Check for updates

Journal of Chemometrics

WILEY

Journal of CHEMOMETRICS

RESEARCH ARTICLE OPEN ACCESS

Population Power Curves in ASCA With Permutation Testing

José Camacho | Michael Sorochan Armstrong

Research Centre for Information and Communication Technologies (CITIC-UGR), University of Granada, Granada, Spain

Correspondence: José Camacho (josecamacho@ugr.es)

Received: 27 February 2024 | **Revised:** 24 July 2024 | **Accepted:** 26 July 2024

Funding: This work was supported by the Agencia Estatal de Investigación in Spain, MCIN/AEI/10.13039/501100011033, grant no. PID2020-113462RB-I00. Michael Sorochan Armstrong has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 101106986. Funding for open access charge: Universidad de Granada/CBUA.

Keywords: ANOVA simultaneous component analysis | effect size | multivariate ANOVA | power curves | sample size

ABSTRACT

In this paper, we revisit the power curves in ANOVA simultaneous component analysis (ASCA) based on permutation testing and introduce the population curves derived from population parameters describing the relative effect among factors and interactions. The relative effect has important practical implications: The statistical power of a given factor depends on the design of other factors in the experiment and not only of the sample size. Thus, understanding the relative power in a specific experimental design can be extremely useful to maximize our capability of success when planning the experiment. In the paper, we derive relative and absolute population curves, where the former represent statistical power in terms of the normalized effect size between structure and noise, and the latter in terms of the sample size. Both types of population curves allow us to make decisions regarding the number and nature (fixed/random) of factors, their relationships (crossed/nested), and the number of levels and replicates, among others, in an multivariate experimental design (e.g., an omics study) during the planning phase of the experiment. We illustrate both types of curves through simulation.

1 | Introduction

Prof. Smilde's research group proposed the ANOVA simultaneous component analysis (ASCA) [1], a powerful framework for analyzing the individual influence of different experimental factors and their interactions in experimental designs with a high number of responses. ASCA represents a natural multivariate extension of the analysis of variance (ANOVA). The theory associated to ANOVA is extensive [2], and several of its developments have been incrementally incorporated into the ASCA framework throughout the years [3–5]. Still, ASCA can be considered a developing technique, and many unresolved questions remain regarding best practices [6, 7].

One such questions is how to derive a power analysis in the context of ASCA. Statistical power is a relevant concept in inferential statistics. The power of a test measures the probability that it correctly rejects the null hypothesis (H_0) when the alternative hypothesis (H_1) is true. The power can be defined as $1-\beta$ for β the Type II error (false negative) probability. Power curves are a form of power analysis where power is represented in terms of the sample size, that is, the number of replicates or experimental runs under the same conditions. This is a recommended analysis prior to any multivariate experiment (e.g., a clinical study with omics responses) to determine the required number of subjects and experimental levels for a desired statistical power.

In standard univariate ANOVA, or when relatively few responses are considered, it is possible to use analytical methods to derive power curves based on typical assumptions (such as normality) [8]. However, in the context of more than just a few responses, or to assess the violation of any possible statistical assumptions,

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Journal of Chemometrics* published by John Wiley & Sons Ltd.

numerical methods [9] are a viable alternative. Arguably, the most popular approach for statistical inference in ASCA is permutation testing [9, 10]: a resampling method that transforms ASCA into a distribution-free approach that is more flexible than parametric ANOVA, where inference is based on analytical distributions. The derivation of power curves based on ASCA's permutation testing necessarily requires numerical approaches.

In previous work [6], we introduced the simulated power curves in ASCA as a strategy to optimize an ASCA pipeline for a specific experimental design in terms of statistical power. Simulated power curves can be used to find the optimal ASCA model in terms of fixed/random factors, crossed/nested relationships, interactions, test statistic, transformations, and others. Our approach was defined to compare several models in terms of relative power, but it cannot be used to make sample size estimations for an entirely new multivariate experiment. In this paper, we revisit the simulated power curve approach to generalize it. In particular, we make the following contributions:

- We generalize ASCA power curves so that any design can be simulated, including complex relationships among factors and their interactions.
- We define the population curves, derived from population parameters (standard deviations) describing the relative effect among factors and interactions.
- We propose two types of population curves:
  - Relative population curves (RPCs) represent statistical power in terms of the relative effect size between structure and noise. They are useful to optimize the ASCA pipeline for an analysis at hand.
  - Absolute population curves (APCs) represent statistical power in terms of the sample size. They are useful to plan ahead the number of replicates and/or levels to use in a designed study.
- We illustrate the behavior of the two types of population curves through simulation.
- We provide open software for the generation of population curves and for the replication of the results in this paper. RPCs and APCs can be computed with "powercurve" routine in the MEDA Toolbox stable release v1.4.[†]

The rest of the paper is organized as follows: Section 2 introduces the ASCA framework. Section 3 discusses the concept of power curves from a theoretical perspective. Sections 4 and 5 present the algorithms to compute RPCs and APCs, respectively. Section 6 discusses simulation results, and Section 7 draws the conclusions of the work.

2 | ANOVA Simultaneous Component Analysis

ASCA, like ANOVA, is mostly concerned with the analysis of data coming from an experimental design. Following ANOVA, a common ASCA pipeline follows three steps: (1) factorization of the data according to the experimental design; (2) significance testing for factors and interactions; (3) visualization of significant factors' and interactions' effects using principal component analysis (PCA) to understand separability among levels.

2.1 | Factorization of the Data

Let $\mathbf{X}$ be an $N \times M$ data matrix with N the number of experimental runs and M the number of responses in a designed experiment. Without loss of generality, we consider here the case of a design with two crossed factors A and B, their interaction, and an additional factor $C(A)$ nested in A. This is a common configuration in multiple omics experiments [4, 5, 11], useful to correct for the (often large) individual variability ($C(A)$) and so increase the statistical power of the test. The data in $\mathbf{X}$ can be decomposed as

$$\mathbf{X} = \mathbf{1}\mathbf{m}^{\mathrm{T}} + \mathbf{X}_A + \mathbf{X}_B + \mathbf{X}_{C(A)} + \mathbf{X}_{AB} + \mathbf{E} \tag{1}$$

where $\mathbf{1}$ is a vector of ones ($N \times 1$), $\mathbf{m}$ ($M \times 1$) denotes a vector containing the intercepts of the M measured variables, $\mathbf{X}_A$, $\mathbf{X}_B$ and $\mathbf{X}_{C(A)}$ represent the factor matrices, $\mathbf{X}_{AB}$ is the interaction matrix, and $\mathbf{E}$ is the residual matrix, all of similar dimensions of $\mathbf{X}$.

To compute the factorization, we use the technique referred to as ASCA+ [3] to account for mild unbalancedness in the data. We intentionally avoid ASCA alternatives based on linear mixed models (LMM) [4, 5] due to the large increase of computational demand, prohibitive in the context of power curves. In ASCA+, the decomposition is derived as the least squares solution of a regression problem, where $\mathbf{X}$ is regressed onto a coding matrix $\mathbf{D}$ as

$$\mathbf{X} = \mathbf{D}\Theta + \mathbf{E} = \mathbf{1}\theta + \mathbf{D}_A\Theta_A + \mathbf{D}_B\Theta_B + \mathbf{D}_{C(A)}\Theta_{C(A)} + \mathbf{D}_{AB}\Theta_{AB} + \mathbf{E} \tag{2}$$

and $\mathbf{D}$ is constructed using sum coding [3] or another alternative coding schemes [5] and Θ and $\mathbf{E}$ are obtained from

$$\Theta = (\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}\mathbf{D}\mathbf{X} \tag{3}$$

$$\mathbf{E} = \mathbf{X} - \mathbf{D}\Theta \tag{4}$$

The encoding of experimental factors in the design matrix is an especially important consideration, as it affects what information is passed through to variance explained by the model, versus residual variance—this in turn affects the apparent evidence for statistical significance.

