NeuroImage 56 (2011) 1338–1352

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Optimal design of multi-subject blocked fMRI experiments

Bärbel Maus [a,*], Gerard J.P. van Breukelen [b], Rainer Goebel [c], Martijn P.F. Berger [a]

[a] Maastricht University, Faculty of Health, Medicine and Life Sciences, Department of Methodology and Statistics, P.O. Box 616, 6200 MD Maastricht, The Netherlands
[b] Maastricht University, Faculty of Psychology and Neuroscience, Department of Methodology and Statistics, P.O. Box 616, 6200 MD Maastricht, The Netherlands
[c] Maastricht University, Faculty of Psychology and Neuroscience, Department of Cognitive Neuroscience, P.O. Box 616, 6200 MD Maastricht, The Netherlands

ARTICLE INFO

Article history:
Received 19 October 2010
Revised 4 March 2011
Accepted 8 March 2011
Available online 13 March 2011

Keyword:
Optimal design
Blocked design
Number of subjects
Number of cycles
Cost function

ABSTRACT

The design of a multi-subject fMRI experiment needs specification of the number of subjects and scanning time per subject. For example, for a blocked design with conditions *A* or *B*, fixed block length and block order *ABN*, where *N* denotes a null block, the optimal number of cycles of *ABN* and the optimal number of subjects have to be determined. This paper presents a method to determine the optimal number of subjects and optimal number of cycles for a blocked design based on the A-optimality criterion and a linear cost function by which the number of cycles and the number of subjects are restricted. Estimation of individual stimulus effects and estimation of contrasts between stimulus effects are both considered. The mixed-effects model is applied and analytical results for the A-optimal number of subjects and A-optimal number of cycles are obtained under the assumption of uncorrelated errors. For correlated errors with a first-order autoregressive (AR1) error structure, numerical results are presented. Our results show how the optimal number of cycles and subjects depend on the within- to between-subject variance ratio. Our method is a new approach to determine the optimal scanning time and optimal number of subjects for a multi-subject fMRI experiment. In contrast to previous results based on power analyses, the optimal number of cycles and subjects can be described analytically and costs are considered.

Introduction

There exists a trade-off between number of subjects and number of measurements in time per subject in terms of statistical efficiency and experimental costs in any kind of experiment. Statistical efficiency increases with increasing number of subjects and increasing number of measurements in time per subject, but the same holds for experimental costs. Worsley et al. (2002) pointed out that if the variance between subjects is higher than the variance within subjects, it is better to scan more subjects with a shorter scanning time per subject than to scan fewer subjects with a longer scanning time per subject. However, Worsley et al. (2002) concluded that due to additional costs per subject a compromise between number of subjects and scanning time has to be chosen. This paper presents a method which will be useful for fMRI researchers in choosing the optimal number of subjects and scanning time. Given the high costs for fMRI experiments, such a method is relevant for these experiments.

Generally, the optimization of fMRI experiments has been studied for single-subject fMRI experiments while less attention has been given to the optimization of multi-subject fMRI experiments. For single-subject fMRI experiments, it has been shown that blocked designs are statistically optimal for detection of stimulus related activation, called detection power, whereas event-related designs are statistically optimal for estimation of the hemodynamic response function (HRF), called estimation efficiency. Bianciardi et al. (2004) studied the effect of filtering and experimental design on second-level group analyses. Furthermore, the statistical power for multi-subject fMRI experiments was studied and sample size recommendations based on power calculations were given by Desmond and Glover (2002), Hayasaka et al. (2007), Mumford and Nichols (2008) and Murphy and Garavan (2004). In this paper, a new approach is given for the optimization of a multi-subject fMRI experiment, taking into account the costs per individual subjects, for scanning time and for the total experiment by a linear cost function.

This paper focuses on the A-optimality criterion (Atkinson et al., 2007; Dale, 1999; Liu et al., 2001) which selects as optimal the design for which the sum of the variances of the estimators for the unknown group parameters is minimized. A linear mixed-effects model is used to describe the relationship between unknown parameters and the fMRI signal. This type of model is often also referred to as random-effects model in fMRI literature (Penny and Holmes, 2004). The number of subjects and the scanning time are restricted by a linear cost function describing the experimental costs due to the number of subjects and scanning time. For a blocked design with task blocks $A_1,\ldots,A_Q$, with null blocks denoted by N and the block order $A_1\ldots A_QN$, the scanning time depends on the number of cycles of $A_1\ldots A_QN$, the length of stimulus blocks and null blocks and the number of stimulus types Q. Assuming uncorrelated errors, we derived analytically the optimal number of subjects and number of cycles for a

* Corresponding author. Fax: +31 43 36 18388.
E-mail address: baerbel.maus@maastrichtuniversity.nl (B. Maus).

1053-8119/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2011.03.019

blocked design assuming fixed block lengths and fixed number of stimulus types. For a linear mixed-effects model with a first-order autoregressive (AR1) error structure (Bullmore et al., 1996; Chatfield, 2003), numerical calculations were performed to find the optimal number of subjects and cycles. The optimal number of subjects and cycles were found to depend on the ratio of within- to between-subject variance for uncorrelated errors as well as correlated errors. Furthermore for both error structures, the ratio of costs per subject and costs per scanning time is crucial for the optimal choice of number of subjects and cycles. The results are valid for any number of stimulus type.

The outline of this paper is as follows. In the Methodology section the linear mixed-effects model, the A-optimality criterion and the cost function are presented and the numerical calculations for correlated errors are explained. In the section about optimal designs for a given budget, the analytical results for the optimal number of subjects and cycles assuming uncorrelated errors and the results of the numerical calculations for correlated errors are given. Further details on the derivations are provided in Appendices A and B. The consequences of the analytical derivations and numerical calculations are illustrated by examples with common values for fMRI experiments in the section about optimal designs for a given budget. Furthermore, we study in the Results section the effect on design efficiency when the number of cycles and subjects vary slightly from the optimal design. In the Discussion and conclusions section the results are discussed and conclusions are drawn based on the results.

Methodology

Model and estimators

Each subject is measured at N_T time points with a given repetition time (TR) and the data vector Y_i is obtained for each subject from these measurements. It is assumed that Q different stimulus types are presented in blocks to each subject. One block contains trials of the same stimulus type. The following first-level model is considered for subject i:

$$Y_i = Z\beta_i + S\gamma_i + \epsilon_i, \tag{1}$$

where Z of dimension $N_T \times Q$ models the hemodynamic responses to the Q different stimulus types and S of dimension $N_T \times r$ models r nuisance terms, e.g., intercept, linear or quadratic trends. The errors ϵ_i are assumed to be normally distributed with expectation zero and covariance matrix $\sigma_\epsilon^2 \sum$. It is assumed that all subjects have the same error covariance matrix $\sigma_\epsilon^2 \sum$.

The interest at the first level is in estimation of the stimulus effects β_i. The generalized least squares (GLS) estimator for $(\beta_i^T \quad \gamma_i^T)^T$ is given by

$$\widehat{\begin{pmatrix}\beta_i \\ \gamma_i\end{pmatrix}} = \left([Z \quad S]^T \sum^{-1} [Z \quad S]\right)^{-1} [Z \quad S]^T \sum^{-1} Y_i. \tag{2}$$

The matrix $[Z \quad S]$ is a concatenation of the matrices Z and S. The generalized least squares (GLS) estimator of $\hat{\beta}_i$ in Eq. (2) is given by

$$\hat{\beta}_i = \left(Z^T V^T (I - P_{VS}) VZ\right)^{-1} Z^T V^T (I - P_{VS}) VY_i, \tag{3}$$

where $V = \sum^{-1/2}$ is the inverse square root matrix of the error correlation matrix $\sum$ and $P_{VS} = VS(S^T \sum^{-1} S)^{-1} S^T V^T$ is the projection matrix onto the space spanned by the columns of the matrix VS (column space) (Liu and Frank, 2004; Lovell, 1963; Seber, 1977). The projection matrix P_{VS} is needed to obtain a closed form expression for $\hat{\beta}_i$. The covariance matrix of $\hat{\beta}_i$ is given by

$$Cov\left(\hat{\beta}_i\right) = \sigma_\epsilon^2 \left(Z^T V^T (I - P_{VS}) VZ\right)^{-1}. \tag{4}$$

We define the detrended and decorrelated design matrix $Z^* = (I - P_{VS})VZ$, which is the decorrelated design matrix VZ after removal of decorrelated trends, e.g., constant, linear and quadratic trends, by application of the projection matrix $(I - P_{VS})$. The following equations are then obtained:

$$\hat{\beta}_i = \left(Z^{*T} Z^*\right)^{-1} Z^{*T} VY_i \tag{5}$$

$$Cov\left(\hat{\beta}_i\right) = \sigma_\epsilon^2 \left(Z^{*T} Z^*\right)^{-1}. \tag{6}$$

The matrix $(Z^{*T} Z^*)/\sigma_\varepsilon^2$ is the Fisher information matrix which can be used to measure the amount of information present in the data about the unknown parameters (Liu and Frank, 2004) and to determine the optimal design (Atkinson et al., 2007).

The second-level model is given by

$$\beta_i = \beta_G + b_i, \tag{7}$$

where β_G are the group stimulus effects and b_i are the random stimulus effects for subject i. The random effects b_i are assumed to be normally distributed with expectation 0 and covariance matrix $\sigma_q^2 D$. The matrix D is the correlation matrix of the random-effects vector b_i. All random effects in the random-effects vector b_i are assumed to have the same random-effects variance, also called between-subject variance, σ_q^2. Combining Eq. (1) with Eq. (7) gives the following model:

$$Y_i = Z\beta_G + Zb_i + S\gamma_i + \epsilon_i. \tag{8}$$

The covariance matrix of the GLS estimator $\hat{\beta}_G$ for β_G in Eq. (8) is given by

$$\begin{aligned} Cov\left(\hat{\beta}_G\right) &= \frac{1}{N}\left(\sigma_\epsilon^2 \left(Z^T V^T (I - P_{VS}) VZ\right)^{-1} + \sigma_q^2 D\right), \\ &= \frac{1}{N}\left(\sigma_\epsilon^2 \left(Z^{*T} Z^*\right)^{-1} + \sigma_q^2 D\right), \end{aligned} \tag{9}$$

where N is the number of subjects. The term $\sigma_\epsilon^2 (Z^{*T} Z^*)^{-1}$ is the within-subject covariance matrix of $\hat{\beta}_i$ assuming the same error covariance $\sigma_\epsilon^2 \sum$ for all subjects, and $\sigma_q^2 D$ is the between-subject covariance matrix of the stimulus effects. Eq. (9) can be seen from Section 4 in Beckmann et al. (2001).

Optimality criterion and relative efficiency

Several optimality criteria exist, each with their own advantages and disadvantages (Atkinson et al., 2007). The two mainly applied criteria are the A- and D-optimality criteria. In this paper, the A-optimality criterion is applied and the number of cycles and the number of subjects are optimized subject to a linear cost function given in the Cost function section. The A-optimality criterion has been applied before for optimization of single-subject fMRI experiments by Dale (1999), Liu et al. (2001) and Wager and Nichols (2003). Assuming equal frequency for all Q stimulus types, Liu and Frank (2004) showed that the A-optimal stimulus frequency for estimation of stimulus effects is equal to $p = 1/(Q + \sqrt{Q})$. For a blocked design, this will result in longer null blocks than stimulus blocks. For estimation of pairwise contrasts, $p = 1/Q$ is the optimal stimulus frequency and results in no null blocks.

The D-optimality criterion minimizes the simultaneous confidence ellipsoid of the estimators for the effects of interest whereas the A-optimality criterion minimizes the sum of the variances of the estimators. This renders the A-optimality criterion more easy to understand for practitioners. One strength of the A-optimality criterion in our specific example is the possibility to derive analytical results. Another strength is the fact that off-diagonal elements of the random-effects covariance matrix $\sigma_q^2 D$ do not have to be taken into account when considering individual stimulus effects by the A-optimality criterion. One general advantage of the D-optimality criterion is that the D-optimal design is unaffected by linear transformation of the design matrix Z, i.e., linear transformation of the columns of the design matrix, whereas this is generally not the case for the A-optimality criterion. However, no linear transformations of the design matrix Z seem to be relevant for the present model except scaling of the HRF which would not affect the A-optimal design.

We will compare blocked designs with block order $A_1 \ldots A_Q N$ which means that a block of stimulus type A_1 is followed by a block of stimulus type A_2 etc. Finally, A_Q is followed by a null block denoted by N. This block order has been shown to be statistically optimal for estimation of individual stimulus effects (Maus et al., 2010a). By assuming null block length equal to 0 s, block order $A_1 \ldots A_Q$ is obtained which is efficient for estimation of pairwise contrasts. A cycle $A_1 \ldots A_Q N$, consisting of the individual stimulus blocks and a null block, can be repeated several times for the considered blocked designs. The compared blocked designs differ in the number of cycles N_C and number of subjects N. The number of stimulus types and block lengths are not optimized as these parameters will be fixed by the researchers for his or her specific experiment. Furthermore, optimal values for block length have already been given in literature (Aguirre and D'Esposito, 1999; Chein and Schneider, 2003; Maus et al., 2010a,b). This optimal value will be valid for all scanning times as the optimal value is influenced by the power spectrum of the HRF and the power spectrum of the error noise (Aguirre and D'Esposito, 1999).

