Submitted to Bernoulli

On some geometric identities involving the sample covariance matrix and its adjugate

ALEXANDER DÜRRE[a] and DAVY PAINDAVEINE[b]

Mathematisch Instituut, Universiteit Leiden, Leiden, The Netherlands [a]*a.m.durre@math.leidenuniv.nl*
ECARES and Department of Mathematics, Université libre de Bruxelles, Brussels, Belgium
[b]*Davy.Paindaveine@ulb.be*

Several identities, involving the Lebesgue measure of data-based simplices or parallelotopes, have been obtained for functionals involving the sample covariance matrix S or its population counterpart Σ. This is the case in particular for Wilks' generalized variance $\det(S)$, which allowed one to obtain an explicit expression for $\mathrm{E}[\det(S)]$ whenever observations are randomly sampled from a distribution with finite second-order moments. To date, however, all such results are limited to scalar functionals. In this paper, we obtain geometric identities for the adjugate $\mathrm{adj}(S)$ of S and for other functionals involving $\mathrm{adj}(S)$ and the sample mean vector $\bar{X}$. This allows us in particular to define uniformly minimum risk unbiased (UMRU) estimators of the corresponding population quantities. Just as the results from Drton, Massam and Olkin (2008) find applications when conditional independence is of interest, our results are relevant in an elliptical framework (resp., in a general framework) where conditional independence is replaced with partial uncorrelatedness (resp., with an original concept of partial median-uncorrelatedness).

Keywords: Adjugate matrices; graphical models; partial uncorrelatedness; random polytopes; random projections; sample covariance matrices; UMRU estimators

1. Introduction

Let $X_1, \ldots, X_n$ form a random sample from a probability measure over $\mathbb{R}^d$ admitting finite second-order moments; throughout, we will denote the mean vector of X_1 as μ and its covariance matrix as Σ. It is well-known (and easy to check) that the sample covariance matrix of $X_1, \ldots, X_n$, namely

$$S = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})', \qquad \text{with } \bar{X} := \frac{1}{n} \sum_{i=1}^{n} X_i, \tag{1.1}$$

rewrites

$$S = \frac{1}{n^2} \sum_{1 \le i_1 < i_2 \le n} (X_{i_1} - X_{i_2})(X_{i_1} - X_{i_2})'. \tag{1.2}$$

This U-statistic expression has several advantages and in particular it directly provides the classical result

$$\mathrm{E}[S] = \frac{1}{n^2} \binom{n}{2} (2\Sigma) = \frac{n-1}{n} \Sigma. \tag{1.3}$$

Similarly, denoting as

$$m_d(\mathrm{Simpl}(x_1, x_2, \ldots, x_{d+1})) := \frac{1}{d!} \left| \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_{d+1} \end{pmatrix} \right|$$

the Lebesgue measure of the simplex with vertices $x_1, \ldots, x_{d+1}$ in $\mathbb{R}^d$ (that is, the convex hull of these $d+1$ points in $\mathbb{R}^d$), Theorem 2 in van der Vaart (1965) derived the expressions

$$\det(S) = \frac{(d!)^2}{n^{d+1}} \sum_{1 \le i_1 < i_2 < \ldots < i_{d+1} \le n} m_d^2(\mathrm{Simpl}(X_{i_1}, X_{i_2}, \ldots, X_{i_{d+1}})) \tag{1.4}$$

$$= \frac{(d!)^2}{n^d} \sum_{1 \le i_1 < i_2 < \ldots < i_d \le n} m_d^2(\mathrm{Simpl}(X_{i_1}, X_{i_2}, \ldots, X_{i_d}, \bar{X})) \tag{1.5}$$

for the *generalized variance* $\det(S)$ from Wilks (1932)[1]; here, it is tacitly assumed that $n > d$. Just like the U-statistic expression of S in (1.2) allows one to obtain $\mathrm{E}[S]$ in (1.3), the expression in (1.4) can be exploited to show that

$$\mathrm{E}[\det(S)] = \frac{d!}{n^d} \binom{n-1}{d} \det(\Sigma); \tag{1.6}$$

see Corollary 2.1 in van der Vaart (1965). When the common distribution of $X_1, \ldots, X_n$ is multinormal, it is well-known that $nS \sim W_d(n-1, \Sigma)$ (see, e.g., Theorem 4.1(a) from Bodnar, Mazur and Podgórski (2016) for the general case where Σ may be singular[2]), and the identities (1.3) and (1.6) can then be recovered from standard results on Wishart matrices (for (1.6), the result follows e.g. from Theorem 3.2.15 in Muirhead, 2005).

By considering measures of lower-dimensional simplices, Pronzato, Wynn and Zhigljavsky (2017) extended the results (1.4) and (1.6) from van der Vaart (1965) to other functionals of S; see also Pronzato, Wynn and Zhigljavsky (2018). To date, however, except for the trivial case of S itself in (1.2)–(1.3), only *real-valued* functionals of S could be considered. In the present paper, we consider new functionals that include matrix-valued and vector-valued functionals of S and $\bar{X}$. These include the adjugate adj(S) of S, a natural estimator for the co-factor matrix $\mathrm{adj}(\Sigma)$ that is an object of interest in applications involving conditional independence; see, e.g., Drton, Massam and Olkin (2008). As we will show, our corresponding results, that do not require Gaussianity, are relevant in applications where conditional independence is replaced with partial uncorrelatedness or, more generally, with an original concept of partial median-uncorrelatedness.

For the various functionals of S and $\bar{X}$ we consider, we will provide geometric representations that allow one to obtain unbiased estimators under minimal assumptions. Measures of random simplices can of course describe real-valued functionals only, which will lead us to consider random projections, too. Trivially, such random projections already appear in the case of S itself since (1.2) rewrites

$$S = \frac{1}{n^2} \sum_{1 \le i_1 < i_2 \le n} m_1^2(\mathrm{Simpl}(X_{i_1}, X_{i_2})) \Psi_{X_{i_1}, X_{i_2}},$$

[1] Incidentally, we mention that a simple proof of (1.5) relying on a double application of the Cauchy–Binet formula is obtained by taking $\ell = d$ and $\mu = \bar{X}$ in the proof of Theorem 2.4(b) from Dürre and Paindaveine (2022a).

[2] Note that this is unrelated to *singular Wishart distributions*, that are obtained with $n < d$; see, e.g., Srivastava (2003) and Bodnar and Okhrin (2008).

where $m_1(\mathrm{Simpl}(x_1, x_2)) := \|x_1 - x_2\|$ is the length of the one-dimensional simplex (a line segment) with vertices x_1 and x_2, and where

$$\Psi_{x_1,x_2} := \frac{(x_1 - x_2)(x_1 - x_2)'}{\|x_1 - x_2\|^2}$$

is the matrix of the orthogonal projection onto the one-dimensional vector space parallel to that simplex. Our results will involve higher-dimensional simplices and more complex orthogonal projections.

The outline of the paper is as follows. In Section 2, we introduce the notation we will use throughout the paper and provide our representation results, that, parallel to (1.2) and (1.4), provide U-statistic expressions for the statistics we will consider. In Section 3, we exploit these results to define (optimal) unbiased estimators of the corresponding population quantities. Our derivations yield in particular alternative ways to establish (1.6) and a determinant identity obtained in Pronzato (1998). In Section 4, we provide a more direct proof of (an extension of) the determinant identity from Pronzato (1998), which is actually required to establish one of the results from Section 3. In Section 5, we show the relevance of our U-statistic representation results when testing for partial uncorrelatedness. In Section 6, we derive, for the statistics already considered in Section 2, representation results that are relative to $\bar{X}$ as in (1.1) and (1.5) rather than U-statistic expressions as in (1.2) and (1.4). In Section 7, we provide conclusions and perspectives for future research. Finally, an appendix collects some technical proofs. Further auxiliary material/proofs are provided in the online supplement Dürre and Paindaveine (2025).

2. U-statistic representation results of functionals involving the adjugate matrix

We introduce the following notation. For $x_1, \ldots, x_d \in \mathbb{R}^d$, we denote as $\mathrm{Simpl}(x_1, \ldots, x_d)$ the simplex with vertices $x_1, \ldots, x_d$, that is still defined as the convex hull of these d points in $\mathbb{R}^d$. The $(d-1)$-measure of this simplex is

$$m_{d-1}(\mathrm{Simpl}(x_1, \ldots, x_d)) := \frac{1}{(d-1)!}\sqrt{\det((x_2 - x_1 \ \ldots \ x_d - x_1)'(x_2 - x_1 \ \ldots \ x_d - x_1))}; \tag{2.1}$$

here, the $(d-1)$-measure is just the length for $d = 2$, the area for $d = 3$, etc. Clearly, the quantity in (2.1) is positive if and only if there is a unique $(d-1)$-dimensional hyperplane $\pi_{x_1,\ldots,x_d}$ containing $x_1, \ldots, x_d$; we then denote as

$$\Gamma_{x_1,\ldots,x_d} := I_d - \Psi(\Psi'\Psi)^{-1}\Psi', \quad \text{with } \Psi := (x_1 - x_d \ \ldots x_{d-1} - x_d),$$

the matrix of the orthogonal projection onto the one-dimensional vector space that is orthogonal to $\pi_{x_1,\ldots,x_d}$. If the quantity in (2.1) is zero, then we let $\Gamma_{x_1,\ldots,x_d} := 0$ (this is for the sake of simplicity as all our results below would still hold if $\Gamma_{x_1,\ldots,x_d}$ would be defined differently when $m_{d-1}(\mathrm{Simpl}(x_1, \ldots, x_d)) = 0$). Finally, $\mathrm{adj}(s)$ is the adjugate of $s := \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})'$, where $\bar{x} := \frac{1}{n}\sum_{i=1}^n x_i$. We then have the following result.

Theorem 2.1. *For any $x_1, \ldots, x_n \in \mathbb{R}^d$,*

$$(i) \quad \frac{((d-1)!)^2}{n^d} \sum_{1 \le i_1 < \ldots < i_d \le n} m^2_{d-1}(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}))\Gamma_{x_{i_1},\ldots,x_{i_d}} = \mathrm{adj}(s),$$

$$(ii) \quad \frac{((d-1)!)^2}{n^d} \sum_{1 \le i_1 < \ldots < i_d \le n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}))\Gamma_{x_{i_1},\ldots,x_{i_d}} x_{i_d} = \mathrm{adj}(s)\bar{x},$$

and

$$(iii) \quad \frac{((d-1)!)^2}{n^d} \sum_{1 \le i_1 < \ldots < i_d \le n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}))x_{i_d}'\Gamma_{x_{i_1},\ldots,x_{i_d}} x_{i_d} = \det(s) + \bar{x}'\mathrm{adj}(s)\bar{x},$$

where sums over an empty collection of indices are defined as being equal to zero.