2.2 | Statistical Significance Testing

In this step, we test the statistical significance of factors and interactions in a similar way as performed in multi-way ANOVA. A widely used approach for ASCA inference is permutation testing [9, 10]. Permutation testing can be performed by randomly shuffling the rows of $\mathbf{X}$ in Equation (3), yielding a new set of regression coefficients, and so a new factorization:

$$\Theta^* = (\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{X}^* \tag{5}$$

$$\mathbf{E}^* = \mathbf{X}^* - \mathbf{D}\Theta^* \tag{6}$$

Permutation tests are carried out by comparing a given statistic, computed after the ASCA factorization, with the corresponding statistic computed from hundreds or more permutations of the rows in the observational data, $\mathbf{X}$. The p value is obtained as[‡]

$$p = \frac{\#\left\{S_k^* \geq S; k = 1, \ \ldots, K\right\} + 1}{K + 1} \tag{7}$$

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where S refers to the statistic computed from the true factorization, S_k^* is the statistic corresponding to the k-th random permutation, $\#\{cond\}$ refers to the number of times condition $cond$ is met, and K is the total number of permutations. See [6] for a recent review on the permutation approach and the relevance of the chosen statistic.

The permutation approach in the rows of $\mathbf{X}$ allows to test all the factors and interactions at the same time. Alternatively, one may be interested in testing an individual factor/interaction, which can be done by permuting the corresponding coding rather than the data [6] or by orthogonalizing the data prior to permutation. Additional considerations may be taken into account for imbalanced designs [12].

2.3 | Post Hoc Visualization

Significance testing in ANOVA/ASCA reveals the statistical significance of factors' and interactions' effects, but not of the specific levels (or combination thereof). To identify significant differences across levels, post hoc tests are typically employed. The equivalent to a post-hoc visualization in ASCA is the use of PCA following Zwanenburg et al. [13], where the PCA loadings are computed from the factor/interaction matrix, and score plots are built from the sum of this matrix and the residuals. Following this approach, score plots provide a visual comparison between the main/interaction effects and the natural variability in the residuals. Other approaches besides PCA can also be used [14].

ASCA is a supervised method, and as such, it can suffer of "overfitting" like any other regression and/or classification approaches [15]. ASCA relies in permutation testing to avoid overfitting. Thus, only statistically significant factors/interactions should be visualized post hoc with PCA.

3 | Power Curves

Figure 1 illustrates the concept of a power curve as a comparison between two distributions: the null distribution $P(t|H_0)$ and the alternative distribution $P(t|H_1)$. In the figure, both distributions are assumed to be normal, but in a real situation, this may not be the case. The shaded areas represent the probability of error, associated with an observation incorrectly rejecting the null hypothesis through random chance (type I error), with probability α (in blue), versus the probability of an observation incorrectly not rejecting the null hypothesis (type II error), with probability β (in red). Larger effect sizes θ, manifesting as wider separations between both distributions, increase the statistical power $1-\beta$ (reducing the type II error) for a fixed value of α. Alternatively, for both fixed α and the effect size θ, increasing the sampling size reduces the variance of the distributions $P(t|H_0)$ and $P(t|H_1)$, which leads to a reduction of overlapping areas as $1-\beta$ increases.

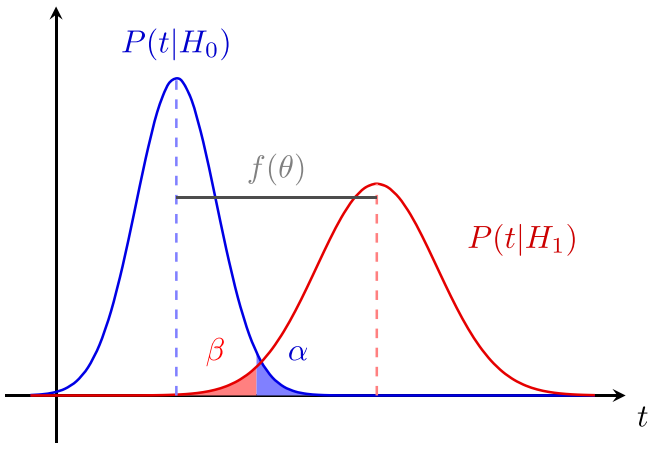


FIGURE 1 | Illustration of effect size ($f(\theta)$) versus the probability of a type-I error, α, and the probability of a type II error, β. Statistical power is defined as 1 − the false negative (type-II) error rate.

4 | RPCs for ASCA

RPCs are derived from population parameters describing the standard deviation in factors and interactions and represent statistical power in terms of the relative effect size between structure and noise.

To generate RPCs, we generalize the approach by Camacho et al. [6]. We simulate data that progressively increases the relative effect size θ and therefore the power (Figure 1). In the following, we use model (1) to showcase the RPCs, since this model includes crossed relationships, interactions (between A and B) and nested relationships (between A and $C(A)$), so that almost any other model can be derived from it. We can also think of this model as an illustration of factors/interactions organized in different orders (Figure 2), as discussed by Anderson and Ter Braak [9] and in the Hasse diagrams by Marini et al. [16].

Our approach to generate RPCs follows these steps:

0. INPUT:
   - $\mathbf{F}$ the design matrix
   - H a model hierarchical structure (as illustrated by a Hassel diagram) with information about factors, crossed/nested relationships and their interactions. In this model structure, a factor f can have descendants (nested factors and associated interactions) and ancestors (factors in which f is nested). An interaction i can

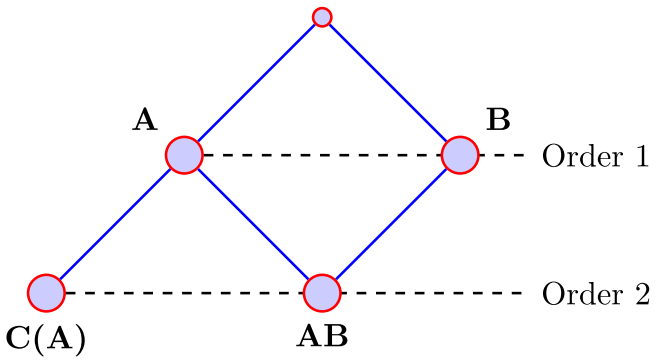


FIGURE 2 | Ordering structure among factors and interaction in model (1).

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

also have descendants (other higher-order interactions) and ancestors (other lower-order interactions or the factors included in the interaction). In model (1) and Figure 2, A is ancestor of $C(A)$ and A and B are ancestors of AB.
- M the number of responses
- k_f for $f \in \{1 \dots F\}$, k_i for $i \in \{1 \dots I\}$, and k_e the coefficients with the standard deviation[§] for factors and the interactions, and the residuals
- R the number of repetitions to generate a statistically representative RPC
- P the number of permutations (in each repetition)
- δ incremental steps in the effect size
- α the imposed probability of falsely rejecting the null hypothesis (i.e., the significance level)

1. Set $power_f(\theta) = 0$ and $power_i(\theta) = 0$ for each factor f and interaction i, respectively, and for θ from 0 to 10δ in δ steps; and N is set to the number of rows of $\mathbf{F}$.

