Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

Constructing all determinantal sampling designs

by Vincent Loonis

Release date: January 3, 2024

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, December 2023
Vol. 49, No. 2, pp. 411-441
Statistics Canada, Catalogue No. 12-001-X

Constructing all determinantal sampling designs

Vincent Loonis[1]

Abstract

In this article, we use a slightly simplified version of the method by Fickus, Mixon and Poteet (2013) to define a flexible parameterization of the kernels of determinantal sampling designs with fixed first-order inclusion probabilities. For specific values of the multidimensional parameter, we get back to a matrix from the family P^{Π} from Loonis and Mary (2019). We speculate that, among the determinantal designs with fixed inclusion probabilities, the minimum variance of the Horvitz and Thompson estimator (1952) of a variable of interest is expressed relative to P^{Π}. We provide experimental R programs that facilitate the appropriation of various concepts presented in the article, some of which are described as non-trivial by Fickus et al. (2013). A longer version of this article, including proofs and a more detailed presentation of the determinantal designs, is also available.

Key Words: Determinantal process; Balanced sampling; Semidefinite optimization.

1. Introduction

In sampling theory, a random sample $\mathbb{S}$ is a random variable whose realizations are elements of the set 2^U of the parts of a finite population U of size N, indexed by $k = 1, \ldots, N$. Each part s of 2^U is called a sample. The probability law of $\mathbb{S}$ is called sampling design (Tillé, 2001). Apart from the terminology, these concepts are the same as those used in point processes on a finite population, in probability or statistics. Among point processes, determinantal processes have been the subject of many studies because they appear in various fields: random matrices, mathematical physics or machine learning. Using these processes in the context of sampling theory and finite populations leads to the concept of determinantal sampling design studied by Loonis and Mary (2019).

Determinantal sampling designs directly inherit properties from determinantal processes established in different frameworks. They are parameterized by Hermitian matrices whose eigenvalues lie between 0 and 1 (Macchi, 1975; Soshnikov, 2000). These are called contracting Hermitian matrices. Such a matrix will then be notated K, called the kernel, and the associated determinantal sampling design will be notated $\mathrm{DSD}(K)$. The inclusion probabilities of such designs are known at all orders and are parameterized by K. This feature differentiates determinantal designs from most of the usual, somewhat complex, designs giving them a real practical interest, beyond their theoretical curiosity. Usually, if the first-order inclusion probabilities are known, the second-order inclusion probabilities are often approximated and the others are most often unknown. For determinantal designs, the first- and second-order inclusion probabilities are given, respectively, by the diagonal and non-diagonal terms of the kernel. Higher-order inclusion probabilities are also expressed directly in terms of K.

Moreover, the distribution of the size of a determinantal random sample is that of a sum of independent Bernoulli variables, where the parameters are the eigenvalues of K (Hough, Krishnapur, Peres and Virág,

1. Vincent Loonis, Insee (Institut national de la statisitque et des études économiques), Paris, France. E-mail: vincent.loonis@insee.fr.

2006). There are algorithms for selecting determinantal random samples (Hough, Krishnapur, Peres and Virág, 2006; Scardicchio, Zachary and Torquato, 2009; Lavancier, Møller and Rubak, 2015). Determinantal designs are negatively associated (Lyons, 2003). In other words, if A and B are two disjointed parts of U, then $\mathrm{pr}(A \cup B \subseteq \mathbb{S}) \leq \mathrm{pr}(A \subseteq \mathbb{S})\,\mathrm{pr}(B \subseteq \mathbb{S})$. As a result, determinantal designs satisfy the Sen-Yates-Grundy conditions (Sen, 1953; Yates and Grundy, 1953), which correspond to the case where A and B are singletons. They also satisfy the strong Rayleigh property (Yuan, Su and Hu, 2003; Brändén and Jonasson, 2012; Pemantle and Peres, 2014). From this technical property, stronger than negative association, a central limit theorem as well as deviation and concentration inequalities (Soshnikov, 2000, 2002; Pemantle et Peres, 2014) directly follow for the Horvitz-Thompson estimator of the total of a variable of interest y (Horvitz and Thompson, 1952).

The specifics of sampling theory lead Loonis and Mary (2019) to focus on novel properties of determinantal designs, such as the necessary and sufficient conditions for the perfect estimation of a total with the Horvitz-Thompson estimator, the search for optimal kernels, or the explicit construction of families of particular contracting Hermitian matrices, such as fixed-diagonal orthogonal projection matrices. These matrices are associated with the important practical case of fixed-size determinantal sampling designs with given first-order inclusion probabilities. More practically, the authors observe that the simplest determinantal designs can be mobilized for populations of several thousand individuals. The size of the population that is compatible with more sophisticated constructions, particularly those associated with optimization issues, is several hundred.

Constructing a fixed-diagonal orthogonal projection matrix is a specific case of the more general problem of constructing Hermitian matrices, whose diagonal and spectrum are fixed. This problem has been the subject of many studies in the literature, some of which are recent. Schur (1911), Horn (1954) and Kadison (2002) studied the necessary and sufficient conditions for the existence of such matrices in complex and real cases. However, these studies do not specify how to construct such matrices. Dhillon, Heath Jr., Sustik and Tropp (2005) propose algorithms for constructing some of these matrices. The algorithm by Fickus et al. (2013) makes it possible to construct them all. In the previous two cases, the results are not known analytically. Loonis and Mary (2019) explicitly show several examples of projection matrices, whose diagonal terms are constant. Building on the work of Kadison (2002), they present formulae for specifically constructing a family P^{Π} of real orthogonal projection matrices of any diagonal Π, provided that $\sum_{k=1}^{N} \Pi_k$ is an integer.

In this article, we focus specifically on the method by Fickus et al. (2013), in the context of determinantal designs, and draw connections with the family P^{Π}. In the first section, we introduce some algebraic notations and concepts that are useful for the overall understanding of the article. In the second section, we reiterate the main properties of determinantal sampling designs. More precisely, we present the properties of the family of designs associated with matrices P^{Π}. The results of this section are taken directly, without proof, from Loonis and Mary (2019). Readers who wish to understand the foundations of these results or are interested in extensions, such as asymptotic properties, are encouraged to refer to the original article. Algorithm 3.1 and theorem 3.1 are new, and their proofs are provided in the longer version of the article.

In the third section, we present the main principles and parameters of the method by Fickus et al. (2013). The latter are not easy to handle directly. We show that, at the cost of simplification, it is possible to modify them to describe the semidefinite positive Hermitian matrices K^{Π} of diagonal Π with a triplet (M, Ω, ρ). M is an integer giving the number of strictly positive eigenvalues of K^{Π}. Ω and ρ are matrices with respective sizes $(M \times N)$ and $(M \times (N-1))$, all of whose coefficients have value in $[0, 1]$. In terms of sampling theory, the columns of matrix Ω directly influence the variability of the sample size falling into domains of the form $D^k = \{1, \ldots, k\}$, $k = 1, \ldots, N$. As for the matrix ρ, it determines the value of the non-diagonal coefficients of the constructed matrices and, as such, the second-order inclusion probabilities of the associated determinantal designs.

In the fourth section, for a probability vector Π such that $\sum_{k=1}^{N} \Pi_k = n \in \mathbb{N}^*$, we examine the choices $(M = n, \Omega = 0^{(n \times N)}, \rho)$ and $(M = n, \Omega = 1^{(n \times N)}, \rho)$. We show that the first is directly related to the family P^{Π} and explain the coefficients of the matrices obtained for the second choice.

In the final section, we carry out simulations and applications. We limit ourselves to real fixed-diagonal orthogonal projection matrices. In this set, we minimize the variance of the Horvitz-Thompson estimator of a variable's total. Given the specificity of the eligible set, we rewrite the problem in the form of an optimization problem on manifolds and mobilize adapted algorithms (Absil, Mahony and Sepulchre, 2009; Boumal, Mishra, Absil and Sepulchre, 2014; Townsend, Koep and Weichwald, 2016). We find that the result is also associated with the family P^{Π}. From this we deduce a conjecture defining the lower bound of the variance of the Horvitz-Thompson estimator of the total of variable y, among the determinantal designs with fixed first-order inclusion probabilities.

In addition to appropriating and programming the method by Fickus et al. (2013) for use in determinantal sampling designs, the main contributions of this document are algorithm 3.1, theorems 3.1, 5.1 and 5.2, and conjecture 6.1. The points considered important for interpreting certain results are presented in the form of remarks or examples. The proofs, which are long and technical, are provided in a long version of the article. The latter also provides a comparison of the performances, in terms of balancing, of the determinantal designs with those of equivalent designs (Deville and Tillé, 2004; Chauvet and Tillé, 2006; Leuenberger, Eustache, Jauslin and Tillé, 2022).

2. Algebraic notations and reminders

In the rest of the article, K is a contracting Hermitian matrix of size $(N \times N)$. A random variable, whose distribution is a determinantal design with kernel K, will be notated $\mathbb{S} \sim \mathrm{DSD}(K)$. The coefficients of K are complex. The conjugate of the complex number z is $\bar{z}$ and its modulus is $|z|$. K is such that $K = \bar{K}^{\top}$ (*Hermitian*) and its eigenvalues are in the interval $[0, 1]$ (*contracting*). If all eigenvalues are 0 or 1, K is an orthogonal projection matrix. The strictly positive eigenvalues of K are M in number. The main submatrix of order $k = 1, \ldots, N$ of K is the square matrix made up of the upper left corner of size k of K. The submatrix of order $k = N$ is K itself.

The set of eigenvalues, called spectrum, of the main submatrix of order k is represented by the vector λ^k. The spectrum of K is λ^N. Each λ^k has at most M strictly positive eigenvalues, which are also less than or equal to 1. Each main submatrix is also Hermitian and contracting. By convention, λ^k is *truncated* to its own strictly positive coefficients at first. If applicable, it is supplemented with 0s to be of constant size M. The coefficients of λ^k are notated λ_j^k, $j = 1, \ldots, M$, $k = 1, \ldots, N$ and are, again by convention, sorted from lowest to highest: $0 \leq \lambda_1^k \leq \ldots \leq \lambda_j^k \leq \ldots \leq \lambda_M^k \leq 1$ with $\lambda_1^N > 0$.

The λ^k, as spectra of the main submatrices, satisfy the Cauchy interlacing conditions (Horn and Johnson, 1991). By placing $\lambda_0^k = 0$ for $k = 1, \ldots, N$, these conditions are written

$$\forall j = 1, \ldots, M,\ \forall k = 1, \ldots, N-1: \lambda_{j-1}^{k+1} \leq \lambda_j^k \leq \lambda_j^{k+1}. \tag{2.1}$$

The letter k is used to index both the individuals of the population and the steps of the algorithm by Fickus et al. (2013). This choice is justified by the fact that step k determines individual k's contribution to the inclusion probabilities of all orders involving it. In general, the first-order inclusion probability of individual $k = 1, \ldots, N$ is π_k, the coefficient of vector π. In some cases, which leave no room for doubt, the letter π is used based on its conventional usage. The notations Π_k and Π will be used when we want π_k to take the value Π_k, set in advance. For example, $\Pi_k = n/N$ or Π_k is proportional to a size criterion, like the number of employees for companies. The various inclusion probabilities contained in vector Π are not necessarily ordered. This is true for constructing matrices P^{Π} from Loonis and Mary (2019). For the method by Fickus et al. (2013), these probabilities will be sorted from highest to lowest, without losing generality. In this case, we will use the notation $\Pi^{\triangleright}$, specifying that $1 > \Pi_1^{\triangleright} \geq \ldots \geq \Pi_k^{\triangleright} \geq \ldots \geq \Pi_N^{\triangleright} > 0$. The notation $\Sigma^{\Pi^{\intercal}}$ refers to a permutation matrix such that $\Pi^{\triangleright} = \Sigma^{\Pi^{\intercal}} \Pi$.

To be consistent with sampling theory, the notations N and M are inverted, compared with those from Fickus et al. (2013). The ordering convention of λ^k also differs from the one chosen by these authors. This seemed to simplify the appropriation of the methods from a programming perspective. This choice can occasionally make the wording of the Schur (1911) and Horn (1954) theorem, as well as some formulae that derive from it, less intuitive. In this context, this theorem stipulates that there is a Hermitian matrix $K_N^{M, \Pi^{\triangleright}, \lambda^N}$ with diagonal $\Pi^{\triangleright}$ and spectrum λ^N if and only if

$$\begin{cases} \sum_{k=1}^{N} \Pi_k^{\triangleright} = \sum_{j=1}^{M} \lambda_j^N = \mu, \\ \sum_{s=M-j+1}^{N} \Pi_s^{\triangleright} = \mu - \sum_{s=1}^{M-j} \Pi_s^{\triangleright} \geq \sum_{s=1}^{j} \lambda_s^N,\ j = 1, \ldots, M-1. \end{cases} \tag{2.2}$$

By adapting the notations, this theorem also applies to all the main submatrices of a Hermitian matrix because they are also Hermitian.

Finally, a matrix A of size $(P \times P)$ is unitary if $\overline{A}^{\intercal} A = I_P$, where I_P is the identity matrix. The square matrix of size $(P \times P)$, all terms of which equal 1, is notated J_P. The symbol $\odot$ indicates Hadamard's matrix product: $A = B \odot C \Rightarrow A_{kl} = B_{kl} C_{kl}$. The notation $x^{(P \times Q)}$ indicates a matrix of size $(P \times Q)$, all coefficients of which equal x. The notation x^P is used for vectors of size P, all coefficients of which equal x. D_x is a diagonal matrix, the diagonal of which is x.

