Maximum independent set on diluted triangular lattices

C. W. Fay IV, J. W. Liu, and P. M. Duxbury*
Dept. of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48823, USA
(Received 10 January 2006; published 16 May 2006)

Core percolation and maximum independent set on random graphs have recently been characterized using the methods of statistical physics. Here we present a statistical physics study of these problems on bond diluted triangular lattices. Core percolation critical behavior is found to be consistent with the standard percolation values, though there are strong finite size effects. A transfer matrix method is developed and applied to find accurate values of the density and degeneracy of the maximum independent set on lattices of limited width but large length. An extrapolation of these results to the infinite lattice limit yields high precision results, which are tabulated. These results are compared to results found using both vertex based and edge based local probability recursion algorithms, which have proven useful in the analysis of hard computational problems, such as the satisfiability problem.

DOI: 10.1103/PhysRevE.73.056112 PACS number(s): 64.60.Ak, 05.70.Jk, 61.43.Bn

I. INTRODUCTION

Relationships between statistical physics and computational complexity have been of interest for some time [1] and have recently proven fruitful [2], particularly the application of the cavity and replica methods to NP-complete problems [3–5] and extensions of computer science probability recursion methods, such as belief propagation, to glassy problems [3,6,7]. The broad NP-complete class of problems [8] lie at the nexus of interdisciplinary discussions, and includes problems such as the traveling salesman problem and the spin glass problem. Exact solvers for these NP-complete graph problems are often restricted to a few hundred nodes or less, which severely limits their utility. Quantum computing [9] offers hope that this new paradigm will significantly improve the efficiency with which NP-complete problems can be solved, though progress in algorithm development has been dissappointing.

Finding a maximum independent set (MIS), or the minimum vertex cover(MVC) to which it is trivially related, is one of the six fundamental NP-complete problems [8], and is NP-complete even on planar graphs. An independent set (IS) in a graph is a set of vertices such that no two sites share an edge. The MIS is an independent set that contains the maximum number of sites [8]. In statistical physics, an MIS corresponds to the maximum packing state of a hard core lattice gas [10]. The hard core lattice gas Hamiltonian for the MIS is

$$H = J\sum_{ij} \epsilon_{ij} n_i n_j - \mu \sum_i n_i, \tag{1}$$

where $n_i=1$ if a site is part of the MIS and $n_i=0$ if a site is part of the minimum vertex cover (MVC). In order to find the MIS, we take the limit $J\to\infty$ to ensure that no bond has an MIS site at both of it's ends and $\epsilon_{ij}=1$ if a bond exists between sites i and j. The chemical potential μ weighs the cardinality of an independent set. In order to find the MIS, we take the limit $\beta\mu\to\infty$, with $J/\mu\to\infty$ to ensure that the independent set (hard core) condition is preserved. We study MIS on diluted triangular lattices as a function of bond concentration p, which is then the dense packing density of the hard core lattice gas, Eq. (1), as a function of p. The average connectivity of a graph, c, is related to p through $p=c/z$, where z is the lattice co-ordination which is six for a triangular lattice.

Recently, statistical physics methods have been applied to the MIS problem on random graphs, which is a classic problem in graph theory [11]. Bauer and Golinelli [12] showed that a nontrivial percolation process called core percolation occurs at connectivity $c=e\approx 2.718$ and heralds the failure of the replica symmetric solution [5,10] for finding the MIS. They also suggested that though the core percolation exponents are close to those of conventional percolation, there are significant deviations. More recently replica symmetry-breaking calculations and other analytic methods have been introduced to go beyond the replica symmetric solution [13,14]. In a broader context local probability recursion (LOPR) algorithms [3,6,7] based on the cavity method and message passing algorithms have been developed for several problems in the NP-complete class.

In this contribution we analyse the MIS problem on bond diluted triangular lattices. First, we analyze the core percolation process on these lattices and find that the data is consistent with the conventional percolation universality class. We then study the MIS problem itself by carrying out transfer matrix calculations in strip geometries. These calculations yield accurate results for the average density of the MIS and also its degeneracy. We develop and apply LOPR algorithms to find approximate results for the MIS and compare them to the transfer matrix results and with a replica symmetric solution.

The paper is organized as follows. Section II introduces the methods that we use to analyze core percolation and the MIS on diluted triangular graphs. A numerical study of core percolation is presented in Sec. III, while in Sec. IV the results of transfer matrix and LOPR calculations of the MIS on diluted triangular lattices are presented. Section V contains a brief summary.

*Email: duxbury@pa.msu.edu

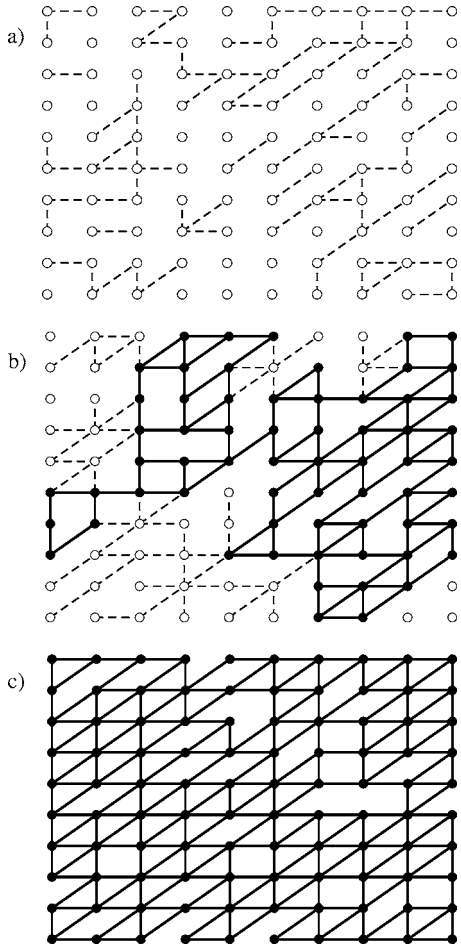

FIG. 1. 100 Node planar triangular graphs before and after core percolation. The dashed edges were originally present, but were removed during the leaf removal process. (a) Initial bond concentration of $c=1.0$ (i.e., $p=1/6$). Leaf removal removes all bonds from the graph. (b) Initial concentration of $c=3.0$ (i.e., $p=1/2$). After leaf removal, a percolating core remains; thus, $p>p_c$. (c) Initial concentration of $c=4.5$ (i.e., $p=3/4$). There are no singly connected bonds; thus, leaf removal is ineffective in reducing the graph.

II. METHODS

A. Iterative leaf removal

It is always possible to find an MIS on a graph by first placing all disconnected and singly connected vertices in the graph in the MIS [11,15]. Once a singly connected vertex has been placed in the MIS, we must exclude its connected neighbor from the MIS. The original singly connected site and its connected neighbor, and all of the edges incident on these two vertices, may then be removed from the graph [15]. The deletion of these vertices may produce a new singly connected site, which may then be treated in the same way. This process of iteratively removing singly connected vertices and their neighbors is called leaf removal, as illustrated in Fig. 1. Once leaf removal has been carried out to completion, the graph that remains is called the core of the graph. The iterative leaf-removal process can be carried out in polynomial time, so that if the core is small, the MIS can be found efficiently. The core percolation threshold is the bond concentration at which an extensive core cluster emerges (as illustrated in Fig. 1) and at this point finding the MIS on this extensive cluster becomes computationally hard, using conventional methods. In Sec. III, we present an analysis of the scaling behavior of the core percolation process on bond diluted triangular lattices.

B. Transfer matrix method

Finding the cardinality of a MIS is NP-complete [8] while counting MIS is in $\#P$ [16,17], even on planar graphs. The transfer matrix method we develop may be used to solve both the cardinality and counting variants of MIS on arbitrary graphs. To illustrate the method, we first solve these problems analytically on one-dimensional bond-diluted chains. We then extend the method to the triangular lattice, focusing on the strip geometry that is the most efficient for transfer matrix methods.

