Partitioning the phenotypic variance of reaction norms

Pierre de Villemereuil[1] and Luis-Miguel Chevin[2]

[1] *Institut de Systématique, Évolution, Biodiversité (ISYEB), École Pratique des Hautes Études PSL, MNHN, CNRS, SU, UA, Paris, France*

[2] *CEFE, CNRS, Université de Montpellier, Université Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France*

Keywords: phenotypic plasticity, quantitative genetics, character-state approach, polynomial approach, non-linear modelling

Corresponding author: Pierre de Villemereuil, E-mail: pierre.devillemereuil@ephe.psl.eu

Abstract

Many phenotypic traits vary in a predictable way across environments, as captured by their norms of reaction. These reaction norms may be discrete or continuous, and can substantially vary in shape across organisms and traits, making it difficult to compare amounts and types of plasticity among (and sometimes even within) studies. In addition, genetic variation and evolutionary potential in heterogeneous environments critically depends on how reaction norms vary genetically, but there is no consensus on how this should be quantified. Here, we propose a partitioning of phenotypic variance across genotypes and environments that jointly address these challenges. We first derive components of phenotypic variance arising from the average reaction norm across genotypes, genetic variation in reaction norms (including additive genetic variance), and a residual that cannot be predicted by reaction norms. We then further partition the first two terms into contributions from parameters of reaction norm shape, such as the mean and variance of reaction norm slope and curvature. We show how to implement this approach in practice in various contexts, including the character-state approach, polynomial functions, or arbitrary non-linear models. We also show how the combination of character-state and curve-parameter approaches can provide a metric of goodness of fit of a given model of reaction norm shape. Overall the toolbox we develop, summarized in an online tutorial, should serve as a base for more robust comparative studies of plasticity across organisms and traits.

Introduction

The phenotype of a given genotype can vary in response to its environment of development or expression, and such phenotypic plasticity is currently attracting considerable interest in the context of rapidly changing

natural environments (Gienapp et al. 2008; Chevin et al. 2010; Merilä & Hendry 2014). While the mere existence (and even prevalence) of phenotypic plasticity is uncontroversial, its relative contribution to observed or predicted phenotypic change in the wild (Teplitsky et al. 2008; Gienapp et al. 2008; Merilä & Hendry 2014; Bonamour et al. 2019), as well as the extent of its interplay with population-level processes such as natural selection and population dynamics (Reed et al. 2010; Vedder et al. 2013; Schaum & Collins 2014; de Villemereuil et al. 2020), are very active research areas. Answering these questions requires being able to quantify phenotypic plasticity at broad taxonomic, ecological, and phenotypic scales.

The relationship between the phenotype and the environment is captured by the reaction norm (or norm of reaction), which is defined at the level of genotypes (Woltereck 1909; Schlichting & Pigliucci 1998). Reaction norms encompass phenotypic responses to both continuous environments (such as temperature, salinity, etc.) and categorical/discrete ones (such as host plant for a phytophagous insect). Within a simple model of reaction norm, quantifying plasticity may be straightforward. For instance when a linear reaction norm is assumed, the reaction norm slope is generally used as a metric of plasticity in both empirical (Charmantier et al. 2008; Nussey et al. 2005) and theoretical (Gavrilets & Scheiner 1993b; Lande 2009) work, since it quantifies how much phenotypic change is induced per unit environmental change. However, regression slopes are signed and have units of trait per environment, so even in this simple case some standardization is needed in order to compare the magnitude of plasticity among studies. Beyond this simple scenario, drawing robust conclusions about phenotypic plasticity requires being able to quantify and compare its magnitude across organisms, traits and environments, in a way that does not depend on reaction norm shape, and can be applied even when shape cannot be simply defined (for instance because environments have no intrinsic order). Such unified measure of plasticity seems to be currently lacking.

How phenotypes change with the environment can also be of importance, beyond *how much* they change. First, different reaction norm shapes may come with different biological interpretations. For instance, a bell-shaped (eg quadratic, Gaussian) reaction norm may indicate that some mechanism underlying a measured trait is maximized at an intermediate value of the environment. This is often expected for traits that are direct components of fitness, or that can be interpreted as proxys for performance, for which the reaction norms are generally described as tolerance or performance curves (Lynch & Gabriel 1987; Deutsch et al. 2008; Angilletta 2009). A sigmoid shape, on the other hand, may indicate that plasticity is directional but that the range of possible phenotypes is constrained, or that selection favors discrete-like variation (Moczek & Emlen 1999; Suzuki & Nijhout 2006; Hammill et al. 2008; Chevin et al. 2013). Second, most theoretical models on the evolution of plasticity, especially those based on quantitative genetics, which are most directly comparable to data on phenotypic plasticity, assume a given reaction norm shape - often linear for simplicity (Scheiner 1993b; Tufto 2000; Lande 2009). The extent to which theoretical predictions on the evolution of plasticity apply

to any particular empirical system thus depends on how well the reaction norm shape assumed in the models conforms to observations in this system. In other words, we need some metric for whether a reaction norm is "mostly linear" or "mostly curved", for instance. In addition, when fitting a particular model of reaction norm shape to an empirical dataset, we would like to know how well this model captures the overall plastic variation of the trait across environments.

A third crucial question regarding reaction norms is how they vary genetically. It has long been recognized that plasticity can evolve if reaction norms vary genetically (Bradshaw 1965), and theory has predicted how different aspects of reaction norm shape are expected to respond to selection in a variable environment (De Jong 1990; Gomulkiewicz & Kirkpatrick 1992; Gavrilets & Scheiner 1993b). However this theory has been little applied empirically, except for predictions about the slope of linear reaction norms (or equivalently, phenotypic differences between two environments), which directly quantifies the degree of plasticity. But beyond this, it should also be of interest to find out which aspects of reaction norm shape are more likely to evolve, based on how they vary genetically. For instance, a reaction norm may be highly curved (e.g. quadratic) but have little genetic variability in curvature, instead mostly varying in position, height, or local slope. There is thus a need to compare genetic variation in different components of reaction norm slope, as previously done in a meta-analysis (Murren et al. 2014, but see Appendix D). However, comparing genetic variation in the slope versus curvature of a reaction norm, for instance, is not straightforward, as these parameters have different scales and even units (trait per environment, vs trait per squared environment). Genetic variation in reaction reaction norm shape can be analyzed by estimating variation in the parameters of a continuous function of the environment (e.g. polynomial), possibly using the flexible framework of function-valued traits (Kirkpatrick & Heckman 1989; Gomulkiewicz & Kirkpatrick 1992; Stinchcombe et al. 2012). But even this flexible approach generally "makes the restrictive assumption that all individuals or genotypes are fully characterized by the chosen parametric model" (Stinchcombe et al. 2012), and the degree to which the overall plastic variance in the trait is explained by this model is rarely evaluated. In addition, it would be useful to be able to compare the relative contributions of variation in different aspects of reaction norm shape to the overall variance in plasticity of a trait.

We herein propose a simple framework to estimate and partition the phenotypic variance of reaction norms, towards three main goals: (i) quantify plasticity across reaction norm shapes and types; (ii) evaluate the contribution of different aspects of reaction norm shape, and of the full assumed reaction norm model, to overall plastic phenotypic variation; and (iii) quantify heritable variation in different aspects of reaction norm shape. Our hope is that this study will stimulate more quantitative investigations of the ways in which phenotypic plasticity contributes to phenotypic variation and evolutionary change.

Alternative models of reaction norms

In the broadest sense, a reaction norm is a decomposition of phenotypic variation among known (often controlled) versus unknown sources of environmental variation. We can write the measure i of phenotypic trait z for genotype g developing in environment k as

$$z_{gki} = \hat{z}_{gk} + \tilde{z}_i. \tag{1}$$

The first term $\hat{z}_{gk}$ is the reaction norm, that is, the component of phenotypic variation that can be predicted (hence the hat notation) from knowing both the genotype of an individual and the environment in which it developed. The second term $\tilde{z}_i$ is the component of the measured phenotype that cannot be predicted from genotype and environment, and arises from unknown environmental factors (usually described as micro-environmental variation), developmental noise, and measurement error.

The reaction norm $\hat{z}_{gk}$ can be further categorized according to the type of environmental variation. The environment may be inherently categorical and unordered, such as host plant for a herbivore insect. It may be ordered but with no (or unknown) quantitative value, such as low, medium, and high treatments. Or it may be ordered quantitatively, with values that are either intrinsically discrete (such as number of resource items), or continuous (even if sampled at discrete intervals), such as temperature or salinity.

When environments are purely categorical, the reaction norm can be studied by treating phenotypic values in different environments as alternative 'character states', considered as different traits in a multivariate framework (Via & Lande 1985; Falconer 1952). The mean character state may differ among environment if the trait is plastic; phenotypic and genetic variation may be larger in some environments; and phenotypes may be more or less correlated across environments (Via & Lande 1985; Falconer 1952). Such a modelling framework is readily described by Equation 1 for a discrete genotype g and environment k. In practice, such approach would correspond to an ANOVA (or a mixed model) with discrete environment and genotype-within-environment as (random) effects of the model. In its most compact form, such a statistical model can be framed as a multivariate Gaussian distribution, with a number of dimensions corresponding to the number of categories in the environment,

$$\hat{\boldsymbol{z}} \sim \mathcal{N}\left(\boldsymbol{\mu}, \mathrm{G}_z\right), \tag{2}$$

where $\boldsymbol{\mu}$ is the vector of expected phenotypic values (across genotypes) within each environment, and G_z is the genetic variance-covariance matrix of the phenotype. Note that when the environment is quantitative but discrete, one may still use the character state approach, but structuring correlations in G_z by environmental distance, in effect treating the phenotype as a stochastic process characterized by its autocovariance function

across environments (Pletcher & Geyer 1999).

For quantitative environments (both discrete and continuous), the most common approach is to use a function f to model the reaction norm,

$$\hat{z}_{gk} = f(\varepsilon_k, \boldsymbol{\theta}_g), \tag{3}$$

where $\boldsymbol{\theta}_g$ is a vector that contains the parameters of the function (e.g. coefficients associated to each exponent for a polynomial) for each genotype g; these parameters are thus genetically variable. In practice, such approach is implemented through (possibly non-linear) mixed models (Morrissey & Liefting 2016), in which genetic variation in f is modelled through random effects on its parameters $\boldsymbol{\theta}$ (omitting the subscript g for simplicity). The $\boldsymbol{\theta}$ are generally assumed to be polygenic and thus follow a multivariate Gaussian distribution,

$$\boldsymbol{\theta} \sim \mathcal{N}(\bar{\boldsymbol{\theta}}, \Theta), \tag{4}$$

where $\bar{\boldsymbol{\theta}}$ is the vector of average parameter values across genotypes and Θ is the additive genetic variance-covariance matrix of the parameters $\boldsymbol{\theta}$. This approach has been described alternatively as the "reaction norm" approach, the "polynomial approach", or a parametric version of function-values traits. To keep it general here and avoid confusion with the general concept of reaction norm as defined in Equation 1, we will describe it as the "curve parameter" approach. Note that, for a given reaction norm, some parameters in $\boldsymbol{\theta}$ (and/or their genetic variation) may depend on how ε was defined (e.g. whether it was mean-centered or not). For instance, changing what environment is chosen as the reference (where $\varepsilon = 0$) will change the intercept of a linear reaction norm and its genetic variance (as explained in more detail in Lande 2009).

We show below that these modelling choices can be unified under a common framework, following the spirit of de Jong (1995). More specifically, common metrics of variance partitioning can be computed regardless of the approach used, and translated from one approach to another, allowing for broad comparison of plasticity across organisms, traits, and environments. This also allows highlighting complementary strengths and weaknesses of the character-state and curve parameter approaches, when both are available.