Thus for fixed block lengths, stimulus onset asynchrony (SOA) and number of stimulus types, the design space Ξ, from which the optimal design will be chosen, consists of these blocked designs ξ with N_C cycles and N subjects:

$$\Xi = \Big\{ \xi | \xi \text{ is a blocked design with block order } A_1 \ldots A_Q N, \; N_C \text{ cycles and } N \text{ subjects}, N_C, N \in \mathbb{N} \Big\}. \tag{10}$$

The stimulus onset asynchrony (SOA) refers here to the time between successive trials in a stimulus block or null events in a null block.

Estimation of all individual stimulus effects β_G and estimation of stimulus effects contrasts $C\beta_G$ with $c \times Q$ contrast matrix C are considered. The contrast matrix C indicates c contrasts of interest. The A-optimal design minimizes the trace of the covariance matrix $Cov\left(C\hat{\beta}_G\right)$ over all possible designs in the design space. This $trace\left(Cov\left(C\hat{\beta}_G\right)\right)$ will vary for designs with different number of cycles and subjects. We denote by $C\hat{\beta}_G^{\xi}$ the estimator obtained for design ξ. The trace of $Cov\left(C\hat{\beta}_G^{\xi}\right)$ is equal to the sum of the variances for each contrast in $C\hat{\beta}_G$ and gives a measure of how precisely these contrasts (or individual stimulus effects if C is an identity matrix) are estimated. It follows from Eq. (9) that the A-optimal design minimizes

$$trace\left(Cov\left(C\hat{\beta}_G^{\xi}\right)\right) = \frac{\sigma_q^2}{N}\left(\frac{\sigma_\epsilon^2}{\sigma_q^2} trace\left(C\left(Z^{*T}Z^{*}\right)^{-1}C^T\right) + trace\left(CDC^T\right)\right) \tag{11}$$

over all designs ξ in Ξ subject to the linear cost function described in the Cost function section. We can simplify Eq. (11) for uncorrelated errors, SOA equal to TR and a nuisance matrix S equal to 1_{N_T}, where 1_{N_T} is a vector with N_T 1s as entries and models an intercept. In Appendix A.1 it is shown the matrix $Z^{*T}Z^{*}$ can then be expressed as $N_C \cdot M$, where N_C is the number of cycles of $A_1 \ldots A_Q N$ and M/σ_ϵ^2 is the $Q \times Q$ information matrix of one cycle $A_1 \ldots A_Q N$. Thus, the trace of the covariance matrix of $C\hat{\beta}_G$ can be rewritten for uncorrelated error as

$$trace\left(Cov\left(C\hat{\beta}_G\right)\right) = \frac{\sigma_q^2}{N}\left(\frac{\sigma_\epsilon^2}{\sigma_q^2}\frac{1}{N_C} \cdot trace\left(CM^{-1}C^T\right) + trace\left(CDC^T\right)\right). \tag{12}$$

The matrix M is only invertible when null blocks are included in the design (see Appendix A.1 for further explanation), but the general expression in Eq. (11) can always be used. From Eqs. (11) and (12) it can be seen that the optimal number of subjects N and the optimal number of cycles N_C, which minimize the A-optimality criterion, depend on the variance ratio $\sigma_\epsilon^2/\sigma_q^2$, but not on the specific values for σ_ϵ^2 or σ_q^2.

The relative efficiency, which can be used to compare a design ξ with the optimal design ξ^*, is given by the following ratio

$$RE\left(\xi|\xi^*\right) = \frac{trace\left(Cov\left(C\hat{\beta}_G^{\xi^*}\right)\right)}{trace\left(Cov\left(C\hat{\beta}_G^{\xi}\right)\right)}. \tag{13}$$

The relative efficiency is a value between 0 and 1 with values close to 0 indicating a very low efficiency of the design ξ versus the optimal design ξ^* and values close to 1 indicating a high efficiency of the design ξ versus the optimal design ξ^*. The inverse of the relative efficiency $RE(\xi|\xi^*)$ gives the number of times that design ξ has to be repeated to be as efficient as the optimal design ξ^*, i.e., to have the same precision for estimation of $C\beta_G$ and the same power if $C\beta_G$ is a contrast or single stimulus effect.

Cost function

A cost function for a blocked fMRI experiment with one run and one session is considered. For the cost function we are denoting the total costs of the experiment with C_T, the costs per subjects with C_1 and the costs per unit scanning time T_S with C_2. C_1 includes the subject fee, recruitment costs and equipment per subject, and costs due to preparation of a subject or the scanner before the subject's session. The costs of C_1 might rise if a special population is considered, e.g., subjects with a rare illness. It should be noted that the value associated with C_2 depends on the unit of time used, e.g., $C_2 =$ €400/h ≈ €6.66/min ≈ €0.11/s. The effective scanning time T_S is here the length of one run, i.e., one continuous period of scanning. For a blocked design with Q stimulus types, the scanning time T_S is given by

$$T_S = N_C \cdot \left(Q \cdot N_{BL_A} + N_{BL_0}\right) \cdot \mathrm{SOA}, \tag{14}$$

where N_C is the number of cycles of $A_1 \ldots A_Q N$, N_{BL_A} is the number of events per stimulus block for all stimulus types, N_{BL_0} is the number of null events per null block and Q is the number of stimulus types. The cost function of the whole experiment expresses the total costs C_T as sum of the costs for all subjects and the costs for the scanner time of all subjects:

$$C_T = NC_1 + N \cdot N_C \cdot \left(Q \cdot N_{BL_A} + N_{BL_0}\right) \cdot \mathrm{SOA} \cdot C_2. \tag{15}$$

Note that C_2 is given above as costs per second, i.e., costs for one second scanning time. The number of subjects N and the scanning time T_S are restricted by the cost function and the assumed total costs. As explained in the previous section block lengths, number of stimulus types Q and SOA are assumed to be fixed by the researcher. Thus, restriction of the scanning time is equivalent to restriction of the number of cycles N_C as can be seen from Eq. (14). We assume fixed total costs C_T and want to find the combination of number of subjects and scanning time T_S which minimizes $trace\left(C\hat{\beta}_G\right)$ in Eq. (11).

Numerical calculations

In the following we give an outline of the numerical calculations which were performed in MATLAB (The Mathworks Inc, Natick, MA) to find optimal designs for a model with correlated errors. The code for correlated errors is available on request from the first author. An AR1 covariance structure with AR1 autocorrelation ρ was assumed for the covariance matrix $\sigma_\epsilon^2\sum$ of all within-subject errors ε_i ($i=1,\ldots,N$). The optimal number of subjects and cycles were chosen as the values which minimize Eq. (11) for a specified design (SOA, null block length, stimulus block length, number of stimulus types Q), contrast matrix C, autocorrelation ρ and variance ratio $\sigma_\epsilon^2/\sigma_q^2$ of within-subject variance σ_ϵ^2 to between-subject variance σ_q^2. As the variance σ_q^2 is the same for all designs, it does not influence the values for which the minimum of Eq. (11) is obtained. It is therefore sufficient to focus on the variance ratio.

The number of cycles was varied from a lower boundary to an upper boundary. If no highpass filtering was assumed, the optimal number of cycles for uncorrelated errors to be derived in the Uncorrelated errors section was chosen as a lower boundary as it was expected and confirmed by preliminary results that the optimal number of cycles for correlated errors is higher than for uncorrelated errors. The reason for this relationship is explained in the Correlated errors section. If highpass filtering was assumed, the lower border was set equal to 1. The upper boundary was raised if the obtained optimal number of cycles for correlated errors turned out to be equal to the upper boundary.

The number of cycles was thus varied within a certain range and the number of subjects was calculated for each number of cycles with the cost function in Eq. (15) for fixed costs. The block lengths, SOA, TR, the number of stimulus types, the autocorrelation and the variance ratio were likewise fixed to constants, and the hemodynamic response function h was given. The combination of number of cycles and subjects, for which $trace\left(Cov\left(\hat{\beta}_G\right)\right)/\sigma_q^2$ as given by Eq. (11) reached its minimum, was the optimal number of cycles and optimal number of subjects.

Results

The optimal number of subjects and cycles for uncorrelated errors are determined in the Uncorrelated errors section analytically. We present some numerical results in the Uncorrelated errors section for the derived optimal number of cycles and subjects. For correlated errors with AR1 error structure (Bullmore et al., 1996; Chatfield, 2003), a numerical calculation in MATLAB (The Mathworks INc, Natick, MA) was performed as described in the Numerical calculations section to determine the optimal number of cycles and subjects. For the numerical results for uncorrelated errors and correlated errors, the hemodynamic response function with default parameters from the MATLAB toolbox BVQXtools_v08d (Brain Innovation, Maastricht, The Netherlands) was applied, sampled with rate 2 s over a period of 32 s and scaled to have a peak amplitude of 1. The toolbox can be downloaded for free from http://support.brainvoyager.com/available-tools/52-matlab-tools-bvxqtools.html. The default HRF has a positive peak at 5 s, a negative peak (undershoot) at 15 s, a positive and negative dispersion of 1 and a positive-to-negative ratio of 6. The exact function and parameters are further described in Maus et al. (2010a). For all calculations, it is assumed that the matrix S equals 1_{N_T} except in the Correlated errors section when the effect of highpass filtering is studied. In Table 1 a list of the applied symbols in all figures is given.

For ease of interpretation the cost ratio C_1/C_2 is given in the figures based on C_2 being the costs per minute scanning time. Then, a cost ratio of $C_1/C_2=12$ means that a subject costs as much as 12 min scanning time.

Optimal design for a given budget

Uncorrelated errors

From Eq. (15) it follows that the number of subjects N can be expressed as the ratio of the total experimental costs and the total

Table 1
List of symbols.

Symbol	Meaning
Q	Number of stimulus types (excluding null events)
N_{BL_0}	Number of events per null block
N_{BL_A}	Number of events per stimulus block
C_T	Total experimental costs
C_1	Costs per subject
C_2	Costs per scanning time, depending on context given per h, min or s
σ_ϵ^2	Error variance (within-subject variance)
σ_q^2	Random-effects variance (between-subject variance)
ρ	Autocorrelation coefficient for AR1 error structure
N	Number of subjects
N_{Copt}	Optimal number of cycles
N_{opt}	Optimal number of subjects
C	Contrast matrix

costs per subject. The obtained expression for N can be substituted in Eq. (12), the trace of $Cov\left(C\hat{\beta}\right)$, for N. In the next step, the trace of $Cov\left(C\hat{\beta}\right)$ is minimized with respect to N_C. This leads to the following optimal number of cycles

$$N_{C_{opt}} = \sqrt{\frac{C_1}{C_2}}\sqrt{\frac{\sigma_\epsilon^2}{\sigma_q^2}}\sqrt{\frac{1}{trace(CDC^T)}}\sqrt{\frac{trace\left(CM^{-1}C^T\right)}{\mathrm{SOA}\left(N_{BL_0}+QN_{BL_A}\right)}}. \tag{16}$$

Using the expression for N obtained by the cost function in Eq. (15) and substituting $N_{C_{opt}}$ for N_C in this expression results in the following optimal number of subjects

$$N_{opt} = \frac{C_T}{C_1+\sqrt{C_1C_2}\sqrt{\frac{\sigma_\epsilon^2}{\sigma_q^2}}\sqrt{\frac{1}{trace(CDC^T)}}\sqrt{trace(CM^{-1}C^T)\mathrm{SOA}\left(N_{BL_0}+QN_{BL_A}\right)}}. \tag{17}$$

If the ratio $\sigma_\epsilon^2/\sigma_q^2$ increases, the optimal number of cycles increases, and as a consequence the optimal number of subjects decreases. Higher fixed costs C_1 per subject or lower costs C_2 per scanning time result in a higher ratio C_1/C_2 so that more cycles are optimal for the A-optimality criterion given costs C_T. Furthermore, it can be seen that the optimal number of cycles is independent of the total costs C_T. Eqs. (16) and (17) are comparable to results in Moerbeek et al. (2008, p. 183, Table 4.1) for optimal nested designs with regard to the effects of cost ratio, variance ratio and costs. The factors $trace(CM^{-1}C^T)$, N_{BL_0}, N_{BL_A} and SOA describe properties of fMRI experiments.

