

# Belief propagation algorithm for computing correlation functions in finite-temperature quantum many-body systems on loopy graphs

David Poulin<sup>1</sup> and Ersen Bilgin<sup>2</sup>

<sup>1</sup>Center for the Physics of Information, California Institute of Technology, Pasadena, California 91125, USA

<sup>2</sup>Physics Department, California Institute of Technology, Pasadena, California 91125, USA

(Received 23 October 2007; published 16 May 2008)

Belief propagation—a powerful heuristic method to solve inference problems involving a large number of random variables—was recently generalized to quantum theory. Like its classical counterpart, this algorithm is exact on trees when the appropriate independence conditions are met and is expected to provide reliable approximations when operated on loopy graphs. In this paper, we benchmark the performances of loopy quantum belief propagation (QBP) in the context of finite-temperature quantum many-body physics. Our results indicate that QBP provides reliable estimates of the high-temperature correlation function when the typical loop size in the graph is large. As such, it is suitable, e.g., for the study of quantum spin glasses on Bethe lattices and the decoding of sparse quantum error correction codes.

DOI: [10.1103/PhysRevA.77.052318](https://doi.org/10.1103/PhysRevA.77.052318)

PACS number(s): 03.67.Lx, 05.50.+q, 61.43.-j, 89.75.Hc

## I. INTRODUCTION

Belief propagation is a powerful algorithm designed to solve inference problems involving a large number of random variables. It operates on graphical models, where variables are located at the vertices of a graph and edges encode dependence relations between the variables. The algorithm is exact when the underlying graph is a tree; but most importantly, it performs remarkably well in circumstances where it is not proven to converge to the exact solutions, i.e., when the graphical model contains loops. It is also highly parallelizable in the sense that each random variable can be associated with a different processor, and messages are exchanged between processors that are joined by an edge [1–4].

These features have made belief propagation an important tool in numerous scientific and technological fields ranging from information theory to image recognition, and from artificial intelligence to statistical physics. Indeed, it is one of the most powerful heuristic algorithms to solve problems such as decoding of low-density and turbo error correction codes [5–7], determining the phase diagram of quenched disordered systems [3,8], and random satisfiability problems [9,10].

Recently, belief propagation and graphical models were generalized to the quantum setting [11,12]. In this paper, we characterize the performances of quantum belief propagation (QBP) when used as a heuristic algorithm to solve inference problems—e.g., compute correlation functions—in the context of finite-temperature quantum many-body physics.

## II. GRAPHICAL MODELS

We consider quantum graphical models  $(G, \rho)$  that consist of a graph  $G$  and an  $n$ -bifactor state  $\rho$ . The graph  $G=(V, E)$  has a set of vertices  $V$  and a set of edges  $E$ . Each  $v \in V$  is a quantum system, with Hilbert space  $\mathcal{H}_v$ . A  $n$ -bifactor state  $\rho$  is a positive operator on  $\mathcal{H} = \otimes_{v \in V} \mathcal{H}_v$  that can be expressed as

$$\rho = \frac{1}{Z} \left( \prod_{v \in V} \mu_v \right) \star^{(n)} \left( \bigodot_{(u,v) \in E} \nu_{u:v} \right), \quad (1)$$

where  $Z$  is some normalization factor, and  $\mu_v$  and  $\nu_{u:v}$  are positive operators on  $\mathcal{H}_u$  and  $\mathcal{H}_u \otimes \mathcal{H}_v$ , respectively. The operators  $\nu_{u:v}$  are required to mutually commute when  $n$  is finite. The product  $\star^{(n)}$  is defined as  $X \star^{(n)} Y \equiv [X^{1/2n} Y^{1/n} X^{1/2n}]^n$ , and has the property of producing a positive operator when both  $X$  and  $Y$  are positive. This product is noncommutative except in the limit  $n \rightarrow \infty$ , which defines the  $\odot$  product:  $X \odot Y \equiv \lim_{n \rightarrow \infty} X \star^{(n)} Y = e^{(\log X + \log Y)}$ . Both products  $\star^{(n)}$  and  $\odot$  reduce to normal matrix product when  $X$  and  $Y$  commute.

