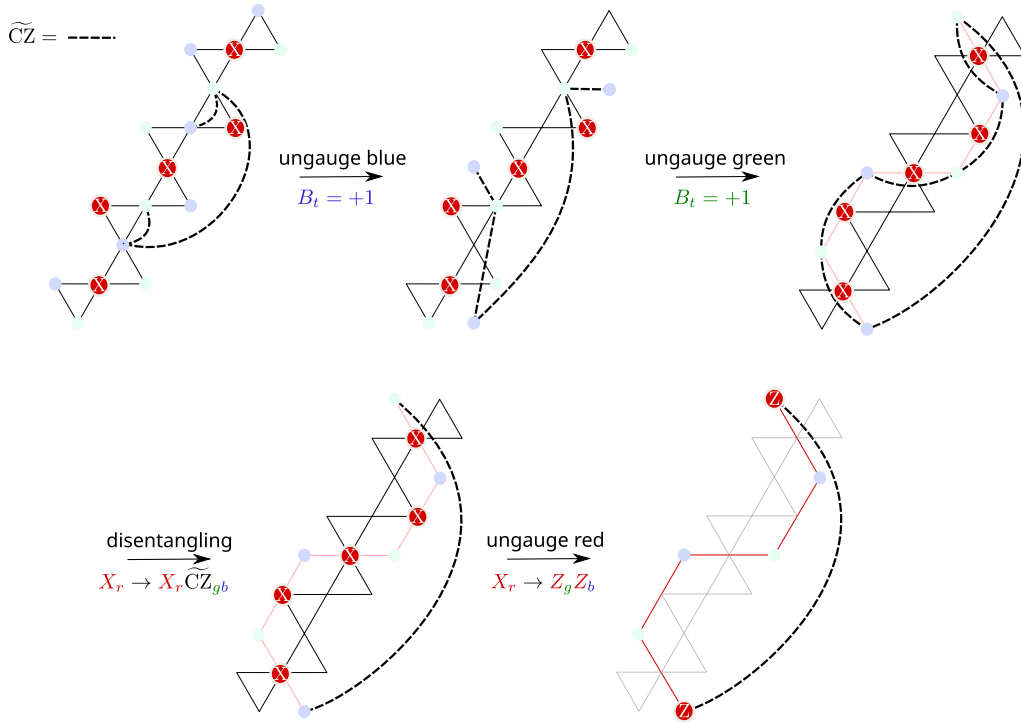


FIG. 14. Ungauging $X_{t_i}^{t_f}$ in Eq. (2).

APPENDIX C: FROM D_4 TOPOLOGICAL ORDER TO $O(N)$ LOOP MODELS: ANALYTICAL DERIVATIONS

This appendix contains the analytical derivations showing that the action of pure wave-function deformations and quantum channels acting on the D_4 topological ordered ground state of Hamiltonian in Eq. (1) with all logical operators $\mathcal{Z} = +1$ can be exactly mapped to $O(N)$ loop models. We show that N is given by the quantum dimension d_a (respectively, d_a^2) of the anyon the said deformation (quantum channel) can proliferate. We provide additional details to the derivations presented in the main text.

1. Non-Abelian error X

In this section, we compute the expectation value $\langle D_4 | \prod_{j \in L_R} X_j | D_4 \rangle$. To do so, we make use of the ungauging and disentangling maps introduced in Appendix B, leading to

$$\langle D_4 | \prod_{j \in L_R} X_j | D_4 \rangle = \langle \tilde{+} | \langle \text{TC} | \prod_{e \in L_R} X_e \widetilde{\text{CZ}}_{e_+, e_-} | \tilde{+} \rangle | \text{TC} \rangle, \quad (\text{C1})$$

where, in the first equality, e corresponds to a bond of the honeycomb lattice as in Fig. 1 with vertex boundaries denoted by e_{\pm} , the latter hosting spin-1/2's on the state $|\tilde{+}\rangle$ along X , and with $\widetilde{\text{CZ}}_{i,j} = \frac{1}{2}(\mathbb{1} + \tilde{Z}_i + \tilde{Z}_j - \tilde{Z}_i \tilde{Z}_j)$ coupling B and G sublattices. Notice that this operator is left invariant under the exchange $\tilde{Z}_i \leftrightarrow \tilde{Z}_j$. Using

$$\begin{aligned} & \langle \tilde{+} | \langle \text{TC} | \prod_{e \in L_R} X_e \widetilde{\text{CZ}}_{e_+, e_-} | \tilde{+} \rangle | \text{TC} \rangle \\ &= \langle \text{TC} | \prod_{e \in L_R} X_e | \text{TC} \rangle \langle \tilde{+} | \prod_{e \in L_R} \widetilde{\text{CZ}}_{e_+, e_-} | \tilde{+} \rangle, \quad (\text{C2}) \end{aligned}$$

we again find that only closed-loop configurations L_R (now on the honeycomb lattice) lead to nonvanishing $\langle \text{TC} | \prod_{e \in L_R} X_e | \text{TC} \rangle = 1$. Everything that remains is to evaluate $W_{L_R} \equiv \langle \tilde{+} | \prod_{e \in L_R} \widetilde{\text{CZ}}_{e_+, e_-} | \tilde{+} \rangle$ on such closed-loop configurations. In the following, we prove that $W_{L_R} = 2^{C_{L_R}} / \sqrt{2}^{|L_R|}$, where C_{L_R} corresponds to the number of disconnected components in L_R . Let us denote by $\ell \in L_R$ a connected component of L_R , and define $\text{CZ}_{\ell} = \prod_{e \in \ell} \text{CZ}_{e_+, e_-}$. Then, one finds

$$W_{L_R} = \prod_{\ell \in L_R} \frac{1}{2^{|\ell|}} \text{tr}(\text{CZ}_{\ell}). \quad (\text{C3})$$

Hence, we need only to evaluate $\text{tr}(\text{CZ}_{\ell})$ on every connected component ℓ . First of all, CZ being a diagonal (in the Z basis) two-body gate, it can be written as

$$\begin{array}{c} \sigma_i \quad \sigma_j \\ \boxed{\text{CZ}} \\ \sigma_i \quad \sigma_j \end{array} = \sqrt{2} \begin{array}{c} \sigma_i \quad \sigma_j \\ \boxed{H} \\ \sigma_i \quad \sigma_j \end{array} \quad (\text{C4})$$

since when $\sigma_i = \sigma_j = -1$ the result is the same configurations multiply by an overall -1 phase and by $+1$ in any other configuration. Here, the symbol

$$\begin{array}{c} \sigma_i \\ | \\ \sigma_j \\ | \\ \sigma_k \end{array} = \delta_{\sigma_i \sigma_j \sigma_k} \quad (\text{C5})$$

corresponds to the identity tensor $\delta_{\sigma_i \sigma_j \sigma_k}$ which is non-vanishing and equals 1 only if $\sigma_i = \sigma_j = \sigma_k$. For example, multiplying two overlapping CZ gates $\text{CZ}_{ij} \text{CZ}_{jk}$ corresponds to

$$\sqrt{2}^2 \begin{array}{c} \sigma_i \quad \sigma_j \quad \sigma_k \\ \boxed{H} \quad \boxed{H} \\ \sigma_i \quad \sigma_j \quad \sigma_k \end{array} \quad (\text{C6})$$

Therefore, we just found that $\text{tr}(\text{CZ}_{\ell_R}) = \sqrt{2}^{|L_R|} \text{tr}(H^{|L_R|})$ with $H = (X + Z)/\sqrt{2}$ the Hadamard matrix. Since the Hadamard transformation H has eigenvalues $\lambda_{\pm} = \pm 1$, we then find that $\text{tr}(\text{CZ}_{\ell}) = 2\sqrt{2}^{|\ell_R|}$ if ℓ_R has even length $|\ell|$ but vanishing otherwise. However, ℓ_R being a closed loop on the honeycomb lattice, its length is always even. Notice that this also implies $(\pm 1)^{|L_R|} = +1$ when summing over closed-loop configurations. Hence,

$$W_{L_R} = \prod_{\ell_R \in L_R} \frac{1}{2^{|\ell_R|}} 2\sqrt{2}^{|\ell_R|} = 2^{C_{L_R}} 2^{-|L_R|/2}. \quad (\text{C7})$$

An alternative way to obtain the value of $\langle D_4 | \prod_{j \in L_R} X_j | D_4 \rangle$ is by mapping the D_4 TO to a \mathbb{Z}_2^3 symmetry-protected topological (SPT) phase following Ref. [78]. This corresponds to applying the “ungauging map” not only to operators commuting with green and blue plaquettes terms $B_i = Z^{\otimes 3}$, but also on the red sublattice, instead of following Appendix B 4. Let us consider a single-loop configuration L_R . First, notice that the product of $X_j \in \mathcal{R}_R$ along the boundary of L_R can be written as the product of the six-body operators $X^{\otimes 6}$ with X acting on the tips of a star \star lying on \mathcal{R}_R :

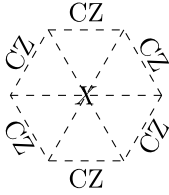
$$\prod_{j \in L_R} X_j = \prod_{\star \in \text{int}(L_R)} X^{\otimes 6}. \quad (\text{C8})$$

Then, following Ref. [78], we map $|D_4\rangle \rightarrow |\text{SPT}\rangle$, and

$$\prod_{\star \in \text{int}(L_R)} X^{\otimes 6} \longrightarrow \prod_{\star \in \text{int}(L_R)} X_{\star} \quad (\text{C9})$$

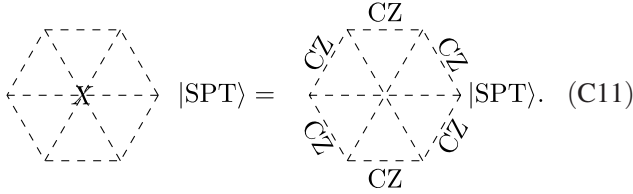
with X_{\star} lying at the center of \star . In fact, this SPT is defined

on a triangular lattice (which is also a trivalent lattice), with the stabilizers on the red sublattice given by



$$|SPT\rangle = |SPT\rangle, \quad (C10)$$

and, hence,



$$|SPT\rangle = |SPT\rangle. \quad (C11)$$

Then, we find that

$$\langle D_4 | \prod_{j \in L_R} X_j | D_4 \rangle = \langle SPT | \prod_{b \in L_R} CZ_b | SPT \rangle, \quad (C12)$$

where we have used the fact that $CZ_b^2 = 1$ and with $b \in L_R$ a bond of the honeycomb lattice lying on L_R . Finally, the SPT can be mapped to a trivial product state $|+\rangle$ on every site of the triangular lattice by applying a CCZ unitary transformation

$$\langle D_4 | \prod_{j \in L_R} X_j | D_4 \rangle = \langle + | \prod_{b \in L_R} CZ_b | + \rangle \quad (C13)$$

and then obtaining the same expression for W_{L_R} as in Eq. (C2).