To calculate the optimal number of cycles in Eq. (16) and the optimal number of subjects in Eq. (17) for individual stimulus effects, we need to determine $trace(M^{-1})$ as C equals in this case the identity matrix I_Q of dimension Q. The trace of the matrix M^{-1} in Eqs. (16) and (17) however depends on the block length N_{BL_0}, the stimulus frequency p and the hemodynamic response function and no closed form expression is available for $trace(M^{-1})$ depending on these factors. Thus, we will replace M^{-1} by an approximation to derive more direct and insightful formulae for the optimal number of cycles and subjects. It is shown in Appendix A.2 that $trace(M^{-1})$ can be approximated by

$$trace\left(M^{-1}\right)\approx\frac{Q-Q^2p+Qp}{pN_{BL_0}H}, \tag{18}$$

where $H=\sum_{m=1}^{N_h}\sum_{n=1}^{N_h}h_mh_n$ depends on the hemodynamic response vector h. The hemodynamic response vector h is the hemodynamic response function sampled with rate TR. Based on

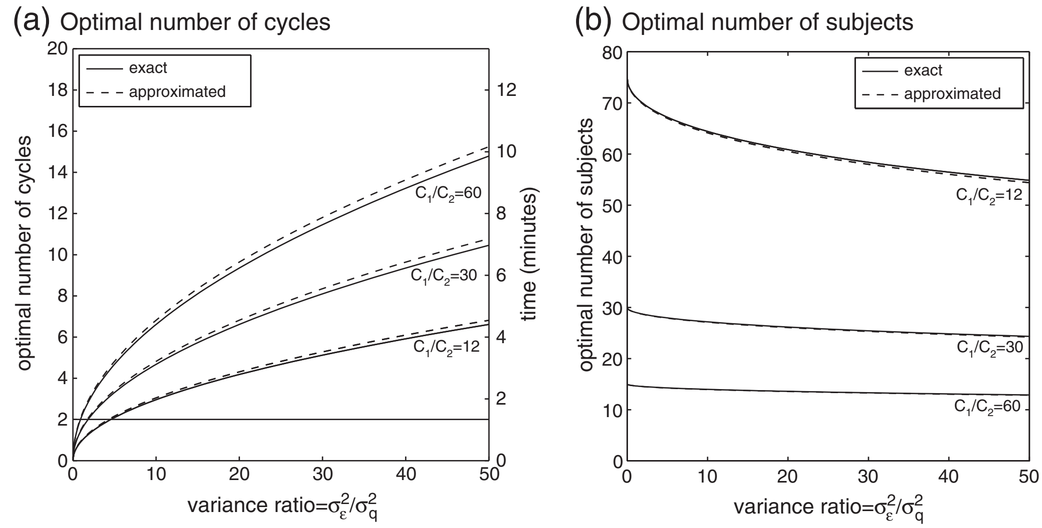

Fig. 1. Optimal number of cycles and subjects for uncorrelated errors, $C=I_2$, $Q=2$, TR = 2 s, SOA = 2 s, $N_{BL_0}=8$ (16 s) and $N_{BL_A}=6$ (12 s), $C_T=€6000$, $C_2=€400/\text{h}\approx€6.66/\text{min}$ and $C_1=€80$, $C_1=€200$ or $C_1=€400$. The horizontal line indicates the minimum number of cycles of 2. The exact values are based on Eqs. (16) and (17) while the approximated values are based on Eqs. (19) and (20). On the right *y*-axis for the optimal number of cycles the corresponding scanning time in minutes is given.

the approximation of $trace(M^{-1})$ in Eq. (18), the following optimal number of cycles and optimal number of subjects are obtained for minimization of $trace\left(Cov\left(\hat{\beta}_G\right)\right)$

$$N_{C_{opt}} = \sqrt{\frac{C_1}{C_2}}\sqrt{\frac{\sigma^2_\epsilon}{\sigma^2_q}}\sqrt{\frac{(1-pQ)(1-pQ+p)}{p}}\frac{1}{\sqrt{HN^2_{BL_0}\text{SOA}}}, \tag{19}$$

$$N_{opt} = \frac{C_T}{C_1+\sqrt{C_1C_2}\sqrt{\frac{\sigma^2_\epsilon}{\sigma^2_q}}\sqrt{\frac{(1-pQ+p)}{p(1-pQ)}}\sqrt{\frac{\text{SOA}}{H}}}. \tag{20}$$

Eqs. (19) and (20) can be seen by substituting $trace(M^{-1})$ with its approximation in Eq. (18) and by substituting N_{BL_A} with $[p/(1-pQ)]N_{BL_0}$ in Eqs. (16) and (17). The equation $N_{BL_A}=[p/(1-pQ)]N_{BL_0}$ follows from the definition of the stimulus frequency $p=N_{BL_A}/(Q\cdot N_{BL_A}+N_{BL_0})$. The approximation of M^{-1} in Appendix A.2 was tested on its precision for $trace(M^{-1})$ and results indicate that the approximation can be imprecise for the off-diagonal elements in M^{-1}. These off-diagonal elements were not needed for $trace(M^{-1})$ but would be needed for $trace(CM^{-1}C^T)$. Therefore, only approximated formulae for N_{Copt} and N_{opt} for the case $C=I_Q$ are presented.

In Fig. 1 the optimal number of cycles and subjects for a design example and costs example are given. The optimal estimation of two individual stimulus effects, i.e., $C=I_2$, is considered. The A-optimal stimulus frequency of $p=1/\left(\sqrt{2}+2\right)\approx0.30$ for estimation of stimulus effects was applied. It can be seen in Fig. 1 that the optimal number of cycles and subjects based on Eqs. (16) and (17) were similar to the optimal number of cycles and subjects based on the approximation as given by Eqs. (19) and (20). The approximated results were also similar to the results based on Eqs. (19) and (20) for

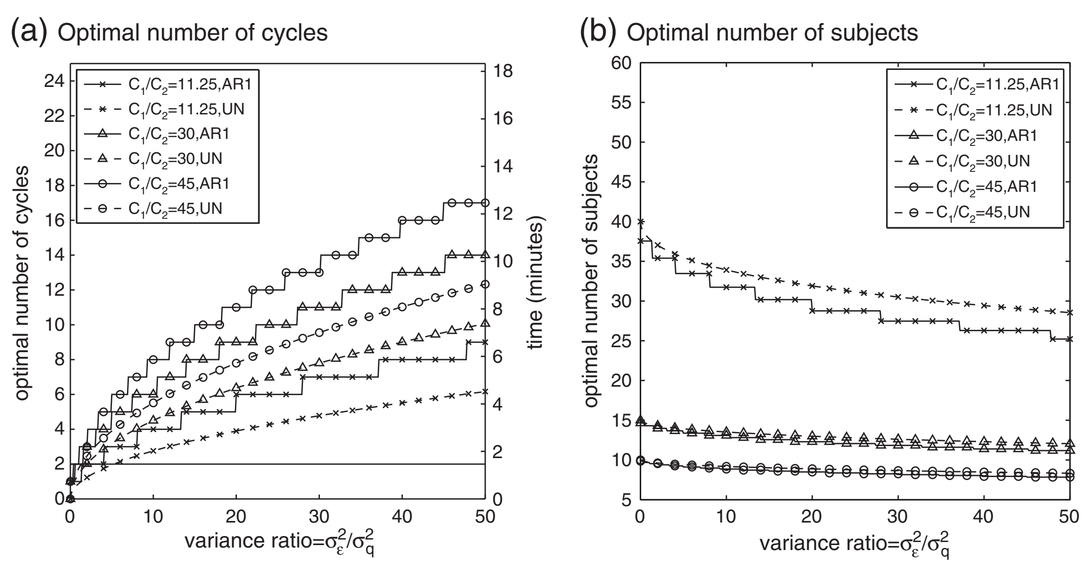

Fig. 2. Optimal number of cycles and subjects for uncorrelated errors (UN) and AR1 ($\rho=0.3$). The results for uncorrelated errors are based on Eqs. (16) and (17) with contrast matrix $C=I_3$ and C_2 in costs per second. Costs were $C_T=€6000$, $C_2=€800/\text{h}\approx€13.33/\text{min}$ and $C_1=€150$, €400 or €600. The number of stimulus types was $Q=3$, null block length was 14 s, stimulus block length was 10 s, TR and SOA were 2 s. The horizontal lines indicate the minimum number of cycles of 2. On the right *y*-axis for the optimal number of cycles the corresponding scanning time in minutes is given.

different costs, different number of stimulus types and stimulus frequency $p = 1/(Q+1)$ which results in stimulus block lengths equal to the null block length, while the A-optimal stimulus frequency results in longer block lengths for null blocks. Fig. 1 illustrates that for increasing cost ratio the optimal number of cycles increases whereas the optimal number of subjects decreases. This is generally true for the optimal number of cycles whereas for the optimal number of subjects this relationship holds because C_2 is fixed. Furthermore, the optimal number of cycles increases and the optimal number of subject decreases for an increasing ratio of within-to-between-subject variance. The described relationship between optimal number of cycles/subjects and cost ratio was also found for $p = 1/(Q+1)$. In addition, the described relationship between the optimal number of cycles/subjects and variance ratio was the same for $p = 1/(Q+1)$ as for $p = 1/(Q + \sqrt{Q})$.

The optimal number of cycles and subjects were calculated in Fig. 1 for given total costs C_T = €6000, costs per scanning time C_2 = €400/h ≈ €6.66/min ≈ €0.11/s and varying costs per subject C_1. The optimal number of subjects and cycles were calculated based on Eqs. (16), (17), (19) and (20) for different cost ratios C_1/C_2 and C_2 given as costs per second, i.e., C_2 = €400/(3600 s) ≈ 0.11/s. A design with two stimulus types, null block length equal to 16 s and stimulus block length equal to 12 s was assumed. This corresponds to the A-optimal stimulus frequency of $p = 1/\left(\sqrt{2} + 2\right) \approx 0.30$ for estimation of stimulus effects. The variance ratio $\sigma_\epsilon^2/\sigma_q^2$ varied in a range from 0 to 50 as this covers a wide range of variance ratios. For values close to 0, σ_q^2 dominates the variance ratio whereas for values close to 50, σ_ϵ^2 dominates the variance ratio.

The number of cycles should be at minimum 2 so that linear trends can be estimated. The maximum run time can be restricted by an assumed realistic maximum length time, e.g., 15 min to avoid boredom and fatigue of subjects. If the optimal number of cycles is unrealistic because the assumed realistic maximum run time is exceeded, the number of cycles should be decreased to the highest possible value for the number of cycles and the number of subjects has to be adapted according to this value for the number of cycles such that the total costs C_T remain the same. In the given example with scanning time of up to 10 min, we assumed the scanning time to be realistic. Generally, as the optimal number of cycles and subjects obtained by Eqs. (16) and (17) will be mostly no integer values, the optimal values have to be rounded.

Correlated errors

In Fig. 2a the optimal number of cycles for correlated errors and uncorrelated errors is displayed for a particular design with three stimulus types, null block length equal to 14 s and stimulus block length equal to 10 s (Maus et al., 2010b). The autocorrelation ρ was set to 0.3 for Fig. 2 as this value is a middle value of autocorrelation for

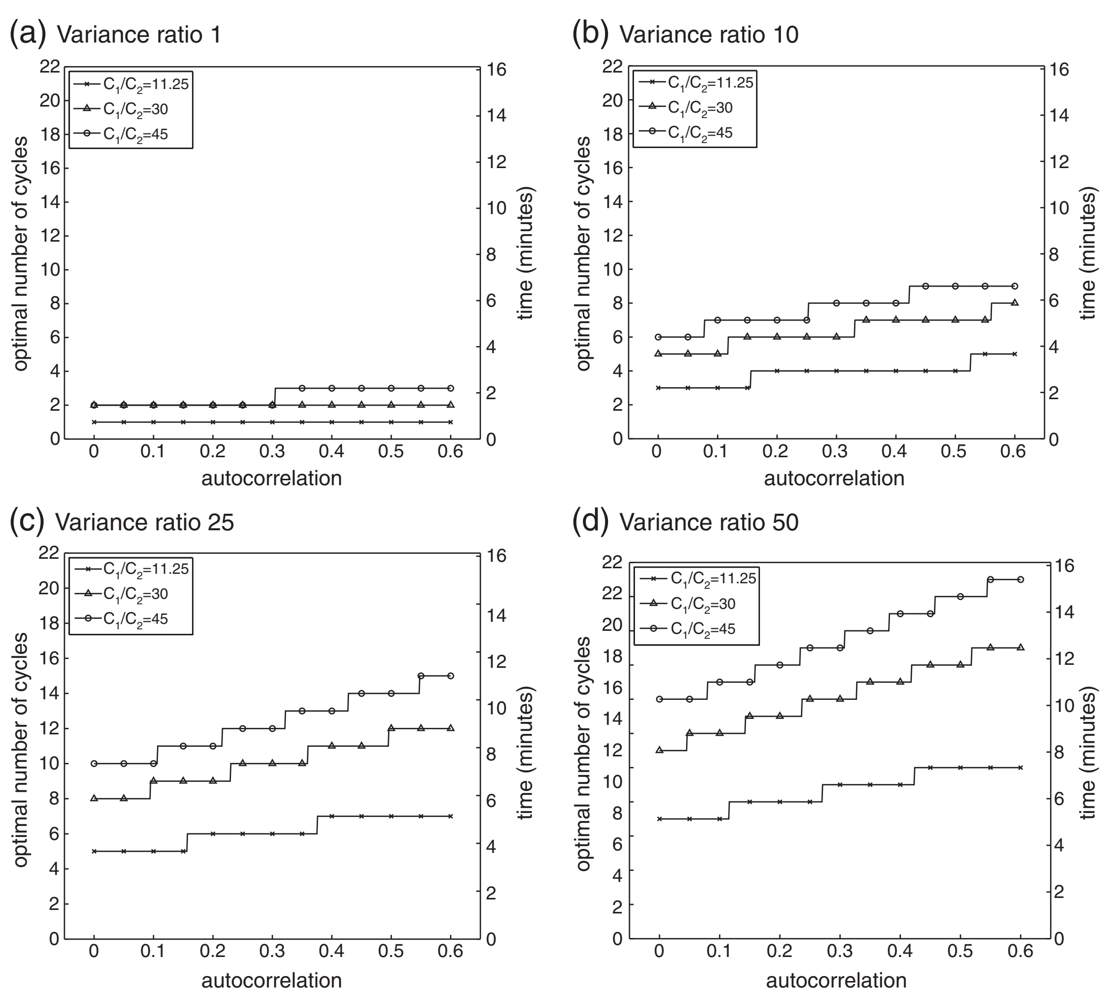

Fig. 3. Optimal number of cycles for different variance ratios $\sigma_\epsilon^2/\sigma_q^2$ depending on autocorrelation ρ. The following costs were assumed: C_T = €6000, C_2 = €800/h ≈ €13.33/min, C_1 = €150, €400 or €600. The number of stimulus types was $Q = 3$, contrast matrix C was equal to I_3, null block length was 14 s, stimulus block length was 10 s, TR and SOA were 2 s.

