DETERMINANTAL POINT PROCESSES CONDITIONED ON RANDOMLY INCOMPLETE CONFIGURATIONS

TOM CLAEYS AND GABRIEL GLESNER

Abstract. For a broad class of point processes, including determinantal point processes, we construct associated marked and conditional ensembles, which allow to study a random configuration in the point process, based on information about a randomly incomplete part of the configuration. We show that our construction yields a well behaving transformation of sufficiently regular point processes. In the case of determinantal point processes, we explain that special cases of the conditional ensembles already appear implicitly in the literature, namely in the study of unitary invariant random matrix ensembles, in the Its-Izergin-Korepin-Slavnov method to analyze Fredholm determinants, and in the study of number rigidity. As applications of our construction, we show that a class of determinantal point processes induced by orthogonal projection operators, including the sine, Airy, and Bessel point processes, satisfies a strengthened notion of number rigidity, and we give a probabilistic interpretation of the Its-Izergin-Korepin-Slavnov method.

Pour une large classe de processus de points, contenant les processus de points déterminantaux, nous construisons les ensembles marqués et conditionnés associés qui permettent d'étudier une configuration aléatoire dans le processus de points, à partir d'information sur une partie aléatoirement incomplète de cette configuration. Nous montrons que notre construction produit une transformation qui se comporte bien pour des processus de points suffisamment réguliers. Dans le cas des processus de points déterminantaux, nous expliquons que certains cas particuliers de ces ensembles conditionnés sont déjà apparus implicitement dans la littérature, à savoir dans l'étude des ensembles de matrices aléatoires unitairement invariant, dans la méthode de Its, Izergin, Korepin et Slavnov pour analyser les déterminants de Fredholm, et dans l'étude de la rigidité des nombres. Comme application de notre construction, nous montrons qu'une classe de processus de points déterminantaux induits par des projections orthogonales, comprenant les processus de points sinus, Airy et Bessel, satisfait une propriété plus forte que la rigidité des nombres, et nous donnons une interprétation probabiliste de la méthode de Its, Izergin, Korepin et Slavnov.

1. Introduction

1.1. Background and motivation. Determinantal point processes (DPPs) are point processes whose correlation functions can be written as determinants of a correlation kernel, and for which average multiplicative statistics are Fredholm determinants. Prominent examples of DPPs are the eigenvalue distributions of a large class of random matrix ensembles, distributions of particles in asymmetric exclusion processes and tiling models, distributions of non-intersecting random paths, and the zeros of Gaussian analytic functions. They are special cases of repulsive point processes, in which one can study relevant probabilistic quantities through the analysis of the correlation kernel and associated Fredholm determinants [1, 46, 48, 55, 56, 60].

A groundbreaking discovery for the development of random matrix theory and more generally the study of DPPs has been the observation of Wigner and Dyson and their collaborators in the 1960s that energy levels of heavy nuclei can be accurately modelled by eigenvalues of random matrices. Despite his spectacular contributions, when Dyson looked back at his work on heavy nuclei in 2002 during the MSRI program *Recent Progress in Random Matrix Theory and Its Applications*, he explained [38] that the practical implications of his work on random matrices in nuclear physics were disappointing, because detectors were imperfect, and missing or spurious energy levels corrupted the data. Inspired by this, Dyson raised the question to develop error-correcting code for random matrix eigenvalues: given an imperfect observed spectrum of a random matrix, can one detect missing or spurious eigenvalues? This would not be possible for point processes with independent points, because the positions of a fraction of the points in the process do not carry any information about the other points. In strongly correlated point configurations such as random matrix eigenvalues or DPPs, one can however hope to extract information based on incomplete data. According to [38], Dyson did not suggest this direction of research because of its importance in nuclear physics, but purely because he believed it would lead to interesting mathematics. This question has been explored by Bohigas and Pato [8, 9] using randomly thinned random matrix eigenvalues, and has been picked up in the mathematics literature with the study of random thinnings of DPPs [5, 15, 16, 17, 18, 28, 29, 30, 31, 41], but a general mathematical theory for extracting information from the observation of randomly thinned DPPs has not been developed so far.

However, in the same spirit of attempting to extract information about DPPs from a partial observation, the remarkable property of number rigidity has recently been investigated. Informally, a point process is said to be number rigid if the configuration of points outside any bounded set determines almost surely the number of points inside the set. Important DPPs like the sine, Airy, and Bessel point processes arising in random matrix theory, are known to be number rigid [21, 43, 45, 55], and in the case of the sine process, the distribution of the points inside a bounded set, conditioned on the configuration of points outside the set, has been studied and proved to converge to the sine process when the size of the interval grows [52].

In this work, for any sufficiently regular point process, and in particular for any DPP, we introduce a family of marked and conditional point processes which allow to formalize the following question: *Given a randomly incomplete sample of the point process, what can we say about the missing points?* Although these point processes have, to the best of our knowledge, not been introduced and studied on a general basis, special cases of them do already appear in the literature in various contexts, as we will explain in more detail later; firstly, unitarily invariant Hermitian random matrix ensembles are a special case of conditional ensembles associated to the Gaussian Unitary Ensemble (GUE); secondly, special cases of the conditional ensembles arise naturally in the Its-Izergin-Korepin-Slavnov (IIKS) [47] method to characterize Fredholm determinants via Riemann-Hilbert problems; and finally, special cases of the conditional ensembles have been studied in relation to number rigidity.

Our objectives are:

(1) to construct the marked and conditional ensembles rigorously;
(2) to prove that the conditional ensembles define a well-behaving transformation which preserves the structure of DPPs and of several interesting subclasses of DPPs;
(3) to introduce a refined notion of number ridigity and to show that important DPPs like the sine, Airy, and Bessel DPPs satisfy this notion of rigidity;
(4) to illustrate that the IIKS method provides an effective framework to study the conditional ensembles via Riemann-Hilbert methods.

1.2. **DPPs: generalities and main examples.** Consider a measure space $(\Lambda, \mathcal{B}_\Lambda, \mu)$, with Λ a complete separable metric space, $\mathcal{B}_\Lambda$ the Borel σ-algebra, and μ a locally[1] finite positive Borel measure on Λ, i.e. satisfying $\mu(B) < \infty$ for any bounded $B \in \mathcal{B}_\Lambda$. We will be mainly interested in $\Lambda = \mathbb{R}$ with the Lebesgue measure or Λ the unit circle in the complex plane with the arc length measure, and the reader may prefer to keep only these examples in mind for the sake of simplicity. Let $\mathbb{P}$ be a simple point process on Λ, i.e. a probability measure on the set $\mathcal{N}(\Lambda)$ of locally finite point configurations in Λ (see Section 2 for a more precise definition of the probability space), such that there are a.s. no points with multiplicity > 1. We can represent such a configuration $\xi \in \mathcal{N}(\Lambda)$ as a locally finite counting measure

$$\xi = \sum_{j \in J} \delta_{x_j},$$

[1]Here and for the rest of this paper, whenever we say that a property holds locally, we mean that it holds for any bounded Borel set.

where J is a countable index set, and $x_j \in \Lambda$, $x_i \neq x_j$ when $i \neq j$. Recall (see e.g. [34, Section 9.4]) that a simple point process on Λ is characterized uniquely by its Laplace functional

$$\mathcal{L} : B_+(\Lambda) \to \mathbb{R}^+ : f \mapsto \mathcal{L}[f], \qquad \mathcal{L}[f] = \mathbb{E}e^{-\sum_{x \in \operatorname{supp}\xi} f(x)} = \mathbb{E}e^{-\int_\Lambda f \mathrm{d}\xi},$$

where $B_+(\Lambda)$ is the space of bounded non-negative measurable functions $f : \Lambda \to [0, +\infty)$ with bounded support.

Some of our results hold for any sufficiently regular point process, but our main focus will be on DPPs, for which the correlation functions $\rho_k : \Lambda^k \to [0, +\infty)$ (see again Section 2 for details) of all orders exist and can be written in terms of a correlation kernel $K(x_i, x_j)$ in determinantal form:

$$\rho_k(x_1, \dots, x_k) = \det\left(K(x_i, x_j)\right)_{i,j=1}^k. \tag{1.1}$$

If $K : \Lambda^2 \to \mathbb{C}$ is the kernel of a locally trace class operator K on $L^2(\Lambda, \mu)$, then the Laplace functional is a Fredholm determinant:

$$\mathcal{L}[f] = \det\left(1 - \mathrm{M}_{\sqrt{1-e^{-f}}}\mathrm{K}\mathrm{M}_{\sqrt{1-e^{-f}}}\right), \tag{1.2}$$

with M_g the multiplication operator with $g \in L^\infty(\Lambda, \mu)$ on $L^2(\Lambda, \mu)$, and the determinant is given by Fredholm's formula

$$\det\left(1 - \mathrm{M}_{\sqrt{g}}\mathrm{K}\mathrm{M}_{\sqrt{g}}\right) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \det\left(\sqrt{g(x_i)}K(x_i, x_j)\sqrt{g(x_j)}\right)_{i,j=1}^k \prod_{j=1}^k \mathrm{d}\mu(x_j). \tag{1.3}$$

Note that the kernel K might not be well defined on the diagonal of Λ^2, however we can always assume that $K(x, x)$ is chosen such that for any bounded Borel set B the following holds (see [60]):

$$\operatorname{Tr}\ \mathrm{K}|_{L^2(B,\mu)} = \int_B K(x, x)\mathrm{d}\mu(x).$$

For notational convenience, let us introduce a change of variable in the Laplace functional and define the *average multiplicative functional*

$$L[\phi] := \mathbb{E} \prod_{x \in \operatorname{supp}\xi} (1 - \phi)(x) = \mathcal{L}[-\log(1 - \phi)], \tag{1.4}$$

for $\phi : \Lambda \to [0, 1]$ measurable and with bounded support, such that $L[\phi] = \det\left(1 - \mathrm{M}_{\sqrt{\phi}}\mathrm{K}\mathrm{M}_{\sqrt{\phi}}\right)$ if $\mathbb{P}$ is the DPP with kernel of the operator K.

Besides DPPs, it will be insightful to keep in mind the example of a Poisson point process with bounded locally integrable intensity $\rho : \Lambda \to [0, +\infty)$, for which

$$\rho_k(x_1, \dots, x_k) = \prod_{j=1}^k \rho(x_j). \tag{1.5}$$

In Sections 3–5, we will consider some important subclasses of DPPs, which we already define now.

Example 1.1. Orthogonal polynomial ensembles (OPEs). *Let N be a positive integer and consider the point process consisting of configurations of N real points $x_1, \dots, x_N$ with joint probability distribution*

$$\frac{1}{Z_N}\Delta(x_1, \dots, x_N)^2 \prod_{j=1}^N w(x_j)\mathrm{d}x_j, \qquad \Delta(x_1, \dots, x_N) = \prod_{1 \le i < j \le N}(x_j - x_i), \tag{1.6}$$

where Z_N is a normalization constant, and $w(x)$ is a non-negative integrable weight function decaying sufficiently fast as $x \to \pm\infty$, such that all the moments $\int_{\mathbb{R}} x^k w(x)dx$, $k \in \mathbb{N}$, exist. If $w(x) = e^{-2Nx^2}$, this is the distribution of the (re-scaled, such that the eigenvalues follow a semi-circle law on $[-1, 1]$) eigenvalues of a random matrix from the GUE. If $w(x) = x^\alpha e^{-Nx}1_{(0,+\infty)}(x)$ with $\alpha > -1$, it is the distribution of the eigenvalues of a random matrix in the Laguerre-Wishart ensemble. More generally, *if w takes the form $w(x) = e^{-NV(x)}$ with V real analytic and growing sufficiently fast at $\pm\infty$, (1.6) is the eigenvalue distribution of a random matrix in the unitary invariant ensemble*

$$\frac{1}{\widehat{Z}_N} e^{-N\mathrm{Tr}V(M)}\mathrm{d}M,$$

with $\mathrm{d}M$ the Lebesgue measure on the space of $N \times N$ Hermitian matrices, and $\widehat{Z}_N$ a normalization constant.

Similarly, let N be a positive integer and consider the point process consisting of configurations of N points $e^{it_1}, \dots, e^{it_N}$ on the unit circle in the complex plane with joint probability distribution

$$\frac{1}{Z_N} \prod_{1\leq j<k\leq N} |\Delta(e^{it_1}, \dots, e^{it_N})|^2 \prod_{j=1}^{N} w(e^{it_j})\mathrm{d}t_j, \qquad t_j \in [0, 2\pi), \tag{1.7}$$

where Z_N is a normalization constant, and $w(e^{it})$ is a non-negative integrable weight function. If $w(e^{it}) = 1$, this is the distribution of the eigenvalues of a random matrix from the Circular Unitary Ensemble (CUE), or in other words a Haar distributed $N \times N$ unitary matrix.

It is well-known that the above OPEs are DPPs, with correlation kernel K_N built out of orthogonal polynomials on the real line or on the unit circle. We will study these ensembles in more detail in Section 4.

Example 1.2. DPPs induced by orthogonal projection operators. *Consider a DPP with correlation kernel K whose associated integral operator K on $L^2(\Lambda, \mu)$, defined by*

$$\mathrm{K}f(x) = \int_\Lambda K(x, y)f(y)\mathrm{d}\mu(y), \tag{1.8}$$

is a locally trace class orthogonal projection onto a closed subspace H of $L^2(\Lambda, \mu)$. As we will see, the OPEs from Example 1.1 are of this form, and the associated projection operators are then of rank N. We recall from [60] that a DPP defined by the kernel of a Hermitian locally trace class operator K has the property that the number of particles is a.s. equal to N, i.e. $\mathbb{P}(\xi(\Lambda) = N) = 1$, if and only if K is a projection operator of rank N. We will also consider DPPs induced by infinite rank projection operators. Such DPPs arise for instance when taking scaling limits of the kernels K_N from Example 1.1: we mention the DPPs defined by the sine kernel, the Airy kernel, the edge Bessel kernel, and the bulk Bessel kernel [35, 51]. More complicated kernels associated to Painlevé equations and hierarchies (see [37] for an overview), arising as double scaling limits of OPEs, are also of this form. We will consider such DPPs and derive rigidity results for some of them in Section 3.

Example 1.3. DPPs with integrable kernels. *In line with the terminology of Its, Izergin, Korepin, and Slavnov [47], we say that a kernel $K(x, y)$ is k-integrable if it can be written in the form*

$$K(x, y) = \frac{\sum_{j=1}^{k} f_j(x)g_j(y)}{x - y} \qquad \text{with } \sum_{j=1}^{k} f_j(x)g_j(x) = 0, \tag{1.9}$$

for some functions $f_j, g_j : \Lambda \to \mathbb{C}$, $j = 1, \dots, k$. The previous examples of OPEs on the real line and on the unit circle are 2-integrable, and so are the sine point process, the Airy point process, and the Bessel point processes.

There are however many DPPs with integrable kernels that are not induced by projection operators. Indeed, if a kernel $K(x, y)$ defines a DPP on Λ, then any kernel of the form $\phi(x)K(x, y)$ with $\phi : \Lambda \to [0, 1]$ measurable also defines a DPP, namely the random thinning of the original DPP realized by removing each particle x in the support of a random point configuration ξ independently with probability $1 - \phi(x)$ [53]. If $K(x, y)$ is of integrable form, it is easy to see that the same is true for $\phi(x)K(x, y)$, but even if $K(x, y)$ defines an orthogonal projection operator, $\phi(x)K(x, y)$ in general does not define a projection operator. DPPs with integrable kernels will be our topic of interest in Section 5.

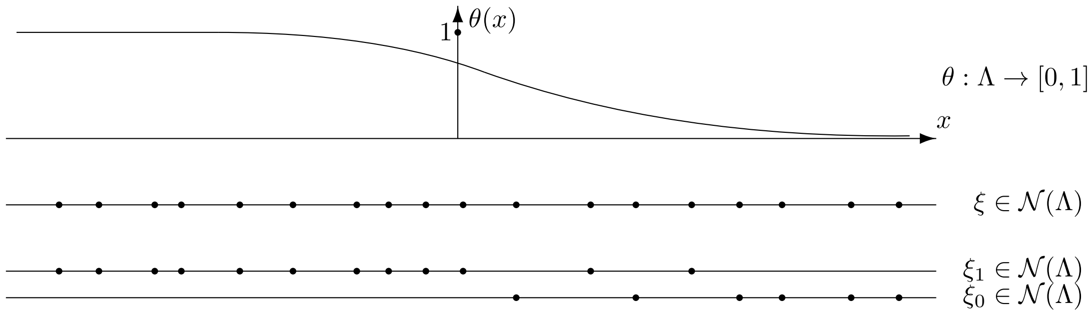

FIGURE 1. Illustration of the marked point process $\mathbb{P}^\theta$: at the top, we see the graph of a possible marking function θ; in the middle, a possible configuration ξ corresponding to the point process $\mathbb{P}$; at the bottom, possible associated mark 0 and mark 1 configurations ξ_0 and ξ_1 corresponding to $\mathbb{P}^\theta$.

1.3. **Marking and conditioning: informal construction and statement of results.** For any sufficiently regular point process $\mathbb{P}$, we can construct an associated marked point process in which we assign a random mark to each point independently. If the random mark is a Bernoulli random variable taking the value 0 or 1, then the marked point process is a point process on $\Lambda \times \{0,1\}$, in which we interpret the points with mark 1 as visible or observed particles, and the points with mark 0 as invisible or unobserved particles. Concretely, we mark the points in the DPP by introducing a measurable marking function $\theta : \Lambda \to [0,1]$, and by assigning mark 1 to particle x in a configuration of the DPP with probability $\theta(x)$, and mark 0 with probability $1-\theta(x)$. We denote the resulting marked point process as $\mathbb{P}^\theta$. The random marking splits a configuration ξ on Λ into configurations ξ_0 and ξ_1, where ξ_b is the configuration ξ restricted to the points with mark b. We denote $\mathbb{P}^\theta_b$, $b=0,1$, for the marginal probability distribution of ξ_b, which is a random position-dependent thinning of the ground process $\mathbb{P}$. We will introduce these marked point processes in detail in Section 2, and gather some of their general properties in Proposition 2.2. The point processes in which we are most interested here, are point processes obtained as conditional ensembles of this marked point process, by conditioning on the (observed) configuration of mark 1 points.

In the remaining part of this section, for the sake of simplicity, we will present our main results about these conditional ensembles only in the case where $\mathbb{P}$ is a DPP. We note however that most of our results hold for more general point processes. The theorems stated below are thus special cases of more general results, stated in full generality and proved in later sections.

In the simplest case, we condition on the event that no points have mark 1 (in other words, there are no observed particles). If this event has non-zero probability, then the resulting conditional point process, which we will denote as $\mathbb{P}^\theta_{|\emptyset}$, is defined in the classical sense, and configurations in this point process have support in $\Lambda \times \{0\}$. Hence, by omitting the marks, we can identify configurations in this point process with configurations on Λ, and identify $\mathbb{P}^\theta_{|\emptyset}$ with a point process on Λ. The following result about the point process transformation $\mathbb{P} \mapsto \mathbb{P}^\theta_{|\emptyset}$, which is part of the more general Theorem 2.4 in Section 2, will be fundamental for our concerns.

