Journal of Neuroscience Methods 219 (2013) 285–291

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Computational Neuroscience

Block-bootstrapping for noisy data

Malenka Mader[a,b,c,*], Wolfgang Mader[b,c], Linda Sommerlade[b,c,d], Jens Timmer[b,c,d], Björn Schelter[b,c,d,e,f]

[a] Department of Neuropediatrics and Muscular Disease, University Medical Center of Freiburg, Mathildenstrasse 1, 79106 Freiburg, Germany
[b] Freiburg Center for Data Analysis and Modeling (FDM), University of Freiburg, Eckerstrasse 1, 79104 Freiburg, Germany
[c] Institute for Physics, University of Freiburg, Hermann-Herder-Strasse 3a, 79104 Freiburg, Germany
[d] Freiburg Institute of Advanced Studies, University of Freiburg, Albertstrasse 19, 79104 Freiburg, Germany
[e] Division of Functional Brain Imaging (FBI), Department of Neurology, University Medical Center of Freiburg, Breisacher Strasse 64, 79106 Freiburg, Germany
[f] Institute for Complex Systems and Mathematical Biology, Scottish Universities Physics Alliance, University of Aberdeen, King's College, Old Aberdeen AB24 3UE, United Kingdom

HIGHLIGHTS

- For the first time, robust statistics is possible due to our new method.
- The proposed method outperforms conventional methods. It renders confidence intervals as small as possible.
- The new method precludes false positive conclusions.

ARTICLE INFO

Article history:
Received 18 May 2013
Received in revised form 30 July 2013
Accepted 31 July 2013

Keywords:
Statistics
Distribution estimation
Autocorrelation
Measurement noise
Dependent data
Tremor

ABSTRACT

Background: Statistical inference of signals is key to understand fundamental processes in the neurosciences. It is essential to distinguish true from random effects. To this end, statistical concepts of confidence intervals, significance levels and hypothesis tests are employed. Bootstrap-based approaches complement the analytical approaches, replacing the latter whenever these are not possible.
New method: Block-bootstrap was introduced as an adaption of the ordinary bootstrap for serially correlated data. For block-bootstrap, the signals are cut into independent blocks, yielding independent samples. The key parameter for block-bootstrapping is the block length. In the presence of noise, naïve approaches to block-bootstrapping fail. Here, we present an approach based on block-bootstrapping which can cope even with high noise levels. This method naturally leads to an algorithm of block-bootstrapping that is immediately applicable to observed signals.
Results: While naïve block-bootstrapping easily results in a misestimation of the block length, and therefore in an over-estimation of the confidence bounds by 50%, our new approach provides an optimal determination of these, still keeping the coverage correct.
Comparison with existing methods: In several applications bootstrapping replaces analytical statistics. Block-bootstrapping is applied to serially correlated signals. Noise, ubiquitous in the neurosciences, is typically neglected. Our new approach not only explicitly includes the presence of (observational) noise in the statistics but also outperforms conventional methods and reduces the number of false-positive conclusions.
Conclusions: The presence of noise has impacts on statistical inference. Our ready-to-apply method enables a rigorous statistical assessment based on block-bootstrapping for noisy serially correlated data.

* Corresponding author at: Freiburg Center for Data Analysis and Modeling (FDM), University of Freiburg, Eckerstrasse 1, 79104 Freiburg, Germany. Tel.: +49 7612037710.
E-mail addresses: Malenka.Mader@fdm.uni-freiburg.de (M. Mader), Wolfgang.Mader@fdm.uni-freiburg.de (W. Mader), Linda.Sommerlade@fdm.uni-freiburg.de (L. Sommerlade), JeTi@fdm.uni-freiburg.de (J. Timmer), B.Schelter@abdn.ac.uk (B. Schelter).

