ALEA, Lat. Am. J. Probab. Math. Stat. **20**, 497–504 (2023)
DOI: 10.30757/ALEA.v20-17

On the mean projection theorem for determinantal point processes

Adrien Kassel and Thierry Lévy

CNRS – UMPA, ENS de Lyon
E-mail address: adrien.kassel@ens-lyon.fr
URL: http://perso.ens-lyon.fr/adrien.kassel/

LPSM, Sorbonne Université, Paris
E-mail address: thierry.levy@sorbonne-universite.fr
URL: https://www.lpsm.paris/users/levyt/index

Abstract. In this short note, we extend to the continuous case a mean projection theorem for discrete determinantal point processes associated with a finite range projection, thus strengthening a known result in random linear algebra due to Ermakov and Zolotukhin. We also give a new formula for the variance of the exterior power of the random projection.

1. Introduction

Kirchhoff's work on electrical networks Kirchhoff (1847) seems to be one of the earliest works in the literature where linear algebra and graph-theoretical combinatorial methods were put together. Later on, linear algebra problems, and classical determinantal methods for solving them, gave rise to various statistical approaches, notably linked to the so-called determinantal point processes (introduced by Macchi (1975), and named like this by Borodin, only around 2000 which saw a blossoming of results on those processes from various authors, see Soshnikov (2000); Shirai and Takahashi (2003); Lyons (2003); Johansson (2006); Borodin (2011)). These methods recently became an active field in randomized numerical linear algebra Dereziński and Mahoney (2021).

In his work, Kirchhoff solved a linear algebra system on an electrical network seen as a finite graph, by expressing the current induced by an external battery hooked on the network, as an average over spanning trees of a certain current associated to the tree. In modern terms, he expressed an orthogonal projection as the expectation of a certain random projection associated to a random spanning tree. Such a mean projection theorem appeared in several guises in the literature, and more or less independently, in works of Maurer (1976), Lyons (2003), Catanzaro et al. (2013), and probably others that we are unaware of.

In our work Kassel and Lévy (2022, Theorem 5.9), we extended the mean projection formula for determinantal point processes on finite sets, thus putting the statements of Kirchhoff (1847); Maurer (1976); Lyons (2003); Catanzaro et al. (2013) in a unified geometric framework, and strengthening

Received by the editors October 24th, 2022; accepted February 17th, 2023.
2020 *Mathematics Subject Classification.* 60G55, 15A75.
Key words and phrases. determinantal point processes, random projection, exterior algebra.

the result by proving a mean projection theorem for the exterior powers of the projections, that is, for minors of their matrices in a fixed basis.

Let us quickly recall our statement. Let $\mathbb{K}$ be $\mathbb{R}$ or $\mathbb{C}$, let E be a finite dimensional Euclidean space on $\mathbb{K}$ of dimension d, and let $(e_i)_{1\leq i\leq d}$ be an orthonormal basis of E. We let $S = \{1, \dots, d\}$ and consider H a subspace of E of dimension n. Let X be the determinantal point process on S associated to the matrix $K = \big(\langle e_i, \Pi^H e_j\rangle\big)_{1\leq i,j\leq d}$, where Π^H is the orthogonal projection on H. For each $X \subseteq S$, let $E_X = \bigoplus_{x\in X} \mathbb{K}e_x$ be the corresponding coordinate subspace of E.

Theorem 1.1. *Almost surely, the equality $E = H \oplus E_{\mathsf{X}}^{\perp}$ holds, and denoting by P_{X} the projection on H parallel to $E_{\mathsf{X}}^{\perp}$, we have*

$$\mathbb{E}\big[\textstyle\bigwedge \mathsf{P}_{\mathsf{X}}\big] = \textstyle\bigwedge \Pi^H .$$

In words, in a fixed basis of E, the expectation of any minor of the matrix of P_{X} is equal to the same minor of Π^H.

A short while ago, it came to our attention while reading the recent statistics paper Gautier et al. (2019) on Monte–Carlo integration methods, that such a mean projection formula had also appeared in Ermakov and Zolotukhin (1960) in the case of $S = \mathbb{R}$, in a different guise, although the relation to the above-cited works was not mentioned there.