A generic inference problem on a graphical model is to compute the reduced density operator on a subset  $W \subset V$  of the quantum systems conditioned on the fact that a measurement was performed on a disjoint subset  $U \subset V$ , where both  $W$  and  $U$  are of constant size. This problem turns out to be equivalent to the seemingly simpler problem of computing the reduced state on any subset  $W$  of constant size, i.e.,  $\rho_W = \text{Tr}_{V-W} \{\rho\}$ , where  $\text{Tr}_X$  denotes the partial trace over a systems in set  $X$ . Without additional assumptions on the structure of  $\rho$ , solving this problem requires resources that grow exponentially with the number of quantum systems  $|V|$ . However, the solution can sometimes be obtained or approximated by QBP in a time polynomial in  $|V|$ .

## III. QBP ALGORITHM

Given a graphical model  $(G, \rho)$ , QBP consists of a sequence of exchanges of operator-valued messages between neighboring vertices, which carry information about the state at other locations in the graph. More precisely, for  $(u, v) \in E$ , the message passed from vertex  $u$  to vertex  $v$  at time  $t$  is an operator on  $\mathcal{H}_v$  given by

$$m_{u \rightarrow v}(t) \propto \text{Tr}_u \left\{ \mu_u \star^{(n)} \left( \nu_{u:v} \odot \bigodot_{v' \in n(u)-v} m_{v' \rightarrow u}(t-1) \right) \right\}, \quad (2)$$

where  $n(u)$  denotes the neighbors of  $u$ . The proportionality factor can be chosen so that  $\text{Tr}\{m_{u \rightarrow v}\} = 1$ , and the messages

are initialized  $m_{u \rightarrow v}(0) = I$ . At time  $t$  the belief  $b_{uv}(t)$ —which is meant to represent some approximation of the state  $\rho_{uv} = \text{Tr}_{V-uv}\{\rho\}$  for  $(u, v) \in E$ —is given by

$$b_{uv}(t) \propto (\mu_u \mu_v) \star^{(n)} \left( \nu_{u:v} \odot \bigcirc_{w \in n(u)-v} m_{w \rightarrow u} \bigcirc_{y \in n(v)-u} m_{y \rightarrow v} \right), \quad (3)$$

where all messages are taken at time  $t$ . When all operators defining the bifactor state commute, QBP reduces to the standard belief propagation algorithm [1–4].

Since the message update rule Eq. (2) at vertex  $u$  depends only on the incoming messages at that vertex, the algorithm can be operated in a highly parallel fashion where each quantum system  $u$  is associated with a processor, and messages are exchanged between processors  $u$  and  $v$  iff  $(u, v) \in E$ . Similarly, the beliefs on the pair  $(u, v)$ , Eq. (3), can be computed by combining the messages received at those vertices.

### Convergence

In [12], it was shown that when  $G$  is a tree and  $(G, \rho)$  is either (i) a 1-bifactor state [cf. Eq. (1) with  $n=1$ ] or (ii) a quantum Markov network, QBP yields the exact solution in a time proportional to the graph's diameter—i.e.,  $b_{uv}(t) = \rho_{uv}$  for  $t \geq \text{diam}(G)$ . Intuitively, this means that the algorithm must run for a time sufficiently long to allow messages to travel between any pair of vertices. When operated on loopy graphs, the beliefs do not necessarily converge to the correct density operators. A good heuristic in that case is to halt the algorithm when  $b_{uv}(t)$  become almost time-independent, which also happens in a time roughly equal to the graph's diameter in all the models we have investigated.