The transfer matrix (TM) method is a standard technique in statistical physics. It was used by Onsager to analyze the finite temperature phase transition of the two-dimensional Ising model [18]. To analyze NP-complete and #P-complete MIS problems we need to take the zero temperature limit of the usual transfer matrix of statistical physics. To do this exactly, we use the transfer matrix to construct a generating function $G(z)$, which keeps account of the number of independent sets with a given cardinality. Using the TM method, we find

$$G(z)=\sum_l a_l z^l, \tag{2}$$

where $z=e^{\beta\mu}$ and a_l is the degeneracy of an independent set and l is its cardinality. The MIS is the highest order term in the "independence polynomial" $G(z)$, and the prefactor of this term is the degeneracy of the MIS. The transfer matrix for a chain with Hamiltonian (1) is

$$T_i(n,n')=\exp[-\beta J\epsilon_i nn'+\beta\mu n] \tag{3}$$

where $n=0,1$, $n'=0,1$. $\epsilon_i=1$ if the bond between sites $i+1$ and i is present while $\epsilon_i=0$ otherwise. The transfer matrix T_i takes the calculation from site i to site $i+1$. The partition function for an L site chain is given by

$$Z(\beta,\mu,J)=\left\langle f\left|\prod_{i=1}^{L-1}T_i\right|1\right\rangle, \tag{4}$$

where $|1\rangle$ is determined by the boundary conditions at the end of the chain and the state $\langle f|$ is determined by the boundary conditions at the beginning of the chain. For a one-dimensional bond-diluted chain, only two matrices occur, corresponding to the cases $\epsilon_i=1$ and $\epsilon_i=0$. In appropriate limits, i.e., $\beta\mu\to\infty$, $J/\mu\to\infty$, and defining $z=Exp[\beta\mu]$ the two matrices are

$$T=\begin{pmatrix}1 & 1\\ z & 0\end{pmatrix} \tag{5}$$

for the case $\epsilon_i=1$ and

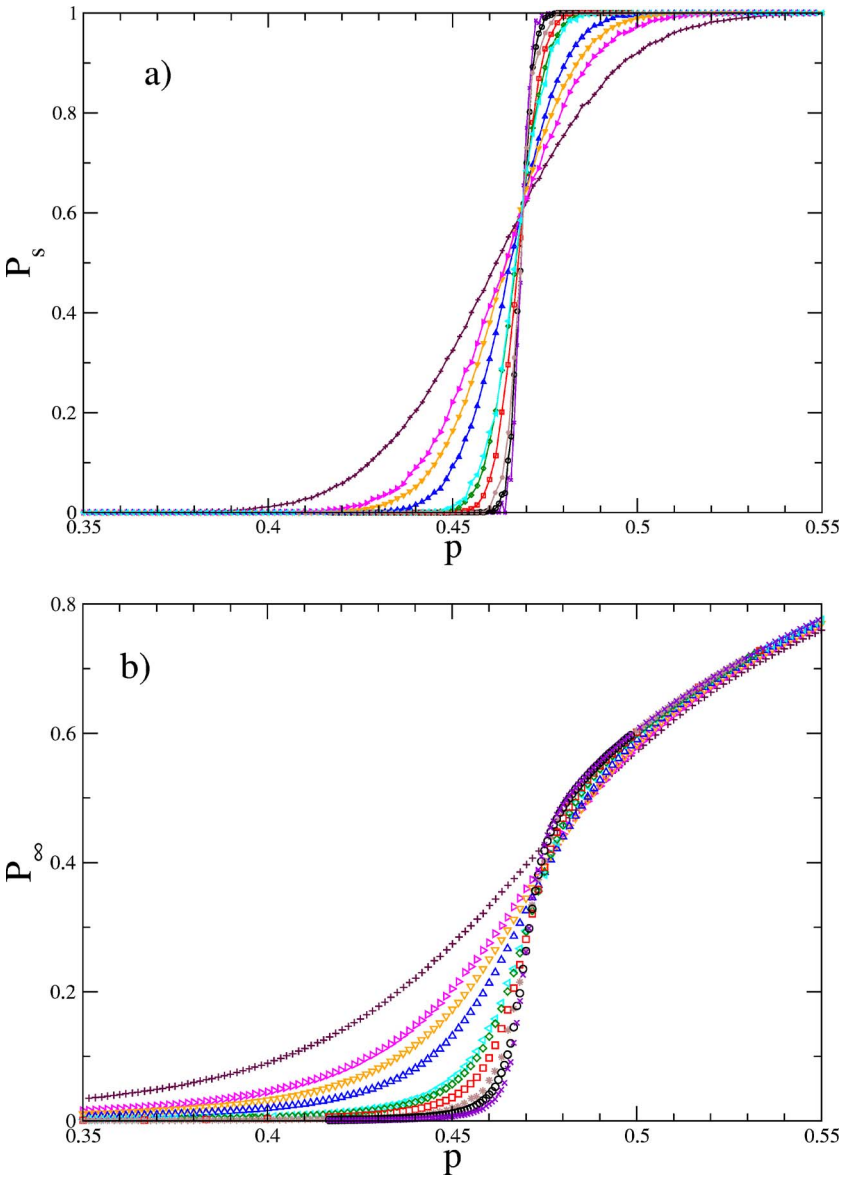

FIG. 2. (Color online) Plots of the raw data for core percolation on a bond diluted triangular lattice. (a) The spanning probability, P_s as a function of the bond concentration, p, for a range of values of number of sites in the graph, N. (b) The infinite cluster probability, P_∞ as a function of p for a range of values of the sample size N. In this figure the data for sample size (number of configurations) are as follows: $10^6(100)$, $5\times10^5(1000)$, $2.5\times10^5(100)$, $10^5(10000)$, $5\times10^4(10000)$, $4\times10^4(1000)$, $10^4(2\times10^4)$, $5000(10^4)$, $3025(10^4)$, and $1000(2\times10^4)$. The data are ordered according to increasing sample size, starting from the top, on the low concentration sides of the graph.

$$T' = \begin{pmatrix} 1 & 1 \\ z & z \end{pmatrix} \tag{6}$$

for $\epsilon_i=0$. The initial vector for a chain with free boundaries includes two possibilities: the last site is unoccupied (not part of MIS), and the last site occupied (part of MIS) so that

$$|1\rangle = (1,z). \tag{7}$$

The final state is $|f\rangle=(1,1)$ as we seek the maximum independent set over all possible configurations of the first site in the chain.

As an example consider a chain consisting of four connected sites. We then expect $G(z)=\langle f|T^3|1\rangle$. The generating function is found iteratively by acting on the state $|1\rangle$ with matrix T. For example the first iteration $T|1\rangle$ yields, $|2\rangle=(1+z,z)$. A key reduction in complexity occurs in searching for the MIS as in that case, we need only keep the highest-order terms in each element of $T|1\rangle$. Using this reduction, we find $|2\rangle=(z,z)$. This reduction preserves both the exact calculation of the MIS and its degeneracy. Continuing this procedure, we find, $|3\rangle=T|2\rangle=(2z,z^2)$, $|4\rangle=T|3\rangle=(z^2,2z^2)$. In carrying out this process, we find a generating function that has the correct highest-order term but that omits many of the lower-order terms. This reduced generating function, $G_R(z)$, for the case of four sites in a chain is

$$G_R(z) = \langle f|4\rangle = 3z^2. \tag{8}$$

From this we find that the four site chain has MIS of cardinality two, with degeneracy three. From the vector $|4\rangle$, we deduce that two of the three degenerate states corresponds to the site $m=4$ being part of the MIS, while one of the three degenerate states corresponds to the site $m=4$ *not* being part of the MIS. With this information it is possible to iterate backwards to find the three degenerate states: (1010), (0101),(1001).

In order to check the calculations of Sec. II C, it is useful to also calculate the average cardinality and entropy of diluted chains analytically. This is carried out as follows. For