2. For each repetition from 1 to R
   2.1. Generate random matrices to represent level/cell averages
      2.1.1. For each factor f_1 of order 1 in H, and so with no ancestors (e.g., A and B in model (1)), count the number of levels L_{f_1} in matrix $\mathbf{F}$ and simulate:

$$\overline{\mathbf{X}}_{f_1}(L_{f_1}, M) \sim \mathcal{D}_{f_1} \tag{8}$$

      where $\mathcal{D}_{f_1}$ is a pseudo-random number generator (PRNG), potentially based on the normal distribution or other distribution that may deviate moderately (e.g., uniform) or severely (e.g., $\exp^3$) from normality [9].
      2.1.2. For each factor f in H with ancestor factor(s) f_a, for $a = 1 \dots A_f$ (e.g., $C(A)$ nested on A in model (1)):

$$L_f = r_f \cdot \prod_{a=1}^{A_f} L_{f_a} \tag{9}$$

$$\overline{\mathbf{X}}_f(L_f, M) \sim \mathcal{D}_f \tag{10}$$

      with r_f the number of replicates in each unique combination of levels (cell), and $\mathcal{D}_f$ the chosen PRNG.
      2.1.3. For each interaction i in H with ancestor factor(s) f_a for $a = 1 \dots A_i$ (e.g., AB in model (1)):

$$L_i = \prod_{a=1}^{A_i} L_{i_a} \tag{11}$$

$$\overline{\mathbf{X}}_i(L_i, M) \sim \mathcal{D}_i \tag{12}$$

      with $\mathcal{D}_i$ the chosen PRNG.
   2.2. Generate background variability for a chosen PRNG:

$$\mathbf{X}_E(N, M) \sim \mathcal{D}_E \tag{13}$$

   2.3. Normalize each matrix $\overline{\mathbf{X}}_f$ for each factor f, $\overline{\mathbf{X}}_i$ for each interaction i, and $\mathbf{X}_E$ so that the Frobenius norm equals the squared root of the number of rows.[¶]
   2.4. For each observation n from 1 to N:
      2.4.1. For each factor f, we build $\mathbf{X}_f(N, M)$ from $\overline{\mathbf{X}}_f(L_f, M)$ and the design matrix $\mathbf{F}$:

$$\mathbf{x}_f^n = \overline{\mathbf{x}}_f^l \tag{14}$$

      with $\mathbf{x}_f^n$ the nth row of $\mathbf{X}_f(N, M)$ and $\overline{\mathbf{x}}_f^l$ the lth row of $\overline{\mathbf{X}}_f(L_f, M)$, with l determined according to matrix $\mathbf{F}$.
      2.4.2. For each interaction i, we build $\mathbf{X}_i(N, M)$ from $\overline{\mathbf{X}}_i(L_i, M)$ and the design matrix $\mathbf{F}$:

$$\mathbf{x}_i^n = \overline{\mathbf{x}}_i^l \tag{15}$$

      with $\mathbf{x}_i^n$ the nth row of $\mathbf{X}_i(N, M)$ and $\overline{\mathbf{x}}_i^l$ the lth row of $\overline{\mathbf{X}}_i(L_i, M)$, with l determined according to matrix $\mathbf{F}$.
   2.5. Compute the matrices with the structural and residual part with the standard deviation coefficients:

$$\mathbf{X}_S = \sum_f k_f \mathbf{X}_f + \sum_i k_i \mathbf{X}_i \tag{16}$$

$$\mathbf{X}_E = k_e \mathbf{X}_E \tag{17}$$

   2.6. For θ from 0 to 10δ in δ steps:
      2.6.1. Yield the simulated data:
      2.6.2. Compute ASCA+ partition and the F-ratio

$$\mathbf{X} = \theta \mathbf{X}_S + \mathbf{X}_E \tag{18}$$

      for factors and interactions for both the simulated data and P permutations. For (high-order) factors and interactions in H with no descendants:

$$F_f = (SS_f / DoF_f) / (SS_E / DoF_E) \tag{19}$$

$$F_i = (SS_i / DoF_i) / (SS_E / DoF_E) \tag{20}$$

      where SS refers to the sum-of-squares (the Frobenius norm) of a factor/interaction/residual matrix in the factorization with ASCA+, and $DoFs$ represents the corresponding degrees of freedom [2]. The DoFs of a factor is the number of levels minus one, the DoFs of an interaction is the product of the DoFs of its factors, and the DoFs of the residuals is the total (number of observations minus 1 in the data) minus the DoFs of all factors and interactions in the model. The ratio between the SS and the DoF is often called the mean sum-of-squares (MS). For any factors and interactions in H with descendants:

$$F_f = \frac{(SS_f / DoF_f)}{\left(\sum_{d=1}^{D_f} SS_{f_d} + \sum_{d=1}^{D_i} SS_{i_d}\right) / \left(\sum_{d=1}^{D_f} DoF_{f_d} + \sum_{d=1}^{D_i} DoF_{i_d}\right)} \tag{21}$$

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$F_i = \frac{(SS_i/DoF_i)}{\left(\sum_{d=1}^{D_i} SS_{i_d}\right)/\left(\sum_{d=1}^{D_i} DoF_{i_d}\right)} \quad (22)$$

with descendant factor(s) f_d for $d = 1 \ldots D_f$ and descendant interaction(s) f_i for $d = 1 \ldots D_i$.

2.6.3. For each factor f, if the associated p value computed using Equation (7) is below α do:

$$power_f(\theta) = power_f(\theta) + 1 \quad (23)$$

2.6.4. For each factor i, if the associated p value computed using Equation (7) is below α do:

$$power_i(\theta) = power_i(\theta) + 1 \quad (24)$$

3. Normalize $power_f(\theta) = power_f(\theta)/R$ and $power_i(\theta) = power_i(\theta)/R$ for each factor f/interaction i

The algorithm works as follows. In step 0, we set the general characteristics of the data simulation. In step 1, the algorithm initializes the values in the RPC. In step 2, we iterate through a number of repetitions to compute the RPC. Each repetition consists on the simulation of a structural part (in $\mathbf{X}_S$) and a residual part (in $\mathbf{X}_E$). The inner loop builds the data for intermediate cases between the absence of effect ($\mathbf{X} = \mathbf{X}_E$) and absence of residuals ($\mathbf{X} = \mathbf{X}_S$), factorizes it with ASCA+ and performs the statistical inference through permutation testing. If the computed significance is below the imposed significance level, the power is increased by one. In step 3, the power is normalized by the number of repetitions.

Some alternative configurations for a RPC that are relevant in practice can be straightforwardly implemented by modifying specific parameters or small parts of the algorithm:

- Non-balanced designs can be easily integrated in the design matrix $\mathbf{F}$, see [6].
- By properly choosing $\mathcal{D}_E$, we can emulate different distributions in the residuals to assess robustness to deviations from normality in the manner of Anderson and Ter Braak [9] but for multivariate responses. We can also generate the level averages using different distributions.
- Both fixed and random factors can be simulated in the same manner.
- We can integrate complex designs by adding multiple crossed and nested relationships as well as interactions.
- We can generate RPCs for a specific factor or interaction in an experimental design, or for all of them. The algorithm provides the solution for a RPC that considers a simultaneous incremental effect in all factors and interactions of the model, but some factors may be deactivated by setting the corresponding standard deviation coefficient to 0. We can also maintain the effect of a significant factor/interaction fixed along the curve by adding its contribution directly in Equation (18), for example:

$$\mathbf{X} = \theta\mathbf{X}_S + \mathbf{X}_E + k_f\mathbf{X}_f \quad (25)$$

- In a similar way, we can add a covariate $\mathbf{X}_{cv}(N, M) \sim \mathcal{D}_{cv}$ directly in Equation (18), for example:

$$\mathbf{X} = \theta\mathbf{X}_S + \mathbf{X}_E + k_{cv}\mathbf{X}_{cv} \quad (26)$$

- We can generate RPCs for alternative statistics to the F-ratio in Equations (19)–(22), see [6].

5 | APCs for ASCA

APCs differ to RPCs in that the power is shown in terms of the sampling size, rather than the relative effect size. APCs have the same applications as the RPCs, but with the additional advantage that they provide information about the number of replicates and/or factor levels that one may use in a multivariate experiment in order to attain a given probability of success, that is, the probability of rejecting the null hypothesis when the alternative hypothesis is true: the power $1 - \beta$.

An APC is built by simulating data that progressively enlarges the number of levels in a specific factor or the whole experimental design. The consequence of this enlargement is a reduction of the variance of both the null and the alternative distributions (Figure 1), with a subsequent increase of statistical power. For instance, if we take model (1), we can apply APCs to investigate how the statistical power of the factors and the interaction is affected when:

- We iteratively enlarge the number of levels in A. Often A is a factor that controls the number of groups of individuals in a clinical study, for example, with a number of disease subtypes. This APC gives useful information about whether incorporating more or less subtypes can impact our probability of success.
- We iteratively enlarge the number of levels in B. Often B is a factor with several repeated measures over the same individual, for example, in time or in biological samples. Then, the APC will allow us to investigate the effect of adding more time points/biological samples in our probability of success.
- We iteratively enlarge the number of levels in $C(A)$. Often $C(A)$ models the individual variability. The APC will give us an idea about the number of individuals per group we should choose for a certain probability of success.
- We iterative enlarge the whole experiment. The APC will give us an idea about the general number of replicates we should choose for a certain probability of success. This is often a good alternative choice when a factor like $C(A)$ is not in the design.

