Supplementary material for manuscript "Inferring the finest pattern of mutual independence from data"

G. Marrelec and A. Giron

Contents

1 The lattice of partitions

We here quickly review the main features of $\Omega(N)$ as a lattice based on Birkhoff (1973, Chap. 1), Aigner (1979, Chap. I, §2.B) and Knuth (2004).

On the set of partitions $\Omega(N)$, one can define the following relation: π is a finer partition than (or a refinement of) ω, denoted $\pi \leqslant \omega$, if every block of π is contained in a block of ω. If π is finer than ω, then ω is said to be coarser than π. The relation "$\leqslant$" has 3 properties:

- reflexivity: for all π, $\pi \leqslant \pi$;
- antisymmetry: if $\pi \leqslant \omega$ and $\omega \leqslant \pi$ then $\pi = \omega$;
- transitivity: if $\pi \leqslant \omega$ and $\omega \leqslant \rho$, then $\pi \leqslant \rho$.

$\Omega(N)$ is therefore called a partially ordered set, or poset.

If $\pi \leqslant \omega$ and $\pi \neq \omega$, we write $\pi < \omega$. If $\pi < \omega$ and there does not exist a ρ such that $\pi < \rho < \omega$, we say that ω covers π, written $\pi \prec \omega$. ω is then a direct successor to π in the hierarchy induced by $\leqslant$. In $\Omega(N)$, ω covers π if ω is obtained by merging two blocks of π.

In $\Omega(N)$, the partition $O = 1 \mid 2 \mid \cdots \mid N$ is such that $O \leqslant \pi$ for all $\pi \in \Omega(N)$. It is called the least element, or bottom, of $\Omega(N)$. Dually, $I = 12 \dots N$ is such that $\pi \leqslant I$ for all $\pi \in \Omega(N)$. It is called the greatest element, or top, of $\Omega(N)$. O and I are called universal bounds of $\Omega(N)$ as $O \leqslant \pi \leqslant I$ for all $\pi \in \Omega(N)$. The elements that cover the bottom are called the atoms. In $\Omega(N)$, there are $N(N-1)/2$ atoms, each one with $N-2$ blocks of size 1 and 1 block of size 2, e.g., $12 \mid 3 \mid \cdots \mid N$ or $1 \mid 23 \mid 4 \mid \cdots \mid N$.

Given a pair of partitions π and ω, their upper bound is defined as the set of all ρ such that $\pi \leqslant \rho$ and $\omega \leqslant \rho$. In $\Omega(N)$, a unique least upper bound exists, called the join and written $\pi \vee \omega$. Similarly, we can define the lower bound as the set of all ρ such that $\rho \leqslant \pi$ and $\rho \leqslant \omega$. In $\Omega(N)$, a unique greatest lower bound exists, called the meet and written $\pi \wedge \omega$. In particular, we have for the top partition I, bottom partition O and all $\pi \in \Omega(N)$,

$$O \wedge \pi = O, \quad O \vee \pi = \pi, \quad I \wedge \pi = \pi, \quad \text{and} \quad I \vee \pi = I. \tag{1}$$

Besides, the consistency relationship expresses the relationship between the ordering and the join and meet operators:

$$\pi \leqslant \omega \qquad \Leftrightarrow \qquad \left\{ \begin{array}{rcl} \pi \wedge \omega & = & \pi \\ \pi \vee \omega & = & \omega. \end{array} \right. \tag{2}$$

Practically, the meet $\pi \wedge \omega$ of $\pi = a_1 \mid \cdots \mid a_k$ and $\omega = b_1 \mid \cdots \mid b_l$ is the partition that has blocks $a_i \cap b_j$ for any $1 \leq i \leq k$ and $1 \leq j \leq l$ such that $a_i \cap b_j \neq \emptyset$. In this definition, "$\cap$" stands for the usual set intersection. In other words, two elements of N belong to the same block of $\pi \wedge \omega$ if and only if they do for both π and ω.

The concept of covering can be used to provide a graphical representation of $\Omega(N)$. If $\pi \leqslant \omega$, then ω is drawn higher than π in the diagram. If $\pi \prec \omega$ then π and ω are connected through a line. See for instance Figure 1 for a representation of $\Omega([3])$ and $\Omega([4])$.