2. Combined errors

In this section, we combine both Abelian Z and non-Abelian errors X acting on one or several sublattices. As the previous section has made explicit, the difference will come from the evaluation of expectation of the corresponding errors on the $|D_4\rangle$ To ground state.

a. Abelian plus non-Abelian errors on different sublattices

We start by combining an X deformation acting on the red sublattice and Z acting on blue and green sublattices. If $\beta_G^z = 0$, one finds that the norm of the deformed wave function

$$|\psi(\beta_R^x, \beta_B^z)\rangle = e^{\beta_R^x/2} \sum_{j \in \mathcal{R}_R} X_j e^{\beta_B^z/2} \sum_{j \in \mathcal{R}_B} Z_j |D_4\rangle \quad (C14)$$

is given by

$$\begin{aligned} \mathcal{Z}_{|\psi\rangle}(\beta_R^x, \beta_B^z) &= \mathcal{Z}_{0,R} \mathcal{Z}_{0,B} \sum_{\text{closed } L_R} \sum_{\gamma_B} \tanh(\beta_R^x)^{|L_R|} \\ &\times \tanh(\beta_B^z)^{|\gamma_B|} \langle D_4 | \prod_{j \in L_R} X_j \prod_{i \in \gamma_B} Z_i | D_4 \rangle, \end{aligned} \quad (C15)$$

where now the errors are coupled through the last factor. As in Appendix C 1, we know that only closed red loop configurations L_R on the honeycomb lattice lead to a finite contribution, and, hence, we need to evaluate

$$\begin{aligned} W_{L_R}(\gamma_B) &\equiv \langle D_4 | \prod_{j \in L_R} X_j \prod_{i \in \gamma_B} Z_i | D_4 \rangle \\ &= \langle \tilde{+} | \prod_{e \in L_R} \tilde{CZ}_{e_+, e_-} \prod_{b \in \gamma_B} \tilde{Z}_{b_+} \tilde{Z}_{b_-} | \tilde{+} \rangle, \end{aligned} \quad (C16)$$

where once again γ_B collects the coordinates of the bonds on the (blue) triangular lattice—or, equivalently, blue sites of the original kagome lattice—which corresponds to a sublattice of the honeycomb. Now we need to distinguish two possible scenarios. The case where L_R and γ_B do not intersect reduces to the discussion in the previous sections, with γ_B required to be a closed-loop configuration. On the other hand, one can wonder whether γ_B needs to be closed if they intersect, given that an L_R loop contains e_B Abelian anyons as coming from the fusion channel $m_R \times m_R = 1 + e_B + e_G + e_B e_G$. To compute $W_{L_R}(\gamma_B)$, we recall that $\text{tr}(CZ_{L_R}) = \sqrt{2}^{|L_R|} \text{tr}(H^{|L_R|})$, and, hence, the weight requires calculating quantities of the form

$$\text{tr}(HH \dots HZH \dots HZH \dots HH \dots H) \quad (C17)$$

with an even number of H for any closed-loop configuration L_R on the honeycomb lattice. Using the cyclicity of the trace, that $H^2 = 1$, and finally that $ZH = HX$, one finds that Eq. (C17) is nonvanishing only if there are an even number of Z 's. Hence, γ_B needs to intersect (on the blue sublattice) each single connected component ℓ_R of $L_R = \cup \ell_R$ an even number of times. Let us assume the coordinates of these points along one of these components are given by $\{b_j\}_{j=1}^{2n}$ with $b_1 < b_2 < \dots < b_{2n}$. Then,

$$\begin{aligned} &\langle \tilde{+} | \prod_{e \in \ell_R} \tilde{CZ}_{e_+, e_-} \prod_{j=1}^{2n} \tilde{Z}_{b_j} | \tilde{+} \rangle \\ &= \frac{\sqrt{2}^{|\ell_R|}}{2^{|\ell_R|}} \text{tr}(ZH^{b_2-b_1} ZH^{b_3-b_2} \dots ZH^{|\ell_R|-(b_{2n}-b_1)}), \end{aligned} \quad (C18)$$

where H lies on the links of the honeycomb lattice. This calculation can be efficiently performed by introducing a graphical notation shown in Fig. 15. Notice that there are always an even number of H in between two consecutive Z 's, and using $H^2 = \mathbb{1}$ we find that

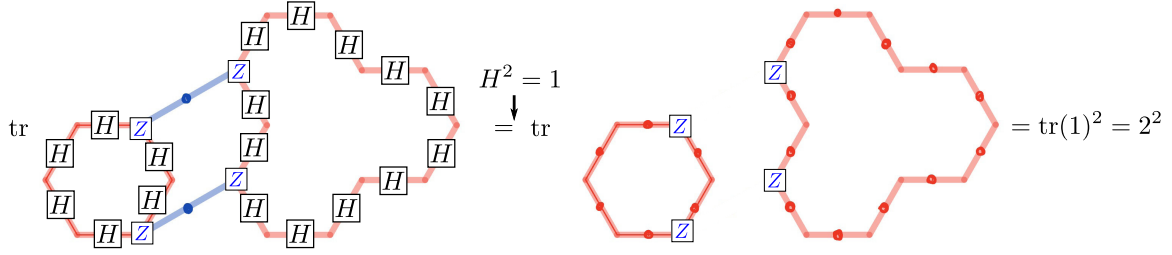


FIG. 15. Example of the computation of the weight in Eq. (C18).

$\langle \tilde{\tau} | \prod_{e \in \ell_R} \tilde{CZ}_{e_+, e_-} \prod_{j=1}^{2n} \tilde{Z}_{b_j} | \tilde{\tau} \rangle = 2\sqrt{2}^{-|\ell_R|}$. Hence, we find the weight

$$W_{L_R}(\gamma_B) = \frac{2^{C_{L_R}}}{\sqrt{2}^{|\ell_R|}}, \quad (\text{C19})$$

whenever either closed loops in γ_B do not intersect with L_R or if components of γ_B end on L_R . See Fig. 1. This result can be understood from the fact that indeed e_B anyons are contained in the fusion channel of $m_R \times m_R$, and, hence, a loop γ_B can close on an L_R loop. Putting it all together, we find

$$\begin{aligned} \mathcal{Z}_{|\psi\rangle}(\beta_R^x, \beta_B^z) &= \mathcal{Z}_{0,R} \mathcal{Z}_{0,B} \sum_{L_R} \sum_{\gamma_B} \tanh(\beta_B^z)^{|\gamma_B|} \\ &\times \left(\frac{\tanh(\beta_R^x)}{\sqrt{2}} \right)^{|\ell_R|} 2^{C_{L_R}}, \end{aligned} \quad (\text{C20})$$

where $\sum_{\text{"closed"} \gamma_B}$ is a sum over all loop configurations in the triangular blue sublattice that either are closed or end up on a L_R loop. These correspond to e_B anyons lying on a closed L_R loop. As a next step, we can add Abelian Z errors on the remaining (green) sublattice. A similar calculation to the one above leads to the classical partition function

$$\mathcal{Z}_{|\psi\rangle}(\beta_R^x, \beta_B^z, \beta_G^z) = \mathcal{Z}_{0,R} \mathcal{Z}_{0,B} \mathcal{Z}_{0,G} \sum_{L_R} \sum_{\gamma_B} \sum_{\gamma_G} \sigma_{L_R}(\gamma_B, \gamma_G) \tanh(\beta_B^z)^{|\gamma_B|} \tanh(\beta_G^z)^{|\gamma_G|} \left(\frac{\tanh(\beta_R^x)}{\sqrt{2}} \right)^{|\ell_R|} 2^{C_{L_R}}, \quad (\text{C21})$$

where now both γ_B and γ_G are loop configurations that can either be closed or end up on an L_R loop configuration on the blue and green sublattices, respectively. Moreover, unlike for vanishing β_G^z , there is an additional sign contribution $\sigma_{L_R}(\gamma_B, \gamma_G)$ which comes from evaluating a weight analogous to that in Eq. (C18) for every connected component γ of L_R , which now includes both Z as well as \bar{Z} insertions. For a given γ_B , γ_G , and L_R , this sign is then given by

$$\sigma_{L_R}(\gamma_B, \gamma_G) = \text{sgn} \left[\prod_{\ell_R \in L_R} \text{tr}(ZHZZ \cdots HZHZZ) \right], \quad (\text{C22})$$

where every nonvanishing factor contains in total an even number of Z and \bar{Z} insertions. However, as explained in the main text, a closed expression for this sign is not necessary to characterize the phase diagram of the deformed wave function. Moreover, notice that by construction an H needs to appear in between insertions of different color, as they lie on different sublattices. The analogous result for the decohered mixed state reads

$$\begin{aligned} \text{tr}[\rho(r_R^x, r_B^z, r_G^z)^2] &\propto \sum_{L_R} \sum_{\gamma_B} \sum_{\gamma_G} (r_B^z)^{|\gamma_B|} (r_G^z)^{|\gamma_G|} \\ &\times \left(\frac{r_R^x}{2} \right)^{|\ell_R|} 4^{C_{L_R}}, \end{aligned} \quad (\text{C23})$$

i.e., three coupled $O(N)$ loop models with $N = 1, 4$.