3. Reminders about determinantal sampling designs

3.1 Definition and inclusion probabilities

A random variable $\mathbb{S}$ over 2^U has as law a determinantal sampling design if there is a contracting Hermitian matrix K such that

$$\forall s \in 2^U, \mathrm{pr}(s \subseteq \mathbb{S}) = \det(K_{|s}), \tag{3.1}$$

where $K_{|s}$ is the submatrix constructed by extracting the rows and columns of K indexed by the elements of s. This definition directly results in the calculation of the inclusion probabilities for all orders, particularly those of orders 1 and 2:

$$\begin{cases} \pi_k = \mathrm{pr}(\{k\} \subseteq \mathbb{S}) = \det(K_{|\{k\}}) = K_{kk}, \\ \pi_{kl} = \mathrm{pr}(\{k,l\} \subseteq \mathbb{S}) = \det(K_{|\{k,l\}}) = K_{kk}K_{ll} - K_{kl}K_{lk} = K_{kk}K_{ll} - K_{kl}\bar{K}_{kl} = K_{kk}K_{ll} - |K_{kl}|^2. \end{cases} \tag{3.2}$$

The diagonal terms K_{kk} correspond to the first-order inclusion probabilities of design $\mathrm{DSD}(K)$, whereas the modulus of the non-diagonal terms is used in the expression of the second-order inclusion probabilities. The properties of the *trace* matrix application give rise to a relationship between inclusion probabilities and eigenvalues, which frequently appear in the proofs:

$$\mathrm{Tr}(K) = \sum_{k=1}^{N} K_{kk} = \sum_{k=1}^{N} \pi_k = \sum_{j=1}^{M} \lambda_j^N. \tag{3.3}$$

Matrix Δ, of size $(N \times N)$, whose coefficients are $\Delta_{kk} = \pi_k(1-\pi_k)$ and $\Delta_{kl} = \pi_{kl} - \pi_k \pi_l$, is expressed in terms of K:

$$\begin{cases} \Delta_{kk} & = & K_{kk}(1-K_{kk}), \\ \Delta_{kl} & = & -|K_{kl}|^2, \\ \Delta & = & K \odot (I_N - \bar{K}). \end{cases} \tag{3.4}$$

This matrix is used in precision calculations. It also makes it possible to see that the determinantal designs confirm the Sen-Yates-Grundy conditions (Sen, 1953; Yates and Grundy, 1953) because $\Delta_{kl} = \pi_{kl} - \pi_k \pi_l = -|K_{kl}|^2 \leq 0$ and, therefore, $\pi_{kl} \leq \pi_k \pi_l$. As such, the Sen-Yates-Grundy conditions provide an upper bound for the second-order inclusion probabilities of determinantal designs, function of single-order inclusion probabilities.

3.2 Sample size, fixed-size determinantal designs and selection algorithm

Let $\sharp\mathbb{S}$ be the size of random sample $\mathbb{S} \sim \mathrm{DSD}(K)$. The distribution of $\sharp\mathbb{S}$ is that of a sum of M independent Bernoulli variables, whose parameters are the M strictly positive eigenvalues λ_j^N, $j = 1, \ldots, M$ of K (Hough et al., 2006). The moments of order 1 and 2 of $\sharp\mathbb{S}$ are deduced and equal

$$\begin{cases} \mathbb{E}(\sharp\mathbb{S}) &= \sum_{j=1}^{M} \lambda_j^N = \sum_{k=1}^{N} \pi_k, \\ \operatorname{var}(\sharp\mathbb{S}) &= \sum_{j=1}^{M} \lambda_j^N (1-\lambda_j^N). \end{cases} \tag{3.5}$$

Remark 3.1 $\mathbb{S} \sim \operatorname{DSD}(K)$ *will be of fixed size,* $\mathbb{E}(\sharp\mathbb{S}) = \sharp\mathbb{S} = n \in \mathbb{N}^*$, *if and only if* $\operatorname{var}(\sharp\mathbb{S}) = 0$. *In other words, if* λ_j^N *equals 1 for all* $j = 1, \ldots, M$. *Because* K *is Hermitian,* K *is an orthogonal projection matrix.*

Lavancier et al. (2015) propose an algorithm for selecting a random sample whose distribution is a determinantal design of fixed size. This algorithm requires a spectral decomposition of K as input, which, when N is large, can be computationally intensive. One of the challenges of then constructing kernels K will be to directly provide this decomposition. If K is not a projection matrix, Hough et al. (2006) show that $\operatorname{DSD}(K)$ can be written as a mixture of fixed-size determinantal designs. Thus, it is possible to refer to the case of projection matrices to select a random-sized determinantal sample.

3.3 Estimating a total, variance of the estimators, balanced designs

For any sampling design, whose first-order inclusion probabilities are strictly positive, the unknown total $t_y = \sum_{k=1}^{N} y_k$ of a variable of interest y, assimilated to a vector of $\mathbb{R}^N$, is estimated without bias by the Horvitz-Thompson estimator (Horvitz and Thompson, 1952) such that $\hat{t}_y = \sum_{k\in\mathbb{S}} y_k / \pi_k$ and whose variance is

$$\operatorname{var}(\hat{t}_y) = \sum_{k\in U}\sum_{l\in U} \frac{y_k}{\pi_k}\frac{y_l}{\pi_l}\Delta_{kl} = y^\top D_\pi^{-1} \Delta D_\pi^{-1} y. \tag{3.6}$$

If the design is determinantal, a consequence of (3.4) is that the variance of $\hat{t}_y$ is a function of K:

$$\operatorname{var}(\hat{t}_y) = y^\top (I_N \odot K)^{-1} [K \odot (I_N - \bar{K})] (I_N \odot K)^{-1} y. \tag{3.7}$$

For a set of inclusion probabilities fixed *a priori*, a balanced design is such that

$$\forall q = 1, \ldots, Q, \quad \hat{t}_{x^q} = \sum_{k\in\mathbb{S}} \frac{x_k^q}{\Pi_k} = t_{x^q} \Leftrightarrow \operatorname{var}(\hat{t}_{x^q}) = 0, \tag{3.8}$$

where the x^q, $q = 1, \ldots, Q$, are a set of auxiliary variables, in other words, variables whose value is known for all individuals in the population, particularly via the sampling frame. Deville and Tillé (2004) propose an efficient approach for approximately resolving this problem. These authors use algebraic and probability methods. By staying within the determinantal family, another possibility is to use optimization techniques and to find K^{opt} such that

$$K^{\text{opt}} = \underset{K}{\operatorname{argmin}} \sum_{q=1}^{Q} \alpha_q \operatorname{var}(\hat{t}_{x^q}), \text{ s.c.} \begin{cases} K = \bar{K}^\top \\ \operatorname{diag}(K) = \Pi \\ 0 \preceq K \preceq I_N, \end{cases} \tag{3.9}$$

where $\operatorname{var}(\hat{t}_{x^q})$ is given by (3.7), Π is fixed *a priori*, and $0 \preceq K \preceq I_N$ indicates that the eigenvalues of K are between 0 and 1. The coefficients α_q make it possible to manage the relative importance of the variables.

The choice $\alpha_q = 1 \big/ t_{x_q}^2$ leads to minimizing the sum of the squares of the variation coefficients of the various estimators. Resolving such problems involves non-linear positive semidefinite optimization. Loonis and Mary (2019) propose heuristics to find approximate solutions. We will in the following use optimization methods on manifolds (Absil et al., 2009; Boumal et al., 2014; Townsend et al., 2016).

3.4 A few examples of determinantal designs

3.4.1 Constructing one design from another

Below we present some general properties for constructing new determinantal designs from a given determinantal design. Let $\mathbb{S} \sim \mathrm{DSD}(K)$, then

1. the complement $\mathbb{S}^c$ of $\mathbb{S}$ in U is determinantal and has as distribution $\mathrm{DSD}(I_N - K)$,
2. the restriction $\mathbb{S} \cap D$ of $\mathbb{S}$ to the domain D included in U is determinantal of distribution $\mathrm{DSD}(K_{|D})$,
3. $\mathrm{DSD}(K)$ is stratified if and only if K is a block diagonal, up to a permutation of the rows and the columns,
4. if U_1 is a unitary matrix, the matrix $U_1 K \bar{U}_1^\top$ has the same eigenvalues as K. Therefore, there is a determinantal design associated with it. Among the unitary transformations, rotations prove useful for defining optimization heuristics (Loonis and Mary, 2019).

3.4.2 A family of fixed-size determinantal designs and any inclusion probabilities fixed *a priori*: The family P^Π

Loonis and Mary (2019) reiterate that the Poisson design is determinantal, whereas the simple random design is determinantal only in the cases $n = 1$ and $n = N - 1$. The authors provide examples of determinantal designs with constant inclusion probabilities. They construct a family of fixed-size determinantal designs for any inclusion probabilities fixed *a priori*. To the extent that it will reappear in various contexts in the following sections, we specifically present the properties of the designs associated with this family in this section.

Let Π be a vector of $]0, 1[^N$, such that $\sum_{k=1}^N \Pi_k = n, n \in \mathbb{N}^*$. Loonis and Mary (2019) mobilize a result from Kadison (2002) to exhibit an explicit formula of the coefficients of a real orthogonal projection matrix P^Π with diagonal Π (Table 3.1). The determinantal design associated with this matrix is of fixed size and has inclusion probabilities Π. The formula is based on the integers k_r $(r = 1, \ldots, n-1)$ such that $\sum_{k=1}^{k_r - 1} \Pi_k < r$ and $\sum_{k=1}^{k_r} \Pi_k \geq r$, the real numbers $\alpha_{k_r} = r - \sum_{k=1}^{k_r - 1} \Pi_k$ and $\gamma_r^{r'}$ such that

$$\gamma_r^{r'} = \sqrt{\prod_{j=r+1}^{r'} \frac{(\Pi_{k_j} - \alpha_{k_j})\, \alpha_{k_j}}{(1 - \alpha_{k_j})\,(1 - (\Pi_{k_j} - \alpha_{k_j}))}},$$

for $r < r'$, $\gamma_r^{r'} = 1$ otherwise.

Table 3.1
Coefficients P_{kl}^{Π} of P^{Π}: $k < l$.

	l	
k	$l = k_{r'+1}$	$k_{r'} < l < k_{r'+1}$
$k_r < k < k_{r+1}$	$\sqrt{\Pi_k}\sqrt{\frac{(1-\Pi_l)\alpha_l}{1-\alpha_l)}}\gamma_r^{r'}$	$\sqrt{\Pi_k \Pi_l}\gamma_r^{r'}$
$k = k_r$	$-\sqrt{\frac{(1-\Pi_l)\,\alpha_l}{1-\alpha_l}}\sqrt{\frac{(1-\Pi_k)\,(\Pi_k-\alpha_k)}{1-(\Pi_k-\alpha_k)}}\gamma_r^{r'}$	$-\sqrt{\Pi_l}\sqrt{\frac{(1-\Pi_k)\,(\Pi_k-\alpha_k)}{1-(\Pi_k-\alpha_k)}}\gamma_r^{r'}$

Knowledge of the coefficients P_{kl}^{Π} makes it possible to deduce information about the inclusion probabilities of order greater than 1 of design $\operatorname{DSD}(P^{\Pi})$:

1. If $\{k, l\}$ is an element of $]k_r, k_{r+1}[^2$ then $\pi_{kl} = 0$ because $P_{kl}^{\Pi} = \sqrt{\Pi_k \Pi_l}$,
2. If $|k - l|$ is *large* $P_{kl}^{\Pi} \simeq 0$ and $\pi_{kl} \simeq \Pi_k \Pi_l$ is maximal under the Sen-Yates-Grundy constraint,
3. If $j \in]k_{r-1}, k_r[$, $k = k_r$, $l \in]k_r, k_{r+1}[$, then $\pi_{jkl} = 0$,
4. If there are integers $r_1, \ldots, r_H$ such that $\sum_{k=1}^{k_{r_h}} \Pi_k = r_h$, $h = 1, \ldots, H$, then the design $\operatorname{DSD}(P^{\Pi})$ is stratified based on the strata $]k_{r_{h-1}}, k_{r_h}]$,
5. If n divides N and $\Pi_k = n/N$, then $\operatorname{DSD}(P^{\Pi})$ is a 1-per-stratum design. It selects an individual from each of the n groups of size N/n taken consecutively within the population: N/n first, and so on.

Example 3.1 illustrates some of these properties and also shows that the construction of P^{Π} depends on the order of the individuals.

