Electronic Journal of Statistics
Vol. 4 (2010) 486–524
ISSN: 1935-7524
DOI: 10.1214/09-EJS503

Estimation for a longitudinal linear model with measurement errors

Laura Dumitrescu*

University of Ottawa
Department of Mathematics and Statistics
585 King Edward Ave., Ottawa, Ontario, K1N 6N5, Canada
e-mail: ldumi030@uottawa.ca

Abstract: In this article we introduce a multivariate structural linear error-in-variables model which is suitable for longitudinal data. We construct estimators of the regression parameters, which correspond to the modified least squares estimators used in the univariate case. We show that these estimators are consistent. We prove a central limit theorem, which is completely data-based, under the assumption that the vector of latent variables belongs to the generalized domain of attraction of the normal law. Our results can be viewed as an extension of the results of [12] to include the longitudinal case.

AMS 2000 subject classifications: Primary 62H12, 62J99; secondary 60F05, 60E07.
Keywords and phrases: Measurement errors, longitudinal data, central limit theorems, generalized domain of attraction of the normal law.

Received October 2009.

1. Introduction

Error-in-variables (also known as measurement error) models are regression models where the covariates cannot be measured directly or without error. Thus, the response variable, y is assumed to depend on a variable ξ (called the latent variable), which is measured by $\mathbf{x}$. A simple linear error-in-variable model can be written in the following form:

$$\begin{aligned} y &= \alpha + \beta\xi + \varepsilon, \\ x &= \xi + \delta. \end{aligned}$$

If ξ is a random variable, the model is called a structural error-in-variable model. We also mention here that the latent variable ξ is assumed to be independent of the errors ε and δ. Thus, if an n dimensional sample is available, the collected data are $(x_i, y_i)_{1\leq i\leq n}$ and the unknown parameters of the model are α and β.

The error-in-variables model cannot be reformulated as a classical regression model with random design since the regressor x is correlated with the error:

$$y = \alpha + \beta x + (\varepsilon - \beta\delta).$$

*This article is based on a portion of the author's doctoral thesis.

The reader may refer to the monograph [5], which is a compendium of up-to-that-date theoretical methods and practical applications associated with measurement error models. The recent monograph [3] contains further developments in this area, viewed from a contemporary perspective.

It is known that the error-in-variables models are, in general, not identifiable and therefore, it becomes impossible to consistently estimate the parameters from the data. In the univariate case, several additional assumptions that make the model identifiable can be found in the literature (for further details, see [4], Section 1.2.1). Under the assumption that the variance of δ is known, the authors of [4] have obtained the modified least squares estimators. However, their asymptotic covariance matrix is a function of unknown parameters.

The author of [12] obtains CLTs for the univariate linear structural models when the explanatory variables are in the domain of attraction of the normal law (DAN), which constitutes a new approach in the study of measurement error models. The concept of DAN has also been used in a regression context by other researchers (for instance, see [10]). This approach has two advantages. Firstly, the assumption about the finiteness of the variance of the latent variable can be relaxed. Secondly, Studentized and self-normalized CLTs (that depend only on the data) for distributions which are in DAN, are already available due to [7].

In the multivariate measurement error regression context, the author of [8] assumes the covariance of the error δ to be known, or estimated to acquire identifiability of the model. The estimation of parameters is suggested to be performed in two steps: firstly, the reliability matrix (the correspondent of the reliability ratio in the univariate case) needs to be estimated. Secondly, classical estimation methods, such as least squares are used to obtain estimators of the unknown parameters.

Longitudinal models are extensively used in biostatistics, sociology and psychology to express the evolution in time, or the occurrence of a response variable, in terms of a number of significant covariates. Since some of these covariates cannot be recorded directly it is important to propose and study longitudinal models which reflect this reality. Data for longitudinal studies are often collected on the same individual on different occasions. The collected measurements are considered to be independent across individuals and correlated within each individual.

The authors of [2] consider a longitudinal linear mixed model with measurement error. Identifiability of their model follows from assuming a restricted model for the covariance matrix of the latent variables. They describe extensions to the case when replicate or validation data is available. For estimation of the regression parameters, they use a regression calibration method, with a substitution for the unknown covariates, which is corrected for estimation of the variance parameters.

In this article, we consider the case of a longitudinal error-in-variable model with no subject specific random effects. The true predictor, $\boldsymbol{\xi}$ is assumed to have the same effect for each repeated measurement within the individual. We employ the method of moments, as used in [4] to obtain consistent estimators appropriate to the multivariate structure of the data. They correspond to the modified least squares estimators of a univariate structural measurement error model. The estimators are constructed under the assumption that the diagonal of the covariance matrix of the error δ is known. This additional assumption ensures the identifiability of our model. We do not discuss here the case when this information is not available. (In this situation, an estimator of the diagonal of the covariance matrix should be provided; for additional details see Section 2 in [8]).

We use the technique developed in [12] to obtain the CLTs for the estimators of the parameters of interest of our model. In our multivariate context, we use the assumption that the vector formed by the latent variables belongs to the generalized domain of attraction of the normal law (GDAN). The concept of GDAN was defined in [9] and it does not constitute a trivial extension of DAN to the multivariate case, as it had been previously assumed (see [11] or [13]). By applying the results obtained in [7], this approach enables us to obtain Studentized and self-normalized CLTs for our estimators.

Thus, our results apply to the case of multivariate data where the components of the vectors formed by the explanatory variables are not necessarily independent, allowing for some degree of correlation between the components of latent vector. As in [12], this approach allows us to obtain CLTs which are data-based and do not involve unknown parameters. Our results can be viewed as a generalization of the results obtained in [12] to include the case of longitudinal data. This generalization is not trivial, the technical difficulty arising from the use of GDAN concept. For the sake of consistency with the literature and to facilitate a comparison of results, we employ the notation and a structure of proofs layed out in [12].

The article is organized as follows. In Section 2 we state our model assumptions and obtain the estimators of the parameters of interest. In Section 3 we introduce the concept of GDAN and give preliminary results, which will be used in the next sections. The main result of this section is Theorem 3.4, which states that the inner product between a vector in GDAN and a vector whose components have moments of order four is in DAN. In Section 4 we obtain the CLTs for our estimators. The Appendix contains the proofs of some technical results. We mention that, as in [12], our proofs cover two distinct cases: the case when all the components of the covariance matrix of the latent vector are finite, and the case when at least one of these components is infinite.

2. Model assumptions and estimators of the parameters

2.1. The model

We consider a sample of n individuals, whose responses are recorded, at fixed moments of time (which are denoted for simplicity by $1, 2, \ldots, m$). For any $i \in \{1, 2, \ldots, n\}$, let $\mathbf{y}_i = (y_{i1}, \ldots, y_{im})^T$ be the collection of m responses supplied by the i-th individual. More precisely, y_{ij} represents the response of the i-th individual at time j. Alternatively, we can think of n as a sample of independent clusters of equal size m. Within each cluster i, the observations $(y_{i1}, \dots, y_{im})$ might be correlated.

Each response y_{ij} depends on a covariate ξ_{ij}, which is *unobservable* (or latent): instead of the covariate ξ_{ij}, one observes a surrogate variable x_{ij}. This happens for any individual $i \in \{1, 2, \dots, n\}$, and for any occasion $j \in \{1, 2, \dots, m\}$. For any $i \in \{1, 2, \dots, n\}$, we denote by $\boldsymbol{\xi}_i = (\xi_{i1}, \dots, \xi_{im})^T$ the collection of m unobservable covariates which correspond to the i-th individual, and by $\mathbf{x}_i = (x_{i1}, \dots, x_{im})^T$ the collection of m respective surrogate variables.

Both variables y_{ij} and x_{ij} are observed with errors. More precisely, we consider the following error-in-variables model: for any $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$

$$\begin{aligned} y_{ij} &= \alpha + \beta\xi_{ij} + \varepsilon_{ij}, \\ x_{ij} &= \xi_{ij} + \delta_{ij}, \end{aligned} \tag{2.1}$$

where α, β are unknown parameters (of dimension 1), and $\varepsilon_{ij}, \delta_{ij}$ are the random error terms. For any $i \in \{1, \dots, n\}$, $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \dots, \varepsilon_{im})^T$ and $\boldsymbol{\delta}_i = (\delta_{i1}, \dots, \delta_{im})^T$ represent the error random vectors.

We assume that $\{\boldsymbol{\xi}_i\}_{1 \le i \le n}$ is a sequence of i.i.d. random vectors with mean $\boldsymbol{\mu} = (\mu_1, \dots, \mu_m)^T$, i.e. $\mu_j := E(\xi_{ij})$, for $j \in \{1, \dots, m\}$ and $i \in \{1, \dots, n\}$. We denote for any $i \in \{1, \dots n\}$, and $j, k \in \{1, \dots m\}$

$$\sigma_{\boldsymbol{\xi},jj} := \mathrm{Var}(\xi_{ij}), \text{ and } \sigma_{\boldsymbol{\xi},jk} := \mathrm{Cov}(\xi_{ij}, \xi_{ik}).$$

In matrix notation, we write

$$\boldsymbol{\Sigma}_{\boldsymbol{\xi}} = (\sigma_{\boldsymbol{\xi},jk})_{1 \le j,k \le m}.$$

We assume that the errors $\{(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i)\}_{1 \le i \le n}$ form a sequence of i.i.d. random vectors, and denote

$$\sigma_{\boldsymbol{\varepsilon},jk} := E(\varepsilon_{ij}\varepsilon_{ik}), \quad \sigma_{\boldsymbol{\delta},jk} := E(\delta_{ij}\delta_{ik}), \quad \sigma_{\boldsymbol{\varepsilon}\boldsymbol{\delta},jk} := E(\varepsilon_{ij}\delta_{ik}),$$

for any $i \in \{1, \dots, n\}$ and $j, k \in \{1, \dots, m\}$. In matrix notation,

$$\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} := (\sigma_{\boldsymbol{\varepsilon},jk})_{1 \le j,k \le m}, \quad \boldsymbol{\Sigma}_{\boldsymbol{\delta}} := (\sigma_{\boldsymbol{\delta},jk})_{1 \le j,k \le m}, \quad \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\delta}} := (\sigma_{\boldsymbol{\varepsilon}\boldsymbol{\delta},jk})_{1 \le j,k \le m}.$$

2.2. Estimation

We are interested in the estimation of α and β. To obtain consistent estimators for α and β, we assume that the diagonal elements of $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\delta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\delta}}$ are known. This assumption was used in the univariate case (i.e. $m = 1$) to construct the modified least squares estimators of α and β (see [4]) and also in a multivariate regression context (see [8]).

For any $i \le n, j \le m$, we have:

$$\begin{aligned} E(x_{ij}) &= E(\xi_{ij}) + E(\delta_{ij}) = \mu_j, \\ E(y_{ij}) &= \alpha + \beta E(\xi_{ij}) + E(\varepsilon_{ij}) = \alpha + \beta\mu_j, \\ \mathrm{Var}(x_{ij}) &= \mathrm{Var}(\xi_{ij}) + \mathrm{Var}(\delta_{ij}) = \sigma_{\boldsymbol{\xi},jj} + \sigma_{\boldsymbol{\delta},jj}, \\ \mathrm{Var}(y_{ij}) &= \beta^2\mathrm{Var}(\xi_{ij}) + \mathrm{Var}(\varepsilon_{ij}) = \beta^2\sigma_{\boldsymbol{\xi},jj} + \sigma_{\boldsymbol{\varepsilon},jj}, \\ \mathrm{Cov}(x_{ij}, y_{ij}) &= \beta\mathrm{Var}(\xi_{ij}) + \mathrm{Cov}(\varepsilon_{ij}, \delta_{ij}) = \beta\sigma_{\boldsymbol{\xi},jj} + \sigma_{\boldsymbol{\varepsilon\delta},jj}. \end{aligned}$$

Using the method of moments (see also [4], Section 1.3.1), we obtain the following system of equations, for any $j \in 1, \dots, m$:

$$\widehat{\mu}_j = \frac{1}{n}\sum_{i=1}^n x_{ij} := \overline{x}_j, \tag{2.2}$$

$$\widehat{\alpha} + \widehat{\beta}\widehat{\mu}_j = \frac{1}{n}\sum_{i=1}^n y_{ij} := \overline{y}_j, \tag{2.3}$$

$$\widehat{\sigma}_{\boldsymbol{\xi},jj} + \sigma_{\boldsymbol{\delta},jj} = \frac{1}{n}\sum_{i=1}^n (x_{ij} - \overline{x}_j)^2, \tag{2.4}$$

$$\widehat{\beta}^2\widehat{\sigma}_{\boldsymbol{\xi},jj} + \widehat{\sigma}_{\boldsymbol{\varepsilon},jj} = \frac{1}{n}\sum_{i=1}^n (y_{ij} - \overline{y}_j)^2, \tag{2.5}$$

$$\widehat{\beta}\widehat{\sigma}_{\boldsymbol{\xi},jj} + \sigma_{\boldsymbol{\varepsilon\delta},jj} = \frac{1}{n}\sum_{i=1}^n (x_{ij} - \overline{x}_j)(y_{ij} - \overline{y}_j). \tag{2.6}$$

By taking the sum of the m equations in (2.2) and (2.3), we obtain the following estimator of α:

$$\widehat{\alpha}_n = \overline{\overline{y}} - \widehat{\beta}_n\overline{\overline{x}}, \tag{2.7}$$

where

$$\overline{\overline{x}} = \frac{1}{nm}\sum_{i=1}^n\sum_{j=1}^m x_{ij} \quad \text{and} \quad \overline{\overline{y}} = \frac{1}{nm}\sum_{i=1}^n\sum_{j=1}^m y_{ij}.$$

For the purpose of the present article, it is useful to develop a vector notation. Let

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i \quad \text{and} \quad \overline{\mathbf{y}} = \frac{1}{n}\sum_{i=1}^n \mathbf{y}_i$$

be the average of $(\mathbf{x}_1, \dots, \mathbf{x}_n)$, respectively $(\mathbf{y}_1, \dots, \mathbf{y}_n)$. Note that $\overline{\mathbf{x}}$ is an m-dimensional random vector whose components are $\overline{x}_j, 1 \le j \le m$. Similarly, $\overline{\mathbf{y}} = (\overline{y}_1, \dots, \overline{y}_m)^T$.

By taking the sum of the m equations in (2.4) and (2.6), we obtain that:

$$\widehat{\beta}_n = \frac{\sum_{i=1}^n[(\mathbf{x}_i - \overline{\mathbf{x}})^T(\mathbf{y}_i - \overline{\mathbf{y}}) - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}})]}{\sum_{i=1}^n[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}, \tag{2.8}$$

where $\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\varepsilon\delta}}) = \sum_{j=1}^{m} \sigma_{\boldsymbol{\varepsilon\delta},jj}$ and $\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}}) = \sum_{j=1}^{m} \sigma_{\boldsymbol{\delta},jj}$. The estimator $\widehat{\beta}_n$ is obtained under the additional assumption $\sum_{i=1}^{n}[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})] > 0$. We remark that a similar assumption is used in the univariate case in order to obtain the modified least square estimator (see [4]).

Remark 2.1. A different estimator of β can be obtained using (2.5) and (2.6), when the diagonal elements of $\mathbf{\Sigma}_{\boldsymbol{\varepsilon\delta}}$ and $\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}$ (instead of $\mathbf{\Sigma}_{\boldsymbol{\delta}}$) are known. Its form is similar to (2.8) and the asymptotic results are similar to those derived in this article.

Remark 2.2. The error-in-variables model with replications is a particular case of model (2.1), obtained when $\xi_{i1} = \xi_{i2} = \ldots = \xi_{im}$. In this case, the form of the estimators of α and β are also given by (2.7) and (2.8), respectively.

The goal of this article is to evaluate the asymptotic properties of the estimators defined by (2.7) and (2.8). This purpose requires the introduction of additional model assumptions as follows.

The distribution of $\boldsymbol{\xi}_1$ is assumed to be *full*, i.e. for each m-dimensional vector of norm 1, $\mathbf{u}$, the random variable $\mathbf{u}^T\boldsymbol{\xi}_1$ is not a constant, almost surely. This is a standard assumption in the context of the generalized domain of attraction of the normal law (GDAN) and allows us to use results obtained in [11] and [7].

Recall that $\boldsymbol{\xi}_1 - \boldsymbol{\mu}$ has a *symmetric* distribution if its distribution is equal to the distribution of $-(\boldsymbol{\xi}_1 - \boldsymbol{\mu})$.

Our model assumptions are the following:

$(A1)$ $\boldsymbol{\xi}_1$ lies in GDAN,
$\boldsymbol{\xi}_1 - \boldsymbol{\mu}$ has a symmetric distribution, whenever the matrix $\mathbf{\Sigma}_{\boldsymbol{\xi}}$ has at least one element on the diagonal which is ∞

$(A2)$ $E(\varepsilon_{1j}) = 0,\ E(\delta_{1j}) = 0,\ E(\varepsilon_{1j}^4) < \infty,\ E(\delta_{1j}^4) < \infty$ with $1 \le j \le m$,
and $\mathbf{\Sigma}_{\text{error}} := \begin{pmatrix} \mathbf{\Sigma}_{\boldsymbol{\varepsilon}} & \mathbf{\Sigma}_{\boldsymbol{\varepsilon\delta}} \\ \mathbf{\Sigma}_{\boldsymbol{\varepsilon\delta}}^T & \mathbf{\Sigma}_{\boldsymbol{\delta}} \end{pmatrix}$ is positive definite

$(A3)$ $(\boldsymbol{\xi}_i)_{1\le i\le n}$ and $\{(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i)\}_{1\le i\le n}$ are independent.

We note that assumptions $(A1)$–$(A4)$ are the multidimensional versions of the assumptions used in [12].

The main results of this article are the consistency theorem (Theorem 4.1) and the CLT (Theorem 4.12). They are stated as follows:

1. Assume that $\boldsymbol{\xi}_1 \in GDAN$ and $(A3)$ holds. Then, the estimators $\widehat{\beta}_n$ and $\widehat{\alpha}_n$ are weakly consistent.

2. Assume that $(A1)$, $(A2)$ and $(A3)$ hold. Then, as $n \to \infty$, we have:

$$(a) \qquad \frac{\sum_{i=1}^{n}\left[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right](\widehat{\beta}_n - \beta)}{\sqrt{\sum_{i=1}^{n}\tilde{u}_i(n)^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

$$(b) \qquad \frac{\sqrt{n}(\widehat{\alpha}_n - \alpha)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\tilde{v}_i(n) - \overline{\tilde{v}(n)})^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

where $\tilde{u}_i(n) = (\mathbf{x}_i - \overline{\mathbf{x}})^T(\mathbf{y}_i - \overline{\mathbf{y}}) - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}}) - \widehat{\beta}_n[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]$,

$$\tilde{v}_i(n) = \frac{1}{m}\sum_{j=1}^{m} y_{ij} - \alpha - \widehat{\beta}_n \frac{1}{m}\sum_{j=1}^{m} x_{ij} - \frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^{n}[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}\tilde{u}_i(n)$$

and $\overline{\tilde{v}(n)} = \frac{1}{n}\sum_{i=1}^{n}\tilde{v}_i(n)$.

3. Preliminary results for GDAN

In this section we present the concept of the generalized domain of attraction of the normal law (GDAN), introduced in [9]. For additional properties of GDAN, see [11] or [13]. We begin with a definition.