b. Only non-Abelian errors on different sublattices

We now consider the scenario of two non-Abelian errors acting on different sublattices (e.g., red and blue):

$$|\psi(\beta_R^x, \beta_B^x)\rangle = e^{\beta_R^x/2 \sum_{r \in \mathcal{R}_R} X_r} e^{\beta_B^x/2 \sum_{b \in \mathcal{R}_B} X_b} |D_4\rangle. \quad (\text{C24})$$

In this case, the classical partition function reads

$$\begin{aligned} \mathcal{Z}_{|\psi\rangle}(\beta_R^x, \beta_B^x) &= \mathcal{Z}_{0,R} \mathcal{Z}_{0,B} \sum_{L_R} \sum_{L_B} \tanh(\beta_R^x)^{|\ell_R|} \\ &\times \tanh(\beta_B^x)^{|\ell_B|} \langle D_4 | \prod_{r \in L_R} X_r \prod_{b \in L_B} X_b | D_4 \rangle, \end{aligned} \quad (\text{C25})$$

where we have used the fact that only closed-loop configurations L_R and L_B in either the red or blue sublattices, respectively, have nonvanishing contributions. Nonetheless, we still need to evaluate

$$W(L_R, L_B) = \langle D_4 | \prod_{r \in L_R} X_r \prod_{b \in L_B} X_b | D_4 \rangle. \quad (\text{C26})$$

To do so, we notice that products of X 's around a closed red (blue) loop can be written as a product of six-body operators $X^{\otimes 6}$ acting on the vertices of stars \star and lying within the loop. Hence, we can follow the same approach shown at the end of Appendix C 1, map $|D_4\rangle$ to a \mathbb{Z}_2^3 SPT defined on a triangular lattice, and finally apply the CCZ disentangling transformation in Appendix B 3. One then finds

$$W(L_R, L_B) = \langle + | \text{CZ}_{L_R} \text{CZ}_{L_B} | + \rangle, \quad (\text{C27})$$

similar to the result in Appendix C 1, where CZ_{L_R} is defined as the product of two-qubit CZ gates along the loop configuration L_R . However, unlike in that section, the weight includes two different loops L_R and L_B which can now overlap as shown in Fig. 12. Hence, we should rather consider the global loop configuration $L_R \oplus L_B$ given by the union. Let us denote by ℓ a connected component of $L_R \oplus L_B$ lying on the intertwined red and blue honeycomb lattices as in Fig. 12. Then, $W(L_R, L_B) = \prod_{\ell \in L_R \oplus L_B} W_\ell$, with W_ℓ given by

$$W_\ell = \frac{1}{2^{|\ell|-|V|}} \text{tr} \left(\prod_{e \in \ell} \text{CZ}_e \right) = \frac{2^{|V|}}{\sqrt{2}^{|\ell|}} \text{tr} \left(\prod_{e \in \ell} H_e \right), \quad (\text{C28})$$

where $|V(\ell)|$ is the number of vertices of the resulting planar graph drawn by ℓ (namely, vertices where L_R and L_B intersect). The difference to the calculation for a single type of error is that now ℓ is a closed-loop configuration on the intersection of two honeycomb lattices. However, we can then use the fact that, on every connected component ℓ , only an even number of Hadamards can appear between intersections of two loops and, hence, can be paired up, resulting in a factor of 2 for every (colorless) connected component. Hence, we find

$$W_\ell = \frac{2^{|V(\ell)|+1}}{\sqrt{2}^{|\ell|}} = \frac{2^{|V(\ell)|+1}}{\sqrt{2}^{|\ell|}}, \quad (\text{C29})$$

where we have used that for a planar graph Euler's theorem applies together with the fact that the number of vertices is twice the number of edges, leading to the relation $|V(\ell)| + 1 = |F(\ell)| - 1 \equiv C_\ell$, with C_ℓ the number of loops or also called cyclomatic number, namely, the minimum number of edges that must be removed from ℓ to

break all its cycles, making it into a tree. In the example in Fig. 12, $C_{L_R \oplus L_B} = 4$.

APPENDIX D: STAT-MECH FORMULATION OF $\text{tr}(\rho^n)$ FOR NON-ABELIAN NOISE

1. Coupled O(2) loop model representation

Let us consider the decohered density matrix

$$\rho = (1-p)^{|\mathcal{R}_R|} \sum_E t^{|E|} \left(\prod_{r \in E} X_r \right) |D_4\rangle \langle D_4| \left(\prod_{r \in E} X_r \right), \quad (\text{D1})$$

appearing as a result of applying an X Pauli channel on the sublattice \mathcal{R}_R . The goal of this appendix is to compute $\text{tr}(\rho^n)$ for any n and provide representation in terms of a local stat-mech model. Here, $t \equiv [p_R/(1-p_R)]$, and E is a collection of sites in \mathcal{R}_R . To compute $\text{tr}(\rho^n)$, we need to evaluate the overlaps $\langle D_4 | \prod_{r \in E \oplus E'} X_r | D_4 \rangle$ that we already encountered in the main text to give

$$\langle D_4 | \prod_{r \in E \oplus E'} | D_4 \rangle = f(E \oplus E') \equiv \frac{2^{C(E \oplus E')}}{\sqrt{2}^{|E \oplus E'|}}, \quad (\text{D2})$$

when $E \oplus E'$ is a contractible closed-loop configuration and zero otherwise. To lighten the notation, let us denote $|E\rangle = \prod_{r \in E} X_r |D_4\rangle$. Then, the n th power of ρ reads

$$\rho^n = (1-p)^{n|\mathcal{R}_R|} \sum_{\{E^{(s)}\}_{s=1}^n} \prod_{s=1}^n t^{|E^{(s)}|} |E^{(s)}\rangle \langle E^{(s)}|. \quad (\text{D3})$$

From here, we can compute any moment $\text{tr}(\rho^n)$, which requires one to evaluate the overlaps $\langle E^{(s)} | E^{(s+1)} \rangle$ with $E^{(n+1)} = E^{(1)}$. This explicitly shows that each pair of layers contribute with a topological weight that relates to the quantum dimension $d_{m_R} = 2$ of m_R . From here, we find

$$\text{tr}(\rho^n) = (1-p)^{n|\mathcal{R}_R|} \sum_{\{E^{(s)}\}_{s=1}^n} \prod_{s=1}^n t^{|E^{(s)}|} \prod_{s=1}^n f(E^{(s)} \oplus E^{(s+1)}), \quad (\text{D4})$$

where the sum over $E^{(s)}$ is constrained to those for which $E^{(s)} \oplus E^{(s+1)}$ forms a closed-loop configuration. Notice that this condition implies that all combinations $E^{(s)} \oplus E^{(s')}$ for any $s, s' = 1, \dots, n$ correspond to a contractible loop. This also implies that any $E^{(s)}$ with $s > 1$ is related to $E^{(1)}$ via a contractible closed-loop configuration $L^{(s-1)}$ as $E^{(s)} = E^{(1)} \oplus L^{(s-1)}$. Hence, Eq. (D4) admits the equivalent representation

$$\text{tr}(\rho^n) = (1-p)^{n|\mathcal{R}_R|} \sum_{E^{(1)}} t^{|E^{(1)}|} \sum_{\{L^{(s)}\}_{s=1}^{n-1}} \prod_{s=1}^{n-1} t^{|E^{(1)} \oplus L^{(s)}|} f(L^{(1)}) \left(\prod_{s=1}^{n-2} f(L^{(s)} \oplus L^{(s+1)}) \right) f(L^{(n-1)}). \quad (\text{D5})$$

2. Local spin model representation

For convenience, we start from the “ungauged” representation of the channel as in Sec. VI:

$$\mathcal{E}_{i,j}^X(\rho_0) = (1-p_R)\rho_0 + p_R \hat{h}_{ij} \rho_0 \hat{h}_{ij} \quad \text{where } \hat{h}_{ij} = Z_i Z_j \widetilde{\mathbf{C}} Z_{ij} \quad (\text{D6})$$

and where the initial state is a trivial product state $|\psi_0\rangle\langle\psi_0|$ with $|\psi_0\rangle = |+\rangle|\tilde{+}\rangle|\tilde{+}\rangle$.

Hence, up to an irrelevant prefactor, we have

$$\text{tr}(\rho^n) = \sum_{\{E^{(s)}\}_{s=1}^n} t^{|E^{(1)}|+|E^{(2)}|+\dots+|E^{(n)}|} \langle\psi_0| \prod_{\langle ij \rangle \in E^{(1)} \oplus E^{(2)}} \hat{h}_{ij} |\psi_0\rangle\langle\psi_0| \prod_{\langle ij \rangle \in E^{(2)} \oplus E^{(3)}} \hat{h}_{ij} |\psi_0\rangle\langle\psi_0| \dots \langle\psi_0| \prod_{\langle ij \rangle \in E^{(n)} \oplus E^{(1)}} \hat{h}_{ij} |\psi_0\rangle\langle\psi_0| \quad (\text{D7})$$

with $t = [p_R/(1-p_R)]$. Since $|\psi_0\rangle$ is a superposition of all classical states, we can replace the expectation values by a sum over all classical $h_{ij}^{(s)} = \sigma_i^{(s)} \sigma_j^{(s)} \widetilde{\mathbf{C}} Z_{ij}^{(s)}$ for each of the $s = 1, 2, \dots, n$ expectation values, namely,

$$\langle\psi_0| \prod_{\langle ij \rangle \in E_s \oplus E_{s+1}} \hat{h}_{ij} |\psi_0\rangle \rightarrow \sum_{\{\sigma_j^{(s)}\}} \sum_{\{\tilde{\sigma}_j^{(s)}\}} \prod_{\langle ij \rangle \in E_s \oplus E_{s+1}} h_{ij}^{(s)} \quad (\text{D8})$$

with $E_{n+1} = E^{(1)}$. This gives the following classical stat-mech model:

$$\text{tr}(\rho^n) = \sum_{\{\sigma_j^{(s)}\}_{s=1}^n} \sum_{\{\tilde{\sigma}_j^{(s)}\}_{s=1}^n} \sum_{\{E^{(s)}\}_{s=1}^n} t^{|E^{(1)}|+|E^{(2)}|+\dots+|E^{(n)}|} \prod_{\langle ij \rangle \in E^{(1)} \oplus E^{(2)}} h_{ij}^{(1)} \times \prod_{\langle ij \rangle \in E^{(2)} \oplus E^{(3)}} h_{ij}^{(2)} \times \dots \times \prod_{\langle ij \rangle \in E^{(n)} \oplus E^{(1)}} h_{ij}^{(n)} \quad (\text{D9})$$