Example 3.1 *Construction of* P^{Π} *and* $P^{\Pi^{\triangleright}}$ *for* $N = 7, n = 4$ *and* $\Pi = \left(\frac{1}{2}, \frac{3}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}\right)^{\top}$. $\Pi^{\triangleright}$ *is the vector of probabilities sorted from highest to lowest:* $\Pi^{\triangleright} = \left(\frac{4}{5}, \frac{3}{4}, \frac{3}{4}, \frac{3}{5}, \frac{1}{2}, \frac{2}{5}, \frac{1}{5}\right)^{\top}$.

$$(a):\ P^{\Pi} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & 0 & 0 & 0 & 0 \\ \frac{1}{2\sqrt{2}} & \frac{3}{4} & -\frac{1}{4} & 0 & 0 & 0 & 0 \\ \frac{1}{2\sqrt{2}} & -\frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{5} & \frac{\sqrt{2}}{5} & \frac{2}{5\sqrt{3}} & \frac{\sqrt{2}}{5\sqrt{3}} \\ 0 & 0 & 0 & \frac{\sqrt{2}}{5} & \frac{2}{5} & \frac{2\sqrt{2}}{5\sqrt{3}} & \frac{2}{5\sqrt{3}} \\ 0 & 0 & 0 & \frac{2}{5\sqrt{3}} & \frac{2\sqrt{2}}{5\sqrt{3}} & \frac{3}{5} & -\frac{\sqrt{2}}{5} \\ 0 & 0 & 0 & \frac{\sqrt{2}}{5\sqrt{3}} & \frac{2}{5\sqrt{3}} & -\frac{\sqrt{2}}{5} & \frac{4}{5} \end{pmatrix},$$

$$(b):\ P^{\Pi^{\rhd}} = \begin{pmatrix} \frac{4}{5} & \frac{1}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} & \frac{3}{5\sqrt{7}} & \frac{-1}{\sqrt{210}} & \frac{2}{15\sqrt{7}} & \frac{1}{15}\sqrt{\frac{2}{7}} \\ \frac{1}{2\sqrt{5}} & \frac{3}{4} & -\frac{1}{4} & \frac{-3}{2\sqrt{35}} & \frac{-1}{2\sqrt{42}} & \frac{-1}{3\sqrt{35}} & \frac{-1}{3\sqrt{70}} \\ \frac{1}{2\sqrt{5}} & -\frac{1}{4} & \frac{3}{4} & \frac{-3}{2\sqrt{35}} & \frac{-1}{2\sqrt{42}} & \frac{-1}{3\sqrt{35}} & \frac{-1}{3\sqrt{70}} \\ \frac{3}{5\sqrt{7}} & \frac{-3}{2\sqrt{35}} & \frac{-3}{2\sqrt{35}} & \frac{3}{5} & \frac{1}{\sqrt{30}} & \frac{2}{15} & \frac{\sqrt{2}}{15} \\ \frac{-1}{\sqrt{210}} & \frac{-1}{2\sqrt{42}} & \frac{-1}{2\sqrt{42}} & \frac{1}{\sqrt{30}} & \frac{1}{2} & -\sqrt{\frac{2}{15}} & \frac{-1}{\sqrt{15}} \\ \frac{2}{15\sqrt{7}} & \frac{-1}{3\sqrt{35}} & \frac{-1}{3\sqrt{35}} & \frac{2}{15} & -\sqrt{\frac{2}{15}} & \frac{2}{5} & \frac{\sqrt{2}}{5} \\ \frac{1}{15}\sqrt{\frac{2}{7}} & \frac{-1}{3\sqrt{70}} & \frac{-1}{3\sqrt{70}} & \frac{\sqrt{2}}{15} & \frac{-1}{\sqrt{15}} & \frac{\sqrt{2}}{5} & \frac{1}{5} \end{pmatrix}.$$

If the initial order of the population, $\{1, 2, 3, 4, 5, 6, 7\}$ *corresponds to vector* $\Pi,$ *then the determinantal design associated with* $P^{\Pi^{\rhd}}$ *applies to the population in the order* $(7, 2, 3, 6, 1, 5, 4)$ *(or* $(7, 3, 2, 6, 1, 5, 4)$*). In connection with the previous properties, we see that* $\Pi_1 + \Pi_2 + \Pi_3 = 2.$ *The determinantal design associated with* P^{Π} *is indeed stratified based on the strata* $(1, 2, 3)$ *and* $(4, 5, 6, 7).$ *For example, we have* $P_{25}^{\Pi} = 0$ *and* $\pi_{25} = \Pi_2\Pi_5 - P_{25}^{\Pi^2} = \Pi_2\Pi_5,$ *according to (3.2). The design associated with* $P^{\Pi^{\rhd}}$ *is not stratified because it is impossible to rearrange the rows and columns of this matrix to obtain a block diagonal matrix. The designs associated with* P^{Π} *and* $P^{\Pi^{\rhd}}$ *are different, even up to a permutation of the rows and the columns.*

We conclude this part with two new results pertaining to matrices P^{Π}. The first provides a basis of eigenvectors of P^{Π}. This basis can be used as input for the algorithm of Lavancier et al. (2015).

Algorithm 3.1 (Construction of $\Phi^{N^{\top}}$, orthonormal eigenbasis of P^{Π})

1. *Set* $k_0 = 0.$ *Define the integers* k_r $(r = 1, \ldots, n-1)$ *such that* $\sum_{k=1}^{k_r-1} \Pi_k < r$ *and* $\sum_{k=1}^{k_r} \Pi_k \geq r$ *and the real numbers* $\alpha_{k_r} = r - \sum_{k=1}^{k_r-1} \Pi_k,$
2. *For every* $k \in U,$ *calculate* s_k *and* c_k *such that*
   - *if there is* r *such that* $k = k_r,$ $s_{k_r} = \pm\sqrt{\frac{1-\Pi_{k_r}}{1-\alpha_{k_r}}}$
   - *otherwise for* $k_r < k < k_{r+1},$ $s_k = \pm\sqrt{\frac{\Pi_k}{r+1-\sum_{i=1}^{k-1}\Pi_i}}$
   - $c_k = \pm\sqrt{1-s_k^2},$ *for every* $k,$
3. *Construct* $\Phi^N,$ *matrix of size* $(n \times N)$ *of which all coefficients are nil except in* $(r+1, k_r+1)$ *(for* $r = 0, \ldots, n-1$*) where they equal 1. Let* φ_k *be the column* k *of* $\Phi^N,$
4. *Incrementally update the columns of* Φ^N *as follows, for* $k = 1, \ldots, N-1$*:*

   (a) calculate $C_1 = s_k\varphi_k - c_k\varphi_{k+1}$

(b) *calculate* $C_2 = c_k \varphi_k + s_k \varphi_{k+1}$

(c) *respectively replace* φ_k *and* φ_{k+1} *with* C_1 *and* C_2,

5. $K^\Pi = \Phi^{N^\intercal} \Phi^N$ *is such that* $K^\Pi = H_N \odot P^\Pi$, *where* H_N *is a real symmetric matrix whose diagonal terms equal* 1 *and non-diagonal terms equal* 1 *or* -1. $\Phi^{N^\intercal}$ *is an orthonormal eigenbasis of* K^Π,
6. *The matrix proposed by Loonis and Mary (2019) corresponds to the systematic choice of* $+$ *in the choices* $\pm$ *of step 2. Another choice leads to the same coefficients as those in Table 3.1, up to the sign. Regardless of the choices, all matrices* K^Π *have the same first- and second-order inclusion probabilities for a given vector* Π.

The new second result, below, shows that the matrix P^Π guarantees the lowest variability, among the determinantal designs, of the number of units sampled in the domains of the form $D^k = \{1, \ldots, k\}$ for $k = 1, \ldots, N$. This optimality property of family P^Π is one of those on which conjecture 6.1 will be based.

Theorem 3.1 *Let* Π *be a vector of* $]0,1[^N$ *such that* $\sum_{k=1}^N \Pi_k = n \in \mathbb{N}^*$ *and* P^Π, *the Loonis and Mary matrix (2019). Let* P_k^Π *be the main submatrix of order* k *of* P^Π, $k = 1, \ldots, N$, *and* $\lambda^k = \{\lambda_j^k\}_{j=1}^n$, *the vector of its strictly positive eigenvalues, supplemented where applicable by 0s to be of size* n, *then*

1. $\lambda^k = \{\lambda_j^k\}_{j=1}^n$ *is composed of the eigenvalue* 1 *with multiplicity* $\left\lfloor \sum_{s=1}^k \Pi_s \right\rfloor$, *the eigenvalue* $\left\{ \sum_{s=1}^k \Pi_s \right\}$ *with multiplicity 1 and the eigenvalue* 0 *with multiplicity* $n - \left\lfloor \sum_{s=1}^k \Pi_s \right\rfloor - 1$, *where* $\lfloor x \rfloor$ *and* $\{x\}$ *indicate the whole and decimal parts of* x.
2. *let* $\mathbb{S} \sim \mathrm{DSD}(K)$ *with inclusion probabilities given by* Π, D_k *be the domain* $D_k = \{1, \ldots, k\}$ *and* $\sharp \mathbb{S} \cap D_k$ *be the random number of individuals of* $\mathbb{S}$ *that are in* D_k, *then, for every* k.

$$\underset{\substack{K \\ \text{s.c.diag}(K)=\Pi}}{\mathrm{Min}} \ \mathrm{var}(\sharp \mathbb{S} \cap D_k) = \left\{ \sum_{s=1}^k \Pi_s \right\} \left(1 - \left\{ \sum_{s=1}^k \Pi_s \right\} \right).$$

For every k, *the minimum is reached, particularly for* $K = P^\Pi$.

4. Constructing all determinantal sampling designs with fixed first-order inclusion probabilities

4.1 Introduction

The method of Fickus et al. (2013) makes it possible to construct all Hermitian matrices with given diagonal and spectrum. For each of them, it provides an orthonormal eigenbasis. The method is described by the authors as non-trivial. The goal of this section is not to understand the whole process; we attempt to understand its broad strokes and identify its parameters and their constraints. We reformulate the latter to achieve parameterization involving mutually independent parameters. This section is technical and makes it possible to justify and understand the notations of theorems 5.1 and 5.2, which introduce a new property of matrices P^{Π} and a new family of fixed-diagonal projection matrices, the coefficients of which are explicitly known.

Given our topic, we limit ourselves to the case of contracting Hermitian matrices, whose diagonal is Π and spectrum, λ^N, is a known element of $]0,1]^M$, with $M \in \mathbb{N}^*$. The mobilization of λ^N can seem less intuitive to statisticians than that of Π. In practice, provided that $\sum_{k=1}^{N} \Pi_k = n \in \mathbb{N}^*$, the natural choice will be $M = n$ and $\lambda^N = 1^n$, which leads to orthogonal projection matrices and thus to fixed-size determinantal designs.

Remark 4.1 *An important point is that the algorithm directly constructs only matrices whose diagonal terms are ordered from highest to lowest. It will be able to directly construct matrix (b) of example 3.1, but not matrix (a) of the same example. For the latter, the algorithm will provide the matrix of example 4.1. The two matrices are the same up to a permutation* $\Sigma^{\intercal}$ *of their rows and columns. The properties of the associated designs are the same, but for a differently sorted population.*

Next, we assume that the population is sorted in a way that the inclusion probabilities are given by the vector $\Pi^{\triangleright}$. The matrices that we are attempting to create are notated $K_N^{M,\Pi^{\triangleright},\lambda^N}$.

Example 4.1 *The algorithm of Fickus et al. (2013) will be able to construct matrix* P^{Π} *of example 3.1 up to a permutation. It will provide matrix* $K^{\Pi^{\triangleright}}$ *such that*

$$K^{\Pi^{\triangleright}} = \Sigma^{\Pi^{\intercal}} P^{\Pi} \Sigma^{\Pi} = \begin{pmatrix} \frac{4}{5} & 0 & 0 & -\frac{\sqrt{2}}{5} & 0 & \frac{2}{5\sqrt{3}} & \frac{\sqrt{2}}{5\sqrt{3}} \\ 0 & \frac{3}{4} & -\frac{1}{4} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{3}{4} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{\sqrt{2}}{5} & 0 & 0 & \frac{3}{5} & 0 & \frac{2\sqrt{2}}{5\sqrt{3}} & \frac{2}{5\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{2}{5\sqrt{3}} & 0 & 0 & \frac{2\sqrt{2}}{5\sqrt{3}} & 0 & \frac{2}{5} & \frac{\sqrt{2}}{5} \\ \frac{\sqrt{2}}{5\sqrt{3}} & 0 & 0 & \frac{2}{5\sqrt{3}} & 0 & \frac{\sqrt{2}}{5} & \frac{1}{5} \end{pmatrix},$$

where $\Sigma^{\Pi^{\intercal}}$ *is a permutation matrix, which transforms* Π *into* $\Pi^{\triangleright}$. *If* $\mathrm{DSD}(P^{\Pi})$ *applies to the population indexed by* $(1,2,3,4,5,6,7)$, *then* $\mathrm{DSD}(K^{\Pi^{\triangleright}})$ *applies to the population in the order* $(7,2,3,6,1,5,4)$ (*or*

$(7,3,2,6,1,5,4)$ *because the value* $3/4$ *appears twice in the inclusion probabilities*). *The designs* $\mathrm{DSD}(P^{\Pi})$ *and* $\mathrm{DSD}(K^{\Pi^{\triangleright}})$, *applied to the same population, based on a tailored order, are equivalent.*

4.2 Sequential construction of main submatrices

Fickus et al. (2013) sequentially create all square matrices $K_k^{M,\Pi^{\triangleright},\lambda^N}$ of size $(k\times k)$, $k=1,\ldots,N$, which are main submatrices of order k of at least one matrix of type $K_N^{M,\Pi^{\triangleright},\lambda^N}$. The algorithm is put into the form

$$K_1^{M,\Pi^{\triangleright},\lambda^N}=(\Pi_1^{\triangleright}),\, K_k^{M,\Pi^{\triangleright},\lambda^N}=\begin{pmatrix} K_{k-1}^{M,\Pi^{\triangleright},\lambda^N} & b_k \\ \bar{b}_k^{\top} & \Pi_k^{\triangleright}\end{pmatrix},\ k=2,\ldots,N,$$

where b_k is a vector of size $k-1$. This vector b_k is constructed in such a way that the spectrum of $K_k^{M,\Pi^{\triangleright},\lambda^N}$ is equal to λ^k, which is a construction parameter. The value of this parameter, chosen by the statistician, must be compatible with a process that ultimately yields, in $k=N$, a matrix $K_N^{M,\Pi^{\triangleright},\lambda^N}$ with the initially desired diagonal $\Pi^{\triangleright}$ and spectrum λ^N.