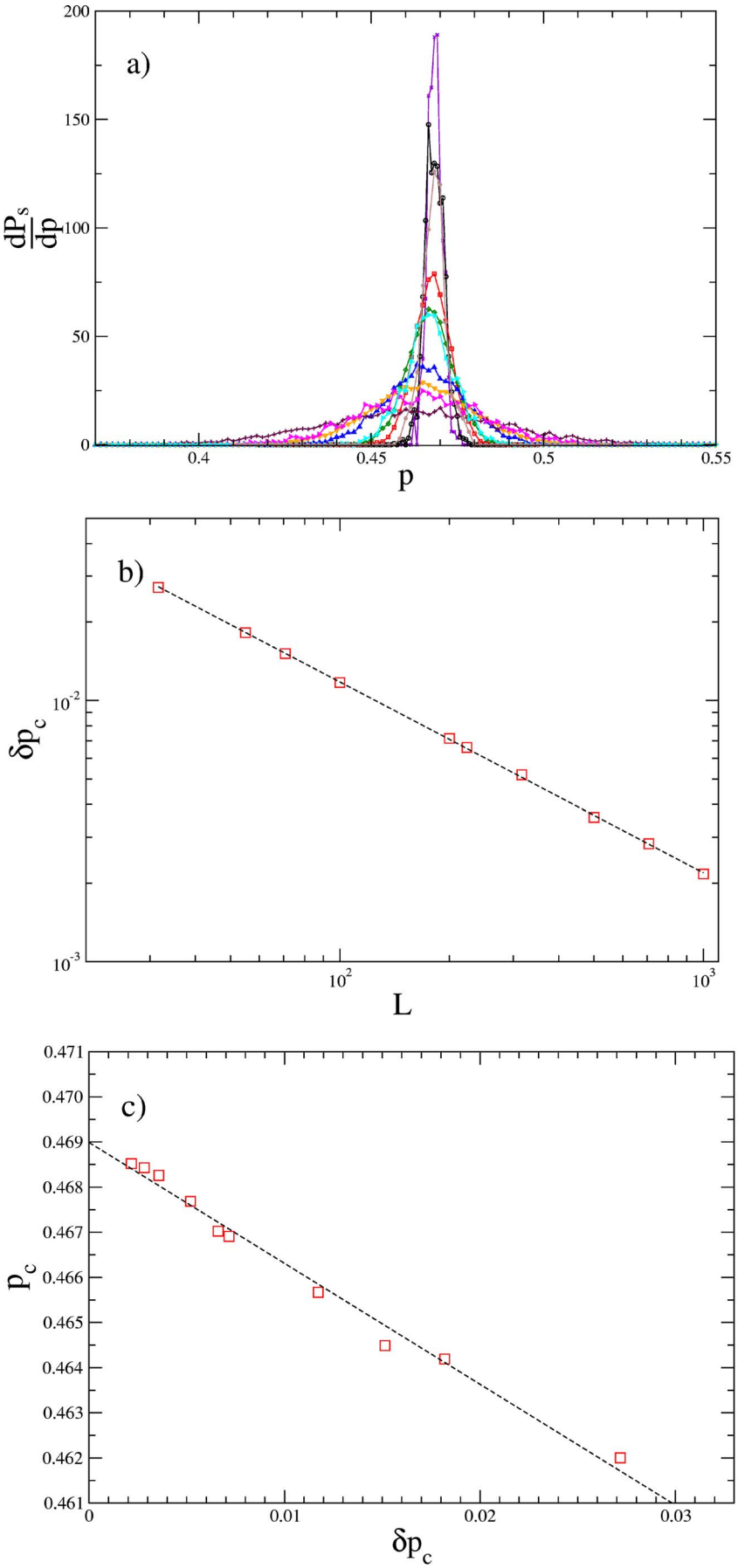

FIG. 3. (Color online) Analysis of spanning probability. (a) A plot of the quantity dP_s/dp as a function of p using the data of Fig. 2. (b) A plot of $\delta p_c^2 = \langle p_c^2 \rangle - \langle p_c \rangle^2$ as a function of lattice size L on a double logarithmic graph. The dashed line is a line of best to all of the data, from which we extract the estimate $\nu = 1.37(3)$. A fit to the largest five sample sizes yields $\nu = 1.34(2)$. (c) Plot of the average value of p_c as a function of δp_c. That is, for each sample size we find $p_c(L)$ and $\delta p_c(L)$ from the data in (a). Extrapolation to find the intercept of this graph yields an unbiased estimate of the percolation threshold and from this analysis we find, $p_c = 0.4690(4)$.

bond probability p, the number of times $n(s,p,L)$ a sequence of s connected sites occurs in a system containing L sites is

$$n(s,p,L) = Lp^{s-1}(1-p)^2. \tag{9}$$

The maximum independent set for s even and s odd are given by

$$i_s = s/2; \quad i_s = \frac{(s-1)}{2} + 1, \tag{10}$$

respectively. The degeneracies of the MIS in each of these cases is easily seen to be,

$$d_s = \frac{s}{2} + 1; \quad d_s = 1. \tag{11}$$

The average MIS for a bond-diluted linear chain is then

$$\langle i \rangle = \sum_{s=1}^{\infty} n(s,p,L) i_s = \frac{L}{(1+p)}. \tag{12}$$

The average entropy of the diluted chain is given by $\langle s \rangle = \langle \ln(d_s) \rangle$ so that

$$\langle s \rangle = (1-p)^2 L \sum_{s=1}^{\infty} p^{2s-1} \ln[s+1]. \tag{13}$$

Equations (12) and (13) are the average MIS and the average entropy for the bond-diluted chain.

In two-dimensional systems, the transfer matrix is studied in "strips," which consist of lattices of relatively small width W, but that have a large length L. This geometry is useful in exact calculations [18], as well as in numerical studies. The triangular lattice is amenable to this analysis and is an interesting MIS problem because it has odd loops, which lead to geometrical frustration. We construct a strip of the triangular lattice by taking a square lattice and placing a 45° diagonal edge across each square. The bonds in this lattice are then removed with probability $1-p$. We have considered both periodic and free boundary conditions in the width direction and used free boundaries in the length direction. With this geometry, the elements of the transfer matrix are

$$T_k(\boldsymbol{n},\boldsymbol{n}') = e^{-\beta J(\sum_{ij} \epsilon_{ij} n_i n_j + \sum'_{ij} \epsilon_{ij} n_i n'_j) + \beta\mu \sum_i n_i}. \tag{14}$$

The vectors $\boldsymbol{n}$ (corresponding to configurations at layer $k+1$) and $\boldsymbol{n}$' (corresponding to configurations at layer k) now have dimension 2^W. The first and last sum in Eq. (14) are over sites that are within the $(k+1)$th layer, whereas the second sum is over the edges which join sites in the kth layer to those in the $(k+1)$th layer.

The generating function is still of the form given in Eq. (4),

$$G(z) = \left\langle f \left| \prod_{k=1}^{m-1} T_k \right| 1 \right\rangle, \tag{15}$$

where for a given instance, the matrices T_k may be different. The final state is simply $\langle f| = (1,1,1\ldots)$, while the initial state depends on the bonds that are present in the first column of the lattice. If there are no bonds present, the initial state is most simply written as, $|1\rangle = \otimes_{k=1}^{l}(1,z)$, which is the direct product of the one-dimensional initial condition $(1,z)$. If there are bonds in the first column of the lattice, all states with two occupied sites at the end of a present bond have an entry of zero in the initial state. Section IV presents results for the MIS cardinality and degeneracy found using this transfer matrix method.

C. Local probability recursion (LOPR) algorithms

LOPR algorithms work with continuous probabilities at each site, instead of with the occupancy variables n_i, which are discrete, i.e., zero or one, in the hard core lattice gas. We define the probability P_i that a site is part of the MIS and $V_i = 1 - P_i$ is the probability that a site is not part of the MIS, which is equivalent to the probability that a site is part of the minimum vertex cover (MVC). If a lattice gas particle is present $P_i = 1$, while if the site is empty, $P_i = 0$. However, in the ensemble of possible ground states, there is the additional possibility that some sites are occupied in some ground states and unoccupied in others. LOPR algorithms allow this possibility by allowing the continuum of possibilities $0 \leqslant P_i \leqslant 1$. If $P_i = 0$, the site i is frozen vacant because there are no ground states in which this site is part of the MIS. Similarly if $P_i = 1$, site i is always part of the MIS. LOPR algorithms are an improved approximation as they take into account the possibility of degeneracy; however, they are not exact as they treat the correlations in the ground state using local procedures which may converge to a local optimum instead of the global optimum that we seek.

The simplest LOPR algorithm for maximum independent set is based on a simple update rule: A site is part of the MIS provided all of its connected neighbors are not part of the MIS, implying that

$$P_i = \prod_{j=1}^{v(i)} [1 - P_{n(j)}] \tag{16}$$

where i is the site which is being updated, $v(i)$ is the number of sites to which it is connected, and $n(i)$ is the set of neighboring sites. The vertex LOPR algorithm then consists of simple iteration of Eq. (16). The computational time required is then $O(N v_{\max} n_{it})$, where N is the number of nodes in the graph, $v_{\max}$ is the number of neighbors of the most highly connected node in the graph, and n_{it} is the number of sweeps of the lattice required for convergence of the site probabilities P_i. We find that n_{it} is at most a few thousand even for lattices of $N = 50\,000$ sites.

Our implementation of the vertex LOPR algorithm is as follows. We generate a graph and initialize the algorithm by assigning continuous random values of $0 \leqslant P_i \leqslant 1$ to each of the sites of the graph. We then sweep through all of the sites of the graph, in a randomized order, updating P_i at each site using Eq. (16). We find that after several hundred sweeps of the lattice, the LOPR procedure leads to a steady-state value of P_i on each site. For small lattice sizes, this procedure finds the exact MIS; however, for larger lattices and particularly beyond the "core percolation" threshold metastability is more likely.