Definition 3.1. *Let $\boldsymbol{\xi}$ be an m-dimensional random vector. We say that $\boldsymbol{\xi}$ belongs to the generalized domain of attraction of the normal law (and we write $\boldsymbol{\xi} \in GDAN$) if there exists a sequence $(\mathbf{B}_n)_{n\geq 1}$ of non-stochastic $m\times m$ matrices and a sequence $(\mathbf{A}_n)_{n\geq 1}$ of non-stochastic m-dimensional vectors, such that:*

$$\mathbf{B}_n\left(\sum_{i=1}^{n}\boldsymbol{\xi}_i - \mathbf{A}_n\right) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I}), \tag{3.1}$$

where $(\boldsymbol{\xi}_i)_{1\leq i\leq n}$ is a sequence of i.i.d. random vectors with the same distribution as $\boldsymbol{\xi}$.

Remark 3.2. (i) If $m = 1$, the above definition coincides with the definition of the domain of attraction of the normal law (DAN). It is known that $\boldsymbol{\xi} \in GDAN$ implies that each component $\xi_j \in DAN$ for $j \leq m$. In general, the converse of this statement is not true. However, Remark (ii) of [11] points out that in the case when $\boldsymbol{\xi}$ has a spherically symmetric distribution, the condition $\boldsymbol{\xi} \in GDAN$ becomes equivalent with the condition that each $\xi_j \in DAN$ for $j \leq m$ (or to $\|\boldsymbol{\xi}\| \in DAN$).

(ii) If $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)^T \in GDAN$, then $\mathrm{E}(\|\boldsymbol{\xi}\|^r) < \infty$, for $0 \leq r < 2$ (in particular, $\mu_j = \mathrm{E}(\xi_j) < \infty$, for all $j \leq m$) and the sequence $(\mathbf{A}_n)_{n\geq 1}$ can be taken as $\mathbf{A}_n = n\boldsymbol{\mu}$, where $\boldsymbol{\mu} = \mathrm{E}(\boldsymbol{\xi})$ (see also Remark (ii), p. 193 in [11]).
Note that, if $\boldsymbol{\xi} \in GDAN$, since $\mathrm{E}(\|\boldsymbol{\xi}\|) < \infty$, the condition $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$ is equivalent to $\mathrm{E}(\|\boldsymbol{\xi}\|^2) < \infty$ (or $\mathrm{E}(|\xi_j|^2) < \infty$ for all $j \leq m$).

(iii) If $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$, by the classical CLT, the sequence $(\mathbf{B}_n)_{n\geq 1}$ of (3.1) can be taken to be $\mathbf{B}_n = \sqrt{n}\,\boldsymbol{\Sigma}_{\boldsymbol{\xi}}^{-1/2}$.

(iv) The matrices $\mathbf{B}_n$ can be taken to be nonsingular and symmetric. For a complete discussion, see Remark (ii) in [11].

This result is a consequence of Remarks (ii) and (iii), p. 193–194 in [11].

Lemma 3.3. *If $\boldsymbol{\xi} \in GDAN$ and $\mathrm{E}(\boldsymbol{\xi}) = \boldsymbol{\mu}$, then $\|\boldsymbol{\xi} - \boldsymbol{\mu}\| \in DAN$,*

$$\frac{1}{b_n}\left(\sum_{i=1}^{n} \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\| - a_n\right) \xrightarrow{\mathcal{D}} N(0,1), \text{ and } \frac{\sum_{i=1}^{n} \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^2}{b_n^2} \xrightarrow{P} 1, \tag{3.2}$$

where $(\boldsymbol{\xi}_i)_{1\le i\le n}$ are i.i.d copies of $\boldsymbol{\xi}$, $b_n^2 = \mathrm{trace}(\mathbf{B}_n^{-2})$, and $(\mathbf{B}_n)_{n\ge 1}$ is a sequence of matrices for which (3.1) holds.

Hereafter, the notation $\mathbf{A} > 0$ will be used to denote a positive definite matrix and the notation $\mathbf{A}^{1/2}$ will be used for its square root. The eigenvalues of an m-dimensional matrix $\mathbf{A}$ will be denoted by $\lambda_j(\mathbf{A})$ and $\lambda_{\max}(\mathbf{A}) = \max_{j\le m} \lambda_j(\mathbf{A})$. Also, trace($\mathbf{A}$) will be used for the sum of its diagonal elements (or, equivalently, the sum of its eigenvalues).

The following theorem is an essential tool for the development of our asymptotic results. The result is important in its own right since it provides sufficient conditions for the inner product between a random vector in GDAN and a random vector with finite fourth order moments to be in the domain of attraction of the normal law. The case of $m = 1$ was fully resolved under more general conditions in [10]. In the case of infinite variance, we assume that $\boldsymbol{\xi} - \boldsymbol{\mu}$ has a symmetric full distribution, as we rely on a result in [6]. We use the method of the proof of Lemma 4, in [12] adapted to a multivariate setting.

Theorem 3.4. *Let $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m)^T \in GDAN$ be an m-dimensional random vector with a full distribution. If $\mathrm{Var}(\|\boldsymbol{\xi}\|) = \infty$ we assume, in addition, that the distribution of $\boldsymbol{\xi} - \boldsymbol{\mu}$ is symmetric. Let $\boldsymbol{\varepsilon}$ be am m-dimensional random vector, such that $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{E}(|\varepsilon_j|^4) < \infty$, for all $j \le m$. Assume that $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} > 0$, where $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ is the covariance matrix of $\boldsymbol{\varepsilon}$. If $\boldsymbol{\xi}$ and $\boldsymbol{\varepsilon}$ are independent, then,*

$$\boldsymbol{\xi}^T\boldsymbol{\varepsilon} = \sum_{j=1}^{m} \xi_j\varepsilon_j \in \mathrm{DAN}.$$

Proof. Case 1. Assume that $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$. Then, the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\xi}}$ has only finite entries and it follows that $\mathrm{Var}(\boldsymbol{\xi}^T\boldsymbol{\varepsilon}) < \infty$. We will prove that $\mathrm{Var}(\boldsymbol{\xi}^T\boldsymbol{\varepsilon}) > 0$ and the conclusion will follow by applying the CLT.

Since $\boldsymbol{\xi}$ is independent of $\boldsymbol{\varepsilon}$,

$$\mathrm{Var}(\boldsymbol{\xi}^T\boldsymbol{\varepsilon}) = \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\xi}}\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}).$$

Since the distribution of $\boldsymbol{\xi}$ is full, $\boldsymbol{\Sigma}_{\boldsymbol{\xi}}$ is positive definite and hence $\boldsymbol{\Sigma}_{\boldsymbol{\xi}}^{1/2}\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\boldsymbol{\Sigma}_{\boldsymbol{\xi}}^{1/2}$ is positive definite. Therefore, $\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\xi}}\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) = \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\xi}}^{1/2}\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\boldsymbol{\Sigma}_{\boldsymbol{\xi}}^{1/2}) > 0$, since $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} > 0$.

Case 2. Assume that $\mathrm{Var}(\|\boldsymbol{\xi}\|) = \infty$ (at least one component of $\boldsymbol{\xi}$ has infinite variance). In this case, we assume in addition that $\boldsymbol{\xi} - \boldsymbol{\mu}$ has a symmetric distribution. Let $\{\boldsymbol{\xi}_i\}_{1\le i\le n}$ and $\{\boldsymbol{\varepsilon}_i\}_{1\le i\le n}$ be i.i.d copies of $\boldsymbol{\xi}$ and $\boldsymbol{\varepsilon}$, respectively. We denote $\boldsymbol{\xi}_i = (\xi_{i1}, \dots, \xi_{im})^T$ and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \dots, \varepsilon_{im})^T$.

(a) Assume that $\mathbf{\Sigma}_{\boldsymbol{\varepsilon}} = \mathbf{I}$. First, we consider the case when $\mathrm{E}(\boldsymbol{\xi}) = \mathbf{0}$.

Since $\boldsymbol{\xi} \in GDAN$, there exist a sequence of symmetric and non-singular nonstochastic matrices $\mathbf{B}_n$ such that $\mathbf{B}_n \sum_{i=1}^n \boldsymbol{\xi}_i \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I})$. Denote $b_n^2 := \mathrm{trace}[(\mathbf{B}_n)^{-2}]$ and note that $b_n^2 \to \infty$. To show that $\boldsymbol{\xi}^T \boldsymbol{\varepsilon} \in DAN$, by [1], it suffices to prove that

$$b_n^{-2} \sum_{i=1}^n (\boldsymbol{\xi}_i^T \boldsymbol{\varepsilon}_i)^2 \xrightarrow{P} 1 \tag{3.3}$$

We have $b_n^{-2} \sum_{i=1}^n (\boldsymbol{\xi}_i^T \boldsymbol{\varepsilon}_i)^2 = \gamma_{n,1}\gamma_{n,2}$, where:

$$\gamma_{n,1} := \frac{\sum_{i=1}^n (\boldsymbol{\xi}_i^T \boldsymbol{\varepsilon}_i)^2}{\sum_{i=1}^n \boldsymbol{\xi}_i^T \boldsymbol{\xi}_i} \text{ and } \gamma_{n,2} := \frac{\sum_{i=1}^n \boldsymbol{\xi}_i^T \boldsymbol{\xi}_i}{b_n^2}.$$

By Lemma 3.3, $\gamma_{n,2} \xrightarrow{P} 1$.

In what follows we show that $\gamma_{n,1} \xrightarrow{P} 1$. Let $\epsilon > 0$ be fixed. Using Chebychev's inequality, we have the following:

$$\begin{aligned} P(|\gamma_{n,1} - 1| > \epsilon) &\leq \epsilon^{-2} \mathrm{E}\left[\left(\sum_{i=1}^n \frac{\boldsymbol{\xi}_i^T (\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T - \mathbf{I}) \boldsymbol{\xi}_i}{\sum_{i=1}^n \boldsymbol{\xi}_i^T \boldsymbol{\xi}_i}\right)^2\right] \\ &= \epsilon^{-2}(I_1 + I_2), \end{aligned} \tag{3.4}$$

where:

$$\begin{aligned} I_1 &= \mathrm{E}\left[\sum_{i=1}^n \left(\frac{\boldsymbol{\xi}_i^T (\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T - \mathbf{I}) \boldsymbol{\xi}_i}{\sum_{i=1}^n \boldsymbol{\xi}_i^T \boldsymbol{\xi}_i}\right)^2\right], \\ I_2 &= \mathrm{E}\left[\sum_{i \neq l}^n \frac{\boldsymbol{\xi}_i^T (\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T - \mathbf{I}) \boldsymbol{\xi}_i \boldsymbol{\xi}_l^T (\boldsymbol{\varepsilon}_l \boldsymbol{\varepsilon}_l^T - \mathbf{I}) \boldsymbol{\xi}_l}{(\sum_{i=1}^n \boldsymbol{\xi}_i^T \boldsymbol{\xi}_i)^2}\right]. \end{aligned}$$

We treat I_1 first. Let $\mathbf{C}_n := \sum_{i=1}^n \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T$ which can be assumed by Lemma 2.3 of [11] to be non-singular. We have the following evaluation:

$$\begin{aligned} [\boldsymbol{\xi}_1^T (\boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^T - \mathbf{I}) \boldsymbol{\xi}_1]^2 &\leq \lambda_{\max}(\boldsymbol{\xi}_1 \boldsymbol{\xi}_1^T) \lambda_{\max}[(\boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^T - \mathbf{I})^2] \boldsymbol{\xi}_1^T \boldsymbol{\xi}_1 \\ &\leq \mathrm{trace}(\boldsymbol{\xi}_1 \boldsymbol{\xi}_1^T) \lambda_{\max}[(\boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^T - \mathbf{I})^2] \boldsymbol{\xi}_1^T \boldsymbol{\xi}_1 \\ &= \lambda_{\max}[(\boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^T - \mathbf{I})^2] (\boldsymbol{\xi}_1^T \boldsymbol{\xi}_1)^2 \\ &\leq \lambda_{\max}[(\boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^T - \mathbf{I})^2] \lambda_{\max}^2(\mathbf{C}_n) (\boldsymbol{\xi}_1^T \mathbf{C}_n^{-1} \boldsymbol{\xi}_1)^2. \end{aligned}$$

We have

$$\lambda_{\max}(\mathbf{C}_n) = \lambda_{\max}\left(\sum_{i=1}^n \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T\right) \leq \sum_{i=1}^n \lambda_{\max}(\boldsymbol{\xi}_i \boldsymbol{\xi}_i^T) \leq \sum_{i=1}^n \mathrm{trace}(\boldsymbol{\xi}_i \boldsymbol{\xi}_i^T) = \sum_{i=1}^n \boldsymbol{\xi}_i^T \boldsymbol{\xi}_i.$$

Using the independence between $\boldsymbol{\xi}$ and $\boldsymbol{\varepsilon}$ we obtain that

$$\begin{aligned} I_1 &= n\mathrm{E}\left[\left(\frac{\boldsymbol{\xi}_1^T(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^T - \mathbf{I})\boldsymbol{\xi}_1}{\sum_{i=1}^n \boldsymbol{\xi}_i^T\boldsymbol{\xi}_i}\right)^2\right] \\ &\leq n\mathrm{E}(\lambda_{\max}[(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^T - \mathbf{I})^2])\mathrm{E}[(\boldsymbol{\xi}_1^T\mathbf{C}_n^{-1}\boldsymbol{\xi}_1)^2]. \end{aligned}$$

Using the fact that $\mathrm{E}(\varepsilon_j^4) < \infty$ for all $j \leq m$ and $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} = \mathbf{I}$, we obtain:

$$\begin{aligned} \mathrm{E}\{\lambda_{\max}[(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^T - \mathbf{I})^2]\} \leq m^2 \sum_{j,k=1}^m \mathrm{E}[(\varepsilon_{1j}\varepsilon_{1k} - \delta_{jk})^2] \leq m^2 \sum_{j,k=1}^m \mathrm{E}(\varepsilon_{1j}^2\varepsilon_{1k}^2) \\ \leq m^2 \sum_{j=1}^m \mathrm{E}(\varepsilon_{1j}^4) < \infty, \end{aligned}$$

Here we let $\delta_{jk} = 1$, if $j = k$ and $\delta_{jk} = 0$, if $j \neq k$.

From the proof of Theorem 3.4 of [6], since $\boldsymbol{\xi} \in GDAN$, it follows that $n\mathrm{E}[(\boldsymbol{\xi}_1^T\mathbf{C}_n^{-1}\boldsymbol{\xi}_1)^2]\longrightarrow 0$. (This theorem is proved under the assumption that $\boldsymbol{\xi}$ has a symmetric distribution.) Hence,

$$I_1 \longrightarrow 0, \text{ as } n \to \infty. \tag{3.5}$$

We now treat I_2. Using the fact that $(\boldsymbol{\xi}_i)_{i\leq n}$ and $(\boldsymbol{\varepsilon}_i)_{i\leq n}$ are identical distributed, the independence between $(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ and $(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)$ and the independence between $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$, we obtain:

$$\begin{aligned} I_2 &= (n^2 - n)\mathrm{E}\left[\frac{\boldsymbol{\xi}_1^T(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^T - \mathbf{I})\boldsymbol{\xi}_1\boldsymbol{\xi}_2^T(\boldsymbol{\varepsilon}_2\boldsymbol{\varepsilon}_2^T - \mathbf{I})\boldsymbol{\xi}_2}{(\sum_{i=1}^n \boldsymbol{\xi}_i^T\boldsymbol{\xi}_i)^2}\right] \\ &= (n^2 - n) \sum_{j,k,l,p=1}^m \mathrm{E}(\varepsilon_{1j}\varepsilon_{1k} - \delta_{jk})\mathrm{E}(\varepsilon_{2l}\varepsilon_{2p} - \delta_{lp})\mathrm{E}\left[\frac{\xi_{1j}\xi_{1k}\xi_{2l}\xi_{2p}}{(\sum_{i=1}^n \boldsymbol{\xi}_i^T\boldsymbol{\xi}_i)^2}\right] \\ &= 0, \end{aligned} \tag{3.6}$$

since, $\mathrm{E}(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^T) = \mathbf{I}$. Note that by Cauchy-Schwarz inequality:

$$\begin{aligned} \left|\mathrm{E}\left[\frac{\xi_{1j}\xi_{1k}\xi_{2l}\xi_{2p}}{(\sum_{i=1}^n \boldsymbol{\xi}_i^T\boldsymbol{\xi}_i)^2}\right]\right| &\leq \left\{\mathrm{E}\left[\frac{\xi_{1j}}{(\sum_{i=1}^n \|\boldsymbol{\xi}_i\|^2)^{1/2}}\right]^4 \mathrm{E}\left[\frac{\xi_{1k}}{(\sum_{i=1}^n \|\boldsymbol{\xi}_i\|^2)^{1/2}}\right]^4 \cdot \right. \\ &\cdot \left. \mathrm{E}\left[\frac{\xi_{2l}}{(\sum_{i=1}^n \|\boldsymbol{\xi}_i\|^2)^{1/2}}\right]^4 \mathrm{E}\left[\frac{\xi_{2p}}{(\sum_{i=1}^n \|\boldsymbol{\xi}_i\|^2)^{1/2}}\right]^4\right\}^{1/4} \\ &\leq 1. \end{aligned}$$

From (3.4), (3.5) and (3.6), it follows that $\gamma_{n,1} \xrightarrow{P} 1$. This concludes the proof of (3.3), when $\mathrm{E}(\boldsymbol{\xi}) = \mathbf{0}$.

Now, consider the case when $\mathrm{E}(\boldsymbol{\xi}) = \boldsymbol{\mu}$. Since $\boldsymbol{\xi} \in GDAN$, it follows that $(\boldsymbol{\xi} - \boldsymbol{\mu}) \in GDAN$ and $\mathrm{E}(\boldsymbol{\xi} - \boldsymbol{\mu}) = \mathbf{0}$. Hence, by applying the first part of the proof, we obtain $(\boldsymbol{\xi} - \boldsymbol{\mu})^T\boldsymbol{\varepsilon} \in DAN$. We write:

$$\boldsymbol{\xi}^T\boldsymbol{\varepsilon} = (\boldsymbol{\xi} - \boldsymbol{\mu})^T\boldsymbol{\varepsilon} + \boldsymbol{\mu}^T\boldsymbol{\varepsilon},$$

and note that $\mathrm{Var}(\boldsymbol{\mu}^T\boldsymbol{\varepsilon}) < \infty$. We apply Lemma 5 in [12] to obtain $\boldsymbol{\xi}^T\boldsymbol{\varepsilon} \in DAN$.