The proof requires the following classical result in linear algebra, that is often referred to as the *matrix determinant lemma*— or, more rarely, as *Cauchy's formula for the determinant of a rank-one perturbation* (see, e.g., Equation (0.8.5.11) in Horn and Johnson, 2013).

Lemma 2.1. *For any $x, y \in \mathbb{R}^d$ and any $d \times d$ matrix A,* $\det(A + uv') = \det(A) + v'\mathrm{adj}(A)u$.

Proof of Theorem 2.1. If $n \in \{1, \ldots, d-1\}$, then s has at most rank $n-1 \le d-2$, so that its adjugate matrix is the zero matrix. Since the result then trivially holds, we may therefore restrict to $n \ge d$ in the rest of the proof. For $n \ge d$, we proceed by computing in two different ways the quantity

$$r_n(x) := \sum_{1 \le i_1 < \ldots < i_d \le n} m_d^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}, x))$$

for any $x \in \mathbb{R}^d$. Lemma A.1 provides

$$m_d(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}, x)) = \frac{1}{d} m_{d-1}(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}))\|\Gamma_{x_{i_1},\ldots,x_{i_d}}(x - x_{i_d})\|.$$

Irrespective of whether $m_{d-1}(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}))$ is zero or not, there exists a unique (up to an unimportant sign) d-vector $v_{i_1,\ldots,i_d}$ such that

$$v_{i_1,\ldots,i_d} v_{i_1,\ldots,i_d}' = m_{d-1}^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}))\Gamma_{x_{i_1},\ldots,x_{i_d}}. \tag{2.2}$$

Therefore,

$$\begin{aligned} m_d^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}, x)) &= \frac{1}{d^2} m_{d-1}^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_d}))\|\Gamma_{x_{i_1},\ldots,x_{i_d}}(x - x_{i_d})\|^2 \\ &= \frac{1}{d^2}(x - x_{i_d})' v_{i_1,\ldots,i_d} v_{i_1,\ldots,i_d}'(x - x_{i_d}), \end{aligned}$$

which yields

$$\begin{aligned} r_n(x) = \frac{1}{d^2} x' \Bigg(\sum_{1 \le i_1 < \ldots < i_d \le n} v_{i_1,\ldots,i_d} v_{i_1,\ldots,i_d}' \Bigg) x \\ - \frac{2}{d^2} x' \Bigg(\sum_{1 \le i_1 < \ldots < i_d \le n} v_{i_1,\ldots,i_d} v_{i_1,\ldots,i_d}' x_{i_d} \Bigg) + \frac{1}{d^2} \Bigg(\sum_{1 \le i_1 < \ldots < i_d \le n} \|v_{i_1,\ldots,i_d}' x_{i_d}\|^2 \Bigg). \end{aligned} \tag{2.3}$$

Now, letting $c_{i_1,\ldots,i_d} = (x_{i_1} - x \ \ldots \ x_{i_d} - x)$, we also have

$$m_d(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_d},x)) = \frac{1}{d!}|\det c_{i_1,\ldots,i_d}| = \frac{1}{d!}\sqrt{\det(c_{i_1,\ldots,i_d}c'_{i_1,\ldots,i_d})},$$

so that the Cauchy–Binet formula provides

$$\begin{aligned} r_n(x) &= \frac{1}{(d!)^2}\sum_{1\le i_1<\ldots<i_d\le n} m_d^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_d},x)) \\ &= \frac{1}{(d!)^2}\sum_{1\le i_1<\ldots<i_d\le n} \det\left(\sum_{\ell=1}^{d}(x_{i_\ell}-x)(x_{i_\ell}-x)'\right) \\ &= \frac{1}{(d!)^2}\det\left(\sum_{i=1}^{n}(x_i-x)(x_i-x)'\right) \\ &= \frac{n^d}{(d!)^2}\det(s+(x-\bar{x})(x-\bar{x})'). \end{aligned}$$

Applying Lemma 2.1, we obtain

$$r_n(x) = \frac{n^d}{(d!)^2}\{\det(s) + (x-\bar{x})'\mathrm{adj}(s)(x-\bar{x})\}.$$

Since x is arbitrary, comparing with (2.3) yields

$$\frac{1}{d^2}\sum_{1\le i_1<\ldots<i_d\le n} v_{i_1,\ldots,i_d}v'_{i_1,\ldots,i_d} = \frac{n^d}{(d!)^2}\mathrm{adj}(s),$$

$$\frac{1}{d^2}\sum_{1\le i_1<\ldots<i_d\le n} v_{i_1,\ldots,i_d}v'_{i_1,\ldots,i_d}x_{i_d} = \frac{n^d}{(d!)^2}\mathrm{adj}(s)\bar{x},$$

and

$$\frac{1}{d^2}\sum_{1\le i_1<\ldots<i_d\le n} x'_{i_d}v_{i_1,\ldots,i_d}v'_{i_1,\ldots,i_d}x_{i_d} = \frac{n^d}{(d!)^2}\{\det(s)+\bar{x}'\mathrm{adj}(s)\bar{x}\},$$

which, in view of (2.2), establishes the result. □

Note that since

$$\Gamma_{x_{i_1},\ldots,x_{i_d}}x_{i_d} = \Gamma_{x_{i_1},\ldots,x_{i_d}}\{x_{i_r} + (x_{i_d} - x_{i_r})\} = \Gamma_{x_{i_1},\ldots,x_{i_d}}x_{i_r}$$

for any $r = 1,\ldots,d$, the U-statistics in Theorem 2.1 actually all involve a symmetric kernel.

3. Unbiased estimation of the considered functionals

In this section, we show how Theorem 2.1 allows us to achieve (optimal) unbiased estimation of $\mathrm{adj}(\Sigma)$, $\mathrm{adj}(\Sigma)\mu$, $\mu'\mathrm{adj}(\Sigma)\mu$, and $\det(\Sigma)$. We will need the following lemma (see the appendix for a proof).

Lemma 3.1. *Let $X_1, \ldots, X_d$ form a random sample from a probability measure over $\mathbb{R}^d$ admitting finite second-order moments. Then,*

$$(i) \quad \mathrm{E}[m^2_{d-1}(\mathrm{Simpl}(X_1, \ldots, X_d))\Gamma_{X_1,\ldots,X_d}] = \frac{d}{(d-1)!}\,\mathrm{adj}(\Sigma),$$

$$(ii) \quad \mathrm{E}[m^2_{d-1}(\mathrm{Simpl}(X_1, \ldots, X_d))\Gamma_{X_1,\ldots,X_d}X_d] = \frac{d}{(d-1)!}\,\mathrm{adj}(\Sigma)\mu,$$

and

$$(iii) \quad \mathrm{E}[m^2_{d-1}(\mathrm{Simpl}(X_1, \ldots, X_d))X_d'\Gamma_{X_1,\ldots,X_d}X_d] = \frac{d}{(d-1)!}\{\det(\Sigma) + \mu'\mathrm{adj}(\Sigma)\mu\},$$

where we wrote $\mu = \mathrm{E}[X_1]$ and $\Sigma = \mathrm{E}[(X_1 - \mu)(X_1 - \mu)']$.

Once it is shown that the expectations in Lemma 3.1(i)–(iii) exist and are finite under finite second-order moment assumptions (which we will do in the proof of this lemma), the result follows by applying in Theorem 2.1 a law of large numbers for U-statistics[3] and the continuous mapping theorem. The proof we provide in the appendix is actually a direct one, that avoids using any limit theorem. More importantly, the following result is now a trivial consequence of Theorem 2.1 and Lemma 3.1.

Theorem 3.1. *Let $X_1, \ldots, X_n$, with $n \geq d$, form a random sample from a probability measure over $\mathbb{R}^d$ admitting finite second-order moments. Then,*

$$(i) \quad \mathrm{E}[\mathrm{adj}(S)] = \frac{d!}{n^d}\binom{n}{d}\mathrm{adj}(\Sigma)$$

and

$$(ii) \quad \mathrm{E}[\mathrm{adj}(S)\bar{X}] = \frac{d!}{n^d}\binom{n}{d}\mathrm{adj}(\Sigma)\mu,$$

so that $\mathrm{E}[\mathrm{adj}(S)\bar{X}] = \mathrm{E}[\mathrm{adj}(S)]\mathrm{E}[\bar{X}]$. Moreover,

$$(iii) \quad \mathrm{E}[\det(S) + \bar{X}'\mathrm{adj}(S)\bar{X}] = \frac{d!}{n^d}\binom{n}{d}\{\det(\Sigma) + \mu'\mathrm{adj}(\Sigma)\mu\}.$$

In the particular case where the common distribution of $X_1, \ldots, X_n$ is multinormal, Theorem 4.1 from Bodnar, Mazur and Podgórski (2016) yields that $nS \sim W_d(n-1, \Sigma)$ and that $\bar{X}$ and S are mutually independent, which allows one to obtain the results in Theorem 3.1 from known results on Wishart matrices. In particular, applying with $(r, m) = (d, d-1)$ the second identity in Corollary 4.2 from Drton, Massam and Olkin (2008) provides

$$\mathrm{E}[\mathrm{adj}(nS)] = \frac{(n-1)!}{(n-d)!}\mathrm{adj}(\Sigma),$$

which, in line with Theorem 3.1(i), rewrites

$$\mathrm{E}[\mathrm{adj}(S)] = \frac{(n-1)!}{n^{d-1}(n-d)!}\mathrm{adj}(\Sigma) = \frac{d!}{n^d}\binom{n}{d}\mathrm{adj}(\Sigma).$$

[3] For instance, the strong law of large numbers stated in Theorem A on page 190 of Serfling (1980).

Of course, the mutual independence between S and $\bar{X}$ then allows one to deduce Theorem 3.1(ii) and to evaluate $\mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}] = \mathrm{tr}(\mathrm{E}[\mathrm{adj}(S)]\mathrm{E}[\bar{X}\bar{X}'])$, which in turn will provide Theorem 3.1(iii) (recall that we indeed explained in the introduction how to obtain $\mathrm{E}[\det(S)]$ in the present Gaussian case from the Wishart distribution of S). It is remarkable that the results in Theorem 3.1 extend from the Gaussian case to an arbitrary distribution with finite second-order moments.

While Theorem 3.1(iii) provides an expression for $\mathrm{E}[\det(S) + \bar{X}'\mathrm{adj}(S)\bar{X}]$, Lemma 3.1 actually allows us to compute both $\mathrm{E}[\det(S)]$ and $\mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}]$. We have the following result.

