Computing monomer-dimer systems through matrix permanent

Yan Huo, Heng Liang, Si-Qi Liu, and Fengshan Bai*
Department of Mathematical Sciences, Tsinghua University, Beijing, China, 100084
(Received 17 August 2007; revised manuscript received 17 October 2007; published 15 January 2008)

The monomer-dimer model is fundamental in statistical mechanics. However, it is #P-complete in computation, even for two-dimensional problems. A formulation for the partition function of the monomer-dimer system is proposed in this paper by transforming the number of all matchings of a bipartite graph into the number of perfect matchings of an extended bipartite graph, which can be given by a matrix permanent. Sequential importance sampling algorithm is applied to compute the permanents. For two-dimensional lattice with periodic condition, the monomer-dimer constant is known as \( h_2 = 0.662798972834 \). We obtain \( 0.6627 \pm 0.0002 \) for our approximation, which shows the robustness and the efficiency of the algorithm. For three-dimensional problem, our numerical result is \( 0.7847 \pm 0.0014 \), which agrees with the best known bounds.

DOI: 10.1103/PhysRevE.77.016706 PACS number(s): 02.70.Uu, 05.50.+q, 02.10.Ox, 02.50.--r

I. INTRODUCTION

A monomer-dimer model is considered in which the set of sites in a lattice is covered by a nonoverlapping arrangement of monomers (molecules occupying one site) and dimers (molecules occupying two sites that are neighbors in the lattice). It is fundamental in lattice statistical mechanics [1,2]. A two-dimensional monomer-dimer model with size \( m=(m_1, m_2) \) is a rectangle lattice with \( m_1 \times m_2 \) sites. The two-dimensional monomer-dimer systems are used to investigate the properties of adsorbed diatomic molecules on a crystal surface [3]; the three-dimensional systems occur classically in the theory of mixtures of molecules of different sizes [4] as well as the cell cluster theory of the liquid state [5]. A more complete description of the history and the significance of the monomer-dimer model can be found in Ref. [1], and references therein.

All possible monomer-dimer coverings for a given lattice defines the configuration space of a monomer-dimer system. A fundamental question for such a statistical mechanics model is to determine the cardinal number of the configuration space. Practically, most of the thermodynamic properties of physical systems can be obtained from the number of all possible ways that a given lattice can be covered. Thus a considerable attention has been devoted to such a counting problem. For a \( d \)-dimensional cubic lattice with size \( m=(m_1, m_2, \ldots, m_d) \), this cardinal number is denoted by \( Z(m,d) \). It is proved that the following limit exists:

\[
  h_d = \lim_{m \to \infty} \frac{\ln Z(m,d)}{m_1 m_2 \cdots m_d}.
\]

The limit \( h_d \) is called monomer-dimer constant [6].

Even for the simplest two-dimensional models, there are very few closed form results on the monomer-dimer constant. Baxter and Gaunt [7,8] give estimates of the constants using the asymptotic expansions. Hammersley and Menon [9] estimate the \( h_2 \) by calculating lower and upper bounds. Numerical simulation should play an important role. However, it has been proved that computing the monomer-dimer constant is a #P-complete problem even for two-dimensional problems [10], which shows the hardness of the computation. The Monte Carlo methods are applied to study the problem in Refs. [2,6,11], which is a natural consideration.

Recently, Friedland and Peled [12] proposed methods for the estimation of multidimensional topological entropy. Applying the methods to the computation of monomer-dimer constant, they obtain \( h_2 = 0.662798979 \), which agrees with the heuristic estimation \( e^{0.2} = 1.940215351 \) of Baxter [7], and \( 0.7653 \leq h_3 \leq 0.7682 \). Two-dimensional model with fixed dimer density was studied intensively by Kong [13]. The monomer-dimer constant with 12 digits accuracy for two-dimensional problem is given as \( h_2 = 0.662798972834 \).