A graphical model  $(G, \rho)$  is a quantum Markov network when the conditional independence conditions  $I(U: (V-n(U)-U)|n(U))=0$  are met for all  $U \subset V$ . The quantity  $I(A: B|C) = S(AC) + S(BC) - S(C) - S(ABC)$  is the quantum conditional mutual information [13, 14], and  $S(A) = \text{Tr}\{\rho_A \log_2 \rho_A\}$  is the von Neumann entropy. As explained in [12, 15], the vanishing of  $I(A: C|B)$  is equivalent to the condition  $\rho_{ABC} = \rho_B^{-1} \odot \rho_{AB} \odot \rho_{BC}$ . This equality is not verified in general, and the Kullback-Leibler distance between the right- and left-hand side is precisely the conditional mutual information  $D(\rho_{ABC} \| \rho_B^{-1} \odot \rho_{AB} \odot \rho_{BC}) = I(A: C|B)$ .

To understand the workings of QBP, consider a bifactor state  $\rho_{uvw}$  on the line  $u-v-w$ . The reduced state on  $w$  is  $\rho_w \propto \text{Tr}_{uv}\{(\mu_u \mu_v \mu_w) \star^{(n)}(\nu_{u:v} \odot \nu_{v:w})\}$ . When  $n=1$ , basic algebra implies that  $\rho_w \propto \text{Tr}_v\{(\mu_v \mu_w) \star^{(1)}(\text{Tr}_u\{\mu_u \star^{(1)} \nu_{u:v}\} \odot \nu_{v:w})\}$ ; the operations  $\text{Tr}_u$  and  $\star^{(1)}$  commute so to say. The computation of  $\rho_w$  can thus be broken into two steps: (i) Compute  $m_{u \rightarrow v} = \text{Tr}_u\{\mu_u \star^{(1)} \nu_{u:v}\}$  and (ii) compute  $\rho_w \propto \text{Tr}_v\{(\mu_v \mu_w) \star^{(1)}(m_{u \rightarrow v} \odot \nu_{v:w})\}$ . When  $n \rightarrow \infty$  on the other hand, the operations  $\text{Tr}_u$  and  $\odot$  do not commute in general, but they do precisely when  $I(u: v|w) = 0$  [12]. QBP is based on a generalization of these observations to arbitrary graphs.

QBP does not rely on the vanishing of the normalized connected correlation functions  $\mathcal{C}(\sigma_A, \sigma_C) = \langle \sigma_A \sigma_C \rangle - \langle \sigma_A \rangle \langle \sigma_C \rangle$  [16], or equivalently [17, 18] on the vanishing of the mutual information  $I(A: C) = S(A) + S(C) - S(AC)$ . In many systems, the mutual informa-

tion is not *a priori* short range. For instance, in the  $T \rightarrow 0$  limit, the thermal state of the  $1-d$  Ising model in zero transverse field is an equal mixture of all spins up and all spins down, which has  $I(A: B) = 1$  between any two disjoint regions, whereas  $I(A: C|B) = 0$  for any three disjoint regions. To compute thermodynamical quantities, one generally introduces a symmetry-breaking field that randomly singles out either the all-up or all-down state, which both have  $I(A: B) = 0$ . Symmetry breaking can be a delicate issue—for instance, on Cayley trees where a constant fraction of vertices live on the boundary [8]—and is circumvented by QBP.

## IV. QBP FOR QUANTUM MANY BODY

In the context of quantum many-body physics, the inference problem consists of computing correlation functions for the thermal state of a system of interacting particles. Given a graph  $G = (V, E)$ , we consider the generic Hamiltonian

$$H = \sum_{v \in V} h_v + \sum_{(u,v) \in E} h_{uv}. \quad (4)$$

The thermal state at inverse temperature  $\beta = 1/T$  is given by  $\rho = \frac{1}{Z} e^{-\beta H}$  where  $Z = \text{Tr}\{e^{-\beta H}\}$  is the partition function. Defining  $\mu_v = e^{-\beta h_v}$  and  $\nu_{u:v} = e^{-\beta h_{uv}}$  enables us to express any such thermal state as an  $\infty$ -bifactor state, cf. Eq. (1).