Our approach to generate APCs follows these steps (the algorithm is summarized not to replicate detailed explanations of the RPC algorithm)

0. INPUT:
   - INPUTs in the RPC algorithm
   - θ the fixed effect size
   - f_{rep} the index of the factor that is replicated, or 0 if all the whole experiment is replicated
   - δ incremental steps in the sampling size η

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1. Set $power_f(\eta) = 0$ and $power_i(\eta) = 0$ for each factor f and interaction i, respectively, and for η from 1 to 10δ in δ steps.
2. For each repetition from 1 to R
   2.1. For σ from 1 to 10δ in δ steps:
      2.1.1 Create $\mathbf{F}_\theta$ from $\mathbf{F}$ and f_{rep}; and N is set to the number of rows of $\mathbf{F}$.
      2.1.2 Generate random matrices to represent level averages in factors and interactions.
      2.2.3 Generate background variability for a chosen PRNG.
      2.2.4 Normalize each matrix $\overline{\mathbf{X}}_f$ for each factor f, $\overline{\mathbf{X}}_i$ for each interaction i, and $\mathbf{X}_E$ so that the Frobenius norm equals the squared root of the number of rows.
      2.2.5 Compute final factor and interaction matrices of N rows.
      2.2.6 Compute the matrices with the structural and residual part with the standard deviation coefficients.
      2.2.7 Yield the simulated data.
      2.2.8 Compute ASCA+ partition and the F-ratio for factors and interactions for both the simulated data and P permutations.
      2.2.9 Update the power of factors and interactions.
3. Normalize $power_f(\eta) = power_f(\eta)/R$ and $power_i(\eta) = power_i(\eta)/R$ for each factor f/interaction i

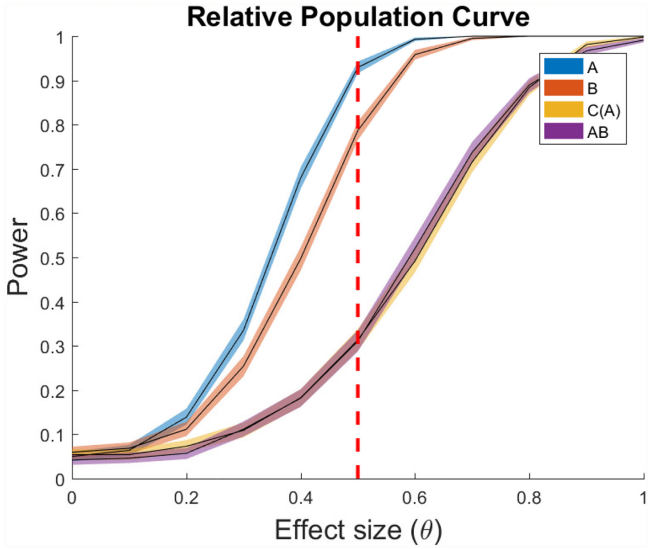


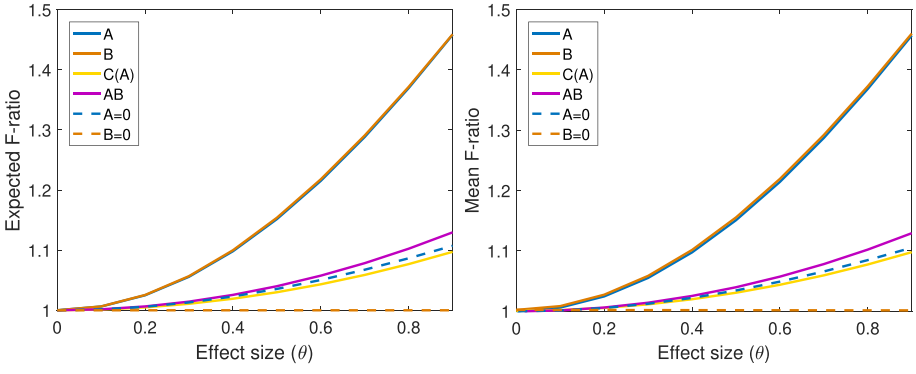
FIGURE 3 | Example of relative population curve for model (1). The design matrix $\mathbf{F}$ contains a full factorial design with four levels in A, three levels in B, and four individuals in each cell of $C(A)$. Other inputs are $M = 400$, $k_A = k_B = k_{C(A)} = k_{AB} = 0.2$, and $k_E = 1$, $R = 1000$, $P = 200$, $\delta = 0.1$, and $\alpha = 0.05$. We marked $\theta = 0.5$ as a reference for the following figures.

The algorithm works as follows: In step 0, we set the general characteristics of the data simulation, and in step 1, the algorithm initializes the values in the APC. In step 2, we iterate through a number of repetitions to compute the APC. Each repetition consists of an inner loop that progressively increases the sampling size by either increasing in one the number of levels of a given factor (for $f_{rep} > 0$) or by adding one complete set of experimental runs. Subsequently, the same simulation approach as in an RPC is followed to generate the simulated matrix with both structural and residuals parts. This matrix is then factorized using ASCA+ and the statistical inference is performed through permutation testing. If the computed significance is below the imposed significance level, the power is increased in one. In step 3, the power is normalized by the number of repetitions.

6 | Simulation Examples

6.1 | RPCs

Figure 3 presents an example of the RPCs for model (1) and for a full factorial design with four levels in A, three levels in B and four individuals in each cell of $C(A)$. The RPCs are shown with 95% confidence intervals computed by bootstrapping. In the example, all standard deviation coefficients are fixed to 1/5 of the standard deviation in the residuals ($k_A = k_B = k_{C(A)} = k_{AB} = 0.2$ and $k_E = 1$). The behavior of the RPCs is the one expected for a correct power curve [6]: (a) in the absence of effect (i.e., at $\theta = 0$), all curves adjust to the significance level of $\alpha = 0.05$, and (b) at some given effect size, the curves start gaining power until they reach 1.



(A) Expected F-ratio (B) Mean F-ratio

FIGURE 4 | Theoretical expected F-ratio (A) and mean simulated F-ratio (B) in terms of the effect size for the RPC in Figure 3.

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Interestingly, the order in which the curves in Figure 3 start rising and finally reach 1, that is, the relative power of the curves for the different factors and the interaction, is not intuitive given the same standard deviation of 0.2 was used for all of them: A is the most statistically powerful factor, followed by B, and $C(A)$ and AB are the least powerful. We found that the relative power is a complex function of the ordering structure among factors/interactions (as depicted in Figure 2) and the number of levels thereof. The relative power has important practical implications, which are well-known in the area of design of experiments but not so widely understood by some experimenters: the statistical power of a given factor depends on the design of other factors in the experiment. To give an example, our ability to determine biological differences between a disease and a control group (factor A) depends on the number of individuals we include in the experiment (factor $C(A)$), but also on the number of repeated measures we take for each of them (factor B and interaction AB). Thus, understanding the relative power in a specific experimental design can be extremely useful to maximize our capability of success. Generally speaking, the relative power is complicated to derive mathematically, especially in the presence of complex and varying null distributions across multiple responses, missing data, and other practicalities. Thus, RPCs are an interesting computational alternative for such derivation which can be made as specific to the problem at hand as desired.