Partitioning variation in reaction norms

The terms in Equation 1 are assumed to be independent, such that the total phenotypic variance $\mathrm{V}(z)$ (usually noted V_P) is the sum of the variance predicted by the genotype and the environment $\mathrm{V}(\hat{z})$, plus a residual component of variance $\mathrm{V}(\tilde{z}_i)$, which we will note V_{Res}. The predicted variance component $\mathrm{V}(\hat{z})$ can be furthered partitioned using the law of total variance across genotypes and environments, leading to

$$\mathrm{V}(\hat{z}) = \mathrm{V}_{\varepsilon}(\mathrm{E}_{g|\varepsilon}(\hat{z})) + \mathrm{E}_{\varepsilon}(\mathrm{V}_{g|\varepsilon}(\hat{z})), \tag{5}$$

where E_x and V_x denote expectation and variance along variable x (either the environment ε, or the genotype-within-environment $g|\varepsilon$). Figure 1 illustrates this variance partitioning for a quadratic reaction norm. The first term captures how much phenotypic variance across environments results from plasticity in the mean reaction norm averaged over genotypes (see Figure 1), so we denote it as V_{Plas}. The second term is the phenotypic variance among genotypes within environment averaged across environments, *i.e.* the variance arising from genetic variation around the average reaction norm (Figure 1), so we denote as as V_{Gen}. Overall, we thus have for the total phenotypic variance

$$V_P = V_{\mathrm{Plas}} + V_{\mathrm{Gen}} + V_{\mathrm{Res}} \tag{6}$$

This differs from the classical partitioning into genetic, environmental, and genotype-by-environment interaction effects in quantitative genetics (Falconer & Mackay 1996; Lynch & Walsh 1998; Des Marais et al. 2013). The environmental component from this classical partitioning is here split between the V_{Plas} and V_{Res} component, while our V_{Gen} component accounts for both the genetic and genotype-by-environment effects. Note that this is in contrast to another view, where the genotype-by-environment interaction is instead associated with the environmental component, e.g. as *plastic variance* (Scheiner & Lyman 1989; Scheiner 1993a; Falconer & Mackay 1996; Lynch & Walsh 1998). Each variance partitioning is relevant in what it can unveil and limited by what it hides. We explore here what the partitioning in Equation 6 can bring, both conceptually and methodologically. A more detailed and nuanced comparison, with a worked example, is provided in Appendix A. The genetic variance can further be decomposed into an additive (heritable) component V_{A} and a non-additive component V_{NA}, with the latter comprising the dominance and epistasis variance, which are not our focus here.

Contributions from the average plasticity

We can now proceed to refine the definition of V_{Plas} and analyze its dependency on reaction norm shape. In the character-state approach, the variance partitioning in Equation 5 readily follows from Equation 2 since $\mathrm{E}_{g|\varepsilon_k}(\hat{z}) = \mu_k$, and we have

$$V_{\mathrm{Plas}} = \mathrm{V}_{\varepsilon}(\mu), \tag{7}$$

i.e. the plastic variance is the variance in the expected character state μ_k across environmental levels k.

In the curve parameter approach, the first step is to compute the mean phenotypic conditional on the

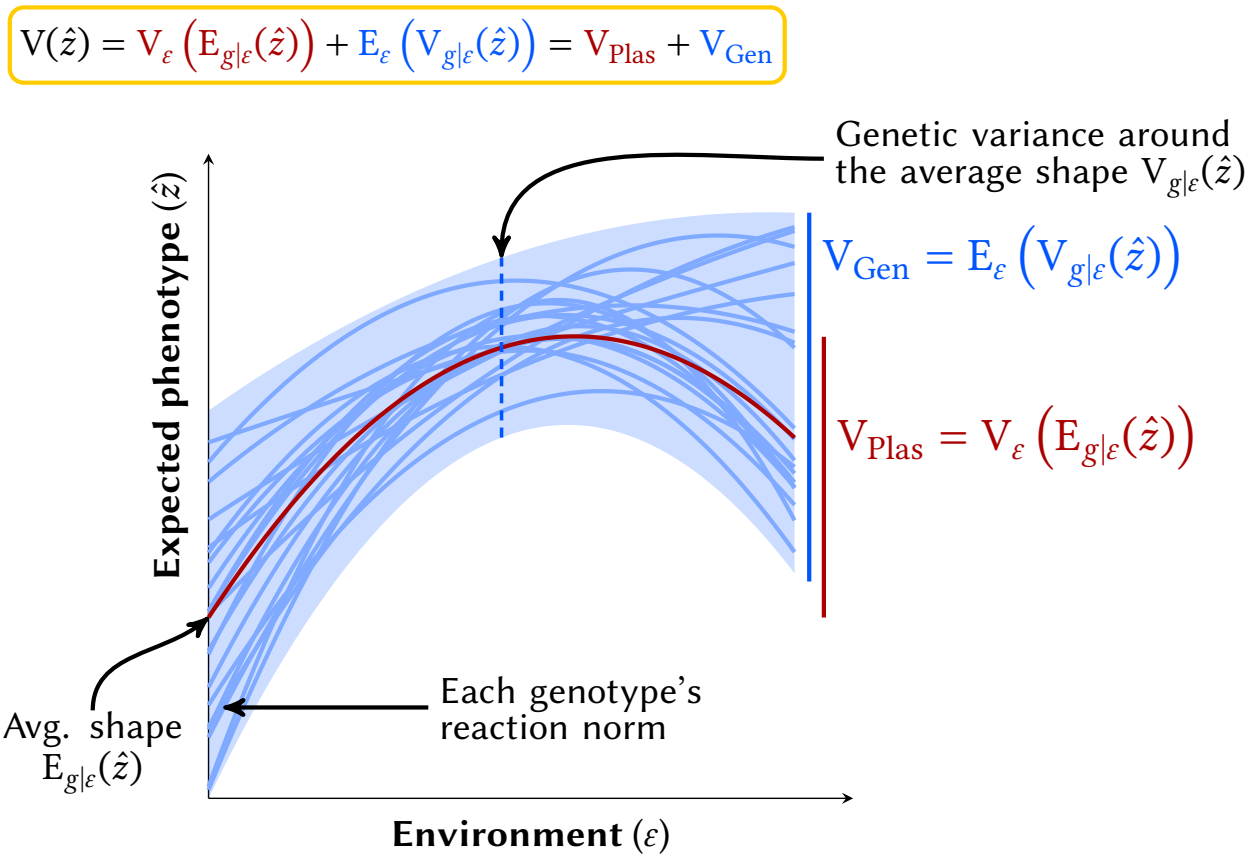

Figure 1: Schematic illustration of our variance partitioning in the case of a quadratic reaction norm, using the curve parameter approach. The variance of the expected phenotype according to each genotype's reaction norm (light blue lines) is partitioned into the component due to the average plasticity shape (V_{Plas}, in red) and the component due to genetic variation around this average shape (V_{Gen}, in blue and corresponding to the blue area). For example, if the genetic variation (blue area) is small comparatively to the "trajectory" of the average shape (red line), then V_{Gen} will be small compared to V_{Plas}, meaning that most of the phenotypic variation comes from the direct effect of plasticity, rather than from genetic variation in plasticity.

environment,

$$\mathrm{E}_{g|\varepsilon}(\hat{z}) = \int f(\varepsilon, \boldsymbol{\theta}_g) p(\boldsymbol{\theta}_g) \mathrm{d}g, \tag{8}$$

where $p(\boldsymbol{\theta}_g)$ is the probability density function of the parameters $\boldsymbol{\theta}_g$ due to the variability across genotypes. From this, V_{Plas} can be computed as

$$V_{\mathrm{Plas}} = \int (\mathrm{E}_{g|\varepsilon}(\hat{z}) - \bar{z})^2 p(\varepsilon) \mathrm{d}\varepsilon, \tag{9}$$

where $p(\varepsilon)$ is the probability density function of the environmental variable ε, and $\bar{z}$ is the average phenotype among genotypes and environments (*i.e.*, the grand mean phenotype). If the reaction norm function f is linear in its parameters $\boldsymbol{\theta}$ (not to be confused with linearity with respect to the environment ε, *i.e.* a linear reaction norm) then $\mathrm{E}_{g|\varepsilon}(\hat{z}) = f(\varepsilon, \bar{\boldsymbol{\theta}})$ (noted simply as $f(\varepsilon)$ below), which simplifies the computation.

Although the shape of the true reaction norm function f cannot be known with certainty and may be complex, it is often of interest to fit relatively simple functions with interpretable parameters. For instance, first- or second-order approximations to the reaction norm provide information on its slope or curvature. More generally, polynomial functions allow fitting reaction norms with potentially complex shapes while retaining linearity in their parameters, making them popular in studies of reaction norms, both theoretically (Scheiner 1993b) and empirically (Morrissey & Liefting 2016). To exemplify how different components of reaction norm shape contribute to phenotypic variance, let us first focus on the quadratic case,

$$f(\varepsilon) = a + b\varepsilon + c\varepsilon^2, \tag{10}$$

which includes linear reaction norms as a subcase when $c = 0$. In this model, the variance arising from the average reaction norm is

$$V_{\text{Plas}} = \bar{b}^2 \mathrm{V}_\varepsilon(\varepsilon) + \bar{c}^2 \mathrm{V}_\varepsilon\left(\varepsilon^2\right) + 2\bar{b}\bar{c}\,\mathrm{Cov}_\varepsilon(\varepsilon, \varepsilon^2), \tag{11}$$

where bars denote averages over genetic variation. If the environmental variable ε has been mean-centered and is symetrical (*e.g.* Gaussian), then $\mathrm{cov}(\varepsilon, \varepsilon^2) = 0$ and the third term vanishes. We may then compute the relative contributions of reaction norm slope and curvature to the total variance attributable to the average reaction norm as

$$\pi_b = \frac{\bar{b}^2 \mathrm{V}_\varepsilon(\varepsilon)}{V_{\text{Plas}}}, \qquad \pi_c = \frac{\bar{c}^2 \mathrm{V}_\varepsilon\left(\varepsilon^2\right)}{V_{\text{Plas}}}. \tag{12}$$

An important point arising from Equation 12 is that the relative importances of the linear and quadratic components of the curves depends on variation in the environment, respectively $\mathrm{V}_\varepsilon(\varepsilon)$ and $\mathrm{V}_\varepsilon\left(\varepsilon^2\right)$. Figure 2 show the values of π_b and π_c for various quadratic reaction norms, assuming ε follows either a normal or uniform distribution, with same mean 0 and variance 1. The values for π_b and π_c translate well the perceived "trendiness" (for large π_b) or "curviness" (for large π_c) of reaction norms, but they may also strongly depend on the statistical distribution of the environmental variable ε, as shown especially in the third example of Figure 2. In this example, the difference arises because the assumed environmental distributions have different kurtosis (the scaled fourth central moment, related to $V_\varepsilon(\varepsilon^2)$ in Equation 12). Because $V_\varepsilon(\varepsilon^2)$ is larger for the Gaussian, this distribution leads to larger π_c than the uniform.