Theorem 1.1. *Let $\mathbb{P}$ be the DPP with kernel K of a locally trace class operator* K *and let $\theta : \Lambda \to [0,1]$ be measurable and such that $\mathrm{M}_{\sqrt{\theta}+1_B}\mathrm{K}\mathrm{M}_{\sqrt{\theta}+1_B}$ is trace class for any bounded Borel set B, and*

$$\mathbb{P}^\theta(\xi_1(\Lambda)=0) = \det(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}}) > 0.$$

Then $\mathbb{P}^\theta_{|\emptyset}$ is also a DPP, defined by the kernel of the $L^2(\Lambda,\mu)$-operator

$$\mathrm{M}_{1-\theta}\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}. \tag{1.10}$$

Remark 1.2. *If the locally trace-class operator K is self-adjoint, it induces a DPP if and only if* $0 \le \mathrm{K} \le 1$ [60]*. In this case, the condition that $\mathrm{M}_{\sqrt{\theta}+1_B}\mathrm{K}\mathrm{M}_{\sqrt{\theta}+1_B}$ is trace class for any bounded Borel* *set B, is equivalent to $\int(\sqrt{\theta(x)}+1_B(x))^2K(x,x)\mathrm{d}\mu(x)<\infty$, which is automatically satisfied whenever $\int\theta(x)K(x,x)d\mu(x)<\infty$, if $K(x,x)$ is locally integrable. The trace class condition is then practical to verify in concrete situations. However, for non self-adjoint operators K, $\mathrm{Tr}\,\mathrm{K}<\infty$ does not imply K being trace class, and then the condition that $\mathrm{M}_{\sqrt{\theta}+1_B}\mathrm{K}\mathrm{M}_{\sqrt{\theta}+1_B}$ is trace class cannot be verified directly by computing a trace. In such cases, one rather tries to prove that an operator is a composition of Hilbert-Schmidt operators, to prove that it is trace class.*

Remark 1.3. *Since*

$$\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}=(1-\mathrm{K}\mathrm{M}_\theta)^{-1}\mathrm{K}=\mathrm{K}+\mathrm{K}\mathrm{M}_{\sqrt{\theta}}(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}})^{-1}\mathrm{M}_{\sqrt{\theta}}\mathrm{K},$$

the operator $\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}$ indeed exists provided that $\det(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}})>0$. If K is self-adjoint, the operator (1.10) is in general not self-adjoint, however the operator

$$\mathrm{M}_{\sqrt{1-\theta}}\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}\mathrm{M}_{\sqrt{1-\theta}}$$

is self-adjoint, and it is readily verified that this operator induces the same DPP $\mathbb{P}^\theta_{|\emptyset}$. If K is a projection, then it is easily seen that (1.10) is equal to the conjugation $(1-\mathrm{M}_\theta\mathrm{K})\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}$ of K.

The probability to observe a given non-empty finite configuration of points in the marked point process will typically be zero, but we can still, $\mathbb{P}^\theta$-a.s., condition on such events by making use of disintegration and reduced Palm measures (see Section 2 for details). Given a mark 1 configuration $\mathbf{v}=\{v_1,\dots,v_m\}$, we will denote this conditional ensemble, which we will define properly in Section 2.4 below, as $\mathbb{P}^\theta_{|\mathbf{v}}$. Before stating our main result about $\mathbb{P}^\theta_{|\mathbf{v}}$ in the case where $\mathbb{P}$ is a DPP, we need to introduce the reduced Palm measure $\mathbb{P}_v$ of $\mathbb{P}$ associated to a point $v\in\Lambda$. This represents the conditional ensemble obtained by first conditioning $\mathbb{P}$ on the event $v\in\operatorname{supp}\xi$, and then removing the point v from the configuration. If $\mathbb{P}$ is the DPP with kernel K and if $K(v,v)>0$, then [58] $\mathbb{P}_v$ is also a DPP, with kernel

$$K_v(x,y)=\frac{\det\begin{pmatrix}K(x,y) & K(x,v)\\ K(v,y) & K(v,v)\end{pmatrix}}{K(v,v)}. \tag{1.11}$$

Similarly, we can condition $\mathbb{P}$ on the presence of a finite number of distinct points $\mathbf{v}=\{v_1,\dots,v_m\}$. This is consistent in the sense that the reduced Palm measure $\mathbb{P}_{\mathbf{v}}=\mathbb{P}_{v_1,\dots,v_m}$ is, for $\mu^{\otimes m}$-a.e. $\mathbf{v}\in\Lambda^m$ such that $\det(K(v_\ell,v_k))_{\ell,k=1}^m>0$, equal to the measure $\left((\mathbb{P}_{v_1})_{v_2}\cdots\right)_{v_m}$ obtained by iteratively conditioning on $v_1,\dots,v_m$, for any chosen order of the points. Let us for notational convenience write $K(\mathbf{v},\mathbf{v})$ for the $m\times m$ matrix $\left(K(v_\ell,v_k)\right)_{\ell,k=1}^m$, $K(x,\mathbf{v})$ for the row vector$(K(x,v_k))_{k=1}^m$, and $K(\mathbf{v},y)$ for the column vector $(K(v_\ell,y))_{\ell=1}^m$. If $\mathbb{P}$ is a DPP with kernel K and if $\det K(\mathbf{v},\mathbf{v})>0$, then $\mathbb{P}_{\mathbf{v}}$ is the DPP with kernel given by

$$K_{\mathbf{v}}(x,y)=\frac{\det\begin{pmatrix}K(x,y) & K(x,\mathbf{v})\\ K(\mathbf{v},y) & K(\mathbf{v},\mathbf{v})\end{pmatrix}}{\det K(\mathbf{v},\mathbf{v})}, \tag{1.12}$$

which defines a finite rank perturbation of K. Let us also set for consistency the convention that when $\mathbf{v}=\emptyset$, $\mathbb{P}_\emptyset=\mathbb{P}$ and $K_\emptyset=K$.

In analogy to and as a generalization of Theorem 1.1, we have the following result, which is part of the more general Theorem 2.7 below.

Theorem 1.4. *If $\mathbb{P}$ is the DPP with locally trace class operator K and $\theta\in L^\infty(\Lambda,\mu)$ is such that $\mathrm{M}_{\sqrt{\theta}+1_B}\mathrm{K}\mathrm{M}_{\sqrt{\theta}+1_B}$ is trace class for any bounded Borel set B, then for $\mathbb{P}^\theta$-a.e. ξ_1, writing $\mathbf{v}=\operatorname{supp}\xi_1$, we have $\det\left(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}_{\mathbf{v}}\mathrm{M}_{\sqrt{\theta}}\right)\neq0$, and $\mathbb{P}^\theta_{|\mathbf{v}}$ is also a DPP, defined by the $L^2(\Lambda,\mu)$-operator*

$$\mathrm{M}_{1-\theta}\mathrm{K}_{\mathbf{v}}\left(1-\mathrm{M}_\theta\mathrm{K}_{\mathbf{v}}\right)^{-1}. \tag{1.13}$$

Remark 1.5. *This result implies that the class of DPPs is stable under the transformation $\mathbb{P}\mapsto\mathbb{P}^\theta_{|\mathbf{v}}$. More is actually true: as we will see, each of the subclasses of DPPs defined in Examples 1.1–1.3 are also stable, and in Assumptions 2.1 below, we will define a larger class of (not necessarily determinantal) point processes which is stable under this transformation.*

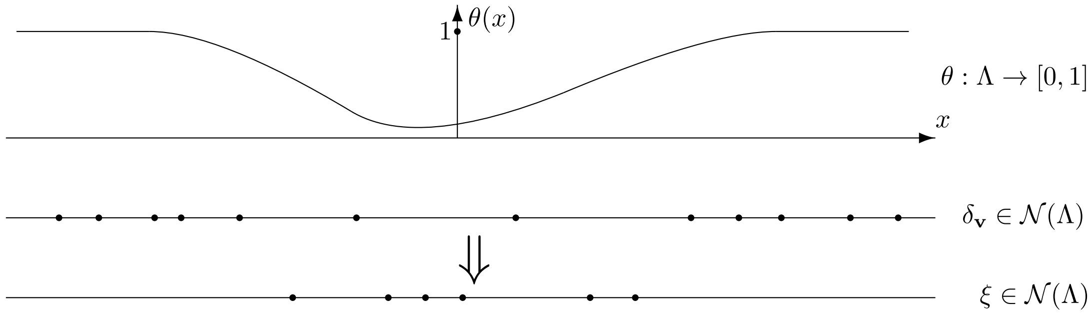

FIGURE 2. Illustration of marking rigidity: at the top, we see the graph of a possible marking function θ; at the bottom, a possible configuration of observed points $\delta_{\mathbf{v}}$ and a possible configuration ξ in the conditional ensemble $\mathbb{P}^{\theta}_{|\mathbf{v}}$. If $\mathbb{P}$ is marking rigid, then the marking function θ and the observed configuration $\delta_{\mathbf{v}}$ a.s. determine the number of points (6 in the picture) in the unobserved configuration ξ.

Section 2 will be devoted to the rigorous construction of the marked and conditional point processes $\mathbb{P}^{\theta}, \mathbb{P}^{\theta}_{|\emptyset}, \mathbb{P}^{\theta}_{|\mathbf{v}}$, and to the proofs of (generalizations of) the results stated above.

We should note that in the case where θ is the indicator function of a subset of Λ, all the above results are well-known, see e.g. [12, 22, 23].

1.4. **Rigidity.** In Section 3, we will study conditional ensembles corresponding to infinite configurations of mark 1 points $\delta_{\mathbf{v}} := \sum_j \delta_{v_j} \in \mathcal{N}(\Lambda)$. In such cases, the disintegration theorem implies that one can still define $\mathbb{P}^{\theta}_{|\mathbf{v}}$. If $B \in \mathcal{B}_{\Lambda}$ is bounded and if $\theta = 1_{B^c}$ is the indicator function of the complement of B, then $\mathbb{P}^{\theta}_{|\mathbf{v}} = \mathbb{P}^{1_{B^c}}_{|\mathbf{v}}$ is connected to the notion of number rigidity in the following manner. A point process $\mathbb{P}$ is said to be (number) rigid if for any bounded $B \in \mathcal{B}_{\Lambda}$, the conditional ensemble $\mathbb{P}^{1_{B^c}}_{|\mathbf{v}}$ has for $\mathbb{P}^{1_{B^c}}_{1}$-a.e. $\delta_{\mathbf{v}}$ a deterministic number of points, or in other words if there exists a $\mathcal{C}(\Lambda)$-measurable function

$$\ell : \mathcal{N}(\Lambda) \to \mathbb{N} \cup \{0, \infty\} : \delta_{\mathbf{v}} \mapsto \ell_{\mathbf{v}} \quad \text{such that} \quad \mathbb{P}^{1_{B^c}}_{|\mathbf{v}} \left(\xi(B) = \ell_{\mathbf{v}}\right) = 1.$$

This property is trivially satisfied for DPPs defined by kernels of finite rank orthogonal projections, since the number of particles in these DPPs is deterministic. Remarkably, a wide class of DPPs defined by kernels of infinite rank locally trace class orthogonal projections are also known to be number rigid [43, 45, 21]. Conversely, it is known [44] that a DPP can only be number rigid if it is defined by a projection. The construction of marked and conditional ensembles naturally suggests the following stronger notion of rigidity, which requires that given a.e. configuration of mark 1 points, the number of mark 0 point is deterministic.

Definition 1.6. A point process $\mathbb{P}$ is *marking rigid* if for any Borel measurable $\theta : \Lambda \to [0,1]$, there exists a Borel measurable function

$$\ell : \mathcal{N}(\Lambda) \to \mathbb{N} \cup \{0, \infty\} : \delta_{\mathbf{v}} = \sum_i \delta_{v_i} \mapsto \ell_{v_1, v_2, \ldots} =: \ell_{\mathbf{v}}$$

such that the following holds: for $\mathbb{P}^{\theta}_{1}$-a.e. $\delta_{\mathbf{v}}$,

$$\mathbb{P}^{\theta}_{|\mathbf{v}} \left(\xi(\Lambda) = \ell_{\mathbf{v}}\right) = 1.$$

Here $\xi(\Lambda)$ denotes the number of points of a random configuration ξ in the set Λ.
The following result is a special case of Theorem 3.5.

Theorem 1.7. *Let $\mathbb{P}$ be a DPP induced by a locally trace class orthogonal projection K such that the following holds: for any $\epsilon > 0$ and for any bounded $B \in \mathcal{B}_{\Lambda}$, there exists a bounded measurable function*

$f : \Lambda \to [0, +\infty)$ with bounded support such that

$$f|_B = 1, \qquad \mathrm{Var} \int_\Lambda f \mathrm{d}\xi < \epsilon,$$

where Var *denotes the variance with respect to $\mathbb{P}$. Then, $\mathbb{P}$ is marking rigid.*

Remark 1.8. *It is well-known that the existence of a function f as in the above statement for any bounded $B \in \mathcal{B}_\Lambda$ and $\epsilon > 0$, implies number rigidity of the point process $\mathbb{P}$ [43, 45], and it is also known that such f exists if $\mathbb{P}$ is a DPP with sufficiently regular 2-integrable kernel defining an orthogonal projection, such as the sine, Airy, and Bessel point processes [21]. We thus prove that these point processes are marking rigid.*

Remark 1.9. *The above result is trivial for DPPs induced by finite rank orthogonal projections, which a.s. have a deterministic number of points. For DPPs associated to infinite rank orthogonal projections, which have a.s. configurations with an infinite number of points, it is striking that the observation of a random (possibly infinite) part of a configuration determines a.s. the number of unobserved points.*

1.5. **Orthogonal polynomial ensembles.** In Section 4, we will focus on the OPEs from Example 1.1, and we will show that conditional ensembles of OPEs are also OPEs, but with a deformed weight function, see Proposition 4.1. As a consequence, for $\Lambda = \mathbb{R}$, we show that a large class of OPEs on the real line, which are eigenvalue distributions of unitarily invariant Hermitian random matrices, are in fact conditional ensembles of the GUE. We also give explicit expressions for the marginal distribution of the mark 0 points, given the number of mark 1 points. These are in general not DPPs, but do have a special structure involving Hankel determinants.

1.6. **DPPs with integrable kernels and Riemann-Hilbert problems.** In Section 5, we will consider DPPs associated to integrable kernels. We will show how we can characterize the kernels of the associated conditional ensembles in terms of Riemann-Hilbert problems via the IIKS method, and explain how this opens the door for asymptotic analysis and for deriving integrable differential equations associated to the conditional ensembles $\mathbb{P}^\theta_{|\mathbf{v}}$. We will also be able to interpret Jacobi's identity for Fredholm determinants in terms of the conditional ensembles $\mathbb{P}^\theta_{|\emptyset}$.

2. Construction of marked and conditional processes

2.1. **Preliminaries.** We consider a measurable space $(\Lambda, \mathcal{B}_\Lambda)$, where Λ is a complete separable metric space and $\mathcal{B}_\Lambda$ its Borel σ-algebra. We denote by $\mathcal{N}(\Lambda)$ the set of locally finite Borel counting measures on Λ, and by $\mathcal{C}(\Lambda)$ the σ-algebra generated by cylinder sets of the form

$$C = \bigcap_{i=1}^n \{\xi \in \mathcal{N}(\Lambda) \ : \ \xi(B_i) = k_i\},$$

where $B_1, \dots, B_n \in \mathcal{B}_\Lambda$ are disjoint and $n, k_1, \dots, k_n$ are non-negative integers. Note that we can identify $\mathcal{N}(\Lambda)$ with the space of locally finite sets of points, counted with multiplicity. For configurations of distinct points, this means that we identify the counting measure ξ with its support. We consider a point process $\mathbb{P}$ on Λ, i.e. a probability measure on the complete separable metric space $(\mathcal{N}(\Lambda), \mathcal{C}(\Lambda))$.

For disjoint sets $B_1, \dots, B_n \in \mathcal{B}_\Lambda$ and non negative integers $k_1, \dots, k_n$ such that $\sum_{j=1}^n k_j = m$, the m-th *factorial moment measure* M_m of $\mathbb{P}$ is the symmetric measure on Λ^m given by

$$M_m(B_1^{k_1} \times \cdots \times B_n^{k_n}) = \mathbb{E}\xi(B_1)^{[k_1]} \dots \xi(B_n)^{[k_n]}, \quad \text{with} \quad l^{[k]} = l(l-1)\cdots(l-k+1), \tag{2.1}$$

if the average exists. Similarly, the m-th *Janossy measure* of $\mathbb{P}$ (encoding its finite dimensional distributions) associated to $B \in \mathcal{B}_\Lambda$ is the symmetric measure on B^m given by

$$J_m^B(B_1^{k_1} \times \cdots \times B_n^{k_n}) = k_1! \cdots k_n! \mathbb{P}\big(\xi(B) = m, \ \xi(B_j) = k_j \text{ for } j = 1, ..., n\big),$$

where $\sum_{j=1}^n k_j = m$ and $\sqcup_{j=1}^n B_j = B$.

Throughout this section, we will impose the following regularity assumptions on the point process $\mathbb{P}$ on Λ.

Assumptions 2.1.
There exists a locally finite positive Borel measure μ on Λ such that:

(1) *the point process $\mathbb{P}$ is simple, i.e. for μ-a.e. $x \in \Lambda$, $\mathbb{P}(\xi(\{x\}) \le 1) = 1$;*
(2) *$\mathbb{P}$ admits correlation functions of all orders, i.e. for any positive integer m there exists a (symmetric) locally integrable function $\rho_m : \Lambda^m \to [0, +\infty)$ with respect to the measure $\mu^{\otimes m}$ on Λ^m such that*
$$\mathrm{d}M_m = \rho_m \mathrm{d}^m \mu;$$
(3) *for any bounded $B \in \mathcal{B}_\Lambda$, there exists $\epsilon_B > 0$ such that*
$$\sum_{m=1}^{\infty} \frac{(1+\epsilon_B)^m}{m!} M_m(B^m) < \infty.$$

Under these assumptions, it is a classical fact [54, 60] that the correlation functions ρ_m uniquely determine the point process $\mathbb{P}$. We also have [56] that for every bounded $B \in \mathcal{B}_\Lambda$, there exist locally integrable Janossy densities $j_m^B : \Lambda^m \to [0, +\infty)$ such that $\mathrm{d}J_m^B = j_m^B \mathrm{d}^m\mu$. Note that j_m^B is only defined on B^m, however under Assumptions 2.1, we have the identity

$$j^B(\mathbf{x}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{B^n} \rho(\mathbf{x} \sqcup \mathbf{y}) \mathrm{d}^n \mu(\mathbf{y}), \tag{2.2}$$

which allows to extend j_m^B to Λ^m, since the series converges in the space of locally integrable functions on Λ^m. Here we abbreviated

$$j^B(\mathbf{x}) := j_m^B(x_1, \ldots, x_m), \qquad \rho(\mathbf{x}) := \rho_m(x_1, \ldots, x_m),$$

because we interpret $\mathbf{x}$ either as a vector with m components $x_1, \ldots, x_m$ or as a configuration $\{x_1, \ldots, x_m\}$ of m (not necessarily distinct) points; $\rho(\mathbf{x} \sqcup \mathbf{y})$ then means $\rho_{m+n}(x_1, ..., x_m, y_1, ..., y_n)$ with $\mathbf{x} = (x_1, \ldots, x_m)$, $\mathbf{y} = (y_1, \ldots, y_n)$. This notation in which we neglect the order of the variables is justified because ρ_m and j_m^B are symmetric in their variables. Moreover, if Assumptions 2.1 (3) holds also globally, i.e. for $B = \Lambda$, we have

$$\rho(\mathbf{x}) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\Lambda^n} j^\Lambda(\mathbf{x} \sqcup \mathbf{y}) \mathrm{d}^n \mu(\mathbf{y}). \tag{2.3}$$

The above formulas continue to hold for $m = 0$ by adopting the conventions

$$\Lambda^0 = B^0 = \{\emptyset\}, \quad J_0^B(\emptyset) = j^B(\emptyset) = \mathbb{P}(\xi(B) = 0), \quad M_0(\emptyset) = \rho(\emptyset) = 1, \quad \mu^{\otimes 0} = \delta_\emptyset.$$

Let us note first that the Poisson point process with locally bounded intensity $\rho : \Lambda \to [0, +\infty)$ on (Λ, μ) satisfies Assumptions 2.1 if μ is non-atomic, with correlation functions given by (1.5). Our interest goes in particular to DPPs, characterized by the kernel $K : \Lambda^2 \to \mathbb{C}$ of a locally trace class operator K on $L^2(\Lambda, \mu)$. These point processes are simple [60], and the correlation functions are locally integrable and given by

$$\rho(x_1, \ldots, x_m) = \det\left(K(x_j, x_k)\right)_{j,k=1}^m.$$

The average multiplicative functional is a Fredholm determinant, recall (1.2)–(1.4). In particular, by (1.3), we have for any $\epsilon > 0$ and bounded $B \in \mathcal{B}_\Lambda$ that

$$\sum_{m=0}^{\infty} \frac{(1+\epsilon)^m}{m!} M_m(B^m) = \det(1 + (1+\epsilon)\mathrm{M}_{1_B} \mathrm{K} \mathrm{M}_{1_B}) < \infty.$$

Hence, we can conclude that Assumptions 2.1 are satisfied when $\mathbb{P}$ is a DPP induced by a locally trace class operator K.