Despite the fact that thermal states are bifactor states, the result from [12] cited above does not imply that correlation functions can be evaluated exactly and efficiently with QBP. This is primarily because  $G$  is not necessarily a tree, but also because thermal states are neither 1-bifactor nor quantum Markov networks in general. There is no general remedy to the first hurdle, unless the loops happen to be very small and can be eliminated by merging some vertices. Thus QBP will need to be executed on a loopy graph and it is the primary goal of this paper to determine the effects of such loops on the performance of QBP. Two pragmatic solutions, named the replica method and sliding window QBP, have been proposed to overcome the second set of obstacles [12].

### A. Replica

The general idea of the replica method is to approximate the thermal state by a 1-bifactor state on which QBP can be executed directly and is guaranteed to converge in the absence of loops. In a first step, a Trotter-Suzuki (TS) decomposition is used to approximate a thermal state by an  $N_\tau$ -bifactor state with finite  $N_\tau$ . This produces a systematic error that scales as  $\beta/N_\tau$ . Then, in a fashion reminiscent of the replica trick used in the study of spin glass, the  $N_\tau$ -bifactor state is replaced by a 1-bifactor state at the expense of substituting the quantum system at each vertex by  $N_\tau$  replicas:

$$\mu_v \rightarrow (\mu_v^{1/N_\tau})^{\otimes N_\tau} T_v^{(N_\tau)} \quad \text{and} \quad \nu_{u:v} \rightarrow (\nu_{u:v}^{1/N_\tau})^{\otimes N_\tau}, \quad (5)$$

where  $T_v^{(N_\tau)}$  is the operator that cyclicly permutes the  $N_\tau$  replicas of  $v$ . The operators  $\nu_{u:v} = e^{-\beta h_{uv}}$  do not commute in general, but this can be fixed in practice on sparse graphs by merging some vertices. On a tree, the TS decomposition is the only source of error, so accuracy  $\epsilon$  can be achieved at a

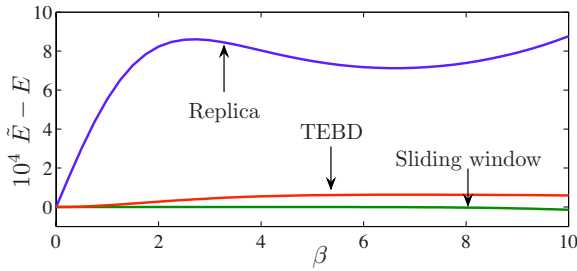


FIG. 1. (Color online) Critical Ising model on infinite line. Energy density estimate  $\tilde{E}$  using the method of replicas with  $N_\tau=10$ , sliding window QBP with  $\ell=6$ , and TEBD with  $\chi=150$ , compared to the exact energy density  $E$  obtained from fermionization.

computational cost that is exponential in  $\beta/\epsilon$ . This method is particularly useful as it allows for a direct computation of correlation functions at arbitrary distances, see [12].

### B. Sliding window

While all quantum Markov networks are thermal states of some local Hamiltonian on  $G$  [12], the converse is not true in general. Sliding window QBP is motivated by the fact that quantum Markov networks are fixed points of coarse graining procedures. Thermal states, regarded as  $\infty$ -bifactor states, are used directly to implement the message passing rule in Eq. (2) with  $n=\infty$ , except that messages are computed not just using the nearest neighbors but with all vertices within a distance  $\leq \ell$ . On a line, for instance, vertex  $j$  receives a message from  $(j-1, j-2, \dots, j-\ell)$  and one from  $(j+1, j+2, \dots, j+\ell)$ . In that case, sliding window QBP produces the exact solution efficiently if the conditional mutual information dies off at a finite distance.