To generalise this reasoning to any polynomial order n, it is convenient to use linear algebra, in line with theoretical work by Gavrilets & Scheiner (1993a). A polynomial reaction norm can be written as

$$\hat{z} = \mathrm{x}^T \boldsymbol{\theta}, \tag{13}$$

where the column-vector $\mathrm{x} = \left(1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^n\right)^T$ (where ^{T} denotes transposition) includes all exponentiation

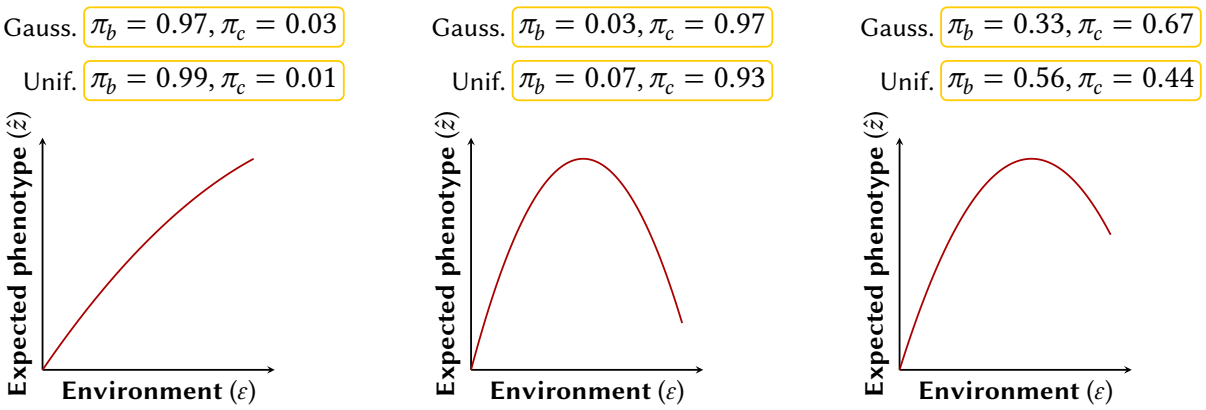

Figure 2: Computation of π_b and π_c, the relative contributions of linear and quadratic terms to phenotypic variation caused by the mean reaction norm, for different shapes of reaction norms, and two distributions of the environmental variable ε: a standard Gaussian (of mean 0 and variance 1), and a uniform distribution between $-\sqrt{3}$ and $\sqrt{3}$ (of mean 0 and variance 1).

levels (up to n) of the environmental variable ε. The variance component due to plasticity in the average reaction norm is then

$$V_{\text{Plas}} = \bar{\boldsymbol{\theta}}^T \mathrm{X} \bar{\boldsymbol{\theta}}, \tag{14}$$

where X is the variance-covariance matrix of x, recalling that $\bar{\boldsymbol{\theta}}$ is the average of reaction norm parameters across the genotypes. The relative contribution of a given exponent m to the variance caused by the mean plasticity becomes

$$\pi_m = \frac{\bar{\theta}_m^2 \mathrm{X}_{m,m}}{V_{\text{Plas}}} = \frac{\bar{\theta}_m^2 \mathrm{V}(\varepsilon^m)}{V_{\text{Plas}}}, \tag{15}$$

and the contribution of the covariance between exponents l and m is

$$\pi_{lm} = \frac{2\bar{\theta}_l \bar{\theta}_m \mathrm{X}_{l,m}}{V_{\text{Plas}}} = \frac{2\bar{\theta}_l \bar{\theta}_m \mathrm{Cov}(\varepsilon^l, \varepsilon^m)}{V_{\text{Plas}}}. \tag{16}$$

Note that even with a symmetrical and mean-centered environment, the covariance between higher-up order exponents will not be zero in general, contrary to ε and ε^2 in the quadratic case.

Contributions from genetic variation

We now turn to how genetic variation in reaction norms translates into genetic variance of the trait across environments. In the character-state approach, the genetic variance within each environment is given by the diagonal elements of G_z, so we simply have

$$V_{\text{Gen}} = \mathrm{E}(\mathrm{diag}(\mathrm{G}_z)), \tag{17}$$

that is, V_{Gen} is the average genetic variance of character states across environments. Note however that this cannot be directly used to predict the mean response to selection in a variable environment, as the latter are also influenced by genetic correlations in character state across environments (Via & Lande 1985; Gomulkiewicz & Kirkpatrick 1992). In addition, whether Equation 17 actually outputs V_{Gen}, or rather its heritable component V_{A}, entirely depends on whether the matrix G_z is defined as containing the *total* genetic (co)variances or only the *additive* genetic (co)variances.

In the curve parameter approach, expanding the second term in Equation 5 we get

$$V_{\text{Gen}} = \int V_{g|\varepsilon}(\varepsilon) p(\varepsilon) \mathrm{d}\varepsilon. \tag{18}$$

From the reaction norm function in Equation 3 and under multivariate Gaussian distribution assumed in Equation 4, the genetic variance conditional on environment becomes

$$V_{g|\varepsilon}(\varepsilon) = \int \left(f(\varepsilon, \boldsymbol{\theta}) - \mathrm{E}_{g|\varepsilon}\left(f(\varepsilon, \boldsymbol{\theta})\right)\right)^2 p_{\mathcal{N}}(\boldsymbol{\theta}) \mathrm{d}g. \tag{19}$$

Numerical integration of Equation 19 can be used in any case to obtain V_{Gen}. However, further analytical progress can be made when focusing more specifically on the additive genetic variance V_{A}, which more directly influences responses to selection (Lynch & Walsh 1998). Using the property of additivity of breeding values, and relying on a multivariate extension of the framework in de Villemereuil et al. (2016), it is shown in Appendix B that the additive genetic variance in environment ε is

$$V_{\text{A}|\varepsilon} = \boldsymbol{\psi}_\varepsilon^T \Theta \boldsymbol{\psi}_\varepsilon. \tag{20}$$

where $\boldsymbol{\psi}_\varepsilon$ is the vector of mean partial derivatives of the reaction norm function f with respect to each of its parameters. The total additive genetic variance is then obtained by averaging over environments: $V_{\text{A}} = \mathrm{E}_\varepsilon(V_{\text{A}|\varepsilon})$. Terms in the quadratic form of Equation 20 can be expanded to yield a decomposition of the additive genetic variance into contributions from (co)variances of different parameters of the reaction norm function,

$$\gamma_i = \frac{\mathrm{E}_\varepsilon\left(\psi_{\varepsilon,i}^2\right) \mathrm{V}_g(\theta_i)}{V_{\text{A}}}, \qquad \gamma_{ij} = \frac{2\mathrm{E}_\varepsilon\left(\psi_{\varepsilon,i}\psi_{\varepsilon,j}\right) \mathrm{Cov}_g(\theta_i, \theta_j)}{V_{\text{A}}}, \qquad \sum_i \gamma_i + \sum_{i<j} \gamma_{ij} = 1 \tag{21}$$

Importantly, when the reaction norm function f (and thus $\hat{z}$) is linear in its parameters (which again covers many cases of non-linear reaction norms with respect to the environment, including polynomial functions), it can be shown that $V_{\text{Gen}} = V_{\text{A}}$ (see Appendix B), so Equation 21 and Equation 20 apply directly to V_{Gen}, providing a simpler way to compute it in this case.

We can illustrate the general decomposition of V_{Gen} in the case of polynomial reaction norms (following

De Jong 1990; Gavrilets & Scheiner 1993a,b), as done above for V_{Plas}. When the reaction norm is a polynomial function of the environment, then the gradient of $\hat{z}$ with respect to reaction norm parameters is simply the vector of exponents of the environment defined below Equation 13, $\boldsymbol{\psi} = \mathrm{x}$. Then using Equation 20, we have

$$V_{\text{Gen}} = V_{\text{A}} = \mathrm{E}_{\boldsymbol{\varepsilon}}(\mathrm{x}^T \Theta \mathrm{x}) = \bar{\mathrm{x}}^T \Theta \bar{\mathrm{x}} + \mathrm{Tr}(\Theta \mathrm{X}) \tag{22}$$

where $\bar{\mathrm{x}}$ is the vector of the average of the exponentiated environments, X their covariance matrix defined in Equation 14 and Tr stands for the trace of a matrix. Note that the trace of a matrix product is the sum of element-wise products of their terms. With a quadratic reaction norm as in Equation 10, this becomes

$$V_{\text{Gen}} = V_a + 2C_{ab}\mathrm{E}(\varepsilon) + 2C_{ac}\mathrm{E}(\varepsilon^2) + V_b\mathrm{E}(\varepsilon^2) + 2C_{bc}\mathrm{E}(\varepsilon^3) + V_c\mathrm{E}(\varepsilon^4), \tag{23}$$

where terms in V and C denote additive genetic variances and covariances of the reaction norm parameters defined in Equation 10. If the environmental variable is symmetrical and has been mean-centred, then $\mathrm{E}(\varepsilon) = \mathrm{E}(\varepsilon^3) = 0$, such that

$$V_{\text{Gen}} = V_a + 2C_{ac}\mathrm{E}(\varepsilon^2) + V_b\mathrm{E}(\varepsilon^2) + V_c\mathrm{E}(\varepsilon^4) \tag{24}$$

Note the importance of the genetic covariance between the intercept and the curvature component C_{ac}, which can have a critical evolutionary role (Gavrilets & Scheiner 1993b). From Equation 24, we can compute the contribution of each component of genetic variance in reaction norm to the total genetic variance (averaged across environments):

$$\gamma_a = \frac{V_a}{V_{\text{A}}}, \quad \gamma_b = \frac{V_b\mathrm{E}(\varepsilon^2)}{V_{\text{A}}}, \quad \gamma_c = \frac{V_c\mathrm{E}(\varepsilon^4)}{V_{\text{A}}}, \quad \gamma_{ac} = \frac{2C_{ac}\mathrm{E}(\varepsilon^2)}{V_{\text{A}}}. \tag{25}$$

As noted above for components of V_{Plas} in Equation 12, the components of V_{Gen} in Equation 25 depend on the distribution of environments, through its moments $\mathrm{E}(\varepsilon^n)$.

Parameter estimation and variance partitioning in practice

Estimating the parameters

All the parameters mentioned above can be estimated through commonly used statistical frameworks. A tutorial is available at github.com/devillemereuil/TutoPartReacNorm showing how to implement such models using e.g. the frequentist lme4 (Bates et al. 2015) and Bayesian brms R packages (Bürkner 2017). For the character-state approach (Equation 2), a random-intercept model can be used, or alternatively a "multi-trait"

model (Rovelli et al. 2020; Mitchell & Houslay 2021). We will focus here on the former, which is more easily implemented while seemingly scarcely used in the literature on plasticity. In a random-intercept model, the environment is considered as a categorical variable, to which a random effect is added using the genotype as the grouping factor. In the curve parameter approach, the appropriate models will be random-slope models for a polynomial approach (as mentioned in Morrissey & Liefting 2016), or non-linear mixed models. Such a model is based on the reaction norm function $f(\varepsilon, \boldsymbol{\theta})$, possibly written as a linear model (e.g. for a polynomial function), to which random effects (with the genotype as grouping factor) are added for all of its parameters, e.g. the intercept, slope, and any higher-order effects for a polynomial function.

Since the parameters are estimated with noise, it is important to account for the impact of estimation uncertainty when computing variance components. In particular, while variances directly obtained using random effects (e.g. variances related to V_{Gen}) are expected to be unbiased, the variances arising from fixed effects (e.g. variances related to V_{Plas}) should be corrected for biases due to uncertainty. For example, the unbiased estimator of V_{Plas} in a polynomial model would be:

$$\hat{V}_{\text{Plas}} = \hat{\boldsymbol{\theta}}^T \mathrm{X} \hat{\boldsymbol{\theta}} - \mathrm{Tr}(\mathrm{S}_\theta \mathrm{X}), \tag{26}$$

where S_θ is the variance-covariance matrix of errors around the $\hat{\boldsymbol{\theta}}$ estimators (see Appendix C). The unbiased estimator for a character-state model would be:

$$\hat{V}_{\text{Plas}} = \mathrm{V}_\varepsilon(\mu) - \mathrm{E}_\varepsilon(s^2). \tag{27}$$

where s_k is the standard-error of μ_k at environment k (see Appendix C).

Perfect modelling of polynomial curves

We simulated phenotypic data conforming to a quadratic reaction norm, to evaluate the performance of the proposed approach when the true reaction norm is correctly modeled. We considered an environmental gradient of 10 values, equally spaced from -2 to 2, over which we defined a quadratic curve with average parameters $\bar{\boldsymbol{\theta}} = (1.5, 0.5, -0.5)$ for intercept, slope and curvature. We then drew 20 different genotype-specific vectors of curve parameter $\boldsymbol{\theta}$ from a multivariate normal distribution with mean $\bar{\boldsymbol{\theta}}$ and (genotypic) variance-covariance matrix

$$\Theta = \begin{pmatrix} 0.090 & -0.024 & -0.012 \\ -0.024 & 0.160 & 0.008 \\ -0.012 & 0.008 & 0.040 \end{pmatrix}.$$

Figure 1 displays examples of curves resulting from these parameters. Finally, we sampled 20 individual measures for each genotype with a residual variance $V_{Res} = 0.25$. This scenario corresponds to expected values $V_{Plas} = 0.92$ and $V_{Gen} = 0.5$, for a total phenotypic variance of 1.67. Our simulated conditions resulted in $20 \times 10 \times 20 = 4000$ data points per simulation, which is on the higher-end of the realm of practical datasets, since the aim was not to perform a power analysis, but to evaluate the soundness of the approach in practice. However the results were qualitatively unchanged when using 4 instead of 10 environments. The simulation process was repeated 100 times in R, and for each simulated dataset, we ran estimations using the lme4 R package (Bates et al. 2015) under both the curve parameter and character-state approaches, in order to check how these approaches compare in practice.

