Exact expressions and numerical evaluation of average evolvability measures for characterizing and comparing $\mathbf{G}$ matrices

Junya Watanabe*

Abstract

Theory predicts that the additive genetic covariance ($\mathbf{G}$) matrix determines a population's short-term (in)ability to respond to directional selection—evolvability in the Hansen–Houle sense—which is typically quantified and compared via certain scalar indices called evolvability measures. Often, interest is in obtaining the averages of these measures across all possible selection gradients, but explicit formulae for most of these average measures have not been known. Previous authors relied either on approximations by the delta method, whose accuracy is generally unknown, or Monte Carlo evaluations (including the random skewers analysis), which necessarily involve random fluctuations. This study presents new, exact expressions for the average conditional evolvability, average autonomy, average respondability, average flexibility, average response difference, and average response correlation, utilizing their mathematical structures as ratios of quadratic forms. The new expressions are infinite series involving top-order zonal and invariant polynomials of matrix arguments, and can be numerically evaluated as their partial sums with, for some measures, known error bounds. Whenever these partial sums numerically converge within reasonable computational time, they will replace the previous approximate methods. In addition, new expressions are derived for the average measures under a general normal distribution for the selection gradient, extending the applicability of these measures into a substantially broader class of selection regimes.

Key words: evolutionary constraint; phenotypic integration; quadratic form; quantitative genetics; random skewers; zonal polynomial.

1 Introduction

The quantitative genetic theory of multivariate trait evolution provides a powerful framework to analyze and predict phenotypic evolution (Steppan *et al.*, 2002; Blows, 2007; Walsh & Blows, 2009). At the core of the theory is the Lande equation, which describes a population's response to directional selection under certain simplifying conditions (Lande, 1979; Lande & Arnold, 1983):

$$\Delta\bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta},$$

where $\boldsymbol{\beta}$ is a p-dimensional selection gradient vector (partial regression coefficients of trait values to relative fitness), $\mathbf{G}$ is a $p \times p$ (additive) genetic covariance matrix, and $\Delta\bar{\mathbf{z}}$ is a p-dimensional vector of per-generation change of the population's mean trait values. (Throughout the paper,

*Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge, CB2 3EQ, United Kingdom. Email: jw2098@cam.ac.uk. ORCID: https://orcid.org/0000-0002-9810-5286

p denotes the (nominal) number of the traits/variables analyzed, and all vectors are column vectors.) In light of this theory, the $\mathbf{G}$ matrix is supposed to represent the population's (in)ability to respond to directional selection—what has been termed evolvability by Hansen, Houle and colleagues (Houle, 1992; Hansen, 2003; Hansen *et al.*, 2003a,b, 2011; Hansen & Houle, 2008; Hansen & Pélabon, 2021).[1] The theory has motivated extensive investigations into aspects of the $\mathbf{G}$ matrix or its substitutes, ranging from theoretical and simulation-based analyses (e.g., Pavlicev & Hansen, 2011; Chevin, 2013; Hansen *et al.*, 2019) to empirical inter-population/specific comparisons (e.g., Cheverud, 1982, 1989; Porto *et al.*, 2009; Marroig *et al.*, 2009; Rolian, 2009; Bolstad *et al.*, 2014; Haber, 2016; Machado, 2020) of trait covariation.

Based on this framework, and following theoretical developments in univariate cases (Hansen, 2003; Hansen *et al.*, 2003a,b), Hansen & Houle (2008) proposed several scalar indices to capture certain aspects of the multivariate evolvability represented in $\mathbf{G}$ (see also Kirkpatrick, 2009; Marroig *et al.*, 2009). Of these, (unconditional) evolvability e, conditional evolvability c, autonomy a, integration i, respondability r, and flexibility f (due to Marroig *et al.*, 2009) concern a single $\mathbf{G}$ matrix (Fig. 1A), whereas response difference d concerns comparison between two $\mathbf{G}$ matrices. Related to the latter is the use of response correlation ρ for matrix comparison (Cheverud, 1996; Cheverud & Marroig, 2007) (Fig. 1B). Notably, all these indices are simple or multiple ratios of quadratic forms in $\boldsymbol{\beta}$ (see below). In this paper, they are collectively called evolvability measures. Many of them can be related to the rate of adaptation or evolutionary constraint/bias (e.g., Hansen & Houle, 2008; Chevin, 2013; Bolstad *et al.*, 2014; Hansen *et al.*, 2019), providing means to characterize and compare $\mathbf{G}$ matrices in biologically meaningful ways.

The evolvability measures are primarily defined with respect to a given selection gradient $\boldsymbol{\beta}$. In this sense, they represent aspects of $\mathbf{G}$ matrices under a fixed, deterministic selection. In many theoretical and comparative analyses of evolvability (see references above), however, interest is often in characterizing $\mathbf{G}$ matrices without referring to $\boldsymbol{\beta}$. For this purpose, it is sensible to take the expectation of an evolvability measure assuming a certain probability distribution of $\boldsymbol{\beta}$. Most typically, although not necessarily, the uniform distribution on a unit hypersphere is assumed for this purpose, representing a completely randomly directed and uncorrelated selection regime. The resultant quantities are called average measures hereafter.[2] Biologically speaking, the average measures may be regarded as representations of general evolvability (in the sense of Riederer *et al.*, 2022), as compared to specific evolvability which is represented by the evolvability measures as functions of a fixed $\boldsymbol{\beta}$.

However, there is a major obstacle to using average measures for practical or even theoretical investigations. That is, most of the average measures lack known explicit, exact expressions beyond as expectations, except for the average evolvability $\bar{e}$ and, under certain special conditions, the average conditional evolvability $\bar{c}$ and the average respondability $\bar{r}$ (Hansen & Houle, 2008;

[1] The term evolvability bears various meanings in the current evolutionary biology (e.g., Brown, 2014; Crother & Murray, 2019; Feiner *et al.*, 2021; Jablonski, 2022; Love *et al.*, 2022; Riederer *et al.*, 2022). In the present paper, the use of this term is restricted to certain mathematical aspects of the $\mathbf{G}$ matrix and response vectors—the Hansen–Houle evolvability—unless stated otherwise.

[2] A potential alternative to the average measures in the above sense is the average of the evolvability measures calculated from the $\boldsymbol{\beta}$'s aligned with all individual traits (Armbruster *et al.*, 2009; Haber, 2011; Pavlicev & Hansen, 2011; Machado *et al.*, 2018). Although that would be valid for descriptive purposes, its interpretability in the context of multivariate selection might be limited. In any case, the calculation of such averages is straightforward and not concerned herein.

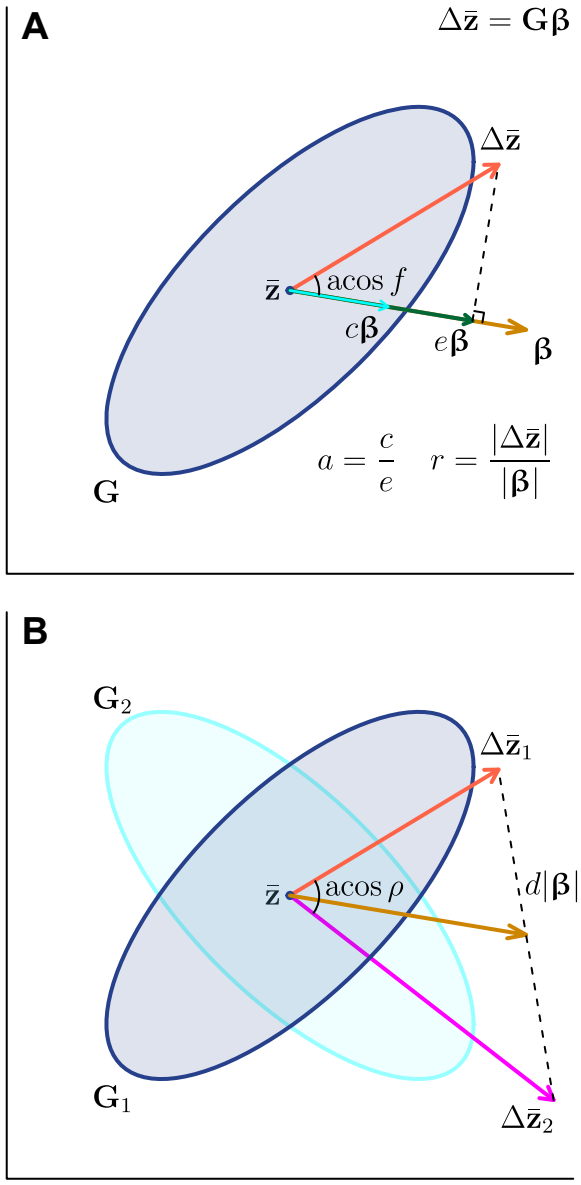

Figure 1. Geometric interpretations of evolvability measures. In a hypothetical bivariate trait space, genetic covariance matrices $\mathbf{G}$ are schematically represented by equiprobability ellipses around the mean phenotype $\bar{\mathbf{z}}$, and selection gradients $\boldsymbol{\beta}$ by acute-headed brown arrows. Response vectors $\Delta\bar{\mathbf{z}}$ are matrix products of $\mathbf{G}$ and $\boldsymbol{\beta}$, represented by broad-headed red/magenta arrows. **A**) One-matrix measures: evolvability e is the norm of $\Delta\bar{\mathbf{z}}$ projected on $\boldsymbol{\beta}$, standardized by the norm of the latter, $|\boldsymbol{\beta}|$; conditional evolvability c is the same but when $\Delta\bar{\mathbf{z}}$ is constrained along $\boldsymbol{\beta}$; autonomy a is the ratio of c to e; respondability r is $|\Delta\bar{\mathbf{z}}|$ standardized by $|\boldsymbol{\beta}|$; and flexibility f is the similarity in direction between $\Delta\bar{\mathbf{z}}$ and $\boldsymbol{\beta}$. **B**) Two-matrix comparison measures: response difference d is the distance between the end points of the two $\Delta\bar{\mathbf{z}}$'s standardized by $|\boldsymbol{\beta}|$, and response correlation ρ is their similarity in direction.

Kirkpatrick, 2009). Consequently, previous studies had to rely on approximations of average measures. Presumably the most widely used approximation method is Monte Carlo evaluation (e.g., Marroig *et al.*, 2009; Bolstad *et al.*, 2014; Haber, 2016; Grabowski & Porto, 2017), in which a large number of $\boldsymbol{\beta}$ is randomly generated on a computer and the resultant values calculated with the primary definition are literally averaged to yield an estimate of average measure. This has long been done in the so-called random skewers analysis for matrix comparison using response correlation ρ (e.g., Cheverud, 1996; Cheverud & Marroig, 2007; Revell, 2007; see also Rohlf, 2017), and is implemented in the `R` package `evolqg` (Melo *et al.*, 2016) for calculating various average measures. This method necessarily involves random fluctuations in the estimate and can take a large computational time, although the latter is becoming less of a concern with modern-day computers. On the other hand, Hansen & Houle (2008, 2009) themselves provided approximate expressions for average measures based on the delta method (which they called "analytical" or "standard" approximations). Apart from ill-defined notations used therein, a practical problem there is that the accuracy of this sort of approximation is generally unknown, so a separate Monte Carlo evaluation is usually required to ascertain whether the delta method approximation has an acceptable accuracy. This approximation has been used in some subsequent studies (e.g., Hansen & Voje, 2011; Brommer, 2014; Houle & Meyer, 2015) and is implemented in the `R` package `evolvability` (Bolstad *et al.*, 2014). Apart from these methods, Kirkpatrick (2009) used numerical integration to evaluate his average selection response, a measure equivalent to respondability (below), but this method has not been widely used in evaluating average measures.

This technical paper provides exact expressions for the following average measures: average evolvability $\bar{e}$, average conditional evolvability $\bar{c}$, average autonomy $\bar{a}$, average respondability $\bar{r}$, average flexibility $\bar{f}$, average response difference $\bar{d}$, and average response correlation $\bar{\rho}$. These expressions are derived from existing and new results on the moments of simple or multiple ratios of quadratic forms in normal variables. They are expressed as infinite series involving zonal or invariant polynomials of matrix arguments, and in most cases can be numerically evaluated as their partial sums to yield quasi-exact values. For some of the expressions, an upper bound for the truncation error is available. In addition to the special condition where $\boldsymbol{\beta}$ is spherically distributed as assumed in most previous studies (but see also Chevin, 2013), this study also concerns the general condition that $\boldsymbol{\beta}$ is normally distributed with potentially nonzero mean and nonspherical covariance. The latter condition can be of substantial biological interest, as it can model a fairly wide range of random directional and/or correlated selection regimes.

The paper is structured as follows. Section 2 first reviews the definition of the evolvability measures; some of them are redefined to accommodate potentially singular $\mathbf{G}$ matrices. After reviewing some known results on the moment of a ratio of quadratic forms, it then provides new expressions for average evolvability measures under the spherical distribution of $\boldsymbol{\beta}$. Section 3 presents a new `R` implementation of the analytic results and evaluate its performance by numerical experiments. Section 4 concludes the main body of the paper by adding some theoretical and practical considerations on the use of average measures. As the zonal and invariant polynomials appear to have rarely been used in the biological literature, Appendix A gives a brief overview on their theories, providing a basis for the present mathematical results. Appendix B states a new result on the moment of a multiple ratio of quadratic forms in normal variables, and

Appendix C presents new expressions for average measures under the general normal distribution of $\boldsymbol{\beta}$. Appendix D clarifies connections to previous results derived under special conditions (Hansen & Houle, 2008; Kirkpatrick, 2009).

2 Theory

2.1 Notations

In the following discussion, it is convenient to define the linear combination of the variables (traits) along the direction of $\boldsymbol{\beta}$. For this purpose, consider the decomposition $\boldsymbol{\beta} = |\boldsymbol{\beta}|\mathbf{u}$, where $|\cdot|$ is the vector norm or length, $|\mathbf{a}| = \sqrt{\mathbf{a}^T\mathbf{a}}$ for an arbitrary vector $\mathbf{a}$ (the superscript ^{T} denotes matrix transpose), and $\mathbf{u}$ is a unit vector: $\mathbf{u}^T\mathbf{u} = 1$. Define the $p \times (p-1)$ matrix $\mathbf{U}_{(-1)}$ so that the matrix $\mathbf{U} = (\mathbf{u}, \mathbf{U}_{(-1)})$ is orthogonal. With these, the orthogonal linear transform of the variables $\mathbf{z}^* = \mathbf{U}^T\mathbf{z}$ will be considered; the first entry of this new vector $\mathbf{u}^T\mathbf{z}$ represents the score along the direction of $\boldsymbol{\beta}$. Note that the covariance matrix of the new variables $\mathbf{z}^*$ can be written as

$$\mathbf{G}^* = \mathbf{U}^T\mathbf{G}\mathbf{U} = \begin{pmatrix} \mathbf{u}^T\mathbf{G}\mathbf{u} & \mathbf{u}^T\mathbf{G}\mathbf{U}_{(-1)} \\ \mathbf{U}_{(-1)}^T\mathbf{G}\mathbf{u} & \mathbf{U}_{(-1)}^T\mathbf{G}\mathbf{U}_{(-1)} \end{pmatrix}.$$

Hansen & Houle (2008) defined some of the evolvability measures under the constraint $|\boldsymbol{\beta}| = 1$, or $\boldsymbol{\beta} = \mathbf{u}$ in the notation here. For the discussions that follow, both forms in terms of $\boldsymbol{\beta}$ and $\mathbf{u}$ are presented. Throughout this paper, $\mathbf{G}$ is assumed to be given rather than random and validly constructed as a covariance matrix, that is, symmetric and nonnegative definite (i.e., either positive definite or positive semidefinite, in the terminology of Schott, 2016).

Although Hansen & Houle (2008) assumed $\mathbf{G}$ to be positive definite, this assumption is not imposed here to accommodate potentially singular (and positive semidefinite) $\mathbf{G}$ where possible. This is done by using the generalized inverse; for a matrix $\mathbf{A}$, a generalized inverse $\mathbf{A}^-$ is such a matrix that satisfies $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$. If $\mathbf{A}$ is nonsingular, $\mathbf{A}^- = \mathbf{A}^{-1}$.

For notational simplicity, the range or column space and the null space of a nonzero $p \times q$ matrix $\mathbf{A}$ is defined: $R(\mathbf{A}) = \{\mathbf{y} : \mathbf{y} = \mathbf{A}\mathbf{x}, \mathbf{x} \in \mathbb{R}^q\}$, and $N(\mathbf{A}) = \{\mathbf{x} : \mathbf{0}_p = \mathbf{A}\mathbf{x}, \mathbf{x} \in \mathbb{R}^q\}$, where $\mathbf{0}_p$ is the p-dimensional vector of zeros. When $\mathbf{A}$ is nonnegative definite, these are the spaces spanned by the eigenvectors corresponding to its nonzero and zero eigenvalues, respectively. The p-dimensional identity matrix is denoted by $\mathbf{I}_p$.

2.2 Evolvability measures: fixed selection

Evolvability $e(\boldsymbol{\beta})$ is the variance in $\mathbf{G}$ along the direction of selection $\mathbf{u}$, or equivalently the norm of the response vector $\Delta\bar{\mathbf{z}}$ projected onto $\boldsymbol{\beta}$, standardized by $|\boldsymbol{\beta}|$ (Hansen & Houle, 2008) (Fig. 1):

$$\begin{aligned} e(\boldsymbol{\beta}) =& \frac{\boldsymbol{\beta}^T \Delta\bar{\mathbf{z}}}{\boldsymbol{\beta}^T\boldsymbol{\beta}} \\ =& \frac{\boldsymbol{\beta}^T \mathbf{G}\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}} \qquad (1) \\ =& \mathbf{u}^T\mathbf{G}\mathbf{u}. \qquad (2) \end{aligned}$$

The numerator in (1) can be related to the rate of adaptation (change in mean fitness) under simple directional selection (Agrawal & Stinchcombe, 2009; Chevin, 2013).

Conditional evolvability $c(\boldsymbol{\beta})$ is conceptually defined as the variance along $\mathbf{u}$ when evolution in other traits/directions is not allowed due to stabilizing selection (Hansen, 2003; Hansen *et al.*, 2003a, 2019). With above notations, it is the conditional variance of the first element of $\mathbf{z}^*$ given the other elements (redefined here after Hansen & Houle, 2008):

$$\begin{aligned} c(\boldsymbol{\beta}) =& \mathbf{u}^T\mathbf{G}\mathbf{u} - \mathbf{u}^T\mathbf{G}\mathbf{U}_{(-1)}\left(\mathbf{U}_{(-1)}^T\mathbf{G}\mathbf{U}_{(-1)}\right)^{-}\mathbf{U}_{(-1)}^T\mathbf{G}\mathbf{u} \\ =& (\mathbf{u}^T\mathbf{G}^{-}\mathbf{u})^{-} \quad \text{if } \boldsymbol{\beta}, \mathbf{u} \in R(\mathbf{G}), \qquad (3) \\ =& \frac{\boldsymbol{\beta}^T\boldsymbol{\beta}}{\boldsymbol{\beta}^T\mathbf{G}^{-}\boldsymbol{\beta}}. \qquad (4) \end{aligned}$$

Here, the equation of (3) is from well-known results on the (generalized) inverse of a partitioned matrix (e.g., Schott, 2016, theorems 7.1 and 7.12), for which Hansen & Houle (2008) restated a proof in nonsingular $\mathbf{G}$. A lack of conditional variance along $\boldsymbol{\beta}$ results in $c(\boldsymbol{\beta}) = 0$, which happens when $\mathbf{G}$ is singular and $\boldsymbol{\beta} \notin R(\mathbf{G})$; the expressions (3) and (4) do not hold in this case.

Autonomy $a(\boldsymbol{\beta})$ is defined as the complement of the squared multiple correlation of the first element of $\mathbf{z}^*$ with respect to the other elements:

$$\begin{aligned} a(\boldsymbol{\beta}) =& 1 - (\mathbf{u}^T\mathbf{G}\mathbf{u})^{-1}\mathbf{u}^T\mathbf{G}\mathbf{U}_{(-1)}\left(\mathbf{U}_{(-1)}^T\mathbf{G}\mathbf{U}_{(-1)}\right)^{-}\mathbf{U}_{(-1)}^T\mathbf{G}\mathbf{u} = \frac{c(\boldsymbol{\beta})}{e(\boldsymbol{\beta})} \\ =& (\mathbf{u}^T\mathbf{G}\mathbf{u}\mathbf{u}^T\mathbf{G}^{-}\mathbf{u})^{-1} \quad \text{if } c(\boldsymbol{\beta}) \neq 0, \qquad (5) \\ =& \frac{(\boldsymbol{\beta}^T\boldsymbol{\beta})^2}{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{G}^{-}\boldsymbol{\beta}}. \qquad (6) \end{aligned}$$

This quantity is undefined when $e(\boldsymbol{\beta}) = \mathbf{u}^T\mathbf{G}\mathbf{u} = 0$. Hansen & Houle (2008) also defined the integration $i(\boldsymbol{\beta})$, which is the squared correlation just mentioned: $i(\boldsymbol{\beta}) = 1 - a(\boldsymbol{\beta})$.

Respondability $r(\boldsymbol{\beta})$ is the magnitude of response standardized by that of selection, or the

ratio of the norms of $\Delta\bar{\mathbf{z}}$ and $\boldsymbol{\beta}$:

$$
\begin{aligned}
r(\boldsymbol{\beta}) =& \frac{|\Delta\bar{\mathbf{z}}|}{|\boldsymbol{\beta}|} \\
=& \sqrt{\frac{\boldsymbol{\beta}^T\mathbf{G}^2\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}}} && (7) \\
=& \sqrt{\mathbf{u}^T\mathbf{G}^2\mathbf{u}}. && (8)
\end{aligned}
$$

The definition and terminology of r here follows Hansen & Houle (2008), but it should be noted that Kirkpatrick (2009) independently devised an equivalent measure as the (relative) selection response (R in the latter author's notation) with a slightly different standardization.

Flexibility $f(\boldsymbol{\beta})$ quantifies similarity in direction between $\Delta\bar{\mathbf{z}}$ and $\boldsymbol{\beta}$ in terms of vector correlation (arccosine of the angle formed by two vectors), and is also the ratio of $e(\boldsymbol{\beta})$ to $r(\boldsymbol{\beta})$ (Hansen & Houle, 2008; Marroig *et al.*, 2009; Rolian, 2009):

$$
\begin{aligned}
f(\boldsymbol{\beta}) =& \frac{e(\boldsymbol{\beta})}{r(\boldsymbol{\beta})} = \frac{\boldsymbol{\beta}^T\Delta\bar{\mathbf{z}}}{|\boldsymbol{\beta}||\Delta\bar{\mathbf{z}}|} \\
=& \frac{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^T\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{G}^2\boldsymbol{\beta}}} && (9) \\
=& \frac{\mathbf{u}^T\mathbf{G}\mathbf{u}}{\sqrt{\mathbf{u}^T\mathbf{G}^2\mathbf{u}}}. && (10)
\end{aligned}
$$

This quantity is undefined if $|\Delta\bar{\mathbf{z}}| = 0$, i.e., $\boldsymbol{\beta} \in N(\mathbf{G})$. Otherwise, $0 < f(\boldsymbol{\beta}) \leq 1$ from the nonnegative definiteness of $\mathbf{G}$ and the definition of vector correlation. The term flexibility was coined by Marroig *et al.* (2009) due to Hansen's suggestion.