One of the referees of this paper kindly pointed out to us that results in the spirit of Theorem 1.1 have also been obtained in the context of the resolution of singular linear systems of equations, for instance in Berg (1986); Ben-Tal and Teboulle (1990) and more recently in the context of active sampling for linear regression Dereziński and Warmuth (2018, Thms 5, 6 and 7), see also Avron and Boutsidis (2013); Mariet and Sra (2017); Dereziński et al. (2022). In Dereziński et al. (2020, Def. 4), the authors define the class of random matrices for which the expectation of any minor equals the same minor of the expectation, give basic properties, and provide a few examples. Theorem 1.1 and Kassel and Lévy (2022, Thm 5.9) give families of examples of such random matrices, namely the matrices P_{X}. A systematic study of this class of random matrices would certainly be interesting.

The goal of this short note is to extend Theorem 1.1 to the case of a determinantal point process associated to a finite rank orthogonal projection on any Polish space S, so that it applies for instance to any orthogonal polynomial ensemble, see Lyons (2014, Section 3.8). This extension is the content of Theorem 2.2. An extension of Theorem 1.1 to the case of a projection with infinite range (both in the case where S is discrete or continuous) would be interesting. An example of this situation is investigated in Bufetov and Qiu (2022), where the author study among other things the continuous analogue of P_{X} in the case of the Bergman kernel.

2. The mean projection theorem

Let S be a Polish space and λ a positive Radon measure on S. Let us consider the space $E = L^2(S, \lambda)$ and the space $\mathcal{C}(S)$ of continuous functions on S.[1] Let $H \subseteq E \cap \mathcal{C}(S)$ be a linear subspace of finite dimension n.

Let $\mathsf{Conf}_n(S)$ be the set of collections of n distinct points in S, and let μ be the determinantal probability measure on $\mathsf{Conf}_n(S)$ associated with the orthogonal projection on H. This means that if we choose an orthonormal basis $(\varphi_j)_{1\leq j\leq n}$ of H, then we have for any bounded continuous symmetric test function $T : S^n \to \mathbb{C}$ the equality

$$\int_{\mathsf{Conf}_n(S)} T(X)\ \mathrm{d}\mu(X) = \frac{1}{n!}\int_{S^n} T(x_1, \dots, x_n)\big|\det\big(\varphi_j(x_i)_{1\leq i,j\leq n}\big)\big|^2\ \mathrm{d}\lambda^{\otimes n}(x_1, \dots, x_n), \qquad (2.1)$$

[1] The space of continuous functions plays for us the role usually devoted to a reproducing kernel Hilbert space (RKHS), namely that of a space of functions that can be evaluated at points. However, we do not need this extra structure, because we do not need evaluation at a point to be a continuous linear form. Moreover, it seems that in many examples of interest, the RKHS is a subspace of continuous functions, so that our result applies.

in which the right-hand side does not depend on the choice of the orthonormal basis. We will denote by X a random subset of S distributed according to μ, and use the notation $\mathbb{E}[T(\mathsf{X})]$ for either of the two sides of the equality above.

It follows from (2.1) that μ-almost every X is a *uniqueness set* for H, in the sense that two elements of H that coincide on X are equal.[2] This fact can be used to define a random projection onto H, as follows. For every $X \in \mathsf{Conf}_n(S)$, let us define $\mathcal{C}(S;X) = \{f \in \mathcal{C}(S) : f_{|X} = 0\}$.

