

Accurate Lower Bounds on Two-Dimensional Constraint Capacities From Corner Transfer Matrices ¹

Yao-ban Chan and Andrew Rechnitzer

Abstract

We analyse the capacity of several two-dimensional constraint families — the exclusion, colouring, parity and charge model families. Using Baxter’s corner transfer matrix formalism combined with the corner transfer matrix renormalisation group method of Nishino and Okunishi, we calculate very tight lower bounds and estimates on the growth rates of these models. Our results strongly improve previous known lower bounds, and lead to the surprising conjecture that the capacity of the even and charge(3) constraints are identical.

Index Terms

Channel capacity, min-max principle, corner transfer matrices, multi-dimensional constraints.

I. INTRODUCTION

IN this paper, we analyse the capacity of several two-dimensional constraints. A constraint induces a model of magnetic spins, each of which can take a small number of spin values (typically 2, representing ‘up’ or ‘down’ magnetic spins). These spins are arranged in an ordered fashion on a two-dimensional plane — in all the models we analyse, they are positioned at the vertices of the square lattice \mathbb{Z}^2 . Each constraint limits in some way the potential arrangements of the spins on the lattice.

An application where constraints arise is that of data storage. Here, the values of the spins contain data stored (in some encoding) on a magnetic disk. Due to properties of the disk, it may be undesirable for certain local arrangements of spins to exist, and this leads to a corresponding constraint. For example, in the hard squares model we define below, the spins can take the values \oplus and \ominus . However, the presence of two \ominus spins next to each other may cause magnetic interference, so we do not allow this to happen.

Given a constraint such as this, we wish to know what the storage capacity of the disk is. In the language of statistical mechanics, the *partition function* is the number of different configurations that can be written on N spins under the constraint:

$$Z_N = (\# \text{ of legal configurations on } N \text{ spins}). \quad (1)$$

We are interested in the behaviour of this number as the size of the storage array grows to infinity, and so define the *partition function per site*

$$\kappa = \lim_{N \rightarrow \infty} Z_N^{1/N}. \quad (2)$$

We also call κ the *growth rate* of the model. The growth rate is directly related to the *capacity* of a constraint, which is defined to be the number of bits of information per spin that is possible to be written on a disk with that constraint:

$$\text{cap}(C) = \lim_{N \rightarrow \infty} \frac{1}{N} \log_2 Z_N = \log_2 \kappa. \quad (3)$$

In other words, a disk with N spins can store approximately $\text{cap}(C) \times N$ bits of information under the constraint C . A ‘free’ disk with no constraints has a capacity of 1, since there are 2^N possible configurations on N spins,

Manuscript submitted 23/10/2012.

YBC is at Universität Wien; yao-ban.chan@univie.ac.at

AR is at the University of British Columbia; andrewr@math.ubc.ca

and as such each spin represents one bit of information. As we approach this problem from a combinatorial point of view, we shall limit our terminology to growth rates rather than capacity (of course, they are trivial to derive from each other).

This problem is also studied in symbolic dynamics and dynamical systems, but with different terminology. Here, the capacity is also known as the (topological) *entropy* of the model, while the spins are *letters* (of an alphabet). Constraints are called *sofic shifts* (or systems), with a slight distinction: in constrained coding, the focus is on configurations of finite systems of arbitrary size, while in symbolic dynamics (and this paper), the focus is on configurations on the entire lattice. The exclusion and colouring models below are sub-classes of so-called ‘shifts of finite type’ or ‘finite memory systems’.

The capacity of constraints in one dimension is, in general, very easily computed using a transfer matrix. Unfortunately, capacities of two-dimensional constraints are significantly more difficult to compute, and because of this we seek estimates and rigorous lower and upper bounds for their values. Of interest here are the methods of Engel [1], and Calkin and Wilf [2], who provide a method to calculate such bounds using transfer matrices, illustrating with the hard squares model. Louidor and Marcus [3] generalised this method and applied it to the exclusion, parity and charge families of constraints discussed here. Additional references are provided for individual models as we introduce them below.

In this paper, we use corner transfer matrix (CTM) formalism to derive very precise lower bounds on the growth rates of the models that we study. The corner transfer matrix method was first devised by Baxter in 1968 [4] as a way of numerically estimating the growth rate of fully-packed dimers on a rectangular lattice. In later studies in the late 70s [5], [6], it was developed into method to calculate series expansions of the partition function and magnetisation. A particularly noteworthy triumph of this method was Baxter’s famous solution of the hard hexagons model [7], [8], which was derived by noticing a pattern in the eigenvalues of the corner transfer matrices.

As far as we know, CTM formalism has never been used before to calculate lower bounds, rather than estimates (although in fact, the theoretical machinery required to produce lower bounds has long been in place). When used for this purpose, the accuracy of the bound depends on how close we can come to the solution of the model.

For this purpose, we use the corner transfer matrix renormalisation group (CTMRG) method of Nishino and Okunishi [9], [10]. This method combines the original CTM method with ideas from the density matrix renormalisation group method (see, for example, [11], [12]), and also produces estimates of the growth rate. The CTMRG is more general and easier to apply numerically than Baxter’s original method, and has been applied, among other things, to self-avoiding walks [13] and the Ising model scaling function [14]. Recently, it too has been extended by one of the authors [15] to calculate series expansions.

In Section II, we define the models that we study and discuss some reformulations needed to make them amenable to CTM analysis. In Section III, we explain the corner transfer matrix formalism that we use to derive our lower bounds and estimates. These numerical results are presented in Section IV. In Section V, we discuss the (conjectured) equivalency between two of our models, and conclude in Section VI.

II. THE MODELS

We calculate the growth rate of several models in this paper. All of these models have *spins* which lie on the vertices of a square lattice, which usually take the values \oplus and \ominus . The configurations will be constrained in certain ways that we describe below.

A. Exclusion models

In the exclusion models, we consider configurations of \oplus, \ominus spins which are forbidden to contain certain local configurations of \ominus s (see Fig. 1).

Definition 1. Consider a configuration of \oplus, \ominus spins on the vertices of the square lattice.

- The configuration satisfies the *hard squares* constraint if it does not contain two \ominus spins joined by a single bond. If one considers the \ominus spins to be the centres of rotated square blocks, then this constraint requires that these blocks do not overlap.

This model is a special case of a more general hard squares model where the constraint is enforced, but each \ominus spin is also given a weight z . Our model corresponds to $z = 1$. In the more general model, questions of

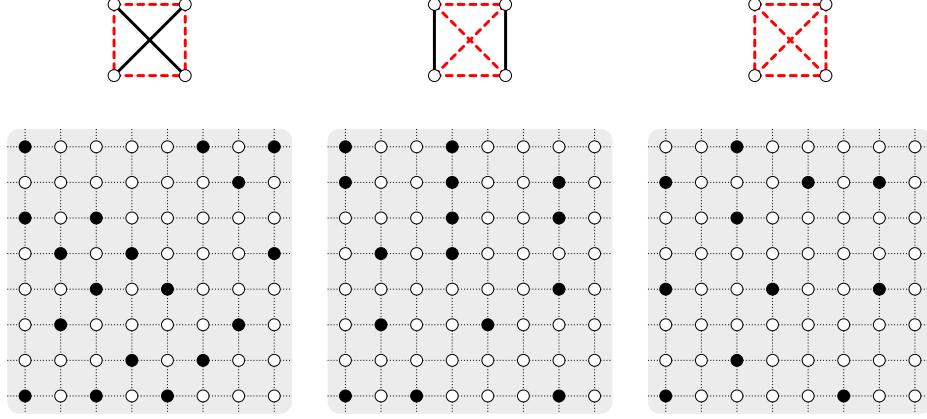


Fig. 1. (bottom) Examples of the three exclusion models we consider (from left to right): hard squares, RWIM and NAK. We represent \ominus spins by black vertices and \oplus spins by white vertices. (top) The diagrams show the allowed configurations around a face for each model; two vertices joined by a solid line may be simultaneously occupied by \ominus s, while vertices joined by a dashed line may not.

interest include not only this special case, but also the location of the phase transitions (singularities in the free energy). These questions have been studied using corner transfer matrices in [6], [15]. Placing the spins on the vertices of a triangular lattice results in the hard hexagon model, which has been exactly solved using corner transfer matrices [7], [8]. Studies on the $z = 1$ case include [2], [16].

- The configuration satisfies the *read-write isolated memory* (RWIM) constraint [17], [18] if it does not contain two \ominus spins joined by a horizontal bond, or lying diagonally across a single face. This constraint can be used to model one-dimensional memory states with the hard square constraint, which are sequentially updated in such a way that no two adjacent spins are changed in the same update.
- The configuration satisfies the *non-attacking kings* (NAK) constraint if it does not contain two \ominus spins joined by a single bond, or lying diagonally across a single face. Equivalently every face of the lattice contains at most one \ominus spin. If one considers each \ominus spin to be a king on a chessboard, then the NAK constraint requires that no king can attack another. This has been studied, along with the hard squares and other models, in [19].