All the above indices are one-matrix measures that primarily concern characterization of a single $\mathbf{G}$ matrix. On the other hand, the following two indices concern pairwise comparisons between two $\mathbf{G}$ matrices in terms of (dis)similarity of the response vectors given the same $\boldsymbol{\beta}$. The two $\mathbf{G}$ matrices and $\Delta\bar{\mathbf{z}}$ are denoted here with subscripts: $\mathbf{G}_i$, $\Delta\bar{\mathbf{z}}_i$. Response difference $d(\boldsymbol{\beta})$ is a measure of dissimilarity, defined as a standardized difference between the two response vectors (Hansen & Houle, 2008):

$$
\begin{aligned}
d(\boldsymbol{\beta}) =& \frac{|\Delta\bar{\mathbf{z}}_1 - \Delta\bar{\mathbf{z}}_2|}{|\boldsymbol{\beta}|} \\
=& \sqrt{\frac{\boldsymbol{\beta}^T\left(\mathbf{G}_1 - \mathbf{G}_2\right)^2\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}}} && (11) \\
=& \sqrt{\mathbf{u}^T\left(\mathbf{G}_1 - \mathbf{G}_2\right)^2\mathbf{u}}. && (12)
\end{aligned}
$$

Response correlation $\rho(\boldsymbol{\beta})$ is a measure of similarity in direction, defined as the vector

correlation between the two response vectors (e.g., Cheverud, 1996; Cheverud & Marroig, 2007):

$$\rho(\boldsymbol{\beta}) = \frac{\Delta\bar{\mathbf{z}}_1^T \Delta\bar{\mathbf{z}}_2}{|\Delta\bar{\mathbf{z}}_1||\Delta\bar{\mathbf{z}}_2|}$$

$$= \frac{\boldsymbol{\beta}^T \mathbf{G}_1 \mathbf{G}_2 \boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^T \mathbf{G}_1^2 \boldsymbol{\beta} \boldsymbol{\beta}^T \mathbf{G}_2^2 \boldsymbol{\beta}}} \tag{13}$$

$$= \frac{\mathbf{u}^T \mathbf{G}_1 \mathbf{G}_2 \mathbf{u}}{\sqrt{\mathbf{u}^T \mathbf{G}_1^2 \mathbf{u} \mathbf{u}^T \mathbf{G}_2^2 \mathbf{u}}} \tag{14}$$

This quantity is undefined when at least one of $|\Delta\bar{\mathbf{z}}_1|$ and $|\Delta\bar{\mathbf{z}}_2|$ is 0. Otherwise, $-1 \leq \rho(\boldsymbol{\beta}) \leq 1$.[3]

2.3 Ratio of quadratic forms

A key to derive expressions for average measures is that the measures defined above are simple or multiple ratios of quadratic forms in $\boldsymbol{\beta}$ or $\mathbf{u}$. Ratios of quadratic forms, especially those in normal random variables, play a pivotal role in statistics, as many practically important test statistics can be written as a ratio of quadratic forms. Naturally, there is a vast body of literature regarding the distribution and moments of quadratic forms and ratios thereof (e.g., Magnus, 1986; Jones, 1986, 1987; Smith, 1989, 1993; Mathai & Provost, 1992; Hillier, 2001; Forchini, 2002; Hillier *et al.*, 2009, 2014; Bao & Kan, 2013).

Regarding the ratio of quadratic forms in normal variables, there are several equivalent ways to derive its moments. One starts from the joint moment-generation function (e.g., Cressie *et al.*, 1981; Meng, 2005) of quadratic forms, and typically results in integrations which does not have simple closed-form expressions (e.g., Magnus, 1986; Jones, 1986, 1987; Gupta & Kabe, 1998). Another way is to expand a ratio into infinite series, which in turn is integrated using the zonal and invariant polynomials (e.g., Smith, 1989, 1993; Hillier, 2001; Hillier *et al.*, 2009, 2014). (See Bao & Kan (2013) for connections between these two approaches.) The latter way typically yields an infinite series including zonal or invariant polynomials, which can be evaluated with reasonable speed and accuracy with the aid of recent algorithmic developments (Hillier *et al.*, 2009, 2014; Bao & Kan, 2013). This is the way followed in the present paper. The zonal polynomials are a special form of homogeneous polynomials in eigenvalues of a matrix which generalize powers of scalars into symmetric matrices (e.g., James, 1960, 1964; Muirhead, 1982, 2006; Gross & Richards, 1987; Mathai *et al.*, 1995). The invariant polynomials are a generalization of the zonal polynomials into multiple matrix arguments (Davis, 1979, 1980; Chikuse, 1980, 2003; Chikuse & Davis, 1986b). A brief overview on these mathematical tools is provided in Appendix A.

A general result for the moments of a simple ratio of quadratic forms in normal variables is restated in the following proposition.

Proposition 1 (Smith (1989), Bao & Kan (2013))**.** *Let* $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \mathbf{I}_p)$*,* $\mathbf{A}$ *be a* $p \times p$ *symmetric matrix,* $\mathbf{B}$ *be a* $p \times p$ *nonnegative definite matrix, and* m, n *be positive real numbers. When the*

[3] Rohlf (2017, p. 544) asserted that ρ cannot be negative because he assumed, inappropriately in this context, the response vectors to have no polarity. Negative ρ can occur when, e.g., $\boldsymbol{\beta}$ bisects the wider of the two angles formed by the major axes of highly integrated $\mathbf{G}$ matrices. Except for this subtle fallacy, his criticisms on previous interpretations of ρ or $\bar{\rho}$ (e.g., Melo *et al.*, 2016) seem largely pertinent (see also Discussion).

expectation of $\frac{(\mathbf{x}^T\mathbf{A}\mathbf{x})^m}{(\mathbf{x}^T\mathbf{B}\mathbf{x})^n}$ *exists, it can be written as*

$$\mathrm{E}\left(\frac{(\mathbf{x}^T\mathbf{A}\mathbf{x})^m}{(\mathbf{x}^T\mathbf{B}\mathbf{x})^n}\right) = 2^{m-n}\alpha_{\mathbf{A}}^{-m}\alpha_{\mathbf{B}}^{n}e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-m)_i\,(n)_j\,\Gamma\left(\frac{p}{2}+m-n+k\right)}{2^k\left(\frac{1}{2}\right)_k\Gamma\left(\frac{p}{2}+i+j+k\right)} \cdot \frac{\left(\frac{1}{2}\right)_{i+j+k}}{i!j!k!}C_{[i+j+k]}^{[i],[j],[k]}\left(\mathbf{I}_p-\alpha_{\mathbf{A}}\mathbf{A},\mathbf{I}_p-\alpha_{\mathbf{B}}\mathbf{B},\boldsymbol{\mu}^T\boldsymbol{\mu}\right), \tag{15}$$

where $\alpha_{\mathbf{A}}$ *and* $\alpha_{\mathbf{B}}$ *are any positive constants that satisfy* $0 < \alpha_{\mathbf{A}} < 2/\lambda_{\max}(\mathbf{A})$ *and* $0 < \alpha_{\mathbf{B}} < 2/\lambda_{\max}(\mathbf{B})$*, with* $\lambda_{\max}(\cdot)$ *denoting the largest eigenvalue of the argument matrix,* $(a)_k = a(a+1)\ldots(a+k-1)$*, and* $C_{[i+j+k]}^{[i],[j],[k]}(\cdot,\cdot,\cdot)$ *are the* (i,j,k)*-th top-order invariant polynomials (whose explicit form is presented in Appendix A).*

Remark. Conditions for the existence of the moment are stated in Bao & Kan (2013, proposition 1 therein): when $\mathbf{B}$ is positive definite, the moment exists if and only if $\frac{p}{2}+m>n$; and when $\mathbf{B}$ is positive semidefinite, slightly different conditions apply, depending on the rank of $\mathbf{B}$ and structures of $\mathbf{A}$ and $\mathbf{B}$. When m is an integer, the expressions can be simplified so that the summation for i disappears. Also, when $\boldsymbol{\mu}=\mathbf{0}_p$, they substantially simplify as all terms with $k>0$ are zero. If $\mathbf{A}$ or $\mathbf{B}$ is asymmetric, it can be symmetrized by using $\mathbf{x}^T\mathbf{A}\mathbf{x}=\mathbf{x}^T\mathbf{A}^T\mathbf{x}=\mathbf{x}^T(\mathbf{A}+\mathbf{A}^T)\mathbf{x}/2$.

In practice, the above series can be approximated by its partial sum, with recursive algorithms to calculate $d_{i,j,k}(\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3)=\frac{\left(\frac{1}{2}\right)_{i+j+k}}{i!j!k!}C_{[i+j+k]}^{[i],[j],[k]}(\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3)$ (Chikuse, 1987; Hillier *et al.*, 2009, 2014; Bao & Kan, 2013). General covariance structures for $\mathbf{x}$ can be accommodated under certain conditions by transforming the variables and quadratic forms, as described in Appendix C. Appendix C also provides a similar result for multiple ratios of the form $\frac{(\mathbf{x}^T\mathbf{A}\mathbf{x})^l}{(\mathbf{x}^T\mathbf{B}\mathbf{x})^m(\mathbf{x}^T\mathbf{D}\mathbf{x})^n}$.

2.4 Average measures for uniformly distributed selection

With the above theories, average measures can be readily derived as expectations with respect to random $\boldsymbol{\beta}$ under the assumption of normality. Previous treatments of evolvability measures (Hansen & Houle, 2008; Kirkpatrick, 2009) and random skewers analysis (Revell, 2007) typically assumed that $\mathbf{u}$ is uniformly distributed on the unit hypersphere in the p-dimensional space. This condition is entailed in spherical multivariate normal distributions of $\boldsymbol{\beta}$: $\boldsymbol{\beta}\sim N_p\left(\mathbf{0}_p,\sigma^2\mathbf{I}_p\right)$ for aribtrary positive σ^2. This section provides expressions of average measures under this condition. (Note that the normality of $\boldsymbol{\beta}$ is not necessary for $\mathbf{u}$ to be uniformly distributed; i.e., all results in this section hold as long as the distribution of $\boldsymbol{\beta}$ is spherically symmetric.) Expressions for a general normal case $\boldsymbol{\beta}\sim N_p\left(\boldsymbol{\eta},\boldsymbol{\Sigma}\right)$ are given in Appendix C. Most of the results below can be derived either by applying Proposition 1 or Proposition 4 in Appendix B, or directly as done in Appendix A.

The average evolvability $\bar{e}$ is straightforward to obtain:

$$\begin{aligned}\bar{e} &= \mathrm{E}_{\mathbf{u}}\left(\mathbf{u}^T\mathbf{G}\mathbf{u}\right)\\ &= \mathrm{E}_{\mathbf{u}}\left(\mathrm{tr}\left(\mathbf{G}\mathbf{u}\mathbf{u}^T\right)\right)\\ &= \mathrm{tr}\left(\mathbf{G}\right)/p,\end{aligned} \tag{16}$$

as derived by Hansen & Houle (2008). This is because $\mathrm{E}_{\mathbf{u}}\left(\mathbf{u}\mathbf{u}^T\right)=\mathbf{I}_p/p$ under this condition.

Conditional evolvability $c(\boldsymbol{\beta})$ and autonomy $a(\boldsymbol{\beta})$ can be nonzero only when $\boldsymbol{\beta} \in R(\mathbf{G})$ (see above). When $\mathbf{G}$ is singular, this happens with probability 0 when $\boldsymbol{\beta}$ is continuously distributed across the entire space. Hence, the average conditional evolvability $\bar{c}$ and the average autonomy $\bar{a}$ are 0 when $\mathbf{G}$ is singular. Although the following expressions can operationally be applied to singular $\mathbf{G}$ by using $\mathbf{G}^{-}$ in place of $\mathbf{G}^{-1}$, such a result is invalid because the ratio expressions do not hold when $\boldsymbol{\beta} \notin R(\mathbf{G})$.

When $\mathbf{G}$ is nonsingular, the average conditional evolvability $\bar{c}$ is:

$$\begin{aligned}
\bar{c} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\frac{\boldsymbol{\beta}^T\boldsymbol{\beta}}{\boldsymbol{\beta}^T\mathbf{G}^{-1}\boldsymbol{\beta}}\right) \\
=& \alpha \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\right)_k}{\left(\frac{p}{2}\right)_k} C_{[k]}\left(\mathbf{I}_p - \alpha\mathbf{G}^{-1}\right) \\
=& \alpha \sum_{k=0}^{\infty} \frac{(1)_k}{\left(\frac{p}{2}\right)_k} d_k\left(\mathbf{I}_p - \alpha\mathbf{G}^{-1}\right) \\
=& \alpha\, {}_2F_1\left(\frac{1}{2}, 1; \frac{p}{2}; \mathbf{I}_p - \alpha\mathbf{G}^{-1}\right),
\end{aligned} \tag{17}$$

where α is a positive constant that satisfies $0 < \alpha < 2/\lambda_{\max}\left(\mathbf{G}^{-1}\right)$, $C_{[k]}(\cdot)$ are the kth top-order zonal polynomials, and ${}_2F_1\left(a, b; c; \cdot\right)$ is an alternative expression using the hypergeometric function of matrix argument with the three parameters a, b, and c (see Appendix A). When $p = 2$, this expression simplifies to the geometric mean of the two eigenvalues of $\mathbf{G}$ as mentioned by Hansen & Houle (2008) (Appendix D).

The average autonomy $\bar{a}$ is, when $\mathbf{G}$ is nonsingular,

$$\begin{aligned}
\bar{a} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\frac{(\boldsymbol{\beta}^T\boldsymbol{\beta})^2}{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{G}^{-1}\boldsymbol{\beta}}\right) \\
=& \alpha_1\alpha_2 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\left(\frac{1}{2}\right)_{i+j}}{\left(\frac{p}{2}\right)_{i+j}} C_{[i+j]}^{[i],[j]}\left(\mathbf{I}_p - \alpha_1\mathbf{G}, \mathbf{I}_p - \alpha_2\mathbf{G}^{-1}\right) \\
=& \alpha_1\alpha_2 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(1)_i\,(1)_j}{\left(\frac{p}{2}\right)_{i+j}} d_{i,j}\left(\mathbf{I}_p - \alpha_1\mathbf{G}, \mathbf{I}_p - \alpha_2\mathbf{G}^{-1}\right),
\end{aligned} \tag{18}$$

where α_1 and α_2 are to satisfy $0 < \alpha_1 < 2/\lambda_{\max}(\mathbf{G})$ and $0 < \alpha_2 < 2/\lambda_{\max}\left(\mathbf{G}^{-1}\right)$. The average integration $\bar{i}$ is simply $1 - \bar{a}$ (and equals 1 when $\mathbf{G}$ is singular).

The average respondability $\bar{r}$ is

$$
\begin{aligned}
\bar{r} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\sqrt{\frac{\boldsymbol{\beta}^T\mathbf{G}^2\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}}}\right) \\
=& \alpha^{-\frac{1}{2}}\sum_{k=0}^{\infty}\frac{\left(-\frac{1}{2}\right)_k\left(\frac{1}{2}\right)_k}{\left(\frac{p}{2}\right)_k k!}C_{[k]}\left(\mathbf{I}_p-\alpha\mathbf{G}^2\right) \\
=& \alpha^{-\frac{1}{2}}\sum_{k=0}^{\infty}\frac{\left(-\frac{1}{2}\right)_k}{\left(\frac{p}{2}\right)_k}d_k\left(\mathbf{I}_p-\alpha\mathbf{G}^2\right) \\
=& \alpha^{-\frac{1}{2}}\,{}_2F_1\left(-\frac{1}{2},\frac{1}{2};\frac{p}{2};\mathbf{I}_p-\alpha\mathbf{G}^2\right),
\end{aligned}
\tag{19}
$$

where α is to satisfy $0<\alpha<2/\lambda_{\max}\left(\mathbf{G}^2\right)$ (note that $\lambda_{\max}\left(\mathbf{G}^2\right)=\lambda_{\max}(\mathbf{G})^2$ since $\mathbf{G}$ is nonnegative definite). Kirkpatrick (2009) discussed on an equivalent quantity, the average (relative) selection response ($\bar{R}$ there), but did not provide a closed-form expression for arbitrary $\mathbf{G}$. It is possible to show his expression for the special case with complete integration is entailed in the present result (Appendix D).

The average flexibility $\bar{f}$ is

$$
\begin{aligned}
\bar{f} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\frac{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^T\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{G}^2\boldsymbol{\beta}}}\right) \\
=& \alpha^{\frac{1}{2}}\sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}\right)_k\left(\frac{1}{2}\right)_{k+1}}{\left(\frac{p}{2}\right)_{k+1}k!}C_{[k+1]}^{[k],[1]}\left(\mathbf{I}_p-\alpha\mathbf{G}^2,\mathbf{G}\right) \\
=& \alpha^{\frac{1}{2}}\sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}\right)_k}{\left(\frac{p}{2}\right)_{k+1}}d_{k,1}\left(\mathbf{I}_p-\alpha\mathbf{G}^2,\mathbf{G}\right),
\end{aligned}
\tag{20}
$$

where α is to satisfy $0<\alpha<2/\lambda_{\max}\left(\mathbf{G}^2\right)$. In general, $\bar{f}$ is well defined even if $\mathbf{G}$ is singular, unless $\boldsymbol{\beta}\in N\left(\mathbf{G}\right)$ with nonzero probability (which of course does not happen here).

The average response difference $\bar{d}$ has a form essentially identical to $\bar{r}$ except for the argument matrix

$$
\begin{aligned}
\bar{d} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\sqrt{\frac{\boldsymbol{\beta}^T\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}}}\right) \\
=& \alpha^{-\frac{1}{2}}\sum_{k=0}^{\infty}\frac{\left(-\frac{1}{2}\right)_k\left(\frac{1}{2}\right)_k}{\left(\frac{p}{2}\right)_k k!}C_{[k]}\left(\mathbf{I}_p-\alpha\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\right) \\
=& \alpha^{-\frac{1}{2}}\sum_{k=0}^{\infty}\frac{\left(-\frac{1}{2}\right)_k}{\left(\frac{p}{2}\right)_k}d_k\left(\mathbf{I}_p-\alpha\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\right) \\
=& \alpha^{-\frac{1}{2}}\,{}_2F_1\left(-\frac{1}{2},\frac{1}{2};\frac{p}{2};\mathbf{I}_p-\alpha\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\right),
\end{aligned}
\tag{21}
$$

where α is to satisfy $0<\alpha<2/\lambda_{\max}\left(\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\right)$.

The average response correlation $\bar{\rho}$ is

$$\begin{aligned}
\bar{\rho} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\frac{\boldsymbol{\beta}^T\mathbf{G}_1\mathbf{G}_2\boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^T\mathbf{G}_1^2\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{G}_2^2\boldsymbol{\beta}}}\right)\\
=&(\alpha_1\alpha_2)^{\frac{1}{2}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_{i+j+1}}{\left(\frac{p}{2}\right)_{i+j+1}i!j!}C_{[i+j+1]}^{[i],[j],[1]}\left(\mathbf{I}_p-\alpha_1\mathbf{G}_1^2,\mathbf{I}_p-\alpha_2\mathbf{G}_2^2,\frac{\mathbf{G}_1\mathbf{G}_2+\mathbf{G}_2\mathbf{G}_1}{2}\right)\\
=&(\alpha_1\alpha_2)^{\frac{1}{2}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_j}{\left(\frac{p}{2}\right)_{i+j+1}}d_{i,j,1}\left(\mathbf{I}_p-\alpha_1\mathbf{G}_1^2,\mathbf{I}_p-\alpha_2\mathbf{G}_2^2,\frac{\mathbf{G}_1\mathbf{G}_2+\mathbf{G}_2\mathbf{G}_1}{2}\right),
\end{aligned}\tag{22}$$

where α_1 and α_2 are to satisfy $0<\alpha_1<2/\lambda_{\max}\left(\mathbf{G}_1^2\right)$ and $0<\alpha_2<2/\lambda_{\max}\left(\mathbf{G}_2^2\right)$.

2.5 Truncation error

The above series are known to absolutely converge (Smith, 1989; Bao & Kan, 2013), thus a partial sum up to certain higher-order terms can in principle be used as an approximation of the exact value. Although each term can be easily evaluated with a recursive algorithm (Bao & Kan, 2013; Hillier *et al.*, 2014), accurate numerical evaluation involving higher-order terms is practically a computer-intensive problem. The speed of numerical convergence and computational time will be examined in the next section.

The choice of α is arbitrary within the constraint $0<\alpha<2/\lambda_{\max}$ with respect to the relevant matrix (see above), but influences the signs of the terms as well as the speed of convergence. When $0<\alpha\leq 1/\lambda_{\max}$, all the zonal and invariant polynomials in the above expressions are positive because the argument matrices have only nonnegative eigenvalues. Therefore, all terms in the summations in $\bar{c}$ (17), $\bar{a}$ (18), $\bar{f}$ (20), and $\bar{\rho}$ (22) are positive, so truncation results in (slight) underestimation. On the other hand, all terms in $\bar{r}$ (19) and $\bar{d}$ (21), except those for $k=0$, are negative in the same condition, so truncation results in overestimation. If $1/\lambda_{\max}<\alpha<2/\lambda_{\max}$, the signs of the polynomials possibly, though not always, fluctuate, rendering the behavior of the series less predictable.

It is possible to derive bounds for the truncation errors in $\bar{c}$, $\bar{r}$, $\bar{f}$, and $\bar{d}$ by applying the method of Hillier *et al.* (2009, theorem 6) with slight modifications, if $\mathbf{G}$ is nonsingular and the constants α are taken within $0<\alpha\leq 1/\lambda_{\max}$. Let $\bar{c}_M$, $\bar{r}_M$, $\bar{f}_M$, and $\bar{d}_M$ be the relevant partial sums up to $k=M$. The truncation error bounds are:

$$0\leq\bar{c}-\bar{c}_M\leq\frac{(1)_{M+1}}{\left(\frac{p}{2}\right)_{M+1}}\left(\alpha^{-\frac{p-2}{2}}\left|\mathbf{G}\right|^{\frac{1}{2}}-\alpha\sum_{k=0}^{M}d_k\left(\mathbf{I}_p-\alpha\mathbf{G}^{-1}\right)\right),\tag{23}$$

$$0\geq\bar{r}-\bar{r}_M\geq\frac{\left(-\frac{1}{2}\right)_{M+1}}{\left(\frac{p}{2}\right)_{M+1}}\left(\alpha^{-\frac{p+1}{2}}\left|\mathbf{G}\right|^{-1}-\alpha^{-\frac{1}{2}}\sum_{k=0}^{M}d_k\left(\mathbf{I}_p-\alpha\mathbf{G}^2\right)\right),\tag{24}$$

$$0\leq\bar{f}-\bar{f}_M\leq\frac{\left(\frac{1}{2}\right)_{M+1}}{\left(\frac{p}{2}\right)_{M+2}}\left(\alpha^{-\frac{p+1}{2}}\left|\mathbf{G}\right|^{-1}\frac{\operatorname{tr}\left(\mathbf{G}^{-1}\right)}{2}-\alpha^{\frac{1}{2}}\sum_{k=0}^{M}d_{k,1}\left(\mathbf{I}_p-\alpha\mathbf{G}^2,\mathbf{G}\right)\right),\tag{25}$$

$$0\geq\bar{d}-\bar{d}_M\geq\frac{\left(-\frac{1}{2}\right)_{M+1}}{\left(\frac{p}{2}\right)_{M+1}}\left(\alpha^{-\frac{p+1}{2}}\left|\mathbf{G}_1-\mathbf{G}_2\right|^{-1}-\alpha^{-\frac{1}{2}}\sum_{k=0}^{M}d_k\left(\mathbf{I}_p-\alpha\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\right)\right),\tag{26}$$

where the constants α are as in the corresponding definitions (but to satisfy $0 < \alpha \leq 1/\lambda_{\max}$). Unfortunately, the same method are not applicable to $\bar{a}$ or $\bar{\rho}$. Also, if $\mathbf{G}$ (or $\mathbf{G}_1 - \mathbf{G}_2$ in case of $\bar{d}$) is singular, these error bounds do not hold.