Theorem 3.2. *Let $X_1, \ldots, X_n$, with $n \geq d$, form a random sample from a probability measure over $\mathbb{R}^d$ admitting finite second-order moments. Then,*

$$(i) \quad \mathrm{E}[\det(S)] = \frac{d!}{n^d}\binom{n-1}{d}\det(\Sigma)$$

and

$$(ii) \quad \mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}] = \frac{d!}{n^d}\binom{n-1}{d-1}\det(\Sigma) + \frac{d!}{n^d}\binom{n}{d}\mu'\mathrm{adj}(\Sigma)\mu.$$

Proof of Theorem 3.2. We start with the proof of (ii) in the particular case where $\mu = 0$. Since

$$\mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}] = \frac{1}{n}\mathrm{E}[X_1'\mathrm{adj}(S)X_1] + \frac{n-1}{n}\mathrm{E}[X_1'\mathrm{adj}(S)X_2],$$

Theorem 2.1(i) yields

$$\frac{n^{d+1}}{((d-1)!)^2}\mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}] = \sum_I \mathrm{E}[m_I^2 X_1'\Gamma_I X_1] + (n-1)\sum_I \mathrm{E}[m_I^2 X_1'\Gamma_I X_2],$$

where the sum is over all $I = (i_1, \ldots, i_d)$ with $1 \leq i_1 < \ldots < i_d \leq n$. If it is not so that $\{1,2\} \subset I$, then the mutual independence of the X_i's and the fact that $\mu = 0$ yield $\mathrm{E}[m_I^2 X_1'\Gamma_I X_2] = 0$, so that

$$\frac{n^{d+1}}{((d-1)!)^2}\mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}] = \sum_{I \ni 1}\mathrm{E}[m_I^2 X_1'\Gamma_I X_1] + \sum_{I \not\ni 1}\mathrm{E}[m_I^2 X_1'\Gamma_I X_1] + (n-1)\sum_{I \supset \{1,2\}}\mathrm{E}[m_I^2 X_1'\Gamma_I X_2].$$

Letting $I_a = (1, \ldots, d)$ and $I_b = (2, \ldots, d+1)$, we then have

$$\begin{aligned}
&\frac{n^{d+1}}{((d-1)!)^2}\mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}] \\
&= \binom{n-1}{d-1}\mathrm{E}[m_{I_a}^2 X_1'\Gamma_{I_a} X_1] + \binom{n-1}{d}\mathrm{E}[m_{I_b}^2 X_1'\Gamma_{I_b} X_1] + (n-1)\binom{n-2}{d-2}\mathrm{E}[m_{I_a}^2 X_1'\Gamma_{I_a} X_2] \\
&= d\binom{n-1}{d-1} \times \frac{d}{(d-1)!}\det(\Sigma) + \binom{n-1}{d}\mathrm{E}[m_{I_b}^2 X_1'\Gamma_{I_b} X_1],
\end{aligned} \tag{3.1}$$

where the last equality follows by using the identity $n\binom{n-1}{k-1} = k\binom{n}{k}$, the fact that

$$X_1'\Gamma_{I_a}X_j = \{X_d + (X_1 - X_d)\}'\Gamma_{I_a}\{X_d + (X_j - X_d)\} = X_d'\Gamma_{I_a}X_d$$

for $j = 1, 2$, then Lemma 3.1(iii). Moreover, Lemma 3.1(i) provides

$$\mathrm{E}[m_{I_b}^2 X_1' \Gamma_{I_b} X_1] = \mathrm{E}[\mathrm{tr}(m_{I_b}^2 X_1 X_1' \Gamma_{I_b})] = \mathrm{tr}(\mathrm{E}[X_1 X_1']\mathrm{E}[m_{I_b}^2 \Gamma_{I_b}])$$
$$= \frac{d}{(d-1)!}\,\mathrm{tr}(\Sigma\,\mathrm{adj}(\Sigma)) = \frac{d}{(d-1)!}\det(\Sigma)\,\mathrm{tr}(I_d) = \frac{d^2}{(d-1)!}\det(\Sigma).$$

Plugging this into (3.1) yields

$$\frac{n^{d+1}}{((d-1)!)^2}\mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}] = \frac{d^2}{(d-1)!}\binom{n-1}{d-1}\det(\Sigma) + \binom{n-1}{d}\frac{d^2}{(d-1)!}\det(\Sigma)$$
$$= \frac{d^2}{(d-1)!}\binom{n}{d}\det(\Sigma)$$
$$= \frac{nd}{(d-1)!}\binom{n-1}{d-1}\det(\Sigma),$$

hence

$$\mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}] = \frac{d!}{n^d}\binom{n-1}{d-1}\det(\Sigma),$$

which indeed establishes (ii) for $\mu = 0$. For a general mean vector μ, applying the result for $\mu = 0$ jointly with Theorem 3.1(i)–(ii) then provides

$$\mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}] = \mathrm{E}[(\bar{X}-\mu)'\mathrm{adj}(S)(\bar{X}-\mu)] - 2\mu'\mathrm{E}[\mathrm{adj}(S)(\bar{X}-\mu)] + \mu'\mathrm{E}[\mathrm{adj}(S)]\mu$$
$$= \frac{d!}{n^d}\binom{n-1}{d-1}\det(\Sigma) - 0 + \frac{d!}{n^d}\binom{n}{d}\mu'\mathrm{adj}(\Sigma)\mu,$$

which establishes (ii). Turning to Part (i), Theorem 3.1(iii) and Part (ii) of the result now yield

$$\mathrm{E}[\det(S)] = \mathrm{E}[\det(S) + \bar{X}'\mathrm{adj}(S)\bar{X}] - \mathrm{E}[\bar{X}'\mathrm{adj}(S)\bar{X}]$$
$$= \frac{d!}{n^d}\binom{n}{d}\{\det(\Sigma) + \mu'\mathrm{adj}(\Sigma)\mu\} - \left\{\frac{d!}{n^d}\binom{n-1}{d-1}\det(\Sigma) + \frac{d!}{n^d}\binom{n}{d}\mu'\mathrm{adj}(\Sigma)\mu\right\}$$
$$= \frac{d!}{n^d}\left\{\binom{n}{d} - \binom{n-1}{d-1}\right\}\det(\Sigma)$$
$$= \frac{d!}{n^d}\binom{n-1}{d}\det(\Sigma),$$

which ends the proof. □

As mentioned in the introduction, the identity in Theorem 3.2(i) was obtained in van der Vaart (1965); see also Pronzato, Wynn and Zhigljavsky (2017) and Pronzato, Wynn and Zhigljavsky (2018). To the best of our knowledge, Theorem 3.2(ii) is original. More importantly, we have the following result.

Corollary 3.1. *Let λ be a σ-finite measure over $\mathbb{R}^d$ and $\mathcal{P}$ be the collection of all probability measures of $\mathbb{R}^d$ that are equivalent to λ and admit finite second-order moments. Consider the statistical model where one observes a random sample $X_1, \ldots, X_n$ from a probability measure $P \in \mathcal{P}$. Then, irrespective of the convex loss function that is used to define the risk of estimators,*

$$\frac{n^d}{d!\binom{n}{d}}\mathrm{adj}(S), \quad \frac{n^d}{d!\binom{n}{d}}\mathrm{adj}(S)\bar{X}, \quad \frac{n^d}{d!\binom{n}{d}}\left\{\bar{X}'\mathrm{adj}(S)\bar{X} - \frac{d}{n-d}\det(S)\right\}, \quad \textit{and} \quad \frac{n^d}{d!\binom{n-1}{d}}\det(S)$$

are uniformly minimum risk unbiased (UMRU) estimators of $\mathrm{adj}(\Sigma)$, $\mathrm{adj}(\Sigma)\mu$, $\mu'\mathrm{adj}(\Sigma)\mu$, *and* $\det(\Sigma)$, *respectively.*

PROOF OF COROLLARY 3.1. Unbiasedness directly follows from Theorems 3.1–3.2. Now, under the assumptions we adopted in the statement of the result, Corollary 1.5.11 from Pfanzagl (1994) entails that the considered statistical model involves a family of probability measures that is *symmetrically complete*. In other words, the order statistic[4] is complete for this family of probability measures. Since the four estimators in the statement of the corollary are symmetric functions of $X_1, \ldots, X_n$ (see the remark at the end of Section 2), the result therefore follows from the Lehmann–Scheffé theorem. □

For the quadratic loss (UMVU estimation), the result for $\det(\Sigma)$ was stated in Theorem 3.2 from Pronzato, Wynn and Zhigljavsky (2017). In Section **??** of the online supplement, we conduct a numerical exercise that supports unbiasedness of the estimators considered in Corollary 3.1.

4. An extended determinant result

Let us come back to Theorem 3.1. Note that applying Lemma 2.1 to the statement in Part (iii) of this theorem provides

$$\mathrm{E}\left[\det\left(\frac{1}{n}\sum_{i=1}^n X_i X_i'\right)\right] = \frac{d!}{n^d}\binom{n}{d}\det(\mathrm{E}[X_1 X_1']). \tag{4.1}$$

Applying this identity to $X_i - \mu$ rather than to X_i thus yields

$$\mathrm{E}\left[\det\left(\frac{1}{n}\sum_{i=1}^n (X_i - \mu)(X_i - \mu)'\right)\right] = \frac{n!}{n^d(n-d)!}\det(\Sigma),$$

which is to be compared to the result in Theorem 3.2(i) stating that

$$\mathrm{E}\left[\det\left(\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})'\right)\right] = \frac{(n-1)!}{n^d(n-d-1)!}\det(\Sigma).$$

The remarkable identity in (4.1) was obtained in Pronzato (1998) under stronger assumptions; see Theorem 1 in Pronzato (1998). There, the result was proved by establishing it first for finitely discrete distributions, then by approximating an arbitrary distribution with a finitely discrete distribution. We provide here an original, direct, proof that not only avoids such a density argument but also requires

[4]Here, the order statistic may be defined as $T(x_1, \ldots, x_n) = \{x_1, \ldots, x_n\}$, or more generally, as any other statistic that would be maximal invariant with respect to the group of permutations of $x_1, \ldots, x_n$.

minimal assumptions. The result is also extended to a random matrix that has the flavor of a covariance matrix between a random vector X and a random vector Y.

Theorem 4.1. *Let (X,Y) be a couple of random d-vectors for which both $\mathrm{E}[\|X\|^2]$ and $\mathrm{E}[\|Y\|^2]$ are finite. Let $(X_1,Y_1),\dots,(X_n,Y_n)$ be mutually independent copies of (X,Y). Then,*

$$\mathrm{E}\left[\det\left(\frac{1}{n}\sum_{i=1}^{n} X_i Y_i'\right)\right] = c_{n,d}\det(\mathrm{E}[XY']),$$

with $c_{n,d} = n!/\{n^d(n-d)!\}$ if $n \geq d$ and $c_{n,d} = 0$ otherwise.

We first establish the following preliminary result.

Lemma 4.1. *Let (X,Y) be a couple of random d-vectors for which both $\mathrm{E}[\|X\|^2]$ and $\mathrm{E}[\|Y\|^2]$ are finite. Let $(X_1,Y_1),\dots,(X_n,Y_n)$ be mutually independent copies of (X,Y). Then,*

$$\mathrm{E}\left[\det\left(\frac{1}{n}\sum_{i=1}^{n} X_i Y_i'\right)\right]$$

exists and is finite.

