Statistical Inference for Curved Fibrous Objects in 3D - based on Multiple Short Observations of Multivariate Autoregressive Processes

Gerd Gaiselmann[1], Rafal Kulik[2], Volker Schmidt[1]

[1]Institute of Stochastics, Ulm University, 89069 Ulm, Germany
[2]Department of Mathematics and Statistics, University of Ottawa, K1N 6N5 Ottawa, Ontario, Canada

Abstract

This paper deals with statistical inference on the parameters of a stochastic model describing curved fibrous objects in three dimensions that is based on multivariate autoregressive processes. The model is fitted to experimental data consisting of a large number of short, independently sampled, trajectories of multivariate autoregressive processes. We discuss relevant statistical properties (e.g., asymptotic behaviour if the number of trajectories tends to infinity) of the maximum likelihood (ML) estimators for such processes. Numerical studies are also performed to analyse some of the more intractable properties of the ML estimators. Finally the whole methodology, i.e., the fibre model and its statistical inference, is applied to appropriately describe the tracking of fibres in real materials.

Keywords: *asymptotics; bootstrap; fibre model; multivariate autoregressive processes; maximum likelihood*

1 Introduction

This paper deals with statistical inference related to the stochastic modelling of curved fibrous objects in 3D. In particular, it focuses on statistical inference for parameters used in stochastic models of single fibres. These models were introduced in Gaiselmann *et al.* (2012, 2013) and combine, in a novel way, multivariate time-series analysis with methods from spatial statistics. When fitting the stochastic fibre model to real data, the parameters are estimated using a large number of short, independently sampled, trajectories of multivariate autoregressive processes. For this reason, we consider the asymptotic properties of maximum-likelihood (ML) estimators for such types of multivariate autoregressive processes. These properties can be used to derive confidence intervals for parameters of the underlying time-series model. In addition, we carry out numerical investigations of some less tractable properties of the ML estimators.

The stochastic models developed in Gaiselmann *et al.* (2012, 2013) are primarily designed for use in materials science, where fibre-based materials play an important role. The applications of fibre-based materials include thermal insulation, aircraft and car bodies, and fuel cell technology. Materials produced using different composition and fabrication scenarios have different 3D morphologies (i.e., arrangements of the component fibres). In order to improve the

functional properties of fibre-based materials, it is important to have a good understanding of the relationship between material morphology and the resulting macroscopic behaviour of the material. A parametric stochastic model of the morphology of a given material allows practitioners to investigate this relationship by carrying out numerical simulations of physical processes on realizations of the morphology under different parameter configurations. In addition, by systematic modification of model parameters, practitioners can use stochastic models of the morphology of a material to find production parameters that improve functional properties. This process of using stochastic processes to assist in materials design is called virtual materials design.

In order to develop a stochastic model for the geometry of a fibre-based material, it is necessary to describe single fibres using a parametric stochastic model. There are several stochastic models describing straight single fibres in the literature, including Provatas *et al.* (2000); Redenbach *et al.* (2011); Schulz *et al.* (2007); Thiedmann *et al.* (2008). Stochastic models for curved fibres in 2D and 3D, respectively, were introduced in Gaiselmann *et al.* (2012, 2013). These models use a class of autoregressive processes, with a suitable correlation structure, to describe the tracks of single fibres. In this manner, they are able to describe a large variety of differently shaped 3D fibres. This approach has been successfully used to describe fibres occurring in non-woven gas-diffusion layers (GDL), a key component in proton exchange membrane fuel cells (PEMFC). The well-developed theory of time series means that it is easy to fit these stochastic models to experimental data. In addition, it is possible to perform statistical inference on the parameters of the models, including computing confidence intervals and carrying out goodness-of-fit tests. To the best of our knowledge, there are no other fibre models of this kind in the literature. The only comparable model which is able to reproduce curved fibres was introduced in Altendorf *et al.* (2011). This model uses random walks based on von Mises-Fisher distributions (generalizations of Gaussian distributions defined on the unit sphere). Our modelling approach has both advantages and disadvantages in respect of the approach described in Altendorf *et al.* (2011). For example, the model in Altendorf *et al.* (2011) allows for more control over the main direction of a fibre than our model does, as the time series we use may take a long time to return to the initial direction of the fibre. On the other hand, our time-series models can reproduce arbitrarily strong curvatures of fibres (e.g. loops), whereas the model considered in Altendorf *et al.* (2011) is restricted to fibres with much smaller curvatures.

In this paper, we concentrate on statistical inference about the parameters of the stochastic single-fibre model introduced in Gaiselmann *et al.* (2013), with a particular focus on applications to fibre-based materials. However, our approach can be applied, more generally, to situations involving a large number of short realizations of independent time series. Possible applications include DNA analysis, panel data, longitudinal data and financial modelling. Such time-series phenomena were first studied in Anderson (1978) and Azzalini (1981). More recent papers on the topic include Diggle *et al.* (1997); Ledolter *et al.* (1999); Shi *et al.* (2004); Swift *et al.* (2002). In Diggle *et al.* (1997), biomedical data

is studied using spectral analysis. Credibility models in actuarial science are considered in Ledolter *et al.* (1999). In Swift *et al.* (2002), the authors analyse normal tension glaucoma visual field data consisting of many short observations of multivariate time series. They introducing a novel computational method, based on genetic algorithms, for parameter estimation and forecasting. Finally, in Shi *et al.* (2004), time series data is analysed using ordinary linear regression. Numerous studies with similar types of data are carried out in the literature on longitudinal data and panel data, respectively; see Davis (2002); Diggle *et al.* (2002); Lindsey (1999) and the references therein.

Our main goal in this paper is to use relatively standard theory for multivariate time series, combined with spatial statistics, in the non-standard setting of materials science. As such, we do not claim to make significant theoretical innovations. Proofs are included below for completeness. We also note that an alternative modelling approach would be to use functional data analysis or longitudinal data analysis to model single fibres. Such analysis techniques are usually applied in settings where a particular phenomenon is observed in number of different locations over a common period of time. This results in (say i.i.d.) observations $X_i(t)$, $t \in [0, T]$, $i = 1, \ldots, n$ from a model $X_i(t) = m(t) + \varepsilon_i(t)$, where the $\varepsilon_i(t)$ are zero-mean error functions. A significant barrier to using functional data analysis in our setting is that the fibres we consider do not have identical lengths. That is, $t \in [0, T_i]$, where T_i is different for each fibre. We are not aware of procedures that are designed for settings where the observation windows are different for each observation. It is also worth mentioning that functional and longitudinal data analysis typically aim to extract a common trend $m(t) = \mathbb{E}X_1(t)$. It is not clear that it is sensible to estimate a common trend function $m(t)$ in our setting. This is because the shapes of the individual fibres are quite different from one another and it is not obvious that there is a 'common' functional shape (e.g., sinusoidal or polynomial). In addition, it is not clear how knowledge of a common trend $m(t)$ would allow one to simulate individual fibres. In contrast, the parametric time series approach allows for easy simulation procedure and for modelling the chaotic behaviour of the fibres.

The paper is organized as follows. In Section 2 we briefly explain the single-fibre model proposed in Gaiselmann *et al.* (2013). This is based on multivariate autoregressive processes. Then, in Section 3, we describe parameter estimation for this model. In Section 4, we give some asymptotic properties of the estimators. We then examine further properties of the estimators using simulation studies; see Section 5. An example application to real data is discussed in Section 6. We conclude in Section 7.

2 Stochastic single-fibre model

To begin with, we provide the basic idea for stochastic modelling of strongly curved fibres in 3D, as proposed in Gaiselmann *et al.* (2013). The aim of this model is to simulate systems of curved fibres which represent non-woven GDL being a key component of PEMFC, see Figure 1 (left). Such stochastic models

can then be used to perform virtual materials design, i.e., to detect structures with optimized physical properties and thus, enhanced performance.

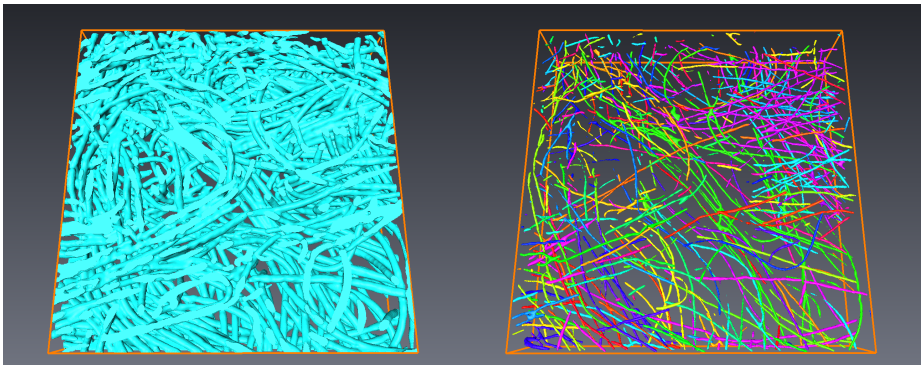

Figure 1: 3D synchrotron data (left) and extracted single fibres (right) of non-woven GDL

The basic modelling idea is to represent single fibres by polygonal tracks $p = (\ell_1 \ldots, \ell_T)$ which are subsequently dilated in 3D, where $\ell_i = (p_i, p_{i+1})$ is the i-th line segment consisting of a starting point p_i and an end point p_{i+1}, and T stands for the number of line segments of the polygonal track p, see Figure 2. Thus, single fibres will be described by a random polygonal track model. It consists of a random (initial) line segment ℓ_1 and a multivariate time-series model to describe the successive line segments $\ell_2, \ldots, \ell_T$, where the spatial correlation between consecutive line segments is taken into account by the time series. When fitting this stochastic fibre model, based on time series, to experimentally measured 3D data, the following challenge occurs: Due to irregularities in the 3D images, e.g. noise, only short systems of line segments can be detected, see Figure 1 (right). Thus, the data basis for estimation of parameters consists of many independently sampled (short) realizations of the time series model instead of one long realization. This makes a discussion on parameter estimation inevitable.

We therefore derive ML estimators for a given class of multivariate time series models (explicitly autoregressive processes) in the context where the observations consist of many independently sampled (short) trajectories of a given model (see Section 3). Subsequently, desirable asymptotic properties of these estimates are demonstrated (see Section 4). To check if the derived parameter estimators are suitable with regard to stochastic modelling of single fibres, we further analyse properties of the estimates by means of numerical experiments which are important in this context (see Section 5).

Note that there exist several possible representations of polygonal tracks, for example: (i) by the starting and end points $(p_1, p_2, \ldots)$ of the line segments, or (ii) by the starting line segment ℓ_1 and a sequence $(\alpha_1, \beta_1, |\ell_2|)^\top, (\alpha_2, \beta_2, |\ell_3|)^\top, \ldots$, where $|\ell_i|$ is the length of the i-th segment, and α_i (resp. β_i) denotes the change of direction from the i-th to the $(i+1)$-th line segment of the polygonal track

with respect to the azimuthal (resp. polar) angle, see also Figure 2.

