J Stat Phys (2007) 129: 1117–1135
DOI 10.1007/s10955-007-9431-9

Non-Colliding Paths in the Honeycomb Dimer Model and the Dyson Process

Cédric Boutillier

Received: 23 March 2007 / Accepted: 11 September 2007 / Published online: 13 October 2007

Abstract In this paper we describe a natural family of random non-intersecting discrete paths in the dimer model on the honeycomb lattice. We show that when the dimer model is going to freeze, this family of paths, after a proper rescaling, converges to the extended sine process, obtained traditionally as the limit of the Dyson model when the number of particles goes to infinity.

Keywords Tilings · Dimer models · Phase transition · Random matrices · Dyson model

1 Introduction

The *Dyson model* is a dynamic version of the Gaussian Unitary Ensemble in random matrix theory. It models a collection of identical charged particles diffusing, but repelling one another by electrostatic interaction. The trajectories of these particles are thus forced not to intersect. If we put these particles in an harmonic potential well, the system reaches an equilibrium. When the number of particles grows to infinity, and in equilibrium, the system is invariant by translation in time and in space. It is described by a determinantal process, known as the *extended sine process*.

The *dimer model* belongs to discrete statistical mechanics. This system was introduced in the 1930's [7] to model the adsorption of diatomic molecules on the surface of a crystal, and consists in a probabilistic study of perfect matchings of a graph. It is remarkable that this model is exactly solvable when no supplementary interaction between edges is imposed. The partition function as well as correlation functions can be expressed explicitly, and have a determinantal form: this is a model of discrete free fermions. We will look at a particular case, when the graph is (a subgraph of) the honeycomb lattice G.

It has been noticed several times that dimer models have some properties very similar to those of random matrices. In presence of boundary conditions, the dimer model may exhibit

C. Boutillier (✉)
Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie, Paris VI,
Case courrier 188, 4 Place Jussieu, 75252 Paris Cedex 05, France
e-mail: boutilli@ccr.jussieu.fr

several behaviors. In some regions (called *liquid regions*, usually far from the boundary) correlations decay polynomially, and in some cases, the kernel used to compute correlations between dimers along a given line is the *discrete sine kernel* (see for example [1, 9, 21]). Under a proper scaling, the continuous determinantal sine process is recovered, describing the bulk of large Gaussian Hermitian matrices [3, 9]. In other regions, the configuration is *frozen*: there the configuration is almost deterministic. The behavior of the system near the frontier between a frozen and a liquid region are related to the *Airy process* and thus to the statistics of the largest eigenvalue of a large Gaussian Hermitian matrix [5, 10, 21]. See also [11, 22] for other interesting results enforcing the relations between these objects. That is why dimer models are sometimes considered as discrete analogues of random matrices [18].

In this article, we add to that list another example by relating the dimer model on the infinite honeycomb lattice with the Dyson model. A dimer configuration on the honeycomb lattice G can be described by the paths obtained by following the non-horizontal dimers. The dimer model on G is endowed with a two-parameter family of ergodic Gibbs measure that are conditionally uniform. The two parameters can be seen as the drift of a path, and the average vertical distance between two successive paths.

We show that the *extended sine process*, obtained at the limit of the Dyson process in the bulk and in the steady regime, when the number of particles becomes infinite, is also obtained as the limit of Gibbs measures on dimer configurations of the infinite honeycomb lattice, after a proper scaling, when the drift is zero and the average distance goes to infinity.

Theorem 1 *The family of discrete non-colliding paths in the honeycomb dimer model converges weakly, after a proper scaling, to the extended sine process.*

The precise meaning of this convergence is explained in the beginning of Sect. 4. A possible strategy is to use Lindström–Gessel–Viennot's approach [8] (see also [12]) to describe the combinatorics of the non-intersecting paths in terms of determinants. However, we will make use of the technology introduced by Kasteleyn [13] and developed in [17, 19] to compute correlations between dimers.

The Dyson model, and thus the extended sine process can be obtained as the scaling limit of another statistical mechanical model, called the *random vicious walks* introduced by Fisher in [6] (see for example [16]). In that model, a finite number of walkers on $\mathbb{Z}$ start at time T_0 from even integers and end at T_1. The measure on these paths is the uniform measure on all configurations without intersections. What differs in this work from the vicious walkers case is that due to the ergodicity property of the Gibbs measures on dimer configurations, the family of path is already at the discrete level infinite, and invariant by (discrete) translations.

The paper is structured as follows: in Sect. 2, we give a brief description of the Dyson model, and the extended sine process, its limit when the number of particles goes to infinity, in the steady regime. In Sect. 3, we present some aspects of the dimer model on the infinite honeycomb lattice and its Gibbs measures, as well as the correspondence with a family of discrete non-colliding paths. In Sect. 4, we prove the weak convergence of the probability measures on the random paths in the dimer model to the extended sine process. The proofs of two propositions used in that section are postponed to Sects. 5 and 6.

2 The Dyson Model and the Sine Process

The *Dyson model* [4] is defined as the motion of n identically charged diffusive particles $\lambda_1(t), \dots, \lambda_n(t)$ repelled one from another by an electrostatic force, but attracted to the origin by an elastic force. This is a Markov process whose evolution kernel $p(t, \lambda, \mu)$ satisfies the following equation

$$\frac{\partial p}{\partial t} = \frac{1}{2}\sum_{j=1}^{k} \frac{\partial^2 p}{\partial \lambda_j^2} - \frac{\partial}{\partial \lambda_j}\left(\frac{\partial \Phi(\lambda)}{\partial \lambda_j} p\right), \tag{1}$$

where the potential $\Phi(\lambda) = -\sum_{j=1}^{n} \lambda_j^2 + \sum_{i<j} \log|\lambda_j - \lambda_i|$. When $n = 1$, this is the Ornstein–Uhlenbeck process.

This process can be interpreted as the evolution of the eigenvalues of a random Hermitian matrix, where the diagonal entries evolve as an Ornstein–Uhlenbeck process with speed 2 and the real and imaginary parts of the entries above the diagonal , as an Ornstein–Uhlenbeck process with speed 1, all these motions being independent. The statistics of the eigenvalues at a fixed time are those of a random matrix from the Gaussian Unitary Ensemble (GUE) [20].

This process is an example of a *determinantal process* where the correlation functions are given by determinants of the *extended Hermite kernel*, a space-time extension of the Hermite kernel appearing in the analysis of the GUE.

If we scale the Dyson process in the bulk when n goes to infinity, we get the *sine process*, a determinantal process associated with the *extended sine kernel* defined as

$$\mathrm{S}(t'-t, y'-y) = \begin{cases} \int_0^1 e^{-(t'-t)\phi^2/2} \cos(\phi(y'-y)) \frac{\mathrm{d}\phi}{\pi}, & \text{if } t' \le t, \\ -\int_1^{+\infty} e^{-(t'-t)\phi^2/2} \cos(\phi(y'-y)) \frac{\mathrm{d}\phi}{\pi}, & \text{if } t' > t. \end{cases} \tag{2}$$

When $t = t'$, this reduces to the usual *sine kernel* $\frac{\sin(y'-y)}{\pi(y'-y)}$.

The sine process describes an infinite system of particles repulsing each other by Coulomb force, and which is homogeneous both in space and time. Such a system has been studied in [4, 15, 24].

Statistics of occupation of the process can be expressed in terms of S: given $\tau_1 < \cdots < \tau_k$ and $(I_j)_{j=1}^k$ a sequence of Borel sets of $\mathbb{R}$, the probability that the number $N_{I_j}(\tau_j)$ of curves crossing I_j equals n_j at time τ_j, for all $j \in \{1, \ldots, k\}$ is given by the following expression

$$\begin{aligned} &\mathbb{P}[\forall j \in \{1, \ldots, k\},\ N_{I_j}(\tau_j) = n_j] \\ &\quad = \sum_{p \in \mathbb{N}^k} \frac{(-1)^{|p|}}{p!\, n!} \int \cdots \int_{(\tau_1, I_1)^{n_1+p_1} \times \cdots \times (\tau_k, I_k)^{n_k+p_k}} \det_{|n+p|} \big(\mathrm{S}(\xi_i - \xi_j)\big) \mathrm{d}^{n+p}\xi, \end{aligned} \tag{3}$$

where ξ_i stands for the variables (τ, y) ranging over the ith factor of the Cartesian product $(\tau_1, I_1)^{n_1+p_1} \times \cdots \times (\tau_k, I_k)^{n_k+p_k}$. In this expression, we used multi-index notations:

$$n! = \prod_{j=1}^{k} n_j!, \qquad |n| = \sum_{j=1}^{k} n_j.$$

It will be convenient in the sequel to use also for $z = (z_1, \ldots, z_k)$

$$z^n = z_1^{n_1} \cdots z_k^{n_k}, \qquad \frac{\partial^n}{\partial z^n} = \frac{\partial^{n_1}}{\partial z_1^{n_1}} \cdots \frac{\partial^{n_k}}{\partial z_k^{n_k}},$$

and

$$\{\tau\} \times I = (\{\tau_1\} \times I_1, \ldots, \{\tau_k\} \times I_k),$$

$$(\{\tau\} \times I)^n = (\{\tau_1\} \times I_1)^{n_1} \times \cdots \times (\{\tau_k\} \times I_k)^{n_k}.$$

The probabilities in (3) can be seen as derivatives of a generating function for the joint distribution of $(N_{I_1}(\tau_1), \ldots, N_{I_k}(\tau_k))$: posing

$$Q(z) = Q_{\{\tau\}\times I}(z) = \mathbb{E}\left[\prod_{j=1}^{k}(1-z_j)^{N_{I_j}(\tau_j)}\right] = \sum_{n\in\mathbb{N}^k} \frac{(-z)^n}{n!}\mathbb{E}\left[\frac{N_{I_j}(\tau_j)!}{(N_{I_j}(\tau_j)-n_j)!}\right]$$

$$= \sum_{n\in\mathbb{N}^k} \frac{(-z)^n}{n!} \int_{(\{\tau\}\times I)^n} \det_{|n|}(\mathrm{S}(\xi_i - \xi_j))\mathrm{d}^n\xi,$$

we have

$$\mathbb{P}[\forall j \in \{1,\ldots,k\},\ N_{I_j}(\tau_j) = n_j] = \frac{(-1)^{|n|}}{n!} \frac{\partial^n}{\partial z^n} Q_{\{\tau\}\times I}(z)\bigg|_{z=(1,\ldots,1)}. \tag{4}$$

Thus, all the information about the $N_{I_j}(\tau_j)$ is encoded in the function $Q_{\{\tau\}\times I}$.