2.2. **Bernoulli marking.** Given a point process satisfying Assumptions 2.1 and a measurable function $\theta : \Lambda \to [0,1]$, we now construct a marked point process $\mathbb{P}^\theta$ on $\Lambda \times \{0,1\}$, by assigning to each point $x \in \Lambda$ independently a random Bernoulli variable which takes the value 1 with probability $\theta(x)$, and the value 0 with probability $1-\theta(x)$. Let us define the measures ν_x^θ and μ^θ respectively on $\{0,1\}$ and $\Lambda_{\{0,1\}} := \Lambda \times \{0,1\}$ as

$$\nu_x^\theta = (1-\theta(x))\delta_0 + \theta(x)\delta_1, \qquad \mathrm{d}\mu^\theta(x;b) = \mathrm{d}\nu_x^\theta(b)\mathrm{d}\mu(x), \qquad x \in \Lambda,\ b \in \{0,1\}. \tag{2.4}$$

This marked point process $\mathbb{P}^\theta$ satisfies Assumptions 2.1 with Λ replaced by $\Lambda_{\{0,1\}}$ and μ by μ^θ. The correlation functions are then simply given by

$$\rho_m^\theta((x_1,b_1),\dots,(x_m,b_m)) = \rho_m(x_1,\dots,x_m), \tag{2.5}$$

with respect to the measure μ^θ, and hence do not depend on the marks. As a direct consequence of the expression for the correlation functions, if the ground process $\mathbb{P}$ is determinantal and induced by the operator K on $L^2(\Lambda,\mu)$ with kernel $K : \Lambda^2 \to \mathbb{C}$, then the marked point process $\mathbb{P}^\theta$ is also determinantal, induced by the operator K^θ on $L^2(\Lambda_{\{0,1\}},\mu^\theta)$ with kernel

$$K^\theta((x,b_x),(y,b_y)) := K(x,y), \tag{2.6}$$

which is independent of the marks.

Now for $b \in \{0,1\}$ and for a marked configuration $\xi_{0,1} \in \mathcal{N}(\Lambda_{\{0,1\}})$, we define $\xi_b \in \mathcal{N}(\Lambda)$ by

$$\xi_b(B) = \xi_{0,1}(B \times \{b\}), \quad B \in \mathcal{B}_\Lambda, \tag{2.7}$$

i.e. ξ_b is the configuration of points with mark b, or equivalently

$$\xi_b = \sum_{j:b_j=b} \delta_{x_j}, \quad \text{when} \quad \xi_{0,1} = \sum_j \delta_{(x_j,b_j)} = \xi_0 \otimes \delta_0 + \xi_1 \otimes \delta_1. \tag{2.8}$$

As explained in the introduction, we interpret ξ_1 as the configuration of observed particles and ξ_0 as the configuration of unobserved particles. If we define the Borel measures μ_b^θ on Λ for $b \in \{0,1\}$ by

$$\mathrm{d}\mu_b^\theta(x) = \theta_b(x)\mathrm{d}\mu(x), \qquad \theta_1 = \theta, \quad \theta_0 = 1-\theta, \tag{2.9}$$

then the point processes $\mathbb{P}_b^\theta$, $b=0,1$, obtained from $\mathbb{P}^\theta$ via transportation through the maps $\xi_{0,1} \mapsto \xi_b$, or in other words the marginal distributions of the mark b configurations, also satisfy Assumptions 2.1 with correlations functions $\rho_b^\theta(\mathbf{x}) = \rho(\mathbf{x})$ with respect to μ_b^θ and Janossy densities for any bounded $B \in \mathcal{B}_\Lambda$ given by

$$j_b^{\theta,B}(\mathbf{x}) = \sum_{n=0}^\infty \frac{(-1)^n}{n!} \int_{B^n} \rho(\mathbf{x} \sqcup \mathbf{y}) \mathrm{d}^n \mu_b^\theta(\mathbf{y}). \tag{2.10}$$

Both point processes $\mathbb{P}_0^\theta$ and $\mathbb{P}_1^\theta$ on Λ are random independent thinnings of the ground point process $\mathbb{P}$. If the ground process is determinantal and induced by the kernel of a locally trace class operator K on $L^2(\Lambda,\mu)$, then so is $\mathbb{P}_b^\theta$ with the same kernel, but now with the corresponding operator acting on $L^2(\Lambda,\mu_b^\theta)$ [53].

Summarizing the above, we have proved the following result.

Proposition 2.2. *Let $\mathbb{P}$ satisfy Assumptions 2.1, and let $\theta : \Lambda \to [0,1]$ be measurable.*

(1) The marked point process $\mathbb{P}^\theta$ satisfies Assumptions 2.1 with Λ replaced by $\Lambda_{\{0,1\}}$ and μ by μ^θ; for $b=0,1$, the component $\mathbb{P}_b^\theta$ satisfies Assumptions 2.1 with μ replaced by μ_b^θ ; in both cases the correlation functions are the same as those of the ground process $\mathbb{P}$.

(2) If $\mathbb{P}$ is the DPP with kernel K on (Λ,μ), then $\mathbb{P}^\theta$ is the DPP with kernel K^θ on $(\Lambda_{\{0,1\}},\mu^\theta)$. For $b=0,1$, the component $\mathbb{P}_b^\theta$ is the DPP with kernel K on (Λ,μ_b^θ).

Remark 2.3. *Observe the analogy with the corresponding result if $\mathbb{P}$ is the Poisson point process with intensity $\rho : \Lambda \to [0,+\infty)$ with respect to μ. Then $\mathbb{P}^\theta$ is the Poisson point process with intensity $\rho^\theta(x,b) = \rho(x)$ on $\Lambda_{\{0,1\}}$ with respect to μ^θ, and $\mathbb{P}_b^\theta$ is the Poisson point process on Λ with intensity ρ with respect to μ_b^θ.*

2.3. **Conditioning on an empty observation.** Let us now assume, in addition to Assumptions 2.1, that the probability to have no mark 1 particles is non-zero, i.e.

$$\mathbb{P}^\theta(\xi_1(\Lambda)=0)=\mathbb{E}\prod_{x\in\operatorname{supp}\xi}(1-\theta(x))=L[\theta]>0, \tag{2.11}$$

where we recall the definition of $L[.]$ from (1.4). Then, we can condition $\mathbb{P}^\theta$ on the event $\xi_1(\Lambda)=0$ in the classical sense and identify it with a point process on Λ by identifying $\xi_{0,1}$ with ξ_0, to obtain the conditional point process $\mathbb{P}^\theta_{|\emptyset}$ on Λ defined by

$$\mathbb{P}^\theta_{|\emptyset}(\xi\in C)=\frac{\mathbb{P}^\theta(\xi_{0,1}\text{ is such that }\xi_0\in C,\ \xi_1(\Lambda)=0)}{\mathbb{P}^\theta(\xi_1(\Lambda)=0)},\qquad C\in\mathcal{C}(\Lambda). \tag{2.12}$$

We write $L^\theta_{|\emptyset}$ for the average multiplicative functional (1.4) corresponding to the probability $\mathbb{P}^\theta_{|\emptyset}$. For K locally trace class on (Λ,μ) such that $\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}}$ is trace class and $\det(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}})>0$, let us introduce $K^\theta_{|\emptyset}$ as the kernel of the integral operator

$$\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}:L^2(\Lambda,\mu)\to L^2(\Lambda,\mu),\qquad \mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}f(x)=\int_\Lambda K^\theta_{|\emptyset}(x,y)f(y)\mathrm{d}\mu(y),$$

and $\mathrm{K}^\theta_{|\emptyset}$ as the operator with the kernel $K^\theta_{|\emptyset}$ on $L^2(\Lambda,\mu^\theta_0)$,

$$\mathrm{K}^\theta_{|\emptyset}:L^2(\Lambda,\mu^\theta_0)\to L^2(\Lambda,\mu^\theta_0),\qquad \mathrm{K}^\theta_{|\emptyset}f(x)=\int_\Lambda K^\theta_{|\emptyset}(x,y)f(y)\mathrm{d}\mu^\theta_0(y).$$

Theorem 2.4. *Let $\theta:\Lambda\to[0,1]$ be measurable and let $\mathbb{P}$ be such that $L[\theta]>0$.*

(1) The point process $\mathbb{P}^\theta_{|\emptyset}$ is well-defined and has average multiplicative functional

$$L^\theta_{|\emptyset}[\phi]=\frac{L[1-(1-\phi)(1-\theta)]}{L[\theta]}.$$

If in addition $\mathbb{P}$ satisfies Assumptions 2.1 and there exists $\epsilon>1$ such that $L[-\epsilon\theta]<\infty$, then so does $\mathbb{P}^\theta_{|\emptyset}$, with correlations functions with respect to μ^θ_0 given by

$$\rho^\theta_{|\emptyset}(\mathbf{x})=\frac{j_1^{\theta,\Lambda}(\mathbf{x})}{\mathbb{P}^\theta(\xi_1(\Lambda)=0)}.$$

(2) If $\mathbb{P}$ is the DPP with kernel K of a locally trace class operator K and $\mathrm{M}_{\sqrt{\theta+1_B}}\mathrm{K}\mathrm{M}_{\sqrt{\theta+1_B}}$ is trace class for any bounded Borel set B, then $\mathbb{P}^\theta_{|\emptyset}$ is the DPP on (Λ,μ) with kernel of the integral operator (1.10) acting on $L^2(\Lambda,\mu)$, or equivalently the DPP on (Λ,μ^θ_0) with kernel $K^\theta_{|\emptyset}$. Moreover, if K is self-adjoint then $\mathrm{K}^\theta_{|\emptyset}$ is self-adjoint.

Proof. (1) By definition of conditional probability, for $\phi:\Lambda\to\mathbb{R}^+$ measurable, we have

$$L^\theta_{|\emptyset}[\phi]=\mathbb{E}^\theta_{|\emptyset}\prod_{u\in\operatorname{supp}\xi_0}(1-\phi(u))=\frac{\mathbb{E}\prod_{x\in\operatorname{supp}\xi}(1-\phi(x))(1-\theta(x))}{\mathbb{P}^\theta(\xi_1(\Lambda)=0)}$$
$$=\frac{\mathbb{E}\prod_{x\in\operatorname{supp}\xi}(1-\phi(x))(1-\theta(x))}{\mathbb{E}\prod_{x\in\operatorname{supp}\xi}(1-\theta(x))}=\frac{L[1-(1-\phi)(1-\theta)]}{L[\theta]}.$$

Now if $\mathbb{P}$ is simple, then so is $\mathbb{P}^\theta$ and a fortiori so is $\mathbb{P}^\theta_{|\emptyset}$, and the inequality for $\phi\geq 0$

$$L^\theta_{|\emptyset}[-\phi]\leq\frac{L[-\phi]}{L[\theta]}$$

shows that $\mathbb{P}^\theta_{|\emptyset}$ satisfies the third of Assumptions 2.1 whenever $\mathbb{P}$ does. It thus remains to compute the correlation functions. Note first that $L[-\epsilon\theta]<\infty$ implies that $\mathbb{P}^\theta_1$ satisfies the third of Assumptions 2.1 with $B=\Lambda$, so that the global Janossy densities $j_1^{\theta,\Lambda}$ are well-defined and given by (2.2). The computations hereafter then involve absolutely convergent series, and all the needed results of integration theory may be applied. Let $\eta = 1-(1-\theta)(1-\phi) = \theta + (1-\theta)\phi$, then

$$L[\eta] = \sum_{n\geq 0} \frac{(-1)^n}{n!} \int_{\Lambda^n} \rho_n(\mathbf{x}) \prod_{j=1}^n \eta(x_j) \mathrm{d}^n\mu(\mathbf{x}).$$

Writing $\mathbf{x} = \mathbf{y} \sqcup \mathbf{z}$ and using the symmetry of the measure $\rho_n(\mathbf{x})\mathrm{d}^n\mu(\mathbf{x})$ yields that each integral is equal to

$$\sum_{l=0}^n \binom{n}{l} \int_{\Lambda^n} \rho_n(\mathbf{y} \sqcup \mathbf{z}) \prod_{j=1}^l (1-\theta(y_j))\phi(y_j) \prod_{i=1}^{n-l} \theta(z_i) \mathrm{d}^n\mu(\mathbf{y} \sqcup \mathbf{z}),$$

so that

$$L[\eta] = \sum_{l\geq 0} \frac{(-1)^l}{l!} \int_{\Lambda^l} \left[\sum_{n\geq l} \frac{(-1)^{n-l}}{(n-l)!} \int_{\Lambda^{n-l}} \rho(\mathbf{y} \sqcup \mathbf{z}) \mathrm{d}^{n-l}\mu_1^\theta(\mathbf{z}) \right] \prod_{j=1}^l (1-\theta(y_j))\phi(y_j) \mathrm{d}^l\mu_0^\theta(\mathbf{y}).$$

We recognize expression (2.2) for $j_1^{\theta,\Lambda}$ in the integral. Dividing the previous equation by $\mathbb{P}^\theta(\xi_1(\Lambda)=0) = L[\theta]$, we get an expression for $L^\theta_{|\emptyset}[\phi]$, and when $\phi = -1_B$, this implies the existence of all factorial moment measure $M^\theta_{m|\emptyset}$ of $\mathbb{P}^\theta_{|\emptyset}$, given the estimate $M^\theta_{m|\emptyset}(B^m) \leq L^\theta_{|\emptyset}[-1_B]$. Replacing ϕ by $w\phi$ for $w \in \mathbb{C}$ with a small enough modulus, we obtain a power series in w and we can read off the expressions for the correlation functions $\mathrm{d}M^\theta_{m|\emptyset} = \rho^\theta_{m|\emptyset}\mathrm{d}^m\mu_0^\theta$ by looking at each power of w.

(2) Let $(B_n)_{n\in\mathbb{N}}$ be an exhausting increasing sequence of bounded Borel subsets of Λ, and let $\mathrm{K}_n = \mathrm{M}_{1_{B_n}}\mathrm{K}\mathrm{M}_{1_{B_n}}$. By (1.2)–(1.4), the associated conditional ensemble $(\mathbb{P}_n)^\theta_{|\emptyset}$ has average multiplicative functional equal to

$$\begin{aligned}\frac{\det\left(1-\mathrm{M}_{\phi+\theta-\phi\theta}\mathrm{K}_n\right)}{\det(1-\mathrm{M}_\theta\mathrm{K}_n)} &= \det\left[\left((1-\mathrm{M}_\theta\mathrm{K}_n) - \mathrm{M}_\phi\mathrm{M}_{1-\theta}\mathrm{K}_n\right)(1-\mathrm{M}_\theta\mathrm{K}_n)^{-1}\right] \\ &= \det\left[1-\mathrm{M}_\phi\mathrm{M}_{1-\theta}\mathrm{K}_n(1-\mathrm{M}_\theta\mathrm{K}_n)^{-1}\right],\end{aligned}$$

and it follows that $(\mathbb{P}_n)^\theta_{|\emptyset}$ is also determinantal on (Λ,μ) with kernel of the integral operator $\mathrm{M}_{1-\theta}\mathrm{K}_n(1-\mathrm{M}_\theta\mathrm{K}_n)^{-1}$. The left hand side in the above identity is equal to $\frac{\det\left(1-\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}\mathrm{K}_n\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}\right)}{\det(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}_n\mathrm{M}_{\sqrt{\theta}})}$ and as $n\to\infty$, it converges to

$$\frac{\det\left(1-\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}\right)}{\det(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}})} = L^\theta_{|\emptyset}[\phi],$$

since $\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}\mathrm{K}_n\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}$ and $\mathrm{M}_{\sqrt{\theta}}\mathrm{K}_n\mathrm{M}_{\sqrt{\theta}}$ converge in trace norm to $\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}$ and $\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}}$, since the latter two operators are trace class. Indeed, $\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}} = \mathrm{M}_{\sqrt{\theta+1_B}}\mathrm{K}\mathrm{M}_{\sqrt{\theta+1_B}}$ with $B=\emptyset$; $\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}$ can be decoomposed, with $B = \operatorname{supp}\phi$, as

$$\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}1_B\mathrm{K}1_B\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}} + 1_{B^c}\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}}1_{B^c} + 1_{B^c}\mathrm{M}_{\sqrt{\theta}}\mathrm{K}1_B\mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}} + \mathrm{M}_{\sqrt{\phi+\theta-\phi\theta}}1_B\mathrm{K}\mathrm{M}_{\sqrt{\theta}}1_{B^c}$$

and it is easy to see that each term is trace class.

Similarly, the right hand side converges as $n\to\infty$ to

$$\det\left[1-\mathrm{M}_{\sqrt{\phi}}\mathrm{M}_{1-\theta}\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}\mathrm{M}_{\sqrt{\phi}}\right],$$

since $\mathrm{M}_{\sqrt{\phi}}\mathrm{M}_{1-\theta}\mathrm{K}_n(1-\mathrm{M}_\theta\mathrm{K}_n)^{-1}\mathrm{M}_{\sqrt{\phi}}$ converges to $\mathrm{M}_{\sqrt{\phi}}\mathrm{M}_{1-\theta}\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}\mathrm{M}_{\sqrt{\phi}}$ in trace norm (note that we need the condition that $\mathrm{M}_{\sqrt{\theta+1_B}}\mathrm{K}\mathrm{M}_{\sqrt{\theta+1_B}}$ is trace class for any bounded Borel set B here again, in order to have $\mathrm{M}_{\sqrt{\phi}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}}$, $\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\phi}}$ trace class). Thus, $\mathbb{P}^\theta_{|\emptyset}$ is the DPP with kernel of the operator $\mathrm{M}_{1-\theta}\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}$ on $L^2(\Lambda,\mu)$, or equivalently the DPP on (Λ,μ_0^θ)

with kernel $K^\theta_{|\emptyset}$.
If K is self-adjoint on $L^2(\Lambda,\mu)$, then so is

$$\mathrm{K}(1-\mathrm{M}_\theta\mathrm{K})^{-1}=\mathrm{K}+\mathrm{K}\mathrm{M}_{\sqrt{\theta}}(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}})^{-1}\mathrm{M}_{\sqrt{\theta}}\mathrm{K},$$

as the sum of two self-adjoint operators, hence the kernel $K^\theta_{|\emptyset}$ defines a self-adjoint operator on $L^2(\Lambda,\mu^\theta_0)$ as well.

□

Remark 2.5. *If $\mathbb{P}$ is the Poisson point process with intensity ρ on Λ with respect to μ, then $\mathbb{P}^\theta_{|\emptyset}$ is the Poisson point process with the same intensity ρ on Λ, but with respect to μ^θ_0. Hence, $\mathbb{P}^\theta_{|\emptyset}$ is equal to $\mathbb{P}^\theta_0$, and as it should be, the fact that there are no mark 1 points does not give any further information about the mark 0 points.*

Remark 2.6. *Theorem 1.1 is a restatement of the second part of the above result.*

2.4. **Conditioning on a finite mark 1 configuration ξ_1.** For non-empty configurations ξ_1 of points with mark 1, the situation is more involved. Here we need to assume that θ is such that there exists $\epsilon>0$ such that

$$L[-(1+\epsilon)\theta]=\mathbb{E}\prod_{x\in\operatorname{supp}\xi}(1+(1+\epsilon)\theta(x))<\infty, \tag{2.13}$$

where the average is with respect to the ground process $\mathbb{P}$. This condition ensures, by (2.1), that $\mathbb{P}^\theta_1$ satisfies Assumptions 2.1 (3) also for $B=\Lambda$, and in particular that

$$\mathbb{E}^\theta\xi_1(\Lambda)=\mathbb{E}\sum_{x\in\operatorname{supp}\xi}\theta(x)\le\mathbb{E}\prod_{x\in\operatorname{supp}\xi}(1+\theta(x))<\infty.$$

This implies that the number of observed particles $\xi_1(\Lambda)$ is finite for $\mathbb{P}^\theta$-a.e. ξ_1. Based on such an observed configuration ξ_1, we would like to obtain information about the configuration ξ_0 of points with mark 0. To this end, we want to define a point process $\mathbb{P}^\theta_{|\mathbf{v}}$ on $\Lambda\times\{0\}$ representing the restriction to $\Lambda\times\{0\}$ of the conditioning of $\mathbb{P}^\theta$ on an observation $\mathbf{v}=\{v_1,\dots,v_m\}$, or more precisely on ξ_1 being equal to $\delta_{\mathbf{v}}:=\sum_{j=1}^m\delta_{v_j}$. We can then identify $\mathbb{P}^\theta_{|\mathbf{v}}$ with a point process on Λ by omitting the marks 0. The probability to observe given points $\mathbf{v}$ with mark 1 will typically be zero, such that we cannot use classical conditional probability to construct the conditional point processes.