(b) Assume that $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ is a positive definite covariance matrix. It follows that $\boldsymbol{\varepsilon}' := \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1/2}\boldsymbol{\varepsilon}$ is such that $\mathrm{E}(\boldsymbol{\varepsilon}') = \mathbf{0}$ and $\mathrm{E}[\boldsymbol{\varepsilon}'(\boldsymbol{\varepsilon}')^T] = \mathbf{I}$. If $\boldsymbol{\xi} \in GDAN$, there exist a sequence of matrices $\mathbf{B}_n$ such that $\mathbf{B}_n(\sum_{i=1}^n \boldsymbol{\xi}_i - n\mu) \xrightarrow{\mathcal{D}} N(0, \mathbf{I})$. It follows that $\boldsymbol{\xi}' := \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}{}^{1/2}\boldsymbol{\xi} \in GDAN$, with $\mathbf{B}_n' := \mathbf{B}_n\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}{}^{-1/2}$ the sequence of normalizing matrices. We apply part (a) of the proof and obtain $(\boldsymbol{\xi}')^T\boldsymbol{\varepsilon}' = \boldsymbol{\xi}^T\boldsymbol{\varepsilon} \in DAN$. □

Lemma 3.5. *Assume that* $\boldsymbol{\xi} \in GDAN$, $\mathrm{E}(\boldsymbol{\xi}) = 0$ *and* $\mathbf{B}_n \sum_{i=1}^n \boldsymbol{\xi}_i \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I})$, *where* $(\boldsymbol{\xi})_{1\le i\le n}$ *are i.i.d copies of* $\boldsymbol{\xi}$. *Let* $\boldsymbol{\Sigma}$ *be a positive definite matrix and denote* $\boldsymbol{\xi}' = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\xi}$. *Then* $\boldsymbol{\xi}' \in GDAN$, *and* $\mathbf{B}_n' \sum_{i=1}^n \boldsymbol{\xi}_i' \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I})$, *where* $\boldsymbol{\xi}_i' = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\xi}_i$ *and* $\mathbf{B}_n' = (\boldsymbol{\Sigma}^{-1/2}\mathbf{B}_n^T\mathbf{B}_n\boldsymbol{\Sigma}^{-1/2})^{1/2}$.

Proof. We denote by $\mathbf{C}_n := \mathbf{B}_n\boldsymbol{\Sigma}^{-1/2}$, and so $\mathbf{C}_n \sum_{i=1}^n \boldsymbol{\xi}_i' \xrightarrow{\mathcal{D}} N(0, \mathbf{I})$. We apply Lemma 2.1 of [11] to conclude that:

$$(\mathbf{C}_n^T\mathbf{C}_n)^{1/2}\sum_{i=1}^n \boldsymbol{\xi}_i' \xrightarrow{\mathcal{D}} N(0, \mathbf{I}).$$

□

The next result gives the rate of convergence for the trace of the sequence $(\mathbf{B}_n)_{n\ge1}$ of normalizing matrices. It will be used frequently in the proofs of the main results.

Lemma 3.6. *Let* $\boldsymbol{\xi} \in GDAN$, $\mathrm{E}(\boldsymbol{\xi}) = \boldsymbol{\mu}$ *and* $(\mathbf{B}_n)_{n\ge1}$ *be a sequence for which* (3.1) *holds, with* $\mathbf{A}_n = n\boldsymbol{\mu}$.

(*a*) (*i*) *If* $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$, $\frac{n}{\mathrm{trace}(\mathbf{B}_n^{-2})} = \frac{1}{\mathrm{Var}(\|\boldsymbol{\xi}-\boldsymbol{\mu}\|)}$, *for any* $n \ge 1$.

(*ii*) *If* $\mathrm{Var}(\|\boldsymbol{\xi}\|) = \infty$, $\frac{n}{\mathrm{trace}(\mathbf{B}_n^{-2})} \to 0$.

(*b*) $\frac{\mathrm{trace}(\mathbf{B}_n^{-2})}{n^2} \to 0$, *as* $n \to \infty$.

Proof. By Lemma 3.3, $\|\boldsymbol{\xi} - \boldsymbol{\mu}\| \in DAN$ and

$$\frac{1}{b_n}\left(\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\| - a_n\right) \xrightarrow{\mathcal{D}} N(0,1),$$

where $(\boldsymbol{\xi}_i)_{i\le n}$ are i.i.d copies of $\boldsymbol{\xi}$ and $a_n = n\mathrm{E}(\|\boldsymbol{\xi}_1 - \boldsymbol{\mu}\|)$.

(i) If $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$, it follows that $\sigma^2 := \mathrm{Var}(\|\boldsymbol{\xi} - \boldsymbol{\mu}\|) < \infty$. Then $b_n^2 = n\sigma^2$ and so $\frac{n}{b_n^2} = \frac{1}{\sigma^2}$ and $\frac{b_n^2}{n^2} = \frac{\sigma^2}{n} \to 0$.

(ii) If $\mathrm{Var}(\|\boldsymbol{\xi}\|) = \infty$, then $\mathrm{Var}(\|\boldsymbol{\xi} - \boldsymbol{\mu}\|) = \infty$ and $b_n^2 = nl^2(n)$, where $l(n)$ is a slowly varying function at infinity. Hence $\frac{n}{b_n^2} = \frac{1}{l^2(n)} \to 0$ and $\frac{b_n^2}{n^2} = \frac{l^2(n)}{n} \to 0$. □

4. Consistency and Central Limit Theorems

In this section we state and prove our main results: consistency and CLTs for our estimators. The consistency property is relatively straight-forward. The CLT requires more work and its proof will be devided into several steps. First, we consider the case $\alpha = 0$ and we prove a CLT in which the normalizing factors depend on β (Theorem 4.4). Next, we remove the condition $\alpha = 0$ (Theorem 4.7). Finally, we substitute the unknown parameter β by its estimator, $\widehat{\beta}_n$.

4.1. Consistency

The assumptions required for consistency are slightly weaker than the assumptions $(A1)$–$(A3)$, which are required for CLTs.

Here, we introduce the following assumption:

$$(A1') \qquad \boldsymbol{\xi}_1 \in GDAN.$$

Theorem 4.1. *Assume that $(A1')$ and $(A3)$ are satisfied. Then, the estimators $\widehat{\beta}_n$ and $\widehat{\alpha}_n$, given by (2.8) and (2.7) are weakly consistent, i.e. $\widehat{\beta}_n \xrightarrow{P} \beta$ and $\widehat{\alpha}_n \xrightarrow{P} \alpha$.*

Proof. Let $b_n^2 = \mathrm{trace}(\mathbf{B}_n^{-2})$, where $(\mathbf{B}_n)_{n\geq 1}$ is a sequence of normalizing matrices for which (3.1) holds, with $\mathbf{A}_n = n\boldsymbol{\mu}$. From (2.8), the consistency of $\widehat{\beta}_n$ follows once we prove that:

$$\frac{\sum_{i=1}^n \left[(\mathbf{x}_i - \overline{\mathbf{x}})^T(\mathbf{y}_i - \overline{\mathbf{y}}) - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}})\right]}{b_n^2} \xrightarrow{P} \beta, \tag{4.1}$$

$$\frac{\sum_{i=1}^n \left[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right]}{b_n^2} \xrightarrow{P} 1, \tag{4.2}$$

We denote $\overline{\boldsymbol{\delta}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\delta}_i$, $\overline{\boldsymbol{\varepsilon}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\varepsilon}_i$, $\overline{\boldsymbol{\xi}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\xi}_i$. Note that $\overline{\boldsymbol{\delta}}$ is an m-dimensional random vector with components $\overline{\delta}_j = \frac{1}{n}\sum_{i=1}^n \delta_{ij}$, where $j \leq m$.

Writing $\mathbf{x}_i$ and $\mathbf{y}_i$ in terms of $\boldsymbol{\xi}_i$, $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\delta}_i$ we obtain:

$$\begin{aligned}
&\frac{\sum_{i=1}^n \left[(\mathbf{x}_i - \overline{\mathbf{x}})^T(\mathbf{y}_i - \overline{\mathbf{y}}) - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}})\right]}{b_n^2} = \beta\frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}}\|^2}{b_n^2} \\
&+\beta\frac{\sum_{i=1}^n (\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}})^T(\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}})}{b_n^2} + \frac{\sum_{i=1}^n (\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}})^T(\boldsymbol{\varepsilon}_i - \overline{\boldsymbol{\varepsilon}})}{b_n^2} \\
&+\frac{\sum_{i=1}^n \left[(\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}})^T(\boldsymbol{\varepsilon}_i - \overline{\boldsymbol{\varepsilon}}) - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}})\right]}{b_n^2} \\
&:= \beta I_1 + \beta I_2 + I_3 + I_4,
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\sum_{i=1}^n \left[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right]}{b_n^2} = \frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}}\|^2}{b_n^2} \\
&+2\frac{\sum_{i=1}^n (\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}})^T(\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}})}{b_n^2} + \frac{\sum_{i=1}^n \left[\|\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right]}{b_n^2} \\
&:= I_1 + 2I_2 + I_5.
\end{aligned}$$

Note that:

$$I_1 = \frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^2}{b_n^2} - \frac{n\|\overline{\boldsymbol{\xi}} - \boldsymbol{\mu}\|^2}{b_n^2},$$

where the first term converges to 1 in probability, by Lemma 3.3 and the second term converges to 0 in probability, by the weak law of large numbers (WLLN) and Lemma 3.6. Hence, $I_1 \xrightarrow{P} 1$. Since $\frac{1}{n}\sum_{i=1}^n \|\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}}\|^2 \xrightarrow{P} \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})$, (by WLLN), using Lemma 3.6, it follows that $I_5 \xrightarrow{P} 0$. Similarly, $I_4 \xrightarrow{P} 0$.

Using the WLLN and the fact that $(\boldsymbol{\xi}_i)_{i\le n}$ and $(\boldsymbol{\delta}_i)_{i\le n}$ are independent, we obtain:

$$\frac{\sum_{i=1}^n (\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}})^T(\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}})}{n} \xrightarrow{P} 0.$$

By Lemma 3.6, it follows that $I_2 \xrightarrow{P} 0$. Similarly, $I_3 \xrightarrow{P} 0$. Hence, (4.1) and (4.2) hold and therefore $\widehat{\beta}_n \xrightarrow{P} \beta$.

By WLLN, we have $\overline{\overline{\mathbf{x}}} \xrightarrow{P} \frac{1}{m}\sum_{j=1}^m \mu_j$ and $\overline{\overline{\mathbf{y}}} \xrightarrow{P} \alpha + \beta\frac{1}{m}\sum_{j=1}^m \mu_j$. Since $\widehat{\beta}_n \xrightarrow{P} \beta$, by Slutsky's theorem, it follows that $\widehat{\alpha}_n = \overline{\overline{\mathbf{y}}} - \widehat{\beta}_n\overline{\overline{\mathbf{x}}} \xrightarrow{P} \alpha$. □

Remark 4.2. If $\text{Var}(\|\boldsymbol{\xi}\|) < \infty$, one can apply the strong law of large numbers (SLLN) to obtain that $\widehat{\alpha}_n$ and $\widehat{\beta}_n$ are strongly consistent, i.e. $\widehat{\beta}_n \xrightarrow{a.s.} \beta$ and $\widehat{\alpha} \xrightarrow{a.s.} \alpha$. (Details are omitted.)

4.2. CLTs in the case $\boldsymbol{\alpha = 0}$

In this case, the estimator of β has the following form:

$$\widehat{\beta}'_n = \frac{\sum_{i=1}^n[\mathbf{x}_i^T\mathbf{y}_i - \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta\varepsilon}})]}{\sum_{i=1}^n[\|\mathbf{x}_i\|^2 - \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}. \tag{4.3}$$

It follows that:

$$\left[\sum_{i=1}^n[\|\mathbf{x}_i\|^2 - \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]\right](\widehat{\beta}'_n - \beta) = \sum_{i=1}^n u_i := n\overline{u}, \tag{4.4}$$

where

$$u_i = \mathbf{x}_i^T\mathbf{y}_i - \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta\varepsilon}}) - \beta[\|\mathbf{x}_i\|^2 - \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})],\ 1 \le i \le n. \tag{4.5}$$

In terms of $\boldsymbol{\xi}_i$, $\boldsymbol{\varepsilon}_i$, $\boldsymbol{\delta}_i$, we obtain, for $1 \le i \le n$

$$u_i \quad = \quad \boldsymbol{\xi}_i^T\boldsymbol{\varepsilon}_i - \beta\boldsymbol{\xi}_i^T\boldsymbol{\delta}_i + \boldsymbol{\delta}_i^T\boldsymbol{\varepsilon}_i - \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}}) - \beta[\|\boldsymbol{\delta}_i\|^2 - \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})].$$

Our goal is to prove that $u_1 \in DAN$. We first present the general idea of the proof. By Theorem 3.4, the first two terms of u_1 are in DAN. The remaining terms are clearly also in DAN. The fact that $u_1 \in DAN$ follows by Lemma 5 of [12]. We write, for $1 \le i \le n$

$$u_i = \mathbf{b}^T\boldsymbol{\zeta}_i, \tag{4.6}$$

where $\mathbf{b} = (1, -\beta, 0, 0, 1, -\beta) \in \mathbb{R}^6$ and

$$\begin{aligned}\boldsymbol{\zeta}_i &= \left(\boldsymbol{\xi}_i^T \boldsymbol{\varepsilon}_i, \boldsymbol{\xi}_i^T \boldsymbol{\delta}_i, \frac{1}{m}\sum_{j=1}^m \varepsilon_{ij}, \frac{1}{m}\sum_{j=1}^m \delta_{ij}, \boldsymbol{\delta}_i^T \boldsymbol{\varepsilon}_i - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\delta}}),\right.\\ &\qquad \left.\|\boldsymbol{\delta}_i\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right)^T. \end{aligned} \tag{4.7}$$

Note that u_i are i.i.d. random variables and that $\mathrm{E}(\boldsymbol{\zeta}_i) = 0$, for any $i \le n$.

Let

$$\begin{aligned}\boldsymbol{\zeta}_i' &= \left((\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\varepsilon}_i, (\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\delta}_i, \frac{1}{m}\sum_{j=1}^m \varepsilon_{ij}, \frac{1}{m}\sum_{j=1}^m \delta_{ij}, \boldsymbol{\delta}_i^T \boldsymbol{\varepsilon}_i - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\delta}}),\right.\\ &\qquad \left.\|\boldsymbol{\delta}_i\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right)^T. \end{aligned} \tag{4.8}$$

The following result is a generalization of Lemma 6, in [12] to the case when $\boldsymbol{\xi} \in GDAN$.

Lemma 4.3. *Assume that (A1), (A2) and (A3) hold. Let* $\mathbf{b} \in \mathbb{R}^6$ *and* $\boldsymbol{\zeta}_i$ *be given by (4.7). If* $|b_1| + |b_2| = 0$*, assume that* $\mathrm{Var}(\mathbf{b}^{\mathrm{T}}\boldsymbol{\zeta}_1) > 0$*. Then,*

$$(a)\ \mathbf{b}^T\boldsymbol{\zeta}_1 \in DAN \text{ and } (b)\ \mathbf{b}^T\boldsymbol{\zeta}_1' \in DAN.$$

In particular, $u_1 \in DAN$*, where* u_1 *is defined by (4.5).*

Proof. (a). Since in this proof there is no risk of confusion, we suppress the index i of the random vectors $\boldsymbol{\zeta}_i$, $\boldsymbol{\xi}_i$, $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\delta}_i$.

We write $\mathbf{b}^T\boldsymbol{\zeta} = \boldsymbol{\xi}^T(b_1\boldsymbol{\varepsilon} + b_2\boldsymbol{\delta}) + f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta})$, where

$$\begin{aligned} f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta}) &= b_3\left(\frac{1}{m}\sum_{j=1}^m \varepsilon_j\right) + b_4\left(\frac{1}{m}\sum_{j=1}^m \delta_j\right) + b_5[\boldsymbol{\delta}^T\boldsymbol{\varepsilon} - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\delta}})]\\ &+ b_6[\|\boldsymbol{\delta}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]. \end{aligned}$$

Case 1. $|b_1| + |b_2| = 0$. By assumption (A2), $0 < \mathrm{Var}(\mathbf{b}^T\boldsymbol{\zeta}) = \mathrm{Var}(f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta})) < \infty$, since $f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta})$ is a function containing powers of maximum order 2 of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$. Hence we apply the CLT and obtain $\mathbf{b}^T\boldsymbol{\zeta} \in DAN$.

Case 2. $|b_1| + |b_2| > 0$.

I. Assume that $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$. Then $\mathrm{Var}(\mathbf{b}^T\boldsymbol{\zeta}) < \infty$ and if $\mathrm{Var}(\mathbf{b}^T\boldsymbol{\zeta}) > 0$, we can apply the CLT to reach the conclusion.

Assume, by contradiction, that $\mathrm{Var}(\mathbf{b}^T\boldsymbol{\zeta}) = 0$. Then $\mathbf{b}^T\boldsymbol{\zeta} = C$ a.s., where C is a constant. Since $\mathrm{E}(\mathbf{b}^T\boldsymbol{\zeta}) = 0$, it follows that $C = 0$, (i.e. $\mathbf{b}^T\boldsymbol{\zeta} = 0$ a.s.).

Now $0 = \mathrm{E}[\mathbf{b}^T\boldsymbol{\zeta}|\boldsymbol{\varepsilon}, \boldsymbol{\delta}] = \mathrm{E}[\boldsymbol{\xi}^T(b_1\boldsymbol{\varepsilon} + b_2\boldsymbol{\delta}) + f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta})|\boldsymbol{\varepsilon}, \boldsymbol{\delta}]$ a.s. Since $\boldsymbol{\xi}$ is independent of $(\boldsymbol{\varepsilon}, \boldsymbol{\delta})$ and the other terms and factors are $(\boldsymbol{\varepsilon}, \boldsymbol{\delta})$- measurable functions, we have $\boldsymbol{\mu}^T(b_1\boldsymbol{\varepsilon} + b_2\boldsymbol{\delta}) + f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta}) = 0$ a.s., so $(\boldsymbol{\xi} - \boldsymbol{\mu})^T(b_1\boldsymbol{\varepsilon} + b_2\boldsymbol{\delta}) = 0$ a.s.

But $\text{Var}[(\boldsymbol{\xi}-\boldsymbol{\mu})^T(b_1\boldsymbol{\varepsilon}+b_2\boldsymbol{\delta})] = \text{trace}[\boldsymbol{\Sigma}_{\boldsymbol{\xi}}\mathbf{V}_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta})] > 0$, where $\mathbf{V}_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta}) := \text{Var}(b_1\boldsymbol{\varepsilon}+b_2\boldsymbol{\delta}) = (b_1)^2\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} + b_1b_2\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}} + b_1b_2\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}}^T + (b_2)^2\boldsymbol{\Sigma}_{\boldsymbol{\delta}} > 0$ and we used assumptions $(A1)$, $(A2)$ and $(A3)$. We reached a contradiction and therefore, $\text{Var}(\mathbf{b}^T\boldsymbol{\zeta}) > 0$.

II. Assume that $\text{Var}(\|\boldsymbol{\xi}\|) = \infty$. By Theorem 3.4 we have $\boldsymbol{\xi}^T(b_1\boldsymbol{\varepsilon}+b_2\boldsymbol{\delta}) \in DAN$. Since $f_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta}) \in DAN$, we apply Lemma 5 in [12] to obtain $\mathbf{b}^T\boldsymbol{\zeta} \in DAN$, provided we checked two conditions:

(i) $\text{E}[\boldsymbol{\xi}^T(b_1\boldsymbol{\varepsilon}+b_2\boldsymbol{\delta})f_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta})] < \infty$ and

(ii) $\text{Var}(\mathbf{b}^T\boldsymbol{\zeta}) > 0$. Denote $\mathbf{V}_{\boldsymbol{\xi}} = \text{E}(\boldsymbol{\xi}\boldsymbol{\xi}^T)$. We have:

$$\begin{aligned}\text{Var}(\mathbf{b}^T\boldsymbol{\zeta}) &= \text{Var}[\boldsymbol{\xi}^T(b_1\boldsymbol{\varepsilon}+b_2\boldsymbol{\delta})] + \text{Var}[f_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta})] \\ &+ 2\text{Cov}[\boldsymbol{\xi}^T(b_1\boldsymbol{\varepsilon}+b_2\boldsymbol{\delta}), f_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta})] \\ &= \text{trace}[\mathbf{V}_{\boldsymbol{\xi}}\mathbf{V}_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta})] + \text{Var}[f_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta})] + 2\boldsymbol{\mu}^T\text{E}[(b_1\boldsymbol{\varepsilon}+b_2\boldsymbol{\delta})f_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta})] \\ &= \infty,\end{aligned}$$

since the last two terms are finite and, by Lemma A.1,

$$\text{trace}[\mathbf{V}_{\boldsymbol{\xi}}\mathbf{V}_{\mathbf{b}}(\boldsymbol{\varepsilon},\boldsymbol{\delta})] \geq \frac{\text{trace}(\mathbf{V}_{\boldsymbol{\xi}})}{\text{trace}[\mathbf{V}_{\mathbf{b}}^{-1}(\boldsymbol{\varepsilon},\boldsymbol{\delta})]} = \infty.$$

(b) Since $\boldsymbol{\xi} \in GDAN$ is equivalent to $\boldsymbol{\xi}-\boldsymbol{\mu} \in GDAN$, we apply part (a) to $\boldsymbol{\xi}' = \boldsymbol{\xi}-\boldsymbol{\mu}$ to obtain the conclusion. □

Using Lemma 4.3 and the self-normalized and Studentized versions of the classical CLT, we obtain the following CLTs for $\widehat{\beta}_n$, when $\alpha = 0$.

