Classification of transformations of equivalent kernels of some determinantal point processes

Harry Sapranidis Mantelos[3]

[3]University of Warwick, Coventry, CV4 7AL, United Kingdom

Abstract

Determinantal point processes (abbr., DPPs) are randomly arranged points whose distribution is characterized via determinants of matrices. The entries of these matrices are given by one fixed function of two variables, $K(x, y)$, called the kernel of the DPP. It is well-known that the kernel K is not unique and that there exist various other functions of two variables that are valid kernels of the same DPP. We refer to such kernels as *equivalent kernels to* K. It was recently shown by Stevens in [Random Matrices: Theory and Applications, 10(03):2150027, 2021] that, restricting to the case of symmetric kernels, all equivalent kernels of some DPP can be transformed into one another by conjugation transformations. This partially solves a conjecture of Bufetov from 2017 which states that all equivalent kernels of some DPP can be transformed into one another by conjugation and transposition transformations. In this work, we completely relax the symmetry assumptions on the kernels. We go through why this conjecture cannot hold in this general setting but that under some surprisingly simple and natural conditions on the kernel it does.

Keywords

determinantal point processes, equivalent correlation kernels, cycle graphs, cocycle functions

1. Introduction and Motivation

Point processes are probabilistic models for random scatterings of points in some mathematical space. These processes are characterized by their correlation function. If a point process is determinantal (a.k.a., the point process is a DPP) then its correlation function is of a determinantal form (see, e.g., [1], [2], [3]).

More precisely, if we denote by ρ_n the n^{th} correlation function of such a point process on some measure space Λ, then there exists a function $K : \Lambda^2 \to \mathbb{F}$ such that for any $n > 0$ and any tuple $(x_1, \ldots, x_n) \in \Lambda^n$, we have

$$\rho_n(x_1, \ldots, x_n) = \det\left(K(x_i, x_j)\right)_{i,j=1}^n,$$

where $\mathbb{F}$ is a suitable field which in most (if not all) cases is either $\mathbb{R}$ or $\mathbb{C}$ (and so this is what we also take $\mathbb{F}$ to be in our work). We then call K a **correlation kernel** of the DPP.

As in [4] (the paper we build on), for our research problem, we neglect the measure space structure of Λ and just consider it as a set. Hence, the kernel K can be regarded as a mere function. If there exists another function $Q : \Lambda^2 \to \mathbb{F}$ such that

$$\det(Q(x_i, x_j))_{i,j=1}^n = \det(K(x_i, x_j))_{i,j=1}^n \qquad \forall x_1, \ldots, x_n \in \Lambda \ \ \forall n \in \mathbb{N}, \tag{1}$$

then we call K and Q **equivalent kernels** ($K \equiv Q$ for short).

Remark 1. *If Q is of the form*

$$Q(x, y) = g(x)g(y)^{-1}K(x, y), \qquad x, y \in \Lambda,$$

*for some non-zero function $g : \Lambda \to \mathbb{F}$, we say that Q is a **conjugation transformation** of K with **conjugation function** g. It is then easy to check that $Q \equiv K$.*

Remark 2. *It is not difficult to see that if $Q(x, y) = K(y, x)$ for all $x, y \in \Lambda$, then $Q \equiv K$. In this case we say that Q is a **transposition transformation** of K.*

Research Twinning Conference on Digitalisation and Digital Transformation 2023, March 27-30, 2023, Liverpool, UK
h.sapranidis-mantelos@warwick.ac.uk (H. S. Mantelos)
https://warwick.ac.uk/fac/sci/statistics/staff/research_students/mantelos (H. S. Mantelos)
0009-0001-0095-5494 (H. S. Mantelos)

CEUR Workshop Proceedings (CEUR-WS.org)

Although the problem of classifying transformations that yield equivalent kernels of DPPs - in the above framework - originates (and is of theoretical interest) in the theory of determinantal point processes, it essentially asks a basic question about a pair of functions of two variables: given that function K is related to Q via equation (1) - and thus $K \equiv Q$ - what are all the possible transformations that transform K into Q? In this setup, it was established in [4] that, restricting to the case of symmetric kernels (a function h of two variables is **symmetric** if $h(x,y) = h(y,x)$ for every x, y), K must be a conjugation transformation of Q. More precisely, the following result was established in [4].

