

# Entanglement negativity and sudden death in the toric code at finite temperature

O. Hart and C. Castelnovo

*T.C.M. Group, Cavendish Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom*

(Dated: April 2018)

We study the fate of quantum correlations at finite temperature in the two-dimensional toric code using the logarithmic entanglement negativity. We are able to obtain exact results that give us insight into how thermal excitations affect quantum entanglement. The toric code has two types of elementary excitations (defects) costing different energies. We show that an  $O(1)$  density of the lower energy defect is required to degrade the zero-temperature entanglement between two subsystems in contact with one another. However, one type of excitation alone is not sufficient to kill all quantum correlations, and an  $O(1)$  density of the higher energy defect is required to cause the so-called sudden death of the negativity. Interestingly, if the energy cost of one of the excitations is taken to infinity, quantum correlations survive up to arbitrarily high temperatures, a feature that is likely shared with other quantum spin liquids and frustrated systems in general, when projected down to their low-energy states. We demonstrate this behaviour both for small subsystems, where we can prove that the negativity is a necessary and sufficient condition for separability, as well as for extended subsystems, where it is only a necessary condition. We further observe that the negativity per boundary degree of freedom at a given temperature increases (parametrically) with the size of the boundary, and that quantum correlations between subsystems with extended boundaries are more robust to thermal fluctuations.

## I. INTRODUCTION

With the modern surge of interest in harvesting the capabilities of quantum mechanical systems to develop new technologies, it has become ever so important to quantify and characterise quantum correlations in physical systems. While this remains a tall order in real systems, substantial progress has been made in recent years at the theoretical level. Several measures of quantum correlations (including, but not limited to, “entanglement”) have been proposed and studied (see Refs. 1–3 for reviews). One of the most successful is the von Neumann entanglement entropy. Similarly to other measures, the entanglement entropy works remarkably well when a system is prepared in a pure state—namely, when the density matrix is a projector onto a single quantum mechanical state. However, it becomes less descriptive once we deal with mixed states, as is the case for systems at finite temperature. Under these circumstances, the von Neumann entropy becomes dominated by classical correlations [4] and it is difficult to distil quantum from classical correlations.

One of the most promising ways to address this issue has come of late from a newly introduced measure of entanglement, the so called (logarithmic) *entanglement negativity*  $\mathcal{E}$  [5, 6] and variants thereof [7, 8]. The entanglement negativity uses the positive partial transpose criterion [9] to detect the separability of a system’s density matrix, and as such it is a proper measure of purely quantum correlations—although it is in general only a necessary but not sufficient condition for separability [7, 10].

Unfortunately, calculating  $\mathcal{E}$  in many-body quantum systems is notoriously difficult and analytic results are few. At zero temperature, the negativity has been evaluated for a limited number of sufficiently simple lattice models in one [11–13] and two [14, 15] dimensions, suggesting universality at quantum critical points and unveiling leading area law behaviour respectively. That the negativity is fully universal and scale invariant at quantum critical points has subsequently been proven rigorously using conformal field theory (CFT) [16, 17]. When temperature is raised from zero, one expects that thermal

mixing of eigenstates (“thermal fluctuations”) will lead to a reduction in the entanglement. This intuition has been verified for a large number of *two*-spin systems [18], initially using the concurrence [19–30] (for a review, see Ref. 2) and more recently using the negativity [31–33]. Generically, there exists a well-defined temperature  $T_c$ , dubbed the “sudden death temperature,” above which the negativity vanishes identically. Beyond two-spin systems, in particular considering the entanglement between two blocks of spins, far less is known. At quantum critical points, CFT results have been extended into the regime of finite temperature [34, 35], while on the numerical side a linked cluster expansion has been employed, alongside exact diagonalisation, to study the negativity at finite temperature in one- and two-dimensional bipartite spin systems [4]. Even though the negativity does not capture *all* quantum correlations, sudden death is nevertheless an intriguing phenomenon that requires further investigation. Further results in  $d > 1$  systems, and in particular exact expressions for the negativity at finite temperature, would be highly beneficial to gain a better understanding about the fate of quantum correlations and the origin of the sudden death behaviour.

The goal of this manuscript is to present the first exact calculation of the negativity in a 2D lattice system in thermal equilibrium at finite temperature. We choose to work with the toric code model [36] whose exact solubility has allowed the negativity at  $T = 0$  [14, 15] and the von Neumann entropy at finite temperature [37] to be calculated exactly. The dynamical approach to thermal equilibrium, and the resulting relaxation of the negativity, is beyond the scope of this work. A review of quantum memories at finite temperature, which highlights the relevant time scales associated with equilibration, can be found in Ref. 38. We consider the entanglement between two subsystems that share a boundary of finite length, which allows us to discern the fate of the zero-temperature area law, and how the sudden death temperature  $T_c$  depends on boundary length.

We find that the zero-temperature area law holds for all temperatures, unlike the von Neumann entropy which transitions to a volume law for sufficiently large temperatures [37, 39].

We also show that, *irrespective of boundary length*, the negativity is killed by an  $O(1)$  density of the most energetically costly defect. The entanglement between larger partitions is however more robust to the effects of thermal fluctuations, with the sudden death temperature of the largest boundaries being approximately twice that of the smallest boundaries. Sending the energy cost of the most expensive defects to infinity removes sudden death, and instead a gradual  $1/T$  demise of the negativity is observed.

The manuscript is arranged as follows. We briefly introduce the toric code model, the entanglement negativity, and the techniques pertaining to its calculation in Sec. IA and Sec. IB. A summary of our main (exact) analytical results, as well as a discussion of their significance, is presented in Sec. II. Details of the relatively long and technical analytical calculations are for convenience postponed until Sec. III. We then briefly comment on the potential experimental relevance of our results in Sec. IV. Finally, in Sec. V, we draw our conclusions and outlook.

### A. Entanglement negativity

Given a tripartite system  $\mathcal{S} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{B}$ , the logarithmic negativity is defined in terms of the reduced state  $\rho_{\mathcal{A}} = \text{Tr}_{\mathcal{B}} \rho$  as

$$\mathcal{E} = \ln \|\rho_{\mathcal{A}}^{T_2}\|_1, \quad (1)$$

where  $T_2$  denotes partial transposition over the  $\mathcal{A}_2$  subsystem and  $\|\cdot\|_1 = \text{Tr}|\cdot|$  is the trace norm [6]. The logarithmic negativity quantifies the entanglement between the subsystems  $\mathcal{A}_1$  and  $\mathcal{A}_2$ . One can verify that it is symmetric in  $\mathcal{A}_1 \leftrightarrow \mathcal{A}_2$  as a good measure of entanglement must be, so that we can equally compute  $\|\rho_{\mathcal{A}}^{T_1}\|_1$ .

Given an (arbitrary) orthonormal basis  $|\psi_i \phi_j\rangle \equiv |\psi_i\rangle \otimes |\phi_j\rangle \in \mathcal{H}_{\mathcal{A}_1} \otimes \mathcal{H}_{\mathcal{A}_2}$ , the operation of partial transposition can be defined in terms of matrix elements as [6]

$$\langle \psi_i \phi_j | \rho_{\mathcal{A}}^{T_2} | \psi_k \phi_l \rangle = \langle \psi_i \phi_l | \rho_{\mathcal{A}} | \psi_k \phi_j \rangle. \quad (2)$$

In order to evaluate the negativity analytically, we employ the replica approach, introduced by the authors of Ref. 16, which has since been used to calculate the negativity in a variety of models [14, 16, 17, 35, 40, 41]. If the operator  $\rho_{\mathcal{A}}^{T_2}$  has eigenvalues  $\lambda_i$ , then

$$\text{Tr} (\rho_{\mathcal{A}}^{T_2})^{n_e} = \sum_{\lambda_i \geq 0} |\lambda_i|^{n_e} + \sum_{\lambda_i < 0} |\lambda_i|^{n_e}, \quad (3)$$

$$\text{Tr} (\rho_{\mathcal{A}}^{T_2})^{n_o} = \sum_{\lambda_i \geq 0} |\lambda_i|^{n_o} - \sum_{\lambda_i < 0} |\lambda_i|^{n_o}, \quad (4)$$

where  $n_e$  ( $n_o$ ) is an even (odd) positive integer. Crucially, in contrast to  $\rho_{\mathcal{A}}$ , the operator  $\rho_{\mathcal{A}}^{T_2}$  can have negative eigenvalues. To obtain the sum of the absolute values of the eigenvalues, i.e., the trace norm, the replica method consists of following the *even* series and taking the analytic continuation  $n_e \rightarrow 1$ . In this way, we are able to evaluate the negativity exactly via the expression

$$\mathcal{E} = \lim_{n_e \rightarrow 1} \ln \text{Tr} (\rho_{\mathcal{A}}^{T_2})^{n_e}. \quad (5)$$

### B. Toric code model

The toric code is defined on a 2D square lattice composed of  $N$  sites with periodic boundary conditions (i.e., on a torus). Spin- $\frac{1}{2}$  degrees of freedom are located on each of the  $2N$  bonds of the lattice and interact via the Hamiltonian

$$H = -\lambda_A \sum_s A_s - \lambda_B \sum_p B_p \equiv -\lambda_A S - \lambda_B P, \quad (6)$$

where the labels  $s$  and  $p$  denote the ‘‘stars’’ and ‘‘plaquettes’’ of the lattice, respectively, and the operators  $A_s \equiv \prod_{i \in s} \sigma_i^x$ ,  $B_p \equiv \prod_{i \in p} \sigma_i^z$  [36].  $\lambda_A, \lambda_B > 0$  are the two coupling constants of the model. As usual  $\sigma_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$  denote the Pauli matrices, which describe the spin on bond  $i$ .

All operators  $A_s, B_p$  in the Hamiltonian mutually commute and so can be diagonalised simultaneously. The property  $A_s^2 = \mathbb{1} = B_p^2 \forall s, p$  implies that the operators  $A_s$  and  $B_p$  have eigenvalues  $\pm 1$ , and the states for which all  $A_s, B_p$  have eigenvalue  $+1$  are the ground states. The model has 4 ground states belonging to different ‘‘topological sectors,’’ which are classified according to the eigenvalues of nonlocal (system-spanning) operators. Such operators can be defined for instance as  $\Gamma_{h,v}^x = \prod_{i \in \mathcal{L}_{h,v}} \sigma_i^x$  [42], where  $\mathcal{L}_{h,v}$  are two noncontractible loops on the dual lattice that span the torus in the horizontal ( $h$ ) and vertical ( $v$ ) directions, respectively (see Fig. 1).

There exist two types of elementary excitations: *electric charges* and *magnetic vortices* generated by open strings of Pauli matrices,  $S^\alpha(\ell) = \prod_{i \in \ell} \sigma_i^\alpha$  with  $\alpha = z, x$  respectively [43] (see Fig. 1).  $S^\alpha(\ell)$  flips the eigenvalues of the *two* stars (plaquettes) at the ends of the path  $\ell$ , which lives on the direct (dual) lattice, costing an energy  $4\lambda_A$  ( $4\lambda_B$ ). With periodic boundary conditions, it is not possible to change the eigenvalue of just one star or plaquette operator as a result of the condition  $\prod_s A_s = \prod_p B_p = 1$ . At finite temperature, excitations are present in the system with a density given approximately by the Fermi-Dirac distribution  $n_F(2\lambda_X) \equiv (e^{2\beta\lambda_X} + 1)^{-1}$ , for  $X = A, B$ ,

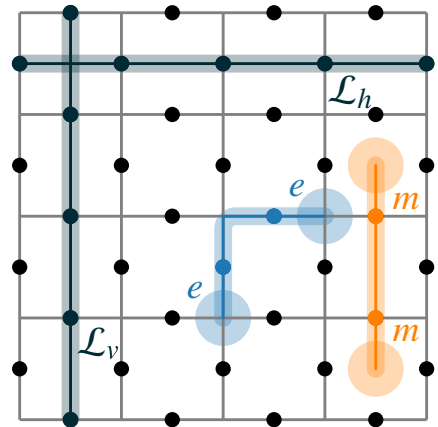


FIG. 1. Illustration of two possible noncontractible loops  $\mathcal{L}_{h,v}$ , and of two open strings generating electric ( $e$ ) and magnetic ( $m$ ) defects at their ends. Recall that periodic boundary conditions are imposed. Each electric (magnetic) defect costs an energy  $2\lambda_A$  ( $2\lambda_B$ ) so that each open string costs  $4\lambda_A$  ( $4\lambda_B$ ).

in the limit of large system size  $N$  (see Appendix A for further details).

## II. MAIN RESULTS

We consider a tripartition of the system  $\mathcal{S}$  into three disjoint subsystems  $\mathcal{S} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{B}$ . Our goal is to quantify the entanglement between subsystems  $\mathcal{A}_1$  and  $\mathcal{A}_2$  using the logarithmic negativity  $\mathcal{E}$ , when the system is in a thermal state  $\rho = e^{-\beta H}/Z$ . To obtain explicit expressions, it is necessary to make some assumptions about the arrangement of the various subsystems. The partition schemes with which we choose to work are shown schematically in Fig. 2. In all cases, subsystems  $\mathcal{A}_1$  and  $\mathcal{A}_2$  have one common edge since it is known for the toric code that the zero-temperature negativity vanishes identically if the subsystems are spatially separated [14, 15]. The length of this boundary is parameterised by  $N_\partial$ , which we define as the number of stars that straddle the  $\mathcal{A}_1$ - $\mathcal{A}_2$  boundary (hereinafter ‘‘boundary stars’’). In Fig. 2b, where subsystem  $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$  spans the system in one direction, we expect to obtain a constant, additive contribution to the zero-temperature negativity, indicative of topological order [14].

### A. Star plaquette pair

It is useful to study first the smallest subsystem  $\mathcal{A}$  that gives a nonvanishing negativity: a neighbouring star plaquette pair (SPP), embedded in an arbitrarily large lattice of  $2N$  spins. This system choice allows us to gain some intuition about the behaviour of the negativity at finite temperatures without any of the additional complications presented by more general partition schemes. The fact that we cannot simply consider a neighbouring pair of spins is a direct consequence of the four-body interactions present in the Hamiltonian (6). The SPP system offers a number of technical advantages. First, the negativity  $\mathcal{E}$  is both a necessary and sufficient condition for separability, which we prove in Sec. III A. This allows us to study the entanglement properties of the system exactly. Also, the SPP is symmetric under interchange of stars with plaquettes, implying that the negativity must obey the symme-

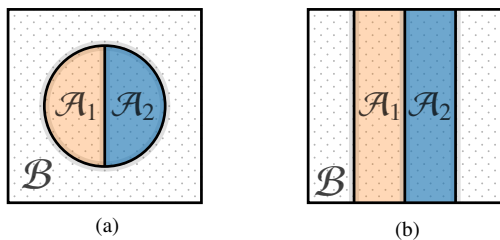


FIG. 2. Schematic depiction of the two types of partitions considered in this manuscript. Note that in (a) the subsystem  $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$  is contractible, while in (b) it is not. As a result, we expect to see a constant topological contribution to the zero-temperature negativity when using partition scheme (b) [14, 15].

try  $\mathcal{E}(\lambda_A, \lambda_B) = \mathcal{E}(\lambda_B, \lambda_A)$ . Finally, one can diagonalise the operator  $\rho_{\mathcal{A}}^{T_2}$  explicitly without having to resort to the replica trick. A short calculation, outlined in Sec. III A, gives

$$\mathcal{E}(T) = \ln \left[ 1 + \frac{1}{2} \max(\eta_A + \eta_B + \eta_A \eta_B - 1, 0) \right], \quad (7)$$

where we have defined

$$\eta_X = \frac{\tanh(\beta \lambda_X) + \tanh(\beta \lambda_X)^{N-1}}{1 + \tanh(\beta \lambda_X)^N}, \quad (8)$$

for  $X = A, B$ . The expression (7) manifestly respects the symmetry under interchange of  $\lambda_A$  with  $\lambda_B$ . Since  $\eta_X$  is a monotonic increasing function of  $\beta$ , the negativity is a monotonically decreasing function of temperature, with a single discontinuity in its first derivative at the sudden death temperature  $T_c$ . Above  $T_c$ , the state of subsystem  $\mathcal{A}$  is separable. Agreement with numerical exact diagonalisation of  $\rho_{\mathcal{A}}^{T_2}$  is perfect (not shown).