Proof of Lemma 4.1. The Cauchy–Binet formula yields

$$\begin{aligned}\det\left(\sum_{i=1}^{n} X_i Y_i'\right) &= \det((X_1\ \dots\ X_n)(Y_1\ \dots\ Y_n)') \\ &= \sum_{1\leq i_1<\dots<i_d\leq n} \det((X_{i_1}\ \dots\ X_{i_d}))\det((Y_{i_1}\ \dots\ Y_{i_d})).\end{aligned}$$

Since the distribution of each term in the last sum is the same as the distribution of $\det((X_1\ \dots\ X_d))$ $\det((Y_1\ \dots\ Y_d))$, it is then enough to show that

$$\mathrm{E}[|\det((X_1\ \dots\ X_d))\det((Y_1\ \dots\ Y_d))|] < \infty.$$

From the Cauchy–Schwarz inequality, it is of course sufficient to prove that

$$\mathrm{E}[\{\det((X_1\ \dots\ X_d))\}^2] < \infty \quad \text{and} \quad \mathrm{E}[\{\det((Y_1\ \dots\ Y_d))\}^2] < \infty.$$

Recall that $|\det((X_1\ \dots\ X_d))|$ is the Lebesgue measure of the parallelotope

$$\left\{\sum_{i=1}^{d} \lambda_i X_i : 0 \leq \lambda_i \leq 1\right\}$$

and note that, trivially, this Lebesgue measure is upper-bounded by $\prod_{i=1}^{d}\|X_i\|$ (the upper-bound is achieved when the parallelotope is a hyper-rectangle). The mutual independence of $X_1,\dots,X_d$ then provides

$$\mathrm{E}[\{\det((X_1 \ldots X_d))\}^2] \leq \prod_{i=1}^{d} \mathrm{E}[\|X_i\|^2] = (\mathrm{E}[\|X\|^2])^d < \infty.$$

Since the same argument shows that $\mathrm{E}[\{\det((Y_1 \ldots Y_d))\}^2] < \infty$, the result is proved. □

We can now prove our extended determinant result.

PROOF OF THEOREM 4.1. Note first that the result is trivial for $n = 1, \ldots, d-1$. Now, if the result holds for $n = d$, then the Cauchy–Binet formula yields

$$\begin{aligned}
&\mathrm{E}[\det((X_1 \ldots X_n)(Y_1 \ldots Y_n)')] \\
&\quad = \sum_{1 \leq i_1 < \ldots < i_d \leq n} \mathrm{E}[\det((X_{i_1} \ldots X_{i_d})(Y_{i_1} \ldots Y_{i_d})')] \\
&\quad = \binom{n}{d} \mathrm{E}[\det((X_1 \ldots X_d)(Y_1 \ldots Y_d)')] \\
&\quad = d! \binom{n}{d} \det(\mathrm{E}[XY']),
\end{aligned}$$

so that it is sufficient to prove the result for $n = d$. To do so, note that the definition of the determinant (involving the signature $\mathrm{sign}(\sigma)$ of any permutation σ in the collection $\mathcal{S}_d$ of all permutations of $\{1, \ldots, d\}$) provides

$$\begin{aligned}
&\mathrm{E}[\det((X_1 \ldots X_d)(Y_1 \ldots Y_d)')] \\
&\quad = \mathrm{E}[\det((Y_1 \ldots Y_d)) \det((X_1 \ldots X_d))] \\
&\quad = \sum_{\sigma \in \mathcal{S}_d} \mathrm{sign}(\sigma) \mathrm{E}\left[\prod_{i=1}^{d} (Y_{\sigma(i)})_i \det(X_1 \ldots X_d)\right] \\
&\quad = \sum_{\sigma \in \mathcal{S}_d} \mathrm{E}\left[\prod_{i=1}^{d} (Y_{\sigma(i)})_i \det(X_{\sigma(1)} \ldots X_{\sigma(d)})\right],
\end{aligned} \tag{4.2}$$

where the last equality is obtained by permuting the columns in the determinant. Exchangeability of the (X_i, Y_i)'s implies that the expectation in (4.2) does not depend on σ, so that

$$\begin{aligned}
&\mathrm{E}[\det((X_1 \ldots X_d)(Y_1 \ldots Y_d)')] \\
&\quad = d! \mathrm{E}\left[\prod_{i=1}^{d} (Y_i)_i \det(X_1 \ldots X_d)\right] \\
&\quad = d! \mathrm{E}\Big[\det((Y_1)_1 X_1 \ldots (Y_d)_d X_d)\Big]
\end{aligned}$$

$$= d! \sum_{\sigma \in \mathcal{S}_d} \operatorname{sign}(\sigma) \mathrm{E}\left[\prod_{i=1}^{d} (Y_{\sigma(i)})_{\sigma(i)} (X_{\sigma(i)})_i \right].$$

Mutual independence of the (X_i, Y_i)'s then gives

$$\begin{aligned}
&\mathrm{E}[\det((X_1 \ldots X_d)(Y_1 \ldots Y_d)')] \\
&\qquad = d! \sum_{\sigma \in \mathcal{S}_d} \operatorname{sign}(\sigma) \prod_{i=1}^{d} \mathrm{E}[(Y_{\sigma(i)})_{\sigma(i)} (X_{\sigma(i)})_i] \\
&\qquad = d! \det(\mathrm{E}[(Y_1)_1 X_1] \ldots \mathrm{E}[(Y_d)_d X_d]) \\
&\qquad = d! \det(\mathrm{E}[Y_1 X] \ldots \mathrm{E}[Y_d X]) \\
&\qquad = d! \det(\mathrm{E}[XY']),
\end{aligned}$$

which establishes the result. □

Another motivation for Theorem 4.1 is that it is needed for the proof of Lemma 3.1 we provide in the appendix.

5. Testing for partial uncorrelatedness

For a random vector $X = (X^{(1)}, \ldots, X^{(d)})$ with a distribution P admitting a positive definite covariance matrix Σ, the *partial correlation* $\rho_{k\ell}$ *between* $X^{(k)}$ *and* $X^{(\ell)}$ is defined as the correlation between the residuals $X^{(k)} - \beta' U$ and $X^{(\ell)} - \gamma' U$ obtained when regressing $X^{(k)}$ and $X^{(\ell)}$ onto the $d-2$ remaining marginals in X. As shown by Corollary 5.8.2 in Whittaker (1990),

$$\rho_{k\ell} = -\frac{\Sigma^{-1}_{k\ell}}{\sqrt{\Sigma^{-1}_{kk} \Sigma^{-1}_{\ell\ell}}}; \tag{5.1}$$

for an invertible $d \times d$ matrix A, we write throughout $A^{-1}_{k\ell}$ instead of $(A^{-1})_{k\ell}$. If P is elliptical with scatter matrix Σ, then (5.1) can be used to extend the concept of partial correlation when second-order moments are infinite[5]. If X is multinormal, then partial uncorrelatedness ($\rho_{k\ell} = 0$; $X^{(k)}$ and $X^{(\ell)}$ are uncorrelated conditional on U) is equivalent to conditional independence ($X^{(k)}$ and $X^{(\ell)}$ are independent conditional on U). This relation is pivotal for Gaussian graphical models; see, e.g., Chapter 6 of Whittaker (1990), or Uhler (2018). Away from the multinormal case, partial uncorrelatedness does not imply conditional independence, but it is still useful to identify sparse models, particularly so under

[5] When considering the elliptical case, we will throughout tacitly assume that P admits a density with respect to the Lebesgue measure of $\mathbb{R}^d$. This elliptical density is of the form $f(x) = c_{\Sigma,g}\, g((x-\mu)'\Sigma^{-1}(x-\mu))$, where Σ is a *scatter matrix* (a positive definite symmetric matrix) and $g : \mathbb{R}^+ \to \mathbb{R}^+$ is a *radial density*. The quantity $\rho_{k\ell}$ in (5.1) obtained from the scatter matrix is then well-defined, even though the scatter matrix is identified up to a positive scalar factor only. Under finite second-order moments, $\rho_{k\ell}$ coincides with the usual partial correlation.

ellipticity; we refer to Vogel and Fried (2011) for a study of the resulting elliptical graphical models. Accordingly, we consider the problem of testing

$$\mathcal{H}_0 : \rho_{k\ell} = 0 \quad \text{against} \quad \mathcal{H}_1 : \rho_{k\ell} \neq 0$$

based on a random sample $X_1, \ldots, X_n$ from P. If we do not make further assumptions on P, then we need to assume finite second-order moments; in contrast, if we assume that P is elliptical, then we avoid any moment assumption by defining $\rho_{k\ell}$ from the underlying scatter matrix.

Now, if P is assumed to admit finite fourth-order moments, then the sample covariance matrix S is asymptotically normal; more precisely, $\sqrt{n}\,\mathrm{vec}(S - \Sigma)$ is asymptotically normal with mean zero and covariance matrix $\mathrm{Var}[\mathrm{vec}\{(X_1 - \mu)(X_1 - \mu)'\}]$. From the delta method, one can show that for some suitable statistics $\hat{\sigma}_\rho$, $\hat{\sigma}_{S^{-1}}$, and $\hat{\sigma}_{\mathrm{adj}}$,

$$T_\rho = \frac{\sqrt{n}\hat{\rho}_{k\ell}}{\hat{\sigma}_\rho}, \quad T_{S^{-1}} = \frac{\sqrt{n}(S^{-1})_{k\ell}}{\hat{\sigma}_{S^{-1}}}, \quad \text{and} \quad T_{\mathrm{adj}} = \frac{\sqrt{n}(\mathrm{adj}(S))_{k\ell}}{\hat{\sigma}_{\mathrm{adj}}}$$

are asymptotically standard normal under the null hypothesis of partial uncorrelatedness; here, $\hat{\rho}_{k\ell}$ is obtained by substituting S for Σ in (5.1). The corresponding tests—ϕ_ρ, $\phi_{S^{-1}}$ and ϕ_{adj}, say—reject the null hypothesis at asymptotic level α whenever $|T_\rho|$, $|T_{S^{-1}}|$, or $|T_{\mathrm{adj}}|$ exceeds the $(1 - \alpha/2)$-quantile, $z_{\alpha/2}$ say, of the standard normal distribution; of course, $\phi_{S^{-1}}$ (resp., ϕ_{adj}) is based on the fact that $\rho_{k\ell} = 0$ holds if and only if $\Sigma^{-1}_{k\ell} = 0$ (resp., if and only if $(\mathrm{adj}(\Sigma))_{k\ell} = 0$). For the sake of completeness, the exact expressions of $\hat{\sigma}_\rho$, $\hat{\sigma}_{S^{-1}}$ and $\hat{\sigma}_{\mathrm{adj}}$, along with a proof that the test statistics above are indeed asymptotically standard normal under the null hypothesis, are provided in Section **??** of the online supplement. While these are natural tests of partial uncorrelatedness, they require finite fourth-order moments and are likely to behave poorly under heavy tails.