Theorem 1.1 (theorem 1.5 from [4]). *Suppose that Λ is a set, let $\mathbb{F}$ be a field and let $K, Q : \Lambda^2 \to \mathbb{F}$ be symmetric kernels. If K and Q are equivalent (i.e., equation (1) holds), then it must be the case that Q and K are conjugation transformations of one another, i.e., there exists a non-zero function $g : \Lambda \to \mathbb{F}$ such that for every $x, y \in \Lambda$,*

$$Q(x,y) = g(x)g(y)^{-1}K(x,y). \tag{2}$$

Of course, in the (symmetric) setting of the above theorem, a transposition transformation would have just been the trivial identity transformation. Thus, [4] partially solves the following conjecture of Bufetov, in the case where **both** equivalent kernels K and Q are symmetric.

Conjecture 1 (Conj. 1.4 from [4]). *If K and Q are equivalent kernels (i.e., (1) is satisfied), then they can be transformed into one another by transposition and conjugation transformations.*

It was then left as an open problem to solve Bufetov's conjecture in the general setting without the symmetry assumptions on the kernels. It is imperative, of course, to first determine whether the conjecture does indeed hold in that general setting. In the next section we show that, in fact, it is possible to find counterexamples to the conjecture. However, the nature of such counterexamples, as we shall see, is rather exceptional. Fortunately, our research findings show that it is possible to impose some simple and natural conditions on the kernels that elegantly rule out such counterexamples. The proof of the conjecture (under these conditions) - briefly outlined in this paper - is essentially revolved around the recurrent use of an elementary lemma (lemma 4.2) and some basic graph theory and combinatorics.

Hereafter we summarize the research conducted in [arXiv:2302.02471].

2. Counterexamples of Bufetov's Conjecture

Firstly, let us consider the case where Λ from Section 1 is any discrete set of cardinality $n \geq 4$. For simplicity, assume $\Lambda = \{1, 2, \dots, n\}$. Then, consider the kernels $K : \Lambda^2 \to \mathbb{F}$ and $Q : \Lambda^2 \to \mathbb{F}$ defined by

$$(K(x,y))_{1 \leq x,y \leq n} = \left[\begin{array}{c|c} \mathrm{C} & 0 \\ \hline 0 & \mathrm{D} \end{array}\right], \qquad (Q(x,y))_{1 \leq x,y \leq n} = \left[\begin{array}{c|c} C^T & 0 \\ \hline 0 & D \end{array}\right],$$

where C and D are some non-symmetric square matrices of the same dimension. In this case $K \equiv Q$ but K is neither a transposition nor a conjugation transformation of Q. Thus, Bufetov's conjectured classification of transformations does not hold: we have just identified one other type of transformation other than the two stated in the conjecture which yields equivalent kernels. Perhaps a suitable name for it would be a "**partial transposition transformation**", since only *part* of the matrix is being transposed.

Interestingly, by the block determinant formula, even if the null sub-matrices of the above pair of matrices were replaced by matrices of ones, it would still be the case that $K \equiv Q$. For the same reason as in the previous paragraph, Bufetov's conjecture would still not stand.