Lemma 2.1. *For μ-almost every $X \in \mathsf{Conf}_n(S)$, the decomposition $\mathcal{C}(S) = H \oplus \mathcal{C}(S;X)$ holds.*

Proof: Let f be an element of $\mathcal{C}(S)$. Let $(\varphi_j)_{1\leq j\leq n}$ be an orthonormal basis of H. For μ-almost every $X = \{x_1, \dots, x_n\}$ in $\mathsf{Conf}_n(S)$, we have $\det(\varphi_j(x_i)_{1\leq i,j\leq n}) \neq 0$, so that the system

$$\alpha_1\varphi_1(x_i) + \ldots + \alpha_n\varphi_n(x_i) = f(x_i)\ , \quad \forall i \in \{1, \dots, n\}$$

admits a unique solution. Then $\mathsf{P}_X f = \alpha_1\varphi_1 + \ldots + \alpha_n\varphi_n$ is the unique element of H which takes the same values as f on X. □

For the rest of this note, we will keep the notation P_X introduced in the previous proof for the projection on H parallel to $\mathcal{C}(S;X)$. Let us emphasize that the decomposition given by Lemma 2.1 depends on H and X, but is independent of the Euclidean structure of E. In particular, the projection P_X is independent of this Euclidean structure.

For example, if $S = \mathbb{R}$, λ is a measure with infinite support which admits moments of all orders, and $\varphi_1, \dots, \varphi_n$ are the first n orthogonal polynomials with respect to λ, then H is the space of polynomial functions of degree at most $n-1$ and $\mathsf{P}_X f$ is the *interpolating polynomial* of the restriction of f to X.

For all $g_1, \dots, g_m \in E \cap \mathcal{C}(S)$, let us define $g_1 \wedge \ldots \wedge g_m \in L^2(S^m, \frac{1}{m!}\lambda^{\otimes m}) \cap \mathcal{C}(S^m)$ by setting, for all $y_1, \dots, y_m \in S$,

$$(g_1 \wedge \ldots \wedge g_m)(y_1, \dots, y_m) = \det\big(g_j(y_i)_{1\leq i,j\leq m}\big). \tag{2.2}$$

We will use several times the Andreieff–Heine identity, which is a continuous analogue of the Cauchy–Binet identity, and can be phrased as follows: if $h_1, \dots, h_m$ belong to $E \cap \mathcal{C}(S)$, then

$$\langle g_1 \wedge \ldots \wedge g_m, h_1 \wedge \ldots \wedge h_m \rangle_{L^2(S^m, \frac{1}{m!}\lambda^{\otimes m})} = \det\big(\langle g_i, h_j\rangle\big)_{1\leq i,j\leq m}. \tag{2.3}$$

This equality justifies, for instance, the fact that the measure μ defined by (2.1) is a probability measure.

Let us write $H^0 = H$ and $H^1 = H^\perp$. The isomorphism of vector spaces $L^2(S^m, \frac{1}{m!}\lambda^{\otimes m}) \simeq L^2(S,\lambda)^{\otimes m}$ is $\sqrt{m!}$ times an isometry, and the orthogonal decomposition $L^2(S) = H^0 \oplus H^1$ gives rise to an orthogonal decomposition

$$L^2(S^m) \simeq L^2(S)^{\otimes m} = \bigoplus_{\varepsilon_1,\dots,\varepsilon_m \in\{0,1\}} H^{\varepsilon_1} \otimes \ldots \otimes H^{\varepsilon_m} = \bigoplus_{k=0}^{m} \Bigg[\bigoplus_{\substack{\varepsilon_1,\dots,\varepsilon_m\in\{0,1\} \\ \varepsilon_1+\ldots+\varepsilon_m=k}} H^{\varepsilon_1} \otimes \ldots \otimes H^{\varepsilon_m} \Bigg]. \tag{2.4}$$

Let us denote by Π_k the orthogonal projection of $L^2(S^m)$ on the k-th summand of the last expression. In order to describe this operator more concretely, recall that we denote by Π^H the orthogonal projection on H in E. For all real t, let us define the linear operator $\mathsf{D}_t = \Pi^H + t\Pi^{H^\perp}$ on E. Then

$$\mathsf{D}_t g_1 \wedge \ldots \wedge \mathsf{D}_t g_m = \sum_{k=0}^{m} t^k\, \Pi_k(g_1 \wedge \ldots \wedge g_m).$$

[2]The uniqueness property is true for all determinantal processes associated with an orthogonal projection of possibly infinite range, that is with infinitely many points ($n = \infty$), as proved in the discrete case by Lyons (2003), and recently by Bufetov et al. (2021) in the general case, following partial results by Ghosh (2015).