Theorem 4.4. *Assume that* $(A1)$, $(A2)$ *and* $(A3)$ *hold and* $\alpha = 0$. *Let* $\widehat{\beta}'_n$ *be the estimator given by* (4.3). *Then, as* $n \to \infty$, *we have:*

$$(a)\ \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n[\|\mathbf{x}_i\|^2 - \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})](\widehat{\beta}'_n-\beta)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n(u_i-\overline{u})^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

$$(b)\ \frac{\sum_{i=1}^n\left[\|\mathbf{x}_i\|^2 - \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right](\widehat{\beta}'_n-\beta)}{\sqrt{\sum_{i=1}^n u_i^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

where u_i *is given by* (4.5) *and* $\overline{u} = \frac{1}{n}\sum_{i=1}^n u_i$.

Proof. By (4.6) and Lemma 4.3, $u_1 \in DAN$. Note that $\text{E}(u_1) = 0$. Hence,

$$\frac{\sqrt{n}\overline{u}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n(u_i-\overline{u})^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

$$\frac{\sum_{i=1}^n u_i}{\sqrt{\sum_{i=1}^n u_i^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

(see e.g. [7]). The conclusion follows by (4.4). □

The next result shows that Theorem 4.4 continues to hold true if we replace β by $\widehat{\beta}'_n$ in the definition of u_i. Its proof is omitted, as it is similar to the proof of Theorem 4.12 (below).

Theorem 4.5. *Assume that (A1), (A2) and (A3) hold and $\alpha = 0$. Let $\widehat{\beta}'_n$ be the estimator given by (4.3). Then, as $n \to \infty$, we have:*

$$\frac{\sum_{i=1}^n [\|\mathbf{x}_i\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})](\widehat{\beta}'_n - \beta)}{\sqrt{\sum_{i=1}^n \tilde{u}_i(n)^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

where $\tilde{u}_i(n) = \mathbf{x}_i^T \mathbf{y}_i - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\delta}}) - \widehat{\beta}_n[\|\mathbf{x}_i\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]$.

4.3. CLTs in the case of α arbitrary

We write:

$$\sum_{i=1}^n \left[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})\right](\widehat{\beta}_n - \beta) = \sum_{i=1}^n u_i(n), \tag{4.9}$$

with

$$u_i(n) = (\mathbf{x}_i - \overline{\mathbf{x}})^T(\mathbf{y}_i - \overline{\mathbf{y}}) - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}\boldsymbol{\varepsilon}}) - \beta[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]. \tag{4.10}$$

Note that $u_i(n)$ depends on n and β.

Relation (4.10) is of crucial importance and lies at the core of our developments. The goal is to obtain CLTs for the right hand side of (4.9), which will give a CLT for $\widehat{\beta}_n$.

We express $u_i(n)$ in terms of $\boldsymbol{\xi}_i$, $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\delta}_i$, for every $1 \le i \le n$.

$$\begin{aligned} u_i(n) &= (\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}})^T(\boldsymbol{\varepsilon}_i - \overline{\boldsymbol{\varepsilon}}) - \beta(\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}})^T(\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}}) + (\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}})^T(\boldsymbol{\varepsilon}_i - \overline{\boldsymbol{\varepsilon}}) \\ &- \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\delta}}) - \beta[\|\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]. \end{aligned}$$

Note that

$$u_i(n) = \mathbf{d}^T \boldsymbol{\eta}_i(n), \tag{4.11}$$

where $\mathbf{d} = (0, 0, 1, -\beta)^T$ and

$$\begin{aligned} \boldsymbol{\eta}_i(n) &= \Bigg(\frac{1}{m}\sum_{j=1}^m y_{ij} - \alpha, \frac{1}{m}\sum_{j=1}^m x_{ij}, (\mathbf{x_i} - \overline{\mathbf{x}})^T(\mathbf{y_i} - \overline{\mathbf{y}}) - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\delta}}), \\ &\quad \|\mathbf{x_i} - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})\Bigg)^T \qquad (4.12) \\ &= \Bigg(\beta\frac{1}{m}\sum_{j=1}^m \xi_{ij} + \frac{1}{m}\sum_{j=1}^m \varepsilon_{ij}, \frac{1}{m}\sum_{j=1}^m \xi_{ij} + \frac{1}{m}\sum_{j=1}^m \delta_{ij}, \\ &\quad \beta\|\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}}\|^2 + (\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}})^T(\boldsymbol{\varepsilon}_i - \overline{\boldsymbol{\varepsilon}}) + \beta(\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}})^T(\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}}) \\ &\quad +(\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}})^T(\boldsymbol{\varepsilon}_i - \overline{\boldsymbol{\varepsilon}}) - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\delta}}), \\ &\quad \|\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}}\|^2 + \|\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}}\|^2 + 2(\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}})^T(\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}}) - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})\Bigg). \end{aligned}$$

(The first two components in the definition of $\boldsymbol{\eta}_i(n)$ are introduced artificially at this point, but they will be used in the proofs of the asymptotic properties of the estimator $\widehat{\alpha}_n$.)

The idea is to replace $\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}}$ by $\boldsymbol{\xi}_i - \boldsymbol{\mu}$ in the expression of $u_i(n)$, i.e. to obtain an expression for $u_i(n)$ which contains an inner product of the form $\mathbf{e}^T\boldsymbol{\zeta}_i'$ (for some $\mathbf{e} \in \mathbb{R}^6$), plus a remainder term. Then, we use the fact that $\mathbf{e}^T\boldsymbol{\zeta}_1' \in DAN$ (which was shown in Lemma 4.3, (b)).

For a sake of a generalization needed later in the sequel, we consider $\mathbf{d} \in \mathbb{R}^4$ such that $\beta d_1 + d_2 = 0$ and $\beta d_3 + d_4 = 0$. We have the following decomposition:

$$\begin{aligned}
\mathbf{d}^T\boldsymbol{\eta}_i(n) &= e_1[(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T\boldsymbol{\varepsilon}_i - \boldsymbol{\varepsilon}_i^T(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu}) - \overline{\boldsymbol{\varepsilon}}^T(\boldsymbol{\xi}_i - \boldsymbol{\mu}) + \overline{\boldsymbol{\varepsilon}}^T(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})] \\
&+ e_2[(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T\boldsymbol{\delta}_i - \boldsymbol{\delta}_i^T(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu}) - \overline{\boldsymbol{\delta}}^T(\boldsymbol{\xi}_i - \boldsymbol{\mu}) + \overline{\boldsymbol{\delta}}^T(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})] \\
&+ e_3\frac{1}{m}\sum_{j=1}^{m}\varepsilon_{ij} + e_4\frac{1}{m}\sum_{j=1}^{m}\delta_{ij} \\
&+ e_5[\boldsymbol{\delta}_i^T\boldsymbol{\varepsilon}_i - \boldsymbol{\delta}_i^T\overline{\boldsymbol{\varepsilon}} - \overline{\boldsymbol{\delta}}^T\boldsymbol{\varepsilon}_i + \overline{\boldsymbol{\delta}}^T\overline{\boldsymbol{\varepsilon}} - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}})] \\
&+ e_6[\|\boldsymbol{\delta}_i\|^2 - 2\boldsymbol{\delta}_i^T\overline{\boldsymbol{\delta}} + \|\overline{\boldsymbol{\delta}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})],
\end{aligned} \tag{4.13}$$

with

$$\mathbf{e} = (d_3, \beta d_3 + 2d_4, d_1, d_2, d_3, d_4)^T. \tag{4.14}$$

Therefore, for any $\mathbf{d} \in \mathbb{R}^4$ which satisfies $\beta d_1 + d_2 = 0$ and $\beta d_3 + d_4 = 0$, we write:

$$\mathbf{d}^T\boldsymbol{\eta}_i(n) = \mathbf{e}^T\boldsymbol{\zeta}_i' + R_i(n), \tag{4.15}$$

where $\boldsymbol{\zeta}_i'$ is given by (4.8) and

$$\begin{aligned}
R_i(n) &:= e_1[-\boldsymbol{\varepsilon}_i^T(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu}) - \overline{\boldsymbol{\varepsilon}}^T(\boldsymbol{\xi}_i - \boldsymbol{\mu}) + \overline{\boldsymbol{\varepsilon}}^T(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})] \\
&+ e_2[-\boldsymbol{\delta}_i^T(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu}) - \overline{\boldsymbol{\delta}}^T(\boldsymbol{\xi}_i - \boldsymbol{\mu}) + \overline{\boldsymbol{\delta}}^T(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})] \\
&+ e_5(-\boldsymbol{\delta}_i^T\overline{\boldsymbol{\varepsilon}} - \overline{\boldsymbol{\delta}}^T\boldsymbol{\varepsilon}_i + \overline{\boldsymbol{\delta}}^T\overline{\boldsymbol{\varepsilon}}) \\
&+ e_6(-2\boldsymbol{\delta}_i^T\overline{\boldsymbol{\delta}} + \|\overline{\boldsymbol{\delta}}\|^2).
\end{aligned}$$

In particular, (4.15) holds for $d = (0, 0, 1, -\beta)^T$, in which case (4.15) gives a representation of $u_i(n)$.

The next result is a self-normalized (and Studentized) CLT for the sequence $\{\mathbf{d}^T\eta_i(n)\}_{i\le n}$, (and in particular, for $\{u_i(n)\}_{i\le n}$).

Lemma 4.6. *Assume that (A1), (A2) and (A3) hold. Let $\mathbf{d} \in \mathbb{R}^4$ which satisfies $\beta d_1 + d_2 = 0$, $\beta d_3 + d_4 = 0$ and $\mathbf{e}$ be given by (4.14). If $|e_1| + |e_2| = 0$, assume that $\mathrm{Var}(\mathbf{e}^{\mathrm{T}}\boldsymbol{\zeta}_{\mathrm{i}}') > 0$, where $\boldsymbol{\zeta}_i'$ is given by (4.8). Then, as $n \to \infty$:*

$$(a)\ \frac{\sqrt{n}\ \overline{\mathbf{d}^T\boldsymbol{\eta}(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}[\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

$$(b)\frac{n\overline{\mathbf{d}^T\boldsymbol{\eta}(n)}}{\sqrt{\sum_{i=1}^{n}(\mathbf{d}^T\boldsymbol{\eta}_i(n))^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

with $\eta_i(n)$ *given by* (4.12), $\overline{\mathbf{d}^T\boldsymbol{\eta}(n)} = \frac{1}{n}\sum_{i=1}^n \mathbf{d}^T\boldsymbol{\eta}_i(n)$ *and* $\overline{\boldsymbol{\eta}(n)} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\eta}_i(n)$.

Proof. By (4.15), it follows that $\overline{\mathbf{d}^T\boldsymbol{\eta}(n)} = \overline{\mathbf{e}^T\boldsymbol{\zeta}'} + \overline{R(n)}$, where

$$\overline{\mathbf{e}^T\boldsymbol{\zeta}'} = \frac{1}{n}\sum_{i=1}^n \mathbf{e}^T\boldsymbol{\zeta}_i', \ \ \overline{R(n)} = \frac{1}{n}\sum_{i=1}^n R_i(n),$$

and therefore:

$$\frac{\sqrt{n}\ \overline{\mathbf{d}^T\boldsymbol{\eta}(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}} = I_1 + I_2,$$

where:

$$I_1 := \frac{\sqrt{n}\ \overline{\mathbf{e}^T\boldsymbol{\zeta}'}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}},$$
$$I_2 := \frac{\sqrt{n}\ \overline{R(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}}.$$

We write:

$$I_1 = \frac{\sqrt{n}\ \overline{\mathbf{e}^T\boldsymbol{\zeta}'}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2}} \cdot \sqrt{\frac{\sum_{i=1}^n[\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2}{\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}},$$
$$I_2 = \frac{\sqrt{n}\ \overline{R(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2}} \cdot \sqrt{\frac{\sum_{i=1}^n[\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2}{\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}},$$

where $\overline{\boldsymbol{\zeta}'} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\zeta}_i'$. We will prove that, as $n \to \infty$:

$$\frac{\sqrt{n}\ \overline{R(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2}} \xrightarrow{P} 0, \tag{4.16}$$
$$\frac{\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}{\sum_{i=1}^n[\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2} \xrightarrow{P} 1. \tag{4.17}$$

By Lemma 4.3, (b) with $\mathbf{b} = \mathbf{e}$, $\mathbf{e}^T\boldsymbol{\zeta}_1' \in DAN$ and so:

$$\frac{\sqrt{n}\ \overline{\mathbf{e}^T\boldsymbol{\zeta}'}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2}} \xrightarrow{\mathcal{D}} N(0,1).$$

This result and (4.17) imply $I_1 \xrightarrow{\mathcal{D}} N(0,1)$, by Slutsky's Theorem. Furthermore, from (4.16) and (4.17), we obtain $I_2 \xrightarrow{P} 0$. Therefore, the convergence in (a) would follow by applying Slutsky's Theorem. The proofs of (4.16) and (4.17) are given in Appendix C.

Part (a) can be used to prove part (b). We have:

$$\frac{n\overline{\mathbf{d}^T\boldsymbol{\eta}(n)}}{\sqrt{\sum_{i=1}^n(\mathbf{d}^T\boldsymbol{\eta}_i(n))^2}} = \frac{\sqrt{n}\ \overline{\mathbf{d}^T\boldsymbol{\eta}(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n)-\overline{\boldsymbol{\eta}(n)})]^2}}\cdot\sqrt{\frac{n}{n-1}} \cdot \sqrt{\frac{\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n)-\overline{\boldsymbol{\eta}(n)})]^2}{\sum_{i=1}^n(\mathbf{d}^T\boldsymbol{\eta}_i(n))^2}}.$$

By (a), the first factor converges in distribution to $N(0,1)$. Hence, by Slutsky's Theorem, to obtain the convergence in (b), it suffices to prove:

$$\frac{\sum_{i=1}^n(\mathbf{d}^T\boldsymbol{\eta}_i(n))^2}{\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n)-\overline{\boldsymbol{\eta}(n)})]^2} \xrightarrow{P} 1. \tag{4.18}$$

Using the result in (a), we obtain:

$$\begin{aligned}\frac{n(\overline{\mathbf{d}^T\boldsymbol{\eta}(n)})^2}{\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n)-\overline{\boldsymbol{\eta}(n)})]^2} &= \frac{1}{n-1}\left\{\frac{\sqrt{n}\ \overline{\mathbf{d}^T\boldsymbol{\eta}(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{d}^T(\boldsymbol{\eta}_i(n)-\overline{\boldsymbol{\eta}(n)})]^2}}\right\}^2\\ &= \frac{O_P(1)}{n-1} = o_p(1).\end{aligned}$$

Relation (4.18) follows by applying Lemma A.2, with $s_i = \mathbf{d}^T\boldsymbol{\eta}_i(n)$, $t_i = \mathbf{d}^T[\boldsymbol{\eta}_i(n)-\overline{\boldsymbol{\eta}(n)}]$. □

The next theorem gives CLTs for $\widehat{\beta}_n$, as a direct consequence of Lemma 4.6. Note that the normalizing factors in the CLTs depend on the parameter β. In Theorem 4.12 we show that the result is still valid with β replaced by $\widehat{\beta}_n$, for large values of n.

Theorem 4.7. *Assume that (A1), (A2) and (A3) hold. Let $\widehat{\beta}_n$ be the estimator given by (2.8). Then, as $n\to\infty$, we have:*

$$(a)\quad \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n\left[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\mathrm{trace}(\boldsymbol{\Sigma_\delta})\right](\widehat{\beta}_n-\beta)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[u_i(n)-\overline{u(n)}]^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

$$(b)\quad \frac{\sum_{i=1}^n\left[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\mathrm{trace}(\boldsymbol{\Sigma_\delta})\right](\widehat{\beta}_n-\beta)}{\sqrt{\sum_{i=1}^n u_i(n)^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

where $u_i(n)$ is given by (4.10) and $\overline{u}(n)=\frac{1}{n}\sum_{i=1}^n u_i(n)$.

Proof. Due to (4.11) and (4.9), the results follow by applying Lemma 4.6, with $\mathbf{e}^T=(1,-\beta,0,0,1,-\beta)$. □

To examine the asymptotic behavior of $\widehat{\alpha}_n$ we need three auxiliary results. We placed their proofs in Appendix to preserve the flow of the reading process.

The first result gives the convergence rate of $\frac{1}{n}\sum_{i=1}^n \mathbf{b}^T\boldsymbol{\zeta}_i$. The proof is given in Appendix B.1.

Lemma 4.8. *Under the same hypotheses as in Lemma 4.3, as $n \to \infty$:*

$$\frac{1}{n}\sum_{i=1}^n \mathbf{b}^T\boldsymbol{\zeta}_i = \begin{cases} \frac{1}{\sqrt{n}}O_P(1), & \text{if } \operatorname{Var}(\|\boldsymbol{\xi}\|) < \infty, \text{ or } b_1 = b_2 = 0 \\ \sqrt{\frac{c_n^2}{n^2}}O_P(1), & \text{if } \operatorname{Var}(\|\boldsymbol{\xi}\|) = \infty, \text{ and } |b_1| + |b_2| > 0, \end{cases}$$

where $c_n^2 = \operatorname{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}\mathbf{B}_n^{-2})$, with $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}} = \mathrm{E}(\gamma_i\gamma_i^T)$, $\gamma_i = b_1\boldsymbol{\varepsilon}_i + b_2\boldsymbol{\delta}_i$, for $i \le n$ and $(\mathbf{B}_n)_{n\ge1}$ is a sequence of symmetric matrices such that $\mathbf{B}_n\sum_{i=1}^n(\boldsymbol{\xi}_i - \boldsymbol{\mu}) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I})$. The sequence $(c_n)_{n\ge1}$ satisfies $\frac{c_n^2}{n^2} \to 0$.

The next result gives the convergence rate of $\overline{\mathbf{d}^T\boldsymbol{\eta}(n)} = \frac{1}{n}\sum_{i=1}^n \mathbf{d}^T\boldsymbol{\eta}_i(n)$. Its proof is given in Appendix B.2 and essentially follows from Lemma 4.8, with $\mathbf{b} = \mathbf{e}$, and (4.17).

