Stockholms universitet

Research Report

Department of Statistics

No. 2011:2

Influence analysis in two-treatment cross-over designs with special reference to the ABBA|BAAB design

Chengcheng Hao
Tatjana von Rosen
Dietrich von Rosen

Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

Influence analysis in two-treatment cross-over designs with special reference to the ABBA|BAAB design

Chengcheng Hao[a], Tatjana von Rosen[a], Dietrich von Rosen[b,c]

[a]*Department of Statistics, Stockholm University, SE-106 91 Stockholm*
[b]*Department of Energy and Technology, Swedish University of Agricultural Sciences, SE-750 07 Uppsala*
[c]*Department of Mathematics, Linköping University, SE-581 83 Linköping*

Abstract

This work is to develop methodology to detect influential observations in linear mixed model for multiple-period two-treatment cross-over designs. Existence of explicit maximum likelihood estimates (MLEs) of variance parameters as well as of mean parameters in the mixed model with treatment, residual, period and sequence effects is proven. Special reference is taken to the four-period ABBA|BAAB design. Case-weighted perturbations are performed. The influence quantities on each parameter estimate and their dispersion matrix are presented as closed-form functions of residuals in the unperturbed model.

Keywords: Delta-beta influence, Explicit maximum likelihood estimate, Mixed linear model, Multiple-period cross-over design, Perturbation scheme, Variance-ratio influence

1. Introduction

Cross-over designs, also mentioned in the literature as change-over, multiple time series or repeated measurements designs, are designs in which each subject receives more than one treatment in certain order (Jones and Kenward, 1989). The cross-over designs can reduce the number of subjects needed in studies, which in many applications may be plots of land, animals or human beings. This is particularly important if there are ethical concerns or with scarce or threatened populations. Therefore, multiple-period cross-over designs are common employed in many fields.

From a statistical point of view, the main advantage of cross-over designs is that they result in an increase of statistical power since each subject can serve as its own control. Due to the fact that subjects in the study are often randomly selected from a large population with unknown variance, subject effects are typically random effects. Recently, interests is to study cross-over designs within the framework of mixed linear models (see e.g. Carrière and Huang, 2000; Hedayat *et al.*, 2006; Yan and Locke, 2010; Hedayat and Zheng, 2010).

Although linear models are extensively applied in studies of cross-over designs, most of the available contributions focus on the associated optimal designs or tests under the assumed models. The sensitivity of the models, which is one of the most important issues when validating the model, is seldom discussed in cross-over design studies. One way to formulate the sensitivity problem is to develop statistics robust to minor perturbations on the cross-over design model (Putt and Chinchilli, 2000). However, there are other formulations. Suppose that a minor perturbation exists in a single or a few observations of the model. Influence analysis evaluates the changes on the estimators or test statistics after a perturbation has been performed and aims to identify the observations that have dramatically large influence. Such observations are defined as influential observations (Belsley *et al.*, 2004). This work aims to carry out influence analysis for multiple-period two-treatment cross-over designs.

Except for Hao *et al.* (2011), no pervious work, by the authors' knowledge, develops methodology to detect influential observations in cross-over design, either in mixed linear models or in fixed-effect linear models. We extend the delta-beta-based local influence approach proposed by Hao *et al.* (2011) for two-sequence two-period cross-over design to multiple-period cross-over designs. An underlying mixed linear model is assumed. Closed-form maximum likelihood estimates (MLEs) of the parameters in the cross-over designs are utilised. Although other influence diagnostics for general linear mixed models are expected to be able to detect the influential observations in cross-over designs, e.g. the methods in Lesaffre and Verbeke (1998) or Christensen *et al.* (1992), the fact that our influential quantities yield explicit expressions as functions of the residuals helps to interpret the data and is computationally more efficient.

In the next section, we start with a mixed linear model for general two-treatment cross-over designs. Examples of its specification in various cross-over designs are provided. Basic tools of influence analysis, e.g. perturbation scheme and objective functions of influence are defined in Section 3 and applied in the coming discussion. Explicit results of the influence analysis for a balanced four-period cross-over designs, which is referred to as the ABBA|BAAB design, are presented in Section 4. Section 5 contains our final conclusions and remarks.

2. Model

Throughout this paper, upper case letters with bold face denote matrices, bold lower case letters denote column vectors and non-bold lower case letters with subscripts are used to show elements of matrices or vectors. Let $\boldsymbol{I}_p$, $\mathbf{1}_p$ and $\boldsymbol{J}_p = \mathbf{1}_p\mathbf{1}_p^{\mathrm{T}}$ denote the $p \times p$ identity matrix, the $p \times 1$ vector and the $p \times p$ matrix with elements equal to 1, respectively. The symbol $\otimes$ represents the Kronecker product of matrices. Moreover, the vector space generated by the columns of the $p \times q$ matrix $\boldsymbol{A}$, $\mathcal{C}(\boldsymbol{A})$, is given by $\mathcal{C}(\boldsymbol{A}) = \{\boldsymbol{a} : \boldsymbol{a} = \boldsymbol{A}\boldsymbol{z}, \boldsymbol{z} \in \mathbf{R}^q\}$. The orthogonal complement to $\mathcal{C}(\boldsymbol{A})$ is denoted by $\mathcal{C}(\boldsymbol{A})^{\perp}$, and a matrix of which columns generate $\mathcal{C}(\boldsymbol{A})^{\perp}$ is denoted by $\boldsymbol{A}^o$. The p-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is denoted $N_p(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$.

In the following discussion, the terminology subject will be mentioned as a unit of experiment, observation as data observed in single period within the subject, and a case is a subject or a observation in general.

2.1. General model for cross-over designs

The key feature of cross-over design modelling is that each response can be affected not only by the "direct" effects of the treatment in the current period, but possibly also by "residual" effects from treatments applied in previous periods.

This work will focus on the comparison of two treatments in a experiment, treatment A and treatment B. It can be studied by a two-treatment cross-over design d with s sequences and p periods. Following the notation of Kershner and Federer (1981), we denote the design $COD(2, s, p)$. Let y_{ijk} represent the response observed during the k-th period on j-th subject within the i-th sequence under the design d, with $i = 1, 2, \dots, s;\ j = 1, 2, \dots, n;\ k = 1, 2, \dots, p$. Kershner and Federer (1981) surveyed a list of frequently used linear models in cross-over designs, which were rewritten by Carrière and Reinsel (1992) for two-treatment CODs as

$$y_{ijk} = \mu + \alpha_k + \phi\Phi_{d(i,k)} + \rho\Phi_{d(i,k-1)} + \lambda_i + \gamma_{ij} + \epsilon_{ijk}, \tag{1}$$

where μ is the general mean, α_k is the effect of the k-th period, and λ_i is the effect of the i-th sequence. The function value of $d(i, k)$ stands for the treatment that is assigned to the i-th sequence during the k-th period by the design d. Let $d(i, k) = 1$ denote treatment A, and $d(i, k) = 2$ treatment B. We define $\Phi_1 = 1/2$, $\Phi_2 = -1/2$ and $\Phi_{d(i,0)} = 0$. The parameter ϕ is the direct treatment effect contrast between treatment A and B, and ρ is the first-order residual effect contrast between treatment A and B. The effect γ_{ij} represents random individual effect of the j-th subject within sequence i, which is assumed to be $\gamma_{ij} \overset{\text{i.i.d.}}{\sim} N(0, \sigma_\gamma^2)$ and independent of the random error $\epsilon_{ijk} \overset{\text{i.i.d.}}{\sim} N(0, \sigma_e^2)$. The variances σ_γ^2 and σ_e^2 are supposed to be unknown.

2.2. Reparametrization

Model (1) is over-parametrized. In order to eliminate the redundancy of the parameters and to obtain unique mean estimators, reparametrization on nuisance parameters, i.e period effects and sequence effects, is commonly done. Examples of reparametrized (1) in various cross-over designs are provided.

COD Example I. AB and BA design.
In the simplest two-sequence two-period cross-over design, where $s = 2$ and $p = 2$, subjects are administered with two sequences of treatments, to receive treatment A followed by treatment B (sequence AB) or to receive treatment B followed by treatment A (sequence BA). It implies a design function $d(i, k)$ given by

$$d(i,k) = \begin{cases} 1, \text{ if } (i,k) \in \{(1,1),(2,2)\}, \\ 2, \text{ if } (i,k) \in \{(1,2),(2,1)\}. \end{cases}$$

In matrix notation and standard mixed models notation, model (1) for the AB|BA design is specified as

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\epsilon}, \tag{2}$$

where the response vector $\boldsymbol{y}=(y_{111}, y_{112}, y_{121}, \ldots, y_{2n2})^{\mathrm{T}}$, the vector of random effects $\boldsymbol{\gamma}=(\gamma_{11}, \gamma_{12}, \ldots, \gamma_{2n})^{\mathrm{T}}$, and the vector of random errors $\boldsymbol{\epsilon}=(\epsilon_{111}, \epsilon_{112}, \ldots, \epsilon_{2n2})^{\mathrm{T}}$. The matrix $\boldsymbol{Z}=\boldsymbol{I}_2\otimes\boldsymbol{I}_n\otimes\mathbf{1}_2$ is the $4n\times 2n$ known incidence matrix for $\boldsymbol{\gamma}$. Since in the AB|BA design, the residual effect ρ is completely confounded with the treatment and sequence effects, without loss of generality, the restrictions

$$\begin{aligned}
&\alpha_1 = -\alpha_2 = \pi/2,\\
&\lambda_1 = -\lambda_2 = \lambda/4,\\
&\rho = 0,
\end{aligned}$$

are set on the original mean parameters space of model (1). Define the parameter $\boldsymbol{\beta} = (\mu, \pi, \phi, \lambda)^{\mathrm{T}}$ to be the vector of reparametrized unknown mean parameters. The matrix $\boldsymbol{X} = (\mathbf{x}_{111}, \mathbf{x}_{112}, \mathbf{x}_{121}, \ldots, \mathbf{x}_{2n2})^{\mathrm{T}}$ is a $4n\times 4$ known design matrix for $\boldsymbol{\beta}$. The column vector $\mathbf{x}_{ijk}$ is a row of $\boldsymbol{X}$ written as a column, which for $j = 1, 2, \ldots, n$, is given by

$$\begin{aligned}
\mathbf{x}_{1j1} &= \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}^{\mathrm{T}},\\
\mathbf{x}_{1j2} &= \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{4} \end{pmatrix}^{\mathrm{T}},\\
\mathbf{x}_{2j1} &= \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{4} \end{pmatrix}^{\mathrm{T}},\\
\mathbf{x}_{2j2} &= \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{4} \end{pmatrix}^{\mathrm{T}}.
\end{aligned}$$

COD Example II. ABB and BAA design.
Consider a cross-over design with $s = 2$ and $p = 3$, where each subject is allocated to the treatment sequence ABB or BAA. It implies a design function $d(i, k)$ given by

$$d(i,k) = \begin{cases} 1, \text{ if } (i,k) \in \{(1,1),(2,2),(2,3)\}, \\ 2, \text{ if } (i,k) \in \{(1,2),(1,3),(2,1)\}. \end{cases}$$