Conditioning on m mark 1 points. Let us assume that $\mathbb{P}^\theta(\xi_1(\Lambda)=m)>0$. Then we can condition $\mathbb{P}^\theta$ on the event $\xi_1(\Lambda)=m$ in the classical sense. Now, we want to construct a family of conditional point processes $\left\{\mathbb{P}^\theta_{|\mathbf{v}}\right\}_{\mathbf{v}\in\mathbf{\Lambda^m}}$, which is consistent in the sense that averaging the $\mathbb{P}^\theta_{|\mathbf{v}}$-probability of an event $\xi_0\in C\in\mathcal{C}(\Lambda)$ over the positions of the m-point configuration $v_1,\dots,v_m$ (with respect to the probability $\mathbb{P}^\theta_1(.|\xi_1(\Lambda)=m)$) is equal to the $\mathbb{P}^\theta(.|\xi_1(\Lambda)=m)$-probability of the event $\xi_0\in C$. In other words, we average $v_1,\dots,v_m$ with respect to the joint probability distribution

$$d\pi^\theta_{1,m}(\mathbf{v})=d\pi^\theta_{1,m}(v_1,\dots,v_m):=\frac{j^{\theta,\Lambda}_{1,m}(v_1,\dots,v_m)}{m!\mathbb{P}^\theta(\xi_1(\Lambda)=m)}\prod_{j=1}^m d\mu^\theta_1(v_j), \tag{2.14}$$

for $v_1,\dots,v_m$, where $j^{\theta,\Lambda}_{1,m}$ is the m-th order global Janossy density of the measure $\mathbb{P}^\theta_1$ for the mark 1 configuration (which exists if (2.13) holds), and we will need consistency in the sense that

$$\int_{\Lambda^m}\mathbb{P}^\theta_{|\mathbf{v}}(\xi\in C)d\pi^\theta_{1,m}(\mathbf{v})=\mathbb{P}^\theta\left(\xi_0\in C|\ \xi_1(\Lambda)=m\right),\qquad C\in\mathcal{C}(\Lambda). \tag{2.15}$$

Preliminaries on reduced Palm measures. As explained in Section 1, to construct $\mathbb{P}^\theta_{|\mathbf{v}}$, we need reduced local Palm distributions. Given a point process $\mathbb{P}$ satisfying Assumptions 2.1 and $m \in \mathbb{N}$, there exists a family of point processes $\{\mathbb{P}_\mathbf{w}\}_{\mathbf{w}\in\Lambda^m}$, which represent the conditioning of $\mathbb{P}$ on m points $\mathbf{w} = \{w_1, \dots, w_m\} \subset \operatorname{supp}\xi$, reduced by mapping $\xi \in \mathcal{N}(\Lambda)$ to its restriction $\xi|_{\Lambda\setminus\mathbf{w}}$. We need the following fundamental properties (see e.g. [34]) of these m-th order reduced Palm measures.

(1) For any $C \in \mathcal{C}(\Lambda)$, the map $\mathbf{w} \in \Lambda^m \mapsto \mathbb{P}_\mathbf{w}(C)$ is $\mathcal{B}_{\Lambda^m}$-measurable.
(2) For $\mu^{\otimes m}$-a.e. $\mathbf{w} \in \Lambda^m$ such that $\rho(\mathbf{w}) > 0$, the reduced Palm measure $\mathbb{P}_\mathbf{w}$ satisfies Assumptions 2.1, and its correlation functions $\rho_\mathbf{w}$ with respect to μ are given by (see [58])

$$\rho_\mathbf{w}(\mathbf{x}) = \frac{\rho(\mathbf{x} \sqcup \mathbf{w})}{\rho(\mathbf{w})}. \tag{2.16}$$

(3) Writing $\delta_\mathbf{w} = \sum_{j=1}^m \delta_{w_j}$, we have for any measurable $\psi : \Lambda^m \times \mathcal{N}(\Lambda) \to \mathbb{R}^+$ that the disintegration

$$\mathbb{E} \sum \psi(\mathbf{w}; \xi - \delta_\mathbf{w}) = \int_{\Lambda^m} \mathbb{E}_\mathbf{w} \psi(\mathbf{w}, \xi) \rho(\mathbf{w}) \mathrm{d}^m \mu(\mathbf{w}) \tag{2.17}$$

holds, where the sum at the left is over all ordered m-tuples $\mathbf{w} = (w_1, \dots, w_m)$ of distinct points in $\operatorname{supp}\xi$ and where $\mathbb{E}_\mathbf{w}$ is the average with respect to $\mathbb{P}_\mathbf{w}$.

In particular, the second property implies that if $\mathbb{P}$ is determinantal with kernel K, then for $\mu^{\otimes m}$-a.e. $\mathbf{w} \in \Lambda^m$ such that $\det K(\mathbf{w}, \mathbf{w}) > 0$, the reduced Palm measure $\mathbb{P}_\mathbf{w}$ is determinantal and induced by the kernel $K_\mathbf{w}$ given by (1.12), or equivalently by

$$K_\mathbf{w}(x, y) = K(x, y) - K(x, \mathbf{w}) K(\mathbf{w}, \mathbf{w})^{-1} K(\mathbf{w}, y), \tag{2.18}$$

where we used the block determinant formula

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det \left(A - BD^{-1}C\right) \det D, \tag{2.19}$$

and where similarly as before, $K(\mathbf{w}, \mathbf{w})$ represents an $m \times m$ matrix, $K(x, \mathbf{w})$ a row vector, and $K(\mathbf{w}, y)$ a column vector.

Construction of the conditional ensembles. We will now apply the above properties of reduced Palm measures to the point process $\mathbb{P}^\theta\left(.|\xi_1(\Lambda) = m\right)$, the marked point process conditioned on observing exactly m particles. If $\mathbb{P}^\theta(\xi_1(\Lambda) = m) > 0$, this point process indeed satisfies Assumptions 2.1. Setting $\mathbf{w} = \{(v_1, 1), \dots, (v_m, 1)\}$ and $\mathbf{v} = \{v_1, \dots, v_m\}$, we define $\mathbb{P}^\theta_{|\mathbf{v}}$ as the m-th order reduced local Palm distribution of $\mathbb{P}^\theta\left(.|\xi_1(\Lambda) = m\right)$ associated to the points $\mathbf{w}$. This is a point process on $\Lambda_{\{0,1\}}$ whose configurations have a.s. no points in $\Lambda \times \{1\}$; hence we can identify $\mathbb{P}^\theta_{|\mathbf{v}}$ with a point process on Λ by omitting the marks. Before we prove some important properties of the conditional ensembles $\mathbb{P}^\theta_{.|\mathbf{v}}$, let us mention that another intuitive way of defining them would be to first take the Palm measure of $\mathbb{P}^\theta$ at $\mathbf{w} = ((v_1, 1), ..., (v_m, 1))$ and then condition on there being no other particles with mark 1. The third item of the next result shows that this is indeed equivalent to our definition, and when $\mathbb{P}$ is a DPP, it has the advantage that it allows us to define the DPP $\mathbb{P}^\theta_{|\mathbf{v}}$ without need to pass via the (in general not determinantal) point process $\mathbb{P}^\theta(.|\xi_1(\Lambda) = m)$. Thus for K a locally trace class operator on (Λ, μ) such that $\mathrm{M}_\theta\mathrm{K}$ is trace class and $\det(1 - \mathrm{M}_\theta\mathrm{K}_\mathbf{v}) > 0$, let us introduce $K^\theta_{|\mathbf{v}}$ as the kernel of the integral operator

$$\mathrm{K}_\mathbf{v}(1 - \mathrm{M}_\theta\mathrm{K}_\mathbf{v})^{-1} : L^2(\Lambda, \mu) \to L^2(\Lambda, \mu), \qquad \mathrm{K}_\mathbf{v}(1 - \mathrm{M}_\theta\mathrm{K}_\mathbf{v})^{-1} f(x) = \int_\Lambda K^\theta_{|\mathbf{v}}(x, y) f(y) \mathrm{d}\mu(y),$$

and $\mathrm{K}^\theta_{|\mathbf{v}}$ as the operator with the kernel $K^\theta_{|\mathbf{v}}$ on $L^2(\Lambda, \mu^\theta_0)$,

$$\mathrm{K}^\theta_{|\mathbf{v}} : L^2(\Lambda, \mu^\theta_0) \to L^2(\Lambda, \mu^\theta_0), \qquad \mathrm{K}^\theta_{|\mathbf{v}} f(x) = \int_\Lambda K^\theta_{|\mathbf{v}}(x, y) f(y) \mathrm{d}\mu^\theta_0(y).$$

Theorem 2.7. *Let $\mathbb{P}$ satisfy Assumptions 2.1, and let $\theta : \Lambda \to [0,1]$ be measurable and such that (2.13) holds. Let $m \geq 0$ be such that $\mathbb{P}^\theta(\xi_1(\Lambda) = m) > 0$. The family of point processes $\left\{\mathbb{P}^\theta_{|\mathbf{v}}\right\}_{\mathbf{v}\in\Lambda^m}$ satisfies the following properties.*

(1) For any $C \in \mathcal{C}(\Lambda)$, the map $\mathbf{v} \in \Lambda^m \mapsto \mathbb{P}^\theta_{|\mathbf{v}}(C)$ is $\mathcal{B}_{\Lambda^m}$-measurable.

(2) For any Borel measurable $\phi : \mathcal{N}(\Lambda_{\{0,1\}}) \to [0,+\infty)$, writing $\delta_{\mathbf{v}} = \sum_{j=1}^m \delta_{v_j}$, we have the disintegration

$$\mathbb{E}^\theta\left[\phi(\xi_{0,1}) \mid \xi_1(\Lambda) = m\right] = \int_{\Lambda^m} \mathbb{E}^\theta_{|\mathbf{v}}\phi(\xi \otimes \delta_0 + \delta_{\mathbf{v}} \otimes \delta_1)\mathrm{d}\pi^\theta_{1,m}(\mathbf{v}), \tag{2.20}$$

where $\pi^\theta_{1,m}$ is given by (2.14).

(3) For $\pi^\theta_{1,m}$-a.e. $\mathbf{v} \in \Lambda^m$, the point process $\mathbb{P}^\theta_{|\mathbf{v}}$ satisfies Assumptions 2.1, its correlation functions $\rho^\theta_{|\mathbf{v}}$ with respect to μ^θ_0 are given by

$$\rho^\theta_{|\mathbf{v}}(\mathbf{x}) = \frac{j_1^{\theta,\Lambda}(\mathbf{x} \sqcup \mathbf{v})}{j_1^{\theta,\Lambda}(\mathbf{v})},$$

and its average multiplicative functional is given by

$$L^\theta_{|\mathbf{v}}[\phi_0] = \frac{L_{\mathbf{v}}[1-(1-\theta)(1-\phi_0)]}{L_{\mathbf{v}}[\theta]}, \tag{2.21}$$

where $L_{\mathbf{v}}$ denotes the average multiplicative functional of the reduced Palm measure $\mathbb{P}_{\mathbf{v}}$ of the ground process $\mathbb{P}$ on Λ associated to the points $\mathbf{v}$.

(4) If $\mathbb{P}$ is the DPP on (Λ,μ) with kernel of the operator K on $L^2(\Lambda,\mu)$, and $\mathrm{M}_{\sqrt{\theta}+1_B}\mathrm{K}\mathrm{M}_{\sqrt{\theta}+1_B}$ is trace class for any bounded Borel set B, then for $\pi^\theta_{1,m}$-a.e. $\mathbf{v} \in \Lambda^m$, $\mathbb{P}^\theta_{|\mathbf{v}}$ is the DPP on (Λ,μ) with kernel $(1-\theta)(x)K^\theta_{|\mathbf{v}}(x,y)$ of the operator $\mathrm{M}_{1-\theta}\mathrm{K}_{\mathbf{v}}\left(1-\mathrm{M}_\theta\mathrm{K}_{\mathbf{v}}\right)^{-1}$ on $L^2(\Lambda,\mu)$, or equivalently the DPP on (Λ,μ^θ_0) with kernel $K^\theta_{|\mathbf{v}}$. Moreover, if K is self-adjoint on $L^2(\Lambda,\mu)$ then the operator $\mathrm{K}^\theta_{|\mathbf{v}}$ with kernel $K^\theta_{|\mathbf{v}}$ on $L^2(\Lambda,\mu^\theta_0)$ is self-adjoint.

Proof.
(1) This follows directly from the corresponding general property of reduced Palm measures.
(2) Applying (2.17) to $\mathbb{P} = \mathbb{P}^\theta(.|\xi_1(\Lambda) = m)$ and

$$\psi : \Lambda^m_{\{0,1\}} \times \mathcal{N}(\Lambda_{\{0,1\}}) \to [0,+\infty) : (\mathbf{w},\xi_{0,1}) \mapsto \begin{cases} \phi(\xi_{0,1}+\delta_{\mathbf{w}}) & \text{if } \xi_1(\Lambda)=0 \text{ and } \mathbf{w} \in (\Lambda\times\{1\})^m, \\ 0 & \text{otherwise,}\end{cases}$$

with $\phi : \mathcal{N}(\Lambda_{\{0,1\}}) \to [0,+\infty)$, and denoting $\mathbf{w} = ((v_1,1),\dots,(v_m,1))$, $\mathbf{v} = (v_1,\dots,v_m)$, we obtain a family of point processes $\left\{\mathbb{P}^\theta_{|\mathbf{v}}\right\}_{\mathbf{v}\in\Lambda^m}$ on Λ which is such that

$$m!\mathbb{E}^\theta\left[\phi(\xi_{0,1}) \mid \xi_1(\Lambda) = m\right] = \int_{\Lambda^m} \mathbb{E}^{\theta|m}_{\mathbf{w}}\phi(\xi\otimes\delta_0+\delta_{\mathbf{v}}\otimes\delta_1)\frac{j^{\theta,\Lambda}_{1,m}(\mathbf{v})}{\mathbb{P}^\theta(\xi_1(\Lambda)=m)}\mathrm{d}^m\mu^\theta_1(\mathbf{v}),$$

where we used the symmetry of ψ and the fact that there are $m!$ ordered m-tuples $\mathbf{w}$ in $\operatorname{supp}\xi_{0,1}$ at the left, and the fact that the m-point correlation function of $\mathbb{P}^\theta\left(.|\xi_1(\Lambda)=m\right)$ evaluated at $\mathbf{w}$ is equal to $\frac{j^{\theta,\Lambda}_{1,m}(\mathbf{v})}{\mathbb{P}^\theta(\xi_1(\Lambda)=m)}$ (by (2.10)) at the right. Using (2.14), we obtain the required disintegration.
(3) Let us apply (2.20) to the multiplicative statistic

$$\phi(\xi_{0,1}) = \prod_{(x,b)\in\operatorname{supp}\xi_{0,1}} (1-\phi_b(x)),$$

where $\phi_0,\phi_1 : \Lambda \to (-\infty,1]$ are Borel measurable and ϕ_0 has bounded support. If $\phi_1 = 0$, the disintegration implies that for $\pi^\theta_{1,m}$-a.e. $\mathbf{v} \in \Lambda^m$, $\mathbb{P}^\theta_{|\mathbf{v}}$ satisfies Assumptions 2.1 *(3)*, thereby justifying the computations hereafter involving series and integrals.

The right hand side of (2.20) is then equal to

$$\int_{\Lambda^m} L^\theta_{|\mathbf{v}}[\phi_0]\prod_{j=1}^m(1-\phi_1(v))\mathrm{d}\pi^\theta_{1,m}(\mathbf{v}) = \frac{1}{m!\mathbb{P}^\theta(\xi_1(\Lambda)=m)}\int_{\Lambda^m} L^\theta_{|\mathbf{v}}[\phi_0]\prod_{j=1}^m(1-\phi_1(v))j_1^{\theta,\Lambda}(\mathbf{v})\mathrm{d}^m\mu_1^\theta(\mathbf{v})$$

$$= \frac{1}{m!\mathbb{P}^\theta(\xi_1(\Lambda)=m)}\int_{\Lambda^m} L^\theta_{|\mathbf{v}}[\phi_0]\prod_{j=1}^m(1-\phi_1(v))\left(\sum_{n=0}^\infty\frac{(-1)^n}{n!}\int_{\Lambda^n}\rho(\mathbf{u}\sqcup\mathbf{v})\mathrm{d}^n\mu_0^\theta(\mathbf{u})\right)\mathrm{d}^m\mu_1^\theta(\mathbf{v})$$

$$= \frac{1}{m!\mathbb{P}^\theta(\xi_1(\Lambda)=m)}\int_{\Lambda^m} L^\theta_{|\mathbf{v}}[\phi_0]L_{\mathbf{v}}[\theta]\prod_{j=1}^m(1-\phi_1(v))\rho(\mathbf{v})\mathrm{d}^m\mu_1^\theta(\mathbf{v}),$$

by (2.14), (2.10), and (2.16).

The left hand side of (2.20) is equal to

$$\sum_{n=0}^\infty\frac{(-1)^n}{n!\ m!\mathbb{P}^\theta(\xi_1(\Lambda)=m)}\int_{\Lambda^m}\int_{\Lambda^n}\prod_{k=1}^n\phi_0(u_k)\rho(\mathbf{u}\sqcup\mathbf{v})\mathrm{d}^n\mu_0^\theta(\mathbf{u})\prod_{j=1}^m(1-\phi_1(v_j))\mathrm{d}^m\mu_1^\theta(\mathbf{v})$$

$$= \frac{1}{m!\mathbb{P}^\theta(\xi_1(\Lambda)=m)}\int_{\Lambda^m}\left(\sum_{n=0}^\infty\frac{(-1)^n}{n!}\int_{\Lambda^n}\prod_{k=1}^n\phi_0(u_k)\rho_{\mathbf{v}}(\mathbf{u})\mathrm{d}^n\mu_0^\theta(\mathbf{u})\right)\prod_{j=1}^m(1-\phi_1(v_j))\rho(\mathbf{v})\mathrm{d}^m\mu_1^\theta(\mathbf{v})$$

$$= \frac{1}{m!\mathbb{P}^\theta(\xi_1(\Lambda)=m)}\int_{\Lambda^m} L_{\mathbf{v}}[1-(1-\theta)(1-\phi_0)]\prod_{j=1}^m(1-\phi_1(v_j))\rho(\mathbf{v})\mathrm{d}^m\mu_1^\theta(\mathbf{v}).$$

Since both sides are equal for any choice of ϕ_1, we can conclude that (2.21) holds. To compute the correlation functions, we note that the transformation under consideration is the composition of taking the Palm measure and then conditioning on observing no particles, as the form of the average multiplicative functional reveals. Since for $\pi^\theta_{1,m}$-a.e $\mathbf{x}\in\Lambda^m$ one has $\rho(\mathbf{x})\geq j_1^{\theta,\Lambda}(\mathbf{x})>0$ by (2.10), the result follows from the corresponding one in Theorem 2.4 after noticing that the Janossy densities of the Palm measure are given by $j^B_{\mathbf{v}}(\mathbf{x})=\frac{j^B(\mathbf{x}\sqcup\mathbf{v})}{\rho(\mathbf{v})}$, while recalling the convention $j^B(\emptyset)=\mathbb{P}(\xi(B)=0)$.

(4) This follows after a straightforward computation from *(3)* and (1.4). If K is self-adjoint then so is $\mathrm{K}_{\mathbf{v}}$, thus the result follows again from *(3)* and Theorem 2.4. □

Remark 2.8. *The disintegration in part (2) of the above result is more general than* (2.15)*: it suffices indeed to take*

$$\phi(\xi_{0,1})=1_C(\xi_0)1_{\{\xi_1(\Lambda)=m\}}(\xi_{0,1}),$$

to recover (2.15)*.*

Remark 2.9. *For the Poisson point process on Λ with intensity ρ, the above result is again trivial. We then have that $\mathbb{P}^\theta_{|\mathbf{v}}=\mathbb{P}^\theta_{|\emptyset}=\mathbb{P}^\theta_0$, in other words the positions of the mark 1 points do not carry any information about the mark 0 points.*

Remark 2.10. *The last part of the above result implies Theorem 1.4.*

3. Number rigidity and DPPs corresponding to projection operators

3.1. **DPPs induced by orthogonal projections.** Let $\mathbb{P}$ be a DPP on Λ, defined by a correlation kernel K with respect to a locally finite positive Borel measure μ which is such that the associated operator $\mathrm{K}:L^2(\Lambda,\mu)\to L^2(\Lambda,\mu)$ is a locally trace class orthogonal projection, i.e. $0\leq\mathrm{K}\leq 1$ and $\mathrm{K}^2=\mathrm{K}$, onto a closed subspace H of $L^2(\Lambda,\mu)$. The rank of K can be finite or infinite, but the results in this section will only be non-trivial in the infinite rank case. We assume here that the kernel $K:\Lambda^2\to\mathbb{C}$ of K is such that $\mathrm{K}f(x)$ is defined for every $x\in\Lambda$ and for every $f\in L^2(\Lambda,\mu)$. Note that this is true whenever $K(x,.)\in L^2(\Lambda,\mu)$ for every $x\in\Lambda$, by the Cauchy-Schwarz inequality. Classical examples of admissible point processes are the sine, Airy, and Bessel point processes on the real line.