In words, $\Pi_k(g_1 \wedge \ldots \wedge g_m)$ is the sum of all the functions obtained from $g_1 \wedge \ldots \wedge g_m$ by replacing k of the g_i's by their projections on $H^\perp$, and the others by their projection on H.

Theorem 2.2. *For all $m \geq 1$, and all $f_1, \ldots, f_m \in E \cap \mathcal{C}(S)$, we have*

$$\mathbb{E}[\mathsf{P}_\mathsf{X} f_1 \wedge \ldots \wedge \mathsf{P}_\mathsf{X} f_m] = \Pi^H f_1 \wedge \ldots \wedge \Pi^H f_m\,, \tag{2.5}$$

$$\mathrm{Var}(\mathsf{P}_\mathsf{X} f_1 \wedge \ldots \wedge \mathsf{P}_\mathsf{X} f_m) = \sum_{k=1}^{m} \tbinom{n-m+k}{k} \|\Pi_k(f_1 \wedge \ldots \wedge f_m)\|^2. \tag{2.6}$$

The variance in the second assertion is that of a random element of $L^2(S^m, \frac{1}{m!}\lambda^{\otimes m})$, that is, to be explicit, and in view of the first assertion,

$$\mathrm{Var}(\mathsf{P}_\mathsf{X} f_1 \wedge \ldots \wedge \mathsf{P}_\mathsf{X} f_m) = \mathbb{E}\Big[\big\|\mathsf{P}_\mathsf{X} f_1 \wedge \ldots \wedge \mathsf{P}_\mathsf{X} f_m - \Pi^H f_1 \wedge \ldots \wedge \Pi^H f_m\big\|^2_{L^2(S^m, \frac{1}{m!}\lambda^{\otimes m})}\Big].$$

Further note that the quadratic identity (2.6) may be polarized to obtain information on covariances.

Given the remark made after Lemma 2.1, one can view Theorem 2.2 as providing a statistical estimator of part of the Euclidean structure of E given H and a realisation X.

When $m = 1$, this is the theorem of Ermakov and Zolotukhin (1960), rephrased by Gautier et al. (2019):

$$\mathbb{E}[\mathsf{P}_\mathsf{X} f] = \Pi^H f \quad\text{and}\quad \mathrm{Var}(\mathsf{P}_\mathsf{X} f) = n\|\Pi^{H^\perp} f\|^2.$$

In order to prove Theorem 2.2, we will use the following generalization of Cramer's formula, which surprisingly enough, we have not encountered in our undergraduate linear algebra class.

For all integers n and m, we denote by $[\![n]\!]$ the set $\{1, \ldots, n\}$ and by $\mathcal{P}_m([\![n]\!])$ the set of its subsets with m elements. Given a $p \times q$ matrix M and two subsets $I \subseteq [\![p]\!]$ and $J \subseteq [\![q]\!]$, we define

$$M^I_J = (M_{ij})_{i \in I, j \in J} \quad\text{and}\quad M^I = M^I_{[\![q]\!]}.$$

Proposition 2.3 (Cramer's identity for minors)**.** *Let $1 \leq m \leq n$ be two integers. Let M be an $n \times n$ invertible square matrix, and F an $n \times m$ rectangular matrix. Let A be the $n \times m$ rectangular matrix solving $MA = F$. Then for all $I \in \mathcal{P}_m([\![n]\!])$, the $m \times m$ submatrix A^I has determinant*

$$\det A^I = (\det M)^{-1} \det M_{[I \leftarrow F]}\,, \tag{2.7}$$

where $M_{[I \leftarrow F]}$ is the $n \times n$ square matrix obtained by replacing in M the columns indexed by I by the columns of the matrix F.

If $I = \{i_1 < \ldots < i_m\}$, then $(M_{[I \leftarrow F]})_{ij} = M_{ij}$ for $j \notin I$, and $(M_{[I \leftarrow F]})_{ij} = F_{ik}$ for $j = i_k$.