Empirically, a larger value of α often improves the speed of convergence, as more eigenvalues of the argument matrix approach 0, but this should be used with caution since the above error bounds are not strictly applicable if $\alpha > 1/\lambda_{\max}$. To be precise, the error bounds should hold even under this condition provided that all the zonal/invariant polynomials above $k = M$ are nonnegative, but it is not obvious in which case that condition can be shown to be true.

3 Numerical evaluation

3.1 Software implementation

There exist two `Matlab` programs distributed by Raymond M. Kan (https://www-2.rotman.utoronto.ca/~kan/) for evaluating the moments of simple ratios of quadratic forms in normal variables (Hillier *et al.*, 2009, 2014; Bao & Kan, 2013), which can be used to evaluate $\bar{c}$, $\bar{r}$, and $\bar{d}$ if a suitable environment is available. Alternatively, these average measures could be evaluated as hypergeometric functions of matrix argument, for which an efficient, general algorithm has been developed by Koev & Edelman (2006) and implemented in the `R` package `HypergeoMat` (Laurent, 2022). However, that seems relatively inefficient in the present applications where the terms corresponding to all non-top-order partitions can be omitted. None of these can be used to evaluate moments of multiple ratios like $\bar{a}$, $\bar{f}$, and $\bar{\rho}$.

In order to make all above expressions available in the popular statistical software environment `R`, the present author has developed an independent implementation of the algorithms of Hillier *et al.* (2014) and Bao & Kan (2013) and published it as the package `qfratio` on the GitHub repository (https://github.com/watanabe-j/qfratio). This package can be used to evaluate general moments of ratios of quadratic forms in normal variables with arbitrary mean and covariance (under certain constraints for singular cases as discussed in Appendix C.1). It uses `RcppEigen` (Bates & Eddelbuettel, 2013) to effectively execute the computationally intensive recursive algorithms.

3.2 Numerical experiment: methods

Using the new `R` package, numerical experiments were conducted in a small set of artificial covariance matrices to evaluate behaviors and efficiency of the present series expressions. For each of $p = 5, 20, 50$, and 200, three different eigenvalue conformations were used; two were set to have single large eigenvalues with different magnitudes of integration, $V_{\mathrm{rel}} = 0.1, 0.5$—corresponding to relatively weak and rather strong integration, respectively—and one was set to have a quadratically decreasing series of eigenvalues (for the algorithms see Watanabe (2022) and the `R` package `eigvaldisp` (https://github.com/watanabe-j/eigvaldisp)). The matrices with these eigenvalue conformations are referred to as $\mathbf{G}_{0.1}$, $\mathbf{G}_{0.5}$, and $\mathbf{G}_{\mathrm{Q}}$. For each of these, two eigenvector conformations were used so that the resultant covariance matrices were either diagonal matrices of the eigenvalues or correlation matrices with the same eigenvalues constructed with Givens rotations (Davies & Higham, 2000). For the one-matrix measures ($\bar{c}$, $\bar{r}$, $\bar{a}$, and $\bar{f}$),

the choice of eigenvectors is inconsequential. For the two-matrix measures ($\bar{d}$ and $\bar{\rho}$), comparisons were made between pairs of matrices with identical and different eigenvectors. The results were also compared to those from the traditional Monte Carlo evaluation using the `qfratio` package and the delta method approximation of Hansen & Houle (2008, 2009) using new codes.

When possible, an average computational time from 10 runs is recorded for the present series evaluation and Monte Carlo evaluation with 10,000 iterations in each condition. This is except for $\bar{\rho}$ in most conditions with $p \geq 50$ where only a single run was timed for the series evaluation as it took a substantially long time to numerically converge (below). Where a truncation error bound is available, an order of series evaluation M was chosen so that the error is below 1.0×10^{-8}. Where an error bound is unavailable, it was aimed to evaluate the series until the result rounded to the digit of 10^{-9} appeared stable. (These values were arbitrarily chosen for benchmarking purposes, and not to be used as a guide for applications.) No formal timing was done for the delta method approximation as it is not computer-intensive. All computational times reported are in elapsed time rather than CPU time. The calculations were executed on a regular desktop PC with Intel® Core i7-8700 CPU and 16 GB RAM, running `R` version 4.2.1 (R Core Team, 2022) with its built-in BLAS. The codes used in the numerical experiments are provided as a Supplementary Material.

3.3 Numerical experiment: results

The present series evaluations as implemented in the `R` package `qfratio` reached numerical convergence within reasonable computational times in all cases except $\bar{\rho}$ with large ($p = 200$) or highly integrated ($\mathbf{G}_{0.5}$) matrices (detailed below). The order of evaluation M required to accomplish a given level of accuracy varied across evolvability measures and eigenvalue conformations, but in most cases increased with p (Figs. 2–5; Tables 1–6). The rate of scaling with respect to p tended to be more rapid in the series evaluation than in Monte Carlo evaluation, when the required accuracy for the former and the number of iteration for the latter are pre-determined as described above.

For $\bar{c}$, the computational time required for the error bound to numerically converge below 1.0×10^{-8} was at most in the order of ∼0.1 s and compared favorably with the Monte Carlo evaluation with 10,000 iterations across the entire range of p examined (Fig. 2; Table 1). For $\bar{r}$ and $\bar{f}$, the computational time varied from ∼1.5 ms in $p = 5$ to several seconds in $p = 200$, and in some eigenvalue conformations the series evaluation was outcompeted by the Monte Carlo evaluation with a fixed number of iterations in terms of plain computational time (Tables 2, 3). However, it is to be remembered that results from Monte Carlo evaluations always involve stochastic error, whose standard error (SE) scales only with the inverse square root of the number of iteration (i.e., 100 as many calculations are required to reduce the SE by the factor of 1/10). It is also notable that the partial sums themselves typically appeared to reach asymptotes much earlier than the error bounds did in these measures (Figs. 2, 3). It took substantially longer time to accurately evaluate $\bar{a}$ than the other one-matrix measures, since it involves a double series so that the computational burden scales roughly with M^2. Consequently, the computational time required for $\bar{a}$ to numerically converge at the order of 10^{-9} varied from ∼5 ms to over 300 s (Fig. 3; Table 4). Technically, this large computational time was because of the necessity to use

Table 1. Results of numerical experiments for $\bar{c}$. Tabulated are combinations of eigenvalue conformation and number of variables p, order of evaluation M required to attain the accuracy specified in text, partial sum from series evaluation $\bar{c}_M$ and its error bound, average computational time of series and Monte Carlo evaluations, t_{S} and t_{MC}, empirical standard error from one Monte Carlo run with 10,000 iterations $\mathrm{SE}_{\mathrm{MC}}$, and Hansen & Houle's delta method approximation $\hat{c}_{\mathrm{HH}}$.

Matrix	p	M	$\bar{c}_M$	Error bound	t_{S} (ms)	t_{MC} (ms)	$\mathrm{SE}_{\mathrm{MC}}$	$\hat{c}_{\mathrm{HH}}$
$\mathbf{G}_{0.1}$	5	33	0.8289	6.637×10^{-9}	0.3	2.5	2.008×10^{-3}	0.8187
	20	19	0.7194	9.622×10^{-9}	0.4	11.1	5.212×10^{-4}	0.7188
	50	9	0.6977	9.309×10^{-9}	0.6	28.5	1.984×10^{-4}	0.6976
	200	5	0.6872	2.978×10^{-9}	12.9	137.5	5.031×10^{-5}	0.6872
$\mathbf{G}_{0.5}$	5	130	0.4016	4.687×10^{-9}	0.3	2.4	2.004×10^{-3}	0.3803
	20	25	0.3097	8.858×10^{-9}	0.4	11.2	2.582×10^{-4}	0.3094
	50	10	0.2991	5.217×10^{-9}	0.6	28.7	9.094×10^{-5}	0.2991
	200	5	0.2944	3.512×10^{-9}	12.9	138.9	2.061×10^{-5}	0.2944
$\mathbf{G}_{\mathrm{Q}}$	5	290	0.4730	7.239×10^{-9}	0.4	2.5	3.179×10^{-3}	0.4456
	20	1800	0.1501	6.932×10^{-9}	0.9	11.1	1.153×10^{-3}	0.1468
	50	7800	0.0632	8.195×10^{-9}	3.0	28.3	4.841×10^{-4}	0.0627
	200	260 000	0.0162	4.445×10^{-10}	113.9	134.7	1.279×10^{-4}	0.0162

the `long double` floating point number (rather than the regular double precision floating point number) in the evaluation of $\bar{a}$ in $\mathbf{G}_{\mathrm{Q}}$ for $p \geq 50$ to avoid numerical overflow and underflow.

For two-matrix comparisons, $\bar{d}$ behaved similarly to $\bar{r}$ with which it shares the same mathematical form (Fig. 4; Table 5). In some comparisons between matrices with the same eigenvalues and different eigenvectors, the error bound of $\bar{d}$ could not be calculated because the argument matrix $((\mathbf{G}_1 - \mathbf{G}_2)^2)$ was singular due to the artificial conformations of eigenvalues in $\mathbf{G}_{0.1}$ and $\mathbf{G}_{0.5}$ where the trailing eigenvalues are identical in magnitude. Series evaluation of $\bar{\rho}$ took substantially longer computational time and memory than the other measures. For $p \geq 20$, the series could not be evaluated until numerical convergence in any comparisons that involve $\mathbf{G}_{0.5}$ due to excessively large memory requirement for the 16 GB RAM machine used here, at least when the accuracy in the order of 10^{-9} is aimed at (Fig. 5; Table 6). This is apparently

Table 2. Results of numerical experiments for $\bar{r}$. See Table 1 for legend.

Matrix	p	M	$\bar{r}_M$	Error bound	t_{S} (ms)	t_{MC} (ms)	$\mathrm{SE}_{\mathrm{MC}}$	$\hat{r}_{\mathrm{HH}}$
$\mathbf{G}_{0.1}$	5	120	1.1176	-6.518×10^{-9}	0.4	2.9	3.895×10^{-3}	1.1082
	20	1300	1.5095	-3.571×10^{-9}	0.7	11.6	7.911×10^{-3}	1.4443
	50	8500	2.0677	-9.517×10^{-9}	3.2	29.2	1.264×10^{-2}	1.9439
	200	360 000	3.7407	-9.967×10^{-10}	151.6	139.8	2.678×10^{-2}	3.4955
$\mathbf{G}_{0.5}$	5	1600	1.4865	-8.769×10^{-9}	0.7	3.1	8.892×10^{-3}	1.4986
	20	24 000	2.6450	-8.893×10^{-9}	7.5	12.2	1.865×10^{-2}	2.5521
	50	200 000	4.0700	-6.271×10^{-9}	63.8	29.2	2.998×10^{-2}	3.8681
	200	9 200 000	8.0216	-9.269×10^{-9}	3692.7	139.5	6.016×10^{-2}	7.5602
$\mathbf{G}_{\mathrm{Q}}$	5	1400	1.2052	-7.504×10^{-9}	0.6	2.9	4.056×10^{-3}	1.2078
	20	540	1.2992	-9.765×10^{-9}	0.4	11.7	2.574×10^{-3}	1.2990
	50	700	1.3237	-7.742×10^{-9}	0.6	29.4	1.703×10^{-3}	1.3237
	200	2100	1.3370	-3.993×10^{-9}	12.2	137.4	8.837×10^{-4}	1.3370

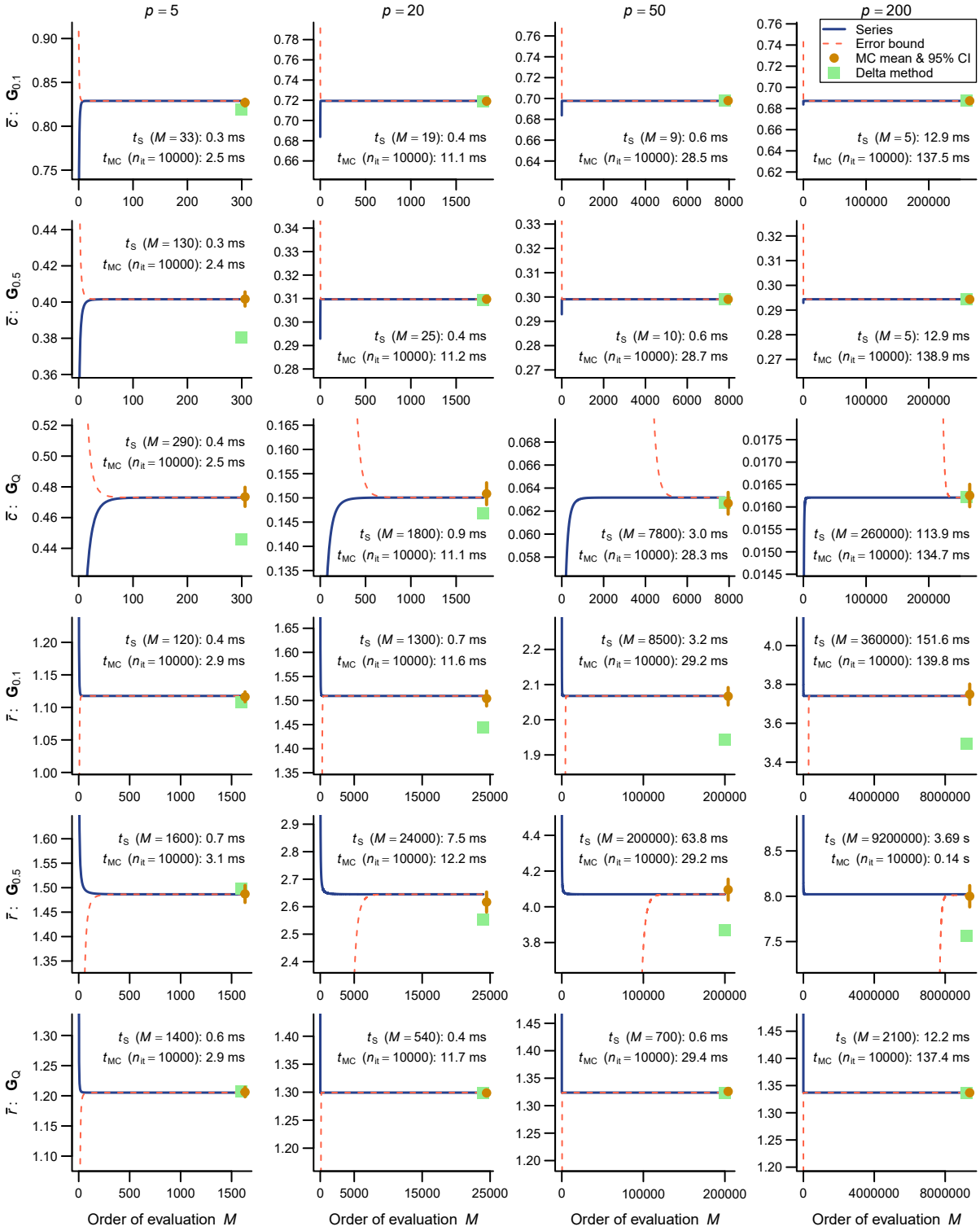

Figure 2. Results of numerical experiments for $\bar{c}$ (top three rows) and $\bar{r}$ (bottom three rows). Partial sums from series expressions are shown in blue solid lines, and associated truncation error bounds are in red broken lines, across varying orders of evaluation M. Mean and its 95% confidence interval from a Monte Carlo run with 10,000 iterations are shown in brown circles and vertical bars. Delta method approximations by Hansen & Houle (2008, 2009) are shown in green squares. Rows correspond to eigenvalue conformations (see text) and columns to number of variables p. Average computational times from 10 runs for the series t_{S} and Monte Carlo evaluations with 10,000 iteration t_{MC} are shown in each panel, along with M required for numerical convergence in the former.

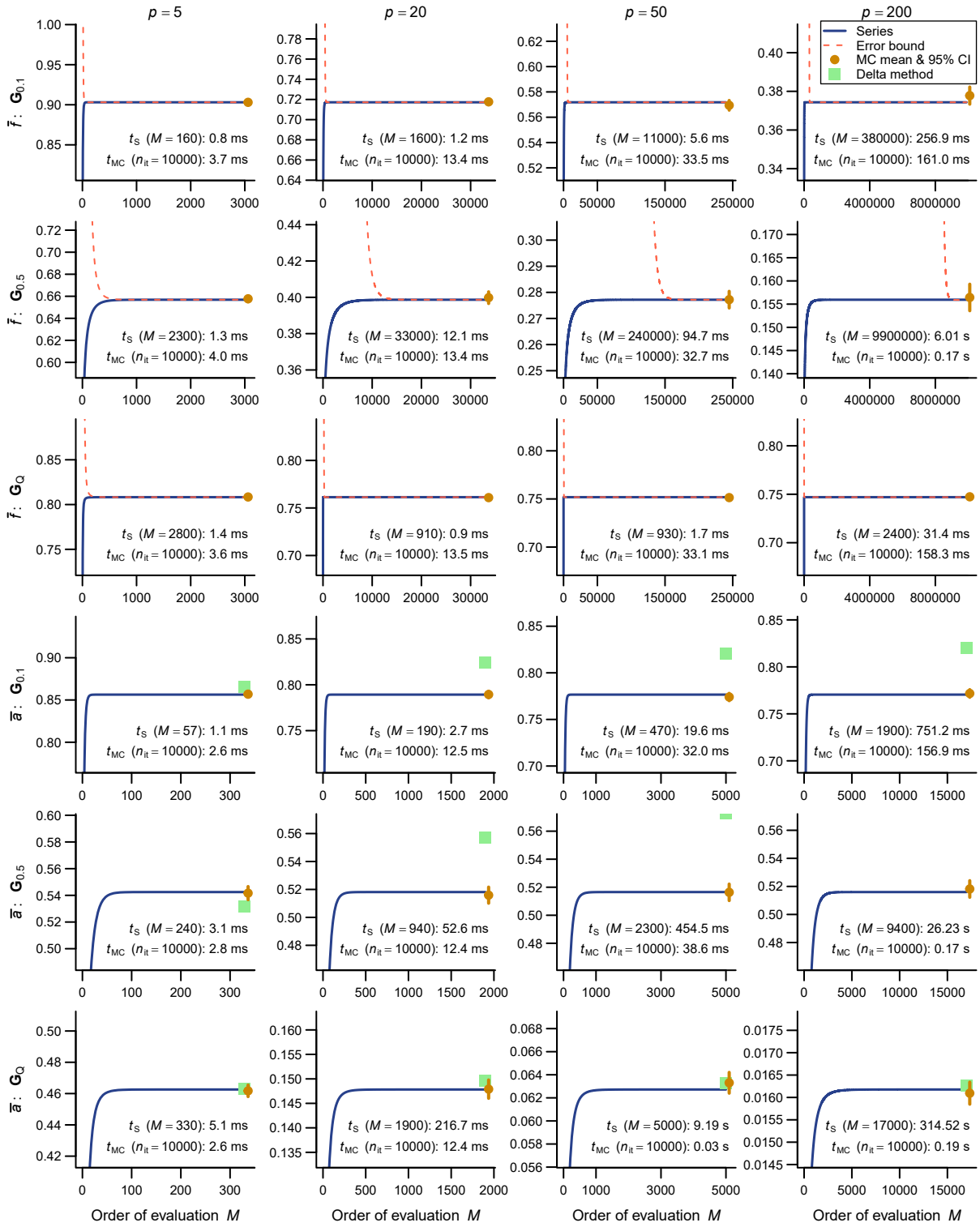

Figure 3. Results of numerical experiments for $\bar{f}$ (top three rows) and $\bar{a}$ (bottom three rows). See Figure 2 for legend. Delta method approximations are not available for $\bar{f}$, and truncation error bounds are not available for $\bar{a}$. In some results for $\bar{a}$, delta method approximation is too inaccurate to be plotted in the panels, which span $\pm 10\%$ of the terminal values of series evaluation.

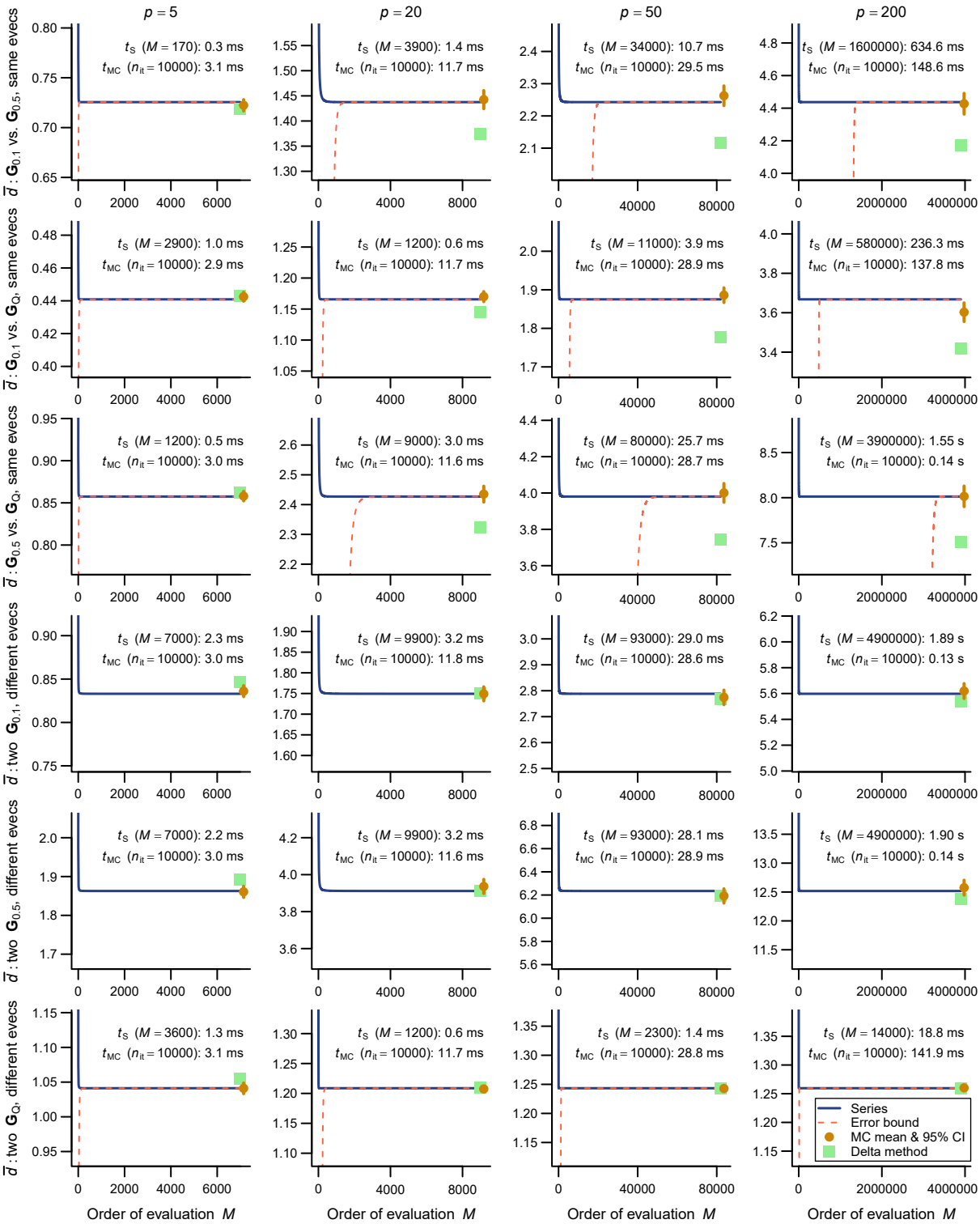

Figure 4. Results of numerical experiments for $\bar{d}$. Top three rows are comparisons between covariance matrices with different eigenvalues and same eigenvectors, whereas the bottom three rows are between matrices with same eigenvalues and different eigenvectors. Comparisons between different eigenvalues and different eigenvectors are not shown for space limitations. See Figure 2 for legend. Error bounds are not available in some same-eigenvalue comparisons because of the singularity of argument matrices (see text).