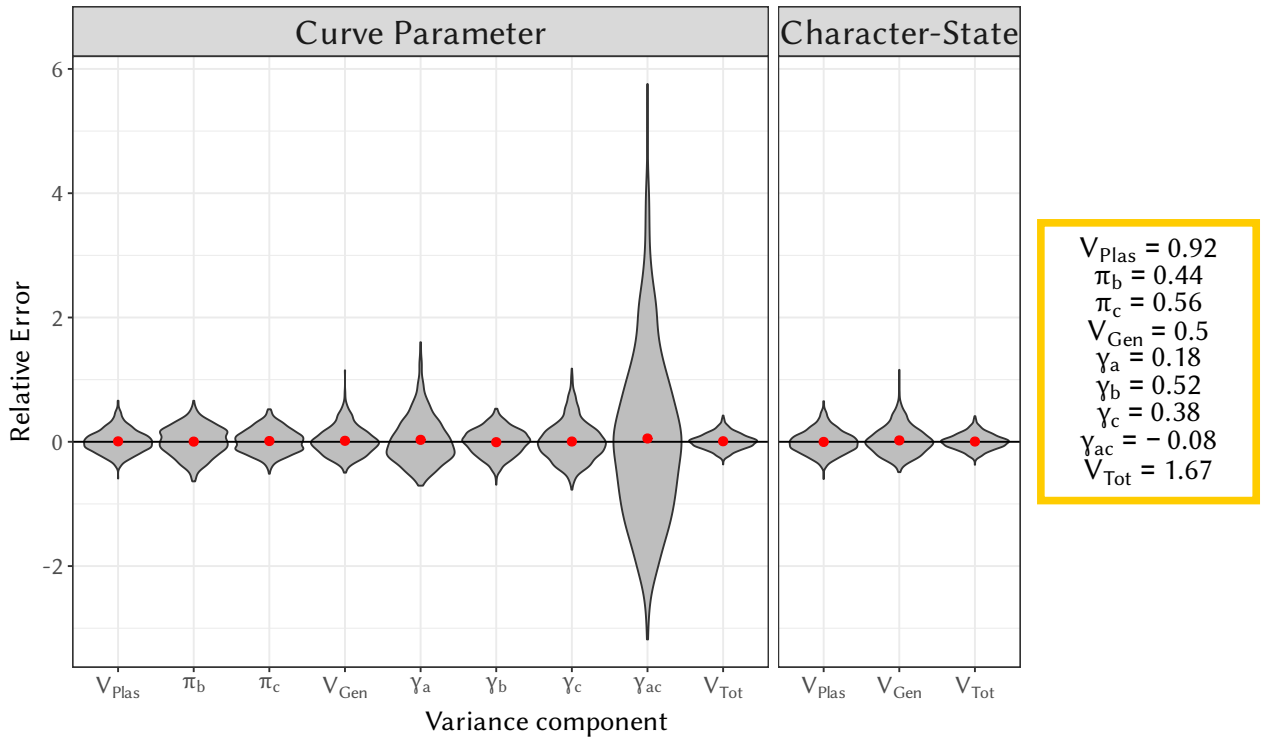

Figure 3: Distribution of the relative error (difference between the inferred and true value, divided by the true value) for each the inferred variance components. Estimates are for $\hat{V}_{Plas}$, $\hat{V}_{Gen}$ and $\hat{V}_{Tot}$ for both the curve parameter and character-state approaches. For the parameter curve, the π-decomposition of $\hat{V}_{Plas}$ into π_b (contribution of the slope) and π_c (contribution of the curvature) and the γ-decomposition of $\hat{V}_{Gen}$ into γ_a (genetic contribution of the intercept), γ_b (genetic contribution of the slope), γ_c (genetic contribution of the curvature) and γ_{ac} (genetic contribution of the covariance between the intercept and the curvature) is also shown. The red dots correspond to the average over the 1000 simulations. The yellow box provides the expected values for all of the estimates.

From the curve parameter models, we computed $\hat{V}_{Plas}$ as in Equation 26, as well as its π-decomposition (Equation 12) into π_b (part explained by the average linear trend) and π_c (part explained by the average curvature). We also computed $\hat{V}_{Gen}$ as in Equation 22 and its γ-decomposition (Equation 25) into γ_a (impact of the genetic variation of the intercept), γ_b (for the slope), γ_c (for of the curvature) and γ_{ac} (for the covariance between the intercept and curvature). From the character-state model, we computed $\hat{V}_{Plas}$ as in Equation 27 and $\hat{V}_{Gen}$ as in Equation 17. Finally for both models, we computed the total inferred variance as the sum

$\hat{V}_{\text{Plas}} + \hat{V}_{\text{Gen}} + V_{\text{Res}}$, and compared it to the sample phenotypic variance, to verify the ability of both approaches to implement the variance partitioning in Equation 6.

The results of the inferences are available in Figure 3. First, they show that both methods allow for unbiased inference (Wilcoxon's rank test, $p > 0.05$ for all components) of all estimates, showing that our variance partitioning is easily implemented with existing tools. There was, however, considerable uncertainty in the estimation of γ_{ac}, as covariances are typically more difficult to estimate. Second, and as a consequence, $\hat{V}_{\text{Tot}}$ retrieved the total phenotypic variance with extreme precision (correlation $> 99\%$). Third, and most interestingly, the results illustrate the equivalence between the curve parameter and character-state approches, as the distributions of $\hat{V}_{\text{Plas}}$ and $\hat{V}_{\text{Gen}}$ were correlated at $> 99\%$ between the two approaches. This means that our variance partitioning is not impacted by which approach is chosen to study plasticity, as long as the curve parameter approach captures the true reaction norm shape. When this does not hold, the differences between estimates from these alternative approaches can be exploited efficiently, as we describe below.

Assessing goodness-of-fit under imperfect modelling

The true shapes of reaction norms are generally unknown and may be complex, such that any curve parameter model is likely to be mis-specified to some extent. The character-state approach is arguable more general, as it does not assume anything about the "true" shape of the reaction norm (as pointed out previously by de Jong 1995). Nonetheless, having access to curve parameters is often very interesting and more actionable and interpretable, especially to predict evolution of phenotypic plasticity (see also de Jong 1995). To get the best of both worlds, we offer to rely on the robust ability of the character-state approach to recover V_{Plas}, using it as an "anchor" to test the goodness-of-fit of an assumed curve.

In order to demonstrate the soundness and usefulness of this approach, we simulated datasets following relatively common curves that are not well-captured by a second order polynomial: a logistic sigmoid, or a Gompertz-Gaussian performance curve (see Figure 4). We assumed that the environment is sampled at either 10 or 4 values. For each of these conditions, we simulated 1000 datasets, with 10 measures *per* environment (for the sake of simplicity, and given the focus on $\hat{V}_{\text{Plas}}$ here, we did not include different genotypes in these simulations). We estimated the parameters of a polynomial model, and computed the relative contributions of the slope and curvature using Equation 12. In addition, we computed the variance explained by our polynomial model as in Equation 26 (here specifically termed $\hat{V}_{\text{mod}}$), and compared it to $\hat{V}_{\text{Plas}}$ estimated from a character-state model (here a simple ANOVA, since genotypes are not modelled).

As a measure of goodness-of-fit, we computed the ratio of the variance explained by the polynomial curve

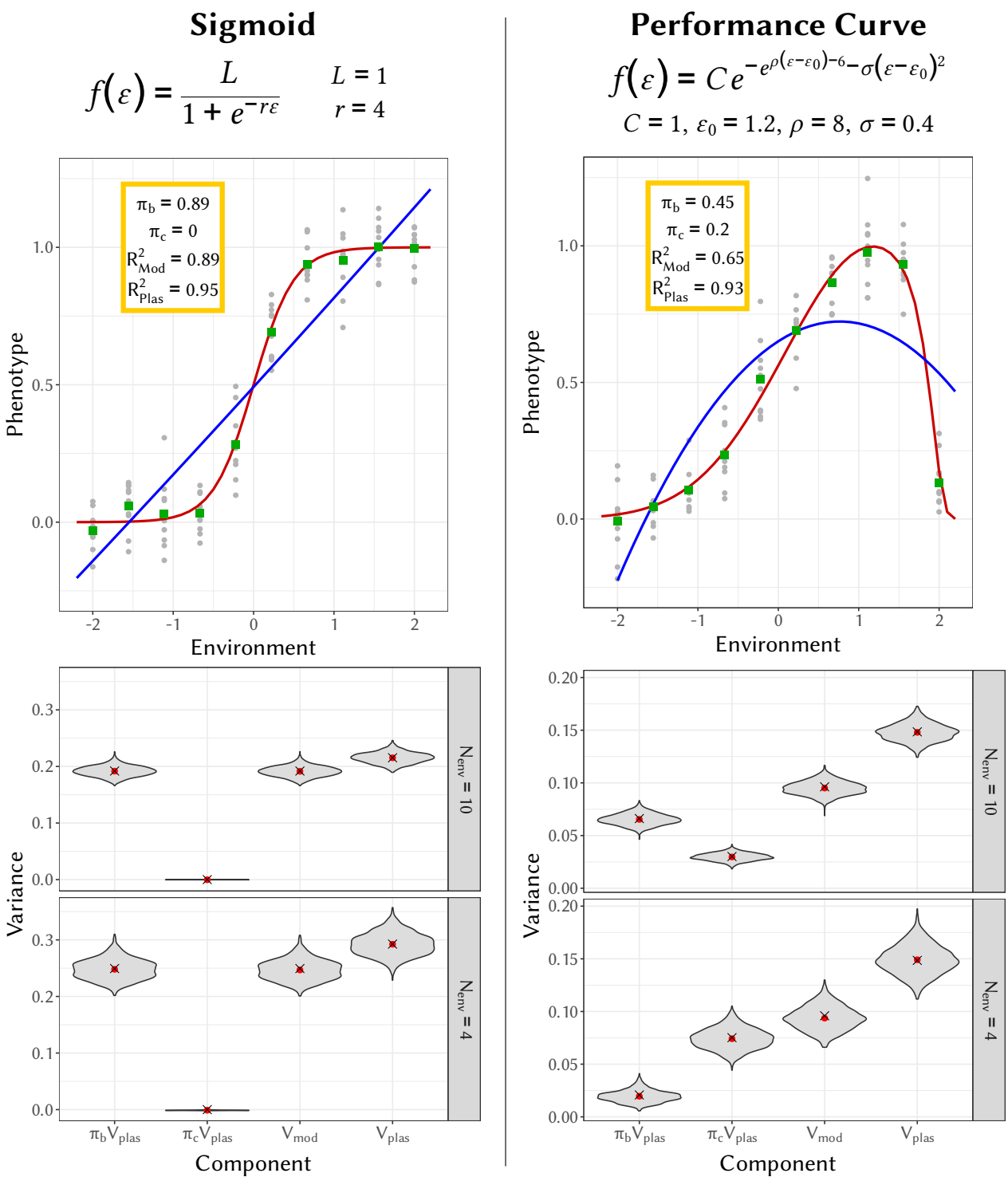

Figure 4: Estimation of the variance of the reaction norm when the true shape (sigmoid on the left, Gompertz-Gaussian performance curve on the right, red lines on top graphs) is unknown and approximated from a polynomial function. The estimated reaction norms using a polynomial function (blue line, top graphs) only account for a part of the reaction norm shape, while the ANOVA estimation (green dots, top graphs) fit the true shape more accurately. As a result, the model is expected to explain only a part $\hat{V}_{mod}$ of phenotypic variance due to plasticity (see R^2_{Mod}). The part of the total phenotypic variance explained by overall plasticity, $R^2_{Plas} = \hat{V}_{Plas}/V(z)$, is also provided for information. Replicating the simulation 1000 times shows that our estimation process is without bias (red dots: average estimated values; black crosses: expected values) and produce reasonable sampling variance, even if only 4 environment values are used (bottom graphs). For better readability, the π-decomposition of $\hat{V}_{mod}$ is provided on the scale of the original variance as the products $\pi_b \hat{V}_{mod}$ and $\pi_c \hat{V}_{mod}$.

to the total variance due to phenotypic plasticity:

$$R^2_{\text{mod}} = \hat{V}_{\text{mod}}/\hat{V}_{\text{plas}}. \tag{28}$$

Our results show that, as expected, the polynomial function is an imperfect proxy of our complex shapes (Figure 4, $R^2_{\text{mod}} = 0.89$ for the sigmoid and $R^2_{\text{mod}} = 0.65$ for the performance curve), but using the character-state approach allows retrieving the total plastic variance without bias. The approach described here is thus useful to compute a measure of goodness-of-fit of a given reaction norm model (e.g. a polynomial function) to an unknown true shape of the reaction norm. Here, while a linear function might be acceptable for the sigmoid curve, with $R^2_{\text{mod}} = 0.89$, even a quadratic function can be considered as a bad fit to the Gompertz-Gaussian performance curve ($R^2_{\text{mod}} = 0.65$). In more details, the average slope was the most important component to explain the phenotypic variation for the sigmoid curve ($\pi_b = 0.89$, same as the total model). This was because, as the average curvature of a sigmoid is zero, the quadratic component was always estimated close to zero ($< 10^{-3}$), resulting in no variance explained by the curvature in this case ($\pi_c = 0$). Of course, the sigmoid is not a straight line either, and some remaining variance unexplained by the polynomial curve ($1 - 0.89 = 0.11$) could have been explained by higher-order effects (e.g. cubic effect). By contrast, for the Gompertz-Gaussian performance curve, while the average slope was an important factor ($\pi_b = 0.47$), the average curvature also explained quite a lot of the variance as well ($\pi_c = 0.2$). Again, higher-order effect, including at least a cubic effect, would have explained more of the variance arising from the average shape of plasticity.