In standard mixed model notation, model (1) for the ABB|BAA design is specified as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}, \tag{3}$$

where the response vector $\boldsymbol{y}=(y_{111}, y_{112}, y_{113}, y_{121}, \ldots, y_{2n3})^{\mathrm{T}}$, the random effects $\boldsymbol{\gamma}=(\gamma_{11}, \gamma_{12}, \ldots, \gamma_{2n})^{\mathrm{T}}$, and the random errors $\boldsymbol{\epsilon}=(\epsilon_{111}, \epsilon_{112}, \epsilon_{113}, \epsilon_{121}, \ldots, \epsilon_{2n3})^{\mathrm{T}}$. The matrix $\boldsymbol{Z} = \boldsymbol{I}_2 \otimes \boldsymbol{I}_n \otimes \mathbf{1}_3$ is the $6n \times 2n$ known incidence matrix for $\boldsymbol{\gamma}$. The difference between the ABB|BAA and the AB|BA cross-over design model is that the residual effect is not confounded. Without loss of generality, the restrictions

$$\begin{aligned}
&\alpha_1 = \pi_1/2 + \pi_2/3,\\
&\alpha_2 = -\pi_1/2 + \pi_2/3,\\
&\alpha_3 = -2/3\phi_2,\\
&\lambda_1 = -\lambda_2 = (\lambda + \phi)/6,
\end{aligned}$$

are set on the original mean parameter space of (1). Define the parameter $\boldsymbol{\beta} = (\mu, \pi_1, \pi_2, \phi, \rho, \lambda)^{\mathrm{T}}$ to be the vector of reparametrized unknown mean parameters. The matrix $\boldsymbol{X} = (\mathbf{x}_{111}, \mathbf{x}_{112}, \mathbf{x}_{113}, \mathbf{x}_{121}, \ldots, \mathbf{x}_{2n3})^{\mathrm{T}}$ is a $6n \times 6$ known design matrix for $\boldsymbol{\beta}$, where the vector $\mathbf{x}_{ijk}$ for $j = 1, 2, \ldots, n$, is given by

$$\begin{aligned}
\mathbf{x}_{1j1} &= \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{2}{3} & 0 & \frac{1}{6} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{1j2} &= \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{3} & -\frac{1}{3} & \frac{1}{2} & \frac{1}{6} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{1j3} &= \begin{pmatrix} 1 & 0 & -\frac{2}{3} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{6} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{2j1} &= \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & -\frac{2}{3} & 0 & -\frac{1}{6} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{2j2} &= \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{3} & \frac{1}{3} & -\frac{1}{2} & -\frac{1}{6} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{2j3} &= \begin{pmatrix} 1 & 0 & -\frac{2}{3} & \frac{1}{3} & \frac{1}{2} & -\frac{1}{8} \end{pmatrix}^{\mathrm{T}}.
\end{aligned}$$

COD Example III. ABBA and BAAB design.
Consider a cross-over design with $s = 2$ and $p = 4$, where each subject is allocated to the treatment sequence ABBA or BAAB. It implies a design function $d(i, k)$ given by

$$d(i,k) = \begin{cases} 1, \text{ if } (i,k) \in \{(1,1),(1,4),(2,2),(2,3)\}, \\ 2, \text{ if } (i,k) \in \{(1,2),(1,3),(2,1),(2,4)\}. \end{cases}$$

In standard mixed model notation, model (1) for the ABBA|BAAB design is specified as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}, \tag{4}$$

where $\boldsymbol{y} = (y_{111}, y_{112}, y_{113}, y_{114}, y_{121}, \ldots, y_{2n4})^{\mathrm{T}}$, $\boldsymbol{\gamma} = (\gamma_{11}, \gamma_{12}, \ldots, \gamma_{2n})^{\mathrm{T}}$, and $\boldsymbol{\epsilon} = (\epsilon_{111}, \epsilon_{112}, \epsilon_{113}, \epsilon_{114}, \epsilon_{121}, \ldots, \epsilon_{2n4})^{\mathrm{T}}$. The matrix $\boldsymbol{Z} = \boldsymbol{I}_2 \otimes \boldsymbol{I}_n \otimes \mathbf{1}_4$ is the $8n \times 2n$ known incidence matrix for $\boldsymbol{\gamma}$. Without loss of generality, the restrictions

$$\begin{aligned}
\alpha_1 &= \pi_1/2 + \pi_2/3 + \pi_3/4, \\
\alpha_2 &= -\pi_1/2 + \pi_2/3 + \pi_3/4, \\
\alpha_3 &= -2/3\pi_2 + \pi_3/4, \\
\alpha_4 &= -3/4\phi_3, \\
\lambda_1 &= -\lambda_2 = (\lambda + \rho)/8,
\end{aligned}$$

are set on the original mean parameter space of (1). Define the parameter $\boldsymbol{\beta} = (\mu, \pi_1, \pi_2, \pi_3, \phi, \rho, \lambda)^{\mathrm{T}}$ to be the vector of reparametrized unknown mean parameters. The matrix $\boldsymbol{X} = (\mathbf{x}_{111}, \ldots, \mathbf{x}_{114}, \mathbf{x}_{121}, \ldots, \mathbf{x}_{2n4})^{\mathrm{T}}$ is a $8n \times 7$ known design matrix for $\boldsymbol{\beta}$, where the vector $\mathbf{x}_{ijk}$ is for $j = 1, 2, \ldots, n$, given by

$$\begin{aligned}
\mathbf{x}_{1j1} &= \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{1j2} &= \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{3} & \frac{1}{4} & -\frac{1}{2} & \frac{5}{8} & \frac{1}{8} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{1j3} &= \begin{pmatrix} 1 & 0 & -\frac{2}{3} & \frac{1}{4} & -\frac{1}{2} & -\frac{3}{8} & \frac{1}{8} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{1j4} &= \begin{pmatrix} 1 & 0 & 0 & -\frac{3}{4} & \frac{1}{2} & -\frac{3}{8} & \frac{1}{8} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{2j1} &= \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & -\frac{1}{2} & -\frac{1}{8} & -\frac{1}{8} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{2j2} &= \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{2} & -\frac{5}{8} & -\frac{1}{8} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{2j3} &= \begin{pmatrix} 1 & 0 & -\frac{2}{3} & \frac{1}{4} & \frac{1}{2} & \frac{3}{8} & -\frac{1}{8} \end{pmatrix}^{\mathrm{T}}, \\
\mathbf{x}_{2j4} &= \begin{pmatrix} 1 & 0 & 0 & -\frac{3}{4} & -\frac{1}{2} & \frac{3}{8} & -\frac{1}{8} \end{pmatrix}^{\mathrm{T}}.
\end{aligned}$$

2.3. Explicit maximum likelihood estimates

There are alternative model setups for mean parameters of cross-over designs. For example, the model without sequence effects is the most frequently used; Kershner and Federer (1981) mention that the treatment-by-period interaction model is frequently applied for $COD(t, t, p)$, where the numbers of treatments and sequences are equal; Afsarinejad and Hedayat (2002) propose a model with self and mixed carry-over effects; Park *et al.* (2010) introduce the interaction terms of direct effects and residual effects to model.

Model (1) is preferred to its alternatives without sequence effects because it ensures the existence of the explicit maximum likelihood estimators (MLEs) in general $COD(2, s, p)$, given that the variance parameters σ_γ^2 and σ_e^2 are unknown. One important finding is that model (1) can always be represented as two randomly independent homoscedastic linear models with independent sets of parameters. This is shown in the following theorem where the explicit MLEs in (4) for the ABBA|BAAB design are derived.

Theorem 2.1. *In the two-sequence four-period cross-over design, where each subject is allocated to a treatment sequence ABBA or BAAB, model (4) is equivalent to two independent homoscedastic models with functionally independent mean and variance parameters given by*

$$\begin{cases} \boldsymbol{y}_s = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\eta}_1, & \boldsymbol{\eta}_1 \sim N_{2n}(\boldsymbol{0}, \sigma_1^2\boldsymbol{I}_{2n}), \\ \boldsymbol{y}_d = \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\eta}_2, & \boldsymbol{\eta}_2 \sim N_{6n}(\boldsymbol{0}, \sigma_2^2\boldsymbol{I}_{6n}), \end{cases} \tag{5}$$

for some responses vectors $\boldsymbol{y}_s$ *and* $\boldsymbol{y}_d$*, and design matrices* $\boldsymbol{X}_1$ *and* $\boldsymbol{X}_2$ *of proper sizes, where the parameters*

$$\boldsymbol{\beta}_1 = (\mu,\ \lambda)^{\mathrm{T}}, \quad \boldsymbol{\beta}_2 = (\pi_1,\ \pi_2,\ \pi_3,\ \phi,\ \rho)^{\mathrm{T}},$$

contain separate sets of mean parameters, and the two random-error vectors $\boldsymbol{\eta}_1$ *and* $\boldsymbol{\eta}_2$ *are mutually independent, with separate variance parameters*

$$\sigma_1^2 = \sigma_e^2 + 4\sigma_\gamma^2, \quad \sigma_2^2 = \sigma_e^2.$$

Proof. The result can be proven by pre-multiplying with an orthogonal matrix

$$\boldsymbol{T} = \boldsymbol{I}_{2n} \otimes (\boldsymbol{T}_s : \boldsymbol{T}_d)^{\mathrm{T}} \tag{6}$$

to both sides of model (4) which satisfy

$$\mathcal{C}(\boldsymbol{T}_s) = \mathcal{C}\left(\tfrac{1}{4}\boldsymbol{J}_4\right) \text{ and } \mathcal{C}(\boldsymbol{T}_d) = \mathcal{C}(\boldsymbol{T}_s)^{\perp}.$$

Since the transformation matrix $\boldsymbol{T}$ is of full rank and orthogonal, a transformed model can be inverted into (4) by the transformation $\boldsymbol{T}^{\mathrm{T}}$. The two model systems with respect to the transformation $\boldsymbol{T}$ are equivalent.

Let us denote the subvector of the responses and the submatrix of the design matrix in (4) for the ij-th subject by

$$\begin{aligned}\boldsymbol{y}_{ij} &= (y_{ij1},\ y_{ij2},\ y_{ij3},\ y_{ij4})^{\mathrm{T}},\\ \mathbf{X}_{ij} &= (\mathbf{x}_{ij1},\ \mathbf{x}_{ij2},\ \mathbf{x}_{ij3},\ \mathbf{x}_{ij4})^{\mathrm{T}},\end{aligned}$$

and the within-subject covariance matrix

$$\boldsymbol{\Sigma} = Var(\boldsymbol{y}_{ij}) = \sigma_\gamma^2 \boldsymbol{J}_4 + \sigma_e^2 \boldsymbol{I}_4,$$

for $i = 1, 2,\ j = 1, 2, \ldots, n$. It can be verified that the transformation matrix $\boldsymbol{T}$ has two effects on (4):

(i): On the variance parameter space of (4), i.e.

$$\boldsymbol{\Sigma} = \left(\sigma_e^2 + 4\sigma_\gamma^2\right) \boldsymbol{P}_{T_s} + \sigma_e^2 \boldsymbol{P}_{T_d}, \tag{7}$$

where

$$\boldsymbol{P}_{T_s} = \boldsymbol{T}_s \boldsymbol{T}_s^{\mathrm{T}} = \tfrac{1}{4}\boldsymbol{J}_4, \quad \boldsymbol{P}_{T_d} = \boldsymbol{T}_d \boldsymbol{T}_d^{\mathrm{T}} = \boldsymbol{I}_4 - \tfrac{1}{4}\boldsymbol{J}_4, \tag{8}$$

are orthogonal projections on $\mathcal{C}(\boldsymbol{T}_s)$ and $\mathcal{C}(\boldsymbol{T}_d)$, respectively.