For reasons that will be explained in the next section, our analysis is restricted to non-zero kernels (i.e., kernels K such that $K(x,y) \neq 0$ for every $x, y \in \Lambda$). To this end, let us further analyze the counterexample specified in the preceding paragraph with the additional condition that the matrices C and D are entrywise non-zero. We seek for a condition on K that would make kernels K and Q of the previously described form a non-admissible pair satisfying (1). We first need to deal with the case when $n = 4$, in which case kernels K

and Q from before satisfy

$$(K(x,y))_{1\le x,y\le n} = \begin{pmatrix} a & b & 1 & 1 \\ c & d & 1 & 1 \\ 1 & 1 & e & f \\ 1 & 1 & g & h \end{pmatrix}, \qquad (Q(x,y))_{1\le x,y\le n} = \begin{pmatrix} a & c & 1 & 1 \\ b & d & 1 & 1 \\ 1 & 1 & e & f \\ 1 & 1 & g & h \end{pmatrix}, \tag{3}$$

for some $a, b, c, d, e, f, g, h \in \mathbb{F} \setminus \{0\}$ such that $c \neq b$ and $f \neq g$.

Clearly, the sub-matrices of ones in the upper-right and lower-left block regions of the above two block matrices are the ones that bring about the block structure that enables this "partial transposition" transformation discussed previously to be a viable transformation that transforms equivalent kernels (of a special form) into one another. And so by far the most intuitive condition to impose on our pair of matrices that we would hope would be enough to make this type of transformation non-permissible would be to require non-zero determinants in the upper-right and lower-left 2×2 sub-matrix regions, that is, to require

$$\begin{vmatrix} K(1,3) & K(1,4) \\ K(2,3) & K(2,4) \end{vmatrix}, \quad \begin{vmatrix} K(3,1) & K(3,2) \\ K(4,1) & K(4,2) \end{vmatrix} \neq 0,$$

which - in the case of $n = 4$ - is the same as requiring for every $x, y, z, w \in \Lambda$ distinct,

$$\begin{vmatrix} K(x,y) & K(x,w) \\ K(z,y) & K(z,w) \end{vmatrix} \neq 0.$$

Our research findings not only show that this condition, in fact, is sufficient in solving Bufetov's conjecture in the above setting with $\Lambda = \{1, 2, \ldots, n\}$ and $n = 4$ (or for any discrete set Λ of any cardinality, for that matter); but even for a **general** abstract set Λ (of any kind of cardinality).

3. Main result

Precisely, our main result in [5] reads as follows.

Theorem 3.1. *Suppose that Λ is a set, let $\mathbb{F}$ be a field and let $K, Q : \Lambda^2 \to \mathbb{F}$ be non-zero (not necessarily symmetric) kernels such that for every $x, y, z, w \in \Lambda$ distinct,*

$$\begin{vmatrix} K(x,y) & K(x,w) \\ K(z,y) & K(z,w) \end{vmatrix} \neq 0. \tag{4}$$

*If K and Q are equivalent (i.e., equation (1) holds - in which case, equation (4) with K replaced by Q also holds), then the following two transformations are the **only** possible transformations that transform K into Q:*

- *Conjugation transformations, i.e., there exists a non-zero function $g : \Lambda \to \mathbb{F}$ such that (2) holds for every $x, y \in \Lambda$.*
- *Transposition transformations followed by conjugation transformations, i.e., there exists a non-zero function $g : \Lambda \to \mathbb{F}$ such that for every $x, y \in \Lambda$,*

$$Q(x,y) = g(x)g(y)^{-1}K(y,x). \tag{5}$$

In the last section, the significance of the (4) condition from theorem 3.1 was explained and illustrated; but the reason why we require non-zero kernels was not. The reason we insist on the latter is because, had it not been so, then we may have been presented with quite a few problems in solving Bufetov's conjecture in this general (not necessarily symmetric) setting. Indeed, equation (1) with $n = 1$ and $n = 2$ can easily be seen to yield the identity

$$K(x,y)K(y,x) = Q(x,y)Q(y,x), \qquad x, y \in \Lambda. \tag{6}$$

This identity, in the symmetric setting of [4] (where both K and Q are symmetric), implies that for every $x, y \in \Lambda$,