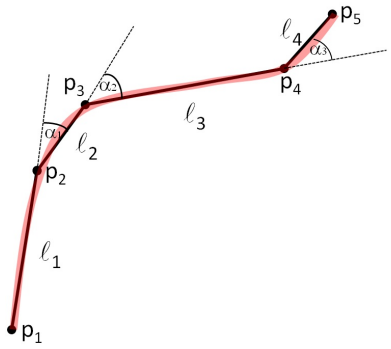

Figure 2: Planar fibre (red) approximated by a 2D polygonal track

In the following we will consider the latter (so-called incremental) representation of polygonal tracks. We thus model curved fibres in 3D by a random line segment ℓ_1 and a sequence of random vectors $(\alpha_1, \beta_1, |\ell_2|), (\alpha_2, \beta_2, |\ell_3|), \ldots$ representing the change of directions and lengths of line segments. This sequence of random vectors is described by a vectorial autoregressive process $\{\boldsymbol{Y}_i,\, i \in \mathbb{Z}\}$ of some order $q \geq 1$ (see e.g Lütkepohl (2006, p. 13)), which is given by

$$\boldsymbol{Y}_i - \boldsymbol{\mu} = \boldsymbol{A}_1\,(\boldsymbol{Y}_{i-1} - \boldsymbol{\mu}) + \ldots + \boldsymbol{A}_q\,(\boldsymbol{Y}_{i-q} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_i \qquad \text{for each } i \in \mathbb{Z}, \quad (2.1)$$

where $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$ is the set of integers, $\boldsymbol{\mu} = \mathbb{E}\boldsymbol{Y}_i \in \mathbb{R}^3$ denotes the process mean, and $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_q \in \mathbb{R}^{3\times 3}$ are coefficient matrices. The 'errors' $\{\boldsymbol{\varepsilon}_i\,, i \in \mathbb{Z}\}$ are assumed to form a sequence of three-dimensional random vectors which are independent and identically distributed with vanishing mean vector $\mathbb{E}\,\boldsymbol{\varepsilon}_i = \boldsymbol{o}$ and some (non-singular) covariance matrix $\boldsymbol{\Sigma} = \mathbb{E}\,(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top)$.

The usage of multivariate time series as a modelling approach for single fibres, allows the inclusion of (cross)correlations of angles and lengths of consecutive line segments. This is essential in order to describe a large variety of differently shaped 3D fibres. Note that the first line segment ℓ_1 can be chosen in either a random or a deterministic manner, see Gaiselmann *et al.* (2013).

3 Statistical inference

In this section, we derive ML estimators for multivariate autoregressive processes when many independently sampled (short) trajectories are observed. Furthermore, the estimation of the process order q for autoregressive processes $\{\boldsymbol{Y}_i,\, i \in \mathbb{Z}\}$ is discussed.

The parameters of the model consists of $(q, \boldsymbol{\mu}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_q, \boldsymbol{\Sigma})$, where $q > 0$, $\boldsymbol{\mu} \in \mathbb{R}^3$, $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_q \in \mathbb{R}^{3\times 3}$, and $\boldsymbol{\Sigma} \in \mathbb{R}^{3\times 3}$, see (2.1).

Additionally, note that statistical inference in case of observing many short (multivariate) time series is not restricted to single-fibre modelling. It can be

applied to other settings involving a similar type of data such as occur, e.g. in longitudinal problems, DNA analysis, financial modelling, etc. Furthermore, in the following we always consider multivariate time series in 3D, but the same methodology can be used for an arbitrary dimension.

3.1 Data basis for model fitting

The aim of stochastic modelling of single fibres is to ensure that realizations sampled from the model adequately represent (real) fibres which are observed, e.g. in experimental image data of fibre-based materials, see Figure 1.

The first step towards stochastic modelling of single fibres is extracting representations of these fibres from image data Gaiselmann *et al.* (2012) and subsequently post-processing the representations i.e. approximating each extracted fibre representation by a polygonal track represented by its starting line segment, sequences of azimuthal and polar angles describing the change of direction of consecutive line segments, and a sequence of segment lengths, see Figure 2. These sequences of angles and lengths of line segments constitute the data basis for model fitting; they are considered to be realizations of a multivariate autoregressive process. Mathematically speaking, we consider a large number K of (short) independently sampled trajectories $\{\boldsymbol{y}_{1,-q+1},\dots,\boldsymbol{y}_{1T_1}\}$, ..., $\{\boldsymbol{y}_{K,-q+1},\dots,\boldsymbol{y}_{KT_K}\}$ of the autoregressive process $\{\boldsymbol{Y}_i,\, i\in\mathbb{Z}\}$. This stands in strong contrast to classical time series theory, where parameters are estimated given a (long) single trajectory of a multivariate time series, see Lütkepohl (2006, chap. 3). In the present context we deal with a (large) number of (short and independent) time series, that make the classical maximum likelihood (ML) estimators not directly applicable. Thus, we adopt the standard ML estimator to our specific data situation.

3.2 Estimation of model parameters

ML estimators for autoregressive processes when observing many independently sampled (short) trajectories can be derived in the following way. We assume for simplicity that the errors $\boldsymbol{\varepsilon}_i$ in (2.1) are normally distributed with vanishing mean vector and some (non-singular) covariance matrix $\boldsymbol{\Sigma}$, i.e., $\boldsymbol{\varepsilon}_i \sim N(\boldsymbol{o},\boldsymbol{\Sigma})$. This implies that also the $3k$-dimensional random vector $(\boldsymbol{Y}_1,\dots,\boldsymbol{Y}_k)$ given in (2.1) is normally distributed for each $k\geq 1$. Note that in principle it is possible to consider errors following some other probability distribution where the analytical form of the probability density function is known. Moreover, we assume that the lengths of the trajectories T_n are independent and identically distributed random variables, i.e., the T_n are independently sampled from a parametric family of length distributions $\{F(\cdot,\upsilon),\upsilon\in\mathbb{R}\}$.

Let $\boldsymbol{a} = vec(\boldsymbol{A})$ with $\boldsymbol{A} = (\boldsymbol{A}_1,\dots,\boldsymbol{A}_q)$, where the *vec* operator stacks together the columns of a matrix. The estimation of the process order q is accomplished by means of the Akaike information criterion which is discussed in Section 3.3 in detail. Under the assumption of a normal distribution for the $\boldsymbol{\varepsilon}_i$, the log-likelihood function can be easily determined. More precisely,

we consider the (conditional) log-likelihood function under the condition that the lengths $(T_1, \ldots, T_K)$ of individual time series independently sampled from $\{F(\cdot, \upsilon), \upsilon \in \mathbb{R}\}$ are given. The conditional log-likelihood function depends on the time series $\{\boldsymbol{Y}_{1,-q+1}, \ldots, \boldsymbol{Y}_{1T_1}\}$, ..., $\{\boldsymbol{Y}_{K,-q+1}, \ldots, \boldsymbol{Y}_{KT_K}\}$ as well as on the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{a}, \boldsymbol{\Sigma})$ (see e.g. Lütkepohl (2006, pp. 88-89)) and is given by

$$\begin{aligned} & \log\left(L(\boldsymbol{Y}_{11}, \ldots, \boldsymbol{Y}_{KT_K}; \boldsymbol{\theta})\right) = \log\left(\prod_{n=1}^{K} f_{\boldsymbol{Y}_{n1}, \ldots, \boldsymbol{Y}_{nT_n}}(\boldsymbol{\theta})\right) \\ = & \sum_{n=1}^{K}\left[-\frac{3T_n}{2}\log(2\pi) - \frac{T_n}{2}\log\left(\det(\boldsymbol{\Sigma})\right)\right] - \frac{1}{2}\sum_{i=1}^{T_n}\left[\boldsymbol{Y}_{ni} - \boldsymbol{\mu} - \sum_{j=1}^{q}\boldsymbol{A}_j(\boldsymbol{Y}_{n,i-j} - \boldsymbol{\mu})\right]^{\top} \\ & \boldsymbol{\Sigma}^{-1}\left[\boldsymbol{Y}_{ni} - \boldsymbol{\mu} - \sum_{j=1}^{q}\boldsymbol{A}_j(\boldsymbol{Y}_{n,i-j} - \boldsymbol{\mu})\right], \end{aligned}$$

where $f_{\boldsymbol{Y}_{n1}, \ldots, \boldsymbol{Y}_{nT_n}}$ denotes the density of the random vector $(\boldsymbol{Y}_{n1}, \ldots, \boldsymbol{Y}_{nT_n})$. The product form of the (conditional) likelihood function follows from the independence assumption of the sampled time series. The solution $\widetilde{\boldsymbol{\theta}} = (\widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{a}}, \widetilde{\boldsymbol{\Sigma}})$ of the likelihood equation

$$\frac{d}{d\boldsymbol{\theta}}\log\left(L(\boldsymbol{Y}_{11}, \ldots, \boldsymbol{Y}_{KT_K}; \boldsymbol{\theta})\right) = 0$$

is given by

$$\widetilde{\boldsymbol{\mu}} = \left(\sum_{n=1}^{K} T_n\right)^{-1}\left(\boldsymbol{I}_3 - \sum_{j=1}^{q}\widetilde{\boldsymbol{A}}_j\right)^{-1}\sum_{n=1}^{K}\sum_{i=1}^{T_n}\left(\boldsymbol{Y}_{ni} - \sum_{j=1}^{q}\widetilde{\boldsymbol{A}}_j\boldsymbol{Y}_{n,i-j}\right) \tag{3.1}$$

$$\widetilde{\boldsymbol{a}} = \left(\sum_{n=1}^{K}\boldsymbol{X}_n\boldsymbol{X}_n^{\top}\right)^{-1}\left(\sum_{n=1}^{K}(\boldsymbol{X}_n \otimes \boldsymbol{I}_3)\left((\boldsymbol{Y}_{n1} - \widetilde{\boldsymbol{\mu}})^{\top}, \ldots, (\boldsymbol{Y}_{nT_n} - \widetilde{\boldsymbol{\mu}})^{\top}\right)^{\top}\right) \tag{3.2}$$

$$\widetilde{\boldsymbol{\Sigma}} = \left(\sum_{n=1}^{K} T_n\right)^{-1}\sum_{n=1}^{K}\left(\boldsymbol{X}_n^0 - (\widetilde{\boldsymbol{A}}_1, \ldots, \widetilde{\boldsymbol{A}}_q)\boldsymbol{X}_n\right)\left(\boldsymbol{X}_n^0 - (\widetilde{\boldsymbol{A}}_1, \ldots, \widetilde{\boldsymbol{A}}_q)\boldsymbol{X}_n\right)^{\top} \tag{3.3}$$

where $\boldsymbol{X}_{ni} = \left((\boldsymbol{Y}_{ni} - \widetilde{\boldsymbol{\mu}})^{\top}, \ldots, (\boldsymbol{Y}_{n,i-q+1} - \widetilde{\boldsymbol{\mu}})^{\top}\right)^{\top}$, $\boldsymbol{X}_n = \left(\boldsymbol{X}_{n0}, \ldots, \boldsymbol{X}_{n,T_n-1}\right)$ and $\boldsymbol{X}_n^0 = \left(\boldsymbol{Y}_{n1} - \widetilde{\boldsymbol{\mu}}, \ldots, \boldsymbol{Y}_{nT_n} - \widetilde{\boldsymbol{\mu}}\right)$. In the foregoing, $\otimes$ denotes the Kronecker product and $\boldsymbol{I}_3$ the 3×3 unit matrix. Note that we consider the conditional likelihood function and treat the lengths of individual time series T_n as given numbers. If we consider the (unconditional) likelihood function by assuming that the T_n are sampled from $\{F(\cdot, \upsilon), \upsilon \in \mathbb{R}\}$, the likelihood function has the form of a mixture and the estimators must be determined numerically. Hence, for simplicity, we choose to work with the conditional likelihood function.