Recall that when  $N \gg 1$ , the defect density of species  $X$ , when the system is in thermodynamic equilibrium, is approximately given by the Fermi-Dirac distribution  $n_F(2\lambda_X) = (e^{2\beta\lambda_X} + 1)^{-1}$ . The condition for sudden death is then written most succinctly in terms of the density of holes  $n_h(2\lambda_X) \equiv 1 - n_F(2\lambda_X)$ , i.e., the density of nondefective stars (plaquettes) for  $X = A(B)$ . In the thermodynamic limit, this condition becomes

$$n_h(2\lambda_A)n_h(2\lambda_B) = \frac{1}{2}. \quad (9)$$

That is, at  $T_c$ , the geometric mean of the hole densities must equal  $1/\sqrt{2}$ . When the coupling constants are well separated, one can obtain an approximate explicit expression for the critical temperature

$$\frac{T_c}{\lambda_A} \simeq \frac{2\lambda_B/\lambda_A}{W\left(2\lambda_B/\lambda_A\right)} \sim \frac{2\lambda_B/\lambda_A}{\ln\left(2\lambda_B/\lambda_A\right)} \quad \text{for } \lambda_B \gg \lambda_A, \quad (10)$$

where  $W(x)$  is the product-log function. Exploiting the duality, the expression for  $\lambda_A \gg \lambda_B$  is found by exchanging  $\lambda_A \leftrightarrow \lambda_B$ . The energy scale for sudden death is set by the larger of the two coupling constants, closely mirroring the perturbative expressions of Ref. 4. In the classical limit  $\lambda_A \rightarrow 0$ , keeping  $\lambda_B$  fixed,  $T_c$  vanishes logarithmically.

The result (7) allows for the following physical interpretation. When the temperature is finite but low compared to  $\lambda_A$ , the densities of both defect species are exponentially suppressed ( $n_h \simeq 1 - e^{-2\beta\lambda_X}$ ). As can be seen in Fig. 3, over the temperature range  $T \lesssim \lambda_A$ , the negativity remains essentially undegraded from its zero-temperature value,  $\mathcal{E} \simeq \mathcal{E}(0)$ . Namely, a density of order one defect per site of the lattice of the least energetically costly defects is needed to start affecting the quantum correlations between  $\mathcal{A}_1$  and  $\mathcal{A}_2$ . Above the temperature threshold  $T \sim \lambda_A$ , electric charge defects proliferate in the system and correspondingly the negativity decays like  $\mathcal{E} \sim \lambda_A/T$ . This decay is cut off at a temperature  $T \sim \lambda_B$  (up to logarithmic factors), corresponding to an  $\mathcal{O}(1)$  density of the most energetically costly (magnetic) defects in the system, where the negativity vanishes.

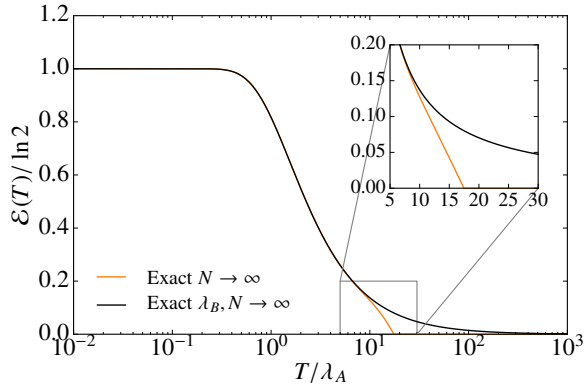


FIG. 3. The analytic result (7), and its limit with  $\lambda_B \rightarrow \infty$  (18), for the SPP embedded in an infinite system  $N \rightarrow \infty$  with  $\lambda_B/\lambda_A = 25$ . The  $\lambda_B \rightarrow \infty$  limit provides an excellent approximation up to a temperature of order  $\lambda_B$ , where it becomes possible to thermally excite magnetic vortices and the assumption of undegraded plaquette magnetisation  $M(\alpha) = N$  is no longer valid.

### B. Extended boundary

The SPP results are simple and intuitive, but do not tell us how the negativity depends on the length of the  $\mathcal{A}_1$ – $\mathcal{A}_2$  boundary. To address this, we consider the subsystem  $\mathcal{A}$  depicted in Fig. 2b, i.e., a “strip” that spans the torus in one direction. One can imagine constructing this partition scheme by tiling neighbouring SPPs. Choosing such a noncontractible subsystem makes the calculation slightly simpler than for a strip of finite length, by eliminating its end points (although, as shown in Sec. III B, the two results are indeed in agreement).

In the thermodynamic limit  $N \rightarrow \infty$ , keeping the length of the boundary  $N_\partial$  finite (namely, for a system wrapped around a cylinder of infinite length), the replica method gives (see Sec. III B)

$$\mathcal{E}(T) = \ln \langle |Z(\{\gamma_s\}, \{k_s\}, T)| \rangle. \quad (11)$$

Here  $Z$  is proportional to the partition function of a disordered 1D Ising chain with periodic boundary conditions in a complex magnetic field,

$$Z = \frac{1}{(\cosh \beta \lambda_B)^{N_\partial}} \sum_{\{\tau_s = \pm 1\}} e^{-H(\{\tau_s\})}, \quad (12)$$

with the classical “Hamiltonian”

$$H(\{\tau_s\}) = \sum_{s=1}^{N_\partial} \left[ \gamma_s \tau_s \tau_{s+1} + \frac{1}{2} (K_A + i\pi k_s) (\tau_s + 1) \right]. \quad (13)$$

The angled brackets in Eq. (11) refer to a disorder average over the variables  $\{\gamma_s\}$  and  $\{k_s\}$ , with  $\gamma_s \in \{-\beta\lambda_B, \beta\lambda_B\}$  and  $k_s \in \{0, 1\}$  with equal probability.

Note that because the magnetic field is complex, the partition function is not necessarily positive. Details of the derivation are given in Sec. III B, and we have confirmed that the finite-size version of Eq. (11) [that is, Eq. (58) in Sec. III B]

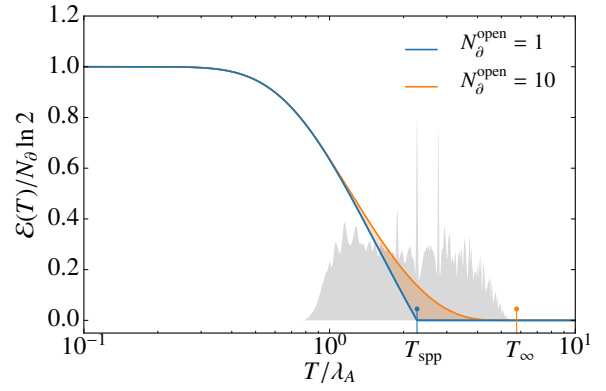


FIG. 4. Negativity for a contractible strip partition, as discussed in the text, with boundary length  $N_\partial^{\text{open}}$  and  $\lambda_A = \lambda_B$ . As the boundary between  $\mathcal{A}_1$  and  $\mathcal{A}_2$  is made longer,  $\mathcal{E}$  develops an increasing number of discontinuities (at temperatures  $T_i$ ) in its first derivative whose effect is to “smooth” the function  $\mathcal{E}(T)$  and push  $T_c$  to higher temperatures. In the background is plotted the “density of discontinuities”  $\sum_i \delta(T - T_i)$  for  $N_\partial^{\text{open}} = 7$ . The vertical lines indicate the positions of  $T_c$  for the SPP and for an infinite boundary.

agrees precisely with numerical exact diagonalisation results for numerically accessible system sizes (not shown).

Because the subsystem  $\mathcal{A}$  spans the torus in one direction, in addition to leading area law behaviour, the zero-temperature negativity  $\mathcal{E}(0) = (N_\partial - 1) \ln 2$  contains a constant topological contribution, consistent with Refs. 14 and 15. The reader is referred to Sec. III B for a derivation of this result. At zero temperature, the result remains unchanged for more general partitions in which subsystems  $\mathcal{A}_1$  and  $\mathcal{A}_2$  have larger widths. This reflects that the boundary is the source of entanglement, and that the toric code is a zero-range correlated model [14]. It is worth noting that although  $\lim_{T \rightarrow 0} \rho$  is not pure, owing to the ground state degeneracy, each topological sector gives an independent and equal contribution to  $\mathcal{E}(0)$ .

One can show that the disorder average without taking the absolute value is identically unity, i.e.,  $\langle Z \rangle \equiv 1$ , leading to a vanishing negativity. If at least one disorder realisation satisfies  $Z(\{\gamma_s\}, \{k_s\}) < 0$ , then the average  $\langle |Z| \rangle > 1$  and consequently  $\mathcal{E} > 0$ , indicating the presence of entanglement. For any particular disorder realisation, there exists some temperature  $T_*(\{\gamma_s\}, \{k_s\}) \geq 0$  above which  $Z(\{\gamma_s\}, \{k_s\}) > 0$  for all temperatures  $T \geq T_*$ . Such a temperature must exist since  $Z \rightarrow 1$  for all disorder realisations in the limit of infinite temperature. The largest of these temperatures corresponds to the sudden death temperature  $T_c = \max_{\{\gamma_s, k_s\}} T_*$ .

As temperature is increased from zero, similarly to the SPP case in Sec. II A, the negativity remains undegraded,  $\mathcal{E} \simeq \mathcal{E}(0)$ , until the density of thermally excited electric defects becomes of  $\mathcal{O}(1)$ . It then vanishes identically above some temperature  $T_c$ , discussed in more depth below. The SPP and extended case however differ in that the latter exhibits a large number of discontinuities. These discontinuities correspond to the (Lee-Yang [44, 45]) zeros of the complex partition functions, whose effect is to “smooth out”  $\mathcal{E}(T)$ , and to push  $T_c$  to higher temperatures (see Fig. 4). This number grows rapidly

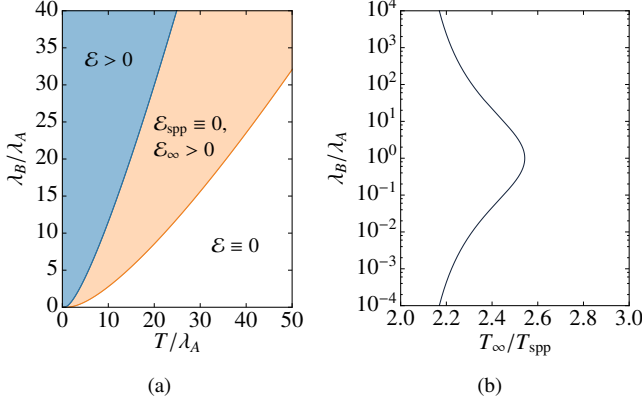


FIG. 5. (a) Entanglement phase diagram for the two-dimensional toric code, as predicted by the negativity. As temperature is increased, defects are thermally excited and the negativity is degraded until, at  $T_{\text{SPP}}$ , the SPP negativity vanishes identically. However, entanglement between subsystems that share a boundary with  $N_\partial > 1$  is more robust, and able to survive above  $T_{\text{SPP}}$ . When  $T_\infty$  is reached, the negativity vanishes identically even for strip partitions that span the entire system. (b) Ratio  $T_\infty/T_{\text{SPP}}$  of the two limiting temperatures.

(exponentially for  $N_\partial \gg 1$ ) with increasing boundary length, e.g., 1, 19, 78, ... for  $N_\partial = 2, 3, 4, \dots$  and  $\lambda_A = \lambda_B$ .

The disorder realisation that maximises  $T_*$ , i.e., the partition function with the largest zero, corresponds to the special ordered configuration of parameters  $\gamma_s = \beta\lambda_B$  and  $k_s = 1, \forall s$ . Since the resulting Ising chain is translationally invariant, one can use a transfer matrix approach to calculate  $Z$  analytically. The critical temperature  $T_c$  is then found by solving  $Z = 0$ , which in general has multiple solutions as the partition function oscillates. The largest one corresponds to  $T_c$ , while the others correspond to the positions of discontinuities at lower temperatures. Following this argument, one arrives at the following implicit expression for  $T_c$ :

$$e^{4\beta_c(\lambda_A + \lambda_B)} - e^{4\beta_c\lambda_A} - e^{4\beta_c\lambda_B} = \tan^2\left(\frac{\pi}{2N_\partial}\right), \quad (14)$$

which manifestly respects the duality in  $\lambda_A \leftrightarrow \lambda_B$ , as expected. In contrast to the SPP system of Sec. II A, this equation remains the same when  $N$  is taken to be finite, and even when subsystems  $\mathcal{A}_1$  and  $\mathcal{A}_2$  have nontrivial widths (not shown).

Writing the sudden death condition in terms of the number density of each type of excitation, and defining for convenience  $\Delta_X \equiv n_h(2\lambda_X) - n_F(2\lambda_X) > 0$ , one obtains

$$\frac{n_F(2\lambda_A)n_F(2\lambda_B)}{(\Delta_A\Delta_B)^{1/2}} = \cos\left(\frac{\pi}{2N_\partial}\right). \quad (15)$$

That is, the negativity vanishes when the geometric mean of the number densities of both types of excitations, each weighted by  $\Delta_X^{-1/2}$ , equals or exceeds an  $\mathcal{O}(1)$  number (which tends towards unity in the limit of a long boundary). The similarity between Eq. (9) and Eq. (15) is apparent. Arguably the most important conclusion we can draw from Eq. (15) is that the critical temperature  $T_c$  remains finite in the limit of

large boundaries  $N_\partial \rightarrow \infty$ : indeed  $T_c \sim \lambda_B$ , as in the SPP case. For a system-spanning strip, the limiting behaviour of  $T_c$  is

$$\frac{T_c}{\lambda_A} \simeq \frac{4\frac{\lambda_B}{\lambda_A}}{W\left(\frac{\lambda_B}{\lambda_A}\right)} \quad \text{for } \lambda_B \gg \lambda_A, N_\partial \rightarrow \infty, \quad (16)$$

which is to be contrasted with Eq. (10).

