Diagnosing deconfinement and topological order

K Gregor¹, David A Huse², R Moessner³,⁴ and S L Sondhi²

¹ Courant Institute, New York University, New York, NY 10003, USA
² Department of Physics, Princeton University, Princeton, NJ 08544 USA
³ Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany
E-mail: moessner@pks.mpg.de

New Journal of Physics 13 (2011) 025009 (21pp)
Received 22 November 2010
Published 11 February 2011
Online at http://www.njp.org/
doi:10.1088/1367-2630/13/2/025009

Abstract. Topological or deconfined phases are characterized by emergent, weakly fluctuating, gauge fields. In condensed matter settings, they inevitably come coupled to excitations that carry the corresponding gauge charges which invalidate the standard diagnostic of deconfinement—the Wilson loop. Inspired by a mapping between symmetric sponges and the deconfined phase of the $Z_2$ gauge theory, we construct a diagnostic for deconfinement that has the interpretation of a line tension. One operator version of this diagnostic turns out to be the Fredenhagen–Marcu order parameter known to lattice gauge theorists and we show that a different version is best suited to condensed matter systems. We discuss generalizations of the diagnostic, use it to establish the existence of finite temperature topological phases in $d \geq 3$ dimensions and show that multiplets of the diagnostic are useful in settings with multiple phases, such as $U(1)$ gauge theories with charge $q$ matter. (Additionally, we present an exact reduction of the partition function of the toric code in general dimensions to a well-studied problem.)

4 Author to whom any correspondence should be addressed.
1. Introduction

There is currently much interest in condensed matter systems that exhibit ordering captured by a gauge theory lacking a local order parameter. Such systems offer a contrast to the classical broken symmetry paradigm built around the notion of an order parameter introduced by Landau and these phases are said to be topological or deconfined \([1, 7]\) as they also exhibit quasiparticles with fractional quantum numbers \([2]\). These phases are of interest in the study of strongly correlated quantum systems—where the quantum Hall states \([3]\) and resonating valence bond liquids \([4, 5]\) are canonical examples. Interest in them has been further enhanced by Kitaev’s \([6]\) and Freedman and coworkers’ \([7]\) proposal for utilizing their subset composed of non-abelian braiding statistics for quasiparticles, for the construction of a quantum computer.

The theoretical description of such phases is in terms of a gauge theory, possibly with low-energy matter. In the simplest cases, the low-energy theory is a purely topological gauge theory, such as the Chern–Simons theory in the case of the quantum Hall states \([3]\) or the BF theory in the case of resonating valence bond liquids \([8]\). Indeed, the term ‘topological phase’ dates from these early instances; in the absence of a better standard term, we will use it here to refer to all phases with an emergent gauge field. The explosion of recent interest has come from the construction of a wide variety of lattice models that realize a variety of long-wavelength theories, including those with dynamical gauge fields \([1, 9]\).
In idealized models, the existence of such phases is transparent for one can readily show that the gauge field (i) exhibits weak fluctuations and thus a perimeter law for the Wilson loop in ground states, (ii) mediates a non-confining force between its sources (‘quarks’) and (iii) (where the low-energy theory is purely topological) leads to a topology-dependent ground state degeneracy. For more realistic Hamiltonians, unfortunately, this transparency is lost. The problem with the first two characterizations on our list is that, as is well known from the lattice gauge theory literature, they fail in the presence of dynamical sources (‘matter’) for the gauge field.

In the condensed matter setting, the gauge field is only emergent and necessarily comes with higher-energy excitations that carry a gauge charge, making this a generically fatal flaw. The situation only gets worse at finite temperatures, where such matter excitations must enter the statistical sum. The third characterization can continue to work in strictly topological phases precisely at \( T = 0 \) but lacks a clear meaning beyond that limit and outside that subset of deconfined phases.

All of this is an unsatisfactory state of affairs. As in the Landau theory, it would be nice to have a fixed time diagnostic operator that one can compute when handed a ground state or a Gibbs state to decide whether it exhibits the ordering characteristic of a given topological phase. In this paper, we report the construction of such a diagnostic that teases out the underlying weakly fluctuating behavior of the gauge field. We do so via a detour into the theory of sponge phases, where our diagnostic has the interpretation of a line tension—picking up a line of analysis begun by Huse and Leibler \([12]\) a while back. Remarkably, this diagnostic turns out to be a spacetime generalization of the so-called Fredenhagen–Marcu order parameter \([13, 14]\) known to lattice gauge theorists.

In this paper, we also begin the process of applying these ideas to a wide class of systems. While the bulk of our paper is concerned with the \( Z_2 \) gauge theory with matter (or Kitaev’s toric code \([6]\) with generic perturbations), we sketch the generalization to different low-energy gauge structures, including cases where it is necessary to use a multiplet of diagnostics. Notably, we use the diagnostic to establish the survival of topological phases to finite temperatures in \( d \geq 3 \) \([15]\).

Turning to the contents of the paper, we begin in sections 2 and 3 by reviewing the formulation of the \( Z_2 \) gauge theory with \( Z_2 \) matter and its reformulation as a statistical mechanics of surfaces with edges. This leads us to the identification of the line tension diagnostic in section 4 and its operator formulations in section 5. In section 6, we discuss why condensed matter systems prefer a particular operator formulation. Sections 7 and 8 deal with the applications of the ideas to finite-temperature phases and more complex phase diagrams. The conclusion flags some open questions. Finally, two appendices establish some useful results on Kitaev’s toric code: firstly, we obtain its partition function in general dimension via a reduction.

---

5 There is a spectral test—the deconfined phase has particles that carry the charge of the deconfined gauge field and also characteristically exhibit fractionalization of the microscopic quantum numbers. For example, in our workhorse example later in this paper, we can ask if the Hamiltonian leads to isolable particles with \( Z_2 \) \([10]\). However, this test requires a fair amount of information and in the presence of bound states of the fractionalized objects even more so. Note also that fractionalized quantum numbers are not quantized signatures of a deconfined phase \([11]\).

6 Of the existing proposals for diagnosing topological phases, the one that is closest in spirit to such a prescription is the computation of an entanglement entropy, using a combination of appropriately defined reduced density matrices for a system subjected to different partitionings \([31, 32]\).
to a classical pure $Z_2$ gauge theory; secondly, we demonstrate the perturbative stability of its topological order [16, 17] and find that it persists to finite temperature only for $d \geq 3$, in accordance with [18].

Before proceeding, we should point out an important antecedent to our work which also addressed the challenge of exhibiting topological order away from the idealized models. Hastings and Wen [19] gave a $T = 0$ continuity construction starting with idealized Hamiltonians that generates a Hamiltonian-dependent gauge field operator that exhibits a perfect perimeter law (with zero coefficient) for the Wilson loop in deconfined phases. By contrast, our construction is Hamiltonian independent in the spirit of an order parameter.