By Proposition 2.2, the marked point process associated to $\mathbb{P}$ with marking function θ is the DPP on $(\Lambda_{\{0,1\}}, \mu^\theta)$ with correlation kernel $K^\theta((x,b),(x',b')) = K(x,x')$, where we recall that μ^θ is given by (2.4). The induced operator K^θ acting on $L^2(\Lambda_{\{0,1\}}, \mu^\theta)$ is the orthogonal projection operator onto the space

$$H^\theta := \left\{h_{0,1} \in L^2(\Lambda_{\{0,1\}}, \mu^\theta) : h_{0,1}(.,0) = h_{0,1}(.,1) \in H\right\},$$

and it is straightforward to verify that $\dim H^\theta = \dim H$.

As in Section 2, let us consider the conditional measure obtained by conditioning the marked point process on a configuration of mark 1 points. Under the assumptions that $\mathbb{P}^\theta(\xi_1(\Lambda) = m) > 0$ and $\mathrm{M}_{\sqrt{\theta}+1_B}\mathrm{K}\mathrm{M}_{\sqrt{\theta}+1_B}$ is trace class for any bounded Borel set B, we know from Theorem 2.7 (4) that for $\pi^\theta_{1,m}$-a.e. $\mathbf{v} \in \Lambda^m$, the conditional measure $\mathbb{P}^\theta_{|\mathbf{v}}$ is the DPP induced by the operator

$$\mathrm{M}_{1-\theta}\mathrm{K}_{\mathbf{v}}(1 - \mathrm{M}_\theta \mathrm{K}_{\mathbf{v}})^{-1} = (1 - \mathrm{M}_\theta)\mathrm{K}_{\mathbf{v}}(1 - \mathrm{M}_\theta \mathrm{K}_{\mathbf{v}})^{-1}$$

on $L^2(\Lambda, \mu)$. Moreover, from (1.12), it is straightforward to verify that $\mathrm{K}_{\mathbf{v}}$ is the orthogonal projection on the subspace

$$H_{\mathbf{v}} = \overline{\{h \in H : h(v) = 0 \ \forall v \in \mathbf{v}\}}. \tag{3.1}$$

Consequently, since $\mathrm{K}_{\mathbf{v}}^2 = \mathrm{K}_{\mathbf{v}}$, the $L^2(\Lambda,\mu)$-operator $\mathrm{M}_{1-\theta}\mathrm{K}_{\mathbf{v}}(1 - \mathrm{M}_\theta \mathrm{K}_{\mathbf{v}})^{-1}$ inducing $\mathbb{P}^\theta_{|\mathbf{v}}$ is equal to a conjugation of $\mathrm{K}_{\mathbf{v}}$,

$$(1 - \mathrm{M}_\theta \mathrm{K}_{\mathbf{v}})\mathrm{K}_{\mathbf{v}}(1 - \mathrm{M}_\theta \mathrm{K}_{\mathbf{v}})^{-1},$$

and this implies that it is a (not necessarily self-adjoint) projection onto the subspace

$$H^\theta_{\mathbf{v}} := \overline{(1 - \mathrm{M}_\theta \mathrm{K}_{\mathbf{v}})H_{\mathbf{v}}} = \overline{(1 - \mathrm{M}_\theta)H_{\mathbf{v}}}, \tag{3.2}$$

with dimension equal to that of $H_{\mathbf{v}}$, and that the $L^2(\Lambda, \mu^\theta_0)$-operator $\mathrm{K}^\theta_{|\mathbf{v}}$ is the orthogonal projection onto $H_{\mathbf{v}}$. Indeed, $\mathrm{K}^\theta_{|\mathbf{v}}$ is Hermitian, and for $h \in H_{\mathbf{v}}$, we have

$$\mathrm{K}^\theta_{|\mathbf{v}} h = \mathrm{K}_{\mathbf{v}}(1 - \mathrm{M}_\theta \mathrm{K}_{\mathbf{v}})^{-1}\mathrm{M}_{1-\theta} h = \mathrm{K}_{\mathbf{v}}(1 - \mathrm{M}_\theta \mathrm{K}_{\mathbf{v}})^{-1}\mathrm{M}_{1-\theta}\mathrm{K}_{\mathbf{v}} h = \mathrm{K}_{\mathbf{v}} h = h. \tag{3.3}$$

Let us now consider the more general case where $\mathbb{P}$ is induced by a not necessarily Hermitian projection operator, say $\mathrm{K} = \mathrm{P}_{H,J}$ is the unique linear projection with range H and kernel $J^\perp$, where H, J are closed subspaces of $L^2(\Lambda, \mu)$ such that $H \oplus J^\perp = L^2(\Lambda, \mu)$. Note that the adjoint projection is given by $\mathrm{P}^*_{H,J} = \mathrm{P}_{J,H}$. Since $\phi \in L^2(\Lambda, \mu)$ can be identified with $\phi \in L^2(\Lambda, \mu^\theta_0)$, we can also see H, J as subspaces of $L^2(\Lambda, \mu^\theta_0)$. Examples of DPPs induced by non-Hermitian projections are biorthogonal ensembles and their scaling limits like the Pearcey DPP.

Proposition 3.1. *If* $\mathrm{K} = \mathrm{P}_{H,J}$*, then*

$$\mathrm{K}^\theta_{|\mathbf{v}} = \mathrm{P}_{H_{\mathbf{v}}, J_{\mathbf{v}}},$$

where $H_{\mathbf{v}}, J_{\mathbf{v}}$ *are seen as closed subspaces of* $L^2(\Lambda, \mu^\theta_0)$.

Proof. First we recall that the transformation $\mathrm{K}^\theta_{|\mathbf{v}}$ is obtained by first taking the reduced Palm measure and then conditioning on $\xi_1 = \emptyset$ in view of Theorem 2.7, so that it suffices to prove the result separately in the cases $\theta = 0$ and $\mathbf{v} = \emptyset$. The case $\theta = 0$ is straightforward from (1.12), while for $\mathbf{v} = \emptyset$, $\mathrm{K}^\theta_{|\emptyset}$ is a projection with range H by (3.3) (with $\mathbf{v} = \emptyset$; observe that these equalities continue to hold when K is not self-adjoint). Finally, to identity the kernel, it suffices to apply the previous reasoning to $(\mathrm{K}^\theta_{|\emptyset})^* = (\mathrm{K}^*)^\theta_{|\emptyset}$. □

DPPs induced by projections have the property that the number of points in a configuration is almost surely equal to the rank of the projection [60]. If $\mathbb{P}$ has configurations with a deterministic number of points, it is obvious that the same must hold for $\mathbb{P}^\theta_{|\mathbf{v}}$, for any finite configuration $\mathbf{v}$. Since the projection $P_{H_{\mathbf{v}}, J_{\mathbf{v}}}$ is also defined for infinite configurations $\mathbf{v}$, it is natural to ask whether the DPP induced by this projection can in such a situation still be interpreted as the conditional DPP $\mathbb{P}^\theta_{|\mathbf{v}}$. This is not true in general, see e.g. [24], but we will see below that $\mathbb{P}^\theta_{|\mathbf{v}}$ is under suitable assumptions induced by an orthogonal projection, albeit not necessarily equal to $P_{H_{\mathbf{v}}, J_{\mathbf{v}}}$.

3.2. **Disintegration.** We first show that the family of conditional ensembles $\left\{\mathbb{P}^\theta_{|\mathbf{v}}\right\}_{\delta_\mathbf{v}\in\mathcal{N}(\Lambda)}$ exists under general conditions, and then we rely on results from [25] to prove that $\mathbb{P}^\theta_{|\mathbf{v}}$ is a DPP induced by a Hermitian operator $\mathrm{K}^\theta_{|\mathbf{v}}$ if $\mathbb{P}$ is a DPP induced by an orthogonal projection.

Proposition 3.2. *Let $\theta:\Lambda\to[0,1]$ be measurable, and let $\mathbb{P}$ satisfy Assumptions 2.1. There exists a family of point processes $\left\{\mathbb{P}^\theta_{|\mathbf{v}}\right\}_{\delta_\mathbf{v}\in\mathcal{N}(\Lambda)}$ such that the following conditions hold.*

(1) The map $\delta_\mathbf{v}\in\mathcal{N}(\Lambda)\mapsto\mathbb{P}^\theta_{|\mathbf{v}}(C)$ is $\mathcal{C}(\Lambda)$-measurable for any $C\in\mathcal{C}(\Lambda)$, and the disintegration

$$\mathbb{P}^\theta_0(C)=\int_{\mathcal{N}(\Lambda)}\mathbb{P}^\theta_{|\mathbf{v}}(C)\mathrm{d}\mathbb{P}^\theta_1(\delta_\mathbf{v})$$

holds for any $C\in\mathcal{C}(\Lambda)$.

(2) If $\mathbb{P}$ is a DPP induced by an orthogonal projection with kernel $K:\Lambda^2\to\mathbb{R}$ such that $\mathrm{K}f(x)$ is defined for every $x\in\Lambda$ and for every $f\in L^2(\Lambda,\mu)$, then for $\mathbb{P}^\theta_1$-a.e. $\delta_\mathbf{v}$, $\mathbb{P}^\theta_{|\mathbf{v}}$ is a DPP induced by a Hermitian locally trace class operator $\mathrm{K}^\theta_{|\mathbf{v}}$.

Proof. Let us define $\widehat{\mathbb{P}}^\theta_{|\mathbf{v}}$ by disintegrating $\mathbb{P}^\theta$ with respect to the surjective mapping

$$r:\mathcal{N}(\Lambda_{\{0,1\}})\to\mathcal{N}(\Lambda\times\{1\}):\xi_0\otimes\delta_0+\xi_1\otimes\delta_1\mapsto\xi_1\otimes\delta_1.$$

The disintegration theorem then implies that the map $\delta_\mathbf{v}\mapsto\widehat{\mathbb{P}}^\theta_{|\mathbf{v}}(\widetilde{C})$ is Borel measurable for any $\widetilde{C}\in\mathcal{C}(\Lambda_{\{0,1\}})$, and that

$$\mathbb{P}^\theta(\widetilde{C})=\int_{\mathcal{N}(\Lambda_{\{0,1\}})}\widehat{\mathbb{P}}^\theta_{|\mathbf{v}}(\widetilde{C})\mathrm{d}\mathbb{P}^\theta(r^{-1}(\delta_\mathbf{v}\otimes\delta_1)).$$

Taking $\widetilde{C}=C\otimes\{\delta_0\}\subset\mathcal{N}(\Lambda\times\{0\})$ with $C\in\mathcal{C}(\Lambda)$ and defining $\mathbb{P}^\theta_{|\mathbf{v}}$ on Λ as $\mathbb{P}^\theta_{|\mathbf{v}}(C):=\widehat{\mathbb{P}}^\theta_{|\mathbf{v}}(\widetilde{C})$, this becomes

$$\mathbb{P}^\theta_0(C)=\int_{\mathcal{N}(\Lambda)}\mathbb{P}^\theta_{|\mathbf{v}}(C)\mathrm{d}\mathbb{P}^\theta_1(\delta_\mathbf{v}),$$

and part *(1)* of the theorem is proved.

Part (2) follows directly upon applying [25, Lemma 1.11] to the marked point process $\mathbb{P}^\theta$ and $W=\Lambda\times\{1\}$. ☐

Remark 3.3. *It is important to note that the operator $\mathrm{K}^\theta_{|\mathbf{v}}$ is not necessarily a projection in part (2) of the above result.*

3.3. **Marking rigidity.** We will now further refine our assumptions on $\mathbb{P}$, in order to obtain a sufficient condition for $\mathbb{P}$ to be marking rigid. Let us emphasize that we will not need $\mathbb{P}$ to be a DPP. However, DPPs induced by integrable orthogonal projection operators will provide our main example of point processes which satisfy the assumption below.

Assumptions 3.4. *$\mathbb{P}$ satisfies Assumptions 2.1 and is such that the following holds: for any $\epsilon>0$ and for any bounded $B\in\mathcal{B}_\Lambda$, there exists a bounded measurable function $f:\Lambda\to[0,+\infty)$ with bounded support such that*

$$f|_B=1,\qquad \mathrm{Var}\int_\Lambda f\mathrm{d}\xi<\epsilon,$$

where Var denotes the variance with respect to $\mathbb{P}$.

Theorem 3.5. *Let $\mathbb{P}$ satisfy Assumptions 3.4.*

(1) If for any measurable $\theta:\Lambda\to[0,1]$, $\mathbb{P}^\theta(\xi_0(\Lambda)<\infty)$ is either 0 or 1, then $\mathbb{P}$ is marking rigid.

(2) Let $\mathbb{P}$ be a DPP induced by a locally trace class orthogonal projection with kernel $K:\Lambda^2\to\mathbb{C}$ such that $\mathrm{K}f(x)$ is defined for every $x\in\Lambda$ and for every $f\in L^2(\Lambda,\mu)$. Then $\mathbb{P}$ is marking rigid, and for any measurable $\theta:\Lambda\to[0,1]$ such that $\mathrm{M}_{1-\theta}\mathrm{K}$ is trace class, the conditional ensemble $\mathbb{P}^\theta_{|\mathbf{v}}$ is for $\mathbb{P}^\theta_1$-a.e. $\delta_\mathbf{v}$ induced by a finite rank orthogonal projection $\mathrm{K}^\theta_{|\mathbf{v}}$.

Proof. Let us first consider the case where θ is such that $\mathbb{P}^\theta(\xi_0(\Lambda) < \infty) = 0$, when $\mathbb{P}^\theta$-a.s., we have $\xi_0(\Lambda) = \infty$. Then, by Proposition 3.2,

$$1 = \mathbb{P}_0^\theta(\xi(\Lambda) = \infty) = \int_{\mathcal{N}(\Lambda)} \mathbb{P}_{|\mathbf{v}}^\theta(\xi(\Lambda) = \infty) \mathrm{d}\mathbb{P}_1^\theta(\delta_\mathbf{v}),$$

and consequently $\mathbb{P}_{|\mathbf{v}}^\theta(\xi(\Lambda) = \infty) = 1$ for $\mathbb{P}_1^\theta$-a.e. $\delta_\mathbf{v}$.

We now assume that $\mathbb{P}^\theta(\xi_0(\Lambda) < \infty) = 1$. Let $\Lambda_1 \subset \Lambda_2 \subset \cdots$ be an exhausting sequence of bounded Borel subsets of Λ. First, we observe that for $\mathbb{P}^\theta$-a.s. ξ_0, $\xi_0(\Lambda \setminus \Lambda_n) = 0$ for n sufficiently large. Secondly, we take a sequence of positive numbers $\epsilon_1, \epsilon_2, \ldots$ which converges to 0 as $n \to \infty$, and we observe that by Assumptions 3.4, there exists a sequence of bounded measurable functions $f_1, f_2, \ldots$ with bounded support such that $f_n|_{\Lambda_n} = 1$, and such that $\mathrm{Var} \int_\Lambda f_n \mathrm{d}\xi < \epsilon_n$. This implies by Chebyshev's inequality that

$$\mathbb{P}\left(\left| \int_\Lambda f_n \mathrm{d}\xi - \mathbb{E} \int_\Lambda f_n \mathrm{d}\xi \right| \geq \delta \right) \leq \delta^{-2} \mathrm{Var} \int_\Lambda f_n \mathrm{d}\xi \leq \delta^{-2} \epsilon_n \to 0,$$

as $n \to \infty$ for any $\delta > 0$, and hence that there exists a subsequence $(f_{n_j})_{j \in \mathbb{N}}$ such that

$$\lim_{j \to \infty} \left(\int_\Lambda f_{n_j} \mathrm{d}\xi - \mathbb{E} \int_\Lambda f_{n_j} \mathrm{d}\xi \right) = 0, \qquad \text{for } \mathbb{P}\text{-a.e. } \xi,$$

or equivalently with $\hat{f}_n(x, b) = f_n(x)$

$$\lim_{j \to \infty} \left(\int_{\Lambda_{\{0,1\}}} \hat{f}_{n_j} \mathrm{d}\xi_{0,1} - \mathbb{E}^\theta \int_{\Lambda_{\{0,1\}}} \hat{f}_{n_j} \mathrm{d}\xi_{0,1} \right) = 0, \qquad \text{for } \mathbb{P}^\theta\text{-a.e. } \xi_{0,1}.$$

For any $\xi_{0,1} \in \mathcal{N}(\Lambda_{\{0,1\}})$, we can write $\xi_0(\Lambda)$ as

$$\begin{aligned} \xi_0(\Lambda) &= \int_{\Lambda_{\{0,1\}}} 1_{\{b=0\}} \mathrm{d}\xi_{0,1}(x, b) = \int_{\Lambda_{\{0,1\}}} \hat{f}_{n_j} \mathrm{d}\xi_{0,1} + \int_\Lambda (1 - f_{n_j}) \mathrm{d}\xi_0 - \int_\Lambda f_{n_j} \mathrm{d}\xi_1 \\ &= \left(\int_{\Lambda_{\{0,1\}}} \hat{f}_{n_j} \mathrm{d}\xi_{0,1} - \mathbb{E}^\theta \int_{\Lambda_{\{0,1\}}} \hat{f}_{n_j} \mathrm{d}\xi_{0,1} \right) + \mathbb{E} \int_\Lambda f_{n_j} \mathrm{d}\xi + \int_\Lambda (1 - f_{n_j}) \mathrm{d}\xi_0 - \int_\Lambda f_{n_j} \mathrm{d}\xi_1. \end{aligned}$$

Taking the limit $j \to \infty$, the part between parentheses at the right converges to 0 $\mathbb{P}^\theta$-a.s., and the term $\int_\Lambda (1 - f_{n_j}) \mathrm{d}\xi_0$ vanishes $\mathbb{P}^\theta$-a.s. for sufficiently large j, since $1 - f_{n_j}$ vanishes on $\operatorname{supp} \xi_0 \subset \Lambda_{n_j}$. The other terms at the right are deterministic or depend only on ξ_1. We can conclude that

$$\begin{aligned} 1 &= \mathbb{P}^\theta \left(\xi_0(\Lambda) = \lim_{j \to \infty} \left(\mathbb{E} \int_\Lambda f_{n_j} \mathrm{d}\xi - \int_\Lambda f_{n_j} \mathrm{d}\xi_1 \right) \right) \\ &= \int_{\mathcal{N}(\Lambda)} \mathbb{P}_{|\mathbf{v}}^\theta \left(\xi(\Lambda) = \lim_{j \to \infty} \left(\mathbb{E} \int_\Lambda f_{n_j} \mathrm{d}\xi - \int_\Lambda f_{n_j} \mathrm{d}\delta_\mathbf{v} \right) \right) \mathrm{d}\mathbb{P}_1^\theta(\delta_\mathbf{v}), \end{aligned}$$

by Proposition 3.2, and it follows that

$$\mathbb{P}_{|\mathbf{v}}^\theta \left(\xi(\Lambda) = \lim_{j \to \infty} \left(\mathbb{E} \int_\Lambda f_{n_j} \mathrm{d}\xi - \int_\Lambda f_{n_j} \mathrm{d}\delta_\mathbf{v} \right) =: \ell_\mathbf{v} \right) = 1 \quad \text{for } \mathbb{P}_1^\theta\text{-a.e. } \delta_\mathbf{v}.$$

By Proposition 3.2 (1), the map $\delta_\mathbf{v} \mapsto \mathbb{P}_{|\mathbf{v}}^\theta(\xi(\Lambda) = \ell)$ is $\mathcal{C}(\Lambda)$-measurable for any $\ell \in \mathbb{N} \cup \{0, \infty\}$, hence the preimage of 1 under this map is in $\mathcal{C}(\Lambda)$. But by definition, this is the same as the preimage of ℓ under the map $\delta_\mathbf{v} \mapsto \ell_\mathbf{v}$, hence $\delta_\mathbf{v} \mapsto \ell_\mathbf{v}$ is $\mathcal{C}(\Lambda)$-measurable. We can conclude that $\mathbb{P}$ is marking rigid, and part (1) of the theorem is proved.