In this paper, we propose a formulation that transforms the counting of all matchings in a bipartite graph to the counting of perfect matchings in an extended bipartite graph. Hence, the monomer-dimer systems in any dimension can be computed by permanents of the matrices. Permanent of a matrix is studied for a long time [14,15]. After Valiant proves that evaluating the permanent of a 0-1 matrix is a #P-complete problem [16], many randomized approximate algorithms are developed [17–19]. They can give reasonable estimations for permanents within acceptable computer time.

We consider cubic lattices with periodic condition, and concentrate on two- and three-dimensional lattices in the computation. The algorithms are applicable to other dimensions and domains other than rectangle. For simplicity of notation, we assume that \( m_1=m_2=\cdots=m_d \). But this is not essential for the algorithms.

In the next section, a formulation of the monomer-dimer model in bipartite graph is presented. The partition function of the system is represented as the permanent of the adjacency matrix of the graph in Sec. III. Computational methods are discussed in Sec. IV. The sequential importance sampling algorithms are applied to compute the matrix permanent. In Sec. V, numerical results are given which clearly show the efficiency of our formulation and the computational methods. Finally in Sec. VI, some discussions and comments are given.
lattice can be divided into two vertex sets monomers and dimers. Figure 1
lattice in Fig. 1 mapped as matchings in the graph
There are edges between the neighbors, and all edges form
sets. Consider the lattice in Fig. 1 of the graph
dimer system can be represented as a covering of the vertices
of the graph
neighbors in the lattice. Hence a graph
a vertex, and an edge between two vertices exists if they are

Consider a lattice with the site number
define

Each site in the lattice is regarded as
V
A
shows a
V
A
1 and all stars give the set
2.

Actually the monomer-dimer configurations can easily be
mapped as matchings in the graph G. The sites of a cubic
lattice can be divided into two vertex sets V₁ and V₂. A site
and its neighbor should always belong to different vertex
sets. Consider the lattice in Fig. 1(a) as an example. All
circles give the vertex set V₁ and all stars give the set V₂.
There are edges between the neighbors, and all edges form
the edge set E. Thus an undirected bipartite graph
G(V₁∪V₂,E) is constructed (see Fig. 2). In terms of the
graph theory, a covering of all vertices with dimers is a per-
fect matching of the bipartite graph, and a covering with k
dimers is a k-matching of the graph. Hence the cardinal
number of the configuration space of the monomer-dimer model
equals to the number of all possible matchings in the
bipartite.

II. FORMULATION IN BIPARTITE GRAPH

Consider a lattice with the site number N (N is even), and
define n=N/2. [Figure 1(a) shows a (4,4) planer lattice, with
N=4×4=16 and n=8.] Each site in the lattice is regarded as
a vertex, and an edge between two vertices exists if they are
neighbors in the lattice. Hence a graph G=(V,E) is naturally
defined. Using the terminology of graph theory, a monomer-
dimer system can be represented as a covering of the vertices
of the graph G=(V,E) by a nonoverlapping arrangement of
monomers and dimers. Figure 1(b) shows a covering of the
lattice in Fig. 1(a) with two monomers and seven dimers.

Actually the monomer-dimer configurations can easily be
mapped as matchings in the graph G. The sites of a cubic
lattice can be divided into two vertex sets V₁ and V₂. A site
and its neighbor should always belong to different vertex
sets. Consider the lattice in Fig. 1(a) as an example. All
circles give the vertex set V₁ and all stars give the set V₂.
There are edges between the neighbors, and all edges form
the edge set E. Thus an undirected bipartite graph
G(V₁∪V₂,E) is constructed (see Fig. 2). In terms of the
graph theory, a covering of all vertices with dimers is a per-
fect matching of the bipartite graph, and a covering with k
dimers is a k-matching of the graph. Hence the cardinal
number of the configuration space of the monomer-dimer model
equals to the number of all possible matchings in the
bipartite.