$$K(x,y) = 0 \text{ if and only if } Q(x,y) = 0. \tag{7}$$

However, when we drop this symmetry assumption on both kernels, we do not have the luxury of concluding the statement in (7). On the contrary, there could very well exist $\bar{x}, \bar{y}, \hat{x}, \hat{y} \in \Lambda$ such that, for example, $K(\bar{x}, \bar{y}) = 0$, $K(\bar{y}, \bar{x}) \neq 0$, $Q(\bar{x}, \bar{y}) \neq 0$, $Q(\bar{y}, \bar{x}) = 0$, $K(\hat{x}, \hat{y}) = 0$, $K(\hat{y}, \hat{x}) \neq 0$, $Q(\hat{x}, \hat{y}) = 0$ and $Q(\hat{y}, \hat{x}) \neq 0$. In this case, both the pair $\bar{x}, \bar{y}$ and the pair $\hat{x}, \hat{y}$ satisfy equation (6) - so there's no contradiction to the (1) equation. However, the pair $\bar{x}, \bar{y}$ could not possibly satisfy equation (2) for some non-zero function g; nor could the pair $\hat{x}, \hat{y}$ satisfy equation (5) for some non-zero function g. Therefore, K and Q in this case could not possibly be transformed into one another through conjugation and transposition transformations: neither equation (2) nor equation (5) is satisfied **for every** $x, y \in \Lambda$.

4. Idea & Outline of the Proof of Theorem 3.1

The main tools we used to prove theorem 3.1 were the classic Leibniz formula for determinants, some (perhaps known) results from graph theory regarding cycles of lengths 2, 3 and 4, the concept of a cocycle function (see [4] for more details), and some basic combinatorics. We provide a brief outline of our proof below. We begin by laying some necessary groundwork and introducing some (perhaps non-standard) terminology.

Definition 1. *We call an $(n+1)$-tuple $(x_0, x_1, \ldots, x_n) \in \Lambda^{n+1}$ such that $x_n = x_0$ a **cycle of length** n. Moreover, if all the x_i's in the tuple are distinct (except for x_0 and x_n - which by definition are equal), then we call $(x_0, x_1, \ldots, x_n)$ a **simple cycle**.*

Definition 2. *We call a function $c : \Lambda^2 \to \mathbb{F}$ a **cocycle function** if for every $r \geq 1$ and for every tuple $(z_1, \ldots, z_r) \in \Lambda^r$,*

$$c(z_1, z_2)c(z_2, z_3) \cdots c(z_{r-1}, z_r)c(z_r, z_1) = 1. \tag{8}$$

Definition 3. *If $Q : \Lambda \to \mathbb{F}$ and $K : \Lambda \to \mathbb{F}$ are functions that satisfy*

$$Q(x, y) = c(x, y)K(x, y) \qquad \forall \; x, y \in \Lambda$$

*for some cocycle function $c : \Lambda^2 \to \mathbb{F}$, then we say that Q is a **cocycle transformation** of K.*

One can easily verify that a cocycle transformation of a kernel yields an equivalent kernel. It is also a straightforward exercise to show that, under our paper's non-zero kernel assumption, a kernel Q is a cocycle transformation of a kernel K if and only if Q is a conjugation transformation of K. This is a critical observation which simplifies our work, since establishing that some function satisfies the cocycle property (8) is objectively easier than constructing some conjugation function from scratch. In particular, the previous observation means that if we are able to prove that either the function

$$S : \Lambda^2 \to \mathbb{F}, \;\; S(x, y) = \frac{Q(x, y)}{K(x, y)},$$

or the function

$$\tilde{S} : \Lambda^2 \to \mathbb{F}, \;\; \tilde{S}(x, y) = \frac{Q(x, y)}{K(y, x)}$$

is a cocycle function, then, since for every $x, y \in \Lambda$,

$$Q(x, y) = S(x, y)K(x, y) \text{ and } Q(x, y) = \tilde{S}(x, y)K(y, x),$$

it would follow that Q must be in one of the two forms from the conclusion of theorem 3.1 - which is exactly the statement of Bufetov's conjecture. And so the proof of theorem 3.1 boils down to proving that one of the functions S or $\tilde{S}$ must be a cocycle function.