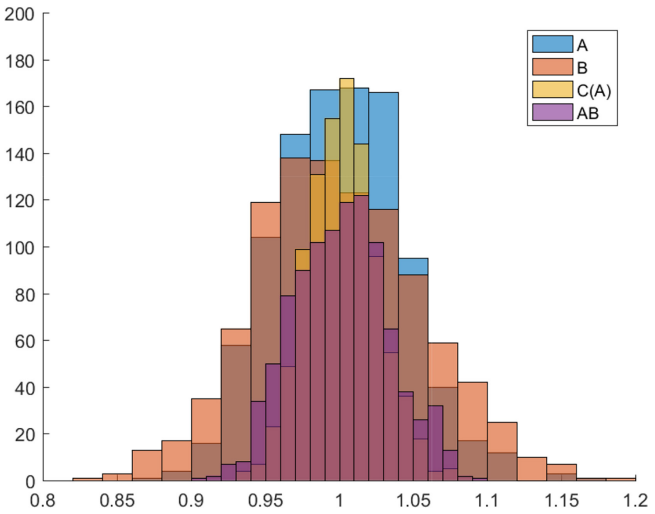


FIGURE 5 | Null distributions for the first simulated dataset in the RPC in Figure 3 and for $\theta = 0.5$.

While the mathematical derivation of multivariate power curves is often hopeless, we can still derive expected values for the variance in factors, interactions and residuals [2], and so of the F-ratios. Appendix A provides such derivation for the example considered in Figure 3. Using Equations (A10)–(A13), we can plot the expected F-ratios for the set of values of θ, as illustrated in Figure 4A and compare them to the averaged F-ratios obtained from the 1000 datasets simulated to compute the RPCs in Figure 3. These averaged F-ratios are in Figure 4B. We can see that the theoretical and numerical results match perfectly, which shows that our simulation approach accurately follows the ANOVA theory. We can also see that the F-ratios alone cannot explain the relative power observed in the RPCs of Figure 3: For instance, the F-ratio profiles of A and B are similar, while the relative power in the RPCs are not.

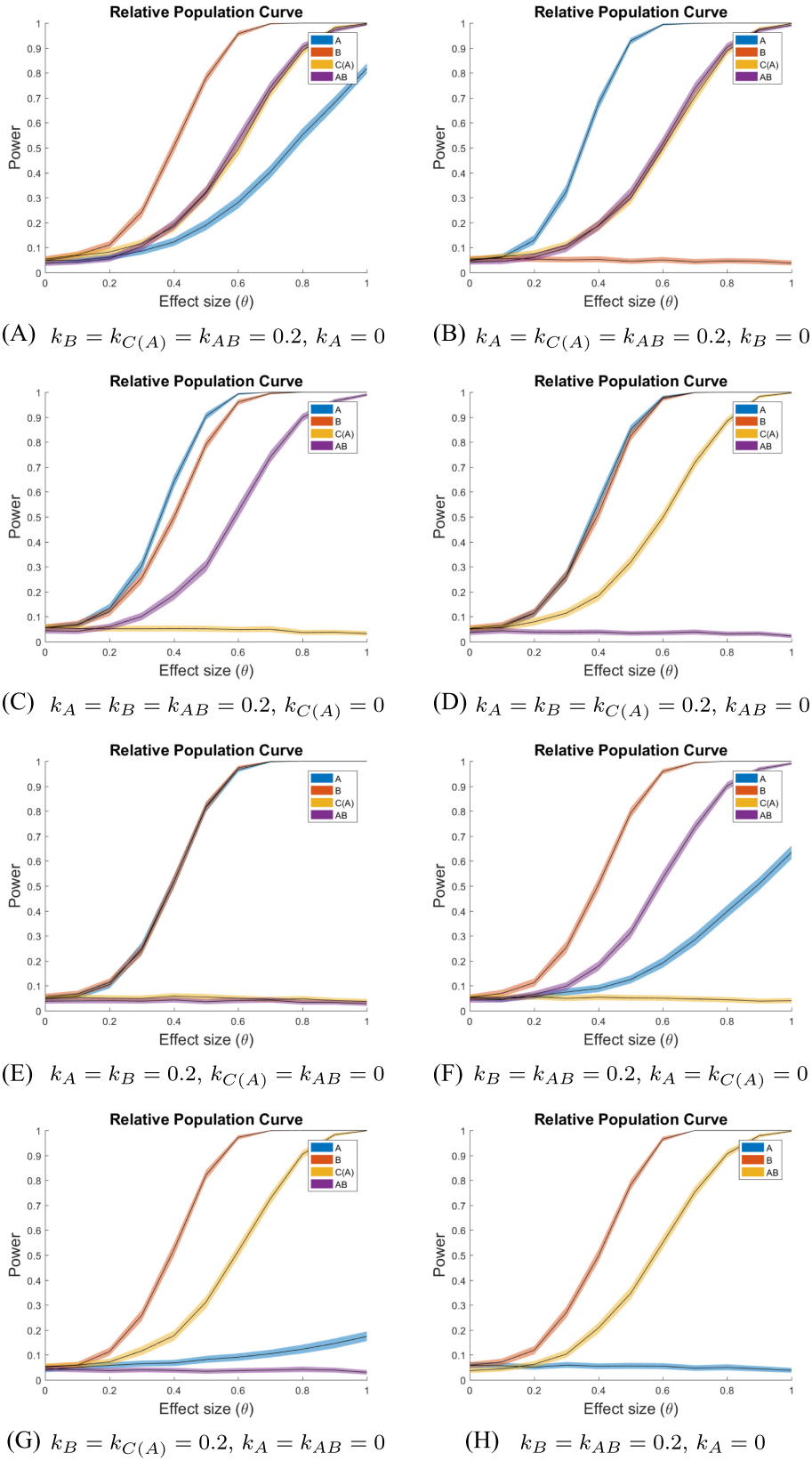
The discrepancy between F-ratio and RPC profiles is caused by the different null distributions of factors and interactions. Actually, it is the complexity to mathematically derive these null distributions which makes our computational approach a suitable tool to compute power curves. We illustrate the null distributions of our example in Figure 5, generated with permutation testing for the first of the 1000 simulated datasets in the RPC and for $\theta = 0.5$. The null distribution of B is significantly wider than the others. Since the p value is obtained by comparing the F-ratio to the null distribution, and A and B show similar F-ratios at $\theta = 0.5$ (Figure 4), the wider null distribution in B makes the power curve to rise slower than that in A in Figure 3. This is because statistical power is associated to lower p values. The ASCA table for the same simulated dataset in Figure 5 is shown in Table 1. The F-ratios and the p values in the table are consistent with what we see in Figures 4 and 3, respectively, for $\theta = 0.5$: A and B present larger F-ratios that the others but still close to 1, A is statistically significant while the rest are not. Note that this selected dataset in Figure 5 and Table 1 is a single instance of the distribution that is averaged in Figures 3 and 4, which is the reason why some discrepancy is expected (e.g., B is expected to be statistically significant 80% of times at $\theta = 0.5$, according to Figure 3, but it is not in this example).

TABLE 1 | ASCA table for the first simulated dataset in the RPC of Figure 3 and for $\theta = 0.5$.

	SumSq	PercSumSq	df	MeanSq	F	p value
Mean	1.3861	2.7832	1	1.3861		
A	3.4739	6.9757	3	1.158	1.1096	0.00999
B	2.184	4.3855	2	1.092	1.0613	0.14585
$C(A)$	12.6117	25.3243	12	1.051	1.0522	0.023976
AB	6.1738	12.397	6	1.029	1.0302	0.17582
Residuals	23.9712	48.1342	24	0.9988		
Total	49.8006	100	48	1.0375		

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Figure 6 presents several RPC examples similar to the first one, but where some of the factors or the interaction are deactivated with a null standard deviation. The figure shows that for any case where $k_B = 0$, $k_{C(A)} = 0$ and/or $k_{AB} = 0$, the corresponding RPC stays at expected type I error of 0.05. This behavior is not found for factor A (so that the RPC does not go to 0.05 even for



(A) $k_B = k_{C(A)} = k_{AB} = 0.2,\ k_A = 0$ (B) $k_A = k_{C(A)} = k_{AB} = 0.2,\ k_B = 0$

(C) $k_A = k_B = k_{AB} = 0.2,\ k_{C(A)} = 0$ (D) $k_A = k_B = k_{C(A)} = 0.2,\ k_{AB} = 0$

(E) $k_A = k_B = 0.2,\ k_{C(A)} = k_{AB} = 0$ (F) $k_B = k_{AB} = 0.2,\ k_A = k_{C(A)} = 0$

(G) $k_B = k_{C(A)} = 0.2,\ k_A = k_{AB} = 0$ (H) $k_B = k_{AB} = 0.2,\ k_A = 0$

FIGURE 6 | Examples of relative population curve for model (1). The design matrix **F** contains a full factorial design with four levels in A, three levels in B and four individuals in each cell of $C(A)$. Other inputs are $M = 400$, $k_E = 1$, $R = 1000$, $P = 200$, $\delta = 0.1$ and $\alpha = 0.05$.