Remark 1. It can be shown that the least squares estimator for $\boldsymbol{a}$ is equal to $\widetilde{\boldsymbol{a}}$, i.e. in particular the estimation of the parameter $\boldsymbol{a}$ does not depend on the normality assumption of the errors $\boldsymbol{\varepsilon}_i$.

Example 1. Let us discuss the connection between the estimators given in (3.1)-(3.3) and estimators considered in classical time series analysis, when we typically observe one single (long) time series. For simplicity, we will compute the estimators given in (3.1)-(3.3) for a 1D autoregressive model of order $q = 1$.

Regarding the 1D case, we obtain from (3.2) that

$$\widetilde{a} = \frac{\sum_{n=1}^{K} \sum_{i=0}^{T_n-1} X_{ni} X_{n,i+1}}{\sum_{n=1}^{K} \sum_{i=0}^{T_n-1} X_{ni}^2}. \tag{3.4}$$

On the other hand, the estimator $\widehat{a}$ for a in case of one single (long) time series of length T_1 is known to be equal to

$$\widehat{a} = \frac{\sum_{i=0}^{T_1-1} X_{1i} X_{1,i+1}}{\sum_{i=0}^{T_1-1} X_{1i}^2}$$

and thus the estimator $\widetilde{a}$ considered in (3.4) is a natural extension of $\widehat{a}$. As an estimator for Σ, we get from (3.3) that

$$\widetilde{\Sigma} = \left(\sum_{n=1}^{K} T_n \right)^{-1} \sum_{n=1}^{K} \sum_{i=1}^{T_n-1} \left(X_{n,i+1} - \widetilde{A}_1 X_{ni} \right)^2 , \tag{3.5}$$

whereas the estimator $\widehat{\Sigma}$ for Σ in case of one single (long) time series of length T_1 is known to be equal to

$$\widehat{\Sigma} = \frac{1}{T_1} \sum_{i=0}^{T_1-1} \left(X_{1,i+1} - \widetilde{A}_1 X_{1i} \right)^2 .$$

Finally, the estimator $\widehat{\mu}$ for μ in case of one single (long) time series of length T_1 is known to be equal to

$$\widehat{\mu} = \frac{1}{T_1} \left(I_3 - \sum_{j=1}^{q} \widetilde{A}_j \right)^{-1} \sum_{i=0}^{T_1-1} \left(Y_{1i} - \sum_{j=1}^{q} \widetilde{A}_j Y_{1,i-j} \right) .$$

3.3 Estimation of process order

So far we have derived ML estimators for autoregressive processes for our specific data basis given a fixed process order q. This section deals with the estimation of the order of an autoregressive model q, where we consider the Akaike information criterion (AIC) which is a widely accepted method in time series analysis. By means of this method, one computes the maximum of a penalized log-likelihood function for a set of $q \in \{0, \dots, Q\}$ with some $Q \geq 1$ and one chooses the order that gives the largest likelihood. The AIC is given by $AIC(q, \widetilde{\boldsymbol{\theta}}) = P - 2 \log \left(L(\boldsymbol{Y}_{11}, \dots, \boldsymbol{Y}_{KT_K}; \widetilde{\boldsymbol{\theta}}) \right)$, where P is the number of parameters in the $AR(q)$ model, and $\tilde{\boldsymbol{\theta}}$ is the value of the parameter vector that

maximises the log likelihood function $\log L$ of the fitted model. The AIC is most effective if the potential values for q are relatively small. For the applications of our fibre model in Gaiselmann *et al.* (2012, 2013), the underlying data sets consist of fibres which are mainly represented by short time series ($T_n \in \{3, \dots, 30\}$). Thus, at least in case of single-fibre modelling it is reasonable to apply the AIC. In practice we compute $AIC(0, \widetilde{\boldsymbol{\theta}}), \dots, AIC(Q, \widetilde{\boldsymbol{\theta}})$ and choose $\widetilde{q} = arg\min_{q \in \{0,\dots,Q\}} AIC(q, \widetilde{\boldsymbol{\theta}})$ for some $Q \geq 1$. Note that typically the AIC overestimates the model order. Thus we take the maximum order Q to be a small number, explicitly $Q = 5$. An alternative to the AIC could be to undertake model estimation for $q \in \{1, 2, 3, 4, 5\}$ and choose the minimum order leading to satisfying results, e.g. with respect to residual analysis or comparison of structural characteristics of simulated and extracted fibres.

4 Asymptotic properties

Note that in classical time series analysis (i.e. observing one long observation) the asymptotic investigations concentrate on the case that the length of an individual time series tends to infinity. In Section 2 we described the stochastic model for single 3D fibres based on the autoregressive process. In the context of experimentally measured image data, many short observations of time series are available rather than one long observation. Hence, this paper deals with the asymptotic properties of the parameter estimates as the number of observed time series K tends to infinity.

In this section we briefly discuss the asymptotic normality of the (conditional) ML estimators given in (3.1)-(3.3) when the number of time series observations K tends to infinity.

We consider K independent copies $\{\boldsymbol{Y}_{1,-q+1}, \dots, \boldsymbol{Y}_{1T_1}\}, \dots, \{\boldsymbol{Y}_{K,-q+1}, \dots, \boldsymbol{Y}_{KT_K}\}$ of an autoregressive process $\{\boldsymbol{Y}_i,\, i \in \mathbb{Z}\}$ with order $0 \leq q < \infty$. We assume the 'errors' $\boldsymbol{\varepsilon}_i$ affecting the $\{\boldsymbol{Y}_i,\, i \in \mathbb{Z}\}$ to be Gaussian, i.e., we consider independent copies of a Gaussian autoregressive process $\{\boldsymbol{Y}_i,\, i \in \mathbb{Z}\}$.

In the following, we slightly modify the representation of the parameter vector $\boldsymbol{\theta}$. We consider $\boldsymbol{\delta} = (\boldsymbol{\mu}^\top, \boldsymbol{a}^\top, \boldsymbol{\sigma}^\top)^\top = vec(\boldsymbol{\theta})$ instead of $\boldsymbol{\theta}$ as the parameter vector on which the (conditional) likelihood function of Section 3.2 depends on, where $\boldsymbol{\sigma} = vech(\boldsymbol{\Sigma})$ and $vech$ is a column stacking operator that stacks only the elements on and below the main diagonal of a matrix, see Lütkepohl (2006). Considering $\boldsymbol{\delta}$ instead of $\boldsymbol{\theta}$ has the advantage that we collect only the potentially different elements of $\boldsymbol{\Sigma}$ since the covariance matrix $\boldsymbol{\Sigma}$ is symmetric. We let $\widetilde{\boldsymbol{\delta}} = (\widetilde{\boldsymbol{\mu}}^\top, \widetilde{\boldsymbol{a}}^\top, \widetilde{\boldsymbol{\sigma}}^\top)^\top$ be the (conditional) ML estimator for $\boldsymbol{\delta} = (\boldsymbol{\mu}^\top, \boldsymbol{a}^\top, \boldsymbol{\sigma}^\top)^\top$ given in (3.1) – (3.3).

In what follows we assume that the limit

$$\gamma = \lim_{K \to \infty} \frac{\sum_{n=1}^{K} T_n}{K} \in (0, \infty)$$

exists. The quantity γ can be interpreted as the average length of a random trajectory $\{\boldsymbol{Y}_{n1}, \dots, \boldsymbol{Y}_{nT_n}\}$. Finally, in order to get consistent estimators, we

have to assume that the number of trajectories $\{\boldsymbol{Y}_{i1},\ldots,\boldsymbol{Y}_{iT_i}\}$ with $T_i > q$ tends to infinity as the total number K of observed trajectories tends to infinity, i.e., we assume that

$$\mathbb{P}(\lim_{K\to\infty} \#\{\{\boldsymbol{Y}_{i1},\ldots,\boldsymbol{Y}_{iT_i}\} : i=1,\ldots,K,\ T_i > q\} = \infty) = 1.$$

Under all these assumptions it holds that

$$\sqrt{K}\begin{pmatrix} \widetilde{\boldsymbol{\mu}}-\boldsymbol{\mu} \\ \widetilde{\boldsymbol{a}}-\boldsymbol{a} \\ \widetilde{\boldsymbol{\sigma}}-\boldsymbol{\sigma} \end{pmatrix} \xrightarrow{D} N\left(0, \begin{pmatrix} \boldsymbol{\Sigma}_{\widetilde{\boldsymbol{\mu}}}, 0, 0 \\ 0, \boldsymbol{\Sigma}_{\widetilde{\boldsymbol{\alpha}}}, 0 \\ 0, 0, \boldsymbol{\Sigma}_{\widetilde{\boldsymbol{\sigma}}} \end{pmatrix}\right) \quad \text{as } K\to\infty\ , \tag{4.1}$$

where

$$\begin{aligned} \boldsymbol{\Sigma}_{\widetilde{\boldsymbol{\mu}}} &= \frac{1}{\gamma}\left(\boldsymbol{I}_3 - \sum_{j=1}^{q}\boldsymbol{A}_j\right)^{-1}\boldsymbol{\Sigma}\left(\boldsymbol{I}_3 - \sum_{j=1}^{q}\boldsymbol{A}_j^{\top}\right)^{-1}\ , \\ \boldsymbol{\Sigma}_{\widetilde{\boldsymbol{a}}} &= \left(\frac{\boldsymbol{\Gamma}_{\boldsymbol{Y}}(0)^{-1}}{\gamma}\otimes\boldsymbol{\Sigma}\right)\ , \\ \boldsymbol{\Sigma}_{\widetilde{\boldsymbol{\sigma}}} &= \frac{1}{\gamma}\boldsymbol{D}_3^{+}\left(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}\right)\boldsymbol{D}_3^{+\top}\ . \end{aligned}$$

In the foregoing $\boldsymbol{\Gamma}_{\boldsymbol{Y}}(h) = \mathbb{E}\left((\boldsymbol{Y}_i-\boldsymbol{\mu})(\boldsymbol{Y}_{i-h}-\boldsymbol{\mu})^{\top}\right)$ with

$$\boldsymbol{\Gamma}_{\boldsymbol{Y}}(0) = \begin{pmatrix} \boldsymbol{\Gamma}_{\boldsymbol{Y}}(0) & \ldots & \boldsymbol{\Gamma}_{\boldsymbol{Y}}(q-1) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Gamma}_{\boldsymbol{Y}}(-q+1) & \ldots & \boldsymbol{\Gamma}_{\boldsymbol{Y}}(0) \end{pmatrix}.$$

$\boldsymbol{D}_3$ is the uniquely determined solution of the linear equation system $vec(\boldsymbol{\Sigma}) = \boldsymbol{D}_3\boldsymbol{\sigma}$ and $\boldsymbol{D}_3^{+}$ is the Moore-Penrose generalized inverse of $\boldsymbol{D}_3$ (see e.g. Ben-Israel *et al.* (2003)).