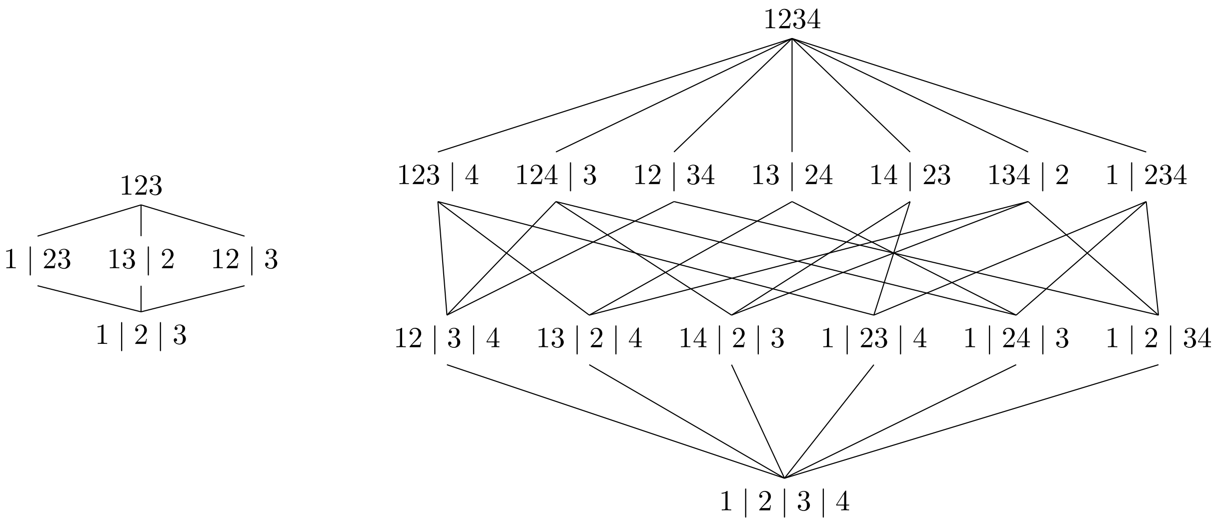

Figure 1: **Examples of partition lattices.** Representation of $\Omega([3])$ (left) and $\Omega([4])$ (right).

Since $\Omega(N)$ is a poset for which the join ($\vee$) and meet ($\wedge$) exist for every pair of partitions, it is called a lattice. In lattices, the join and meet have the following properties:

- idempotency: $\pi \wedge \pi = \pi$ and $\pi \vee \pi = \pi$;
- commutativity: $\pi \wedge \omega = \omega \wedge \pi$ and $\pi \wedge \omega = \omega \vee \pi$;
- associativity: $(\pi \wedge \omega) \wedge \rho = \pi \wedge (\omega \wedge \rho)$ and $(\pi \vee \omega) \vee \rho = \pi \vee (\omega \vee \rho)$;
- absorption: $\pi \wedge (\pi \vee \omega) = \pi \vee (\pi \wedge \omega) = \pi$.

We also have the following properties:

$$\omega \leqslant \rho \qquad \Rightarrow \qquad \begin{cases} \pi \wedge \omega & \leqslant \quad \pi \wedge \rho \\ \pi \vee \omega & \leqslant \quad \pi \vee \rho. \end{cases} \tag{3}$$

Note that $\Omega(N)$ is not distributive.

2 Simulation study: correlation thresholding

We independently tested each sample correlation coefficient r under the null hypothesis that it is equal to 0 using the fact that, under this assumption,

$$\frac{r\sqrt{k-2}}{\sqrt{1-r^2}} \tag{4}$$

has the t-distribution with $k-2$ degrees of freedom (Anderson, 1958, §4.2.1). We then performed multiple comparison using the FDR approach as detailed in §3.3.2 of the manuscript (Benjamini and Hochberg, 1995). The finest pattern of mutual independence was finally obtained by identifying each block with a maximally connected component. Results regarding the ratio of correct inference at a significance level of $\alpha = 0.1$ are summarized in Figure 2.

Note that we did not compute the other quantities presented in Figure 2 of the manuscript, such as AUC, sensitivity and specificity. While we could compute them, their use as a way to compare approaches would be very limited. Indeed, in the case of our approach, these quantities were calculated using the rates of true/false positive/negatives among the set $\Omega_2(N)$ of dichotomic independence relationships, which has a cardinality of $2^{n-1}-1$. By contrast, in the case of correlation threshlding, they could only be calculated using the rates of true/false positive/negatives among the set of correlation coefficients, whic has a cardinality of $n(n-1)/2$.

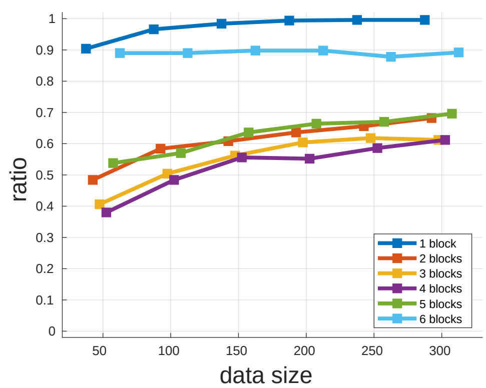

Figure 2: **Simulation study.** For a significance level $\alpha = 0.1$, ratio of patterns of mutual independence correctly detected using correlation thresholding and FDR.