(ii): On the mean parameter space of (4), i.e.

$$\mathcal{C}(\mathbf{X}_{ij}\boldsymbol{L}) \subseteq \mathcal{C}(\boldsymbol{T}_s) \text{ and } \mathcal{C}(\mathbf{X}_{ij}\boldsymbol{L}^o) \subseteq C(\boldsymbol{T}_s)^{\perp} = \mathcal{C}(\boldsymbol{T}_d), \tag{9}$$

where

$$\boldsymbol{L} = \begin{pmatrix} 1&0&0&0&0&0&0\\ 0&0&0&0&0&0&1 \end{pmatrix}^{\mathrm{T}}, \quad \boldsymbol{L}^o = \begin{pmatrix} 0&1&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&0&1&0&0&0\\ 0&0&0&0&1&0&0\\ 0&0&0&0&0&1&0 \end{pmatrix}^{\mathrm{T}}. \tag{10}$$

Now we will show that the transformed model can be formulated as (5). Without loss of generality, let

$$\boldsymbol{T}_s = \begin{pmatrix} 1/2\\ 1/2\\ 1/2\\ 1/2 \end{pmatrix}, \quad \boldsymbol{T}_d = \begin{pmatrix} 1/2 & 1/\sqrt{10} & 2/\sqrt{10}\\ -1/2 & 2/\sqrt{10} & -1/\sqrt{10}\\ -1/2 & -2/\sqrt{10} & 1/\sqrt{10}\\ 1/2 & -1/\sqrt{10} & -2/\sqrt{10} \end{pmatrix}. \tag{11}$$

The response vector of the transformed model can be partitioned into two vectors written as

$$\underset{2n\times 1}{\boldsymbol{y}_s} = (y_{s,11}, y_{s,12}, \ldots, y_{s,2n})^{\mathrm{T}}, \quad \underset{6n\times 1}{\boldsymbol{y}_d} = (\boldsymbol{y}_{d,11}^{\mathrm{T}},\ \boldsymbol{y}_{d,12}^{\mathrm{T}}, \ldots, \boldsymbol{y}_{d,2n}^{\mathrm{T}})^{\mathrm{T}},$$

with

$$y_{s,ij} = \boldsymbol{T}_s^{\mathrm{T}} \boldsymbol{y}_{ij}, \quad \boldsymbol{y}_{d,ij} = \boldsymbol{T}_d^{\mathrm{T}} \boldsymbol{y}_{ij}, \quad \text{for } i = 1, 2,\ j = 1, 2, \ldots, n. \tag{12}$$

Based on (7), the variances satisfy

$$Var\left(y_{s,ij}\right)=\sigma_e^2+4\sigma_\gamma^2,\quad Var\left(\boldsymbol{y}_{d,ij}\right)=\sigma_e^2\,\boldsymbol{I}_3,\quad Cov\left(y_{s,ij},\boldsymbol{y}_{d,ij}\right)=\mathbf{0}.$$

Based on (9), we have

$$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{T}_d=\mathbf{0},\quad \boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{T}_s=\mathbf{0},$$

which imply that

$$\mathcal{C}\left(\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{T}_d\right)\subseteq\mathcal{C}\left(\boldsymbol{L}^o\right)\ \text{and}\ \mathcal{C}\left(\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{T}_s\right)\subseteq\mathcal{C}\left(\boldsymbol{L}\right).\tag{13}$$

Because $\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}$ and $\boldsymbol{L}^o\boldsymbol{L}^{o\mathrm{T}}$ are orthogonal projections on $\mathcal{C}(\boldsymbol{L})$ and $\mathcal{C}(\boldsymbol{L}^o)$, respectively, the expectations satisfy

$$\begin{aligned}E\left(y_{s,ij}\right)&=\boldsymbol{T}_s^{\mathrm{T}}\boldsymbol{X}_{ij}\boldsymbol{\beta}=\boldsymbol{T}_s^{\mathrm{T}}\boldsymbol{X}_{ij}\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{\beta},\\ E\left(\boldsymbol{y}_{d,ij}\right)&=\boldsymbol{T}_d^{\mathrm{T}}\boldsymbol{X}_{ij}\boldsymbol{\beta}=\boldsymbol{T}_d^{\mathrm{T}}\boldsymbol{X}_{ij}\boldsymbol{L}^o\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{\beta},\end{aligned}$$

where

$$\boldsymbol{\beta}=(\mu,\ \pi_1,\ \pi_2,\ \pi_3,\ \phi,\ \rho,\ \lambda)^{\mathrm{T}}.$$

By denoting

$$\begin{gathered}\sigma_1^2=\sigma_e^2+4\sigma_\gamma^2,\quad \sigma_2^2=\sigma_e^2,\\ \boldsymbol{\beta}_1=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{\beta},\quad \boldsymbol{\beta}_2=\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{\beta},\end{gathered}$$

and

$$\underset{2n\times 2}{\boldsymbol{X}_1}=\left(\mathbf{x}_{1,11},\mathbf{x}_{1,12},\ldots,\mathbf{x}_{1,2n}\right)^{\mathrm{T}},\quad \underset{6n\times 5}{\boldsymbol{X}_2}=\left(\boldsymbol{X}_{2,11}^{\mathrm{T}},\boldsymbol{X}_{2,12}^{\mathrm{T}},\ldots,\boldsymbol{X}_{2,2n}^{\mathrm{T}}\right)^{\mathrm{T}},$$

with

$$\mathbf{x}_{1,ij}^{\mathrm{T}}=\boldsymbol{T}_s^{\mathrm{T}}\boldsymbol{X}_{ij}\boldsymbol{L},\quad \boldsymbol{X}_{2,ij}=\boldsymbol{T}_d^{\mathrm{T}}\boldsymbol{X}_{ij}\boldsymbol{L}^o,\quad \text{for } i=1,2,\ j=1,2,\ldots,n,\tag{14}$$

and since normality holds, the theorem is proven. ■

Theorem 2.2. *Consider a balanced ABBA|BAAB cross-over design with n subjects in each sequence. Denote the averages of responses $\overline{y}_{i\cdot k}=\frac{1}{n}\sum_{j=1}^{n}y_{ijk}$, for $i=1,2,\ k=1,2,3,4$.*

(i) *The MLE of $\boldsymbol{\beta}$ in (4) is given by*

$$\widehat{\boldsymbol{\beta}}=\begin{pmatrix}\frac{1}{8}(\overline{y}_{1\cdot1}+\overline{y}_{1\cdot2}+\overline{y}_{1\cdot3}+\overline{y}_{1\cdot4})+\frac{1}{8}(\overline{y}_{2\cdot1}+\overline{y}_{2\cdot2}+\overline{y}_{2\cdot3}+\overline{y}_{2\cdot4})\\ \frac{1}{2}(\overline{y}_{1\cdot1}-\overline{y}_{1\cdot2})+\frac{1}{2}(\overline{y}_{2\cdot1}-\overline{y}_{2\cdot2})\\ \frac{1}{4}(\overline{y}_{1\cdot1}+\overline{y}_{1\cdot2}-2\overline{y}_{1\cdot3})+\frac{1}{4}(\overline{y}_{2\cdot1}+\overline{y}_{2\cdot2}-2\overline{y}_{2\cdot3})\\ \frac{1}{6}(\overline{y}_{1\cdot1}+\overline{y}_{1\cdot2}+\overline{y}_{1\cdot3}-3\overline{y}_{1\cdot4})+\frac{1}{6}(\overline{y}_{2\cdot1}+\overline{y}_{2\cdot2}+\overline{y}_{2\cdot3}-3\overline{y}_{2\cdot4})\\ \frac{1}{20}(6\overline{y}_{1\cdot1}-3\overline{y}_{1\cdot2}-7\overline{y}_{1\cdot3}+4\overline{y}_{1\cdot4})-\frac{1}{20}(6\overline{y}_{2\cdot1}-3\overline{y}_{2\cdot2}-7\overline{y}_{2\cdot3}+4\overline{y}_{2\cdot4})\\ \frac{1}{10}(2\overline{y}_{1\cdot1}+4\overline{y}_{1\cdot2}-4\overline{y}_{1\cdot3}-2\overline{y}_{1\cdot4})-\frac{1}{10}(2\overline{y}_{2\cdot1}+4\overline{y}_{2\cdot2}-4\overline{y}_{2\cdot3}-2\overline{y}_{2\cdot4})\\ (\overline{y}_{1\cdot1}+\overline{y}_{1\cdot2}+\overline{y}_{1\cdot3}+\overline{y}_{1\cdot4})-(\overline{y}_{2\cdot1}+\overline{y}_{2\cdot2}+\overline{y}_{2\cdot3}+\overline{y}_{2\cdot4})\end{pmatrix}.$$

(ii) *The dispersion matrix of* $\widehat{\boldsymbol{\beta}}$ *is given by*

$$D\left[\widehat{\boldsymbol{\beta}}\right]=\frac{1}{n}\begin{pmatrix} \frac{1}{8}(4\sigma_\gamma^2+\sigma_e^2) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3\sigma_e^2}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2\sigma_e^2}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{11\sigma_e^2}{20} & \frac{\sigma_e^2}{5} & 0 \\ 0 & 0 & 0 & 0 & \frac{\sigma_e^2}{5} & \frac{4\sigma_e^2}{5} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8(4\sigma_\gamma^2+\sigma_e^2) \end{pmatrix}.$$

(iii) *Let the residual in the unperturbed model for a single subject be denoted by* $\mathbf{r}_{ij}=\boldsymbol{y}_{ij}-\mathbf{X}_{ij}\widehat{\boldsymbol{\beta}}$, $i=1,2,\ j=1,2,\ldots,n$. *The residual of prediction for the* j*-th subject within sequence ABBA equals*

$$\mathbf{r}_{1j}=\mathbf{r}_{W1j}+\boldsymbol{P}_{T_{d1}}\mathbf{r}_B.$$

The residual of the j*-th subject with sequence BAAB equals*

$$\mathbf{r}_{2j}=\mathbf{r}_{W2j}-\boldsymbol{P}_{T_{d1}}\mathbf{r}_B.$$

We denote $\mathbf{r}_{Wij}$ *the* 4×1 *vector of within-sequence residuals for the* ij*-th subject and* $\mathbf{r}_B$ *the* 4×1 *vector of between-sequence residuals given by*

$$\mathbf{r}_{Wij}=\begin{pmatrix} y_{ij1}-\overline{y}_{i\cdot 1} \\ y_{ij2}-\overline{y}_{i\cdot 2} \\ y_{ij3}-\overline{y}_{i\cdot 3} \\ y_{ij4}-\overline{y}_{i\cdot 4} \end{pmatrix},\quad \mathbf{r}_B=\frac{1}{2}\begin{pmatrix} \overline{y}_{1\cdot 1}-\overline{y}_{2\cdot 1} \\ \overline{y}_{1\cdot 2}-\overline{y}_{2\cdot 2} \\ \overline{y}_{1\cdot 3}-\overline{y}_{2\cdot 3} \\ \overline{y}_{1\cdot 4}-\overline{y}_{2\cdot 4} \end{pmatrix}, \tag{15}$$

and the matrix

$$\boldsymbol{P}_{T_{d1}}=\boldsymbol{T}_{d1}\boldsymbol{T}_{d1}^{\mathrm{T}},\ \textit{with}\ \boldsymbol{T}_{d1}=\left(\tfrac{2}{\sqrt{10}}\ -\tfrac{1}{\sqrt{10}}\ \tfrac{1}{\sqrt{10}}\ -\tfrac{2}{\sqrt{10}}\right)^{\mathrm{T}}, \tag{16}$$

is the orthogonal projection on the column space $\mathcal{C}\left(\boldsymbol{T}_{d1}\right)$.