III. PARTITION FUNCTION VIA PERMANENTS OF
ADJACENCY MATRICES

The partition function of the monomer-dimer system is
defined as

\begin{equation}
Z(\lambda) = Z_G(\lambda) = \sum_{k=0}^{n} m_k \lambda^k,
\end{equation}

where \( m_k = m_k(G) \) is the number of k-matchings in the graph
G, which is equivalent to the number of monomer-dimer
configurations with k dimers. \( Z_G(1) \) enumerates all possible
matchings in G.

Let G be a bipartite graph and A be the adjacency matrix
of the graph G. The number of perfect matchings in G is
equal to the permanent of the matrix A, which is defined as

\begin{equation}
\text{perm}(A) = \sum_{\sigma \in \Pi_n} \prod_{i=1}^{n} a_{i\sigma(i)}.
\end{equation}

Here \( \Pi_n \) is the symmetric group of degree n.

A matrix permanent formulation for enumerating
k-matching in a bipartite graph is proposed by Friedland and
Levy recently [21]. For any given k, the method by Friedland
and Levy can compute \( m_k \). Thus \( Z_G(1) \), all possible match-
ings in the graph G, can be given by Eq. (1). In this way the
monomer-dimer constant can be approximated. Note that the
number of matrix permanents that have to be computed is n,
and the n could not be a big number.

Here in the following we propose a new formulation in
matrix permanent by constructing an auxiliary bipartite
digraph. The number of all possible matchings can be approxi-
mated directly.

Let A be the adjacent matrix of a bipartite graph G. Thus
A is a 0-1 matrix. Denote \( G(A) \) as the bipartite graph with
adjacent matrix A. An auxiliary graph is constructed as an
extended graph of \( G(A) \) (see Fig. 3) as follows. Vertex sets

FIG. 2. The bipartite graph \( G(V_1 \cup V_2,E) \).

FIG. 3. (Color online) The auxiliary graph \( G(B) \).
$V_1^*$ and $V_2^*$ are added to $V_1$ and $V_2$, respectively, hence the vertex sets for the extended graph are $V_1 \cup V_1^*$ and $V_2 \cup V_2^*$. The cardinal numbers of the new sets $V_1^*$ and $V_2^*$ are both $n$. Each vertex in $V_1$ is adjacent to one vertex in $V_2^*$, and different vertex in $V_1$ links to different vertex in $V_2^*$. Hence there are exactly $n$ edges linking $V_1$ and $V_2^*$. Each vertex of $V_1^*$ is adjacent to all vertexes in $V_2 \cup V_2^*$. Let

$$B = \begin{pmatrix} A & I_{n \times n} \\ 1_{n \times n} & 1_{n \times n} \end{pmatrix},$$

(3)

where $1_{n \times n}$ is the $n \times n$ matrix whose entries are all equal to 1, and $I_{n \times n}$ is the identity matrix of order $n$. It is obvious that $B$ is a 0-1 matrix, and it is the adjacent matrix of the auxiliary graph.

Let $M_{all}(A)$ denote the number of all possible matchings in the graph $G(A)$. Now consider the graph $G(B)$. Note that $\text{perm}(B)$ always gives the number of perfect matchings in the graph $G(B)$. In a perfect matching of $G(B)$, each vertex in $V_1$ is assigned to be adjacent to a vertex in $V_2 \cup V_2^*$. The number of all possible assignment between $V_1$ and $V_2 \cup V_2^*$ equals to $M_{all}(A)$. After the adjacent edges between $V_1$ and $V_2 \cup V_2^*$ are chosen, there are $n!$ possibilities for choosing the adjacent edges between $V_1^*$ and $V_2 \cup V_2^*$. So we have $M_{all}(A) \cdot n! = \text{perm}(B)$, that is,

$$M_{all}(A) = \frac{1}{n!} \text{perm} \begin{pmatrix} A & I_{n \times n} \\ 1_{n \times n} & 1_{n \times n} \end{pmatrix}. \quad (4)$$

Moreover we can get the following formulation:

$$Z_G(\lambda) = \frac{1}{n!} \text{perm} \begin{pmatrix} \lambda \cdot A & I_{n \times n} \\ 1_{n \times n} & 1_{n \times n} \end{pmatrix} = \sum_{k=0}^{n} m_k \lambda^k. \quad (5)$$

Hence the partition function of the monomer-dimer system is formulated as the permanent of a matrix. It is important to notice that the matrix $B$ is very special in structure, which will be explored in the efficient numerical algorithms.