$$= \sum_{\{\sigma_j^{(s)}\}_{s=1}^n} \sum_{\{\tilde{\sigma}_j^{(s)}\}_{s=1}^n} \sum_{\{E^{(s)}\}_{s=1}^n} \prod_{\langle ij \rangle \in E^{(1)}} t h_{ij}^{(n)} h_{ij}^{(1)} \times \prod_{\langle ij \rangle \in E^{(2)}} t h_{ij}^{(1)} h_{ij}^{(2)} \times \dots \times \prod_{\langle ij \rangle \in E^{(n)}} t h_{ij}^{(n-1)} h_{ij}^{(n)} \quad (\text{D10})$$

$$= \sum_{\{\sigma_j^{(s)}\}_{s=1}^n} \sum_{\{\tilde{\sigma}_j^{(s)}\}_{s=1}^n} \prod_{\langle ij \rangle \in \bigcirc} \prod_{s=1}^n (1 + t h_{ij}^{(s)} h_{ij}^{(s+1)}) \propto \sum_{\{\sigma_j^{(s)}\}_{s=1}^n} \sum_{\{\tilde{\sigma}_j^{(s)}\}_{s=1}^n} \exp \left(\beta \sum_{\langle i,j \rangle \in \bigcirc} \sum_{s=1}^n h_{ij}^{(s)} h_{ij}^{(s+1)} \right), \quad (\text{D11})$$

where $\tanh(\beta) = t$. In the second step, we have used that

$$\prod_{\langle ij \rangle \in E^{(s)} \oplus E^{(s+1)}} h_{ij}^{(s)} = \prod_{\langle ij \rangle \in E^{(s)}} h_{ij}^{(s)} \prod_{\langle ij \rangle \in E^{(s+1)}} h_{ij}^{(s)}. \quad (\text{D12})$$

This derives Eq. (49) in the main text. We note that tracing out the σ variables in Eq. (D9) would directly rederive the loop model representation in Eq. (D4), thereby establishing a direct link between these two equivalent formulations.

a. Alternative spin model

One can reduce the number of Ising spins $\sigma_i^{(s)}$. Note that if we flip all replicas $\sigma_i^{(s)} \rightarrow -\sigma_i^{(s)}$ for a *fixed* index i , the model is symmetric. This means that, without loss of generality, we can fix the σ spins on the last layer: $\sigma_i^{(n)} = 1$ (up to an irrelevant prefactor of the partition function). There are now $n-1$ σ spins remaining. It is convenient to introduce a change of variables:

$$\tau_i^{(1)} = \sigma_i^{(1)} \sigma_i^{(2)}, \quad \tau_i^{(2)} = \sigma_i^{(2)} \sigma_i^{(3)}, \dots, \quad \tau_i^{(n-2)} = \sigma_i^{(n-2)} \sigma_i^{(n-1)}, \quad \tau_i^{(n-1)} = \sigma_i^{(n-1)}. \quad (\text{D13})$$

Note that this implies $\prod_{s=1}^{n-1} \tau_i^{(s)} = \sigma_i^{(1)} = \sigma_i^{(n)} \sigma_i^{(1)}$. In conclusion, we arrive at

$$\text{tr}(\rho^n) = \sum_{\{\tau_j^{(s)}\}_{s=1}^{n-1}} \sum_{\{\tilde{\sigma}_j^{(s-1/2)}\}_{s=1}^n} \exp\left(\beta \sum_{\langle i,j \rangle_\circ} \left[\sum_{s=1}^{n-1} h_{i,j}^{(s)} + \prod_{s=1}^{n-1} h_{i,j}^{(s)} \right]\right), \quad (\text{D14})$$

where each local term for $s = 1, \dots, n-1$ is given by

$$h_{i,j}^{(s)} = \tau_i^{(s)} \tau_j^{(s)} \widetilde{\text{CZ}}_{ij}^{(s-1/2)} \widetilde{\text{CZ}}_{ij}^{(s+1/2)}; \quad (\text{D15})$$

here, we have labeled the $\tilde{\sigma}$ spins by a half-integer index to emphasize that we can think of them as living in between the τ -spin layers. Note that this formulation closely resembles the structure found for the toric code in Ref. [6]: $n-1$ terms and a last term involving all replicas. In fact, one can recover the same result by setting $\text{CZ} \rightarrow 1$.

APPENDIX E: QUANTUM FIDELITY

The quantum fidelity between two density matrices ρ and σ is defined as

$$F(\rho, \sigma) \equiv (\text{tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}})^2 \quad (\text{E1})$$

and quantifies how distinguishable these density matrices are. Sometimes, $F \rightarrow \sqrt{F}$ is also used as a definition of quantum fidelity. For example, when $\rho = |\psi_\rho\rangle\langle\psi_\rho|$ and $\sigma = |\psi_\sigma\rangle\langle\psi_\sigma|$ are projectors on normalized pure states, then the quantum fidelity agrees with the overlap

$$F(\rho, \sigma) = |\langle\psi_\rho|\psi_\sigma\rangle|^2. \quad (\text{E2})$$

This vanishes when $|\psi_\rho\rangle$ is orthogonal to $|\psi_\sigma\rangle$.

Let us now consider the two initial states $|D_4\rangle$ and $\mathcal{X}|D_4\rangle$ and apply \mathcal{E}^X at maximum error rate $p_R = 1/2$, such that $\rho = \mathcal{E}^X(|D_4\rangle\langle D_4|)$ and $\sigma = \mathcal{E}^X(\mathcal{X}|D_4\rangle\langle D_4|\mathcal{X})$. Using the fact that \mathcal{E}^X can be written as a random projector channel as in Eq. (53), we denote

$$|\eta\rangle = \prod_{r \in \mathcal{R}_R} \frac{1}{2} (1 + \eta_r X_r) |D_4\rangle, \quad (\text{E3})$$

$$|\mathcal{X}, \eta\rangle = \prod_{r \in \mathcal{R}_R} \frac{1}{2} (1 + \eta_r X_r) \mathcal{X} |D_4\rangle. \quad (\text{E4})$$

Then,

$$\rho = \sum_{\{\eta\}} |\eta\rangle\langle\eta|, \quad \sigma = \sum_{\{\eta\}} |\mathcal{X}, \eta\rangle\langle\mathcal{X}, \eta|, \quad (\text{E5})$$

where $\langle\eta|\eta'\rangle = \delta_{\eta,\eta'}$ due to the orthogonality of the projectors $P_{\pm,j}$. Hence,

$$\sqrt{\rho} = \sum_{\{\eta\}} \frac{|\eta\rangle\langle\eta|}{\sqrt{\langle\eta|\eta\rangle}}, \quad (\text{E6})$$

and we, thus, have

$$\sqrt{\rho} \sigma \sqrt{\rho} = \sum_{\{\eta\}} \frac{|\langle\eta|\mathcal{X}, \eta\rangle|^2}{\langle\eta|\eta\rangle} |\eta\rangle\langle\eta|, \quad (\text{E7})$$

which implies

$$\sqrt{\sqrt{\rho} \sigma \sqrt{\rho}} = \sum_{\{\eta\}} \frac{|\langle\eta|\mathcal{X}, \eta\rangle|}{\langle\eta|\eta\rangle} |\eta\rangle\langle\eta|. \quad (\text{E8})$$

One then finds that the (square-root) fidelity equals the average overlap

$$F'(\rho, \sigma) = \sum_{\{\eta\}} |\langle\eta|\mathcal{X}, \eta\rangle| = \sum_{\{\eta\}} P(\eta) \frac{|\langle\eta|\mathcal{X}, \eta\rangle|}{\langle\eta|\eta\rangle} \quad (\text{E9})$$

with

$$P(\eta) = \langle\eta|\eta\rangle = \frac{1}{4^{|\mathcal{R}_R|}} \sum_{L_R} \left(\prod_{e \in L_R} \frac{\eta_e}{\sqrt{2}} \right) 2^{C_{L_R}} \propto \frac{1}{2} [\mathcal{Z}(\eta) - \mathcal{Z}_-(\eta)] \quad (\text{E10})$$

with

$$\mathcal{Z}(\eta) = \sum_{\{\sigma, \tilde{\sigma}\}} \prod_{\langle i,j \rangle} (1 + \eta_{ij} \sigma_i \sigma_j \widetilde{\text{CZ}}_{ij}), \quad (\text{E11})$$

agreeing with the partition function of the disorder O(2) loop model, and $\mathcal{Z}_-(\eta)$ a partition function taking the same form but with a line of flipped antiferromagnetic bonds along a noncontractible loop \mathcal{C} perpendicular to \mathcal{X} . On the other hand, the overlap $\langle\eta|\mathcal{X}, \eta\rangle = \langle\eta|\mathcal{X}|\eta\rangle$ corresponds to a similar partition function but constrained to include an odd number of noncontractible loops on each configuration L_R . This constraint can be naturally accounted for on the local-stat mech model in terms of Ising variables by writing

$$|\langle\eta|\mathcal{X}|\eta\rangle| = \frac{1}{2} |\mathcal{Z}(\eta) - \mathcal{Z}_-(\eta)|. \quad (\text{E12})$$

Indeed, the difference $\mathcal{Z}(\eta) - \mathcal{Z}_-(\eta)$ vanishes whenever an even number of noncontractible closed loops intersects this defect line. Hence, the fidelity becomes

$$\begin{aligned}
F'(\rho, \sigma) &= \frac{1}{2} \sum_{\{\eta\}} P(\eta) \frac{|\mathcal{Z}(\eta) - \mathcal{Z}_-(\eta)|}{\mathcal{Z}(\eta) + \mathcal{Z}_-(\eta)} \\
&= \frac{1}{2} \sum_{\{\eta\}} P(\eta) \frac{|1 - e^{-\Delta F_c}|}{1 + e^{-\Delta F_c}}
\end{aligned} \quad (\text{E13})$$

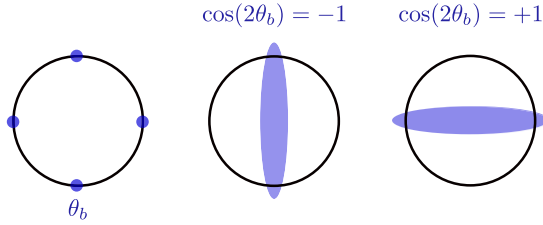
as stated in the main text, with $e^{-\Delta F_c} = \{[\mathcal{Z}_-(\eta)]/[\mathcal{Z}(\eta)]\}$.