Our main findings are embodied by the entanglement ‘‘phase diagram’’ in Fig. 5. The entanglement between subsystems that share an extended boundary is always more robust to the effects of thermal fluctuations than the SPP system, but the two sudden death temperatures only differ by a factor of approximately two, as shown in Fig. 5b. From these results, it seems reasonable to expect that for all finite  $\lambda_B/\lambda_A$ , there exists a temperature  $T_\infty$  above which the negativity vanishes for any tripartition of the lattice. We expect this temperature to be controlled by the largest gap to elementary excitations in the system, up to a prefactor of  $\mathcal{O}(1)$ .

In Sec. III B we verify that these results remain qualitatively unchanged for a *contractible* subsystem, and moreover that both contractible and system-spanning subsystems give rise to the same sudden death temperature  $T_\infty$  in the limit of infinite boundaries  $N_\partial \rightarrow \infty$ .

### C. $\lambda_B \rightarrow \infty$ limit

When one of the two coupling constants in the model is infinite,  $\lambda_B \rightarrow \infty$ , the formation of magnetic vortices is rendered energetically impossible for any finite temperature. Conversely, electric charges are present in the system. This hard constraint means that the subspace of thermally accessible states is substantially smaller than the full Hilbert space, and the calculation is greatly simplified. For the partition scheme in Fig. 2a, we arrive at the expression (see Sec. III C for details)

$$e^{\mathcal{E}(T)} = \frac{\cosh\left(\frac{1}{2}(N - N_\partial)K_A\right)}{\cosh\left(\frac{1}{2}NK_A\right)} \left[2 \cosh\left(\frac{K_A}{2}\right)\right]^{N_\partial}, \quad (17)$$

where the temperature dependence is controlled by the parameter  $K_A \equiv -\ln \tanh \beta\lambda_A$ . In this limit, the negativity does not exhibit any discontinuities. When evaluated for  $N_\partial = 1$ , this expression correctly agrees with the SPP result (7) in the appropriate limit  $\lambda_B \rightarrow \infty$ .

If we take the thermodynamic limit  $N \rightarrow \infty$  while keeping the length of the boundary  $N_\partial$  finite, this expression simplifies further, and its extensivity in the number of boundary stars becomes apparent. Introducing the density of nondefective stars  $n_h$ , the negativity becomes

$$\mathcal{E}(T) = N_\partial \ln(2n_h). \quad (18)$$

For *all* temperatures, the negativity for  $N \rightarrow \infty$  satisfies an exact area law.

At zero temperature, the density of excitations is zero, or  $n_h = 1$ , and the negativity assumes the value  $\mathcal{E}(0) = N_\partial \ln 2$ , in agreement with the results of Ref. 15. Since the subsystem  $\mathcal{A}$  is now considered contractible, the topological contribution

seen in Sec. II B is removed. If the subsystems  $\mathcal{A}_1$  and  $\mathcal{A}_2$  do not share a boundary then the negativity vanishes identically for all temperatures.

The  $1/T$  decay of the negativity above  $T \sim \lambda_A$  that was observed in the SPP is no longer cut off at high temperatures. Sending  $\lambda_B \rightarrow \infty$  prevents the thermal excitation of magnetic defects, thereby locking in the zero-temperature magnetic loop structure. As a result, thermal fluctuations are never sufficient to rid the system of all quantum entanglement.

Since the ground states are independent of the system's parameters, so is  $\mathcal{E}(0)$ . This peculiarity of the toric code means that entanglement is able to survive up to arbitrarily large temperatures, suffering only a slow  $1/T$  demise. In other models this is not always possible as divergence of the sudden death temperature is often accompanied by a vanishing zero-temperature negativity. This is the case in the 1D transverse field Ising model, for example, which we discuss in Appendix B.

### III. CALCULATIONS

We begin this section with some general considerations about the calculation of the entanglement negativity  $\mathcal{E}(T)$  of the toric code. We will then consider specific cases that allow us to streamline the calculations and shape the final results in the relatively simple and insightful formulae discussed in Sec. II.

We consider a system of  $2N$  spins in thermodynamic equilibrium with a thermal reservoir at temperature  $T = 1/\beta$  (we set the Boltzmann constant  $k_B = 1$  throughout). The system is therefore described by the canonical density matrix  $\rho = e^{-\beta H}/Z$ , where  $H$  is the Hamiltonian and  $Z = \text{Tr} e^{-\beta H}$  is the partition function (to ensure proper normalisation,  $\text{Tr} \rho = 1$ ). The density matrix can be written in terms of its matrix elements with respect to complete bases  $\{|\alpha\rangle\}$  and  $\{|\gamma\rangle\}$  as

$$\rho = \frac{1}{Z} \sum_{\alpha, \gamma} \langle \alpha | e^{-\beta H} | \gamma \rangle |\alpha\rangle \langle \gamma|. \quad (19)$$

Following the work in Refs. 14, 37, 46–48, we choose the tensor product basis of eigenstates of the operator  $\otimes_i \sigma_i^z$ . The operators  $B_p \equiv \prod_{i \in p} \sigma_i^z$  are diagonal in this basis:  $P|\alpha\rangle = M(\alpha)|\alpha\rangle$ , where one can interpret the eigenvalue  $M(\alpha)$  of  $P = \sum_p B_p$  as a ‘‘plaquette magnetisation’’ equal to the sum of all local magnetisations  $M_p \equiv \langle \alpha | B_p | \alpha \rangle$ . We then introduce the group  $\mathcal{G}$  generated by products of star operators  $A_s$ . The group elements  $g \in \mathcal{G}$  must be defined modulo the identity for  $g^{-1}$  to be unique, since  $\prod_s A_s = \mathbb{1}$ . Only states that differ by the action of a group element  $g \in \mathcal{G}$  give nonvanishing matrix elements in (19), which allows us to write  $|\gamma\rangle = g|\alpha\rangle$ . Hence [14, 37, 46–48],

$$\rho = \frac{1}{Z} \sum_{\alpha} \sum_{g \in \mathcal{G}} e^{\beta \lambda_B M(\alpha)} \langle \alpha | e^{\beta S} g | \alpha \rangle |\alpha\rangle \langle \alpha | g. \quad (20)$$

Following similar considerations, the partition function  $Z$  can be written as

$$Z = (\cosh \beta \lambda_A)^N [1 + (\tanh \beta \lambda_A)^N] \cdot \sum_{\alpha} e^{\beta \lambda_B M(\alpha)} \quad (21)$$

$$\equiv Z_A \cdot Z_B. \quad (22)$$

Notice that the system is symmetric upon exchanging star and plaquette operators, and  $\lambda_A \leftrightarrow \lambda_B$ . Therefore, it would be completely equivalent to choose the tensor product basis of the operator  $\otimes_i \sigma_i^x$ , and so we must have  $Z_A(\beta \lambda) \propto Z_B(\beta \lambda)$ .

To evaluate the matrix elements in (20), it is helpful to expand the exponential in terms of its constituent star operators. Remembering that  $S = \sum_s A_s$  and that  $A_s^2 = \mathbb{1}$ ,

$$e^{\beta \lambda_A S} = \prod_s [\cosh(\beta \lambda_A) + \sinh(\beta \lambda_A) A_s], \quad (23)$$

and we thus obtain

$$\frac{1}{Z_A} e^{\beta \lambda_A S} = \sum_{\tilde{g} \in \mathcal{G}} \eta_T(\tilde{g}) \tilde{g}. \quad (24)$$

Here we introduced for convenience of notation the weighting factor

$$\eta_T(g) = \frac{e^{-K_A n(g)} + e^{-K_A [N - n(g)]}}{1 + e^{-K_A N}}, \quad (25)$$

where  $K_A \equiv -\ln \tanh \beta \lambda_A$  and  $n(g)$  is the number of star operators  $A_s$  that appear in the decomposition of the group element  $g$ . Note that  $\eta_T(g)$  is invariant under  $n(g) \rightarrow N - n(g)$ , and therefore the ambiguity in the definition of the group elements of  $\mathcal{G}$  modulo the identity is immaterial [14, 37, 46–48]. We then arrive at the expression for the density matrix

$$\rho = \frac{1}{Z_B} \sum_{\alpha} \sum_{g \in \mathcal{G}} e^{\beta \lambda_B M(\alpha)} \eta_T(g) |\alpha\rangle \langle \alpha | g. \quad (26)$$

If we are interested only in a subsystem  $\mathcal{A}$  of the total system  $S$ , then we should trace out the complementary subsystem  $\mathcal{B}$  to form the reduced density matrix  $\rho_{\mathcal{A}} = \text{Tr}_{\mathcal{B}} \rho$ . To achieve this, we decompose  $|\alpha\rangle = |\alpha_{\mathcal{A}}\rangle \otimes |\alpha_{\mathcal{B}}\rangle$  and  $g = g_{\mathcal{A}} \otimes g_{\mathcal{B}}$ ,

$$\rho_{\mathcal{A}} = \frac{1}{Z_B} \sum_{\alpha} \sum_{g \in \mathcal{G}} e^{\beta \lambda_B M(\alpha)} \eta_T(g) \times \langle \alpha_{\mathcal{B}} | g_{\mathcal{B}} | \alpha_{\mathcal{B}} \rangle |\alpha_{\mathcal{A}}\rangle \langle \alpha_{\mathcal{A}} | g_{\mathcal{A}}. \quad (27)$$

Now, the matrix element equals unity if  $g$  acts trivially on the  $\mathcal{B}$  subsystem, and zero otherwise. The resulting constraint on  $g$  is implemented by restricting the summation to group elements belonging to the subgroup  $\{g \in \mathcal{G} \mid g_{\mathcal{B}} = \mathbb{1}_{\mathcal{B}}\} \equiv \mathcal{G}_{\mathcal{A}} \subset \mathcal{G}$ .

Finally, we recall that  $\mathcal{A}$  is further partitioned into two subsystems ( $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$ ), and that we want to take the partial transpose over one of them (say,  $\mathcal{A}_2$ ). This is effected by splitting up the states  $|\alpha_{\mathcal{A}}\rangle = |\alpha_{\mathcal{A}_1}\rangle \otimes |\alpha_{\mathcal{A}_2}\rangle$ , and similarly for the group elements  $g_{\mathcal{A}} = g_{\mathcal{A}_1} \otimes g_{\mathcal{A}_2}$ .

$$\rho_{\mathcal{A}}^{T_2} = \frac{1}{Z_B} \sum_{\alpha} \sum_{g \in \mathcal{G}_{\mathcal{A}}} e^{\beta \lambda_B M(\alpha)} \eta_T(g) \times (|\alpha_{\mathcal{A}_1}\rangle \langle \alpha_{\mathcal{A}_1} | g_{\mathcal{A}_1}) \otimes (g_{\mathcal{A}_2} |\alpha_{\mathcal{A}_2}\rangle \langle \alpha_{\mathcal{A}_2}|). \quad (28)$$

We can then apply the replica approach (5) by taking the trace of the  $n$ th power of this operator,

$$\begin{aligned} \text{Tr}(\rho_{\mathcal{A}}^{T_2})^n &= \frac{1}{Z_B^n} \sum_{\alpha_1, \dots, \alpha_n} \sum_{g_1, \dots, g_n \in \mathcal{G}_{\mathcal{A}}} \left( \prod_{\ell=1}^n e^{\beta \lambda_B M(\alpha_\ell)} \eta_T(g_\ell) \right) \\ &\quad \langle \alpha_{1, \mathcal{A}_1} | g_{1, \mathcal{A}_1} | \alpha_{2, \mathcal{A}_1} \rangle \cdots \langle \alpha_{(n-1), \mathcal{A}_1} | g_{(n-1), \mathcal{A}_1} | \alpha_{n, \mathcal{A}_1} \rangle \langle \alpha_{n, \mathcal{A}_1} | g_{n, \mathcal{A}_1} | \alpha_{1, \mathcal{A}_1} \rangle \\ &\quad \langle \alpha_{1, \mathcal{A}_2} | g_{2, \mathcal{A}_2} | \alpha_{2, \mathcal{A}_2} \rangle \cdots \langle \alpha_{(n-1), \mathcal{A}_2} | g_{(n-1), \mathcal{A}_2} | \alpha_{n, \mathcal{A}_2} \rangle \langle \alpha_{n, \mathcal{A}_2} | g_{1, \mathcal{A}_2} | \alpha_{1, \mathcal{A}_2} \rangle . \end{aligned} \quad (29)$$

It is helpful to notice that the subgroup  $\mathcal{G}_{\mathcal{A}}$  can in general be decomposed as  $\mathcal{G}_{\mathcal{A}} = \mathcal{G}_{\mathcal{A}_1} \mathcal{G}_{\mathcal{A}_2} \mathcal{G}_{\mathcal{A}_1, \mathcal{A}_2}$ . The subgroups  $\mathcal{G}_{\mathcal{A}_i}$  are defined as  $\mathcal{G}_{\mathcal{A}_i} \equiv \{g \in \mathcal{G}_{\mathcal{A}} \mid g_{\bar{\mathcal{A}}_i} = \mathbb{I}_{\bar{\mathcal{A}}_i}\}$  with  $\bar{\mathcal{A}}_1 = \mathcal{A}_2$  and vice versa, while  $\mathcal{G}_{\mathcal{A}_1, \mathcal{A}_2} \equiv \mathcal{G}_{\mathcal{A}} / (\mathcal{G}_{\mathcal{A}_1} \mathcal{G}_{\mathcal{A}_2})$  is the quotient group. Any element  $g \in \mathcal{G}_{\mathcal{A}}$  can therefore be uniquely decomposed into the product of three group elements: one that acts exclusively on  $\mathcal{A}_1$ , one exclusively on  $\mathcal{A}_2$ , and one that acts simultaneously (and exclusively) on  $\mathcal{A}_1$  and  $\mathcal{A}_2$ . Namely,  $g = \bar{g} \bar{g} \theta$ , with  $\bar{g} \in \mathcal{G}_{\mathcal{A}_1}$ ,  $\bar{g} \in \mathcal{G}_{\mathcal{A}_2}$ , and  $\theta \in \mathcal{G}_{\mathcal{A}_1, \mathcal{A}_2}$  [14].