0165-0270/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jneumeth.2013.07.022

1. Introduction

In the neurosciences, concepts, mechanisms, and characteristics often need to be inferred from measured signals. This typically renders powerful means for statistical evaluation necessary to preclude erroneous conclusions. For several data analysis approaches analytic evaluation schemes have been developed. For others, numerical Monte-Carlo based approaches are used to evaluate the statistical significance of the findings. The bootstrap has emerged as a powerful tool for this (Efron, 1979; Arndt et al., 1996; Davison and Hinkley, 1997; Foster and Bischof, 1987; Hentschke and Stüttgen, 2011; Zoubir and Boashash, 1998).

The general idea of bootstrap is to provide the distribution of a statistics based on the measured signals alone, when the analytic derivation of the statistics is not known. The distribution is sampled by randomly drawing with replacement from the measured data (Efron and Gong, 1983; Hall et al., 1995). Once this empirical distribution is obtained from the bootstrap, confidence intervals can be derived and hypothesis tests can be performed based on the empirical α-quantiles. Bootstrapping leads to a valid approximation of the true distribution of the test statistics under some assumptions (Mammen, 1992).

Among others, independence of the sampled data points is one assumption that has to be fulfilled in order to render bootstrapping sensible. When investigating time series as often measured in the neurosciences, this fundamental prerequisite for the applicability of bootstrap is violated. The temporal correlation, which characterises the dependence of the random variables of the time series, can be quantified by the autocorrelation function. This insight has led to the idea of block-bootstrapping (Carlstein, 1986; Hall et al., 1995; Künsch, 1989).

For temporally correlated data, block-bootstrapping draws with replacement from a set of independent blocks, i.e. snippets of the data (Davison and Hinkley, 1997). The appropriate choice of the block length is a key parameter and does not only depend on the measured time series but also on the analysis technique that is applied. "Optimality" in case of block-bootstrapping refers to the minimal squared distance between the true and the estimated quantity, yielding a trade-off between squared bias and variance (Peifer et al., 2005; Percival and Walden, 1993; Schelter et al., 2007).

The decay rate of the autocorrelation function as a measure of dependence in the data is the vital parameter in block-bootstrapping, as has been shown for the variance (Peifer et al., 2005) and mean phase coherence (Schelter et al., 2007), explicitly. It needs to be estimated as reliably as possible in order to render the segments as short as possible but long enough to guarantee independence. The decay rate in the autocorrelation can either be estimated by fitting an exponential function to the envelope of the empirical autocorrelation function or alternatively it can be estimated by modelling the process as an autoregressive one. While the latter is sensible only for small orders of the autoregressive model, the former provides a robust means for more complicated, potentially nonlinear dynamics as well.

A naïve choice of block length yields a bias or high variance of the statistics, and eventually fails in providing an optimal estimate for the block length. This is due to the influence of noise, becoming particularly important in the case of measurement noise or in cases in which the signal itself is modulated noise. While for the former the electroencephalography (EEG) is a prototypical example, the electromyography (EMG) is genuine for the latter.

As we demonstrate in this manuscript, the presence of these two types of noise strongly influences the determination of the optimal block length. A modification of conventional methods is necessary otherwise sub-optimal or even anti-conservative statistics can be obtained due to an underestimation of the block lengths. As we demonstrate here by both analytic calculations and simulations, the amount to which the length is underestimated is a function of the noise to signal ratio. We provide a modified block length selection approach that is robust with respect to the presence of a range of noise levels. We demonstrate the benefit of this new approach not only in model systems but also by investigating confidence intervals for the tremor amplitude based on EMG activity. The amplitude provides a measure for tremor severity and is therefore key to support physicians in the various tasks, such as the diagnosis or treatment of tremor. However, we emphasise that the proposed approach is not confined to EMG recordings and tremor data. Neither does it depend on recording modalities, such as EEG, EMG or fMRI, nor on scientific fields, such as tremor or epilepsy. Examples of its applicability are seizure detection as initiated by Gotman (1982), various studies concerned with network estimation, e.g. by the mean phase coherence (Schelter et al., 2007), or resting state studies (Bellec et al., 2010).