(iv) *The MLEs of* σ_e^2 *and* σ_γ^2 *equal*

$$\widehat{\sigma}_e^2=\frac{1}{6n}\sum_{ij}\mathbf{r}_{Wij}^{\mathrm{T}}\left(\boldsymbol{I}_4-\tfrac{1}{4}\boldsymbol{J}_4\right)\mathbf{r}_{Wij}+\frac{1}{3}\mathbf{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\mathbf{r}_B,$$

$$\widehat{\sigma}_\gamma^2=\frac{1}{24n}\sum_{ij}\mathbf{r}_{Wij}^{\mathrm{T}}\left(\boldsymbol{J}_4-\boldsymbol{I}_4\right)\mathbf{r}_{Wij}-\frac{1}{12}\mathbf{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\mathbf{r}_B.$$

Proof. In the proof of Theorem 2.1, the transformation is invertible. Thus, MLEs in (4) can be obtained from the MLEs in (5), and vice versa. According to (9), we have

$$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{P}_{T_s}=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}},\quad \boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{P}_{T_d}=\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}}.$$

Therefore,

$$\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1 \;=\; \boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\left(\boldsymbol{I}_{2n}\otimes\boldsymbol{P}_{T_s}\right)\boldsymbol{X}\boldsymbol{L} = \; \boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L} = \; n\begin{pmatrix} 8 & 0 \\ 0 & \frac{1}{8}\end{pmatrix},$$

$$\left(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1\right)^{-1} = \left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{I}_{2n}\otimes\boldsymbol{P}_{T_s})\boldsymbol{X}\boldsymbol{L}\right)^{-1} = \left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1} = \frac{1}{n}\begin{pmatrix} \frac{1}{8} & 0 \\ 0 & 8\end{pmatrix},$$

$$\begin{aligned}\left(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1\right)^{-1}\boldsymbol{X}_{1,1j}^{\mathrm{T}}\boldsymbol{T}_s^{\mathrm{T}} &= \left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\left(\boldsymbol{I}_{2n}\otimes\boldsymbol{P}_{T_s}\right)\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}}\boldsymbol{P}_{T_s} \\ &= \left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}} = \frac{1}{n}\begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ 1 & 1 & 1 & 1\end{pmatrix},\end{aligned}$$

$$\begin{aligned}\left(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1\right)^{-1}\boldsymbol{X}_{1,2j}^{\mathrm{T}}\boldsymbol{T}_s^{\mathrm{T}} &= \left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\left(\boldsymbol{I}_{2n}\otimes\boldsymbol{P}_{T_s}\right)\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}}\boldsymbol{P}_{T_s} \\ &= \left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}} = \frac{1}{n}\begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ -1 & -1 & -1 & -1\end{pmatrix},\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2 &= \boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\left(\boldsymbol{I}_{2n}\otimes\boldsymbol{P}_{T_d}\right)\boldsymbol{X}\boldsymbol{L}^o \\ &= \boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o = n\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{4}{3} & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{2} & 0 & 0 \\ 0 & 0 & 0 & 2 & -\frac{1}{2} \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{11}{8}\end{pmatrix},\end{aligned}$$

$$\begin{aligned}\left(\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2\right)^{-1} &= \left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\left(\boldsymbol{I}_{2n}\otimes\boldsymbol{P}_{T_d}\right)\boldsymbol{X}\boldsymbol{L}^o\right)^{-1} \\ &= \left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1} = \frac{1}{n}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{3}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{11}{20} & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{5} & \frac{4}{5}\end{pmatrix},\end{aligned}$$

$$\begin{aligned}\left(\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2\right)^{-1}\boldsymbol{X}_{2,1j}^{\mathrm{T}}\boldsymbol{T}_d^{\mathrm{T}} &= \left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\left(\boldsymbol{I}_{2n}\otimes\boldsymbol{P}_{T_d}\right)\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}}\boldsymbol{P}_{T_d} \\ &= \left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}} = \frac{1}{n}\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{2} \\ \frac{3}{10} & -\frac{3}{20} & -\frac{7}{20} & \frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} & -\frac{2}{5} & -\frac{1}{5}\end{pmatrix},\end{aligned}$$

$$\left(\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2\right)^{-1}\boldsymbol{X}_{2,2j}^{\mathrm{T}}\boldsymbol{T}_d^{\mathrm{T}} = \left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\left(\boldsymbol{I}_{2n}\otimes\boldsymbol{P}_{T_d}\right)\boldsymbol{X}\boldsymbol{L}^{o}\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}}\boldsymbol{P}_{T_d}$$

$$= \left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^{o}\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}} = \frac{1}{n}\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{2} \\ -\frac{3}{10} & \frac{3}{20} & \frac{7}{20} & -\frac{1}{5} \\ -\frac{1}{5} & -\frac{2}{5} & \frac{2}{5} & \frac{1}{5} \end{pmatrix}.$$

For each separate model in (5), the homoscedastic setup is satisfied. Then, the MLEs of the mean parameters are identical with the ordinary least squares estimators given by

$$\widehat{\boldsymbol{\beta}}_1 = \left(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1\right)^{-1}\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{y}_s = \left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$$

$$= \begin{pmatrix} \frac{1}{8}(\overline{y}_{1\cdot1}+\overline{y}_{1\cdot2}+\overline{y}_{1\cdot3}+\overline{y}_{1\cdot4}) + \frac{1}{8}(\overline{y}_{2\cdot1}+\overline{y}_{2\cdot2}+\overline{y}_{2\cdot3}+\overline{y}_{2\cdot4}) \\ (\overline{y}_{1\cdot1}+\overline{y}_{1\cdot2}+\overline{y}_{1\cdot3}+\overline{y}_{1\cdot4}) - (\overline{y}_{2\cdot1}+\overline{y}_{2\cdot2}+\overline{y}_{2\cdot3}+\overline{y}_{2\cdot4}) \end{pmatrix},$$

$$\widehat{\boldsymbol{\beta}}_2 = \left(\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2\right)^{-1}\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{y}_d = \left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^{o}\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$$

$$= \begin{pmatrix} \frac{1}{2}(\overline{y}_{1\cdot1}-\overline{y}_{1\cdot2}) + \frac{1}{2}(\overline{y}_{2\cdot1}-\overline{y}_{2\cdot2}) \\ \frac{1}{4}(\overline{y}_{1\cdot1}+\overline{y}_{1\cdot2}-2\overline{y}_{1\cdot3}) + \frac{1}{4}(\overline{y}_{2\cdot1}+\overline{y}_{2\cdot2}-2\overline{y}_{2\cdot3}) \\ \frac{1}{6}(\overline{y}_{1\cdot1}+\overline{y}_{1\cdot2}+\overline{y}_{1\cdot3}-3\overline{y}_{1\cdot4}) + \frac{1}{6}(\overline{y}_{2\cdot1}+\overline{y}_{2\cdot2}+\overline{y}_{2\cdot3}-3\overline{y}_{2\cdot4}) \\ \frac{1}{20}(6\overline{y}_{1\cdot1}-3\overline{y}_{1\cdot2}-7\overline{y}_{1\cdot3}+4\overline{y}_{1\cdot4}) - \frac{1}{20}(6\overline{y}_{2\cdot1}-3\overline{y}_{2\cdot2}-7\overline{y}_{2\cdot3}+4\overline{y}_{2\cdot4}) \\ \frac{1}{10}(2\overline{y}_{1\cdot1}+4\overline{y}_{1\cdot2}-4\overline{y}_{1\cdot3}-2\overline{y}_{1\cdot4}) - \frac{1}{10}(2\overline{y}_{2\cdot1}+4\overline{y}_{2\cdot2}-4\overline{y}_{2\cdot3}-2\overline{y}_{2\cdot4}) \end{pmatrix},$$

with dispersion matrices

$$D\left[\widehat{\boldsymbol{\beta}}_1\right] = \sigma_1^2\left(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1\right)^{-1} = \frac{\sigma_e^2+4\sigma_\gamma^2}{n}\begin{pmatrix} \frac{1}{8} & 0 \\ 0 & 8 \end{pmatrix},$$

$$D\left[\widehat{\boldsymbol{\beta}}_2\right] = \sigma_2^2\left(\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2\right)^{-1} = \frac{\sigma_e^2}{n}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{3}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{11}{20} & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{5} & \frac{4}{5} \end{pmatrix}.$$

Thus, the results in (i) and (ii) are proven.

It follows that the subject residual in (4) equals

$$\boldsymbol{r}_{1j} = \boldsymbol{y}_{1j} - \boldsymbol{X}_{1j}\widehat{\boldsymbol{\beta}}$$

$$= \begin{pmatrix} y_{1j1} \\ y_{1j2} \\ y_{1j3} \\ y_{1j4} \end{pmatrix} - \begin{pmatrix} \frac{4}{5} & \frac{1}{10} & -\frac{1}{10} & \frac{1}{5} & \frac{1}{5} & -\frac{1}{10} & \frac{1}{10} & -\frac{1}{5} \\ \frac{1}{10} & \frac{19}{20} & \frac{1}{20} & -\frac{1}{10} & -\frac{1}{10} & \frac{1}{20} & -\frac{1}{20} & \frac{1}{10} \\ -\frac{1}{10} & \frac{1}{20} & \frac{19}{20} & \frac{1}{10} & \frac{1}{10} & -\frac{1}{20} & \frac{1}{20} & -\frac{1}{10} \\ \frac{1}{5} & -\frac{1}{10} & \frac{1}{10} & \frac{4}{5} & -\frac{1}{5} & \frac{1}{10} & -\frac{1}{10} & \frac{1}{5} \end{pmatrix} \begin{pmatrix} \overline{y}_{1\cdot 1} \\ \overline{y}_{1\cdot 2} \\ \vdots \\ \overline{y}_{2\cdot 4} \end{pmatrix},$$

$$\boldsymbol{r}_{2j} = \boldsymbol{y}_{2j} - \boldsymbol{X}_{2j}\widehat{\boldsymbol{\beta}}$$

$$= \begin{pmatrix} y_{2j1} \\ y_{2j2} \\ y_{2j3} \\ y_{2j4} \end{pmatrix} - \begin{pmatrix} \frac{1}{5} & -\frac{1}{10} & \frac{1}{10} & -\frac{1}{5} & \frac{4}{5} & \frac{1}{10} & -\frac{1}{10} & \frac{1}{5} \\ -\frac{1}{10} & \frac{1}{20} & -\frac{1}{20} & \frac{1}{10} & \frac{1}{10} & \frac{19}{20} & \frac{1}{20} & -\frac{1}{10} \\ \frac{1}{10} & -\frac{1}{20} & \frac{1}{20} & -\frac{1}{10} & -\frac{1}{10} & \frac{1}{20} & \frac{19}{20} & \frac{1}{10} \\ -\frac{1}{5} & \frac{1}{10} & -\frac{1}{10} & \frac{1}{5} & \frac{1}{5} & -\frac{1}{10} & \frac{1}{10} & \frac{4}{5} \end{pmatrix} \begin{pmatrix} \overline{y}_{1\cdot 1} \\ \overline{y}_{1\cdot 2} \\ \vdots \\ \overline{y}_{2\cdot 4} \end{pmatrix}.$$

Thus, the result in (iii) is established.

The homoscedastic setups also imply that the MLEs of the variance parameters in (5) equal

$$\widehat{\sigma}_1^2 = \frac{1}{2n}\sum_{ij}\left(\boldsymbol{T}_s^{\mathrm{T}}\boldsymbol{r}_{ij}\right)^{\mathrm{T}}\left(\boldsymbol{T}_s^{\mathrm{T}}\boldsymbol{r}_{ij}\right) = \frac{1}{2n}\sum_{ij}\boldsymbol{r}_{ij}^{\mathrm{T}}\boldsymbol{P}_{T_s}\boldsymbol{r}_{ij},$$

$$\widehat{\sigma}_2^2 = \frac{1}{6n}\sum_{ij}\left(\boldsymbol{T}_d^{\mathrm{T}}\boldsymbol{r}_{ij}\right)^{\mathrm{T}}\left(\boldsymbol{T}_d^{\mathrm{T}}\boldsymbol{r}_{ij}\right) = \frac{1}{6n}\sum_{ij}\boldsymbol{r}_{ij}^{\mathrm{T}}\boldsymbol{P}_{T_d}\boldsymbol{r}_{ij}.$$