2. $Z_2$ Gauge theory with matter

Consider the Hamiltonian of the $Z_2$ lattice gauge theory,

$$-H = K \sum_p \prod_{l \in \partial p} \sigma^z_l + \Gamma \sum_l \sigma^z_l + J \sum_{s \in \partial l} \tau^z_s + \Gamma_M \sum_s \tau^x_s,$$

supplemented by the constraint that we restrict its action to ‘gauge invariant’ states defined by

$$G_s|\psi\rangle = |\psi\rangle, \quad G_s = \tau^x_s \prod_{l:x \in \partial l} \sigma^z_l,$$

where the gauge ($\sigma^i_l$) and matter ($\tau^i_s$) operators act in spin-1/2 Hilbert spaces that live on the links $l$ and sites $s$, respectively, of a (hyper) cubic lattice in $d$ dimensions. The subscripts $s$, $l$ and $p$ denote sites, links and plaquettes. $\partial p$ and $\partial l$ are the boundaries of the corresponding objects.

For $\Gamma = J = 0$, our Hamiltonian is equivalent to Kitaev’s toric code as noted in the original paper [6]. In that case, the degenerate $2^{N_{nc}}$ ground states ($N_{nc}$ is the number of independent non-contractible loops on the lattice) are easily constructed and exhibit a perfect ‘perimeter’ law for the contractible Wilson loop,

$$W[C] = \left( \prod_{l \in C} \sigma^z_l \right) = 1.$$  

Consistent with this ground state description, one sees that the excited states contain non-interacting charges, free to sit at arbitrary locations in this particular model. The remaining part of the excitation spectrum is described by purely gauge excitations. These are vortices of the gauge field or, visons, which are isolated plaquettes for which $W[C] = -1$. Both charged and gauge excitations are separated from the set of ground states by a finite gap.

The full $T = 0$ phase diagram of the Hamiltonian (1) has two distinct phases. The first is the deconfined phase that exists for small perturbations of the Kitaev point. (For a more detailed analysis of the toric code phase diagram, including $T > 0$, see the appendix.) The remaining phase is the confined-Higgs phase, which takes its name from the limiting regions that were shown to be smoothly connected by Fradkin and Shenker [10].

How do we distinguish between the two phases? Unfortunately, except along the line $J / \Gamma_M = 0$, the Wilson loop $W[C]$ does not readily detect the phase transition—it exhibits a perimeter law everywhere (and also at $T > 0$). As previously remarked, the potential between two test charges is also not a sharp diagnostic due to the presence of dynamical matter.
3. Surfaces, edges and the symmetric sponge

In order to get to our diagnostic, let us now reformulate the Hamiltonian gauge theory considered above as the classical statistical mechanics of a system of membranes. To do this, we first write down the path integral formulation with a discretized imaginary time. This is governed by the action

$$-S = K \sum_p \prod_{l \in \partial p} \sigma_l + J \sum_l \prod_{s \in \partial l} \tau_s,$$

(4)

where $\sigma_l$ are the gauge fields living on the links, and $\tau_s$ the matter fields residing on the sites of a $d + 1$-dimensional hypercubic lattice, with $d \geq 2$ so that we can discuss a deconfined phase. In writing this spacetime symmetric form, we have moved away from the anisotropic limit that is needed for precise reproduction of the Hamiltonian problem but which is not important for our purposes. The reader should also bear in mind that $K$ and $J$ refer to different quantities, with different units, in the Hamiltonian (1) and the action (4). However, as their physical import is very similar, it is useful to keep this notation.

We next rewrite the partition function as

$$Z(K, J) = (\cosh K)^{N_p} (\cosh J)^{N_l} \times \text{Tr}_{\sigma, \tau} \left[ \prod_p (1 + \tanh K \prod_{l \in \partial p} \sigma_l) \right] \left[ \prod_l (1 + \tanh J \prod_{s \in \partial l} \tau_s) \right],$$

(5)

where $N_{p,l}$ are the number of plaquettes and links, respectively, in the lattice. Multiplying out the products in the brackets and performing the trace annihilates all terms in which any $\tau_s, \sigma_l$ appears an odd number of times. For $J = 0$, this allows only for the presence of closed surfaces, a surface being defined as containing a given plaquette $p$ if the factor of $\prod_{l \in \partial p} \sigma_l$ appears in the sum. Switching on a non-zero $J$ allows the surfaces to have free edges. Thus we have rewritten our gauge theory as a statistical mechanics of surfaces with edges as promised. In this formulation, each area element of the surface has a bare surface tension $\tanh K$, and each edge a bare line tension $\tanh J$. The actual geometrical properties of the different regimes and phases are determined by the renormalization of these quantities due to the entropy of the surfaces and edges.

The phase diagram of the problem is sketched in figure 1 and includes more structure than is significant for the gauge theory, as we explain below. We direct the reader’s attention to the following features.

3.1. The pure gauge theory

As already noted, for $J = 0$, the system excludes edges. All surfaces are closed, which allows a consistent definition of an inside and an outside. For small $K$, the total surface area remains small, and the volume of the inside of the surfaces is tiny, whereas the outside fills most of the space. In this regime, there is a macroscopic surface tension that is in one-to-one correspondence with an area law for the Wilson loop, as the reader can check by tracking the latter through our mapping.

At a larger $K = K_g$, the inside of the surfaces percolates as well. This first transition is purely geometrical and thus does not produce thermodynamic or local operator singularities,
e.g., the surface tension is analytic across the boundary. In the language of membranes, the percolating phase is called the asymmetric sponge, as the inside and outside have unequal volumes. The transition is dual to the one encountered in the Ising models in $d = 3$, where the minority phase starts percolating at a temperature below the thermodynamic $T_c$.

At $K_c > K_g$, there is a thermodynamic transition that corresponds to deconfinement in the pure gauge theory. This is signaled by the vanishing of the macroscopic surface tension: the Wilson loop acquires a perimeter law, while for the sponge the inside and outside volumes become equal whence the new, deconfined phase is labeled a ‘symmetric sponge’.

### 3.2. Regime with heavy dynamical matter

Turning on $J$ causes edges to appear. Even at small $J$, where the bare line tension is large, edges have the qualitatively important effect of causing the macroscopic surface tension to vanish whence the Wilson loop exhibits a perimeter law for any $K$ and $J > 0$. However, the phase transition between the asymmetric and symmetric sponges is not destroyed at small $J$ even if the labeling is now problematic as the inside and outside are not distinct. There continues to be an intuitive distinction in the tension associated with surfaces between the two phases, as illustrated in the bottom panels of figure 2. In the symmetric sponge phase, the Wilson loop on the right follows a perimeter law despite the large area present, whereas on the left, a perimeter law arises because dynamical matter allows the surface area to be proportional to the perimeter of the loop by introducing an edge that runs along it.