This example illustrates the usefulness of a combined curve parameter and character-state approach to study the shape of reaction norms. While the character-state approach provides a robust estimation of V_{Plas}, the curve parameter approach provides interpretable information about the average slope and curvature (and higher-orders if needed) of the reaction norm, which helps describing where most phenotypic variance lies. Using our measure of goodness-of-fit R^2_{mod}, this analysis can be performed to assess how well a chosen polynomial function models an actual reaction norm. Note that R^2_{mod} is not penalised for the number of parameters, and thus should not be used for model selection.

Estimation of non-linear models

Although we have focused so far on models that are linear in the parameters to estimate (*e.g.*, the coefficients associated to each exponent of the environment for a polynomial reaction norm), the approach we propose can also be applied to arbitrary functions. This requires numerically computing the integrals in the most general definitions of V_{Plas} and V_{Gen} above, but this can be solved with efficient algorithms. We illustrate this here using the sigmoid and performance curve shapes above, introducing genetic variation in the parameters,

beyond the mean curves illustrated in Figure 4 (top panels). Instead of fitting polynomials as in Figure 5, we estimated the actual functions used to generate the datasets, using the non-linear modelling function of nlme package (Pinheiro et al. 2009). We used the cubature package (Narasimhan et al. 2023), as in the QGglmm package (de Villemereuil et al. 2016), to compute $\boldsymbol{\psi}_\varepsilon$ and $V_{A|\varepsilon}$. We simulated 1000 datasets for each scenario, consisting of 100 individuals (i.e. the "genotype") measured in each of 10 environments (say at 10 different temperatures).

We retrieved our simulated parameters without bias using the nlme function. As a result, we successfully recovered all the variance components defined in Equation 6 (Figure 5, bottom panels). This includes the estimation of the total additive genetic variance of the trait V_A. Indeed, almost all components of variance were unbiased (Wilcoxon's rank test, all $p > 0.05$ but one). The only exceptions ($p < 0.05$) were V_{Gen} and V_A in the Performance Curve case, although the relative bias is extremely small (resp. 1.20% and 1.13%), especially with regard to the uncertainty surrounding the estimates. This results from a slight bias in the estimation of the Θ matrix by the nlme function. Because of this, there is a slighter bias in V_{Tot} (0.39%).

Moreover, the sum of variance components (V_{Tot} in Figure 5) successfully reflects the total phenotypic variance, with a correlation between the two quantities $> 99.9\%$. One unfortunate aspect of running a non-linear model is that the correction method offered in Equation 26 no longer holds, precisely because of non-linearity in the model. However, this bias is generally small provided the standard error is small for most parameters, and the resulting bias in V_{Plas} is extremely small, especially with regard to the imprecision, as can be seen in Figure 5 and the non-significant result of Wilcoxon's rank test. In general, this bias will be small in regards to other sources of imprecision, unless the standard error of the estimates is extremely large (e.g. for very small sample size). An important distinction here is the difference between the curve defined by the average parameters $f(\varepsilon, \bar{\boldsymbol{\theta}})$ (Figure 5, top panel, black curve) and the one defined by the local average phenotype $\mathrm{E}_{g|\varepsilon}(\hat{z})$ (Figure 5, top panel, red curve), recalling that V_{Plas} is linked to the latter. While the two are very close for the sigmoid case, their differ quite strongly for the performance curve one.

Although the variation between individuals (i.e. genotypes in this simulation) in the top panel of Figure 5 seems quite large, the variance due to the average plasticity V_{Plas} is two to four times higher than the genetic variance V_{Gen} (Figure 5, yellow box in top panels). This occurs because the genetic variance is actually very low in most environments (Figure 5, blue violins of the middle panels), and scarcely as high as V_{Plas}. This illustrates how our variance partitioning can quantify and objectify variations that may be counter-intuitive for the human eye, notably because of non-linearities.

An important aspect of such modelling of the reaction norm is that there is no longer an equivalence between the genetic variance V_{Gen} and the additive genetic variance V_A, due to the non-linearity of the system (de Villemereuil et al. 2016). In this regard the sigmoid model does unexpectedly yield extremely close values

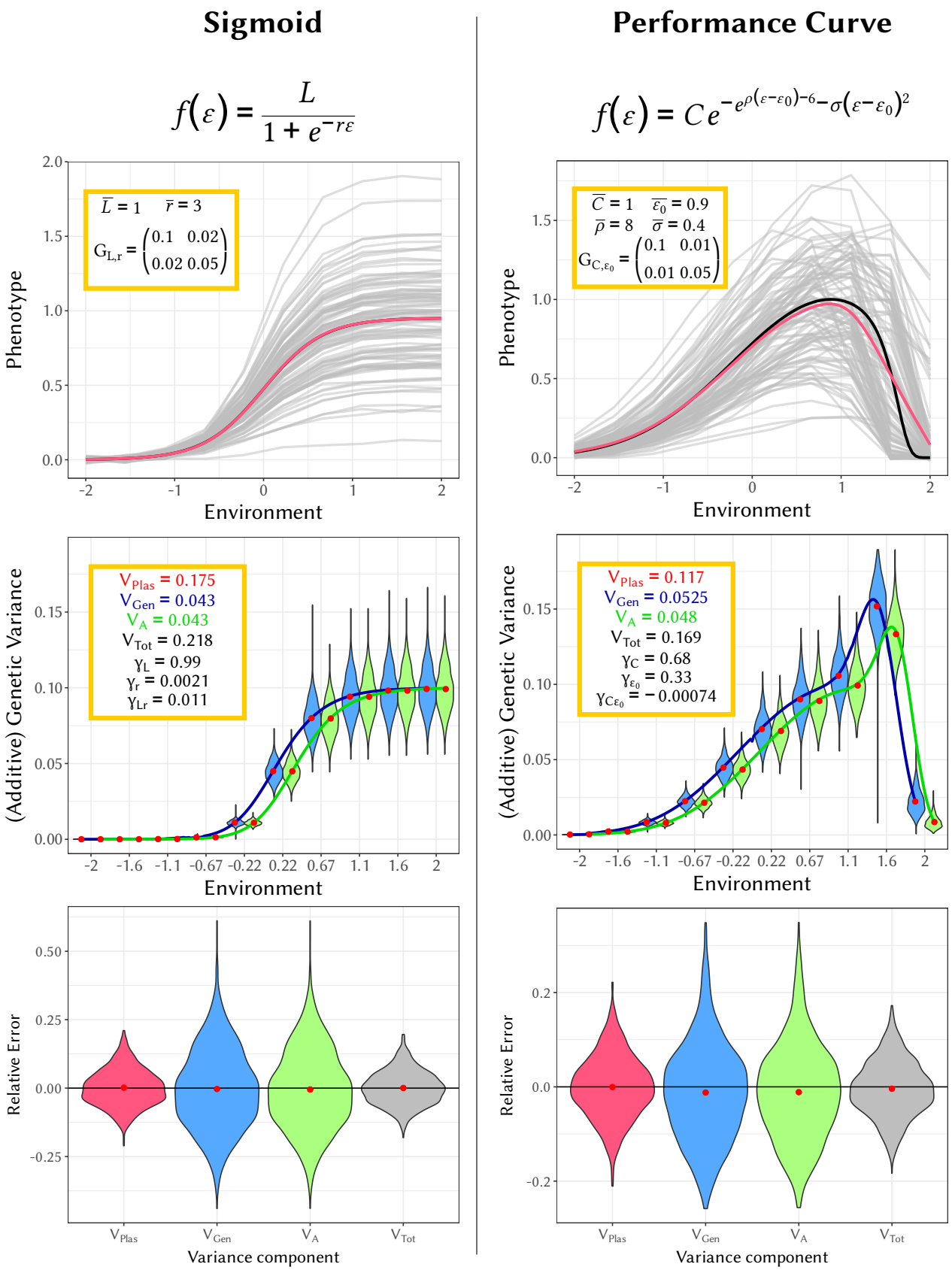

Figure 5: Scenarios and results of non-linear modelling of phenotypic plasticity. On the left: results corresponding to a sigmoid curve scenario; on the right: results corresponding to a performance curve scenario. Top panels: example of the individual curves (each curve corresponds to one individual) simulated in each scenario; yellow box: true parameters; black curve : $f(\varepsilon, \overline{\theta})$; red curve : $E_{g|\varepsilon}(\hat{z})$. Middle panels: distribution of the estimations of V_{Gen} (blue) and V_A (green), for each environment; red dot is the average of estimates over all simulations; blue and green solid lines are the true values for V_{Gen} and V_A in each environment (the lines are shifted horizontally for more clarity); yellow box: expected values for the variance partition. Bottom panels: distribution of the relative error (error divided by the expected value) for each component of our variance partition and the total variance, red dot is the average of estimates over all simulations.

for V_{Gen} and V_A (Figure 5, yellow box in top panels, blue and green violins in middle panels). This is the result of the disproportionate importance of the genetic variation in the L parameter is this model ($\gamma_L = 0.99$), even though the genetic variance in L is only twice that in r in the Θ matrix. Since L is only a mere scaling factor for the model, its relation with the phenotype is linear and thus $V_{Gen} \simeq V_A$. On the contrary, V_{Gen} and V_A differ for the performance curve model, especially in parts of the model where the local shape differs strongly between individuals (e.g. the two last environmental values, Figure 5, right middle panel). In this case, V_A depends less exclusively on variation in the scaling factor C ($\gamma_C = 0.68$), with $\gamma_{\varepsilon_0} = 0.33$. Hence in this model, the non-linearity due to the exponential function of ε_0 causes more substantial difference between V_{Gen} and V_A.

Despite being slightly more complex to implement, this non-linear approach can be highly relevant in practice, as it offers an in-depth analysis of the shape and genetic features of phenotypic plasticity. Moreover, although the environment simulated here was discretised for the sake of simplicity (and to favour good convergence in nlme), this approach would be most relevant when the (measured) environment is continuous rather than discretised, as in analysis of natural, uncontrolled environments.

Discussion

The variance partitioning that we implement here has several conceptual and practical advantages. First, being based on the law of total variance, it is very general and does not rely on any particular assumptions, such as Independence between the genotype and the environment. Note that contrary to the common genotype/environment/genotype-by-environment partition, the law of total variance is not symmetrical. Indeed, Equation 5 takes averages and variances first over genotypes, and then over environments. This allows recovering intuitive metrics of the influence of the average reaction norm (V_{Plas}), and the average genetic variance (V_{Gen}).