B. Colouring models

We also consider vertex colourings of the lattice; each vertex is coloured using labels from 0 to $q - 1$ for some fixed q , and again we prohibit certain local configurations.

Definition 2. A configuration of $\{0, \dots, q - 1\}$ spins satisfies the *q-colouring* constraint if it does not contain two nearest-neighbour spins with the same colour.

This is related to the chromatic polynomial and the anti-ferromagnetic Potts model (see for example [20], [21], [22], [23]). When $q = 2$, there are only two possible configurations, while for $q = 3$ it is known that the growth rate is exactly $(4/3)^{3/2} = 1.539\,600\,7178\dots$ [24]. We study this model for higher q , where the growth rate is not known exactly.

C. Parity models

Again we place the spins \oplus and \ominus at the vertices of the square grid. However unlike the models above, the constraints are no longer local. The two parity models arise by constraining the parity of the number of \oplus spins lying between two \ominus spins along any horizontal or vertical line.

Definition 3. Consider a configuration of \oplus, \ominus spins on the vertices of the square grid. Take any two \ominus spins lying along a horizontal or vertical line so that there is no other \ominus spin between them.

- If every such pair is separated by an even number of \oplus spins, then the configuration satisfies the *even* constraint.

- Conversely, if every such pair is separated by an odd number of \oplus spins, then the configuration satisfies the *odd* constraint.

While the above spin constraints are not local, it is possible to map them to other models of states with local constraints, which makes them amenable to CTM analysis. We explain this mapping below.

The growth rate of the odd model is known to be $\sqrt{2}$ [3]. To prove that the growth rate is at least $\sqrt{2}$, populate the lattice with \oplus in a checkerboard pattern. The remaining spins, which constitute half the total spins, can then be assigned \oplus or \ominus independently. This always produces a valid configuration, so the bound follows.

To obtain the opposite bound, consider configurations of spins which obey the odd constraint along horizontal lines, while having no constraint along vertical lines. Being less constrained, the growth rate of this new model is greater than that of the odd model and is equal to the growth rate of the odd model in 1 dimension. It is straightforward to show that this is exactly $\sqrt{2}$.

Since the growth rate of the odd model is known exactly, we do not discuss it further.

D. Charge(3) model

A final family of constraints is parameterised by a positive integer corresponding to the maximum cumulative electric ‘charge’ along a horizontal or vertical line.

Definition 4. A configuration of \oplus, \ominus spins satisfies the *charge(3)* constraint if for any finite horizontal or vertical line segment, the sum of the spins along that segment lies between -3 and $+3$.

The above definition means that starting at any horizontal bond in the lattice, there exists an initial charge between 0 and 3 so that scanning to the right, the cumulative charge always lies between 0 and 3. The same holds for any vertical bond, scanning downwards.

Clearly the above definition can be generalised by replacing 3 by any other positive integer. There are only 2 valid charge(1) configurations — checkerboard patterns of \oplus and \ominus — so this system has growth rate 1. The growth rate of charge(2) configurations on \mathbb{Z}^d was proved in [3] to be 2^{2-d} . Charge(3) is the first of this family of models for which the growth rate is unknown — indeed, previous works have been unable to exactly determine its first digit. Our results strongly support the following surprising conjecture:

Conjecture 1. *The charge(3) and even models have equal growth rate on \mathbb{Z}^2 .*

E. Model transformations: states on bonds

The exclusion and colouring models all share the property that their constraint is local, acting entirely within the confines of a single cell. This is precisely what we need for the model to be amenable to analysis using corner transfer matrices. As such, we can apply this analysis to these models directly.

However, the even and charge(3) constraints are non-local. Fortunately, we can transform these models so that the constraints can be expressed in a local manner. To do this, we will map configurations of \oplus and \ominus spins on the vertices of the lattice to configurations of *states* that lie on the bonds of the lattice.

- For the even model, we place states on the bonds of the lattice as follows. First label all bonds adjacent to \ominus spins by ‘*p*’. Then for any \oplus spin, bonds on opposite sides of the vertex must take different labels — so a bond with a ‘*p*’ state must lie opposite a bond with an ‘*i*’ state and vice versa.

This means that for any given \ominus spin, if we scan the bonds to its right towards the next \ominus spin, we will pass an even number of \oplus spins and so read an odd-length alternating sequence of bond states $pipi \dots ip$. In particular we can interpret the ‘*p*’ state to mean that we have passed an even number of \oplus spins since the \ominus spin we started from. Similarly the ‘*i*’ state indicates that we have passed an odd number of \oplus spins since the \ominus spin. The same will occur if we scan downwards. The notation *p* and *i* is taken from the French *pair* and *impair* for even and odd.

- Consider a valid charge(3) configuration of spins. To each bond in the lattice we assign a state between 0 and 3. Around a \oplus spin, the state of the south and east bonds are 1 more than those of the north and west bonds respectively. Around a \ominus spin we do similarly, except that the south and east bond states are 1 less than the north and west bond states. In this way the bond states represent the cumulative charge reading left to right or top to bottom along the lines of the lattice.

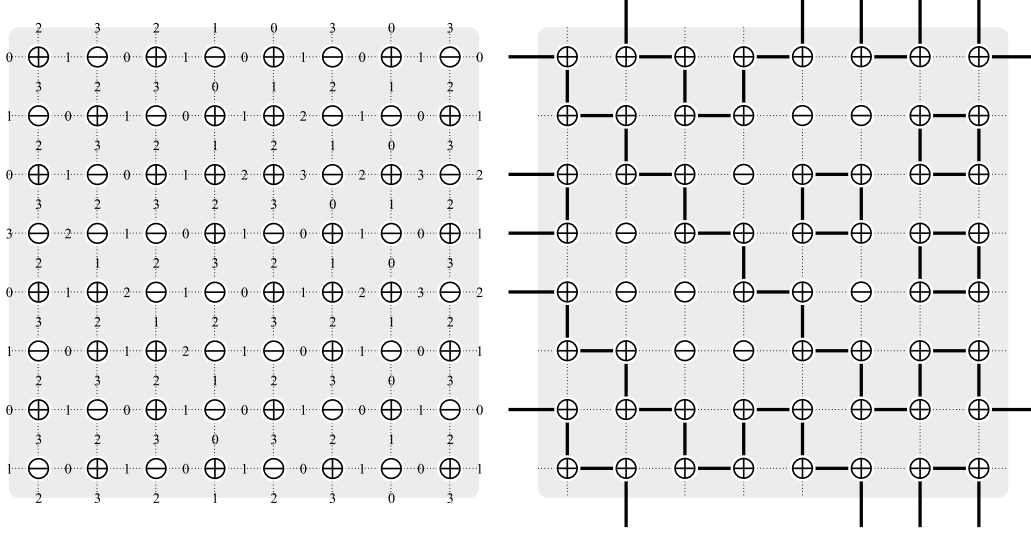


Fig. 2. An example of the Q -charge (left) and p - i (right) models. For the p - i model we have represented bonds in state p as being vacant, while those in state i are drawn as occupied bonds.

The mapping for the even model defines a ‘ p - i ’ model on the bonds of the lattice. A labelling of the bonds by p and i will be valid when there are no two adjacent i states along any horizontal or vertical line, and further any vertex must be adjacent to zero or two i states. Thus the mapping transforms the non-local even constraint into a local p - i constraint.

For the charge(3) model the mapping defines what we will refer to as the ‘ Q -charge’ model on the square lattice (Q being traditional notation for charge in electrostatics). A labelling of the bonds by 0, 1, 2, 3 will be a valid Q -charge configuration when around any vertex, the difference between the north and south bounds is the same as that between the west and east bonds and equal to ± 1 . That is,

$$b \begin{matrix} a \\ \otimes \\ c \\ d \end{matrix} = \begin{cases} \text{valid} & \text{if } d - a = c - b = \pm 1 \\ \text{invalid} & \text{otherwise.} \end{cases} \quad (4)$$

Again, this mapping transforms the non-local charge(3) constraint into a local Q -charge constraint.

The mappings to p - i and Q -charge configurations are not bijective. In each case a single valid configuration of \ominus and \oplus may map to many different p - i or Q -charge configurations. However, it is easy to show that this factor does not affect the growth rate.

Lemma 2. *The growth rate of the even model is equal to that of the p - i model. Similarly, the growth rate of the charge(3) and Q -charge models are equal.*

Proof: The mapping defined above shows that $\kappa_{\text{even}} \leq \kappa_{p-i}$. To prove the reverse inequality, consider both constraints on the lattice $[1, N]^2$ (including boundary bonds).

Starting with a valid even configuration, if each row and column contains at least one \ominus spin then the p - i configuration is uniquely determined. On the other hand, if a given row or column has no \ominus then the bonds it contains can be labelled $pipipi\dots$ or $ipipip\dots$. So on this finite lattice every even configuration corresponds to at most 2^{2N} p - i configurations. Since there are N^2 vertices in this lattice, the growth rates are equal in the $N \rightarrow \infty$ limit.