TR = 2 s. Integer values were used for the number of cycles in the numerical calculation for correlated errors. Therefore, the optimal number of cycles and the optimal number of subjects in Fig. 2 change in steps. Fig. 2 illustrates that the behavior of the optimal number of cycles and the optimal number of subjects as a function of the variance ratio and the cost ratio is similar for uncorrelated and correlated errors. For example, the relationship between the optimal number of cycles and the variance ratio is curvilinear (square root) for uncorrelated and correlated errors. Another example is that for uncorrelated and correlated errors the optimal number of subjects decreases with increasing variance ratio and stabilizes for high variance ratios.

In Fig. 3 the optimal number of cycles and in Fig. 4 the optimal number of subjects are shown for varying autocorrelation ρ, different variance ratios $\sigma_\epsilon^2/\sigma_q^2$ and different cost ratios C_1/C_2. Fig. 3 illustrates that for increasing autocorrelation the optimal number of cycles increases and the optimal number of subjects decreases except for the smallest variance ratio $\sigma_\epsilon^2/\sigma_q^2 = 1$ in combination with the lower cost ratios. This increasing and decreasing effect of the autocorrelation can also be seen in Fig. 2 by comparing the results for uncorrelated error ($\rho = 0$) with correlated error ($\rho = 0.3$). Further just as in Fig. 2, is it seen in Figs. 3 and 4 that for a given autocorrelation and a given variance ratio the optimal number of cycles increases with increasing cost ratio whereas the optimal number of subjects decreases. Furthermore, the optimal number of cycles increases and the optimal number of subjects decreases in Figs. 3 and 4 for increasing variance ratio. This generalizes the relationships in Fig. 2 for $\rho = 0$ and $\rho = 0.3$ to other autocorrelation values.

An increase of optimal number of cycles for increasing autocorrelation was expected since for stronger correlated errors the different scans are stronger correlated and as a consequence each scan will provide less additional information than a scan in a model with lower autocorrelation. Therefore, more cycles are necessary for strongly correlated errors to obtain the same information as for weakly correlated errors. As a result of the increased optimal number of cycles, the optimal number of subjects decreases for increasing autocorrelation (see Fig. 4 for an example).

In Table 2 the effect of highpass filtering on the optimal design is studied. A highpass filter was included into the matrix S by a discrete cosine transform (DCT) set. We considered the effect of number of DCT basis functions from 1 to 5 basis functions for different autocorrelation coefficients ρ. The first basis function is a constant and corresponds to the vector $S = 1_{N_T}$ which was used for the analytical derivations for uncorrelated errors. We considered here the same experimental parameters (TR = 2 s, SOA = 2 s, number of stimulus types $Q = 3$, block length null equal to 14 s and stimulus block length equal to 10 s) as before, total costs of $C_T = €6000$, costs per subject of $C_1 = €200$ and costs per hour scanning time $C_2 = €400/$h. The individual stimulus effects were of interest, thus $C = I_3$. Different variance ratios $\sigma_\epsilon^2/\sigma_q^2$ and different autocorrelations ρ were

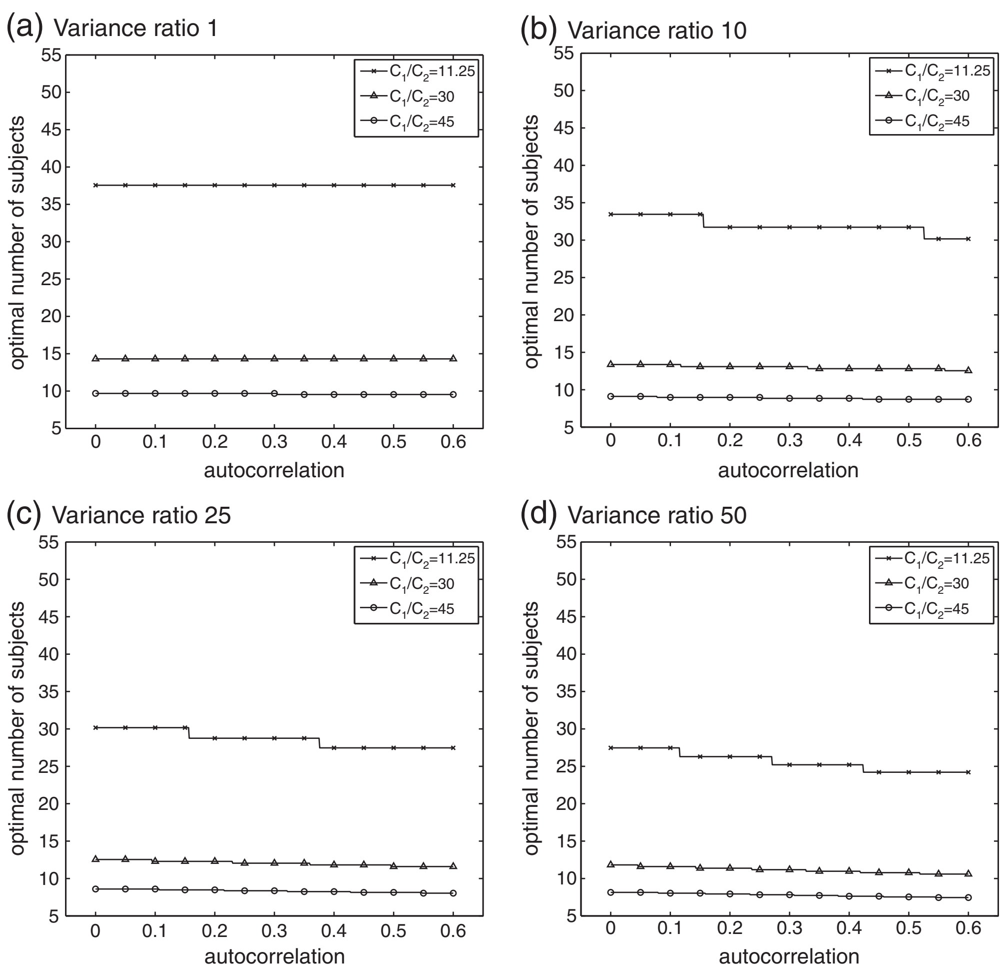

Fig. 4. Optimal number of subjects for different variance ratios $\sigma_\epsilon^2/\sigma_q^2$ depending on autocorrelation ρ. The following costs were assumed: $C_T = €6000$, $C_2 = €800/\text{h} \approx €13.33/\text{min}$, $C_1 = €150$, €400 or €600. The number of stimulus types was $Q = 3$, contrast matrix C was equal to I_3, null block length was 14 s, stimulus block length was 10 s, TR and SOA were 2 s.

Table 2
Optimal number of cycles for different numbers of DCT basis functions, (bf = basis function(s)).

	(a) Variance ratio = 1					(b) Variance ratio = 10				
	1 bf	2 bf	3 bf	4 bf	5 bf	1 bf	2 bf	3 bf	4 bf	5 bf
$\rho = 0$	2	2	2	2	3	5	5	5	5	5
$\rho = 0.2$	2	2	2	2	3	6	6	6	6	6
$\rho = 0.4$	2	2	2	3	3	7	7	7	7	7
$\rho = 0.6$	2	2	2	3	3	8	8	8	8	8

considered. It can be seen in Table 2 that highpass filtering had in some cases an increasing effect on the number of cycles, i.e., the optimal number of cycles increased in some cases if the number of basis functions increased. As a consequence the optimal number of subjects decreased for an increasing number of basis functions. Further calculations showed that for a fixed number of basis functions and $\rho = 0.3$, the optimal number of cycles increased and the optimal number of subjects decreased for increasing variance ratio. This is consistent with the previous results for uncorrelated errors and correlated errors without highpass filtering. The results for a highpass filter using Legendre polynomials were similar.

Effect of deviations from the optimal design for a given budget

It can help researchers to plan their experiment and budget more efficiently if they have knowledge about the effect on design efficiency of deviations in the number of cycles and subjects from the optimal design. We will focus here on the effect of small deviations from the optimal design. The effect of deviations in the number of cycles and number of subjects from the optimal number of cycles and subjects

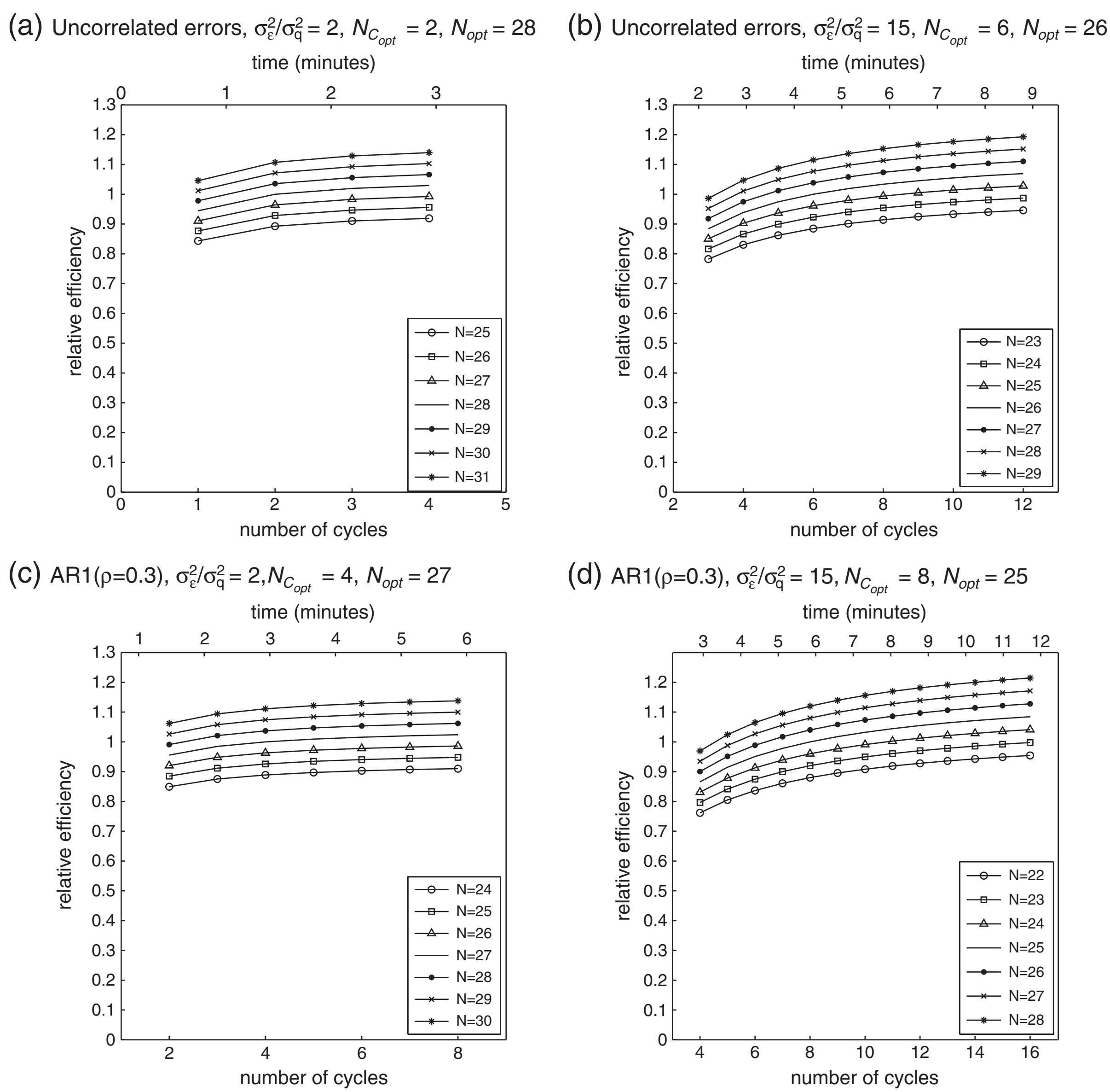

Fig. 5. Effect of small deviations from the optimal design for individual stimulus effects ($C = I_3$). A TR of 2 s and the following design characteristics were assumed: $Q = 3$, SOA = 2 s, $N_{BL_0} = 7$ (14 s), $N_{BL_A} = 5$ (10 s). The total costs C_T were set to €6000. The costs per subject were $C_1 =$ €200 and the costs per scanning time were $C_2 =$ €400/h ≈ 6.67/min.

can be evaluated by the relative efficiency. Assuming uncorrelated errors, it follows with Eqs. (12) and (13) that the relative efficiency of a design ξ with N_C cycles and N subjects versus the optimal design ξ^* with N_{Copt} and N_{opt} is given by:

$$RE(\xi|\xi^*) = \frac{N}{N_{opt}} \cdot \frac{\frac{\sigma_\epsilon^2}{\sigma_q^2}\frac{1}{N_{C_{opt}}} \cdot trace\left(CM^{-1}C^T\right) + trace\left(CDC^T\right)}{\frac{\sigma_\epsilon^2}{\sigma_q^2}\frac{1}{N_C} \cdot trace\left(CM^{-1}C^T\right) + trace\left(CDC^T\right)}. \tag{21}$$

Keeping N fixed and increasing or decreasing N_C towards $N_{C_{opt}}$, the relative efficiency approaches N/N_{opt}. Furthermore, the relative efficiency increases for increasing N_C but this increase attenuates for higher N_C as $(\sigma_\epsilon^2/N_C)trace(CM^{-1}C^T)$, the trace of the within-subject part of the covariance matrix of $C\hat{\beta}_G$, decreases and $trace(CDC^T)$, the trace of the between-subject part of the correlation matrix of $C\hat{\beta}_G$, then dominates the denominator of the second part in Eq. (21).