3 The Dimer Model on the Honeycomb Lattice

Let G be the honeycomb lattice. Since all the cycles are even, there is a bipartite coloring of the vertices of G in white and black. The group $\mathbb{Z}^2$ acts by color-preserving translations on G, generated by $\hat{x}$ and $\hat{y}$, and the fundamental domain contains one white vertex $\mathbf{w}_0$ and one black vertex $\mathbf{b}_0$. They are represented in Fig. 1. All the vertices of G are images of vertices of the fundamental domain under some translation of $\mathbb{Z}^2$. We label them by their color and element of $\mathbb{Z}^2$. We thus denote by

$$\mathbf{w}_{x,y} = \mathbf{w}_0 + (x, y), \qquad \mathbf{b}_{x,y} = \mathbf{b}_0 + (x, y).$$

The white vertex and the black vertex in the fundamental domain with coordinates (x, y).

The edges of G are classified in 3 types, a, b and c, according to their orientations around a white vertex (see Fig. 1b).

A *dimer configuration* $\mathcal{C}$ of G is a subset of edges such that every vertex of G is incident with exactly one edge of $\mathcal{C}$. An example of a dimer configuration on a piece of the honeycomb lattice is presented in Fig. 2. We now define probability measures on the set of all dimer configurations of G.

3.1 Gibbs Measures and Correlations in the Honeycomb Dimer Model

When the graph is finite, one can easily define the uniform probability measure on dimer configurations. But since the honeycomb lattice is an infinite periodic graph, it is not possible to construct directly a uniform measure. However, it was proved in [19, 23] that there exists a two-parameter family of Gibbs measures on the set of dimer configurations, having the following properties:

- they are ergodic under the action of $\mathbb{Z}^2$ by translation,
- they are conditionally uniform: if the configuration is fixed in some annular region of the graph, the configurations outside and inside the annulus are independent, and the measure induced inside is the uniform measure.

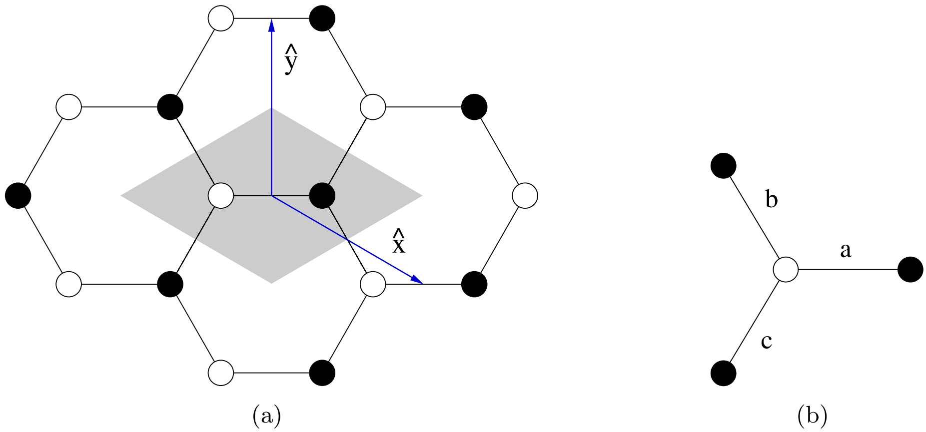
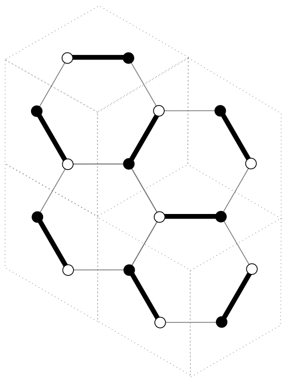


Fig. 1 **a** A piece of the honeycomb lattice. The fundamental domain is represented by the *shaded region*. The two vectors $\hat{x}$ and $\hat{y}$ generate the group of translations preserving the graph G and its bipartite coloring. **b** Weights a, b, c are assigned to edges according to their orientation around a *white vertex*

Fig. 2 An example of a dimer configuration on a piece of the honeycomb lattice G. The corresponding tiling with rhombi is drawn in *dotted lines*



These measures are obtained as weak limits of grand-canonical probability measures on dimer configurations on tori $G/n\mathbb{Z}^2$, with different activities a, b and c for the different kinds of edges. The measure $\mathbb{P}_{a,b,c}$ obtained this way depends only on the ratios b/a and c/a.

This model is exactly solvable. The free energy as well as the correlation functions can be computed explicitly. Following Kasteleyn's ideas [13, 14], Kenyon [17] proved that under $\mathbb{P}_{a,b,c}$, the probability that the edges $e_1 = (\mathbf{w}_{x_1,y_1}, \mathbf{b}_{x'_1,y'_1}), \dots, e_n = (\mathbf{w}_{x_n,y_n}, \mathbf{b}_{x'_n,y'_n})$ are covered by dimers in the random configuration has a determinantal form and is given by the following formula

$$\mathbb{P}_{a,b,c}[e_1, \dots, e_n] = \left(\prod_{i=1}^{n} \mathrm{weight}(e_i)\right) \det_{1\le i,j\le n}[\mathsf{K}^{-1}(\mathbf{b}_{x'_i,y'_i}, \mathbf{w}_{x_j,y_j})], \tag{5}$$

where weight(e) equals a, b or c depending of the type of the edge e, and

$$\mathsf{K}^{-1}(\mathbf{b}_{x',y'},\mathbf{w}_{x,y})=\mathsf{K}^{-1}(\mathbf{b}_{x'-x,y'-y},\mathbf{w}_0)=\iint_{|z|=|w|=1}\frac{z^{-(y'-y)}w^{x'-x}}{a+\frac{1}{w}(b+cz)}\frac{\mathrm{d}z}{2i\pi z}\frac{\mathrm{d}w}{2i\pi w}. \tag{6}$$

K^{-1} is the inverse of the weighted bipartite adjacency matrix K, whose rows are indexed by white vertices and columns by black vertices.

$$\mathsf{K}_{\mathbf{w},\mathbf{b}}=\begin{cases}\text{weight}(e) & \text{if } e=(\mathbf{w},\mathbf{b}),\\ 0 & \text{otherwise.}\end{cases}$$

Kasteleyn discovered that this matrix K encodes all the combinatorics of the model. K is named after him the *Kasteleyn operator*. K^{-1} is called the *inverse Kasteleyn operator*. The entries of K^{-1} are computed using Fourier transform where z (resp. w) is the multiplicator associated with translations in the y direction (resp. $-x$).

The decay of the entries of K^{-1} is directly related to the presence of singularities of the integrand on the unit torus $\mathbb{T}^2=\{(z,w)\in\mathbb{C}^2;\ |z|=|w|=1\}$. Depending on the weights, the behavior of the integral defining K^{-1}, and thus properties of the model can be drastically different:

- If a weight is greater than the sum of the two others, then with probability 1, we only see edges of one type (the type with highest weight). We say that the system is *frozen*. There are three frozen phases, each one corresponding to a type of edges.
- If (a,b,c) satisfy the triangle inequalities, then the density of each type of dimers is positive with probability 1, and correlations between dimers decay polynomially with the distance. The system is said to be in the *liquid phase*.

We will be interested in the phase transition between the liquid phase and the frozen a-phase, when $b=c$ and a approaches $b+c$ from below.

3.2 Non-Colliding Paths in the Honeycomb Dimer Model

The dual graph of the honeycomb lattice is the triangular lattice. This duality allows to define a bijection between dimer configurations of the honeycomb lattice and tilings of the plane by rhombi with a $\frac{\pi}{3}$ angle: one can associate to each edge the rhombus obtained by the union of the two dual faces corresponding to the endpoints of the edge. It will be convenient to use those two points of view.

The three types of rhombi, corresponding to the three types of edges, are also called a, b and c. To describe a tiling, it is sufficient to give the positions of tiles of type, for example, b and c, forming left-to-right non-intersecting paths[1] (see Fig. 3). A measure on dimer configurations can be viewed equivalently as a measure on tilings on the plane as well as a measure on non-colliding left-to-right paths.

When a approaches $b+c$ from below, one sees more and more a-tiles, and the b/c-paths are more and more dilute: the average distance between two paths goes to infinity. The probability of seeing a piece of such a path in a finite window goes to zero as we approach the phase transition. To see something non-trivial in the limit, one needs to rescale correctly the lattice in the same time as we let a go to $b+c$. To find the correct regime, it can be enlightening to look at the behavior of one single path.