Interestingly, the U-statistic representation result in Theorem 2.1(i) provides another approach. Rather than using the U-statistic itself as a test statistic, which would provide again the test ϕ_{adj}, we can use the corresponding *U-quantiles*; see, e.g., Serfling (1984) and Janssen, Serfling and Veraverbeke (1984). At the population level, the U-median associated to the kernel in Theorem 2.1(i) is

$$\rho^M_{k\ell} := \inf\left\{t \in \mathbb{R} : H_{k\ell}(t) \geq 1/2\right\},$$

where $H_{k\ell}$ is the cumulative distribution function defined by

$$H_{k\ell}(t) = P\left[m^2_{d-1}(\mathrm{Simpl}(X_1, \ldots, X_d))(\Gamma_{X_1,\ldots,X_d})_{k\ell} \leq t\right]. \tag{5.2}$$

We then consider the problem of testing for *partial median-uncorrelatedness*, that is, the problem of testing

$$\mathcal{H}_0 : \rho^M_{k\ell} = 0 \quad \text{against} \quad \mathcal{H}_1 : \rho^M_{k\ell} \neq 0.$$

Letting $Z := m^2_{d-1}(\mathrm{Simpl}(X_1, \ldots, X_d))(\Gamma_{X_1,\ldots,X_d})_{k\ell}$, where the X_i's form a random sample from P, note that partial uncorrelatedness corresponds to $\mathrm{E}[Z] = 0$, whereas (under assumptions ensuring that Z has a distribution that is absolutely continuous with respect to the Lebesgue measure) partial median-uncorrelatedness rather corresponds to $\mathrm{Median}[Z] = 0$. In general, none of these concepts implies the other one; under ellipticity, however, partial uncorrelatedness implies partial median-uncorrelatedness. More precisely, we have the following result (see the appendix for a proof).

Proposition 5.1. *Let P be an elliptical probability measure over $\mathbb{R}^d$ with (positive definite) scatter matrix Σ. Then, $\rho_{k\ell}(:= \Sigma^{-1}_{k\ell}/\sqrt{\Sigma^{-1}_{kk}\Sigma^{-1}_{\ell\ell}}) = 0$ implies $\rho^M_{k\ell} = 0$.*

In the elliptical model, a test for partial median-uncorrelatedness therefore qualifies as a test for partial uncorrelatedness. We will introduce two such tests, that are based on the empirical counterpart

$$\hat{H}_{k\ell}(t) = \frac{1}{\binom{n}{d}} \sum_{1\leq i_1<\ldots<i_d\leq n} \mathbb{I}\big[m^2_{d-1}(\mathrm{Simpl}(X_{i_1},\ldots,X_{i_d}))(\Gamma_{X_{i_1},\ldots,X_{i_d}})_{k\ell} \leq t\big]$$

of $H_{k\ell}(t)$ and on the resulting sample U-median

$$\hat{\rho}^M_{k\ell} := \inf\big\{t : \hat{H}_{k\ell}(t) \geq 1/2\big\}.$$

The construction of the tests is based on the following result.

Proposition 5.2. *Let P be a probability measure over $\mathbb{R}^d$ and $X_1,\ldots,X_n$ be a random sample from P. Assume that $t \mapsto H_{k\ell}(t)$ is differentiable at $\rho^M_{k\ell}$, with derivative $h_{k\ell}(\rho^M_{k\ell}) > 0$. Then,*

$$(i) \quad \sqrt{n}(\hat{\rho}^M_{k\ell} - \rho^M_{k\ell}) = \sqrt{n}\,\frac{\frac{1}{2} - \hat{H}_{k\ell}(\rho^M_{k\ell})}{h_{k\ell}(\rho^M_{k\ell})} + o_{\mathrm{P}}(1),$$

as n diverges to infinity. Moreover, (ii) if $\xi_{k\ell} := \mathrm{Var}[L_{k\ell}(X_1)] > 0$, where we let

$$L_{k\ell}(x) := P\big[m^2_{d-1}(\mathrm{Simpl}(X_1,\ldots,X_d))(\Gamma_{X_1,\ldots,X_d})_{k\ell} \leq \rho^M_{k\ell} | X_1 = x\big],$$

then, as n diverges to infinity,

$$\sqrt{n}\Big(\hat{H}_{k\ell}(\rho^M_{k\ell}) - \frac{1}{2}\Big) \xrightarrow{\mathcal{D}} N(0, d^2\xi_{k\ell}),$$

so that $\sqrt{n}(\hat{\rho}^M_{k\ell} - \rho^M_{k\ell}) \xrightarrow{\mathcal{D}} N\big(0, d^2\xi_{k\ell}/h^2_{k\ell}(\rho^M_{k\ell})\big)$.

The Bahadur representation in Proposition 5.2(i) follows from Lemma 3.5 of Serfling (1984), whereas the asymptotic normality result in Proposition 5.2(ii) is a direct consequence of Theorem A on page 192 of Serfling (1980). Of course, defining the corresponding tests of partial median-uncorrelatedness requires consistent estimation of $\xi_{k\ell}$ and $h_{k\ell}(\rho^M_{k\ell})$ under the null hypothesis. For the latter, one can use kernel density estimation, which requires suitable choices of a kernel function and a bandwidth. We have the following results (proofs are provided in Section **??** of the online supplement).

Lemma 5.1. *Let P be a probability measure over $\mathbb{R}^d$ and $X_1,\ldots,X_n$ form a random sample from P. Assume that $\rho^M_{k\ell} = 0$ and that $t \mapsto H_{k\ell}(t)$ is continuous at 0. Then,*

$$\hat{\xi}_{k\ell} := \frac{1}{n}\sum_{i=1}^n \Big(\hat{L}_{i;k\ell} - \frac{1}{n}\sum_{i=1}^n \hat{L}_{i;k\ell}\Big)^2 \xrightarrow{\mathrm{P}} \xi_{k\ell}$$

as n diverges to infinity, where we let

$$\hat{L}_{i;k\ell} := \frac{1}{\binom{n-1}{d-1}} \sum_{(i_1,\ldots,i_{d-1})\in\mathcal{I}^{d-1}_{n,-i}} \mathbb{I}\big[m^2_{d-1}(\mathrm{Simpl}(X_i, X_{i_1},\ldots,X_{i_{d-1}}))(\Gamma_{X_i,X_{i_1},\ldots,X_{i_{d-1}}})_{k\ell} \leq 0\big],$$

with $\mathcal{I}^{d-1}_{n,-i} := \{(i_1,\ldots,i_{d-1}) \in (\mathbb{N}\setminus\{i\})^{d-1} : 1 \le i_1 < \ldots < i_{d-1} \le n\}$.

Lemma 5.2. *Let P be a probability measure over $\mathbb{R}^d$ and $X_1,\ldots,X_n$ form a random sample from P. Let $(b_n)_{n\in\mathbb{N}}$ be a sequence of bandwidths with $b_n \to 0$ and $nb_n \to \infty$. Assume furthermore that the kernel $K:\mathbb{R}\to[0,\infty)$ has a bounded support, and is such that $\int_{-\infty}^{\infty} K(t)\,dt = 1$ and $\int_{-\infty}^{\infty}(K(t))^2\,dt < \infty$. Assume that $t \mapsto H_{k\ell}(t)$ is differentiable in a neighborhood of 0, and that the corresponding derivative $h_{k\ell}$ is continuous at zero. Then,*

$$\hat{h}_{k\ell}(0) := \frac{1}{b_n\binom{n}{d}} \sum_{1\le i_1<\ldots<i_d\le n} K\left(\frac{m^2_{d-1}(\mathrm{Simpl}(X_{i_1},\ldots,X_{i_d}))(\Gamma_{X_{i_1},\ldots,X_{i_d}})_{k\ell}}{b_n}\right) \xrightarrow{\mathrm{P}} h_{k\ell}(0)$$

as n diverges to infinity.

Quite naturally, the results above lead to the tests for partial median-uncorrelatedness ϕ^{M}_H and ϕ^{M}_ρ rejecting the null hypothesis $\mathcal{H}_0 : \rho^M_{k\ell} = 0$ at asymptotic level α whenever

$$\left|\frac{\sqrt{n}(\hat{H}_{k\ell}(0)-\frac{1}{2})}{d\sqrt{\hat{\xi}_{k\ell}}}\right| > z_{\alpha/2} \quad \text{and} \quad \left|\frac{\sqrt{n}\hat{h}_{k\ell}(0)\hat{\rho}^M_{k\ell}}{d\sqrt{\hat{\xi}_{k\ell}}}\right| > z_{\alpha/2},$$

respectively. These tests do not require any finite moment assumption. They do not require ellipticity either, but, as explained below Proposition 5.1, it is in the elliptical model that they qualify as tests of partial uncorrelatedness. It is therefore in this framework that we will perform a Monte Carlo study in order to compare them with the tests ϕ_ρ, $\phi_{S^{-1}}$ and ϕ_{adj} introduced at the beginning of this section.

The Monte Carlo exercise we conducted is as follows. For each combination of $n \in \{30, 100\}$ and $\ell \in \{0,1,2,3,4\}$, we generated $M = 10{,}000$ random samples of size n from the trivariate normal distribution with mean vector $\mu = 0$ and covariance matrix $\Sigma = (1-\ell\tau/\sqrt{n})I_3 + (\ell\tau/\sqrt{n})1_3 1_3'$, where $1_3 = (1,1,1)'$ and $\tau = .8$. In each sample, we applied the five tests considered in this section, all at asymptotic level $\alpha = 5\%$. For ϕ^M_ρ, we used the Epanechnikov kernel $K(t) = (3/4)(1-t^2)\mathbb{I}[|t| \le 1]$ and $b_n = b_{n0}/8$, where b_{n0} is the default bandwidth provided by the R function `density`. We repeated this exercise for three heavy-tailed elliptical distributions, namely the t_5 distribution (finite fourth-order moments), the t_3 distribution (finite second-order moments, but infinite fourth-order moments) and the t_1 distribution (infinite first-order moments), each with symmetry center μ and scatter matrix Σ as above (for these three distributions, we took $\tau = 1$, $\tau = 1.1$ and $\tau = 1.3$, respectively, in order to achieve powers at $\ell = 4$ that are roughly equal to those obtained in the multinormal case). The resulting rejection frequencies are reported in Figure 1.

In the multinormal case, all tests behave in a very similar way, even though the proposed median-based tests show slightly lower powers than their competitors based on the empirical covariance matrix. Under the heaviest tails considered in this simulation (t_3 and t_1), the median-based tests outperform their competitors: ϕ_{adj} is very conservative, with a clear cost in terms of power, whereas, on the contrary, ϕ_ρ and $\phi_{S^{-1}}$ are extremely liberal. The same happens under the t_5 distribution, but in a less severe fashion. Irrespective of the underlying distribution, the median-based tests behave very well, particularly so the test ϕ^M_H, that does not involve any tuning parameter (further simulations revealed that the finite-sample performances of ϕ^M_ρ crucially depend on the chosen bandwidth).