## V. NUMERICAL RESULTS

We have numerically implemented the QBP algorithm on various graphs for the Ising and Heisenberg model whose Hamiltonians are

$$H_I = \sum_{v \in V} \vec{g} \cdot \vec{\sigma}_v + \sum_{(u,v) \in E} J_{uv} \sigma_u^z \sigma_v^z \quad \text{and} \quad H_H = \sum_{(u,v) \in E} \vec{\sigma}_u \cdot \vec{\sigma}_v, \quad (6)$$

respectively, and  $\vec{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$  are the usual Pauli matrices normalized so that  $\sigma^2 = 1/4$ . On a line, the homogeneous ( $J_{uv}=1$ ) Ising model has a zero temperature phase transition at the critical transverse field  $\vec{g} = (\frac{1}{2}, 0, 0)$ . Most of our simulations were performed at this critical value, as it is expected to represent the “hardest case.” Unless otherwise specified, it is henceforth assumed that  $\vec{g} = (\frac{1}{2}, 0, 0)$  and  $J_{uv}=1$ .

We used QBP to compute the energy density of the Ising model on an infinite line. This model can be solved exactly by means of a Jordan-Wigner transform that maps the interacting spin chain to a collection of free fermions [19]. Figure 1 shows the difference between the energy density computed with QBP and its exact value as a function of inverse temperature. Also shown are results obtained from a superoperator version of time-evolving block decimation (TEBD),

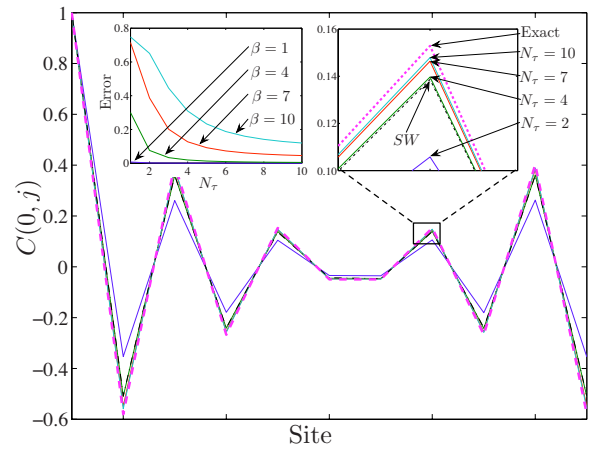


FIG. 2. (Color online) Correlations for  $H_I$  on a 11-site circle at  $\beta=6$ . Exact numerical solution (dashed), sliding window with  $\ell=5$  (dashed-dotted), and the replica method for various values of  $N_\tau$  (full). Left inset: Error Eq. (7) vs the  $N_\tau$  for different  $\beta$ .

which combines ideas from [20,21]. Since a line is a tree, the error in the results obtained from the replica method is entirely caused by the TS decomposition. The results obtained for sliding window QBP are in remarkably good agreement with the exact value, and can be systematically improved by increasing  $\ell$ . This reflects the fact that correlations are short-ranged in finite-temperature one-dimensional (1D) models. As expected the agreement improves for noncritical  $g$ . Results obtained for the Heisenberg model on the infinite line (not shown) are similar in all aspects.

To characterize the performance of QBP on more general graphs, we restrict our attention to systems with less than 12 spins, allowing comparison to direct brute-force numerical solutions. Figure 2 shows the correlation function  $C(0,j) = \text{Tr}\{\sigma_0^z \sigma_j^z \rho\}$  for  $H_I$  on a frustrated 11-site circle. We assess the quality of the approximation  $\tilde{C}$  to the exact correlation  $C$  by the average relative error

$$(\text{Error}) = \frac{\sum_j |C(0,j) - \tilde{C}(0,j)|}{\sum_j |C(0,j)|}. \quad (7)$$

Sliding window is again in very good agreement with the exact value for a relatively small window size. For the values of  $N_\tau \leq 10$  accessible with modest computational resources, the replica QBP reproduces the exact correlation function within a few percents at sufficiently high temperatures  $\beta \lesssim 6$ , which is consistent with the systematic error due to the TS decomposition.