Using this decomposition, we make the following relabelling of the states  $|\alpha_k\rangle$ , for all  $k > 1$ :

$$|\alpha'_k\rangle = \left( \prod_{\ell=1}^{k-1} \bar{g}_\ell \right) \left( \prod_{\ell=2}^k \bar{g}_\ell \right) |\alpha_k\rangle \rightarrow |\alpha_k\rangle . \quad (30)$$

This removes all dependence on  $\bar{g}_\ell$  and  $\bar{g}_\ell$  from the matrix elements, and we obtain a more compact expression:

$$\text{Tr}(\rho_{\mathcal{A}}^{T_2})^n = \frac{1}{Z_B^n} \sum_{\alpha_1, \dots, \alpha_n} \sum_{\substack{\bar{g}_1, \dots, \bar{g}_n \\ \in \mathcal{G}_{\mathcal{A}_1}}} \sum_{\substack{\bar{g}_1, \dots, \bar{g}_n \\ \in \mathcal{G}_{\mathcal{A}_2}}} \sum_{\substack{\theta_1, \dots, \theta_n \\ \in \mathcal{G}_{\mathcal{A}_1, \mathcal{A}_2}}} \langle 0 | \prod_{\ell=1}^n \bar{g}_\ell \bar{g}_\ell \theta_\ell | 0 \rangle \left( \prod_{\ell=1}^n e^{\beta \lambda_B M(\alpha_\ell)} \eta_T(\bar{g}_\ell \bar{g}_\ell \theta_\ell) \langle \alpha_{\ell, \mathcal{A}_1} | \theta_{\ell, \mathcal{A}_1} \otimes \theta_{\ell+1, \mathcal{A}_2} | \alpha_{\ell+1, \mathcal{A}_2} \rangle \right), \quad (31)$$

where  $\theta_{n+1} \equiv \theta_1$  and similarly  $\alpha_{n+1} \equiv \alpha_1$ . Notice that the matrix elements in Eq. (31) impose constraints on which terms  $\bar{g}_\ell$  and  $\bar{g}_\ell$  in the summation give a nonvanishing contribution. Furthermore, they implicitly impose the constraint  $\prod_{\ell} \theta_\ell = \mathbb{I}$ . This is because, given  $\alpha_1$  (say), it takes only  $n-1$  of the  $n$  matrix elements to uniquely determine all other  $\alpha_\ell$ ; the final matrix element then evaluates to  $\langle \alpha_1 | \prod_{\ell} \theta_\ell | \alpha_1 \rangle$ , which indeed is nonvanishing only if the aforementioned constraint is satisfied.

Equation (31) represents the general form of  $\text{Tr}(\rho_{\mathcal{A}}^{T_2})^n$  for a tripartition of the system, and is the starting point of all replica method calculations discussed later.

### A. Star plaquette pair

It is convenient to consider first the case of the smallest non-trivial subsystem: a single star plaquette pair, shown in Fig. 6a. Notice that subsystem  $\mathcal{A}$  is symmetric under interchange of stars with plaquettes, which has the important implication that  $\mathcal{E}$  must be symmetric upon exchanging  $\lambda_A \leftrightarrow \lambda_B$ . This property will be used at the end of this section: given the dependence of  $\mathcal{E}$  on  $\lambda_A$ , say, one can then deduce the dependence on  $\lambda_B$  from the relation  $\mathcal{E}(\lambda_A, \lambda_B) = \mathcal{E}(\lambda_B, \lambda_A)$ .

Rather than proceeding via the replica method, it is more convenient in this case to diagonalise the partial transpose  $\rho_{\mathcal{A}}^{T_2}$  of the reduced density matrix explicitly. We start with the general expression for the partially transposed reduced density matrix (28) derived earlier:

$$\begin{aligned} \rho_{\mathcal{A}}^{T_2} &= \sum_{\alpha_{\mathcal{A}}} \sum_{g \in \mathcal{G}_{\mathcal{A}}} e^{\beta \lambda_B M_{\partial}(\alpha_{\mathcal{A}})} \xi_T(M_{\partial}) \eta_T(g) \\ &\quad (|\alpha_{\mathcal{A}_1}\rangle \langle \alpha_{\mathcal{A}_1}| g_{\mathcal{A}_1}) \otimes (g_{\mathcal{A}_2} |\alpha_{\mathcal{A}_2}\rangle \langle \alpha_{\mathcal{A}_2}|) . \end{aligned} \quad (32)$$

The sum over spins  $i \in \mathcal{B}$  has been absorbed into the factors  $\xi_T(M_{\partial}) \equiv \frac{1}{Z_B} \sum_{\alpha_{\mathcal{B}}} e^{\beta \lambda_B M_{\partial}(\alpha_{\mathcal{A}})}$  by splitting the magnetisation into a boundary component  $M_{\partial} \equiv \langle B_{p \in \partial} \rangle$  and its complement  $M_{\bar{\partial}} = \sum_{p \notin \partial} \langle B_p \rangle$  so that  $M(\alpha) = M_{\partial}(\alpha_{\mathcal{A}}) + M_{\bar{\partial}}(\alpha)$ . Although at first sight it may appear that  $\xi_T$  could assume a different value for each spin configuration  $\alpha_{\mathcal{A}}$ , it turns out to depend only on the magnetisation  $M_{\partial}$  of the boundary plaquette, as the notation indicates. For brevity we ask the reader to take this result for granted until Sec. III B, where we provide a proof.

The subsystem contains only one star which acts simultaneously on  $\mathcal{A}_1$  and  $\mathcal{A}_2$ . Therefore the group  $\mathcal{G}_{\mathcal{A}}$  coincides with  $\mathcal{G}_{\mathcal{A}_1, \mathcal{A}_2}$  and contains only one nontrivial element: the boundary

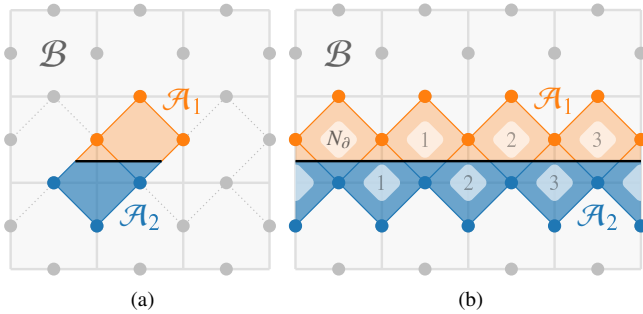


FIG. 6. Illustration of (a) a star plaquette pair and (b) a strip partition chosen for the calculation of  $\mathcal{E}$  when the boundary length  $N_{\partial} > 1$ . In both cases, the subsystem  $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$  is symmetric under interchange of stars and plaquettes, and in (b) it spans the torus in one direction.

star operator  $A_s$ . Decomposing it into components that act separately on subsystems  $\mathcal{A}_1$  and  $\mathcal{A}_2$ ,  $A_s = A_s^{(1)} \otimes A_s^{(2)}$ , Eq. (32) can be written as

$$\rho_{\mathcal{A}}^{T_2} = \sum_{\alpha_{\mathcal{A}}} e^{\beta \lambda_B M_{\partial}(\alpha_{\mathcal{A}})} \xi_T(M_{\partial}) \left( |\alpha_{\mathcal{A}}\rangle \langle \alpha_{\mathcal{A}}| + \eta_A |\alpha_{\mathcal{A}_1} \otimes A_s^{(2)} \alpha_{\mathcal{A}_2}\rangle \langle A_s^{(1)} \alpha_{\mathcal{A}_1} \otimes \alpha_{\mathcal{A}_2}| \right), \quad (33)$$

where  $\eta_A$  was defined in Eq. (8). Relabelling  $\alpha_{\mathcal{A}_1} \otimes A_s^{(2)} \alpha_{\mathcal{A}_2} = \alpha'_{\mathcal{A}}$  in the second term, we obtain

$$|\alpha_{\mathcal{A}_1} \otimes A_s^{(2)} \alpha_{\mathcal{A}_2}\rangle \langle A_s^{(1)} \alpha_{\mathcal{A}_1} \otimes \alpha_{\mathcal{A}_2}| = |\alpha'_{\mathcal{A}}\rangle \langle A_s \alpha'_{\mathcal{A}}|. \quad (34)$$

We can then change the mute index  $\alpha'_{\mathcal{A}} \rightarrow \alpha_{\mathcal{A}}$  and denote  $|\bar{\alpha}_{\mathcal{A}}\rangle = A_s |\alpha_{\mathcal{A}}\rangle$  to obtain

$$\rho_{\mathcal{A}}^{T_2} = \sum_{\alpha_{\mathcal{A}}} e^{\beta \lambda_B M_{\partial}(\alpha_{\mathcal{A}})} \xi_T(M_{\partial}) |\alpha_{\mathcal{A}}\rangle \langle \alpha_{\mathcal{A}}| + \eta_A e^{-\beta \lambda_B M_{\partial}(\alpha_{\mathcal{A}})} \xi_T(-M_{\partial}) |\alpha_{\mathcal{A}}\rangle \langle \bar{\alpha}_{\mathcal{A}}|. \quad (35)$$

Each state  $|\alpha_{\mathcal{A}}\rangle$  of the subsystem  $\mathcal{A}$  is coupled only to itself and its counterpart with the spins  $i \in s$  flipped. The expression (35) implies that  $\rho_{\mathcal{A}}^{T_2}$  can be written as a block diagonal matrix with respect to the eigenstates of  $\otimes_{i \in \mathcal{A}} \sigma_i^z$ , with the  $2 \times 2$  matrices

$$\rho_{\pm} = \begin{pmatrix} \Xi^{\pm} & \eta_A \Xi^{\mp} \\ \eta_A \Xi^{\mp} & \Xi^{\pm} \end{pmatrix} \quad (36)$$

along the diagonal (note that these are not density matrices). For convenience we introduced the shorthand  $\Xi^{\pm}$  to represent  $\xi_T(M_{\partial}) e^{\beta \lambda_B M_{\partial}}$  evaluated with boundary magnetisation  $M_{\partial} = \pm 1$ . Since the number of spin configurations  $\alpha_{\mathcal{A}}$  with  $M_{\partial} = +1$  equals the number with  $M_{\partial} = -1$ , there exist an equal number of  $\rho_+$  and  $\rho_-$  blocks along the diagonal of  $\rho_{\mathcal{A}}^{T_2}$ . Therefore the normalisation condition  $\text{Tr} \rho_{\mathcal{A}}^{T_2} = 1$  gives  $\Xi^+ + \Xi^- = 2^{1-N_{\mathcal{A}}}$ , where  $N_{\mathcal{A}} \equiv 6$  equals the number of spins  $i \in \mathcal{A}$ .

After calculating the eigenvalues of  $\rho_{\pm}$ , invoking the symmetry  $\mathcal{E}(\lambda_A, \lambda_B) = \mathcal{E}(\lambda_A, \lambda_B)$  allows us to conjecture that  $\Xi^+ - \Xi^- = 2^{1-N_{\mathcal{A}}} \eta_B$ . This implies that  $\Xi^{\pm} = (1 \pm \eta_B)/2^{N_{\mathcal{A}}}$ , and

$$e^{\mathcal{E}(T)} = 1 + \frac{1}{2} \max(\eta_A + \eta_B + \eta_A \eta_B - 1, 0). \quad (37)$$

Alternatively, one can follow the (lengthier) calculations in Sec. III B to evaluate  $\Xi^{\pm}$  explicitly; we find indeed that the latter approach reproduces (37) exactly, without relying on the symmetry argument.

That the negativity for the SPP is both a necessary and sufficient condition for separability follows almost trivially from the fact that  $\mathcal{E} = 0$  is a sufficient condition for separability when  $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$  [49]. In a similar manner to (33), the reduced density matrix of the SPP is

$$\rho_{\mathcal{A}} = \sum_{\alpha_{\mathcal{A}}} e^{\beta \lambda_B M_{\partial}(\alpha_{\mathcal{A}})} \xi_T(M_{\partial}) \left( |\alpha_{\mathcal{A}}\rangle \langle \alpha_{\mathcal{A}}| + \eta_A |\alpha_{\mathcal{A}}\rangle \langle \alpha_{\mathcal{A}}| A_s \right). \quad (38)$$

We then partition the Hilbert space,  $\mathcal{H}_{\mathcal{A}}$ , of the SPP into a direct sum of  $(16 \equiv k)$  4-dimensional subspaces  $\mathcal{H}_{\mathcal{A}} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \dots \oplus \mathcal{H}_k$ . Each  $\mathcal{H}_i$  is taken to be of the form

$$\text{span} \left\{ |\alpha_{\mathcal{A}}\rangle, A_s^{(1)} |\alpha_{\mathcal{A}}\rangle, A_s^{(2)} |\alpha_{\mathcal{A}}\rangle, A_s |\alpha_{\mathcal{A}}\rangle \right\}, \quad (39)$$

for some state  $\alpha$ . Inspection of (38) shows that one can also write the density matrix as a direct sum

$$\rho_{\mathcal{A}} = \frac{1}{k} \bigoplus_{i=1}^k \varrho, \quad (40)$$

where we have introduced the density matrix

$$\varrho = k \begin{pmatrix} \Xi^+ & 0 & 0 & \eta_A \Xi^+ \\ 0 & \Xi^- & \eta_A \Xi^- & 0 \\ 0 & \eta_A \Xi^- & \Xi^- & 0 \\ \eta_A \Xi^+ & 0 & 0 & \Xi^+ \end{pmatrix}. \quad (41)$$

By partitioning  $\mathcal{H}_{\mathcal{A}}$  according to (39), partial transposition acts on each  $\varrho$  separately:

$$\rho_{\mathcal{A}}^{T_2} = \frac{1}{k} \bigoplus_{i=1}^k \varrho^{T_2}, \quad (42)$$

and consequently  $\mathcal{E}(\rho_{\mathcal{A}}) = \mathcal{E}(\varrho)$ . We are now able to make use of the result that the negativity  $\mathcal{E}$  is both a necessary and sufficient condition for separability when  $\dim \mathcal{H} = 2 \times 2$  [49]. If  $\mathcal{E}(\varrho) = 0$ , then the state  $\varrho$  is separable and, by definition, can be written  $\varrho = \sum_j p_j (\varrho_j^{\mathcal{A}_1} \otimes \varrho_j^{\mathcal{A}_2})$  with  $\sum_j p_j = 1$ . Substituting into (40), after some manipulation, one arrives at

$$\begin{aligned} \rho_{\mathcal{A}} &= \sum_j p_j \left( \frac{1}{\sqrt{k}} \bigoplus_{i=1}^{\sqrt{k}} \varrho_j^{\mathcal{A}_1} \right) \otimes \left( \frac{1}{\sqrt{k}} \bigoplus_{i=1}^{\sqrt{k}} \varrho_j^{\mathcal{A}_2} \right) \\ &\equiv \sum_j p_j \rho_j^{\mathcal{A}_1} \otimes \rho_j^{\mathcal{A}_2}. \end{aligned} \quad (43)$$

That is, the state  $\rho_{\mathcal{A}}$  is separable if its negativity  $\mathcal{E}(\rho_{\mathcal{A}})$  vanishes. We remark that this is a consequence of the very special form of the reduced density matrix described by Eq. (40).

## B. Extended boundary

Let us now consider the more general case where the  $\mathcal{A}_1$ - $\mathcal{A}_2$  boundary spans the entire system, as illustrated schematically in Fig. 6b, leading to the result (11) discussed in Sec. II B. Both  $\mathcal{A}_1$  and  $\mathcal{A}_2$  span the torus in one direction and share one edge. The subsystem  $\mathcal{A} \equiv \mathcal{A}_1 \cup \mathcal{A}_2$  consists only of boundary stars ( $A_{s \in \partial}$ ) and boundary plaquettes ( $B_{p \in \partial}$ ) which act simultaneously (and exclusively) on both the  $\mathcal{A}_1$  and  $\mathcal{A}_2$  subsystems. In addition,  $\mathcal{A}$  is symmetric under interchange of stars and plaquettes so, as for the SPP system, the final result must obey the symmetry  $\mathcal{E}(\lambda_A, \lambda_B) = \mathcal{E}(\lambda_B, \lambda_A)$ .