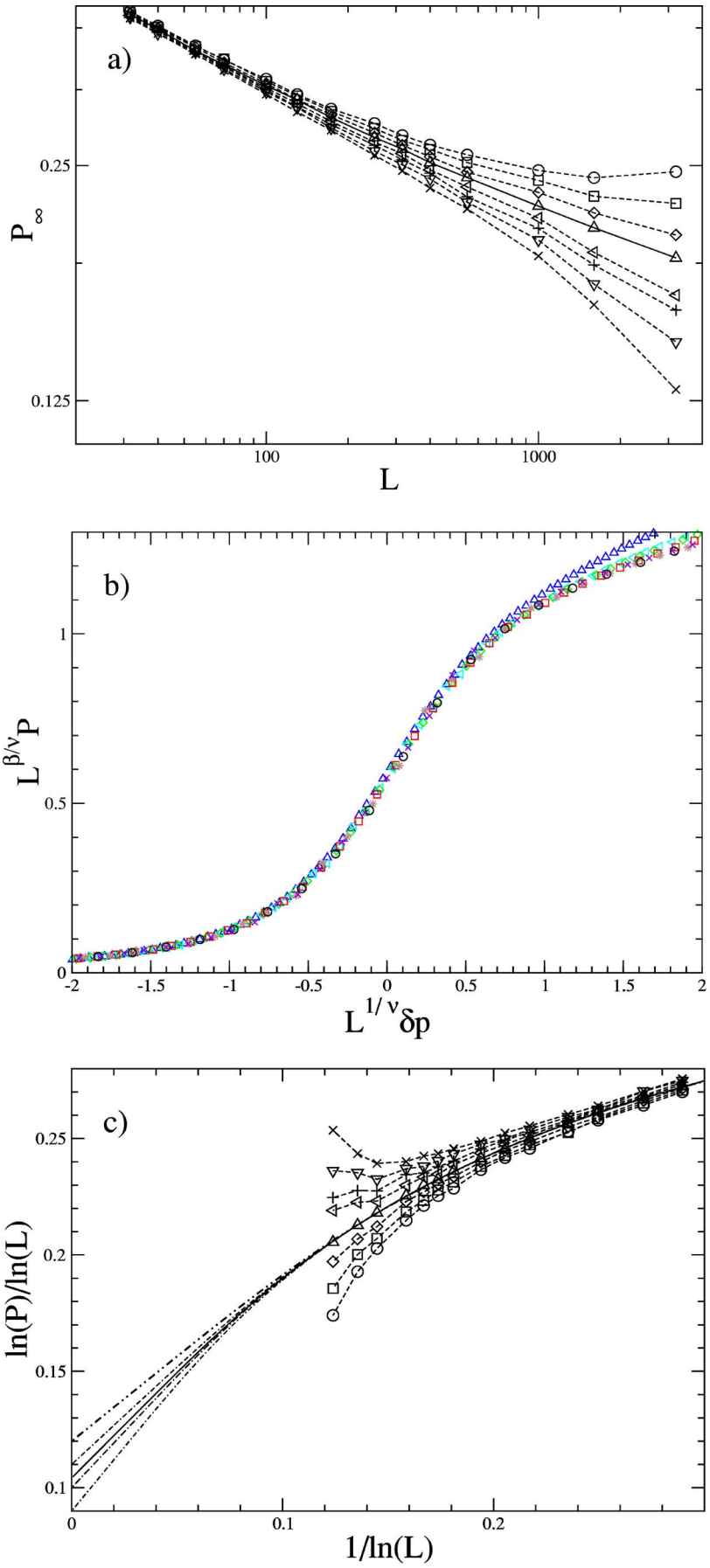

FIG. 4. (Color online) Analysis of infinite cluster probability; (a) A finite size scaling analysis for a range of closely spaced values of $p=c/z$, where $c=2.811, 2.812, \ldots, 2.818$, where the bottom curve is for $c=2.811$. From this data, we find that critical value of c is 2.815(2) (solid line in the figure), so that $p_c=0.4692(4)$, and $\beta/\nu=0.135(10)$ from fitting to the largest five sample sizes, without use of corrections to scaling; (b) A scaling plot of the infinite cluster probability. The best data collapse, as presented in the figure, was found for $p_c=0.4692$, $\beta/\nu=0.135$, $\nu=1.35$. (c) Analysis of the infinite cluster probability data using a next to leading order finite size correction, as given by Eq. (29), for $\beta/\nu=0.08, 0.1, 0.12, 0.14$ for data at $p_c=0.4692$.

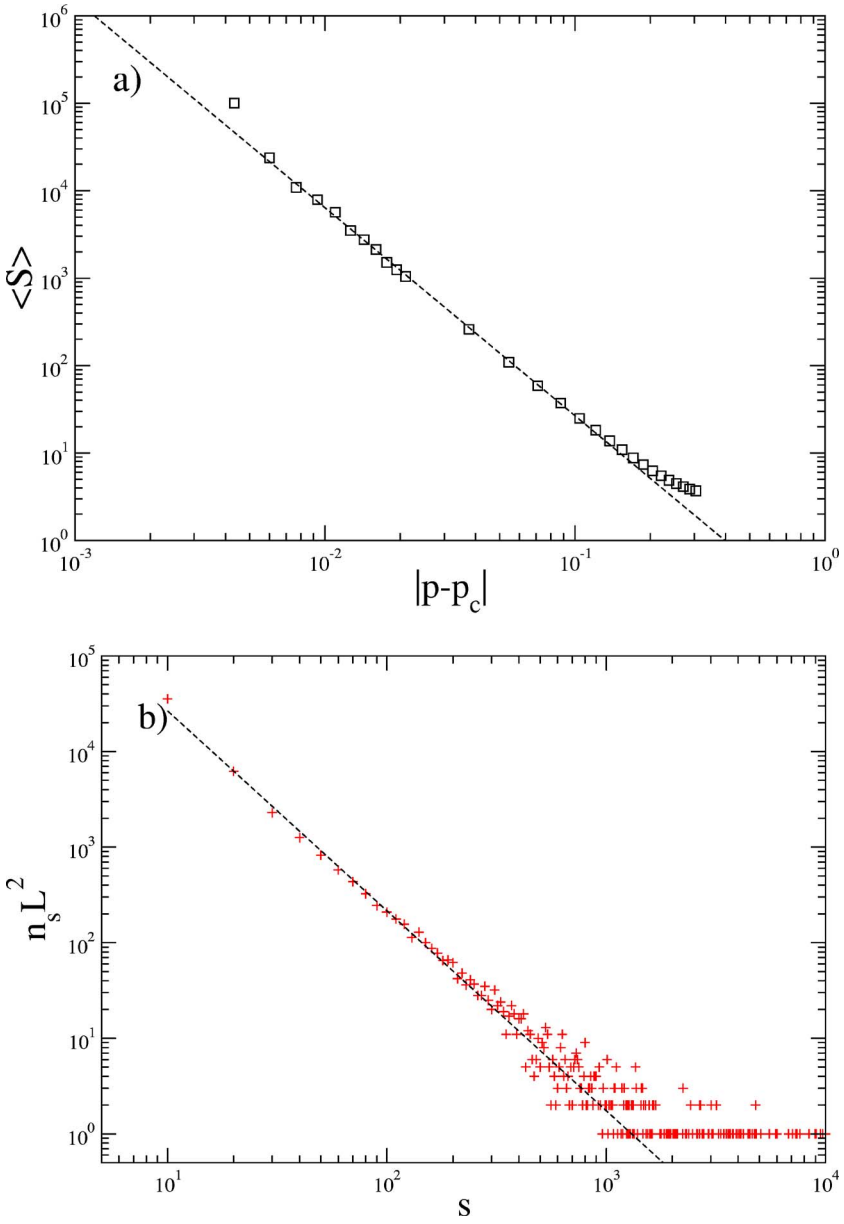

FIG. 5. (Color online) (a) A plot of the average cluster size on approach to p_c from below for lattices of size $N=10\,240\,000$ sites. From this plot we extract the exponent $\gamma=2.16(3)$. (b) A double logarithmic plot of the cluster size distribution at 0.4692 for a lattice of size $N=10\,240\,000$ sites. The dashed line has slope −2.09(5).