---

7 ‘Heavy’ refers to the large creation cost, $2\Gamma_m$, rather than a small bandwidth, $\propto J$, of the matter particle.
3.3. Regime with light dynamical matter

As the bare line tension, \( \tanh J \), is increased further, edges proliferate. Of most interest to us, this has the effect of terminating the symmetric sponge phase and driving the system into the 'sponge with free edges', which is the Higgs region of the gauge theory in the present formulation. This transition is most vivid at \( K = \infty \), where surfaces drop out and we are left with the statistical mechanics of loops alone. In fact, these loops are those of the high-temperature expansion of the Ising model, with a vanishing line tension at \( J_c \) signaling the onset of ferromagnetic order, a thermodynamic transition that separates the deconfined from the Higgs phase\(^8\).

The sponge with free edges is accessible, without crossing a thermodynamic phase transition, from both the asymmetric sponge and the vesicles and disc regions, in keeping with the continuity between the Higgs and confinement regions. However, the appearance of the surfaces nonetheless varies drastically. Deep in the Higgs phase, long edges frame large surfaces. By contrast, in the ‘seaweed’ regime, edges are still cheap and the tradeoff is between edge entropy and the cost of the surfaces attached to the edges, leading to lines joined by (lattice-dependent) surfaces close to the minimal area.

4. The Huse–Leibler horseshoe

Having described the phase diagram of the \( Z_2 \) gauge theory in terms of our system of surfaces with edges, let us now turn to the task of distinguishing the symmetric sponge/deconfined phase from all others. Intuitively, we need a quantity that teases out whether the perimeter law of the Wilson loop is fundamentally due to a vanishing surface tension, or instead due to the absence of a surface spanning it and whether there is an underlying large line tension. Somewhat non-intuitively, this can be accomplished by simply defining an appropriate line tension.

\(^8\) Note that there is no additional geometrical transition for the edges at a \( J_g > J_c \) even in the self-dual case \( d = 3 \); rather, it will show up as a transition in the appropriate dual variables.
To this end, consider the geometrical object inscribed in figure 1. It is a Wilson loop broken open, so that its total length remains fixed but so that a gap of size $g$ opens up. To keep the quantity gauge invariant, it is necessary to terminate the open strings by a matter term, so that the Huse–Leibler horseshoe [12] is defined by the quantity

$$H(L, g) = \left\langle \tau_s \left( \prod_{l \in C_H(s, s')} \sigma_l \right) \tau_{s'} \right\rangle,$$

where the product is taken over the links $C_H(s, s')$ constituting the horseshoe between $s$ and $s'$.

The basic point is that if surfaces are cheap and edges are expensive—the situation prevailing in the symmetric sponge—the system will cover most of the horseshoe with a surface and then close the gap with an edge at a cost in free energy of $\exp(-\lambda g)$ due to the tension of the line closing the gap (red line in figure 1). By contrast, if surfaces are expensive and/or edges are cheap, which is the situation elsewhere in the phase diagram, no such dependence on $g$ arises. As a practical matter, to separate out the dependence on $g$ from that on the linear size of the horseshoe, $L$, we need to study this quantity in the limit $L \gg g \gg 1$ and then check for an effective line tension $\lambda$ in the expectation value of the horseshoe.

An elegant alternative is to remove the $L$-dependence by dividing the horseshoe by the Wilson loop of the same size $L$, as the leading $L$-dependence of the two is set by the thermodynamic properties of the surface and edges. Once this has been done, the resulting ratio will either vanish as $\exp(-\lambda g)$ in the symmetric sponge, or approach a constant. Indeed, one does not even need to introduce a separate length scale $g$ altogether. One can instead decide to cut the Wilson loop in two. For a square loop of side length $L$, as shown in figure 2, this yields two rectangular loops of size $L/2 \times L$, one long side of which is open and now takes over the role of the gap; again, a $\tau$ needs to be attached to each open end. The result is the line tension diagnostic ratio,

$$R(L) = \frac{W_{1/2}(L)}{W(L)} = \frac{\left\langle \tau_s \left( \prod_{l \in C_{1/2}} \sigma_l \right) \tau_{s'} \right\rangle}{\sqrt{\left\langle \prod_{l \in C} \sigma_l \right\rangle}},$$

where the curves $C_{1/2}$ and $C$ define the half-Wilson loop $W_{1/2}$ and the full Wilson loop $W$, respectively. $R(L)$ has the desired asymptotic behavior,

$$\lim_{L \to \infty} R(L) = 0 \quad \text{deconfined phase},$$

$$\lim_{L \to \infty} R(L) \neq 0 \quad \text{otherwise}.$$  

Before we work further with $R$, three comments are in order. Firstly, along the $K = \infty$ axis, this ratio is simply the spin–spin correlation function of the Ising model evaluated between the endpoints of the half-Wilson loop and it distinguishes the two phases as asserted. Secondly, this diagnostic fails along exactly one line—when $J = 0$, but fortunately the Wilson loop itself is available exactly in that case. Thirdly, we have checked by Monte Carlo, and in the small $J$ expansion, that $R$ has the claimed behavior.

5. Operator formulations and the Fredenhagen–Marcu order parameter

The line tension diagnostic introduced in the last section can be oriented in an arbitrary direction for the Lorentz invariant example studied so far. To obtain an operator formulation,
we must pick a particular orientation. Here we consider three and gain some insight into the operation of the diagnostic.

5.1. Equal time formulation

The simplest option is to orient the variables in the ratio so that they all lie in the same time slice of the path integral as in figure 3(a). Then the transcription to the Hamiltonian formulation is straightforward and we obtain from the definition in equation (7),

\[
R(L) = \frac{W_{1/2}^{1/2}(L)}{\sqrt{W(L)}} = \frac{\langle G| \tau^z_s (\prod_{l \in C_1} \sigma^z_l) \tau^z_s' |G \rangle}{\sqrt{\langle G| \prod_{l \in C} \sigma^z_l |G \rangle}},
\]

(9)

where \(|G\rangle\) is the ground state. As written, at \(T = 0\), \(R(L)\) clearly measures a property of the ground state wavefunction but it is straightforward to use it at \(T > 0\) with thermal
averages replacing ground state ones. Transcribing our previous analysis, \( R(\infty) \) vanishes if the equilibrium gauge field fluctuations are weak and the matter is uncondensed but not otherwise. This turns out to be the form of the diagnostic which can be used most generally, for reasons that we explain later.