Proof: Let us write $A = M^{-1}F$ and use the Cauchy–Binet formula:

$$\det A^I = \sum_{J \in \mathcal{P}_m([\![n]\!])} \det(M^{-1})^I_J \det F^J\,.$$

Now, by Jacobi's complementary minor formula,

$$\det(M^{-1})^I_J = (-1)^{\sum_{i \in I} i + \sum_{j \in J} j}\,(\det M)^{-1} \det M^{J^c}_{I^c}\,.$$

Combining the two previous equations and checking signs, we now recognize the Laplace expansion of $\det M_{[I \leftarrow F]}$ with respect to all columns in I:

$$\det A^I = (\det M)^{-1} \sum_{J \in \mathcal{P}_m([\![n]\!])} (-1)^{\sum_{i \in I} i + \sum_{j \in J} j}\, \det M^{J^c}_{I^c} \det F^J = (\det M)^{-1} \det M_{[I \leftarrow F]}\,,$$

which concludes the proof. □

Proof of Theorem 2.2: Let $(\varphi_i)_{1 \leq i \leq n}$ be an orthonormal basis of H. Let $X = (x_1, \ldots, x_n) \in S^n$ be such that $\det(\varphi_j(x_i)_{1 \leq i,j \leq n}) \neq 0$. Let us introduce the following matrices:

- $M = (\varphi_j(x_i)) \quad 1 \leq i, j \leq n$,

- $F = (f_j(x_i)) \quad 1 \le i \le n, 1 \le j \le m$,
- $A = (\alpha_{ij}) \quad 1 \le i \le n, 1 \le j \le m$, the solution to $MA = F$,
- $G = (\langle \varphi_i, f_j \rangle) \quad 1 \le i \le n, 1 \le j \le m$.

For each $I = \{i_1 < \ldots < i_k\} \subseteq \{1, \ldots, n\}$, let us write $\varphi_I = \varphi_{i_1} \wedge \ldots \wedge \varphi_{i_k}$.

For each $1 \le i \le m$, we have

$$\mathsf{P}_X f_i = \sum_{k=1}^{n} A_{ki}\, \varphi_k \quad \text{and} \quad \Pi^H f_i = \sum_{k=1}^{n} G_{ki}\, \varphi_k,$$

so that

$$\mathsf{P}_X f_1 \wedge \ldots \wedge \mathsf{P}_X f_m = \sum_{I \in \mathcal{P}_m([\![n]\!])} \det A^I \varphi_I \quad \text{and} \quad \Pi^H f_1 \wedge \ldots \wedge \Pi^H f_m = \sum_{I \in \mathcal{P}_m([\![n]\!])} \det G^I \varphi_I. \tag{2.8}$$

In order to prove the first assertion of the theorem, namely (2.5), we are thus left to show that for all $I \in \mathcal{P}_m([\![n]\!])$, we have

$$\mathbb{E}\big[\det A^I\big] = \det G^I, \tag{2.9}$$

where we view A as a function of the subset $X \subseteq S$ and the expectation is with respect to μ.

By Proposition 2.3, we can write $\det A^I = (\det M)^{-1} \det M_{[I \leftarrow F]}$. Using the form (2.1) of the density of μ and the Andreieff–Heine identity (2.3), we find

$$\begin{aligned}
\mathbb{E}\big[\det A^I\big] &= \frac{1}{n!} \int_{S^n} \det A^I \, |\det M|^2 \, d\lambda^{\otimes n} \\
&= \frac{1}{n!} \int_{S^n} \det M_{[I \leftarrow F]} \, \overline{\det M} \, d\lambda^{\otimes n} \\
&= \det \big(\langle \varphi_a, \psi_{I,b} \rangle \big)_{1 \le a,b \le n} ,
\end{aligned}$$

where $(\psi_{I,1}, \ldots, \psi_{I,n})$ is the list $(\varphi_1, \ldots, \varphi_n)$ in which the terms labelled by elements of I have been replaced by $f_1, \ldots, f_m$. In symbols, $\psi_{I,b} = \varphi_b$ if $b \notin I$ and $\psi_{I,b} = f_k$ if $I = \{i_1 < \ldots < i_m\}$ and $b = i_k$.