At chosen λ_k, the authors show that various b_k are possible. They therefore introduce a second matrix parameter: V^{k-1}, of size $(M\times M)$, which reflects the variability of the b_k. The index $k-1$ refers to the fact that the structure of matrix V^{k-1} depends on λ^{k-1}. The nature of the constraints on the parameters is explained later. The way b_k is derived from λ_k and V^{k-1} is detailed in the appendix (Section A.5). Example 4.2 shows how the algorithm works.

Example 4.2 (How the algorithm works) *In example 4.1, we have* $M=4=n$, $N=7$. *Because matrix* $K^{\Pi^{\triangleright}}$ *is a projection matrix, it follows that* $\lambda^7=1^4$. *According to Fickus et al. (2013), there are two sequences of multidimensional parameters* $\{\lambda^k\}_{k=1}^6$ *and* $\{V^k\}_{k=1}^6$ *that we do not intend to explain here, and that, in six steps, lead to the matrix renamed* $K^{\Pi^{\triangleright}}=K_7^{4,\Pi^{\triangleright},1^4}$. *The first steps are*

$$\underbrace{\overset{\text{spectre}=\lambda^1}{\begin{pmatrix}\frac{4}{5}\end{pmatrix}}}_{K_1^{4,\Pi^{\triangleright},1^4}} \underset{V^1}{\overset{\lambda^2}{\rightarrow}} \underbrace{\overset{\text{spectre}=\lambda^2}{\begin{pmatrix}\frac{4}{5} & 0\\ 0 & \frac{3}{4}\end{pmatrix}}}_{K_2^{4,\Pi^{\triangleright},1^4}} \underset{V^2}{\overset{\lambda^3}{\rightarrow}} \underbrace{\overset{\text{spectre}=\lambda^3}{\begin{pmatrix}\frac{4}{5} & 0 & 0\\ 0 & \frac{3}{4} & -\frac{1}{4}\\ 0 & -\frac{1}{4} & \frac{3}{4}\end{pmatrix}}}_{K_3^{4,\Pi^{\triangleright},1^4}} \underset{V^3}{\overset{\lambda^4}{\rightarrow}} \underbrace{\overset{\text{spectre}=\lambda^4}{\begin{pmatrix}\frac{4}{5} & 0 & 0 & -\frac{\sqrt{2}}{5}\\ 0 & \frac{3}{4} & -\frac{1}{4} & 0\\ 0 & -\frac{1}{4} & \frac{3}{4} & 0\\ -\frac{\sqrt{2}}{5} & 0 & 0 & \frac{3}{5}\end{pmatrix}}}_{K_4^{4,\Pi^{\triangleright},1^4}} \underset{V^4}{\overset{\lambda^5}{\rightarrow}} \cdots$$

Moving from step $k-1$ *to* k *depends, in practice, on parameters* λ^k *and* V^{k-1}. *A different choice at one of the steps would lead to a different final matrix, but it would still be a projection matrix with diagonal* $\Pi^{\triangleright}$.

Remark 4.2 *In sampling theory, vector* λ^k *parameterizes the variance of the sample size in the domain* $D_k^{\triangleright}=\{1,\ldots,k\}$, *based on (3.5). The term* b_k, *a function of* λ_k *and* V^{k-1}, *is used in the expression of second-order inclusion probabilities and, ultimately, in that of the variance of the Horvitz-Thompson estimator based on (3.7).*

4.3 Reformulation of the constraints on vector parameters $\{\lambda^k\}_{k=1}^N$

For $1 \leq k \leq N-1$, every spectrum λ^k is subject to two types of constraints. As a spectrum of a main submatrix of a Hermitian matrix, λ^k is subject to the Cauchy interlacing constraints (2.1) with vectors λ^{k-1} and λ^{k+1}. As a spectrum of a Hermitian matrix with diagonal $(\Pi_1^{\triangleright}, \ldots, \Pi_k^{\triangleright})^{\intercal}$, λ^k satisfies the Schur-Horn theorem (2.2). Fickus et al. (2013) show that this set of constraints will be respected if and only if, for $k = 1, \ldots, N-1$, $j = 1, \ldots, M$, $\lambda_j^k \in [A_j^k, B_j^k]$, where the formulae explaining A_j^k and B_j^k are provided in the appendix (equations [A.1] and [A.2], Section A.3).

For $k = N$, Fickus et al. (2013) consider that λ^N is given exogenously. Noting that, for fixed $\Pi^{\triangleright}$, the only constraints imposed on λ^N are those of the Schur-Horn theorem, we show that λ^N will be the spectrum of a Hermitian matrix with diagonal $\Pi^{\triangleright}$ if and only if for $j = 1, \ldots, M$, $\lambda_j^N \in [A_j^N, B_j^N]$, where the formulae explaining A_j^N and B_j^N are provided in the appendix (proposition A.1, Section A.4).

A feature of A_j^k and B_j^k, for $k \leq N-1$ and $j > 0$, is that these two bounds are a function only of $\Pi^{\triangleright}$, λ^{k+1} and $\{\lambda_s^k\}_{s=1}^{j-1}$. Likewise, for $k = N$ and $j > 0$, A_j^N and B_j^N depend only on $\Pi^{\triangleright}$ and $\{\lambda_s^N\}_{s=1}^{j-1}$, with the convention $\lambda_0^k = 0$, for $k = 1, \ldots, N$. These remarks introduce a new parameterization of all the eigenvalues used in the method of Fickus et al. (2013).

Proposition 4.1

1. *By setting* $\lambda_j^k = A_j^k + \Omega_{jk}(B_j^k - A_j^k)$, *with* $\Omega_{jk} \in [0, 1]$, *all* $\{\{\lambda^k\}_{k=1}^{k=N}\}$ *of the spectra of the main submatrices of the contracting Hermitian matrices with diagonal* $\Pi^{\triangleright}$, *having* M *strictly positive eigenvalues, can be parameterized by all the matrices of size* $(M \times N)$, *whose coefficients have value in* $[0, 1]$. *Such a matrix is notated* Ω.
2. *According to this parameterization, the eigenvalue* λ_j^k *is a function only of* $\Pi^{\triangleright}$ *and*

$$\{\Omega_{1N}, \ldots, \Omega_{MN}, \Omega_{1(N-1)}, \ldots, \Omega_{M(N-1)}, \ldots, \Omega_{1(k+1)}, \ldots, \Omega_{M(k+1)}, \Omega_{1k}, \ldots, \Omega_{jk}\}.$$

3. *The value* $\Omega_{jk} = 0$ (*resp.* $\Omega_{jk} = 1$) *leads to the smallest* (*resp. largest*) *possible value* λ_j^k, *conditional on*

$$\{\lambda_{1N}, \ldots, \lambda_{MN}, \lambda_{1(N-1)}, \ldots, \lambda_{M(N-1)}, \ldots, \lambda_{1(k+1)}, \ldots, \lambda_{M(k+1)}, \lambda_{1k}, \ldots, \lambda_{(j-1)k}\}.$$

4. *The matrix* $\Omega = 0^{(M \times N)}$ (*resp.* $\Omega = 1^{(M \times N)}$) *leads to systematically retaining the smallest* (*resp. largest*) *possible eigenvalue* λ_j^k, *conditional on*

$$\{\lambda_{1N}, \ldots, \lambda_{MN}, \lambda_{1(N-1)}, \ldots, \lambda_{M(N-1)}, \ldots, \lambda_{1(k+1)}, \ldots, \lambda_{M(k+1)}, \lambda_{1k}, \ldots, \lambda_{(j-1)k}\}.$$

5. *If* $\sum_{k=1}^N \Pi_k^{\triangleright} = M = n \in \mathbb{N}^*$, *this parameterization will lead to a projection matrix for any matrix* $\Omega : \forall \Omega : \lambda^N(\Omega) = 1^n$.

An implementation of this parameterization appears in the appendix (Section A.3).

4.4 Reformulating the constraints on matrix parameters $\{V^k\}_{k=1}^{N-1}$

According to Fickus et al. (2013), the matrix V^k can be arbitrarily chosen from the matrices that satisfy the following constraints:

1. V^k is of size $(M \times M)$,
2. V^k is a block diagonal,
3. The number of blocks of V^k is equal to the number of distinct eigenvalues of λ^k,
4. A block's size is equal to the order of multiplicity of the corresponding eigenvalue,
5. Each block is a unitary matrix of some kind.

The parameterization of matrices $\{V^k\}_{k=1}^{N-1}$ occurs through the parameterization of the blocks that constitute them and, therefore, through the parameterization of unitary matrices. We did not identify any easily workable parameterization of this type of matrix in the literature (Dita, 1982, 1994; Jarlskog, 2005; Spengler, Huber, and Hiesmayr, 2010). We therefore propose a simplification. We assume that each element of λ^k is of multiplicity 1. As a result, matrix V^k is unitary diagonal. Its j^{th} diagonal term can be written $V_{jj}^k = \exp(2i\pi\rho_{jk})$, with ρ_{jk} element of $[0, 1]$. Thus, all matrices $\{V^k\}_{k=1}^{N-1}$ can be obtained from a matrix ρ of size $(M \times (N-1))$, whose coefficients are independent and have value in $[0, 1]$. The k^{th} column of ρ is used to construct the diagonal of V^k.

Example 4.3 *In example 4.2, the matrix V^2 is constructed from the spectrum λ^2 of matrix $K_2^{4,\Pi^{\triangleright},1^4}$, which has two strictly positive eigenvalues: $3/4$ and $4/5$. Because $M = 4$, λ^2 is supplemented with 0s so it is size 4 and $\lambda^2 = (0, 0, 3/4, 4/5)^{\mathsf{T}}$. V^2 has three blocks of respective sizes 2.1 and 1, corresponding to the multiplicities seen in λ^2. Thus, we have*

$$V^2 = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & e & 0 \\ 0 & 0 & 0 & f \end{pmatrix} \underset{\text{simplification}}{\rightarrow} V^2 = \begin{pmatrix} \exp(2i\pi\rho_{12}) & 0 & 0 & 0 \\ 0 & \exp(2i\pi\rho_{22}) & 0 & 0 \\ 0 & 0 & \exp(2i\pi\rho_{32}) & 0 \\ 0 & 0 & 0 & \exp(2i\pi\rho_{42}) \end{pmatrix}.$$

The simplification is inconsequential because the block associated with the eigenvalue 0 is not, in practice, used in the calculations of Fickus et al. (2013). The simplification could be consequential if, for a λ^k, there were an eigenvalue of multiplicity greater than 1 in the interval $]0, 1[$. If Ω is random, with each coefficient following a uniform distribution on $[0, 1]$, an intuition is that this event is of zero measurement. The values of the coefficients ρ_{32} and ρ_{42} that led to $K_3^{4,\Pi^{\triangleright},1^4}$ are not given. To find them, it would be necessary to use the reciprocal of Fickus et al. (2013), the principle of which is presented in the longer version of the article.

Remark 4.3 *In the algorithm of Fickus et al. (2013), the nature of $K_N^{M,\Pi^{\triangleright},\lambda^N}$, complex or real, depends on the choice of V^k. If at least one matrix V^k is complex, $K_N^{M,\Pi^{\triangleright},\lambda^N}$ will be complex. Conversely, $K_N^{M,\Pi^{\triangleright},\lambda^N}$ will be real if and only if all matrices V^k are real. Based on our parameterization, this case corresponds to a choice of ρ with value in $\{0; \frac{1}{2}; 1\}^{(M\times(N-1))}$.*

Remark 4.4 *In the following, the notation* $K_N^{\Pi^{\triangleright}}(M, \Omega, \rho)$ *will refer to a Hermitian matrix with diagonal* $\Pi^{\triangleright}$ *and constructed according to the method of Fickus et al. (2013) with the parameters* (M, Ω, ρ). *The reference to* λ^N *is omitted because this quantity is considered a parameter in the same way as the other spectra.*

5. Matrices $K_N^{\Pi^{\triangleright}}(M, \Omega, \rho)$ for specific values of Ω

5.1 Where we get back to $P^{\Pi^{\triangleright}}$

The construction of P^{Π}, in Loonis and Mary (2019), relies the sequential modification of the columns of a matrix of size $(M(=n) \times N)$ using rotation matrices, which are also unitary matrices (algorithm 3.1). This approach is consistent with the spirit of the Fickus et al. method. (2013). However, the formalism of the two methods does not appear to be directly equivalent. In the theorem below, we specify some links between the two approaches.