In Fig. 5 the relative efficiencies for several designs versus the optimal design for estimation of individual effects ($C=I_Q$) are given. The optimal number of cycles and subjects were calculated with the numerical program for correlated errors setting ρ to 0 for uncorrelated errors and ρ to 0.3 for AR1. The designs with small deviations in the number of cycles and subjects from the optimal design were chosen such that the number of cycles varied between $N_{C_{opt}}/2$ and $N_{C_{opt}} \cdot 2$, and the number of subjects varied between $N_{opt}-3$ and $N_{opt}+3$ in Fig. 5.

It can be seen in Fig. 5 that several designs have a high relative efficiency, e.g., between 0.9 and 1.1. Thus, these designs with small deviations from the optimal design perform well in comparison to the optimal design but may not make full use of the budget of €6000 or exceed the budget. Fig. 5 further illustrates that for a fixed number of subjects the relative efficiency attenuates if the number of cycles increases. This has consequences for designs with fewer subjects than the optimal design. For example in Fig. 5b for a variance ratio of 15, the design with 23 subjects does not reach an efficiency very close to 1, even when the number of cycles is twice the optimal value 6. The attenuation was explained analytically for uncorrelated errors by Eq. (21). For correlated errors, it can also be seen in Fig. 5 that the relative efficiency flattens off for higher numbers of cycles.

Another important comparison is between designs which employ the full budget C_T and the optimal design for that budget. In Fig. 6 we present the relative efficiencies versus the optimal design of designs with the same budget C_T as the optimal design. It can be seen that several designs for both variance ratios and both error structures have a high relative efficiency above 0.9 in Fig. 6. For a given value of N_C, the number of subjects was calculated for costs of $C_T=€6000$ using the expression for N in dependence of N_C from Eq. (15) and rounded down to the nearest integer. In this way, the maximum possible number of subjects for a given number of cycles and costs was obtained and as a consequence, $trace\left(Cov\left(\hat{\beta}_G\right)\right)$ was minimized. As the number of cycles and the number of subjects were integer values, the full costs were often not completely employed. The costs varied in the range between €5740 and €6000 for both variance ratios. The number of cycles was chosen close to the optimal number of cycles for the different variance ratios of 2 and 15.

Optimal design for a given power

We will now explain how a given power can be obtained for a minimum budget. Firstly, the optimal number of cycles and subjects for a given budget have to be calculated. In our example a fixed budget $C_T=€4000$ is used together with the following parameters: $Q=2$, TR = SOA = 2 s, null block length 14 s, stimulus block length 10 s, $C=[1\ -1]$, $C_1=€200$ and $C_2=€400/h\approx 6.67/min\approx 0.11/s$, $\rho=0.3$, $\sigma_\epsilon^2/\sigma_q^2=2$ and between random-effects correlation $D_{12}=0$. The interest is in the parameter contrast $\theta=\beta_{G1}-\beta_{G2}$. Using numerical computations as outlined in the Numerical calculations section, the optimal number of cycles is then equal to 3 and the optimal number of subjects is equal to $18.93\approx 19$.

Secondly, the power of this design with the derived optimal number of cycles and subjects is determined. If the power is too low or too high, the variance $Var\left(\hat{\theta}\right)$, which guarantees the given level of power, has to be calculated. The desired power in our example is 80% and the null-hypothesis is $H_0:\theta=0$ versus the alternative hypothesis

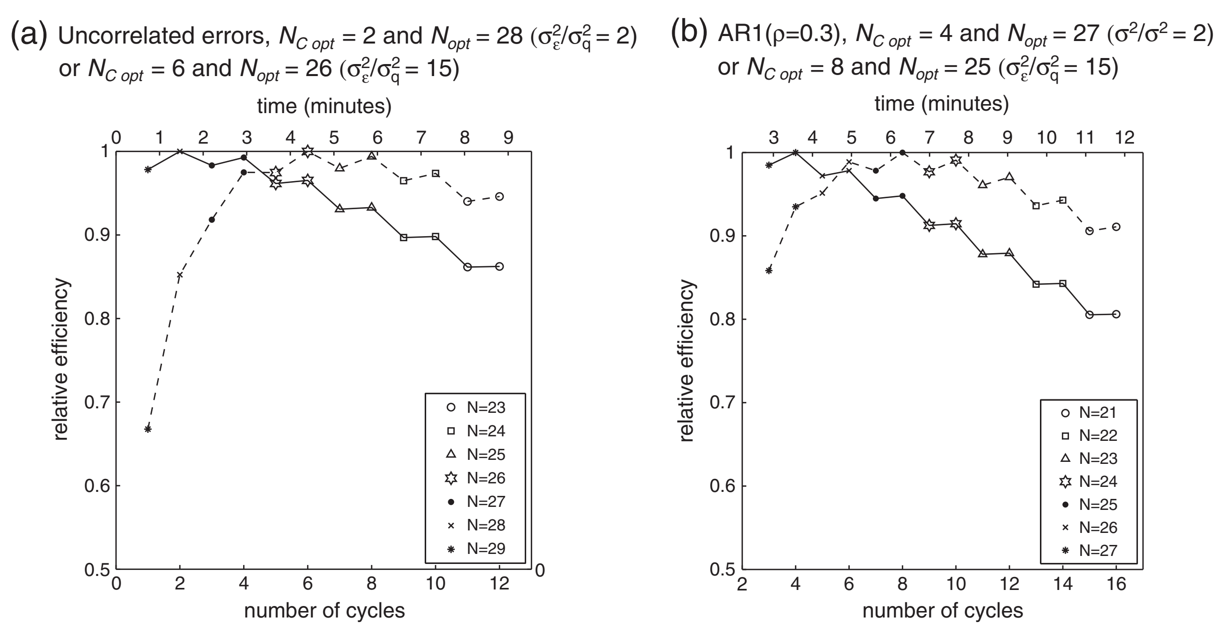

Fig. 6. Relative efficiency versus the optimal design of designs employing the full costs. The solid line is for variance ratio $\sigma_\epsilon^2/\sigma_q^2=2$ and the dashed line is for variance ratio $\sigma_\epsilon^2/\sigma_q^2=15$. The following assumptions were made: $Q=3$, $C=I_3$, TR = 2 s, SOA = 2 s, $N_{BL_0}=7$ (14 s), $N_{BL_A}=5$ (10 s), $C_1=€200$, $C_2=€400/h=€6.66/min$.

$H_A: \theta = \Delta$. Specifying the effect size in percent signal change $\Delta = 1\%$, $\sigma_\epsilon^2 = 2$, $\sigma_q^2 = 1$ and a conservative significance level $\alpha = 0.005$ to correct for testing at multiple voxels, a power of 63.48% is obtained which is too low. The following formula calculates the variance which gives a power of approximatively $(1-\delta) = 0.80$:

$$Var\left(\hat{\theta}\right) = \left(\frac{\Delta}{z_{1-\delta} + z_{1-\alpha}}\right)^2. \tag{22}$$

The $100(1-\alpha)$ percentile respectively the $100(1-\delta)$ percentile of the standard normal distribution is given by $z_{1-\alpha}$ respectively by $z_{1-\delta}$. In our example, $Var\left(\hat{\theta}\right)$ must be 0.0856 to obtain 80% power for effect size $\Delta = 1\%$.

Thirdly, the budget for this new variance has to be calculated. If the power is too low, the budget has to be raised and if the power is unnecessarily high, the budget has to be decreased. The new budget is given by

$$C_T = \frac{C_1 + C_2 \cdot T_S}{Var\left(\hat{\theta}\right)} \cdot \sigma_q^2 \left(\frac{\sigma_\epsilon^2}{\sigma_q^2} trace\left(C\left(Z^{*T}Z^*\right)^{-1}C^T\right) + trace\left(CDC^T\right)\right). \tag{23}$$

This equation follows from Eqs. (11) and (15). The number of cycles is fixed in Eq. (23) to the previously calculated optimal value of 3 as Eq. (16) indicates that the optimal number of cycles does not depend on the budget C_T. Eq. (23) results in a new budget of €5130.74. The optimal number of subjects can then be calculated by a new numerical computation which results in $N_{opt} = 24.28 \approx 24$. This example shows how the optimal design with a minimum budget for a given power can be obtained.

Discussion and conclusions

The aim of this paper was to find the optimal design for multi-subject fMRI studies with respect to the scanning time and the number of subjects. We found the combination of number of subjects and number of cycles which was optimal according to the A-optimality criterion when both the number of subjects and the scanning time were restricted by experimental costs. Thus, our optimal design is the design which minimizes the trace of the group effects estimators' covariance matrix, respectively the trace of the contrasts estimators' covariance matrix, but simultaneously does not exceed experimental costs for a given cost function. Analytical solutions for the optimal number of cycles and optimal number of subjects were presented for blocked designs and uncorrelated errors. The analytical results are useful to describe the general relationships between the optimal number of subjects and cycles and the factors variance ratio, cost ratio and costs. These relationships were also valid for the numerical results for correlated errors with and without highpass filtering.

For correlated errors more cycles and fewer subjects are optimal than for uncorrelated errors. This can be explained by the decrease of information from individual scans when correlated errors are assumed instead of uncorrelated errors. The higher the correlation between scans is, the stronger the increase in optimal number of cycles and the stronger the decrease in optimal number of subjects from uncorrelated errors to correlated errors will be. The results for uncorrelated errors and correlated errors further show that for an increasing cost ratio, higher number of cycles become optimal. The explanation is that for higher cost ratios, cycles are becoming cheaper in relation to subjects and thus more cycles are optimal within the same total experimental costs.

The results for both error structures also illustrate that the optimal number of cycles increases and the optimal number of subjects decreases with increasing variance ratio. This can be explained by considering separately the three causes for an increasing variance ratio. Firstly, the variance ratio increases when σ_ϵ^2 increases while σ_q^2 is fixed or increases less strongly than σ_ϵ^2. In this first case a higher number of cycles is needed to estimate subject-specific effects efficiently from each subject which leaves less budget for subjects. Therefore, the optimal number of subjects decreases. Secondly, the variance ratio increases when σ_q^2 decreases while σ_ϵ^2 is fixed or decreases less strongly than σ_q^2. In this second case a lower number of subjects becomes optimal so that more budget is available for cycles and the optimal number of cycles increases. Thirdly, the variance ratio increases when σ_ϵ^2 increases and σ_q^2 decreases. As in the first case, more cycles are needed to estimate subject-specific effects efficiently and as in the second case, fewer subjects are needed so that more budget is available for cycles.

The effect of small deviations from the optimal design was studied. It was seen that small deviations from the optimal number of cycles and subjects had little effect on the relative efficiency. Keeping the number of subjects fixed and increasing the number of cycles, increases the relative efficiency, but this increase levels off with increasing number of cycles (see Fig. 5). For uncorrelated errors, this behavior of the relative efficiency was explained by the analytical formula for the relative efficiency in the section about effects of deviations from the optimal design for a given budget. Such a point of diminishing return has also been found by Desmond and Glover (2002) and Mumford and Nichols (2008) for power analyses. When the number of subjects and the number of cycles were chosen such that the total costs of €6000 were used, it was seen that several designs had high relative efficiencies. Similarly, Mumford and Nichols (2008) found several designs with a power of 80% and costs of around $7600.