Lemma 4.9. *Under the same hypotheses as in Lemma 4.6, as $n \to \infty$:*

$$\overline{\mathbf{d}^T\boldsymbol{\eta}(n)} = \begin{cases} \frac{1}{\sqrt{n}}O_P(1), & \text{if } \operatorname{Var}(\|\boldsymbol{\xi}\|) < \infty \text{ or } e_1 = e_2 = 0 \\ \sqrt{\frac{c_n^2}{n^2}}O_P(1), & \text{if } \operatorname{Var}(\|\boldsymbol{\xi}\|) = \infty \text{ and } |e_1| + |e_2| > 0, \end{cases}$$

where $c_n^2 = \operatorname{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}\mathbf{B}_n^{-2})$, with $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}} = \mathrm{E}(\gamma_i\gamma_i^T)$, $\boldsymbol{\gamma}_i = e_1\boldsymbol{\varepsilon}_i + e_2\boldsymbol{\delta}_i$, for $i \le n$ and $(\mathbf{B}_n)_{n\ge1}$ is a sequence of symmetric matrices such that $\mathbf{B}_n\sum_{i=1}^n(\boldsymbol{\xi}_i - \boldsymbol{\mu}) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I})$.

The next result gives the asymptotic convergence rate of the estimator $\widehat{\beta}_n$ to β. Its proof is presented in Appendix B.3.

Lemma 4.10. *Assume that (A1), (A2) and (A3) hold. Let $\widehat{\beta}_n$ be the estimator given by (2.8). Then, as $n \to \infty$:*

$$(a)\ \frac{b_n^2}{\sqrt{n}}(\widehat{\beta}_n - \beta) \xrightarrow{\mathcal{D}} N(0, \lambda^2), \text{ if } \operatorname{Var}(\|\boldsymbol{\xi}\|) < \infty,$$

$$(b)\ \frac{b_n^2}{\sqrt{c_n^2}}(\widehat{\beta}_n - \beta) \xrightarrow{\mathcal{D}} N(0, 1), \text{ if } \operatorname{Var}(\|\boldsymbol{\xi}\|) = \infty,$$

where $b_n^2 = \operatorname{trace}(\mathbf{B}_n^{-2})$, $(\mathbf{B}_n)_{n\ge1}$ is a sequence of $m \times m$ non-stochastic matrices such that $\mathbf{B}_n\sum_{i=1}^n(\boldsymbol{\xi}_i - \boldsymbol{\mu}) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I})$, $(c_n)_{n\ge1}$ is the sequence of constants defined as in Lemma 4.8, with $\boldsymbol{\gamma}_i = \boldsymbol{\varepsilon}_i - \beta\boldsymbol{\delta}_i$, for $1 \le i \le n$ and $\lambda > 0$ is a constant.

The next theorem gives a CLT for $\widehat{\alpha}_n$.

Theorem 4.11. *Assume that* (A1), (A2) *and* (A3) *hold. Let* $\widehat{\alpha}_n$ *be the estimator given by* (2.7). *Then, as* $n \to \infty$, *we have:*

$$\frac{\sqrt{n}(\widehat{\alpha}_n - \alpha)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n (v_i(n) - \overline{v(n)})^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

where $v_i(n) = \frac{1}{m}\sum_{j=1}^m y_{ij} - \alpha - \frac{\beta}{m}\sum_{j=1}^m x_{ij} - \frac{\overline{\overline{\mathbf{x}}}\; u_i(n)}{\frac{1}{n}\sum_{i=1}^n \|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})}$, $u_i(n)$ *is given by* (4.10) *and* $\overline{v(n)} = \frac{1}{n}\sum_{i=1}^n v_i(n)$.

Proof. We divide the proof into two cases and use the notation $\overline{\mu} = \frac{1}{m}\sum_{j=1}^m \mu_j$.

Case 1. Assume that $\text{Var}(\|\boldsymbol{\xi}\|) < \infty$. We use (4.9) in the following decomposition for $\widehat{\alpha}_n - \alpha$:

$$\begin{aligned}
\widehat{\alpha}_n - \alpha &= \overline{\overline{\mathbf{y}}} - \alpha - \widehat{\beta}_n \overline{\overline{\mathbf{x}}} \\
&= \overline{\overline{\mathbf{y}}} - \alpha - \overline{\overline{\mathbf{x}}}\frac{\sum_{i=1}^n [u_i(n) + \beta(\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}}))]}{\sum_{i=1}^n [\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]} \\
&= \overline{\overline{\mathbf{y}}} - \alpha - \beta\overline{\overline{\mathbf{x}}} - \overline{u(n)}\frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^n [\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]}.
\end{aligned}$$

We introduce $v_i'(n) = \frac{1}{m}\sum_{j=1}^m y_{ij} - \alpha - \beta\frac{1}{m}\sum_{j=1}^m x_{ij} - \frac{\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})}u_i(n)$ and note that $\overline{v'(n)} = \frac{1}{n}\sum_{i=1}^n v_i'(n) = \overline{\overline{\mathbf{y}}} - \alpha - \beta\overline{\overline{\mathbf{x}}} - \overline{u(n)}\frac{\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})}$. Hence:

$$\widehat{\alpha}_n - \alpha = \overline{v'(n)} + \overline{u(n)}\left(\frac{\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})} - \frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^n [\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]}\right).$$

By (4.11) and Lemma 4.9, with $\mathbf{d} = (0,0,1,-\beta)^T$, $\sqrt{n}\overline{u(n)} = O_P(1)$.

By (4.2), $\frac{\sum_{i=1}^n [\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]}{b_n^2} \xrightarrow{P} 1$, where $b_n^2 = n\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})$ and so

$$\frac{1}{n}\sum_{i=1}^n \|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}}) \xrightarrow{P} \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}}). \tag{4.19}$$

Using the WLLN, $\overline{\overline{\mathbf{x}}} \xrightarrow{P} \overline{\mu}$ and it follows that

$$\sqrt{n}\overline{u(n)}\left(\frac{\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})} - \frac{\overline{\overline{\mathbf{x}}}}{\frac{1}{n}\sum_{i=1}^n \|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \text{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})}\right) = o_P(1).$$

Therefore, we have

$$\sqrt{n}(\widehat{\alpha}_n - \alpha) = \sqrt{n}\;\overline{v'(n)} + o_P(1). \tag{4.20}$$

Note that $v_i'(n) = \mathbf{d}^T\boldsymbol{\eta}_i(n)$, where $\mathbf{d} = \left(1, -\beta, -\frac{\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})}, \frac{\beta\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})}\right)^T$, and so using (B.6), it follows that

$$\frac{1}{n}\sum_{i=1}^n [v_i'(n) - \overline{v_i'(n)}]^2 \xrightarrow{P} \text{Var}(\mathbf{e}^T\boldsymbol{\zeta}_1') > 0, \tag{4.21}$$

where $\mathbf{e} = \left(-\frac{\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})}, \frac{\beta\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})}, 1, -\beta, -\frac{\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})}, \frac{\beta\overline{\mu}}{\text{trace}(\mathbf{\Sigma}_{\boldsymbol{\xi}})}\right)^T$.
Therefore, we apply Slutsky's Theorem, to obtain

$$\frac{\sqrt{n}(\widehat{\alpha}_n - \alpha)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n [v_i'(n) - \overline{v'(n)}]^2}} \xrightarrow{\mathcal{D}} N(0,1). \tag{4.22}$$

Slutsky's Theorem applied again completes the proof once we proved that:

$$\frac{\sum_{i=1}^n [v_i(n) - \overline{v(n)}]^2}{\sum_{i=1}^n [v_i'(n) - \overline{v'(n)}]^2} \xrightarrow{P} 1. \tag{4.23}$$

This convergence is proved in Appendix D.

Case 2. Assume that $\text{Var}(\|\boldsymbol{\xi}\|) = \infty$. In this case we use a different decomposition of $\widehat{\alpha}_n - \alpha$. We write:

$$\widehat{\alpha}_n - \alpha \quad = \quad \overline{\overline{\mathbf{y}}} - \alpha - \widehat{\beta}_n\overline{\overline{\mathbf{x}}} = \overline{\overline{\mathbf{y}}} - \alpha - \beta\overline{\overline{\mathbf{x}}} - (\widehat{\beta}_n - \beta)\overline{\overline{\mathbf{x}}} = \overline{v''} - (\widehat{\beta}_n - \beta)\overline{\overline{\mathbf{x}}},$$

where $\overline{v''} = \frac{1}{n}\sum_{i=1}^n v_i''$ and $v_i'' = \frac{1}{m}\sum_{j=1}^m y_{ij} - \alpha - \beta\frac{1}{m}\sum_{j=1}^m x_{ij} = \frac{1}{m}\sum_{j=1}^m (\varepsilon_{ij} - \beta\delta_{ij})$.
Note that $v_i'' = \mathbf{d'}^T\boldsymbol{\eta}_i(n)$, with $\mathbf{d'} = (1, -\beta, 0, 0)^T$.

We have:

$$\begin{aligned}\frac{\sqrt{n}(\widehat{\alpha}_n - \alpha)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n [v_i(n) - \overline{v(n)}]^2}} \quad &= \quad \frac{\sqrt{n}\;\overline{v''} - \sqrt{n}\;\overline{\overline{\mathbf{x}}}(\widehat{\beta}_n - \beta)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n [v_i'' - \overline{v''}]^2}} \\ &\cdot \quad \sqrt{\frac{\sum_{i=1}^n [v_i'' - \overline{v''}]^2}{\sum_{i=1}^n [v_i(n) - \overline{v(n)}]^2}}. \end{aligned} \tag{4.24}$$

We use Lemma 4.10, Lemma A.1 and Lemma 3.6, (a) to obtain $\sqrt{n}(\widehat{\beta}_n - \beta) = o_P(1)$. By WLLN we have: $\overline{\overline{\mathbf{x}}} \xrightarrow{P} \overline{\mu}$ and:

$$\frac{1}{n-1}\sum_{i=1}^n [v_i'' - \overline{v''}]^2 \xrightarrow{P} \text{Var}(v_1'') > 0, \tag{4.25}$$

where the inequality follows from assumption $(A2)$.

Hence, we obtain $\frac{\sqrt{n}\;\overline{\overline{\mathbf{x}}}(\widehat{\beta}_n - \beta)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n [v_i'' - \overline{v''}]^2}} = o_P(1)$, which by applying Lemma 4.6, with $\mathbf{d'} = (1, -\beta, 0, 0)^T$ and Slutsky's Theorem implies that the first factor in (4.24) converges in distribution to $N(0,1)$. To finish the proof, it remains to show:

$$\frac{\sum_{i=1}^n [v_i(n) - \overline{v(n)}]^2}{\sum_{i=1}^n [v_i'' - \overline{v''}]^2} \xrightarrow{P} 1, \tag{4.26}$$

which is done in Appendix D. □

The normalizers of the CLTs in Theorem 4.7 and Theorem 4.11 depend on β. The next result proves that, under the same assumptions, we can substitute the unknown parameter β by its estimator, $\widehat{\beta}_n$, for large values of n. The idea is to show that the normalizer $\sum_{i=1}^n u_i(n)^2$ has the same asymptotic behavior as $\sum_{i=1}^n \tilde{u}_i(n)^2$, where $\tilde{u}_i(n)$ has the same expression as $u_i(n)$, with β replaced by $\widehat{\beta}_n$.

Theorem 4.12. *Assume that* (A1), (A2) *and* (A3) *hold. Then, as* $n \to \infty$, *we have:*

$$(a) \qquad \frac{\sum_{i=1}^n [\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})](\widehat{\beta}_n - \beta)}{\sqrt{\sum_{i=1}^n \tilde{u}_i(n)^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

$$(b) \qquad \frac{\sqrt{n}(\widehat{\alpha}_n - \alpha)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n (\tilde{v}_i(n) - \overline{\tilde{v}(n)})^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

where $\tilde{u}_i(n) = (\mathbf{x}_i - \overline{\mathbf{x}})^T(\mathbf{y}_i - \overline{\mathbf{y}}) - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\varepsilon\delta}}) - \widehat{\beta}_n[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]$, $\widehat{\beta}_n$ *is given by* (2.8), $\widehat{\alpha}_n$ *is given by* (2.7),

$$\tilde{v}_i(n) = \frac{1}{m}\sum_{j=1}^m y_{ij} - \alpha - \widehat{\beta}_n \frac{1}{m}\sum_{j=1}^m x_{ij} - \frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^n [\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]}\tilde{u}_i(n),$$

and $\overline{\tilde{v}(n)} = \frac{1}{n}\sum_{i=1}^n \tilde{v}_i(n)$.

Proof. (a) Note that the numerator of the ratio is $\sum_{i=1}^n u_i(n)$, by (4.9). We write:

$$\frac{\sum_{i=1}^n u_i(n)}{\sqrt{\sum_{i=1}^n \tilde{u}_i(n)^2}} = \frac{\sum_{i=1}^n u_i(n)}{\sqrt{\sum_{i=1}^n u_i(n)^2}} \cdot \sqrt{\frac{\sum_{i=1}^n u_i(n)^2}{\sum_{i=1}^n \tilde{u}_i(n)^2}}. \tag{4.27}$$

The first factor in (4.27) converges in distribution to $N(0,1)$, by Theorem 4.7, (b). The proof will be complete once we show that:

$$\frac{\sum_{i=1}^n \tilde{u}_i(n)^2}{\sum_{i=1}^n u_i(n)^2} \xrightarrow{P} 1. \tag{4.28}$$

By Lemma A.2, it is enough to prove:

$$\frac{\sum_{i=1}^n [\tilde{u}_i(n) - u_i(n)]^2}{\sum_{i=1}^n u_i(n)^2} \xrightarrow{P} 0. \tag{4.29}$$

Using (4.9) and the definitions of $u_i(n)$ and $\tilde{u}_i(n)$, we have

$$\sum_{i=1}^n [\tilde{u}_i(n) - u_i(n)]^2 = \left[\sum_{i=1}^n u_i(n)\right]^2 \frac{\sum_{i=1}^n [\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]^2}{\left\{\sum_{i=1}^n [\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]\right\}^2}.$$

Since, $\frac{\sum_{i=1}^n u_i(n)}{\sqrt{\sum_{i=1}^n u_i(n)^2}} \xrightarrow{\mathcal{D}} N(0,1)$, by Lemma 4.6 and (4.11), to prove (4.29), it suffices to show:

$$\frac{\sum_{i=1}^n \left[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})\right]^2}{\left\{\sum_{i=1}^n \left[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})\right]\right\}^2} \xrightarrow{P} 0. \tag{4.30}$$

The convergence in (4.30) will follow from (4.2) and:

$$\frac{\sum_{i=1}^n \left[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right]^2}{(b_n^2)^2} \xrightarrow{P} 0, \tag{4.31}$$

where $b_n^2 = \mathrm{trace}(\mathbf{B}_n^{-2})$ and $(\mathbf{B}_n)_{n\geq 1}$ is a sequence of matrices such that

$$\mathbf{B}_n \sum_{i=1}^n (\boldsymbol{\xi}_i - \boldsymbol{\mu}) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I}).$$

We write

$$\begin{aligned}\frac{\sum_{i=1}^n \left[\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right]^2}{(b_n^2)^2} &= \frac{\sum_{i=1}^n \|\mathbf{x}_i - \overline{\mathbf{x}}\|^4}{(b_n^2)^2} + n\frac{[\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]^2}{(b_n^2)^2} \\ &- 2\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\frac{\sum_{i=1}^n \|\mathbf{x}_i - \overline{\mathbf{x}}\|^2}{(b_n^2)^2} \\ &:= J_1 + J_2 + J_3.\end{aligned}$$

We use (4.2) to obtain $J_3 \xrightarrow{P} 0$ and Lemma 3.6, (a) to obtain $J_2 \to 0$. It remains to show that $J_1 \xrightarrow{P} 0$. We have:

$$\begin{aligned}\frac{\sum_{i=1}^n \|\mathbf{x}_i - \overline{\mathbf{x}}\|^4}{(b_n^2)^2} &\leq 4\frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \overline{\boldsymbol{\xi}}\|^4}{(b_n^2)^2} + 4\frac{\sum_{i=1}^n \|\boldsymbol{\delta}_i - \overline{\boldsymbol{\delta}}\|^4}{(b_n^2)^2} \\ &\leq 16\left(\frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^4}{(b_n^2)^2} + \frac{\sum_{i=1}^n \|\boldsymbol{\delta}_i\|^4}{(b_n^2)^2}\right. \\ &+ \left. n\frac{\|\overline{\boldsymbol{\xi}} - \boldsymbol{\mu}\|^4}{(b_n^2)^2} + n\frac{\|\overline{\boldsymbol{\delta}}\|^4}{(b_n^2)^2}\right) \\ &:= 16(T_1 + T_2 + T_3 + T_4).\end{aligned}$$

We prove that $T_i \xrightarrow{P} 0$, for all $i = 1, 2, 3, 4$. We first treat T_1 and write:

$$T_1 = \frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^4}{(b_n^2)^2} = \frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^4}{\left(\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^2\right)^2} \cdot \left(\frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^2}{b_n^2}\right)^2,$$

where the last factor converges to 1, in probability, by Lemma 3.3 and therefore is bounded, in probability. Hence, $T_1 \xrightarrow{P} 0$, if we prove that

$$\frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^4}{\left(\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^2\right)^2} \xrightarrow{P} 0.$$

Let $\epsilon > 0$ be arbitrary. By Markov's inequality and the fact that $(\boldsymbol{\xi}_i)_{1\leq i\leq n}$ are i.i.d. random vectors:

$$\begin{aligned}P\left(\left|\frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^4}{\left(\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^2\right)^2}\right| > \epsilon\right) &\leq \epsilon^{-1} n \mathrm{E}\left[\frac{\|\boldsymbol{\xi}_1 - \boldsymbol{\mu}\|^4}{\left(\sum_{i=1}^n \|\boldsymbol{\xi}_i - \boldsymbol{\mu}\|^2\right)^2}\right] \\ &= \epsilon^{-1} n \mathrm{E}\left[\left(\frac{\mathbf{Z}_1^T \mathbf{Z}_1}{\sum_{i=1}^n \mathbf{Z}_i^T \mathbf{Z}_i}\right)^2\right],\end{aligned} \tag{4.32}$$

where we have denoted $\mathbf{Z}_i = \boldsymbol{\xi}_i - \boldsymbol{\mu}$. We have the following inequalities:

$$\begin{aligned}\frac{\mathbf{Z}_1^T\mathbf{Z}_1}{\sum_{i=1}^n \mathbf{Z}_i^T\mathbf{Z}_i} &\leq \frac{1}{\sum_{i=1}^n \lambda_{\max}(\mathbf{Z}_i\mathbf{Z}_i^T)}\mathbf{Z}_1^T\mathbf{Z}_1 \leq \lambda_{\min}\left[\left(\sum_{i=1}^n \mathbf{Z}_i\mathbf{Z}_i^T\right)^{-1}\right]\mathbf{Z}_1^T\mathbf{Z}_1 \\ &\leq \mathbf{Z}_1^T\left(\sum_{i=1}^n \mathbf{Z}_i\mathbf{Z}_i^T\right)^{-1}\mathbf{Z}_1. \end{aligned} \tag{4.33}$$

From the proof of Theorem 3.4 of [6], since $\mathbf{Z}_1 \in GDAN$, it follows that

$$n\mathrm{E}\left\{\left[\mathbf{Z}_1^T\left(\sum_{i=1}^n \mathbf{Z}_i\mathbf{Z}_i^T\right)^{-1}\mathbf{Z}_1\right]^2\right\} \to 0,$$

and hence by (4.33) and (4.32), we obtain $\frac{\sum_{i=1}^n \|\boldsymbol{\xi}_i-\boldsymbol{\mu}\|^4}{\left(\sum_{i=1}^n \|\boldsymbol{\xi}_i-\boldsymbol{\mu}\|^2\right)^2} = o_P(1)$.