Since $\boldsymbol{r}_{1j} = \boldsymbol{r}_{W1j} + \boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B$, $\boldsymbol{r}_{2j} = \boldsymbol{r}_{W2j} - \boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B$, and the column spaces

$$\mathcal{C}\left(\boldsymbol{T}_{d1}\right) \subset \mathcal{C}\left(\boldsymbol{T}_d\right) = \mathcal{C}\left(\boldsymbol{T}_s\right)^{\perp},$$

we get

$$\widehat{\sigma}_1^2 = \frac{1}{2n}\sum_j\left(\boldsymbol{r}_{W1j}^{\mathrm{T}}\boldsymbol{P}_{T_s}\boldsymbol{r}_{W1j}\right) + \frac{1}{2n}\sum_j\left(\boldsymbol{r}_{W2j}^{\mathrm{T}}\boldsymbol{P}_{T_s}\boldsymbol{r}_{W2j}\right) = \frac{1}{2n}\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}\boldsymbol{P}_{T_s}\boldsymbol{r}_{Wij},$$

$$\begin{aligned}
\widehat{\sigma}_2^2 &= \frac{1}{6n}\sum_j\left(\boldsymbol{r}_{W1j}^{\mathrm{T}}\boldsymbol{P}_{T_d}\boldsymbol{r}_{W1j} + 2\boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_{W1j} + \boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B\right) \\
&\quad + \frac{1}{6n}\sum_j\left(\boldsymbol{r}_{W2j}^{\mathrm{T}}\boldsymbol{P}_{T_d}\boldsymbol{r}_{W2j} - 2\boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_{W2j} + \boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B\right) \\
&= \frac{1}{6n}\sum_j\left(\boldsymbol{r}_{W1j}^{\mathrm{T}}\boldsymbol{P}_{T_d}\boldsymbol{r}_{W1j} + \boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B\right) + \frac{1}{3n}\boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\sum_j\boldsymbol{r}_{W1j} \\
&\quad + \frac{1}{6n}\sum_j\left(\boldsymbol{r}_{W2j}^{\mathrm{T}}\boldsymbol{P}_{T_d}\boldsymbol{r}_{W2j} + \boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B\right) - \frac{1}{3n}\boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\sum_j\boldsymbol{r}_{W2j} \\
&= \frac{1}{6n}\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}\boldsymbol{P}_{T_d}\boldsymbol{r}_{Wij} + \frac{1}{3}\boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B.
\end{aligned}$$

The MLEs $\widehat{\sigma}_\gamma^2$ and $\widehat{\sigma}_e^2$ in (2) equal

$$\begin{aligned}\widehat{\sigma}_e^2 = \widehat{\sigma}_2^2 &= \frac{1}{6n}\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}\boldsymbol{P}_{T_d}\boldsymbol{r}_{Wij} + \frac{1}{3}\boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B \\ &= \frac{1}{6n}\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}\left(\boldsymbol{I}_4 - \tfrac{1}{4}\boldsymbol{J}_4\right)\boldsymbol{r}_{Wij} + \frac{1}{3}\boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B,\end{aligned}$$

and

$$\begin{aligned}\widehat{\sigma}_\gamma^2 = \frac{1}{4}\left(\widehat{\sigma}_1^2 - \widehat{\sigma}_2^2\right) &= \frac{1}{24n}\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}\left(3\boldsymbol{P}_{T_s} - \boldsymbol{P}_{T_d}\right)\boldsymbol{r}_{Wij} - \frac{1}{12}\boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B \\ &= \frac{1}{24n}\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}\left(\boldsymbol{J}_4 - \boldsymbol{I}_4\right)\boldsymbol{r}_{Wij} - \frac{1}{12}\boldsymbol{r}_B^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B.\end{aligned}$$

■

3. Delta-beta-based local influence

3.1. Basic concepts

The principal idea associated with local influence is assuming a small perturbation on the interested model and aims to evaluate the changes of this perturbation on key statistics, e.g. on the observed likelihood or on the maximum likelihood estimates of parameters. According to the statistics of interest, the local influence analysis can be categorised into two classes: the likelihood-based local influence approach Cook (1986) and the delta-beta-based local influence approach Hao *et al.* (2011).

To identify the influential observations in the ABBA|BAAB design, we extend the methodology proposed by Hao *et al.* (2011) for 2× 2 cross-over design, namely delta-beta-based local influence approach, to multiple-period cross-over designs.

Three important concepts used in the work of Hao *et al.* (2011) for 2×2 cross-over design are the case-weighted perturbation scheme, the delta-beta influence function and the variance-ratio influence function. We express the general definitions for them as follow.

Definition 3.1. *Suppose that a perturbation scheme $P(\boldsymbol{\omega})$ exists such that the response vector is modified from $\boldsymbol{y}$ to $\boldsymbol{y}_{P(\boldsymbol{\omega})}$, and the design matrix from $\boldsymbol{X}$ to $\boldsymbol{X}_{P(\boldsymbol{\omega})}$. With respect to a subset I of observations, $P(\boldsymbol{\omega})$ is the case-weighted perturbation scheme if and only if it satisfies the following two criteria.*

(i) *The subset I of observations is analogous to be removed when $\boldsymbol{\omega}=\boldsymbol{0}$;*
(ii) $\boldsymbol{y}_{P(\boldsymbol{\omega}_0)} = \boldsymbol{y}$ *and* $\boldsymbol{X}_{P(\boldsymbol{\omega}_0)} = \boldsymbol{X}$ *for some null perturbation weight* $\boldsymbol{\omega}_0$.

Let us call the model

$$\boldsymbol{y}_{P(\boldsymbol{\omega})} = \boldsymbol{X}_{P(\boldsymbol{\omega})}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon} \tag{17}$$

the perturbed model of (4), which assumes $\boldsymbol{\gamma} \sim N_{2n}(\mathbf{0}, \sigma_\gamma^2 \boldsymbol{I}_{2n})$, $\boldsymbol{\epsilon} \sim N_{8n}(\mathbf{0}, \sigma_e^2 \boldsymbol{I}_{8n})$, and $Cov(\boldsymbol{\gamma}, \boldsymbol{\epsilon}) = \mathbf{0}$. The influence of the perturbation with respect to the set I on mean parameters in (4) can be measured by the delta-beta influence.

Definition 3.2. *Let $\widehat{\boldsymbol{\beta}}(\boldsymbol{\omega})$ be the MLE of $\boldsymbol{\beta}$ and $D\left[\widehat{\boldsymbol{\beta}}(\boldsymbol{\omega})\right]$ be the associated dispersion matrix under the perturbed model. The delta-beta influence contains two statistics*

(i) *The statistic $\Delta\widehat{\boldsymbol{\beta}}$ with respect to a perturbation $P(\boldsymbol{\omega})$ on the subset I of observations is defined by*

$$\Delta_I \widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}}(\boldsymbol{\omega}) - \widehat{\boldsymbol{\beta}}(\boldsymbol{\omega}_0). \tag{18}$$

(ii) *The statistic $\Delta D\left[\widehat{\boldsymbol{\beta}}\right]$ with respect to a perturbation $P(\boldsymbol{\omega})$ on the subset I of observations is defined by*

$$\Delta_I D\left[\widehat{\boldsymbol{\beta}}\right] = \widehat{D}\left[\widehat{\boldsymbol{\beta}}(\boldsymbol{\omega})\right] - \widehat{D}\left[\widehat{\boldsymbol{\beta}}(\boldsymbol{\omega}_0)\right], \tag{19}$$

where $\widehat{D}\left[\widehat{\boldsymbol{\beta}}(\boldsymbol{\omega})\right]$ and $\widehat{D}\left[\widehat{\boldsymbol{\beta}}(\boldsymbol{\omega}_0)\right]$ are estimators of $D\left[\widehat{\boldsymbol{\beta}}(\boldsymbol{\omega})\right]$ and $D\left[\widehat{\boldsymbol{\beta}}(\boldsymbol{\omega}_0)\right]$, respectively, when the MLEs of σ_γ^2 and σ_e^2 are inserted.

The influence of the perturbation with respect to the set I on variance parameters in (4) can be measured by the variance-ratio influence.

Definition 3.3. *Let $\sigma_e^2(\boldsymbol{\omega})$ and $\sigma_\gamma^2(\boldsymbol{\omega})$ be the MLEs of the variance parameters under the perturbed model. The variance ratio for random errors (VRE) and the variance ratio for random effects (VRR) with respect to the perturbation $P(\boldsymbol{\omega})$ on the set I of observations are defined by*

$$\text{VRE}_I = \frac{\widehat{\sigma}_e^2(\boldsymbol{\omega})}{\widehat{\sigma}_e^2(\boldsymbol{\omega}_0)}, \tag{20}$$

$$\text{VRR}_I = \frac{\widehat{\sigma}_\gamma^2(\boldsymbol{\omega})}{\widehat{\sigma}_\gamma^2(\boldsymbol{\omega}_0)}. \tag{21}$$

A natural example of a case-weighted perturbation scheme with respect to subset I is that all the observations within the subset are scaled by the same perturbation weight ω. A perturbation defined by the following perturbation scheme to the ij-th subject in the ABBA|BAAB design will be used through the next section. For other possible perturbation schemes, we refer to Hao *et al.* (2011) and Beckman *et al.* (1987).

Example. Let

$$\boldsymbol{y}_{P(\omega)} = \begin{pmatrix} \omega \boldsymbol{y}_I \\ \boldsymbol{y}_{[I]} \end{pmatrix}, \text{ and } \boldsymbol{X}_{P(\omega)} = \begin{pmatrix} \omega \boldsymbol{X}_I \\ \boldsymbol{X}_{[I]} \end{pmatrix}, \tag{22}$$

for some non-negative ω belonging to the neighbourhood of 1. The vector $\boldsymbol{y}_I$ and the matrix $\boldsymbol{X}_I$ denote the subvector of responses and the rows of design matrix for the set I, respectively. The vector or matrix with index $[I]$ represents the associated vector or matrix with the set I removed. The null perturbation weight mentioned in Definition 3.1 exists at $\omega_0 = 1$ for (22).

3.2. Basic algebra for influence analysis

The result presented in the following auxiliary lemma for the homoscedastic linear model will be extended to the cross-over design models with unknown variance within subjects in the subsequent.

Lemma 3.1. *Let us consider the following homoscedastic linear model*

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma_e^2 \boldsymbol{I}_n), \tag{23}$$

where $\boldsymbol{\beta}$ and σ_e^2 are unknown. Assume that the case-weighted perturbation scheme $P(\omega)$ in (22) is applied to the subset I of observations in the above model. The functions defined by (18) to (20) equal

$$\Delta_I \widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}} \left(\boldsymbol{H}_I - \frac{1}{1-\omega^2}\boldsymbol{I}\right)^{-1} \boldsymbol{r}_I, \tag{24}$$

$$\Delta_I D\left[\widehat{\boldsymbol{\beta}}\right] = \widehat{\sigma}^2 (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}} \left(\frac{1}{1-\omega^2}\boldsymbol{I} - \boldsymbol{H}_I\right)^{-1} \boldsymbol{X}_I (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}, \tag{25}$$

$$VRE_I = 1 - \frac{\boldsymbol{r}_I^{\mathrm{T}} \left(\frac{1}{1-\omega^2}\boldsymbol{I} - \boldsymbol{H}_I\right)^{-1} \boldsymbol{r}_I}{\boldsymbol{r}^{\mathrm{T}}\boldsymbol{r}}, \tag{26}$$

where the matrix $\boldsymbol{H}_I = \boldsymbol{X}_I \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1} \boldsymbol{X}_I^{\mathrm{T}}$ is the submatrix of the hat matrix $\boldsymbol{H} = \boldsymbol{X}\left(\boldsymbol{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}$, and the vector $\boldsymbol{r}_I = \boldsymbol{y}_I - \boldsymbol{X}_I\widehat{\boldsymbol{\beta}}(1)$ is the subvector of residuals $\mathbf{r} = \boldsymbol{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}(1)$ for observations belonging to set I.