[1] These non-intersecting paths can be seen directly in the dimer picture by superimposing on top of the random configuration, the frozen a-configuration.

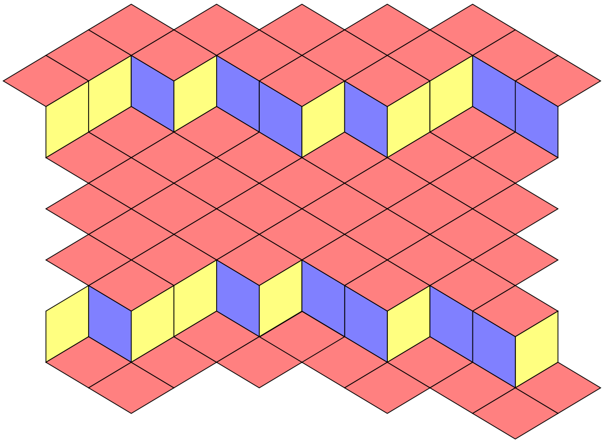


Fig. 3 A piece of a tiling of the plane with rhombi, in bijection with a dimer configuration on G. One can distinguish two left-to-right paths made of rhombi of type b and c corresponding to non-horizontal dimers

3.3 Looking at a Single Path near the Transition

In what follows, we will take $b=c=1$ and $a=2\cos(\frac{\varepsilon}{2})$. We will denote by $\mathbb{P}_\varepsilon$ the Gibbs measure for these weights and $\mathsf{K}_\varepsilon^{-1}$ the corresponding inverse Kasteleyn operator.

For this particular choice of weights, after the change of variables $u=w/\sqrt{z}$ and $e^{i\theta}=z$, $\mathsf{K}_\varepsilon^{-1}$ has the following form:

$$\mathsf{K}_\varepsilon^{-1}(\mathbf{b}_{x,y},\mathbf{w}_0)=\mathsf{K}_\varepsilon^{-1}(x,y)=\int_{-\pi}^{\pi} e^{i\theta(-y+x/2)}\left(\int_{S^1}\frac{u^x}{2\cos(\varepsilon/2)u+2\cos(\theta/2)}\frac{\mathrm{d}u}{2i\pi}\right)\frac{\mathrm{d}\theta}{2\pi}.$$

The pole at $u=-\frac{\cos(\theta/2)}{\cos(\varepsilon/2)}$ is outside of the unit disk if and only if $2\cos(\theta/2)>a$, i.e. if and only if $\theta\in(-\varepsilon,\varepsilon)$. The integral over u in the definition of $\mathsf{K}_\varepsilon^{-1}$ can be computed using residues, and we get the following expressions for the entries of that operator

$$\mathsf{K}_\varepsilon^{-1}(x,y)=\begin{cases}\frac{(-1)^{x+1}}{2\cos(\varepsilon/2)}\int_{[-\varepsilon,\varepsilon]}\left(\frac{\cos(\theta/2)}{\cos(\varepsilon/2)}\right)^x e^{i\theta(-y+x/2)}\frac{\mathrm{d}\theta}{2\pi} & \text{if } x<0,\\ \frac{(-1)^{x}}{2\cos(\varepsilon/2)}\int_{[-\pi,\pi]\setminus[-\varepsilon,\varepsilon]}\left(\frac{\cos(\theta/2)}{\cos(\varepsilon/2)}\right)^x e^{i\theta(-y+x/2)}\frac{\mathrm{d}\theta}{2\pi} & \text{if } x\geq 0.\end{cases} \tag{7}$$

From (5), the probability of seeing an edge of type a is $a\mathsf{K}_\varepsilon^{-1}(0,0)=1-\frac{\varepsilon}{\pi}$. As a goes to 2, the non-intersecting paths become more and more distant, and one can believe that in the limit, these paths will behave almost like simple random walks, at least locally. This is indeed the case, as stated in the following lemma:

Lemma 1 *When $b=c=1$, and in the limit $a\to 2^-$, the left-to-right paths locally behave like simple random walks. More precisely, for any fixed $n>0$, the distribution of n successive steps—up ($\nearrow$) or down ($\searrow$)—in a left-to-right path converges to the uniform measure on $\{\nearrow,\searrow\}^n$.*

Proof Let $(0,y_0),\dots,(n,y_n)$ be a sequence of $\mathbb{Z}^2$ such that $|y_{k+1}-y_k|=1$ for all $k\in\{0,\dots,n-1\}$. For the sake of simplicity, we will write $\mathbf{b}_j$ and $\mathbf{w}_j$ instead of $\mathbf{b}_{j,y_j}$ and $\mathbf{w}_{j,b_j}$. The sequence $(y_{k+1}-y_k)$ can be interpreted as a sequence $\omega=(\omega_n)$ of n steps: $+1$ for up, and -1 for down.

The fact that a left-to-right path of rhombi covers $\mathbf{b}_0$ and then follows the sequence of steps ω can be formulated in terms of dimers as follows: the edge $(\mathbf{w}_0, \mathbf{b}_0)$ is not covered by a dimer, and the edges $(\mathbf{b}_0, \mathbf{w}_1), \dots, (\mathbf{b}_{n-1}, \mathbf{w}_n)$ are covered by dimers.

Thus by formula (5), the probability that, given that a path covers $\mathbf{b}_0$, it follows the succession of steps ω, can be expressed as a determinant:

$$\frac{1}{1 - a\mathsf{K}_\varepsilon^{-1}(\mathbf{b}_0, \mathbf{w}_0)} \det \begin{bmatrix} \mathsf{K}_\varepsilon^{-1}(\mathbf{b}_0, \mathbf{w}_1) & \cdots & \mathsf{K}_\varepsilon^{-1}(\mathbf{b}_{n-1}, \mathbf{w}_1) \\ \vdots & \ddots & \vdots \\ \mathsf{K}_\varepsilon^{-1}(\mathbf{b}_0, \mathbf{w}_n) & \cdots & \mathsf{K}_\varepsilon^{-1}(\mathbf{b}_{n-1}, \mathbf{w}_n) \end{bmatrix}. \tag{8}$$

Using the explicit expressions for $\mathsf{K}_\varepsilon^{-1}$ given in (7), one sees that $\mathsf{K}_\varepsilon^{-1}(\mathbf{b}_j, \mathbf{w}_j) = \frac{1}{2} + o(1)$, and $\mathsf{K}_\varepsilon^{-1}(\mathbf{b}_j, \mathbf{w}_i) = \frac{\varepsilon}{2} + o(\varepsilon)$ for $j < i$. The determinant in (8) is equal up to higher order terms to

$$\det \begin{bmatrix} 0 & \frac{1}{2} & \star & \cdots & \star \\ 0 & 0 & \frac{1}{2} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \star \\ 0 & & & 0 & \frac{1}{2} \\ \frac{\varepsilon}{2} & 0 & \cdots & \cdots & 0 \end{bmatrix} = \frac{\varepsilon}{2^n}$$

and hence, the probability (8) equals

$$\frac{1}{\varepsilon + o(\varepsilon)} \left(\frac{\varepsilon}{2^n} + o(\varepsilon) \right) = \frac{1}{2^n} + o(1). \tag{9}$$

When ε goes to zero, this conditional probability converges to the corresponding probability for a simple random walk. □

Close to the phase transition, we are left with an infinite collection of random paths that individually behave locally like a random walk, but that are conditioned not to intersect with one another.

This individual random walk behavior is an indication that if we want to get a non-trivial model in the limit, we have to choose the same regime as the one giving a non trivial limit for the simple random walk: the diffusive regime.

4 Convergence to the Extended Sine Process

For a fixed ε, we represent our paths of rhombi by continuous functions as follows. We first embed the honeycomb lattice H in the plane correctly. Recall that the vectors defining the coordinates on H are not orthogonal and that the horizontal direction is that of the vector $\hat{y} + 2\hat{x}$. In order to rescale the horizontal direction by a factor ε^2 and the vertical direction by a factor ε, we map every white vertex with coordinates (x, y) to the point with usual cartesian coordinates $(\frac{\varepsilon^2}{8}x, \frac{\varepsilon}{2}(2y - x))$. This diffusive scaling seems natural since the paths of rhombi behave like random walks at small scales.

A path of rhombi is then encoded by the piecewise linear path joining all the white vertices covered by this path. We get a bi-infinite family of paths $(X_n^\varepsilon(t))_{n\in\mathbb{Z}}$, indexed by time $t \in \mathbb{R}$. By convention, the path corresponding to the index 0 is the one passing the closest to $y = 0$ at time 0.