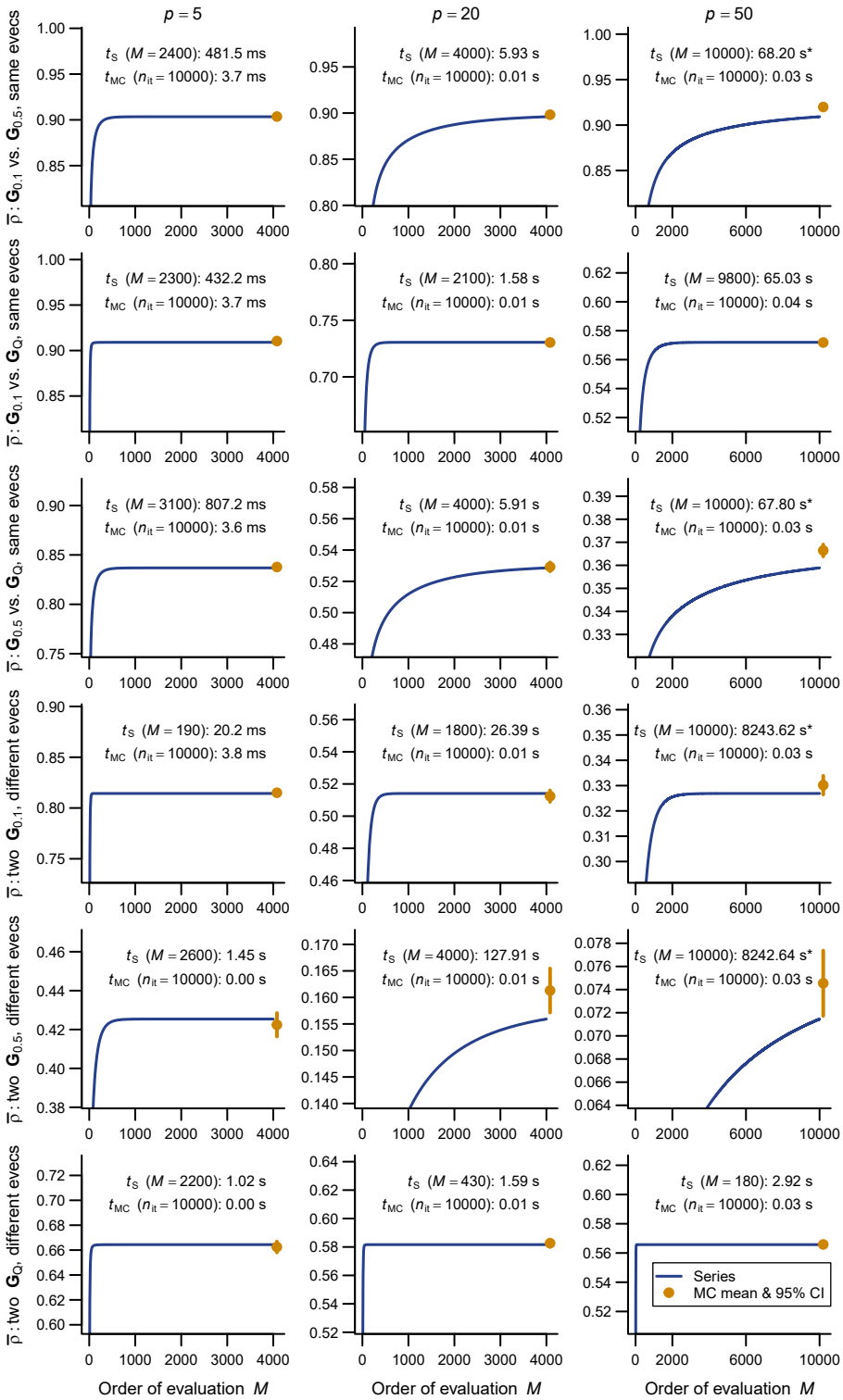

Figure 5. Results of numerical experiments for $\bar{\rho}$. Top three rows are comparisons between covariance matrices with different eigenvalues and same eigenvectors, whereas the bottom three rows are between matrices with same eigenvalues and different eigenvectors. Comparisons between different eigenvalues and different eigenvectors are not shown for space limitations. See Figure 2 for legend. Delta methods approximations or error bounds are not available for $\bar{\rho}$. Asterisks after computational times denote ones from single runs rather than from 10 runs, for non-convergence or excessive computational time required.

Table 3. Results of numerical experiments for $\bar{f}$. See Table 1 for legend.

Matrix	p	M	$\bar{f}_M$	Error bound	t_{S} (ms)	t_{MC} (ms)	$\mathrm{SE}_{\mathrm{MC}}$
$\mathbf{G}_{0.1}$	5	160	0.9031	3.628×10^{-9}	0.8	3.7	5.229×10^{-4}
	20	1600	0.7172	5.471×10^{-9}	1.2	13.4	1.467×10^{-3}
	50	11 000	0.5718	8.857×10^{-10}	5.6	33.5	1.950×10^{-3}
	200	380 000	0.3743	1.267×10^{-9}	256.9	161.0	2.259×10^{-3}
$\mathbf{G}_{0.5}$	5	2300	0.6569	6.730×10^{-9}	1.3	4.0	1.375×10^{-3}
	20	33 000	0.3987	8.228×10^{-9}	12.1	13.4	1.642×10^{-3}
	50	240 000	0.2772	9.889×10^{-9}	94.7	32.7	1.627×10^{-3}
	200	9 900 000	0.1559	3.968×10^{-9}	6005.8	170.5	1.459×10^{-3}
$\mathbf{G}_{\mathrm{Q}}$	5	2800	0.8082	9.271×10^{-9}	1.4	3.6	1.117×10^{-3}
	20	910	0.7614	9.615×10^{-9}	0.9	13.5	7.015×10^{-4}
	50	930	0.7518	8.400×10^{-9}	1.7	33.1	4.624×10^{-4}
	200	2400	0.7470	1.679×10^{-10}	31.4	158.3	2.373×10^{-4}

Table 4. Results of numerical experiments for $\bar{a}$. See Table 1 for legend.

Matrix	p	M	$\bar{a}_M$	t_{S} (ms)	t_{MC} (ms)	$\mathrm{SE}_{\mathrm{MC}}$	$\hat{a}_{\mathrm{HH}}$
$\mathbf{G}_{0.1}$	5	57	0.8563	1.1	2.6	1.012×10^{-3}	0.8656
	20	190	0.7893	2.7	12.5	1.837×10^{-3}	0.8245
	50	470	0.7766	19.6	32.0	2.017×10^{-3}	0.8210
	200	1900	0.7705	751.2	156.9	2.080×10^{-3}	0.8205
$\mathbf{G}_{0.5}$	5	240	0.5425	3.1	2.8	2.518×10^{-3}	0.5319
	20	940	0.5181	52.6	12.4	2.943×10^{-3}	0.5572
	50	2300	0.5165	454.5	38.6	3.002×10^{-3}	0.5728
	200	9400	0.5160	26 233.9	167.6	3.023×10^{-3}	0.5823
$\mathbf{G}_{\mathrm{Q}}$	5	330	0.4626	5.1	2.6	1.878×10^{-3}	0.4631
	20	1900	0.1478	216.7	12.4	9.551×10^{-4}	0.1496
	50	5000	0.0627	9186.8	31.3	4.595×10^{-4}	0.0633
	200	17 000	0.0162	314 523.8	192.5	1.256×10^{-4}	0.0163

Table 5. Results of numerical experiments for $\bar{d}$. Comparisons were made in pairs of covariance matrices with same and different eigenvectors (evecs). Error bound was unavailable for comparisons between two $\mathbf{G}_{0.1}$'s and two $\mathbf{G}_{0.5}$'s with different eigenvectors (see text). See Table 1 for further explanations.

Matrices	p	M	$\bar{d}_M$	Error bound	t_{S} (ms)	t_{MC} (ms)	$\mathrm{SE}_{\mathrm{MC}}$	$\hat{d}_{\mathrm{HH}}$
	5	170	0.7256	-5.603×10^{-9}	0.3	3.1	2.881×10^{-3}	0.7189
$\mathbf{G}_{0.1}$ vs. $\mathbf{G}_{0.5}$,	20	3900	1.4375	-8.787×10^{-9}	1.4	11.7	9.178×10^{-3}	1.3736
same evecs	50	34 000	2.2428	-6.090×10^{-9}	10.7	29.5	1.556×10^{-2}	2.1176
	200	1 600 000	4.4371	-1.852×10^{-9}	634.6	148.6	3.249×10^{-2}	4.1696
	5	2900	0.4409	-9.244×10^{-9}	1.0	2.9	1.375×10^{-3}	0.4429
$\mathbf{G}_{0.1}$ vs. $\mathbf{G}_{\mathrm{Q}}$,	20	1200	1.1657	-5.016×10^{-9}	0.6	11.7	4.122×10^{-3}	1.1456
same evecs	50	11 000	1.8752	-4.953×10^{-9}	3.9	28.9	9.711×10^{-3}	1.7781
	200	580 000	3.6679	-2.875×10^{-9}	236.3	137.8	2.403×10^{-2}	3.4190
	5	1200	0.8575	-6.347×10^{-9}	0.5	3.0	2.875×10^{-3}	0.8620
$\mathbf{G}_{0.5}$ vs. $\mathbf{G}_{\mathrm{Q}}$,	20	9000	2.4270	-9.410×10^{-9}	3.0	11.6	1.358×10^{-2}	2.3229
same evecs	50	80 000	3.9804	-9.196×10^{-9}	25.7	28.7	2.628×10^{-2}	3.7449
	200	3 900 000	8.0124	-2.997×10^{-9}	1547.4	135.7	5.798×10^{-2}	7.5050
	5	7000	0.8330	—	2.3	3.0	3.221×10^{-3}	0.8465
Two $\mathbf{G}_{0.1}$,	20	9900	1.7493	—	3.2	11.8	8.642×10^{-3}	1.7500
different evecs	50	93 000	2.7882	—	29.0	28.6	1.428×10^{-2}	2.7693
	200	4 900 000	5.5980	—	1894.2	135.0	2.926×10^{-2}	5.5357
	5	7000	1.8627	—	2.2	3.0	7.312×10^{-3}	1.8929
Two $\mathbf{G}_{0.5}$,	20	9900	3.9116	—	3.2	11.6	1.937×10^{-2}	3.9131
different evecs	50	93 000	6.2347	—	28.1	28.9	3.135×10^{-2}	6.1923
	200	4 900 000	12.5174	—	1899.1	135.0	6.510×10^{-2}	12.3783
	5	3600	1.0412	-9.597×10^{-9}	1.3	3.1	3.889×10^{-3}	1.0546
Two $\mathbf{G}_{\mathrm{Q}}$,	20	1200	1.2087	-4.817×10^{-9}	0.6	11.7	2.190×10^{-3}	1.2095
different evecs	50	2300	1.2433	-3.719×10^{-9}	1.4	28.8	1.353×10^{-3}	1.2434
	200	14 000	1.2595	-8.004×10^{-11}	18.8	141.9	6.765×10^{-4}	1.2595
	5	450	1.4432	-8.000×10^{-9}	0.3	2.8	5.620×10^{-3}	1.4478
$\mathbf{G}_{0.1}$ vs. $\mathbf{G}_{0.5}$,	20	9900	2.9662	-9.012×10^{-9}	3.2	11.6	1.617×10^{-2}	2.8938
different evecs	50	93 000	4.6991	-9.283×10^{-9}	29.8	28.8	2.711×10^{-2}	4.5298
	200	4 900 000	9.4026	-5.707×10^{-9}	1951.3	144.0	5.573×10^{-2}	8.9991
	5	690	0.9545	-9.978×10^{-9}	0.4	2.8	3.248×10^{-3}	0.9630
$\mathbf{G}_{0.1}$ vs. $\mathbf{G}_{\mathrm{Q}}$,	20	2200	1.4997	-9.404×10^{-9}	1.0	11.6	6.354×10^{-3}	1.4564
different evecs	50	13 000	2.0938	-5.509×10^{-9}	4.8	29.0	1.140×10^{-2}	1.9798
	200	590 000	3.7798	-9.522×10^{-9}	242.9	136.5	2.538×10^{-2}	3.5236
	5	6800	1.5135	-9.579×10^{-9}	2.1	2.8	5.724×10^{-3}	1.5199
$\mathbf{G}_{0.5}$ vs. $\mathbf{G}_{\mathrm{Q}}$,	20	8600	2.7842	-9.060×10^{-9}	2.8	11.6	1.590×10^{-2}	2.6625
different evecs	50	82 000	4.2077	-9.025×10^{-9}	25.7	28.8	2.738×10^{-2}	3.9590
	200	3 900 000	8.1259	-9.530×10^{-9}	1544.9	137.1	5.928×10^{-2}	7.6115

Table 6. Results of numerical experiments for $\bar{\rho}$. Parentheses around $\bar{\rho}_M$ denote values that did not reach numerical convergence with specified M; for these conditions, estimates from single Monte Carlo runs $\hat{\rho}_{\mathrm{MC}}$ are also shown for reference. Asterisks after t_{S} denote times from single runs rather than 10 runs. For $p = 200$, it was infeasible to apply the series evaluation in most combinations for excessive computational time and/or memory required. See Table 5 for further explanations.

Matrices	p	M	$\bar{\rho}_M$	t_{S} (s)	t_{MC} (s)	$\hat{\rho}_{\mathrm{MC}}$	$\mathrm{SE}_{\mathrm{MC}}$
$\mathbf{G}_{0.1}$ vs. $\mathbf{G}_{0.5}$, same evecs	5	2400	0.9034	0.48	0.00	—	6.305×10^{-4}
	20	4000	(0.8960)	5.93	0.01	0.8982	7.731×10^{-4}
	50	10 000	(0.9092)	68.20*	0.03	0.9199	7.805×10^{-4}
$\mathbf{G}_{0.1}$ vs. $\mathbf{G}_{\mathrm{Q}}$, same evecs	5	2300	0.9090	0.43	0.00	—	7.978×10^{-4}
	20	2100	0.7306	1.58	0.01	—	7.113×10^{-4}
	50	9800	0.5720	65.03	0.04	—	8.860×10^{-4}
$\mathbf{G}_{0.5}$ vs. $\mathbf{G}_{\mathrm{Q}}$, same evecs	5	3100	0.8369	0.81	0.00	—	1.138×10^{-3}
	20	4000	(0.5286)	5.91	0.01	0.5293	1.462×10^{-3}
	50	10 000	(0.3589)	67.80*	0.03	0.3664	1.325×10^{-3}
Two $\mathbf{G}_{0.1}$, different evecs	5	190	0.8143	0.02	0.00	—	1.071×10^{-3}
	20	1800	0.5141	26.39	0.01	—	1.826×10^{-3}
	50	9900	0.3269	8103.95*	0.03	—	1.883×10^{-3}
Two $\mathbf{G}_{0.5}$, different evecs	5	2600	0.4254	1.45	0.00	—	3.079×10^{-3}
	20	4000	(0.1559)	127.91	0.01	0.1613	2.120×10^{-3}
	50	10 000	(0.0715)	8242.64*	0.03	0.0745	1.439×10^{-3}
Two $\mathbf{G}_{\mathrm{Q}}$, different evecs	5	2200	0.6644	1.02	0.00	—	2.244×10^{-3}
	20	430	0.5816	0.29	0.01	—	1.307×10^{-3}
	50	180	0.5657	0.62	0.03	—	7.995×10^{-4}
	200	110	0.5581	8.56	0.18	—	4.004×10^{-4}
$\mathbf{G}_{0.1}$ vs. $\mathbf{G}_{0.5}$, different evecs	5	2300	0.5904	1.12	0.00	—	2.036×10^{-3}
	20	4000	(0.2835)	128.03	0.01	0.2844	1.960×10^{-3}
	50	10 000	(0.1532)	8258.69*	0.03	0.1571	1.568×10^{-3}
$\mathbf{G}_{0.1}$ vs. $\mathbf{G}_{\mathrm{Q}}$, different evecs	5	2300	0.7293	1.16	0.00	—	1.717×10^{-3}
	20	1800	0.5435	26.12	0.01	—	1.648×10^{-3}
	50	9200	0.4293	7009.14*	0.03	—	1.693×10^{-3}
$\mathbf{G}_{0.5}$ vs. $\mathbf{G}_{\mathrm{Q}}$, different evecs	5	2300	0.5276	1.12	0.00	—	2.607×10^{-3}
	20	4000	(0.2979)	128.01	0.01	0.3023	1.944×10^{-3}
	50	10 000	(0.2009)	8466.77*	0.03	0.2044	1.533×10^{-3}

because $\mathbf{G}_{0.5}^2$ has rather small eigenvalues, which translate into extremely large magnitudes of higher-order terms, $d_{i,j,1}$ in (22). For $p = 200$, accurate series evaluation of $\bar{\rho}$ seemed infeasible in most combinations for too much computational time and/or memory required. The only exception in the present test cases is the comparison between two $\mathbf{G}_{\mathrm{Q}}$'s with different eigenvectors where the series converged fairly rapidly (Table 6).

As expected, the Monte Carlo evaluation always yielded an unbiased estimate with the confidence intervals having the intended nominal coverage of the converged series where available. As implemented in the `qfratio` package, iterations in the order of 10^4 can be executed within a second, so might be favored over the new series expression when the computational burden is a constraint and stochastic error is tolerated. Interestingly, relative accuracies of the series and Monte Carlo evaluations showed a qualitatively similar pattern of variation across eigenvalue conformations for a given average measure and p; the Monte Carlo evaluation tended to yield a large SE when the series evaluation or its error bound was slow to converge (Figs. 2–4). In other words, eigenvalue conformations seem to largely determine the accuracy at which average measures can be empirically evaluated.

The delta method approximations by Hansen & Houle (2008, 2009) can be evaluated almost instantaneously (from ∼0.1 ms to ∼7 ms in $p = 5$ and 200), as they are straightforward arithmetic operations once eigenvalues of the argument matrix are obtained. For $\bar{c}$ in large p, they showed notable empirical accuracy (Fig. 2; Table 1) (see also Hansen & Houle, 2008). In many other cases, however, they yielded inaccurate values, often deviating from the converged series by >5%, and it is even difficult to discern a consistent trend in the directions of errors (Figs. 2–4; Tables 2, 4, 5). Except in the former cases, it is rather questionable when those approximations can be reliably used.

4 Discussion

The present paper derived new expressions for average evolvability measures using results on the moments of ratios of quadratic forms in normal variables. These series expressions are in theory exact, although for practical evaluation summation must be truncated, yielding truncation error. A great advantage in the present approach is that a hard upper bound for the truncation error is known for some of the measures: namely, average conditional evolvability $\bar{c}$, average respondability $\bar{r}$, average flexibility $\bar{f}$, and average response difference $\bar{d}$ (above). This is unlike the confidence intervals from Monte Carlo evaluations, which always involve some uncertainty. In addition, empirically speaking, evaluation of most of the average measures, excepting average autonomy $\bar{a}$ and average response correlation $\bar{\rho}$, can be faster than Monte Carlo evaluation for modest-sized problems (e.g., $p \leq 50$), depending on the accuracy aimed at.

As a practical guide for calculating average measures, the series evaluation is recommended whenever numerical convergence can be reached within reasonable computational time. Its feasibility largely depends on the eigenvalues of $\mathbf{G}$, but it is always easy to check whether convergence has been reached by inspecting a profile of the partial sum and by using error bounds when available. When the computational time and/or memory is a constraint (for $\bar{a}$ and $\bar{\rho}$ in large p with many small eigenvalues), then an alternative will be Monte Carlo evaluation, which is adequate because the evolvability measures considered herein have finite second moments

(Appendix B). The delta method approximations as proposed by Hansen & Houle (2008, 2009) can yield rather inaccurate values with hardly predictable errors (above), thus will not remain a viable option in estimating average measures, especially in light of immediate availability of powerful computers in recent years.

It should be remembered that, in empirical analyses, $\mathbf{G}$ is almost always estimated with uncertainty, so that estimates of (average) evolvability measures involve error and potentially bias (Houle & Meyer, 2015; Grabowski & Porto, 2017). Although uncertainty may partially be incorporated in a Bayesian sampling of $\mathbf{G}$ (Aguirre *et al.*, 2014) or a parametric-bootstrap-like sampling with multivariate normal approximation (Houle & Meyer, 2015), these method will not by themselves deal with potential estimation bias. The present series expressions are not free from this problem, which need to be addressed in future investigations. It might be possible to derive sampling moments of the (average) measures under simple sampling conditions (e.g., Wishartness), as was recently done for eigenvalue dispersion indices of covariance and correlation matrices (Watanabe, 2022). A potentially promising fact is that the expectations of zonal and invariant polynomials of Wishart matrices are known (Constantine, 1963; Davis, 1980). It remains to be done to derive equivalent results to polynomials in more complicated functions of random matrices like the ones appearing in the present expressions. Apart from that, simulation-based approaches would also be useful in obtaining rough estimates of sampling error and bias (Haber, 2011; Grabowski & Porto, 2017). In any case, the present expressions will greatly facilitate future investigations into sampling properties of average measures as they provide quasi-exact values given (estimates of) $\mathbf{G}$, unlike the approximate methods used in previous studies.