Proof. In the perturbed model, we have

$$\boldsymbol{X}_{P(\omega)}^{\mathrm{T}}\boldsymbol{X}_{P(\omega)} = \boldsymbol{X}_{[I]}^{\mathrm{T}}\boldsymbol{X}_{[I]} + \omega^2 \boldsymbol{X}_I^{\mathrm{T}}\boldsymbol{X}_I = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} + (\omega^2 - 1)\boldsymbol{X}_I^{\mathrm{T}}\boldsymbol{X}_I,$$

and

$$\boldsymbol{X}_{P(\omega)}^{\mathrm{T}}\boldsymbol{y}_{P(\omega)} = \boldsymbol{X}_{[I]}^{\mathrm{T}}\boldsymbol{y}_{[I]} + \omega^2 \boldsymbol{X}_I^{\mathrm{T}}\boldsymbol{y}_I = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{y} + (\omega^2 - 1)\boldsymbol{X}_I^{\mathrm{T}}\boldsymbol{y}_I.$$

If the matrices $\boldsymbol{A}$, $\boldsymbol{A} + \boldsymbol{BCD}$ and $\boldsymbol{C}$ are non-singular, it is well-known that

$$(\boldsymbol{A} + \boldsymbol{BCD})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{DA}^{-1}\boldsymbol{B} + \boldsymbol{C}^{-1})^{-1}\boldsymbol{DA}^{-1}.$$

For the perturbation let $\boldsymbol{A} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X},\ \boldsymbol{B}^{\mathrm{T}} = \boldsymbol{D} = \boldsymbol{X}_I,$ and $\boldsymbol{C} = (\omega^2 - 1)\boldsymbol{I}$. Then,

$$\left(\boldsymbol{X}_{P(\omega)}^{\mathrm{T}}\boldsymbol{X}_{P(\omega)}\right)^{-1} = (\boldsymbol{X}^{\mathrm{T}}\mathbf{X})^{-1} + (\boldsymbol{X}^{\mathrm{T}}\mathbf{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}} \left(\frac{1}{1-\omega^2}\boldsymbol{I} - \boldsymbol{H}_I\right)^{-1} \boldsymbol{X}_I (\boldsymbol{X}^{\mathrm{T}}\mathbf{X})^{-1}.$$

For the observations in set I, denote the predicted values in the unperturbed model

$$\widehat{\boldsymbol{y}}_I = \boldsymbol{X}_I\widehat{\boldsymbol{\beta}}(1) = \boldsymbol{X}_I(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}.$$

Then, the MLE of $\boldsymbol{\beta}$ under the perturbation equals

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}(\omega) =& \left(\boldsymbol{X}_{P(\omega)}^{\mathrm{T}}\boldsymbol{X}_{P(\omega)}\right)^{-1}\left(\boldsymbol{X}_{P(\omega)}^{\mathrm{T}}\boldsymbol{y}_{P(\omega)}\right)\\
=& \widehat{\boldsymbol{\beta}}(1) + (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}}\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\widehat{\boldsymbol{y}}_I + (\omega^2-1)(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}}\boldsymbol{y}_I\\
&+ (\omega^2-1)(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}}\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\boldsymbol{H}_I\boldsymbol{y}_I\\
=& \widehat{\boldsymbol{\beta}}(1) + (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}}\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\widehat{\boldsymbol{y}}_I\\
&+ (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}}\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\Big(\boldsymbol{I}-(\omega^2-1)\boldsymbol{H}_I+(\omega^2-1)\boldsymbol{H}_I\Big)\boldsymbol{y}_I\\
=& \widehat{\boldsymbol{\beta}}(1) - (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}}\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\boldsymbol{r}_I,
\end{aligned}$$

with dispersion matrix

$$\begin{aligned}
D\left[\widehat{\boldsymbol{\beta}}(\omega)\right] =& \left(\boldsymbol{X}_{P(\omega)}^{\mathrm{T}}\boldsymbol{X}_{P(\omega)}\right)^{-1}\sigma^2\\
=& D[\widehat{\boldsymbol{\beta}}(1)] + \sigma^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}_I^{\mathrm{T}}\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\boldsymbol{X}_I(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}.
\end{aligned}$$

The MLE of σ^2 under the perturbation can be calculated by

$$\begin{aligned}
n\widehat{\sigma}^2(\omega) =& \boldsymbol{y}_{P(\omega)}^{\mathrm{T}}\boldsymbol{y}_{P(\omega)} - \boldsymbol{y}_{P(\omega)}^{\mathrm{T}}\boldsymbol{X}_{P(\omega)}\widehat{\boldsymbol{\beta}}(\omega)\\
=& \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} + (\omega^2-1)\boldsymbol{y}_I^{\mathrm{T}}\boldsymbol{y}_I - \boldsymbol{y}^{\mathrm{T}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}(1) - (\omega^2-1)\boldsymbol{y}_I^{\mathrm{T}}\widehat{\boldsymbol{y}}_I\\
&+ \widehat{\boldsymbol{y}}_I^{\mathrm{T}}\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\boldsymbol{r}_I + (\omega^2-1)\boldsymbol{y}_I^{\mathrm{T}}\boldsymbol{H}_I\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\boldsymbol{r}_I\\
=& \Big(\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} - \boldsymbol{y}^{\mathrm{T}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}(1)\Big) + \widehat{\boldsymbol{y}}_I^{\mathrm{T}}\left(\frac{1}{1-\omega^2}\boldsymbol{I}+\boldsymbol{H}_I\right)^{-1}\boldsymbol{r}_I\\
&+ \boldsymbol{y}_I^{\mathrm{T}}\Big(\boldsymbol{I}-(\omega^2-1)\boldsymbol{H}_I+(\omega^2-1)\boldsymbol{H}_I\Big)\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\boldsymbol{r}_I\\
=& n\widehat{\sigma}^2(1) - \boldsymbol{r}_I^{\mathrm{T}}\left(\frac{1}{1-\omega^2}\boldsymbol{I}-\boldsymbol{H}_I\right)^{-1}\boldsymbol{r}_I.
\end{aligned}$$

■

If the number of observations in set I is 1, the above Lemma can be simplified.

Corollary 3.2. *If the subset I is composed of the i-th observation in model (23), the functions calculated in Lemma 3.1 reduce to*

$$\Delta_i\widehat{\boldsymbol{\beta}} = \frac{(\omega^2-1)r_i}{(\omega^2-1)h_{ii}+1}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\mathbf{x}_i,$$

$$\Delta_i D\left[\widehat{\boldsymbol{\beta}}\right] = \frac{(1-\omega^2)\widehat{\sigma}^2}{(\omega^2-1)h_{ii}+1}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1},$$

$$VRE_i = 1 + \frac{\omega^2-1}{(\omega^2-1)h_{ii}+1}\frac{r_i^2}{\sum_{i=1}^n r_i^2},$$

where the column vector $\mathbf{x}_i$ is the i-th row of the design matrix $\boldsymbol{X}$, the scalar $h_{ii} = \mathbf{x}_i^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\mathbf{x}_i$ denotes the i-th diagonal element of the hat matrix for the unperturbed model, and $r_i = y_i - \mathbf{x}_i\widehat{\boldsymbol{\beta}}(1)$ is the residual of the i-th observation.

4. Results for the ABBA|BAAB design

It will be proven in this section that the cross-over design model can be reformulated in a form where influential quantities for deletion or perturbation have closed-form solutions. The influences of the ij-th subject on the estimates of the fixed effects and the variance parameters, which are defined in the previous section, are presented in Theorem 4.1.

Theorem 4.1. *Assume that the case-weighted perturbation scheme $P_{(}\omega)$ in (22) is applied to the observations of the j-th subject within the i-th sequence in model (4). The functions defined by (18) to (21) are*

(i)

$$\begin{aligned}\Delta_{ij}\widehat{\boldsymbol{\beta}} &= \frac{\omega^2-1}{\omega^2+n-1}\boldsymbol{F}_i\boldsymbol{P}_{T_s}\boldsymbol{r}_{Wij} \\ &\quad + \frac{\omega^2-1}{\omega^2+n-1}\boldsymbol{G}_i\boldsymbol{P}_{T_{d2}}\boldsymbol{r}_{Wij} + \frac{2(\omega^2-1)}{\omega^2+2n-1}\boldsymbol{G}_i\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_{Tij},\end{aligned}$$

(ii)

$$\begin{aligned}\Delta_{ij}D\left[\widehat{\boldsymbol{\beta}}\right] &= \frac{1-\omega^2}{n(\omega^2+n-1)}\boldsymbol{F}_i\boldsymbol{F}_i^{\mathrm{T}}(\widehat{\sigma}_e^2+4\widehat{\sigma}_\gamma^2) \\ &\quad + \frac{1-\omega^2}{n(\omega^2+n-1)}\boldsymbol{G}_i\boldsymbol{P}_{T_{d2}}\boldsymbol{G}_i^{\mathrm{T}}\widehat{\sigma}_e^2 + \frac{2(1-\omega^2)}{n(\omega^2+2n-1)}\boldsymbol{G}_i\boldsymbol{P}_{T_{d1}}\boldsymbol{G}_i^{\mathrm{T}}\widehat{\sigma}_e^2,\end{aligned}$$

(iii)

$$\begin{aligned}\mathrm{VRE}_{ij} &= 1 + \frac{(\omega^2-1)n}{\omega^2+n-1}\frac{\boldsymbol{r}_{Wij}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{r}_{Wij}}{\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{r}_{Wij} + \sum_{ij}\boldsymbol{r}_{Tij}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_{Tij}} \\ &\quad + \frac{(\omega^2-1)2n}{\omega^2+2n-1}\frac{\boldsymbol{r}_{Tij}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_{Tij}}{\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{r}_{Wij} + \sum_{ij}\boldsymbol{r}_{Tij}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_{Tij}},\end{aligned}$$

$$\mathrm{VRR}_{ij} = 1+\frac{(\omega^2-1)n}{\omega^2+n-1}\frac{\boldsymbol{r}_{Wij}^{\mathrm{T}}\left(3\boldsymbol{P}_{T_s}-\boldsymbol{P}_{T_{d2}}\right)\boldsymbol{r}_{Wij}}{\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}(3\boldsymbol{P}_{T_s}-\boldsymbol{P}_{T_{d2}})\,\boldsymbol{r}_{Wij}+\sum_{ij}\boldsymbol{r}_{Tij}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_{Tij}}$$
$$-\frac{(\omega^2-1)2n}{\omega^2+2n-1}\frac{\boldsymbol{r}_{Tij}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_{Tij}}{\sum_{ij}\boldsymbol{r}_{Wij}^{\mathrm{T}}\left(3\boldsymbol{P}_{T_s}-\boldsymbol{P}_{T_{d2}}\right)\boldsymbol{r}_{Wij}+\sum_{ij}\boldsymbol{r}_{Tij}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_{Tij}},$$

where the vectors

$$\boldsymbol{r}_{Wij} = \begin{pmatrix} y_{ij1}-\overline{y}_{i\cdot1} \\ y_{ij2}-\overline{y}_{i\cdot2} \\ y_{ij3}-\overline{y}_{i\cdot3} \\ y_{ij4}-\overline{y}_{i\cdot4} \end{pmatrix}, \quad \boldsymbol{r}_{Tij} = \begin{pmatrix} y_{ij1}-\overline{y}_{\cdot\cdot1} \\ y_{ij2}-\overline{y}_{\cdot\cdot2} \\ y_{ij3}-\overline{y}_{\cdot\cdot3} \\ y_{ij4}-\overline{y}_{\cdot\cdot4} \end{pmatrix},$$

are defined as the within-sequence residual and the total residual of the j-th subject in the i-th sequence, respectively. The matrices

$$\boldsymbol{F}_i = n\boldsymbol{L}\left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}}, \tag{27}$$
$$\boldsymbol{G}_i = n\boldsymbol{L}^{o}\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^{o}\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}}, \tag{28}$$

$i=1,2$, *with $\boldsymbol{L}$ and $\boldsymbol{L}^o$ defined by (10), and the orthogonal projections*

$$\boldsymbol{P}_{T_s} = \boldsymbol{T}_s\boldsymbol{T}_s^{\mathrm{T}}, \quad \boldsymbol{P}_{T_{d1}} = \boldsymbol{T}_{d1}\boldsymbol{T}_{d1}^{\mathrm{T}}, \quad \boldsymbol{P}_{T_{d2}} = \boldsymbol{I}_4-\boldsymbol{P}_{T_s}-\boldsymbol{P}_{T_{d1}},$$

with

$$\begin{aligned}\boldsymbol{T}_s &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}^{\mathrm{T}}, \\ \boldsymbol{T}_{d1} &= \begin{pmatrix} \frac{2}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & \frac{1}{\sqrt{10}} & -\frac{2}{\sqrt{10}} \end{pmatrix}^{\mathrm{T}},\end{aligned}$$

are decided only by the cross-over design function $d(i,k)$.