Such a family of paths is in bijection with a random tiling with rhombi. The Gibbs measure $\mathbb{P}_\varepsilon$ on tilings (or equivalently on dimer configurations of G) can therefore be viewed as a probability measure on $\mathcal{C}(\mathbb{R}^\infty)$, the set of continuous functions on $\mathbb{R}$ with values in $\mathbb{R}^\infty = \{\bar{x} = (x_j)_{j\in\mathbb{Z}} \mid \forall j \in \mathbb{Z}\ x_j \in \mathbb{R}\}$. The space $\mathbb{R}^\infty$ is a complete separable metric space with the metric

$$\mathrm{dist}(\bar{x}, \bar{y}) = \sum_{k=-\infty}^{+\infty} \frac{1}{2^{|k|}} \frac{|x_k - y_k|}{1 + |x_k - y_k|}.$$

The topology induced by this distance is the topology of coordinatewise convergence, and the sets

$$B_{k,\varepsilon}(\bar{x}) = \{\bar{y};\ \forall j \in \{-k, \dots, k\}\ |y_j - x_j| < \varepsilon\}$$

form a basis for this topology. One can then define a topology on $\mathcal{C}(\mathbb{R}^\infty)$, using the $\|\cdot\|_\infty$ norm with respect to that distance on $\mathbb{R}^\infty$, which allows to define the notion of weak convergence of probability measures on $\mathcal{C}(\mathbb{R}^\infty)$: a sequence of probability measure $(\mathbb{P}_\varepsilon)$ on $\mathcal{C}(\mathbb{R}^\infty)$ converges weakly to $\mathbb{P}$ if for all functions f on $\mathcal{C}(\mathbb{R}^\infty)$ continuous for the topology defined,

$$\lim_{\varepsilon\to 0} \mathbb{E}_\varepsilon[f(X)] = \mathbb{E}[f(X)],$$

where $\mathbb{E}_\varepsilon$ and $\mathbb{E}$ are the expectations with respect to $\mathbb{P}_\varepsilon$ and $\mathbb{P}$, the laws of the random families of paths X^ϵ and X respectively.

The aim of this section is to prove the main result of this paper.

Theorem 2 *The probability measures* $(\mathbb{P}_\varepsilon)$, *as measures on bi-infinite families of continuous paths on* $\mathbb{R}$, *converge weakly in* $\mathcal{C}(\mathbb{R}^\infty)$ *to the sine process.*

The proof of the convergence of the ensemble of paths to the sine model is based on precise asymptotics of the inverse Kasteleyn operator when ε goes to zero, that are proved in Lemma 2. These asymptotics allow us to prove that the finite dimensional distributions of $\mathbb{P}_\varepsilon$ converge to those for the sine process in Proposition 1. Then we prove that $(\mathbb{P}_\varepsilon)$ is tight in Proposition 2. A key point of the proof of this proposition is a simple comparison between probability for paths of rhombi and for simple random walks given by Lemma 3.

Since the proofs of Propositions 1 and 2 are rather long and technical, they are exposed separately, in Sects. 5 and 6 respectively.

Let us begin with the asymptotics of the kernel in the scaling regime described above, when ε goes to zero.

Lemma 2 *For every* $\varepsilon > 0$, *let* $(x_\varepsilon, y_\varepsilon) \in \mathbb{Z}^2$ *such that*

$$\lim_{\varepsilon\to 0} \frac{1}{8}\varepsilon^2 x_\varepsilon = T \in \mathbb{R}, \qquad \lim_{\varepsilon\to 0} \frac{1}{2}\varepsilon(2y_\varepsilon - x_\varepsilon) = Y \in \mathbb{R}. \tag{10}$$

Recall that $a = 2\cos\frac{\varepsilon}{2}$.

- *If* $T \neq 0$, *then*

$$\lim_{\varepsilon\to 0} \frac{(-1)^{x_\varepsilon+1}}{\varepsilon} a\mathsf{K}_\varepsilon^{-1}(x_\varepsilon, y_\varepsilon) = e^T \mathrm{S}(T, Y). \tag{11}$$

- *If* $x_\varepsilon = 0$ *and* $y_\varepsilon \neq 0$,

$$\lim_{\varepsilon\to 0} \frac{-a}{\varepsilon} \mathsf{K}_\varepsilon^{-1}(0, y_\varepsilon) = \frac{\sin(Y)}{\pi Y} = e^0 \mathrm{S}(0, Y). \tag{12}$$

- *If* $x_\varepsilon = 0$ *and* $y_\varepsilon = 0$,

$$\lim_{\varepsilon\to 0} \frac{1 - a\mathsf{K}_\varepsilon^{-1}(0,0)}{\varepsilon} = \frac{1}{\pi} = e^0 \mathrm{S}(0, 0). \tag{13}$$

Moreover, the quantities $\frac{a\mathsf{K}_\varepsilon^{-1}(x_\varepsilon, y_\varepsilon)}{\varepsilon}$ *(or* $\frac{1-a\mathsf{K}_\varepsilon^{-1}(0,0)}{\varepsilon}$ *if* $(x_\varepsilon, y_\varepsilon) = (0, 0)$*) are uniformly bounded in* ε.

This lemma shows that the discrete kernel for the dimer model on the honeycomb lattice converges, up to a factor e^T (and except for the third case), to the extended sine kernel. It is based on the exact formulae for $\mathsf{K}_\varepsilon^{-1}$ given in (7).

Proof Suppose first that $T < 0$. In this case, for ε small enough, $x_\varepsilon < 0$ and

$$\begin{aligned}
\mathsf{K}_\varepsilon^{-1}(x_\varepsilon, y_\varepsilon) &= \frac{(-1)^{x_\varepsilon+1}}{a} \int_{-\varepsilon}^{\varepsilon} \left(\frac{2\cos(\theta/2)}{a}\right)^{x_\varepsilon} e^{i(-y_\varepsilon + x_\varepsilon/2)\theta} \frac{\mathrm{d}\theta}{2\pi} \\
&= \frac{(-1)^{x_\varepsilon+1}}{a} \int_{-\varepsilon}^{\varepsilon} \left(\frac{\cos(\theta/2)}{\cos(\varepsilon/2)}\right)^{\frac{8T}{\varepsilon^2}(1+o(1))} e^{-i\theta\frac{Y}{\varepsilon}(1+o(1))} \frac{\mathrm{d}\theta}{2\pi} \\
&= \frac{\varepsilon(-1)^{x_\varepsilon+1}}{a} \int_{-1}^{1} \left(\frac{\cos(\frac{\varepsilon\phi}{2})}{\cos(\frac{\varepsilon}{2})}\right)^{\frac{8T}{\varepsilon^2}(1+o(1))} e^{-iY\phi(1+o(1))} \frac{\mathrm{d}\phi}{\pi} \\
&= \frac{\varepsilon(-1)^{x_\varepsilon+1}}{a} \int_{-1}^{1} \left(1 - \frac{\varepsilon^2(\phi^2-1)}{8} + o(\varepsilon^2)\right)^{\frac{8T}{\varepsilon^2}(1+o(1))} e^{-iY\phi + o(1)} \frac{\mathrm{d}\phi}{2\pi}.
\end{aligned}$$

When ε goes to zero, the integrand converges to $e^T e^{-\phi^2 T} e^{-iY\phi}$. A simple application of Lebesgue dominated convergence theorem shows then that

$$\lim_{\varepsilon\to 0} \frac{(-1)^{x_\varepsilon+1}}{\varepsilon} a\mathsf{K}_\varepsilon^{-1}(x_\varepsilon, y_\varepsilon) = e^T \int_{-1}^{1} e^{-\phi^2 T} e^{-iY\phi} \frac{\mathrm{d}\phi}{2\pi} = e^T \int_0^1 e^{-\phi^2 T} \cos(Y\phi) \frac{\mathrm{d}\phi}{\pi} = \mathrm{S}(T, Y). \tag{14}$$

When $T > 0$, then $x_\varepsilon > 0$ for ε small enough, and we have

$$\mathsf{K}_\varepsilon^{-1}(x_\varepsilon, y_\varepsilon) = \frac{(-1)^{x_\varepsilon}}{a} \int_{[-\pi,\pi]\setminus[-\varepsilon,\varepsilon]} \left(\frac{2\cos(\theta/2)}{a}\right)^{x_\varepsilon} e^{i(-y_\varepsilon + x_\varepsilon/2)\theta} \frac{\mathrm{d}\theta}{2\pi} \tag{15}$$

$$= \frac{(-1)^{x_\varepsilon}}{a} \int_{\varepsilon}^{\pi} \left(\frac{2\cos(\theta/2)}{a}\right)^{x_\varepsilon} \cos((-y + x/2)\theta) \frac{\mathrm{d}\theta}{\pi}. \tag{16}$$

If we change the variable in the integral, defining $\phi = \varepsilon\theta$, we get

$$\mathsf{K}_\varepsilon^{-1}(x_\varepsilon, y_\varepsilon) = \frac{\varepsilon(-1)^{x_\varepsilon}}{a} \int_1^{\pi/\varepsilon} \left(\frac{\cos(\frac{\varepsilon\phi}{2})}{\cos(\frac{\varepsilon}{2})}\right)^{\frac{8T}{\varepsilon^2}(1+o(1))} \cos(Y\phi + o(1)) \frac{\mathrm{d}\phi}{\pi}$$

$$= \frac{\varepsilon(-1)^{x_\varepsilon}}{a}\int_1^A \left(\frac{\cos(\frac{\varepsilon\phi}{2})}{\cos(\frac{\varepsilon}{2})}\right)^{\frac{8T}{\varepsilon^2}(1+o(1))}\cos(Y\phi+o(1))\frac{\mathrm{d}\phi}{\pi}$$
$$+\frac{\varepsilon(-1)^{x_\varepsilon}}{a}\int_A^{\pi/\varepsilon} \left(\frac{\cos(\frac{\varepsilon\phi}{2})}{\cos(\frac{\varepsilon}{2})}\right)^{\frac{8T}{\varepsilon^2}(1+o(1))}\cos(Y\phi+o(1))\frac{\mathrm{d}\phi}{\pi} \quad (17)$$

for a given large number A. The second integral can be bounded by

$$\left|\frac{\varepsilon(-1)^{x_\varepsilon}}{a}\int_A^{\pi/\varepsilon} \left(\frac{\cos(\frac{\varepsilon\phi}{2})}{\cos(\frac{\varepsilon}{2})}\right)^{\frac{8T}{\varepsilon^2}(1+o(1))}\cos(Y\phi+o(1))\frac{\mathrm{d}\phi}{\pi}\right| \le Ce^{-T(A-1)+o(1)} \quad (18)$$

uniformly in ε. Lebesgue's dominated convergence theorem shows that the first integral converges to

$$e^T\int_1^A e^{-\phi^2 T}\cos(Y\phi)\frac{\mathrm{d}\phi}{\pi}. \quad (19)$$

As a consequence, letting A go to ∞, we get

$$\lim_{\varepsilon\to 0}\frac{(-1)^{x_\varepsilon+1}}{\varepsilon}a\mathsf{K}_\varepsilon^{-1}(y,x) = -e^T\int_1^{+\infty} e^{-\phi^2 T}\cos(Y\phi)\frac{\mathrm{d}\phi}{\pi}. \quad (20)$$