For part (2), it follows from [60, Theorem 4] that $\mathbb{P}^\theta(\xi_0(\Lambda) < \infty) = 0$ if $\operatorname{Tr} \mathrm{M}_{1-\theta}\mathrm{K} = \infty$, while $\mathbb{P}^\theta(\xi_0(\Lambda) < \infty) = 1$ if $\operatorname{Tr} \mathrm{M}_{1-\theta}\mathrm{K} < \infty$. We then know that $\mathbb{P}$ is marking rigid from part (1), and it follows that, for any measurable $\theta : \Lambda \to [0, 1]$ such that $\operatorname{Tr} \mathrm{M}_{1-\theta}\mathrm{K} < \infty$ and for $\mathbb{P}_1^\theta$-a.e. $\delta_\mathbf{v}$, there exists a finite number $\ell_\mathbf{v}$ such that $\mathbb{P}_{|\mathbf{v}}^\theta(\xi(\Lambda) = \ell_\mathbf{v}) = 1$. By Proposition 3.2 (2), $\mathbb{P}_{|\mathbf{v}}^\theta$ is a DPP induced by a Hermitian locally trace class operator, hence again by [60, Theorem 4], it is induced by an orthogonal projection $\mathrm{K}_{|\mathbf{v}}^\theta$ of rank $\ell_\mathbf{v}$. □

Remark 3.6. *The above proof gives us more information about* $\xi_0(\Lambda)$. *Depending on* θ, *this number is either a.s. infinite, or a.s. equal to*

$$\ell_{\mathbf{v}} = \lim_{j\to\infty} \mathbb{E} \int_\Lambda f_{n_j} \mathrm{d}(\xi - \delta_{\mathbf{v}}).$$

It is surprising that this number does not depend explicitly on the marking function θ. *Of course, the configurations* $\delta_{\mathbf{v}}$ *for which it holds implicitly encode information about* θ.

4. OPEs on the real line or on the unit circle

4.1. **OPEs on the real line.** Let us consider the N-point OPE on the real line defined by (1.6). It is well-known that (1.6) is a DPP on $(\mathbb{R}, w(x)\mathrm{d}x)$, with kernel

$$K_N(x,y) = \sum_{j=0}^{N-1} p_j(x)p_j(y), \tag{4.1}$$

where p_j is the normalized orthogonal polynomial of degree j with positive leading coefficient on the real line with respect to the weight $w(x)$. From the orthogonality of the polynomials, it follows that the integral operator K_N with kernel K_N acting on $L^2(\mathbb{R}, w(x)\mathrm{d}x)$, defined by

$$\mathrm{K}_N f(y) = \int_{\mathbb{R}} K_N(x,y)f(y)w(y)\mathrm{d}y, \tag{4.2}$$

is the orthogonal projection onto the N-dimensional space

$$H_N := \{p : p \text{ is a polynomial of degree } \leq N-1\}.$$

Alternatively, by the Christoffel-Darboux formula, we can write the correlation kernel in 2-integrable form

$$K_N(x,y) = \gamma_N \frac{p_N(x)p_{N-1}(y) - p_N(y)p_{N-1}(x)}{x-y}, \tag{4.3}$$

with $\gamma_N = \frac{\kappa_{N-1}}{\kappa_N}$, where κ_n is the leading coefficient of p_n, or equivalently $\kappa_n^{-1} = \int_{\mathbb{R}} p_n(x)x^n w(x)\mathrm{d}x$. See e.g. [35] for more background and details about these ensembles.

4.2. **OPEs on the unit circle.** For general integrable weight functions w, (1.7) is a DPP on the unit circle $\{z = e^{it}\}$ with respect to $w(e^{it})\mathrm{d}t$, with correlation kernel

$$K_N(e^{it}, e^{is}) = \sum_{j=0}^{N-1} \varphi_j(e^{it})\overline{\varphi_j(e^{is})}, \tag{4.4}$$

where φ_j is the normalized orthogonal polynomial of degree j with positive leading coefficient on the unit circle with respect to the weight $w(e^{it})$. The associated integral operator K_N is the orthogonal projection onto the space

$$H_N := \{\varphi : \varphi \text{ is a polynomial of degree } \leq N-1\}$$

Alternatively, by the Christoffel-Darboux formula for orthogonal polynomials on the unit circle, we have the 2-integrable form

$$K_N(e^{it}, e^{is}) = \frac{e^{iN(t-s)}\varphi_N(e^{is})\overline{\varphi_N(e^{it})} - \varphi_N(e^{it})\overline{\varphi_N(e^{is})}}{1 - e^{i(t-s)}}.$$

For the uniform weight $w = 1$, we have $\varphi_j(z) = (2\pi)^{-\frac{1}{2}} z^j$ and thus, after conjugation of the operator K_N, the kernel can be taken to be

$$K_N(e^{it}, e^{is}) = \frac{1}{2\pi} \frac{\sin \frac{N(t-s)}{2}}{\sin \frac{t-s}{2}}.$$

In the scaling limit where $t-s=\frac{2\pi(u-v)}{N}$ and $N\to\infty$, $\frac{2\pi}{N}K_N(e^{it},e^{is})$ converges to the sine kernel

$$K^{\sin}(u,v)=\frac{\sin\pi(u-v)}{\pi(u-v)}, \tag{4.5}$$

uniformly for u,v in compact subsets of the real line. See e.g. [40, 57] for details.

4.3. **Conditional ensembles associated to OPEs.** From Proposition 3.1, it follows that the conditional ensemble $\mathbb{P}^{\theta}_{|\mathbf{v}}$ is the DPP, on $(\mathbb{R},(1-\theta(x))w(x)\mathrm{d}x)$ or on the unit circle with measure $(1-\theta(e^{it}))w(e^{it})\mathrm{d}t$, with kernel $\widetilde{K}_N$ of the orthogonal projection onto the space

$$H_{N,\mathbf{v}}:=\{p:p\text{ is a polynomial of degree}\leq N-1,\text{ and }p(v)=0\ \forall v\in\mathbf{v}\}.$$

Let us now define

$$w^{\theta}_{|\mathbf{v}}(x):=(1-\theta(x))w(x)\prod_{v\in\mathbf{v}}|x-v|^2$$

for the real line, and

$$w^{\theta}_{|\mathbf{v}}(e^{it}):=(1-\theta(e^{it}))w(e^{it})\prod_{v\in\mathbf{v}}|e^{it}-v|^2$$

for the unit circle. In the case of the real line, we then have $\widetilde{K}_N(x,y)=\prod_{v\in\mathbf{v}}(x-v)(y-v)K_n(x,y)$, with $n:=N-\#\mathbf{v}$ and with K_n the Christoffel-Darboux kernel (4.1) with N replaced by n and w by $w^{\theta}_{|\mathbf{v}}$. It follows that we can also see $\mathbb{P}^{\theta}_{|\mathbf{v}}$ as a DPP on $(\mathbb{R},w^{\theta}_{|\mathbf{v}}(x)\mathrm{d}x)$ with kernel K_n, which implies that it is the n-point OPE

$$\frac{1}{Z_n}\Delta(\mathbf{x})^2\prod_{j=1}^{n}w^{\theta}_{|\mathbf{v}}(x_j)\mathrm{d}x_j,\qquad \Delta(\mathbf{x})=\prod_{1\leq i<j\leq N}(x_j-x_i).$$

In the case of the unit circle, we obtain similarly that $\mathbb{P}^{\theta}_{|\mathbf{v}}$ is the n-point OPE

$$\frac{1}{Z_n}|\Delta(\mathbf{e^{it}})|^2\prod_{j=1}^{n}w^{\theta}_{|\mathbf{v}}(e^{it_j})\mathrm{d}t_j,\qquad \Delta(\mathbf{e^{it}})=\prod_{1\leq l<k\leq N}(e^{it_k}-e^{it_j}),\qquad t_j\in[0,2\pi).$$

Summarizing the above, we have proved the following result.

Proposition 4.1. *If $\mathbb{P}$ is the N-point OPE with weight w on the real line or the unit circle and $n:=N-\#\mathbf{v}>0$, then $\mathbb{P}^{\theta}_{|\mathbf{v}}$ is the n-point OPE with weight $w^{\theta}_{|\mathbf{v}}$ on the real line or the unit circle.*

4.4. **Unitary invariant ensembles and scaling limits.** The above form of the conditional ensembles has the remarkable consequence that any unitary invariant ensemble (1.6) with $V(x)\geq x^2$, can be constructed theoretically from the GUE: to see this, consider the conditional ensemble $\mathbb{P}^{\theta}_{|\emptyset}$ with $w(x)=e^{-Nx^2}$ the Gaussian weight in (1.6), and with $\theta(x)=1-e^{-N(V(x)-x^2)}\in[0,1]$. The latter has the joint probability distribution (1.6), but now with weight

$$w^{\theta}_{|\emptyset}(x):=(1-\theta(x))e^{-Nx^2}=e^{-NV(x)}.$$

This is of course not of any practical use for N large, because the event on which we condition then has very small probability unless $V(x)$ is close to x^2 (note that there exist algorithms to generate DPPs in general, see e.g. [46]). Nevertheless, it is striking that the GUE encodes any of the above unitary invariant ensembles via marking and conditioning.

This becomes even more surprising if we look at scaling limits of the correlation kernels. It is a classical fact that the GUE converges to the sine point process in the bulk scaling limit and to the Airy point process in the edge scaling. It is also understood that conditioning on an eigenvalue and scaling around this eigenvalue leads to the bulk Bessel point process, and that conditioning on a gap leads to the hard edge Bessel kernel. But unitary invariant ensembles admit for special choices of V also more complicated limit processes, associated to Painlevé equations and hierarchies [37]. In fact, it follows from the above that these Painlevé point processes are already encoded in the GUE eigenvalue distribution, if one combines a suitable conditioning with taking scaling limits.

4.5. **Marginal distribution of mark 0 points with known number of mark 1 points.** The construction of the conditional ensembles in Section 2 passed through the marked point process conditioned on having m mark 1 particles. In this case, we can make these ensembles more explicit. Indeed, from (1.6), we obtain that the marginal distribution of the mark 0 particles, conditioned on having exactly m mark 1 particles, is given by

$$\frac{1}{Z_{N,m}}|\Delta(\mathbf{u})|^2 \left(\int_{\mathbb{R}^m} |\Delta(\mathbf{v})|^2 \prod_{k=1}^m \left(\prod_{\ell=1}^n |u_\ell - v_k|^2\right) \theta(v_k)w(v_k)\mathrm{d}v_k\right) \prod_{j=1}^n (1-\theta(u_j))w(u_j)\mathrm{d}u_j,$$

where

$$Z_{N,m} = \int_{\mathbb{R}^n} |\Delta(\mathbf{u})|^2 \left(\int_{\mathbb{R}^m} |\Delta(\mathbf{v})|^2 \prod_{k=1}^m \left(\prod_{\ell=1}^n |u_\ell - v_k|^2\right) \theta(v_k)w(v_k)\mathrm{d}v_k\right) \prod_{j=1}^n (1-\theta(u_j))w(u_j)\mathrm{d}u_j.$$

By Heine's formula, the v-integral can be written as a Hankel determinant in the case of the real line, and as a Toeplitz determinant in the case of the unit circle. For the real line, defining the Hankel determinant as

$$H_m(f) = \det\left(f_{j+k}\right)_{j,k=0}^{m-1}, \qquad f_\ell = \int_{\mathbb{R}} x^\ell f(x)\mathrm{d}x,$$

we have

$$\frac{1}{Z'_{N,m}}|\Delta(\mathbf{u})|^2\ H_m\left(\theta.w.\prod_{j=1}^n(.-u_j)^2\right) \prod_{j=1}^n (1-\theta(u_j))w(u_j)\mathrm{d}u_j, \tag{4.6}$$

with $Z'_{N,m} = \int_{\mathbb{R}^n} |\Delta(\mathbf{u})|^2\ H_m\left(\theta.w.\prod_{j=1}^n(.-u_j)^2\right) \prod_{j=1}^n (1-\theta(u_j))w(u_j)\mathrm{d}u_j$. For the unit circle, defining the Toeplitz determinant as

$$T_m(g) = \det\left(g_{j-k}\right)_{j,k=0}^{m-1}, \qquad g_\ell = \frac{1}{2\pi}\int_0^{2\pi} e^{-i\ell t} g(e^{it})\mathrm{d}t,$$

we obtain

$$\frac{1}{Z'_{N,m}}|\Delta(\mathbf{e^{it}})|^2\ T_m\left(\theta.w.\prod_{j=1}^n|.-e^{it_j}|^2\right) \prod_{j=1}^n (1-\theta(e^{it_j}))w(e^{it_j})\mathrm{d}t_j, \tag{4.7}$$

with $Z'_{N,m} = \int_{(0,2\pi)^n} |\Delta(\mathbf{e^{it}})|^2\ T_m\left(\theta.w.\prod_{j=1}^n|.-e^{it_j}|^2\right) \prod_{j=1}^n (1-\theta(e^{it_j}))w(e^{it_j})\mathrm{d}t_j$. Similar formulas hold for the marginal distributions of the mark 1 points. Alternatively, by [2, Theorem 3.2], we can write both densities, with either $x_j = u_j$, $\mathrm{d}\mu(x_j) = \mathrm{d}x_j$ or $x_j = e^{it_j}$, $\mathrm{d}\mu(x_j) = \mathrm{d}t_j$, as

$$\frac{1}{Z''_{N,m}} \det\left(K_N^\theta(x_l,x_k)\right)_{l,k=1}^n \prod_{j=1}^n (1-\theta(x_j))w(x_j)\mathrm{d}\mu(x_j),$$

with a new normalization constant $Z''_{N,m}$ obtained in a similar manner, and where $K_N^\theta(x,y)$ is the kernel inducing the point processes (1.6)/(1.7) with $N = n+m$ particles and weight function θw. There is no reason to believe that these marginal distributions are in general DPPs, but they do have a special Hankel or Toeplitz determinant structure. In particular, probabilities can be expressed in terms of integrals of Toeplitz or Hankel determinants, which can in some cases be computed asymptotically as $m \to \infty$. Similar integrals of Toeplitz and Hankel determinants appear in the study of moments of moments in random matrix ensembles, connected to the study of extreme values of characteristic polynomials, see e.g. [4, 42].

5. Integrable DPPs

In this section, we will consider DPPs $\mathbb{P}$ on curves Λ in the complex plane with k-integrable kernels of the form (1.9). For simplicity, let us assume that Λ is a smooth closed curve on $\mathbb{C}\cup\{\infty\}$ without self-intersections, that the functions $f_1,\dots,f_k$ and $g_1,\dots,g_k$ are smooth functions on Λ, and that the reference measure is smooth with respect to $\mathrm{d}z$, i.e. $\mathrm{d}\mu(z)=h(z)\mathrm{d}z$ with h smooth (say C^∞, even if one can proceed with less regularity if needed) on Λ. Even if $\mathrm{d}z$ is not a positive measure on Λ, by mapping $K(x,y)$ to $K(x,y)h(y)$, we can then work with a kernel

$$K(x,y)=\frac{\mathbf{f}(x)^T\mathbf{g}(y)}{x-y}=\frac{\mathbf{g}(y)^T\mathbf{f}(x)}{x-y},$$

with column vectors $\mathbf{f}=(f_j)_{j=1}^k,\mathbf{g}=(g_j)_{j=1}^k$, with respect to the complex measure $\mathrm{d}z$, and with the associated integral operator K acting on $L^2(\Lambda,\mathrm{d}z)$.

5.1. General integrable kernels.

Let us first show that the Palm kernels $K_{\mathbf{v}}$ are also of k-integrable form.

Proposition 5.1. *For any $\mathbf{v}=\{v_1,\dots,v_m\}$ such that $\det K(\mathbf{v},\mathbf{v})>0$, the kernel of the reduced Palm measure $\mathbb{P}_{\mathbf{v}}$ is of k-integrable form $K_{\mathbf{v}}(x,y)=\frac{\mathbf{f}_{\mathbf{v}}(x)^T\mathbf{g}_{\mathbf{v}}(y)}{x-y}$, and the j-th entries of $\mathbf{f}_{\mathbf{v}}$ and $\mathbf{g}_{\mathbf{v}}$ are given by*

$$f_{\mathbf{v},j}(x)=\frac{1}{\det K(\mathbf{v},\mathbf{v})}\det\begin{pmatrix}f_j(x) & K(x,\mathbf{v})\\ f_j(\mathbf{v}) & K(\mathbf{v},\mathbf{v})\end{pmatrix},\quad g_{\mathbf{v},j}(y)=\frac{1}{\det K(\mathbf{v},\mathbf{v})}\det\begin{pmatrix}g_j(y) & g_j(\mathbf{v})\\ K(\mathbf{v},y) & K(\mathbf{v},\mathbf{v})\end{pmatrix},$$

where $K(\mathbf{v},\mathbf{v})$ represents the $m\times m$ matrix with (i,j)-entry $K(v_i,v_j)$, $K(x,\mathbf{v}),g_j(\mathbf{v})$ represent m-dimensional row vectors with entries $K(x,v_\ell),g_j(v_\ell)$, and $K(\mathbf{v},y),f_j(\mathbf{v})$ represent m-dimensional column vectors with entries $K(v_i,y),f_j(v_i)$.

Proof. Using the block determinant formula (2.19), we have that $f_{\mathbf{v},j}$ as defined in the statement of the proposition is given by

$$f_{\mathbf{v},j}(x)=f_j(x)-K(x,\mathbf{v})K(\mathbf{v},\mathbf{v})^{-1}f_j(\mathbf{v}).$$

Now let $\mathbf{v}=\mathbf{v}'\sqcup\{v\}$ and assume without loss of generality that $\det K(\mathbf{v}',\mathbf{v}')>0$, then using again the block determinant formula (2.19) one obtains

$$\begin{aligned}f_{\mathbf{v},j}(x)&=\frac{1}{\det K(\mathbf{v},\mathbf{v})}\det\begin{pmatrix}f_j(x) & K(x,v) & K(x,\mathbf{v}')\\ f_j(v) & K(v,v) & K(v,\mathbf{v}')\\ f_j(\mathbf{v}') & K(\mathbf{v}',v) & K(\mathbf{v}',\mathbf{v}')\end{pmatrix}\\
&=\frac{\det K(\mathbf{v}',\mathbf{v}')}{\det K(\mathbf{v},\mathbf{v})}\det\left(\begin{pmatrix}f_j(x) & K(x,v)\\ f_j(v) & K(v,v)\end{pmatrix}-\begin{pmatrix}K(x,\mathbf{v}')\\ K(v,\mathbf{v}')\end{pmatrix}K(\mathbf{v}',\mathbf{v}')^{-1}\begin{pmatrix}f_j(\mathbf{v}') & K(\mathbf{v}',v)\end{pmatrix}\right)\\
&=\frac{1}{K_{\mathbf{v}'}(v,v)}\det\begin{pmatrix}f_{\mathbf{v}',j}(x) & K_{\mathbf{v}'}(x,v)\\ f_{\mathbf{v}',j}(v) & K_{\mathbf{v}'}(v,v)\end{pmatrix},\end{aligned}$$

which implies that $\mathbf{f}_{\mathbf{v}}=(\mathbf{f}_{\mathbf{v}'})_v$, and similarly for $\mathbf{g}_{\mathbf{v}}$. Since also $\mathrm{K}_{\mathbf{v}}=(\mathrm{K}_{\mathbf{v}'})_v$, it now suffices to prove the result for $\mathbf{v}=\{v\}$. We then easily verify by (2.18) that

$$\begin{aligned}(x-y)K_v(x,y)&=\mathbf{f}(x)^T\mathbf{g}(y)-((x-v)+(v-y))\frac{K(x,v)K(v,y)}{K(v,v)}\\
&=\mathbf{f}(x)^T\mathbf{g}(y)-\mathbf{f}(x)^T\mathbf{g}(v)\frac{K(v,y)}{K(v,v)}-\frac{K(x,v)}{K(v,v)}\mathbf{f}(v)^T\mathbf{g}(y)\\
&=\left(\mathbf{f}(x)-\frac{K(x,v)}{K(v,v)}\mathbf{f}(v)\right)^T\left(\mathbf{g}(y)-\frac{K(v,y)}{K(v,v)}\mathbf{g}(v)\right),\end{aligned}$$

since $\mathbf{f}^T\mathbf{g}=0$. To complete the proof, it remains to check that $\mathbf{f}_v^T\mathbf{g}_v=0$, but this follows from a similar computation:

$$\mathbf{f}_v(x)^T\mathbf{g}_v(x)=-\mathbf{f}(x)^T\mathbf{g}(v)\frac{K(v,x)}{K(v,v)}-\frac{K(x,v)}{K(v,v)}\mathbf{f}(v)^T\mathbf{g}(x)=K(x,v)\frac{\mathbf{f}(v)^T\mathbf{g}(x)}{K(v,v)}-\frac{K(x,v)}{K(v,v)}\mathbf{f}(v)^T\mathbf{g}(x)=0.$$

□

Next, we will explain how the kernel of the point process $\mathbb{P}^\theta_{|\mathbf{v}}$ on (Λ, μ_0^θ), with kernel of the operator $\mathrm{K}(1-\mathrm{M}_\theta \mathrm{K})^{-1}$, can be characterized in terms of a RH problem. For this, we rely on the IIKS method developed in [47, 36].