The calculation starts from the most general expression (31) for  $\text{Tr} (\rho_{\mathcal{A}}^{T_2})^n$ . Since  $\mathcal{A}$  contains only boundary stars, the groups  $\mathcal{G}_{\mathcal{A}_1}$  and  $\mathcal{G}_{\mathcal{A}_2}$  become trivial ( $\{\mathbb{I}\}$ ), removing the summations



over  $\bar{g}_\ell$  and  $\bar{g}_\ell$ . In addition, the magnetisation can be decomposed into a boundary component  $M_\partial$ , equal to the sum of the magnetisations of all boundary plaquettes straddling  $\mathcal{A}_1$  and  $\mathcal{A}_2$ , and its complement:  $M(\alpha) = M_\partial(\alpha_{\mathcal{A}}) + M_{\bar{\partial}}(\alpha)$ . Note that the boundary magnetisation  $M_\partial$  depends only on spins contained within subsystem  $\mathcal{A}$  while  $M_{\bar{\partial}}$  depends on all spins. Similarly to the SPP, tracing over subsystem  $\mathcal{B}$  is then entirely contained within a factor  $\xi_T(\alpha_{\mathcal{A}}) \equiv \frac{1}{Z_B} \sum_{\alpha_B} e^{\beta \lambda_B M_\partial(\alpha)}$ , whose explicit calculation is left until later. After making these simplifications, (31) becomes

$$\text{Tr}(\rho_{\mathcal{A}}^{T_2})^n = \sum_{\alpha_{1,\mathcal{A}}, \dots, \alpha_{n,\mathcal{A}}} \sum_{\theta_1, \dots, \theta_n} \langle 0 | \prod_{\ell=1}^n \theta_\ell | 0 \rangle e^{\beta \lambda_B \sum_{\ell} M_\partial(\alpha_{\ell,\mathcal{A}})}$$

$$\prod_{\ell=1}^n \xi_T(\alpha_{\ell,\mathcal{A}}) \eta_T(\theta_\ell) \langle \alpha_{\ell,\mathcal{A}} | \theta_{\ell,\mathcal{A}_1} \otimes \theta_{(\ell+1),\mathcal{A}_2} | \alpha_{(\ell+1),\mathcal{A}} \rangle. \quad (44)$$

To evaluate Eq. (44) to an explicit function of  $n$ , we proceed by introducing classical variables  $\sigma_s^{(\ell)}$  that take value 1 (0) if the star  $A_s$  is present (not present) in the decomposition of the boundary group element  $\theta_\ell$ . We can then write  $\theta_\ell = \prod_s A_s^{\sigma_s^{(\ell)}}$  and

$$\langle 0 | \prod_{\ell=1}^n \theta_\ell | 0 \rangle = \langle 0 | \prod_{s \in \partial} A_s^{\sum_{\ell} \sigma_s^{(\ell)}} | 0 \rangle$$

$$= \prod_{s \in \partial} \delta \left( \sum_{\ell=1}^n \sigma_s^{(\ell)} \text{ mod } 2 \right). \quad (45)$$

The second equality holds because the matrix element equals unity iff an even number of  $A_s$  act on each star. Although this form of the constraint is entirely valid, it is more convenient to represent the Kronecker delta as

$$\delta(y \text{ mod } 2) = \frac{1}{2} \sum_{k=0}^1 e^{i\pi k y}. \quad (46)$$

When we perform the product over boundary stars, the benefit of the alternative representation becomes evident:

$$\langle 0 | \prod_{\ell} \theta_\ell | 0 \rangle = \frac{1}{2^{N_\partial}} \sum_{\{k_s\}} e^{i\pi \sum_{\ell} \sum_s k_s \sigma_s^{(\ell)}}. \quad (47)$$

At the expense of introducing a further  $N_\partial$  variables  $\{k_s\}$ , we have been able to write the constraint (45) in a form that is separable into a product over the different replicas. The number of stars present in the decomposition of  $\theta_\ell$  has a particularly simple representation in terms of  $\sigma_s^{(\ell)}$  variables,  $n(\theta_\ell) = \sum_s \sigma_s^{(\ell)}$ , which allows us to write  $\eta_T(\theta_\ell)$  explicitly as

$$\eta_T(\theta_\ell) = \frac{1}{2 \cosh\left(\frac{NK_A}{2}\right)} \sum_{J_\ell=1} e^{J_\ell K_A (N/2 - \sum_s \sigma_s^{(\ell)})}, \quad (48)$$

where we recall that  $K_A = -\ln \tanh \beta \lambda_A$ .

To evaluate the  $\xi_T$  factor, it is more convenient to work with configurations of the plaquette magnetisations  $M_p = \langle B_p \rangle_\alpha \in$

$\{-1, 1\}$  rather than the physical spin configurations  $\alpha$ . This is possible because  $\xi_T(\alpha_{\mathcal{A}}) = \xi_T(\alpha'_{\mathcal{A}})$  if all boundary plaquette magnetisations  $\{M_{p \in \partial}\}$  of  $\alpha_{\mathcal{A}}$  and  $\alpha'_{\mathcal{A}}$  are the same, which may be shown as follows. If  $M_p(\alpha_{\mathcal{A}}) = M_p(\alpha'_{\mathcal{A}}), \forall p \in \partial$ , then we are able to write

$$\alpha'_{\mathcal{A}} = (\Gamma_{\mathcal{A}} \cdot g_{\mathcal{A}}) \alpha_{\mathcal{A}}, \quad (49)$$

with  $\Gamma = \Gamma_{\mathcal{A}} \otimes \Gamma_{\mathcal{B}} \in \{\mathbb{1}, \Gamma_v^x\}$  and  $g = g_{\mathcal{A}} \otimes g_{\mathcal{B}} \in \mathcal{G}$ . That is, the two spin configurations  $\alpha_{\mathcal{A}}, \alpha'_{\mathcal{A}}$  differ at most by the action of a group element  $g$  and the vertical winding operator  $\Gamma_v^x$  (see Sec. IB for the definition of winding operators). This implies

$$\xi_T(\alpha'_{\mathcal{A}}) = \frac{1}{Z_B} \sum_{\alpha'_B} e^{\beta \lambda_B M_\partial[(\Gamma_{\mathcal{A}} \cdot g_{\mathcal{A}}) \alpha_{\mathcal{A}} \otimes \alpha'_B]}. \quad (50)$$

We can then introduce new summation variables  $\alpha_B = \Gamma_B g_B \alpha'_B$ , giving

$$\xi_T(\alpha'_{\mathcal{A}}) = \frac{1}{Z_B} \sum_{\alpha_B} e^{\beta \lambda_B M_\partial[(\Gamma \cdot g) \alpha]}. \quad (51)$$

Since  $M_\partial[(\Gamma \cdot g) \alpha] = M_\partial(\alpha)$ , we have shown the desired result.

We can treat the magnetisations  $M_p$  as Ising spins that live at the centers of the plaquettes  $p$ . They are subject to the condition that an even number of Ising spins are negative in any given configuration, since they must satisfy  $\prod_p B_p = \mathbb{1}$ . If an even (odd) number of boundary plaquette magnetisations  $M_{p \in \partial}$  are negative, there must be a compensating even (odd) number of negative plaquette magnetisations  $M_{p \notin \partial}$  in the bulk, introducing an even–odd effect.

To evaluate  $\xi_T$  we need to sum over bulk magnetisation configurations  $\{M_{p \notin \partial}\}$ , keeping the boundary configuration  $\{M_{p \in \partial}\}$  fixed. This results in

$$\xi_T(M_\partial) = \frac{1}{2^{3N_\partial}} \frac{\cosh(\beta \lambda_B)^{N-N_\partial} \pm \sinh(\beta \lambda_B)^{N-N_\partial}}{\cosh(\beta \lambda_B)^N + \sinh(\beta \lambda_B)^N}. \quad (52)$$

The upper (lower) sign corresponds to an even (odd) number of negative  $M_{p \in \partial}$ . The factor  $1/2^{3N_\partial}$  comes from two contributions. Firstly, one must account for the fact that there are  $2^{\mathcal{N}_{\mathcal{A}} - N_\partial}$  physical spin configurations per magnetisation configuration of the  $\mathcal{A}$  subsystem, where  $\mathcal{N}_{\mathcal{A}} = 3N_\partial$  is the total number of physical spins contained within  $\mathcal{A}$ , which leads to a prefactor of  $1/2^{2N_\partial}$ . Secondly, a factor of  $1/2^{N_\partial}$  comes from the conversion of exponentials to sinh and cosh.

Since each boundary star is adjacent to two boundary plaquettes, the parity of the number of negative  $M_{p \in \partial}$  is conserved between replicas so that  $\prod_{\ell} \xi_T(M_\partial(\alpha_\ell)) = \xi_T^n(M_\partial(\alpha_1))$ .

Finally, we need to calculate the boundary plaquette magnetisation  $M_\partial = \sum_{p \in \partial} M_p$  for each of the replicas. Given  $\alpha_{1,\mathcal{A}}$ , say, then all  $\alpha_{\ell,\mathcal{A}}$  for  $\ell > 1$  are uniquely determined by (44). Introducing Ising spins  $\tau_s^{(\ell)} = 2\sigma_s^{(\ell)} - 1 \in \{-1, 1\}$ , the recursion relation between boundary magnetisations  $M_s, s = p \in \partial$ , of adjacent replicas is  $M_s(\alpha_\ell) = M_s(\alpha_{\ell-1}) \tau_s^{(\ell)} \tau_{s+1}^{(\ell-1)} \tau_{s+1}^{(\ell-1)}$ . Alternatively, in terms of  $M_s(\alpha_1)$ ,

$$M_s(\alpha_\ell) = M_s(\alpha_1) \tau_s^{(1)} \tau_{s+1}^{(1)} \tau_{s+1}^{(\ell)} \tau_{s+1}^{(\ell-1)}. \quad (53)$$

The notation  $M_s$  is interpreted as the magnetisation of the plaquette with the same index as the star  $s$ , as depicted in

$$\text{Tr} (\rho_{\mathcal{A}}^{T_2})^n = \frac{1}{2^{N_\partial}} \frac{1}{\left[2 \cosh\left(\frac{NK_A}{2}\right)\right]^n} \sum_{\alpha_{\mathcal{A}}} \xi_T^n(M_\partial(\alpha_{\mathcal{A}})) \sum_{\{k_s\}_{s=1}^{N_\partial}} \sum_{\{J_\ell\}_{\ell=1}^n} \prod_{\ell=1}^n \sum_{\{\sigma_s^{(\ell)}\}_{s=1}^{N_\partial}} e^{-\sum_{s=1}^{N_\partial} [\gamma_s \tau_s^{(\ell)} \tau_{s+1}^{(\ell)} + h_s^{(\ell)} \sigma_s^{(\ell)}] + \frac{1}{2} N J_\ell K_A}, \quad (54)$$

where  $h_s^{(\ell)} \equiv J_\ell K_A + i\pi k_s$  and  $\gamma_s \equiv -\beta \lambda_B M_s$ .

Notice that we have performed a relabelling of the physical spin configurations  $\alpha_{\mathcal{A}}$ , for each configuration  $\{\tau_s^{(1)}\}$ , such that  $M_s \tau_s^{(1)} \tau_{s+1}^{(1)} \rightarrow M_s$  [the parity is conserved by this transformation since  $\prod_s M_s \tau_s^{(1)} \tau_{s+1}^{(1)} = \prod_s M_s$ , thus leaving  $\xi_T^n(M_\partial)$  unchanged]. This puts all replicas on an equal footing, making the trace separable into a product over replicas. The problem has therefore been reduced to evaluating the classical partition function

$$Z(\{\gamma_s\}, \{k_s\}, J_\ell K_A) = \sum_{\{\tau_s = \pm 1\}} e^{-\sum_{s=1}^{N_\partial} [\gamma_s \tau_s \tau_{s+1} + \frac{1}{2} h_s^{(\ell)} (\tau_s + 1)]} \quad (55)$$

of a 1D Ising model with reduced couplings  $\{\gamma_s\}$  in a complex reduced magnetic field  $\{h_s\}$ . Using the expression

$$\lim_{n \rightarrow 1} \sum_{m=0}^n \binom{n}{m} x^m y^{n-m} = |x + y|, \quad (56)$$

having followed the even  $n$  series, one can then evaluate the sum over configurations  $\{J_\ell\}$

$$\sum_{\{J_\ell\}_{\ell=1}^n} \prod_{\ell=1}^n e^{\frac{1}{2} N J_\ell K_A} Z(J_\ell K_A) = \left| e^{\frac{1}{2} N K_A} Z(K_A) + e^{-\frac{1}{2} N K_A} Z(-K_A) \right|, \quad (57)$$

where we have suppressed the dependence of  $Z$  on  $\{\gamma_s\}$  and  $\{k_s\}$  to simplify the notation. The final result follows by

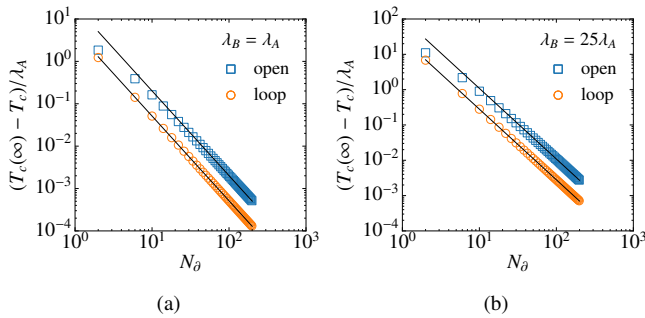


FIG. 7. Convergence of the sudden death temperature  $T_c$  to its limiting value  $T_c(\infty) \equiv T_\infty$  for large boundaries follows a power law  $T_\infty - T_c \sim N_\partial^{-2}$ . Both contractible (“open”) and system-spanning (“loop”) subsystems tend towards the same limiting value.