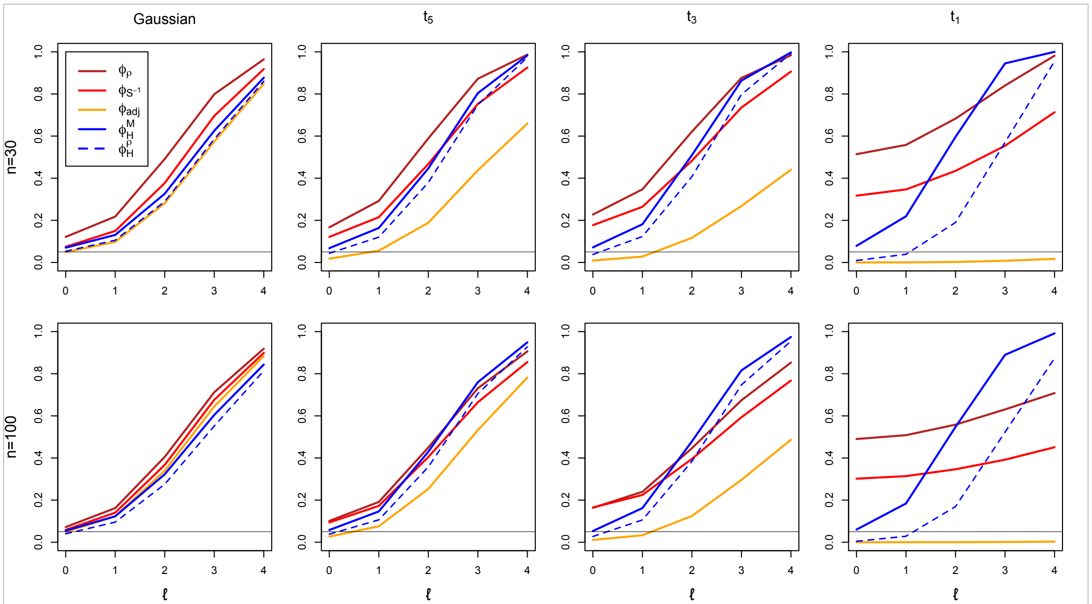

Figure 1. Rejection frequencies of five tests of partial uncorrelatedness, all performed at nominal level $\alpha = 5\%$, based on $M = 10{,}000$ mutually independent random samples of size $n \in \{30, 100\}$ drawn from four trivariate elliptical distributions (Gaussian, t_5, t_3, and t_1 distributions). In each case, $\ell = 0$ corresponds to the null hypothesis and $\ell = 1, 2, 3, 4$ provide increasingly severe alternatives; see Section 5 for details.

6. Lower-order U-statistics involving the sample mean

The representation results in Theorem 2.1 provide U-statistic expressions for $\mathrm{adj}(s)$, $\mathrm{adj}(s)\bar{s}$ and $\det(s) + \bar{x}'\mathrm{adj}(s)\bar{x}$, in the same way (1.2) and (1.4) yield U-statistic expressions for S and $\det(S)$. For S and $\det(S)$, (1.1) and (1.5) rather provide corresponding formulas involving the sample mean $\bar{X}$ and one order less in the respective sums. A natural question is thus whether one can obtain similar formulas for the statistics involved in Theorem 2.1. As the next result shows, the answer is positive.

Theorem 6.1. *For any $x_1, \ldots, x_n \in \mathbb{R}^d$,*

$$(i)\quad \frac{((d-1)!)^2}{n^{d-1}} \sum_{1 \le i_1 < \ldots < i_{d-1} \le n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, \bar{x}))\Gamma_{x_{i_1}, \ldots, x_{i_{d-1}}, \bar{x}} = \mathrm{adj}(s),$$

$$(ii)\quad \frac{((d-1)!)^2}{n^{d-1}} \sum_{1 \le i_1 < \ldots < i_{d-1} \le n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, \bar{x}))\Gamma_{x_{i_1}, \ldots, x_{i_{d-1}}, \bar{x}} x_{i_{d-1}} = \mathrm{adj}(s)\bar{x},$$

and

$$(iii)\quad \frac{((d-1)!)^2}{n^{d-1}} \sum_{1 \le i_1 < \ldots < i_{d-1} \le n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, \bar{x}))x'_{i_{d-1}}\Gamma_{x_{i_1}, \ldots, x_{i_{d-1}}, \bar{x}} x_{i_{d-1}} = \bar{x}'\mathrm{adj}(s)\bar{x},$$

where sums over an empty collection of indices are still defined as being equal to zero.

We will need the following crucial result (its proof, that is quite technical, is deferred to the appendix).

Lemma 6.1. *Let* $x_1,\ldots,x_n \in \mathbb{R}^d$*, with* $n \geq d$*. Then, defining*

$$T_n(x,\mu) := \sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},x))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},x}(x-\mu)$$

and

$$S_n(x,\mu) := \sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},\mu))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\mu}(x-\mu)$$

for any $x,\mu \in \mathbb{R}^d$*, we have* $\sum_{i=1}^n T_n(x_i,\mu) = d\sum_{i=1}^n S_n(x_i,\mu)$*.*

Proof of Theorem 6.1. For $n \leq d-1$, the adjugate matrix $\mathrm{adj}(s)$ is the zero matrix, so that the result trivially holds (for $n \leq d-2$, the sums in the U-statistics expression are over an empty set of indices, hence are equal to zero, whereas for $n = d-1$, each U-statistic in the statement involves a single term, associated with $(i_1,\ldots,i_{d-1}) = (1,\ldots,d-1)$, and this term is zero since $m_{d-1}(\mathrm{Simpl}(x_1,\ldots,x_{d-1},\bar{x})) = 0$).

We may thus focus on the case $n \geq d$. First note that the

$$\begin{aligned}\sum_{i=1}^n T_n(x_i,\mu) &= \sum_{i=1}^n \sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},x_i))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},x_i}(x_i-\mu)\\ &= \frac{1}{(d-1)!}\sum_{1\leq i_1,\ldots,i_d\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},x_{i_d}))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d}}(x_{i_d}-\mu)\\ &= d\sum_{1\leq i_1<\cdots<i_d\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},x_{i_d}))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d}}(x_{i_d}-\mu)\end{aligned}$$

and

$$\begin{aligned}\sum_{i=1}^n S_n(x_i,\mu) &= \sum_{i=1}^n \sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},\mu))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\mu}(x_i-\mu)\\ &= n\sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},\mu))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\mu}(\bar{x}-\mu).\end{aligned}$$

Lemma 6.1 thus entails that

$$\begin{aligned}&\frac{((d-1)!)^2}{n^d}\sum_{1\leq i_1<\cdots<i_d\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},x_{i_d}))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d}}(x_{i_d}-\mu)\\ &= \frac{((d-1)!)^2}{n^{d-1}}\sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},\mu))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\mu}(\bar{x}-\mu),\end{aligned}$$

which, in view of Theorem 2.1, rewrites

$$\mathrm{adj}(S)(\bar{x}-\mu) = \frac{((d-1)!)^2}{n^{d-1}} \sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},\mu))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\mu}(\bar{x}-\mu).$$

Letting $\mu = \lambda\bar{x}$, with $\lambda \neq 1$, we then have

$$\mathrm{adj}(x)\bar{x} = \frac{((d-1)!)^2}{n^{d-1}} \sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},\lambda\bar{x}))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\lambda\bar{x}}\bar{x},$$

which, from continuity, provides

$$\mathrm{adj}(s)\bar{x} = \frac{((d-1)!)^2}{n^{d-1}} \sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},\bar{x}))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\bar{x}}\bar{x}. \tag{6.1}$$

Now, replacing x_i with $x_i + v$, we have

$$\mathrm{adj}(s)(\bar{x}+v) = \frac{((d-1)!)^2}{n^{d-1}} \sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},\bar{x}))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\bar{x}}(\bar{x}+v). \tag{6.2}$$

Subtracting (6.1) from (6.2) then yields

$$\mathrm{adj}(s)v = \frac{((d-1)!)^2}{n^{d-1}} \sum_{1\leq i_1<\ldots<i_{d-1}\leq n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1},\ldots,x_{i_{d-1}},\bar{x}))\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\bar{x}}v.$$

Since this holds for any v, this establishes Part (i) of the result. Parts (ii)–(iii) then follow by using the fact that $\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\bar{x}}(x_{i_{d-1}}-\bar{x}) = 0 = (x_{i_{d-1}}-\bar{x})'\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\bar{x}}$. □

7. Conclusions and perspectives for future research

In this work, we established a series of geometric identities involving the sample covariance matrix and its adjugate, thereby complementing results that were so far restricted to scalar functionals. These identities, that do not impose any restriction on the sample size n and the dimension d, involve U-statistics whose kernels are based on measures of random simplices and associated projection matrices. They find applications both in point estimation and in hypothesis testing. Regarding point estimation, they naturally yield, irrespective of the possible singularity of Σ, unbiased estimators of several population quantities, including $\mathrm{adj}(\Sigma)$, $\mathrm{adj}(\Sigma)\mu$, and $\mu'\mathrm{adj}(\Sigma)\mu$. These estimators were shown to be uniformly minimum risk unbiased (UMRU) under very broad conditions. In hypothesis testing, our U-statistic representations offer in particular new tools for testing partial uncorrelatedness in elliptical models. The proposed median-based tests, that rely on the robust concept of U-quantiles rather than on U-statistics, do not require moment assumptions and perform well under heavy-tailed distributions.

Our work opens several perspectives for future research. One direction is to investigate whether similar geometric representations can be obtained for other matrix functionals, such as the Moore–Penrose inverse or the matrix square root, and to explore their implications for estimation and testing. Another possible direction is to extend our results to structured covariance models, including those arising in high-dimensional settings or under sparsity constraints. Finally, the tests of partial uncorrelatedness we introduced involve a computational burden of order $O(n^d)$, hence can be applied for rather small dimensions d only. For larger d, incomplete U-statistics can potentially solve this issue; we refer to Dürre and Paindaveine (2022a). More generally, the concept of partial median-uncorrelatedness we introduced deserves further study, both from a theoretical perspective and in terms of its applications to graphical modeling and robust inference.

Appendix A: Technical proofs

The following results follow from Lemma A.2 in Paindaveine (2022) and Lemma S.1.2 in Dürre and Paindaveine (2022b), respectively.

Lemma A.1. *Fix a positive integer d. For any $x_1, \ldots, x_{d+1} \in \mathbb{R}^d$,*

$$m_d(\mathrm{Simpl}(x_1, \ldots, x_d, x_{d+1})) = \frac{1}{d} m_{d-1}(\mathrm{Simpl}(x_1, \ldots, x_d)) \|\Gamma(x_{d+1} - x_d)\|,$$

where we wrote $\Gamma = \Gamma_{x_1,\ldots,x_d}$.