Second, in combination with polynomial modelling (or other forms of parametric approaches), this partitioning allows quantifying the impacts of different aspects of reaction norm shape on the mean plastic variance, versus the genetic variance of the trait. This should prove especially relevant with respect to responses to selection. For instance if a given selection episode concerns individuals that all experienced the same plasticity-inducing environment (i.e. when spatial environmental variation is negligible reltive to temporal variation), using the multivariate breeder's equation (Lande 1979) the response to selection for the expressed plastic trait z is

$$\Delta\bar{z} = \sum_i \gamma_i \beta V_A + \sum_{i<j} \gamma_{ij} \beta V_A, \tag{29}$$

where β is the selection gradient on the expressed trait, and the γ_i and γ_{ij} are defined in Equation 21. In

other words, the contributions of responses to selection by different reaction norm parameters (e.g. slope, curvature, etc) to overall response to selection by the plastic trait z is directly proportional to their contribution to its genetic variance. Importantly, these contributions will depend on the environment, as illustrated in Equation 21. In fact, the environment-specific additive genetic variance $V_{A,\varepsilon}$ is a critical piece of information regarding evolutionary potential. For example, in the Performance Curve scenario investigated above, there is a peak of additive genetic variance close to the performance optimum, followed by a sharp decrease at higher temperatures (Figure 5, middle right panel). In the context of predicting eco-evolutionary response to warming, this would mean that a slight temperature rise above the optimum would provide a very short window of higher evolvability, but followed by a sharp decrease thereof if warming persists. Beyond these simple scenarios, how selection acts on reaction norms and plasticity depends on how the environment varies in space and/or time (Scheiner 1993b; De Jong 1999) [add ref: Tufto 2015 Evolution, king & Hadfield 2019 Evol Lett], but an in-depth exploration of how to estimate these selection responses is beyond the scope of the present work.

Third, our general framework treats the curve-parameter and character-state approaches under the same umbrella, allowing evaluation of any chosen parametrical model through the goodness-of-fit parameter R^2_{mod}. This also opens the door to better commensurability and comparatibility across studies, which can be a challenge in meta-analyses of plasticity. Murren et al. (2014) performed such a meta-analysis, comparing genetic variation in different parameters of reaction norm shape across published datasets. However they (i) computed these parameters using only extreme environmental values, instead of the whole range of environments; (ii) did not account for uneven spacing between environments where relevant; (iii) did not account for uncertainty in estimations of reaction norms (as previously highlighted by Morrissey & Liefting 2016); and (iv) assumed the modeled reaction norm shape is true. More detail about the analyses in that study is provided in Appendix D. Our approach overcomes all these issues (some of which had been dealt with already by Morrissey & Liefting 2016). Unfortunately the dataset compiled by Murren et al. (2014) does not provide information on uncertainty of phenotypic estimates (related to V_{Res}), precluding proper meta-analysis of reaction norm shape variation.

Fourth and finally, our variance partitioning can be implemented through commonly used statistical models, notably linear mixed models. Furthermore, we showed that even complex non-linear modelling can perform well, only at the cost of using dedicated libraries to compute integrals numerically. This means that biologists can readily seize all the modelling tools introduced here. In particular, although a character-state approach can be performed using a simple random-intercept model, studies of genetic variance in plasticity seem to rather use a multi-trait model, which offers more control, but is more difficult to implement (but see Stirling & Roff 2000). In order to make the variance partitioning introduced here more accessible, we

provide a tutorial on how to use linear and non-linear modelling to analyse data at the following address: github.com/devillemereuil/TutoPartReacNorm. We have also implemented the computation of V_{plas}, V_{Gen} and V_A for non-linear models as a new feature of the QGglmm R package (de Villemereuil et al. 2016). We hope that this will further stimulate interest in investigating variation and evolutionary potential of reaction norms.

Code availability The code for the data simulation and analyses performed in this article is available at the following repository: github.com/devillemereuil/CodePartReacNorm

References

Angilletta, M. J. (2009) *Thermal adaptation: a theoretical and empirical synthesis*. OUP Oxford, Jan. 29, 2009. 304 pp.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67:(2015), 48.

Bonamour, S., Chevin, L.-M., Charmantier, A., & Teplitsky, C. (2019) Phenotypic plasticity in response to climate change: the importance of cue variation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374:(Mar. 18, 2019), 20180178. doi: 10.1098/rstb.2018.0178.

Bradshaw, A. D. (1965) Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics*. Ed. by E. W. Caspari & J. M. Thoday. Vol. 13. Cambridge (MA, USA): Academic Press, Jan. 1, 1965, pp. 115–155. doi: 10.1016/S0065-2660(08)60048-6.

Brown, G. G. & Rutemiller, H. C. (1977) Means and variances of stochastic vector products with applications to random linear models. *Management Science*, 24:(Oct. 1977), 210–216. doi: 10.1287/mnsc.24.2.210.

Bürkner, P.-C. (2017) Advanced bayesian multilevel modeling with the R package brms. *ArXiv170511123 Stat*:(May 31, 2017).

Charmantier, A., McCleery, R. H., Cole, L. R., Perrins, C., Kruuk, L. E. B., & Sheldon, B. C. (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, 320:(May 9, 2008), 800–803. doi: 10.1126/science.1157174.

Chevin, L.-M., Collins, S., & Lefèvre, F. (2013) Phenotypic plasticity and evolutionary demographic responses to climate change: taking theory out to the field. *Functional Ecology*, 27:(2013), 967–979. doi: 10.1111/j.1365-2435.2012.02043.x.

Chevin, L.-M., Lande, R., & Mace, G. M. (2010) Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLOS Biology*, 8:(Apr. 27, 2010), e1000357. doi: 10.1371/journal.pbio.1000357.

De Jong, G. (1990) Quantitative genetics of reaction norms. *Journal of evolutionary biology*, 3:(1990), 447–468.

De Jong, G. (1999) Unpredictable selection in a structured population leads to local genetic differentiation in evolved reaction norms. *Journal of Evolutionary Biology*, 12:(1999), 839–851.

de Jong, G. (1995) Phenotypic plasticity as a product of selection in a variable environment. *The American Naturalist*, 145:(Apr. 1, 1995), 493–512. doi: 10.1086/285752.

Des Marais, D. L., Hernandez, K. M., & Juenger, T. E. (2013) Genotype-by-environment interaction and plasticity: exploring genomic responses of plants to the abiotic environment. *Annual Review of Ecology, Evolution, and Systematics*, 44:(2013), 5–29. doi: 10.1146/annurev-ecolsys-110512-135806.

Deutsch, C. A., Tewksbury, J. J., Huey, R. B., Sheldon, K. S., Ghalambor, C. K., Haak, D. C., & Martin, P. R. (2008) Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences*, 105:(May 6, 2008), 6668–6672. doi: 10.1073/pnas.0709472105.

de Villemereuil, P., Schielzeth, H., Nakagawa, S., & Morrissey, M. B. (2016) General methods for evolutionary quantitative genetic inference from generalised mixed models. *Genetics*, 204:(Nov. 1, 2016), 1281–1294. doi: 10.1534/genetics.115.186536.

de Villemereuil, P. et al. (2020) Fluctuating optimum and temporally variable selection on breeding date in birds and mammals. *Proceedings of the National Academy of Sciences*, 117:(2020), 31969–31978. doi: 10.1073/pnas.2009003117.

Falconer, D. S. (1952) The problem of environment and selection. *The American Naturalist*, 86:(Sept. 1, 1952), 293–298. doi: 10.1086/281736.

Falconer, D. S. & Mackay, T. F. (1996) *Introduction to quantitative genetics*. 4th ed. Harlow, Essex (UK): Benjamin Cummings, Feb. 16, 1996.

Gavrilets, S. & Scheiner, S. M. (1993a) The genetics of phenotypic plasticity. V. Evolution of reaction norm shape. *Journal of Evolutionary Biology*, 6:(1993), 31–48. doi: 10.1046/j.1420-9101.1993.6010031.x.

Gavrilets, S. & Scheiner, S. M. (1993b) The genetics of phenotypic plasticity. VI. Theoretical predictions for directional selection. *Journal of Evolutionary Biology*, 6:(1993), 49–68.

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A., & Merilä, J. (2008) Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology*, 17:(Jan. 1, 2008), 167–178. doi: 10.1111/j.1365-294X.2007.03413.x.

Gomulkiewicz, R. & Kirkpatrick, M. (1992) Quantitative genetics and the evolution of reaction norms. *Evolution*, 46:(Apr. 1, 1992), 390–411. doi: 10.1111/j.1558-5646.1992.tb02047.x.

Hammill, E., Rogers, A., & Beckerman, A. P. (2008) Costs, benefits and the evolution of inducible defences: a case study with *Daphnia pulex*. *Journal of Evolutionary Biology*, 21:(May 1, 2008), 705–715. doi: 10.1111/j.1420-9101.2008.01520.x.

Kirkpatrick, M. & Heckman, N. (1989) A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *Journal of Mathematical Biology*, 27:(Aug. 1, 1989), 429–450. doi: 10.1007/BF00290638.

Lande, R. (1979) Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution*, 33:(1979), 402–416.

Lande, R. (2009) Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology*, 22:(July 1, 2009), 1435–1446. doi: 10.1111/j.1420-9101.2009.01754.x.

Landsman, Z. & Nešlehová, J. (2008) Stein's Lemma for elliptical random vectors. *Journal of Multivariate Analysis*, 99:(May 1, 2008), 912–927. doi: 10.1016/j.jmva.2007.05.006.

Landsman, Z., Vanduffel, S., & Yao, J. (2013) A note on Stein's lemma for multivariate elliptical distributions. *Journal of Statistical Planning and Inference*, 143:(Nov. 1, 2013), 2016–2022. doi: 10.1016/j.jspi.2013.06.003.

Lynch, M. & Walsh, B. (1998) *Genetics and analysis of quantitative traits*. Sunderland, Massachussets (US): Sinauer Associates, 1998.

Lynch, M. & Gabriel, W. (1987) Environmental tolerance. *The American Naturalist*, 129:(Feb. 1, 1987), 283–303. doi: 10.1086/284635.

Merilä, J. & Hendry, A. P. (2014) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications*, 7:(2014), 1–14. doi: 10.1111/eva.12137.

Mitchell, D. J. & Houslay, T. M. (2021) Context-dependent trait covariances: how plasticity shapes behavioral syndromes. *Behavioral Ecology*, 32:(Jan. 1, 2021), 25–29. doi: 10.1093/beheco/araa115.

Moczek & Emlen (1999) Proximate determination of male horn dimorphism in the beetle Onthophagus taurus (Coleoptera: Scarabaeidae). *Journal of Evolutionary Biology*, 12:(1999), 27–37. doi: 10.1046/j.1420-9101.1999.00004.x.

Morrissey, M. B. & Liefting, M. (2016) Variation in reaction norms: Statistical considerations and biological interpretation. *Evolution*, 70:(Sept. 1, 2016), 1944–1959. doi: 10.1111/evo.13003.

Murren, C. J., Maclean, H. J., Diamond, S. E., Steiner, U. K., Heskel, M. A., Handelsman, C. A., Ghalambor, C. K., Auld, J. R., Callahan, H. S., & Pfennig, D. W. (2014) Evolutionary change in continuous reaction norms. *The American Naturalist*, 183:(2014), 453–467.

Narasimhan, B., Johnson, S. G., Hahn, T., Bouvier, A., & Kiêu, K. (2023) *Cubature: Adaptive multivariate integration over hypercubes*. manual. 2023.

Nussey, D. H., Postma, E., Gienapp, P., & Visser, M. E. (2005) Selection on heritable phenotypic plasticity in a wild bird population. *Science*, 310:(Oct. 14, 2005), 304–306. doi: 10.1126/science.1117004.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & {the R Core team} (2009) *Nlme: Linear and Nonlinear Mixed Effects Models*. 2009.

Pletcher, S. D. & Geyer, C. J. (1999) The genetic analysis of age-dependent traits: modeling the character process. *Genetics*, 153:(Oct. 1, 1999), 825–835. doi: 10.1093/genetics/153.2.825.

Reed, T. E., Waples, R. S., Schindler, D. E., Hard, J. J., & Kinnison, M. T. (2010) Phenotypic plasticity and population viability: the importance of environmental predictability. *Proceedings of the Royal Society B: Biological Sciences*, 277:(Nov. 22, 2010), 3391–3400. doi: 10.1098/rspb.2010.0771.

Rovelli, G. et al. (2020) The genetics of phenotypic plasticity in livestock in the era of climate change: a review. *Italian Journal of Animal Science*, 19:(Dec. 14, 2020), 997–1014. doi: 10.1080/1828051X.2020.1809540.