Theorem 5.1 *Let* $\Pi^{\triangleright}$ *be a vector of size* N, *such that* $0 < \Pi_k^{\triangleright} < 1$, $\sum_{k=1}^{N} \Pi_k^{\triangleright} = n \in \mathbb{N}^*$ *and* $\Pi_1^{\triangleright} \geq \ldots \Pi_k^{\triangleright} \geq \ldots \geq \Pi_N^{\triangleright}$. *For any value of the parameter* ρ,

$$K_N^{\Pi^{\triangleright}}(M = n, \Omega = 0^{(n \times N)}, \rho) \;=\; H^N(\rho) \odot P^{\Pi^{\triangleright}},$$

where $K_N^{\Pi^{\triangleright}}(M = n, \Omega = 0^{n \times N}, \rho)$ *is constructed based on the method of Fickus et al. (2013),* $P^{\Pi^{\triangleright}}$ *is the matrix defined by Loonis and Mary (2019) from the vector* $\Pi^{\triangleright}$ *and* $H^N(\rho)$ *is an Hermitian matrix of size* $(N \times N)$ *where all diagonal terms equal 1 and non-diagonal terms have a modulus of 1.*

For point (b) in example 3.1, this theorem indicates that the method of Fickus et al. (2013), with parameters $(M = n = 4, \Omega = 0^{(4 \times 10)})$, directly provides the matrix $P^{\Pi^{\triangleright}}$, up to H_N for all ρ. According to this theorem, the modulus of the off-diagonal terms of $K^{\Pi^{\triangleright}}(M = n, \Omega = 0^{n \times N}, \rho)$ does not depend on ρ because they equal $P_{kl}^{\Pi^{\triangleright}} H_{kl}^N(\rho) \overline{H_{kl}^N(\rho)} P_{kl}^{\Pi^{\triangleright}} = (P_{kl}^{\Pi^{\triangleright}})^2$. In terms of surveys, the variance of $\hat{t}_y$ is dependent on Δ_{kl} and thus on the modulus of the $K_{kl}^{\Pi^{\triangleright}}$ when the design is determinantal, according to (3.7). The implication is that there is nothing to expect from changing the parameter ρ to change the variance obtained from $P^{\Pi^{\triangleright}}$. This is a very special feature among Hermitian matrices, and among matrices of the family P^{Π}. When Ω is fixed, different from $0^{(M \times N)}$, it is empirically observed that the variance of the estimators is usually affected by variations of ρ.

Remark 5.1 *For a matrix* P^{Π}, *constructed from any vector* Π, *in other words, whose coefficients are not necessarily sorted, there is a matrix* Ω^{Π} *that leads to* $\Sigma^{\Pi^{\top}} P^{\Pi} \Sigma^{\Pi}$ *with the method of Fickus et al. (2013). There is no reason, a priori, for* Ω^{Π} *to be of the form* $0^{(n \times N)}$. *This is particularly true for the matrix* P^{Π} *of point* (a) *in example 3.1 and its reordered version given in example 4.1.*

5.2 Where we discover $Q^{\Pi^{\triangleright}}$

The theorem below shows that it is also possible to find an explicit formulation of $K^{\Pi^{\triangleright}}(M = n, \Omega = 1^{(n\times N)}, \rho),$ except for a few coefficients.

Theorem 5.2 *Let* $\Pi^{\triangleright}$ *be a vector of size* $N,$ *such that* $0 < \Pi_k^{\triangleright} < 1,$ $\sum_{k=1}^{N} \Pi_k^{\triangleright} = n \in \mathbb{N}^*$ *and* $\Pi_1^{\triangleright} \geq \ldots \Pi_k^{\triangleright} \geq \ldots \geq \Pi_N^{\triangleright}.$ *Let* $\{\{\lambda_j^k\}_{j=1}^n\}_{k=n}^N$ *be a sequence of reals such that* $\lambda^N = 1^n$ *and*

$$\begin{cases} \lambda_1^k &= \lambda_n^{k+1} - \Pi_{k+1}^{\triangleright} \quad k = n, \ldots, N-1, \\ \lambda_j^k &= \lambda_{j-1}^{k+1}, \qquad j = 2, \ldots, n, k = n, \ldots, N-1. \end{cases} \tag{5.1}$$

It is assumed that $\Pi^{\triangleright}$ *is such that one of the following two conditions is satisfied:*

- *C1: only 0 and 1 can appear multiple times in the* λ^k, $k = n, \ldots, N.$
- *C2:* $\Pi_k = n/N$, $k = 1, \ldots, N.$

For any value of the parameter $\rho,$ *we then have*

$$K^{\Pi^{\triangleright}}(M = n, \Omega = 1^{n\times N}, \rho) = H^N(\rho) \odot Q^{\Pi^{\triangleright}},$$

where $H^N(\rho)$ *is a Hermitian matrix of size* $(N \times N),$ *where all diagonal coefficients equal 1 and non-diagonal coefficients have a modulus of 1. The matrix* $Q^{\Pi^{\triangleright}}$ *is such that*

1. *for* $k \geq n,$ λ^k *yields the* n *largest eigenvalues of the main submatrix of order* k *of* $Q^{\Pi^{\triangleright}}$, *for which no more than* n *are strictly positive*
2. *for any* $(k, l) \in [n+1, N]^2,$
   (a) under $C1$: $Q_{kl}^{\Pi^{\triangleright}} = \sqrt{\Pi_k^{\triangleright} \Pi_l^{\triangleright}}\, \mathbb{1}(l \equiv k \bmod n)$
   (b) under $C2$: $Q_{kl}^{\Pi} = n/N \mathbb{1}\ (l \equiv k \ \textit{ou}\ k - N \equiv N - (l+1) \bmod 2n)$
3. *under* $C2,$ *if* n *divides* $N,$ *the formulae (5.1) and point 1 are true for* $0 < k < n,$ *the point* $2(b)$ *is true for* $k \leq n$ *or* $l \leq n.$

Figure 5.1 shows an example of matrix $Q^{\Pi^{\triangleright}}$ under $C1.$ Table 5.1 shows the organization of the spectra of the main submatrices of $Q^{\Pi^{\triangleright}}$ based on $\Pi^{\triangleright}.$ The equivalents for $C2,$ including in the case n divides $N,$ are provided in the appendix (Section A.7).

For $C1,$ the constraint on the multiplicities, set out in the theorem, implies that in each column there are different values except, possibly, for the eigenvalues 0 and 1. If that is not the case, inextricable calculation difficulties arise. When $k < n,$ the expression of the eigenvalues is more difficult to find. As a result, we do not arrive at an explicit formula of some coefficients of the matrices $Q^{\Pi^{\triangleright}}$ as a function of $\Pi^{\triangleright},$ except under $C2$ and n divides $N.$

Figure 5.1 $Q^{\Pi^{\triangleright}}$ under $C1$ for $n = 2,\ N = 8.$

$$
Q^{\Pi^{\triangleright}} = \begin{pmatrix}
\Pi_1^{\triangleright} & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \Pi_2^{\triangleright} & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \Pi_3^{\triangleright} & 0 & \sqrt{\Pi_3^{\triangleright}\Pi_5^{\triangleright}} & 0 & \sqrt{\Pi_3^{\triangleright}\Pi_7^{\triangleright}} & 0 \\
\bullet & \bullet & 0 & \Pi_4^{\triangleright} & 0 & \sqrt{\Pi_4^{\triangleright}\Pi_6^{\triangleright}} & 0 & \sqrt{\Pi_4^{\triangleright}\Pi_8^{\triangleright}} \\
\bullet & \bullet & \sqrt{\Pi_5^{\triangleright}\Pi_3^{\triangleright}} & 0 & \Pi_5^{\triangleright} & 0 & \sqrt{\Pi_5^{\triangleright}\Pi_7^{\triangleright}} & 0 \\
\bullet & \bullet & 0 & \sqrt{\Pi_6^{\triangleright}\Pi_4^{\triangleright}} & 0 & \Pi_6^{\triangleright} & 0 & \sqrt{\Pi_6^{\triangleright}\Pi_8^{\triangleright}} \\
\bullet & \bullet & \sqrt{\Pi_7^{\triangleright}\Pi_3^{\triangleright}} & 0 & \sqrt{\Pi_7^{\triangleright}\Pi_5^{\triangleright}} & 0 & \Pi_7^{\triangleright} & 0 \\
\bullet & \bullet & 0 & \sqrt{\Pi_8^{\triangleright}\Pi_4^{\triangleright}} & 0 & \sqrt{\Pi_8^{\triangleright}\Pi_6^{\triangleright}} & 0 & \Pi_8^{\triangleright}
\end{pmatrix}.
$$

The symbol $\bullet$ indicates that the explicit formula of the coefficient in question is unknown.

Table 5.1
Eigenvalues λ_j^k of the main submatrices of $Q^{\Pi^{\triangleright}}$, for $n = M = 5,\ k = N - 11, \ldots, N$.

k					
N-11	**N-10**	**N-9**	**N-8**	**N-7**	**N-6**
$1-\Pi^{\triangleright}_{N-10}-\Pi^{\triangleright}_{N-5}-\Pi^{\triangleright}_{N}$	$1-\Pi^{\triangleright}_{N-9}-\Pi^{\triangleright}_{N-4}$	$1-\Pi^{\triangleright}_{N-8}-\Pi^{\triangleright}_{N-3}$	$1-\Pi^{\triangleright}_{N-7}-\Pi^{\triangleright}_{N-2}$	$1-\Pi^{\triangleright}_{N-6}-\Pi^{\triangleright}_{N-1}$	$1-\Pi^{\triangleright}_{N-5}-\Pi^{\triangleright}_{N}$
$1-\Pi^{\triangleright}_{N-9}-\Pi^{\triangleright}_{N-4}$	$1-\Pi^{\triangleright}_{N-8}-\Pi^{\triangleright}_{N-3}$	$1-\Pi^{\triangleright}_{N-7}-\Pi^{\triangleright}_{N-2}$	$1-\Pi^{\triangleright}_{N-6}-\Pi^{\triangleright}_{N-1}$	$1-\Pi^{\triangleright}_{N-5}-\Pi^{\triangleright}_{N}$	$1-\Pi^{\triangleright}_{N-4}$
$1-\Pi^{\triangleright}_{N-8}-\Pi^{\triangleright}_{N-3}$	$1-\Pi^{\triangleright}_{N-7}-\Pi^{\triangleright}_{N-2}$	$1-\Pi^{\triangleright}_{N-6}-\Pi^{\triangleright}_{N-1}$	$1-\Pi^{\triangleright}_{N-5}-\Pi^{\triangleright}_{N}$	$1-\Pi^{\triangleright}_{N-4}$	$1-\Pi^{\triangleright}_{N-3}$
$1-\Pi^{\triangleright}_{N-7}-\Pi^{\triangleright}_{N-2}$	$1-\Pi^{\triangleright}_{N-6}-\Pi^{\triangleright}_{N-1}$	$1-\Pi^{\triangleright}_{N-5}-\Pi^{\triangleright}_{N}$	$1-\Pi^{\triangleright}_{N-4}$	$1-\Pi^{\triangleright}_{N-3}$	$1-\Pi^{\triangleright}_{N-2}$
$1-\Pi^{\triangleright}_{N-6}-\Pi^{\triangleright}_{N-1}$	$1-\Pi^{\triangleright}_{N-5}-\Pi^{\triangleright}_{N}$	$1-\Pi^{\triangleright}_{N-4}$	$1-\Pi^{\triangleright}_{N-3}$	$1-\Pi^{\triangleright}_{N-2}$	$1-\Pi^{\triangleright}_{N-1}$
N-5	**N-4**	**N-3**	**N-2**	**N-1**	**N**
$1-\Pi^{\triangleright}_{N-4}$	$1-\Pi^{\triangleright}_{N-3}$	$1-\Pi^{\triangleright}_{N-2}$	$1-\Pi^{\triangleright}_{N-1}$	$1-\Pi^{\triangleright}_{N}$	1
$1-\Pi^{\triangleright}_{N-3}$	$1-\Pi^{\triangleright}_{N-2}$	$1-\Pi^{\triangleright}_{N-1}$	$1-\Pi^{\triangleright}_{N}$	1	1
$1-\Pi^{\triangleright}_{N-2}$	$1-\Pi^{\triangleright}_{N-1}$	$1-\Pi^{\triangleright}_{N}$	1	1	1
$1-\Pi^{\triangleright}_{N-1}$	$1-\Pi^{\triangleright}_{N}$	1	1	1	1
$1-\Pi^{\triangleright}_{N}$	1	1	1	1	1

6. Applications

6.1 The data

The samples from Insee (Institut national de la statistique et des études économiques) household surveys are usually selected based on a two-stage design, when the collection method is face-to-face. In the first stage, primary units (PUs) are selected, from which the households to survey are then selected. The PUs consist of groups of the smallest nearby municipalities, with at least 2,000 principal residences. The sampling of PUs is stratified according to the 22 former metropolitan regions. Here we are interested in the sole first-stage selection of the PUs in the Provence-Alpes-Côte d'Azur region. They total $N = 148$ and are described by auxiliary variables yielding, for each PU, the total population by sex, age, the total amount of certain incomes or the total number of dwellings by category (vacant, secondary residence, etc.). Inclusion probabilities are proportional to the number of principal residences of the PUs.