APPENDIX F: CONSISTENT FIELD THEORY

A consistent field theory description characterizing the phase diagrams in Fig. 6(b) is given by

$$\begin{aligned}
\mathcal{H} &= \frac{1}{2\pi} \left(K(\partial_x \theta)^2 + \frac{1}{4K} (\partial_x \varphi)^2 \right) - g \cos(8\theta) - \lambda \cos(\varphi) \\
&\quad - J \sigma_g \sin(2\theta) - J \sigma_b \cos(2\theta) - m \varepsilon_g - m \varepsilon_b,
\end{aligned} \quad (\text{F1})$$

when $\beta_G^z = \beta_B^z$. Here, $J \geq 0$; λ and m can take both positive and negative values; and finally σ_g , σ_b and ε_g , ε_b , respectively, correspond to the order and energy fields for the Ising CFT [175,176]. The motivation to include these new Ising variables is the following: If β_B^z is large, $\cos(2\theta_b - 2\theta_b')$ tries to develop long-range order in $\cos(2\theta_b) = \pm 1$. If we think of the four values that θ_b can take (namely, $\theta_b \in \{0, \pi/2, \pi, 3\pi/2\}$), this implies the following two patterns:

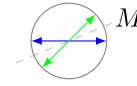


Hence, we expect a type of “nematic” order for large β_B^z , where either the system spontaneously chooses either the horizontal or vertical axis. Note that this nematic ordering would indeed be Ising-like: It would break D_4 down to a remaining $\mathbb{Z}_2 \times \mathbb{Z}_2$, which means only one \mathbb{Z}_2 is spontaneously broken. That is consistent with the solvable limit where $\beta_R^x = \beta_G^z = 0$, because we already know that β_B^z by itself leads to an Ising transition. A similar reasoning applies when β_G^z is large.

For the green sublattice, the \mathbb{Z}_2 Ising symmetry is given by S , which acts as $S: \sigma_g \rightarrow -\sigma_g$ and $\sin(2\theta) \rightarrow -\sin(2\theta)$. For the blue sublattice, this corresponds to RS , which similarly acts on σ_b and $\cos(2\theta)$. The corresponding phase diagram for $\beta_G^z = 0$ is shown in Fig. 6(a). Moreover, under M , σ_g and σ_b are exchanged, and $\sin(2\theta) \leftrightarrow \cos(2\theta)$. When instead $\beta_G^z \neq \beta_B^z$, this field theory includes an additional potential term $\cos(4\theta)$ and the parameters J and m can take a different value on each sublattice, namely, $J \rightarrow J_G, J_B$ and similarly $m \rightarrow m_G, m_B$. Moreover, since the condensation of $\cos(2\theta)$ and $\sin(2\theta)$ gives rise to the Ising

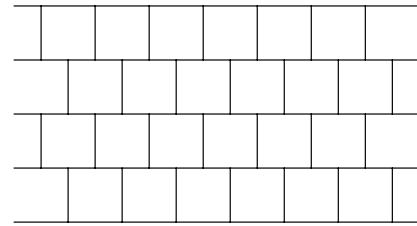
transition that we found in the previous section when tuning $\beta_G^z = \beta_B^z$, we expect that the mass of each species of Ising variables is controlled by $m_G \sim \beta_G^z$, $m_B \sim \beta_B^z$. Once the system undergoes the Ising transition indicated with a horizontal black line in the lower right corner in Fig. 6(b), one can replace $\sigma_b, \sigma_g \rightarrow \text{const}$, and, hence, the terms $\cos(2\theta)$ and $\sin(2\theta)$ are effectively added to the Hamiltonian for the compact boson. Then, \mathcal{W}_{e_b} (and also \mathcal{W}_{e_g} when $\beta_G^z = \beta_B^z$) picks up a finite value, leading to the trivial (toric code) phase in the upper right in Fig. 6(a) [Fig. 6(b)]. At this point, $\cos(\theta)$ has also acquired a finite value.

Hence, once the system undergoes the Ising transition $\sigma_b, \sigma_g \rightarrow \text{const}$, the terms $\cos(2\theta)$ and $\sin(2\theta)$ become relevant perturbations for the compact boson Hamiltonian, pinning the value of θ . Upon minimizing $-\sin(2\theta) - \cos(2\theta)$, one finds that the minimum is attained for $\theta = \pi/8, \pi + \pi/8$, right in between the possible values that θ_g and θ_b can take and along the symmetry axis on which the M symmetry is defined below Eq. (35). Given these constraints, one of the symmetry-breaking patterns consistent with this minimization is given by



APPENDIX G: NUMERICAL SIMULATION USING MONTE CARLO

In this appendix, we provide details about the Monte Carlo simulations that we use to obtain the numerical results shown in Figs. 7–9 when considering pure wave-function deformations and Fig. 11 when dealing with the decohered mixed states characterized by the purity. To simplify the numerical implementation, we coordinate each honeycomb layer with a brick wall structure shown below:



We take even linear system sizes such that we can fix periodic boundary conditions and take the number of sites in the horizontal direction \mathcal{N}_x doubled of that in the vertical one \mathcal{N}_y , i.e., $\mathcal{N}_x = 2\mathcal{N}_y$. For the characterization of the phase diagrams in Figs. 7, 8, and 11, we take $\mathcal{N}_y = 40$, while $\mathcal{N}_x = 260$ in the simulations in Fig. 9.

To address the deformed pure wave-function scenario, we consider a bilayer system composed of two stacked honeycomb layers as shown in Fig. 4(a). No tilde $\sigma_j = \pm 1$ spins lie on the upper layer, while tilde spins $\tilde{\sigma}_j$ lie on the lower one. Moreover, we refer to the two honeycomb

sublattices on each layer as green G and B . To numerically tackle this problem, we consider two equivalent formulations. On the one hand, we utilize the Hamiltonian as presented in Eq. (69) in the main text, to evaluate the order parameter \mathcal{W}_{m_R} . Second, after performing the unitary transformation $\tilde{\sigma}_j \rightarrow \sigma_j \tilde{\sigma}_j, \sigma_j \rightarrow \sigma_j$ explained above Eq. (36), an equivalent presentation of the Hamiltonian reads (including the temperatures)

$$H_{|\psi\rangle} = -\frac{\beta_R^x}{2} \sum_{\langle i,j \rangle_\circ} (\sigma_i \sigma_j + \sigma_i \tilde{\sigma}_j + \tilde{\sigma}_i \sigma_j - \tilde{\sigma}_i \tilde{\sigma}_j) - \beta_G^z \sum_{\langle\langle g,g' \rangle\rangle} \sigma_g \tilde{\sigma}_g \sigma_{g'} \tilde{\sigma}_{g'} - \beta_B^z \sum_{\langle\langle b,b' \rangle\rangle} \sigma_b \tilde{\sigma}_b \sigma_{b'} \tilde{\sigma}_{b'}, \quad (\text{G1})$$

that we observe leads to less fluctuation numerical results for the order parameters \mathcal{W}_{e_G} and \mathcal{W}_{e_B} .

The algorithm then runs as follows:

- (1) Fix parameters β_R^x , β_G^z , and β_B^z .
- (2) Initialize random configuration of ± 1 per site on the bilayer system.
- (3) Perform eqSteps = 7×10^4 number of Metropolis steps—each of them involving $2 \times \mathcal{N}_x \times \mathcal{N}_y$ single-site updates—with the acceptance ratio given by the energy difference with respect to the energy function Eq. (G1). Notice that we assume that single-site updates lead to nonreducible dynamics with the unique stationary distribution given by $e^{-H_{|\psi\rangle}}$.
- (4) Compute average quantities by performing additional mcSteps = 7×10^4 Metropolis steps.

When dealing with the decohered density matrix, as characterized by the purity, we instead use the representation of the stat-mech model given by Eq. (83). Here, we use eqSteps = 6×10^4 and mcSteps = 7×10^4 .

-
- [1] M. B. Hastings, *Topological order at nonzero temperature*, *Phys. Rev. Lett.* **107**, 210501 (2011).
 - [2] A. Coser and D. Pérez-García, *Classification of phases for mixed states via fast dissipative evolution*, *Quantum* **3**, 174 (2019).
 - [3] T.-C. Lu, T. H. Hsieh, and T. Grover, *Detecting topological order at finite temperature using entanglement negativity*, *Phys. Rev. Lett.* **125**, 116801 (2020).
 - [4] C. de Groot, A. Turzillo, and N. Schuch, *Symmetry protected topological order in open quantum systems*, *Quantum* **6**, 856 (2022).
 - [5] Y. Bao, R. Fan, A. Vishwanath, and E. Altman, *Mixed-state topological order and the errorfield double formulation of decoherence-induced transitions*, *arXiv:2301.05687*.
 - [6] R. Fan, Y. Bao, E. Altman, and A. Vishwanath, *Diagnostics of mixed-state topological order and breakdown of quantum memory*, *PRX Quantum* **5**, 020343 (2024).