The vertex LOPR algorithm is the simplest in a series of probability-based algorithms. The next simplest possibility is a bond-based algorithm, where we introduce four bond probabilities, x_{00}^{ij}, x_{01}^{ij}, x_{10}^{ij},x_{D}^{ij}, which correspond, respectively, to occupancies 00, 01, 10, and the degenerate possibility 10 or 01 of the two sites i and j at the ends of the bond. Clearly, configuration 11 is not allowed for the MIS. This bond-based procedure is similar to algorithms introduced recently for KSAT [19,20]. To simplify the formula, we introduce,

$$y_{ij} = \prod_{k \epsilon n(i) \backslash j} (1 - P_k). \tag{17}$$

We then define the pair probabilities,

$$x_{00}^{ij} = (1 - y_{ij})(1 - y_{ji}), \quad x_{10}^{ij} = y_{ij}(1 - y_{ji}),$$

$$x_{01}^{ij} = (1 - y_{ij})y_{ji}, \quad x_{D}^{ij} = y_{ij}y_{ji},$$

where x_{D}^{ij} is the probability that the bond configuration is degenerate as it may be either 10 or 01. The probability that site i is part of the MIS is on average,

$$P_i = \frac{1}{v_i} \sum_{j}^{v_i} \left(x_{10}^{ij} + \frac{1}{2} x_{D}^{ij} \right) \tag{19}$$

In Sec. IV, results found using Eqs. (16)–(19) will be compared to similar calculations using the transfer matrix method described above.

The LOPR algorithms above allow the probabilities to vary from site to site in the graph. It is possible to build analytic approximations based on these algorithms by making a symmetric approximation. In the case of the site algorithm [Eq. (16)], the symmetric approximation is based on

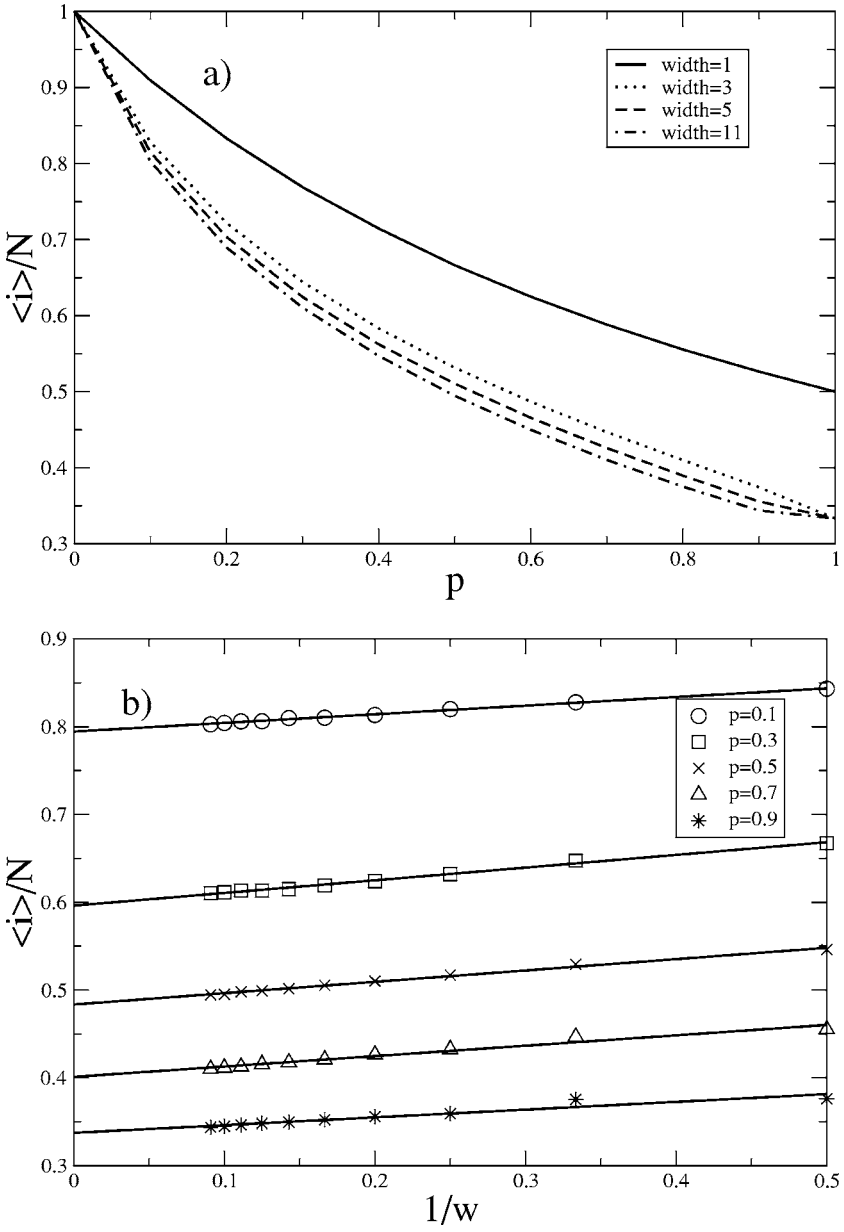

FIG. 6. The average cardinality of the MIS as a function of bond concentration p for strips of triangular lattices of varying width W and of length $L=10\,000$, with $N=LW$. (a) The MIS as a function of p for $W\leqslant 11$. (b) Illustration of the uniform and monotonic behavior as a function of system size $(1/W)$. Solid lines are linear fits to the data.

the recurrence formula for the probability that a site is certainly part of the MIS (i.e., that it is frozen and occupied), P_f, which is given by

$$P_f = (1 - pP_f)^{\alpha}, \tag{20}$$

for a Bethe lattice of coordination z and with $\alpha=z-1$. Similarly, the probability that a site is certainly part of the minimum vertex cover (i.e., frozen and unoccupied in the MIS) is given by

$$V_f = 1 - (1 - pP_f)^{\alpha} - \alpha p P_f (1 - pP_F)^{\alpha-1}. \tag{21}$$

Note that the second term in Eq. (21) is necessary because when one site at the lower level is uncovered, an edge connecting it to the next level up may be covered by covering either of its ends. If we assume that each end is part of the MIS with probability 1/2, we find that the average value of the MIS is given by

$$\langle P \rangle = P_f + \tfrac{1}{2} p \alpha P_f (1 - pP_F)^{\alpha-1}. \tag{22}$$

In the random graph limit, this expression reduces to the replica symmetric result

$$\langle P \rangle = \frac{W(c)}{c} + \frac{W(c)^2}{2c}, \tag{23}$$

where $W(c)=cP$ is the Lambert function and the solution to Eq. (20) in the random graph limit $N\to\infty$, with $p=c/N$ [5,10].

III. CORE PERCOLATION

As outlined in Sec. II A, the core of a graph is found by iterative leaf removal. In this section, we study the percolative properties of the core found in this way. The initial state is produced by random dilution of bonds from a triangular

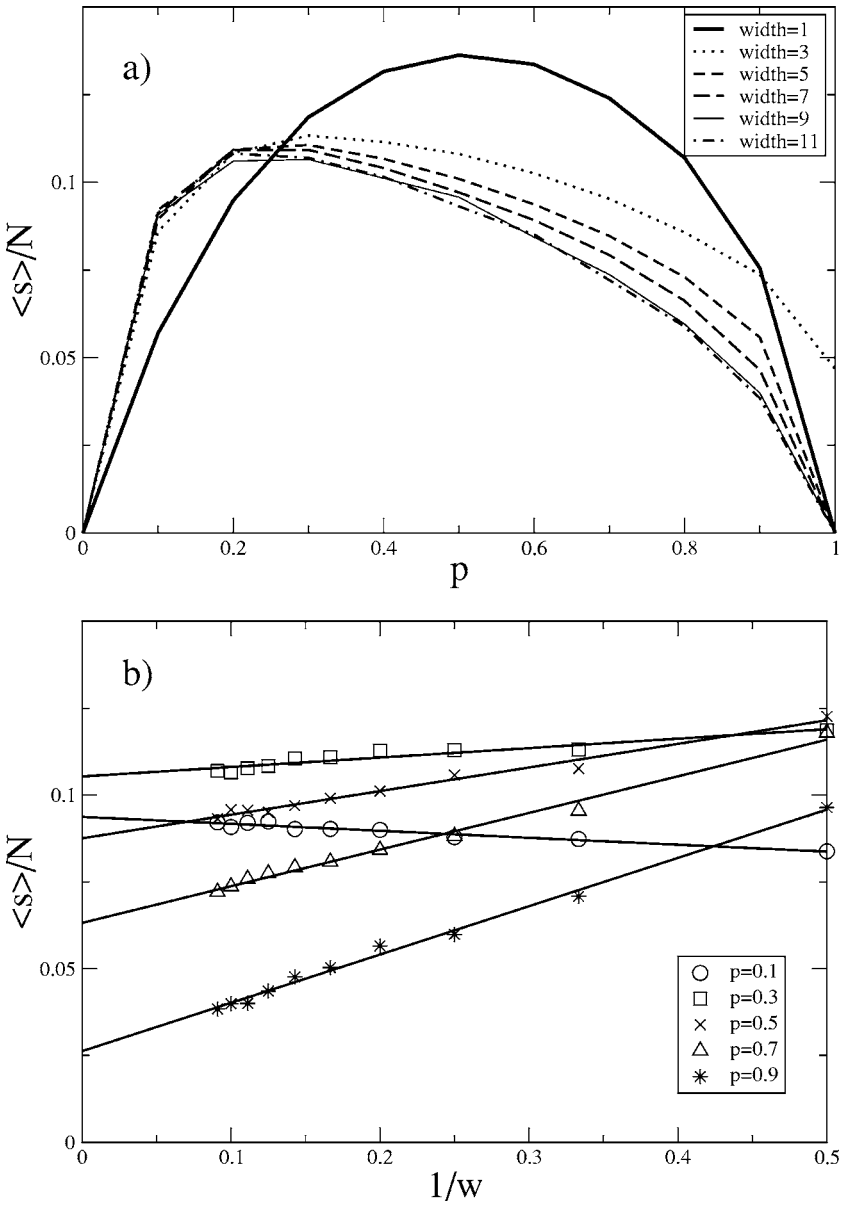

FIG. 7. The average ground state entropy per site of the MIS for strips of length $L=10\,000$. (a) Entropy per site as a function of bond concentration for strips of width $W \leqslant 11$. (b) Extrapolation as a function of $1/W$ to estimate the infinite lattice behavior. The data were generated using the same samples as that presented in Fig. 6.