Lemma A.2. *Fix a positive integer d. Let $X_1, \ldots, X_d$ form a random sample from P with finite moments of order two. Then, for any $r \in \{0, 1, 2\}$,*

$$\mathrm{E}[m_{d-1}^2(\mathrm{Simpl}(X_1, \ldots, X_d)) \|\Gamma_{x_1,\ldots,x_d} X_d\|^r]$$

exists and is finite.

We can now prove Lemma 3.1.

Proof of Lemma 3.1. We first show that the expectations on the lefthand sides of (i)–(iii) exist and are finite. Denote as e_k the kth vector of the canonical basis of $\mathbb{R}^d$. Since $\Gamma_{x_1,\ldots,x_d}$ is a projection matrix, we have that, for any $k, \ell = 1, \ldots, d$, $|(\Gamma_{x_1,\ldots,x_d})_{k\ell}| = |e_k' \Gamma_{x_1,\ldots,x_d} e_\ell| \leq \|\Gamma_{x_1,\ldots,x_d} e_\ell\| \leq 1$, hence also

$$\mathrm{E}[m_{d-1}^2(\mathrm{Simpl}(X_1, \ldots, X_d)) |(\Gamma_{x_1,\ldots,x_d})_{k\ell}|] \leq \mathrm{E}[m_{d-1}^2(\mathrm{Simpl}(X_1, \ldots, X_d))],$$

where the second expectation exists and is finite (see Lemma A.2 with $r = 0$). For any $k = 1, \ldots, d$,

$$\mathrm{E}[m_{d-1}^2(\mathrm{Simpl}(X_1, \ldots, X_d)) |(\Gamma_{x_1,\ldots,x_d} X_d)_k|] \leq \mathrm{E}[m_{d-1}^2(\mathrm{Simpl}(X_1, \ldots, X_d)) \|\Gamma_{x_1,\ldots,x_d} X_d\|],$$

where the second expectation exists and is finite (see Lemma A.2 with $r = 1$). Finally, since $\Gamma_{x_1,\ldots,x_d}$ is symmetric and idempotent,

$$\mathrm{E}[m_{d-1}^2(\mathrm{Simpl}(X_1, \ldots, X_d)) X_d' \Gamma_{x_1,\ldots,x_d} X_d] = \mathrm{E}[m_{d-1}^2(\mathrm{Simpl}(X_1, \ldots, X_d)) \|\Gamma_{x_1,\ldots,x_d} X_d\|^2]$$

exists and is finite (see Lemma A.2 with $r = 2$). We conclude that the expectations on the lefthand sides of (i)–(iii) indeed exist and are finite.

Parallel to the proof of Theorem 2.1, the proof then proceeds by computing in two different ways

$$r(x) := \mathrm{E}[m_d^2(\mathrm{Simpl}(X_1, \ldots, X_d, x))]$$

for any $x \in \mathbb{R}^d$. Note that

$$m_d(\mathrm{Simpl}(X_1,\ldots,X_d,x)) = \frac{1}{d} m_{d-1}(\mathrm{Simpl}(X_1,\ldots,X_d))\|\Gamma_{x_1,\ldots,x_d}(x - X_d)\|.$$

Irrespective of whether $m_{d-1}(\mathrm{Simpl}(X_1,\ldots,X_d))$ is zero or not, there exists a unique (up to an unimportant sign) d-vector V such that

$$VV' = m^2_{d-1}(\mathrm{Simpl}(X_1,\ldots,X_d))\Gamma_{x_1,\ldots,x_d}. \tag{A.1}$$

Since $\Gamma_{x_1,\ldots,x_d}$ is a symmetric idempotent matrix, we then have

$$\begin{aligned}
&m^2_d(\mathrm{Simpl}(X_1,\ldots,X_d,x)) \\
&\quad = \frac{1}{d^2} m^2_{d-1}(\mathrm{Simpl}(X_1,\ldots,X_d))\|\Gamma_{x_1,\ldots,x_d}(X_d - x)\|^2 \\
&\quad = \frac{1}{d^2}(X_d - x)'VV'(X_d - x),
\end{aligned}$$

which yields

$$r(x) = \frac{1}{d^2} x'\mathrm{E}[VV']x - \frac{2}{d^2} x'\mathrm{E}[VV'X_d] + \frac{1}{d^2}\mathrm{E}[\|V'X_d\|^2]. \tag{A.2}$$

Now, letting $C = (X_1 - x \ \ldots \ X_d - x)$, we also have

$$m_d(\mathrm{Simpl}(X_1,\ldots,X_d,x)) = \frac{1}{d!}|\det C|,$$

so that Theorem 4.1 provides

$$\begin{aligned}
r(x) = \frac{1}{(d!)^2}\mathrm{E}[\det(CC')] &= \frac{1}{(d!)^2}\,\mathrm{E}\left[\det\left(\sum_{i=1}^d (X_i - x)(X_i - x)'\right)\right] \\
&= \frac{1}{d!}\det(\mathrm{E}[(X_1 - x)(X_1 - x)']) \\
&= \frac{1}{d!}\det(\Sigma + (x - \mu)(x - \mu)') \\
&= d!\{\det(\Sigma) + (x - \mu)'\mathrm{adj}(\Sigma)(x - \mu)\}.
\end{aligned} \tag{A.3}$$

Since x is arbitrary, (A.2)–(A.3) provide

$$\frac{1}{d^2}\mathrm{E}[VV'] = \frac{1}{d!}\,\mathrm{adj}(\Sigma),$$

$$\frac{1}{d^2}\mathrm{E}[VV'X_d] = \frac{1}{d!}\,\mathrm{adj}(\Sigma)\mu,$$

and

$$\frac{1}{d^2}\mathrm{E}[X_d' V V' X_d] = \frac{1}{d!}\{\det(\Sigma) + \mu' \mathrm{adj}(\Sigma)\mu\},$$

which, in view of (A.1), establishes the result. □

PROOF OF PROPOSITION 5.1. To make the notation simpler, we will restrict to the case $(k,\ell) = (1,2)$, which is of course without any loss of generality. We thus assume that $\rho_{12} = 0$, hence that $\Sigma^{-1}_{12} = 0$. Partition the scatter matrix Σ into

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix},$$

where Σ_{aa} is a 2×2 matrix. Conditional on $U = (X^{(3)}, \ldots, X^{(d)})$, the distribution of $(X^{(1)}, X^{(2)})$ is then elliptically symmetric with shape matrix

$$\Omega = \Sigma_{aa} - \Sigma_{ab}\Sigma^{-1}_{bb}\Sigma_{ba};$$

see, e.g., Corollary 5 in Cambanis, Huang and Simons (1981). The block-inverse formula (see, e.g., Equation (7) on page 12 of Magnus and Neudecker, 2007) yields that $\Omega^{-1}_{12} = \Sigma^{-1}_{12}$ (invertibility of Σ of course guarantees invertibility of Ω). Since $\Sigma^{-1}_{12} = 0$, it follows that Ω^{-1} is a diagonal matrix, hence so is Ω. From ellipticity, this implies that, conditional on U, the random vectors $(-X^{(1)}, X^{(2)})$ and $(X^{(1)}, X^{(2)})$ are equally distributed. It follows that, conditional on $U_1, \ldots, U_d$ (with $U_j := (X^{(3)}_j, \ldots, X^{(d)}_j)$), the distribution of

$$\begin{aligned} Z &:= m^2_{d-1}(\mathrm{Simpl}(X_1, \ldots, X_d))(\Gamma_{X_1,\ldots,X_d})_{12} \\ &= -\frac{1}{((d-1)!)^2}\det(\Psi'\Psi) \times (X^{(1)}_2 - X^{(1)}_1 \ \ldots \ X^{(1)}_d - X^{(1)}_1)(\Psi'\Psi)^{-1}\begin{pmatrix} X^{(2)}_2 - X^{(2)}_1 \\ \vdots \\ X^{(2)}_d - X^{(2)}_1 \end{pmatrix} \qquad \text{(A.4)} \end{aligned}$$

is symmetric about zero[6]. Since symmetry of the conditional distribution implies symmetry of the unconditional distribution, we deduce that $\rho^M_{12} = 0$ (under the assumption of the proposition, the distribution of Z is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$). □

PROOF OF LEMMA 6.1. Fix $x_1, \ldots, x_n \in \mathbb{R}^d$ and consider the function

$$(x,\mu) \mapsto L(x_1, \ldots, x_n; x, \mu) := \frac{d^2}{2}\sum_{1 \le i_1 < \ldots < i_{d-1} \le n} m^2_d(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, x, \mu)).$$

From Lemma A.1, we have

$$\begin{aligned} &\nabla_x L(X_1, \ldots, X_n; x, \mu) \\ &\quad = \frac{1}{2}\nabla_x \sum_{1 \le i_1 < \ldots < i_{d-1} \le n} m^2_{d-1}(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, \mu))\|\Gamma_{x_{i_1},\ldots,x_{i_{d-1}},\mu}(x-\mu)\|^2 \end{aligned}$$

[6] When $\det(\Psi'\Psi) = 0$, we define the righthand side of (A.4) as being zero.

$$= S_n(x, \mu) \tag{A.5}$$

and

$$\nabla_\mu L(X_1, \ldots, X_n; x, \mu)$$

$$= \frac{1}{2} \nabla_\mu \sum_{1 \le i_1 < \ldots < i_{d-1} \le n} m_{d-1}^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, x)) \|\Gamma_{x_{i_1}, \ldots, x_{i_{d-1}}, x}(x - \mu)\|^2$$

$$= -T_n(x, \mu). \tag{A.6}$$

Establishing Lemma 6.1 requires obtaining alternative expressions of these gradients.