- [7] J. Y. Lee, C.-M. Jian, and C. Xu, *Quantum criticality under decoherence or weak measurement*, *PRX Quantum* **4**, 030317 (2023).
- [8] S. Sang, Y. Zou, and T. H. Hsieh, *Mixed-state quantum phases: Renormalization and quantum error correction*, *arXiv:2310.08639*.
- [9] Y.-H. Chen and T. Grover, *Separability transitions in topological states induced by local decoherence*, *Phys. Rev. Lett.* **132**, 170602 (2024).
- [10] Z. Wang, Z. Wu, and Z. Wang, *Intrinsic mixed-state topological order without quantum memory*, *PRX Quantum* **6**, 010314 (2025).
- [11] R. Ma and C. Wang, *Average symmetry-protected topological phases*, *Phys. Rev. X* **13**, 031016 (2023).
- [12] T.-C. Lu, Z. Zhang, S. Vijay, and T. H. Hsieh, *Mixed-state long-range order and criticality from measurement and feedback*, *PRX Quantum* **4**, 030318 (2023).
- [13] J. Y. Lee, Y.-Z. You, and C. Xu, *Symmetry protected topological phases under decoherence*, *Quantum* **9**, 1607 (2025).
- [14] G.-Y. Zhu, N. Tantivasadakarn, A. Vishwanath, S. Trebst, and R. Verresen, *Nishimori's cat: Stable long-range entanglement from finite-depth unitaries and weak measurements*, *Phys. Rev. Lett.* **131**, 200201 (2023).
- [15] A. Lyons, *Understanding stabilizer codes under local decoherence through a general statistical mechanics mapping*, *arXiv:2403.03955*.
- [16] Z. Li and R. S. K. Mong, *Replica topological order in quantum mixed states and quantum error correction*, *arXiv:2402.09516*.
- [17] R. Sohal and A. Prem, *A noisy approach to intrinsically mixed-state topological order*, *PRX Quantum*, **6**, 010313 (2025).
- [18] T. Ellison and M. Cheng, *Towards a classification of mixed-state topological orders in two dimensions*, *PRX Quantum* **6**, 010315 (2025).
- [19] K. Su, Z. Yang, and C.-M. Jian, *Tapestry of dualities in decohered quantum error correction codes*, *Phys. Rev. B* **110**, 085158 (2024).
- [20] Y.-H. Chen and T. Grover, *Unconventional topological mixed-state transition and critical phase induced by self-dual coherent errors*, *arXiv:2403.06553*.
- [21] J. Hauser, Y. Bao, S. Sang, A. Lavasani, U. Agrawal, and M. P. A. Fisher, *Information dynamics in decohered quantum memory with repeated syndrome measurements: A dual approach*, *arXiv:2407.07882*.
- [22] P. Sala, S. Gopalakrishnan, M. Oshikawa, and Y. You, *Spontaneous strong symmetry breaking in open systems: Purification perspective*, *arXiv:2405.02402*.
- [23] S. Sang and T. H. Hsieh, *Stability of mixed-state quantum phases via finite Markov length*, *arXiv:2404.07251*.
- [24] T.-C. Lu, *Disentangling transitions in topological order induced by boundary decoherence*, *arXiv:2404.06514*.
- [25] J. Y. Lee, *Exact calculations of coherent information for toric codes under decoherence: Identifying the fundamental error threshold*, *arXiv:2402.16937*.
- [26] R. Ma, J.-H. Zhang, Z. Bi, M. Cheng, and C. Wang, *Topological phases with average symmetries: The decohered, the disordered, and the intrinsic* *Phys. Rev. X* **15**, 021062 (2025).

- [27] T. Rakovszky, S. Gopalakrishnan, and C. von Keyserlingk, *Defining stable phases of open quantum systems*, [arXiv:2308.15495](#).
- [28] R. Ma and A. Turzillo, *Symmetry protected topological phases of mixed states in the doubled space*, [arXiv:2403.13280](#).
- [29] P.-S. Hsin, Z.-X. Luo, and H.-Y. Sun, *Anomalies of average symmetries: Entanglement and open quantum systems*, [arXiv:2312.09074](#).
- [30] K. Su, N. Myerson-Jain, and C. Xu, *Conformal field theories generated by Chern insulators under decoherence or measurement*, *Phys. Rev. B* **109**, 035146 (2024).
- [31] Z. Wang and L. Li, *Anomaly in open quantum systems and its implications on mixed-state quantum phases*, [arXiv:2403.14533](#).
- [32] K. Kawabata, R. Sohal, and S. Ryu, *Lieb-Schultz-Mattis theorem in open quantum systems*, *Phys. Rev. Lett.* **132**, 070402 (2024).
- [33] Y. Guo and Y. Ashida, *Two-dimensional symmetry-protected topological phases and transitions in open quantum systems*, *Phys. Rev. B* **109**, 195420 (2024).
- [34] Z. Zhang, U. Agrawal, and S. Vijay, *Quantum communication and mixed-state order in decohered symmetry-protected topological states*, [arXiv:2405.05965](#).
- [35] A. Lavasani and S. Vijay, *The stability of gapped quantum matter and error-correction with adiabatic noise*, [arXiv:2402.14906](#).
- [36] L. A. Lessa, R. Ma, J.-H. Zhang, Z. Bi, M. Cheng, and C. Wang, *Strong-to-weak spontaneous symmetry breaking in mixed quantum states*, *PRX Quantum* **6**, 010344 (2025).
- [37] L. A. Lessa, M. Cheng, and C. Wang, *Mixed-state quantum anomaly and multipartite entanglement*, *Phys. Rev. X* **15**, 011069 (2025).
- [38] Z. Li, D. Lee, and B. Yoshida, *How much entanglement is needed for emergent anyons and fermions?*, [arXiv:2405.07970](#).
- [39] Y.-J. Liu and S. Lieu, *Dissipative phase transitions and passive error correction*, *Phys. Rev. A* **109**, 022422 (2024).
- [40] S. Lieu, Y.-J. Liu, and A. V. Gorshkov, *Candidate for a passively protected quantum memory in two dimensions*, *Phys. Rev. Lett.* **133**, 030601 (2024).
- [41] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, England, 2010).
- [42] X.-G. Wen, *Quantum Field Theory of Many-Body Systems: From the Origin of Sound to an Origin of Light and Electrons* (Oxford University Press, New York, 2007).
- [43] S. Sachdev, *Quantum Phases of Matter* (Cambridge University Press, Cambridge, England, 2023).
- [44] J. M. Leinaas and J. Myrheim, *On the theory of identical particles*, *Il Nuovo Cimento B* (1971-1996) **37**, 1 (1977).
- [45] G. A. Goldin, R. Menikoff, and D. H. Sharp, *Representations of a local current algebra in nonsimply connected space and the Aharonov-Bohm effect*, *J. Math. Phys. (N.Y.)* **22**, 1664 (1981).
- [46] F. Wilczek, *Quantum mechanics of fractional-spin particles*, *Phys. Rev. Lett.* **49**, 957 (1982).
- [47] T. Einarsson, *Fractional statistics on a torus*, *Phys. Rev. Lett.* **64**, 1995 (1990).
- [48] A. Kitaev, *Fault-tolerant quantum computation by anyons*, *Ann. Phys. (Amsterdam)* **303**, 2 (2003).
- [49] M. H. Freedman, M. Larsen, and Z. Wang, *A modular functor which is universal for quantum computation*, *Commun. Math. Phys.* **227**, 605 (2002).
- [50] M. Freedman, C. Nayak, and K. Walker, *Towards universal topological quantum computation in the $\nu = 5/2$ fractional quantum Hall state*, *Phys. Rev. B* **73**, 245307 (2006).
- [51] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, *Non-Abelian anyons and topological quantum computation*, *Rev. Mod. Phys.* **80**, 1083 (2008).
- [52] B. M. Terhal, *Quantum error correction for quantum memories*, *Rev. Mod. Phys.* **87**, 307 (2015).
- [53] N. Read and S. Sachdev, *Large- n expansion for frustrated quantum antiferromagnets*, *Phys. Rev. Lett.* **66**, 1773 (1991).
- [54] X. G. Wen, *Mean-field theory of spin-liquid states with finite energy gap and topological orders*, *Phys. Rev. B* **44**, 2664 (1991).
- [55] C. Wang, J. Harrington, and J. Preskill, *Confinement-Higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory*, *Ann. Phys. (Amsterdam)* **303**, 31 (2003).
- [56] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, *Topological quantum memory*, *J. Math. Phys. (N.Y.)* **43**, 4452 (2002).
- [57] Z. Nussinov and G. Ortiz, *Autocorrelations and thermal fragility of anyonic loops in topologically quantum ordered systems*, *Phys. Rev. B* **77**, 064302 (2008).
- [58] S. Bravyi and B. Terhal, *A no-go theorem for a two-dimensional self-correcting quantum memory based on stabilizer codes*, *New J. Phys.* **11**, 043029 (2009).
- [59] O. Landon-Cardinal and D. Poulin, *Local topological order inhibits thermal stability in 2d*, *Phys. Rev. Lett.* **110**, 090502 (2013).
- [60] B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton, *Quantum memories at finite temperature*, *Rev. Mod. Phys.* **88**, 045005 (2016).
- [61] The mapping to the two-dimensional RBIM holds only under the assumption of perfect syndrome measurements. If one incorporates imperfect measurements as part of the physical process, then the underlying model becomes the three-dimensional random-plaquette model [55,56].
- [62] G. A. Goldin, R. Menikoff, and D. H. Sharp, *Comments on "general theory for quantum statistics in two dimensions"*, *Phys. Rev. Lett.* **54**, 603 (1985).
- [63] X. G. Wen, *Non-Abelian statistics in the fractional quantum Hall states*, *Phys. Rev. Lett.* **66**, 802 (1991).
- [64] G. Moore and N. Read, *Non-Abelions in the fractional quantum Hall effect*, *Nucl. Phys. B* **360**, 362 (1991).
- [65] G. Moore and N. Seiberg, *Classical and quantum conformal field theory*, *Commun. Math. Phys.* **123**, 177 (1989).
- [66] J. R. Wootton, J. Burri, S. Iblisdir, and D. Loss, *Error correction for non-Abelian topological quantum computation*, *Phys. Rev. X* **4**, 011051 (2014).
- [67] C. G. Brell, S. Burton, G. Dauphinais, S. T. Flammia, and D. Poulin, *Thermalization, error correction, and memory*