Fig. 6b. The description in terms of classical spin variables is now complete and we are in a position to substitute back into (44). Relabelling  $\alpha_1 \rightarrow \alpha$ , we obtain

substituting back into (54):

$$\|\rho_{\mathcal{A}}^{T_2}\|_1 = \frac{1}{2^{N_\partial}} \frac{1}{2 \cosh\left(\frac{NK_A}{2}\right)} \sum_{\alpha_{\mathcal{A}}} \xi_T(M_\partial(\alpha_{\mathcal{A}})) \sum_{\{k_s\}_{s=1}^{N_\partial}} \left| e^{\frac{1}{2} N K_A} Z(K_A) + e^{-\frac{1}{2} N K_A} Z(-K_A) \right|. \quad (58)$$

Taking the thermodynamic limit  $N \rightarrow \infty$ , the partition function  $Z(-K_A)$  is suppressed by a factor  $e^{-NK_A}$ , and the even-odd effect is removed from  $\xi_T \rightarrow \text{sech}(\beta \lambda_B)^{N_\partial} / 2^{3N_\partial}$ . Absorbing  $\xi_T$  into the partition function as an energy shift and trading the sum over spin configurations  $\alpha_{\mathcal{A}}$  for one over magnetisations  $\{M_s\}$ , we obtain

$$\|\rho_{\mathcal{A}}^{T_2}\|_1 = \frac{1}{2^{2N_\partial}} \sum_{\{M_s\}_{s=1}^{N_\partial}} \sum_{\{k_s\}_{s=1}^{N_\partial}} |Z'(\{\gamma_s\}, \{k_s\}, T)|, \quad (59)$$

where we introduced the modified partition function

$$Z' = \sum_{\{\tau_s = \pm 1\}} e^{-\sum_{s=1}^{N_\partial} [\gamma_s \tau_s \tau_{s+1} + \frac{1}{2} h_s (\tau_s + 1) + \ln \cosh \beta \lambda_B]}. \quad (60)$$

The sums in (59) can be regarded as performing a disorder average over the couplings and the local magnetic fields, leading to the expression  $\mathcal{E} = \ln \langle |Z'| \rangle$  given in Sec. II B.

In the limit  $T \rightarrow 0^+$ , one can verify after a few lines of algebra that half of the magnetisation configurations (those satisfying  $\prod_s M_s = 1$ ), and half of the magnetic field configurations (those satisfying  $\sum_s k_s \bmod 2 = 0$ ), give a nonvanishing contribution of  $2^{N_\partial+1}$ . All other configurations  $\{M_s\}$  and  $\{k_s\}$  give instead a vanishing contribution. Hence, one arrives at the stated zero-temperature result  $\mathcal{E}(0) = (N_\partial - 1) \ln 2$ , in agreement with Refs. 14 and 15.

If the boundary does not span the entire system, the calculation must be adjusted to account for the open ends of the boundary. This results in a slightly more involved calculation, the details of which we hint at in Appendix C. One arrives at an expression which displays qualitatively similar behaviour to the system-spanning strip. The agreement becomes quantitative in the limit of large boundaries, where both system-spanning and contractible boundaries give the same sudden death temperature  $T_\infty$ , as evidenced by Fig. 7. As one would expect, the limiting case  $N_\partial = 1$  reproduces exactly the SPP expression of Sec. III A.

### C. $\lambda_B \rightarrow \infty$ limit

We now consider a limiting case that allows for a substantial simplification of the calculation of the negativity, and tells us something interesting about its behaviour at finite temperature in the presence of hard constraints. Namely, we take the limit  $\lambda_B \rightarrow \infty$  that removes magnetic defects from the system at all temperatures. We also make the following working assumptions about the partition of the system: (i)  $\mathcal{A}$  is contractible, i.e., not system-spanning in either direction, and (ii) for each boundary star there corresponds at least one boundary plaquette.

If these assumptions are satisfied, the matrix elements in (31) imply that  $\theta_\ell = \theta_{\ell+1} \equiv \theta, \forall \ell$ . This is because only terms for which  $M(\alpha_\ell) = N, \forall \ell$ , contribute (due to  $Z_B^n$  diverging as  $e^{nN\beta\lambda_B}$  for  $\beta\lambda_B \gg 1$ ) and we then must have  $\theta_{\ell\mathcal{A}_1} \otimes \theta_{(\ell+1)\mathcal{A}_2} \in \mathcal{G}_{\mathcal{A}_1\mathcal{A}_2}$  to satisfy  $M(\alpha_{\ell+1}) = M(\alpha_\ell)$ .

Given  $\alpha_1 \equiv \alpha$ , say, then each  $\alpha_{\ell\mathcal{A}}$  for  $\ell > 1$  is uniquely determined by the matrix elements in (31):  $|\alpha_{(2m+1)\mathcal{A}}\rangle = |\alpha_{\mathcal{A}}\rangle$  and  $|\alpha_{2m\mathcal{A}}\rangle = \theta |\alpha_{\mathcal{A}}\rangle$  for integer  $m$ . Since  $n$  is assumed even, the constraint  $\prod_\ell \theta_\ell = \theta^n = \mathbb{I}$  is automatically satisfied. On the other hand, the  $\alpha_{\ell\mathcal{B}}$  for  $\ell > 1$  are only subject to the condition that their plaquette magnetisation remains undegraded from  $M(\alpha_\ell) = N$ , and therefore each  $\alpha_\ell$  summation,  $\ell > 1$ , produces a factor of  $4|\mathcal{G}_{\mathcal{B}}|$ . The remaining summation over  $\alpha$ , in the limit  $\lambda_B \rightarrow \infty$ , gives instead

$$\frac{1}{Z_B^n} \sum_\alpha e^{n\beta\lambda_B M(\alpha)} \sim \frac{\sum'_\alpha e^{nN\beta\lambda_B}}{(\sum'_\alpha e^{N\beta\lambda_B})^n} \rightarrow \frac{1}{[4|\mathcal{G}|]^{n-1}}, \quad (61)$$

where the prime on the summation denotes that the constraint  $M(\alpha) = N$  is imposed on the spin configurations  $\alpha$ . We therefore arrive at the expression

$$\begin{aligned} \text{Tr}(\rho_{\mathcal{A}}^{T_2})^n &= \frac{|\mathcal{G}_{\mathcal{B}}|^{n-1}}{|\mathcal{G}|^{n-1}} \sum_{\substack{\bar{g}_1, \dots, \bar{g}_n \\ \in \mathcal{G}_{\mathcal{A}_1}}} \sum_{\substack{\bar{g}_1, \dots, \bar{g}_n \\ \in \mathcal{G}_{\mathcal{A}_2}}} \sum_{\theta \in \mathcal{G}_{\mathcal{A}_1\mathcal{A}_2}} \times \\ &\langle 0 | \prod_{\ell=1}^n \bar{g}_\ell \bar{g}_\ell | 0 \rangle \left( \prod_{\ell=1}^n \eta_T(\bar{g}_\ell \bar{g}_\ell \theta) \right). \quad (62) \end{aligned}$$

Before proceeding further, we evaluate the zero-temperature limit and compare with earlier results in Ref. 15. When  $T \rightarrow 0^+$ ,  $\eta_T(g) \rightarrow 1 \forall g \in \mathcal{G}$ , and one simply needs to count the number of terms in the summation, leading to

$$\text{Tr}(\rho_{\mathcal{A}}^{T_2})^n = |\mathcal{G}_{\mathcal{A}_1\mathcal{A}_2}| \left[ \frac{|\mathcal{G}_{\mathcal{A}_1}| |\mathcal{G}_{\mathcal{A}_2}| |\mathcal{G}_{\mathcal{B}}|}{|\mathcal{G}|} \right]^{n-1}. \quad (63)$$

One must remember that we arrived at this expression following the even  $n$  series. Taking  $n \rightarrow 1$  and decomposing the number

of stars as  $N_{\mathcal{A}} = N_{\mathcal{A}_1} + N_{\mathcal{A}_2} + N_\partial$ , we arrive at  $\mathcal{E}(0) = N_\partial \ln 2$ , in agreement with Ref. 15. Here  $N_{\mathcal{A}_i}$  corresponds to the number of stars that act solely on subsystem  $\mathcal{A}_i$  while  $N_\partial$  equals the number of stars that straddle the  $\mathcal{A}_1$ - $\mathcal{A}_2$  boundary. Had we followed the odd series, the constraint  $\prod_\ell \theta_\ell = \theta = \mathbb{I}$  removes the summation over  $\theta$  and we would obtain instead  $\lim_{n \rightarrow 1} \text{Tr}(\rho_{\mathcal{A}}^{T_2})^n = 1$ , as required for the trace of a density matrix.

Taking  $n \rightarrow 1$  for finite  $T$  is more difficult. In order to carry out the calculation, in a similar manner to Sec. III B, we need to introduce  $2n+1$  sets of Ising spins  $\{\vartheta_s^{(\ell)}\}$ ,  $\{\varphi_s^{(\ell)}\}$  and  $\{\psi_s\}$ . The spins are defined as follows:  $\vartheta_s^{(\ell)} = -1(1)$  if  $A_s$  appears (does not appear) in the decomposition of  $\bar{g}_\ell \in \mathcal{G}_{\mathcal{A}_1}$ , and similarly for  $\varphi_s^{(\ell)}$  and  $\psi_s$  with respect to  $\bar{g}_\ell \in \mathcal{G}_{\mathcal{A}_2}$  and  $\theta \in \mathcal{G}_{\mathcal{A}_1\mathcal{A}_2}$ . For convenience, we define

$$\Delta_\psi \equiv \frac{N - N_{\mathcal{A}}}{N_\partial} > 1, \quad (64)$$

and introduce the notation  $\tilde{\psi}_s = \psi_s + \Delta_\psi$ , so that in terms of the spin variables, the number of stars flipped by the action of the group element  $\bar{g}_\ell \bar{g}_\ell \theta$  can be written as

$$n(\bar{g}_\ell \bar{g}_\ell \theta) = \frac{N}{2} - \frac{1}{2} \left( \sum_{s \in \mathcal{A}_1} \vartheta_s^{(\ell)} + \sum_{s \in \mathcal{A}_2} \varphi_s^{(\ell)} + \sum_{s \in \partial} \tilde{\psi}_s \right). \quad (65)$$

The constraint imposed by the matrix element  $\langle 0 | \prod_\ell \bar{g}_\ell \bar{g}_\ell | 0 \rangle \neq 0$ , which is satisfied iff an even number of star operators acts on each site, becomes

$$\prod_{s \in \mathcal{A}_1} \delta \left( \prod_{\ell=1}^n \vartheta_s^{(\ell)} - 1 \right) \prod_{s \in \mathcal{A}_2} \delta \left( \prod_{\ell=1}^n \varphi_s^{(\ell)} - 1 \right). \quad (66)$$

Denoting a configuration of spins, for a given  $\ell$ , by  $\chi_\ell = \{\vartheta_s^{(\ell)}, \varphi_s^{(\ell)}, \psi_s\}$  we can complete the transformation to the Ising spin language by writing

$$\begin{aligned} 2 \cosh \left( \frac{K_A}{2} N \right) \sum_{\{\vartheta_s^{(\ell)}, \varphi_s^{(\ell)}\}} \eta_T(\chi_\ell) &= \\ \sum_{J_\ell = \pm 1} \left( \prod_{s \in \mathcal{A}_1} \sum_{\vartheta_s^{(\ell)} = \pm 1} e^{(K_A/2) J_\ell \vartheta_s^{(\ell)}} \right) &\left( \prod_{s \in \mathcal{A}_2} \sum_{\varphi_s^{(\ell)} = \pm 1} e^{(K_A/2) J_\ell \varphi_s^{(\ell)}} \right) \\ \left( \prod_{s \in \partial} e^{(K_A/2) J_\ell \tilde{\psi}_s} \right). \quad (67) \end{aligned}$$

Collecting together results from (65) to (67), one obtains an explicit expression for the trace in terms of the Ising spin variables

$$\text{Tr} (\rho_{\mathcal{A}}^{T_2})^n = \frac{|\mathcal{G}_{\mathcal{B}}|^{n-1}}{|\mathcal{G}|^{n-1}} \frac{1}{\left[2 \cosh\left(\frac{K_A}{2} N\right)\right]^n} \times \sum_{\{J_\ell\}_{\ell=1}^n} \sum_{\{\psi_s\}_{s=1}^{N_\partial}} \left( \prod_{s \in \mathcal{A}_1} \sum_{\substack{\{\vartheta_s^{(\ell)}\}_{\ell=1}^n \\ \prod_\ell \vartheta_s^{(\ell)} = +1}} e^{(K_A/2) \sum_\ell J_\ell \vartheta_s^{(\ell)}} \right) \left( \prod_{s \in \mathcal{A}_2} \sum_{\substack{\{\varphi_s^{(\ell)}\}_{\ell=1}^n \\ \prod_\ell \varphi_s^{(\ell)} = +1}} e^{(K_A/2) \sum_\ell J_\ell \varphi_s^{(\ell)}} \right) \left( \prod_{s \in \partial} e^{(K_A/2) \sum_\ell J_\ell \tilde{\psi}_s} \right). \quad (68)$$

The explicit constraint has now been replaced by conditions on the relevant summations and we have exchanged the order of the product over  $\ell$  with the summations over  $J_\ell$  using  $\prod_\ell (\sum_{J_\ell = \pm 1} F(J_\ell)) = \sum_{\{J_\ell\}} \prod_\ell F(J_\ell)$ , for arbitrary  $F$ .

The sums over  $\{\vartheta_s^{(\ell)}\}$  and  $\{\varphi_s^{(\ell)}\}$  can be carried out explicitly by exploiting their analogy to the partition function of a 1D Ising model as follows. First, note that the site labels  $s$  of  $\vartheta_s^{(\ell)}$  and  $\varphi_s^{(\ell)}$  are mute since the contribution from every star  $s \in \mathcal{A}_1$  is identical, and similarly for every star  $s \in \mathcal{A}_2$ . This allows us to substantially simplify the notation  $\vartheta_s^{(\ell)} \rightarrow \vartheta^{(\ell)}$  and  $\varphi_s^{(\ell)} \rightarrow \varphi^{(\ell)}$ . Then we make use of the one-to-two mapping  $\vartheta^{(\ell)} = \tau_\ell \tau_{\ell+1}$ , and similarly for the  $\varphi^{(\ell)}$  spins. The motivation behind this mapping is that the constraint  $\prod_\ell \vartheta^{(\ell)} = +1$  is automatically satisfied if we impose periodic boundary conditions,  $\tau_{n+1} = \tau_1$ , since  $\tau_\ell^2 = 1$ . A factor of  $1/2$  is required to compensate for the double counting of spin configurations in the  $\tau$  spin notation,

$$\begin{aligned} \sum_{\substack{\{\vartheta^{(\ell)}\}_{\ell=1}^n \\ \prod_\ell \vartheta^{(\ell)} = +1}} e^{(K_A/2) \sum_\ell J_\ell \vartheta^{(\ell)}} &= \frac{1}{2} \sum_{\{\tau_\ell\}_{\ell=1}^n} e^{(K_A/2) \sum_\ell J_\ell \tau_\ell \tau_{\ell+1}} \\ &\equiv \frac{1}{2} Z_n^p(K_A, \{J_\ell\}). \end{aligned} \quad (69)$$

This is nothing but the partition function of a 1D Ising chain of length  $n$ , with position-dependent reduced couplings  $K_A J_\ell / 2$ , which we denote for convenience as  $Z_n^p$  (the  $p$  reminding us of the periodic boundary conditions). A straightforward manipulation (see for instance Ref. 37) gives us

$$Z_n^p(K_A, \{J_\ell\}) = \left[ 2 \cosh\left(\frac{K_A}{2}\right) \right]^n + \left( \prod_{\ell=1}^n J_\ell \right) \left[ 2 \sinh\left(\frac{K_A}{2}\right) \right]^n. \quad (70)$$