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

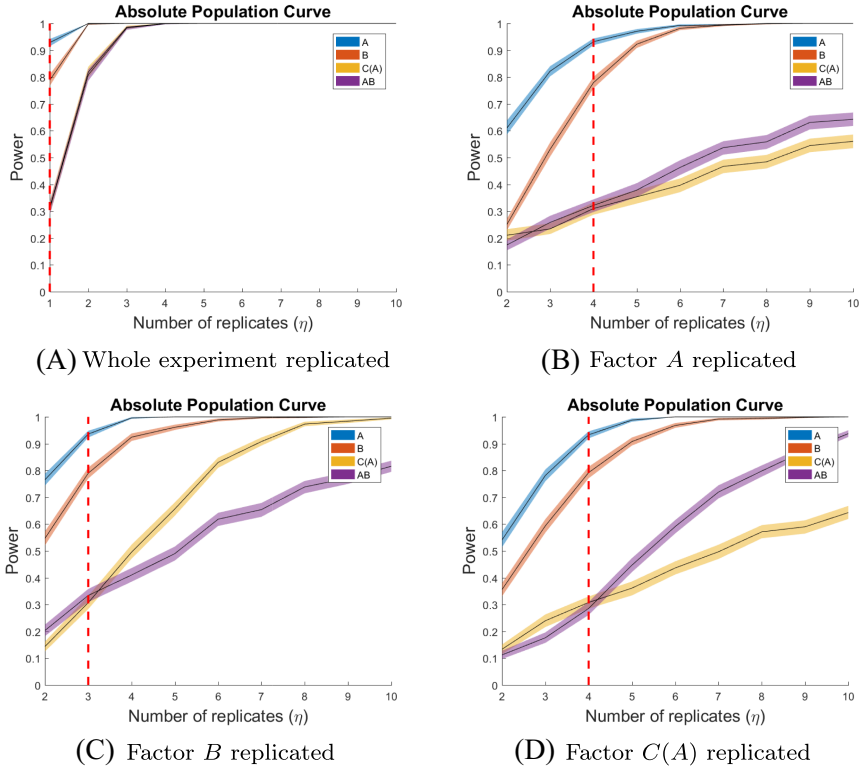
null effect in the factor) because its F-ratio is an approximate test rather than an exact one [2, 9].[#] These profiles in the RPCs can be explained from the theoretical derivation in Appendix A. Using this derivation, we included in Figure 4 the evolution of the expectation for the F-ratio of A and B for a null effect of the corresponding factor, marked with the labels $A = 0$ and $B = 0$, respectively. We can see that the approximate F-ratio in A does not cancel out for a null variance of the factor, which leads to the misleading RPC of A that shows an unrealistic power. This undesirable behavior in the test of factor A, and so in the corresponding power curve, remains even when the variance of $C(A)$ or AB is also cancelled out (Figure 6F,G, respectively), and the power curves only works as expected when both are cancelled or when either $C(A)$ or AB is not considered in the model, so that the F-ratio of A corresponds to an exact test (Figure 6H). The RPC is a very useful tool to identify these situations, that is, when an approximate test provides unrealistic statistical power or lack of thereof, allowing us to avoid false negatives in practice: See [6, 17] for an example.

6.2 | APCs

Figure 7 shows four examples of APCs computed from the same parameters as the RPC in Figure 3 and for $\theta = 0.5$. The first example of APCs iteratively replicates the whole experiment (Figure 7A) and the remaining examples iteratively increase the number of levels of each of the factors (Figure 7B–D for factors A, B and $C(A)$, respectively). For reference, we marked with a red dashed line the same baseline situations in all APCs and the original RPC in Figure 3. Thus, the dashed line at $\theta = 0.5$ in Figure 3 identifies the same simulation point as at $\eta = 1$ in Figure 7A, $\eta = 4$ (for four levels in A) in Figure 7B, $\eta = 3$ (for three levels in B) in Figure 7C, and $\eta = 4$ (for four replicates in $C(A)$, $r_{CA} = 4$) in Figure 7D.

In general, we can see that increasing the replicates enhances the power in all factors and the interaction for all APCs. As discussed in Figure 1, this enhancement is motivated by a reduction in the variance of the null and the alternative distributions of the test. Let us discuss this for each of the four examples.

In Figure 7A, we duplicate the whole experiment, but all levels of the factors remain the same. This makes all variance coefficients in Equations (A2)–(A5), with the exception of the variance of the error, to be multiplied by a factor of 2 (and in general of η if the experiment is duplicated η-wise). This makes the expected MSs, and in turn the expected F-ratios, larger. The null distribution of a single instance simulated with this duplication (Figure 8A) remains similar to the original one with no duplication in Figure 5. However, if we compare the ASCA tables of the case with and without duplication, Tables B1 and 1, respectively, we can see that after duplication all the factors and the interaction are statistically significant as a result of the larger F-ratios. This is correctly depicted by the APC in Figure 7A, where power for $\eta = 2$ is above 0.8 for all factors and the interaction (which means that in more than 80% of the simulated experiments we get statistically significant differences in the factors and the interaction).



(A) Whole experiment replicated (B) Factor A replicated

(C) Factor B replicated (D) Factor $C(A)$ replicated

FIGURE 7 | Examples of absolute population curve for model (1). The original design matrix **F** contains a full factorial design with four levels in A, three levels in B, and four individuals in each cell of $C(A)$. Other inputs are $M = 400$, $k_A = k_B = k_{C(A)} = k_{AB} = 0.2\theta$ for $\theta = 0.5$ and $k_E = 1$, $R = 1000$, $P = 200$, and $\alpha = 0.05$.

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

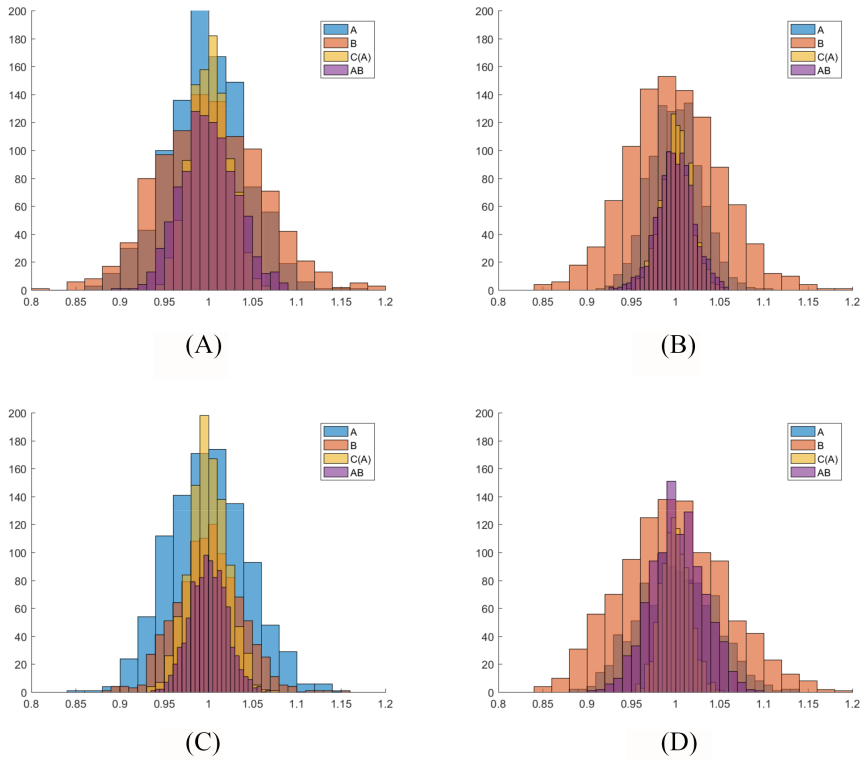


FIGURE 8 | Null distributions for the first simulated dataset in the RPC in Figure 3 and for $\theta = 0.5$, but when the whole experiment replicated (A), the number of levels of A is duplicated (B), the number of levels of B is duplicated (C), and the number of levels of $C(A)$ is duplicated (D).