Proof. Use the matrix $\boldsymbol{T}$ defined in (6) to pre-multiply both sides of the perturbed model. The restriction on the perturbation scheme that the observations within each subject are scaled by the same perturbation weight enables the perturbed model to be splitted into the following two models

$$\begin{cases}\boldsymbol{y}_{s,P(\omega)} = \boldsymbol{X}_{1,P(\omega)}\boldsymbol{\beta}_1+\boldsymbol{\eta}_1, \\ \boldsymbol{y}_{d,P(\omega)} = \boldsymbol{X}_{2,P(\omega)}\boldsymbol{\beta}_2+\boldsymbol{\eta}_2,\end{cases} \tag{29}$$

where the parameters $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, the random terms $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ have the same meaning as those given in (5). Let us denote the perturbed response variables

$$\boldsymbol{y}_{s,P(\omega)} = (\boldsymbol{y}_{s,[ij]}^{\mathrm{T}}:\omega y_{s,ij})^{\mathrm{T}} \text{ and } \boldsymbol{y}_{d,P(\omega)} = (\boldsymbol{y}_{d,[ij]}^{\mathrm{T}}:\omega\boldsymbol{y}_{d,ij}^{\mathrm{T}})^{\mathrm{T}},$$

and the perturbed design matrices

$$\boldsymbol{X}_{1,P(\omega)} = (\boldsymbol{X}_{1,[ij]}^{\mathrm{T}}:\omega\boldsymbol{x}_{1,ij})^{\mathrm{T}} \text{ and } \boldsymbol{X}_{2,P(\omega)} = (\boldsymbol{X}_{2,[ij]}^{\mathrm{T}}:\omega\boldsymbol{X}_{2,ij}^{\mathrm{T}})^{\mathrm{T}},$$

where $\boldsymbol{y}_{d,ij}$, $y_{s,ij}$, $\mathbf{x}_{1,ij}$ and $\boldsymbol{X}_{2,ij}$ are defined in (12) and (14).

By applying similar simplifications as the proof of Theorem 2.2, the submatrices of the hat matrices of the above models before perturbation associated with the ij-th subject equal

$$\begin{aligned} h_{1,ij} &= \mathbf{x}_{1,ij}^{\mathrm{T}}(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1)^{-1}\mathbf{x}_{1,ij} = \boldsymbol{T}_s^{\mathrm{T}}\boldsymbol{X}_{ij}\boldsymbol{L}\left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{T}_s, \\ \boldsymbol{H}_{2,ij} &= \boldsymbol{X}_{2,ij}(\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_{2,ij}^{\mathrm{T}} = \boldsymbol{T}_d^{\mathrm{T}}\boldsymbol{X}_{ij}\boldsymbol{L}^o\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{T}_d, \end{aligned}$$

where $\boldsymbol{T}_s$ and $\boldsymbol{T}_d$ are defined by (6).

Obviously, $h_{1,ij}$ and $\boldsymbol{H}_{2,ij}$ depend on the choices of $\boldsymbol{T}_s$ and $\boldsymbol{T}_d$. However, since $\boldsymbol{T}_s$ in (6) is given uniquely by

$$\boldsymbol{T}_s = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}^{\mathrm{T}},$$

we have

$$h_{1,ij} = \frac{1}{n}, \quad i = 1, 2,\ j = 1, 2, \ldots, n.$$

Let us do a spectral decomposition on $\boldsymbol{H}_{2,ij}$. Without loss generality, a choice of $\boldsymbol{T}_d$ in (11) is utilised through this proof in order to simplify the decomposition, because it follows

$$\boldsymbol{H}_{2,ij} = \frac{1}{n}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad i = 1, 2,\ j = 1, 2, \ldots, n.$$

Later, (11) will be referred to as the canonical transformation for the ABBA|BAAB design because its columns can be divided into three sets of normalized orthogonal vectors which span the column spaces

$$\mathcal{C}(\boldsymbol{T}_s), \quad \mathcal{C}(\boldsymbol{T}_{d1}) \text{ and } \mathcal{C}(\boldsymbol{T}_s : \boldsymbol{T}_{d1})^{\perp},$$

respectively, where $\boldsymbol{T}_{d1}$ is introduced by (16) and determined only by the design matrix of treatment effects and residual effects in (4). The orthogonal projections on the above column spaces are denoted by $\boldsymbol{P}_{T_s}$, $\boldsymbol{P}_{T_{d1}}$ and $\boldsymbol{P}_{T_{d2}}$ in the given order.

It follows that

$$\begin{aligned} \frac{\omega^2 - 1}{(\omega^2 - 1)h_{1,ij} + 1} &= \frac{(\omega^2 - 1)n}{\omega^2 + n - 1}, \\ \left(\boldsymbol{H}_{2,ij} - \frac{1}{1 - \omega^2}\boldsymbol{I}_3\right)^{-1} &= \begin{pmatrix} \frac{(\omega^2-1)n}{\omega^2+n-1} & 0 & 0 \\ 0 & \frac{(\omega^2-1)n}{\omega^2+n-1} & 0 \\ 0 & 0 & \frac{(\omega^2-1)2n}{\omega^2+2n-1} \end{pmatrix}. \end{aligned}$$

According to Lemma 3.1, Corollary 3.2 and Theorem 2.2 (iii), if the perturbation occurs within sequence ABBA, the MLEs in (29) equal,

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_1(\omega) &= \widehat{\boldsymbol{\beta}}_1(1) + \frac{\omega^2-1}{(\omega^2-1)h_{1,1j}+1}(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1)^{-1}\mathbf{x}_{1,1j}\boldsymbol{T}_s^{\mathrm{T}}\boldsymbol{r}_{1j} \\
&= \widehat{\boldsymbol{\beta}}_1(1) + \frac{(\omega^2-1)n}{\omega^2+n-1}\left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}}\boldsymbol{r}_{W1j},
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_2(\omega) &= \widehat{\boldsymbol{\beta}}_2(1) + (\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_{2,1j}^{\mathrm{T}}\left(\boldsymbol{H}_{2,1j} - \frac{1}{1-\omega^2}\boldsymbol{I}_3\right)^{-1}\boldsymbol{T}_d^{\mathrm{T}}\boldsymbol{r}_{1j} \\
&= \widehat{\boldsymbol{\beta}}_2(1) + \left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}}\boldsymbol{T}_d\left(\boldsymbol{H}_{2,1j} - \frac{1}{1-\omega^2}\boldsymbol{I}_3\right)^{-1}\boldsymbol{T}_d^{\mathrm{T}}(\boldsymbol{r}_{W1j}+\boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B) \\
&= \widehat{\boldsymbol{\beta}}_2(1) + \frac{(\omega^2-1)n}{\omega^2+n-1}\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{r}_{W1j} \\
&\qquad + \frac{(\omega^2-1)2n}{\omega^2+2n-1}\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}(\boldsymbol{r}_{W1j}+\boldsymbol{r}_B).
\end{aligned}
$$

with the corresponding dispersion matrices

$$
\begin{aligned}
D\left[\widehat{\boldsymbol{\beta}}_1(\omega)\right] &= \frac{\omega^2-1}{(\omega^2-1)h_{1,1j}+1}(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1)^{-1}\mathbf{x}_{1,1j}\mathbf{x}_{1,1j}^{\mathrm{T}}(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1)^{-1}\sigma_1^2 + D\left[\widehat{\boldsymbol{\beta}}_1(1)\right] \\
&= \frac{(1-\omega^2)n}{\omega^2+n-1}\left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}}\boldsymbol{X}_{1j}\boldsymbol{L}\left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\sigma_1^2 + D\left[\widehat{\boldsymbol{\beta}}_1(1)\right],
\end{aligned}
$$

$$
\begin{aligned}
D\left[\widehat{\boldsymbol{\beta}}_2(\omega)\right] &= (\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_{2,1j}\left(\frac{1}{1-\omega^2}\boldsymbol{I}_3 - \boldsymbol{H}_{2,1j}\right)\boldsymbol{X}_{2,1j}^{\mathrm{T}}(\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2)^{-1}\sigma_2^2 + D\left[\widehat{\boldsymbol{\beta}}_2(1)\right] \\
&= \frac{(1-\omega^2)n}{\omega^2+n-1}\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{X}_{1j}\boldsymbol{L}^o\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\sigma_2^2 \\
&\quad + \frac{(1-\omega^2)2n}{\omega^2+2n-1}\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{1j}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{X}_{1j}\boldsymbol{L}^o\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\sigma_2^2 + D\left[\widehat{\boldsymbol{\beta}}_2(1)\right],
\end{aligned}
$$

and

$$
\begin{aligned}
\widehat{\sigma}_1^2(\omega) &= \widehat{\sigma}_1^2(1) + \frac{\omega^2-1}{(\omega^2-1)h_{1,1j}+1}\frac{1}{2n}\boldsymbol{r}_{1j}^{\mathrm{T}}\boldsymbol{P}_{T_s}\boldsymbol{r}_{1j} \\
&= \widehat{\sigma}_1^2(1) + \frac{\omega^2-1}{2(\omega^2+n-1)}\boldsymbol{r}_{W1j}^{\mathrm{T}}\boldsymbol{P}_{T_s}\boldsymbol{r}_{W1j},
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\sigma}_2^2(\omega) &= \widehat{\sigma}_2^2(1) + \frac{1}{6n}\boldsymbol{r}_{1j}^{\mathrm{T}}\boldsymbol{T}_d\left(\boldsymbol{H}_{2,1j} - \frac{1}{1-\omega^2}\boldsymbol{I}_3\right)^{-1}\boldsymbol{T}_d^{\mathrm{T}}\boldsymbol{r}_{1j} \\
&= \widehat{\sigma}_2^2(1) + \frac{\omega^2-1}{6(\omega^2+n-1)}\boldsymbol{r}_{W1j}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{r}_{W1j} \\
&\qquad + \frac{\omega^2-1}{3(\omega^2+2n-1)}(\boldsymbol{r}_{W1j}^{\mathrm{T}}+\boldsymbol{r}_B^{\mathrm{T}})\boldsymbol{P}_{T_{d1}}(\boldsymbol{r}_{W1j}+\boldsymbol{r}_B).
\end{aligned}
$$

Using the parametrization that $\sigma_e^2 = \sigma_2^2$ and $\sigma_\gamma^2 = \frac{1}{4}(\sigma_1^2 - \sigma_2^2)$,

$$\begin{aligned}\widehat{\sigma}_e^2(\omega) &= \widehat{\sigma}_e^2(1) + \frac{\omega^2-1}{6(\omega^2+n-1)}\boldsymbol{r}_{W1j}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{r}_{W1j} \\ &\quad + \frac{\omega^2-1}{3(\omega^2+2n-1)}(\boldsymbol{r}_{W1j}^{\mathrm{T}} + \boldsymbol{r}_B^{\mathrm{T}})\boldsymbol{P}_{T_{d1}}(\boldsymbol{r}_{W1j} + \boldsymbol{r}_B),\end{aligned}$$

$$\begin{aligned}\widehat{\sigma}_\gamma^2(\omega) &= \widehat{\sigma}_\gamma^2(1) + \frac{\omega^2-1}{24(\omega^2+n-1)}\boldsymbol{r}_{W1j}^{\mathrm{T}}\left(3\boldsymbol{P}_{T_s} - \boldsymbol{P}_{T_{d2}}\right)\boldsymbol{r}_{W1j} \\ &\quad - \frac{\omega^2-1}{12(\omega^2+2n-1)}(\boldsymbol{r}_{W1j}^{\mathrm{T}} + \boldsymbol{r}_B^{\mathrm{T}})\boldsymbol{P}_{T_{d1}}(\boldsymbol{r}_{W1j} + \boldsymbol{r}_B).\end{aligned}$$