Substituting back into (68), we are hence left with the simpler expression

$$\text{Tr} (\rho_{\mathcal{A}}^{T_2})^n = \frac{|\mathcal{G}_{\mathcal{B}}|^{n-1}}{|\mathcal{G}|^{n-1}} \frac{1}{\left[2 \cosh\left(\frac{K_A}{2} N\right)\right]^n} \sum_{\{\psi_s\}_{s=1}^{N_\partial}} \sum_{\{J_\ell\}_{\ell=1}^n} \left[ \frac{1}{2} Z_n^p \right]^{N_{\mathcal{A}_1} + N_{\mathcal{A}_2}} e^{(K_A/2) \sum_\ell J_\ell \sum_s \tilde{\psi}_s}. \quad (71)$$

To obtain an explicit function of  $n$ , as is required to take the limit  $n \rightarrow 1$ , we must perform the summation over configurations  $\{J_\ell\}$ . Thankfully, the summand in (71) only depends on  $J_\ell$

through  $\prod_\ell J_\ell$  and  $\sum_\ell J_\ell$ . We can take advantage of this by trading the sum over configurations  $\{J_\ell\}_{\ell=1}^n$  for one over the number of negative  $J_\ell$ , denoted by  $m$ , with the appropriate combinatoric weighting  $\binom{n}{m}$ . The terms on the second line in (71) become

$$\begin{aligned} \sum_{\substack{m=0 \\ \text{even}}}^n \binom{n}{m} \left[ \frac{1}{2} Z_n^+ \right]^{N_{\mathcal{A}} + N_{\mathcal{B}}} e^{(K_A/2) \sum_s \tilde{\psi}_s (n-2m)} + \\ \sum_{\substack{m=0 \\ \text{odd}}}^n \binom{n}{m} \left[ \frac{1}{2} Z_n^- \right]^{N_{\mathcal{A}} + N_{\mathcal{B}}} e^{(K_A/2) \sum_s \tilde{\psi}_s (n-2m)}. \end{aligned} \quad (72)$$

With the help of the binomial generating function  $\sum_{m=0}^n \binom{n}{m} x^m = (1+x)^n$ , evaluated at  $x$  and  $-x$ , we can then evaluate the sums over even and odd  $m$ ,

$$\lim_{n \rightarrow 1} \sum_{\substack{m=0 \\ \text{even}}}^n \binom{n}{m} x^m = \frac{1}{2} (|1+x| + |1-x|) = \max(1, x), \quad (73)$$

$$\lim_{n \rightarrow 1} \sum_{\substack{m=0 \\ \text{odd}}}^n \binom{n}{m} x^m = \frac{1}{2} (|1+x| - |1-x|) = \min(1, x), \quad (74)$$

where we used the fact that  $(1 \pm x)^n \rightarrow |1 \pm x|$ , for  $n \rightarrow 1$  following the even  $n$  series. Note that the Ising partition function has the limit  $\frac{1}{2} Z_{n \rightarrow 1}^\pm \rightarrow e^{\pm(K_A/2)}$ . It is crucial at this stage to remember that  $\tilde{\psi}_s > 0$ , so that (72) evaluates to

$$\sum_{\sigma = \pm 1} e^{\sigma (K_A/2) (N_{\mathcal{A}_1} + N_{\mathcal{A}_2} + \sum_s \tilde{\psi}_s)}. \quad (75)$$

Both the remaining sums over  $\{\psi_s\}$  and over  $\sigma$  reduce to the partition function of noninteracting Ising spins in a magnetic field, leading to the main result of this section

$$e^{\mathcal{E}(T)} = \frac{\cosh\left(\frac{1}{2}(N - N_\partial)K_A\right)}{\cosh\left(\frac{1}{2}NK_A\right)} \left[ 2 \cosh\left(\frac{K_A}{2}\right) \right]^{N_\partial}, \quad (76)$$

as discussed in Sec. II C.

#### IV. EXPERIMENTAL CONSIDERATIONS

Although the negativity remains an invaluable theoretical tool, allowing us to efficiently understand and quantify the

entanglement content of mixed quantum states, it is not a quantity that is directly accessible in experiments. However, it is possible to obtain the negativity experimentally via explicit reconstruction of the system's density matrix using quantum state tomography [50]. Such an approach is hindered by the exponential scaling in the required resources with system size, which prevents tomography from being applied to larger systems. This shortcoming can be addressed in weakly entangled systems, where more efficient tomography schemes with polynomial scaling exist [51, 52]. Alternatively, it is possible to approximate the negativity by measuring moments of the partially transposed density matrix in a scheme recently proposed by Gray *et al.* [53]. This method also exhibits polynomial scaling with system size.

In this manuscript, we have worked exclusively with Kitaev's toric code. Despite being highly artificial, vast interest in the topological properties of this model, particularly with reference to quantum computation, has led to the proposition of various experimental implementations of the Hamiltonian [54–56] and its ground states [57–60]. The latter approach has proved more fruitful, since directly simulating four-body interactions is experimentally challenging. An alternative method, which only requires the simulation of two-body interactions, consists of cold atoms hopping on a honeycomb lattice [61]. In this setup, the toric code Hamiltonian appears as a perturbative limit [62].

## V. CONCLUSIONS

We performed exact calculations of the logarithmic negativity at finite temperature in the 2D toric code. The exact formulae that we have presented in this manuscript have simple and intuitive physical interpretations that provide us with a deeper understanding of finite temperature entanglement and sudden death in two-dimensional lattice systems.

The smallest nontrivial choice of subsystem—a neighbouring star plaquette pair (SPP)—allows us to extract the general behaviour of the negativity at finite temperatures. We establish that, for this subsystem, the negativity is both a necessary and sufficient condition for separability. Below the gap to the least energetically costly elementary excitations (defects), the negativity remains exponentially close to its zero-temperature value. Above this gap, defects are thermally excited with some  $O(1)$  density, and the negativity correspondingly decays. Sudden death of the negativity, and therefore entanglement, occurs at a temperature of order the largest energy scale in the problem, consistent with previous studies in other models, e.g., Refs. 4 and 63. This corresponds to the thermal energy required to excite an  $O(1)$  density of the most energetically costly defect. We have verified that this physical picture works well for other systems, including the harmonic lattice and Ising chain, which are discussed in Appendices B and D, respectively.

Our main results concern the extension of the SPP to a system consisting of a block of stars connected to a block of plaquettes by a boundary of length  $N_\partial > 1$  stars. We find that extending the length of the boundary leads to some important differences. The number of discontinuities in the first derivative of  $\mathcal{E}$ ,

which arise whenever an eigenvalue of the partially transposed reduced density matrix changes sign, increases exponentially with  $N_\partial$ . Sudden death of the negativity is pushed to higher temperatures, tending towards approximately twice the SPP value in the limit of large boundaries. Therefore a higher density of magnetic defects is required to kill entanglement between subsystems that share a thermodynamically large boundary, but the temperature required to create this density is still  $O(\lambda_B)$  (up to logarithmic factors). We have also shown that, for sufficiently large boundaries, the negativity no longer vanishes abruptly at  $T_c$ ; instead it vanishes with zero gradient asymptotically in the thermodynamic limit.

The interpretation of the sudden death temperature  $T_c$  for subsystems larger than the SPP is less clear-cut, as the negativity is only a necessary condition for separability. It represents a lower bound for the temperature above which *all* entanglement vanishes, including any quantum correlations not picked up by the negativity. Any entanglement that may be present in the system above  $T_c$  is, however, not able to be distilled, and as a result is often referred to as *bound entanglement* [64]. We would welcome the introduction of an entanglement monotone that is both necessary *and* sufficient for separability, but expect on physical grounds that our conclusions would remain unchanged up to  $O(1)$  factors.

Finally, it was shown that imposing the hard constraint  $\lambda_B \rightarrow \infty$  on the system leads to nonzero entanglement at all temperatures. The constraint renders some portion of the full Hilbert space thermally inaccessible, i.e., the span of the states with a nonzero number of magnetic defects, thereby locking in the zero-temperature magnetic loop structure. We observe a slow  $1/T$  demise of the negativity up to arbitrarily large temperatures as thermal fluctuations gradually wash out the zero-temperature quantum correlations. Only at infinite temperature, where the density matrix is proportional to the identity operator, is the system completely separable. One can carry out very similar calculations to obtain the negativity of a hard-constrained bipartite system, but the results are no more insightful and were therefore omitted. This provides a potentially physical example of how a hard constraint applied to a system with a macroscopically degenerate projected manifold can result in thermally robust entanglement. This is likely to be the case for instance in other spin liquids and frustrated magnetic systems in general, when projected down to their low-energy states.

## ACKNOWLEDGEMENTS

We are grateful to Pasquale Calabrese for bringing this problem to our attention in the first place. We also thank him, as well as Claudio Chamon and Attila Szabó, for useful discussions. This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) Grants No. EP/G049394/1 and No. EP/M007065/1.

### Appendix A: Defect density

The defect density of each species  $X = A, B$  can be calculated from the expression

$$\varrho_X = \frac{1}{N} \frac{\partial \ln Z_X}{\partial \epsilon}, \quad (\text{A1})$$

where we have introduced  $\epsilon \equiv 2\beta\lambda_X$  and the partition function

$$Z_X = \sum_{\substack{n=0 \\ \text{even}}}^N \binom{N}{n} e^{-n\epsilon}. \quad (\text{A2})$$

The degeneracy of each energy level has been omitted since it cancels in (A1). Evaluating Eq. (A1), we arrive at the full expression

$$\varrho_X = e^{-2\beta\lambda_X} \frac{(1 + e^{-2\beta\lambda_X})^{N-1} - (1 - e^{-2\beta\lambda_X})^{N-1}}{(1 + e^{-2\beta\lambda_X})^N + (1 - e^{-2\beta\lambda_X})^N}, \quad (\text{A3})$$

which, in the thermodynamic limit  $N \rightarrow \infty$ , reduces to

$$\varrho_X \rightarrow \frac{1}{1 + e^{2\beta\lambda_X}} \equiv n_F(2\lambda_X), \quad (\text{A4})$$

i.e., the Fermi-Dirac distribution. The constraint that electric and magnetic defects must be created in pairs has a vanishing effect in the thermodynamic limit—as far as the density is concerned, each star (plaquette) behaves as an independent two-level system with energy gap  $\lambda_A$  ( $\lambda_B$ ).

### Appendix B: Ising model

A brief study of the one-dimensional transverse field Ising model (TFIM) highlights some of the peculiarities associated with the toric code. The Hamiltonian of an Ising chain with  $L$  sites and periodic boundary conditions is

$$H_L = -J \sum_{i=1}^L \sigma_i^z \sigma_{i+1}^z - h \sum_{i=1}^L \sigma_i^x, \quad (\text{B1})$$

where  $J$  ( $h$ ) controls the strength of nearest-neighbour interactions (magnetic field). For the limiting case of two sites, the Hamiltonian (B1) can be written in matrix form

$$H_2 = - \begin{pmatrix} J & h & h & 0 \\ h & -J & 0 & h \\ h & 0 & -J & h \\ 0 & h & h & J \end{pmatrix}, \quad (\text{B2})$$

with respect to the eigenstates of  $\otimes_i \sigma_i^z$ , and has energy eigenvalues  $\pm\epsilon_1 = \pm J$  and  $\pm\epsilon_2 = \pm\sqrt{J^2 + 4h^2}$ . The density matrix  $\rho = e^{-\beta H}/Z$  and its partial transpose (over the second site) can be evaluated explicitly to give the negativity

$$e^{\mathcal{E}(T)} = \frac{2}{Z} \left[ \cosh(\beta\epsilon_2) + \max \left( \cosh(\beta\epsilon_1), \frac{\epsilon_1}{\epsilon_2} \sinh(\beta\epsilon_2) \right) \right], \quad (\text{B3})$$

with the partition function  $Z = 2 \cosh(\beta\epsilon_1) + 2 \cosh(\beta\epsilon_2)$ .

In the limit of zero temperature,  $T \rightarrow 0^+$ , we obtain

$$\mathcal{E}(0) = \ln \left( 1 + \frac{J}{\sqrt{J^2 + 4h^2}} \right), \quad (\text{B4})$$

for nonzero field  $h$ , whereas the negativity vanishes identically if  $h = 0$  from the outset. In contrast with the toric code, the zero-temperature negativity is a function of the Hamiltonian's parameters. The expression (B4) is maximal ( $\mathcal{E} = \ln 2$ ) for infinitesimal field when the ground state coincides with one of the maximally entangled Bell states  $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ . Increasing the field strength makes it increasingly favourable for the spins to align with the field, eventually leading to  $\mathcal{E}(0) \rightarrow 0$  in the limit of  $h/J \rightarrow \infty$  where the ground state,  $|\sigma_i^x = +1\rangle$ , is separable.

Similarly to the toric code, the sudden death temperature is set by the larger of the two energy scales (up to logarithmic factors)

$$T_c \simeq \begin{cases} \frac{2J}{W\left(\frac{8J^2}{e^2 h^2}\right)} & \text{for } J \gg h, \\ \frac{2h}{\ln\left(\frac{4h}{J}\right)} & \text{for } h \gg J. \end{cases} \quad (\text{B5})$$

However, the toric code and TFIM behave very differently at temperatures  $T < T_c$ . If one takes  $J \gg h$  in an attempt to preserve the largest zero-temperature entanglement  $\mathcal{E}(0) \simeq \ln 2$ , one arrives at  $\mathcal{E}(T) \simeq h^2/(JT)$  for  $h^2/J \ll T \ll T_c$ . That is, for any given nonzero temperature, the negativity vanishes in the limit  $J \rightarrow \infty$ . The slow  $1/T$  decay of the negativity that was observed in the toric code is suppressed by a factor of  $h/J$ . This is because the energy gap  $\Delta$  to the first excited state in the TFIM vanishes as  $\Delta \simeq 2h^2/J$ , while it remains finite in the toric code.

This effect is made more significant when the length  $L$  of the chain is increased, since the energy gap becomes exponentially small in  $L$ , i.e.,  $\Delta \sim h(h/J)^{L-1}$ . Taking the thermodynamic limit  $L \rightarrow \infty$  and then temperature to zero  $T \rightarrow 0^+$ , the negativity of the infinite chain vanishes like  $\mathcal{E}(0) \sim J/h$  for  $h \gg J$  and  $\mathcal{E}(0) \sim h^2/J^2$  for  $J \gg h$ . Therefore, in the infinite 1D TFIM, the divergence of the sudden death temperature is accompanied by the vanishing of the zero-temperature negativity. The toric code is special in that its ground state(s), and therefore its zero-temperature negativity, are independent of the system's parameters. It is this feature that allows quantum correlations to survive up to arbitrarily high temperatures when one takes  $\lambda_B \rightarrow \infty$ .