### IV. COMPUTATIONAL METHODS

A bridge between the computation of permanent and the partition function of the monomer-dimer system is established via the relationships (4) and (5). Thus the partition function of the monomer-dimer model can be computed or approximated by taking the advantage of the efficient algorithms in matrix permanent.

The definition of the permanent $\text{perm}(A)$ looks similar to that of the determinant $\text{det}(A)$. However, it is much harder to be computed, even for 0-1 matrices [16]. Hence the approximate algorithms would be a nature consideration. The well known approximate methods for matrix permanents are Monte Carlo algorithms. One way to do so is to relate matrix permanents to matrix determinants by randomizing the elements of matrices [17,18]. However it is not suit for large scale computation such as monomer-dimer problems.

The Markov chain Monte Carlo approach can give a fully polynomial randomized approximation scheme for the permanent of arbitrary non-negative matrix. This result is important theoretically and obtained by Jerrum, Sinclair, and Vigoda [19]. Beichl, O’Leary, and Sullivan [11] compute the number of $k$-matchings of the monomer-dimer model using Markov chain Monte Carlo method. They improve the Kenyon, Randall, and Sinclair (KRS) method [2] and obtain numerical results for the monomer-dimer systems.

The Monte Carlo methods with sequential importance sampling (SIS), which are a kind of practical algorithms for approximating permanent of relatively large matrices, would be more promising for the monomer-dimer model [20,22,23]. Beichl and Sullivan give the best known numerical result for three-dimensional dimer constant by using the techniques [23]. It would be an efficient method for the monomer-dimer model as well, because of the formulation that we establish through the auxiliary graph in the last section. The framework of sequential importance sampling for the permanent of an $n \times n$ 0-1 matrix $A$ is as follows.

**Algorithm SIS:**

1. Choose a nonzero element from the first row of the matrix $A$ with some probability $p_i$. Suppose the column index of this element be $k_i$. Set all the other entries in the first row and the $k_i$th column to 0’s.
2. Proceed to the next row, applying the same sampling strategy as step 1, recursively. Hence the values $p_2, \ldots, p_n$ can be obtained.
3. Compute $X = \frac{1}{p_1 p_2 \cdots p_n}$.

The output $X$ of algorithm SIS is a random variable. It is an unbiased estimator to the permanent of a 0-1 matrix $A$. Different strategies of choosing the probability distributions would lead to different sequential importance sampling algorithms.

In the procedure of SIS, only nonzero terms are valuable to the computation. It is important to notice that algorithm SIS never produce any zero sample for the matrix $B$ defined in Eq. (3) because of the special structure of the matrix. Assume that the $k$ ($k<n$) nonzero elements in the first $k$ rows of matrix $B$ have been chosen successfully. It is noted that $(n+i)$-th column must keep unchanged for all ($i>k$) up to now. Then at least one nonzero element $b_{i+n,i+k+1}$ in the $(k+1)$-th row of the $B$ can be chosen in the procedure of SIS algorithm. Thus the formulation and computational methods never meet any zero term. This fact is crucial for the efficiency of the algorithm.