Similarly, $\frac{\sum_{i=1}^n \|\boldsymbol{\delta}_i\|^4}{\left(\sum_{i=1}^n \|\boldsymbol{\delta}_i\|^2\right)^2} = o_P(1)$, since $\boldsymbol{\delta}_1 \in GDAN$. We write:

$$T_2 = \frac{\sum_{i=1}^n \|\boldsymbol{\delta}_i\|^4}{(b_n^2)^2} = o_P(1)\cdot\left[\frac{\sum_{i=1}^n \|\boldsymbol{\delta}_i\|^2}{n\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})}\right]^2\cdot\left[\frac{n\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})}{b_n^2}\right]^2.$$

By WLLN, $\frac{\sum_{i=1}^n \|\boldsymbol{\delta}_i\|^2}{n\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})} \to 1$, in probability. In addition, $\frac{n\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})}{b_n^2}$ is bounded, by Lemma 3.6, (a) and hence, $T_2 \to 0$, in probability.

Also, $T_i \to 0$, in probability, where $i = 3, 4$, by WLLN and Lemma 3.6, (a). This finishes the proof of $J_1 \xrightarrow{P} 0$.

For the proof of (b), we use Theorem 4.11 and Slutsky's Theorem. Hence, it suffices to show:

$$\frac{\sum_{i=1}^n [\tilde{v}_i(n) - \overline{\tilde{v}(n)}]^2}{\sum_{i=1}^n [v_i(n) - \overline{v(n)}]^2} \xrightarrow{P} 1. \tag{4.34}$$

By Lemma A.2, it suffices to prove:

$$\frac{\sum_{i=1}^n \{\tilde{v}_i(n) - v_i(n) - [\overline{\tilde{v}(n)} - \overline{v(n)}]\}^2}{\sum_{i=1}^n [v_i(n) - \overline{v(n)}]^2} \xrightarrow{P} 0. \tag{4.35}$$

We have

$$\begin{aligned}&\sum_{i=1}^n \{\tilde{v}_i(n) - v_i(n) - [\overline{\tilde{v}(n)} - \overline{v(n)}]\}^2 \\ &= \sum_{i=1}^n \left\{(\beta - \widehat{\beta}_n)\left(\frac{1}{m}\sum_{j=1}^m x_{ij} - \overline{\overline{\mathbf{x}}}\right) - \frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^n [\|\mathbf{x}_i - \overline{\mathbf{x}}\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}\right. \\ &\left.\cdot\{\tilde{u}_i(n) - u_i(n) - [\overline{\tilde{u}(n)} - \overline{u(n)}]\}\right\}^2\end{aligned}$$

$$
\begin{aligned}
&\leq 2(\beta-\widehat{\beta}_n)^2\sum_{i=1}^n\left(\frac{1}{m}\sum_{j=1}^m x_{ij}-\overline{\overline{\mathbf{x}}}\right)^2\\
+\quad & 2\left(\frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}\right)^2\sum_{i=1}^n\{\tilde{u}_i(n)-u_i(n)-[\overline{\tilde{u}(n)}-\overline{u(n)}]\}^2\\
&=2I_1+2I_2\cdot I_3,
\end{aligned}
$$

where we have denoted by:

$$
\begin{aligned}
I_1 &:= (\widehat{\beta}_n-\beta)^2\sum_{i=1}^n\left(\frac{1}{m}\sum_{j=1}^m x_{ij}-\overline{\overline{\mathbf{x}}}\right)^2,\\
I_2 &:= \left(\frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}\right)^2,\\
I_3 &:= \sum_{i=1}^n\{\tilde{u}_i(n)-u_i(n)-[\overline{\tilde{u}(n)}-\overline{u(n)}]\}^2.
\end{aligned}
$$

By (D.3),

$$
I_2=\frac{n^2}{(b_n^2)^2}O_P(1). \tag{4.36}
$$

We have the following inequality:

$$
I_1\leq 2(\widehat{\beta}_n-\beta)^2\left[\sum_{i=1}^n\left(\frac{1}{m}\sum_{j=1}^m\xi_{ij}-\overline{\overline{\boldsymbol{\xi}}}\right)^2+\sum_{i=1}^n\left(\frac{1}{m}\sum_{j=1}^m\delta_{ij}-\overline{\overline{\boldsymbol{\delta}}}\right)^2\right],
$$

where $\overline{\overline{\boldsymbol{\xi}}}=\frac{1}{nm}\sum_{i=1}^n\sum_{j=1}^m\xi_{ij}$ and $\overline{\overline{\boldsymbol{\delta}}}=\frac{1}{nm}\sum_{i=1}^n\sum_{j=1}^m\delta_{ij}$.

Since $\tilde{u}_i(n)-u_i(n)=(\beta-\widehat{\beta}_n)[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]$, it follows that:

$$
\begin{aligned}
I_3 &\leq 2(\widehat{\beta}_n-\beta)^2\left\{\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]^2\right.\\
&+\left. n\left[\frac{1}{n}\sum_{i=1}^n\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})\right]^2\right\}\\
&= 2(\widehat{\beta}_n-\beta)^2\left\{\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]^2\right.\\
&+\left.\frac{1}{n}\left[\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]\right]^2\right\},\\
&= (\widehat{\beta}_n-\beta)^2\left[(b_n^2)^2o_P(1)+\frac{1}{n}(b_n^2)^2O_P(1)\right]\\
&= 2(\widehat{\beta}_n-\beta)^2(b_n^2)^2o_P(1),
\end{aligned} \tag{4.37}
$$

using (4.31) and (4.2).

From (4.36) and (4.37) we get $I_2 \cdot I_3 = n^2(\widehat{\beta}_n - \beta)^2 o_P(1)$.

Since the asymptotic behavior of $\widehat{\beta}_n - \beta$ is different in the cases $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$ and $\mathrm{Var}(\|\boldsymbol{\xi}\|) = \infty$, we have to consider these cases separately.

Case 1. Assume that $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$.

We apply WLLN and so:

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\xi_{ij} - \overline{\overline{\boldsymbol{\xi}}}\right)^2 \xrightarrow{P} \frac{1}{m^2}\sum_{j,k=1}^{m}\sigma_{\boldsymbol{\xi},jk},$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\delta_{ij} - \overline{\overline{\boldsymbol{\delta}}}\right)^2 \xrightarrow{P} \frac{1}{m^2}\sum_{j,k=1}^{m}\sigma_{\boldsymbol{\delta},jk}.$$

By Lemma 4.10, (a) and Lemma 3.6, (a), we obtain $\frac{1}{n}I_1 \le \frac{n}{(b_n^2)^2}O_P(1) = o_P(1)$ and $\frac{1}{n}I_2 \cdot I_3 = \frac{1}{n}n^2 O_P(1)\frac{n}{(b_n^2)^2}o_P(1) = o_p(1)$.

Hence, $\frac{1}{n}\sum_{i=1}^{n}\{\tilde{v}_i(n) - v_i(n) - [\overline{\tilde{v}(n)} - \overline{v(n)}]\}^2 = o_P(1)$.

By (4.23) and (4.21), the sequence $\left\{\frac{1}{n}\sum_{i=1}^{n}[v_i(n) - \overline{v(n)}]^2\right\}_{n\ge 1}$ has a positive finite limit, in probability. Hence, (4.35) follows in this case.

Case 2. Assume that $\mathrm{Var}(\|\boldsymbol{\xi}\|) = \infty$.

By Assumption $(A1)$, $\boldsymbol{\xi}_1 \in GDAN$ and therefore, by Theorem 1.1 in [13], we obtain $\frac{1}{m}\sum_{j=1}^{m}\xi_{1j} \in DAN$. Let $(\overline{b}_n)_{n\ge 1}$ be a sequence of constants such that

$$\frac{1}{\overline{b}_n}\sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\xi_{ij} - \overline{\mu}\right) \xrightarrow{\mathcal{D}} N(0,1).$$

Hence, $\frac{1}{\overline{b}_n^2}\sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\xi_{ij} - \overline{\mu}\right)^2 \xrightarrow{P} 1$. By applying WLLN, the following term converges to 1, in probability

$$\frac{1}{\overline{b}_n^2}\sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\xi_{ij} - \overline{\overline{\boldsymbol{\xi}}}\right)^2 = \frac{1}{\overline{b}_n^2}\left[\sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\xi_{ij} - \overline{\mu}\right)^2 - n(\overline{\overline{\boldsymbol{\xi}}} - \overline{\mu})^2\right]$$

and

$$\frac{1}{\overline{b}_n^2}\sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\delta_{ij} - \overline{\overline{\boldsymbol{\delta}}}\right)^2 = \frac{n}{\overline{b}_n^2}O_P(1) = o_P(1),$$

using the fact that $\overline{b}_n^2 = nl^2(n)$, where $l(n)$ it is a slowly varying function at ∞. Hence

$$\sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\xi_{ij} - \overline{\overline{\boldsymbol{\xi}}}\right)^2 + \sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m}\delta_{ij} - \overline{\overline{\boldsymbol{\delta}}}\right)^2 = \overline{b}_n^2[O_P(1) + o_P(1)].$$

By Lemma 4.10, $\widehat{\beta}_n - \beta = \frac{\sqrt{c_n^2}}{b_n^2} O_P(1)$, with $c_n^2 = \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}} \mathbf{B}_n^{-2})$ and $\boldsymbol{\gamma}_i = \boldsymbol{\varepsilon}_i - \beta \boldsymbol{\delta}_i$, for $i \le n$. Hence, $(\widehat{\beta}_n - \beta)^2 = \frac{c_n^2}{(b_n^2)^2} O_P(1)$ and so

$$\frac{1}{n-1} I_1 = \frac{\overline{b}_n^2 c_n^2}{(n-1)(b_n^2)^2} O_P(1)[O_P(1) + o_P(1)] = o_P(1).$$

In addition, $\frac{1}{n-1} I_2 \cdot I_3 = \frac{n^2 c_n^2}{(n-1)(b_n^2)^2} O_P(1) o_P(1) = o_P(1)$. Therefore,

$$\frac{1}{n-1} \sum_{i=1}^{n} \{\tilde{v}_i(n) - v_i(n) - [\overline{\tilde{v}(n)} - \overline{v(n)}]\}^2 = o_P(1).$$

From (4.26) and (4.25), it follows that $\left\{ \frac{1}{n-1} \sum_{i=1}^{n} [v_i(n) - \overline{v(n)}]^2 \right\}_{n \ge 1}$ has a positive finite limit, in probability. □

Remark 4.13. (see also Remark 2.2)

In the case of replications, the assumption that $\boldsymbol{\xi}$ has a full distribution does not hold. Nevertheless, the results remain valid, under slightly different conditions:

($A1$) ξ_1 lies in the domain of attraction of the normal law (DAN)

($A2$) $E(\varepsilon_{1j}) = 0,\ E(\delta_{1j}) = 0,\ E(\varepsilon_{1j}^4) < \infty,\ E(\delta_{1j}^4) < \infty$ with $1 \le j \le m$, and $\boldsymbol{\Sigma}^*_{\text{error}} := \begin{pmatrix} \sigma^*_{\boldsymbol{\varepsilon}} & \sigma^*_{\boldsymbol{\varepsilon\delta}} \\ \sigma^*_{\boldsymbol{\varepsilon\delta}} & \sigma^*_{\boldsymbol{\delta}} \end{pmatrix}$ is positive definite

($A3$) $(\xi_i)_{1 \le i \le n}$ and $\{(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i)\}_{1 \le i \le n}$ are independent.

where $\sigma^*_{\boldsymbol{\varepsilon}} = \sum_{j,k=1}^{m} \sigma_{\boldsymbol{\varepsilon},jk}$, $\sigma^*_{\boldsymbol{\delta}} = \sum_{j,k=1}^{m} \sigma_{\boldsymbol{\delta},jk}$, $\sigma^*_{\boldsymbol{\varepsilon\delta}} = \sum_{j,k=1}^{m} \sigma_{\boldsymbol{\varepsilon\delta},jk}$.

In this case,

$$\begin{aligned} \boldsymbol{\zeta}_1 &= \Bigg(\xi_1 \sum_{j=1}^{m} \varepsilon_{1j}, \xi_1 \sum_{j=1}^{m} \delta_{1j}, \frac{1}{m} \sum_{j=1}^{m} \varepsilon_{1j}, \frac{1}{m} \sum_{j=1}^{m} \delta_{1j}, \boldsymbol{\delta}_1^T \boldsymbol{\varepsilon}_1 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}}), \\ &\qquad \|\boldsymbol{\delta}_1\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}}) \Bigg)^T. \end{aligned}$$

One can follow the steps in the proof of Lemma 4.3 and use Lemma 4 in [12] in lieu of Theorem 3.4 to obtain this result. Furthermore, one can prove Theorem 4.4, Theorem 4.7 and Theorem 4.12, using similar arguments.

Note that in this special case of repeated measurements, the converse of Lemma 4.3 also holds true, by Theorem 1 in [10]. Hence, in this case the assumption ($A1^*$) becomes necessary for the CLT in Theorem 4.4 and Theorem 4.7, by Lemma 7 in [12].

Appendix A: Auxiliary results

The first lemma is essential for the development of our results and enables us to find the relationship between the trace of $\Sigma\mathbf{B}_n$ and the trace of $\mathbf{B}_n$.

Lemma A.1. *Assume that* $\mathbf{A}$ *and* $\mathbf{B}$ *are two* $m \times m$ *symmetric matrices such that* $\mathbf{A}$ *is positive definite and* $\mathbf{B}$ *is positive semidefinite. Then:*

$$\frac{\text{trace}(\mathbf{B})}{\text{trace}(\mathbf{A^{-1}})} \leq \text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA}) \leq \text{trace}(\mathbf{A})\text{trace}(\mathbf{B})$$

Proof. The second inequality follows from Theorem 6.5, of [14]. Applying the theorem again, since $\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2}$ is positive definite, we obtain:

$$\begin{aligned} \text{trace}(\mathbf{B}) &= \text{trace}(\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2}\mathbf{A}^{-1}) \leq \text{trace}(\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2})\text{trace}(\mathbf{A}^{-1}) \\ &= \text{trace}(\mathbf{AB})\text{trace}(\mathbf{A}^{-1}). \end{aligned}$$

□

Lemma A.2. *If* $(s_i)_{i\geq 1}$ *and* $(t_i)_{i\geq 1}$ *are sequences of random variables such that:* $\frac{\sum_{i=1}^n (s_i - t_i)^2}{\sum_{i=1}^n t_i^2} \xrightarrow{P} 0$*, then* $\frac{\sum_{i=1}^n s_i^2}{\sum_{i=1}^n t_i^2} \xrightarrow{P} 1$.

Proof. We have $|\sum_{i=1}^n (s_i^2 - t_i^2)| = |\sum_{i=1}^n [(s_i - t_i)^2 + 2t_i(s_i - t_i)]| \leq \sum_{i=1}^n (s_i - t_i)^2 + 2(\sum_{i=1}^n t_i^2)^{1/2}[\sum_{i=1}^n (s_i - t_i)^2]^{1/2}$ and therefore:

$$\left|\frac{\sum_{i=1}^n (s_i^2 - t_i^2)}{\sum_{i=1}^n t_i^2}\right| \leq \frac{\sum_{i=1}^n (s_i - t_i)^2}{\sum_{i=1}^n t_i^2} + 2\left(\frac{\sum_{i=1}^n (s_i - t_i)^2}{\sum_{i=1}^n t_i^2}\right)^{1/2}.$$

The result follows since, by hypothesis, the right hand side converges to 0, in probability. □

Appendix B: Proof of Lemmas 4.8, 4.9 and 4.10

B.1. Proof of Lemma 4.8

By Lemma 4.3, $\mathbf{b}^T\boldsymbol{\zeta}_1 \in DAN$ and so:

$$\frac{\sqrt{n}\ \overline{\mathbf{b}^T\boldsymbol{\zeta}}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n [\mathbf{b}^T(\boldsymbol{\zeta}_i - \overline{\boldsymbol{\zeta}})]^2}} \xrightarrow{\mathcal{D}} N(0,1), \tag{B.1}$$

where $\overline{\mathbf{b}^T\boldsymbol{\zeta}} = \frac{1}{n}\sum_{i=1}^n \mathbf{b}^T\boldsymbol{\zeta}_i$ and $\overline{\boldsymbol{\zeta}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\zeta}_i$.

Case 1. Assume that $\text{Var}(\|\boldsymbol{\xi}\|) < \infty$ or $b_1 = b_2 = 0$. Since $\text{Var}(\mathbf{b}^T\boldsymbol{\zeta}_1) < \infty$, we apply WLLN to obtain:

$$\frac{1}{n}\sum_{i=1}^n [\mathbf{b}^T(\boldsymbol{\zeta}_i - \overline{\boldsymbol{\zeta}})]^2 \xrightarrow{P} \text{Var}(\mathbf{b^T}\boldsymbol{\zeta_1}) > \mathbf{0}. \tag{B.2}$$

(The fact that $\mathrm{Var}(\mathbf{b^T}\boldsymbol{\zeta}_\mathbf{1}) > \mathbf{0}$ follows by hypothesis, if $b_1 = b_2 = 0$, and was shown in the proof of Lemma 4.3, if $|b_1| + |b_2| > 0$.) Using Slutsky's Theorem, (B.1) and (B.2) we obtain $\overline{\mathbf{b}^T\boldsymbol{\zeta}} = \frac{O_P(1)}{\sqrt{n}}$.

Case 2. Assume that $\mathrm{Var}(\|\boldsymbol{\xi}\|) = \infty$ and $|b_1| + |b_2| > 0$.

Let $(c_n)_{n\geq 1}$ be a sequence of positive constants. Then:

$$\sqrt{\frac{n(n-1)}{c_n^2}}\overline{\mathbf{b}^T\boldsymbol{\zeta}} = \frac{\sqrt{n}\overline{\mathbf{b}^T\boldsymbol{\zeta}}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{b}^T(\boldsymbol{\zeta}_i - \overline{\boldsymbol{\zeta}})]^2}}\sqrt{\frac{\sum_{i=1}^n[\mathbf{b}^T(\boldsymbol{\zeta}_i - \overline{\boldsymbol{\zeta}})]^2}{c_n^2}}.$$

By (B.1), the first factor of the product converges in distribution to $N(0, 1)$ and we show

$$\sum_{i=1}^n \frac{[\mathbf{b}^T(\boldsymbol{\zeta}_i - \overline{\boldsymbol{\zeta}})]^2}{c_n^2} \xrightarrow{P} 1. \tag{B.3}$$

From the proof of Lemma 4.3, we recall that, for each $i \leq n$

$$\mathbf{b}^T\boldsymbol{\zeta}_i = \boldsymbol{\xi}_i^T\boldsymbol{\gamma}_i + f_b(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i), \tag{B.4}$$

where $\boldsymbol{\gamma}_i = b_1\boldsymbol{\varepsilon}_i + b_2\boldsymbol{\delta}_i$ and

$$\begin{aligned} f_\mathbf{b}(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i) &= b_3\left(\frac{1}{m}\sum_{j=1}^m \varepsilon_{ij}\right) + b_4\left(\frac{1}{m}\sum_{j=1}^m \delta_{ij}\right) + b_5[\boldsymbol{\delta}_i^T\boldsymbol{\varepsilon}_i - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\delta}})] \\ &+ b_6[\|\boldsymbol{\delta}_i\|^2 - \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]. \end{aligned}$$

Note that $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}} = \mathrm{E}(\boldsymbol{\gamma}_1\boldsymbol{\gamma}_1^T) > 0$, since $\boldsymbol{\Sigma}_{\mathrm{error}} > 0$. We denote $\boldsymbol{\gamma}_i' = \boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1/2}\boldsymbol{\gamma}_i$, and $\boldsymbol{\xi}_i' := \boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{1/2}(\boldsymbol{\xi}_i - \boldsymbol{\mu})$ for all $i \leq n$. It follows that $\mathrm{Var}(\boldsymbol{\gamma}_i') = \mathbf{I}$ and $\mathrm{E}(\boldsymbol{\xi}_i') = \mathbf{0}$, for all $i \leq n$.