In Figure 7B, the APCs show the behavior of the power in terms of the number of levels of A, L_A. From all of the expected MS values, only $E(MS_B)$ in Equation (A7) is affected by L_A. Consequently, the APC with larger slope in terms of L_A is actually B. All the other factors and the interaction are increase their power with L_A, but rather than because of a change in expected MS and/or F-ratio, they do because of a change in their null distribution. This can be seen by comparing Figure 5, for $L_A = 4$, with Figure 8B, for $L_A = 8$. The latter shows a clear reduction in the variance of the null distributions of A, $C(A)$ and AB. This reduction of variance leads to an increase of power. Finally, comparing Tables 1 and B2, we can see that the duplication of the levels of A clearly reduces the p value in A and B, in agreement to what we see in Figure 7B, where the power of these factors for $\eta = 8$ is close to 1 (that is, in almost 100% of the simulated experiments we get statistically significant differences in these factors, but only 50% for $C(A)$ and AB).

Figure 7C,D illustrates that the power in factor $C(A)$ is mostly affected by the increase of levels of B, L_B, and the power of the interaction AB is mostly affected by the replicates in $C(A)$, $r_{C(A)}$. All the other factors and interactions also increase their power. Again, this increase of power is a complex mixture of a modification of the expected MS's and F-ratios, and a reduction of the variance in the null distribution. It is hopeless to predict this behavior without computational means, but easily observed in Figure 8C,D and Tables B3 and B4.

As a general conclusion, we can see that sampling size in the form of replicates and varying number of levels in the factors can have a complex influence on the relative power of factors and their interactions. In complex practical cases, an easy way to understanding how any form of duplication affect the power of each given factor and/or interaction is through the APC algorithm.

7 | Conclusion

In this paper, we introduce the population power curves for ASCA and demonstrate them in simulation, discussing their relation to the theory of ANOVA and derive two useful forms of curve: RPCs and APCs. RPCs are useful to find the optimal ASCA pipeline for the analysis of an experimental design at hand. APCs are useful to determine the sample size and the optimal number of levels for each factor during the planning phase on an experiment. We believe that both tools should be adopted by ASCA practitioners to plan their experimental design (APCs) and analysis pipeline (RPCs) during the planning phase of a multivariate experimental design.

In a sequel of this paper, we will introduce the sample power curves for ASCA, which is an optimized version of a power curve when a small sample of the experiment at hand is available, for instance, obtained by running a reduced number of trials before a larger experiment.

Acknowledgments

This work was supported by the Agencia Estatal de Investigación in Spain, MCIN/AEI/10.13039/501100011033, grant no. PID2020-113462RB-I00. Michael Sorochan Armstrong has received funding

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 101106986. Funding for open access charge: Universidad de Granada /CBUA. We would like to acknowledge the feedback and interesting comments on the paper of Prof. Morten Rasmussen.

Data Availability Statement

All simulated examples in the paper can be reproduced with the code available at https://github.com/CoDaSLab/PopulationCurvesASCA.

Endnotes

[1] Stable release of MEDA Toolbox v1.4 (https://github.com/CoDaSLab/MEDA-Toolbox/releases/tag/v1.4).

[2] Throughout the paper, we assume that the higher the statistic the more significant the effect of the factor/interaction.

[3] Given the multivariate nature of the response, and to simplify notation, we generally refer to the standard deviation σ of each response vector, so that the expected standard deviation of each individual response would be $\sigma/\sqrt{M}$.

[4] This normalization is instrumental for the correspondence of theoretical and numerical results, as shown later on.

[5] Please note an exact test for A in the experimental design of model (1) does not exist.

References

1. A. K. Smilde, J. J. Jansen, H. C. J. Hoefsloot, R.-J. A. N. Lamers, J. Van Der Greef, and M. E. Timmerman, "ANOVA-Simultaneous Component Analysis (ASCA): A New Tool for Analyzing Designed Metabolomics Data," *Bioinformatics* 21, no. 13 (2005): 3043–3048.

2. D. C. Montgomery, *Design and Analysis of Experiments* (Wiley, 2020), https://books.google.es/books?id=kB7zDwAAQBAJ.

3. M. Thiel, B. Feraud, and B. Govaerts, "ASCA+ and APCA+: Extensions of ASCA and APCA in the Analysis of Unbalanced Multifactorial Designs," *Journal of Chemometrics* 31, no. 6 (2017): e2895.

4. M. Martin and B. Govaerts, "LIMM-PCA: Combining ASCA+ and Linear Mixed Models to Analyse High-Dimensional Designed Data," *Journal of Chemometrics* 34, no. 6 (2020): e3232.

5. T. S. Madssen, G. F. Giskeødegård, A. K. Smilde, and J. A. Westerhuis, "Repeated Measures ASCA+ for Analysis of Longitudinal Intervention Studies With Multivariate Outcome Data," *PLoS Computational Biology* 17, no. 11 (2021): e1009585.

6. J. Camacho, C. Díaz, and P. Sánchez-Rovira, "Permutation Tests for ASCA in Multivariate Longitudinal Intervention Studies," *Journal of Chemometrics* 37, no. 7 (2023): e3398.

7. A. M. G. C. C. J. Polushkina, "On Missing Data, Outliers and Transformations in Permutation Testing for ASCA and Related Factorizations," (2023), in preparation.

8. Z. Zhang and K.-H. Yuan, "Practical Statistical Power Analysis Using Webpower and R," (2018).

9. M. Anderson and C. T. Braak, "Permutation Tests for Multi-Factorial Analysis of Variance," *Journal of statistical computation and simulation* 73, no. 2 (2003): 85–113.

10. D. J. Vis, J. A. Westerhuis, A. K. Smilde, and J. van der Greef, "Statistical Validation of Megavariate Effects in ASCA," *BMC Bioinformatics* 8, no. 1 (2007): 1–8.

11. F. Koleini, S. Hugelier, M. A. Lakeh, H. Abdollahi, J. Camacho, and P. J. Gemperline, "On the Complementary Nature of ANOVA Simultaneous Component Analysis (ASCA+) and Tucker3 Tensor Decompositions on Designed Multi-Way Datasets," *Journal of Chemometrics* 37, no. 11 (2023): e3514.

12. M. A. Rasmussen, B. Khakimov, J. Engel, and J. Jansen, "Permutation Strategies for Inference in ANOVA Based Models for Mixed and Non-Orthogonal Designs Including Continuous Covariates," *Journal of Chemometrics* (2024).

13. G. Zwanenburg, H. C. J. Hoefsloot, J. A. Westerhuis, J. J. Jansen, and A. K. Smilde, "ANOVA-Principal Component Analysis and ANOVA-Simultaneous Component Analysis: A Comparison," *Journal of Chemometrics* 25, no. 10 (2011): 561–567.

14. M. S. Armstrong and J. Camacho, "Chapter 4—Replicate Analysis for Uncertainty Estimation in PARAFAC and PARAFASCA Analyses of Factorial Metabolomics Data," in *Fundamentals and Applications of Multiway Data Analysis*, eds. A. C. Olivieri, G. M. Escandar, H. C. Goicoechea, and A. M. de la Peña, Data Handling in Science and Technology, Vol. 33, (Elsevier, 2024), 61–81, https://www.sciencedirect.com/science/article/pii/B9780443132612000199.

15. M. A. Babyak, "What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models," *Psychosomatic Medicine* 66, no. 3 (2004): 411–421.

16. F. Marini, D. de Beer, E. Joubert, and B. Walczak, "Analysis of Variance of Designed Chromatographic Data Sets: The Analysis of Variance-Target Projection Approach," *Journal of Chromatography A* 1405 (2015): 94–102.