Although the present paper dealt with six evolvability measures, this is not to endorse all of them. Some of these measures may be disfavored from a biological or other theoretical ground. For example, Hansen & Houle (2008) criticized the use of (average) response correlation $\bar{\rho}$ in the random skewers analysis for its insensitivity to potentially different magnitudes of response vectors and recommended (average) response difference $\bar{d}$ as an alternative. (To be fair, Cheverud (1996) argued for the use of $\bar{\rho}$ by claiming that there tends to be a large sampling bias in the magnitude of response vectors from empirical covariance matrices, but that remains to be justified.) Rohlf (2017) also pointed out difficulties in interpreting $\bar{\rho}$ without knowledge on the eigenvalues and eigenvectors of the matrices compared, and recommended the use of other methods for matrix comparison. The present author refrains from going into details on biological validity or utility of individual evolvability measures, but leave the decision as to which measures are to be used to individual studies.

It seems worth repeating one of Rohlf's (2017) insightful caveats here. Rohlf (2017) considered the angle between response vectors (arccosine of response correlation ρ) from two $\mathbf{G}$ matrices in bivariate traits and empirically showed that the angle can have a bimodal distribution when $\boldsymbol{\beta}$ is distributed uniformly on the unit circle. The interpretation of the mean of such a multimodal distribution will not be straightforward, although, at least as empirically observed, the multimodality may not be a universal phenomenon (see Machado *et al.*, 2018). At present, the possibility remains that other evolvability measures can have multimodal distributions as well. Generally speaking, it is known that the distribution of the ratio of quadratic forms $\mathbf{x}^T\mathbf{A}\mathbf{x}/\mathbf{x}^T\mathbf{B}\mathbf{x}$, where $\mathbf{B}$ is positive definite, has different functional forms across intervals bordered by eigenvalues

of $\mathbf{B}^{-1}\mathbf{A}$ (Hillier, 2001; Forchini, 2002, 2005), around which distinct density peaks can potentially be observed. Formal evaluation of the conditions for multimodality seems rather tedious (if tractable at all) and is not pursued herein. It should at least be borne in mind that the average measures are insensitive to the potential multimodality and that, generally speaking, a scalar summary index like them can mask those nuanced aspects of the underlying distribution and the multivariate covariance structure.

The present paper suggested slight modifications for some evolvability measures to accommodate potentially singular $\mathbf{G}$ matrices. As originally defined by Hansen & Houle (2008), conditional evolvability c and autonomy a (and integration i) could not be evaluated for singular $\mathbf{G}$ as they required matrix inversion. This restriction can be released by using the generalized inverse as done above. Conditional evolvability c and autonomy a are zero unless $\boldsymbol{\beta} \in R(\mathbf{G})$. The corresponding average measures ($\bar{c}$ and $\bar{a}$) are also zero unless the same condition is satisfied with nonzero probability (i.e., when $\boldsymbol{\beta}$ is continuously distributed, its distribution is to be restricted to $R(\mathbf{G})$), as partially surmised by Hansen & Houle (2008). The new expressions enable to calculate c (4), a (6), i and their averages (56), (57) even when $\mathbf{G}$ is singular, as long as $\boldsymbol{\beta} \in R(\mathbf{G})$. Biologically speaking, the singularity of $\mathbf{G}$ represents the presence of absolute genetic constraint, where no genetic variance is available in certain directions. Although detection of absolute constraint in empirical systems requires different frameworks than evolvability measures (Mezey & Houle, 2005; Hine & Blows, 2006; Meyer & Kirkpatrick, 2008; Pavlicev *et al.*, 2009), the present development will facilitate investigations into systems with absolute constraints. However, it is necessary to distinguish the singularity in a true $\mathbf{G}$ matrix and that in its empirical estimates ($\hat{\mathbf{G}}$, say) due to small sample size. It should be viewed with much caution whether evolvability measures can be meaningfully applied to the latter, because $R\left(\hat{\mathbf{G}}\right)$ as a space fluctuates due to random sampling when the sample size is not large enough to span the entire $R(\mathbf{G})$.

Previous treatments of evolvability measures (Hansen & Houle, 2008; Kirkpatrick, 2009; Marroig *et al.*, 2009; Bolstad *et al.*, 2014) and random skewers (Revell, 2007) almost exclusively considered the simple condition where selection gradients are spherically distributed, $\boldsymbol{\beta} \sim N_p(\mathbf{0}_p, \sigma^2\mathbf{I}_p)$, or uniformly distributed on a unit hypersphere. As detailed in Appendix C, the present results can be extended to a more general condition where the selection gradient has an arbitrary normal distribution, $\boldsymbol{\beta} \sim N_p(\boldsymbol{\eta}, \boldsymbol{\Sigma})$. These extended results can potentially be useful when interest is in characterizing and comparing evolvability of populations under a certain directional and/or correlated selection regime. Of course, the original expressions of evolvability measures for a fixed $\boldsymbol{\beta}$ can be used when the selection is considered deterministic, so the utility of the new extended expressions will be in when the selection is considered random but potentially directional and/or correlated.

Chevin (2013) theoretically investigated the rate of adaptation (change in log average fitness across generations) in the same condition where $\boldsymbol{\beta}$ has a general normal distribution. He notably showed that, when stabilizing selection is assumed to be weak, the rate of adaptation is approximately equal to $\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}$. He called this quantity "evolvability", ascribing it to Hansen & Houle (2008), and discussed at length on its interpretation. However, that quantity should be distinguished from the evolvability e as defined by Hansen *et al.* (2003b) and Hansen & Houle

(2008), which is standardized by the norm of the selection gradient $\boldsymbol{\beta}$.[4] This standardization can also be seen as projection of the selection gradients onto a unit hypersphere (Mardia & Jupp, 1999); directional and correlational selections can be expressed as concentration of density on the hypersphere. Accordingly, Chevin's quantity and the present expressions for evolvability measures under general normal selection regimes have different interpretations and utility. The former would be more appropriate as a measure of absolute rate of adaptation, whereas the latter seems to be more appropriate for characterizing $\mathbf{G}$ matrices independently of the magnitude of selection.

The present expressions for a general normal case would be particularly useful when the selection gradients are assumed to be restricted to a certain subspace rather than distributed across the entire trait space, i.e., when $\boldsymbol{\Sigma}$ is singular. Such restriction can potentially arise from biological reasons, but probably are more often encountered from structural constraints in certain types of traits, e.g., shape variables. When the support of the (nominal) traits is restricted to a certain subspace, it would be sensible to assume that selection is also restricted to the same subspace (e.g., when geometric shapes are concerned, it is senseless to consider the components of selection corresponding to geometric rotation or translocation). Along with the accommodation of potentially singular $\mathbf{G}$ matrices, this extension allows for application of average measures to various types of traits encountered in the current evolutionary biology.

Appendix A Zonal and invariant polynomials

This paper evaluates the expectations of ratios of quadratic forms using the zonal and invariant polynomials, mainly following Smith (1989) and Hillier *et al.* (2009, 2014). These polynomials are widely used in the literature of multivariate distribution theory, primarily to integrate out components affected by orthogonal rotations from functions of random matrices. To be precise, they are not strictly required to derive results relevant to the present applications (see Bao & Kan, 2013), but allow for a conceptually elegant derivation and provide alternative interpretations of some of the results as hypergeometric functions of matrix argument. A brief introduction to these mathematical toolkits is provided in this section, as they appear to have attained little use in the biological literature. For more details, readers are directed to, e.g., James (1964), Muirhead (1982, 2006), Takemura (1984), Gross & Richards (1987), Mathai *et al.* (1995), and Chikuse (2003, appendix A).

A.1 Notations

In this section, $\mathbf{Y}$ denotes a $p \times p$ symmetric matrix with eigenvalues $y_1, y_2, \ldots, y_p$. $\mathbf{H}$ denotes a $p \times p$ orthogonal matrix, and the space of $p \times p$ orthogonal matrices is denoted by $O(p)$.

For indexing polynomials, (ordered) partitions κ of positive integers k are used; $\kappa = (k_1, k_2, \ldots)$, where k_i are nonnegative integers that satisfy $k_1 \geq k_2 \geq \ldots$ and $\sum_i k_i = k$. For instance, the possible partitions of $k = 4$ are $(4), (3,1), (2,2), (2,1,1)$, and $(1,1,1,1)$. The

[4] This confusion seems to have arisen from the fact that $\boldsymbol{\beta}$ is often (sometimes implicitly) standardized as a unit vector in the literature dealing with evolvability measures (Hansen & Houle, 2008; Marroig *et al.*, 2009; Grabowski & Porto, 2017). It seems advisable to use a different symbol for standardized selection gradients, e.g., $\mathbf{u}$ herein.

partition of k with a single nonzero integer part is called the top-order partition of k, and is denoted by $[k]$.

The Pochhammer's symbol is

$$\begin{aligned}(a)_k =& a(a+1)\dots(a+k-1) \\ =& \frac{\Gamma(a+k)}{\Gamma(a)}, \quad a \neq 0, -1, -2, \dots,\end{aligned} \tag{27}$$

where $\Gamma(\cdot)$ is the gamma function, with the understanding $(a)_0 = 1$. A multivariate extension $(a)_\kappa$ (indexed by a partition) is defined as

$$(a)_\kappa = \prod_{i=1}^{m} \left(a - \frac{1}{2}(i-1)\right)_{k_i}. \tag{28}$$

A.2 Definitions of zonal polynomials

The zonal polynomials are a certain class of polynomials in matrix entries which has several useful properties. The primary definition of the zonal polynomials comes from the group representation theory (James, 1960, 1961, 1964), which tends to be too abstract to comprehend for anyone new to the concept. There are alternative, more operational definitions of the zonal polynomials (Muirhead, 1982; Takemura, 1984), which would be more approachable to most biologists. However, those alternative definitions do not generalize into the invariant polynomials of multiple matrix arguments, which are also relevant to the present application. For this reason, the representation-theoretic definition is briefly outlined herein, omitting technical details. See Gross & Richards (1987) for a readable and sufficiently comprehensive introduction to the topic.

Consider the vector space V_k of homogeneous polynomials $\varphi(\mathbf{Y})$ of degree k in the elements of the symmetric matrix $\mathbf{Y}$. The transformation with a $p \times p$ invertible matrix $\mathbf{L}$,

$$\mathbf{Y} \to \mathbf{L}\mathbf{Y}\mathbf{L}^T,$$

transforms the vector space by

$$\varphi \to T(\mathbf{L})(\varphi) : (T(\mathbf{L})(\varphi))(\mathbf{Y}) = \varphi\left(\mathbf{L}^{-1}\mathbf{Y}\mathbf{L}^{-1^T}\right).$$

(Technically, $T(\mathbf{L})$ here is a representation, which abstracts the transformation of the right-hand side by a linear transformation.) It is known that the vector space V_k decomposes into a direct sum of irreducible invariant subspaces V_κ as

$$V_k = \bigoplus_\kappa V_\kappa,$$

where the direct summation runs over all partitions κ of k into not more than p parts. By V_κ being invariant, it is meant that $T(\mathbf{L})V_\kappa \subset V_\kappa$ for any invertible $\mathbf{L}$, and by them being irreducible invariant, it is meant that they do not contain any proper invariant subspace by themselves. It is further known that each V_κ contains a unique one-dimensional subspace that is invariant under $\mathbf{Y} \to \mathbf{H}\mathbf{Y}\mathbf{H}^T$ (restricting the above transformation to $\mathbf{L} \in O(p)$). Then, $[\mathrm{tr}\,(\mathbf{Y})]^k \in V_k$ has a

unique decomposition

$$[\mathrm{tr}\,(\mathbf{Y})]^k = \sum_{\kappa} C_\kappa(\mathbf{Y}) \tag{29}$$

into some polynomials $C_\kappa(\mathbf{Y}) \in V_\kappa$ which satisfy $C_\kappa(\mathbf{Y}) = C_\kappa\big(\mathbf{HYH}^T\big)$. These polynomials $C_\kappa(\mathbf{Y})$ are called the zonal polynomials corresponding to the partitions κ. In short, the zonal polynomials of degree k constitute a basis of the subspace of V_k that is invariant under the transformation $\mathbf{Y} \to \mathbf{HYH}^T$, each corresponding to the irreducible subspace represented by a partition κ of k. In terms of (29), they can be regarded as a generalization of powers of scalars into matrices.

These polynomials can be scaled by arbitrary constants, and there exist several different normalizations in the literature. The normalization used in (29), with all coefficients for $C_\kappa(\mathbf{Y})$ being unity, is primarily for convenience of notation.

A.3 Basic properties of zonal polynomials

A consequence of the above definition with (29) is that $C_\kappa(\mathbf{Y})$ are polynomials in the eigenvalues $y_1, \ldots, y_p$ of $\mathbf{Y}$ (note that $\mathrm{tr}\,(\mathbf{Y}) = \mathrm{tr}\,\big(\mathbf{HYH}^T\big) = \sum_i y_i$). It also entails that they are symmetric against permutation of the eigenvalues. Alternatively, or equivalently, they can be expressed as polynomials in the traces of matrix powers $\mathrm{tr}\,\big(\mathbf{Y}^j\big)$ $(= \sum_i y_i^j)$. Unfortunately, no closed-form expressions are known for the zonal polynomials, except for a few special cases, so the coefficients in these polynomials in general need to be determined recursively (see Muirhead, 1982, 2006).

The zonal polynomials are originally defined for symmetric matrices, but the definition can be extended. If $\mathbf{X}$ is positive definite and $\mathbf{Y}$ is symmetric, the eigenvalues of $\mathbf{XY}$ and $\mathbf{X}^{1/2}\mathbf{YX}^{1/2}$ are the same. Therefore, a zonal polynomial of $\mathbf{XY}$ is defined as $C_\kappa(\mathbf{XY}) = C_\kappa\big(\mathbf{X}^{1/2}\mathbf{YX}^{1/2}\big)$.

Properties of the zonal polynomials have been extensively investigated (Constantine, 1963; James, 1964; Muirhead, 1982). Elementary properties involve the following:

- $C_\kappa(a\mathbf{Y}) = a^k C_\kappa(\mathbf{Y})$ holds for any constant a.
- If the number of parts in κ exceeds the rank of $\mathbf{Y}$, $C_\kappa(\mathbf{Y}) = 0$ (Muirhead, 1982, corollary 7.2.4(i)).
- If $\mathbf{Y}$ is positive definite, $C_\kappa(\mathbf{Y}) > 0$ (Muirhead, 1982, corollary 7.2.4(ii)).
- $C_\kappa(\mathbf{Y}) = C_\kappa(\mathrm{diag}\,(\mathbf{Y}, \mathbf{0}))$, where the right-hand side is with respect to a block diagonal matrix with a square matrix of 0's of arbitrary dimension.

One of the most important properties of the zonal polynomials is the following integral identity.

Proposition 2 (James (1960))**.** *Let* $\mathbf{A}$ *and* $\mathbf{Y}$ *be* $p \times p$ *symmetric matrices. Then,*

$$\int_{O(p)} C_\kappa\big(\mathbf{AHYH}^T\big)\,(\mathrm{d}\mathbf{H}) = \frac{C_\kappa(\mathbf{A})\,C_\kappa(\mathbf{Y})}{C_\kappa(\mathbf{I}_m)}, \tag{30}$$

where $(\mathrm{d}\mathbf{H})$ *is the normalized invariant measure on* $O(p)$ *that satisfy* $\int_{O(p)}(\mathrm{d}\mathbf{H}) = 1$.

Accessible proofs are given in Muirhead (1982, theorem 7.2.5) and Gross & Richards (1987, proposition 5.5).

A.4 Top-order zonal polynomials

The zonal polynomials corresponding to the top-order partition, $[k] = (k)$, are called the top-order zonal polynomials and appear frequently in the multivariate distribution theory. An explicit expression of the top-order zonal polynomials is (James, 1964, (123)):

$$C_{[k]}(\mathbf{Y}) = \frac{k!}{\left(\frac{1}{2}\right)_k} \sum_{\sum_j i_j j = k} \left[\prod_{j=1}^{p} \frac{1}{i_j!} \left(\frac{s_j}{2j} \right)^{i_j} \right], \tag{31}$$

where $s_j = \sum_{i=1}^{p} y_i^j = \operatorname{tr}\left(\mathbf{Y}^j\right)$ and the summation is over all such combinations of the nonnegative integers i_j that satisfy $\sum_{j=0}^{k} i_j j = i_1 + 2i_2 + 3i_3 + \cdots = k$. An alternative expression is given in Gross & Richards (1987, (6.8.1)).

For the identity matrix $\mathbf{I}_p$, the expression simplifies as:

$$C_{[k]}(\mathbf{I}_p) = \frac{\left(\frac{p}{2}\right)_k}{\left(\frac{1}{2}\right)_k}. \tag{32}$$

A.5 Hypergeometric functions of matrix argument

With the zonal polynomials as a generalization of powers into matrices, the hypergeometric functions of matrix argument are defined as

$${}_mF_n\left(a_1, \ldots, a_m; b_1, \ldots, b_n; \mathbf{Y}\right) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \ldots (a_m)_\kappa}{(b_1)_\kappa \ldots (b_n)_\kappa} \frac{C_\kappa(\mathbf{Y})}{k!}, \tag{33}$$

where the summation $\sum_\kappa$ is across all ordered partitions κ of k, and $(x)_\kappa$ is defined as in (28). Conditions for convergence are as follows: (i) when $m \leq n$, the series converges for all $\mathbf{Y}$; (ii) when $m = n+1$, the series converges for all $\|\mathbf{Y}\| < 1$, where $\|\mathbf{Y}\|$ is the maximum of the absolute values of the eigenvalues; (iii) when $m > n+1$, the series diverges for all $\mathbf{Y} \neq 0$, unless the series terminates (when any a_i is a negative integer). The following paragraphs list a few results relevant to the present applications. More extensive treatments on this topic can be found in, e.g., Muirhead (1982), Gross & Richards (1987), and Chikuse (2003).

If any of a_i is $1/2$, the series can be expressed as a weighted sum of top-order zonal polynomials, as $(1/2)_\kappa = 0$ unless $\kappa = [k]$ (see (28)):

$${}_{m+1}F_n\left(a_1, \ldots, a_m, \frac{1}{2}; b_1, \ldots, b_n; \mathbf{Y}\right) = \sum_{k=0}^{\infty} \frac{(a_1)_k \ldots (a_m)_k \left(\frac{1}{2}\right)_k}{(b_1)_k \ldots (b_n)_k} \frac{C_{[k]}(\mathbf{Y})}{k!}. \tag{34}$$

Analogous to the scalar equivalent, ${}_0F_0$ is an exponential series but about the trace:

$$\begin{aligned}{}_0F_0\left(\mathbf{X}\right) &= \sum_{k=0}^{\infty}\sum_{\kappa}\frac{C_\kappa(\mathbf{Y})}{k!}\\ &= \sum_{k=0}^{\infty}\frac{\left(\operatorname{tr}\mathbf{Y}\right)^k}{k!}\\ &= \operatorname{etr}\left(\mathbf{Y}\right).\end{aligned} \tag{35}$$

Provided that all eigenvalues of $\mathbf{Y}$ are less than 1 in absolute values, ${}_1F_0$ is a generalization of the binomial series (Muirhead, 1982, corollary 7.3.5):

$$\begin{aligned}{}_1F_0\left(a;\mathbf{Y}\right) =& \sum_{k=0}^{\infty}\sum_{\kappa}\left(a\right)_\kappa\frac{C_\kappa(\mathbf{Y})}{k!}\\ =&\left|\mathbf{I}_m-\mathbf{Y}\right|^{-a}, \quad \left\|\mathbf{Y}\right\|<1.\end{aligned} \tag{36}$$

A.6 Invariant polynomials of multiple matrix arguments

The invariant polynomials of multiple matrix arguments provide some extension of the zonal polynomials into multiple matrices (Davis, 1979, 1980; Chikuse, 1980; Chikuse & Davis, 1986a). Accessible introductions to the topic are found in, e.g., Chikuse & Davis (1986b) and Mathai *et al.* (1995, appendix).

Let $V_{k_1,\ldots,k_r} = V_{k_1}\otimes\cdots\otimes V_{k_r}$ be the vector space of polynomials of degree $k_1,\ldots,k_r$ in the elements of $p\times p$ symmetric matrices $\mathbf{Y}_1,\ldots,\mathbf{Y}_r$, respectively. Consider the following simultaneous transformation with a $p\times p$ invertible matrix $\mathbf{L}$:

$$\mathbf{Y}_1\to\mathbf{L}\mathbf{Y}_1\mathbf{L}^T,\ldots,\mathbf{Y}_r\to\mathbf{L}\mathbf{Y}_r\mathbf{L}^T.$$

First noting the subspaces in $V_{k_1,\ldots,k_r}$ that are invariant against this transformation and then by decomposing them into irreducible ones, one has the following decomposition:

$$\begin{aligned}V_{k_1,\ldots,k_r}\left(\mathbf{Y}_1,\ldots,\mathbf{Y}_r\right) &= \bigoplus_{\kappa_1,\ldots,\kappa_r}\left\{V_{\kappa_1}\left(\mathbf{Y}_1\right)\otimes\cdots\otimes V_{\kappa_r}\left(\mathbf{Y}_r\right)\right\}\\ &= \bigoplus_{\kappa_1,\ldots,\kappa_r}\bigoplus_{\Phi}V_\Phi^{\kappa_1,\ldots,\kappa_r}\left(\mathbf{Y}_1,\ldots,\mathbf{Y}_r\right).\end{aligned}$$

Here, V_{κ_i} are analogous to V_κ discussed above but for $\mathbf{Y}_i$, and $V_\Phi^{\kappa_1,\ldots,\kappa_r}$ are irreducible invariant subspaces indexed by Φ and $\kappa_1,\ldots,\kappa_r$. Φ are certain ordered partitions of $2f$, with $f = k_1+\cdots+k_r$, into not more than p parts, whose range is determined by $\kappa_1,\ldots,\kappa_r$. κ_i run across partitions of k_i, $i=1,\ldots,r$, into not more than p parts. It is known that only those $V_\Phi^{\kappa_1,\ldots,\kappa_r}$ with partitions taking the form $\Phi = 2\phi$, where ϕ is a partition of f, contain a subspace invariant under the simultaneous transformation $\mathbf{Y}_1\to\mathbf{H}\mathbf{Y}_1\mathbf{H}^T,\ldots,\mathbf{Y}_r\to\mathbf{H}\mathbf{Y}_r\mathbf{H}^T$, which is one-dimensional. The corresponding polynomials are called the invariant polynomials, $C_\phi^{\kappa_1,\ldots,\kappa_r}(\mathbf{Y}_1,\ldots,\mathbf{Y}_r)$, indexed by $\kappa_1,\ldots,\kappa_r$ and $\phi\in\kappa_1\cdots\kappa_r$. (There is a potential issue of nonuniqueness, but this is not concerned in most applications.)

Explicit forms of the invariant polynomials are even less obvious than the zonal polynomials,

but they are polynomials in matrix traces of the form $\operatorname{tr}(\mathbf{Y}_1^{a_1}\mathbf{Y}_2^{a_2}\dots\mathbf{Y}_r^{a_r})$, so that the total degrees in elements of $\mathbf{Y}_1,\dots,\mathbf{Y}_r$ are $k_1,\dots,k_r$, respectively.

An important generalization of Proposition 2 enabled by the invariant polynomials is the following (see also Mathai *et al.*, 1995, theorem A.3).