The proof for charge(3) and Q -charge is similar, except that the state of the first bond in each row/column may have up to 3 possible values. We note that the ‘worst’ possible charge(3) configuration — a checkerboard of \ominus and \oplus — will have 3^{2N} corresponding Q -charge configurations. ■

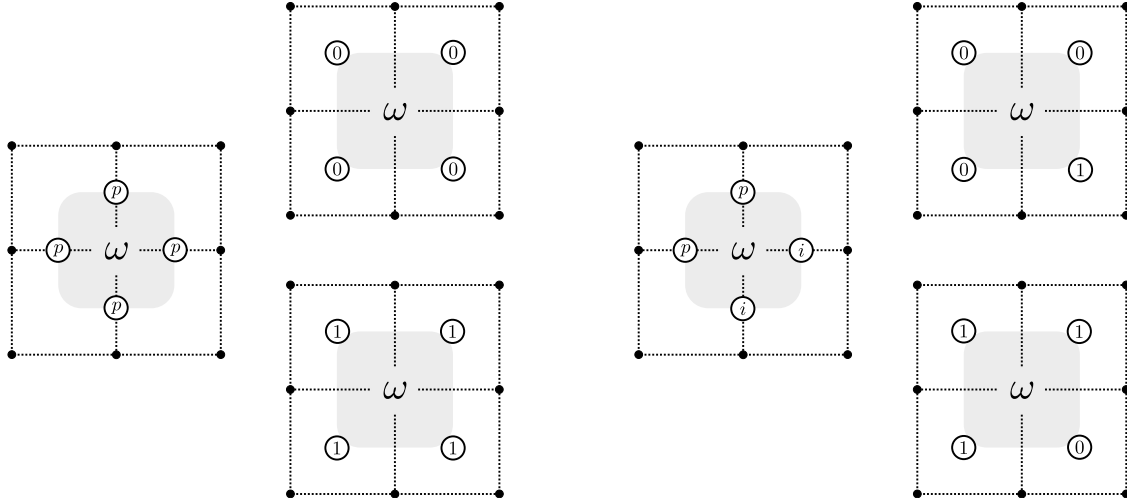


Fig. 3. Possible cells in the p - i and even-face models, up to rotation.

F. Model transformations: states on faces

We make a further transformation from the p - i model to another model (which we call *even-face*), which places states on the centres of the faces of the lattice. These states can take the values 0 or 1. It is not necessary to make this transformation, but empirically it responds better to the corner transfer matrix renormalisation group method.

Given a p - i configuration, two states on neighbouring faces in the corresponding even-face configuration are equal if they are separated by a p bond, and unequal otherwise. A full enumeration of the possible p - i configurations shows that the resulting configuration is internally consistent (see Fig. 3). Now we determine the validity of a configuration according to each cell:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{cases} \text{invalid} & \text{if } a + b + c + d = 2 \\ \text{valid} & \text{otherwise.} \end{cases} \quad (5)$$

Again, this mapping is not a bijection, but determining the value of one state on an arbitrary face in the even-face model will fully determine the rest of the configuration. As such, the number of even-face configurations is exactly twice the number of p - i configurations, and their growth rates are equal.

III. CORNER TRANSFER MATRIX FORMALISM

A. Computing a rigorous lower bound

Consider, for the sake of exposition, a model in which spins lie on the vertices of \mathbb{Z}^2 and can take two values. Further, the model is governed by an ‘interaction round a face’ ω that is symmetric across a vertical line. This includes the exclusion models (hard squares, NAK and RWIM), the colouring models and the even-face model (since the square lattice is self-dual). More work is required for the charge model and we discuss it further below. We wish to count the number of configurations, and so take

$$\omega \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{cases} 1 & \text{if the local configuration is valid} \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

First consider a strip of the square lattice of height m with cylindrical boundary conditions identifying the top and bottom edges. Let V be the $2^m \times 2^m$ column transfer matrix defined by

$$V_{\sigma, \tau} = \prod_{i=1}^m \omega \begin{pmatrix} \sigma_{i+1} & \tau_{i+1} \\ \sigma_i & \tau_i \end{pmatrix} \quad (7)$$

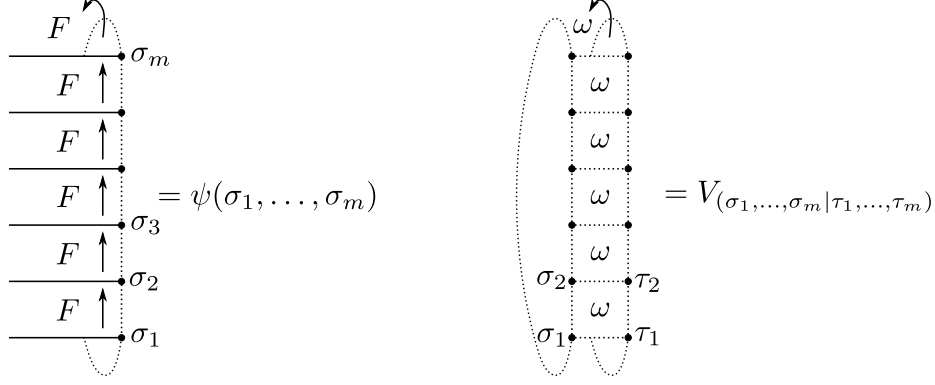


Fig. 4. We define the 2^m elements of the vector ψ to be the trace of the product of F matrices. The elements of the column transfer matrix V are products of the face weight ω . In both cases we apply cylindrical boundary conditions so that $\sigma_{m+1} \equiv \sigma_1$ and $\tau_{m+1} \equiv \tau_1$.

where $\sigma_{m+1} \equiv \sigma_1$ and similarly for τ . See Fig. 4. Note that this is just the traditional column transfer matrix and is symmetric.

Since V is symmetric, we can apply a variational principle — namely, the dominant eigenvalue $\Lambda(m)$ is given by

$$\Lambda(m) = \max_{\psi} \frac{\psi^T V \psi}{\psi^T \psi}. \quad (8)$$

The growth rate of the system is the limit

$$\kappa = \lim_{m \rightarrow \infty} \Lambda(m)^{1/m}. \quad (9)$$

To obtain a lower bound, it suffices to substitute a particular ψ into (8), which we do using Baxter's corner transfer matrix ansatz.

Let $\{F(a, b)\}$ be a set of four $n \times n$ matrices indexed by two spin values $a, b \in \{0, 1\}$ so that $F(a, b) = F^T(b, a)$. While some of these matrices may be zero (e.g. for the hard squares model $F(1, 1) = \mathbf{0}$), they should not all be trivial. Consider the vector of dimension 2^m which satisfies

$$\psi_{\sigma} = \text{Tr} [F(\sigma_1, \sigma_2) F(\sigma_2, \sigma_3) \dots F(\sigma_m, \sigma_1)]. \quad (10)$$

This is shown in Fig. 4. For any particular choice of F , we obtain a vector ψ and so a lower bound for the maximisation problem. Now, given a choice of F , we need to be able to compute $\psi^T V \psi$ and $\psi^T \psi$ and so our lower bound.

Let us fix m and the matrices F . We denote the n^2 entries of $F(a, b)$ as $F(a, b)_{\alpha, \beta}$ for $\alpha, \beta \in \{1, \dots, n\}$. It is possible to rewrite the product of $\psi^T \psi$ as the trace of a power of a new matrix, R . This is extremely helpful because R is not dependent on the height of the strip. Instead, the height dependence is confined to the power of the matrix. In particular, we may write

$$\psi^T \psi = \text{Tr}(R^m) \quad (11)$$

where R is a $(2n^2) \times (2n^2)$ matrix defined by

$$R_{(\alpha|a|\alpha'), (\beta|b|\beta')} = F(a, b)_{\alpha, \beta} F(a, b)_{\alpha', \beta'}. \quad (12)$$

See Fig. 5 for an illustration of this. To see this consider

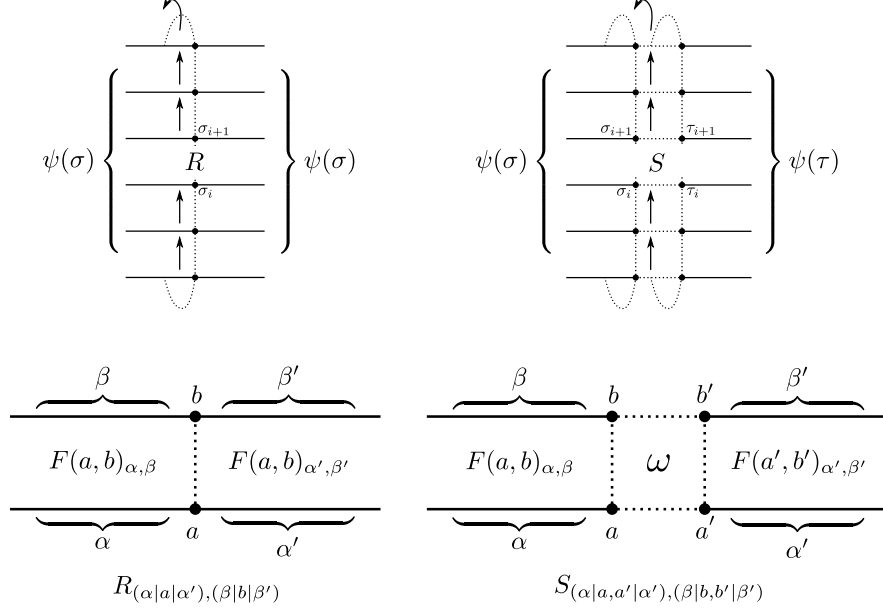


Fig. 5. In order to compute $\psi^T \psi$ (top-left) and $\psi^T V \psi$ (top-right), we define two matrices R and S so that $\psi^T \psi = \text{Tr}(R^m)$ and $\psi^T V \psi = \text{Tr}(S^m)$. The elements of R and S are defined as indicated (bottom-left and bottom-right respectively).