For non-zero functions of two variables (which, under our paper's non-zero kernel assumption, both S and $\tilde{S}$ are), there is a shortcut to establishing the cocycle property (8); we state it below.

Proposition 4.1. *If a non-zero function $c : \Lambda^2 \to \mathbb{F}$ satisfies*

1. $c(x,x) = 1$ *for every* $x \in \Lambda$;
2. $c(x,y)c(y,x) = 1$ *for every* $x, y \in \Lambda$;
3. $c(x,y)c(y,z)c(z,x) = 1$ *for every* $x, y, z \in \Lambda$,

then c is a cocycle function.

In other words, c satisfying the cocycle property for every cycle of length 1, 2 and 3 is both a necessary and sufficient condition for c to be a (full) cocycle function, that is, for it to satisfy the cocycle property for any cycle of any length. This result simplifies/shortens our analysis in that instead of having to prove the cocycle property (8) for every cycle in Λ (of every possible length) for either the function S or $\tilde{S}$; proving it for cycles of lengths 1, 2 and 3 is sufficient.

It is immediate from equation (1) with $n = 1$ and equation (6) that both S and $\tilde{S}$ automatically satisfy the cocycle property for any cycle of length 1 and 2 in Λ. Thus, by Proposition 4.1, it suffices to show that either S or $\tilde{S}$ satisfies the cocycle property for all cycles of length 3 in Λ. This task, however, is not at all trivial in this general setting. In particular, it is no longer enough to just use equation (1) with $n = 3$ in conjunction with the Leibniz formula for determinants (as was the case in [4]). The reason being that in this general setting the aforementioned step will merely yield, for any simple cycle $p = (p_i)_{i=0}^3$ of length 3 in Λ,

$$\begin{gathered}\overbrace{K(p_0,p_1)K(p_1,p_2)K(p_2,p_0)}^{=:\,K[p]} + \overbrace{K(p_0,p_2)K(p_2,p_1)K(p_1,p_0)}^{=:\,K'[p]} \\ = \\ \underbrace{Q(p_0,p_1)Q(p_1,p_2)Q(p_2,p_0)}_{Q[p]\,:=} + \underbrace{Q(p_0,p_2)Q(p_2,p_1)Q(p_1,p_0)}_{Q'[p]\,:=} .\end{gathered} \tag{9}$$

Note that, had both K and Q been symmetric, equation (9) would simplify to $\cancel{2}K[p] = \cancel{2}Q[p]$, which proves the cocycle property for all cycles of length 3 in Λ for the function S from before. In contrast, to be able to proceed with equation (9) in the non-symmetric setting, one needs to make the additional observation that $K[p]K'[p] = Q[p]Q'[p]$ also holds. This is an immediate consequence of equation (6). Thereafter the following elementary lemma can be applied.

Lemma 4.2. *Let $a, b, a', b' \in \mathbb{F}$ be constants satisfying*

$$a + b = a' + b' \qquad \textit{and} \qquad ab = a'b'. \tag{10}$$

Then, it is either the case that $a = a'$ and $b = b'$; or, $a = b'$ and $b = a'$.

This simple lemma is the key ingredient of our proof, as it is what first introduces the "*dichotomy*" that enables us to eventually reach our desired conclusion about S and $\tilde{S}$ (and is repeatedly invoked).

Thereafter we need to extract a bit more information about the relationship between K and Q. For this we make use of equation (1) with $n = 4$ in conjunction with the Leibniz formula for determinants. As a result cycles of length 4 also enter the scene; and so then our task is to find a way to link cycles of lengths 2, 3 and 4 together in some suitable way. In figures 1-3 below we represent graphically the three lemmas from graph theory we use to achieve this. In very broad terms, the way in which we exploited these lemmas was via the following graphical identification we made for functions of two variables evaluated on cycles: for an arbitrary function $h : \Lambda^2 \to \mathbb{F}$, we identify (graphically) the product $h[p] := \prod_{i=1}^n h(p_{i-1}, p_i)$, where $p := (p_i)_{i=0}^n \in \Lambda^{n+1}$ is a cycle of length n in Λ, by precisely the (drawn) cycle graph of p.