lattice until a fraction p of the bonds remain. This corresponds to an average site connectivity of $c=6p$, for a triangular lattice. The scaling behavior of the core, which remains after leaf removal, is analyzed using well-established methods [21].

We focus on three percolative properties: (i) the spanning probability P_s, which is the probability that a core cluster spans the lattice; (ii) the infinite cluster probability P_∞, which is the probability that a site is part of the core cluster that spans the lattice; and (iii) core cluster numbers $n(s,p)$ and the average cluster size $\langle s \rangle$. Near the core percolation threshold p_c, if the transition is second order, the scaling behavior of these quantities is expected to be

$$P_s(p) \sim f_s(\delta p L^{1/\nu}) \tag{24}$$

where $\delta p = |p - p_c|$,

$$P_\infty(p) \sim L^{-\beta/\nu} f_\infty(\delta p L^{1/\nu}), \tag{25}$$

$$n(s,p) \sim s^{-\tau} f_n(s^\sigma \delta p) \quad L \to \infty, \tag{26}$$

and

$$\langle s \rangle \sim \delta p^{-\gamma} f_a(\delta p L^{1/\nu}), \tag{27}$$

where the scaling functions f_s, f_∞, f_n, f_a in Eqs. (24), (25), and (27) include the finite size scaling behaviors and enable determination of β, ν, and γ, while Eq. (26) is the scaling behavior in the infinite lattice limit and enables determination of τ, the cluster size exponent at p_c. Only two of the above exponents are independent, and a check on the calculations is provided by various exponent relations, for example,

$$\nu = \frac{\beta}{2}\frac{\tau - 1}{\tau - 2}, \quad \alpha + 2\beta + \gamma = 2, \quad d\nu = 2 - \alpha. \tag{28}$$

The standard two-dimensional (i.e., $d=2$) percolation values are, $\beta = 5/36(0.139)$, $\nu = 4/3(1.33)$, $\tau = 187/91(2.05)$, and γ

TABLE I. The MIS per site, $\langle i\rangle/N$, and entropy per site, $\langle s\rangle/N$, for 11 values of the bond concentration p of bond diluted triangular lattices. The results were found by linear extrapolation as a function of $1/W$ [see Figs. 6(b) and 7(b)] for strips of up to width $W=11$ and for strip length, $L=10\,000$.

p	$\langle i\rangle/N$	$\langle s\rangle/N$
0.0	1.0	0.0
0.1	0.7944	0.0937
0.2	0.6790	0.1071
0.3	0.5961	0.1053
0.4	0.5339	0.0977
0.5	0.4835	0.0875
0.6	0.4389	0.0766
0.7	0.4012	0.0632
0.8	0.3663	0.0496
0.9	0.3374	0.0262
1.0	0.3333	0.0

$=43/18(2.39)$, and the conventional bond percolation threshold on a triangular lattice occurs at $p_c=0.34729$ [21].

Raw data for core percolation on a triangular lattice are presented in Fig. 2, from which it is clear that the percolation threshold is significantly higher than the conventional percolation value.

A more refined analysis of the spanning probability data is presented in Fig. 3, where Fig. 3(a) is the derivative dP_s/dp, which is the probability that a spanning cluster first appears at p. The average value of p_c is then given by $\langle p_c\rangle = \int pdpdP_s/dp$ and its second moment is $\langle p_c^2\rangle = \int p^2dpdP_s/dp$. An unbiased estimate of the exponent ν is found by using, $\delta p_c^2=\langle p_c^2\rangle-\langle p_c\rangle^2\sim L^{-2/\nu}$. A plot of δp_c vs L is presented on a log-log plot in Fig. 3(b), from which we find that $\nu=1.34(2)$.

A determination of the value of the critical threshold is presented in Fig. 3(c), where we present data for the average value of p_c as a function of the variation in the value of p_c. This plot removes the dependence on the critical exponents and gives a less biased estimate of p_c, which is the intercept of this graph and from which we find that the core percolation threshold is given by $p_c=0.4690(4)$.

A double logarithmic plot of the finite size scaling behavior of the infinite cluster data is presented in Fig. 4(a) for a range of closely spaced values of the bond concentration p. For $p<p_c$, the infinite cluster probability approaches zero exponentially, while for $p>p_c$ it approaches a constant value exponentially. However at $p=p_c$ there is power-law scaling, with $P_\infty(p_c)\sim L^{-\beta/\nu}$. From the double logarithmic plot of Fig. 4(a), we find that power-law scaling occurs at the threshold value $p_c=0.4692(4)$, and $\beta/\nu=0.163$ from a fit to all samples sizes and a smaller value $\beta/\nu=0.135$ from a fit to the largest five sample sizes. Both of these values are significantly higher than the usual percolation value of 0.104; however, the decrease in β/ν with fits to larger L suggests that anomalous finite size corrections may be important. A further test of the scaling behavior of P_∞ is presented in the scaling plot in Fig. 4(b). A very good data collapse if found for $p_c=0.4692$, $\beta/\nu=0.135$, $\nu=1.35$, though considerable variation in these values is possible without affecting the quality of the fit significantly. The value of β determined from the analysis of Figs. 4(a) and 4(b) is significantly higher than the standard percolation value $\beta=5/36=0.139$. However, it is well known that determination of β in correlated percolation problems can be strongly affected by higher-order finite size scaling corrections [22]. An analysis of the data of Fig. 4(a) using an anomalous scaling correction [22,23] is presented in Fig. 4(c). We used the form

$$P_\infty = aL^{-\beta/\nu}(1+bL^{-w}) \tag{29}$$

so that $-\ln P_\infty/\ln L=\beta/\nu-\ln a/\ln L-bL^{-w}/\ln L$, where we have assumed that bL^{-w} is small. The value of the correction to scaling exponent determined from this analysis lies in the range $w=0.25\rightarrow0.42$ for $\beta/\nu=0.08\rightarrow0.14$. This is a relatively small correction to scaling exponent and provides understanding of why the β exponent we observe in a naive fit [Fig. 4(b)] is higher than expected. The data is certainly consistent with a conventional exponent $\beta/\nu=0.104$ but with a correction to scaling exponent $w=0.35(5)$.

An analysis of the cluster statistics is presented in Fig. 5. In figure 5(a), we present an analysis of the γ using data on a large lattice. The value of γ we find from a naive fit is significantly smaller than the conventional percolation value; however, it is consistent with the larger value of β, which occurs on these lattices sizes. The correction to scaling exponent, which affects β, also affects γ (e.g., via the exponent relation $\alpha+2\beta+\gamma=2$), and hence also may reconcile this discrepancy. In Fig. 5(b), a double logarithmic plot of a histogram of the cluster numbers at percolation is presented.

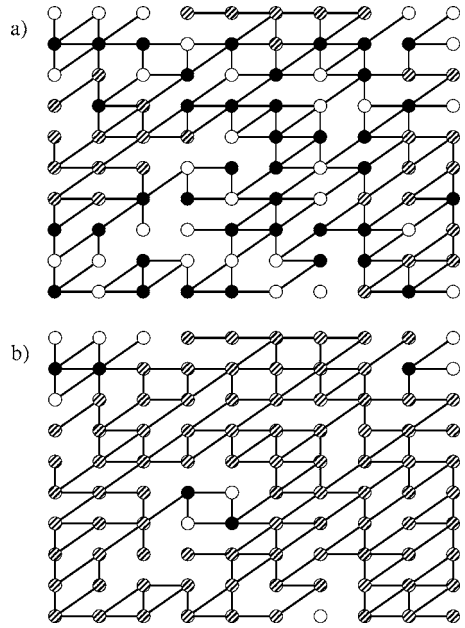

FIG. 8. The maximum independent set on bond diluted triangular lattices. (a) Site LOPR and (b) bond LOPR. Open circles are sites where $P_i=1$, filled circles have $P_i=0$ and the hatched sites have intermediate values of the probability P_i.