$$\psi^T \psi = \sum_{\sigma} \left(\text{Tr} \prod_{i=1}^m F(\sigma_i, \sigma_{i+1}) \right)^2 \quad (13a)$$

$$= \sum_{\sigma} \left(\sum_{\alpha_1, \dots, \alpha_m} \prod_{i=1}^m F(\sigma_i, \sigma_{i+1})_{\alpha_i, \alpha_{i+1}} \right)^2 \quad (13b)$$

$$= \sum_{\sigma} \sum_{\alpha_1, \dots, \alpha_m} \sum_{\alpha'_1, \dots, \alpha'_m} \prod_{i=1}^m F(\sigma_i, \sigma_{i+1})_{\alpha_i, \alpha_{i+1}} F(\sigma_i, \sigma_{i+1})_{\alpha'_i, \alpha'_{i+1}}. \quad (13c)$$

Now rewrite the summands as elements of R

$$= \sum_{\sigma} \sum_{\alpha_1, \dots, \alpha_m} \sum_{\alpha'_1, \dots, \alpha'_m} \prod_{i=1}^m R_{(\alpha_i | \sigma_i | \alpha'_i), (\alpha_{i+1} | \sigma_{i+1} | \alpha'_{i+1})} \quad (13d)$$

$$= \text{Tr}(R^m). \quad (13e)$$

In a similar way,

$$\psi^T V \psi = \text{Tr}(S^m) \quad (14)$$

where S is a $(4n^2) \times (4n^2)$ matrix defined by

$$S_{(\alpha|a, a'|\alpha'), (\beta|b, b'|\beta')} = F(a, b)_{\alpha, \beta} \cdot \omega \begin{pmatrix} b & b' \\ a & a' \end{pmatrix} \cdot F(a', b')_{\alpha', \beta'}. \quad (15)$$

Note that due to the symmetry $F(a, b) = F^T(b, a)$, both R and S are symmetric.

Now let ξ and η be the maximal eigenvalues of R and S respectively. Our lower bound becomes

$$\Lambda(m) \geq \frac{\text{Tr}(S^m)}{\text{Tr}(R^m)} \Rightarrow \kappa = \lim_{m \rightarrow \infty} \Lambda(m)^{1/m} \geq \frac{\eta}{\xi}. \quad (16)$$

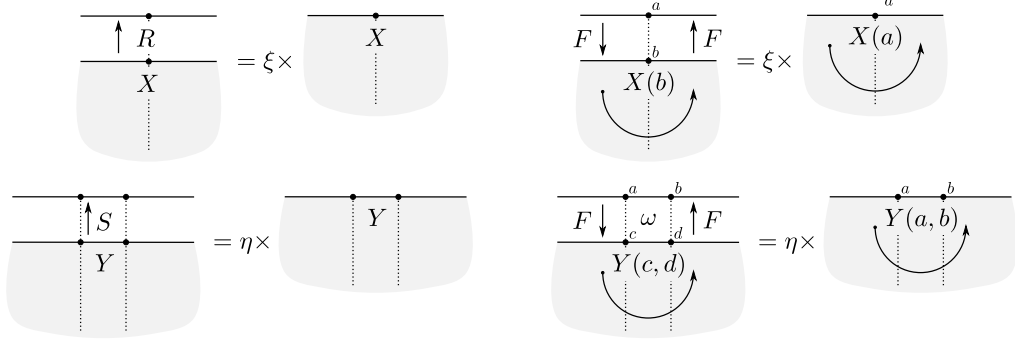


Fig. 6. (left) X and Y are the dominant eigenvectors of the matrices R and S (respectively) with eigenvalues ξ and η (respectively). (right) By recasting X and Y as matrices, we can rewrite the eigenvalue equations as matrix equations in terms of F as shown.

So to compute our bound it suffices to compute the dominant eigenvalues of R and S .

Let X be the dominant eigenvector of R , so that

$$\xi X = RX. \quad (17)$$

See Fig. 6. It is helpful in what follows to recast X as a set of matrices, and the corresponding eigenvalue equation as a matrix equation. The vector X has dimension $2n^2$ and we index its elements as $X_{(\alpha|a|\beta)}$. Define $X(a)_{\alpha,\beta} = X_{(\alpha|a|\beta)}$. Then we can write the eigenvalue equation as

$$\xi X(a)_{\alpha,\alpha'} = (RX)_{(\alpha|a|\alpha')} \quad (18a)$$

$$= \sum_{\beta,b,\beta'} R_{(\alpha|a|\alpha'),(\beta|b|\beta')} X_{(\beta|b|\beta')} \quad (18b)$$

$$= \sum_{\beta,b,\beta'} F(a,b)_{\alpha,\beta} F(a,b)_{\alpha',\beta'} X_{(\beta|b|\beta')} \quad (18c)$$

$$= \sum_{\beta,b,\beta'} F(a,b)_{\alpha,\beta} X_{(\beta|b|\beta')} F(b,a)_{\beta',\alpha'} \quad (18d)$$

$$= \left(\sum_b F(a,b) X(b) F(b,a) \right)_{\alpha,\alpha'} \quad (18e)$$

and so

$$\xi X(a) = \sum_b F(a,b) X(b) F(b,a). \quad (19)$$

Now define Y to be the dominant eigenvector of S , so $\eta Y = SY$. It is of dimension $4n^2$, and we index its elements as $Y_{(\alpha|a,b|\beta)}$. Again, partition these as $Y(a,b)_{\alpha,\beta} = Y_{(\alpha|a,b|\beta)}$. With similar working, we have the eigenvalue equation

$$\eta Y(a,b) = \sum_{c,d} \omega \begin{pmatrix} a & b \\ c & d \end{pmatrix} F(a,c) Y(c,d) F(d,b). \quad (20)$$

To compute ξ and η , one could apply the power method to the matrices R and S using a random initial vector. However, since R and S have dimension $O(n^2)$, each matrix-vector multiplication (and so each iteration) would take $O(n^4)$ operations. Instead it is more efficient to apply the power method implicitly using (19) and (20).

Consider equation (19). Given a prospective eigenvector $X_i(a)$ with a fixed normalisation, we substitute this into the right-hand side of (19). Given this normalisation, the next iteration $\xi_{i+1}, X_{i+1}(a)$ is now given by the left-hand side of the equation. Each iteration of these equations involves multiplying square matrices of dimension n , and so requires $O(n^3)$ operations (or smaller with the use of more sophisticated algorithms).

We can use any value of $X(a)$ to start the power method, but the CTMRG method below will suggest an appropriate starting value. To cover the possibility that this starting value is a sub-dominant eigenvector, we perturb



Fig. 7. In Fig. 6 the eigenvectors X and Y were recast as matrices. Now we rewrite them again as products of the matrix F together with the corner transfer matrix A . The matrix F sweeps through a half-row or half-column and the matrix A sweeps through a quarter of the lattice.

it slightly before starting the power method. η (and Y) is found in an identical manner, except this perturbation is not needed (the result will still give a lower bound on η and so κ).

So our process for computing a rigorous lower bound for κ is this: we take any set of F matrices, apply the power method to (19) and (20) to calculate η and ξ , and then our lower bound is η/ξ .

Of course, the efficacy of this approach depends critically on the set of F matrices — thus we now need a heuristic to find a ‘good’ set of F matrices. This is achieved by use of the corner transfer matrix renormalisation group algorithm, which we detail in the next section.

We note that the method described in Sections 3 and 6 of [3] is similar in spirit to that described above, in that they both replace the eigenvector of the transfer matrix with an expression that simplifies calculations.

B. Corner transfer matrix renormalisation group method

We now wish to choose a set of F matrices which provide a good lower bound, i.e. as large as possible. We do this using a heuristic algorithm known as the corner transfer matrix renormalisation group (CTMRG) method. Note that while the choice of F alters the quality of the bound, it does not affect the fact that it is a lower bound, and as such it is unnecessary to prove any convergence properties of the CTMRG method.