In what follows, we assume that the entries of $\sqrt{\theta}\mathbf{g}$ and $\sqrt{\theta}\mathbf{f}$ are smooth, bounded and integrable functions on Λ which decay as $z \to \infty$, and that their derivatives are also bounded and integrable.

Let us consider the following RH problem.

RH problem for Y.

(a) $Y : \mathbb{C} \setminus \Lambda \to \mathbb{C}^{k\times k}$ is analytic; we mean by this that every entry of the matrix is analytic in $\mathbb{C} \setminus \Lambda$.

(b) Y has continuous boundary values $Y_\pm$ when Λ is approached from the left $(+)$ or right $(-)$, with respect to the orientation chosen for Λ, and they are related by

$$Y_+(z) = Y_-(z) J_Y(z), \qquad J_Y(z) = I_k - 2\pi i \theta(z) \mathbf{f}_{\mathbf{v}}(z) \mathbf{g}_{\mathbf{v}}(z)^T, \qquad z \in \Lambda, \tag{5.1}$$

where I_k is the $k \times k$ identity matrix.

(c) As $z \to \infty$, $Y(z) \to I_k$ uniformly.

The following is a consequence of results from, e.g., [36, Section 2], see also [47] and [6].

Proposition 5.2. *Suppose that* $\mathrm{M}_{\sqrt{\theta}+1_B}\mathrm{K}_{\mathbf{v}}\mathrm{M}_{\sqrt{\theta}+1_B}$ *is trace class on* $L^2(\Lambda, \mathrm{d}z)$ *for any bounded Borel set* B*, and that* $\det(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}_{\mathbf{v}}\mathrm{M}_{\sqrt{\theta}}) \neq 0$.

(1) The RH problem for Y is uniquely solvable, and the solution $Y(z)$ is invertible for any $z \in \mathbb{C}\setminus\Lambda$.

(2) The DPP $\mathbb{P}^\theta_{|\mathbf{v}}$ on $(\Lambda, (1-\theta)\mathrm{d}z)$ is characterized by the k-integrable kernel

$$K^\theta_{|\mathbf{v}}(x,y) = \frac{\mathbf{f}^\theta_{|\mathbf{v}}(x)^T \mathbf{g}^\theta_{|\mathbf{v}}(y)}{x-y} \tag{5.2}$$

where

$$\mathbf{f}^\theta_{|\mathbf{v}} = Y_\pm \mathbf{f}_{\mathbf{v}}, \qquad \mathbf{g}^\theta_{|\mathbf{v}} = Y_\pm^{-T}\mathbf{g}_{\mathbf{v}}, \tag{5.3}$$

and the above expressions are independent of the choice $\pm$ of boundary value, with $Y_\pm^{-T}$ denoting the inverse transpose of the matrix $Y_\pm$. Consequently

$$K^\theta_{|\mathbf{v}}(x,y) = \frac{1}{x-y}\mathbf{g}_{\mathbf{v}}(y)^T Y_\pm(y)^{-1} Y_\pm(x) \mathbf{f}_{\mathbf{v}}(x). \tag{5.4}$$

Proof. Observe first that, because of the assumptions and Proposition 5.1, $\theta(x)\mathbf{f}_{\mathbf{v}}(x)\mathbf{g}_{\mathbf{v}}(x)^T$ is also smooth, in $L^2(\Lambda, \mathrm{d}z)$, and decaying as $x \to \infty$, $x \in \Lambda$. We then set $\mathrm{A} = \mathrm{M}_{\sqrt{\theta}}\mathrm{K}_{\mathbf{v}}\mathrm{M}_{\sqrt{\theta}}$ and $V = I_k - 2\pi i\theta \mathbf{f}_{\mathbf{v}}\mathbf{g}_{\mathbf{v}}^T$, and apply [36, Lemma 2.12]: this result states that

$$(1-\mathrm{A})^{-1} - 1 = \mathrm{A}(1-\mathrm{A})^{-1} = \mathrm{M}_{\sqrt{\theta}}\mathrm{K}_{\mathbf{v}}\left(1-\mathrm{M}_\theta\mathrm{K}_{\mathbf{v}}\right)^{-1}\mathrm{M}_{\sqrt{\theta}}$$

has kernel $\frac{\mathbf{F}(x)^T\mathbf{G}(y)}{x-y}$ with

$$\mathbf{F} = Y_\pm\sqrt{\theta}\mathbf{f}_{\mathbf{v}}, \qquad \mathbf{G} = Y_\pm^{-T}\sqrt{\theta}\mathbf{g}_{\mathbf{v}}.$$

Hence, if θ has no zeros on Λ, the operator $\mathrm{K}_{\mathbf{v}}(1-\mathrm{M}_\theta\mathrm{K}_{\mathbf{v}})^{-1}$ has kernel $\frac{\mathbf{f}^\theta_{|\mathbf{v}}(x)^T\mathbf{g}^\theta_{|\mathbf{v}}(y)}{x-y}$ on $L^2(\Lambda, \mathrm{d}z)$ with

$$\mathbf{f}^\theta_{|\mathbf{v}} = \frac{1}{\sqrt{\theta}}\mathbf{F}, \qquad \mathbf{g}^\theta_{|\mathbf{v}} = \frac{1}{\sqrt{\theta}}\mathbf{G},$$

and the result follows from Theorem 2.7. If θ has zeros on Λ, the result does not directly follow, but it is readily seen that one can follow the proof of [36, Lemma 2.12] to prove the result also in this case. □

Remark 5.3. *The smoothness and decay of the entries of $\theta \mathbf{f}\mathbf{g}^T$ are assumptions we make to avoid technical complications, and which guarantee smooth boundary values $Y_\pm$ and uniform convergence at infinity. One can also proceed with less regularity, but then care must be taken about the sense of the boundary values of Y, which are not necessarily continuous, and about the convergence at infinity, which is not necessarily uniform, see e.g. [35, 36] for general theory of RH problems.*

The above results imply that given $K, \theta, \mathbf{v}$, we obtain $K^\theta_{|\mathbf{v}}$ by first computing $K_\mathbf{v}$, and then solving the RH problem for Y. Next, we explain how to bypass this procedure by characterizing $K^\theta_{|\mathbf{v}}$ directly in terms of a RH problem which depends in a simple explicit way on $K, \theta, \mathbf{v}$, without need to go through the transformation $K \mapsto K_\mathbf{v}$. For that purpose, let us construct a rational matrix-valued function R, which will allow us to connect $\mathbf{f}, \mathbf{g}$ with $\mathbf{f}_\mathbf{v}, \mathbf{g}_\mathbf{v}$.

For a singleton $\mathbf{v} = \{v\}$, we observe that

$$\mathbf{f}_v(x) = \mathbf{f}(x) - \mathbf{f}(v)\frac{K(x,v)}{K(v,v)} = \left(I_k - \frac{R_1}{x-v}\right)\mathbf{f}(x), \qquad R_1 = \frac{\mathbf{f}(v)\mathbf{g}(v)^T}{K(v,v)},$$

and similarly since $R_1^2 = 0$,

$$\mathbf{g}_v(x)^T = \mathbf{g}(x)^T \left(I_k + \frac{R_1}{x-v}\right) = \mathbf{g}(x)^T \left(I_k - \frac{R_1}{x-v}\right)^{-1}.$$

For the general case $\mathbf{v} = \{v_1, ..., v_m\}$, we inductively define the matrices R_j for $j = 1, ..., m$ by

$$R_j = \frac{\mathbf{f}_{v_1,\dots,v_{j-1}}(v_j)\mathbf{g}_{v_1,\dots,v_{j-1}}(v_j)^T}{K_{v_1,\dots,v_{j-1}}(v_j,v_j)},$$

satisfying $R_j^2 = 0$, and

$$R(z) = \left(I_k + \frac{R_1}{z-v_1}\right)\left(I_k + \frac{R_2}{z-v_2}\right)\cdots\left(I_k + \frac{R_m}{z-v_m}\right), \qquad z \in \mathbb{C} \setminus \{v_1, \dots, v_m\}. \tag{5.5}$$

Then R has determinant identically equal to 1, and

$$\mathbf{f}_\mathbf{v}(x) = R(x)^{-1}\mathbf{f}(x), \qquad \mathbf{g}_\mathbf{v}(x)^T = \mathbf{g}(x)^T R(x),$$

so that we can rewrite the jump matrix J_Y as

$$J_Y(x) = I_k - 2\pi i\theta(x)\mathbf{f}_\mathbf{v}(x)\mathbf{g}_\mathbf{v}(x)^T = R(x)^{-1}\left(I_k - 2\pi i\theta(x)\mathbf{f}(x)\mathbf{g}(x)^T\right)R(x). \tag{5.6}$$

Note that although the construction of R uses a certain order of $v_1, ..., v_m$, the result only depends on the unordered set $\mathbf{v}$, as it can be checked that

$$R(z) = I_k + \frac{\mathbf{f}(\mathbf{v})}{z-\mathbf{v}}K(\mathbf{v},\mathbf{v})^{-T}\mathbf{g}(\mathbf{v})^T = \left(I_k - \mathbf{f}(\mathbf{v})K(\mathbf{v},\mathbf{v})^{-T}\left(\frac{\mathbf{g}(\mathbf{v})}{z-\mathbf{v}}\right)^T\right)^{-1},$$

where $\frac{\mathbf{f}(\mathbf{v})}{z-\mathbf{v}}$ and $\mathbf{g}(\mathbf{v})$ are the $k \times m$ matrices whose i-th columns are $\frac{\mathbf{f}(v_i)}{z-v_i}$ and $\mathbf{g}(v_i)$, and similarly for the others. It turns out that R is a rational function which can also be characterized by a discrete RH problem (in fact, R^{-1} is the solution to the below RH problem for U with $\theta = 0$). Let us now define

$$U(z) = Y(z)R(z)^{-1}, \tag{5.7}$$

then U satisfies the following RH problem.

RH problem for U.

1. Each entry of $U : \mathbb{C} \setminus \Lambda \to \mathbb{C}^{k\times k}$ is analytic.
2. On $\Lambda \setminus \mathbf{v}$, U has continuous boundary values $U_\pm$ which satisfy the jump condition

$$U_+ = U_-(I_k - 2\pi i\theta\mathbf{f}\mathbf{g}^T),$$

while for each $v \in \mathbf{v}$, the residue $\rho_U(v) = \lim_{z\to v}(z-v)U(z)$ is well-defined and given by

$$\rho_U(v) = -\lim_{z\to v} U(z)\frac{\mathbf{f}(v)\mathbf{g}(v)^T}{K(v,v)}.$$

(3) As $z \to \infty$, $U(z) \to I_k$ uniformly.

Conditions (1) and (3) are immediately verified. To check the jump relation (2) for U, it suffices to use (5.7) and (5.6). For the residues of U, observe that it is sufficient to verify the condition for $v = v_1$ by the iterative construction of R. We then have by (5.5), (5.7), $R_1^2 = 0$, and the fact that $Y_+(v_1) = Y_-(v_1)$ (since $\mathbf{f}_{\mathbf{v}}(v_1) = \mathbf{g}_{\mathbf{v}}(v_1) = 0$) that

$$\begin{aligned}\lim_{z\to v_1}(z-v_1)U(z) &= -Y_\pm(v_1)\mathrm{Res}_{z=v_1} R^{-1} = -Y_\pm(v_1)\left(I_k - \frac{R_m}{v_1 - v_m}\right)\cdots\left(I_k - \frac{R_2}{v_1-v_2}\right)R_1 \\ &= -\lim_{z\to v_1} Y(z)\left(I_k - \frac{R_m}{z-v_m}\right)\cdots\left(I_k - \frac{R_1}{z-v_1}\right)R_1 = -\lim_{z\to v} U(z)\frac{\mathbf{f}(v)\mathbf{g}(v)^T}{K(v,v)}.\end{aligned}$$

In conclusion, we have the following result.

Proposition 5.4. *Suppose that* $\mathrm{M}_{\sqrt{\theta+1_B}}\mathrm{K}_{\mathbf{v}}\mathrm{M}_{\sqrt{\theta+1_B}}$ *is trace class on* $L^2(\Lambda, \mathrm{d}z)$ *for any bounded Borel set* B*, and that* $\det(1 - \mathrm{M}_{\sqrt{\theta}}\mathrm{K}_{\mathbf{v}}\mathrm{M}_{\sqrt{\theta}}) \neq 0$*. There exists a unique solution* U *to the RH problem for* U *which is furthermore invertible and satisfies*

$$\mathbf{f}^\theta_{|\mathbf{v}} = U_\pm \mathbf{f}, \qquad \mathbf{g}^\theta_{|\mathbf{v}} = U_\pm^{-T}\mathbf{g},$$

and

$$K^\theta_{|\mathbf{v}}(x,y) = \frac{1}{x-y}\mathbf{g}(y)^T U_\pm(y)^{-1}U_\pm(x)\mathbf{f}(x).$$

5.2. **Integrable kernels characterized by a RH problem.** The above RH characterization of $K^\theta_{|\emptyset}$ and $K^\theta_{|\mathbf{v}}$ is particularly useful in cases where the kernel K of the DPP $\mathbb{P}$ itself can also be characterized in terms of a RH problem. In such a case, the IIKS method allows to transform the RH problem to an *undressed* RH problem which is in a form amenable to asymptotic analysis and to derive differential equations [6, 7, 47].

Such a RH characterization is available for many important 2-integrable DPPs, like OPEs and the DPPs characterized by the Airy kernel, the sine kernel, the Bessel kernels, the confluent hypergeometric kernels, and kernels connected to Painlevé equations. Multiple orthogonal polynomial ensembles [50] and their scaling limits like Pearcey and tacnode kernels are examples of k-integrable kernels with $k > 2$, which can also be characterized through a $(k \times k)$ RH problem.

Let us illustrate this in the case $k = 2$.

Suppose that we can write

$$K(x,y) = \frac{\mathbf{f}(x)^T\mathbf{g}(y)}{x-y}, \qquad \mathbf{f}(x) = \frac{w(x)}{2\pi i}\Psi_\pm(x)\begin{pmatrix}1\\0\end{pmatrix}, \qquad \mathbf{g}(y)^T = \begin{pmatrix}0 & 1\end{pmatrix}\Psi_\pm(y)^{-1}, \tag{5.8}$$

for a smooth bounded function $w : \Lambda \to [0, +\infty)$, where Ψ satisfies a RH problem of the following form.

RH problem for Ψ*.*

(1) $\Psi : \mathbb{C} \setminus \Lambda \to \mathbb{C}^{2\times 2}$ is analytic.

(2) Ψ has continuous boundary values $\Psi_\pm$, and they are related by

$$\Psi_+(z) = \Psi_-(z)\begin{pmatrix}1 & w(z)\\ 0 & 1\end{pmatrix}, \qquad z \in \Lambda,$$

for some smooth bounded function $w : \Lambda \to \mathbb{C}$.

(3) For some $\Psi_\infty : \mathbb{C}\setminus\Lambda \to \mathbb{C}^{2\times 2}$ such that $\det \Psi_\infty(z) = 1$, we have

$$\Psi(z) = \left(I_2 + O(z^{-1})\right)\Psi_\infty(z),$$

uniformly as $z \to \infty$.

Then, it is straightforward to show that $\det \Psi(z) \equiv 1$, hence $\Psi(z)$ is an invertible matrix for every $z \in \mathbb{C} \setminus \Lambda$, and that there is only one solution to the RH problem for Ψ.

The third RH condition would be trivially valid with $\Psi_\infty = \Psi$, but as we illustrate in examples below, one usually prefers to specify a simpler explicit function Ψ_∞ to describe the asymptotic behavior of Ψ, in order to facilitate further analysis of the RH problem. Observe that the RH conditions imply that the first column of Ψ and the second row of Ψ^{-1} extend to entire functions in the complex plane, and hence that $\mathbf{f}/w$ and $\mathbf{g}$ extend to entire functions.

Let us define

$$\Psi^\theta_{|\mathbf{v}} = U\Psi. \tag{5.9}$$

Then, $\Psi^\theta_{|\mathbf{v}}$ is invertible and it is the unique solution to the following RH problem.

RH problem for $\Psi^\theta_{|\mathbf{v}}$.

(1) Each entry of $\Psi^\theta_{|\mathbf{v}} : \mathbb{C} \setminus \Lambda \to \mathbb{C}^{2\times 2}$ is analytic.

(2) $\Psi^\theta_{|\mathbf{v}}$ has continuous boundary values $\Psi^\theta_{|\mathbf{v}\pm}$ on $\Lambda \setminus \mathbf{v}$ and they are related by

$$\Psi^\theta_{|\mathbf{v}+}(z) = \Psi^\theta_{|\mathbf{v}-}(z) \begin{pmatrix} 1 & w(z)(1-\theta(z)) \\ 0 & 1 \end{pmatrix},$$

while as $z \to v \in \mathbf{v}$ we have

$$\Psi^\theta_{|\mathbf{v}}(z) = \mathcal{O}(1)(z-v)^{\sigma_3}, \qquad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(3) As $z \to \infty$, we have the uniform asymptotics

$$\Psi^\theta_{|\mathbf{v}}(z) = \left(I_2 + \mathcal{O}(z^{-1})\right) \Psi_\infty(z).$$

The first and the third conditions are immediate from the corresponding ones for U and Ψ. The jump relation is obtained from the jump relation for U and the one for Ψ along with (5.8):

$$U_+\Psi_+ = U_- \left(I_2 - 2\pi i \theta \frac{w}{2\pi i} \Psi_- \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \Psi_+^{-1} \right) \Psi_+ = U_- \Psi_- \left(\begin{pmatrix} 1 & w \\ 0 & 1 \end{pmatrix} - w\theta \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right).$$

The singular behavior near $\mathbf{v}$ is obtained in a similar manner: the second column of $\Psi^\theta_{|\mathbf{v}}(z-v)^{-\sigma_3}$ is obviously $\mathcal{O}(1)$ since the second column of U is $\mathcal{O}((z-v)^{-1})$ as $z \to v \in \mathbf{v}$, while for the first column we notice that for each $v \in \mathbf{v}$, by (5.8),

$$\begin{aligned} \lim_{z \notin \Lambda \to v} \Psi^\theta_{|\mathbf{v}}(z) \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \lim_{z \in \Lambda \to v} \Psi^\theta_{|\mathbf{v}\pm}(z) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = Y_\pm(v) \lim_{z \in \Lambda \to v} R(z)^{-1} \Psi_\pm(z) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= Y_\pm(v) \frac{2\pi i}{w(v)} \lim_{z \in \Lambda \to v} R(z)^{-1} \frac{w(z)}{2\pi i} \Psi_\pm(z) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = Y_\pm(v) \frac{2\pi i}{w(v)} \mathbf{f}_{\mathbf{v}}(v) = 0. \end{aligned}$$

Moreover, we have by (5.4) that the kernel of the conditional ensemble is given by

$$\begin{aligned} K^\theta_{|\mathbf{v}}(x,y) &= \frac{1}{x-y} \begin{pmatrix} 0 & 1 \end{pmatrix} \Psi_\pm(y)^{-1} U_\pm(y)^{-1} U_\pm(x) \Psi_\pm(x) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{w(x)}{2\pi i} \\ &= \frac{w(x)}{2\pi i (x-y)} \left(\Psi^\theta_{|\mathbf{v}}(y)^{-1} \Psi^\theta_{|\mathbf{v}}(x) \right)_{21}. \end{aligned} \tag{5.10}$$

Let us illustrate the above procedure in some examples.