5. Open Problems

Firstly, one could attempt to extend the results from [4] and [5] (summarized in this extended abstract) by taking potential measure space structures of Λ into consideration.

A somewhat related question one could also explore is the following: given that we start with a non-symmetric kernel, say K, of a DPP, does there necessarily exist an equivalent kernel of K, say Q, which is symmetric?

Figure 1: In this figure we seek to write the disjoint union of two distinct simple cycles of length 4, $p = (p_0, p_1, p_2, p_3, p_0)$ and $q = (p_0, p_2, p_3, p_1, p_0)$ (drawn in black and red, respectively, in the LHS directed graph) as the disjoint union of distinct simple cycles of lengths 2 and 3. We do this in the RHS directed graph by considering the 3-cycle $s^{(1)} = (p_2, p_3, p_1, p_2)$, the 3-cycle $s^{(2)} = (p_2, p_3, p_0, p_2)$ and the 2-cycle $r = (p_0, p_1, p_0)$.

Figure 2: In this figure we decompose a simple cycle of length 4, $p = (p_0, p_1, p_2, p_3, p_0)$, into two simple cycles of length 3. There are two possible such decompositions: on the LHS graph, in red is the 3-cycle $s^{(1)} = (p_0, p_1, p_2, p_0)$, and in blue is the 3-cycle $s^{(2)} = (p_0, p_2, p_3, p_0)$; on the RHS graph, in red is the 3-cycle $r^{(1)} = (p_0, p_1, p_3, p_0)$, and in blue is the 3-cycle $r^{(2)} = (p_1, p_2, p_3, p_1)$.

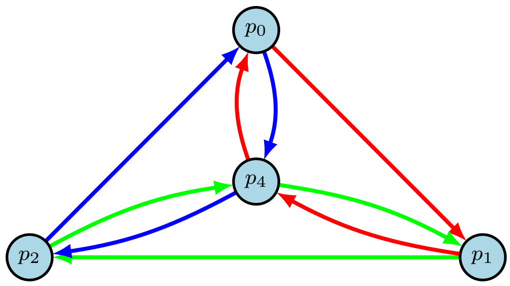

Figure 3: In this figure we decompose a 3-cycle $p = (p_0, p_1, p_2, p_0)$ into three distinct simple cycles of length 3 that each contains the vertex p_4, where $p_4 \neq p_i$ for all $i \in \{0, 1, 2\}$. In red is the 3-cycle $s^{(1)} = (p_0, p_1, p_4, p_0)$, in green is the 3-cycle $s^{(2)} = (p_1, p_2, p_4, p_1)$, and in blue is the 3-cycle $s^{(3)} = (p_2, p_0, p_4, p_2)$.

Acknowledgments

Harry Sapranidis Mantelos is supported by the Warwick Mathematics Institute Centre for Doctoral Training, and gratefully acknowledges funding from the University of Warwick.

References

[1] J. Hough, M. Krishnapur, Y. Peres, B. Virág, Zeros of Gaussian Analytic Functions and Determinantal Point Processes, University lecture series, American Mathematical Society, 2009. URL: https://books.google.co.uk/books?id=KCuCAwAAQBAJ.

[2] A. Soshnikov, Determinantal random point fields, Russian Mathematical Surveys 55 (2000) 923.

[3] A. Kulesza, B. Taskar, Determinantal Point Processes for Machine Learning, Foundations and trends in machine learning, World Scientific, 2012. URL: https://books.google.gr/books?id=FY-6lAEACAAJ.

[4] M. Stevens, Equivalent symmetric kernels of determinantal point processes, Random Matrices: Theory and Applications 10 (2021) 2150027.

[5] C. S. Mantelos, Classification of transformations of equivalent kernels of some determinantal point processes, 2023. URL: https://arxiv.org/abs/2302.02471. doi:10.48550/ARXIV.2302.02471.