Since we wish our lower bound to be as large as possible, we wish to maximise η/ξ with respect to F . The derivation of equations which ensure this is long and complicated, and we shall not repeat it here. It is sufficient [5] to note that if we have a set of $n \times n$ matrices $A(a)$ which satisfy

$$X(a) = A^2(a) \qquad Y(a, b) = A(a) F(a, b) A(b) \qquad (21)$$

then κ is stationary with respect to F . Note that defining A in this way assumes that the model is invariant under rotation by $\pi/2$. For models without this symmetry, we define several different corner transfer matrices — see the next section. Unfortunately (21) does not prove that κ is maximal, merely stationary, though empirically this does appear to be the case. See Fig. 7 for a graphical interpretation of the above equations. Unlike the more familiar column and row transfer matrices, the matrix $A(a)$ (the eponymous corner transfer matrix) adds the weight of one quadrant or corner of the lattice.

The CTMRG method calculates matrices which converge to the finite-size solution of (19), (20) and (21), known collectively as the *CTM equations*. As such, the F matrices that it produces can be used in our calculations above to provide (what appears to be) a very tight lower bound.

To gain an intuition for the CTMRG method, consider $A(a)$ and $F(a, b)$ to be infinite-size transfer matrices (with an appropriate normalisation) which represent a corner and half-row of the entire plane respectively. It is easily seen that with X and Y determined according to (21), they represent (normalised) half-plane transfer matrices. In this case, (19) and (20) are seen to be satisfied by this choice of F , X and Y . So the infinite-dimensional solution of the CTM equations can be thought of as a set of transfer matrices.

In order to approximate this infinite-dimensional solution, we can construct initial A and F matrices as the corner and half-row transfer matrices of a lattice of $(2p + 1) \times (2p + 1)$ spins. These matrices have dimension $2^p \times 2^p$. We can now bootstrap to a larger lattice by constructing the block matrices

$$A_l(c)|_{d,a} = \sum_b \omega \begin{pmatrix} a & b \\ c & d \end{pmatrix} F(d, b) A(b) F(b, a) \qquad (22a)$$

$$F_l(d, c)|_{b,a} = \omega \begin{pmatrix} a & b \\ c & d \end{pmatrix} F(b, a) \qquad (22b)$$

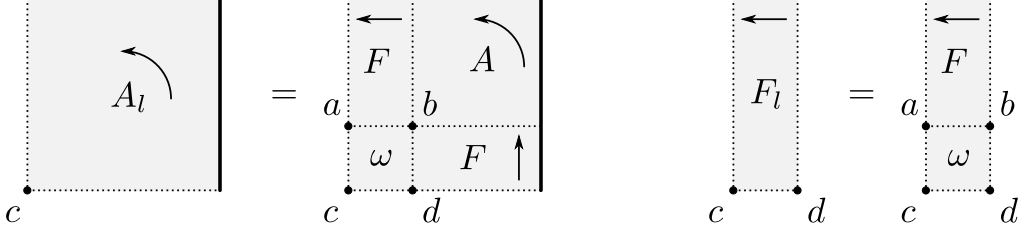


Fig. 8. Expansion of transfer matrices.

each having dimension $2^{p+1} \times 2^{p+1}$. From Fig. 8, it is apparent that A_l and F_l are the corner and half-row transfer matrices of a lattice of $(2p+3) \times (2p+3)$ spins. By repeating this process indefinitely, we approach the infinite-size transfer matrices which we require.

At any p , we can compute the number of configurations on the lattice of size $(2p+1) \times (2p+1)$ as $Z_{2p+1,2p+1} = \sum_a \text{Tr}(A^4(a))$. Unfortunately, since the size of these matrices grows exponentially, this is not a practical approach for even moderate p . It does, however, suggest that we should examine the eigenvalues of A . If the eigenspectrum of A decays quickly, then we can approximate this trace by using only the largest few eigenvalues of A .

To manipulate the eigenvalues directly, we diagonalise the matrix A_l . Observe that the CTM equations are invariant under the similarity transforms

$$A(a) \mapsto P^T(a) A(a) P(a) \quad F(a, b) \mapsto P^T(a) F(a, b) P(b) \quad (23a)$$

$$X(a) \mapsto P^T(a) X(a) P(a) \quad Y(a, b) \mapsto P^T(a) Y(a, b) P(b) \quad (23b)$$

for orthogonal matrices $P(a)$. This means that we can take $A_l(a)$ to be diagonal, with entries ordered from largest to smallest. We throw away the smaller eigenvalues by truncating our large matrices to size $n \times n$. Since the discarded eigenvalues are small, the effect of this truncation on the trace (and so our estimate of $Z_{2p+1,2p+1}$) will also be small. Empirically we also find that the resulting matrices are close to the $n \times n$ solution of the CTM equations.

Ideally, we would find approximate solutions at size $n \times n$ by using (22) to construct very large transfer matrices, then transform and truncate these transfer matrices according to (23). However, since the matrices double in size for each use of (22), this quickly becomes impractical.

Instead, we transform and truncate the matrices every time we expand. Since we may choose exactly how many eigenvalues to discard, we can take matrices of size $n \times n$, expand them to size $2n \times 2n$ and then truncate back down to size $n \times n$. In this way we can produce a sequence of $n \times n$ matrices which appear to converge to the finite-size solution of the CTM equations.

As far as we know, it is not proved that this procedure will converge at all, let alone to the finite-size solution. However, this happens empirically, and as stated above such a proof is not necessary to establish a rigorous lower bound — we can simply observe that the resulting F matrices result in good bounds. In full, the CTMRG method is as follows:

- (1) Start with initial 1×1 matrices $A(a), F(a, b)$ and let $n = 1$.
- (2) Expand A and F using (22).
- (3) If required, increase n by 1.
- (4) Diagonalise $A_l(a)$. Let $P(a)$ be the matrix consisting of the eigenvectors corresponding to the n largest eigenvalues of $A(a)$.
- (5) Reduce A_l and F_l using (23).
- (6) Normalise A and F so that the top-left elements of $A(0)$ and $F(0, 0)$ are both 1.
- (7) Go to step 2.

Once the above procedure has converged sufficiently at a given size n , we can substitute the resulting F into the CTM equations to obtain ξ and η and the lower bound. In fact, we can terminate the procedure at any stage, since any F will result in a bound, but in practice we wait until the procedure has converged to a desired precision.

We calculate our estimates of κ by the formula [5]

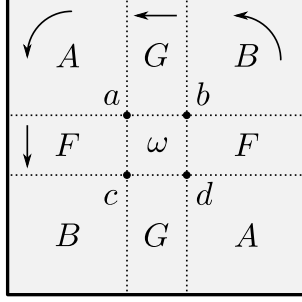


Fig. 9. For a model with symmetry under rotation by π we need two sets of corner transfer matrices A and B as well as half-row and half-column transfer matrices F and G .

$$\kappa = \frac{(\sum_a \text{Tr } A^4(a)) \left(\sum_{a,b,c,d} \text{Tr } \omega \begin{pmatrix} a & b \\ c & d \end{pmatrix} A(a)F(a,c)A(c)F(c,d)A(d)F(d,b)A(b)F(b,a) \right)}{\left(\sum_{a,b} \text{Tr } A^2(a)F(a,b)A^2(b)F(b,a) \right)^2}. \quad (24)$$

Each term in this formula represents the partition function of the entire plane. However, the term in the denominator contains an extra row compared to the first term in the numerator, while the second term in the numerator contains both an extra row and column. The net effect is to cancel out everything but the effect of a single cell. In practice, using this formula is faster and more accurate than the lower bound computation, but of course does not provide a bound.

C. Asymmetry and the CTMRG

The decomposition in (21), and thus the CTMRG, implicitly assumes that the corner transfer matrices are equal under a rotation of $\pi/2$. This naturally occurs if the model itself is invariant under this rotation, which holds true for the majority of the models we analyse. The exceptions are the RWIM and Q -charge models, which are only symmetric under rotation by π . To accommodate this, we define separate half-row and half-column transfer matrices, and two different families of corner transfer matrices — see Fig. 9.

We will use A to denote the corner transfer matrices for the second and fourth quadrants of the lattice, and B for the first and third. Similarly we let F and G denote the half-row and half-column transfer matrices respectively. We expand these matrices in a similar way to (22), taking care with the orientations:

$$A_l(d)|_{b,c} = \sum_a \omega \begin{pmatrix} a & b \\ c & d \end{pmatrix} G(b,a) A(a) F(a,c) \quad F_l(c,a)|_{d,b} = \omega \begin{pmatrix} a & b \\ c & d \end{pmatrix} F(d,b) \quad (25a)$$

$$B_l(c)|_{d,a} = \sum_b \omega \begin{pmatrix} a & b \\ c & d \end{pmatrix} F(d,b) B(b) G(b,a) \quad G_l(d,c)|_{b,a} = \omega \begin{pmatrix} a & b \\ c & d \end{pmatrix} G(b,a). \quad (25b)$$

To reduce the matrices, we observe that we can represent the transfer matrix of the entire plane as either $(AB)^2$ or $(BA)^2$, and these are not necessarily equal. In order to diagonalise both these matrices and simultaneously preserve the CTM equations, we define $P_{AB}(a)$ and $P_{BA}(a)$ to be the matrices consisting of the first n eigenvectors of $A_l(a)B_l(a)$ and $B_l(a)A_l(a)$ respectively. Our reducing transformations then become

$$A(a) \mapsto P_{AB}^T(a) A_l(a) P_{BA}(a) \quad F(a,b) \mapsto P_{BA}^T(a) F_l(a,b) P_{BA}(b) \quad (26a)$$

$$B(a) \mapsto P_{BA}^T(a) B_l(a) P_{AB}(a) \quad G(a,b) \mapsto P_{AB}^T(a) G_l(a,b) P_{AB}(b). \quad (26b)$$

The remainder of the method is unchanged.