Example 5.1. *Let p_k be the normalized degree k orthogonal polynomial with respect to a weight function w on $\Lambda = \mathbb{R}$, with leading coefficient $\kappa_k > 0$. Write*

$$\Psi(z) := \begin{pmatrix} \frac{1}{\kappa_N} p_N(z) & \frac{1}{2\pi i \kappa_N} \int_{-\infty}^{+\infty} \frac{p_N(s) w(s) ds}{s-z} \\ -2\pi i \kappa_{N-1} p_{N-1}(z) & -\kappa_{N-1} \int_{-\infty}^{+\infty} \frac{p_{N-1}(s) w(s) ds}{s-z} \end{pmatrix}. \tag{5.11}$$

This is the solution of the Fokas-Its-Kitaev RH problem [39]*, which is the above RH problem for* Ψ *with*

$$\Lambda=\mathbb{R},\qquad \Psi_\infty(z)=z^{N\sigma_3},\qquad \sigma_3:=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

With $\mathbf{f},\mathbf{g}$ *as in* (5.8)*, the kernel* $K_N(x,y)$ *is then the Christoffel-Darboux kernel (note the factor* $w(x)$ *which was not present in* (4.3)*; this is due to the different reference measures* $\mathrm{d}x$ *here and* $w(x)\mathrm{d}x$ *in* (4.3)*)*

$$K_N(x,y)=\frac{\kappa_{N-1}w(x)}{\kappa_N}\frac{p_N(x)p_{N-1}(y)-p_N(y)p_{N-1}(x)}{x-y}.$$

The RH problem for $\Psi^\theta_{|\emptyset}$ *is then the Fokas-Its-Kitaev RH problem, but with a deformed weight function* $(1-\theta)w$*; for non-empty* $\mathbf{v}$*, the function* $\Psi^\theta_{|\mathbf{v}}(z)\prod_{v\in\mathbf{v}}(z-v)^{-\sigma_3}$ *then satisfies the Fokas-Its-Kitaev RH problem with weight function* $(1-\theta(z))w(z)\prod_{v\in\mathbf{v}}(z-v)^2$ *and with* N *replaced by* N *minus the cardinality of* $\mathbf{v}$*, which is in perfect agreement with Proposition 4.1. This RH problem has been an object of intensive study in the past decades and large* N *asymptotics for its solution have been obtained for a large class of weight functions, see e.g.* [35, 50, 51]*.*

Example 5.2. *Write*

$$\Psi(z)=\begin{cases}\begin{pmatrix}e^{\pi iz}&e^{\pi iz}\\-e^{-\pi iz}&0\end{pmatrix}, & \text{for } \Im z>0,\\ \begin{pmatrix}e^{\pi iz}&0\\-e^{-\pi iz}&e^{-\pi iz}\end{pmatrix}, & \text{for } \Im z<0.\end{cases}$$

This matrix satisfies the RH problem for Ψ *with*

$$\Lambda=\mathbb{R},\qquad w(x)=1,\qquad \Psi_\infty=\Psi.$$

With $\mathbf{f},\mathbf{g}$ *as in* (5.8)*, the kernel* $K(x,y)$ *is then the sine kernel* (4.5)*. The associated RH problem for* $\Psi^\theta_{|\emptyset}$ *for* $\theta=1_B$ *the indicator function of a union of intervals was the RH problem studied originally in* [36]*, and was also analyzed succesfully in* [17] *for* $\theta=\gamma 1_B$ *with* $\gamma\in(0,1)$*.*

Example 5.3. *Write*

$$\Psi(z):=\sqrt{2\pi}e^{\frac{\pi i}{6}}\times\begin{cases}\begin{pmatrix}\mathrm{Ai}(z)&\mathrm{Ai}(\omega^2 z)\\-i\mathrm{Ai}'(z)&-i\omega^2\mathrm{Ai}'(\omega^2 z)\end{pmatrix}e^{-\frac{\pi i}{6}\sigma_3}, & \text{for } \Im z>0,\\ \begin{pmatrix}\mathrm{Ai}(z)&-\omega^2\mathrm{Ai}(\omega z)\\-i\mathrm{Ai}'(z)&i\mathrm{Ai}'(\omega z)\end{pmatrix}e^{-\frac{\pi i}{6}\sigma_3}, & \text{for } \Im z<0,\end{cases}\tag{5.12}$$

with $\omega=e^{\frac{2\pi i}{3}}$ *and* Ai *the Airy function. Using the relation* $\mathrm{Ai}(z)+\omega\mathrm{Ai}(\omega z)+\omega^2\mathrm{Ai}(\omega^2 z)=0$*, one verifies, using the asymptotic behavior of the Airy function, that this matrix satisfies the RH problem for* Ψ *with*

$$\Lambda=\mathbb{R},\quad w(x)=1,\quad \Psi_\infty(z)=\begin{cases}\frac{1}{\sqrt{2}}z^{-\frac{\sigma_3}{4}}\begin{pmatrix}1&i\\i&1\end{pmatrix}e^{-\frac{2}{3}z^{3/2}\sigma_3} & \text{for } |\arg z|<\pi-\delta,\\ \frac{1}{\sqrt{2}}z^{-\frac{\sigma_3}{4}}\begin{pmatrix}1&i\\i&1\end{pmatrix}e^{-\frac{2}{3}z^{3/2}\sigma_3}\begin{pmatrix}1&0\\\pm1&1\end{pmatrix} & \text{for } |\arg z|<\pi-\delta,\ \pm\Im z>0,\end{cases}$$

for any sufficiently small $\delta>0$*, with principal branches of the root functions. With* $\mathbf{f},\mathbf{g}$ *as in* (5.8)*, the kernel* $K(x,y)$ *is then the Airy kernel*

$$K(x,y)=\frac{\mathrm{Ai}(x)\mathrm{Ai}'(y)-\mathrm{Ai}(y)\mathrm{Ai}'(x)}{x-y}.$$

The RH problem for $\Psi^\theta_{|\emptyset}$*, or an equivalent RH problem obtained after opening of the lenses, was then studied in* [32, 26, 27] *for a rather large class of functions* θ*, in order to derive differential equations and asymptotics for Airy kernel Fredholm determinants of the form* $\det(1-\mathrm{M}_{\sqrt{\theta}}\mathrm{K}\mathrm{M}_{\sqrt{\theta}})$*. In particular, these determinants are important in the study of the narrow wedge solution of the Kardar-Parisi-Zhang* *equation and in the study of finite temperature free fermions, and they have a remarkably rich integrable structure: they are connected to the Korteweg-de Vries equation and to an integro-differential version of the second Painlevé equation. The asymptotics resulting from this RH analysis allow also to derive asymptotics for the conditional kernels* $K^\theta_{|\emptyset}$*. Moreover, the density of the pushed Coulomb gas from* [33, 49] *can be interpreted as an approximation of the one-point function* $K^\theta_{|\emptyset}(x,x)$*.*

The conclusion of this section is two-fold. First, we just showed that the IIKS RH problem allows one to characterize the conditional kernels $K^\theta_{|\emptyset}$ and $K^\theta_{|\mathbf{v}}$ in terms of a RH problem, which can potentially be analyzed asymptotically. Secondly, the conditional ensembles $\mathbb{P}^\theta_{|\emptyset}$ enable us to give a natural probabilistic interpretation to the IIKS method, as we explain next.

The starting point of the IIKS method to study Fredholm determinants of the form $\det(1-\mathrm{M}_\theta\mathrm{K})$, is the Jacobi identity: if $\theta(x)=\theta_t(x)$ depends smoothly on a deformation parameter t, we have

$$\partial_t \log\det(1-\mathrm{M}_{\sqrt{\theta_t}}\mathrm{K}\mathrm{M}_{\sqrt{\theta_t}}) = -\mathrm{Tr}\left[\partial_t\mathrm{M}_{\theta_t}\mathrm{K}(1-\mathrm{M}_{\theta_t}\mathrm{K})^{-1}\right] = -\int_\Lambda \partial_t\theta_t(x)K^{\theta_t}_{|\emptyset}(x,x)\mathrm{d}\mu(x).$$

In analytic terms, this implies that one can compute the Fredholm determinant $\det(1-\mathrm{M}_{\theta_t}\mathrm{K})$, or at least its logarithmic derivative, provided that one has sufficiently accurate knowledge of the conditional kernel $K^{\theta_t}_{|\emptyset}(x,x)$.

In probabilistic terms, if $1-\theta_t$ does not vanish, this identity reads

$$\partial_t \log\mathbb{E}\prod_{x\in\operatorname{supp}\xi}(1-\theta_t(x)) = \mathbb{E}^{\theta_t}_{|\emptyset}\int_\Lambda \partial_t\log(1-\theta_t(x))\mathrm{d}\xi(x). \tag{5.13}$$

The logarithmic derivative of an average multiplicative statistic in $\mathbb{P}$ is thus equal to an average linear statistic in the conditional ensemble $\mathbb{P}^\theta_{|\emptyset}$. Moreover, if the function $t\mapsto\theta_t(x)$ is a smooth probability distribution function, then the function

$$h^\theta_x(t) = -\partial_t\log(1-\theta_t(x)) = \frac{\partial_t\theta_t(x)}{1-\theta_t(x)} = \mathrm{Prob}\left(t_x = t \mid t_x\geq t\right), \tag{5.14}$$

has the natural interpretation of a hazard rate likelihood of the random variable t_x with distribution $t\mapsto\theta_t(x)$. We can interpret t_x for instance as the detection time of point x, and then $h^\theta_x(t)$ is the likelihood to detect the particle at position $x\in\operatorname{supp}\xi$ at time t, given that it was not detected before time t.

Acknowledgements. The authors were supported by the Fonds de la Recherche Scientifique-FNRS under EOS project O013018F. They are grateful to Alexander Bufetov for instructive discussions about rigidity, and to Guilherme Silva for useful discussions.

References

[1] J. Baik, P. Deift, and T. Suidan, Combinatorics and random matrix theory, Graduate Studies in Mathematics 172, American Mathematical Society, Providence, RI, 2016.
[2] J. Baik, P. Deift, and E. Strahov, Products and ratios of characteristic polynomials of random Hermitian matrices, Integrability, topological solitons and beyond, J. Math. Phys. 44 (2003), no. 8, 3657–3670.
[3] J. Baik, T. Kriecherbauer, K. McLaughlin, and P. Miller, Discrete orthogonal polynomials. Asymptotics and applications, Annals of Mathematics Studies 164, Princeton University Press, Princeton, NJ, 2007.
[4] E. Bailey and J. Keating, On the moments of the moments of the characteristic polynomials of random unitary matrices, Comm. Math. Phys. 371 (2019), no. 2, 689–726.
[5] T. Berggren and M. Duits, Mesoscopic fluctuations for the thinned circular unitary ensemble, Math. Phys. Anal. Geom. 20 (2017), no. 3.
[6] M. Bertola and M. Cafasso, The transition between the gap probabilities from the Pearcey to the Airy process - a Riemann-Hilbert approach, Int. Math. Res. Not. IMRN 2012, no. 7, 1519–1568.
[7] M. Bertola and M. Cafasso, Darboux transformations and random point processes, Int. Math. Res. Not. IMRN 2015, no. 15, 6211–6266.

[8] O. Bohigas, J.X. de Carvalho, and M.P. Pato, Deformations of the Tracy-Widom distribution, Phys. Rev. E (3) 79 (2009), no. 3.
[9] O. Bohigas and M.P. Pato, Randomly incomplete spectra and intermediate statistics, Phys. Rev. E (3) 74 (2006), no. 3.
[10] A.M. Borodin, Determinantal point processes, in The Oxford Handbook of Random Matrix Theory, Oxford University Press, 2011.
[11] A. Borodin and E. Rains, Eynard-Mehta theorem, Schur measures, and their Pfaffian analogs, J. Stat. Phys. 121 (2005), 291–317.
[12] A. Borodin and A. Soshnikov, Janossy densities. I. Determinantal ensembles, J. Statist. Phys. 113 (2003), no. 3-4, 595–610.
[13] A. Borodin and E. Strahov, Averages of characteristic polynomials in random matrix theory, Comm. Pure Appl. Math. 59 (2006), no. 2, 161–253.
[14] A. Borodin and P. Deift, Fredholm determinants, Jimbo-Miwa-Ueno τ-functions, and representation theory, Comm. Pure Appl. Math. 55 (2002), no. 9, 1160–1230.
[15] T. Bothner, Transition asymptotics for the Painlevé II transcendent, Duke Math. J. 166 (2017), no. 2, 205–324.
[16] T. Bothner and R. Buckingham, Large deformations of the Tracy-Widom distribution I: Non-oscillatory asymptotics, Comm. Math. Phys. 359 (2018), no. 1, 223–263.
[17] T. Bothner, P. Deift, A. Its, and I. Krasovsky, On the asymptotic behavior of a log gas in the bulk scaling limit in the presence of a varying external potential I, Comm. Math. Phys. 337 (2015), no. 3, 1397–1463.
[18] T. Bothner, A. Its, and A. Prokhorov, On the analysis of incomplete spectra in random matrix theory through an extension of the Jimbo-Miwa-Ueno differential, Adv. Math. 345 (2019), 483–551.
[19] A.I. Bufetov, Infinite determinantal measures, Electronic Research Announcements in the Mathematical Sciences, 20 (2013), 8–20.
[20] A.I. Bufetov, Multiplicative functionals of determinantal processes, Uspekhi Mat. Nauk 67 (2012), no. 1 (403), 177–178; translation in Russian Math. Surveys 67 (2012), no. 1, 181–182.
[21] A.I. Bufetov, Rigidity of determinantal point processes with the Airy, the Bessel and the gamma kernel, Bull. Math. Sci. 6 (2016), no. 1, 163–172.
[22] A.I. Bufetov, Quasi-symmetries of determinantal point processes, Ann. Probab. 46 (2018), no. 2, 956–1003.
[23] A.I. Bufetov, Conditional measures of determinantal point processes, (Russian) Funktsional. Anal. i Prilozhen. 54 (2020), no. 1, 11–28; translation in Funct. Anal. Appl. 54 (2020), no. 1, 7–20.
[24] A.I. Bufetov, The sine process has excess one, arXiv:1912.13454.
[25] A.I. Bufetov, Y. Qiu, and A. Shamov, Kernels of conditional determinantal measures and the Lyons-Peres completeness conjecture, J. Eur. Math. Soc. (JEMS) 23 (2021), no. 5, 1477–1519.
[26] M. Cafasso and T. Claeys, A Riemann-Hilbert approach to the lower tail of the KPZ equation, to appear in Comm. Pure Appl. Math., https://doi.org/10.1002/cpa.21978.
[27] M. Cafasso, T. Claeys, and G. Ruzza, Airy kernel determinant solutions of the KdV equation and integro-differential Painlevé equations, Comm. Math. Phys. 386 (2021), 1107–1153.
[28] C. Charlier, Asymptotics of Hankel determinants with a one-cut regular potential and Fisher–Hartwig singularities, Int. Math. Res. Notices 2019 (24), 7515–7576.
[29] C. Charlier, Exponential Moments and Piecewise Thinning for the Bessel Point Process, Int. Math. Res. Notices 2021, no. 21, 16009–16071, doi: 10.1093/imrn/rnaa054.
[30] C. Charlier, Large gap asymptotics for the generating function of the sine point process, Proc. Lond. Math. Soc. 2021 123(2), 103–152.
[31] C. Charlier and T. Claeys, Thinning and conditioning of the Circular Unitary Ensemble, Random Matrices: Theory and Applications 6 (2017), no. 2, 1750007.
[32] C. Charlier and T. Claeys, Large gap asymptotics for Airy kernel determinants with discontinuities, Comm. Math. Phys. (2019), https://doi.org/10.1007/s00220-019-03538-w.
[33] I. Corwin, P. Ghosal, A. Krajenbrink, P. Le Doussal, and L.C. Tsai, Coulomb-gas electrostatics controls large fluctuations of the Kardar-Parisi-Zhang equation, Phys. Rev. Lett. 121, 060201 (2018).
[34] D.J. Daley and D. Vere-Jones, An introduction to the theory of point processes. Vol. II. General theory and structure. Second edition. Probability and its Applications (New York), Springer, New York, 2008.
[35] P. Deift, Orthogonal polynomials and random matrices: a Riemann-Hilbert approach. Courant Lecture Notes in Mathematics, 3. New York University, Courant Institute of Mathematical Sciences, New York, American Mathematical Society, Providence, RI, 1999.
[36] P. Deift, A. Its, A., and X. Zhou, A Riemann-Hilbert approach to asymptotic problems in the theory of random matrix models, and also in the theory of integrable statistical mechanics. Ann. Math. 146 (1997), 149–235.
[37] M. Duits, Painlevé kernels in Hermitian matrix models, Constr. Approx. 39 (2014), no. 1, 173–196.
[38] F. Dyson, Random matrices, neutron capture levels, quasicrystals and zeta-function, lecture notes, MSRI Program Recent Progress In Random Matrix Theory And Its Applications (2002), available on https://www.msri.org/workshops/220/schedules/1385 (September 8, 2021).
[39] A.S. Fokas, A.R. Its, and A.V. Kitaev, The isomonodromy approach to matrix models in 2D quantum gravity, Comm. Math. Phys. 147 (1992), 395–430.

[40] P.J. Forrester, Log-gases and random matrices, London Mathematical Society Monographs Series, 34. Princeton University Press, Princeton, NJ, 2010.
[41] P.J. Forrester and A. Mays, Finite-size corrections in random matrix theory and Odlyzko's dataset for the Riemann zeros, Proc. A. 471 (2015), no. 2182, 20150436, 21 pp.
[42] Y. Fyodorov and J. Keating, Freezing transitions and extreme values: random matrix theory, and disordered landscapes, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 372 (2014), no. 2007, 20120503, 32 pp.
[43] S. Ghosh, Palm measures and rigidity phenomena in point processes, Electron. Commun. Probab. 21 (2016), Paper No. 85, 14 pp.
[44] S. Ghosh and M. Krishnapur, Rigidity hierarchy in random point fields: random polynomials and determinantal processes, to appear in Commun. Math. Phys., arXiv:1510.08814.
[45] S. Ghosh and Y. Peres, Rigidity and tolerance in point processes: Gaussian zeros and Ginibre eigenvalues, Duke Math. J. 166 (2017), no. 10, 1789–1858.
[46] J.B. Hough, M. Krishnapur, Y. Peres, and B. Virag, Determinantal processes and independence. Probab. Surv. 3 (2006), 206–229.
[47] A.R. Its, A.G. Izergin, V.E. Korepin, and N.A. Slavnov, The quantum correlation function as the τ function of classical differential equations. Important developments in soliton theory, 407–417, Springer Ser. Nonlinear Dynam., Springer, Berlin, 1993.
[48] K. Johansson, Random matrices and determinantal processes, Mathematical statistical physics, 1–55, Elsevier B. V., Amsterdam, 2006.
[49] A. Krajenbrink and P. Le Doussal, Linear statistics and pushed Coulomb gas at the edge of β-random matrices: Four paths to large deviations, EPL (Europhysics Letters) 125, 20009 (2019).
[50] A. Kuijlaars, Multiple orthogonal polynomials in random matrix theory, in: Proceedings of the International Congress of Mathematicians, Volume III (R. Bhatia, ed.) Hyderabad, India, 2010, pp. 1417–1432.
[51] A. Kuijlaars, Universality, in: The Oxford Handbook of Random Matrix Theory, Oxford University Press, 2011.
[52] A. Kuijlaars and E. Miña-Díaz, Universality for conditional measures of the sine point process, J. Approx. Theory 243 (2019), 1–24.
[53] E. Lavancier, J. Moller, and E. Rubak, Determinantal point process models and statistical inference, J. R. Stat. Soc. Ser. B. Stat. Methodol. 77 (2015), no. 4, 853–877.
[54] A. Lenard, States of classical statistical mechanical systems of infinitely many particles. I, Arch. Rational Mech. Anal. 59 (1975), no. 3, 219–239.
[55] R. Lyons, Determinantal probability measures, Publ. Math. Inst. Hautes Etudes Sci. No. 98 (2003), 167–212.
[56] O. Macchi, The coincidence approach to stochastic point processes, Advances in Appl. Probability, 7 (1975), 83–122
[57] M.L. Mehta, Random Matrices, 2nd ed., Academic Press, Boston (1991).
[58] T. Shirai and Y. Takahashi, Random point fields associated with certain Fredholm determinants. I. Fermion, Poisson and boson point processes, J. Funct. Anal. 205 (2003), no. 2, 414–463.
[59] T. Shirai and Y. Takahashi, Random point fields associated with certain Fredholm determinants. II. Fermion shifts and their ergodic and Gibbs properties, Ann. Probab. 31 (2003), no. 3, 1533–1564.
[60] A. Soshnikov, Determinantal random point fields, (Russian) Uspekhi Mat. Nauk 55 (2000), no. 5(335), 107–160; translation in Russian Math. Surveys 55 (2000), no. 5, 923–975.