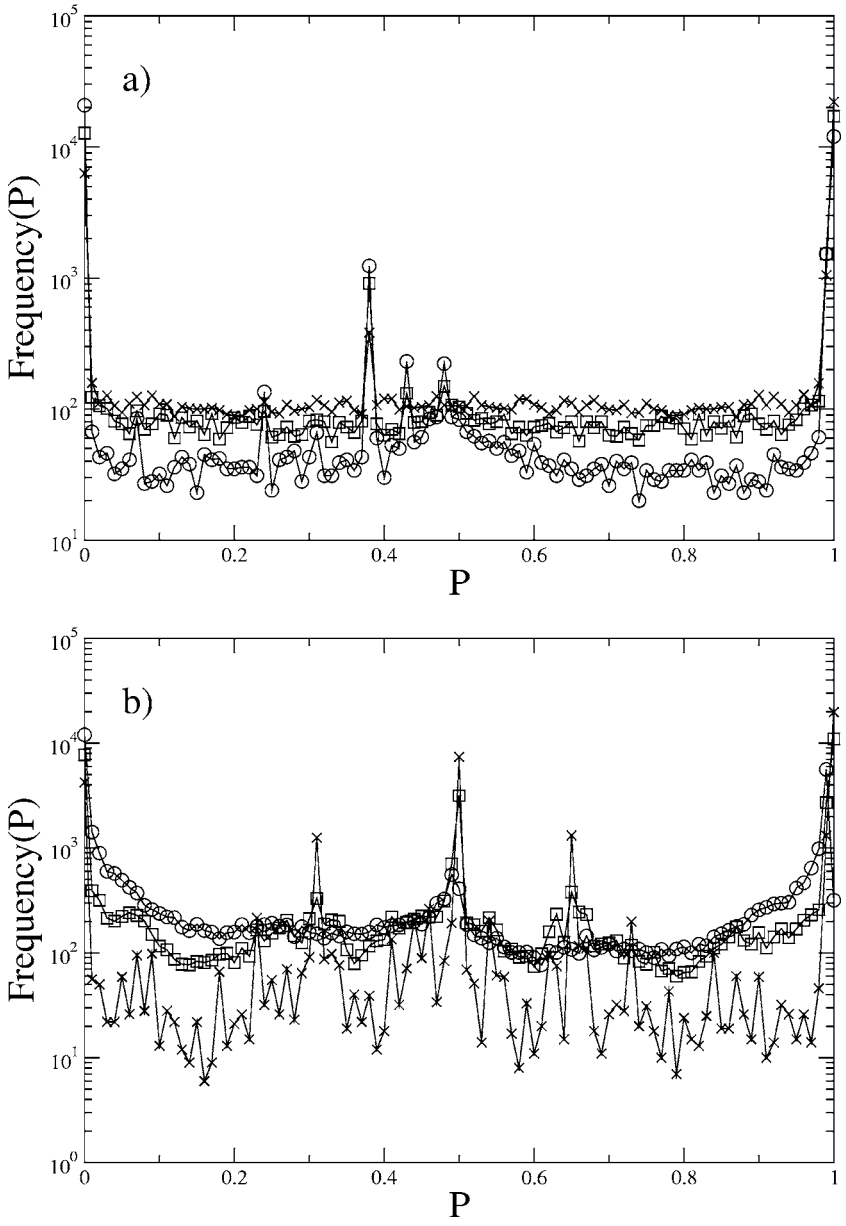

FIG. 9. The distribution of MIS probabilities *Frequency*(P), for $N=50,000$ site triangular lattices at $c=1.0$ ($\times$), $c=2.0$ ($\square$), and $c=4.0$($\circ$): (a) site LOPR and; (b) bond LOPR.

From this data, we find that $\tau=2.09(5)$, which is consistent with the conventional value.

From the analysis presented in Figs. 2–5, we have found that core percolation on bond-diluted triangular lattices occurs at bond concentration 0.4692(4), which is considerably higher than that of conventional percolation, which occurs at $p_c=0.34729$. The transition remains continuous and has the thermal exponent $\nu=1.34(2)$, which is consistent with the conventional value. The exponent τ for core percolation is also consistent with the conventional value; however, a naive analysis of the order parameter exponent β and the susceptibility exponent γ yield values that are considerably different from the conventional values. These deviations may be real in that the fractal dimension of the core may be different from the fractal dimension of the percolative infinite cluster, as occurs for the backbone in conventional percolation. However, the value of β found here is quite close to the conventional value so that the deviations can also be reconciled by the presence of strong corrections to scaling as indicated in Fig. 4(c) for β.

IV. DENSITY AND DEGENERACY OF THE MIS

In this section, we find the MIS for bond-diluted triangular lattices, first using the exact transfer matrix method and then using local probability recursion (LOPR) methods.

A. Transfer matrix results

The transfer matrix method gives exact results for strips of width W and length L, with a total number of sites $N=LW$. From Eqs. (14) and (15), it is evident that the computational complexity of this approach, for the triangular lattices studied here, scales as $L2^{2W}$, while the storage requirement increases as 2^{2W}. Because of this scaling, it is efficient to carry out calculations for long chains (large L) but of

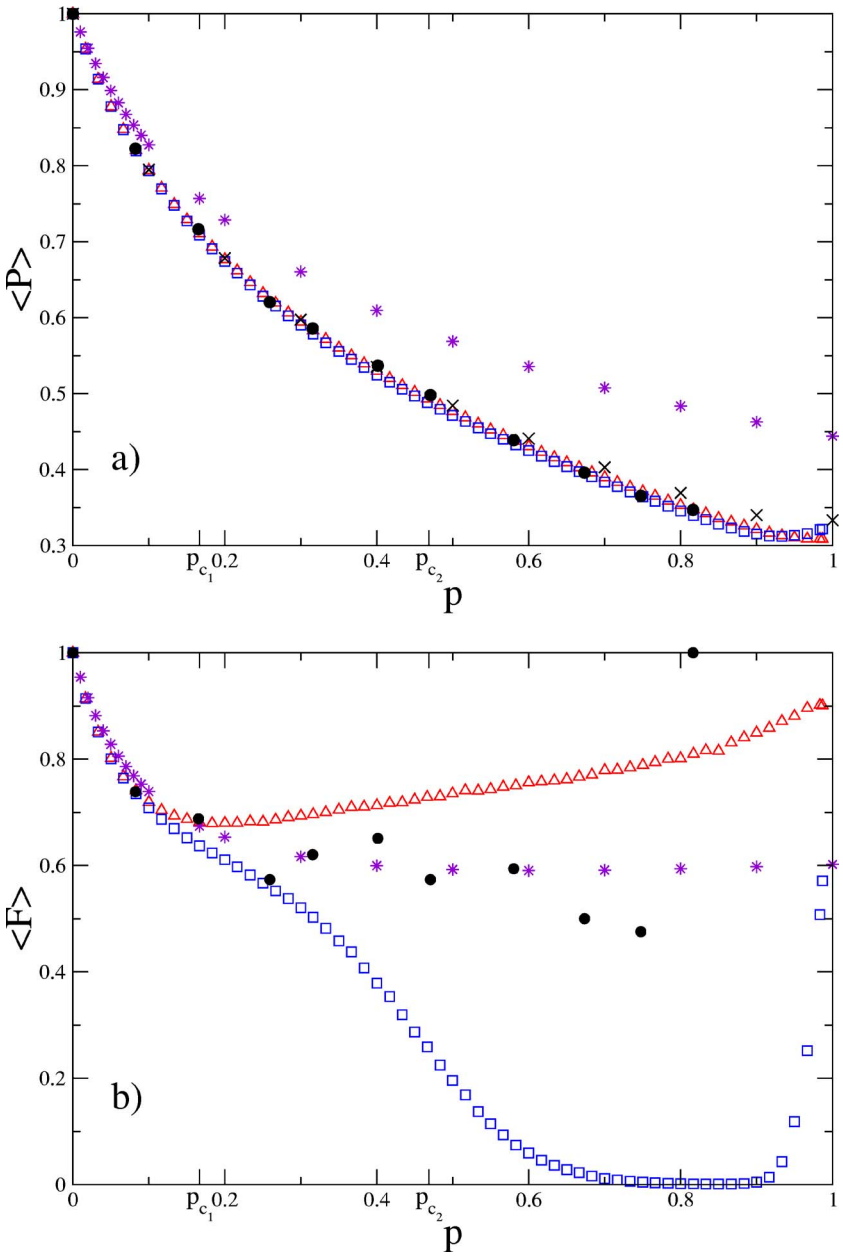

FIG. 10. (Color online) (a) the average MIS per site $\langle P \rangle$ and (b) the average number of frozen sites $\langle F \rangle$ for $N=10\,000$ bond-diluted triangular lattices. The data are for site LOPR (△), bond LOPR (□), transfer matrix results [in (a)] (×), the symmetric Bethe approximation (☆), and exact results on 49 node lattices (•) [10], averaged over 10 configurations. In this figure, p_{c1} is the connectivity percolation threshold while p_{c2} is the core percolation threshold.

limited width, W. However, we find that there is rapid convergence to the infinite lattice value as a function of W, so that accurate values of the average MIS and its degeneracy can be found by extrapolation from relatively small values of W.