The last determinant is, up to conjugation by a permutation matrix, that of a 2×2 block-triangular matrix. One of the diagonal blocks of this matrix is the identity, and the other is G^I. Thus, its determinant is equal to $\det G^I$, which proves (2.9) and thus (2.5).

We now turn to the computation of the variance. An important observation is that the family $\{\varphi_I : I \in \mathcal{P}_m([\![n]\!])\}$ is orthonormal in $L^2(S^m, \frac{1}{m!}\lambda^{\otimes m})$. Thus, using (2.8), Pythagoras' theorem, and (2.5), we find that

$$\mathrm{Var}(\mathsf{P}_{\mathsf{X}} f_1 \wedge \ldots \wedge \mathsf{P}_{\mathsf{X}} f_m) = \sum_{I \in \mathcal{P}_m([\![n]\!])} \mathbb{E}\big[(\det A^I)^2\big] - \big\| \Pi^H f_1 \wedge \ldots \wedge \Pi^H f_m \big\|^2. \tag{2.10}$$

Using the same strategy as before, we compute, for each set I of cardinality m,

$$\begin{aligned}
\mathbb{E}\big[|\det A^I|^2\big] &= \frac{1}{n!} \int_{S^n} |\det A^I|^2 \, |\det M|^2 \, d\lambda^{\otimes n} \\
&= \frac{1}{n!} \int_{S^n} |\det M_{[I \leftarrow F]}|^2 \, d\lambda^{\otimes n} \\
&= \det(\langle \psi_{I,a}, \psi_{I,b} \rangle)_{1 \le a,b \le n}.
\end{aligned}$$

The last matrix has a simple block structure coresponding to the partition $[\![n]\!] = I \sqcup I^c$, in which the block indexed by (I^c, I^c) is the identity. The Schur complement formula thus gives

$$\begin{aligned}\det(\langle \psi_{I,a}, \psi_{I,b}\rangle)_{1\le a,b\le n} &= \det\Big(\langle f_i, f_j\rangle_{1\le i,j\le m} - \big(\langle f_i, \varphi_b\rangle\big)_{1\le i\le m, b\in I^c}\big(\langle \varphi_a, f_j\rangle\big)_{a\in I^c, 1\le j\le m}\Big)\\ &= \det\big(\langle f_i, (\mathrm{Id} - \Pi^{H_{I^c}}) f_j\rangle\big)_{1\le i,j\le m}\\ &= \det\big(\langle f_i, (\Pi^{H^\perp} + \Pi^{H_I}) f_j\rangle\big)_{1\le i,j\le m},\end{aligned}$$

where for all $J \subseteq \{1, \ldots, n\}$, we set $H_J = \mathrm{Vect}(\varphi_j, j \in J)$. Using the Andreieff–Heine identity, we rewrite this determinant as

$$\mathbb{E}\big[|\det A^I|^2\big] = \big\langle f_1 \wedge \ldots \wedge f_m, (\Pi^{H^\perp} + \Pi^{H_I}) f_1 \wedge \ldots \wedge (\Pi^{H^\perp} + \Pi^{H_I}) f_m\big\rangle$$

and what we need now is to sum this quantity over all $I \in \mathcal{P}_m([\![n]\!])$.

For each $i \in \{1, \ldots, m\}$, let us decompose f_i as $f_{i,0} + f_{i,1} + \ldots + f_{i,n}$, where $f_{i,0} = \Pi^{H^\perp} f_i$ and for all $j \in \{1, \ldots, n\}$, $f_{i,j} = \langle \varphi_j, f_i\rangle \varphi_j$. By multilinearity, we find

$$\mathbb{E}\big[|\det A^I|^2\big] = \sum_{j_1,\ldots,j_m=0}^{n} \big\langle f_1 \wedge \ldots \wedge f_m, \underbrace{(\Pi^{H^\perp} + \Pi^{H_I}) f_{1,j_1} \wedge \ldots \wedge (\Pi^{H^\perp} + \Pi^{H_I}) f_{m,j_m}}_{R}\big\rangle.$$