- lifetime for Ising anyon systems, *Phys. Rev. X* **4**, 031058 (2014).
- [68] J. R. Wootton and A. Hutter, *Active error correction for Abelian and non-Abelian anyons*, *Phys. Rev. A* **93**, 022318 (2016).
- [69] S. Burton, C. G. Brell, and S. T. Flammia, *Classical simulation of quantum error correction in a Fibonacci anyon code*, *Phys. Rev. A* **95**, 022309 (2017).
- [70] G. Dauphinais and D. Poulin, *Fault-tolerant quantum error correction for non-Abelian anyons*, *Commun. Math. Phys.* **355**, 519 (2017).
- [71] A. Schotte, L. Burgelman, and G. Zhu, *Fault-tolerant error correction for a universal non-Abelian topological quantum computer at finite temperature*, arXiv:2301.00054.
- [72] A. Schotte, G. Zhu, L. Burgelman, and F. Verstraete, *Quantum error correction thresholds for the universal Fibonacci Turaev-Viro code*, *Phys. Rev. X* **12**, 021012 (2022).
- [73] M. Iqbal, N. Tantivasadakarn, R. Verresen, S. L. Campbell, J. M. Dreiling, C. Figgatt, J. P. Gaebler, J. Johansen, M. Mills, S. A. Moses, J. M. Pino, A. Ransford, M. Rowe, P. Siegfried, R. P. Stutz, M. Foss-Feig, A. Vishwanath, and H. Dreyer, *Non-Abelian topological order and anyons on a trapped-ion processor*, *Nature (London)* **626**, 505 (2024).
- [74] More than for $D(S_3)$ TO [48], which counts six anyons, or double Fibonacci (dFib) [75], which has only four
- [75] M. A. Levin and X.-G. Wen, *String-net condensation: A physical mechanism for topological phases*, *Phys. Rev. B* **71**, 045110 (2005).
- [76] The fact that this TO can be obtained by a single round of measurements is believed to be impossible [77] for smaller TOs such as $D(S_3)$ and dFib TO.
- [77] N. Tantivasadakarn, A. Vishwanath, and R. Verresen, *Hierarchy of topological order from finite-depth unitaries, measurement, and feedforward*, *PRX Quantum* **4**, 020339 (2023).
- [78] B. Yoshida, *Topological phases with generalized global symmetries*, *Phys. Rev. B* **93**, 155131 (2016).
- [79] R. Dijkgraaf and E. Witten, *Topological gauge theories and group cohomology*, *Commun. Math. Phys.* **129**, 393 (1990).
- [80] R. Dijkgraaf, V. Pasquier, and P. Roche, *Quasi hope algebras, group cohomology and orbifold models*, *Nucl. Phys. B, Proc. Suppl.* **18**, 60 (1991).
- [81] Y. Hu, Y. Wan, and Y.-S. Wu, *Twisted quantum double model of topological phases in two dimensions*, *Phys. Rev. B* **87**, 125114 (2013).
- [82] J. Haegeman, K. Van Acoleyen, N. Schuch, J. I. Cirac, and F. Verstraete, *Gauging quantum states: From global to local symmetries in many-body systems*, *Phys. Rev. X* **5**, 011024 (2015).
- [83] G.-Y. Zhu and G.-M. Zhang, *Gapless Coulomb state emerging from a self-dual topological tensor-network state*, *Phys. Rev. Lett.* **122**, 176401 (2019).
- [84] M. Mariën, J. Haegeman, P. Fendley, and F. Verstraete, *Condensation-driven phase transitions in perturbed string nets*, *Phys. Rev. B* **96**, 155127 (2017).
- [85] W.-T. Xu and N. Schuch, *Characterization of topological phase transitions from a non-Abelian topological state and its Galois conjugate through condensation and confinement order parameters*, *Phys. Rev. B* **104**, 155119 (2021).
- [86] W.-T. Xu, J. Garre-Rubio, and N. Schuch, *Complete characterization of non-Abelian topological phase transitions and detection of anyon splitting with projected entangled pair states*, *Phys. Rev. B* **106**, 205139 (2022).
- [87] A. Schotte, J. Carrasco, B. Vanhecke, L. Vanderstraeten, J. Haegeman, F. Verstraete, and J. Vidal, *Tensor-network approach to phase transitions in string-net models*, *Phys. Rev. B* **100**, 245125 (2019).
- [88] P. Fendley, *Topological order from quantum loops and nets*, *Ann. Phys. (Amsterdam)* **323**, 3113 (2008).
- [89] E. Domany, D. Mukamel, B. Nienhuis, and A. Schwimmer, *Duality relations and equivalences for models with $O(n)$ and cubic symmetry*, *Nucl. Phys. B* **190**, 279 (1981).
- [90] B. Nienhuis, *Exact critical point and critical exponents of $O(n)$ models in two dimensions*, *Phys. Rev. Lett.* **49**, 1062 (1982).
- [91] R. Peled and Y. Spinka, *Lectures on the spin and loop $O(n)$ models*, arXiv:1708.00058.
- [92] P. Sala and R. Verresen, *Stability and loop models from decohering non-Abelian topological order*, arXiv:2409.12230.
- [93] M. de Wild Propitius, *Topological interactions in broken gauge theories*, arXiv:hep-th/9511195.
- [94] A. Kitaev, *Anyons in an exactly solved model and beyond*, *Ann. Phys. (Amsterdam)* **321**, 2 (2006).
- [95] L. Lootens, B. Vancraeynest-De Cuiper, N. Schuch, and F. Verstraete, *Mapping between Morita-equivalent string-net states with a constant depth quantum circuit*, *Phys. Rev. B* **105**, 085130 (2022).
- [96] More concretely, it is obtained by gauging a $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry protected topological phase on a triangular lattice [78].
- [97] M. Barkeshli, P. Bonderson, M. Cheng, and Z. Wang, *Symmetry fractionalization, defects, and gauging of topological phases*, *Phys. Rev. B* **100**, 115147 (2019).
- [98] X. Chen, *Symmetry fractionalization in two dimensional topological phases*, *Rev. Phys.* **2**, 3 (2017).
- [99] D. Ben-Zion, D. Das, and J. McGreevy, *Exactly solvable models of spin liquids with spinons, and of three-dimensional topological paramagnets*, *Phys. Rev. B* **93**, 155147 (2016).
- [100] D. T. Stephen, J. Garre-Rubio, A. Dua, and D. J. Williamson, *Subsystem symmetry enriched topological order in three dimensions*, *Phys. Rev. Res.* **2**, 033331 (2020).
- [101] N. Tantivasadakarn, R. Verresen, and A. Vishwanath, *Shortest route to non-Abelian topological order on a quantum processor*, *Phys. Rev. Lett.* **131**, 060405 (2023).
- [102] Calling certain anyons “charges” and others “fluxes” depends on how one describes it as a gauge theory, which is a matter of choice. Here, we follow the convention of this state as a twisted \mathbb{Z}_2^3 gauge theory as in Refs. [73,78–81].
- [103] While the arXiv version of this manuscript includes colored text and formulas that can facilitate the reading, we here include whenever necessary either a subscript or superscript to clarify the color of the corresponding sublattice: red (R or r), green (G or g), and blue (B or b).
- [104] The orientation of the string is set by the fact that the initial triangle t_i lives in either a blue or green plaquette.