6.2 A lower bound for the variance of the estimators in the determinantal situation?

When a sampling design is of a fixed size, the variance of the estimators, given by (3.6), takes the following particular form:

$$\operatorname{var}(\hat{t}_y) = -\frac{1}{2}\sum_k \sum_{l \neq k} \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l} \right)^2 \Delta_{kl}. \tag{6.1}$$

In the determinantal case, (3.7) becomes

$$\operatorname{var}(\hat{t}_y) = \frac{1}{2}\sum_k \sum_{l \neq k} \left(\frac{y_k}{K_{kk}} - \frac{y_l}{K_{ll}} \right)^2 |K_{kl}|^2. \tag{6.2}$$

The determinantal variance will be low if the K_{kl} modulus is low for the values of y_k / K_{kk} distant from those of y_l / K_{ll}. Assuming that the population is sorted according to $y_k / \Pi_k = y_k / K_{kk}$, this situation is observed for the matrix P^{Π} because we have, for $|k - l|$ *large*, $\pi_{kl} \simeq \Pi_k \Pi_l$ and therefore $|K_{kl}|^2 = \Pi_k \Pi_l - \pi_{kl} \simeq 0$ (see Section 3.4.2). Empirical results from Loonis and Mary (2019) suggest that matrices P^{Π}, constructed on a population sorted by y_k / Π_k, perform well in terms of variance. Theorems 3.1 and 5.1 also show that matrices P^{Π} have special properties among all fixed-diagonal projection matrices, even though, in the case of theorem 5.1, the scope is limited to diagonals of the form $\Pi^{\triangleright}$.

In this section, we characterize the performance of the designs associated with P^{Π} among all fixed-size determinantal designs and inclusion probabilities Π. We limit ourselves to the case of real kernels. For this, we conduct the following experiment:

- We create nested subpopulations of size $N = 20, 40, 60, 80, 100, 120$ associated with samples of size $n = 3, 6, 9, 12, 15, 18$.
- We consider three auxiliary variables, representing respectively the total amount of salaries, the number of tenants and the number of homeowners.
- For each variable, population and sample size,
  - we minimize (6.2), in K, among the real projection matrices with diagonal Π. We use optimization algorithms on manifolds (Absil et al., 2009; Boumal et al., 2014). We succinctly present the main principles of these procedures in the Appendix (Section A.8).
  - we calculate (6.2) for a matrix $K_{x^q} = P^{\Pi}_{x^q}$, where $P^{\Pi}_{x^q}$ was constructed on a population that was previously sorted by the values of x^q_k / Π_k.

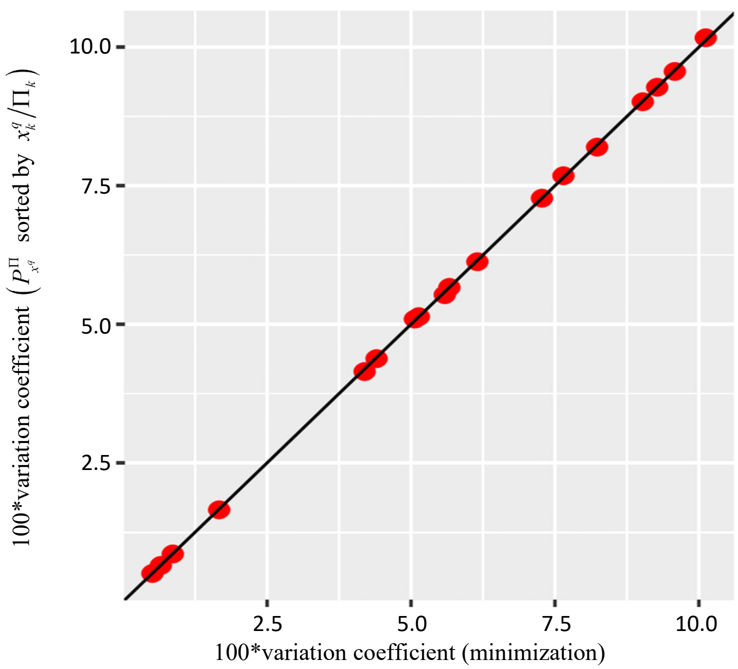
Figure 6.1 shows that the scatterplot obtained by crossing the previous two values, for each variable and sample size, is aligned with the first bisector. From this, we empirically deduce that the variance obtained with $P^{\Pi}_{x^q}$ corresponds to a minimum of the variance of $\hat{t}_{x^q}$, among determinantal designs. The proof of this result, or its refutation, seems to be out of reach. We therefore propose conjecture 6.1 below.

Conjecture 6.1 *Let* y *be a positive variable and* Π *be a vector of* $]0,1[^N$ *such that* $\sum_{k=1}^{N} \Pi_k = n \in \mathbb{N}^*$. *Let* $\mathcal{P}$ *be a determinantal sampling design with inclusion probabilities* $\pi_k = \Pi_k$, $k = 1, \ldots, N$ *and* $\hat{t}_y$ *be the Horvitz-Thompson estimator of the total* t_y *of* y *on that design. Without losing generality,* Π *and* y *are such that* $y_1/\Pi_1 \leq \ldots \leq y_k/\Pi_k \leq \ldots \leq y_N/\Pi_N$, *then*

$$\operatorname{var}(\hat{t}_y) \geq y^{\intercal} D_{\Pi}^{-1} [(I_N - P^{\Pi}) \odot P^{\Pi}] D_{\Pi}^{-1} y,$$

where $y^{\intercal} D_{\Pi}^{-1} [(I_N - P^{\Pi}) \odot P^{\Pi}] D_{\Pi}^{-1} y$ *is the variance of* $\hat{t}_y$ *obtained with design* $\operatorname{DSD}(P^{\Pi})$, *with* P^{Π} *matrix of Loonis and Mary (2019).*

Figure 6.1 Comparison, for three auxiliary variables and six population sizes, of the variation coefficients for variances obtained by minimizing (6.2) into K and for a matrix $P^{\Pi}_{x^q}$ sorted based on the values of x_k^q/Π_k.



In the classic configuration of a fixed-size design with constant inclusion probabilities, the previous conjecture makes it possible to find intuitive interpretations, when n divides N.

According to Section 3.4.2, if $\Pi_k = n/N$ and if n divides N, P^{Π} is a block diagonal with n identical blocks. The coefficients of each block of size $(N/n \times N/n)$ all equal n/N. The associated design involves selecting 1 individual in each of the n strata of size N/n. After sorting based on the values of y, the first stratum U_1 combines the first N/n individuals, the second stratum U_2 combines the next N/n individuals, and so on. Applying the variance formulae of the Horvitz-Thompson estimator for a stratified design (Särndal, Swensson, & Wretman, 2003) leads to the lower bound, in this case, to be equal to

$$y^\top D_\Pi^{-1} [(I_N - P^\Pi) \odot P^\Pi] \, D_\Pi^{-1} y \ = \ \frac{N^2}{n^2}\left(1 - \frac{n}{N}\right) \sum_{h=1}^{n} S_{yU_h}^2,$$

where $S_{yU_h}^2 = N/(N-n)\, \sigma_{yU_h}^2$ and $\sigma_{yU_h}^2$ is the variance of y in stratum U_h.

If the total you want to estimate is N_c, the number of individuals who have a given characteristic c, the underlying variable y equals 1 if the individual has the characteristic, otherwise it equals 0. After sorting this population based on this variable, in $n-1$ strata, the variance of y will be zero because y will still be 0 or 1. In a single stratum, there will be r_c 1 values, where r_c is the remainder of the Euclidean division of N_c by N/n, and $N/n - r_c$ 0 values. The lower bound is then $r_c\,(N/n - r_c)$.

7. Conclusion

In this article, we propose a workable parameterization of the kernels of determinantal sampling designs. We show that the family P^Π, originally constructed by Loonis and Mary (2019) with the sole objective of having an example of a fixed-sized determinantal design with given first-order inclusion probabilities, turns out to have *unexpected* statistical properties. For any $k = 1, \ldots, N$, it minimizes the variability of the sample size that falls within the fields $D^k = \{1, \ldots, k\}$ (theorem 3.1). It is directly associated with a very particular value of the parameter Ω in the construction of Fickus et al. (2013) (theorem 5.1). Finally, it appears empirically as a solution to a problem of minimizing the variance of the Horvitz-Thompson estimator (Section 6.2). This ubiquity leads to the conjecture that the lower bound of the variance of the Horvitz-Thompson estimator, among determinantal designs, is expressed as a function of P^Π (conjecture 6.1).

These results are obtained at the cost of theoretical or computational complexity, which can seem stimulating or bewildering. Bridges have been built between sampling theory and other fields, such as probabilities, algebra or semidefinite optimization. These bridges can be a source of new theoretical studies. However, the concepts and methods used are not part of the ordinary toolbox of survey statisticians of public institutes, even of the author initially. One of the challenges for developing studies around determinantal designs therefore lies in education, which can be facilitated through the provision of R programs that make the various results tangible. Another possibility is, in the future, to focus on more practical applications. As such, indirect sampling, the search for alternatives to the Horvitz-Thompson estimator, or spatial sampling appear to be promising areas.

Acknowledgements

The author thanks Alina Matei and two anonymous referees for the relevance of their comments, which led to substantial improvements in this version of the article. Any remaining errors, omissions, typos and blunders are the responsibility of the author.

Appendix

A.1 Determinantal designs and population order

Matters of population order often arise in the article. Below, we provide a summary of the various findings, as well as some elements to assist in their interpretation. We first reiterate that exchanging two rows (or two columns) of a square matrix leads to multiplying its determinant by -1.

1. **Determinantal designs are not dependent on population order**. We consider a population $U = \{1, \ldots, N\}$ and a determinantal design $\text{DSD}(K)$. Sorting U involves applying a permutation to it, the matrix of which is notated $\Sigma^\top$. The design $\text{DSD}(K)$ on U is then equivalent to the design $\text{DSD}(\Sigma^\top K \Sigma)$ on the sorted population. This property results from the fact that an even number of exchanges of rows and columns are applied to the matrix K. Therefore, the determinants of the extracted matrices are not changed.
2. **An algorithm for constructing particular kernels can depend on the order of the units**. This is true for matrices P^{Π}. Two different population orders will lead to two matrices P^{Π_1} and P^{Π_2}. In general, there will be no permutation matrix such that $P^{\Pi_1} = \Sigma^\top P^{\Pi_2} \Sigma$. The designs $\text{DSD}(P^{\Pi_1})$ and $\text{DSD}(P^{\Pi_2})$ are not equivalent. However, the previous remark applies to each of them.
3. **The algorithm by Fickus et al. (2013) is not dependent on the order**. Let K^{Π} be a contracting Hermitian matrix. For numerical reasons, the algorithm of Fickus et al. (2013) will construct the matrix $K^{\Pi^{\triangleright}} = \Sigma^\top K^{\Pi} \Sigma$, where $\Sigma^\top$ transforms Π into $\Pi^{\triangleright}$. To find the original matrix, simply take $K^{\Pi} = \Sigma K^{\Pi^{\triangleright}} \Sigma^\top$.
4. The connection between the algorithm of Fickus et al. (2013) and the family P^{Π} arises from the fact that $P^{\Pi^{\triangleright}} = K(n, 0^{(n \times N)}, \rho)$, up to matrix $H_N(\rho)$. Theorem 5.1 does not say that there is $P^{\Pi} = \Sigma K^{\Pi^{\triangleright}}(n, 0^{(n \times N)}, \rho)\, \Sigma^\top$, for any Π, even up to matrix H_N. However, there are indeed many unknown parameters such that $P^{\Pi} = \Sigma K^{\Pi^{\triangleright}}(n, \Omega^{\Pi}, \rho^{\Pi})\, \Sigma^\top$.
5. The lower bound of conjecture 6.1 is a matrix P^{Π} constructed on a population sorted in such a way that $y_1/\Pi_1 \leq \ldots \leq y_k/\Pi_k \leq \ldots \leq y_N/\Pi_N$.

A.2 Programs

The applications and simulations in this article were mostly obtained using SAS programs. We are gradually making their retranscription in R available, experimentally, on the Insee Lab at

https://github.com/InseeFrLab/Determinantal-Sampling-Designs.

In particular, these will gradually be made available:

1. The matrix P^{Π} and an eigenvector base constructed with algorithm 3.1

2. The matrix $K^{\Pi^{\triangleright}}(M, \Omega, \rho)$ constructed following the method described in Section A.5 and an eigenvector base as obtained at the very end of algorithm A.2
3. The selection algorithm of Lavancier et al. (2015) for selecting samples based on the determinantal designs associated with the kernels defined in the previous points.

There will be programs for points not covered here, such as the reciprocal of (Fickus et al., 2013) or the construction of *periodic* determinantal designs with constant inclusion probabilities, such as described in Loonis and Mary (2019). A longer version of this article, which presents the determinantal designs in more detail and provides the proofs of algorithm 3.1, theorems 3.1, 5.1, 5.2, and proposition A.1, is available at the same address.

A.3 Constructing the sequence $\{\lambda^k\}_{k=1}^{k=N-1}$ from Ω

According to Fickus et al. (2013), λ_j^k can be arbitrarily chosen in the interval $[A_j^k, B_j^k]$, with

$$A_j^k = \max\left\{\lambda_{j-1}^{k+1}, \sum_{s=1}^{j}\lambda_s^{k+1} - \sum_{s=1}^{j-1}\lambda_s^k - \Pi_{k+1}^{\triangleright}\right\}, \tag{A.1}$$

$$B_j^k = \min\left\{\lambda_j^{k+1}, \min_{i=j,\ldots,M}\left\{\sum_{s=M-i+1}^{k}\Pi_s^{\triangleright} - \sum_{s=j}^{i-1}\lambda_s^{k+1} - \sum_{s=1}^{j-1}\lambda_s^k\right\}\right\}, \tag{A.2}$$

for $k = 1, \ldots, N-1$, $j = 1, \ldots, M$ and where $\lambda_0^k = 0$ and, by convention, the sums over empty sets equal 0.