5.2. Fredenhagen–Marcu order parameter

Now let us orient the operators so that they lie in a plane containing the imaginary time axis with the endpoints of the half-Wilson loop at the same time (see figure 3(b)), which we take to be zero for convenience. Then we can write the operator expression,

\[
W_{1/2} = \langle G | \tau_z^s \tau_z^{s'} \prod_{l \in C_{ss'}} \sigma_z^l (-T/2) | G \rangle,
\]

where \( T = cL \) up to multiplication by a velocity \( c \), here set to 1, \( C_{ss'} \) is the linear path between \( s \) and \( s' \), and

\[
\sigma_z^l (-T/2) = e^{-HT/2} \sigma_z^l e^{-HT/2}
\]

and \( e^{-H} \) is the transfer matrix for one step in the time direction—with the implied meaning in the time continuum limit. (In the process of passing to the Hamiltonian setting, we have, as usual, set \( \sigma_z^l = 1 \) along the temporal links.)

Similarly, we can write the denominator as

\[
W = \langle G | \prod_{l \in C_{ss'}} \sigma_z^l (+T/2) \prod_{l \in C_{ss'}} \sigma_z^l (-T/2) | G \rangle.
\]

Now define

\[
\langle ss' \rangle = \tau_z^s \tau_z^{s'} \prod_{l \in C_{ss'}} \sigma_z^l (-T/2) | G \rangle
\]

as the candidate two-spinon state constructed from the ground state. Thus, we see that

\[
R(L) = \frac{\langle G | ss' \rangle}{\langle ss' | ss' \rangle},
\]

i.e. it is the overlap between the ground state and the normalized two-spinon state. In this form, our diagnostic was first discovered by Fredenhagen and Marcu \[13, 14\] and they proved that it does show the asymptotic behavior \((8)\) in the \( Z_2 \) gauge theory with matter. In subsequent work, they also considered the equal time form \((9)\).

It follows from our previous considerations that

- in the deconfined phase, the candidate two-spinon state is orthogonal to the ground state in the limit \( cL = T \to \infty \) and has finite energy and thus there are free spinons in the spectrum;
- in the confined-Higgs phase, the candidate two-spinon state is not orthogonal to the ground state and thus there are no free spinons in the spectrum.

In this fashion, we see that in this interpretation, \( R(\infty) \) distinguishes between the phases based on the existence of free spinons in the spectrum.

Finally, it is instructive to see how these limits work directly in the Hamiltonian formalism. Consider the sequence
(a) the gauge-non-invariant string $\prod_{l \in C^\prime} \sigma_z^l (-T/2)$ acts on the ground state leading to a state with a modified constraint at $s$ and $s'$.
(b) the system evolves/relaxes for an imaginary time $T/2$,
(c) the gauge-non-invariant product $\tau^z_s \tau^z_{s'}$ acts on the state at the end of the last step and leads to a final gauge invariant state.

Deep in the deconfined phase, (a) and (b) leave an additional electric flux string that meanders between the sites $s$ and $s'$ and (c) adds the matter needed to create a proper two-spinon state. Conversely, deep in the confined phase, (b) eliminates the string introduced by (a) in favor of fluxless charges at $s$ and $s'$, which are then eliminated in (c), leaving the ground state behind.

### 5.3. Spinon delocalization

The two interpretations provided above are sufficient for our purposes in the rest of the paper, but some further insight can be gained by considering a third one: now we orient the half-Wilson loop so that $s$ and $s'$ are separated by imaginary time $T$, as shown in figure 3(c).

With this choice, the numerator becomes the expectation value,

$$ W_{1/2} = \langle G | \tau^z_s (T/2) \prod_{l \in C_0} \sigma_z^l (T/2) \tau^z_s (-T/2) \prod_{l \in C_0} \sigma_z^l (-T/2) | G \rangle, \quad (15) $$

while the denominator is given by the same expression exhibited previously.

To understand why the ratio $W_{1/2} / \sqrt{W}$ behaves differently in the two phases, consider the intermediate states produced by imaginary time evolution. In the deconfined phase
- $W_{1/2}$ couples to a state with one defect site at $s'$ and a delocalized spinon, while
- $W$ couples to a state with two widely separated defect sites.

Hence, the ratio is controlled by intermediate states with different energies and scales as

$$ \frac{e^{- (E_{\text{defect}} + E_{\text{spinon}}) T}}{\sqrt{e^{- (E_{\text{defect}} + E_{\text{defect}}) T}}} \quad (16) $$

and thus vanishes at large $cL = T$.

By contrast, in the confined phase, both operators couple to eigenstates with localized spinons, at $s'$ for $W_{1/2}$ and at both $s$, $s'$ for $W$, whence the ratio tends to a constant. It follows then that in this version the diagnostic is sensitive to the existence of a delocalized spinon state in the deconfined phase.

### 6. Fluctuating constraints and choice of operator formulation

In a condensed matter context, a constraint such as the one encoded in equation (2) is not microscopic (or ‘fundamental’) but rather of emergent origin; that is to say it is a legacy of higher-energy terms in the Hamiltonian. The constraint is thus not inviolable—rather, states outside the ‘gauge invariant’ or ‘physical’ sector are associated with a large but finite energy penalty set by a scale $U$. To put it differently, in the context of gauge theory, it is assumed that ‘unphysical’ states are really unphysical—they are forced on us by the challenge of handling the constraint—but in condensed matter setting they are certainly physical but are instead high-energy states.
This distinction has an important consequence for the operator versions of our diagnostic, e.g. the Fredenhagen–Marcu version (14). As noted in the discussion following its introduction, it contains a two-spinon state creation operator that can be factored into two pieces,

$$\tau_s^z \tau_s^z$$ and $$\prod_{l \in C_{ss}'} \sigma_l^z (-T/2),$$

which are not individually gauge invariant, although their combination is. Now the second piece is sensitive not only to matrix elements of $$H$$ between gauge invariant states but also to matrix elements of $$H$$ involving gauge non-invariant states. Indeed, our verbal explanation of the creation of the two-spinon state referred specifically to imaginary time evolution in the gauge non-invariant sector where the constraint was violated at two sites. For the specific choice of Hamiltonian (1) all is well—the detour into the gauge non-invariant sector introduces no physics inconsistent with the proper operation of the diagnostic. Of course, this had to be the case as we got this form working from a manifestly Lorentz invariant Euclidean path integral.