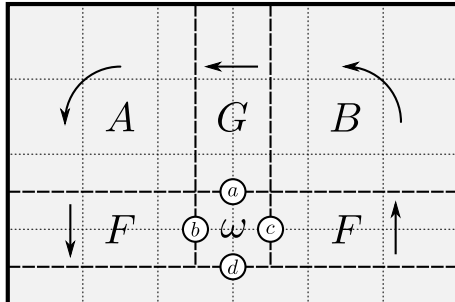


Fig. 10. For a model with bond states we define the corner transfer matrices A and B , and the half-row and half-column matrices F and G as shown. The edges of the lattice are shown in thin dotted lines. The bond-states surrounding the weight ω are shown as labelled white circles.

D. States on bonds

The CTMRG is defined on a model where the states are located on the vertices of the \mathbb{Z}^2 lattice. As this lattice is self-dual, it can be applied without change to a model where the states are on the faces. However, we must make some small adjustments when the states lie on the bonds, as for the Q -charge model. As this model is not invariant under a rotation of $\pi/2$, we will also use the machinery developed in the previous section.

In order to apply the CTMRG to this form of model, we use a form of the method similar to that developed by Foster and Pinettes [25], [13]. The face weight ω becomes a weight around a vertex, and we shift all the matrices by half a length unit (see Fig. 10). This also has the effect that the corner transfer matrices are no longer parameterised by a state — so we only have one A and one B . Similarly the half-row and half-column matrices, $F(a)$ and $G(a)$, are now parameterised by a single state.

We make obvious adjustments to the expansion and contraction of the matrices. These now become

$$A_l|_{c,d} = \sum_{a,b} \omega \begin{pmatrix} a & c \\ b & d \end{pmatrix} G(a) A F(b) \quad F_l(b)|_{d,a} = \sum_c \omega \begin{pmatrix} a & c \\ b & d \end{pmatrix} F(c) \quad (27a)$$

$$B_l|_{d,b} = \sum_{a,c} \omega \begin{pmatrix} a & c \\ b & d \end{pmatrix} F(c) B G(a) \quad G_l(d)|_{c,b} = \sum_a \omega \begin{pmatrix} a & c \\ b & d \end{pmatrix} G(a) \quad (27b)$$

$$A \mapsto P_{AB}^T A_l P_{BA} \quad F(a) \mapsto P_{BA}^T F_l(a) P_{BA} \quad (27c)$$

$$B \mapsto P_{BA}^T B_l P_{AB} \quad G(a) \mapsto P_{AB}^T G_l(a) P_{AB} \quad (27d)$$

where the face weight ω is given by (4).

Note that the corner transfer matrix formalism which leads to the rigorous lower bound only requires that the original column transfer matrix V is symmetric. In the case of the charge model, the symmetry of V follows from the fact that if you invert all the vertex spins in a column, then the column is still legal. However, it is no longer the case that $F^T(a) = F(a)$ and our eigenvalue equations must be changed to reflect this:

$$\xi X = \sum_a F(a) X F^T(a) \quad (28a)$$

$$\eta Y(a) = \sum_{b,c,d} \omega \begin{pmatrix} a & c \\ b & d \end{pmatrix} F(b) Y(d) F^T(c). \quad (28b)$$

In practice, applying the power method using these equations can be unstable for the Q -charge model for reasons we have not been able to determine. In any case, we are able to obtain significantly better bounds for this model through a relation to the even model (see Section V).

IV. RESULTS

The CTMRG algorithm described in the previous section was implemented in C++ using the Eigen numerical linear algebra library [26]. The main advantage of Eigen over other libraries is that it readily supports multiple

precision computations through the MPFR library [27]¹. The eigenvalues of the corner transfer matrices range over many orders of magnitude, so all computations must be done at very high precision. For example, for the hard squares model, the eigenvalues of $A(0)$ and $A(1)$ at size 50×50 range from 2^0 down to approximately 2^{-50} .

All of the computations were run on modest desktop computers running Linux and took between a few hours and a few days to complete. The fastest computation (overall) was hard squares and the slowest was charge(3). Indeed, our most precise results were for hard squares and the least precise were for charge(3).

Matrix diagonalisation requires $O(n^3)$ operations, and so a full iteration of the algorithm described in Section III-B also requires $O(n^3)$ operations. We have observed that for most of the models, the number of iterations required to converge at size n does not increase with n . However, since our computer code waits until convergence at every size before proceeding, we observe that the algorithm is roughly $O(n^4)$. Note however, that this is the time taken to converge at a fixed matrix size, not to a fixed number of digits in κ .

The rate of convergence of our estimates and bounds depends upon the eigenspectrum of the corner transfer matrices. For the purposes of this discussion we will assume that the matrices at each size have converged to their limiting value under the renormalisation procedure. A key observation is that the largest eigenvalues do not change much with matrix size — the 4 eigenvalues of $A(0)$ at size 4 are nearly the same as the 4 dominant eigenvalues of $A(0)$ at size 40, and presumably at size 400.

Thus, what changes as we increase matrix size are the eigenvalues we discard by truncating the matrices. Since each term in our computation of κ (via (24)) is the trace of the product of 4 corner transfer matrices (some interlaced with half-row and half-column transfer matrices), we expect that the contribution of each eigenvalue to the estimate of κ is roughly proportional to its fourth power.

The question of the tightness of the bound is a little more complicated, as it depends on how close our F matrices are to the solution of the CTM equations. If they exactly solve the CTM equations at finite size (for some values of the other matrices), then there is no difference between our bound calculation using (16) and our estimate using (24), and so the error should be the same. As we run the algorithm to convergence at each finite size, we expect that the extra error induced by having a not-quite-optimal F is minimal.

This analysis shows that the utility of the method hinges on the distribution of the eigenvalues of the corner transfer matrices. If the eigenvalues decay slowly, then by truncating the matrices we will throw away quite large contributions to the estimates and bounds of κ . In such a situation, we would expect our estimates and bounds to converge slowly with increasing matrix size. On the other hand, if the eigenvalues decay quickly then the truncated eigenvalues will have little impact.

It is believed [28] that the eigenvalues of the corner transfer matrices scale as

$$\lambda_k \sim \exp(-c(\ln k)^2). \quad (29)$$

Here we have denoted the k^{th} largest eigenvalue by λ_k ; since these do not vary significantly between different matrix sizes, we have not specified the dimension of the underlying matrix.

If the above scaling form holds true — and our observations suggest that it does (see Fig. 11) — then we expect the error in our estimates to have a similar asymptotic form with increasing matrix size. Thus, to achieve an error approximately ε , we need to take matrices of size $k \times k$ with

$$k \sim \exp\left(\alpha\sqrt{-\log(\varepsilon)}\right). \quad (30)$$

Substituting $\varepsilon = 2^{-N}$ gives

$$k \sim \exp\left(\beta\sqrt{N}\right). \quad (31)$$

Since the running time and memory required grow polynomially with matrix size, the above argument supports the assertion that the time and memory required by the algorithm to correctly compute N digits grows subexponentially with N .

In Table IV, we list our lower bounds and estimates for the growth rates of the models we have considered. Due to the instability of the bound calculation for the charge(3) model, we only give the estimate for this model. Note

¹At the time of writing the Eigen and MPFR libraries were available from <http://eigen.tuxfamily.org> and <http://www.mpfr.org/>.

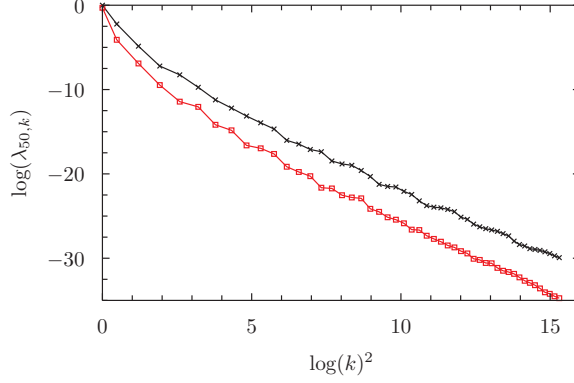


Fig. 11. A plot of $\log(\lambda_k)$ vs $\log(k)^2$, for the eigenvalues of $A(0)$ (upper curve) and $A(1)$ (lower curve) at size 50 for the hard squares model. For larger k , the plots appear approximately linear, supporting the conjectured scaling form of (29).