When $x_\varepsilon = 0$, $\mathsf{K}^{-1}(x_\varepsilon, y_\varepsilon)$ given by

$$\mathsf{K}_\varepsilon^{-1}(0, y_\varepsilon) = \frac{1}{a}\int_\varepsilon^\pi \cos(y_\varepsilon\theta)\frac{\mathrm{d}\theta}{\pi}. \quad (21)$$

If $y_\varepsilon \neq 0$, then

$$\frac{-a}{\varepsilon}\mathsf{K}_\varepsilon^{-1}(0, y_\varepsilon) = -\frac{a}{\varepsilon}\frac{\sin(\pi y_\varepsilon)-\sin(\varepsilon y_\varepsilon)}{a\pi} = \frac{\sin(Y)}{\pi Y}+o(1). \quad (22)$$

However, if $y_\varepsilon = 0$, $\mathsf{K}_\varepsilon^{-1}(0,0) = \frac{\pi-\varepsilon}{a\pi}$ and thus

$$\lim_{\varepsilon\to 0}\frac{1-a\mathsf{K}_\varepsilon^{-1}(0,0)}{\varepsilon} = \frac{1}{\pi}. \quad (23)$$

The uniform bounds are obtained directly form the expression of $\mathsf{K}_\varepsilon^{-1}$ given by (7). When $T<0$,

$$\left|\frac{a}{\varepsilon}\mathsf{K}_\varepsilon^{-1}(x_\varepsilon, y_\varepsilon)\right| = \left|\frac{1}{\varepsilon}\int_{-\varepsilon}^{\varepsilon}\left(\frac{2\cos(\theta/2)}{a}\right)^{x_\varepsilon} e^{i\theta(-y_\varepsilon+x_\varepsilon/2)}\frac{\mathrm{d}\theta}{2\pi}\right| \le \frac{1}{\pi}\left(1-\left(\frac{\varepsilon}{2}\right)^2\right)^{-x_\varepsilon/2} \le C. \quad (24)$$

When $T>0$,

$$\left|\frac{a}{\varepsilon}\mathsf{K}_\varepsilon^{-1}(x_\varepsilon, y_\varepsilon)\right| = \left|\frac{1}{\varepsilon}\int_{[-\pi,\pi]\setminus[-\varepsilon,\varepsilon]}\left(\frac{2\cos(\theta/2)}{a}\right)^{x_\varepsilon} e^{i\theta(-y_\varepsilon+x_\varepsilon/2)}\frac{\mathrm{d}\theta}{2\pi}\right|$$
$$\le \frac{1}{\pi}\left(1-\left(\frac{\varepsilon}{2}\right)^2\right)^{-x_\varepsilon/2}\int_1^{\frac{\pi}{\varepsilon}} e^{-\frac{\varepsilon^2 x_\varepsilon}{8}t^2}\mathrm{d}t \le C. \quad (25)$$

When $x_\varepsilon = 0$, the bound is obtained by a direct evaluation of (7). □

As in the case of the extended sine process, for given times $\tau_1, \dots, \tau_k$ and unions of intervals $I_1, \dots, I_k$, we can compute for every ε the probability that for every j, the number $N^\varepsilon_{I_j}(\tau_j)$ of paths in I_j at time τ_j equals n_j. The following proposition states that these probabilities converge to the corresponding probabilities for the sine process.

Proposition 1 *The finite dimensional distributions of X^ε converge to that of the extended sine process.*

For all $\tau_1, \dots, \tau_m \in \mathbb{R}$, and $I_1, \dots, I_m \subset \mathbb{R}$ finite unions of intervals,

$$\lim_{\varepsilon\to 0} \mathbb{P}_\varepsilon[N^\varepsilon_{I_1}(\tau_1) = n_1, \dots, N^\varepsilon_{I_k}(\tau_k) = n_k] = \mathbb{P}[N_{I_1}(\tau_1) = n_1, \dots, N_{I_k}(\tau_k) = n_k]. \tag{26}$$

We prove also tightness of the family of probability measures $(\mathbb{P}_\varepsilon)$ in $\mathcal{C}(\mathbb{R}^\infty)$. Since $\mathbb{R}^\infty$ is a nice complete separable metric space, tightness is characterized by the following two properties [2] we have to check:

Proposition 2 (i) *The sequence of distributions of $(X^\varepsilon(0))$ is tight: for each $\eta > 0$, there exists a sequence of closed intervals $([-b_k, b_k])_{k\in\mathbb{Z}}$ such that*

$$\forall \varepsilon, \quad \mathbb{P}_\varepsilon[\exists k \in \mathbb{Z},\ X^\varepsilon_k(0) \notin [-b_k, b_k]] < \eta \tag{27}$$

(ii) *For each $L > 0$, for each positive δ and η, there exists an $\alpha \in (0, L)$ such that*

$$\forall \varepsilon, \quad \mathbb{P}_\varepsilon\Big[\sup_{\substack{-L/2<s,t<L/2\\ |s-t|<\alpha}} \mathrm{dist}(X^\varepsilon(s), X^\varepsilon(t)) > \delta\Big] < \eta. \tag{28}$$

Tightness and convergence of finite dimensional distributions of the family $(\mathbb{P}_\varepsilon)$ imply the weak convergence of this family to a probability measure on $\mathcal{C}(\mathbb{R}^\infty)$ whose finite distributions coincide with that of the sine process. The family of discrete random paths therefore converges weakly to the family of random continuous paths described by the sine process. This ends the proof of Theorem 2, modulo the proofs of the two last propositions.

5 Proof of Proposition 1

This section is devoted to the proof of the convergence of finite dimensional distributions of the family $(\mathbb{P}_\varepsilon)$. This is done by studying the convergence of generating functions Q^ε for the distribution of the number of visits of a finite number of intervals at given times when ε goes to zero.

We denote by τ^ε the quantity $\lfloor \frac{\tau}{8\varepsilon^2} \rfloor$ and $(I)^\varepsilon_\tau$ the set of integers

$$(I)^\varepsilon_\tau = \left\{ y \in \mathbb{Z} \mid \varepsilon\left(y - \frac{1}{2}\tau^\varepsilon\right) \in I \right\}.$$

The white vertices with coordinates $(x, y) \in \{\tau^\varepsilon\} \times (I)^\varepsilon_\tau$ are the closest to $\{\tau\} \times I$.

The function Q^ε of the variable $z = (z_1, \dots, z_k) \in \mathbb{C}^k$ defined by

$$Q^\varepsilon(z) = \mathbb{E}_\varepsilon\left[\prod_{j=1}^k (1 - z_j)^{N^\varepsilon_{I_j}(\tau_j)}\right] = \sum_{n\in\mathbb{N}^k} \mathbb{E}_\varepsilon\left[\prod_{j=1}^k \frac{N^\varepsilon_{I_j}(\tau_j)!}{(N^\varepsilon_{I_j}(\tau_j) - n_j)!}\right] \frac{(-z)^n}{n!}, \tag{29}$$

where $z^n = \prod_{j=1}^k z_j^{n_j}$, is a generating function for the probabilities we are interested in. Indeed, from (4), we have:

$$\mathbb{P}_\varepsilon[N_{I_1}^\varepsilon(\tau_1)=n_1,\dots,N_{I_k}^\varepsilon(\tau_k)=n_k] = \sum_{p\in\mathbb{N}^k} \frac{(-1)^p}{n!p!}\mathbb{E}_\varepsilon\left[\prod_{j=1}^k \frac{N_{I_j}^\varepsilon(\tau_j)!}{(N_{I_j}^\varepsilon(\tau_j)-(n_j+p_j))!}\right] \tag{30}$$

$$= \frac{(-1)^n}{n!}\frac{\partial^n}{\partial z^n} Q^\varepsilon(z)|_{z=(1,\dots,1)}. \tag{31}$$

We will show that Q^ε converges to the corresponding generating function for the extended sine process. For this, we first compute the quantities $\mathbb{E}_\varepsilon[\prod_{j=1}^k \frac{N_{I_j}^\varepsilon(\tau_j)!}{(N_{I_j}^\varepsilon(\tau_j)-n_j)!}]$ and their limit when ε goes to 0. These quantities are given by the sum over all distinct n_j-tuples of white vertices in $\{\tau_j^\varepsilon\}\times(I_j)_{\tau_j}^\varepsilon$ for $j\in\{1,\dots,k\}$ of the probability that these white vertices are covered by a path of rhombi of type b and c, or equivalently, that the horizontal edges with weight a incident with these vertices are not present in the random dimer configuration. We have

$$\mathbb{E}_\varepsilon\left[\prod_{j=1}^k \frac{N_{I_j}^\varepsilon(\tau_j)!}{(N_{I_j}^\varepsilon(\tau_j)-n_j)!}\right] = \sum_{\substack{j=1,\dots,k\\ y_{(j,1)},\dots,y_{(j,n_j)}\in(I_j)_{\tau_j}^\varepsilon\\ \text{distinct}}} \mathbb{P}_\varepsilon[\mathbf{a}_{\tau_1^\varepsilon,y_{(1,1)}},\dots,\mathbf{a}_{\tau_1^\varepsilon,y_{(1,n_1)}},\dots,\mathbf{a}_{\tau_k^\varepsilon,y_{(k,n_k)}}\notin\mathcal{C}], \tag{32}$$

where $\mathbf{a}_{x,y}$ is the horizontal edge with weight a incident with the white vertex with coordinates (x,y) in H, and $\mathcal{C}$ is the random dimer configuration.