We now show that this does not have to be the case by explicitly exhibiting an example with a soft constraint where the Fredenhagen–Marcu version of $$R(\Lambda)$$ fails to diagnose deconfinement. The example is somewhat contrived but it is intended only to function as a cautionary counter-example.

Consider modifying our starting $$Z_2$$ gauge theory Hamiltonian (1) to $$H_f$$,

$$H_f = H + \frac{U}{2} \sum_s (G_s - 1)^2 - \lambda \sum_s (G_s - 1) \tau^x$$

and set $$J, \Gamma \rightarrow 0$$ for now. Now we do not implement the constraint (2) on the Hilbert space, instead a suitably large value of $$U$$ does this in the low-energy sector. With that ordering of energy scales, the ground state at small $$\lambda \ll U$$,

$$|G\rangle = \left(\prod_s \frac{1}{2} (G_s - 1)\right) |\{\sigma^z = 1\}\rangle \otimes |\{\tau^x = 1\}\rangle,$$

has precisely the same form as that of the toric code. However, there is now a change of meaning, as microscopically both the link and site variables are physical while in the toric code the site variables can be gauged away. Yet, at low energies in our model, the same reduction takes place and so it has as many degrees of freedom and upon our choices of parameters exhibits the same, but now emergent, $$Z_2$$ order.

Now consider the excitations of our model. The gauge invariant excitations preserve $$G_s = 1$$ at all sites and are visons with an energy $$2K$$ and spinons with an energy $$2\Gamma_M$$. Note that their energies are independent of $$\lambda$$; indeed, all gauge invariant states have (by construction) energies independent of $$\lambda$$. The gauge non-invariant excitations involve violating the constraint by setting $$G_s = -1$$, and also come in two species: spinonic defects with energy $$U + 2\Gamma_M - 2\lambda$$ obtained by setting $$\tau^x = -1$$ at a site and gauge defects with energy $$U + 2\lambda$$ obtained by acting with a string $$\prod \sigma_l^z$$ extending from a given site to infinity. Observe that the energies of the two defect states cross when $$U \gg \lambda > \Gamma_M/2$$, even though the spectrum of the entire set of lower-energy gauge invariant states is unchanged. (For the Hamiltonian (1), the spinonic and gauge defects have energy $$2\Gamma_M$$ and 0, respectively. While there is no explicit energy penalty for gauge non-invariant
states, the separate requirement that the ground state obey (2) eliminates the degeneracy that would be present otherwise.)

In our previous discussion, we argued that the action of \( \prod_{l \in C_{ss'}} \sigma^z_l (-T/2) \) was to create a string stretching between sites \( s \) and \( s' \). This works as long as the state with two gauge defects is the ground state in the sector with two defect sites. Once the defects cross and the two spinonic defects become the ground state, this is no longer true. As a technical matter, the crossing alone is not enough, and we also need to turn \( J \) on to a small value so that imaginary time evolution connects the two states. Once this is done, \( R(\infty) \neq 0 \), even though we are clearly in the deconfined phase and the diagnostic fails in this form.

However, the diagnostic still works in its equal time form, so the lesson is that in general condensed matter settings the formulations presented in the previous section are not fully equivalent: whereas the equal-time formulation (section 5.1) is robust, the other pair of formulations (sections 5.2 and 5.3) need not be.

7. Finite temperature topological order in \( d \geq 3 \)

We now return to our main development and consider the use of our diagnostic to establish the finite temperature phase structure of the \( \mathbb{Z}_2 \) gauge theory where contrasting answers can be found even in the recent literature [15, 18, 21]. We begin by considering the \( T > 0 \) physics of the toric code, \( J = \Gamma = 0 \) in equation (1). It can be shown (see the appendix) that at this point in \( d + 1 \) spacetime dimensions the quantum partition function at all temperatures is proportional, up to a simple multiplicative analytic factor, to that of the pure \( \mathbb{Z}_2 \) gauge theory without matter defined on a \( d \)-dimensional Euclidean lattice.

From what is known about the latter theory, one can immediately read off that topological order will persist to finite temperature for all \( d \geq 3 \), but not below. Further, we see that, despite the presence of matter, the transition between the low-temperature deconfined phase and high-temperature phase can be diagnosed by the Wilson loop, which goes from a perimeter law to an area law between the two phases.

This physics has to do with the gauge excitations. For \( d = 2 \), finite energy point-like visons appear at an exponentially small but non-zero density at any \( T > 0 \) destroying the deconfined phase. As the Wilson loop measures the averaged parity of the enclosed vorticity, it acquires an area law as a result.

By contrast, this mechanism is not operative in higher dimension because the gauge/magnetic excitations now appear in a higher-dimensional version—loops in \( d = 3 \)—whose characteristic size, \( \xi_m \), vanishes as \( T \to 0 \). As the Wilson loop now counts the parity of flux loops that link it, it is sensitive only to those within \( \xi_m \) of its boundary, whence the perimeter law and the deconfined phase survive at sufficiently small temperatures.

At higher temperatures, there is a phase transition, e.g in \( d = 3 \) where the vortex loops unbind. The basic physics of vortex loops discussed above indicates that the finite temperature deconfined phase should be stable beyond the Kitaev point in \( d \geq 3 \), as argued in [15, 18]. However, once matter becomes dynamical, the Wilson loop exhibits a perimeter law at any \( T \) and other standard diagnostics fail as well.

Fortunately, our diagnostic continues to work at \( T > 0 \). At small temperatures, one can work perturbatively at small \( J \), when the dynamical matter is heavy. As at \( T = 0 \), the basic result is that in this limit the behavior of \( R(L) \) follows from that of \( W(L) \) in the absence of dynamical matter. As the latter exhibits a transition between perimeter and area dependence, \( R(\infty) \)
transitions between vanishing and developing an expectation value of $O(L^0)$. In somewhat more detail, one can examine the Euclidean path integral, now finite in the (large) imaginary time direction at low temperatures, much as in section 3. As before, a large line tension between any pair of charges survives, and it continues to be preferable to minimize free edges at the expense of adding the (cheap) surface, leading to a vanishing of $R(\infty)$.

At high temperatures, $T \to \infty$, it is more convenient to work directly in the Hamiltonian formulation and the high-temperature expansion. In this expansion, one obtains polynomials in powers of $\beta = 1/T$ upon expanding $\exp(-\beta H)$. Non-vanishing terms in the expectation values appearing in equation (7) are only obtained if each unpaired $\sigma^z, \tau^z$ is matched by the polynomials. The cheapest (in powers of $\beta$) way of doing this is to retrace the perimeter of $W$ and $W_{1/2}$. This leads to the same leading power of $\beta$ in the expectation value of each, which thence cancel in the expression for $R(\infty)$. Consequently, $R(\infty)$ is non-zero at sufficiently high temperatures and we find that there is a phase transition between the high and low temperature phases.