Schaum, C. E. & Collins, S. (2014) Plasticity predicts evolution in a marine alga. *Proceedings of the Royal Society B: Biological Sciences*, 281:(Oct. 22, 2014), 20141486. doi: 10.1098/rspb.2014.1486.

Scheiner, S. M. (1993a) Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology and Systematics*, 24:(Nov. 1993), 35–68. doi: 10.1146/annurev.es.24.110193.000343.

Scheiner, S. M. (1993b) Plasticity as a selectable trait: reply to Via. *The American Naturalist*, 142:(Aug. 1, 1993), 371–373. doi: 10.1086/285544.

Scheiner, S. M. & Lyman, R. F. (1989) The genetics of phenotypic plasticity I. Heritability. *Journal of Evolutionary Biology*, 2:(Mar. 1989), 95–107. doi: 10.1046/j.1420-9101.1989.2020095.x.

Schlichting, C. D. & Pigliucci, M. (1998) Phenotypic evolution: a reaction norm perspective. *Phenotypic evolution: a reaction norm perspective.*:(1998).

Stinchcombe, J. R., Function-valued Traits Working Group, & Kirkpatrick, M. (2012) Genetics and evolution of function-valued traits: understanding environmentally responsive phenotypes. *Trends in Ecology & Evolution*, 27:(Nov. 1, 2012), 637–647. doi: 10.1016/j.tree.2012.07.002.

Stirling, G. & Roff, D. A. (2000) Behaviour plasticity without learning: phenotypic and genetic variation of naïve *Daphnia* in an ecological trade-off. *Animal Behaviour*, 59:(May 1, 2000), 929–941. doi: 10.1006/anbe.1999.1386.

Suzuki, Y. & Nijhout, H. F. (2006) Evolution of a polyphenism by genetic accommodation. *Science*, 311:(Feb. 3, 2006), 650–652. doi: 10.1126/science.1118888.

Teplitsky, C., Mills, J. A., Alho, J. S., Yarrall, J. W., & Merilä, J. (2008) Bergmann's rule and climate change revisited: Disentangling environmental and genetic responses in a wild bird population. *Proceedings of the National Academy of Sciences*, 105:(Sept. 9, 2008), 13492–13496. doi: 10.1073/pnas.0800999105.

Tufto, J. (2000) The evolution of plasticity and nonplastic spatial and temporal adaptations in the presence of imperfect environmental cues. *The American Naturalist*, 156:(Aug. 1, 2000), 121–130. doi: 10.1086/303381.

572 Vedder, O., Bouwhuis, S., & Sheldon, B. C. (2013) Quantitative assessment of the importance of phenotypic
573 plasticity in adaptation to climate change in wild bird populations. *PLOS Biology*, 11:(2013), e1001605. doi:
574 10.1371/journal.pbio.1001605.
575 Via, S. & Lande, R. (1985) Genotype-environment interaction and the evolution of phenotypic plasticity. *Evo-*
576 *lution*, 39:(May 1, 1985), 505–522. doi: 10.1111/j.1558-5646.1985.tb00391.x.
577 Woltereck, R. (1909) Weitere experimentelle Unter-suchungen uber Artveranderung, speziell uber das Wesen
578 quantitativer Artunterschiede bei Daphniden. *Verh. D. Tsch. Zool. Ges.*, 1909:(1909), 110–172.

Appendix

A Comparison between alternative variance partitionings

A schematic example

To illustrate the difference between the variance partitioning in Equation 6 and the 'classical' variance partitioning between V_G, V_E and $V_{G\times E}$, we will first consider a very schematic example. Let us consider two scenarios with 3 genotypes in 2 environments. For now, we will consider the environmental variable is mean-centered, so that the zero for the environment is exactly at mid-value between the two environments. In the first scenario, all of the reaction norms are parallel between each genotype, such that this is a typical case where there is no genotype-by-environment interaction (Figure S1, left panel). In the second scenario, we invert the values of the most extremes genotypes in the second environment, so that the reaction norms are now crossing with considerable genotype-by-environment interaction (Figure S1, right panel). An interesting feature of such scenarios is that, since we only reassigned values to different genotypes, we conserved the genetic variance within each environments. Note that the reaction norms are directly considered here, so that V_{Res} is ignored in this section. Also, in the second scenario, since all reaction norms cross exactly at the mid-point between environments, there is no variation in the intercept.

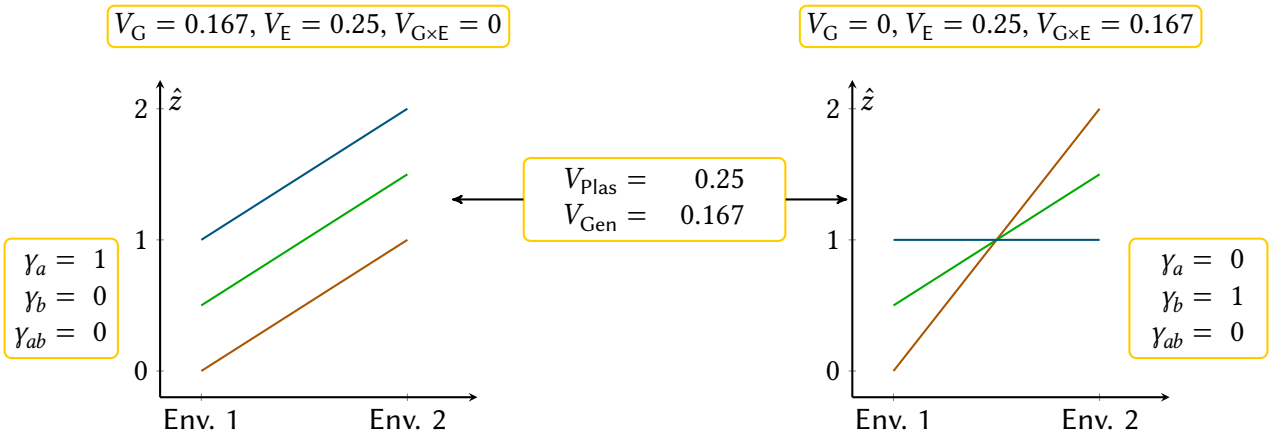

Figure S1: Two different scenarios with the same total variance. On the left: all reaction norms are parallel, so that $V_{G\times E} = 0$, by definition. On the right, the two extreme values on the second environment were switched, resulting in the crossing of reaction norms and thus substantial $V_{G\times E}$, at the full expanse of V_G. Our variance partition in V_{Plas} and V_{Gen} is equal in both scenarios, however, the γ-decomposition (where a stands for the intercept and b for the slope) of the genetic variance V_{Gen} is completely different, reflecting the (co)variation of the intercept and slope of reaction norms on the second scenario (right).

In the first scenario, since all reaction norms are parallel we have $V_{G\times E} = 0$, and there is a perfect correspondence between terms in both partitionings, with that $V_{Plas} = V_E$ and $V_{Gen} = V_G$ (Figure S1, left and

center). All of the genetic variance in the trait comes from variation in the intercept of reaction norms, which is reflected by the γ-decomposition from Equation 21 (Figure S1, left).

In contrast in the second scenario, all genotypes have the same mean phenotype averaged across environments, leading to $V_{\mathrm{G}} = 0$ in the classical partitioning. However, V_{Gen} is not zero, and is in fact exactly equal to that in the first scenario, in this example. In other words, both scenarios lead to the same amount of genetic variation available for responding to selection across the two environments where phenotypes have been measured. The only thing that differs between these scenarios is the constraints they impose on evolution of reaction norms. Scenario 1 facilitates responses to phenotypic selection that goes in the same direction in both environments, while scenario 2 facilitates responses to selection in opposite directions across environments. Although the value for V_{Gen} is unchanged, these constraints are adequately reflected by the γ-decomposition of V_{Gen}, for which we now have $\gamma_a = 0$ and $\gamma_b = 1$. Note that in this scenario, we instead have $V_{\mathrm{Plas}} = V_{\mathrm{E}}$ and $V_{\mathrm{Gen}} = V_{\mathrm{G\times E}}$.

As a final note on this example, let us imagine that, instead of choosing the mid-point between environments as reference (set to zero), we choose the first environment. In this case, the intercept is defined in this first environment, and there is now considerable variation in the intercept. Such arbitrary choice has no impact on the values of neither V_{Plas} and V_{Gen}, nor on V_{G}, V_{E} and $V_{\mathrm{G\times E}}$. However, this new definition of the intercept and its variation leads to a different γ-decomposition: $\gamma_a = 1$, $\gamma_b = 4$ and $\gamma_{ab} = -4$. In other words, redefining the zero in the scale of the environment changed the definition of the parameter "intercept", and made apparent the negative genetic correlation between the intercept and slope (a perfect one in this scenario), whereby steeper negative slopes are associated with higher intercept (phenotype in environment 1). Nevertheless, the evolutionary dynamics are not sensitive to the arbitrary choice of a zero in the environmental scale, as the distribution of genetic variation along environments is the same in both versions of the second scenario.

General comparison

This example illustrates how our variance partitioning differs from the classical one with genotype, environment, and genotype-by-environment interaction effects.

In particular, there is no distinction between V_{G} and $V_{\mathrm{G\times E}}$ in our partitioning, as $V_{\mathrm{Gen}} = V_{\mathrm{G}} + V_{\mathrm{G\times E}}$. This is due to our use of the total variance, which integrates over genotypes in each environment, before integrating over environments. However, this does not mean that our framework is degenerate and looses information on how genetic variance is distributed across environments, and how this constrains evolution of reaction norm shape. Instead, these aspects are captured by two things. The first is the γ-decomposition in Equation 21, which provides an explicit measure of genetic variation in different components of reaction norm shape. The

second is the environment-specific amount of genetic variance, as detailed in our worked example of non-linear reaction norm models (Figure 5).

Regarding the environmental variance V_E and V_{Plas}, there were considered equal on our example, but this is because we considered directly the reaction norms, and thus ignored V_{Res}. In many contexts, we can consider V_{Plas} and V_{Res} as respectively measuring the *general* (environment shared by a group of individuals) and *specific* (environment specific to an individual) environmental variance as defined by Falconer (Falconer & Mackay 1996; Lynch & Walsh 1998). A complication is that, in reality, V_{Plas} is not defined relative to groups of individuals (or genotype), but rather to a singled-out environmental variable. In that regard, V_{Res} contains the part of what could be considered the general environment, which results from the influence of other environmental variable. In any case, if no distinction is made between general and specific environment components in V_E and the phenotypic trait is under consideration rather than reaction norms themselves, then we can write $V_E = V_{Plas} + V_{Res}$.

B Computation of the additive genetic variance

Multiple regression from variance-covariance matrix Let us assume a multiple regression between a random variable y and a series a random variables $\mathrm{x} = (x_1, \ldots, x_n)$ such that:

$$y = \mu + \mathrm{x}^T \beta + e, \tag{S1}$$

where μ is the intercept and e is the residual of the model. Note that in practical regression, the realised sampling of x will be contained in the design matrix of the model. If it exists and is unique, the solution for β can be formulated in terms variance-covariance matrices (see e.g. p.179, Lynch & Walsh 1998):

$$\beta = \mathrm{V}(\mathrm{x})^{-1} \mathrm{cov}(\mathrm{x}, y), \tag{S2}$$

where $\mathrm{V}(\mathrm{x})$ is the variance-covariance matrix of x and $\mathrm{cov}(\mathrm{x}, y)$ is the column-vector of covariances between the x_i and y.

Multivariate version of Stein's lemma Let us assume that $\mathrm{y} = (x_1, \ldots, x_{p_y})$ follows a multivariate normal distribution, that $\mathrm{x} = (x_1, \ldots, x_{p_x})$ follows a multivariate normal distribution and that g is a differentiable, $\mathbb{R}^{p_x} \to \mathbb{R}$ function such that $\mathrm{E}(\nabla g)$, where ∇g is the gradient of g (the vector of partial differentials), is a vector of finite values, then (Landsman & Nešlehová 2008; Landsman et al. 2013):

$$\mathrm{cov}\left(g(\mathrm{x}), \mathrm{y}\right) = \mathrm{cov}(\mathrm{x}, \mathrm{y}) \mathrm{E}\left(\nabla g\right). \tag{S3}$$

In the case where $p_y = 1$, then $\mathrm{y} = y$ follows a normal distribution and:

$$\mathrm{cov}\left(g(\mathrm{x}), y\right) = \mathrm{cov}(y, \mathrm{x})\mathrm{E}\left(\nabla g\right). \tag{S4}$$

Note that $\mathrm{cov}(y, \mathrm{x})$ is a row-vector and $\mathrm{cov}(\mathrm{x}, y)$ is a column-vector by convention.