3 Simulation study: noncentral chi-squared distribution

We used the same procedure as detailed in §4.2 of the manuscript, except for the fact that we used the noncental chi-squared approximation developped in §3.3.1 of the manuscript. Results are reported in Figure 3.

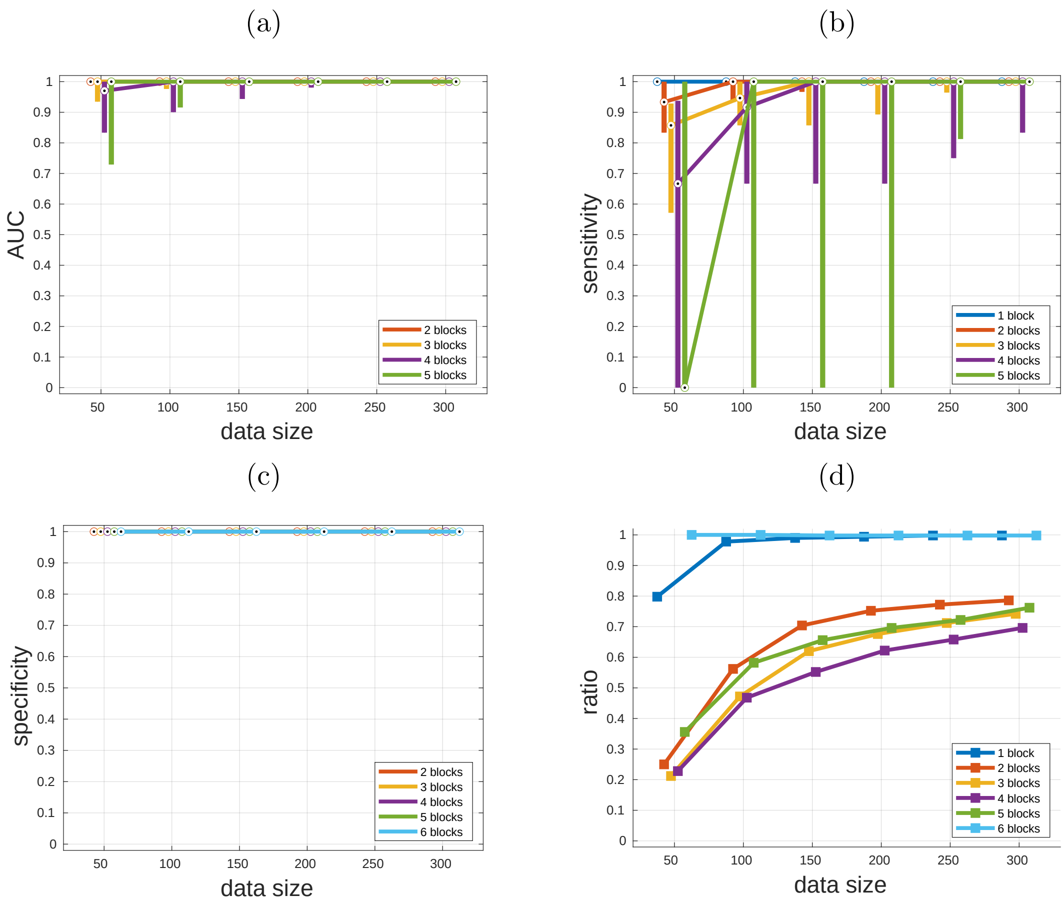

Figure 3: **Simulation study.** (a) AUC. For a significance level $\alpha = 0.1$: (b) Sensitivity; (c) Specificity. (d) Ratio of patterns of mutual independence correctly detected. (a), (b) and (c) are boxplot (median and $[25\%, 75\%]$ frequency interval). p-values are computed with noncentral chi-squared distribution.

References

Aigner, M., 1979. Combinatorial Theory. volume 234 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin.

Anderson, T.W., 1958. An Introduction to Multivariate Statistical Analysis. Wiley Publications in Statistics, John Wiley and Sons, New York.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 57, 289–300.

Birkhoff, G., 1973. Lattice Theory. volume XXV of *American Mathematical Society Colloquium Publications*. 3rd ed., American Mathematical Society, Providence, Rhode Island.

Knuth, K.H., 2004. Lattice duality: the origin of probability and entropy. Neurocomputing 67, 245–274.