TABLE I

LOWER BOUNDS AND ESTIMATES FOR THE GROWTH RATES OF THE VARIOUS MODELS. FOR EACH MODEL WE GIVE THE FULL DECIMAL EXPANSION OF THE LOWER BOUND, AND THEN WE GIVE THE LAST FEW DIGITS OF OUR ESTIMATE IN BRACKETS. THESE REPLACE THAT MANY DIGITS IN THE LOWER BOUNDS — THE PRECEDING DIGITS ARE IDENTICAL. WE UNDERLINE THE DIGITS WHICH AGREE WITH THE PREVIOUS BEST KNOWN LOWER BOUNDS (REFERENCES GIVEN UNDERNEATH THE MODEL NAME).

Model	Matrix size	Lower bound on (and estimate of) κ
Hard squares [29]	256	<u>1.503 048 082 475 332 264 322 066 329 475</u> 553 689 385 781 038 610 305 062 028 101 735 933 850 396 923 440 380 463 299 47 (65)
NAK [3]	256	<u>1.342 643 951</u> 124 601 297 851 730 161 875 740 395 719 438 196 938 393 943 434 885 455 0 (1)
RWIM [3]	128	<u>1.448 957 371 775 608 489 872 231 406 108</u> 136 686 434 371 (7)
Even [3]	128	<u>1.357 587 502 184 123</u> (5)
Charge(3) [3]	74	(<u>1.357 587 50</u>)
4-Colouring [22]	96	<u>2.336 056 641 041 133 656 814 01</u> (4)
5-Colouring [22]	64	<u>3.250 404 923 167 119 143 819 73</u> (6)

that since RWIM is asymmetric, we computed the bounds and estimates both using (19) and (20) as stated, and also with F replaced by G ; they gave the same bounds (to the precision stated in the table).

To show how much our bounds are an improvement on existing bounds, we have underlined the digits which agree with previous best known bounds (after which, our bounds are greater). Of course, rigorously measuring the accuracy of a bound is impossible without exact knowledge of the quantity being estimated, or at least a comparable upper bound. However, comparing the number of digits with which our estimates agree with our bounds and previously known bounds, we are confident that our bounds are a great improvement in all cases.

We note that our estimates agree with all 43 digits of the estimate for the hard squares growth rate given in [16], and agree up to our last 2 digits with a naïve evaluation of the q -colouring series given in [23] with $q = 4, 5$. We have attempted to identify all of these constants using the Inverse Symbolic Calculator [30]. Unfortunately, we were not able to identify any of them.

Our estimates lead to the rather surprising observation that the growth rates of the charge(3) and even models are identical — at least to the 8 significant digits we have computed. This observation led us to Conjecture 1. We discuss this further below.

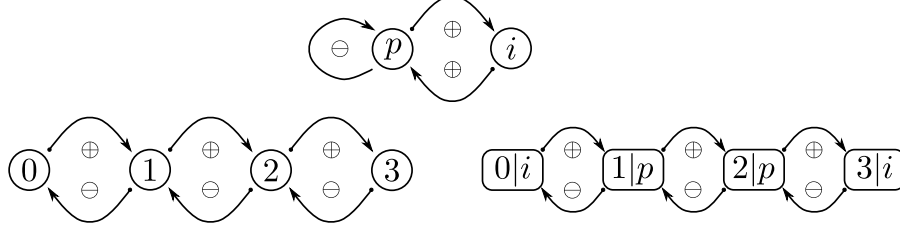


Fig. 12. The allowed transitions for the one-dimensional p - i (top) and Q -charge (left) models. There is a 2-1 graph homomorphism from the Q -charge finite-state graph to that of the p - i model — namely map the nodes 0, 3 to the node i and the nodes 1, 2 to the node p .

V. THE CHARGE(3) AND EVEN MODELS

As noted above, our estimates of the growth rates of the charge(3) and even models agree to 8 significant digits, which is the limit of our accuracy for the charge(3) model. Because of this we conjecture that they are equal. Should this conjecture be true, our lower bound and estimate for the charge(3) model can be extended by a few digits, using values from the even model.

The one-dimensional counterpart to this conjecture can be proved quite readily:

Theorem 3. *For the one-dimensional lattice,*

$$\kappa_{\text{charge}(3)} = \kappa_{\text{even}} = \frac{1 + \sqrt{5}}{2}. \quad (32)$$

Proof: Perhaps the most obvious way to demonstrate this is to construct the appropriate transfer matrices for the Q -charge and p - i models (see Fig. 12). These are

$$V_{p-i} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad V_{Q\text{-charge}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (33)$$

A quick computation shows that the dominant eigenvalue (and hence the growth rate) of each is $\frac{1+\sqrt{5}}{2}$.

It is more instructive, however, to construct a mapping between the two states-on-bonds models. This mapping is a 2-1 graph homomorphism between the two underlying state graphs as illustrated in Fig. 12. Simply map the vertices labelled by 0, 3 to the vertex labelled i , and the vertices labelled 1, 2 to the vertex labelled p .

Now consider any valid Q -charge configuration in one dimension. Using the above mapping we can transform it into a valid p - i configuration. The reverse mapping is 1 to 2. To see this pick a single p in the p - i configuration — we may choose whether this maps to a 1 or a 2 in the corresponding Q -charge configuration. Once this choice is made, the rest of the mapping is fixed. The result follows. ■

The mapping used in the above proof can also be used to show that the growth rate of the charge(3) model in two dimensions is at least that of the even model.

Theorem 4. *For the two-dimensional lattice,*

$$\kappa_{\text{charge}(3)} \geq \kappa_{\text{even}} (\geq 1.357587502184123). \quad (34)$$

Proof: Choose any bond on the lattice, and draw a line parallel to $y = x$ through this bond. Choose all bonds intersecting this line to have Q -charge states 0 or 2, as shown in Fig. 13. Now consider a valid Q -charge configuration conforming to these restrictions. The set of such configurations is a (strict) subset of all valid Q -charge configurations. However, since the bond-states must increase or decrease by 1 as we read across horizontal and vertical lines, we have determined the parity of every bond state on the lattice.

Now in every cell, the parities of the N and W bonds are determined and identical. A full enumeration of the (five) possible cell configurations shows that in this case, the above mapping between the Q -charge and p - i models is a bijection between all valid Q -charge and p - i cell configurations. Thus the number of configurations for each model

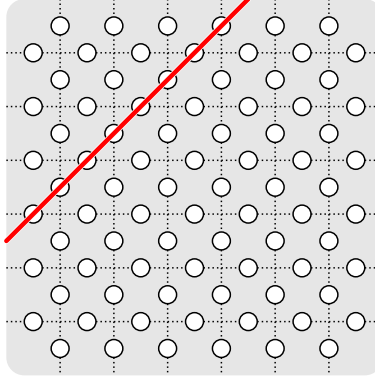


Fig. 13. Restriction for the Q -charge $\leftrightarrow p$ - i mapping. All states on the red line must have state 0 or 2 in the Q -charge model.

is the same. Since we have only taken a subset of valid Q -charge configurations, but all valid p - i configurations, this provides the inequality. \blacksquare

Unfortunately, we have not been able to prove the reverse inequality, though we can make a non-rigorous argument as to why this should be so. In every Q -charge configuration, there will be a number of faces where the N and W bonds have identical parity (the ‘good’ faces), and a number of faces where the N and E bonds have the same parity (the ‘bad’ faces). As mentioned above, it is possible to take configurations which consist entirely of good faces, and these are equivalent to the p - i configurations.

It is also possible to take configurations which consist entirely of bad faces, and a similar analysis of cell configurations shows that these faces map onto (invalid) p - i faces containing exactly one i and three p 's. By considering an i as an ‘occupied’ bond and a p as an ‘vacant’ bond, we see that these configurations are in fact fully-packed dimer configurations. The growth rate of such configurations has been proved [31] to be

$$\kappa = e^{\text{Catalan}/\pi} = 1.338\,515\,15\dots \quad (35)$$

which is less than our proved lower bound for κ_{even} . Therefore, these configurations are exponentially infrequent in the set of Q -charge configurations.

Another possibility is to take configurations where the good faces and bad faces are arranged in a checkerboard pattern. We call this model the ‘elbow’ model, due to the characteristic shape of the individual bond configurations. The growth rate of this model can be rigorously bounded above using the method of Calkin and Wilf [2]. Using transfer matrices of width 8, we calculate $\kappa_{\text{elbow}} \leq 1.333\,521\dots$, which is again less than the proved lower bound for κ_{even} .

Now choose any fixed set of parities for the rows and columns of the lattice. In these configurations, the good faces and bad faces form a ‘tartan’-like configuration (see Fig. 14). The growth rate of such configurations is some combination of the growth rates of the even model (the good faces), the fully-packed dimer model (the bad faces) and the elbow model (an interfacial term between good and bad faces). Out of these three models, the even model has the largest growth rate and so the growth rate of such configurations cannot be greater than κ_{even} .

Now there are $2^{2\sqrt{N}}$ ways to choose these parities on a lattice of N spins, a sub-exponential factor. Each of these ways results in a growth rate not more than κ_{even} , and therefore the set of all valid Q -charge configurations has a growth rate not greater than κ_{even} .

This (admittedly non-rigorous) argument, combined with the accuracy of our numerical results and the above theorem, presents what we find to be a compelling case that $\kappa_{\text{charge}(3)} = \kappa_{\text{even}}$.