An outline of the proof of the asymptotic normality (4.1) of the ML estimators (3.1)-(3.3) is given in the Appendix. The proof is standard, however, we provide some details in order to deduce the diagonal form of the limiting covariance matrix.

In the following, some consequences of this asymptotic behaviour of the ML estimators given in (3.1)-(3.3) are discussed. The lengths of trajectories T_n appear only via the average length γ in the limiting distribution of the (conditional) ML estimator for $\boldsymbol{\delta}$ under the condition that $T_1,\ldots,T_K$ are given. Due to the strong law of large numbers and the independent sampling of $T_1,\ldots,T_K$, we have $\gamma = \mathbb{E}T_n$ almost surely. Therefore the limiting distribution of the (conditional) ML estimator for $\boldsymbol{\delta}$ is almost surely the same for all given $T_1,\ldots,T_K$. In other words, the limiting distribution does not depend on the condition that the numbers $T_1,\ldots,T_K$ are given.

From formula (4.1) it follows directly that the (conditional) ML estimators given in (3.1)-(3.3) are weakly consistent. Since the covariance matrix in (4.1)

exhibits the specified diagonal form, we can conclude that the estimators $\widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{a}}, \widetilde{\boldsymbol{\Sigma}}$ are asymptotically pairwise independent.

Furthermore, regarding the entries of the asymptotic covariance matrix displayed in (4.1), we can conclude that they are influenced by the total length of the processes $M = \sum_{n=1}^{K} T_n$, i.e., the larger M is the smaller are the entries of the asymptotic covariance matrix. Thus formula (4.1) indicates that it is valid to use all trajectories for parameter estimation, regardless of their lengths T_n.

In practical time series analysis the ML estimator for the process mean $\boldsymbol{\mu}$ is often replaced by the sample mean of all observed data, i.e., the estimator $\widetilde{\boldsymbol{\mu}}$ given in (3.1) is replaced by

$$\bar{\boldsymbol{Y}} = \frac{1}{M} \sum_{n=1}^{K} \sum_{i=1}^{T_n} \boldsymbol{Y}_{ni}$$

everywhere in (3.1)-(3.3). This has the advantage that $\bar{\boldsymbol{Y}}, \widetilde{\boldsymbol{a}}$ and $\widetilde{\boldsymbol{\Sigma}}$ can be computed easily without solving the complex system of equations (3.1)-(3.3).

If we consider parameter estimation for a single long time series we find that the asymptotic distributions of $\sqrt{K}(\widetilde{\boldsymbol{\mu}} - \boldsymbol{\mu})$ and $\sqrt{K}(\bar{\boldsymbol{Y}} - \boldsymbol{\mu})$ as $K \to \infty$ are the same and thus for one long time series $\widetilde{\boldsymbol{\mu}}$ and $\bar{\boldsymbol{Y}}$ can be exchanged without changing the goodness of parameter estimation. In the present situation, since the $\boldsymbol{Y}_{ij}$ are normal, we can easily conclude that

$$\sqrt{K}(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}) \xrightarrow{D} N(\boldsymbol{o}, \boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}}) \ , \text{ for } K \to \infty,$$

where $\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}} = \left[\mathrm{Cov}\left(\bar{\boldsymbol{Y}}_i, \bar{\boldsymbol{Y}}_j\right)\right]$ and $\bar{\boldsymbol{Y}} = \left(\bar{\boldsymbol{Y}}_1, \bar{\boldsymbol{Y}}_2, \bar{\boldsymbol{Y}}_3\right)^\top$. Unfortunately, it is difficult to compute the covariance matrix $\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}}$ whence we cannot compare the limiting distributions of $\sqrt{K}(\widetilde{\boldsymbol{\mu}} - \boldsymbol{\mu})$ and $\sqrt{K}(\bar{\boldsymbol{Y}} - \boldsymbol{\mu})$. However, the numerical results discussed in Section 5 indicate that the estimator $\bar{\boldsymbol{Y}}$ has approximately the same asymptotic properties as $\widetilde{\boldsymbol{\mu}}$.

Note that the asymptotic normality of the (conditional) ML estimator $\widetilde{\boldsymbol{\delta}}$ as described in Section 4 is useful for constructing asymptotic confidence intervals. However, the covariance matrix of the asymptotic normal distribution, see (4.1), has a complex form. Thus, it is preferable to use the classical bootstrap to derive an estimator of the covariance matrix. By means of the bootstrap approach we are able to compute confidence intervals using the estimates given in (3.1)-(3.3). Note that for our data structure (which consists of many independent short trajectories) using the classical bootstrap means that we independently resample with replacement whole trajectories from the complete set of trajectories in each instance of forming a bootstrap sample.

5 Simulation Study

In this section, we will numerically investigate further properties of the ML estimators given in (3.1)-(3.3). Due to the complex forms of these estimators, it is difficult to investigate their properties analytically. We consider a Gaussian

$AR(1)$-model with predefined parameters and analyse the goodness of parameter estimation for this specific setting. In particular we investigate numerically the following issues:

- The parameter estimation of $\{\boldsymbol{Y}_i,\, i \in \mathbb{Z}\}$ when observing many (short) multivariate time series can be accomplished in two different ways: The parameters can be estimated by solving the equation system given in (3.1)-(3.3) or by replacing $\widetilde{\boldsymbol{\mu}}$ in (3.1)-(3.3) by $\bar{\boldsymbol{Y}}$. We compare these two different methods.
- From (4.1), the weak consistency of the ML estimates is obtained, i.e., for a sufficiently large number of trajectories, the ML estimators will be arbitrarily close to the correct values with a high probability. We therefore analyse the quality of parameter estimation for small samples. This also reflects realistic settings since only a relatively small number of trajectories is usually available from experimental data.
- We check numerically if classical bootstrap is applicable for the ML estimators given in (3.1)-(3.3).
- We compare the asymptotic behaviour of the least squares estimator $\bar{\boldsymbol{a}}$ for $\boldsymbol{a}$ when the errors of the autoregressive process are Gaussian with the asymptotic behaviour when the errors are non-Gaussian. This comparison serves to check whether the fibre model is appropriate when the angle-length representations are non-Gaussian.
- We check whether the method of parameter estimation proposed in this paper is adequate for modelling 3D fibres, i.e., we test the sensitivity of geometric properties of simulated fibres with respect to the parameters of the $AR(q)$ model. This is an important aspect since it is not clear a priori how accurately the parameters have to be estimated in order to model 3D fibres with geometric properties similar to those of the observed fibres.

In our simulation study, we consider the following $AR(q)$ model, where we put $q = 1$ and set

$$\boldsymbol{A}_1 = \begin{pmatrix} -0.5 & 0.4 & 0.1 \\ -0.2 & 1.0 & -0.5 \\ 1.2 & 0.8 & -0.7 \end{pmatrix},\ \boldsymbol{\Sigma} = \begin{pmatrix} 0.5 & -0.1 & 0.1 \\ -0.1 & 0.5 & -0.1 \\ 0.1 & -0.1 & 5 \end{pmatrix},\ \boldsymbol{\mu} = \begin{pmatrix} 0.5 \\ 0.3 \\ 1.3 \end{pmatrix}. \quad (5.1)$$

Although this autoregressive process $\{\boldsymbol{Y}_i,\, i \in \mathbb{Z}\}$ is described by 21 parameters, we will only provide results and plots for three of them, namely $\boldsymbol{A}_1(1,1), \boldsymbol{\Sigma}(1,3)$ and $\boldsymbol{\mu}(2)$. To ensure comparability with the experimental data considered in Sections 2 and 6, we independently sample T_n from F, i.e., $T_n \sim F$, where F is a discrete probability distribution of the form $\sum_{i=q+1}^{W} r_i \delta_i$, W is the maximal length, r_i is a relative frequency associated with length i, and δ_i is the Dirac function. The relative frequencies r_i and the maximal length W are chosen according to the histogram of Figure 8 displaying the length of trajectories which corresponds to fibres observed in experimental 3D image data. Note that our simulation study supports the asymptotic theory considered in Section 4.

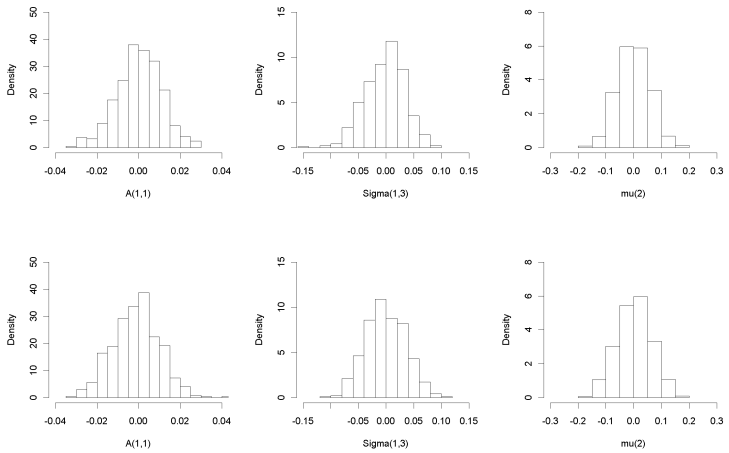

Figure 3: Histograms of estimated parameters for 1000 repetitions for $\widetilde{\boldsymbol{A}}_1(1,1) - \boldsymbol{A}_1(1,1)$, $\widetilde{\boldsymbol{\Sigma}}(1,3) - \boldsymbol{\Sigma}(1,3)$ and $\widetilde{\boldsymbol{\mu}}(2) - \boldsymbol{\mu}(2)$ (resp. $\bar{\boldsymbol{Y}} - \boldsymbol{\mu}(2)$) based on solving the equation system given in (3.1)-(3.3) (first row) and based on $\bar{\boldsymbol{Y}}$ (second row)

5.1 Comparison of mean estimators

We investigate the influence of the two different methods of estimating $\boldsymbol{\mu}$ on the estimates of $\boldsymbol{a}$ and $\boldsymbol{\Sigma}$. We simulate a set of time series of size $K = 1000$ with individual lengths $T_n \sim F$ for each $n \in \{1, \ldots, K\}$. This simulation framework is repeated 1000 times and for each set of time series we estimate the parameters of the $AR(1)$ model using the two different methods. In Figure 3, the histograms of the estimated and centred parameters are plotted. In the first row the estimators are based on solving the equation system and in the second row the estimators are based on $\bar{\boldsymbol{Y}}$. We observe that the histograms for both estimators are quite similar. Thus, we can conclude that the asymptotic distributions are almost the same for the two estimation methods. Since estimation based on $\bar{\boldsymbol{Y}}$ is more time-efficient, it is used in the following.