**Remark.** Looking at the monomer-dimer model directly, each stochastic experiment by the SIS algorithm simply generates a monomer-dimer configuration in random. Consider the vertexes in $V_1$ in turn. For the current vertex, one can choose to either combine it with one of its unoccupied neighbors in $V_2$ so that the current vertex and the chosen neighbor are covered by a dimer, or just keep itself isolated. The decision above is made in random. Repeating this procedure, an admissible configuration of the monomer-dimer system will be obtained without interruption. Viewing the procedure above in the SIS algorithm, forming a dimer corresponds to choosing 1 in the matrix $A$, otherwise to choosing 1 in the matrix $I_{n \times n}$. Thus the SIS algorithm would always produce nonzero term.

Now we apply algorithm SIS to compute the permanent of the matrix $B$ in Eq. (3). The matrix structure is so special that all the elements in the $(n+1)$th to $(2n)$th rows of $B$ are 1. Hence only the first $n$ rows of $B$ are needed to be considered.
TABLE I. The convergence rates of the three algorithms for expectation of $X$. Liu and BS give good results, and Liu runs faster in the dimensional lattice.

<table>
<thead>
<tr>
<th>$m$</th>
<th>Ras</th>
<th>Liu</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>time (s)</td>
<td>time (s)</td>
<td>time (s)</td>
</tr>
<tr>
<td>2</td>
<td>6.9999</td>
<td>7.0000</td>
<td>1.99</td>
</tr>
<tr>
<td>4</td>
<td>41198</td>
<td>41071</td>
<td>5.55</td>
</tr>
</tbody>
</table>

Assume that one sampling gets probability values $p_1, p_2, \ldots, p_n$. The sampling value should be assigned as

$$\frac{1}{n!} \cdot \frac{1}{p_1} \cdot \frac{1}{p_2} \cdots \frac{1}{p_n}.$$

If $N$ samples are obtained by Algorithm SIS, the number of all matchings can be approximated by

$$M_{all}(A) = \frac{\text{perm}(B)}{n!} = \frac{n!}{\prod_{j=1}^{n} p_j^{j}}.$$

Three different importance sampling methods Ras by [22], Liu by [20], and Beichl and Sullivan (BS) by [23] are used, respectively, to compute the number of the cardinal number of the configuration space. The results are given in Table I. The convergence rates of the three algorithms for $m=4$ are also shown in Fig. 4. Simple examples show that both Liu and BS give good results, and Liu runs faster in the computation of monomer-dimer model.

Notice that the probability distribution of the random variable $Y = \ln X$ looks similar to the normal distribution. If the probability distribution of $Y$ is normal with $N(\mu, \sigma)$, then the expectation of $X$ would be

$$E(Y) = \mu = \frac{1}{n} \sum_{j=1}^{n} p_j^{j}.$$

Other than computing the sample mean of $X$ directly, we can estimate the sample mean $\bar{X}$ and sample standard deviation $\sigma$ of the random variable $Y$ first.


time II denotes that in 400 MHz IBM PC. A-PRE presents the values given in Ref. [11].

<table>
<thead>
<tr>
<th>$m$</th>
<th>SIS</th>
<th>Time I (s)</th>
<th>Time II (s)</th>
<th>A-PRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.66366</td>
<td>0.0006</td>
<td>0.0050</td>
<td>0.6611</td>
</tr>
<tr>
<td>6</td>
<td>0.66251</td>
<td>0.0017</td>
<td>0.0110</td>
<td>0.6629</td>
</tr>
<tr>
<td>8</td>
<td>0.662897</td>
<td>0.0028</td>
<td>0.0220</td>
<td>0.6611</td>
</tr>
<tr>
<td>10</td>
<td>0.66291</td>
<td>0.0038</td>
<td>0.0330</td>
<td>0.6663</td>
</tr>
<tr>
<td>12</td>
<td>0.662990</td>
<td>0.0055</td>
<td>0.0770</td>
<td>0.6646</td>
</tr>
<tr>
<td>14</td>
<td>0.662852</td>
<td>0.0072</td>
<td>0.0820</td>
<td>0.6638</td>
</tr>
<tr>
<td>16</td>
<td>0.662644</td>
<td>0.0100</td>
<td>0.1320</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.663390</td>
<td>0.0138</td>
<td>0.1920</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.662960</td>
<td>0.0181</td>
<td>0.2530</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0.663031</td>
<td>0.0237</td>
<td>0.3300</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0.662893</td>
<td>0.0307</td>
<td>0.4230</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>0.663754</td>
<td>0.0398</td>
<td>0.5490</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>0.663013</td>
<td>0.0507</td>
<td>0.7030</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.663062</td>
<td>0.0710</td>
<td>0.8790</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>0.662587</td>
<td>0.0769</td>
<td>1.0980</td>
<td></td>
</tr>
</tbody>
</table>