Using multilinearity of the determinant and an inclusion-exclusion argument, one can write each of these probabilities

$$\mathbb{P}_\varepsilon[\mathbf{a}_{\tau_1^\varepsilon,y_{(1,1)}},\dots,\mathbf{a}_{\tau_1^\varepsilon,y_{(1,n_1)}},\dots,\mathbf{a}_{\tau_k^\varepsilon,y_{(k,n_k)}}\notin\mathcal{C}] \tag{33}$$

as a determinant

$$\det(\delta_{i,j} - a\mathsf{K}_\varepsilon^{-1}(\tau_{\langle i\rangle}^\varepsilon-\tau_{\langle j\rangle}^\varepsilon, y_i-y_j))$$

$$= \det_{|n|}\begin{bmatrix} 1-a\mathsf{K}_\varepsilon^{-1}(\tau_1^\varepsilon-\tau_1^\varepsilon, y_{(1,1)}-y_{(1,1)}) & \cdots & -a\mathsf{K}_\varepsilon^{-1}(\tau_k^\varepsilon-\tau_1^\varepsilon, y_{(k,n_k)}-y_{(1,1)}) \\ \vdots & & \vdots \\ \vdots & \ddots & \vdots \\ \vdots & & \vdots \\ -a\mathsf{K}_\varepsilon^{-1}(\tau_1^\varepsilon-\tau_k^\varepsilon, y_{(1,n_1)}-y_{(k,n_k)}) & \cdots & 1-a\mathsf{K}_\varepsilon^{-1}(\tau_k^\varepsilon-\tau_k^\varepsilon, y_{(k,n_k)}-y_{(k,n_k)}) \end{bmatrix}, \tag{34}$$

where i and j belong to the following list of indices

$$((1,1),\dots,(1,n_1),\dots,(k,1),\dots,(k,n_k)), \tag{35}$$

and the angular brackets $\langle i\rangle$ of a couple $i=(i_1,i_2)$ represent its first factor i_1.

The determinant is unchanged if we multiply each column i by $e^{-\tau_{\langle i\rangle}}(-1)^{\tau_{\langle i\rangle}^\varepsilon}$ and each line by $e^{\tau_{\langle j\rangle}}(-1)^{-\tau_{\langle j\rangle}^\varepsilon}$. It follows from Lemma 2 that this determinant is asymptotic to

$$\varepsilon^{|n|}\det(\mathsf{S}(\tau_{\langle i\rangle}-\tau_{\langle j\rangle}, Y_i^\varepsilon - Y_j^\varepsilon)), \tag{36}$$

where $Y_j^\varepsilon = \varepsilon(y_j - \tau_{\langle j \rangle}^\varepsilon) \in I_j$. The sum (32) becomes a Riemann sum converging as ε goes to zero to the following integral

$$\int \cdots \int_{(\tau_1, I_1)^{n_1} \times \cdots \times (\tau_k, I_k)^{n_k}} \det(\mathrm{S}(\xi_i - \xi_j)) \mathrm{d}^n \xi. \tag{37}$$

Therefore, the coefficients of $Q^\varepsilon(z)$ converge to those of $Q(z)$, the generating function for the sine process. By Lemma 2, the entries of the matrix in (34) are uniformly bounded, say by M. By applying Hadamard's lemma , we bound the sum (32) by

$$\left| \mathbb{E}_\varepsilon \left[\prod_{j=1}^k \frac{N_{I_j}^\varepsilon(\tau_j)!}{(N_{I_j}^\varepsilon(\tau_j) - n_j)!} \right] \right| \le \left(\prod_{j=1}^k |I_j|^{n_j} \right) (|n| M^2)^{|n|/2} \tag{38}$$

uniformly in ε. Thanks to the additional factor $\frac{1}{n!}$, the general term of the series in (30) is uniformly bounded for ε sufficiently small and z in a compact set, by the general term of an absolutely convergent series. Thus, by Lebesgue's dominated convergence theorem, all the derivatives of the series $Q^\varepsilon(z)$ converge uniformly on compact sets of $\mathbb{C}^k$ to the corresponding derivatives of the generating function $Q(z)$ for the sine process. In particular,

$$\lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon[N_{I_1}^\varepsilon(\tau_1) = n_1, \ldots, N_{I_k}^\varepsilon(\tau_k) = n_k] = \mathbb{P}[N_{I_1}(\tau_1) = n_1, \ldots, N_{I_k}(\tau_k) = n_k],$$

where $\mathbb{P}$ is the probability measure for the extended sine process, with the kernel S given by (2).

6 Proof of Proposition 2

This section is devoted to the proof of the tightness of the family of probability measures $(\mathbb{P}_\varepsilon)_{\varepsilon > 0}$.

First, we check the point (i) of Proposition 2. Let $\eta > 0$. We look for a sequence of segments such that

$$\forall \varepsilon, \quad \mathbb{P}_\varepsilon[\exists k \in \mathbb{Z},\ X_k^\varepsilon(0) \notin [-b_k, b_k]] < \eta$$

is satisfied. If we take a sequence verifying that

$$\forall k \in \mathbb{N}, \quad b_k = b_{-k} \quad \text{and} \quad b_{k+1} > b_k \tag{39}$$

then the condition (27) is equivalent to

$$\forall \varepsilon > 0, \quad \mathbb{P}_\varepsilon[\exists k \in \mathbb{N}, N_{b_k} < 2k + 1] < \eta, \tag{40}$$

where $N_{b_k} = N_{[-b_k, b_k]}^\varepsilon(0)$ is the number of paths crossing the interval $[-b_k, b_k]$ at time 0. We will use Chebyshev inequality to get a bound on the probability (40). For this, we need first to estimate and bound the average and the variance of the random variable $N_L = N_{[-L,L]}^\varepsilon(0)$ for any $L > 0$.

The average of N_L under $\mathbb{P}_\varepsilon$ is given by

$$\mathbb{E}_\varepsilon[N_L] = \sum_{y \in [-\frac{L}{\varepsilon}, \frac{L}{\varepsilon}] \cap \mathbb{Z}} \mathbb{P}_\varepsilon \text{ [there is a path crossing at } (0, \varepsilon y)] \tag{41}$$

$$= \sum_{y\in[-\frac{L}{\varepsilon},\frac{L}{\varepsilon}]\cap\mathbb{Z}} (1-\mathbb{P}_\varepsilon[\mathbf{a}_y \in \mathcal{C}]), \tag{42}$$

where $\mathbf{a}_y = \mathbf{a}_{0,y}$ is the horizontal edge whose white end has coordinates $(0, y)$ in H. There are $2\lfloor \frac{L}{\varepsilon} \rfloor + 1$ integer points in $[-\frac{L}{\varepsilon}, \frac{L}{\varepsilon}]$, and $\mathbb{P}_\varepsilon[\mathbf{a}_y \in \mathcal{C}] = 1 - \frac{\varepsilon}{\pi}$. Therefore,

$$\mathbb{E}_\varepsilon[N_L] = \left(2\left\lfloor \frac{L}{\varepsilon}\right\rfloor + 1\right)\frac{\varepsilon}{\pi} = \frac{2L}{\pi}(1+o(1)). \tag{43}$$

Moreover, the variance of N_L is given by

$$\mathrm{Var}_\varepsilon(N_L) = \mathbb{E}_\varepsilon[N_L^2] - \mathbb{E}_\varepsilon[N_L]^2 \tag{44}$$

$$= \mathbb{E}_\varepsilon[N_L(N_L-1)] - \mathbb{E}_\varepsilon[N_L]^2 + \mathbb{E}_\varepsilon[N_L] < \mathbb{E}_\varepsilon[N_L]. \tag{45}$$