We first prove that

$$\frac{1}{c_n^2}\sum_{i=1}^n(\boldsymbol{\xi}_i^T\boldsymbol{\gamma}_i)^2 \xrightarrow{P} 1.$$

Let $(\mathbf{B}_n)_{n\geq 1}$ be a sequence of symmetric matrices such that $\mathbf{B}_n\sum_{i=1}^n(\boldsymbol{\xi}_i - \boldsymbol{\mu}) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I})$.

By Lemma 3.5, $\boldsymbol{\xi}_1' \in GDAN$, and $\mathbf{B}_n'\sum_{i=1}^n \boldsymbol{\xi}_i' \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I})$, where $\mathbf{B}_n' = (\mathbf{C}_n^T\mathbf{C}_n)^{1/2}$, and $\mathbf{C}_n = \mathbf{B}_n\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}{}^{-1/2}$. From the proof of Theorem 3.4, *Case 2*, (a) (applied to $\boldsymbol{\xi}_1'$ and $\boldsymbol{\gamma}_1'$) it follows that:

$$\frac{1}{c_n^2}\sum_{i=1}^n[\boldsymbol{\xi}_i'^T\boldsymbol{\gamma}_i']^2 \xrightarrow{P} 1, \tag{B.5}$$

where

$$\begin{aligned} c_n^2 &= \mathrm{trace}[(\mathbf{B}_n')^{-2}] = \mathrm{trace}[(\mathbf{C}_n^T\mathbf{C}_n)^{-1}] = \mathrm{trace}[\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{1/2}\mathbf{B}_n^{-2}\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{1/2}] \\ &= \mathrm{trace}[\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}(\mathbf{B}_n)^{-2}]. \end{aligned}$$

Since $(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\gamma}_i = (\boldsymbol{\xi}'_i)^T \boldsymbol{\gamma}'_i$, (B.5) becomes:

$$\frac{1}{c_n^2} \sum_{i=1}^n [(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\gamma}_i]^2 \xrightarrow{P} 1.$$

We write:

$$\begin{aligned}\frac{\sum_{i=1}^n (\boldsymbol{\xi}_i^T \boldsymbol{\gamma}_i)^2}{c_n^2} &= \frac{\sum_{i=1}^n [(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\gamma}_i]^2}{c_n^2} + \frac{\sum_{i=1}^n (\boldsymbol{\mu}^T \boldsymbol{\gamma}_i)^2}{c_n^2} \\ &+ 2\frac{\sum_{i=1}^n [(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\gamma}_i](\boldsymbol{\mu}^T \boldsymbol{\gamma}_i)}{c_n^2},\end{aligned}$$

and prove that the last two terms of the sum converge to 0 in probability.

Using $(A2)$, we have $\frac{\sum_{i=1}^n \|\boldsymbol{\gamma}_i\|^2}{d_n^2} \xrightarrow{P} 1$, where $d_n^2 := n\text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}) > 0$. By the Cauchy-Scwharz inequality $\frac{\sum_{i=1}^n (\boldsymbol{\mu}^T \boldsymbol{\gamma}_i)^2}{c_n^2} \leq \|\boldsymbol{\mu}\|^2 \frac{\sum_{i=1}^n \|\boldsymbol{\gamma}_i\|^2}{d_n^2} \frac{d_n^2}{c_n^2}$. We have to show that $\lim_{n\to\infty} \frac{d_n^2}{c_n^2} = 0$. By Lemma A.1, $c_n^2 \geq \frac{\text{trace}[(\mathbf{B}_n)^{-2}]}{\text{trace}[(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}})^{-1}]}$ and hence

$$\frac{d_n^2}{c_n^2} \leq \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}})\text{trace}[(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}})^{-1}] \frac{n}{\text{trace}[(\mathbf{B}_n)^{-2}]} \to 0,$$

where we used Lemma 3.6, (a).

By the Cauchy-Schwarz inequality, it also follows that

$$\left| \frac{\sum_{i=1}^n [(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\gamma}_i](\boldsymbol{\mu}^T \boldsymbol{\gamma}_i)}{c_n^2} \right| \xrightarrow{P} 0,$$

and so $\frac{\sum_{i=1}^n (\boldsymbol{\xi}_i^T \boldsymbol{\gamma}_i)^2}{c_n^2} \xrightarrow{P} 1$.

We now return to the study of $\overline{\mathbf{b}^T \boldsymbol{\zeta}}$. Using (B.4), we obtain:

$$\begin{aligned}\sum_{i=1}^n \frac{[\mathbf{b}^T(\boldsymbol{\zeta}_i - \overline{\boldsymbol{\zeta}})]^2}{c_n^2} &= \sum_{i=1}^n \frac{(\boldsymbol{\xi}_i^T \boldsymbol{\gamma}_i - \overline{\boldsymbol{\xi}^T \boldsymbol{\gamma}})^2}{c_n^2} + \sum_{i=1}^n \frac{[f_{\mathbf{b}}(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i) - \overline{f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta})}]^2}{c_n^2} \\ &+ 2\sum_{i=1}^n \frac{(\boldsymbol{\xi}_i^T \boldsymbol{\gamma}_i - \overline{\boldsymbol{\xi}^T \boldsymbol{\gamma}})[f_{\mathbf{b}}(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i) - \overline{f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta})}]}{c_n^2} \\ &:= I_1 + I_2 + I_3,\end{aligned}$$

where $\overline{\boldsymbol{\xi}^T \boldsymbol{\gamma}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\xi}_i^T \boldsymbol{\gamma}_i$ and $\overline{f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta})} = \frac{1}{n}\sum_{i=1}^n f_{\mathbf{b}}(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i)$.

We have $I_1 = \sum_{i=1}^n \frac{(\boldsymbol{\gamma}_i^T \boldsymbol{\xi}_i - \overline{\boldsymbol{\gamma}^T \boldsymbol{\xi}})^2}{c_n^2} = \sum_{i=1}^n \frac{(\boldsymbol{\gamma}_i^T \boldsymbol{\xi}_i)^2}{c_n^2} - \frac{n}{c_n^2} \overline{\boldsymbol{\gamma}^T \boldsymbol{\xi}}^2 \xrightarrow{P} 1$, since by applying the WLLN and using the fact that $\lim_{n\to\infty} \frac{n}{c_n^2} = 0$, the second term converges to 0, in probability.

We write

$$\begin{aligned}I_2 &= \sum_{i=1}^n \frac{[f_{\mathbf{b}}(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i) - \overline{f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta})}]^2}{c_n^2} \\ &= \frac{\sum_{i=1}^n [f_{\mathbf{b}}(\boldsymbol{\varepsilon}_i, \boldsymbol{\delta}_i) - \overline{f_{\mathbf{b}}(\boldsymbol{\varepsilon}, \boldsymbol{\delta})}]^2}{(n-1)\text{Var}[f_{\mathbf{b}}(\boldsymbol{\varepsilon}_1, \boldsymbol{\delta}_1)]} \frac{(n-1)\text{Var}[f_{\mathbf{b}}(\boldsymbol{\varepsilon}_1, \boldsymbol{\delta}_1)]}{c_n^2}.\end{aligned}$$

By the WLLN, the first factor converges in probability to 1 whereas the second factor of the product converges to 0, in probability, since $\lim_{n\to\infty} \frac{n}{c_n^2} = 0$.

By the Cauchy-Schwarz inequality, $I_3 \le \sqrt{I_1 I_2}$ and hence $I_3 \xrightarrow{P} 0$. Therefore, (B.3) holds.

It remains to prove that $\lim_{n\to\infty} \frac{c_n^2}{n^2} = 0$. Using Lemma A.1 and Lemma 3.6, (b) we obtain:

$$\frac{c_n^2}{n^2} = \frac{\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}} \mathbf{B}_n^{-2})}{n^2} \le \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}) \frac{\mathrm{trace}(\mathbf{B}_n^{-2})}{n^2} \to 0.\square$$

B.2. Proof of Lemma 4.9

By Lemma 4.6, (a)

$$\frac{\sqrt{n}\,\overline{\mathbf{d}^T \boldsymbol{\eta}(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n [\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}} \xrightarrow{\mathcal{D}} N(0,1).$$

Case 1. Assume that $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$ or $e_1 = e_2 = 0$. We write:

$$\begin{aligned} \overline{\mathbf{d}^T \boldsymbol{\eta}(n)} &= \frac{\sqrt{n}\,\overline{\mathbf{d}^T \boldsymbol{\eta}(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n [\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}} \\ &\quad \cdot \sqrt{\frac{1}{n-1}} \sqrt{\frac{1}{n}\sum_{i=1}^n [\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}. \end{aligned}$$

From the proof of Lemma 4.8, with the particular choice of $\mathbf{b} = \mathbf{e}$ and $\boldsymbol{\zeta}_i$ replaced by $\boldsymbol{\zeta}_i'$ (see relation (B.2)) we have $\frac{1}{n}\sum_{i=1}^n [\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2 \xrightarrow{P} \mathrm{Var}(\mathbf{e}^T \boldsymbol{\zeta}_1') > 0$, which together with (4.17) implies:

$$\frac{1}{n}\sum_{i=1}^n [\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2 \xrightarrow{P} \mathrm{Var}(\mathbf{e}^T \boldsymbol{\zeta}_1') > 0. \tag{B.6}$$

To complete the proof we apply Lemma 4.6, (a).

Case 2. Assume that $\mathrm{Var}(\|\boldsymbol{\xi}\|) = \infty$ and $|e_1| + |e_2| > 0$. We write:

$$\begin{aligned} \overline{\mathbf{d}^T \boldsymbol{\eta}(n)} &= \frac{\sqrt{n}\,\overline{\mathbf{d}^T \boldsymbol{\eta}(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n [\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}} \\ &\quad \cdot \sqrt{\frac{c_n^2}{n(n-1)}} \sqrt{\frac{1}{c_n^2}\sum_{i=1}^n [\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)})]^2}. \end{aligned}$$

From the proof of Lemma 4.8, with $\mathbf{b} = \mathbf{e}$ and $\boldsymbol{\zeta}_i$ replaced by $\boldsymbol{\zeta}_i'$ (see relation (B.3)), it follows that $\frac{1}{c_n^2}\sum_{i=1}^n [\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2 \xrightarrow{P} 1$. Hence, we use (4.17) to obtain:

$$\frac{1}{c_n^2}\sum_{i=1}^{n}[\mathbf{d}^T(\boldsymbol{\eta}_i(n)-\overline{\boldsymbol{\eta}(n)})]^2 \xrightarrow{P} 1. \tag{B.7}$$

From (B.7) and Lemma 4.6, (a), the conclusion follows. Here, the convergence rate depends on the sequence $c_n^2 = \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}\mathbf{B}_n^{-2})$, where $\boldsymbol{\gamma}_i = e_1\boldsymbol{\varepsilon}_i + e_2\boldsymbol{\delta}_i$, $1 \le i \le n$ and $(\mathbf{B}_n)_{n\ge 1}$ is a sequence of symmetric matrices such that

$$\mathbf{B}_n\sum_{i=1}^{n}(\boldsymbol{\xi}_i-\boldsymbol{\mu}) \xrightarrow{\mathcal{D}} N(\mathbf{0},\mathbf{I}).\square$$

B.3. Proof of Lemma 4.10

Recall that $u_i(n) = \mathbf{d}^T\boldsymbol{\eta}_i(n)$, where $\mathbf{d} = (0,0,1,-\beta)^T$ (see (4.11)). By Lemma 4.6, (a):

$$\frac{\sqrt{n}\ \overline{u(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}[u_i(n)-\overline{u(n)}]^2}} \xrightarrow{\mathcal{D}} N(0,1), \tag{B.8}$$

where $\overline{u(n)} = \frac{1}{n}\sum_{i=1}^{n}u_i(n)$.

Case 1. Assume that $\text{Var}(\|\boldsymbol{\xi}\|) < \infty$. Then, by (B.6), with $\mathbf{d}^T = (0,0,1,-\beta)$, it follows that:

$$\frac{1}{n-1}\sum_{i=1}^{n}[u_i(n)-\overline{u(n)}]^2 \xrightarrow{P} \lambda = \text{Var}(\mathbf{e}^T\boldsymbol{\zeta}_1') > 0,$$

where $\mathbf{e} = (1,-\beta,0,0,1,-\beta)^T$. We use (4.9) and write

$$\begin{aligned}\frac{b_n^2}{\sqrt{n}}(\widehat{\beta}_n-\beta) &= \frac{b_n^2}{\sqrt{n}}\cdot\frac{n\overline{u(n)}}{\sum_{i=1}^{n}[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}\\ &= \frac{\sqrt{n}\overline{u(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}[u_i(n)-\overline{u(n)}]^2}}b_n^2\frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}[u_i(n)-\overline{u(n)}]^2}}{\sum_{i=1}^{n}[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}.\end{aligned}$$

We apply Slutsky's theorem, (B.8) and (4.2) to obtain the conclusion

Case 2. Assume that $\text{Var}(\|\boldsymbol{\xi}\|) = \infty$. By (B.7), with $\mathbf{d} = (0,0,1,-\beta)^T$, it follows that:

$$\frac{1}{c_n^2}\sum_{i=1}^{n}[u_i(n)-\overline{u(n)}]^2 \xrightarrow{P} 1,$$

where $c_n^2 = \text{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}\mathbf{B}_n^{-2})$, and $\gamma_i = e_1\boldsymbol{\varepsilon}_i + e_2\boldsymbol{\delta}_i = \boldsymbol{\varepsilon}_i - \beta\boldsymbol{\delta}_i$, since $\mathbf{e} = (1,-\beta,0,0,1,-\beta)^T$. Using (4.9) again, we obtain:

$$\begin{aligned}\frac{b_n^2}{\sqrt{c_n^2}}(\widehat{\beta}_n-\beta) &= \frac{b_n^2}{\sqrt{c_n^2}}\cdot\frac{n\overline{u(n)}}{\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}\\ &= \frac{\sqrt{n}\overline{u(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[u_i(n)-\overline{u(n)}]^2}}\cdot\sqrt{\frac{1}{c_n^2}\sum_{i=1}^n[u_i(n)-\overline{u(n)}]^2}\\ &\cdot\ \sqrt{\frac{n}{n-1}}\cdot\frac{b_n^2}{\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}.\end{aligned}$$

The conclusion follows by Slutsky's theorem (B.8) and (4.2).□

Appendix C: Proof of (4.16) and (4.17)

Recall that $\overline{\boldsymbol{\delta}}=(\overline{\delta}_1,\dots,\overline{\delta}_m)^T$, where $\overline{\delta}_j=\frac{1}{n}\sum_{i=1}^n\delta_{ij}$. Note that:

$$\begin{aligned}\overline{R(n)} &= e_1[-\overline{\boldsymbol{\varepsilon}}^T(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})]+e_2[-\overline{\boldsymbol{\delta}}^T(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})]\\ &+\ e_5(-\overline{\boldsymbol{\delta}}^T\overline{\boldsymbol{\varepsilon}})+e_6(-\|\overline{\boldsymbol{\delta}}\|^2). \end{aligned}\tag{C.1}$$

To prove (4.16), we consider two cases.

Case 1. Assume that $\mathrm{Var}(\boldsymbol{\xi})<\infty$ or $|e_1|+|e_2|=0$.
In this case, we prove:

$$\sqrt{n}\overline{R(n)}\xrightarrow{P}0.\tag{C.2}$$

For each $1\le j\le m$, by applying the CLT, we obtain $\sqrt{n}\overline{\delta}_j=O_P(1)$ and so $\sqrt{n}\overline{\delta}_j^2=o_P(1)$. Hence, $\sqrt{n}\|\overline{\boldsymbol{\delta}}\|^2=o_P(1)$. Similarly, we have $\sqrt{n}\|\overline{\boldsymbol{\varepsilon}}\|^2=o_P(1)$. By Cauchy-Schwarz inequality, $n|\overline{\boldsymbol{\delta}}^T\overline{\boldsymbol{\varepsilon}}|^2\le n\|\overline{\boldsymbol{\varepsilon}}\|^2\cdot\|\overline{\boldsymbol{\delta}}\|^2$ and therefore, we also have $\sqrt{n}|\overline{\boldsymbol{\delta}}^T\overline{\boldsymbol{\varepsilon}}|=o_P(1)$.

Similarly, we obtain $\sqrt{n}|\overline{\boldsymbol{\varepsilon}}^T(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})|=o_P(1)$ and $\sqrt{n}|\overline{\boldsymbol{\delta}}^T(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})|=o_P(1)$, so (C.2) follows.

To prove (4.16) in this case, we write:

$$\frac{\sqrt{n}\overline{R(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{e}^T(\boldsymbol{\zeta}_i'-\overline{\boldsymbol{\zeta}'})]^2}}=\frac{\sqrt{n}\overline{R(n)}}{\sqrt{\mathrm{Var}(\mathbf{e}^T\boldsymbol{\zeta}')}}\cdot\sqrt{\frac{\mathrm{Var}(\mathbf{e}^T\boldsymbol{\zeta}')}{\frac{1}{n-1}\sum_{i=1}^n[\mathbf{e}^T(\boldsymbol{\zeta}_i'-\overline{\boldsymbol{\zeta}'})]^2}}.$$

By (B.2), (with $\mathbf{b}=\mathbf{e}$ and $\boldsymbol{\zeta}_i$ replaced by $\boldsymbol{\zeta}_i'$) the second factor of the product converges to 1, in probability. Since $0<\mathrm{Var}(\mathbf{e}^T\boldsymbol{\zeta}')<\infty$, (C.2) implies (4.16).

Case 2. Assume that $\mathrm{Var}(\|\boldsymbol{\xi}\|)=\infty$ and $|e_1|+|e_2|>0$.

Let $(\mathbf{B}_n)_{n\ge1}$ be a sequence of (nonsingular, symmetric) non-stochastic matrices such that $\mathbf{B}_n\sum_{i=1}^n(\boldsymbol{\xi}_i-\boldsymbol{\mu})\xrightarrow{\mathcal{D}}N(\mathbf{0},\mathbf{I})$. Let $b_n^2=\mathrm{trace}(\mathbf{B}_n^{-2})$. Define $c_n^2=\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}\mathbf{B}_n^{-2})$, where $\boldsymbol{\gamma}_i=e_1\boldsymbol{\varepsilon}_i+e_2\boldsymbol{\delta}_i$ and $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}=\mathrm{Var}(\boldsymbol{\gamma}_i)$. In this case, in lieu of (C.2), we prove:

$$\frac{n}{\sqrt{c_n^2}}\overline{R(n)}\xrightarrow{P}0.\tag{C.3}$$

As in the proof of Lemma 4.8, $\frac{n}{c_n^2} \to 0$. Therefore, the last three terms in (C.1), multiplied by $\frac{n}{\sqrt{c_n^2}}$ converge to 0, in probability.