5.2 Small samples investigations

Formula (4.1) states that the limiting distributions of the ML estimators for $\boldsymbol{\delta}$ given in (3.1)-(3.3) are normal, where the asymptotic covariance matrix of the estimators depends on $M = \sum_{n=1}^{K} T_n$. However, it is not a priori clear how large the total length of observations M has to be in order to deliver adequate estimates of the parameters. Thus, we repeat 1000 times the estimation of parameters for a set of simulated time series with varying M, where we choose $M = 100$, $M = 1000$, $M = 5000$, $M = 10000$ and $T_n \sim F$. The mean values and standard deviations of the centred parameters, i.e., $\widetilde{\boldsymbol{A}}_1(1,1) - \boldsymbol{A}_1(1,1)$, $\widetilde{\boldsymbol{\Sigma}}(1,3) - \boldsymbol{\Sigma}(1,3)$ and $\bar{\boldsymbol{Y}}(2) - \boldsymbol{\mu}(2)$ for the different values of M are listed in Table 1, where

Table 1: Mean values (standard deviations) of the centred parameters computed for repeated parameter estimations with different total length of observations M

	$\bar{\boldsymbol{Y}}(2) - \boldsymbol{\mu}(2)$	$\widetilde{\boldsymbol{A}}_1(1,1) - \boldsymbol{A}_1(1,1)$	$\widetilde{\boldsymbol{\Sigma}}(1,3) - \boldsymbol{\Sigma}(1,3)$
$M = 100$	-6.8 $\times 10^{-3}$ (0.252)	-2.4 $\times 10^{-3}$ (0.046)	-3.3 $\times 10^{-2}$ (0.1541)
$M = 1000$	-5.1 $\times 10^{-3}$ (0.082)	3.2 $\times 10^{-4}$ (0.015)	-3.6 $\times 10^{-3}$ (0.05)
$M = 5000$	-1.9 $\times 10^{-4}$ (0.038)	6.8 $\times 10^{-4}$ (0.0066)	-8.5 $\times 10^{-4}$ (0.023)
$M = 10000$	3.4 $\times 10^{-5}$ (0.027)	3.9 $\times 10^{-5}$ (0.005)	-9.2 $\times 10^{-4}$ (0.016)
$M = 50000$	-3.6 $\times 10^{-6}$ (0.012)	-3.3 $\times 10^{-5}$ (0.002)	4.3 $\times 10^{-4}$ (0.007)

we can see clearly that the mean value and the standard deviation decrease for increasing M. The mean value is already small for $M \geq 1000$ in contrast to the standard deviation which is significantly reduced for $M \geq 10000$. Thus, a size of $M \geq 10000$ is desirable for adequate parameter estimation. In the application that we consider the total length M is ≥ 50000. So this restriction is met by the data in which we are interested.

5.3 Bootstrapping

Since the covariance matrix of the asymptotic normal distribution stated in (4.1) has a complex form, it is preferable to use the classical bootstrap in order to obtain an estimator of the covariance matrix. By means of the bootstrap approach we are able to compute, e.g. confidence intervals of the estimates. Note that in the context under consideration where there are many independent and short trajectories, use of the classical bootstrap means that we independently resample with replacement whole trajectories from the complete set of trajectories.

In this section we check numerically if bootstrapping is also applicable to the estimators $\widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{a}}$ and $\widetilde{\boldsymbol{\Sigma}}$. The applicability of bootstrapping to these estimators is supported by the good agreement of the histograms in Figure 4, where the first row of histograms are parameter estimates for a simulated set of 1000 time series with $M = 10000$ and $T_n \sim F$. In the second row of Figure 4 the histograms show the empirical bootstrap distribution of the parameter estimates for a single simulated set of time series with $M = 10000$ and $T_n \sim F$. In addition we computed 95% confidence intervals (CI) for the parameters from ML-estimators based on bootstrapping and on repeated Monte-Carlo simulations. The results are shown in Table 2. The intervals from the two approaches are almost the same. Hence the bootstrap method seems to be also applicable to the ML estimators.

5.4 Comparison of asymptotic behaviour of least squares for non-Gaussian errors

In Remark 1 it was mentioned that the least squares estimator $\bar{\boldsymbol{a}}$ of $\boldsymbol{a}$ is equal to the ML estimator $\widetilde{\boldsymbol{a}}$. Since in real-world applications it often occurs that

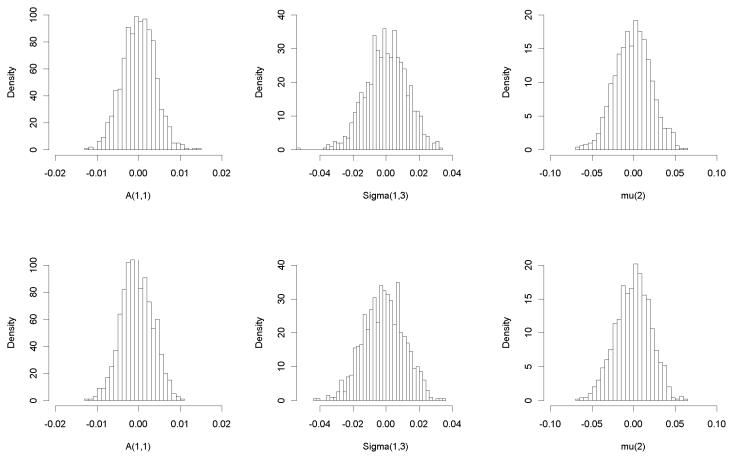

Figure 4: Histograms of centred estimates for 1000 repetitions of simulated time series with $M = 10000$ and $T_n \sim F$ by means of repeated Monte-Carlo simulations (first row) and bootstrapping (second row)

Table 2: Comparison of confidence intervals computed by repeated Monte-Carlo simulations and bootstrapping for $\boldsymbol{A}_1(1,1), \boldsymbol{\Sigma}(1,3)$ and $\boldsymbol{\mu}(2)$

	Monte-Carlo	bootstrapping
$\boldsymbol{A}_1(1,1)$	[-0.5074, -0.4924]	[-0.5078, -0.4932]
$\boldsymbol{\Sigma}(1,3)$	[0.0752, 0.1222]	[0.0718, 0.1195]
$\boldsymbol{\mu}(2)$	[0.2562, 0.3427]	[0.2559, 0.3377]

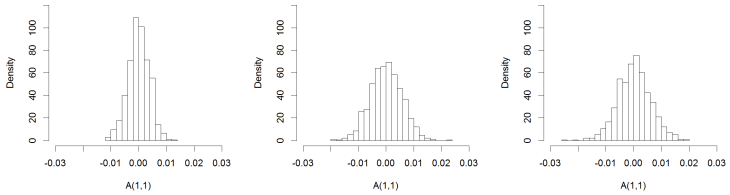

Figure 5: Histograms of estimated and centred parameter $\boldsymbol{A}_1(1,1)$ for 1000 repetitions based on the least squares estimator $\bar{\boldsymbol{a}}$ where the distribution function of the errors $\boldsymbol{\varepsilon}_n$ is taken to be $\boldsymbol{\varepsilon}_n \sim N(\boldsymbol{o}, \boldsymbol{\Sigma})$ (left), $\boldsymbol{\varepsilon}_n = (R_1^n, R_2^n, R_3^n)^\top$ with $R_i^n \sim U(-500, 500)$ (centre) and $\boldsymbol{\varepsilon}_n = (S_1^n, S_2^n, S_3^n)^\top$ with $S_i^n + 5 \sim Exp(1/5)$ (right).

the errors are not Gaussian distributed, we are interested in the asymptotic behaviour of the least squares estimator $\bar{\boldsymbol{a}}$ for non-Gaussian errors. To check this, we conducted a simulation study based on the $AR(1)$ model specified in (5.1) where the distribution function of the errors was taken to be two non-Gaussian distributions. Subsequently, the resulting limiting distributions of $\bar{\boldsymbol{a}}$ for the different distribution functions are investigated. Explicitly we simulated a set of $K = 1000$ time series with individual lengths $T_n \sim F$, $n \in \{1, \ldots, K\}$. The simulations were based on the $AR(1)$ model specified in (5.1) where we change the underlying distribution function of the errors $\boldsymbol{\varepsilon}_n$ from $\boldsymbol{\varepsilon}_n \sim N(\boldsymbol{o}, \boldsymbol{\Sigma})$ to $\boldsymbol{\varepsilon}_n = (R_1^n, R_2^n, R_3^n)^\top$ with $R_i^n \sim U(-500, 500)$ (where $U(-500, 500)$ denotes the uniform distribution on the interval $[-500, 500]$) and to $\boldsymbol{\varepsilon}_n = (S_1^n, S_2^n, S_3^n)^\top$ with $S_i^n + 5 \sim Exp(1/5)$ (where $Exp(1/5)$ denotes the exponential distribution with mean value 5). This simulation framework was repeated 1000 times for each of the three considered scenarios. Subsequently we computed for each set of time series the estimator $\bar{\boldsymbol{a}}$. In Figure 5 histograms of the values of the estimates for the first entry of $\boldsymbol{a}$ are plotted for the three different scenarios. All of the histograms look Gaussian which indicates that the least squares estimator $\bar{\boldsymbol{a}}$ is asymptotic normally distributed regardless the distributions of the errors. Moreover, we observe that the histograms for the non-Gaussian cases are almost identical. In contrast, a substantial difference between the Gaussian and non-Gaussian cases can be observed. It seems that $\bar{\boldsymbol{a}}$ has a smaller variance if the errors are Gaussian.

5.5 Curvature characteristics - Two-sample tests

The motivation for the methodology considered in the present paper i.e., the investigation of estimators for parameters of $AR(q)$ models based on a large number of short trajectories, is to adequately model 3D fibres. It is not however a priori clear that 3D fibres simulated according to an $AR(q)$ model, using estimated parameters, have 'correct' (or at least sufficiently similar) geometric

properties compared with 3D fibres simulated with exact parameters. Thus there is a need to test the sensitivity of geometric properties of simulated fibres to the values of the parameters of the $AR(q)$ model.

In this section, we provide some numerical results supporting the claim that the confidence intervals determined in Section 5.3, have a close connection with the goodness-of-fit of curvature properties of fibres generated by the fibre model. In other words, if the true parameters are within the 95% confidence interval, the geometric properties of simulated fibres from the true and estimated parameters are quite similar.