Proposition 3 (Davis (1980), Chikuse (1980))**.** *Let $\mathbf{A}_1,\dots,\mathbf{A}_r$ and $\mathbf{Y}_1,\dots,\mathbf{Y}_r$ be $p\times p$ symmetric matrices, and $\kappa_1,\dots,\kappa_r$ be ordered partitions of $k_1,\dots,k_r$ into not more than p parts.*

$$\int_{O(p)}\prod_{i=1}^{r}C_{\kappa_i}\left(\mathbf{A}_i\mathbf{H}\mathbf{Y}_i\mathbf{H}^T\right)(\mathrm{d}\mathbf{H})=\sum_{\phi\in\kappa_1\cdots\kappa_r}\frac{C_\phi^{\kappa_1,\dots,\kappa_r}(\mathbf{A}_1,\dots,\mathbf{A}_r)\,C_\phi^{\kappa_1,\dots,\kappa_r}(\mathbf{Y}_1,\dots,\mathbf{Y}_r)}{C_\phi(\mathbf{I}_p)},\tag{37}$$

where the summation runs across all permissible partitions ϕ of $f=k_1+\dots+k_r$ corresponding to $\kappa_1,\dots,\kappa_r$.

One notable corollary from this result is (Davis, 1980; Chikuse, 1980)

$$C_\phi^{\kappa_1,\dots,\kappa_r}(\mathbf{A},\dots,\mathbf{A})=\frac{C_\phi^{\kappa_1,\dots,\kappa_r}(\mathbf{I}_p,\dots,\mathbf{I}_p)}{C_\phi(\mathbf{I}_p)}C_\phi(\mathbf{A})\,.\tag{38}$$

The top-order invariant polynomials, in which $\phi=[f]$, the top-order partition of f, are of particular importance here. This partition occurs only when $\kappa_i=[k_i]$ for all i. An explicit form of the top-order invariant polynomials is (Chikuse, 1987; Hillier *et al.*, 2009):

$$C_{[f]}^{[k_1],\dots,[k_r]}(\mathbf{Y}_1,\dots,\mathbf{Y}_r)=\frac{\prod_{i=1}^{r}k_i!}{\left(\frac{1}{2}\right)_f}d_{k_1,\dots,k_r}.\tag{39}$$

Here, $d_{k_1,\dots,k_r}$ is defined as:

$$d_{k_1,\dots,k_r}=\sum_{\substack{\sum_{i_l}i_l\nu_{i_1,\dots,i_r}=k_l\\ l=1,\dots,r}}\left[\prod_{i_1=0}^{k_1}\cdots\prod_{i_r=0}^{k_r}\frac{1}{\nu_{i_1,\dots,i_r}!}\left(\frac{p_{i_1,\dots,i_r}}{2\sum_j i_j}\right)^{\nu_{i_1,\dots,i_r}}\right],\tag{40}$$

where i_j run from 0 to k_j (but at least one of i_j is nonzero), and $\nu_{i_1,\dots,i_r}$ are nonnegative integers; these collectively satisfy the r linear constraints $\sum_{\nu_1=0}^{k_1}\cdots\sum_{\nu_r=0}^{k_r}i_l\nu_{i_1,\dots,i_r}=k_l$ for $l=1,\dots,r$.[5] And $p_{i_1,\dots,i_r}$ are the coefficients for $\prod_{j=1}^{r}t_j^{i_j}$ in the expansion of the following:

$$\operatorname{tr}\left[(t_1\mathbf{Y}_1+\dots+t_r\mathbf{Y}_r)^{\sum_j i_j}\right].\tag{41}$$

Note that (39) simplifies into (31) when $r=1$. In addition, it holds that (Chikuse, 1987)

$$C_{[f]}^{[k_1],\dots,[k_r]}(\mathbf{I}_m,\dots,\mathbf{I}_m)=C_{[f]}(\mathbf{I}_m)\,,\tag{42}$$

regardless of individual values of k_i (see also (32)).

[5] A few examples will clarify the indexing in (40) (semicolons separate terms of the summation). For $(k_1,k_2)=(1,1)$: $(i_1,j_2)=(1,1),\nu_{1,1}=1$; $(\nu_1,\nu_2)=(1,0),(0,1),\nu_{1,0}=1,\nu_{0,1}=1$. For $(k_1,k_2)=(2,1)$: $(i_1,i_2)=(2,1),\nu_{2,1}=1$; $(i_1,i_2)=(2,0),(0,1),\nu_{2,0}=1,\nu_{0,1}=1$; $(i_1,i_2)=(1,1),(1,0),\nu_{1,1}=1,\nu_{1,0}=1$; $(i_1,i_2)=(1,0),(0,1),\nu_{1,0}=2,\nu_{0,1}=1$.

A.7 Example application to evolvability measures

As a demonstration for how above mathematical toolkits act in the present problem of evaluating moments of (multiple) ratios of quadratic forms, an expression of the average autonomy $\bar{a}$ is derived here, under the simplest case that $\boldsymbol{\beta} \sim N_p(\mathbf{0}_p, \sigma^2 \mathbf{I}_p)$, where σ^2 is an arbitrary positive constant, and $\mathbf{G}$ is nonsingular. The derivation here is a simplified version of that in Smith (1989). When $\mathbf{G}$ is singular, the approach of Bao & Kan (2013) can be used instead.

As discussed in the text, $\mathbf{u}$ is uniformly distributed on the unit hypersphere S^{p-1} under this assumption. Thus, the average autonomy $\bar{a}$, from the definition in (5), is to be evaluated as the following integral:

$$\bar{a} = \int_{S^{p-1}} \left(\mathbf{u}^T \mathbf{G} \mathbf{u}\right)^{-1} \left(\mathbf{u}^T \mathbf{G}^{-1} \mathbf{u}\right)^{-1} (\mathrm{d}\mathbf{u}), \tag{43}$$

where $(\mathrm{d}\mathbf{u})$ denotes the normalized uniform measure on S^{p-1} (see, e.g. Muirhead, 1982). To evaluate this, consider the following transformation:

$$\begin{aligned}
\left(\mathbf{u}^T \mathbf{G} \mathbf{u}\right)^{-1} =& \alpha_1 \left(1 - \mathbf{u}^T \left(\mathbf{I}_p - \alpha_1 \mathbf{G}\right) \mathbf{u}\right)^{-1} \\
=& \alpha_1 \sum_{k=0}^{\infty} \left(\mathbf{u}^T \left(\mathbf{I}_p - \alpha_1 \mathbf{G}\right) \mathbf{u}\right)^k \\
=& \alpha_1 \sum_{k=0}^{\infty} \sum_{\kappa} C_\kappa\left(\left(\mathbf{I}_p - \alpha_1 \mathbf{G}\right) \mathbf{u}\mathbf{u}^T\right),
\end{aligned} \tag{44}$$

where α_1 is a positive constant that satisfy $0 < \alpha_1 < 2/\lambda_{\max}(\mathbf{G})$ (so that $|\mathbf{u}^T \left(\mathbf{I}_p - \alpha_1 \mathbf{G}\right) \mathbf{u}| < 1$), the second equation is from the fact $(1-x)^{-1} = \sum_{k=0}^{\infty} x^k$ when $|x| < 1$, and the last equation is from (29) and the fact $\mathbf{u}^T \left(\mathbf{I}_p - \alpha_1 \mathbf{G}\right) \mathbf{u} = \mathrm{tr}\left(\left(\mathbf{I}_p - \alpha_1 \mathbf{G}\right) \mathbf{u}\mathbf{u}^T\right)$. Note that (44) can be written as ${}_1F_0\left(1; \left(\mathbf{I}_p - \alpha_1 \mathbf{G}\right) \mathbf{u}\mathbf{u}^T\right)$ from (36) and is uniformly convergent under the assumptions. (This series does not converge when $\mathbf{G}$ is singular.) By similar transformations, $\left(\mathbf{u}^T \mathbf{G}^{-1} \mathbf{u}\right)^{-1} = \alpha_2 \sum_{l=0}^{\infty} \sum_{\lambda} C_\lambda\left(\left(\mathbf{I}_p - \alpha_2 \mathbf{G}^{-1}\right) \mathbf{u}\mathbf{u}^T\right)$, where $0 < \alpha_2 < 2/\lambda_{\max}\left(\mathbf{G}^{-1}\right)$. Then, the above integral becomes:

$$\bar{a} = \alpha_1 \alpha_2 \int_{S^{p-1}} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{\kappa} \sum_{\lambda} C_\kappa\left(\left(\mathbf{I}_p - \alpha_1 \mathbf{G}\right) \mathbf{u}\mathbf{u}^T\right) C_\lambda\left(\left(\mathbf{I}_p - \alpha_2 \mathbf{G}^{-1}\right) \mathbf{u}\mathbf{u}^T\right) (\mathrm{d}\mathbf{u}). \tag{45}$$

Here, term-by-term integration is permissible because the series is uniformly convergent (see also Bao & Kan, 2013).

Note that integration of $\mathbf{u}$ over S^{p-1} is equivalent to that of $\mathbf{H}$ over $O(p)$ (e.g., Mardia & Jupp, 1999), whence the change of variables $\mathbf{u} = \mathbf{H}\mathbf{e}$ is possible, where $\mathbf{e} = (1, 0, \ldots 0)^T$ (or $\mathbf{e}$ could be any unit vector). By denoting $\mathbf{e}\mathbf{e}^T = \mathbf{E}$, the integral is now

$$\begin{aligned}
\bar{a} =& \alpha_1 \alpha_2 \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{\kappa} \sum_{\lambda} \int_{O(p)} C_\kappa\left(\left(\mathbf{I}_p - \alpha_1 \mathbf{G}\right) \mathbf{H}\mathbf{E}\mathbf{H}^T\right) C_\lambda\left(\left(\mathbf{I}_p - \alpha_2 \mathbf{G}^{-1}\right) \mathbf{H}\mathbf{E}\mathbf{H}^T\right) (\mathrm{d}\mathbf{H}) \\
=& \alpha_1 \alpha_2 \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{\kappa} \sum_{\lambda} \sum_{\phi \in \kappa \cdot \lambda} \frac{C_\phi^{\kappa,\lambda}\left(\mathbf{I}_p - \alpha_1 \mathbf{G}, \mathbf{I}_p - \alpha_2 \mathbf{G}^{-1}\right) C_\phi^{\kappa,\lambda}(\mathbf{E}, \mathbf{E})}{C_\phi(\mathbf{I}_p)},
\end{aligned} \tag{46}$$

by applying Proposition 3. Furthermore, using (38), $C_{\phi}^{\kappa,\lambda}(\mathbf{E},\mathbf{E}) = C_{\phi}^{\kappa,\lambda}(\mathbf{I}_p,\mathbf{I}_p)\, C_{\phi}(\mathbf{E})\,/C_{\phi}(\mathbf{I}_p)$. However, $C_{\phi}(\mathbf{E})$ is nonzero only if $\phi = [k+l]$, the top-order partition, because $\mathbf{E}$ has only one nonzero eigenvalue, which is unity. Hence, the summations for κ, λ, ϕ disappear, leaving only the terms corresponding to the top-order partitions, $[k]\,,[l]\,,[k+l]$. Also note that, from (42), $C_{[k+l]}^{[k],[l]}(\mathbf{I}_p,\mathbf{I}_p)\,/C_{[k+l]}(\mathbf{I}_p) = 1$. The integral finally becomes

$$\bar{a} = \alpha_1\alpha_2 \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{C_{[k+l]}^{[k],[l]}\left(\mathbf{I}_p - \alpha_1\mathbf{G}, \mathbf{I}_p - \alpha_2\mathbf{G}^{-1}\right)}{C_{[k+l]}(\mathbf{I}_p)}. \tag{47}$$

Inserting (32) yields (18) in the text. The expressions for all average measures presented in the main text under the same assumption can be derived in a similar manner.

A.8 Generating functions for top-order zonal and invariant polynomials

For practical evaluation of top-order zonal and invariant polynomials, it is more efficient to handle $d_{k_1,\ldots,k_r}$ in (39) and (40) than the polynomials themselves. Chikuse (1987) showed that the same $d_{k_1,\ldots,k_r}(\mathbf{A}_1,\ldots,\mathbf{A}_r)$ appear as coefficients in the following, which is the joint moment generating function of $\mathbf{x}^T\mathbf{A}_1\mathbf{x}/2,\ldots,\mathbf{x}^T\mathbf{A}_r\mathbf{x}/2$ for $\mathbf{x}\sim N_p(\mathbf{0}_p,\mathbf{I}_p)$:

$$|\mathbf{I}_p - t_1\mathbf{A}_1 - \cdots - t_r\mathbf{A}_r|^{-\frac{1}{2}} = \sum_{k_1=0}^{\infty}\cdots\sum_{k_r=0}^{\infty} d_{k_1,\ldots,k_r}(\mathbf{A}_1,\ldots,\mathbf{A}_r)\, t_1^{k_1}\ldots t_r^{k_r}. \tag{48}$$

Chikuse (1987) and Hillier *et al.* (2009, 2014) presented a few recursive algorithms to evaluate $d_{k_1,\ldots,k_r}$.

Bao & Kan (2013) and Hillier *et al.* (2014) presented further improvements for evaluating moments of ratios of quadratic forms in potentially noncentral normal variables, $\mathbf{x}\sim N_p(\boldsymbol{\mu},\mathbf{I}_p)$. Following those papers, the coefficients $h_{k_1,k_2}, \tilde{h}_{k_1;k_2}, \tilde{h}_{k_1;k_2,k_3}$ that appear in the following expansions are used in the present paper:

$$\begin{aligned}
&|\mathbf{I}_p - t_1\mathbf{A}_1 - t_2\mathbf{A}_2|^{-\frac{1}{2}} \exp\left(\frac{(1-t_1-t_2)\boldsymbol{\mu}^T\left(\mathbf{I}_p - t_1\mathbf{A}_1 - t_2\mathbf{A}_2\right)^{-1}\boldsymbol{\mu} - \boldsymbol{\mu}^T\boldsymbol{\mu}}{2}\right)\\
&\quad = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} h_{k_1,k_2}(\mathbf{A}_1,\mathbf{A}_2)\, t_1^{k_1}t_2^{k_2},
\end{aligned} \tag{49}$$

$$\begin{aligned}
&|\mathbf{I}_p - t_1\mathbf{A}_1 - t_2\mathbf{A}_2|^{-\frac{1}{2}} \exp\left(\frac{(1-t_2)\boldsymbol{\mu}^T\left(\mathbf{I}_p - t_1\mathbf{A}_1 - t_2\mathbf{A}_2\right)^{-1}\boldsymbol{\mu} - \boldsymbol{\mu}^T\boldsymbol{\mu}}{2}\right)\\
&\quad = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} \tilde{h}_{k_1;k_2}(\mathbf{A}_1;\mathbf{A}_2)\, t_1^{k_1}t_2^{k_2},
\end{aligned} \tag{50}$$

$$\begin{aligned}
&|\mathbf{I}_p - t_1\mathbf{A}_1 - t_2\mathbf{A}_2 - t_3\mathbf{A}_3|^{-\frac{1}{2}} \exp\left(\frac{(1-t_2-t_3)\boldsymbol{\mu}^T\left(\mathbf{I}_p - t_1\mathbf{A}_1 - t_2\mathbf{A}_2 - t_3\mathbf{A}_3\right)^{-1}\boldsymbol{\mu} - \boldsymbol{\mu}^T\boldsymbol{\mu}}{2}\right)\\
&\quad = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty} \tilde{h}_{k_1;k_2,k_3}(\mathbf{A}_1;\mathbf{A}_2,\mathbf{A}_3)\, t_1^{k_1}t_2^{k_2}t_3^{k_3}.
\end{aligned} \tag{51}$$

(49) and (50) are from Bao & Kan (2013), and (51) is a slight extension from (50) used in appendices B and C. All these simplify to (48) when $\boldsymbol{\mu} = \mathbf{0}_p$. Note that the first arguments in $\tilde{h}_{k_1;k_2}$ and $\tilde{h}_{k_1;k_2,k_3}$ are not interchangeable with the other arguments, unlike those in d_{k_1,k_2} or h_{k_1,k_2}.

Appendix B Moment of multiple ratio of quadratic forms

Among the average evolvability measures mentioned in the text, the expressions for $\bar{e}$, $\bar{c}$, $\bar{r}$, and $\bar{d}$ can be directly derived from Proposition 1. For $\bar{a}$, $\bar{f}$, and $\bar{\rho}$, an expression is required for the expectation of the form $\frac{(\mathbf{x}^T\mathbf{A}\mathbf{x})^l}{(\mathbf{x}^T\mathbf{B}\mathbf{x})^m(\mathbf{x}^T\mathbf{D}\mathbf{x})^n}$; it suffices here to concern integer-valued l. Smith (1989) derived such an expression for $\mathbf{B}, \mathbf{D}$ being positive definite. However, it is possible to derive the following, equivalent expression with a slightly general condition that $\mathbf{B}, \mathbf{D}$ being nonnegative definite, following a similar line of arguments as in Bao & Kan (2013).

Proposition 4. *Let $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \mathbf{I}_p)$, $\mathbf{A}$ be a symmetric $p \times p$ matrix, $\mathbf{B}, \mathbf{D}$ be symmetric, nonnegative definite $p \times p$ matrices, l be a positive integer, and m, n be positive real numbers. When the expectation of $\frac{(\mathbf{x}^T\mathbf{A}\mathbf{x})^l}{(\mathbf{x}^T\mathbf{B}\mathbf{x})^m(\mathbf{x}^T\mathbf{D}\mathbf{x})^n}$ exists, it is*

$$
\begin{aligned}
\mathrm{E}\left(\frac{(\mathbf{x}^T\mathbf{A}\mathbf{x})^l}{(\mathbf{x}^T\mathbf{B}\mathbf{x})^m(\mathbf{x}^T\mathbf{D}\mathbf{x})^n}\right) =& 2^{l-m-n}\alpha_{\mathbf{B}}^m\alpha_{\mathbf{D}}^n l! e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{(m)_i (n)_j \Gamma\left(\frac{p}{2}+l-m-n+k\right)}{2^k \left(\frac{1}{2}\right)_k \Gamma\left(\frac{p}{2}+l+i+j+k\right)} \\
& \cdot \frac{\left(\frac{1}{2}\right)_{i+j+k+l}}{i!j!k!} C_{[i+j+k+l]}^{[i],[j],[k],[l]}\left(\mathbf{I}_p - \alpha_{\mathbf{B}}\mathbf{B}, \mathbf{I}_p - \alpha_{\mathbf{D}}\mathbf{D}, \boldsymbol{\mu}^T\boldsymbol{\mu}, \mathbf{A}\right) \qquad (52) \\
=& 2^{l-m-n}\alpha_{\mathbf{B}}^m\alpha_{\mathbf{D}}^n l! e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{(m)_i (n)_j \Gamma\left(\frac{p}{2}+l-m-n+k\right)}{2^k \left(\frac{1}{2}\right)_k \Gamma\left(\frac{p}{2}+l+i+j+k\right)} \\
& \cdot d_{i,j,k,l}\left(\mathbf{I}_p - \alpha_{\mathbf{B}}\mathbf{B}, \mathbf{I}_p - \alpha_{\mathbf{D}}\mathbf{D}, \boldsymbol{\mu}^T\boldsymbol{\mu}, \mathbf{A}\right) \qquad (53) \\
=& 2^{l-m-n}\alpha_{\mathbf{B}}^m\alpha_{\mathbf{D}}^n l! \frac{\Gamma\left(\frac{p}{2}+l-m-n\right)}{\Gamma\left(\frac{p}{2}+l\right)} \\
& \cdot \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(m)_i (n)_j}{\left(\frac{p}{2}+l\right)_{i+j}} \tilde{h}_{l;i,j}(\mathbf{A}; \mathbf{I}_p - \alpha_{\mathbf{B}}\mathbf{B}, \mathbf{I}_p - \alpha_{\mathbf{D}}\mathbf{D}), \qquad (54)
\end{aligned}
$$

where $\alpha_{\mathbf{B}}, \alpha_{\mathbf{D}}$ are to satisfy $0 < \alpha_{\mathbf{B}} < 2/\lambda_{\max}(\mathbf{B})$ and $0 < \alpha_{\mathbf{D}} < 2/\lambda_{\max}(\mathbf{D})$.

Proof is omitted because it is almost identical to that of the proposition 3 in Bao & Kan (2013, pp. 237–243), but by using (51) in place of (49).

Remark. Conditions for the existence of the above moment were stated by Smith (1989) for positive definite $\mathbf{B}, \mathbf{D}$; in the notation here, the moment exists if and only if $\frac{p}{2} + l > m + n$. For positive semidefinite $\mathbf{B}, \mathbf{D}$, dealing with their null spaces is critical (see Roberts, 1995). Deriving general existence conditions in that case does not seem straightforward, so only the following simple cases are noted here. *(i) When $N(\mathbf{D}) \subseteq N(\mathbf{B})$ (or $N(\mathbf{B}) \subseteq N(\mathbf{D})$):* In the former case, it is possible to find a positive constant ϵ that always satisfies $(\mathbf{x}^T\mathbf{B}\mathbf{x})^m(\mathbf{x}^T\mathbf{D}\mathbf{x})^n \geq (\epsilon\mathbf{x}^T\mathbf{B}\mathbf{x})^{m+n}$, so that $(\mathbf{x}^T\mathbf{A}\mathbf{x})^l / \left[(\mathbf{x}^T\mathbf{B}\mathbf{x})^m(\mathbf{x}^T\mathbf{D}\mathbf{x})^n\right] \leq \epsilon^{-(m+n)}(\mathbf{x}^T\mathbf{A}\mathbf{x})^l/(\mathbf{x}^T\mathbf{B}\mathbf{x})^{m+n}$. The desired moment of the left-hand side exists if and only if that of the right-hand side does, and the necessary and sufficient condition for the latter is given in the proposition 1 of Bao & Kan (2013), which

encompasses the positive definite case above. The same argument applies when $N(\mathbf{B}) \subseteq N(\mathbf{D})$ with obvious modifications. *(ii) When* $N(\mathbf{D}) \nsubseteq N(\mathbf{B})$ *and* $N(\mathbf{B}) \nsubseteq N(\mathbf{D})$: Consider the space $F = \{\mathbf{y} : \mathbf{y} \perp N(\mathbf{B}), \mathbf{y} \perp N(\mathbf{D}), \mathbf{y} \in \mathbb{R}^p\}$. Assuming $F \neq \varnothing$, it is possible to find a positive semi-definite matrix $\mathbf{F}$ that always satisfies $(\mathbf{x}^T\mathbf{A}\mathbf{x})^l / \left[(\mathbf{x}^T\mathbf{B}\mathbf{x})^m(\mathbf{x}^T\mathbf{D}\mathbf{x})^n\right] \leq (\mathbf{x}^T\mathbf{A}\mathbf{x})^l/(\mathbf{x}^T\mathbf{F}\mathbf{x})^{m+n}$, by choosing arbitrary small positive numbers and a basis of F as its nonzero eigenvalues and eigenvectors, respectively. The proposition 1 of Bao & Kan (2013) applied to the right-hand side defines a sufficient condition for the existence of the moment. This is admittedly a restrictive condition, and the existence condition in the case $F = \varnothing$ is not always obvious.