As for the first term, by applying Cauchy-Schwarz inequality, we have:

$$\frac{n}{\sqrt{c_n^2}}|(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \overline{\boldsymbol{\varepsilon}}| \leq \frac{n}{\sqrt{c_n^2}}\|\overline{\boldsymbol{\xi}} - \boldsymbol{\mu}\| \cdot \|\overline{\boldsymbol{\varepsilon}}\|.$$

Recall that $\|\overline{\boldsymbol{\varepsilon}}\| = \frac{O_P(1)}{\sqrt{n}}$.

Since $\|\mathbf{B}_n \sum_{i=1}^n (\boldsymbol{\xi}_i - \boldsymbol{\mu})\| = O_P(1)$, we have the following relations:

$$\begin{aligned}
\|\overline{\boldsymbol{\xi}} - \boldsymbol{\mu}\|^2 &= \frac{1}{n^2}\left\|\sum_{i=1}^n (\boldsymbol{\xi}_i - \boldsymbol{\mu})\right\|^2 = \frac{1}{n^2}\left\|\mathbf{B}_n^{-1}\mathbf{B}_n \sum_{i=1}^n (\boldsymbol{\xi}_i - \boldsymbol{\mu})\right\|^2 \\
&\leq \frac{1}{n^2}\|\mathbf{B}_n^{-1}\|^2 \cdot \left\|\mathbf{B}_n \sum_{i=1}^n (\boldsymbol{\xi}_i - \boldsymbol{\mu})\right\|^2 = \frac{1}{n^2} b_n^2 O_P(1).
\end{aligned}$$

Therefore $\frac{n}{\sqrt{c_n^2}}|(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \overline{\boldsymbol{\varepsilon}}| \leq \frac{n}{\sqrt{c_n^2}} \frac{\sqrt{b_n^2}}{n} \frac{O_P(1)}{\sqrt{n}} = \frac{O_P(1)}{\sqrt{n}} \sqrt{\frac{b_n^2}{c_n^2}}$.

From Lemma A.1 we obtain $\frac{b_n^2}{c_n^2} \leq \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1})$, so $\frac{n}{\sqrt{c_n^2}}|(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \overline{\boldsymbol{\varepsilon}}| = o_P(1)$.

Similarly, $\frac{n}{\sqrt{c_n^2}}|(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \overline{\boldsymbol{\delta}}| = o_P(1)$ and hence the convergence in (C.3) is proved.

Therefore, if we write:

$$\frac{\sqrt{n}\overline{R(n)}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n [\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2}} = \frac{n\overline{R(n)}}{\sqrt{c_n^2}} \cdot \sqrt{\frac{n-1}{n}} \cdot \sqrt{\frac{1}{\frac{1}{c_n^2}\sum_{i=1}^n [\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2}},$$

the last factor converges to 1, in probability by (B.3) (with $\mathbf{b} = \mathbf{e}$ and $\boldsymbol{\zeta}_i$ replaced by $\boldsymbol{\zeta}_i'$). This together with (C.3) concludes the proof of (4.16) in this case.

Now we prove (4.17). Using (4.13), we write:

$$\begin{aligned}
\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)}) &= e_1[(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\varepsilon}_i - (\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \boldsymbol{\varepsilon}_i - (\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \overline{\boldsymbol{\varepsilon}} - \overline{(\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\varepsilon}} \\
&+ 2(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \overline{\boldsymbol{\varepsilon}}] \\
&+ e_2[(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\delta}_i - (\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \boldsymbol{\delta}_i - (\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \overline{\boldsymbol{\delta}} - \overline{(\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\delta}} \\
&+ 2(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \overline{\boldsymbol{\delta}}] \\
&+ e_3\left(\frac{1}{m}\sum_{j=1}^m \varepsilon_{ij} - \overline{\overline{\varepsilon}}\right) + e_4\left(\frac{1}{m}\sum_{j=1}^m \delta_{ij} - \overline{\overline{\delta}}\right) \\
&+ e_5(\boldsymbol{\delta}_i^T \boldsymbol{\varepsilon}_i - \boldsymbol{\delta}_i^T \overline{\boldsymbol{\varepsilon}} - \overline{\boldsymbol{\delta}}^T \boldsymbol{\varepsilon}_i - \overline{\boldsymbol{\delta}^T \boldsymbol{\varepsilon}} + 2\overline{\boldsymbol{\delta}}^T \overline{\boldsymbol{\varepsilon}}) \\
&+ e_6(\|\boldsymbol{\delta}_i\|^2 - 2\boldsymbol{\delta}_i^T \overline{\boldsymbol{\delta}} - \overline{\|\boldsymbol{\delta}\|^2} + 2\|\overline{\boldsymbol{\delta}}\|^2),
\end{aligned}$$

where $\overline{\overline{\varepsilon}} = \frac{1}{nm}\sum_{i=1}^n \sum_{j=1}^m \varepsilon_{ij}$, $\overline{\overline{\delta}} = \frac{1}{nm}\sum_{i=1}^n \sum_{j=1}^m \delta_{ij}$, $\overline{(\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\varepsilon}} = \frac{1}{n}\sum_{i=1}^n (\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\varepsilon}_i$, $\overline{(\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\delta}} = \frac{1}{n}\sum_{i=1}^n (\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\delta}_i$, $\overline{\boldsymbol{\delta}^T \boldsymbol{\varepsilon}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\delta}_i^T \boldsymbol{\varepsilon}_i$ and $\overline{\|\boldsymbol{\delta}\|^2} = \frac{1}{n}\sum_{i=1}^n \|\boldsymbol{\delta}_i\|^2$.

Using (4.8), we obtain

$$\begin{aligned}
\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'}) &= e_1[(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\varepsilon}_i - \overline{(\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\varepsilon}}] \\
&+ e_2[(\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \boldsymbol{\delta}_i^T - \overline{(\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\delta}}] \\
&+ e_3\left(\frac{1}{m}\sum_{j=1}^m \varepsilon_{ij} - \overline{\overline{\varepsilon}}\right) + e_4\left(\frac{1}{m}\sum_{j=1}^m \delta_{ij} - \overline{\overline{\delta}}\right) + e_5(\boldsymbol{\delta}_i^T \boldsymbol{\varepsilon}_i - \overline{\boldsymbol{\delta}^T \boldsymbol{\varepsilon}}) \\
&+ e_6(\|\boldsymbol{\delta}_i\|^2 - \overline{\|\boldsymbol{\delta}\|^2}),
\end{aligned}$$

and therefore

$$\begin{aligned}
\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)}) - \mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'}) &= e_1[-(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \boldsymbol{\varepsilon}_i - (\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \overline{\boldsymbol{\varepsilon}} \\
&+ 2(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \overline{\boldsymbol{\varepsilon}}] \\
&+ e_2[-(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \boldsymbol{\delta}_i - (\boldsymbol{\xi}_i - \boldsymbol{\mu})^T \overline{\boldsymbol{\delta}} \\
&+ 2(\overline{\boldsymbol{\xi}} - \boldsymbol{\mu})^T \overline{\boldsymbol{\delta}}] \\
&+ e_5(-\boldsymbol{\delta}_i^T \overline{\boldsymbol{\varepsilon}} - \overline{\boldsymbol{\delta}}^T \boldsymbol{\varepsilon}_i + 2\overline{\boldsymbol{\delta}}^T \overline{\boldsymbol{\varepsilon}}) \\
&+ e_6(-2\boldsymbol{\delta}_i^T \overline{\boldsymbol{\delta}} + 2\|\overline{\boldsymbol{\delta}}\|^2). \qquad \text{(C.4)}
\end{aligned}$$

By Lemma A.2, to prove (4.17), it is enough to prove

$$\frac{\sum_{i=1}^n [\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)}) - \mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2}{\sum_{i=1}^n [\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2} \xrightarrow{P} 0.$$

If $\mathrm{Var}(\|\boldsymbol{\xi}\|) < \infty$ or $|e_1|+|e_2| = 0$, $\frac{1}{n}\sum_{i=1}^n [\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2 = O_P(1)$, by the WLLN (see also (B.2)). Therefore, it is enough to prove:

$$\frac{1}{n}\sum_{i=1}^n [\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)}) - \mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2 \xrightarrow{P} 0. \qquad \text{(C.5)}$$

If $\mathrm{Var}(\|\boldsymbol{\xi}\|) = \infty$ and $|e_1| + |e_2| > 0$, by (B.3), with $\mathbf{b} = \mathbf{e}$, we have

$$\frac{1}{c_n^2}\sum_{i=1}^n [\mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2 = O_P(1).$$

Therefore, if we prove (C.5), the convergence (4.17) follows since $\lim_{n\to\infty} \frac{n}{c_n^2} = 0$.

Hence, it remains to prove (C.5). By (C.4), we have the following

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n [\mathbf{d}^T(\boldsymbol{\eta}_i(n) - \overline{\boldsymbol{\eta}(n)}) - \mathbf{e}^T(\boldsymbol{\zeta}_i' - \overline{\boldsymbol{\zeta}'})]^2 &\leq 4(e_1)^2 T_1 + 4(e_2)^2 T_2 + 4(e_5)^2 T_3 \\
&+ 4(e_6)^2 T_4,
\end{aligned}$$

where:

$$T_1 := \frac{1}{n}\sum_{i=1}^{n}[-(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})^T\boldsymbol{\varepsilon}_i-(\boldsymbol{\xi}_i-\boldsymbol{\mu})^T\overline{\boldsymbol{\varepsilon}}+2(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})^T\overline{\boldsymbol{\varepsilon}}]^2,$$
$$T_2 := \frac{1}{n}\sum_{i=1}^{n}[-(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})^T\boldsymbol{\delta}_i-(\boldsymbol{\xi}_i-\boldsymbol{\mu})^T\overline{\boldsymbol{\delta}}+2(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})^T\overline{\boldsymbol{\delta}}]^2,$$
$$T_3 := \frac{1}{n}\sum_{i=1}^{n}(-\boldsymbol{\delta}_i^T\overline{\boldsymbol{\varepsilon}}-\overline{\boldsymbol{\delta}}^T\boldsymbol{\varepsilon}_i+2\overline{\boldsymbol{\delta}}^T\overline{\boldsymbol{\varepsilon}})^2,$$
$$T_4 := \frac{1}{n}\sum_{i=1}^{n}(-2\boldsymbol{\delta}_i^T\overline{\boldsymbol{\delta}}+2\|\overline{\boldsymbol{\delta}}\|^2)^2.$$

Using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} T_1 &\leq 4\frac{1}{n}\sum_{i=1}^{n}[(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})^T\boldsymbol{\varepsilon}_i]^2+4\frac{1}{n}\sum_{i=1}^{n}[(\boldsymbol{\xi}_i-\boldsymbol{\mu})^T\overline{\boldsymbol{\varepsilon}}]^2+8\frac{1}{n}\sum_{i=1}^{n}[(\overline{\boldsymbol{\xi}}-\boldsymbol{\mu})^T\overline{\boldsymbol{\varepsilon}}]^2 \\ &\leq 4\frac{1}{n}\sum_{i=1}^{n}\left[\|\overline{\boldsymbol{\xi}}-\boldsymbol{\mu}\|^2\|\boldsymbol{\varepsilon}_i\|^2+\|\overline{\boldsymbol{\varepsilon}}\|^2\|\boldsymbol{\xi}_i-\boldsymbol{\mu}\|^2\right]+8\|\overline{\boldsymbol{\xi}}-\boldsymbol{\mu}\|^2\|\overline{\boldsymbol{\varepsilon}}\|^2. \end{aligned} \tag{C.6}$$

By WLLN, $\|\overline{\boldsymbol{\xi}}-\boldsymbol{\mu}\|^2 \xrightarrow{P} 0$ and $\frac{1}{n}\sum_{i=1}^{n}\|\boldsymbol{\varepsilon}_i\|^2 \xrightarrow{P} \mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$. It follows that the first term of (C.6) converges in probability to 0.

Recall that $n\|\overline{\boldsymbol{\varepsilon}}\|^2 = O_P(1)$, and $\frac{\sum_{i=1}^{n}\|\boldsymbol{\xi}_i-\boldsymbol{\mu}\|^2}{b_n^2} \xrightarrow{P} 1$, by Lemma 3.3. Hence, we have $\|\overline{\boldsymbol{\varepsilon}}\|^2\frac{1}{n}\sum_{i=1}^{n}\|\boldsymbol{\xi}_i-\boldsymbol{\mu}\|^2 = O_P(1)\frac{b_n^2}{n^2} = o_P(1)$, by Lemma 3.6, (b).

Since $\|\overline{\boldsymbol{\xi}}-\boldsymbol{\mu}\|^2\|\overline{\boldsymbol{\varepsilon}}\|^2 = o_P(1)\frac{O_P(1)}{n} = o_P(1)$, we conclude that $T_1 \to 0$, in probability.

By replacing $\boldsymbol{\varepsilon}$ with $\boldsymbol{\delta}$ we obtain $T_2 \to 0$, in probability. Using the same technique as above, one can prove $T_i \to 0$, in probability, for $i = 3, 4$. This concludes the proof of (C.5). □

Appendix D: Proof of (4.23) and (4.26)

By Lemma A.2, to prove (4.23), it is enough to show

$$\left[\frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^{n}[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\delta}})]}-\frac{\overline{\mu}}{\mathrm{trace}(\boldsymbol{\Sigma}_{\boldsymbol{\xi}})}\right]^2\cdot\frac{\sum_{i=1}^{n}[u_i(n)-\overline{u(n)}]^2}{\sum_{i=1}^{n}[v_i'(n)-\overline{v'(n)}]^2} = o_P(1).$$

By (4.19) the first factor converges in probability to 0.

By (B.6), with $\mathbf{d}^* = (0,0,1,-\beta)^T$ and $\mathbf{e}^* = (1,-\beta,0,0,1,-\beta)^T$, we have:

$$\frac{1}{n}\sum_{i=1}^{n}[u_i(n)-\overline{u(n)}]^2 \xrightarrow{P} \mathrm{Var}(\mathbf{e}^{*T}\boldsymbol{\zeta}_1') > 0. \tag{D.1}$$

Hence, using (4.21), we complete the proof of (4.23).

We turn to the proof of (4.26). Using Lemma A.2, it is enough to prove:

$$\left[\frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]}\right]^2\cdot\frac{\sum_{i=1}^n[u_i(n)-\overline{u(n)}]^2}{\sum_{i=1}^n[v_i''-\overline{v''}]^2}=o_P(1). \tag{D.2}$$

By (4.2), $\frac{1}{n}\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]=\frac{b_n^2}{n}O_P(1)$, where $b_n^2=\operatorname{trace}(\mathbf{B}_n^{-2})$ and $(\mathbf{B}_n)_{n\geq 1}$ is a sequence of matrices such that $\mathbf{B}_n\sum_{i=1}^n(\boldsymbol{\xi}_i-\boldsymbol{\mu})\xrightarrow{\mathcal{D}}N(\mathbf{0},\mathbf{I})$. Since $\overline{\overline{\mathbf{x}}}\to\overline{\mu}$, in probability, by WLLN:

$$\frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]}=\frac{n}{b_n^2}O_P(1)=o_P(1), \tag{D.3}$$

using Lemma 3.6, (a) for the second equality.

By (B.7),

$$\frac{1}{c_n^2}\sum_{i=1}^n[u_i(n)-\overline{u(n)}]^2=O_P(1), \tag{D.4}$$

where $c_n^2=\operatorname{trace}(\mathbf{\Sigma}_{\boldsymbol{\gamma}}\mathbf{B}_n^{-2})$ and $\boldsymbol{\gamma}_i=\boldsymbol{\varepsilon}_i-\beta\boldsymbol{\delta}_i$, for $i\leq n$.

Also, (4.25) implies $\frac{\sum_{i=1}^n[u_i(n)-\overline{u(n)}]^2}{\sum_{i=1}^n[v_i''-\overline{v''}]^2}=\frac{c_n^2}{n}O_P(1)$.

Finally, using Lemma A.1:

$$\left[\frac{n\overline{\overline{\mathbf{x}}}}{\sum_{i=1}^n[\|\mathbf{x}_i-\overline{\mathbf{x}}\|^2-\operatorname{trace}(\mathbf{\Sigma}_{\boldsymbol{\delta}})]}\right]^2\cdot\frac{\sum_{i=1}^n[u_i(n)-\overline{u(n)}]^2}{\sum_{i=1}^n[v_i''-\overline{v''}]^2}$$
$$=\left(\frac{n}{b_n^2}\right)^2\frac{c_n^2}{n}O_P(1)=n\frac{c_n^2}{(b_n^2)^2}O_P(1)\leq\frac{n}{b_n^2}\operatorname{trace}(\mathbf{\Sigma}_{\boldsymbol{\gamma}})O_P(1)=o_P(1),$$

where we used Lemma 3.6, (a). This proves (D.2).□

Acknowledgements. The author would like to thank Raluca M. Balan and Ioana Schiopu-Kratina for the valuable comments and suggestions which improved the presentation of this article.

References

[1] Breiman, L. (1965). On some limit theorems similar to the arcsine law. *Theory of Probability and its Applications.* **10**, 323–331. MR0184274

[2] Buonaccorsi, J., Demidenko, E. and Tosteson, T. (2000). Estimation in longitudinal random effects models with measurement error. *Statistica Sinica.* **10**, 885–903. MR1787784

[3] Carroll, R.J., Ruppert, D., Stefanski, L. and Crainiceanu, C.M. (2006). Measurement Error in Nonlinear Models: A Modern Perspective. Chapman and Hall. MR2243417

[4] Cheng, C.-L. and Van Ness, J.W. (1999). *Statistical Regression with Measurement Error.* Arnold, London. MR1719513

[5] Fuller, W.A. (1987). *Measurement Error Models*. Wiley. MR0898653

[6] Giné, E. and Götze, F. (2004). On standard normal convergence of the multivariate student t-statistic for symmetric random vectors. *Electronic Communications in Probability.* **9**, 162–171. MR2108862

[7] Giné, E., Götze, F. and Mason, D.M. (1997). When is the Student t-statistic asymptotically standard normal? *Annals of Probability.* **25**, 1514–1531. MR1457629

[8] Gleser, L.J. (1992). The importance of assessing measurement reliability in multivariate regression. *Journal of the American Statistical Association.* **87**, 696–707. MR1185191

[9] Hahn, M. and Klass, M. (1980). Matrix normalization of sums of random vectors in the domain of attraction of the multivariate normal. *The Annals of Probability.* **8**, 262–280. MR0566593

[10] Maller, R.A. (1981). A theorem on products of random variables, with application to regression. *Australian Journal of Statistics.* **23**, 177–185. MR0636133

[11] Maller, R.A. (1993). Quadratic negligibility and the asymptotic normality of operator normed sums. *Journal of Multivariate Analysis.* **44**, 191–219. MR1219203

[12] Martsynyuk, Y.V. (2007). Central limit theorems in linear structural error-in-variables models with explanatory variables in the domain of attraction of the normal law. *Electronic Journal of Statistics.* **1**, 195–222. MR2312150

[13] Sepanski, S.J. (1994). Probabilistic characterizations of the generalized domain of attraction of the multivariate normal law. *Journal of Theoretical Probability.* **7**, 857–866. MR1295543

[14] Zhang, F. (1999). *Matrix Theory: Basic Results and Techniques*. Springer. MR1691203