The experimental set-up is as follows: We consider sampled time series from the $AR(1)$ model given in (5.1) as incremental representations of polygonal tracks (see Section 2) and calculate two different curvature characteristics of these tracks. The direction of the starting line segment was chosen to be uniformly distributed on the unit sphere. An illustration of the first curvature characteristic is shown in Figure 6. More precisely, let

$$\beta_0(p) = \beta(p)/|\ell|^2 \ ,$$

where $\beta(p)$ denotes the area circumscribed by the polygonal track $p = (\ell_1, \ldots, \ell_n)$ and $|\ell|$ is the length of the (straight) line segment connecting the starting point of ℓ_1 and the end point of ℓ_n where ℓ_i is the i-th line segment of the polygonal track. Notice that in Figure 6, the curvature characteristic $\beta_0(p)$ is only illustrated in 2D but computed in three dimensions, where $\beta(p)$ (red area) is then the area of a curved surface.

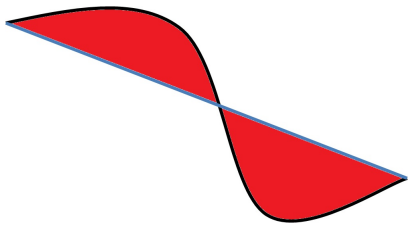

Figure 6: Curvature characteristic $\beta_0(p)$: Circumscribed area (red) $\beta(p)$ divided by the square length $|\ell|^2$ of the blue line

The second characteristic is called the tortuosity of the polygonal track p. It is defined by

$$\tau(p) = \frac{\sum_{i=1}^{n} |\ell_i|}{|\ell|}$$

and describes the 'winding tendency' of the fibres.

The issue that we need to check is whether the curvature characteristics of single fibres (polygonal tracks) generated from the $AR(1)$ model defined in (5.1) are reflected sufficiently well if the true parameters are located in the relevant 95% confidence interval.

First, we check whether the distributions of the curvature characteristics coincide for fibres generated by the $AR(1)$ process with correct and estimated

Table 3: Relative frequency of rejections with respect to Wilcoxon two-sample test for estimated parameters, parameters located in CI and parameters located outside of CI

	estimated parameters	parameters in CI	parameters outside CI
β_0	0.046	0.048	0.83
τ	0.036	0.05	0.96

parameters. We therefore perform a Wilcoxon two-sample test on the hypothesis that the distributions of β_0 (resp. τ) computed for fibres which are generated with correct and estimated parameters, are the same.

In more detail we simulated $K = 500$ trajectories (fibres) from the $AR(1)$ process given in equation (5.1) where $T_n \sim F$. We then calculated the empirical distributions F_{β_0} and F_τ of the 500 resulting values of the estimates of each of the two curvature characteristics. Next we estimated the model parameters from the observations consisting of the 500 simulated fibres. We then simulated another 500 fibres using these estimated parameters and computed the empirical distributions $F_{\widetilde{\beta}_0}$ and $F_{\widetilde{\tau}}$ of the estimates of the two curvature characteristics for this new sample. In addition the empirical distributions $F_{\beta_0^{inCI}}$ and $F_{\tau^{inCI}}$ of the two curvature characteristics were computed for 500 fibres drawn from a fibre model with parameters that were uniformly distributed on the derived 95% confidence intervals. Finally empirical distributions $F_{\beta_0^{outCI}}$ and $F_{\tau^{outCI}}$ were calculated for 500 fibres drawn from a fibre model with parameters that were uniformly distributed on sets contained in the complements of the derived confidence intervals. Explicitly if the derived CI was $[c_1, c_2]$, the parameter in question was chosen to be uniformly distributed on the set $[c_1 - |c_2 - c1|] \cup [c_2 + |c_2 - c1|]$. The distributions $(F_{\widetilde{\beta}_0}, F_{\widetilde{\tau}}), (F_{\beta_0^{inCI}}, F_{\tau^{inCI}}), (F_{\beta_0^{outCI}}, F_{\tau^{outCI}})$ were then pairwise compared by means of Wilcoxon two-sample test with F_{β_0}, F_τ, i.e., we computed the Wilcoxon two-sample test for the pairs of distributions

$$(F_{\beta_0}, F_{\widetilde{\beta}_0}),\ (F_{\beta_0}, F_{\beta_0^{inCI}}),\ (F_{\beta_0}, F_{\beta_0^{outCI}}),\ (F_\tau, F_{\widetilde{\tau}}),\ (F_\tau, F_{\tau^{inCI}}) \text{ and } (F_\tau, F_{\tau^{outCI}}).$$

The whole procedure was repeated 1000 times and we counted the number of events for which the null hypothesis of the two-sample test was rejected. In Table 3 the relative frequencies of rejections are listed. These estimated parameters and the parameters within CI lead to rejection with a probability of approximately 0.05 which is exactly the level at which the two-sample test was applied. Moreover the rejection rate for parameters outside of CI-s is very large. These results support the claim that the estimated confidence intervals of Section 5.3 bear a close relationship to the goodness-of-fit of the fibre model.

In Figure 7, histograms of the computed curvature characteristics are shown for the four scenarios of fibre simulations (real parameters, estimated parameters, parameters within CI, parameters outside of CI). We observe that the model with parameters chosen outside of CI fails to re-create correctly the distributions of curvature β_0 and tortuosity τ. In particular, the 'incorrect' model produces too few fibres with small curvature and tortuosity.

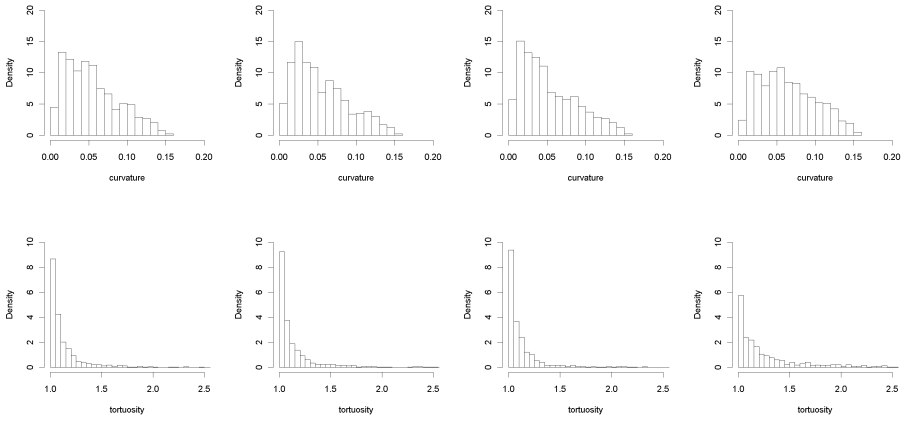

Figure 7: Curvature characteristics β_0, τ computed for the four scenarios of fibre simulations: real parameters (first column), estimated parameters (second column), parameters within CI (third column), parameters outside of CI (fourth column))

6 Application to experimental data

To show that the proposed model for single fibres and the proposed methodology for parameter estimation deliver a powerful tool for describing curved objects in 3D, we give an example of a real-world application of the single-fibre model with specific emphasis on its statistical inference. See Gaiselmann *et al.* (2013) for more background on this example. Explicitly we apply our procedure to the problem of modelling a system of 3D single fibres that exhibits the microstructure of non-woven GDL as shown in the left panel of Figure 1. The modelling procedure is based on the description given in Section 2. To estimate the parameters of the AR(q) model, we extract single fibres from the experimental image data as described in Gaiselmann *et al.* (2012). These fibres are available as 3D polygonal tracks and are subsequently transformed to a set of time series trajectories as described in Section 3.1. Unfortunately, due to irregularities in the 3D grey-scale images, only relatively short sections of fibres could be extracted. We thus ended up with a large number $K \approx 20000$ of (short) trajectories $\{\boldsymbol{y}_{1,-q+1}, \ldots, \boldsymbol{y}_{1T_1}\}$, …, $\{\boldsymbol{y}_{K,-q+1}, \ldots, \boldsymbol{y}_{KT_K}\}$. The histogram of the lengths of trajectories T_n is plotted in Figure 8, from which it can be seen that there are many short trajectories and a few longer ones. Subsequently, the order q of the $AR(q)$ process was estimated using the AIC criterion as described in Section 3.3. In this case the AIC criterion chose order $q = 2$. The parameters of the $AR(2)$ model were estimated by estimators given in (3.1)-(3.3) and we

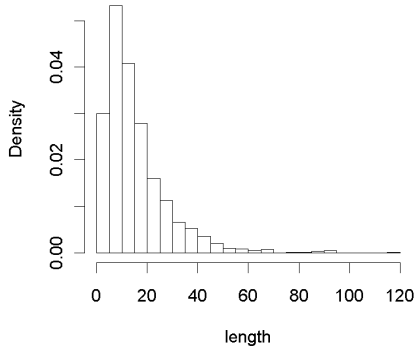

Figure 8: Histogram of the lengths of trajectories T_n

obtained

$$\boldsymbol{Y}_i - \begin{pmatrix} 0.005 \\ 0.00008 \\ 23.5 \end{pmatrix} = \begin{pmatrix} 0.214 & 0.061 & -0.0002 \\ -0.00004 & -0.091 & 0.00002 \\ -0.53 & 1.569 & 0.114 \end{pmatrix} \left(\boldsymbol{Y}_{i-1} - \begin{pmatrix} 0.005 \\ 0.00008 \\ 23.5 \end{pmatrix} \right) + \begin{pmatrix} 0.106 & 0.039 & 0.0002 \\ -0.001 & -0.11 & -0.00005 \\ -0.321 & 3.846 & 0.025 \end{pmatrix} \left(\boldsymbol{Y}_{i-2} - \begin{pmatrix} 0.005 \\ 0.00008 \\ 23.5 \end{pmatrix} \right) + \varepsilon_i \,,$$

where $\boldsymbol{\varepsilon}_i \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.08 & 0.00007 & -0.02 \\ 0.00007 & 0.0018 & -0.016 \\ -0.02 & -0.016 & 207 \end{pmatrix} \right)$.

To investigate the accuracy of the normality assumption of the presented stochastic single fibre model, we computed the residuals and examined the empirical marginal distributions and the corresponding qq-plots. The results shown in Figure 9 indicate that the residuals are not perfectly normal distributed but the normality assumption seems to be quite acceptable.

To check the goodness-of-fit we plotted the histograms of the two curvature characteristics β_0 and τ, introduced in Section 5.5, for (experimental) polygonal tracks extracted from the 3D synchrotron image and for simulated polygonal tracks drawn from the fitted single-fibre model. Figure 10 shows that these histograms coincide nicely which supports the numerical results obtained in Section 5.5.

Finally, in Figure 11 we check visually how accurately the observed fibres are described by the single-fibre model. This figure indicates good agreement.