- [105] For completeness, $A_v^B \rightarrow X_v^B \prod_{\langle v,v' \rangle} (CB)_{v',t}$, where we are controlling the action of the three-body B_t^R term on the state of the three neighboring green vertices on the honeycomb lattice—clearly, if $B_t = 1$, one can forget about this controlled term.
- [106] If we had instead put it in the state, then Eq. (5) would have an additional finite-depth “CCZ” circuit along the bonds of the honeycomb lattice, and, e.g., Eq. (3) would simplify to just mapping $X \rightarrow X$.
- [107] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, *Criticality, the area law, and the computational power of projected entangled pair states*, *Phys. Rev. Lett.* **96**, 220601 (2006).
- [108] For loop models with intersections, one has to take care with how to define components; for $N = 1$, this issue does not arise.
- [109] This implies the trivialization of the linear-depth circuit necessary for the Wilson operators $\mathcal{X}_{t_i}^{t_f}$ and $\mathcal{X}_{t_i}^{t_f}$ to produce a pair of non-Abelian fluxes in the undeformed D_4 state, since they now become Abelian anyons.
- [110] In the present work, we do not focus on the details of these beyond-threshold phases, but see Ref. [111] for subtleties that arise.
- [111] R. Sahay, C. von Keyserlingk, R. Verresen, and C. Zhang, *Enforced gaplessness from states with exponentially decaying correlations*, [arXiv:2503.01977](https://arxiv.org/abs/2503.01977).
- [112] R. Houtappel, *Order-disorder in hexagonal lattices*, *Physica* **16**, 425 (1950).
- [113] We note this physical mechanism has also recently been explored from a Hamiltonian perspective in Ref. [114], where it was found to lead to an infinitely robust spin liquid.
- [114] R. Verresen and A. Vishwanath, *Unifying Kitaev magnets, kagomé dimer models, and ruby Rydberg spin liquids*, *Phys. Rev. X* **12**, 041029 (2022).
- [115] G. H. Wannier, *Antiferromagnetism. the triangular Ising net*, *Phys. Rev.* **79**, 357 (1950).
- [116] H. W. J. Blöte and M. P. Nightingale, *Antiferromagnetic triangular Ising model: Critical behavior of the ground state*, *Phys. Rev. B* **47**, 15046 (1993).
- [117] B. Nienhuis, H. J. Hilhorst, and H. W. J. Blöte, *Triangular SOS models and cubic-crystal shapes*, *J. Phys. A* **17**, 3559 (1984).
- [118] R. Moessner, S. L. Sondhi, and P. Chandra, *Two-dimensional periodic frustrated Ising models in a transverse field*, *Phys. Rev. Lett.* **84**, 4457 (2000).
- [119] D. S. Rokhsar and S. A. Kivelson, *Superconductivity and the quantum hard-core dimer gas*, *Phys. Rev. Lett.* **61**, 2376 (1988).
- [120] N. Read and S. Sachdev, *Large- n expansion for frustrated quantum antiferromagnets*, *Phys. Rev. Lett.* **66**, 1773 (1991).
- [121] M. E. Fisher and J. Stephenson, *Statistical mechanics of dimers on a plane lattice. II. Dimer correlations and monomers*, *Phys. Rev.* **132**, 1411 (1963).
- [122] Y.-H. Chen and T. Grover, *Symmetry-enforced many-body separability transitions*, *PRX Quantum* **5**, 030310 (2024).
- [123] Y. Ashida, S. Furukawa, and M. Oshikawa, *System-environment entanglement phase transitions*, *Phys. Rev. B* **110**, 094404 (2024).
- [124] Specht’s theorem says that two Hermitian matrices A and B are unitarily equivalent if and only if $\text{tr}(A^n) = \text{tr}(B^n)$ for all $n \in \mathbb{N}$.
- [125] M. Schwartz and S. Fishman, *Real space renormalization group study of the random bond Ising model*, *Physica (Amsterdam)* **104A**, 115 (1980).
- [126] A. B. Harris, *Effect of random defects on the critical behaviour of Ising models*, *J. Phys. C* **7**, 1671 (1974).
- [127] H. Nishimori, *Internal energy, specific heat and correlation function of the bond-random Ising model*, *Prog. Theor. Phys.* **66**, 1169 (1981).
- [128] Y. Ozeki and H. Nishimori, *Phase diagram of gauge glasses*, *J. Phys. A* **26**, 3399 (1993).
- [129] P. Le Doussal and A. B. Harris, *Location of the Ising spin-glass multicritical point on Nishimori’s line*, *Phys. Rev. Lett.* **61**, 625 (1988).
- [130] A. Honecker, M. Picco, and P. Pujol, *Universality class of the Nishimori point in the 2D $\pm J$ random bond Ising model*, *Phys. Rev. Lett.* **87**, 047201 (2001).
- [131] I. A. Gruzberg, N. Read, and A. W. W. Ludwig, *Random-bond Ising model in two dimensions: The Nishimori line and supersymmetry*, *Phys. Rev. B* **63**, 104422 (2001).
- [132] S. L. A. de Queiroz, *Multicritical point of Ising spin glasses on triangular and honeycomb lattices*, *Phys. Rev. B* **73**, 064410 (2006).
- [133] Only contractible loops contribute, as otherwise the eigenvalue of the logical \mathcal{Z} operator is toggled when considering the initial state $|D_4\rangle$ —this condition is relaxed when one instead considers, e.g., the superposition $|D_4\rangle + \mathcal{X}|D_4\rangle$.
- [134] The reason we find a loop model instead of a trivalent graph is that m_R does not appear in the fusion $m_R \times m_R$.
- [135] J.-H. Zhang, Y. Qi, and Z. Bi, *Strange correlation function for average symmetry-protected topological phases*, [arXiv:2210.17485](https://arxiv.org/abs/2210.17485).
- [136] B. Shi, K. Kato, and I. H. Kim, *Fusion rules from entanglement*, *Ann. Phys. (Amsterdam)* **418**, 168164 (2020).
- [137] J. Preskill, *Chapter 9. Topological quantum computation* (2004), <https://www.preskill.caltech.edu/ph219/topological.pdf>.
- [138] H. Duminil-Copin, A. Glazman, R. Peled, and Y. Spinka, *Macroscopic loops in the loop $O(n)$ model at Nienhuis’ critical point*, [arXiv:1707.09335](https://arxiv.org/abs/1707.09335).
- [139] Notice that this high-temperature expansion includes both contractible and noncontractible loop configurations.
- [140] Intriguingly, the first equality can be interpreted as a mixed model where blue sites have “face-cubic” spins and green sites have “corner-cubic” spins—it is known that such spins lead to local spin models for $O(N)$ honeycomb loop models [89,141], but usually this covers only the range of tension $t \leq (1/n)$; our mixed model gives a manifestly positive stat-mech model for $t \leq (1/\sqrt{n})$.
- [141] L. Chayes, L. P. Pryadko, and K. Shtengel, *Intersecting loop models on \mathbb{Z}^d : Rigorous results*, *Nucl. Phys.* **B570**, 590 (2000).
- [142] Interestingly, there is a phase transition involving crystalline symmetry breaking for very large string tension [141,143], which is far outside the physical regime under consideration here.

- [143] W. Guo, H. W. J. Blöte, and F. Y. Wu, *Phase transition in the $n > 2$ honeycomb $O(n)$ model*, *Phys. Rev. Lett.* **85**, 3874 (2000).
- [144] Notice that the quantum channel resulting from applying the channel \mathcal{E}^X with the error strength p_R m times is a quantum channel with the same form but error strength $p_R^{(m)} = \frac{1}{2} - (\frac{1}{2} - p_R)(1 - 2p_R)^m \in [0, \frac{1}{2}]$.
- [145] More precisely, it holds for any “flux-free” configuration where $\prod_{(i,j) \in \square} \eta_{ij} = 1$ for each hexagon, which means we can write $\eta_{ij} = \tau_i \tau_j$; by a change of variables $\sigma_i \rightarrow \sigma_i \tau_i$, this leads to the same partition function as $\eta_{ij} = 1$.
- [146] If $p_c^{(\infty)} < \frac{1}{2}$, it implies we have entered a phase with algebraic correlations.
- [147] Notice the difference with the analysis of the toric code [6,56], where the paramagnetic phase of the RBIM corresponds to breakdown of the quantum memory.
- [148] In particular, the initial $|D_4\rangle$ and the state obtained by applying the logical \mathcal{X} string operator [Eq. (2)] around the torus.
- [149] We define $\Delta F_C = -\ln[Z_-(\eta)/Z(\eta)]$, where $Z(\eta) \propto P(\eta)$ and $Z_-(\eta)$ is the partition function with a line C of flipped antiferromagnetic bonds.
- [150] A. Kubica, M. E. Beverland, F. Brandão, J. Preskill, and K. M. Svore, *Three-dimensional color code thresholds via statistical-mechanical mapping*, *Phys. Rev. Lett.* **120**, 180501 (2018).
- [151] H. Shimada, J. L. Jacobsen, and Y. Kamiya, *Phase diagram and strong-coupling fixed point in the disordered $O(n)$ loop model*, *J. Phys. A* **47**, 122001 (2014).
- [152] V. Alba, A. Pelissetto, and E. Vicari, *Quasi-long-range order in the 2d xy model with random phase shifts*, *J. Phys. A* **42**, 295001 (2009).
- [153] Notice that X and Z pure wave-function deformations acting on the same lattice do not commute. One can show that the resulting stat-mech model for a given ordering of the deformations corresponds to coupled self-avoiding $O(2)$ and $O(1)$ loop models.
- [154] E. Witten, *Constraints on supersymmetry breaking*, *Nucl. Phys.* **B202**, 253 (1982).
- [155] J. Wouters, H. Katsura, and D. Schuricht, *Interrelations among frustration-free models via Witten’s conjugation*, *SciPost Phys. Core* **4**, 027 (2021).
- [156] Here, the factor of 2 relates to the degeneracy of domain wall configurations leading to the same loop configuration.
- [157] F. Burnell, *Anyon condensation and its applications*, *Annu. Rev. Condens. Matter Phys.* **9**, 307 (2018).
- [158] While this is a nontrivial assumption, notice that it correctly reproduces the known BKT critical point of the $O(2)$ loop model.
- [159] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, England, 1995).
- [160] C. Dasgupta and B. I. Halperin, *Phase transition in a lattice model of superconductivity*, *Phys. Rev. Lett.* **47**, 1556 (1981).
- [161] Our convention for the normalization of the fields connects to that in Ref. [162] by sending $\varphi \rightarrow 2\varphi$. In this convention, the canonical commutator becomes $[\varphi(x), \partial_y \theta(y)] = i\pi\delta(x-y)$, and the local Hamiltonian density $\mathcal{H} = [1/(2\pi)](K(\partial_x \theta)^2 + (1/K)(\partial_x \varphi)^2)$.
- [162] M. P. A. Fisher and L. I. Glazman, *Transport in a one-dimensional Luttinger liquid*, *arXiv:cond-mat/9610037*.
- [163] B. Nienhuis, *Critical behavior of two-dimensional spin models and charge asymmetry in the Coulomb gas*, *J. Stat. Phys.* **34**, 731 (1984).
- [164] C. Domb and J. Lebowitz, *Phase Transitions and Critical Phenomena*, *Phase Transitions and Critical Phenomena* Vol. 18 (Elsevier Science, New York, 2000).
- [165] We thank Sagar Vijay for an insightful discussion on this point.
- [166] I. Affleck, *Nonlinear σ model at $\theta = \pi$: Euclidean lattice formulation and solid-on-solid models*, *Phys. Rev. Lett.* **66**, 2429 (1991).
- [167] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller, *Preparation of entangled states by quantum Markov processes*, *Phys. Rev. A* **78**, 042307 (2008).
- [168] F. Verstraete, M. M. Wolf, and J. Ignacio Cirac, *Quantum computation and quantum-state engineering driven by dissipation*, *Nat. Phys.* **5**, 633 (2009).
- [169] B. Buča, J. Tindall, and D. Jaksch, *Non-stationary coherent quantum many-body dynamics through dissipation*, *Nat. Commun.* **10**, 1730 (2019).
- [170] J. Dubois, U. Saalmann, and J. M. Rost, *Symmetry-induced decoherence-free subspaces*, *Phys. Rev. Res.* **5**, L012003 (2023).
- [171] X. Mi *et al.*, *Stable quantum-correlated many body states via engineered dissipation*, *Science* **383**, 1332 (2024).
- [172] Y. Li, P. Sala, and F. Pollmann, *Hilbert space fragmentation in open quantum systems*, *Phys. Rev. Res.* **5**, 043239 (2023).
- [173] Y. Li, F. Pollmann, N. Read, and P. Sala, *Highly-entangled stationary states from strong symmetries*, *Phys. Rev. X* **15**, 011068 (2025).
- [174] F. J. Wegner, *Duality in generalized Ising models and phase transitions without local order parameters*, *J. Math. Phys. (N.Y.)* **12**, 2259 (1971).
- [175] P. Ginsparg, *Applied conformal field theory*, *arXiv:hep-th/9108028*.
- [176] P. Francesco, P. Mathieu, and D. Sénéchal, *Conformal Field Theory*, *Graduate Texts in Contemporary Physics* (Springer, New York, 1997).