Our results are based on minimizing the sum of the variances of the effects of interests. For one effect of interest, e.g., a single stimulus effect or a contrast between several stimulus effects, minimizing the variance of the estimator for this effect will lead to maximum power to test the null-hypothesis of zero effect or of a given effect value via a *t*-test. The value of this maximum power will of course depend on the effect size, the significance level α, the within-subject and between-subject variance and might not reach a power of for example 80%. However, instead of maximizing the power for a given budget our approach can also be used to find a design for a given power minimizing the budget. For multiple effects of interest, our method minimizes the sum of the variances of the estimators for the effects and the chosen optimal design is thus efficient for multiple univariate tests. Our method is thus not a replacement for power analyses but complementary to power analyses.

Mumford and Nichols (2008) consider the power of *t*-tests for single hypotheses and of *F*-tests for simultaneous testing of multiple hypotheses. While our approach gives the optimal number of subjects and cycles for a given budget and maximum power of a *t*-test, it is not guaranteed that it will result in maximum power of an *F*-test. That can be seen as a limitation of our approach. By the cost function we are taking costs directly into account and find a specific optimal number of cycles and subjects minimizing the trace of the estimator's covariance matrix for a given budget. In contrast, Mumford and Nichols (2008) first calculate several designs with 80% power and afterwards perform costs calculations for these designs. We are focusing on blocked design while Mumford and Nichols (2008) consider event-related and blocked designs.

Desmond and Glover (2002) also focused on blocked designs and calculated the power of *t*-tests for comparison of different conditions in a blocked design. In contrast to our work and Mumford and Nichols (2008), Desmond and Glover (2002) did not consider correlated errors for their power calculations and no convolution with a hemodynamic response function. However, Mumford and Nichols (2008) showed that power calculations are incorrect if temporal correlation and HRF convolution are ignored. Likewise, our results illustrate that the optimal number of subjects and cycles can differ between correlated and uncorrelated errors. Desmond and Glover (2002) provided a more time-intensive simulation based method whereas our approach results in direct analytical formulae for uncorrelated errors and typically less time-intensive numerical computations for correlated errors.

To apply our results, it is important to remember that the cost ratio depends on the unit in time. In Eqs. (16) and (17) the costs C_2 are given per second scanning time. Furthermore, the random-effects variance σ_q^2 depends on the scaling of the HRF. For example, if the amplitude of the HRF is m at its peak and the HRF is scaled by multiplication with $1/m$ to obtain an amplitude of 1, the random-effects variance will become m^2 as large as every random effect b_i will become m times as large. The variance ratio of within- to between-subject variance has to be determined from previous studies. It would further be interesting to find a design which is efficient for a possible range of parameter values, e.g., for a possible range of the variance ratio $\sigma_\epsilon^2/\sigma_q^2$. For example, a maximin design (Atkinson et al., 2007), which is robust against misspecification of $\sigma_\epsilon^2/\sigma_q^2$ in a prespecified range, could be determined. In the following we will present more possible extensions of our work and discuss the assumptions which were made.

It was assumed that the error covariance matrix is the same for all subjects. While this is a simplification, it is a common assumption in optimal design as a parsimonious model is needed as a basis for further optimal design calculations. When the error covariance matrix $\sigma_\epsilon^2\sum$ is estimated from a previous study, heterogeneity can be taken into account. A different error covariance matrix can be estimated per subject and subsequently an average error covariance matrix can be determined for use in our equations. For example for an AR1 structure, the average correlation parameter and average error variance can be used for further calculations. The assumption of equal random-effects variance for all stimulus effects is also useful to have a parsimonious model as basis for optimal design calculation. The analytical results for uncorrelated errors were based on the following assumptions: TR = SOA, only filtering of constant trend ($S = 1_{N_T}$) and inshifting. The assumption of inshifting is explained in Appendix A.1. Further, it is discussed in Appendix A.1 why this assumption has little to no effect on the results. If the other assumptions do not hold, the numerical program can be used to determine the optimal number of cycles and subjects as no assumption on the TR and SOA or the degree of highpass filtering are necessary for the MATLAB code.

Another often used optimality criterion is the D-optimality criterion which minimizes the determinant of the covariance matrix of the estimators for the unknown parameters (Atkinson et al., 2007). One advantage of the D-optimality criterion is that the D-optimal design is invariant to linear transformation of the regressors in the model. However, the regressors for the stimulus effects β_{Gi} all depend on the scaling of the hemodynamic response function and the A-optimal design is, like the D-optimal design, invariant to scaling of the HRF. Future work could include the application of the D-optimality criterion. For this extension the off-diagonal elements of the random-effects correlation matrix D in Eq. (9) have to be considered not only for contrasts but also for individual stimulus effects. For the A-optimality criterion, consideration of the off-diagonal elements is only necessary for contrasts. Furthermore, an extension to multiple runs per session is important. Our current results consider experiments with one run and one session. Increasing the number of runs will have a decreasing effect on the trace of the covariance matrix of the estimators but likely this effect will attenuate for higher number of runs similar to the effect of the number of cycles on the relative efficiency.

Another extension is the consideration of across-subject contrasts. For example, Suckling et al. (2008) studied the power for comparison of two groups in a two-center repeated-measures pharmacological fMRI study and Schouten (1999) presents the optimal allocation ratio of subjects into two groups depending on the costs in a clinical trial. The difference in means between two independent groups can be considered via an unpaired t-test. The variance of such a between-subject contrast is a pooled sum of the variances of the group-specific estimates. Assuming equal subjects costs, scanning time costs, error variance, autocorrelation and random-effects variances for both groups, the costs per group will be half of the total costs and equal sample sizes, i.e., equal number of subjects and equal scanning time, per group are optimal. The variances of the group-specific estimates can then be minimized by our approach using $C_T/2$ per group and the optimal number of cycles and subjects per group can be determined. By minimizing the variances of the group-specific estimates, we minimize the variance of the between-subject contrast. Unequal costs, autocorrelation or variances between groups affect the optimal budget split between groups, but given a budget split, optimization can be performed per group to optimize between-group estimation.

Summarizing, we presented a method to determine the optimal number of cycles and subjects for a blocked multi-subject fMRI experiment. This method takes design efficiency and costs into consideration. Therefore, efficiency can be gained and costs can be reduced by using this method to determine the optimal number of cycles and optimal number of subjects. Typically, fMRI noise is temporally correlated and highpass filtering is applied. Therefore, we recommend to use numerical computations for determination of the optimal number of subjects and cycles. The analytical formulae are useful to describe the general relationships between the optimal number of subjects and cycles and other parameters.

Acknowledgment

The first author is supported by the Netherlands Organization for Scientific Research (NWO, grant number 400-05-063).

Appendix A. Expression for $Z^{*T}Z^*$ in Eq. (11) assuming uncorrelated errors

Appendix A.1. General expression

We will show that the matrix $Z^{*T}Z^*$ can be expressed as $N_C \cdot M$, where N_C is the number of cycles of $A_1 \ldots A_Q N$ and M/σ_ϵ^2 is the $Q \times Q$ information matrix of one cycle $A_1 \ldots A_Q N$. The matrix $Z^{*T}Z^*$ is given by

$$Z^{*T}Z^* = Z^T(I - P_{VS})Z. \tag{A.1}$$

For uncorrelated errors, it holds that $P_{VS} = P_S$ and $V = I_{N_T}$. Furthermore, we assume that $S = (1, \ldots, 1)^T$. The projection matrix P_S projects onto the space spanned by the vector S. The matrix P_S equals $(1/N_T) \cdot J_{N_T \times N_T}$ with $J_{N_T \times N_T}$ being the $N_T \times N_T$ matrix with all elements equal to 1. Thus, $I - P_S$ is a centering matrix here, and $Z^* = (I - P_S)Z$ is the centered design matrix.

The design matrix $Z = X(I_Q \otimes h)$ is obtained by multiplication of the matrix X with the Kronecker product $I_Q \otimes h$. The matrix X contains 0s and 1s indicating a stimulus type at a certain time point and will be explained in detail below. The vector $h = (h_1, \ldots, h_{N_h})^T$ is the hemodynamic response function sampled at N_h time points with sampling rate TR. To derive Eq. (A.6) and as a consequence the analytical results for uncorrelated errors in the Uncorrelated errors section, the time between trials in one block, the stimulus onset asynchrony (SOA), was assumed to be equal to TR. Therefore, the hemodynamic response function is here sampled with rate TR. The hemodynamic response function, e.g., the canonical double gamma function (Friston et al., 1998; Henson, 2004), is assumed to be known.

The $N_T \times (QN_h)$ dimensional matrix X is a block matrix containing the stimulus convolution matrices X_q (Dale, 1999) for the different stimulus types A_q: $X = (X_1 \ldots X_Q)$. Each $N_T \times N_h$ matrix X_q consists of 0 and 1. The first column of X_q is the actual stimulus sequence s^q where 1 indicates an occurrence of stimulus A_q and 0 indicates no occurrence of stimulus A_q. The other columns of X_q are shifted versions of the stimulus sequence to model convolution of the stimulus sequence s^q with the hemodynamic response function. When the stimulus sequence is shifted, zeros are filled in:

$$X_q = \begin{pmatrix} s_1^q & 0 & \dots & 0 \\ s_2^q & s_1^q & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ s_{N_T}^q & s_{(N_T-1)}^q & \dots & s_{(N_T-N_h+1)}^q \end{pmatrix}. \tag{A.2}$$

We call this zero padding. However, for our following derivations we will assume that the outshifted elements in the columns of the matrices X_q are shifted in again:

$$X_q = \begin{pmatrix} s_1^q & s_{N_T}^q & \dots & s_{(N_T-N_h+2)}^q \\ s_2^q & s_1^q & \ddots & s_{(N_T-N_h+3)}^q \\ \vdots & \vdots & \dots & \vdots \\ s_{N_T}^q & s_{(N_T-1)}^q & \dots & s_{(N_T-N_h+1)}^q \end{pmatrix}. \tag{A.3}$$

This will be called “inshifting”. The assumption of inshifting is a theoretical assumption which is necessary to derive the optimal number of cycles for uncorrelated errors. Due to the assumption of inshifting, the matrices M and $Z^{*T}Z^*$ are only invertible if null blocks are included in the design. We define the centered matrix X^* by $X^* = (I-P_S)X$. The assumption of inshifting had little to no effect on the results when we compared results under the assumption of inshifting to results under the assumption of zero padding. As the null block is assumed at the end of a cycle $A_1 \dots A_Q N$, both assumptions often lead to the same design matrices. Furthermore, the more scans are obtained, the less the matrix $X^{*T}X^*$ under the assumption of inshifting will differ from $X^{*T}X^*$ under the assumption of zero padding.

It follows with the definition of $X^* = (I-P_S)X$ that

$$Z^{*T}Z^* = (I_Q \otimes h)^T X^{*T} X^* (I_Q \otimes h). \tag{A.4}$$

For all stimulus types equal stimulus frequency p is assumed. This means that each stimulus occurs equally often in the experiment and that all blocks for all different stimulus types have the same block length. We denote by N_A the number of times that stimulus A_q occurs in the design for one subject. So, $p = N_A/N_T$ and $X^* = X - p1_{N_T \times (QN_h)}$. The matrix $X^{*T}X^*$ consists of the sub-matrices $X_{q_1}^{*\,T} X_{q_2}^*$:

$$X^{*T}X^* = \begin{pmatrix} X_1^{*T}X_1^* & X_1^{*T}X_2^* & \dots & X_1^{*T}X_Q^* \\ X_2^{*T}X_1^* & X_2^{*T}X_2^* & & \vdots \\ \vdots & & \ddots & \vdots \\ X_Q^{*T}X_1^* & \dots & \dots & X_Q^{*T}X_Q^* \end{pmatrix}. \tag{A.5}$$

In Appendix B it is shown that the entry $\left(X_{q_1}^{*\,T} X_{q_2}^*\right)_{mn}$ can be expressed as

$$\left(X_{q_1}^{*\,T} X_{q_2}^*\right)_{mn} = N_C \cdot (1-p)^2 \cdot P_1 + N_C \cdot (-p)^2 \cdot P_2 + N_C \cdot (1-p)(-p) \cdot P_3. \tag{A.6}$$