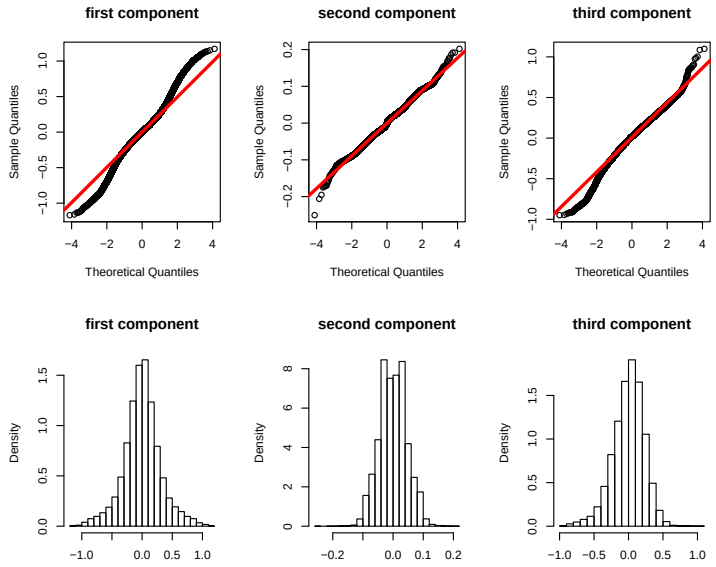

Figure 9: QQ-plots (top) and empirical marginal distributions (bottom) of the first (left), second (centre) and third (right) components of the residual vector

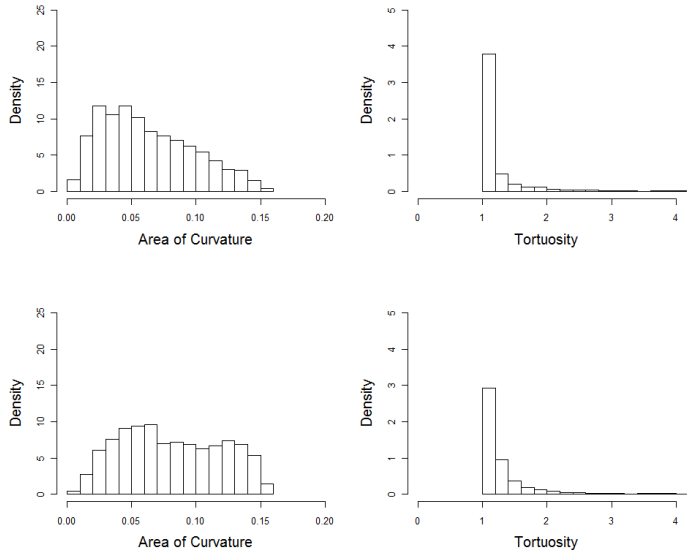

Figure 10: Histograms of $\widehat{\beta}_0$ (left) and $\widehat{\tau}$ (right) for extracted (top) and simulated (bottom) tracks

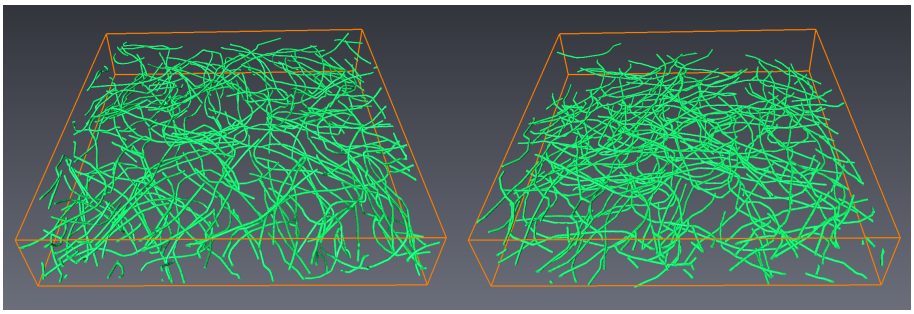

Figure 11: Observed (left) and simulated (right) single fibres.

7 Conclusions

Statistical inference for the parameters of the stochastic model describing curved fibrous objects in 3D was discussed. In particular, we statistically analysed the estimation of parameters of a stochastic single-fibre model. This model has proved its applicability by satisfactorily reproducing the appearance of single fibres in non-woven materials. The model deals with multivariate autoregressive processes, where parameter estimation is based on a large number of short and independently sampled trajectories rather than one long trajectory. Statistical properties of ML estimators were considered for autoregressive processes in this particular context. Simulation studies supported the theoretical results that were obtained. Further properties of the ML estimators which are complicated to handle analytically were investigated. In summary, the methodology proposed in this paper provides an extensive toolbox of time-series analysis methods and places them on a solid theoretical basis. We expect these methods to find wide application.

Acknowledgements

This research has been supported by the German Academic Exchange Service (DAAD) in the framework of the project 'Analysis, modelling and simulation of complex spatio-temporal data' jointly investigated by the Universities of Ulm and Ottawa. The authors are grateful to Tim Brereton for a critical reading of the manuscript.

Appendix

We discuss briefly the basic idea of the proof of the asymptotic normality (cf. (4.1)) of the ML estimator (3.1)-(3.3). In order to make a discussion clear some background from general maximum likelihood theory must also be provided. We consider the setting that was introduced in Section 3. In particular, let

$\{\boldsymbol{Y}_{1,-q+1},\ldots,\boldsymbol{Y}_{1T_1}\},\ \ldots\ ,\{\boldsymbol{Y}_{K,-q+1},\ldots,\boldsymbol{Y}_{KT_K}\}$ be random trajectories of K independent copies of a Gaussian autoregressive process $\{\boldsymbol{Y}_i,\ i\in\mathbb{Z}\}$ and let $\widetilde{\boldsymbol{\delta}}=(\widetilde{\boldsymbol{\mu}}^\top,\widetilde{\boldsymbol{a}}^\top,\widetilde{\boldsymbol{\sigma}}^\top)^\top$ be the ML estimator for $\boldsymbol{\delta}=(\boldsymbol{\mu}^\top,\boldsymbol{a}^\top,\boldsymbol{\sigma}^\top)^\top$ given in (3.1)−(3.3), etc. Under some assumptions (cf. Lütkepohl (2006) which can be seen to be fulfilled in the present setting), the following expression holds:

$$\sqrt{K}(\widetilde{\boldsymbol{\delta}}-\boldsymbol{\delta})\to N\left(\boldsymbol{o},\lim_{K\to\infty}\left(\frac{\boldsymbol{I}(\boldsymbol{\delta})}{K}\right)^{-1}\right)\ \text{as}\ K\to\infty\ . \tag{7.1}$$

In the foregoing $\boldsymbol{I}(\boldsymbol{\delta})=-\mathbb{E}\left(\frac{\partial^2\log L}{\partial\boldsymbol{\delta}\partial\boldsymbol{\delta}^\top}\right)$ denotes the Fisher information matrix. The proof of (7.1) follows directly from classical ML estimation theory and can be found, e.g. in Lütkepohl (2006, p. 693-694).

In view of (7.1) we only have to compute the limiting covariance matrix $\lim_{K\to\infty}(\frac{\boldsymbol{I}(\boldsymbol{\delta})}{K})^{-1}$. In order to do this, we make use of computations of the likelihood function and its derivatives from Lütkepohl (2006) modifying them appropriately to fit our particular context. Setting $\boldsymbol{w}=vec(\boldsymbol{\Sigma})$ we see that the second-order derivatives of the log-likelihood function can be written as follows:

$$\begin{aligned}
\frac{\partial^2\log L}{\partial\boldsymbol{a}\partial\boldsymbol{a}^\top} &= -\sum_{n=1}^{K}\left(\boldsymbol{X}_n\boldsymbol{X}_n^\top\right)\otimes\boldsymbol{\Sigma}^{-1}\,,\\
\frac{\partial^2\log L}{\partial\boldsymbol{w}\partial\boldsymbol{w}^\top} &= \sum_{n=1}^{K}\frac{T_n}{2}\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\right)-\frac{1}{2}\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_n\boldsymbol{U}_n^\top\boldsymbol{\Sigma}^{-1}\right)-\\
&\quad\ \frac{1}{2}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_n\boldsymbol{U}_n^\top\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\right),\\
\frac{\partial^2\log L}{\partial\boldsymbol{\mu}\partial\boldsymbol{\mu}^\top} &= -\sum_{n=1}^{K}T_n\left(\boldsymbol{I}_3-\sum_{j=1}^{q}\boldsymbol{A}_j\right)^\top\Sigma^{-1}\left(\boldsymbol{I}_3-\sum_{j=1}^{q}\boldsymbol{A}_j\right),\\
\frac{\partial^2\log L}{\partial\boldsymbol{\mu}\partial\boldsymbol{a}^\top} &= -\sum_{n=1}^{K}\left(\boldsymbol{I}_3-\left(\boldsymbol{l}^\top\otimes\boldsymbol{I}_3\right)\boldsymbol{A}^\top\right)\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{T_n}\left(\boldsymbol{X}_{n,i-1}^\top\otimes\boldsymbol{I}_3\right)\\
&\quad\ -\left(\sum_{i=1}^{T_n}\boldsymbol{\varepsilon}_{n,i}^\top\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{I}_3\right)\left(\boldsymbol{I}_3\otimes\boldsymbol{l}^\top\otimes\boldsymbol{I}_3\right)\frac{\partial vec(\boldsymbol{A}^\top)}{\partial\boldsymbol{a}^\top},\\
\frac{\partial^2\log L}{\partial\boldsymbol{w}\partial\boldsymbol{a}^\top} &= \sum_{n=1}^{K}-\frac{1}{2}\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\right)\left[\left(\boldsymbol{I}_3\otimes\boldsymbol{U}_n\boldsymbol{X}_n^\top\right)\frac{\partial vec\left(\boldsymbol{A}^\top\right)}{\partial\boldsymbol{a}^\top}+\left(\boldsymbol{U}_n\boldsymbol{X}_n^\top\otimes\boldsymbol{I}_3\right)\right],\\
\frac{\partial^2\log L}{\partial\boldsymbol{w}\partial\boldsymbol{\mu}^\top} &= \sum_{n=1}^{K}\frac{1}{2}\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\right)\left[\left(\boldsymbol{I}_3\otimes\boldsymbol{U}_n\right)\frac{\partial vec(\boldsymbol{U}_n^\top)}{\partial\boldsymbol{\mu}^\top}+\left(\boldsymbol{U}_n\otimes\boldsymbol{I}_3\right)\frac{\partial vec(\boldsymbol{U}_n)}{\partial\boldsymbol{\mu}^\top}\right],
\end{aligned}$$

where $\boldsymbol{l}=(1,\ldots,1)^\top$ is a vector of dimension q, $\boldsymbol{\varepsilon}_{ni}$ is the i-th error term in the n-th time series and $\boldsymbol{U}_n=(\boldsymbol{\varepsilon}_{n1},\ldots,\boldsymbol{\varepsilon}_{nT_n})$.