Similarly, if the perturbed subject from sequence BAAB, we obtain

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_1(\omega) &= \widehat{\boldsymbol{\beta}}_1(1) + \frac{\omega^2-1}{(\omega^2-1)h_{1,2j}+1}(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1)^{-1}\mathbf{x}_{1,2j}\boldsymbol{T}_s^{\mathrm{T}}\boldsymbol{r}_{2j} \\ &= \widehat{\boldsymbol{\beta}}_1(1) + \frac{(\omega^2-1)n}{\omega^2+n-1}\left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}}\boldsymbol{r}_{W2j},\end{aligned}$$

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_2(\omega) &= \widehat{\boldsymbol{\beta}}_2(1) + (\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_{2,2j}^{\mathrm{T}}\left(\boldsymbol{H}_{2,2j} - \frac{1}{1-\omega^2}\boldsymbol{I}_3\right)^{-1}\boldsymbol{T}_d^{\mathrm{T}}\boldsymbol{r}_{2j} \\ &= \widehat{\boldsymbol{\beta}}_2(1) + \left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}}\boldsymbol{T}_d \\ &\qquad \left(\boldsymbol{H}_{2,2j} - \frac{1}{1-\omega^2}\boldsymbol{I}_3\right)^{-1}\boldsymbol{T}_d^{\mathrm{T}}(\boldsymbol{r}_{W2j} - \boldsymbol{P}_{T_{d1}}\boldsymbol{r}_B) \\ &= \widehat{\boldsymbol{\beta}}_2(1) + \frac{(\omega^2-1)n}{\omega^2+n-1}\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{r}_{W2j} \\ &\quad + \frac{(\omega^2-1)2n}{\omega^2+2n-1}\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}(\boldsymbol{r}_{W2j} - \boldsymbol{r}_B).\end{aligned}$$

with the corresponding dispersion matrices

$$\begin{aligned}D\left[\widehat{\boldsymbol{\beta}}_1(\omega)\right] &= \frac{\omega^2-1}{(\omega^2-1)h_{1,2j}+1}(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1)^{-1}\mathbf{x}_{1,2j}\mathbf{x}_{1,2j}^{\mathrm{T}}(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1)^{-1}\sigma_1^2 + D\left[\widehat{\boldsymbol{\beta}}_1(1)\right] \\ &= \frac{(1-\omega^2)n}{\omega^2+n-1}\left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}}\boldsymbol{X}_{2j}\boldsymbol{L}\left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\right)^{-1}\sigma_1^2 + D\left[\widehat{\boldsymbol{\beta}}_1(1)\right],\end{aligned}$$

$$\begin{aligned}D\left[\widehat{\boldsymbol{\beta}}_2(\omega)\right] &= (\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_{2,2j}\left(\frac{1}{1-\omega^2}\boldsymbol{I}_3 - \boldsymbol{H}_{2,2j}\right)^{-1}\boldsymbol{X}_{2,2j}^{\mathrm{T}}(\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2)^{-1}\sigma_2^2 + D\left[\widehat{\boldsymbol{\beta}}_2(1)\right] \\ &= \frac{(1-\omega^2)n}{\omega^2+n-1}\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{X}_{2j}\boldsymbol{L}^o\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\sigma_2^2 \\ &\quad + \frac{(1-\omega^2)2n}{\omega^2+2n-1}\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}_{2j}^{\mathrm{T}}\boldsymbol{P}_{T_{d1}}\boldsymbol{X}_{2j}\boldsymbol{L}^o\left(\boldsymbol{L}^{o\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}^o\right)^{-1}\sigma_2^2 + D\left[\widehat{\boldsymbol{\beta}}_2(1)\right],\end{aligned}$$

and

$$\begin{aligned}\widehat{\sigma}_e^2(\omega) &= \widehat{\sigma}_e^2(1) + \frac{\omega^2-1}{6(\omega^2+n-1)}\boldsymbol{r}_{W2j}^{\mathrm{T}}\boldsymbol{P}_{T_{d2}}\boldsymbol{r}_{W2j} \\ &\quad + \frac{\omega^2-1}{3(\omega^2+2n-1)}(\boldsymbol{r}_{W2j}^{\mathrm{T}}-\boldsymbol{r}_B^{\mathrm{T}})\boldsymbol{P}_{T_{d1}}(\boldsymbol{r}_{W2j}-\boldsymbol{r}_B),\end{aligned}$$

$$\begin{aligned}\widehat{\sigma}_\gamma^2(\omega) &= \widehat{\sigma}_\gamma^2(1) + \frac{\omega^2-1}{24(\omega^2+n-1)}\boldsymbol{r}_{W2j}^{\mathrm{T}}\left(3\boldsymbol{P}_{T_s}-\boldsymbol{P}_{T_{d2}}\right)\boldsymbol{r}_{W2j} \\ &\quad - \frac{\omega^2-1}{12(\omega^2+2n-1)}(\boldsymbol{r}_{W2j}^{\mathrm{T}}-\boldsymbol{r}_B^{\mathrm{T}})\boldsymbol{P}_{T_{d1}}(\boldsymbol{r}_{W2j}-\boldsymbol{r}_B).\end{aligned}$$

The estimators replacing $\widehat{\sigma}_e^2(1)$ and $\widehat{\sigma}_\gamma^2(1)$ are obtained via Theorem 2.2 (iv), and

$$\boldsymbol{r}_{Wij} = \begin{pmatrix} y_{ij1}-\overline{y}_{i\cdot1} \\ y_{ij2}-\overline{y}_{i\cdot2} \\ y_{ij3}-\overline{y}_{i\cdot3} \\ y_{ij4}-\overline{y}_{i\cdot4} \end{pmatrix}, \ \boldsymbol{r}_{W1j}+\boldsymbol{r}_B = \begin{pmatrix} y_{1j1}-\overline{y}_{\cdot\cdot1} \\ y_{1j2}-\overline{y}_{\cdot\cdot2} \\ y_{1j3}-\overline{y}_{\cdot\cdot3} \\ y_{1j4}-\overline{y}_{\cdot\cdot4} \end{pmatrix}, \ \boldsymbol{r}_{W2j}-\boldsymbol{r}_B = \begin{pmatrix} y_{2j1}-\overline{y}_{\cdot\cdot1} \\ y_{2j2}-\overline{y}_{\cdot\cdot2} \\ y_{2j3}-\overline{y}_{\cdot\cdot3} \\ y_{2j4}-\overline{y}_{\cdot\cdot4} \end{pmatrix},$$

where $\overline{y}_{i\cdot k} = \frac{1}{n}\sum_{j=1}^{n} y_{ijk}$ and $\overline{y}_{\cdot\cdot k} = \frac{1}{2n}\sum_{i=1}^{2}\sum_{j=1}^{n} y_{ijk}$, $i=1,2$, $k=1,2,3,4$. Hence, we get explicit solutions of VRR_{ij} and VRE_{ij}.

Next, using the parametrization

$$\boldsymbol{\beta} = \boldsymbol{L}\boldsymbol{\beta}_1 + \boldsymbol{L}^o\boldsymbol{\beta}_2,$$

we have

$$\Delta_{ij}\widehat{\boldsymbol{\beta}} = \boldsymbol{L}\Delta_{ij}\widehat{\boldsymbol{\beta}}_1 + \boldsymbol{L}^o\boldsymbol{\beta}_2\Delta_{ij}\widehat{\boldsymbol{\beta}}_2, \tag{30}$$

$$\Delta_{ij}D\left[\widehat{\boldsymbol{\beta}}\right] = \boldsymbol{L}D\left[\widehat{\boldsymbol{\beta}}_1\right]\boldsymbol{L}^{\mathrm{T}} + \boldsymbol{L}^oD\left[\widehat{\boldsymbol{\beta}}_2\right]\boldsymbol{L}^{o\mathrm{T}}. \tag{31}$$

By replacing (27) and (28) to (31) and (31), the explicit solutions of $\Delta_{ij}\widehat{\boldsymbol{\beta}}$ and $\Delta_{ij}D\left[\widehat{\boldsymbol{\beta}}\right]$ are achieved.

■

5. Discussion

The results in Theorem 4.1 show that the subject-level influence quantities mentioned in Section 3 are only decided by the within-sequence residual $\boldsymbol{r}_{Wij}$, total residual of the perturbed subject $\boldsymbol{r}_{Tij}$, and a column space $\mathcal{C}(\boldsymbol{T}_{d1})$.

Although this work has put special insight into the ABBA|BAAB design, the methodology for influence analysis proposed by this work can be easily generalised to other multiple-period two-treatment cross-over designs as long as the column space $\mathcal{C}(\boldsymbol{T}_{d1})$ is identified. For example, in the ABAB|BABA design, a similar $\mathcal{C}(\boldsymbol{T}_{d1})$ is defined by

$$\boldsymbol{T}_{d1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}^{\mathrm{T}}.$$

Deviation and expression of $\mathcal{C}(\boldsymbol{T}_{d1})$ and the closed-form influence quantities in general form for two-treatment cross-over designs are beyond this paper and are required further efforts.

References

Afsarinejad, K. and Hedayat, A. S. (2002). Repeated measurements designs for a model with self and simple mixed carryover effects. *Journal of Statistical Planning and Inference*, **106**, 449–459.

Beckman, R. J., Nachtsheim, C. and Cook, R. D. (1987). Diagnostics for mixed-model analysis of variance. *Technometrics*, **29**, 413–426.

Belsley, D. A., Kuh, E. and Welsch, R. E. (2004). *Regression diagnostics: Identifying influential data and sources of collinearity.* John Wiley & Sons, Inc., Hoboken, New Jersey.

Carrière, K. C. and Huang, R. (2000). Crossover designs for two-treatment clinical trials. *Journal of Statistical Planning and Inference*, **87**, 125–134.

Carrière, K. C. and Reinsel, G. C. (1992). Investigation of dual-balanced crossover designs for two treatments. *Biometrics*, **48**, 1157–1164.

Christensen, R., Pearson, L. M. and Johnson, W. (1992). Case-deletion diagnostics for mixed models. *Technometrics*, **34**, 38–45.

Cook, R. D. (1986). Assessment of local influence. *Jounal of the Royal Statistical Society: Series B*, **48**, 133–169.

Hao, C., von Rosen, T. and von Rosen, D. (2011). Local influence analysis in 2×2 cross-over designs. Tech. Rep. 2011:1, Department of Statistics, Stockholm Univeristy.

Hedayat, A. S., Stufken, J. and Yang, M. (2006). Optimal and efficient crossover designs when subject effects are random. *Journal of the American Statistical Association*, **101**, 1031–1038.

Hedayat, A. S. and Zheng, W. (2010). Optimal and efficient crossover designs for test-control study when subject effects are random. *Journal of the American Statistical Association*, **105**, 1581–1592.

Jones, B. and Kenward, M. G. (1989). *Design and analysis of cross-over trials.* Chapman & Hall, London and New York.

Kershner, R. and Federer, W. (1981). Two-treatment crossover designs for estimating a variety of effects. *Journal of the American Statistical Association*, **76**, 612–619.

Lesaffre, E. and Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, **54**, 570–582.

Park, D. K., Bose, M., Notz, W. I. and Dean, A. M. (2010). Efficient crossover designs in the presence of interactions between direct and carry-over treatment effects. *Journal of Statistical Planning and Inference*, **141**, 846–860.

Putt, M. E. and Chinchilli, V. M. (2000). A robust analysis of crossover designs using multisample generalized l-statistics. *Journal of the American Statistical Association*, **95**, 1256–1262.

Yan, Z. and Locke, C. (2010). Crossover designs for comparing test treatments to a control treatment when subject effects are random. *Journal of Statistical Planning and Inference*, **140**, 1214–1224.