Figures 6 and 7 present the results of the transfer matrix calculations of the average cardinality, $\langle i \rangle / N$, and the average entropy, $\langle s \rangle / N$ of the MIS for a range of values of W, for strips of length $L \gg W$. The entropy of the MIS is $\langle s \rangle = \ln(D)$, where D is the degeneracy of the MIS and is calculated exactly from the transfer matrix method for finite strips. It is evident that the transfer matrix results converge rapidly and uniformly to the infinite lattice limit and that a linear extrapolation to this limit is possible [see Figs. 6(b) and 7(b)]. One notable feature of Fig. 6(b) is that for small concentrations p the entropy of the MIS increases with increasing system size, whereas for larger p, the entropy decreases with increasing systems size. The change in behavior between these two regimes occurs for $p < p_c$.

The results of linear extrapolations of data, such as that presented in Figs. 6 and 7, are captured in Table I, for $L = 10\,000$. We found that for $L \gg W$, there is little dependence of either $\langle i \rangle / N$ or $\langle s \rangle / N$ on the strip length. The results presented in Table I are compared to approximate LOPR results in Sec. IV B.

B. Local probability recursion (LOPR) results

The results presented below were found using Eq. (16) for the site algorithm and Eqs. (17)–(19) for the edge algorithm using double precision on 32-bit linux PCs. The site probabilities, P_i were converged to an accuracy of 5×10^{-8}. We wrote two different codes, one in Fortran and the other in C++. These codes give identical results, for the same set of

graphs, initial conditions, and convergence criteria.

An illustration of the results for the probabilities P_i using the site algorithm [Eq. (16)] are presented in Fig. 8(a), while those found using the edge algorithm [Eqs. (17)–(19)] for the same sample are presented in Fig. 8(b). In these figures, the open circles indicate a site that has probability 1 of being in the MIS, while those with filled circles have probability, $P_i=0$. The hatched sites have a probability between 0 and 1, i.e., $0<P_i<1$. The most notable feature of these graphs is the fact that the vertex based algorithm [Fig. 8(a)] has a strong tendency to freeze the probabilities to their binary values while the bond algorithm freezes very few sites [see Fig. 8(b)].

A histogram of the probabilities for large lattices for three values of the bond concentration are presented in Figs. 9(a) (site algorithm) and 9(b) (bond algorithm). The graphs of Fig. 9 illustrate the variability in the site probabilities that is typical of hard computational problems and that there is a nontrivial order parameter distribution. There are some pronounced peaks in these probabilities at low-order rational values, in addition to the sharp peaks at the binary values $P=1$ and $P=0$. The strong peaks at 0 and 1 show that the site algorithm finds a far greater number of integer probability "frozen" sites as compared to the edge algorithm. As the average coordination c of the graph increases the degenerate continuum decreases for the site algorithm; but for the bond algorithm, the degeneracy increases with increasing c, indicating a higher degeneracy. The average value of the MIS found from the site and bond algorithms are presented in Fig. 10(a), along with the average MIS found using the transfer matrix method, and the symmetric Bethe lattice result of Eqs. (20)–(22). Despite the differences in the number of frozen sites, the site and bond LOPR results for the average MIS are remarkably similar. In addition, the LOPR results are quite close to the transfer matrix results and are a very significant improvement over the symmetric Bethe approximation. The average fraction of frozen sites is presented in Fig. 10(b), where results from the site and bond LOPR methods are compared to the symmetric Bethe approximation and to exact results on small lattices [10]. It is evident from these results that the LOPR algorithms do not correctly capture the fracture of frozen sites, with the site algorithm being better but significantly higher than the true values while the bond algorithm is remarkably poor in this regard.

V. SUMMARY

We have explored the behavior of core percolation and the maximum independent set on bond-diluted triangular lattices. The core that remains after iterative leaf removal undergoes a continuous percolative phase transition at a bond threshold of $p_c=0.4692(4)$ with thermal exponent consistent with conventional percolation. The order parameter and susceptibility exponents are significantly different from the conventional values even on lattice of ten million sites. For example a naive fit yields $\beta=0.18(1)$, which is significantly higher than the conventional value of $\beta=5/36\approx 0.138$, but it is much smaller than the value for the backbone in percolation, which is 0.48(1). Moreover, the small deviation of the core exponent found from the conventional percolation value for β may be reconciled by the presence of strong corrections to scaling as illustrated in Fig. 4(c).

We introduced a transfer matrix method for finding the exact cardinality and degeneracy of the maximum independent set. Using this method, we carried out detailed calculations of the average cardinality per site $\langle i\rangle/N$ and the average entropy per site $\langle s\rangle/N$ of bond diluted triangular lattices, and by extrapolation found very accurate results for the infinite lattice values of these quantities, as presented in Table I. The computational complexity of the transfer matrix method is $O(L\exp(aW))$, where a is $O(1)$ for strips of width W and length L, making the strip geometry favorable computationally.

The transfer matrix results were compared to the results of local probability recursion (LOPR) methods based on vertex and bond schemes. LOPR methods yield values for the average MIS that are precise for $p<p_c$, where the symmetric Bethe approximation is expected to be exact [10], and are accurate at the 1–2 % level for $p>p_c$ where replica symmetry breaking is important. LOPR methods perform poorly in calculating the fraction of frozen sites with the bond algorithm being particularly erroneous [see Fig. 10(b)]. The fact that the site method overestimates the frozen fraction while the bond method underestimates it suggests that hybrid site and/or bond schemes may be promising.

ACKNOWLEDGMENTS

This work has been supported by the DOE under Contracts No. DE-FG02-90ER45418 and No. DE-FG02-97ER45651. Alexandar Hartmann kindly provided his code for exact solution of small samples.

[1] G. P. M. Mézard and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[2] M. Mézard, Science **301**, 1685 (2003).
[3] M. Mézard, G. Parisi, and R. Zecchina, Science **297**, 812 (2002).
[4] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, Nature (London) **297**, 812 (1999).
[5] M. Weigt and A. K. Hartmann, Phys. Rev. Lett. **84**, 6118 (2000).
[6] A. Braunstein, R. Mulet, A. Pagnani, M. Weigt, and R. Zecchina, Phys. Rev. E **68**, 036702 (2003).
[7] A. Braunstein and R. Zecchina, Lect. Notes Comput. Sci. **2919**, 519 (2004).
[8] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1979).

[9] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[10] M. Weigt and A. K. Hartmann, Phys. Rev. E **63**, 056127 (2001).
[11] R. Karp and M. Sipser, *Proceedings of the 22nd IEEE Symposium on Foundations of Computing* (IEEE, New York, 1981), pp. 364–375.
[12] M. Bauer and O. Golinelli, Eur. Phys. J. B **24**, 339 (2001).
[13] H. Zhou, Eur. Phys. J. B **32**, 265 (2003).
[14] H. Zhou, Phys. Rev. Lett. **94**, 217203 (2005).
[15] R. E. Tarjan and A. E. Trojanowski, SIAM J. Comput. **6**, 537 (1977).
[16] L. G. Valiant, Theor. Comput. Sci. **8**, 189 (1979).
[17] S. P. Vadhan, SIAM J. Comput. **31**, 398 (2001).
[18] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, New York, 1982).
[19] M. Mezard, G. Parisi, and R. Zecchina, Science **297**, 812 (2002).
[20] A. Braunstein and R. Zecchina, J. Stat. Mech.: Theory Exp. 2004, P06007.
[21] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor and Francis, New York, 1994).
[22] C. Moukarzel and P. M. Duxbury, Phys. Rev. E **59**, 2614 (1999).
[23] J. Adler, M. Moshe, and V. Privman, Phys. Rev. B **26**, 1411 (1982).