8. Beyond $Z_2$

8.1. Gauge fields with fundamental matter

So far, our discussion has focused on the simplest case, that of standard $Z_2$ gauge theory with $Z_2$ matter where there are two phases to distinguish—the deconfined phase and the confined-Higgs phase. Our deconfinement diagnostic generalizes straightforwardly to lattice gauge theories based on other gauge groups where there are similarly only two phases to distinguish, to $Z_n$, $U(1)$ and to Yang–Mills theories based on non-abelian groups. Indeed, the last named case was the primary spur to Fredenhagen and Marcu’s work. In all of these cases, $R$ has the same structure with appropriate replacement of the matter and gauge fields.

For example, a compact $U(1)$ gauge theory with charge $q = 1$ matter is defined by a spacetime action analogous to (4)

$$-S = K \sum_p \prod_{l \in \partial p} U_{\tilde{l}(l)} + J \sum_{l, s, s' \in \partial l} \tau_s U_{l} \tau_{s'} + \text{h.c.},$$

(19)

where the $U_{l} = \exp(iA_{l}) = U_{-l}^*$ are now $U(1)$ valued gauge fields living on the links defined in the standard orientation from one sublattice to the other, and the $\tau_s = \exp(i\phi_s)$ are the corresponding matter fields residing on the sites. Unlike the $Z_2$ theory, we now need to pay attention to the orientation of various link variables. The links in the second term are traversed in the standard orientation while in the first term the loops are taken anticlockwise with $\tilde{l}(l) = \pm l$ if the link is traversed in, or opposite to, its standard orientation, respectively.

At zero temperature in spatial dimension $d \geq 3$, this theory exhibits both a confined and a deconfined phase [10] (see figure 4). Absent matter, $J = 0$, the oriented Wilson loop,

$$W_l[C] = \left( \prod_{l \in C} U_{\tilde{l}(l)} \right),$$

(20)

is capable of distinguishing between these two phases, exhibiting an area law in the former and a perimeter law in the latter, with logarithmic corrections signaling the Coulomb interaction [20].
Figure 4. Sketch of the phase diagram of the U(1) gauge theory with charge 1 matter fields. The behavior of the line tension diagnostic $R_1$ and the Wilson loop $W_1$ is indicated for each phase.

This phase structure persists when matter with charge $q = 1$ is turned on. However, as in the case of the $Z_2$ theory above, the Wilson loop fails as a diagnostic in this setting. By analogy to equation (7), we construct the appropriately generalized diagnostic ratio,

$$R_1(L) = \frac{W_{1/2}(L)}{\sqrt{W(L)}} = \frac{\langle \tau_s^\dagger \prod_{l \in C} U_{l_{(i)}} \tau_s \rangle}{\sqrt{\langle \prod_{l \in C} U_{l_{(i)}} \rangle}}.$$  

As before, the phases are distinguished by the asymptotic behavior of $R_1(L)$,

$$\lim_{L \to \infty} R_1(L) = 0 \quad \text{deconfined phase},$$

$$\lim_{L \to \infty} R_1(L) \neq 0 \quad \text{otherwise}. \quad (22)$$

We now turn to the application of the diagnostic to cases with more than one phase and to problems phrased in more standard condensed matter settings.

8.2. U(1) gauge theory with charge $q$ matter

To obtain a richer phase structure, we consider a U(1) gauge theory coupled to matter fields that carry one or more higher (non-fundamental) charges $q \geq 2$. In this setting, it is useful to consider the set of charge-$q$ operators,

$$W_q(L) = \left\langle \prod_{l \in C} U_{l_{(i)}}^q \right\rangle; \quad R_q(L) = \frac{\langle \tau_s^\dagger \prod_{l \in C} U_{l_{(i)}}^q \tau_s \rangle}{\sqrt{\langle \prod_{l \in C} U_{l_{(i)}}^q \rangle}}, \quad (23)$$

whose structure indicates that $q$ electric flux quanta emanate from a given charge.

As our first example, we consider the well-known phase diagram of a single-matter field of charge $q = 2$. This contains, as shown in figure 5, in addition to the deconfined phase of the $U(1)$ field, a Higgs phase that corresponds to the deconfined phase of the appropriate $Z_2$ gauge theory and a distinct completely confined phase [10]. Due to the lack of $q = 1$ matter, we can make some progress using Wilson loops alone.
Figure 5. Sketch of the phase diagram of the $U(1)$ gauge theory with charge 2 matter fields. Charge-$q$ line tension diagnostic $R_q$ and Wilson loop $W_q$ are defined in equation (23).

Table 1. Diagnostics for $U(1)$ gauge theory with only $q = 2$ matter.

| Phase        | $W_1(L)$ | $R_2(L)$ |
|--------------|----------|----------|
| Deconfined   | Perimeter 0 |          |
| Higgs        | Perimeter $L^0$ |          |
| Confined     | Area $L^0$     |          |

Of the three phases, two ($U(1)$ and $Z_2$ deconfined) exhibit a perimeter law in $W_1$, while one ($U(1)$ confined) exhibits an area law. The key point is that the $q = 2$ dynamical matter cannot break up the confining string that stretches between two test $q = 1$ sources in the confined phase. By contrast, $W_2$ exhibits a perimeter law everywhere and distinguishing the two deconfined phases now requires that we evaluate the diagnostic $R_2$. As this tests for the presence of doubly charged states in the spectrum, it vanishes only in the fully deconfined phase.

In terms of the nature of gauge-field fluctuations, these distinct behaviors can be seen most intuitively by gauge fixing $\phi_s \equiv 0$. Then, deep in the deconfined phase, an appropriate projection of the state $\otimes |A_i = 0\rangle$ is a ground state. In contrast, the gauge field fluctuates strongly in the confined phase, and the $A_i$ ranges over the full interval $[0, 2\pi)$. In the Higgs phase, fluctuations are restricted: here, $A_1 = 0$ or $A_1 = \pi$ satisfy the second term of (19). By construction, $R_2$ is sensitive to the former but not the latter fluctuations and it can thus diagnose the Higgs-confinement transition. This state of affairs is summarized in table 1.

As a second example, we consider the more generic and complex situation where charges with both $q = 1$ and $q = 2$ are present simultaneously. The number of parameters in the Hamiltonian makes for a large phase diagram, and we restrict ourselves to the instructive case where the $q = 1$ matter is very heavy, i.e. very sparse. This then has the same physics as our last example and thus the same phase diagram, displayed in figure 6, as figure 5. The new feature is that now $W_1$ obeys a perimeter law in all three phases so that we may no longer use it to test for the deconfinement of charge-1 objects. The solution to this is to evaluate $R_1$ instead. The diagnostic behavior of the combination of $R_1$ and $R_2$ is summarized in table 2.
We note that this combination manages to distinguish all three phases, whereas the \( W_{1,2} \) fail to do so.