Here we describe how to construct the sequence $\{\lambda^k\}_{k=1}^{k=N-1}$ from the formulae (A.1) and (A.2) and proposition 4.1. The vector $\Pi^{\triangleright}$, of size $N = 10$, is such that $\Pi_k^{\triangleright} = \frac{6(11-k)}{10*11}$, $k = 1, \ldots, 10$ and $\sum_{k=1}^{N}\Pi_k^{\triangleright} = 3$. We choose $M = 7$ and set the value of λ^{10} at $\left(\frac{3}{10}, \frac{3}{10}, \frac{3}{10}\frac{3}{10}, \frac{6}{10}\frac{6}{10}\frac{6}{10}\right)^{\top}$. The matrix Ω is of size (7×10). We choose to have all its coefficients equal 0.5 for the first nine columns, which are the ones used in the calculations of $\{\lambda^k\}_{k=1}^{9}$. Table A.1 gives the final result for this sequence of eigenvalues. We then show how to arrive at the particular value $357/704$ for λ_5^5. According to proposition 4.1, this value results from the calculation $A_5^5 + \Omega_{5,5}(B_5^5 - A_5^5) = 0{,}5\,(A_5^5 + B_5^5)$ with $A_5^5 = 219/400$ and $B_5^5 = 909/1{,}760$. Table A.2 shows how $A_5^5 = 219/400$ is arrived at by applying formula (A.1). Table A.3 shows how to arrive at $B_5^5 = 909/1{,}760$ by applying formula (A.2). Since the logic is the same, we don't show how to apply the formulae from proposition (A.1), which would make it possible to construct λ^{10} from the 10th column of Ω. However, we specify that the values in this latter column do not necessarily equal 0.5.

Table A.1
Final values of λ_j^k.

	k									
j	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
1	0	0	0	0	0	0	$\frac{3}{22}$	$\frac{12}{55}$	$\frac{3}{11}$	$\frac{3}{10}$
2	0	0	0	0	0	$\frac{93}{440}$	$\frac{111}{440}$	$\frac{63}{220}$	$\frac{3}{10}$	$\frac{3}{10}$
3	0	0	0	0	$\frac{213}{880}$	$\frac{3}{11}$	$\frac{129}{440}$	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{3}{10}$
4	0	0	0	$\frac{927}{3,520}$	$\frac{501}{1,760}$	$\frac{261}{880}$	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{3}{10}$
5	0	0	$\frac{5,583}{14,080}$	$\frac{669}{1,408}$	$\frac{357}{704}$	$\frac{909}{1,760}$	$\frac{21}{40}$	$\frac{6}{11}$	$\frac{63}{110}$	$\frac{3}{5}$
6	0	$\frac{13,623}{28,160}$	$\frac{909}{1,760}$	$\frac{477}{880}$	$\frac{3,969}{7,040}$	$\frac{2,001}{3,520}$	$\frac{63}{110}$	$\frac{129}{220}$	$\frac{3}{5}$	$\frac{3}{5}$
7	$\frac{6}{11}$	$\frac{15,561}{28,160}$	$\frac{7,881}{14,080}$	$\frac{4,041}{7,040}$	$\frac{4,113}{7,040}$	$\frac{2,073}{3,520}$	$\frac{261}{440}$	$\frac{3}{5}$	$\frac{3}{5}$	$\frac{3}{5}$
Total	$\frac{6}{11}$	$\frac{57}{55}$	$\frac{81}{55}$	$\frac{102}{55}$	$\frac{24}{11}$	$\frac{27}{11}$	$\frac{147}{55}$	$\frac{156}{55}$	$\frac{162}{55}$	3
$\Pi_k^{\triangleright}$	$\frac{6}{11}$	$\frac{27}{55}$	$\frac{24}{55}$	$\frac{21}{55}$	$\frac{18}{55}$	$\frac{3}{11}$	$\frac{12}{55}$	$\frac{9}{55}$	$\frac{6}{55}$	$\frac{3}{55}$

Table A.2
How to obtain $\lambda_5^5 = \frac{1}{2}(A_5^5 + B_5^5) = \frac{357}{704}$ in Table A.1 calculating A_5^5 with formula (A.1).

λ_4^6 (a)	$\sum_{s=1}^{5}\lambda_s^6$ (b)	$\sum_{s=1}^{4}\lambda_s^5$ (c)	$\Pi_6^{\triangleright}$ (d)	b-c-d (e)	A_5^5 $\max(a, e)$
$\frac{261}{880}$	$\frac{93}{440}+\frac{3}{11}+\frac{261}{880}+\frac{909}{1,760}$	$\frac{213}{880}+\frac{501}{1,760}$	$\frac{3}{11}$	$\frac{219}{440}$	$\frac{219}{440}$

Table A.3
How to obtain $\lambda_5^5 = \frac{1}{2}(A_5^5 + B_5^5) = \frac{357}{704}$ in Table A.1 calculating B_5^5 with formula (A.2).

i	λ_5^6 (f)	$\sum_{s=8-i}^{5}\Pi_s^{\triangleright}$ (g)	$\sum_{s=5}^{i-1}\lambda_s^6$ (h)	g-h-c
5	$\frac{909}{1,760}$	$\frac{24}{55}+\frac{21}{55}+\frac{18}{55}$	0	$\frac{99}{160}$
6	$\frac{909}{1,760}$	$\frac{27}{55}+\frac{24}{55}+\frac{21}{55}+\frac{18}{55}$	$\frac{909}{1,760}$	$\frac{261}{440}$
7	$\frac{909}{1,760}$	$\frac{6}{11}+\frac{27}{55}+\frac{24}{55}+\frac{21}{55}+\frac{18}{55}$	$\frac{909}{1,760}+\frac{2,001}{3,520}$	$\frac{2,007}{3,520}$

$B_5^5 = \min\left(\frac{909}{1,760}, \frac{2,007}{3,520}, \frac{261}{440}, \frac{99}{160}\right) = \frac{909}{1,760}, \ \lambda_5^5 = \frac{1}{2}\left(\frac{219}{440} + \frac{909}{1,760}\right) = \frac{357}{704}$

A.4 Expression of the constraints on λ^N

Here we give an expression of the constraints applying to λ^N in the same form as those applying to $\{\lambda^k\}_{k=1}^{N-1}$. The difference is that λ^N is subject only to the Schur-Horn theorem.

Proposition A.1 *Let $\Pi^{\triangleright}$ be a vector of $]0,1[^N$ such that $\Pi_1^{\triangleright} \geq \ldots \geq \Pi_k^{\triangleright} \geq \ldots \geq \Pi_N^{\triangleright}$ and $\sum_{k=1}^N \Pi_k^{\triangleright} = \mu$. Let M be an integer such that $\mu \leq M \leq N$ and λ^N be a vector of $]0,1]^M$. There is a positive semidefinite Hermitian matrix with diagonal $\Pi^{\triangleright}$ and whose strictly positive eigenvalues are given by λ^N if and only if*

$$\begin{cases} \lambda_j^N \in [A_j^N, B_j^N],\ j = 1, \ldots, M \\ A_j^N = \max\left\{\lambda_{j-1}^N,\ \mu - \sum_{s=1}^{j-1} \lambda_s^N - (M-j)\right\} \\ B_j^N = \min\limits_{i=1,\ldots,M-j+1} \left\{\dfrac{\mu - \sum_{s=1}^{j-1} \lambda_s^N - \sum_{s=1}^{M-j-i+1} \Pi_s^{\triangleright}}{i}\right\}, \end{cases} \tag{A.3}$$

where $\lambda_0^N = 0$ and, by convention, the sums over empty sets equal 0.

A.5 The method by Fickus et al. (2013)

The general method is based on two subalgorithms for constructing the various quantities used in equations (A.7) and (A.8). For a given vector $\Pi^{\triangleright}$, it is assumed that values were set for the parameters $M, \Omega, \rho,$ which made it possible to construct the sequences $\{\lambda^k\}_{k=1}^N$ and $\{V^k\}_{k=1}^{N-1}$. In its original form, the method of Fickus et al. (2013) also depends on any unitary matrix U^1 of size $(M \times M)$. This matrix does not directly influence the final matrices. It influences only the choice of one of their eigenvector bases. $U^1 = I_M$ will be taken in practice.

Algorithm A.1 (Fickus et al. [2013])

For $k = 2$ to N,

1. *set $E_1^k = E_2^{k-1} = \{1, \ldots, M\}$*
2. *for $j = 1$ to M,*
   - *if $\lambda_j^{k-1} \in \{\lambda_{E_1^k}^k\}$, where $\{\lambda_{E_1^k}^k\}$ is the set of distinct values of the subvector of λ^k indexed by E_1^k, then*
     - *$E_2^{k-1} = E_2^{k-1} \setminus \{j\}$ ($\setminus$ meaning "deprived of" here)*
     - *$E_1^k = E_1^k \setminus \{j'\}$, where $j' = \min\{j'' \in E_1^k \mid \lambda_{j''}^k = \lambda_j^{k-1}\}$*
3. *construct $\overline{E}_2^{k-1}$ and $\overline{E}_1^k$, complementary in $\{1, \ldots, M\}$ of E_2^{k-1} and E_1^k.*

Algorithm A.2 (Fickus et al. [2013])

1. *Set $\varphi^1 = \sqrt{\Pi_1^{\triangleright}} u^1$ where u^1 is the first column of U^1.*
2. *for $k = 2$ to N,*

- *construct the sets* $E_1^k, E_2^{k-1}, \bar{E}_1^k, \bar{E}_2^{k-1}$, *such that* $E_1^k \cup \bar{E}_1^k = E_2^{k-1} \cup \bar{E}_2^{k-1} = \{1, \ldots, M\}$ *and* $E_1^k \cap \bar{E}_1^k = E_2^{k-1} \cap \bar{E}_2^{k-1} = \varnothing$ *based on the principles of algorithm A.1*
- *let* $r_k = \mathrm{card}(E_1^k) = \mathrm{card}(E_2^{k-1})$
- *let* $E_1'^k = (M+1)1^{r_k} - E_1^k$ *(resp.* $E_2'^{k-1}$*),* $\bar{E}_1'^k = (M+1)\,1^{M-r_k} - \bar{E}_1^k$ *(resp.* $\bar{E}_2'^{k-1}$*)*
- *let* σ_1^k *(resp.* σ_2^k*) be the unique permutation of* $\{1, \ldots, M\}$ *increasing in* $E_1'^k$ *and* $\bar{E}_1'^k$ *(resp.* $E_2'^{k-1}$ *and* $\bar{E}_2'^{k-1}$*) and such that* $\sigma_1^k(j) \in \{1, \ldots, r_k\}$ *for any* $j \in E_1'^k$ *(resp.* $\sigma_2^{k-1}(j) \in \{1, \ldots, r_k\}$ *for any* $j \in E_2^{k-1}$*). Let* Σ_1^k *(resp.* Σ_2^{k-1}*) be the associated permutation matrices*
- *let* $R^k = \mathcal{J}_{r_k}(\lambda_{|E_2^{k-1}}^{k-1}, \lambda_{|E_1^k}^k)$ *be a matrix of size* $(r_k \times 2)$*, where* $\lambda_{|E_2^{k-1}}^{k-1}$ *(resp.* $\lambda_{|E_1^k}^k$*), the vector extracted from* λ^{k-1} *(resp.* λ^k*) whose rows are indexed by* E_2^{k-1} *(resp.* E_1^k*), and* $\mathcal{J}_{r_k}$ *refers to the anti-diagonal matrix of size* r_k*:*

$$\mathcal{J}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

- *let* v^k, w^k *be two vectors of size* r_k *and* W^k *be a matrix* $r_k \times r_k$ *such that*

$$v_i^k = \sqrt{-\frac{\prod_{i'=1}^{r_k}(R_{i1}^k - R_{i'2}^k)}{\prod_{\substack{i'=1 \\ i' \neq i}}^{i'=r_k}(R_{i1}^k - R_{i'1}^k)}} \tag{A.4}$$

$$w_i^k = \sqrt{\frac{\prod_{i'=1}^{r_k}(R_{i2}^k - R_{i'1}^k)}{\prod_{\substack{i'=1 \\ i' \neq i}}^{i'=r_k}(R_{i2}^k - R_{i'2}^k)}}; \tag{A.5}$$

$$W^k = (e_{r_k} \odot R_{.2}^{k^\intercal} - e_{r_k}^\intercal \odot R_{.1}^k)^{[-1]} \odot (v^k w^{k^\intercal}); \tag{A.6}$$

where $[-1]$ *refers to the matrix inverse in the sense of the Hadamard product*

- *set*

$$\varphi^k = U^{k-1} V^{k-1} \Sigma_2^{k-1^\intercal} \begin{bmatrix} v^k \\ 0_{M-r_k} \end{bmatrix}, \tag{A.7}$$

$$U^k = U^{k-1} V^{k-1} \Sigma_2^{k-1^\intercal} \begin{bmatrix} W^k & 0_{(r_k, M-r_k)} \\ 0_{(M-r_k, r_k)} & I_{(M-r_k, M-r_k)} \end{bmatrix} \Sigma_1^k, \tag{A.8}$$

and construct Φ^k *the matrix whose columns are the* $\{\varphi^s\}_{s=1}^{s=k}$.