VI. CONCLUSION

In this paper, we have constructed very accurate rigorous lower bounds on the growth rate of several families of data storage models: exclusion, colouring, even, and charge(3) models. The process we use also gives us very precise estimates of the growth rates. This is achieved using Baxter’s corner transfer matrix formalism combined with the corner transfer matrix renormalisation group method of Nishino and Okunishi.

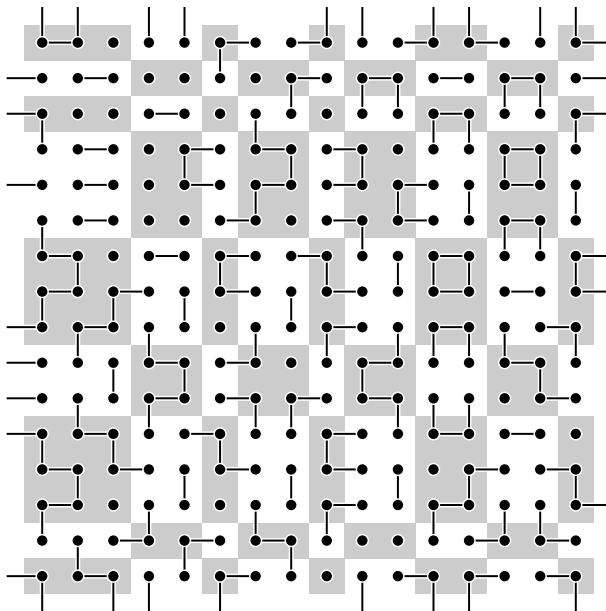


Fig. 14. An example of the Q -charge model mapped to a ‘tartan’ combination of p - i and fully-packed dimer constraints. The shaded regions obey the p - i constraint, while the unshaded regions obey the dimer constraint.

The precision of our bounds allows us to conjecture the equality of the growth rate of two of our models, the charge(3) and even models. We have proved one inequality for this relation, and have a non-rigorous (but overall convincing) argument for equality.

For future work, we would like to be able to construct upper bounds of similar quality. The quality of proved lower bounds have typically been much higher than that of upper bounds. It may be possible to use linear algebra results to obtain a bound on the error of our estimates, and so derive an upper bound for the growth rate. It certainly is possible to bound the distance between our estimate and the nearest eigenvalue of the transfer matrix using a theorem of Wilkinson [32]. However, there is no guarantee that this eigenvalue is the dominant eigenvalue, and so more work needs to be done to establish a provable upper bound.

In addition, there are still a number of constraints which we have not analysed in this paper, for example higher-order charge or colouring models. These are amenable to identical analysis, and are the subject of future work.

ACKNOWLEDGEMENTS

AR acknowledges financial support from NSERC and thanks the University of Vienna for their hospitality during his visit there. YBC acknowledges financial support from the University of Vienna and thanks the University of British Columbia for their hospitality during his visit there. We both thank Brian Marcus for many helpful discussions and proof-reading.

REFERENCES

- [1] K. Engel, “On the Fibonacci number of an $M \times N$ lattice,” *Fibonacci Quarterly*, vol. 28, no. 1, pp. 72–78, 1990.
- [2] N. Calkin and H. Wilf, “The number of independent sets in a grid graph,” *SIAM J. Disc. Math.*, vol. 11, no. 1, pp. 54–60, 1998.
- [3] E. Loidor and B. Marcus, “Improved lower bounds on capacities of symmetric 2D constraints using Rayleigh quotients,” *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1624–1639, 2010.
- [4] R. J. Baxter, “Dimers on a rectangular lattice,” *J. Math. Phys.*, vol. 9, p. 650, 1968.
- [5] —, “Variational approximations for square lattice models in statistical mechanics,” *J. Stat. Phys.*, vol. 19, no. 5, pp. 461–478, 1978.
- [6] R. J. Baxter, I. G. Enting, and S. K. Tsang, “Hard-square lattice gas,” *J. Stat. Phys.*, vol. 22, no. 4, pp. 465–489, 1980.
- [7] R. J. Baxter, “Hard hexagons: exact solution,” *J. Phys. A: Math. Gen.*, vol. 13, p. L61, 1980.
- [8] —, *Exactly solved models in statistical mechanics*. Academic Press London, 1982.
- [9] T. Nishino and K. Okunishi, “Corner transfer matrix renormalization group method,” *J. Phys. Soc. Jpn.*, vol. 65, no. 4, pp. 891–894, 1996.
- [10] —, “Corner transfer matrix algorithm for classical renormalization group,” *J. Phys. Soc. Jpn.*, vol. 66, no. 10, pp. 3040–3047, 1997.
- [11] S. White, “Density matrix formulation for quantum renormalization groups,” *Physical Review Letters*, vol. 69, no. 19, p. 2863, 1992.

- [12] U. Schollwöck, “The density-matrix renormalization group,” *Rev. Mod. Phys.*, vol. 77, no. 1, p. 259, 2005.
- [13] D. P. Foster and C. Pinettes, “A corner transfer matrix renormalization group investigation of the vertex-interacting self-avoiding walk model,” *J. Phys. A: Math. Gen.*, vol. 36, pp. 10 279–10 298, 2003.
- [14] V. V. Mangazeev, M. T. Batchelor, V. V. Bazhanov, and M. Y. Dudalev, “Variational approach to the scaling function of the 2d ising model in a magnetic field,” *J. Phys. A: Math. Theor.*, vol. 42, p. 042005, 2009.
- [15] Y. Chan, “Series expansions from the corner transfer matrix renormalization group method: the hard squares model,” *J. Phys. A: Math. Theor.*, vol. 45, p. 085001, 2012.
- [16] R. J. Baxter, “Planar lattice gases with nearest-neighbor exclusion,” *Ann. Comb.*, vol. 3, no. 2, pp. 191–203, 1999.
- [17] M. Cohn, “On the channel capacity of read/write isolated memory,” *Disc. Appl. Math.*, vol. 56, no. 1, pp. 1–8, 1995.
- [18] M. Golin, X. Yong, Y. Zhang, and L. Sheng, “New upper and lower bounds on the channel capacity of read/write isolated memory,” in *Information Theory, 2000. Proceedings. IEEE International Symposium on*. IEEE, 2000, p. 280.
- [19] W. Weeks IV and R. E. Blahut, “The capacity and coding gain of certain checkerboard codes,” *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1193–1203, 1998.
- [20] N. L. Biggs, “Colouring square lattice graphs,” *Bull. London Math. Soc.*, vol. 9, pp. 54–56, 1977.
- [21] R. Shrock and S.-H. Tsai, “Lower bounds and series for the ground-state entropy of the Potts antiferromagnet on Archimedean lattices and their duals,” *Phys. Rev. E*, vol. 56, pp. 4111–4124, Oct 1997. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevE.56.4111>
- [22] P. H. Lundow and K. Markström, “Exact and approximate compression of transfer matrices for graph homomorphisms,” *LMS J. Comput. Math.*, vol. 11, pp. 1–14, 2008.
- [23] J. Jacobsen, “Bulk, surface and corner free-energy series for the chromatic polynomial on the square and triangular lattices,” *J. Phys. A: Math. Theor.*, vol. 43, p. 315002, 2010.
- [24] E. H. Lieb, “Exact solution of the problem of the entropy of two-dimensional ice,” *Phys. Rev. Lett.*, vol. 18, pp. 692–694, Apr 1967. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevLett.18.692>
- [25] D. P. Foster and C. Pinettes, “Corner-transfer-matrix renormalization-group method for two-dimensional self-avoiding walks and other $O(n)$ models,” *Phys. Rev. E*, vol. 67, no. 045105, 2003.
- [26] G. Guennebaud, B. Jacob *et al.*, “Eigen v3,” <http://eigen.tuxfamily.org>, 2010.
- [27] L. Fousse, G. Hanrot, V. Lefèvre, P. Pélissier, and P. Zimmermann, “MPFR: A multiple-precision binary floating-point library with correct rounding,” *ACM Transactions on Mathematical Software*, vol. 33, no. 2, pp. 13:1–13:15, Jun. 2007. [Online]. Available: <http://doi.acm.org/10.1145/1236463.1236468>
- [28] K. Okunishi, Y. Hieida, and Y. Akutsu, “Universal asymptotic eigenvalue distribution of density matrices and corner transfer matrices in the thermodynamic limit,” *Phys. Rev. E*, vol. 59, no. 6, pp. R6227–R6230, 1999.
- [29] S. Friedland, P. H. Lundow, and K. Markström, “The 1-vertex transfer matrix and accurate estimation of channel capacity,” *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3692–3699, 2010.
- [30] J. Borwein, P. Borwein, and S. Plouffe, “Inverse symbolic calculator,” <http://isc.carma.newcastle.edu.au/>.
- [31] P. Kasteleyn, “Dimer statistics and phase transitions,” *J. Math. Phys.*, vol. 4, no. 2, p. 287, 1963.
- [32] J. H. Wilkinson, “Rigorous error bounds for computed eigensystems,” *The Computer Journal*, vol. 4, no. 3, pp. 230–241, 1961.