Linear relationship between breeding values The additive genetic variance V_A is the variance of the breeding values a_z of the phenotypic trait z. Let us note $a_{\theta,i}$ as the breeding value of the parameter θ_i. Here, we will assume that we are working within a given (and fixed) environment ε. We will follow the same demonstration as in de Villemereuil et al. (2016), which starts from the point that, by definition, breeding values are linked through a linear relationship. More precisely, the breeding value the trait a_z of an individual linearily depends on a linear combination of the breeding values of the parameters $a_{\theta,i}$ of the same individual, so that:

$$z = a_z + e = \mu_a + \boldsymbol{a}_{\boldsymbol{\theta}}^T \boldsymbol{\psi} + e \tag{S5}$$

where e is the residual variance of the regression (assumed independent of the breeding values), $\boldsymbol{\psi}$ is a vector containing the slopes and $\boldsymbol{a}_{\boldsymbol{\theta}}$ is a vector containing the breeding values for all parameters of the reaction norm.

Defining the value of $\boldsymbol{\psi}$ To compute the value of $\boldsymbol{\psi}$, we can solve the linear equation in Equation S5 using Equation S2:

$$\boldsymbol{\psi} = \Theta^{-1}\mathrm{cov}(\boldsymbol{a}_{\boldsymbol{\theta}}, z) \tag{S6}$$

Noting that $z = f(\varepsilon, \boldsymbol{\theta})$, we can apply the multivariate version of Stein's lemma (Equation S3):

$$\boldsymbol{\psi} = \Theta^{-1}\mathrm{cov}(\boldsymbol{a}_{\boldsymbol{\theta}}, \boldsymbol{\theta})\mathrm{E}(\nabla_\theta f) = \Theta^{-1}\Theta\mathrm{E}(\nabla_\theta f) = \mathrm{E}(\nabla_\theta f). \tag{S7}$$

Additive genetic variance From Equation S5, the additive genetic variance of the trait V_A is given by:

$$V_\mathrm{A} = \mathrm{V}(\boldsymbol{a}_{\boldsymbol{\theta}}^T \boldsymbol{\psi}) = \boldsymbol{\psi}^T \Theta \boldsymbol{\psi}. \tag{S8}$$

We worked at a given environment, and to reflect this, these quantities are named $\boldsymbol{\psi}_\varepsilon$ and $V_{\mathrm{A}|\varepsilon}$ in the main text.

C Correcting for uncertainty in the estimation of fixed effects

Character-state approach It is easier to start with the character-state approach and the ANOVA model it is based on. We want to compute V_{Plas} as the variance of the group-level effects μ (see Equation 2 and Equation 7 in the main text) :

$$V_{\text{Plas}} = \text{V}(\mu) \tag{S9}$$

However, we do not have access to the real-world values for μ, instead, we have access to the estimated $\hat{\mu}$ from the model. Such estimates, if unbiased, have an expected value of μ_k at environment k and a standard-error (i.e. the estimation of the sampling standard deviation) s_k. In other words, we can state that $\hat{\mu_k}$ is equal to μ_k up to an additive error:

$$\hat{\mu_k} = \mu_k + \tilde{\mu_k} \tag{S10}$$

where $\tilde{\mu}$ is of mean 0 and variance s_k^2. Considering each sampling r, we can apply the law of total variance, although in a different context than in the main text:

$$\text{V}(\hat{\mu}) = \text{V}_\varepsilon(\text{E}_{r|\varepsilon}(\hat{\mu})) + \text{E}_\varepsilon(\text{V}_{r|\varepsilon}(\hat{\mu})) = \text{V}_\varepsilon(\mu) + \text{E}_\varepsilon(s^2). \tag{S11}$$

We thus have:

$$V_{\text{Plas}} = \text{V}_\varepsilon(\mu) = \text{V}_\varepsilon(\hat{\mu}) - \text{E}_\varepsilon(s^2) \tag{S12}$$

This result is equivalent to e.g. the classical computation of the "sire variance" in sire models in quantitative genetics (Lynch & Walsh 1998), although this later is generally expressed using sums-of-squares.

Parameter curve approach There is unfortunately no simple solution to the problem of accounting for the uncertainty of fixed effects in the general context of non-linear modelling. However, for the particular case where the model can be framed as a linear model, as is the case for the polynomial function (see Equation 13, $\hat{z} = \text{X}\boldsymbol{\theta}$). In this case, we can define V_{Plas} as (Equation 14):

$$V_{\text{Plas}} = \text{V}(\text{x}^T\bar{\boldsymbol{\theta}}) = \bar{\boldsymbol{\theta}}^T\text{X}\bar{\boldsymbol{\theta}}. \tag{S13}$$

Again, the problem is that $\boldsymbol{\theta}$ is unknown, we only have access to the estimated values of the parameters, $\hat{\boldsymbol{\theta}}$, that are inferred with an error provided by the variance-covariance matrix of standard errors, S_θ. We can

write again:

$$\hat{\boldsymbol{\theta}} = \bar{\boldsymbol{\theta}} + \tilde{\boldsymbol{\theta}}, \tag{S14}$$

where $\tilde{\boldsymbol{\theta}}$ has a null mean and a variance-covariance matrix S_θ. Noting that the error is independent from the true value, we have:

$$\mathrm{V}(\mathrm{x}^T\hat{\boldsymbol{\theta}}) = \hat{\boldsymbol{\theta}}^T\mathrm{X}\hat{\boldsymbol{\theta}} = \mathrm{V}(\mathrm{x}^T\bar{\boldsymbol{\theta}}) + \mathrm{V}(\mathrm{x}^T\tilde{\boldsymbol{\theta}}), \tag{S15}$$

To express the variance $\mathrm{V}(\mathrm{X}\tilde{\boldsymbol{\theta}})$, it is important to note that $S_{\theta,ij} = \mathrm{E}(\tilde{\theta}_i\tilde{\theta}_j)$, since $\mathrm{E}(\tilde{\boldsymbol{\theta}}) = 0$. Then, we can note that, the error being unknown, we actually want to compute $\mathrm{E}_r(\mathrm{V}(\mathrm{x}^T\tilde{\boldsymbol{\theta}}))$ taken across all possible sampling r:

$$\mathrm{E}_r(\mathrm{V}(\mathrm{x}^T\tilde{\boldsymbol{\theta}})) = \mathrm{E}_r(\tilde{\boldsymbol{\theta}}^T\mathrm{X}\tilde{\boldsymbol{\theta}}) = \mathrm{E}_r(\sum_{ij}\tilde{\theta}_i\tilde{\theta}_j X_{i,j}) = \sum_{ij}\mathrm{E}_r(\tilde{\theta}_i\tilde{\theta}_j)X_{i,j} = \sum_{ij}S_{\theta,ij}V_{\mathrm{X},ij} = \mathrm{Tr}(\mathrm{S}_\theta\mathrm{X}) \tag{S16}$$

This is similar to the result of Brown & Rutemiller (1977). Finally, we have proven Equation 26:

$$V_{\mathrm{Plas}} = \hat{\boldsymbol{\theta}}^T\mathrm{X}\hat{\boldsymbol{\theta}} - \mathrm{Tr}(\mathrm{S}_\theta\mathrm{X}). \tag{S17}$$

D Comparison with the approach from Murren *et al.* (2014)

The first step in the approach of Murren et al. (2014) is to choose a reference reaction norm in each of the studies and compute contrasts to that particular reaction norm. The contrasts are then analysed, rather than the norms themselves. For the sake of simplicity, and because this does not (or marginally) impact our comments on this approach, we will overlook that step and consider reaction norms directly.

For each genotype k and from its given reaction norm (or contrast) $\mathrm{z}_k = \{z_{k,1}, \ldots, z_{k,n}\}$, Murren et al. (2014) compute four statistics (we removed the absolute values for the sake of simplicity here):

1. The offset, O_{M}, measures the "location" of the reaction norm, i.e. its mean. Comparison of the offsets allows detecting wether reaction norms are "shifted" toward higher or lower values. It is computed, for each genotype k, as the absolute value of the average of the norm across environments:

$$O_{\mathrm{M},k} = \frac{\sum_i^n z_{k,i}}{n}. \tag{S18}$$

2. The slope, S_{M}, measures the linear trend of the norms. Formally, it is the absolute sum of the differences between two consecutive environments, divided by the number of intervals $(n-1)$:

$$S_{\mathrm{M},k} = \frac{\sum_i^{n-1} z_{k,i+1} - z_{k,i}}{n-1}. \tag{S19}$$

3. The curvature, C_{M}, is computed as the absolute value of the average change in norms between two consecutive couples of environments:

$$C_{\mathrm{M},k} = \frac{\sum_i^{n-2}(z_{k,i+2} - z_{k,i+1}) - (z_{k,i+1} - z_{k,i})}{n-2}. \tag{S20}$$

4. The wiggle, W_{M}, is, according to the authors the "the variability in shape not described by any of the previous three measures":

$$W_{\mathrm{M},k} = \frac{\sum_i^{n-2}\left|(z_{k,i+2} - z_{k,i+1}) - (z_{k,i+1} - z_{k,i})\right|}{n-2} - C_{\mathrm{M},k}. \tag{S21}$$

   Given the lower interest in this statistics, we will not comment on it any further. Most of the comments on the other statistics also apply to this one.

One strong assumption underlying the calculations above is that environmental values $\mathrm{x} = \{x_1, \ldots, x_n\}$ on which the reaction norms were evaluated are evenly spaced, e.g. that the differences $x_{i+1} - x_i$ are equal for all possible values of i. More, this calculation assumes that the space between two measures is equal to 1 (which, admittedly, is only a matter of rescaling when evenly-spaced values are already assumed). If this is case, then there is indeed no loss in generality in using the number of components (n, $n-1$ and $n-2$) rather than actual values of x in the denominator. Although it is common for studies on reaction norms to use evenly-spaced environmental values, it is an unnecessary assumption that shall not be satisfied by all studies.

Another issue does not specifically stems from assumptions underlying the estimators, but rather from the fact that these estimators are applied to the estimated values themselves, rather than on a fitted function for the reaction norms. Indeed, developing the sums in S_{M} and C_{M} above show that the intermediate values cancel each other out, leaving only the values at each extreme of the environmental range in the estimate:

$$\begin{aligned} S_{\mathrm{M},k} &= \frac{z_{k,n} - z_{k,1}}{n-1}, \\ C_{\mathrm{M},k} &= \frac{(z_{k,n} - z_{k,n-1}) - (z_{k,2} - z_{k,1})}{n-2}. \end{aligned} \tag{S22}$$

The issue here is double. First, the estimation is highly sensitive to the random noise coming from a small number of values (two or three/four). Second, the intermediate values in the reaction norm are simply thrown out and not used for a more robust estimation. In other words, it would have been exactly the same to not measure the reaction norm at these intermediate values, since they are not accounted for in the calculation.

A final issue, closely related to the second one, is that using the measured values of the reaction norms without accounting for the uncertainty in their estimation (i.e. standard-deviation and sample size for each genotype and environmental value) poses the well-known issue of non-propagation of the error when doing "statistics

on statistics".

Although we also provide estimators of the impact of the intercept, slope and curvature of reaction norms on the phenotypic variation, our approach differs from the one from Murren *et al.* (2014) by many aspects. First, using the law of total variance, we make the explicit distinction between the average shape of the reaction norm and the genetic variance surrounding it. As such, to O_M, S_M and C_M corresponds not only the genetic component r^2_{ga}, r^2_{gb} and r^2_{gc}, but also the average plasticity components (r^2_{pb} and r^2_{pc}). We also account for possible genetic correlation between components. Second, we use the whole of the statistical inference to define our estimates of contribution of intercept, slope and curvature to the phenotypic variance. Third, we explicitly account for the uncertain estimation of reaction norms.