17. C. Díaz, C. González-Olmedo, L. Díaz-Beltrán, et al., "Predicting Dynamic Response to Neoadjuvant Chemotherapy in Breast Cancer: A Novel Metabolomics Approach," *Molecular Oncology* 16, no. 14 (2022): 2658–2671.

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix A: Theoretical Derivation of Expected F-Ratios

Given that all factors are derived from a (pseudo-)random distribution, and following Montgomery [2], we can derive the relationships between population variances and expected MSs in model (1):

$$E(MS_E) = \sigma_E^2 \tag{A1}$$

$$E(MS_{AB}) = \sigma_E^2 + r_{C(A)} \cdot \sigma_{AB}^2 \tag{A2}$$

$$E(MS_{C(A)}) = \sigma_E^2 + L_B \cdot \sigma_{C(A)}^2 \tag{A3}$$

$$E(MS_B) = \sigma_E^2 + r_{C(A)} \cdot \sigma_{AB}^2 + L_A \cdot r_{C(A)} \cdot \sigma_B^2 \tag{A4}$$

$$E(MS_A) = \sigma_E^2 + r_{C(A)} \cdot \sigma_{AB}^2 + L_B \cdot \sigma_{C(A)}^2 + L_B \cdot r_{C(A)} \cdot \sigma_A^2 \tag{A5}$$

where σ_A^2, σ_B^2, $\sigma_{C(A)}^2$, σ_{AB}^2, and σ_E^2 are the population variance of the factors, the interaction and the residuals, respectively, and in our example, we have $r_{C(A)} = 4$ replicates, and number of levels $L_A = 4$ and $L_B = 3$.

Combining previous equations and Equations (19)–(21), the inference statistics follow:

$$\begin{aligned} E(F_A) &= \frac{E(MS_A)}{(DoF_{C(A)}E(MS_{C(A)}) + DoF_{AB}E(MS_{AB}))/(DoF_{C(A)} + DoF_{AB})} \\ &= \frac{(\sigma_E^2 + r_{C(A)} \cdot \sigma_{AB}^2 + L_B \cdot \sigma_{C(A)}^2 + L_B \cdot r_{C(A)} \cdot \sigma_A^2)(r_{C(A)} + L_B - 1)}{r_{C(A)}\left(\sigma_E^2 + L_B \cdot \sigma_{C(A)}^2\right) + (L_B - 1)\left(\sigma_E^2 + r_{C(A)} \cdot \sigma_{AB}^2\right)} \end{aligned} \tag{A6}$$

$$E(F_B) = \frac{E(MS_B)}{E(MS_{AB})} = 1 + \frac{L_A \cdot r_{C(A)} \cdot \sigma_B^2}{\sigma_E^2 + r_{C(A)} \cdot \sigma_{AB}^2} \tag{A7}$$

$$E(F_C(A)) = \frac{E(MS_{C(A)})}{E(MS_E)} = 1 + \frac{L_B \cdot \sigma_{C(A)}^2}{\sigma_E^2} \tag{A8}$$

$$E(F_{AB}) = \frac{E(MS_{AB})}{E(MS_E)} = 1 + \frac{r_{C(A)} \cdot \sigma_{AB}^2}{\sigma_E^2} \tag{A9}$$

Adjusting the population variances to the square of the standard deviation coefficients used in the RPCs of Figure 3, that is, $\sigma_A^2 = \sigma_B^2 = \sigma_{C(A)}^2 = \sigma_{AB}^2 = 0.04\theta^2$ and $\sigma_E^2 = 1$, it holds:

$$E(F_A) = \frac{6(1 + 0.76\theta^2)}{4(1 + 0.12\theta^2) + 2(1 + 0.16\theta^2)} \tag{A10}$$

$$E(F_B) = 1 + \frac{0.64\theta^2}{1 + 0.16\theta^2} \tag{A11}$$

$$E(F_C(A)) = 1 + 0.12\theta^2 \tag{A12}$$

$$E(F_{AB}) = 1 + 0.16\theta^2 \tag{A13}$$

From Equations (A7), (A8), and (A9), we can see that if we set $\sigma_B^2 = 0$, $\sigma_{C(A)}^2 = 0$ and/or $\sigma_{AB}^2 = 0$, the corresponding expected F-ratio equals 1 regardless the variance of the error. This makes the RPC to adjust to the expected type I error regardless of θ. This behavior is not found for factor A. The reason can also be found in the corresponding equation of the F-ratio; see Equation (A6). This equation represents an approximate test rather than an exact one [2, 9]. If we set $\sigma_A^2 = 0$ and adjust the remaining population variances to the square of the standard deviation coefficients in terms of θ, it now holds:

$$E(F_A) = \frac{6(1 + 0.24\theta^2)}{4(a + 0.12\theta^2) + 2(1 + 0.16\theta^2)} \tag{A14}$$

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix B: Tables

ASCA tables for a single instance (dataset) simulate with the same parameters of the RPC in Figure 3 and for $\theta = 0.5$, but (i) when the whole experiment is duplicated (Table B1); (ii) when the number of levels of A, L_A, is duplicated (Table B2); (iii) when the number of levels of AB, L_B, is duplicated (Table B3); and (iv) when the number of replicates in $C(A)$, $r_{C(A)}$, is duplicated (Table B4).

TABLE B1 | ASCA table for the first simulated dataset in the RPC of Figure 3 and for $\theta = 0.5$, when the whole experiment is duplicated.

	SumSq	PercSumSq	df	MeanSq	*F*	*p* value
Mean	1.7243	1.7257	1	1.7243		
A	3.8883	3.8914	3	1.2961	1.1885	0.000999
B	2.7994	2.8017	2	1.3997	1.2888	0.000999
C(*A*)	13.1133	13.1239	12	1.0928	1.0946	0.000999
AB	6.5162	6.5215	6	1.086	1.0879	0.001998
Residuals	71.8774	71.9358	72	0.9983		
Total	99.9188	100	96	1.0408		

TABLE B2 | ASCA table for the first simulated dataset in the RPC of Figure 3 and for $\theta = 0.5$, when the number of levels of *A* is duplicated.

	SumSq	PercSumSq	df	MeanSq	*F*	*p* value
Mean	1.6805	1.6809	1	1.6805		
A	8.1368	8.1388	7	1.1624	1.1314	0.000999
B	2.6052	2.6058	2	1.3026	1.2794	0.000999
C(*A*)	24.7872	24.7931	24	1.0328	1.0219	0.1049
AB	14.2541	14.2575	14	1.0182	1.0074	0.37363
Residuals	48.5123	48.5239	48	1.0107		
Total	99.9762	100	96	1.0414		

TABLE B3 | ASCA table for the first simulated dataset in the RPC of Figure 3 and for $\theta = 0.5$, when the number of levels of *B* is duplicated.

	SumSq	PercSumSq	df	MeanSq	*F*	*p* value
Mean	1.537	1.5426	1	1.537		
A	3.9359	3.9501	3	1.312	1.2495	0.000999
B	5.8933	5.9146	5	1.1787	1.1578	0.000999
C(*A*)	13.0792	13.1262	12	1.0899	1.0913	0.000999
AB	15.2703	15.3253	15	1.018	1.0193	0.17283
Residuals	59.9256	60.1413	60	0.99876		
Total	99.6414	100	96	1.0379		

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE B4 | ASCA table for the first simulated dataset in the RPC of Figure 3 and for $\theta = 0.5$, when the number of replicates in $C(A)$ is duplicated.

	SumSq	PercSumSq	df	MeanSq	*F*	*p* value
Mean	1.6665	1.6705	1	1.6665		
A	4.0614	4.0712	3	1.3538	1.3083	0.000999
B	2.6655	2.672	2	1.3328	1.2989	0.000999
C(*A*)	29.0265	29.0969	28	1.0367	1.0333	0.021978
AB	6.1566	6.1716	6	1.0261	1.0228	0.24276
Residuals	56.1816	56.3178	56	1.0032		
Total	99.758	100	96	1.0391		

1099128x, 0, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.3596 by Universidad De Granada, Wiley Online Library on [11/09/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License