The factors P_1, P_2 and P_3 give the number of products of form $(1-p)^2$, $(-p)^2$ and $(-p)(1-p)$. The number of these products depends further on the shift size d which is the shift between column m in $X_{q_1}^*$ and column n in $X_{q_2}^*$:

$$P_1 = max\left(N_{BL_A} - d, 0\right) + max\left(0, d - N_{BL_{0T}}\right) \tag{A.7}$$

$$P_2 = max\left(N_{BL_{0T}} - d, 0\right) + max\left(0, d - N_{BL_A}\right) \tag{A.8}$$

$$P_3 = min\left(d, N_{BL_T} - d, N_{BL_{0T}}\right). \tag{A.9}$$

The shift size d is given as $|(m-n) + |q_1 - q_2| N_{BL_A}|$. The total length of one cycle $A_1 \dots A_Q N$ is $N_{BL_T} = QN_{BL_A} + N_{BL_0}$ and the total length of one null block in sequence s^q is given by $N_{BL_{0T}} = (Q-1)N_{BL_A} + N_{BL_0}$. It follows that

$$X^{*T}X^* = N_C K, \tag{A.10}$$

where K is a $(QN_h) \times (QN_h)$ dimensional matrix dependent on N_{BL_0} and N_{BL_A} but independent of N_C. We define

$$M = (I_Q \otimes h)^T K (I_Q \otimes h). \tag{A.11}$$

The matrix M depends on the block lengths N_{BL_0}, N_{BL_A} and the vector h but is independent of N_C. The vector h itself depends on the chosen hemodynamic response function, the number of sampling points N_h for the HRF and the sampling rate for the HRF which is assumed to be equal to TR. It follows that

$$Z^{*T}Z^* = N_C M. \tag{A.12}$$

Appendix A.2. Approximated expression for matrix M in the Uncorrelated errors section

In this section, we give the following approximation for the $Q \times Q$ matrix M in Eq. (A.11):

$$M = C_Q \cdot N_{BL_0} \cdot f(p, Q) \cdot H, \tag{A.13}$$

where we denote by H the sum $\sum_{m=1}^{N_h} \sum_{n=1}^{N_h} h_m h_n$ and $f(p,Q) = p(1-p)/(1-Qp)$ which is defined for a stimulus frequency p in the interval $[0, 1/Q)$. This approximation for the information matrix M can be seen as an extension to the approximation for the information matrix $X^{*T}X^*$ in Liu and Frank (2004). The $Q \times Q$ matrix C_Q is given by

$$C_Q = \left(1 + \frac{p}{1-p}\right) I_Q - \frac{p}{1-p} 1_Q 1_Q^T. \tag{A.14}$$

The matrix C_Q is of format $Q \times Q$ and gives the correlation between the detrended stimulus sequences s^{q*}. Thus, the diagonals of C_Q contain 1 and the off-diagonal elements of C_Q are equal to $-p/(1-p)$.

We will derive this approximation by assuming $N_h = 1$ and $h = 1$ (first simplification) or $Q = 1$ and $N_{BL_0} = N_{BL_A}$ (second simplification). Let us assume that $N_h = 1$ and $h = 1$. Then

$$\begin{aligned} \left(Z^{*T}Z^*\right)_{qq} &= h^T \left(X_q^{*T} X_q^*\right) h \\ &= N_C \cdot \left((1-p)^2 N_{BL_A} + (Q-1)(-p)^2 N_{BL_A} + (-p)^2 N_{BL_0}\right). \end{aligned} \tag{A.15}$$

It follows from the definition of the stimulus frequency $p = N_{BL_A}/(QN_{BL_A} + N_{BL_0})$ that $N_{BL_A} = p/(1-Qp) N_{BL_0}$. We define $g(p,Q) = p/(1-Qp)$ and it follows that

$$\begin{aligned} \left(Z^{*T}Z^*\right)_{qq} &= N_C \cdot \left((1-p)^2 g(p,Q) N_{BL_0} + (Q-1)(-p)^2 g(p,Q) N_{BL_0} + (-p)^2 N_{BL_0}\right) \\ &= N_C \cdot N_{BL_0} \underbrace{\left((1-p)^2 g(p,Q) + (Q-1)(-p)^2 g(p,Q) + (-p)^2\right)}_{=:f(p,Q).} \end{aligned} \tag{A.16}$$

We calculate that $f(p,Q)=p(1-p)/(1-pQ)$. Furthermore, it can be calculated that

$$\begin{aligned}\left(Z^{*T}Z^{*}\right)_{q_1q_2} &= h^T\left(X_{q_1}^{*\,T}X_{q_2}^{*}\right)h \\ &= N_C\cdot\left((1-p)(-p)N_{BL_A} + (1-p)(-p)N_{BL_A} + (-p)^2(Q-2)N_{BL_A} + (-p)^2N_{BL_0}\right) \\ &= N_C\cdot N_{BL_0}\cdot\left(-2p(1-p)g(p,Q) + p^2(Q-2)g(p,Q) + p^2\right) \\ &= N_C\cdot N_{BL_0}\cdot\left(\frac{-p^2}{1-pQ}\right) \\ &= N_C\cdot N_{BL_0}f(p,Q)C_{Q\,q_1q_2}.\end{aligned} \tag{A.17}$$

This shows Eq. (A.13) for the first simplification ($N_h=1$ and $h=1$).

For the second simplification, we assume that $Q=1$ and $N_{BL_A}=N_{BL_0}$ so that $Z^{*T}Z^{*}$ is a scalar and $p=1/2$. By h_i the ith element in the vector h is denoted. It holds that

$$\begin{aligned}\left(Z^{*T}Z^{*}\right) &= h^TX^{*T}X^{*}h \\ &= \sum_{n=1}^{N_h}\sum_{m=1}^{N_h} h_m\left(X^{*T}X^{*}\right)_{mn}h_n.\end{aligned} \tag{A.18}$$

We know that

$$\begin{aligned}&\left(X^{*T}X^{*}\right)_{mn} \\ &= N_C\Big[(1-p)^2|N_{BL}-shift\ \ size| + (-p)^2|N_{BL}-shift\ \ size| \\ &\quad + 2(-p)(1-p)shift\ \ size\Big] \\ &= N_C\cdot N_{BL}\cdot\left[(1-p)^2\frac{|N_{BL}-shift\ \ size|}{N_{BL}} + (-p)^2\frac{|N_{BL}-shift\ \ size|}{N_{BL}}\right. \\ &\quad \left. + 2(-p)(1-p)\frac{shift\ \ size}{N_{BL}}\right].\end{aligned} \tag{A.19}$$

We make the following approximations which will be more precise for small shift sizes and/or a larger number of trials within a block

$$\frac{|N_{BL}-shift\ \ size|}{N_{BL}}\approx 1 \tag{A.20}$$

$$\frac{shift\ \ size}{N_{BL}}\approx 0. \tag{A.21}$$

We obtain then

$$\left(Z^{*T}Z^{*}\right) = \sum_{n=1}^{N_h}\sum_{m=1}^{N_h} h_m^T\left(X^{*T}X^{*}\right)_{mn}h_n = N_C\cdot N_{BL}\cdot\frac{1}{2}\sum_{m=1}^{N_h}\sum_{n=1}^{N_h} h_mh_n. \tag{A.22}$$

Note that for $p=1/2$ and $Q=1$ it holds that $f(p,Q)=1/2$. This shows Eq. (A.13) for the second simplification.

By the derivations above we motivated our approximation in Eq. (A.13) based on two simplifications. We calculate now the inverse of M from Eq. (A.13). The inverse of M is given by

$$M^{-1} = \frac{1}{N_{BL_0}f(p,Q)H}C_Q^{-1}. \tag{A.23}$$

Using Woodbury's formula (Harville, 1997, p. 424) it follows that for $p<1/Q$

$$C_Q^{-1} = (1-p)I_Q + \frac{p(1-p)}{1-Qp}1_Q1_Q^T. \tag{A.24}$$

Finally, the following expression, which will be useful for simplification of Eqs. (16) and (17), is obtained

$$\begin{aligned}trace\left(M^{-1}\right) &= \frac{1}{N_{BL_0}f(p,Q)H}trace\left(C_Q^{-1}\right) \\ &= \frac{1}{N_{BL_0}f(p,Q)H}\left(Q(1-p) + \frac{Qp(1-p)}{1-Qp}\right) \\ &= \frac{Q-Q^2p + Qp}{pN_{BL_0}H}.\end{aligned} \tag{A.25}$$

We tested this form of the approximation for various combinations of p ($p=1/(Q+\sqrt{Q})$, the A-optimal stimulus frequency for estimation of individual stimulus effects, or $p=1/(Q+1)$), TR respectively SOA (1 s, 2 s), N_h (1, 17, 33), Q (1, 2, 3) and block length of null blocks (10 s, 11 s, …, 30 s). For the A-optimal frequency, the block length of stimulus blocks was given by SOA$\cdot[(1/\sqrt{Q})\cdot(BL_0/\text{SOA})]$, where $[x]$ denotes the nearest integer number to the real number x and BL_0 is the length of null blocks in seconds. For most combinations the approximation was accurate, i.e., the trace of the real matrix M^{-1} was 92% to 107% of the trace of the approximated matrix M^{-1}. For smaller block lengths the approximation was less accurate. In particular, for stimulus frequency $p=1/(Q+1)$ the given range was 92% to 104% for all block lengths higher than or equal to 12 s. For $p=1/(Q+\sqrt{Q})$, the range for $Q=2$ was 93% to 106% only for block lengths higher than or equal to 11 s and the range for $Q=3$ was 94% to 107% only for block lengths higher than or equal to 16 s. Note that for $Q=1$ the stimulus frequencies $p=1/(Q+\sqrt{Q})$ and $p=1/(Q+1)$ are equal and the range is as given above for $p=1/(Q+1)$.

Appendix B. Calculation of $\mathbf{X}^{*T}\mathbf{X}^{*}$ in Eq. (A.6)

The product $X^{*T}X^{*}$ in Eq. (A.6) will be calculated by making firstly two simplifications in Appendix B.1, i.e., the assumption of one stimulus type and the assumption of stimulus block length being equal to null block length. Based on the results in Appendix B.1, the product $X^{*T}X^{*}$ can be calculated generally for one stimulus type in Appendix B.2 without the assumption of stimulus block length being equal to null block length. Finally in Appendix B.3, $X^{*T}X^{*}$ for more than one stimulus type without any assumption on block lengths is given.

Appendix B.1. Calculation for one stimulus type and stimulus block length equal to null block length

We consider blocked designs with equal number of trials N_{BL} per block for null events and stimulus events and with block order AN, i.e., an A block is always followed by a null block N so that the design has form $ANAN\ldots AN$. We will speak of one cycle AN, when we refer to an A block and a null block together. It will be shown that for all $m=1,\ldots,N_h$ and for all $n=1,\ldots,N_h$:

$$\begin{aligned}&\left(X^{*T}X^{*}\right)_{mn} \\ &= N_C\Big[(1-p)^2\cdot|N_{BL}-mod(|m-n|,2N_{BL})| \\ &\quad + (-p)^2\cdot|N_{BL}-mod(|m-n|,2N_{BL})| \\ &\quad + 2(-p)(1-p)\cdot min(mod(|m-n|,2N_{BL}),2N_{BL}-mod(|m-n|,2N_{BL}))\Big] \\ &=: N_CK_{mn}.\end{aligned} \tag{B.1}$$

The entry $(X^{*T}X^*)_{mn}$ of the matrix $X^{*T}X^*$ is given by the scalar product of the mth column with the nth column in X^*. The centered matrix X^* contains the entries $-p$ and $1-p$. Thus, $(X^{*T}X^*)_{mn}$ contains products of the form $(1-p)^2$, $(-p)^2$ and $-p(1-p)$. The shift between column m and n in X^* is given by $mod(|m-n|, 2N_{BL})$. The scalar product of two columns $X^*_{m_1}$and $X^*_{n_1}$ is the same as the scalar product of the columns $X^*_{m_2}$ and $X^*_{n_2}$ if the shift between column m_1 and n_1 is the same as the shift between column m_2 and column n_2. Thus for derivation of Eq. (B.1), it is sufficient to focus on the scalar product of the first column $X^*_{.1}$ of X^* with all columns of X^* (including the first column).

The first column of X^* is the centered stimulus sequence s^*. As there are blocks of A followed by blocks of N in the fMRI experiment, there are also blocks of $(1-p)$ followed by blocks of $(-p)$ in the stimulus sequence s^*. A block of $(1-p)$ followed by a block of $(-p)$ in the stimulus sequence s^* will also be called a cycle. The scalar product $s^{*T}s^*$ equals:

$$N_C \cdot (1-p)^2 \cdot N_{BL} + N_C \cdot (-p)^2 \cdot N_{BL}. \tag{B.2}$$

Per shift of the centered stimulus sequence, the scalar product of s^* and the nth column $X^*_{.n}$ looses one product of the form $(1-p)^2$ and one product of the form $(-p)^2$ and gains one product of the form $(-p)(1-p)$ in Eq. (B.2) per cycle. However, it has to be corrected for higher shift sizes when products of the form $(1-p)^2$ and $(-p)^2$ are regained and products of the form $(-p)(1-p)$ are lost again. By these considerations and observations the formula in Eq. (B.1) is obtained.