The theorem of Fickus et al. (2013) states that, for $k = 2, \ldots, N$,

- U^k is a basis of orthonormalized eigenvectors of $\Phi^k \bar{\Phi}^{k^\intercal}$ whose spectrum is λ^k
- $K_k^{\Pi^\triangleright} = \bar{\Phi}^{k^\intercal} \Phi^k$ is a positive semidefinite Hermitian matrix whose diagonal is $(\Pi_1^\triangleright, \ldots, \Pi_k^\triangleright)^\intercal$ and whose strictly positive eigenvalues are given by λ^k.

In practice, we are interested in the result at the end of the process, in other words, for $k = N$. In this case, the matrix $K_N^{\Pi^{\triangleright}}$ has the desired properties. The authors demonstrate that, by varying the different parameters, all the matrices with a fixed diagonal and spectrum λ^N are obtained. Conversely, any matrix of this type can be constructed based on this procedure.

It is inferred from these results that a basis of orthonormalized eigenvectors of $K_N^{\Pi^{\triangleright}}$ is given by $\bar{\Phi}^{N^{\intercal}} U^N D_{\lambda^{\triangleright N}}^{-\frac{1}{2}}$. This basis can be used directly as input to the selection algorithm of Lavancier et al. (2015) in the case of projection matrices, and with a slight adaptation for determinantal designs of random size (Loonis and Mary 2019).

A.6 Constructing the elements of algorithms A.1 and A.2

We show here how to obtain the various quantities of algorithms A.1 and A.2 from the example in Table A.1 for $k = 9$ and $M = 7$.

$$E_2^8 = \{1, 2, 5, 6\} \quad E_1^9 = \{1, 4, 5, 7\}$$
$$E_2'^8 = \{2, 3, 6, 7\} \quad E_1'^9 = \{1, 3, 4, 7\}$$

$$k = 9, r_9 = 4, R^9 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{12}{55} & \frac{3}{11} \\ \frac{63}{220} & \frac{3}{10} \\ \frac{6}{11} & \frac{63}{110} \\ \frac{129}{220} & \frac{3}{5} \end{pmatrix} = \begin{pmatrix} \frac{129}{220} & \frac{3}{5} \\ \frac{6}{11} & \frac{63}{110} \\ \frac{63}{220} & \frac{3}{10} \\ \frac{12}{55} & \frac{3}{11} \end{pmatrix}$$

$$v_1^9 = \sqrt{-\frac{\left(\frac{3}{5} - \frac{129}{220}\right)\left(\frac{63}{110} - \frac{129}{220}\right)\left(\frac{3}{10} - \frac{129}{220}\right)\left(\frac{3}{11} - \frac{129}{220}\right)}{\left(\frac{6}{11} - \frac{129}{220}\right)\left(\frac{63}{220} - \frac{129}{220}\right)\left(\frac{12}{55} - \frac{129}{220}\right)}}$$

$$\bar{E}_2'^8 = \{1, 4, 5\}, \bar{E}_1'^9 = \{2, 5, 6\}$$

$$\sigma_2^8 \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 2 \\ 6 \\ 7 \\ 3 \\ 4 \end{pmatrix}, \sigma_1^9 \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 2 \\ 3 \\ 6 \\ 7 \\ 4 \end{pmatrix}.$$

It should be noted that σ_2^8 (resp. σ_1^9) is indeed increasing in $E_2'^8$ (resp. $E_1'^9$). In addition, for any $j \in E_2'^8$ (resp. $j \in E_1'^9$), we indeed have $\sigma_2^8(j) \in \{1, \ldots, r_k\}$ (resp. $\sigma_1^9(j) \in \{1, \ldots, r_k\}$). For a given permutation σ, the associated matrix is such that

$$\Sigma_{ij} = \begin{cases} 1 & \text{if } i = \sigma(j) \\ 0 & \text{otherwise} \end{cases}$$

$$\Sigma_2^8 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \Sigma_1^9 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

A.7 Matrices $Q^{\Pi^{\triangleright}}$ for $\Pi^{\triangleright} = n/N$

Table A.4
Eigenvalues λ_j^k of the main sub-matrices of $Q^{\Pi^{\triangleright}}$, for $\Pi_k = n/N$ and n does not divide N: Example of $n = 5$ and $N = 12$.

						k					
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
0	0	0	0	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	1
0	0	0	•	$1-\frac{2n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	1	1
0	0	•	•	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	1	1	1
0	•	•	•	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	1	1	1	1
$\frac{n}{N}$	•	•	•	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	1	1	1	1	1

Table A.5
Eigenvalues of the main submatrices of $Q^{\Pi^{\triangleright}}$, for $\Pi_k = n/N$ and n divides N: Example of $n = 4$ and $N = 12$.

						k					
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
0	0	0	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	1
0	0	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	1	1
0	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	1	1	1
$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{2n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	$1-\frac{n}{N}$	1	1	1	1

Figure A.1 $Q^{\Pi^{\triangleright}}$ under $C2$.

(a) General case: $n = 3, N = 10,\ \Pi_k = n/N$

$$Q^{\Pi^{\triangleright}} = \begin{pmatrix} n/N & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & n/N & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & n/N & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & n/N & 0 & 0 & n/N & 0 \\ \bullet & \bullet & \bullet & 0 & n/N & n/N & 0 & 0 \\ \bullet & \bullet & \bullet & 0 & n/N & n/N & 0 & 0 \\ \bullet & \bullet & \bullet & n/N & 0 & 0 & n/N & 0 \\ \bullet & \bullet & \bullet & 0 & 0 & 0 & 0 & n/N \end{pmatrix}$$

(b) Specific case: n divides N: $n = 2, N = 10,\ \Pi_k = n/N$

$$Q^{\Pi^{\triangleright}} = \begin{pmatrix} n/N & 0 & 0 & n/N & n/N & 0 & 0 & n/N \\ 0 & n/N & n/N & 0 & 0 & n/N & n/N & 0 \\ 0 & n/N & n/N & 0 & 0 & n/N & n/N & 0 \\ n/N & 0 & 0 & n/N & n/N & 0 & 0 & n/N \\ n/N & 0 & 0 & n/N & n/N & 0 & 0 & n/N \\ 0 & n/N & n/N & 0 & 0 & n/N & n/N & 0 \\ 0 & n/N & n/N & 0 & 0 & n/N & n/N & 0 \\ n/N & 0 & 0 & n/N & n/N & 0 & 0 & n/N \end{pmatrix}.$$

The symbol $\bullet$ indicates that the explicit formula of the coefficient in question is unknown.

A.8 Optimization among real fixed-diagonal projection matrices

Among real matrices of size $(N \times N)$, we attempt to solve a problem of the following type:

$$\underset{K}{\text{Min}} \sum_{q=1}^{Q} \frac{x^{q^\top} (I_N \odot K)^{-1} [K \odot (I_N - K)] (I_N \odot K)^{-1} x^q}{t_{x^q}^2} \quad \text{s.c.} \begin{cases} K = K^\top \\ \text{diag}(K) = \Pi \\ K^2 = K. \end{cases}$$

K is a projection matrix, so there is a basis V of orthonormalized vectors of size $(N \times n)$ such that $K = VV^\top$ and $V^\top V = I_n$. It is possible to rewrite the problem in the following penalized form:

$$\underset{\substack{V \\ V^\top V = I_n}}{\text{Min}} \sum_{q=1}^{Q} \frac{x^{q^\top} (I_N \odot VV^\top)^{-1} [VV^\top \odot (I_N - \overline{VV}^\top)] (I_N \odot VV^\top)^{-1} x^q}{t_{x^q}^2} + r\,\text{Trace}((VV^\top - D_\Pi) \odot (VV^\top - D_\Pi)).$$

The objective of penalization via the *trace* function is for the diagonal of the optimal matrix to indeed be equal to Π. The set of real matrices V such that $V^\top V = I_n$ is the Grassmannian manifold. The problem becomes a penalized optimization problem on varieties for which there are algorithms and powerful problem-solving software, as long as the population size remains reasonable, around a few hundred (Absil et al., 2009; Boumal et al., 2014; Townsend et al., 2016). For the penalization parameter r, we set $r = 10^i$, for $i = 0, \ldots, 10$ and retained the one that minimizes, to the optimum set at i, the function $10^i \operatorname{Trace}((VV^\top - D_\Pi) \odot (VV^\top - D_\Pi))$.

References

Absil, P.-A., Mahony, R. and Sepulchre, R. (2009). Optimization algorithms on matrix manifolds. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press.

Boumal, N., Mishra, B., Absil, P.-A. and Sepulchre, R. (2014). Manopt, a matlab toolbox for optimization on manifolds. *The Journal of Machine Learning Research*, 15(1), 1455-1459.

Brändén, P., and Jonasson, J. (2012). Negative dependence in sampling. *Scandinavian Journal of Statistics*, 39(4), 830-838.

Chauvet, G., and Tillé, Y. (2006). A fast algorithm for balanced sampling. *Computational Statistics*, 21(1), 53-62.

Deville, J.-C., and Tillé, Y. (2004). Efficient balanced sampling: The cube method. *Biometrika*, 91(4), 893-912.

Dhillon, I.S., Heath Jr, R.W., Sustik, M.A. and Tropp, J.A. (2005). Generalized finite algorithms for constructing hermitian matrices with prescribed diagonal and spectrum. *SIAM Journal on Matrix Analysis and Applications*, 27(1), 61-71.

Dita, P. (1982). Parametrisation of unitary matrices. *Journal of Physics A: Mathematical and General*, 15(11), 3465.

Dita, P. (1994). On the parametrisation of unitary matrices by the moduli of their elements. *Communications in Mathematical Physics*, 159(3), 581-591.

Fickus, M., Mixon, D.G. and Poteet, M.J. (2013). Constructing finite frames with a given spectrum. *Finite Frames*, Springer, 55-107.

Horn, A. (1954). Doubly stochastic matrices and the diagonal of a rotation matrix. *American Journal of Mathematics*, 76(3), 620-630.

Horn, R.A., and Johnson, C.R. (1991). *Topics in Matrix Analysis*. Cambridge Univ. Press.

Horvitz, D.G., and Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260), 663-685.

Hough, J.B., Krishnapur, M., Peres, Y. and Virág, B. (2006). Determinantal processes and independence. *Probability Surveys*, 3, 206-229.

Jarlskog, C. (2005). A recursive parametrization of unitary matrices. *Journal of Mathematical Physics*, 46(10), 103 508.

Kadison, R.V. (2002). The pythagorean theorem: I. the finite case. *Proceedings of the National Academy of Sciences*, 99(7), 4178-4184.

Lavancier, F., Møller, J. and Rubak, E. (2015). Determinantal point process models and statistical inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(4), 853-877.

Leuenberger, M., Eustache, E., Jauslin, R. and Tillé, Y. (2022). Balancing a sample almost perfectly. *Statistics & Probability Letters*, 180, 109 229.

Loonis, V., and Mary, X. (2019). Determinantal sampling designs. *Journal of Statistical Planning and Inference*, 199, 60-88.

Lyons, R. (2003). Determinantal probability measures. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 98, 167-212.

Macchi, O. (1975). The coincidence approach to stochastic point processes. *Advances in Applied Probability*, 83-122.

Pemantle, R., and Peres, Y. (2014). Concentration of lipschitz functionals of determinantal and other strong rayleigh measures. *Combinatorics, Probability and Computing*, 23(01), 140-160.

Särndal, C.-E., Swensson, B. and Wretman, J. (2003). *Model Assisted Survey Sampling*. Springer Science & Business Media.

Scardicchio, A., Zachary, C.E. and Torquato, S. (2009). Statistical properties of determinantal point processes in high-dimensional euclidean spaces. *Physical Review E*, 79(4), 041 108.

Schur, J. (1911). Bemerkungen zur theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. *Journal für die reine und angewandte Mathematik*, 140, 1-28.

Sen, A.R. (1953). On the estimate of the variance in sampling with varying probabilities. *Journal of the Indian Society of Agricultural Statistics*, 5(1194), 127.

Soshnikov, A. (2000). Determinantal random point fields. *Russian Mathematical Surveys*, 55(5), 923.

Soshnikov, A. (2002). Gaussian limit for determinantal random point fields. *Annals of Probability*, 171-187.

Spengler, C., Huber, M. and Hiesmayr, B.C. (2010). A composite parameterization of unitary groups, density matrices and subspaces. *Journal of Physics A: Mathematical and Theoretical*, 43(38), 385 306.

Tillé, Y. (2001). Théorie des sondages. *Dunod,* Paris, 13, 23-26.

Townsend, J., Koep, N. and Weichwald, S. (2016). Pymanopt: A python toolbox for optimization on manifolds using automatic differentiation. *arXiv preprint arXiv:1603.03236*.

Yates, F., and Grundy, P.M. (1953). Selection without replacement from within strata with probability proportional to size. *Journal of the Royal Statistical Society: Series B (Methodological)*, 15(2), 253-261.

Yuan, M., Su, C. and Hu, T. (2003). A central limit theorem for random fields of negatively associated processes. *Journal of Theoretical Probability*, 16(2), 309-323.