In the present applications, the moment existence conditions can be easily confirmed for $\bar{e}$, $\bar{c}$, $\bar{r}$, and $\bar{d}$ by directly applying the proposition 1 of Bao & Kan (2013). $\bar{a}$ and $\bar{f}$ always fall within the case (i) above, and the moments exist as long as they are defined (in case of $\bar{f}$, $\boldsymbol{\beta}$ should not be restricted onto $N(\mathbf{G})$). On the other hand, $\bar{\rho}$ may fall within either (i) or (ii) depending on the structures of the two $\mathbf{G}$'s. Even in the latter case and if $F = \varnothing$, which is a pathologic condition in this application, this specific moment exists as long as ρ is defined with probability 1, since $-1 \leq \rho \leq 1$ whenever defined (above). In addition, the same criteria can be used to confirm that these evolvability measures have finite second moments, which formally validate the use of Monte Carlo evaluations.

Appendix C Evolvability measures in general normal cases

C.1 Conditions considered

This section considers a general p-variate normality condition for the selection gradient, $\boldsymbol{\beta} \sim N_p(\boldsymbol{\eta}, \boldsymbol{\Sigma})$. Regarding ratios of quadratic forms, however, tractable results are available only for normal variables with a spherical covariance structure, i.e., $N_q(\boldsymbol{\mu}, \mathbf{I}_q)$ for some q and arbitrary $\boldsymbol{\mu}$. Let $\mathbf{A}$ be an arbitrary $p \times p$ symmetric matrix and $\mathbf{K}$ be a $p \times q$ matrix that satisfies $\mathbf{K}\mathbf{K}^T = \boldsymbol{\Sigma}$, where q $(\leq p)$ is the rank of $\boldsymbol{\Sigma}$. Writing $\boldsymbol{\beta} = \boldsymbol{\eta} + \mathbf{K}\boldsymbol{\gamma}_0$ with $\boldsymbol{\gamma}_0 \sim N_q(\mathbf{0}_q, \mathbf{I}_q)$, the quadratic form in $\boldsymbol{\beta}$ with respect to $\mathbf{A}$ can be written as $\boldsymbol{\beta}^T\mathbf{A}\boldsymbol{\beta} = (\boldsymbol{\eta} + \mathbf{K}\boldsymbol{\gamma}_0)^T\mathbf{A}(\boldsymbol{\eta} + \mathbf{K}\boldsymbol{\gamma}_0)$ (see also Mathai & Provost, 1992, chapter 3). Certain assumptions are required to express this quadratic form as one in normal variables with a spherical covariance structure.

(1) When $\boldsymbol{\Sigma}$ is nonsingular In this case, $p = q$ and $\mathbf{K}$ is invertible. It suffices to consider the transformation $\boldsymbol{\gamma} = \mathbf{K}^{-1}\boldsymbol{\beta} \sim N_p(\boldsymbol{\mu}, \mathbf{I}_p)$, where $\boldsymbol{\mu} = \mathbf{K}^{-1}\boldsymbol{\eta}$. The quadratic form can be rewritten as $\boldsymbol{\beta}^T\mathbf{A}\boldsymbol{\beta} = \boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{A}\mathbf{K}\boldsymbol{\gamma}$, a new quadratic form in $\boldsymbol{\gamma}$ with respect to $\mathbf{K}^T\mathbf{A}\mathbf{K}$, to which results below apply.

(2) When $\boldsymbol{\Sigma}$ is singular In this case, $\mathbf{K}$ is not invertible and at least one of the following additional conditions need to be satisfied for the results below to be applicable. *(i)* $\boldsymbol{\eta} \in R(\boldsymbol{\Sigma})$ *(including* $\boldsymbol{\eta} = \mathbf{0}_p$*)*: Under this condition, it holds that $\boldsymbol{\eta} = \mathbf{K}\mathbf{K}^-\boldsymbol{\eta}$ (e.g., Schott, 2016, theorem 5.25) since $R(\boldsymbol{\Sigma}) = R(\mathbf{K})$. It is then possible to consider $\boldsymbol{\beta}^T\mathbf{A}\boldsymbol{\beta} = \boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{A}\mathbf{K}\boldsymbol{\gamma}$ with $\boldsymbol{\gamma} = \mathbf{K}^-\boldsymbol{\beta} \sim N_p(\boldsymbol{\mu}, \mathbf{I}_p)$, where $\boldsymbol{\mu} = \mathbf{K}^-\boldsymbol{\eta}$, as a generalization of the nonsingular case. *(ii)* $R(\mathbf{A}) \subseteq R(\boldsymbol{\Sigma})$: From the symmetry of $\mathbf{A}$ and the fact that $R(\mathbf{K}) = R\left(\mathbf{K}^{-T}\right)$, it holds that $\mathbf{A} = \mathbf{K}^{-T}\mathbf{K}^T\mathbf{A}\mathbf{K}\mathbf{K}^-$ under this condition (Schott, 2016, theorem 5.25). Then the same

transformation as for *(i)* applies. (Utility of this condition in the present applications is relatively limited, since all evolvability measures except ρ involve at least one $\mathbf{I}_p$ in place of $\mathbf{A}$.) *(iii)* $\mathbf{A}\boldsymbol{\eta} = \mathbf{0}_p$: Under this condition, the above quadratic form simplifies into $\boldsymbol{\gamma}_0^T\mathbf{K}^T\mathbf{A}\mathbf{K}\boldsymbol{\gamma}_0$, to which results below apply with $\boldsymbol{\mu} = \mathbf{0}_q$.

In practice, the conditions (1) or (2)(i) above seem to cover most cases of biological interest. That is, if the variation of selection gradient $\boldsymbol{\beta}$ is restricted to a certain subspace (condition (2)), then it seems plausible that the mean selection gradient $\boldsymbol{\eta}$ is also in the same subspace (condition (2)(i)). Otherwise, $\boldsymbol{\beta}$ ought to contain a nonzero deterministic component ($\boldsymbol{\eta}$) that is strictly unaffected by (linearly independent of) its potential variability (represented by $\boldsymbol{\Sigma}$).

C.2 Average evolvability measures

Expressions for average evolvability measures for $\boldsymbol{\beta} \sim N_p(\boldsymbol{\eta}, \boldsymbol{\Sigma})$ or equivalently $\boldsymbol{\gamma} \sim N_q(\boldsymbol{\mu}, \mathbf{I}_q)$ can be derived from Propositions 1 and 4.

The average evolvability $\bar{e}$ is

$$\begin{aligned}
\bar{e} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\frac{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}}\right) = \mathrm{E}_{\boldsymbol{\gamma}}\left(\frac{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}\mathbf{K}\boldsymbol{\gamma}}{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{K}\boldsymbol{\gamma}}\right) \\
=& \alpha e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}} \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\right)_{j+k+1}}{2^k\left(\frac{1}{2}\right)_k\left(\frac{q}{2}+k\right)_{j+1}k!} C_{[j+k+1]}^{[1],[j],[k]}\left(\mathbf{K}^T\mathbf{G}\mathbf{K}, \mathbf{I}_q - \alpha\mathbf{K}^T\mathbf{K}, \boldsymbol{\mu}\boldsymbol{\mu}^T\right) \\
=& \alpha \sum_{j=0}^{\infty} \frac{(1)_j}{\left(\frac{q}{2}\right)_{j+1}} \tilde{h}_{1;j}\left(\mathbf{K}^T\mathbf{G}\mathbf{K}; \mathbf{I}_q - \alpha\mathbf{K}^T\mathbf{K}\right),
\end{aligned} \tag{55}$$

where α is a positive constant that satisfies $0 < \alpha < 2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{K}\right)$.

As discussed in the text (2.4), the average conditional evolvability $\bar{c}$ and average autonomy $\bar{a}$ can be nonzero only if the distribution of $\boldsymbol{\beta}$ is strictly within $R(\mathbf{G})$. This happens (i) when $\mathbf{G}$ is nonsingular; or (ii) when $\mathbf{G}$ is singular, $\boldsymbol{\eta} \in R(\mathbf{G})$ and $R(\boldsymbol{\Sigma}) \subseteq R(\mathbf{G})$.

Under these conditions, the following expression can be obtained for the average conditional evolvability $\bar{c}$:

$$\begin{aligned}
\bar{c} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\frac{\boldsymbol{\beta}^T\boldsymbol{\beta}}{\boldsymbol{\beta}^T\mathbf{G}^-\boldsymbol{\beta}}\right) = \mathrm{E}_{\boldsymbol{\gamma}}\left(\frac{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{K}\boldsymbol{\gamma}}{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}^-\mathbf{K}\boldsymbol{\gamma}}\right) \\
=& \alpha e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}} \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\right)_{j+k+1}}{2^k\left(\frac{1}{2}\right)_k\left(\frac{q}{2}+k\right)_{j+1}k!} C_{[j+k+1]}^{[1],[j],[k]}\left(\mathbf{K}^T\mathbf{K}, \mathbf{I}_q - \alpha\mathbf{K}^T\mathbf{G}^-\mathbf{K}, \boldsymbol{\mu}\boldsymbol{\mu}^T\right) \\
=& \alpha \sum_{j=0}^{\infty} \frac{(1)_j}{\left(\frac{q}{2}\right)_{j+1}} \tilde{h}_{1;j}\left(\mathbf{K}^T\mathbf{K}; \mathbf{I}_q - \alpha\mathbf{K}^T\mathbf{G}^-\mathbf{K}\right),
\end{aligned} \tag{56}$$

where α is to satisfy $0 < \alpha < 2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{G}^-\mathbf{K}\right)$.

Under the same conditions, the average autonomy $\bar{a}$ is

$$
\begin{aligned}
\bar{a} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\frac{(\boldsymbol{\beta}^T\boldsymbol{\beta})^2}{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{G}^{-}\boldsymbol{\beta}}\right) = \mathrm{E}_{\boldsymbol{\gamma}}\left(\frac{(\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{K}\boldsymbol{\gamma})^2}{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}\mathbf{K}\boldsymbol{\gamma}\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}^{-}\mathbf{K}\boldsymbol{\gamma}}\right) \\
=& \alpha_1\alpha_2 e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\right)_{i+j+k+2}}{2^k\left(\frac{1}{2}\right)_k\left(\frac{q}{2}+k\right)_{i+j+2}k!} \\
& \cdot C_{[i+j+k+2]}^{[2],[i],[j],[k]}\left(\mathbf{K}^T\mathbf{K}, \mathbf{I}_q - \alpha_1\mathbf{K}^T\mathbf{G}\mathbf{K}, \mathbf{I}_q - \alpha_2\mathbf{K}^T\mathbf{G}^{-}\mathbf{K}, \boldsymbol{\mu}\boldsymbol{\mu}^T\right) \\
=& 2\alpha_1\alpha_2 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(1)_i\,(1)_j}{\left(\frac{q}{2}\right)_{i+j+2}} \tilde{h}_{2;i,j}\left(\mathbf{K}^T\mathbf{K}; \mathbf{I}_q - \alpha_1\mathbf{K}^T\mathbf{G}\mathbf{K}, \mathbf{I}_q - \alpha_2\mathbf{K}^T\mathbf{G}^{-}\mathbf{K}\right),
\end{aligned} \tag{57}
$$

where α_1 and α_2 are to satisfy $0 < \alpha_1 < 2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{G}\mathbf{K}\right)$ and $0 < \alpha_2 < 2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{G}^{-}\mathbf{K}\right)$. As before, the average integration $\bar{i}$ is $1 - \bar{a}$.

The average respondability $\bar{r}$ is

$$
\begin{aligned}
\bar{r} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\sqrt{\frac{\boldsymbol{\beta}^T\mathbf{G}^2\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}}}\right) = \mathrm{E}_{\boldsymbol{\gamma}}\left(\sqrt{\frac{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}^2\mathbf{K}^T\boldsymbol{\gamma}}{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{K}\boldsymbol{\gamma}}}\right) \\
=& \alpha_1^{-\frac{1}{2}}\alpha_2^{\frac{1}{2}} e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\left(-\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_j\left(\frac{1}{2}\right)_{i+j+k}}{2^k\left(\frac{1}{2}\right)_k\left(\frac{q}{2}+k\right)_{i+j}i!j!k!} \\
& \cdot C_{[i+j+k]}^{[i],[j],[k]}\left(\mathbf{I}_q - \alpha_1\mathbf{K}^T\mathbf{G}^2\mathbf{K}, \mathbf{I}_q - \alpha_2\mathbf{K}^T\mathbf{K}, \boldsymbol{\mu}\boldsymbol{\mu}^T\right) \\
=& \alpha_1^{-\frac{1}{2}}\alpha_2^{\frac{1}{2}} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\left(-\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_j}{\left(\frac{q}{2}\right)_{i+j}} h_{i,j}\left(\mathbf{I}_q - \alpha_1\mathbf{K}^T\mathbf{G}^2\mathbf{K}, \mathbf{I}_q - \alpha_2\mathbf{K}^T\mathbf{K}\right),
\end{aligned} \tag{58}
$$

where α_1 and α_2 are to satisfy $0 < \alpha_1 < 2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{G}^2\mathbf{K}\right)$ and $0 < \alpha_2 < 2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{K}\right)$.

The average flexibility $\bar{f}$ is

$$
\begin{aligned}
\bar{f} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\frac{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^T\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{G}^2\boldsymbol{\beta}}}\right) = \mathrm{E}_{\boldsymbol{\gamma}}\left(\frac{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}\mathbf{K}\boldsymbol{\gamma}}{\sqrt{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{K}\boldsymbol{\gamma}\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}^2\mathbf{K}\boldsymbol{\gamma}}}\right) \\
=& (\alpha_1\alpha_2)^{\frac{1}{2}} e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_j\left(\frac{1}{2}\right)_{i+j+k+1}}{2^k\left(\frac{1}{2}\right)_k\left(\frac{q}{2}+k\right)_{i+j+1}i!j!k!} \\
& \cdot C_{[i+j+k+1]}^{[1],[i],[j],[k]}\left(\mathbf{K}^T\mathbf{G}\mathbf{K}, \mathbf{I}_q - \alpha_1\mathbf{K}^T\mathbf{K}, \mathbf{I}_q - \alpha_2\mathbf{K}^T\mathbf{G}^2\mathbf{K}, \boldsymbol{\mu}\boldsymbol{\mu}^T\right) \\
=& (\alpha_1\alpha_2)^{\frac{1}{2}} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\left(\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_j}{\left(\frac{q}{2}\right)_{i+j+1}} \tilde{h}_{1;i,j}\left(\mathbf{K}^T\mathbf{G}\mathbf{K}; \mathbf{I}_q - \alpha_1\mathbf{K}^T\mathbf{K}, \mathbf{I}_q - \alpha_2\mathbf{K}^T\mathbf{G}^2\mathbf{K}\right),
\end{aligned} \tag{59}
$$

where α_1 and α_2 are to satisfy $0 < \alpha_1 < 2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{K}\right)$ and $0 < \alpha_2 < 2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{G}^2\mathbf{K}\right)$.

The average response difference $\bar{d}$ is

$$
\begin{aligned}
\bar{d} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\sqrt{\frac{\boldsymbol{\beta}^T\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}}}\right) = \mathrm{E}_{\boldsymbol{\gamma}}\left(\sqrt{\frac{\boldsymbol{\gamma}^T\mathbf{K}^T\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\mathbf{K}\boldsymbol{\gamma}}{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{K}\boldsymbol{\gamma}}}\right) \\
=& \alpha_1^{-\frac{1}{2}}\alpha_2^{\frac{1}{2}}e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{\left(-\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_j\left(\frac{1}{2}\right)_{i+j+k}}{2^k\left(\frac{1}{2}\right)_k\left(\frac{q}{2}+k\right)_{i+j}i!j!k!} \\
& \cdot C_{[i+j+k]}^{[i],[j],[k]}\left(\mathbf{I}_q-\alpha_1\mathbf{K}^T\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\mathbf{K},\mathbf{I}_q-\alpha_2\mathbf{K}^T\mathbf{K},\boldsymbol{\mu}\boldsymbol{\mu}^T\right) \\
=& \alpha_1^{-\frac{1}{2}}\alpha_2^{\frac{1}{2}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(-\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_j}{\left(\frac{q}{2}\right)_{i+j}}h_{i,j}\left(\mathbf{I}_q-\alpha_1\mathbf{K}^T\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\mathbf{K},\mathbf{I}_q-\alpha_2\mathbf{K}^T\mathbf{K}\right),
\end{aligned}
\tag{60}
$$

where α_1 and α_2 are to satisfy $0<\alpha_1<2/\lambda_{\max}\left(\mathbf{K}^T\left(\mathbf{G}_1-\mathbf{G}_2\right)^2\mathbf{K}\right)$ and $0<\alpha_2<2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{K}\right)$.

The average response correlation $\bar{\rho}$ is

$$
\begin{aligned}
\bar{\rho} =& \mathrm{E}_{\boldsymbol{\beta}}\left(\frac{\boldsymbol{\beta}^T\mathbf{G}_1\mathbf{G}_2\boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^T\mathbf{G}_1^2\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{G}_2^2\boldsymbol{\beta}}}\right) = \mathrm{E}_{\boldsymbol{\gamma}}\left(\frac{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}_1\mathbf{G}_2\mathbf{K}\boldsymbol{\gamma}}{\sqrt{\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}_1^2\mathbf{K}\boldsymbol{\gamma}\boldsymbol{\gamma}^T\mathbf{K}^T\mathbf{G}_2^2\mathbf{K}\boldsymbol{\gamma}}}\right) \\
=& (\alpha_1\alpha_2)^{\frac{1}{2}}e^{-\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_j\left(\frac{1}{2}\right)_{i+j+k+1}}{2^k\left(\frac{1}{2}\right)_k\left(\frac{q}{2}+k\right)_{i+j+1}i!j!k!} \\
& \cdot C_{[i+j+k+1]}^{[1],[i],[j],[k]}\left(\mathbf{K}^T\frac{\left(\mathbf{G}_1\mathbf{G}_2+\mathbf{G}_2\mathbf{G}_1\right)}{2}\mathbf{K},\mathbf{I}_q-\alpha_1\mathbf{K}^T\mathbf{G}_1^2\mathbf{K},\mathbf{I}_q-\alpha_2\mathbf{K}^T\mathbf{G}_2^2\mathbf{K},\boldsymbol{\mu}\boldsymbol{\mu}^T\right) \\
=& (\alpha_1\alpha_2)^{\frac{1}{2}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{1}{2}\right)_i\left(\frac{1}{2}\right)_j}{\left(\frac{q}{2}\right)_{i+j+1}}\tilde{h}_{1;i,j}\left(\mathbf{K}^T\frac{\left(\mathbf{G}_1\mathbf{G}_2+\mathbf{G}_2\mathbf{G}_1\right)}{2}\mathbf{K};\mathbf{I}_q-\alpha_1\mathbf{K}^T\mathbf{G}_1^2\mathbf{K},\mathbf{I}_q-\alpha_2\mathbf{K}^T\mathbf{G}_2^2\mathbf{K}\right)
\end{aligned}
\tag{61}
$$

where α_1 and α_2 are to satisfy $0<\alpha_1<2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{G}_1^2\mathbf{K}\right)$ and $0<\alpha_2<2/\lambda_{\max}\left(\mathbf{K}^T\mathbf{G}_2^2\mathbf{K}\right)$.

C.3 Practical evaluation and truncation error

The above series expressions for average measures in general cases can be evaluated by one of the recursive algorithms described by Hillier *et al.* (2009, 2014) and Bao & Kan (2013). When the mean $\boldsymbol{\mu}$ is nonzero, the series can be most efficiently evaluated by recursions for $h_{i,j}$ and $\tilde{h}_{i,j}$ described in Hillier *et al.* (2014, section 5.4) and Bao & Kan (2013, section 5). When possible, it is recommended to inspect a profile of the partial sum across varying orders as the signs of the terms can in general fluctuate in those series.

At present, truncation error bounds for general cases are available only for $\bar{e}$ and $\bar{c}$. From the theorem 7 of Hillier *et al.* (2014), a truncation error bound for these average measures evaluated up to $k=M$ is obtained from the following, by inserting $(\mathbf{A},\mathbf{B})=(\mathbf{K}^T\mathbf{G}\mathbf{K},\mathbf{K}^T\mathbf{K})$ and $(\mathbf{K}^T\mathbf{K},\mathbf{K}^T\mathbf{G}^-\mathbf{K})$ for $\bar{e}$ and $\bar{c}$, respectively:

$$
\alpha^n\frac{(n)_{M+1}}{\left(\frac{q}{2}\right)_{M+2}}\left(|\alpha\mathbf{B}|^{-\frac{1}{2}}\left[\frac{\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}\right)+t\alpha^{-1}\boldsymbol{\mu}^T\mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}\boldsymbol{\mu}}{2}\right]\exp\left(\frac{t\alpha^{-1}\boldsymbol{\mu}^T\mathbf{B}^{-1}\boldsymbol{\mu}-\boldsymbol{\mu}^T\boldsymbol{\mu}}{2}\right)-S_M\right).
\tag{62}
$$

Here, $n = 1$, $t = 2$, and $S_M = \sum_{j=0}^{M} \hat{h}_{1;j}(\mathbf{A}; \mathbf{I}_q - \alpha\mathbf{B})$, with $\hat{h}_{1;j}(\mathbf{A}_1; \mathbf{A}_2)$ being the coefficient of $t_1 t_2^j$ in the expansion

$$|\mathbf{I}_q - t_1\mathbf{A}_1 - t_2\mathbf{A}_2| \exp\left(\frac{(1+t_2)\boldsymbol{\mu}^T \left(\mathbf{I}_q - t_1\mathbf{A}_1 - t_2\mathbf{A}_2\right)^{-1} \boldsymbol{\mu} - \boldsymbol{\mu}^T\boldsymbol{\mu}}{2}\right) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \hat{h}_{i;j}(\mathbf{A}_1; \mathbf{A}_2)\, t_1^i t_2^j,$$

which can be evaluated by a recursive algorithm similar to the one presented in Hillier *et al.* (2014, theorem 6).[6] This expression bounds the truncation error in absolute values unlike those in (23)–(26) which are one-sided. This is because $\tilde{h}_{i,j}$ can take both positive and negative values when $\boldsymbol{\mu}$ is nonzero, so that the partial sums are not always increasing.

Under the restriction $\mathbf{K}^T\mathbf{K} = \mathbf{I}_q$, a similar error bound can be derived for $\bar{f}$, by letting $\mathbf{A} = \mathbf{K}^T\mathbf{G}\mathbf{K}$, $\mathbf{B} = \mathbf{K}^T\mathbf{G}^2\mathbf{K}$, $n = 1/2$, $t = 3$, and $S_M = \sum_{k=0}^{M}\sum_{l=0}^{k} \hat{h}_{1;l,k-l}(\mathbf{A}; \mathbf{I}_q - \alpha\mathbf{B}, \mathbf{0}_{q\times q})$ in (62). Here, $\hat{h}_{1;j,k}(\mathbf{A}_1; \mathbf{A}_2, \mathbf{A}_3)$ is the coefficient of $t_1 t_2^j t_3^k$ in the expansion of

$$|\mathbf{I}_q - t_1\mathbf{A}_1 - t_2\mathbf{A}_2 - t_3\mathbf{A}_3| \exp\left(\frac{(1+t_2+t_3)\boldsymbol{\mu}^T \left(\mathbf{I}_q - t_1\mathbf{A}_1 - t_2\mathbf{A}_2 - t_3\mathbf{A}_3\right)^{-1} \boldsymbol{\mu} - \boldsymbol{\mu}^T\boldsymbol{\mu}}{2}\right)$$
$$= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \hat{h}_{i;j,k}(\mathbf{A}_1; \mathbf{A}_2, \mathbf{A}_3)\, t_1^i t_2^j t_3^k,$$

which can be evaluated by the same recursive algorithm for $\hat{h}_{i;j}$ with obvious modifications.