These deconfined phases persist to finite temperatures in dimension \( d \geq 4 \). The known behavior in lower dimensions is more complicated. The \( U(1) \) deconfined phase is absent in \( d = 3 \) at \( T > 0 \) and in \( d = 2 \) even at \( T = 0 \), while the \( Z_2 \) deconfined phase is present in both of these cases. This complexity is correctly captured by our doublet of diagnostic ratios.

### 8.3. Gauge theories of quantum magnets

The formalism developed above can be directly applied to two classes of Hamiltonians that arise naturally in discussions of quantum magnetism. The first are the ‘slave particle’ or Schwinger boson/fermion Hamiltonians [22]–[24]. Here the spin Hamiltonian of a quantum magnet is reformulated as a matter gauge system with spinons coupled to a gauge field. Here our construction of \( R(L) \) carries over pretty much directly.

The second class are quantum dimer models [25, 26], now with a Hilbert space expanded to include gapped, spin-carrying monomers to represent spinons [27]. In this setting, the microscopic gauge structure arises from the hard core dimer constraint, and the Wilson loop is defined as a dimer rearrangement on a loop that respects the hard core constraint [28]. Accordingly, we can again define our diagnostic.

---

**Figure 6.** Sketch of the phase diagram of the \( U(1) \) gauge theory with charge 2 as well as heavy charge 1 matter fields.

**Table 2.** Diagnostics for \( U(1) \) gauge theory with \( q = 1 \) and \( q = 2 \) matter.

| Phase       | \( R_1(L) \) | \( R_2(L) \) |
|-------------|--------------|--------------|
| Deconfined  | 0            | 0            |
| Higgs       | \( L^0 \)    | 0            |
| Confined    | \( L^0 \)    | \( L^0 \)    |
9. Conclusion

We have substantially met the challenge posed at the start of this paper, that of finding a diagnostic for deconfinement that has the form of a ground state expectation value, i.e. which is a property of a single ground state of the system. It is also nice to find that our reasoning, which comes from thinking about line and surface tensions in the statistical mechanics of surfaces and edges, leads to an object discovered by lattice gauge theorists thinking about creating states with (non)deconfined quarks.

The next steps are to apply this formulation to problems other than those addressed or reviewed in this paper, and two come to mind. The first is using it to detect spin liquids in systems with Heisenberg symmetry where the gauge field involves valence bonds and there is a set of orthogonality issues to contend with. The second is to generalize the construction to non-abelian topological phases captured by the models constructed by Levin and Wen [29].

Acknowledgments

This work was supported in part by NSF grant numbers DMR-1006608 (SLS) and PHY05-51164 (RM at KITP, Santa Barbara, CA).

Appendix. Kitaev’s toric code

The model describing Kitaev’s toric code is obtained from the Ising gauge Hamiltonian, equation (1), straightforwardly. Firstly, set $\Gamma = J = 0$. Secondly, replace the matter term proportional to $\Gamma_M$ by utilizing the constraint, equation (2),

$$G_s \equiv 1 \implies \tau_s^x \equiv \prod_{l \in \partial} \sigma_l^z,$$

(A.1)

to obtain

$$-H_{TC} = K \sum_p \prod_{l \in \partial p} \sigma_l^z + \Gamma_M \sum_s \prod_{l \in \partial} \sigma_l^z,$$

(A.2)

which now acts in the subspace spanned solely by the link degrees of freedom.

One can see that the model is special by noting that all terms commute with all others. This allows the full spectrum to be readily obtained and one sees that the model exhibits topological order. This is most easily done before gauge fixing (equation (A.1)), in the basis of eigenstates of $\{\sigma_l^z\}$ and $\{\tau_s^x\}$. In this formulation, sites where $\tau_s^x = -1$ are occupied by ‘charges’ and links where $\sigma_l^z = -1$ carry ‘electric flux’: $G_s = 1$ is the $Z_2$ version of Gauss’s law. The dual, magnetic, flux passing through loops $C$ is measured by the Wilson loops $W[C] = \prod_{l \in C} \sigma_l^z$.

The ground states of $H_{TC}$ lack charges, $\tau_s^x = 1$, $\forall s$, and are characterized by flux expulsion, $\prod_{l \in \partial p} \sigma_l^z = 1$, $\forall p$.

The complete lack of gauge field fluctuations, diagnosed by the Wilson loop $W = 1$, signals a deconfined phase. Consistent with this ground state signature, the excited states contain non-interacting charges, free to sit at arbitrary locations.
A.1. Mapping to classical $Z_2$ gauge theory and $T > 0$ phase transition

We now turn to a computation of the partition function of Kitaev’s model at $T > 0$. As gauge and matter sectors of the theory only interact via the constraint, $G_s \equiv 1$, one can use the basis states $|\{\sigma_l\}\rangle \otimes |\{\tau_s\}\rangle$, where $\sigma_l^z|\{\sigma_l\}\rangle = \sigma_l^z|\{\sigma_l\}\rangle$ and $\tau_s^x|\{\tau_s\}\rangle = \tau_s^x|\{\tau_s\}\rangle$, to obtain

$$Z = Z_L \cdot Z_S,$$

(A.3)

where

$$Z_L = \sum_{\sigma_l = \pm 1} e^{\beta K \sum_p \prod_{l \in \partial p} \sigma_l}$$

(A.4)

is the partition function of the classical or discretized Euclidean lattice $Z_2$ gauge theory in $d$ dimensions and

$$Z_S = \sum_{\tau_s = \pm 1} \delta_{\Pi_1 \tau_1, 1} e^{\beta \Gamma_m \sum_s \tau_s}$$

(A.5)

is the partition function of independent spins in a magnetic field $\Gamma_m$ up to a single, global constraint that the number of flipped spins (in gauge language: charges) be even. All of the singularities in $Z$ thus are contained entirely in $Z_L$.

The finite-temperature phase structure of $Z_L$ is well known. In $d \geq 3$, it exhibits two phases. The deconfined phase, at large $\beta K$, is characterized by a perimeter law for the Wilson loop at large loop sizes. The confined phase, at small $\beta K$, exhibits an area law. By contrast, in $d = 2$, $Z_L$ exhibits only a confined phase, but with a confinement scale (area law coefficient) $A_0$ that diverges exponentially as $\beta K \to \infty$.