First we show that in this setting the information matrix is block diagonal. We have

$$\lim_{K\to\infty} -\frac{1}{K}\mathbb{E}\frac{\partial^2 \log L}{\partial \boldsymbol{\mu}\partial \boldsymbol{a}^\top} = \boldsymbol{o}\,,$$

because $\mathbb{E}\boldsymbol{X}_{n,i-1}^\top = \boldsymbol{o}$ and $\mathbb{E}\boldsymbol{\varepsilon}_{n,i}^\top = \boldsymbol{o}$. Furthermore,

$$\lim_{K\to\infty} -\frac{1}{K}\mathbb{E}\frac{\partial^2 \log L}{\partial \boldsymbol{w}\partial \boldsymbol{\mu}^\top} = \boldsymbol{o}\,,$$

because $\boldsymbol{U}_n$ does not depend on $\boldsymbol{\mu}$, so that $\frac{\partial vec(\boldsymbol{U}_n^\top)}{\partial \boldsymbol{\mu}^\top} = \boldsymbol{o}$ and $\mathbb{E}\boldsymbol{U}_n = \boldsymbol{o}$. Since $\boldsymbol{U}_n\boldsymbol{X}_n^\top = (\sum_{i=1}^{T_n}\boldsymbol{\varepsilon}_{ni}(\boldsymbol{Y}_{n,i-1}-\boldsymbol{\mu}),\ldots,\sum_{i=1}^{T_n}\boldsymbol{\varepsilon}_{ni}(\boldsymbol{Y}_{n,i-q+1}-\boldsymbol{\mu}))$ and $\boldsymbol{\varepsilon}_{ni}$ is independent of $\boldsymbol{Y}_{nj}$ for $j<i$, we can conclude that $\mathbb{E}\boldsymbol{U}_n\boldsymbol{X}_n^\top = \boldsymbol{o}$ and thus

$$\lim_{K\to\infty} -\frac{1}{K}\mathbb{E}\frac{\partial^2 \log L}{\partial \boldsymbol{w}\partial \boldsymbol{a}^\top} = \boldsymbol{o}\,.$$

This shows that, $\lim_{K\to\infty}\left(\frac{\boldsymbol{I}(\boldsymbol{\delta})}{K}\right)^{-1}$ is block diagonal. Next we investigate the asymptotic behaviour of the diagonal elements of the covariance matrix. To do so we consider the matrix $\boldsymbol{X}_n\boldsymbol{X}_n^\top = (x_{kl})$ with $x_{kl} = \sum_{i=0}^{T_n-1}(\boldsymbol{Y}_{n,i}-\boldsymbol{\mu})(\boldsymbol{Y}_{n,-|k-l|+i}-\boldsymbol{\mu})^\top$. The expectation of this matrix is given by $\mathbb{E}\boldsymbol{X}_n\boldsymbol{X}_n^\top = T_n\boldsymbol{\Gamma}_{\boldsymbol{Y}}(0)$ since $\mathbb{E}x_{kl} = \boldsymbol{\Gamma}_{\boldsymbol{Y}}(|k-l|)$. It then follows that

$$\begin{aligned}
\lim_{K\to\infty} -\frac{1}{K}\mathbb{E}\frac{\partial^2 \log L}{\partial \boldsymbol{a}\partial \boldsymbol{a}^\top} &= \lim_{K\to\infty}\frac{1}{K}\mathbb{E}\sum_{n=1}^{K}(\boldsymbol{X}_n\boldsymbol{X}_n^\top)\otimes\boldsymbol{\Sigma}^{-1}\\
&= \lim_{K\to\infty}\frac{1}{K}\sum_{n=1}^{K}T_n\boldsymbol{\Gamma}_{\boldsymbol{Y}}(0)\otimes\boldsymbol{\Sigma}^{-1}\\
&= \gamma\boldsymbol{\Gamma}_{\boldsymbol{Y}}(0)\otimes\boldsymbol{\Sigma}^{-1}.
\end{aligned}$$

Hence, from (7.1) we get that

$$\sqrt{K}(\widetilde{\boldsymbol{a}}-\boldsymbol{a}) \xrightarrow{d} N(\boldsymbol{o}, \frac{\boldsymbol{\Gamma}_{\boldsymbol{Y}}(0)^{-1}}{\gamma}\otimes\boldsymbol{\Sigma}).$$

Moreover,

$$\begin{aligned}
\lim_{K\to\infty} -\frac{1}{K}\mathbb{E}\frac{\partial^2 \log L}{\partial \boldsymbol{\mu}\partial \boldsymbol{\mu}^\top} &= \lim_{K\to\infty}\frac{1}{K}\mathbb{E}\sum_{n=1}^{K}T_n(\boldsymbol{I}_3-\sum_{j=1}^{q}\boldsymbol{A}_j)^\top\Sigma^{-1}(\boldsymbol{I}_3-\sum_{j=1}^{q}\boldsymbol{A}_j)\\
&= \left(\lim_{K\to\infty}\frac{\sum_{n=1}^{K}T_n}{K}\right)(\boldsymbol{I}_3-\sum_{j=1}^{q}\boldsymbol{A}_j)^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{I}_3-\sum_{j=1}^{q}\boldsymbol{A}_j)\\
&= \gamma(\boldsymbol{I}_3-\sum_{j=1}^{q}\boldsymbol{A}_j)^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{I}_3-\sum_{j=1}^{q}\boldsymbol{A}_j).
\end{aligned}$$

We thus obtain

$$\sqrt{K}(\widetilde{\boldsymbol{\mu}}-\boldsymbol{\mu}) \xrightarrow{D} N(\boldsymbol{o}, \frac{1}{\gamma}(\boldsymbol{I}_3 - \sum_{j=1}^{q} \boldsymbol{A}_j)^{-1}\boldsymbol{\Sigma}(\boldsymbol{I}_3 - \sum_{j=1}^{q} \boldsymbol{A}_j^\top)^{-1}).$$

Finally we obtain

$$\begin{aligned}
\lim_{K\to\infty} -\frac{1}{K}\mathbb{E}\frac{\partial^2 \log L}{\partial \boldsymbol{w}\partial \boldsymbol{w}^\top} &= \lim_{K\to\infty} \frac{1}{K}\mathbb{E}\sum_{n=1}^{K} -[\frac{T_n}{2}\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\right) + \frac{1}{2}\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_n\boldsymbol{U}_n^\top\boldsymbol{\Sigma}^{-1}\right) \\
&\quad + \frac{1}{2}(\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_n\boldsymbol{U}_n^\top\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1})] \\
&= \lim_{K\to\infty}\left(\sum_{n=1}^{K} -\frac{T_n}{2K}(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}) + \frac{1}{2}(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\mathbb{E}(\boldsymbol{U}_n\boldsymbol{U}_n^\top)\boldsymbol{\Sigma}^{-1})\right. \\
&\quad \left. + \frac{1}{2}(\boldsymbol{\Sigma}^{-1}\mathbb{E}(\boldsymbol{U}_n\boldsymbol{U}_n^\top)\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1})\right) \\
&= (\lim_{K\to\infty}\frac{\sum_{n=1}^{K} T_n}{2K})(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}) \\
&= \frac{\gamma}{2}(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}).
\end{aligned}$$

The last equality follows from $\mathbb{E}(\boldsymbol{U}_n\boldsymbol{U}_n^\top) = T_n\boldsymbol{\Sigma}$ (see Lütkepohl (2006)). This gives

$$\sqrt{K}(\widetilde{\boldsymbol{w}}-\boldsymbol{w}) \xrightarrow{d} N\left(\boldsymbol{o}, \frac{2}{\gamma}\left(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}\right)\right)$$

and thus

$$\sqrt{K}(\widetilde{\boldsymbol{\sigma}}-\boldsymbol{\sigma}) \xrightarrow{d} N\left(\boldsymbol{o}, \frac{2}{\gamma}\boldsymbol{D}_3^+\left(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}\right)\boldsymbol{D}_3^{+\top}\right),$$

where $\boldsymbol{D}_3^+$ is the Moore-Penrose generalized inverse of the matrix $\boldsymbol{D}_3$ which is the uniquely determined solution of the linear equation system $\boldsymbol{\omega} = \boldsymbol{D}_3\boldsymbol{\sigma}$ (see Lütkepohl (2006)).

References

Altendorf, H. & Jeulin, D. (2011). Random walk based stochastic modeling of 3D fiber systems. *Phys. Rev. E*, **83**, 041804.

Anderson, T. W. (1978). Repeated measurements in autoregressive processes. *J. Amer. Statist. Assoc.*, **73**, 371-378.

Azzalini, A. (1981). Replicated observations of low order autoregressive time series. *J. Time Series Anal.*, **2**, 63-70.

Ben-Israel, A. & Greville, T. N. E. (2003). Generalized Inverses - Theory and Applications. 2. ed., New York: Springer.

Davis, C.S (2002). Statistical Methods for the Analysis of Repeated Measurements. New York: Springer.

Diggle, P.J. & Al Wasel, I. (1997). Spectral analysis of replicated biomedical time series. *J. Appl. Stat.*, **46**, 31-71.

Diggle, P.J., Heagerty, P., Liang, K.Y. & Zeger, S.L. (2002). Analysis of Longitudinal Data. 2. ed., Oxford: Oxford University Press.

Fuller, W. A. (1996). Introduction to Statistical Time Series. 2. ed., New York: J. Wiley & Sons.

Gaiselmann, G., Froning, D., Tötzke, C., Quick, C., Manke, I., Lehnert, W. & Schmidt, V. (2013). Stochastic 3D modeling of non-woven GDL with wet-proofing agent. *Int. J. Hydrogen Energy*, **38**, 8448-8460.

Gaiselmann, G., Manke, I., Lehnert, W. & Schmidt, V. (2012). Extraction of curved fibers from 3D data. *Image Anal. Stereol.*, **32**, 57-63.

Gaiselmann, G., Thiedmann, R., Manke, I., Lehnert, W. & Schmidt, V. (2012). Stochastic 3D modeling of fiber-based materials. *Comp. Mater. Sci.*, **59**, 75-86.

Ledolter, J. & Lee, C. S. (1999). Analysis of many short time sequences: Forecast improvements achieved by shrinkage. *J. Forecast.*, **12**, 1-11.

Lindsey, J. K. (1999). Models for Repeated Measurements. 2. ed., Oxford: Oxford University Press.

Lütkepohl, H. (2006). New Introduction to Multiple Time Series Analysis. Berlin: Springer.

Provatas, N., Haataja, M., Asikainen, J., Majaniemi, S., Alava, M. & Ala-Nissila, T. (2000). Fiber deposition models in two and three spatial dimensions. *Colloids Surf. A*, **165**, 209–229.

Redenbach, C. & Vecchio, I. (2011). Statistical analysis and stochastic modelling of fibre composites. *Compos. Sci. Technol.*, **71**, 107-112.

Schulz, V. P., Becker, J., Wiegmann, A., Mukherjee, P. P. & Wang, C.-Y. (2007). Modeling of two-phase behavior in the gas diffusion medium of PEFCs via full morphology approach. *J. Electrochem. Soc.*, **154**, B419–B426.

Shi, G. & Chaganty, N. R. (2004). Application of quasi-least squares to analyse replicated autoregressive time series regression models. *J. Appl. Stat.*, **31**, 1147-1156.

Swift, S. & Liu, X. (2002). Predicting glaucomatous visual fields deterioration through short multivariate time series modeling. *Artif. Intell. Med.*, **24**, 5-24.

Thiedmann, R., Fleischer, F., Hartig, C., Lehnert, W. & Schmidt, V. (2008). Stochastic 3D modelling of the GDL structure in PEM fuel cells, based on thin section detection. *J. Electrochem. Soc.*, **155**, B391–B399.