Indeed, $\mathbb{E}_\varepsilon[N_L(N_L-1)] - \mathbb{E}_\varepsilon[N_L]^2$, given by the following expression

$$\sum_{\substack{y,y'\in[-\frac{L}{\varepsilon},\frac{L}{\varepsilon}]\cap\mathbb{Z}\\ y\neq y'}} \mathbb{P}_\varepsilon[\mathbf{a}_y, \mathbf{a}_{y'} \notin \mathcal{C}] - \left(\sum_{y\in[-\frac{L}{\varepsilon},\frac{L}{\varepsilon}]\cap\mathbb{Z}} \mathbb{P}_\varepsilon[\mathbf{a}_y \notin \mathcal{C}]\right)^2 \tag{46}$$

$$= \sum_{\substack{y,y'\in[-\frac{L}{\varepsilon},\frac{L}{\varepsilon}]\cap\mathbb{Z}\\ y\neq y'}} \det\begin{bmatrix} 1-a\mathsf{K}_\varepsilon^{-1}(0,0) & -a\mathsf{K}_\varepsilon^{-1}(0,y-y') \\ -a\mathsf{K}_\varepsilon^{-1}(0,y'-y) & 1-a\mathsf{K}_\varepsilon^{-1}(0,0)\end{bmatrix}$$

$$-\left(\sum_{y\in[-\frac{L}{\varepsilon},\frac{L}{\varepsilon}]\cap\mathbb{Z}} 1-a\mathsf{K}_\varepsilon^{-1}(0,0)\right)^2 \tag{47}$$

$$= -\sum_{\substack{y,y'\in[-\frac{L}{\varepsilon},\frac{L}{\varepsilon}]\cap\mathbb{Z}\\ y\neq y'}} a^2\mathsf{K}^{-1}(0,y-y')^2 - (1-a\mathsf{K}_\varepsilon^{-1}(0,0))^2\left(2\left\lfloor\frac{L}{\varepsilon}\right\rfloor+1\right) < 0 \tag{48}$$

is a sum of negative terms. By Chebyshev inequality, we get

$$\mathbb{P}_\varepsilon\left[N_L < \frac{L}{\pi}\right] \le \mathbb{P}_\varepsilon\left[|N_L - \mathbb{E}_\varepsilon[N_L]| < \frac{L}{\pi}\right] \le \frac{\mathrm{Var}_\varepsilon(N_L)}{(L/\pi)^2} \le \frac{\mathbb{E}_\varepsilon[N_L]}{(L/\pi)^2} \le \frac{2\pi}{L}(1+o(1)). \tag{49}$$

Thus, for $\eta \le \frac{3\pi^2}{2}$, taking $b_k = \frac{2\pi^3(k+1)^2}{3\eta}$, using the fact that for $k \ge 0$,

$$2k+1 \le (k+1)^2 \le \frac{b_k}{\pi},$$

we finally get that

$$\mathbb{P}_\varepsilon[\exists k \in \mathbb{N}, N_{b_k} < 2k+1] \le \sum_{k=0}^{+\infty} \mathbb{P}_\varepsilon[N_{b_k} < 2k+1] \tag{50}$$

$$\le \sum_{k=0}^{+\infty} \mathbb{P}_\varepsilon\left[N_{b_k} < \frac{b_k}{\pi}\right] \tag{51}$$

$$\leq \sum_{k=0}^{+\infty} \frac{2\pi}{b_k}(1+o(1)) = \frac{\eta}{2}(1+o(1)) \leq \eta, \tag{52}$$

for ε small enough, uniformly in η. Therefore, the family of distributions of $(X^\varepsilon(0))$ under $\mathbb{P}_\varepsilon$ is tight.

The proof of point (ii) goes as follows. First, as the topology on $\mathbb{R}^\infty$ is that of coordinate-wise convergence, the condition (28) is equivalent to

$$\forall L>0,\ \forall \eta>0,\ \forall k\in\mathbb{N},\ \exists \alpha\in(0,L),\ \forall\varepsilon>0,\ \forall j\in\{-k,\dots,k\},$$
$$\mathbb{P}_\varepsilon\Bigg[\sup_{\substack{s,t\in(-\frac{L}{2},\frac{L}{2})\\ |s-t|<\alpha}} |X_j^\varepsilon(s)-X_j^\varepsilon(t)|>\delta\Bigg]<\eta. \tag{53}$$

It is therefore sufficient to check this condition for each path separately. It is well-known that this property is true for simple random walks: for any $\delta>0$, there exists an $\alpha>0$ such that the number of properly rescaled walks (Y^ε) with $N=\lfloor L/\varepsilon\rfloor$ steps such that

$$\sup_{\substack{s,t\in(-\frac{L}{2},\frac{L}{2})\\ |s-t|<\alpha}} |Y^\varepsilon(s)-Y^\varepsilon(t)|>\delta \tag{54}$$

is less than $\eta 2^N$ uniformly in ε. Therefore, by the following, rather technical lemma 3, since the probability of a sequence of steps of X_j is bounded by a constant times the same probability for a simple random walk, we get by decomposing the event realization by realisation of the path X_j from 0 to L,

$$\mathbb{P}_\varepsilon\Bigg[\sup_{\substack{s,t\in(-\frac{L}{2},\frac{L}{2})\\ |s-t|<\alpha}} |X_j^\varepsilon(s)-X_j^\varepsilon(t)|>\delta\Bigg]<\mathrm{C}_L\cdot\eta, \tag{55}$$

where the constant C_L depends only in L.

Lemma 3 *Let $L>0$. There exists a $c_L>0$ such that for every ε small enough and every $(\omega_n)_{0\leq n\leq\lfloor L/\varepsilon\rfloor}$ sequence of upward and downward steps, the probability that the first $\lfloor L/\varepsilon\rfloor$ steps of a path of rhombi X_j^ε, conditioned on its starting position (x_0,y_0), coincide with (ω_n), is bounded by*

$$\mathbb{P}_\varepsilon\left[\textit{first}\left\lfloor\frac{L}{\varepsilon}\right\rfloor \textit{steps of } X_j^\varepsilon \textit{ coincide with } (\omega_n)\mid X_j^\varepsilon \textit{ starts at } (x_0,y_0)\right]\leq\frac{c_L}{2^{\lfloor L/\varepsilon\rfloor}}. \tag{56}$$

If the number of steps was finite, one could have proceeded as in Lemma 1. However, here we want a control for a number of steps going to infinity as ε goes to zero.

Proof Define

$$g(x,y)=\frac{1}{2}\int_{-\varepsilon}^{\varepsilon}\left(\cos\frac{\theta}{2}\right)^x e^{i\theta(-y+x/2)}\frac{\mathrm{d}\theta}{2\pi}. \tag{57}$$

Since

$$\int_{-\pi}^{\pi}(2\cos(\theta/2))^x e^{i\theta(-y+x/2)}\frac{\mathrm{d}\theta}{2\pi}=\sum_{k=0}^{x}\binom{x}{k}\int_{-\pi}^{\pi}e^{i\frac{\theta}{2}(2k-x)}e^{i\frac{\theta}{2}(-2y+x)}\frac{\mathrm{d}\theta}{2\pi}=\binom{x}{y}, \tag{58}$$

$\mathsf{K}^{-1}(x,y)$ can be rewritten as

$$\mathsf{K}_\varepsilon^{-1}(x,y)=\begin{cases}\left(-\frac{2}{a}\right)^{x+1}g(x,y) & \text{if } x<0,\\ \left(-\frac{2}{a}\right)^{x+1}\left(g(x,y)-\frac{1}{2^{x+1}}\binom{x}{y}\right) & \text{if } x\geq 0.\end{cases} \tag{59}$$

Let $N=\lfloor\frac{L}{\varepsilon}\rfloor$. Suppose that the sequence (ω_n) corresponds to a path covering the vertices

$$\mathbf{b}_0=\mathbf{b}_{y_0}^{(0)},\quad \mathbf{w}_1=\mathbf{w}_{y_1}^{(1)},\quad \mathbf{b}_1=\mathbf{b}_{y_1}^{(1)},\quad \dots,\quad \mathbf{b}_{N-1}=\mathbf{b}_{y_{N-1}}^{(N-1)},\quad \mathbf{w}_N=\mathbf{w}_{y_N}^{(N)}, \tag{60}$$

with $y_{i+1}-y_i=\frac{1+\omega_i}{2}$.

The probability that a path goes through these vertices is given by

$$\begin{aligned}&\mathbb{P}_\varepsilon[\text{there exists a path through } \mathbf{b}_0,\mathbf{w}_1,\dots,\mathbf{b}_{N-1},\mathbf{w}_N]\\ &\quad=\det(\mathsf{K}_\varepsilon^{-1}(\mathbf{b}_{j-1},\mathbf{w}_i))=\det(\mathsf{K}_\varepsilon^{-1}((j-1)-i,y_{j-1}-y_i))\end{aligned} \tag{61}$$

where the indices i and j refer to white and black vertices respectively and range from 1 to N. Multiplying simultaneously each column i of the matrix (61) by $(\frac{-a}{2})^{-i}$ and each row j by $(\frac{-a}{2})^j$ does not affect the determinant. This manipulation cancels the coefficients $(\frac{-2}{a})^{(j-1)-i+1}$ in the expression of $\mathsf{K}^{-1}(\mathbf{b}_{j-1},\mathbf{w}_i)$, leading to a simpler expression for the determinant (61).

$$\begin{aligned}&\mathbb{P}_\varepsilon[\text{there is path through } \mathbf{b}_0,\mathbf{w}_1,\dots,\mathbf{b}_{N-1},\mathbf{w}_N]\\ &\quad=\det\left[g((j-1)-i,y_{j-1}-y_i)-\frac{\delta_{j>i}}{2^{j-i}}\binom{j-1-i}{y_{j-1}-y_i}\right],\end{aligned} \tag{62}$$

where $\delta_{j>i}=1$ is $j>i$ and 0 otherwise. The matrix, the determinant of which we want to compute, is the sum of a matrix with entries $g(x,y)=O(\varepsilon)$, and a strictly upper triangular matrix.

We have now to find an upper bound for the value of this determinant to get the estimate stated in the lemma. We will use the Hadamard inequality, but a direct application of it would give too approximative a bound. We will have to make some more manipulations on the rows and columns of this matrix.