Indeed, both the TS decomposition and the loopy QBP contribute to the total error Eq. (7). By brute force computation, it is possible to determine exactly what fraction of the error is caused by each of these approximations, and in almost all cases we have studied at critical  $g$ , both contributions were comparable. Figure 3(a) shows each contribution to the total error as a function of the transverse field  $\vec{g} = (g, 0, 0)$ .

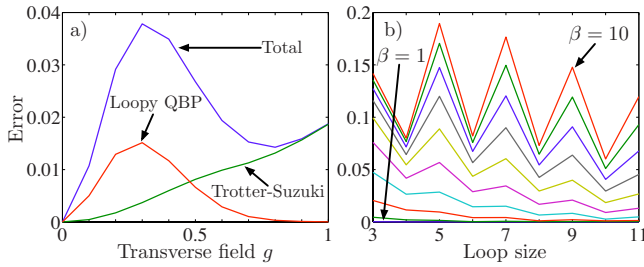


FIG. 3. (Color online) Replica method with  $N_\tau=10$ . (a) Different contributions to the error Eq. (7) vs transverse field for  $H_I$  on a 11-sites circle at  $\beta=6$ . (b) Error vs loop size, for  $\beta=1, 2, \dots, 10$ .

The most successful applications of the classical belief propagation algorithm are on graphs whose typical loop size is very large. This is the case, for instance, of low density parity check codes [6,7] and spin glasses on Bethe lattices [8]. Intuitively, one expects a local algorithm such as belief propagation to be relatively insensitive to the large-scale structure of the graph. We expect QBP to share this feature, and Fig. 3(b) illustrates the effect of the loop size on the average relative error of the correlation function. The oscillatory behavior of the error is explained by the frustration present in odd-size circles. Save from these oscillations, the results show a global improvement as the loop size increases. Errors obtained from the sliding window (not shown) also show a clear improvement as the loop size increases, but tend to have higher errors on even-size loops.

We have tested QBP on a variety of graphs depicted in Figs. 4(a)–4(d). The resulting errors in the correlation functions are shown in Fig. 4. The computational cost is slightly higher for the Heisenberg model because  $h_{uv}$  do not mutually commute. This restricts the computation to lower values of  $N_\tau$  and consequently yields larger errors. Modulo this difference, the error is most prominent for Figs. 4(c) and 4(d) which contain loops of size 3. In those cases, we found that the QBP algorithm was not converging: the magnitude of the errors is consistent with the magnitude of the time fluctuations of  $b_{uv}(t)$ , cf. Eq. (3). As expected, the predicted correlation function is in much better agreement with its exact value in Figs. 4(a) and 4(b) that have only relatively large loops.

## VI. CONCLUSION

We have numerically characterized the performance of the recently proposed QBP algorithm. In the high temperature phase, both the replica and the sliding window QBP algorithms perform remarkably well on a tree with modest com-

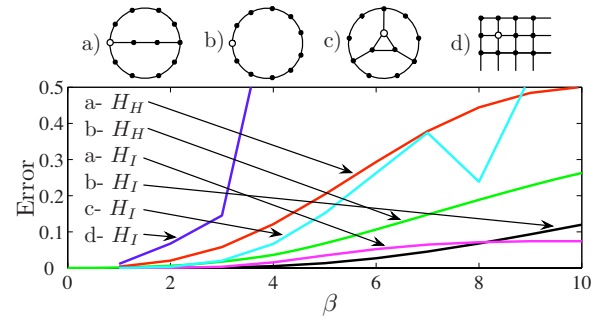


FIG. 4. (Color online) Error Eq. (7) for Ising and Heisenberg models on various loopy graphs using the replica method. For the Ising model  $N_\tau=10$  and for the Heisenberg model  $N_\tau=4$ . [(d) is a torus.]

putational resources, cf. Fig. 1, and offer performances similar to TEBD. On loopy graphs, we found that the algorithm gives reliable approximations when the loop size is large. Most importantly, when the results deviated from the exact value, e.g., in the presence of small loops, the algorithm did not reach a steady state, i.e., the beliefs Eq. (3) were highly fluctuating as a function of time. This provides an indirect way of assessing the validity of the results.