Let us call R the function in the right-hand side of the scalar product. If among the integers $j_1, \ldots, j_m$ two are positive and equal, then R vanishes, and so does the corresponding term of the sum. Let us now assume that the positive indices among $j_1, \ldots, j_m$ are pairwise distinct, and let us list them as $\{l_1, \ldots, l_{m-k}\}$, where $k = \mathbb{1}_{\{j_1=0\}} + \ldots + \mathbb{1}_{\{j_m=0\}}$. We make three observations. Firstly, for R not to be zero, it is necessary that $\{l_1, \ldots, l_{m-k}\} \subseteq I$. Secondly, if this condition is satisfied, then $R = f_{1,j_1} \wedge \ldots \wedge f_{m,j_m}$, and in particular does not depend on I. Finally, the condition $\{l_1, \ldots, l_{m-k}\} \subseteq I$ is verified for $\binom{n-m+k}{k}$ subsets I of $\{1, \ldots, n\}$ with m elements. Putting these observations together, we find

$$\sum_{I\in\mathcal{P}_m([\![n]\!])} \mathbb{E}\big[|\det A^I|^2\big] = \sum_{j_1,\ldots,j_m} \binom{n-m+k}{k} \big\langle f_1 \wedge \ldots \wedge f_m, f_{1,j_1} \wedge \ldots \wedge f_{m,j_m}\big\rangle.$$

The sum runs over those $j_1, \ldots, j_m$ between 0 and n among which no two are positive and equal, but lifting this condition only adds null terms to the sum. Therefore, we let $j_1, \ldots, j_m$ run freely between 0 and n, and k is the number of them that are zero.

Let us sort the terms of the last sum according to which of the indices $j_1, \ldots, j_m$ are zero and which are not: calling B the set $\{p \in [\![m]\!] : j_p = 0\}$ and with the notation $H^0 = H$ and $H^1 = H^\perp$, this resummation yields

$$\sum_{I\in\mathcal{P}_m([\![n]\!])} \mathbb{E}\big[|\det A^I|^2\big] = \sum_{k=0}^{m} \binom{n-m+k}{k} \Big\langle f_1 \wedge \ldots \wedge f_m, \sum_{B\in\mathcal{P}_k([\![m]\!])} \Pi^{H^{\mathbb{1}_B(1)}} f_1 \wedge \ldots \wedge \Pi^{H^{\mathbb{1}_B(m)}} f_m\Big\rangle. \quad (2.11)$$

The sum over B yields exactly the function $\Pi_k(f_1 \wedge \ldots \wedge f_m)$. The result follows from the orthogonality of the decomposition (2.4) and the observation that the term corresponding to $k = 0$ is exactly the last term of (2.10). □

Acknowledgements

We thank the two referees for interesting and helpful comments.

References

Avron, H. and Boutsidis, C. Faster subset selection for matrices and applications. *SIAM J. Matrix Anal. Appl.*, **34** (4), 1464–1499 (2013). MR3121759.

Ben-Tal, A. and Teboulle, M. A geometric property of the least squares solution of linear equations. *Linear Algebra Appl.*, **139**, 165–170 (1990). MR1071706.

Berg, L. Three results in connection with inverse matrices. In *Proceedings of the symposium on operator theory (Athens, 1985)*, volume 84, pp. 63–77 (1986). MR872276.

Borodin, A. Determinantal point processes. In *The Oxford handbook of random matrix theory*, pp. 231–249. Oxford Univ. Press, Oxford (2011). MR2932631.

Bufetov, A. I. and Qiu, Y. The Patterson-Sullivan reconstruction of pluriharmonic functions for determinantal point processes on complex hyperbolic spaces. *Geom. Funct. Anal.*, **32** (2), 135–192 (2022). MR4408430.

Bufetov, A. I., Qiu, Y., and Shamov, A. Kernels of conditional determinantal measures and the Lyons-Peres completeness conjecture. *J. Eur. Math. Soc. (JEMS)*, **23** (5), 1477–1519 (2021). MR4244512.