First, note that for every $x\in\{-N,\dots,N\}$, and $y'=y$ or $y+1$,

$$g(x+1,y')-g(x,y)=\frac{1}{2}\int_{-\varepsilon}^{\varepsilon}\left(\cos\frac{\theta}{2}\right)^x e^{i\frac{\theta}{2}(-2y+x)}\left(e^{\pm i\frac{\theta}{2}}\cos\frac{\theta}{2}-1\right)\frac{d\theta}{2\pi}=O(\varepsilon^3), \tag{63}$$

uniformly in x. Therefore, replacing for $j\in\{2,\dots,N\}$, column C_j by C_j-C_{j-1}, we get on the diagonal and under it entries that are $O(\varepsilon^3)$ (except in the first column). It is important to observe that after these operations, the modulus of the entries (i,j) strictly above the diagonal does not increase, at least at the leading order in ε, and are bounded from above by a constant times $(j-i)^{-1/2}$, since

$$\left|\frac{1}{2^{x+1}}\binom{x}{y}-\frac{1}{2^x}\binom{x-1}{y'}\right|\leq\frac{1}{2^{x+1}}\binom{x}{y}\leq\frac{1}{2^{x+1}}\binom{x}{\lfloor x/2\rfloor}\leq\frac{1}{\sqrt{2\pi x}}. \tag{64}$$

Now, we use the second column to put 0 instead of the first entries of columns C_j, $j>2$ by making the substitution $C_j\leftarrow C_j+\alpha_{j,2}C_2$ for a suitable value of $\alpha_{j,2}$ in the determinant.

From the bound on coefficients, we see that $|\alpha_{j,2}| = O((j-2)^{-1/2})$ and that therefore, the modulus of the entries of column C_j, $j \le 3$ does not increase more than by

$$O((j-2)^{-1/2}\varepsilon^3) = o(1). \tag{65}$$

We then continue this procedure, and for $3 \le k \le N-1$, we use column C_k to eliminate the entries of the $(k-1)$th row on columns C_j, $j > k$.

After this succession of operations, the entries (i, j) of the matrix with $j > i+1$ are 0. The modulus of the other entries of column C_j, $j \le 2$ increased at most by

$$O\left(\frac{\varepsilon^3}{\sqrt{j-2}}\right) + O\left(\frac{\varepsilon^3}{\sqrt{j-3}}\right) + \cdots + O\left(\frac{\varepsilon^3}{\sqrt{j-(j-1)}}\right) = O(\varepsilon^3\sqrt{j}). \tag{66}$$

The probability we are interested is given by the determinant

$$\det\begin{bmatrix} O(\varepsilon) & \frac{1}{2}+O(\sqrt{2}\varepsilon^3) & 0 & \cdots & 0 \\ O(\varepsilon) & O(\sqrt{2}\varepsilon^3) & \frac{1}{2}+O(\sqrt{3}\varepsilon^3) & & \vdots \\ \vdots & & & \ddots & 0 \\ \vdots & \vdots & \ddots & & \frac{1}{2}+O(\sqrt{N}\varepsilon^3) \\ O(\varepsilon) & O(\sqrt{2}\varepsilon^3) & \cdots & & O(\sqrt{N}\varepsilon^3) \end{bmatrix}. \tag{67}$$

Hadamard inequality states that a determinant is bounded by the product of the ℓ^2-norm of its rows. In this particular case, it gives

$$\begin{aligned} &\mathbb{P}_\varepsilon[\text{there is path through } \mathbf{b}_0, \mathbf{w}_1, \ldots, \mathbf{b}_{N-1}, \mathbf{w}_N] \\ &\quad \le \prod_{j=1}^{N-1}(O(\varepsilon)^2 + \sum_{k=2}^{j} O(\sqrt{k}\varepsilon^3)^2 + (\frac{1}{2}+O(\sqrt{j+1}\varepsilon^3))^2)^{\frac{1}{2}} \times (O(\varepsilon)^2 + \sum_{k=1}^{N} O(\sqrt{k}\varepsilon^3)^2)^{\frac{1}{2}}. \end{aligned} \tag{68}$$

Recall that $N = \lfloor \frac{L}{\varepsilon} \rfloor = O(\frac{1}{\varepsilon})$. Thus, from the fact that, for $j \in \{1, \ldots, N-1\}$,

$$\left(O(\varepsilon)^2 + \sum_{k=2}^{j} O(\sqrt{k}\varepsilon^3)^2 + \left(\frac{1}{2}+O(\sqrt{j+1}\varepsilon^3)\right)^2\right)^{\frac{1}{2}} = \frac{1}{2} + O(\varepsilon)^2 \tag{69}$$

and

$$\left(O(\varepsilon)^2 + \sum_{k=1}^{N} O(\sqrt{k}\varepsilon^3)^2\right)^{\frac{1}{2}} = O(\varepsilon), \tag{70}$$

it follows that finally

$$\mathbb{P}_\varepsilon[\text{there is path through } \mathbf{b}_0, \mathbf{w}_1, \ldots, \mathbf{b}_{N-1}, \mathbf{w}_N] \le \left(\frac{1}{2}+O(\varepsilon^2)\right)^{N-1} \times O(\varepsilon) = O\left(\frac{\varepsilon}{2^N}\right). \tag{71}$$

Since the probability that there is a path of rhombi starting from $\mathbf{b}_0$ is of order ε, we obtain the wanted bound for the conditional probability

$$\mathbb{P}_\varepsilon\left[\text{the first } \left\lfloor\frac{L}{\varepsilon}\right\rfloor \text{ steps of } X_j^\varepsilon \text{ coincide with } (\omega_n) \mid X_j^\varepsilon \text{ starts at } (x_0, y_0)\right] \le \frac{c_L}{2^N} \tag{72}$$

for some constant c_L depending on L. This ends the proof of the last lemma, giving the last argument to end the proof of Theorem 2. □

Acknowledgements We are grateful to Richard Kenyon for enlightening discussions on dimer models, which have motivated this work.

References

1. Baik, J., Kriecherbauer, T., McLaughlin, K.D.T.-R., Miller, P.D.: Uniform asymptotics for polynomials orthogonal with respect to a general class of discrete weights and universality results for associated ensembles: announcement of results. Int. Math. Res. Not. 821–858 (2003)
2. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
3. Boutillier, C.: The bead model & limit behaviors of dimer models (2006)
4. Dyson, F.J.: A Brownian-motion model for the eigenvalues of a random matrix. J. Math. Phys. **3**, 1191–1198 (1962)
5. Ferrari, P.L., Spohn, H.: Step fluctuations for a faceted crystal. J. Stat. Phys. **113**, 1–46 (2003)
6. Fisher, M.E.: Walks, walls, wetting, and melting. J. Stat. Phys. **34**, 667–729 (1984)
7. Fowler, R.H., Rushbrooke, G.S.: Statistical theory of perfect solutions. Trans. Faraday Soc. **33**, 1272–1294 (1937)
8. Gessel, I., Viennot, G.: Binomial determinants, paths, and hook length formulae. Adv. Math. **58**, 300–321 (1985)
9. Johansson, K.: Non-intersecting paths, random tilings and random matrices. Probab. Theory Relat. Fields **123**, 225–280 (2002)
10. Johansson, K.: The arctic circle boundary and the Airy process. Ann. Probab. **33**, 1–30 (2005)
11. Johansson, K., Nordenstam, E.: Eigenvalues of gue minors (2006)
12. Karlin, S., McGregor, J.: Coincidence probabilities. Pac. J. Math. **9**, 1141–1164 (1959)
13. Kasteleyn, P.W.: Dimer statistics and phase transitions. J. Math. Phys. **4**, 287–293 (1963)
14. Kasteleyn, P.W.: Graph theory and crystal physics. In: Graph Theory and Theoretical Physics, pp. 43–110. Academic Press, London (1967)
15. Katori, M., Nagao, T., Tanemura, H.: Infinite systems of non-colliding Brownian particles. In: Stochastic Analysis on Large Scale Interacting Systems. Advanced Studies in Pure Mathematics, vol. 39, pp. 283–306. Math. Soc. Japan, Tokyo (2004)
16. Katori, M., Tanemura, H.: Functional central limit theorems for vicious walkers. Stoch. Stoch. Rep. **75**, 369–390 (2003)
17. Kenyon, R.: Local statistics of lattice dimers. Ann. Inst. H. Poincaré Probab. Stat. **33**, 591–618 (1997)
18. Kenyon, R., Okounkov, A.: What is … a dimer? Not. Am. Math. Soc. **52**, 342–343 (2005)
19. Kenyon, R., Okounkov, A., Sheffield, S.: Dimers and amoebae. Ann. Math. (2) **163**, 1019–1056 (2006)
20. Mehta, M.L.: Random Matrices. Pure and Applied Mathematics, vol. 142. Elsevier/Academic Press, Amsterdam (2004)
21. Okounkov, A., Reshetikhin, N.: Random skew plane partitions and the Pearcey process (2005)
22. Okounkov, A., Reshetikhin, N.: The birth of a random matrix (2006)
23. Sheffield, S.: Random surfaces: large deviations principles and gradient Gibbs measure classifications. PhD thesis, Stanford University (2004)
24. Spohn, H.: Interacting Brownian particles: a study of Dyson's model. In: Hydrodynamic Behavior and Interacting Particle Systems. IMA Volumes in Mathematics and Its Applications, vol. 9, pp. 151–179. Springer, New York (1987)