Appendix D Connections to previous results

D.1 Average conditional evolvability in two traits

Hansen & Houle (2008) stated that the average conditional evolvability, under the uniform distribution of $\mathbf{u}$ on S^{p-1}, equals to the geometric mean of the eigenvalues of $\mathbf{G}$ when $p = 2$. It is possible to confirm that the present result is consistent with their argument as follows. From (17),

$$\begin{aligned}
\bar{c} =& \alpha\, {}_2F_1\left(\frac{1}{2}, 1; \frac{2}{2}; \mathbf{I}_p - \alpha\mathbf{G}^{-1}\right) \\
=& \alpha\, {}_1F_0\left(\frac{1}{2}; \mathbf{I}_p - \alpha\mathbf{G}^{-1}\right) \\
=& \alpha \left|\mathbf{I}_p - \left(\mathbf{I}_p - \alpha\mathbf{G}^{-1}\right)\right|^{-\frac{1}{2}} \\
=& \left|\mathbf{G}\right|^{\frac{1}{2}},
\end{aligned}$$

where in the second equation the denominator parameter cancels with one of the numerator parameters, and the third equation is from (36). Since the determinant is equal to the product of the eigenvalues, this proves the argument.

[6] The recursive algorithm for evaluating $\hat{h}_{i;j}$ was said to be "very similar" to that for $h_{i;j}$ but was not explicitly stated in Hillier *et al.* (2014). This is obtained by, in the definition and updating equation for $g_{i,j}$ in the theorem 6 therein, replacing all the subtractions of terms (but not indices) by additions.

D.2 Average respondability under complete integration

Kirkpatrick (2009) considered the average selection response $\bar{R}$, which equals $\bar{r}/\lambda_{\max}(\mathbf{G})$ here when $\mathbf{u}$ is uniformly distributed on S^{p-1}. This section shows that the present expression for $\bar{r}$ (19) entails his result for $\bar{R}$ under the case of complete integration, $\mathbf{G} = \operatorname{diag}(\lambda_{\max}, 0, \ldots, 0)$, the only situation for which an analytic expression was given. The expression is (Kirkpatrick, 2009, (6) and (A7) therein)

$$\begin{aligned}\bar{R} =& \frac{2\Gamma\left(\frac{p}{2}\right)}{(p-1)\sqrt{\pi}\Gamma\left(\frac{p-1}{2}\right)} \\ =& \frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p+1}{2}\right)}.\end{aligned} \tag{63}$$

Taking $\alpha = \lambda_{\max}^{-2}$, (19) becomes

$$\begin{aligned}\bar{r} =& \lambda_{\max} \sum_{k=0}^{\infty} \frac{\left(-\frac{1}{2}\right)_k \left(\frac{1}{2}\right)_k}{\left(\frac{p}{2}\right)_k k!} C_{[k]}(\mathbf{I}_p - \operatorname{diag}(1, 0, \ldots, 0)) \\ =& \lambda_{\max} \sum_{k=0}^{\infty} \frac{\left(-\frac{1}{2}\right)_k \left(\frac{1}{2}\right)_k}{\left(\frac{p}{2}\right)_k k!} \frac{\left(\frac{p-1}{2}\right)_k}{\left(\frac{1}{2}\right)_k} \\ =& \lambda_{\max} \sum_{k=0}^{\infty} \frac{\left(-\frac{1}{2}\right)_k \left(\frac{p-1}{2}\right)_k}{\left(\frac{p}{2}\right)_k k!} \\ =& \lambda_{\max} \frac{\Gamma\left(\frac{p}{2}\right)\Gamma(1)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p+1}{2}\right)},\end{aligned}$$

where the second equation comes from the fact $C_{[k]}(\operatorname{diag}(0, 1, \ldots, 1)) = C_{[k]}(\mathbf{I}_{p-1})$ and (32), and the last equation is from the Gauss hypergeometric theorem: ${}_2F_1(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c-b-a)}{\Gamma(c-b)\Gamma(c-a)}$. The equivalence to (63) is obvious.

References

Agrawal, A. F., Stinchcombe, J. R., 2009. How much do genetic covariances alter the rate of adaptation? *Proceedings of the Royal Society B: Biological Sciences*, **276**: 1183–1191. doi:10.1098/rspb.2008.1671.

Aguirre, J. D., Hine, E., McGuigan, K., Blows, M. W., 2014. Comparing **G**: multivariate analysis of genetic variation in multiple populations. *Heredity*, **112**: 21–29. doi:10.1038/hdy.2013.12.

Armbruster, W. S., Hansen, T. F., Pélabon, C., Pérez-Barrales, R., Maad, J., 2009. The adaptive accuracy of flowers: measurement and microevolutionary patterns. *Annals of Botany*, **103**: 1529–1545. doi:10.1093/aob/mcp095.

Bao, Y., Kan, R., 2013. On the moments of ratios of quadratic forms in normal random variables. *Journal of Multivariate Analysis*, **117**: 229–245. doi:10.1016/j.jmva.2013.03.002.

Bates, D., Eddelbuettel, D., 2013. Fast and elegant numerical linear algebra using the RcppEigen package. *Journal of Statistical Software*, **52**(5): 1–24. doi:10.18637/jss.v052.i05.

Blows, M. W., **2007**. A tale of two matrices: multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology*, **20**: 1–8. doi:10.1111/j.1420-9101.2006.01164.x.

Bolstad, G. H., **Hansen, T. F.**, **Pélabon, C.**, **Falahati-Anbaran, M.**, **Pérez-Baralles, R.**, **Armbruster, W. S.**, **2014**. Genetic constraints predict evolutionary divergence in *Dalechampia* blossoms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **369**: 20130255. doi:10.1098/rstb.2013.0255.

Brommer, J. E., **2014**. Using average autonomy to test whether behavioral syndromes constrain evolution. *Behavioral Ecology and Sociobiology*, **68**: 691–700. doi:10.1007/s00265-014-1699-6.

Brown, R. L., **2014**. What evolvability really is. *British Journal for the Philosophy of Science*, **65**: 549–572. doi:10.1093/bjps/axt014.

Cheverud, J. M., **1982**. Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution*, **36**: 499–516. doi:10.1111/j.1558-5646.1982.tb05070.x.

Cheverud, J. M., **1989**. A comparative analysis of morphological variation patterns in the papionins. *Evolution*, **43**: 1737–1747. doi:10.1111/j.1558-5646.1989.tb02623.x.

Cheverud, J. M., **1996**. Quantitative genetic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S. fuscicollis*) tamarins. *Journal of Evolutionary Biology*, **9**: 5–42. doi:10.1046/j.1420-9101.1996.9010005.x.

Cheverud, J. M., **Marroig, G.**, **2007**. Comparing covariance matrices: random skewers method compared to the common principal components model. *Genetics and Molecular Biology*, **30**: 461–469. doi:10.1590/S1415-47572007000300027.

Chevin, L.-M., **2013**. Genetic constraints on adaptation to a changing environment. *Evolution*, **67**: 708–721. doi:10.1111/j.1558-5646.2012.01809.x.

Chikuse, Y., **1980**. Invariant polynomials with matrix arguments and their applications. *In:* Gupta, R. P. (ed.), *Multivariate Statistical Analysis*. North-Holland, Amsterdam, pp. 53–68.

Chikuse, Y., **1987**. Methods for constructing top order invariant polynomials. *Econometric Theory*, **3**: 195–207. doi:10.1017/S026646660001029X.

Chikuse, Y., **2003**. *Statistics on Special Manifolds*. Springer Science+Business Media, New York, New York.

Chikuse, Y., **Davis, A. W.**, **1986**a. Some properties of invariant polynomials with matrix arguments and their applications in econometrics. *Annals of the Institute of Statistical Mathematics*, **38**: 109–122. doi:10.1007/BF02482504.

Chikuse, Y., **Davis, A. W.**, **1986**b. A survey on the invariant polynomials with matrix arguments in relation to econometric distribution theory. *Econometric Theory*, **2**: 232–248. doi:10.1017/S0266466600011518.

Constantine, A. G., **1963**. Some non-central distribution problems in multivariate analysis. *Annals of Mathematical Statistics*, **34**: 1270–1285. doi:10.1214/aoms/1177703863.

Cressie, N., **Davis, A. S.**, **Folks, J. L.**, **Policello, G. E., II**, **1981**. The moment-generating function and negative integer moments. *American Statistician*, **35**: 148–150. doi:10.2307/2683982.

Crother, B. I., **Murray, C. M.**, **2019**. Early usage and meaning of evolvability. *Ecology and Evolution*, **9**: 3784–3793. doi:10.1002/ece3.5002.

Davies, P. I., **Higham, N. J.**, **2000**. Numerically stable generation of correlation matrices and their factors. *BIT Numerical Mathematics*, **40**: 640–651. doi:10.1023/A:1022384216930.

Davis, A. W., **1979**. Invariant polynomials with two matrix arguments extending the zonal polynomials: applications to multivariate distribution theory. *Annals of the Institute of Statistical Mathematics*, **31**: 465–485. doi:10.1007/BF02480302.

Davis, A. W., **1980**. Invariant polynomials with two matrix arguments, extending the zonal polynomials. *In:* Krishnaiah, P. R. (ed.), *Multivariate Analysis—V*. North-Holland, Amsterdam, pp. 287–299.

Feiner, N., **Jackson, I. S. C.**, **Van der Cruyssen, E.**, **Uller, T.**, **2021**. A highly conserved ontogenetic limb allometry and its evolutionary significance in the adaptive radiation of *Anolis* lizards. *Proceedings of the Royal Society B: Biological Sciences*, **288**: 20210226. doi:10.1098/rspb.2021.0226.

Forchini, G., **2002**. The exact cumulative distribution function of a ratio of quadratic forms in normal variables, with application to the AR(1) model. *Econometric Theory*, **18**: 823–852. doi:10.1017/s0266466602184015.

Forchini, G., **2005**. The distribution of a ratio of quadratic forms in noncentral normal variables. *Communications in Statistics—Theory and Methods*, **34**: 999–1008. doi:10.1081/STA-200056855.

Grabowski, M., **Porto, A.**, **2017**. How many more? Sample size determination in studies of morphological integration and evolvability. *Methods in Ecology and Evolution*, **8**: 592–603. doi:10.1111/2041-210X.12674.

Gross, K. I., **Richards, D. S. P.**, **1987**. Special functions of matrix argument. I: algebraic induction, zonal polynomials, and hypergeometric functions. *Transactions of the American Mathematical Society*, **301**: 781–811. doi:10.1090/S0002-9947-1987-0882715-2.

Gupta, A. K., **Kabe, D. G.**, **1998**. Moments of ratios of quadratic forms. *Statistics and Probability Letters*, **38**: 69–71. doi:10.1016/S0167-7152(97)00155-7.

Haber, A., **2011**. A comparative analysis of integration indices. *Evolutionary Biology*, **38**: 476–488. doi:10.1007/s11692-011-9137-4.

Haber, A., **2016**. Phenotypic covariation and morphological diversification in the ruminant skull. *American Naturalist*, **187**: 576–591. doi:10.1086/685811.

Hansen, T. F., **2003**. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. *BioSystems*, **69**: 83–94. doi:10.1016/S0303-2647(02)00132-6.

Hansen, T. F., **Armbruster, W. S.**, **Carlson, M. L.**, **Pélabon, C.**, **2003**a. Evolvability and genetic constraint in *Dalechampia* blossoms: genetic correlations and conditional evolvability. *Journal of Experimental Zoology, Part B: Molecular Development and Evolution*, **296B**: 23–39. doi:10.1002/jez.b.14.

Hansen, T. F., **Houle, D.**, **2008**. Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology*, **21**: 1201–1219. doi:10.1111/j.1420-9101.2008.01573.x.

Hansen, T. F., **Houle, D.**, **2009**. Corrigendum [of "Measuring and comparing evolvability and constraint in multivariate characters"]. *Journal of Evolutionary Biology*, **22**: 913–915. doi:10.1111/j.1420-9101.2009.01715.x.

Hansen, T. F., **Pélabon, C.**, **2021**. Evolvability: a quantitative-genetics perspective. *Annual Review of Ecology, Evolution, and Systematics*, **52**: 153–175. doi:10.1146/annurev-ecolsys-011121-021241.

Hansen, T. F., **Pélabon, C.**, **Armbruster, W. S.**, **Carlson, M. L.**, **2003**b. Evolvability and genetic constraint in *Dalechampia* blossoms: components of variance and measures of evolvability. *Journal of Evolutionary Biology*, **16**: 754–766. doi:10.1046/j.1420-9101.2003.00556.x.

Hansen, T. F., **Pélabon, C.**, **Houle, D.**, **2011**. Heritability is not evolvability. *Evolutionary Biology*, **38**: 258–277. doi:10.1007/s11692-011-9127-6.

Hansen, T. F., **Solvin, T. M.**, **Pavlicev, M.**, **2019**. Predicting evolutionary potential: a numerical test of evolvability measures. *Evolution*, **73**: 689–703. doi:10.1111/evo.13705.

Hansen, T. F., **Voje, K. L.**, **2011**. Deviation from the line of least resistance does not exclude genetic constraints: a comment on Berner et al. (2010). *Evolution*, **65**: 1821–1822. doi:10.1111/j.1558-5646.2011.01281.x.

Hillier, G., **2001**. The density of a quadratic form in a vector uniformly distributed on the n-sphere. *Econometric Theory*, **17**: 1–28. doi:10.1017/S026646660117101X.

Hillier, G., **Kan, R.**, **Wang, X.**, **2009**. Computationally efficient recursions for top-order invariant polynomials with applications. *Econometric Theory*, **25**: 211–242. doi:10.1017/S0266466608090075.

Hillier, G., **Kan, R.**, **Wang, X.**, **2014**. Generating functions and short recursions, with applications to the moments of quadratic forms in noncentral normal vectors. *Econometric Theory*, **30**: 436–473. doi:10.1017/S0266466613000364.

Hine, E., **Blows, M. W.**, **2006**. Determining the effective dimensionality of the genetic variance–covariance matrix. *Genetics*, **173**: 1135–1144. doi:10.1534/genetics.105.054627.

Houle, D., **1992**. Comparing evolvability and variability of quantitative traits. *Genetics*, **130**: 195–204. doi:10.1093/genetics/130.1.195.

Houle, D., **Meyer, K.**, **2015**. Estimating sampling error of evolutionary statistics based on genetic covariance matrices using maximum likelihood. *Journal of Evolutionary Biology*, **28**: 1542–1549. doi:10.1111/jeb.12674.

Jablonski, D., **2022**. Evolvability and macroevolution: overview and synthesis. *Evolutionary Biology*, **49**: 265–291. doi:10.1007/s11692-022-09570-4.

James, A. T., **1960**. The distribution of the latent roots of the covariance matrix. *Annals of Mathematical Statistics*, **31**: 151–158. doi:10.1214/aoms/1177705994.

James, A. T., **1961**. Zonal polynomials of the real positive definite symmetric matrices. *Annals of Mathematics*, **74**: 456–469. doi:10.2307/1970291.

James, A. T., **1964**. Distributions of matrix variates and latent roots derived from normal samples. *Annals of Mathematical Statistics*, **35**: 475–501. doi:10.1214/aoms/1177703550.

Jones, M. C., **1986**. Expressions for inverse moments of positive quadratic forms in normal variables. *Australian Journal of Statistics*, **28**: 242–250. doi:10.1111/j.1467-842X.1986.tb00604.x.

Jones, M. C., **1987**. On moments of ratios of quadratic forms in normal variables. *Statistics and Probability Letters*, **6**: 129–136. doi:10.1016/0167-7152(87)90086-1.

Kirkpatrick, M., **2009**. Patterns of quantitative genetic variation in multiple dimensions. *Genetica*, **136**: 271–284. doi:10.1007/s10709-008-9302-6.

Koev, P., **Edelman, A.**, **2006**. The efficient evaluation of the hypergeometric function of a matrix argument. *Mathematics of Computation*, **75**: 833–846. doi:10.1090/S0025-5718-06-01824-2.

Lande, R., **1979**. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution*, **33**: 402–416. doi:10.1111/j.1558-5646.1979.tb04694.x.

Lande, R., **Arnold, S. J.**, **1983**. The measurement of selection on correlated characters. *Evolution*, **37**: 1210–1226. doi:10.1111/j.1558-5646.1983.tb00236.x.

Laurent, S., **2022**. HypergeoMat: hypergeometric function of a matrix argument. URL: `https://CRAN.R-project.org/package=HypergeoMat`.

Love, A. C., **Grabowski, M.**, **Houle, D.**, **Liow, L. H.**, **Porto, A.**, **Tsuboi, M.**, **Voje, K. L.**, **Hunt, G.**, **2022**. Evolvability in the fossil record. *Paleobiology*, **48**: 186–209. doi:10.1017/pab.2021.36.

Machado, F. A., **2020**. Selection and constraints in the ecomorphological adaptive evolution of the skull of living Canidae (Carnivora, Mammalia). *American Naturalist*, **196**: 197–215. doi:10.1086/709610.

Machado, F. A., **Zahn, T. M. G.**, **Marroig, G.**, **2018**. Evolution of morphological integration in the skull of Carnivora (Mammalia): changes in Canidae lead to increased evolutionary potential of facial traits. *Evolution*, **72**: 1399–1419. doi:10.1111/evo.13495.

Magnus, J. R., **1986**. The exact moments of a ratio of quadratic forms in normal variables. *Annales d'Économie et de Statistique*, **4**: 95–109. doi:10.2307/20075629.

Mardia, K. V., **Jupp, P. E.**, **1999**. *Directional Statistics*. John Wiley & Sons, Chichester, UK. doi:10.1002/9780470316979.

Marroig, G., **Shirai, L. T.**, **Porto, A.**, **de Oliveira, F. B.**, **De Conto, V.**, **2009**. The evolution of modularity in the mammalian skull II: evolutionary consequences. *Evolutionary Biology*, **36**: 136–148. doi:10.1007/s11692-009-9051-1.

Mathai, A. M., **Provost, S. B.**, **1992**. *Quadratic Forms in Random Variables: Theory and Applications*. Marcel Dekker, New York, New York.

Mathai, A. M., **Provost, S. B.**, **Hayakawa, T.**, **1995**. *Bilinear Forms and Zonal Polynomials*. Springer-Verlag, New York, New York.

Melo, D., **Garcia, G.**, **Hubbe, A.**, **Assis, A. P.**, **Marroig, G.**, **2016**. EvolQG—an R package for evolutionary quantitative genetics [version 3; referees: 2 approved, 1 approved with reservations]. *F1000Research*, **4**: 925. doi:10.12688/f1000research.7082.3.

Meng, X.-L., **2005**. From unit root to Stein's estimator to Fisher's k statistics: if you have a moment, I can tell you more. *Statistical Science*, **20**: 414–162. doi:10.1214/088342304000000279.

Meyer, K., **Kirkpatrick, M.**, **2008**. Perils of parsimony: properties of reduced-rank estimates of genetic covariance matrices. *Genetics*, **180**: 1153–1166. doi:10.1534/genetics.108.090159.

Mezey, J. G., **Houle, D.**, **2005**. The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. *Evolution*, **59**: 1027–1038. doi:10.1111/j.0014-3820.2005.tb01041.x.

Muirhead, R. J., **1982**. *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, Hoboken, New Jersey. doi:10.1002/9780470316559.

Muirhead, R. J., **2006**. Zonal polynomials. *In:* Kotz, S., Balakrishnan, N., Read, C., Vidakovic, B., Johnson, N. L. (eds.), *Encyclopedia of Statistical Sciences*, volume 15, 2nd edition. John Wiley & Sons, Hoboken, New Jersey, pp. 9225–9231.

Pavlicev, M., **Hansen, T. F.**, **2011**. Genotype-phenotype maps maximizing evolvability: modularity revisited. *Evolutionary Biology*, **38**: 371–389. doi:10.1007/s11692-011-9136-5.

Pavlicev, M., **Wagner, G. P.**, **Cheverud, J. M.**, **2009**. Measuring evolutionary constraints through the dimensionality of the phenotype: adjusted bootstrap method to estimate rank of phenotypic covariance matrix. *Evolutionary Biology*, **36**: 339–353. doi:10.1007/s11692-009-9066-7.

Porto, A., **de Oliveira, F. B.**, **Shirai, L. T.**, **De Conto, V.**, **Marroig, G.**, **2009**. The evolution of modularity in the mammalian skull I: morphological integration patterns and magnitudes. *Evolutionary Biology*, **36**: 118–135. doi:10.1007/s11692-008-9038-3.

R Core Team, **2022**. R: a language and environment for statistical computing, version 4.2.1. URL: `https://www.R-project.org/`.

Revell, L. J., **2007**. The **G** matrix under fluctuating correlational mutation and selection. *Evolution*, **61**: 1857–1872. doi:10.1111/j.1558-5646.2007.00161.x.

Riederer, J. M., **Tiso, S.**, **van Eldijk, T. J. B.**, **Weissing, F. J.**, **2022**. Capturing the facets of evolvability in a mechanistic framework. *Trends in Ecology and Evolution*, **37**: 430–439. doi:10.1016/j.tree.2022.01.004.

Roberts, L. A., **1995**. On the existence of moments of ratios of quadratic forms. *Econometric Theory*, **11**: 750–774. doi:10.1017/S0266466600009725.

Rohlf, F. J., **2017**. The method of random skewers. *Evolutionary Biology*, **44**: 542–550. doi:10.1007/s11692-017-9425-8.

Rolian, C., **2009**. Integration and evolvability in primate hands and feet. *Evolutionary Biology*, **36**: 100–117. doi:10.1007/s11692-009-9049-8.

Schott, J. R., **2016**. *Matrix Analysis for Statistics*. 3rd edition. John Wiley & Sons, Hoboken, New Jersey.

Smith, M. D., **1989**. On the expectation of a ratio of quadratic forms in normal variables. *Journal of Multivariate Analysis*, **31**: 244–257. doi:10.1016/0047-259X(89)90065-1.

Smith, M. D., **1993**. Expectations of ratios of quadratic forms in normal variables: evaluating some top-order invariant polynomials. *Australian Journal of Statistics*, **35**: 271–282. doi:10.1111/j.1467-842X.1993.tb01335.x.

Steppan, S. J., **Phillips, P. C.**, **Houle, D.**, **2002**. Comparative quantitative genetics: evolution of the **G** matrix. *Trends in Ecology and Evolution*, **17**: 320–327. doi:10.1016/S0169-5347(02)02505-3.

Takemura, A., **1984**. *Zonal Polynomials*, *Institute of Mathematical Statistics Lecture Notes—Monograph Series*, volume 4. Institute of Mathematical Statistics, Hayward, California.

Walsh, B., **Blows, M. W.**, **2009**. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. *Annual Review of Ecology, Evolution, and Systematics*, **40**: 41–59. doi:10.1146/annurev.ecolsys.110308.120232.

Watanabe, J., **2022**. Statistics of eigenvalue dispersion indices: quantifying the magnitude of phenotypic integration. *Evolution*, **76**: 4–28. doi:10.1111/evo.14382.