Catanzaro, M. J., Chernyak, V. Y., and Klein, J. R. On Kirchhoff's theorems with coefficients in a line bundle. *Homology Homotopy Appl.*, **15** (2), 267–280 (2013). MR3138380.

Dereziński, M., Liang, F., and Mahoney, M. W. Exact Expressions for Double Descent and Implicit Regularization via Surrogate Random Design. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS'20. Curran Associates Inc. (2020). ISBN 9781713829546. Available at https://proceedings.neurips.cc/paper/2020/file/37740d59bb0eb7b4493725b2e0e5289b-Paper.pdf.

Dereziński, M. and Mahoney, M. W. Determinantal point processes in randomized numerical linear algebra. *Notices Amer. Math. Soc.*, **68** (1), 34–45 (2021). MR4202314.

Dereziński, M. and Warmuth, M. K. Reverse iterative volume sampling for linear regression. *J. Mach. Learn. Res.*, **19**, Paper No. 23, 39 (2018). MR3862430.

Dereziński, M., Warmuth, M. K., and Hsu, D. Unbiased estimators for random design regression. *J. Mach. Learn. Res.*, **23**, Paper No. 167, 46 (2022). Available at http://jmlr.org/papers/v23/19-571.html.

Ermakov, S. M. and Zolotukhin, V. G. Polynomial Approximations and the Monte-Carlo Method. *Theory Probab. Appl.*, **5** (4), 428–431 (1960). DOI: 10.1137/1105046.

Gautier, G., Bardenet, R., and Valko, M. On two ways to use determinantal point processes for Monte Carlo integration. In Wallach, H. et al., editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc. (2019). Available at https://proceedings.neurips.cc/paper/2019/file/1d54c76f48f146c3b2d66daf9d7f845e-Paper.pdf.

Ghosh, S. Determinantal processes and completeness of random exponentials: the critical case. *Probab. Theory Related Fields*, **163** (3-4), 643–665 (2015). MR3418752.

Johansson, K. Random Matrices and Determinantal Processes. In Bovier, A. et al., editors, *Mathematical Statistical Physics*, volume 83 of *Les Houches*, pp. 1–56. Elsevier (2006). DOI: 10.1016/S0924-8099(06)80038-7.

Kassel, A. and Lévy, T. Determinantal probability measures on Grassmannians. *Ann. Inst. Henri Poincaré D*, **9** (4), 659–732 (2022). MR4525143.

Kirchhoff, G. Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Annalen der Physik*, **148** (12), 497–508 (1847). DOI: 10.1002/andp.18471481202.

Lyons, R. Determinantal probability measures. *Publ. Math. Inst. Hautes Études Sci.*, **98**, 167–212 (2003). MR2031202.

Lyons, R. Determinantal probability: basic properties and conjectures. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV*, pp. 137–161. Kyung Moon Sa, Seoul (2014). MR3727606.

Macchi, O. The coincidence approach to stochastic point processes. *Advances in Appl. Probability*, **7**, 83–122 (1975). MR380979.

Mariet, Z. E. and Sra, S. Elementary Symmetric Polynomials for Optimal Experimental Design. In Guyon, I. et al., editors, *Advances in Neural Information Processing Systems*, volume 30, pp. 2136–2145. Curran Associates, Inc. (2017). ISBN 9781510860964. Available at https://proceedings.neurips.cc/paper/2017/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf.

Maurer, S. B. Matrix generalizations of some theorems on trees, cycles and cocycles in graphs. *SIAM J. Appl. Math.*, **30** (1), 143–148 (1976). MR392635.

Shirai, T. and Takahashi, Y. Random point fields associated with certain Fredholm determinants. I. Fermion, Poisson and boson point processes. *J. Funct. Anal.*, **205** (2), 414–463 (2003). MR2018415.

Soshnikov, A. Determinantal random point fields. *Uspekhi Mat. Nauk*, **55** (5(335)), 107–160 (2000). MR1799012.