In [22], a technique similar to what we have called the replica method was used to investigate the phase diagram of quantum spin-glasses on Cayley trees. Based on the results we have presented, QBP should be suitable to study this phase diagram for more general Bethe lattices whose typical loop size scales as  $\log|V|$ . In the classical setting, it has been argued that the physics of random Bethe lattices and Cayley trees is greatly different [8]. We note that the randomness in quenched disordered systems should not affect the performances of QBP. In fact, our results obtained for random couplings  $J_{uv}$  and random local fields  $\vec{g}$  are typically in better agreement than the ones we have presented.

Finally, the low temperature phase of these models may be accessible using QBP as part of a variational approach based on projected entangled-pair states [23], which are a form of 1-bifactor states. QBP can be used to approximately compute the reduced state on pairs of sites and minimize their energy. We leave the characterization of this approach for a future study.

## ACKNOWLEDGMENTS

We thank Matt Leifer for stimulating discussions on graphical models and QBP. D.P. was supported in part by the Gordon and Betty Moore Foundation, by the NSF under Grant No. PHY-0456720, and by NSERC.

- [1] S. Aji and R. McEliece, *IEEE Trans. Inf. Theory* **46**, 325 (2000).  
 [2] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms* (Cambridge University Press, Cambridge, England, 2003).

- [3] M. Mézard and A. Montanari, *Constraint Satisfaction Networks in Physics and Computation* (Clarendon Press, Oxford, 2007).  
 [4] J. S. Yedidia, *Advanced Mean Field Methods: Theory and Practice* (MIT Press, Cambridge, MA, 2001), Chap. An idio-

- syncratic journey beyond mean field theory, p. 21.
- [5] C. Berrou, A. Glavieux, and P. Thitimajshima, in *ICC'93* (IEEE, Genève, Switzerland, 1993), pp. 1064–1070.
- [6] R. G. Gallager, *Low Density Parity Check Codes* (MIT Press, Cambridge, MA, 1963).
- [7] T. Richardson and R. Urbanke, *Modern Coding Theory* (Cambridge University Press, Cambridge, England, 2008).
- [8] M. Mézard and G. Parisi, *Eur. Phys. J. B* **20**, 217 (2001).
- [9] M. Mézard, G. Parisi, and R. Zecchina, *Science* **297**, 812 (2002).
- [10] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, *Nature (London)* **400**, 133 (1999).
- [11] M. B. Hastings, *Phys. Rev. B* **76**, 201102(R) (2007).
- [12] M. Leifer and D. Poulin, *Ann. Phys.* (in press).
- [13] P. Hayden, R. Jozsa, D. Petz, and A. Winter, *Commun. Math. Phys.* **246**, 359 (2004).
- [14] E. Lieb and M. Ruskai, *J. Math. Phys.* **14**, 1938 (1973).
- [15] M. B. Ruskai, *J. Math. Phys.* **43**, 4358 (2002).
- [16] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, e-print arXiv:0704.3906.
- [17] M. Fannes, *Commun. Math. Phys.* **31**, 291 (1973).
- [18] C. A. Fuchs and J. van de Graaf, *IEEE Trans. Inf. Theory* **45**, 1216 (1999).
- [19] P. Pfeuty, *Ann. Phys.* **57**, 79 (1970).
- [20] G. Vidal, *Phys. Rev. Lett.* **98**, 070201 (2007).
- [21] M. Zwolak and G. Vidal, *Phys. Rev. Lett.* **93**, 207205 (2004).
- [22] C. Laumann, A. Scardicchio, and S. Sondhi, e-print arXiv:0706.4391.
- [23] F. Verstraete and J. I. Cirac, e-print arXiv:cond-mat/0407066.