All of this ties in with our previous results on Ising gauge theories: (i) Kitaev’s model exhibits a $T > 0$ continuation of the $T = 0$ topological phase in $d \geq 3$ but not in $d = 2$, (ii) the presence of the topological phase can be detected by the (spatial) Wilson loop exhibiting a perimeter law.

A.2. Stability of topological order under perturbations of the toric code

It is often convenient to perform calculations and point-of-principle demonstrations for topological phases at a ‘Kitaev point’, where terms generically present in the Hamiltonian are set to vanish by hand, leaving behind a soluble model. This has some obvious drawbacks—for instance, a ‘zero-law’ for the Wilson loop, $W \equiv 1 = e^{-L}$, is obviously somewhat non-generic, like a vanishing rather than finite correlation length in a disordered system, and any result thus obtained needs to be supplemented at least by a discussion of its fate under perturbations.

We now show that the deconfined phase is characterized by a persistence of the $2^{N_{nc}}$-fold topological degeneracy that signals topological order, where $N_{nc}$ is the number of independent non-contractible loops. More precisely, it exhibits a cluster of states with splitting of $O(e^{-L})$ for a system of linear dimension $L$ which are separated from all other states by a gap of $O(L^0)$. To this end, we consider adding an arbitrary local perturbation $V$ to the toric code [30], where locality now implies that the range of variables coupled is strictly bounded, say by a distance $b$. For example, we may reintroduce the two terms in equation (1) that we had set to zero on our way from the $Z_2$ gauge theory to the toric code, although the argument to follow is completely general. The easy part is to notice that the topologically distinct, degenerate ground states of $H_0$
will not mix up to an order $L/b \sim O(L)$ because they differ by the creation of a vortex loop (or appropriate dimensional analogue) that has to be created and stretched across the system in as many applications of the perturbation. The less obvious part is that the states remain degenerate in energy up to that order.

To see this, observe that the energy shifts $\delta E$ for two ground states, $|\psi_0\rangle = |G\rangle$ (equation (18)) and $|\psi_1\rangle = \Phi^x|\psi_0\rangle$, differing by the flux through a non-contractible loop, $C_{nc}$, are sums of terms of the form

$$\delta E_0 = \langle \psi_0 | VPV \ldots | \psi_0 \rangle,$$

$$\delta E_1 = \langle \psi_0 | \Phi^x VPV \ldots \Phi^x | \psi_0 \rangle.$$  \hspace{1cm} (A.6) (A.7)

Here, $V$ is the local perturbation, $P = \left(1 - \sum_i |\psi_i\rangle \langle \psi_i|\right)^{-1} \prod_{\ell \in C_{nc}} \sigma^x_{\ell}$ projects onto the excited state manifold and $\Phi^x = \prod_{\ell \in C_{nc}} \sigma^x_{\ell}$ inserts a flux by operating on all links crossed by the non-contractible loop, $C_{nc}$.

Now, we claim that $\delta E_0 = \delta E_1$ to $o(L)$ in perturbation theory. To see this, note that we can permute one $\Phi^x$ through the operator products of $Vs$ and $Ps$ to annihilate the other. This is possible because (i) $\Phi^x$ commutes with both the numerator and the denominator of $P$ and (ii) any topologically equivalent deformation of the non-contractible loop along which the $\sigma^x$s act produces the same state. Thus at an order below $L/b \sim O(L)$ in perturbation theory, there is always a choice of loop that avoids the locations of all of the perturbing terms.

References

[1] Wen X G 2004 Quantum Field Theory of Many Body Systems (Oxford: Oxford University Press)
[2] Rajaraman R 2001 arXiv:cond-mat/0103366
[3] Wen X G and Niu Q 1990 Phys. Rev. B 41 9377
[4] Anderson P W 1987 Science 235 1196
[5] Moessner R and Sondhi S L 2001 Phys. Rev. Lett. 86 1881
[6] kitaev A Yu 2003 Ann. Phys. 303 2
[7] Nayak C, Simon S H, Stern A, Freedman M and Sarma S D 2008 Rev. Mod. Phys. 80 1083
[8] Hansson T H, Oganessyan V and Sondhi S L 2004 Ann. Phys. 313 497
[9] Henley C L 2010 Annu. Rev. Condens. Matter Phys. 1 179
[10] Fradkin E and Shenker S H 1979 Phys. Rev. D 19 3682
[11] Moessner R and Sondhi S L 2010 Phys. Rev. Lett. 105 166401
[12] Huse D A and Leibler S 1991 Phys. Rev. Lett. 66 437
[13] Fredenhagen K and Marcu M 1986 Phys. Rev. Lett. 56 223
[14] Fredenhagen K and Marcu M 1988 Nucl. Phys. Proc. Suppl. 4 352
[15] Senthil T and Fisher M P A 2000 Phys. Rev. B 62 7850
[16] Bravyi S, Hastings M and Michalakis S 2010 arXiv:1001.0344
[17] Trebst S, Werner P, Troyer M, Shtengel K and Nayak C 2007 Phys. Rev. Lett. 98 070602
[18] Castelnovo C and Chamon C 2008 Phys. Rev. B 78 155120
[19] Hastings M B and Wen X-G 2005 Phys. Rev. B 72 045141
[20] Kogut J B 1979 Rev. Mod. Phys. 51 659
[21] Genovese L, Gliozzi F, Rago A and Torrero C 2003 Nucl. Phys. Proc. Suppl. 119 894
[22] Coleman P 1984 Phys. Rev. B 29 3035
[23] Baskaran G, Zou Z and Anderson P W 1987 Solid State Commun. 63 973
[24] Arovas D P and Auerbach A 1988 Phys. Rev. B 38 316

New Journal of Physics 13 (2011) 025009 (http://www.njp.org/)
[25] Moessner R and Raman K 2008 arXiv:0809.3051
[26] Rokhsar D S and Kivelson S A 1988 Phys. Rev. Lett. 61 2376
[27] Balents L, Bartosch L, Burkov A, Sachdev S and Sengupta K 2005 Phys. Rev. B 71 144509
[28] Moessner R, Sondhi S L and Fradkin E 2001 Phys. Rev. B 65 024504
[29] Levin M A and Wen X-G 2005 Phys. Rev. B 71 045110
[30] For numerical studies of this issuesee Trebst S, Werner P, Troyer M, Shtengel K and Nayak C 2007 Phys. Rev. Lett. 98 070602
    Dusuel S, Kamfor M, Orus R, Schmidt K P and Vidal J 2010 arXiv:1012.1740
[31] Kitaev and Preskill 2006 Phys. Rev. Lett. 96 110404
[32] Levin and Wen 2006 Phys. Rev. Lett. 96 110405