To do so, write

$$m_d^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, x, \mu)) = \frac{1}{(d!)^2} \{\det(D_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu})\}^2,$$

where $D_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu}$ is the $d \times d$ matrix whose columns are $X_{i_1} - \mu, \ldots, X_{i_{d-1}} - \mu$ and $x - \mu$. Recall Jacobi's formula

$$\frac{d}{dt} \det(A(t)) = \mathrm{tr}\left(\mathrm{adj}(t) \frac{d}{dt} A(t)\right);$$

see, e.g., (1) on page 169 of Magnus and Neudecker (2007). Since

$$\frac{\partial}{\partial x_k} D_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu} = e_k e_d',$$

Jacobi's formula yields

$$\frac{\partial}{\partial x_k} m_d^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, x, \mu))$$

$$= \frac{2}{(d!)^2} \det(D_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu}) \mathrm{tr}\left[\mathrm{adj}(D_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu}) \frac{\partial}{\partial x_k} D_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu}\right]$$

$$= \frac{2}{(d!)^2} \det(D_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu}) e_k' \mathrm{adj}(D'_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu}) e_d, \tag{A.7}$$

where we have used the identity $(\mathrm{adj}(A))' = \mathrm{adj}(A')$. From (A.5), we thus have

$$S_n(x, \mu) = \frac{1}{((d-1)!)^2} \sum_{1 \le i_1 < \ldots < i_{d-1} \le n} \det(D_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu}) \mathrm{adj}(D'_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu}) e_d. \tag{A.8}$$

Aiming at obtaining a corresponding expression for $T_n(x, \mu)$, note that

$$\frac{\partial}{\partial \mu_k} D_{x_{i_1}, \ldots, x_{i_{d-1}}, x; \mu} = -e_k 1_d',$$

which, using Jacobi's formula, yields

$$\frac{\partial}{\partial \mu_k} m_d^2(\mathrm{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, x, \mu))$$

$$= \frac{2}{(d!)^2} \det(D_{x_{i_1},\ldots,x_{i_{d-1}},x;\mu}) \operatorname{tr}\Big[\operatorname{adj}(D_{x_{i_1},\ldots,x_{i_{d-1}},x;\mu}) \frac{\partial}{\partial \mu_k} D_{x_{i_1},\ldots,x_{i_{d-1}},x;\mu}\Big]$$

$$= -\frac{2}{(d!)^2} \det(D_{x_{i_1},\ldots,x_{i_{d-1}},x;\mu}) e_k' \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x;\mu}) 1_d.$$

From (A.6), this provides

$$T_n(x,\mu) = \frac{1}{((d-1)!)^2} \sum_{1\le i_1<\ldots<i_{d-1}\le n} \det(D_{x_{i_1},\ldots,x_{i_{d-1}},x;\mu}) \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x;\mu}) 1_d.$$

Thus, we have

$$\sum_{i=1}^n T_n(X_i,\mu) = \frac{1}{((d-1)!)^2} \sum_{i=1}^n \sum_{1\le i_1<\ldots<i_{d-1}\le n} \det(D_{x_{i_1},\ldots,x_{i_{d-1}},x_i;\mu}) \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x_i;\mu}) 1_d$$

$$= \frac{1}{((d-1)!)^3} \sum_{1\le i_1,\ldots,i_d\le n} \det(D_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) 1_d.$$

Now, with Π_k the $d\times d$ orthogonal matrix that permutes the kth and dth components of any d-vector (that is, the matrix obtained by swapping the kth and dth columns of I_d),

$$\sum_{i=1}^n T_n(X_i,\mu) = \frac{1}{((d-1)!)^3} \sum_{k=1}^d \sum_{1\le i_1,\ldots,i_d\le n} \det(D_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) e_k$$

$$= \frac{1}{((d-1)!)^3} \sum_{k=1}^d \sum_{1\le i_1,\ldots,i_d\le n} \det(D_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) \Pi_k e_d.$$

Since

$$\begin{aligned}
\operatorname{adj}((D_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}\Pi_k)') &= \operatorname{adj}(\Pi_k' D'_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) \\
&= \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) \operatorname{adj}(\Pi_k') \\
&= \det(\Pi_k') \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) (\Pi_k')^{-1} \\
&= \det(\Pi_k) \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) \Pi_k,
\end{aligned}$$

this yields

$$\sum_{i=1}^n T_n(X_i,\mu) = \frac{1}{((d-1)!)^3} \sum_{k=1}^d \sum_{1\le i_1,\ldots,i_d\le n} \det(D_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}\Pi_k) \operatorname{adj}((D_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}\Pi_k)') e_d$$

$$= \frac{1}{((d-1)!)^3} \sum_{k=1}^d \sum_{1\le i_1,\ldots,i_d\le n} \det(D_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) \operatorname{adj}(D'_{x_{i_1},\ldots,x_{i_{d-1}},x_{i_d};\mu}) e_d$$

$$= \frac{d}{((d-1)!)^3} \sum_{1 \le i_1, \ldots, i_d \le n} \det(D_{x_{i_1}, \ldots, x_{i_{d-1}}, x_{i_d}; \mu}) \operatorname{adj}(D'_{x_{i_1}, \ldots, x_{i_{d-1}}, x_{i_d}; \mu}) e_d$$

(the permutation matrix Π_k may be dropped since all permutations are considered anyway in the sum over $i_1, \ldots, i_d$). Finally, in view of (A.8),

$$\begin{aligned} \sum_{i=1}^{n} T_n(X_i, \mu) &= \frac{d}{((d-1)!)^3} \sum_{i_d=1}^{n} \sum_{1 \le i_1, \ldots, i_{d-1} \le n} \det(D_{x_{i_1}, \ldots, x_{i_{d-1}}, x_{i_d}; \mu}) \operatorname{adj}(D'_{x_{i_1}, \ldots, x_{i_{d-1}}, x_{i_d}; \mu}) e_d \\ &= \frac{d}{((d-1)!)^2} \sum_{i=1}^{n} \sum_{1 \le i_1 < \ldots < i_{d-1} \le n} \det(D_{x_{i_1}, \ldots, x_{i_{d-1}}, x_i; \mu}) \operatorname{adj}(D'_{x_{i_1}, \ldots, x_{i_{d-1}}, x_i; \mu}) e_d \\ &= d \sum_{i=1}^{n} S_n(X_i, \mu), \end{aligned}$$

where we used the fact that, for any i, the quantity

$$\det(D_{x_{i_1}, \ldots, x_{i_{d-1}}, x_i; \mu}) \operatorname{adj}(D'_{x_{i_1}, \ldots, x_{i_{d-1}}, x_{i_d}; \mu}) e_d$$

is symmetric in $i_1, \ldots, i_{d-1}$, which follows from the fact that this quantity is the gradient at x_i of the function

$$x \mapsto \frac{(d!)^2}{2} m_d^2(\operatorname{Simpl}(x_{i_1}, \ldots, x_{i_{d-1}}, x, \mu))$$

(see (A.7)) that is symmetric in $i_1, \ldots, i_{d-1}$. The result is proved. □

Acknowledgments

Davy Paindaveine is also affiliated at the Toulouse School of Economics, Université Toulouse 1 Capitole. The authors would like to thank the Editor-in-Chief Kengo Kato, the Associate Editor and two anonymous referees for their insightful comments and suggestions, that led to a substantial improvement of a previous version of this work.

Funding

Davy Paindaveine is supported by a “Projet de Recherche” grant from the FNRS (Fonds National pour la Recherche Scientifique), Communauté Française de Belgique. Alexander Dürre is supported by an Open Competition Domain Science - M grant from the NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek).

References

BODNAR, T., MAZUR, S. and PODGÓRSKI, K. (2016). Singular inverse Wishart distribution and its application to portfolio theory. *J. Multivariate Anal.* **143** 314–326. https://doi.org/10.1016/j.jmva.2015.09.021 MR3431434

BODNAR, T. and OKHRIN, Y. (2008). Properties of the singular, inverse and generalized inverse partitioned Wishart distributions. *J. Multivariate Anal.* **99** 2389–2045. https://doi.org/10.1016/j.jmva.2008.02.024 MR2463397

CAMBANIS, S., HUANG, S. and SIMONS, G. (1981). On the theory of elliptically contoured distributions. *J. Multivariate Anal.* **11** 368–385. https://doi.org/10.1016/0047-259X(81)90082-8 MR0629795

DRTON, M., MASSAM, H. and OLKIN, I. (2008). Moments of minors of Wishart matrices. *Ann. Statist.* **36** 2261–2283. https://doi.org/10.1214/07-AOS522 MR2458187

DÜRRE, A. and PAINDAVEINE, D. (2022a). Affine-equivariant inference for multivariate location under L_p loss functions. *Ann. Statist.* **50** 2616–2640. https://doi.org/10.1214/22-AOS2199 MR4500620

DÜRRE, A. and PAINDAVEINE, D. (2022b). Supplement to "Affine-equivariant inference for multivariate location under L_p loss functions". *Ann. Statist.* Online supplement. https://doi.org/10.1214/22-AOS2199SUPP

DÜRRE, A. and PAINDAVEINE, D. (2025). Supplement to "On some geometric identities involving the sample covariance matrix and its adjugate". Online supplement to the present submission.

HORN, R. A. and JOHNSON, C. R. (2013). *Matrix Analysis*, 2nd ed. Cambridge Univ. Press, New York.

JANSSEN, P., SERFLING, R. and VERAVERBEKE, N. (1984). Asymptotic normality for a general class of statistical functions and applications to measures of spread. *Ann. Statist.* **12** 1369–1379. https://doi.org/10.1214/aos/1176346797 MR0760694

MAGNUS, J. R. and NEUDECKER, H. (2007). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 3rd edition ed. John Wiley & Sons, Chichester.

MUIRHEAD, R. J. (2005). *Aspects of Multivariate Statistical Theory*. Wiley, Hoboken, New Jersey.

PAINDAVEINE, D. (2022). On the measure of anchored Gaussian simplices, with applications to multivariate medians. *Bernoulli* **28** 965–996. https://doi.org/10.3150/21-BEJ1373 MR4388926

PFANZAGL, J. (1994). *Parametric Statistical Theory*. Walter de Gruyter, Berlin.

PRONZATO, L. (1998). On a property of the expected value of a determinant. *Statist. Probab. Lett.* **39** 161–165. https://doi.org/10.1016/S0167-7152(98)00057-1 MR1652548

PRONZATO, L., WYNN, H. P. and ZHIGLJAVSKY, A. A. (2017). Extended generalised variances, with applications. *Bernoulli* **23** 2617–2642. https://doi.org/10.3150/16-BEJ821 MR3648040

PRONZATO, L., WYNN, H. P. and ZHIGLJAVSKY, A. A. (2018). Simplicial variances, potentials and Mahalanobis distances. *J. Multivariate Anal.* **168** 276–289. https://doi.org/10.1016/j.jmva.2018.08.002 MR3858365

SERFLING, R. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.

SERFLING, R. J. (1984). Generalized L-, M-, and R-statistics. *Ann. Statist.* **12** 76–86. https://doi.org/10.1214/aos/1176346393 MR0733500

SRIVASTAVA, M. S. (2003). Singular Wishart and multivariate beta distributions. *Ann. Statist.* **31** 1537–1560. https://doi.org/10.1214/aos/1065705118 MR2012825

UHLER, C. (2018). Gaussian graphical models. In *Handbook of Graphical Models* 217–238. CRC Press.

VAN DER VAART, H. R. (1965). A note on Wilks' internal scatter. *Ann. Math. Statist.* **36** 1308–1312. https://doi.org/10.1214/aoms/1177700006 MR178533

VOGEL, D. and FRIED, R. (2011). Elliptical graphical modelling. *Biometrika* **98** 935–951. https://doi.org/10.1093/biomet/asr037 MR2860334

WHITTAKER, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley Publishing, Chichester.

WILKS, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika* **24** 471–494. https://doi.org/10.2307/2331979