### Appendix C: Extended, contractible boundary

When the boundary is not system spanning, one must change the magnetisation recursion relations from Eq. (53) to

$$M_s(\alpha_\ell) = M_s(\alpha_{\ell-1}) \begin{cases} \tau_s^{(\ell)} \tau_{s+1}^{(\ell)} \tau_s^{(\ell-1)} \tau_{s+1}^{(\ell-1)} & \text{for } s < N_\partial, \\ \tau_s^{(\ell)} \tau_s^{(\ell-1)} & \text{for } s = N_\partial. \end{cases} \quad (\text{C1})$$

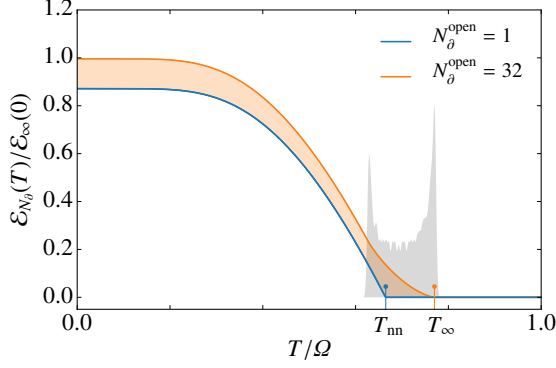


FIG. 8. Negativity for a contractible strip partition (i.e., width two sites) in a harmonic lattice with  $\Omega = \omega$ . The vertical lines correspond to  $T_c$  for a nearest-neighbour pair ( $T_{nn}$ ) and for an infinite boundary ( $T_\infty$ ). The density of discontinuities in gray exhibits bunching at its edges.

These relations embody the fact that boundary plaquettes  $p = s < N_\partial$  are adjacent to two boundary stars, while the plaquette at the end of the boundary,  $p = s = N_\partial$ , is adjacent to only one boundary star. As long as we work in the thermodynamic limit  $N \rightarrow \infty$ , so that the even-odd effect is removed from  $\xi_T$  in Eq. (52), all other aspects of the calculation remain unchanged with respect to the system-spanning case. The effect of the relations (C1) is to change the “Hamiltonian” in  $Z' = \sum_{\{\tau_s\}} e^{-H}$  to

$$H = \gamma_{N_\partial} \tau_{N_\partial} + \sum_{s < N_\partial} \gamma_s \tau_s \tau_{s+1} + \sum_s \left[ \frac{1}{2} h_s (\tau_s + 1) + \ln \cosh \beta \lambda_B \right]. \quad (\text{C2})$$

The SPP result is reobtained for the special case  $N_\partial = 1$  in the above expression. Since the boundary is not system spanning, the zero-temperature negativity is given by  $\mathcal{E}(0) = N_\partial \ln 2$ . The partition function with the largest zero, thereby giving rise to sudden death, corresponds to the disorder realisation  $k_s = 1 \forall s$  and  $\gamma_s = \beta \lambda_B$  for  $s < N_\partial$  and  $M_{N_\partial} = -\beta \lambda_B$ . This makes the equivalent of Eq. (14), which defines  $T_c$  implicitly, more complex and therefore less transparent.

#### Appendix D: Harmonic lattice

In order to test the generality of our results, we compute the negativity in the harmonic lattice numerically. We consider  $N = N_{\mathcal{A}} + N_{\mathcal{B}}$  harmonic oscillators arranged on a two-dimensional square lattice with the Hamiltonian

$$H = \sum_i \frac{1}{2M} p_i^2 + \frac{1}{2} M \omega^2 u_i^2 + \sum_{(i,j)} \frac{1}{2} K (u_i - u_j)^2, \quad (\text{D1})$$

where the momenta  $p_i$  and displacements  $u_i$  satisfy the usual commutation relations  $[u_i, p_j] = i\delta_{ij}$ . The strength of the confining potential at each site (nearest-neighbour coupling) is parameterised by  $\omega$  ( $K$ ), and the mass of each oscillator is  $M$ . The negativity can be computed efficiently using the correlation matrix techniques outlined in Refs. 35, 63, 65, and 66. Given the correlation matrices  $\mathbb{Q}_{ij} = \langle u_i u_j \rangle_\beta$  and  $\mathbb{P}_{ij} = \langle p_i p_j \rangle_\beta$  with  $i, j \in \mathcal{A}$ , the negativity is given in terms of the positive eigenvalues  $\{v_1^2, \dots, v_{N_{\mathcal{A}}}^2\}$  of the matrix  $\mathbb{Q} \cdot \mathbb{P}^{T_2}$ . Partial transposition is effected by reversing the momenta of the oscillators belonging to subsystem  $\mathcal{A}_2$  so that  $\mathbb{P}^{T_2} = \mathbb{R}_{\mathcal{A}_2} \cdot \mathbb{P} \cdot \mathbb{R}_{\mathcal{A}_2}$ , with  $\mathbb{R}_{\mathcal{A}_2}$  the  $N_{\mathcal{A}} \times N_{\mathcal{A}}$  diagonal matrix with entries  $+1$  ( $-1$ ) corresponding to sites in  $\mathcal{A}_1$  ( $\mathcal{A}_2$ ). It can then be shown that (see, e.g., Ref. 17)

$$\mathcal{E} = \sum_{i=1}^{N_{\mathcal{A}}} \ln \max[1, (2v_i)^{-1}]. \quad (\text{D2})$$

At finite temperature, a short calculation gives the elements of the correlation matrices

$$\mathbb{Q}_{ij} = \frac{1}{N} \sum_{\mathbf{k} \in \text{BZ}} \frac{1}{2M\omega_{\mathbf{k}}} e^{-i\mathbf{k} \cdot (\mathbf{x}_i - \mathbf{x}_j)} \coth\left(\frac{1}{2}\beta\omega_{\mathbf{k}}\right), \quad (\text{D3})$$

$$\mathbb{P}_{ij} = \frac{1}{N} \sum_{\mathbf{k} \in \text{BZ}} \frac{M\omega_{\mathbf{k}}}{2} e^{i\mathbf{k} \cdot (\mathbf{x}_i - \mathbf{x}_j)} \coth\left(\frac{1}{2}\beta\omega_{\mathbf{k}}\right), \quad (\text{D4})$$

where the dispersion relation is

$$\omega_{\mathbf{k}}^2 = \omega^2 + 4\Omega^2 \sum_{\alpha=1}^2 \sin^2\left(\frac{1}{2}k_\alpha\right). \quad (\text{D5})$$

We have introduced the natural frequency of the nearest-neighbour interactions  $\Omega \equiv \sqrt{K/M}$ , and all sums over wave vectors  $\mathbf{k}$  are restricted to the first Brillouin zone (BZ).

The two Figs. 8 and 4, for the harmonic lattice and toric code respectively, share many features. Namely,  $\mathcal{E}(T)$  remains undegraded below the gap  $T \ll \omega$  in the dispersion relation, and at higher temperatures exhibits sudden death. As the length of the boundary is increased,  $\mathcal{E}(T)$  develops more discontinuities in its first derivative and sudden death is pushed to higher temperatures. The harmonic lattice differs from the toric code in that the zero-temperature area law has subleading corrections  $\sim N_\partial^{-2}$  and the number of discontinuities scales linearly, as opposed to exponentially, with  $N_\partial$ .

[1] M. B. Plenio and S. Virmani, *Quant. Info. Comput.* **7**, 1 (2007).

[2] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, *Rev. Mod. Phys.*

- 80**, 517 (2008).
- [3] N. Laflorencie, *Phys. Rep.* **646**, 1 (2016).
- [4] N. E. Sherman, T. Devakul, M. B. Hastings, and R. R. P. Singh, *Phys. Rev. E* **93**, 022128 (2016).
- [5] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, *Phys. Rev. A* **58**, 883 (1998).
- [6] G. Vidal and R. F. Werner, *Phys. Rev. A* **65**, 032314 (2002).
- [7] H. Shapourian, K. Shiozaki, and S. Ryu, *Phys. Rev. B* **95**, 165101 (2017).
- [8] J. Eisert, V. Eisler, and Z. Zimborás, (2016), arXiv:1611.08007 [quant-ph].
- [9] A. Peres, *Phys. Rev. Lett.* **77**, 1413 (1996).
- [10] J. Eisert, *Entanglement in quantum information theory*, Ph.D. thesis, University of Potsdam (2001), arXiv:quant-ph/0610253v1.
- [11] H. Wichterich, J. Vidal, and S. Bose, *Phys. Rev. A* **81**, 032311 (2010).
- [12] R. A. Santos, V. Korepin, and S. Bose, *Phys. Rev. A* **84**, 062307 (2011).
- [13] R. A. Santos and V. Korepin, *Quantum Inf. Process.* **15**, 4581 (2016).
- [14] C. Castelnovo, *Phys. Rev. A* **88**, 042319 (2013).
- [15] Y. A. Lee and G. Vidal, *Phys. Rev. A* **88**, 042318 (2013).
- [16] P. Calabrese, J. Cardy, and E. Tonni, *Phys. Rev. Lett.* **109**, 130502 (2012).
- [17] P. Calabrese, J. Cardy, and E. Tonni, *J. Stat. Mech.* **2013**, P02008 (2013).
- [18] In fact, these studies show that entanglement is *not* always a monotonic function of temperature. Particularly in the presence of a large magnetic field, it is possible for entanglement, as measured by the concurrence, to initially *increase* with increasing temperature before eventually vanishing at  $T_c$ , see e.g. Refs. 21 and 30.
- [19] D. Gunlycke, V. M. Kendon, V. Vedral, and S. Bose, *Phys. Rev. A* **64**, 042302 (2001).
- [20] T. J. Osborne and M. A. Nielsen, *Phys. Rev. A* **66**, 032110 (2002).
- [21] M. C. Arnesen, S. Bose, and V. Vedral, *Phys. Rev. Lett.* **87**, 017901 (2001).
- [22] X. Wang, H. Fu, and A. I. Solomon, *J. Phys. A: Math. Gen.* **34**, 11307 (2001).
- [23] X. Wang and P. Zanardi, *Physics Letters A* **301**, 1 (2002).
- [24] X. Wang, *Phys. Rev. A* **66**, 034302 (2002).
- [25] G.-F. Zhang and S.-S. Li, *Phys. Rev. A* **72**, 034302 (2005).
- [26] D.-C. Li, X.-P. Wang, and Z.-L. Cao, *J. Phys.: Condens. Matter* **20**, 325229 (2008).
- [27] L. Zhou, H. S. Song, Y. Q. Guo, and C. Li, *Phys. Rev. A* **68**, 024301 (2003).
- [28] G. Rigolin, *Int. J. Quantum Inform.* **02**, 393 (2004).
- [29] G. Lagmago Kamta and A. F. Starace, *Phys. Rev. Lett.* **88**, 107901 (2002).
- [30] E. Mehran, S. Mahdaviifar, and R. Jafari, *Phys. Rev. A* **89**, 042306 (2014).
- [31] X. Wang and Z. D. Wang, *Phys. Rev. A* **73**, 064302 (2006).
- [32] M. Xiao-San and W. An-Min, *Commun. Theor. Phys.* **52**, 825 (2009).
- [33] Y.-L. Xu, L.-S. Wang, and X.-M. Kong, *Phys. Rev. A* **87**, 012312 (2013).
- [34] V. Eisler and Z. Zimborás, *New J. Phys.* **16**, 123020 (2014).
- [35] P. Calabrese, J. Cardy, and E. Tonni, *J. Phys. A: Math. Theor.* **48**, 015006 (2015).
- [36] A. Kitaev, *Annals of Physics* **303**, 2 (2003).
- [37] C. Castelnovo and C. Chamon, *Phys. Rev. B* **76**, 184442 (2007).
- [38] B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton, *Rev. Mod. Phys.* **88**, 045005 (2016).
- [39] J. Eisert, M. Cramer, and M. B. Plenio, *Rev. Mod. Phys.* **82**, 277 (2010).
- [40] P. Calabrese, L. Tagliacozzo, and E. Tonni, *J. Stat. Mech.* **2013**, P05002 (2013).
- [41] P. Ruggiero, V. Alba, and P. Calabrese, *Phys. Rev. B* **94**, 195121 (2016).
- [42] S. Trebst, P. Werner, M. Troyer, K. Shtengel, and C. Nayak, *Phys. Rev. Lett.* **98**, 070602 (2007).
- [43] C. Castelnovo, S. Trebst, and M. Troyer, “Understanding quantum phase transitions,” (CRC Press, Boca Raton, 2010) Chap. Topological Order and Quantum Criticality.
- [44] C. N. Yang and T. D. Lee, *Phys. Rev.* **87**, 404 (1952).
- [45] T. D. Lee and C. N. Yang, *Phys. Rev.* **87**, 410 (1952).
- [46] C. Castelnovo and C. Chamon, *Phys. Rev. B* **77**, 054433 (2008).
- [47] C. Castelnovo and C. Chamon, *Phys. Rev. B* **76**, 174416 (2007).
- [48] A. Hamma, R. Ionicioiu, and P. Zanardi, *Phys. Rev. A* **71**, 022315 (2005).
- [49] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Lett. A* **223** (1996), 10.1016/S0375-9601(96)00706-2.
- [50] K. Vogel and H. Risken, *Phys. Rev. A* **40**, 2847 (1989).
- [51] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, *Nature Communications* **1**, 149 (2010).
- [52] B. P. Lanyon *et al.*, *Nature Physics* **13**, 1158 (2017).
- [53] J. Gray, L. Banchi, A. Bayat, and S. Bose, (2017), arXiv:1709.04923 [quant-ph].
- [54] H. Weimer, M. Müller, I. Lesanovsky, P. Zoller, and H. P. Büchler, *Nature Physics* **6**, 382 (2010).
- [55] H. Weimer, M. Müller, H. P. Büchler, and I. Lesanovsky, *Quant. Inf. Process.* **10**, 885 (2011).
- [56] M. Sameti, A. Potočník, D. E. Browne, A. Wallraff, and M. J. Hartmann, *Phys. Rev. A* **95**, 042330 (2017).
- [57] Y.-J. Han, R. Raussendorf, and L.-M. Duan, *Phys. Rev. Lett.* **98**, 150404 (2007).
- [58] M. Aguado, G. K. Brennen, F. Verstraete, and J. I. Cirac, *Phys. Rev. Lett.* **101**, 260501 (2008).
- [59] C.-Y. Lu, W.-B. Gao, O. Gühne, X.-Q. Zhou, Z.-B. Chen, and J.-W. Pan, *Phys. Rev. Lett.* **102**, 030502 (2009).
- [60] J. K. Pachos, W. Wieczorek, C. Schmid, N. Kiesel, R. Pohlner, and H. Weinfurter, *New J. Phys.* **11**, 083010 (2009).
- [61] L.-M. Duan, E. Demler, and M. D. Lukin, *Phys. Rev. Lett.* **91**, 090402 (2003).
- [62] A. Kitaev, *Ann. Phys.* **321**, 2 (2006).
- [63] J. Anders, *Phys. Rev. A* **77**, 062102 (2008).
- [64] A. Ferraro, D. Cavalcanti, A. García-Saez, and A. Acín, *Phys. Rev. Lett.* **100**, 080502 (2008).
- [65] K. Audenaert, J. Eisert, M. B. Plenio, and R. F. Werner, *Phys. Rev. A* **66**, 042327 (2002).
- [66] C. D. Nobili, A. Coser, and E. Tonni, *J. Stat. Mech.* **2016**, 083102 (2016).