FIG. 4. (Color online) A $(4, 4)$ lattice is considered, and thus the adjacent matrix is $16 \times 16$. The x axis denotes the number of samples and the y axis denotes the error of the approximate cardinal number of configuration space.
order to compare the algorithms more adequately, we run our MATLAB program on a 400 MHz IBM PC to rule out the factor of hardware. This machine is inferior to Sun-Ultra60 with 450 MHz that was used in Ref. [11]. The computational times on the 400 MHz IBM PC are listed in Table II. The results show that the computational times are roughly proportional between CPU times on 400 MHz IBM PC and that on 2.8 GHz Dell, which suggests that the SIS algorithm is robust in hardware. Hence the method SIS can be applied to relatively large monomer-dimer systems.

In order to fit the limit of \( \ln Z(m,2)/m^2 \) as \( m \) goes to infinity, we apply regression to the computed values. The regression function is the same as [23]

\[
y = \frac{p_1}{x^3} + p_2,
\]

where \( x \) denotes the lattice size \( m \), \( y \) denotes \( \ln Z(m,2) \) and \( p_2 \) is the monomer-dimer constant. The monomer-dimer constant of two-dimensional problem with periodic boundary can be obtained from the regression

\[
h_2 = 0.6627 \pm 0.0002 \text{ with } 95\% \text{ confidence}.
\]

The approximate results of the monomer-dimer constant coincides with the value \( h_2 = 0.662798972834 \) in Ref. [13] very well.

**B. Experiments on three-dimensional lattices**

For three-dimensional problem with periodic condition, computational results are shown in Table III. The time complexity for algorithm SIS in three-dimensional problems is about \( O(m^6) \).

To fit the limit of \( \ln Z(m,3)/m^3 \) as \( m \) goes to infinity, we apply regression again. The function we use is

\[
y = \frac{p_1}{x} + p_2,
\]

where \( x \) denotes the lattice size \( m \), \( y \) denotes \( \ln Z(m,3)/m^3 \) and \( p_2 \) is the monomer-dimer constant. Notice that the regression function here is different from the case of two-dimensional. The regression result is

\[
h_3 = 0.7847 \pm 0.0014 \text{ with } 95\% \text{ confidence}.
\]

This agrees well with the best known bound \( 0.7653 < h_3 \leq 0.7862 \) [12].

### VI. DISCUSSIONS AND COMMENTS

The construction of the auxiliary bipartite graph is the key step in our formulation. Hence the permanent of the matrix \( B \) in Eq. (3) gives the total number of matchings in the original bipartite graph \( G(A) \). The size of the matrix \( B \) doubles that of \( A \). However, since the special structure of the matrix \( B \) can be exploited in the algorithm, the computational cost does not really increase.

The regression function (6) for two-dimensional is discussed and used by other authors, for example, Ref [23]. However, Eq. (7) for three-dimensional is just a result of statistical experiments. We are unable to give it any physical reasoning.

The basic contribution of this paper is the formulation and computational methods for approximating the number of all matchings in bipartite graphs. In this way, larger monomer-dimer systems can be studied.

### ACKNOWLEDGMENTS

The work was supported by the National Science Foundation of China, Grant No. 10501030.