Appendix B.2. Calculation of $X^{*T}X^*$ in Eq. (A.6) for one stimulus

In this section, blocked designs with one stimulus type will be considered without any assumption on block lengths. By N_{BL_0} the number of trials in a null block will be denoted and by N_{BL_A} the number of trials in a stimulus block will be denoted. It will be shown that

$$\begin{aligned}\left(X^{*T}X^*\right)_{mn} &= (1-p)^2 \cdot \left(max\left(N_{BL_A}-d, 0\right) + max\left(0, d-N_{BL_0}\right)\right) \cdot N_C \\ &+ (-p)^2 \cdot \left(max\left(N_{BL_0}-d, 0\right) + max\left(0, d-N_{BL_A}\right)\right) \cdot N_C \\ &+ 2(-p)(1-p) \cdot min\left(d, N_{BL_0} + N_{BL_A}-d, N_{BL_0}, N_{BL_A}\right) \cdot N_C,\end{aligned} \tag{B.3}$$

where $d = mod(|m-n|, N_{BL_0} + N_{BL_A})$ gives the shift size between column m and column n in X^*. In Eq. (B.1) $2N_{BL}$ has to be replaced with $N_{BL_0} + N_{BL_A}$ and N_{BL} has to be replaced either with N_{BL_A} or N_{BL_0} to adapt Eq. (B.1) for unequal block length. With d as defined above, we obtain:

$$\begin{aligned}\left(X^{*T}X^*\right)_{mn} &= (1-p)^2 \cdot |N_{BL_A}-d| \cdot N_C + (-p)^2 \cdot |N_{BL_0}-d| \cdot N_C \\ &+ 2(-p)(1-p) min\left(d, N_{BL_0} + N_{BL_A}-d\right) \cdot N_C.\end{aligned} \tag{B.4}$$

In addition, other adaptations are necessary. It holds again that the scalar product of two columns $X^*_{.m_1}$ and $X^*_{.n_1}$is the same as the scalar product of the columns $X^*_{.m_2}$ and $X^*_{.n_2}$ when the shift size of the two pairs is the same. Thus, we focus again on the scalar product of the first column $X^*_{.1}$ with the other columns in X^*. The scalar product of $X^*_{.1}$ with $X^*_{.1}$ equals:

$$\begin{aligned}\left(X^{*T}X^*\right)_{11} &= X^{*T}_{.1}X^*_{.1} \\ &= (1-p)^2 \cdot N_{BL_A} \cdot N_C + (-p)^2 \cdot N_{BL_0} \cdot N_C.\end{aligned} \tag{B.5}$$

Per cycle, the number of products of the form $(1-p)^2$ in the scalar product $(X^{*T}X^*)_{1n}$ is given by

$$max\left(N_{BL_A}-d, 0\right) + max\left(0, d-N_{BL_0}\right). \tag{B.6}$$

For explanation of Eq. (B.6), we assume in the following that $N_{BL_0} < N_{BL_A}$. For a shift size smaller than or equal to N_{BL_0}, one product of the form $(1-p)^2$ is lost per shift of the column $X^*_{.n}$ for each cycle. Thus, the number of products per cycle is totally explained by the loss of products given by $(N_{BL_A} - shift\ size)$. When the shift size is between N_{BL_0} and N_{BL_A} the number of products is explained by a combination of loosing products and obtaining products as expressed by $(N_{BL_A} - shift\ size) + (shift\ size - N_{BL_0})$ when shifting. For a shift size higher than or equal to N_{BL_A}, the number of products is explained by the gain of products given by $(shift\ size - N_{BL_0})$. Generalizing these observations to a formula for all shift sizes leads to Eq. (B.6).

The same considerations as for the number of products of the form $(1-p)^2$ can be applied for the number of products of the form $(-p)^2$. Thus, the number of products of the form $(-p)^2$ is

$$\left(max\left(N_{BL_0}-d, 0\right) + max\left(0, d-N_{BL_A}\right)\right) \cdot N_C. \tag{B.7}$$

There cannot be more products of the form $(1-p)(-p)$ per cycle than the smallest block length. Thus, the number of products of form $(1-p)(-p)$ per cycle is given by

$$min\left(d, N_{BL_0} + N_{BL_A}-d, N_{BL_0}\right). \tag{B.8}$$

If $N_{BL_A} < N_{BL_0}$, the same observations as above hold when N_{BL_A} is replaced with N_{BL_0} and vice versa. It follows Eq. (B.3) for all possible combinations of N_{BL_A} and N_{BL_0}. The formula Eq. (B.3) reduces to Eq. (B.1) when $N_{BL_0} = N_{BL_A}$. Note that Eq. (B.3) also holds for $N_{BL_0} = 0$ and the right half of Eq. (B.3) reduces then to $N_C(1-p)^2 N_{BL_A}$.

Appendix B.3. Calculation of $X^{*T}X^*$ in Eq. (A.6) for any number of stimulus types

As before, the number of events in a stimulus block is denoted by N_{BL_A} and the number of events in a null block is denoted by N_{BL_0}. The sequence s^q is equal to a stimulus sequence s' with only one stimulus type and null events. The matrix X'^* is the centered matrix for sequence s' and the product $X'^{*T}X'^*$ is equal to $X^{*T}_q X^*_q$ for all stimuli A_1, …,A_Q. As the length of null blocks in s' and s^q is $(Q-1)N_{BL_A} + N_{BL_0}$ and the total length of one cycle $(A_1 \dots A_Q N)$ is $QN_{BL_A} + N_{BL_0}$, we have to replace, N_{BL_0} in Eq. (B.3) with $(Q-1)N_{BL_A} + N_{BL_0}$ and $N_{BL_A} + N_{BL_0}$ in Eq. (B.3) with $QN_{BL_A} + N_{BL_0}$. We will use the notation $N_{BL_{0T}} = (Q-1) N_{BL_A} + N_{BL_0}$ and $N_{BL_T} = QN_{BL_A} + N_{BL_0}$ (T for total). Thus,

$$\begin{aligned}\left(X^{*T}_q X^*_q\right)_{mn} &= (1-p)^2 \cdot \left(max\left(N_{BL_A}-d, 0\right) + max\left(0, d-N_{BL_{0T}}\right)\right) \cdot N_C \\ &+ (-p)^2 \cdot \left(max\left(N_{BL_{0T}}-d, 0\right) + max\left(0, d-N_{BL_A}\right)\right) \cdot N_C \\ &+ 2(-p)(1-p) min\left(d, N_{BL_T}-d, N_{BL_{0T}}\right) N_C,\end{aligned} \tag{B.9}$$

where $d = mod(|m-n|, N_{BL_T})$ is the shift size between column m and n in X^*_q. In the next step, we need to determine what the off-diagonal products of $X^{*T}X^*$ in Eq. (A.5) are. As block lengths are equal for all stimulus blocks, sequence s^q is only a shifted version of s^1. The shift is given by $(q-1)N_{BL_A}$. The shift between column m in X^*_1 and column n in X^*_q is thus given by $(q-1)N_{BL_A} + (m-n)$. The product $X^{*T}_1 X^*_q$ is as in Eq. (B.9) but $|m-n|$ is replaced by $(q-1)N_{BL_A} + (m-n)$. Due to the symmetry of $X^{*T}X^*$ it holds that $X^{*T}_q X^*_1 = X^{*T}_1 X^*_q$. We determined the products $X^{*T}_1 X^*_q$ and $X^{*T}_q X^*_1$ but we still need to determine the products $X^{*T}_{q_1} X^*_{q_2}$ for $q_1 \neq 1$ and $q_2 \neq 1$.

Because of the symmetry of $X^{*T}X^{*}$ we will focus on $q_1 < q_2$. The shift between sequence s^{q_1} and s^{q_2} is the same as the shift between sequence s^1 and sequence $s^{q_2-q_1+1}$. The shift is given by $(q_2-q_1)N_{BL_A}$. Thus, the shift between column m in $X^{*}_{q_1}$ and column n in $X^{*}_{q_2}$ is given by $(q_2-q_1)N_{BL_A}+(m-n)$. The product $X^{*T}_{q_1}X^{*}_{q_2}$ is thus as in Eq. (B.9) but we need to replace $|m-n|$ with $|(q_2-q_1)N_{BL_A}+(m-n)|$. Because of the symmetry of $X^{*T}X^{*}$ it holds that $X^{*T}_{q_1}X^{*}_{q_2}=X^{*T}_{q_2}X^{*}_{q_1}$. It follows Eq. (A.6) for general q_q and q_2. Note that this result also holds for $N_{BL_0}=0$ and of course for $N_{BL_A}=N_{BL_0}$.

References

Aguirre, G.K., D'Esposito, M., 1999. Experimental design for brain fMRI. In: Moonen, C.T.W., Bandettini, P.A. (Eds.), Functional MRI. Springer-Verlag, Berlin, pp. 369–380.
Atkinson, A.C., Donev, A.N., Tobias, R.D., 2007. Optimum Experimental Designs with SAS. Oxford University Press, Oxford.
Beckmann, C.F., Jenkinson, M., Smith, S.M., 2001. General multi-level linear modelling for group analysis in FMRI, Technical report, Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB). http://www.fmrib.ox.ac.uk/analysis/techrep/.
Bianciardi, M., Cerasa, A., Patria, F., Hagberg, G.E., 2004. Evaluation of mixed effects in event-related fMRI studies: impact of first-level design and filtering. Neuroimage 22, 1351–1370.
Bullmore, E., Brammer, M., Williams, S.C.R., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., Sham, P., 1996. Statistical methods of estimation and inference for functional MR image analysis. Magn. Reson. Med. 35, 261–277.
Chatfield, C., 2004. The Analysis of Time Series: An Introduction (6th ed). Chapman and Hall/CRC, Boca Raton, FL.
Chein, J.M., Schneider, W., 2003. Designing effective fMRI experiments. In: Grafman, J., Robertson, I.H. (Eds.), The Handbook of Neuropsychology. Elsevier, Amsterdam, pp. 299–326.
Dale, A.M., 1999. Optimal experimental design for event-related fMRI. Hum. Brain Mapp. 8, 109–114.
Desmond, J.E., Glover, G.H., 2002. Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analysis. J. Neurosci. Methods 118, 115–128.
Friston, K.J., Fletcher, P., Josephs, O., Holmes, A.P., Rugg, M.D., Turner, R., 1998. Event-related fMRI: characterizing differential responses. Neuroimage 7, 30–40.
Harville, D.A., 1997. Matrix Algebra from a Statistician's Perspective. Springer-Verlag, New York.
Hayasaka, S., Peiffer, A.M., Hugenschmidt, C.E., Laurienti, P.J., 2007. Power and sample size calculation for neuroimaging studies by non-central random field theory. Neuroimage 37, 721–730.
Henson, R.N.A., 2004. Analysis of fMRI time series. In: Frackowiak, R.S.J., Friston, K.J., Frith, C.D., Dolan, R.J., Price, C.J., Zeki, S., Ashburner, J.T., Penny, W.D. (Eds.), Human Brain Function. Elsevier, Amsterdam, pp. 793–822.
Liu, T.T., Frank, L.R., 2004. Efficiency, power, and entropy in event-related fMRI with multiple trial types, Part I: theory. Neuroimage 21, 387–400.
Liu, T.T., Frank, L.R., Wong, E.C., Buxton, R.B., 2001. Detection power, estimation efficiency, and predictability in event-related fMRI. Neuroimage 13, 759–773.
Lovell, M.C., 1963. Seasonal adjustment of economic time series and multiple regression analysis. J. Am. Stat. Assoc. 58 (304), 993–1010.
Maus, B., van Breukelen, G.J.P., Goebel, R., Berger, M.P.F., 2010a. Optimization of blocked designs in fMRI studies. Psychometrika 75 (2), 373–390.
Maus, B., van Breukelen, G.J.P., Goebel, R., Berger, M.P.F., 2010b. Robustness of optimal design of fMRI experiments with application of a genetic algorithm. Neuroimage 49, 2433–2443.
Moerbeek, M., van Breukelen, G.J.P., Berger, M.P.F., 2008. Optimal designs for multilevel studies. In: de Leeuw, J., Meijer, E. (Eds.), Handbook of Multilevel Analysis. Springer, New York, pp. 177–205.
Mumford, J.A., Nichols, T.E., 2008. Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation. Neuroimage 39, 261–268.
Murphy, K., Garavan, H., 2004. An empirical investigation into the number of subjects required for an event-related fMRI study. Neuroimage 22, 879–885.
Penny, W.D., Holmes, A.P., 2004. Random-effects analysis. In: Frackowiak, R.S.J., Friston, K.J., Frith, C.D., Dolan, R.J., Price, C.J., Zeki, S., Ashburner, J.T., Penny, W.D. (Eds.), Human Brain Function. Elsevier, Amsterdam, pp. 843–850. Chapter 42.
Schouten, H.J.A., 1999. Sample size formula with a continuous outcome for unequal group sizes and unequal variances. Stat. Med. 18, 87–97.
Seber, G.A.F., 1977. Linear Regression Analysis. Wiley, New York.
Suckling, J., Ohlssen, D., Andrew, C., Johnson, G., Williams, S.C.R., Graves, M., Chen, C.-H., Spiegelhalter, D., Bullmore, E., 2008. Components of variance in a multicentre functional MRI study and implications for calculation of statistical power. Hum. Brain Mapp. 29, 1111–1122.
Wager, T.D., Nichols, T.E., 2003. Optimization of experimental design in fMRI: a general framework using a genetic algorithm. Neuroimage 293–309.
Worsley, K.J., Liao, C.H., Aston, J., Petre, V., Duncan, G.H., Morales, F., Evans, A.C., 2002. A general statistical analysis for fMRI data. Neuroimage 15, 1–15.