The manuscript is structured as follows. In Section 2 we present the EMG data of a tremor patient and the systems used to model them. We first (Section 2.1) introduce the EMG data that we aim to analyse and specify the impacts of noise onto the analysis. Based on the EMG data, parameters of the model system are adapted. As a second step (Section 2.2) we demonstrate the weaknesses of conventional block length selections in the presence of noise. We analytically derive how this can be overcome with our new robust block-bootstrapping. In Section 3 we apply block-bootstrapping to both the model system and the measured EMG signals of a tremor patient, deriving confidence intervals. We compare the confidence intervals obtained from the modified block length selection to the unmodified version, showing the superior efficiency of our method.

2. Material and methods

To motivate and illustrate the new approach to block-bootstrapping, we use an example of a representative recording of the wrist muscle activity of a tremor patient measured by EMG (see Fig. 1a). We model the EMG by autoregressive processes (Fig. 1b) in order to analytically show the effect of noise onto the block length selection.

2.1. Tremor example

Tremor is characterised by an involuntary oscillating movement of extremities. In essential tremor, a hereditarian form of pathological tremor, typically, the hands tremble at a frequency at around

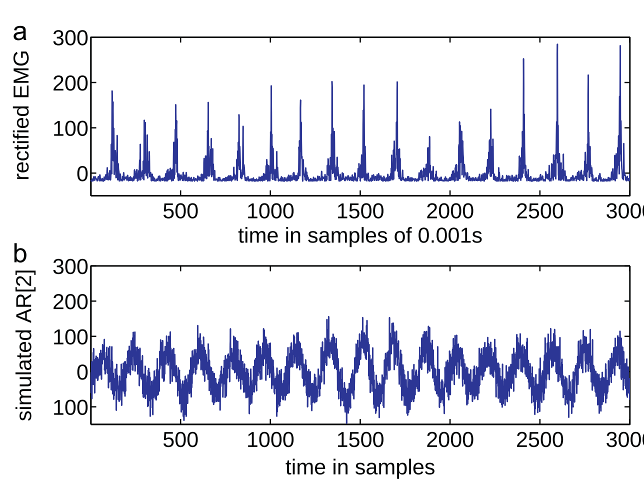

Fig. 1. Short section of data from a rectified EMG (a), and an autoregressive process (b) with parameters $a_1 = 1.9975$ and $a_2 = -0.9987$, intrinsic noise variance $\sigma^2 = 0.05^2$ and measurement noise variance $\Sigma^2 = 30^2$ (see Eqs. (1) and (2)).

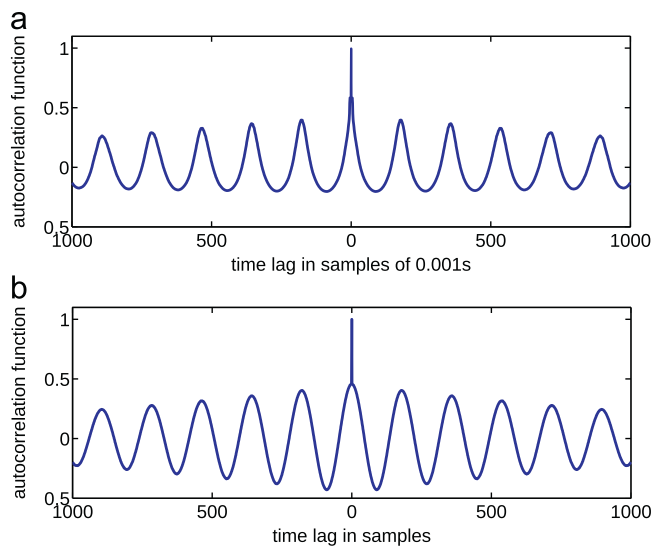

Fig. 2. Estimated autocorrelation function of the EMG data (a) and the AR[2] (b), which models the EMG recordings. Both signals are influenced by noise and therefore have a discontinuity at time lag 0.

5–7 Hz. The corresponding tremor amplitude, which is quantified by the variance of the signal, is far stronger than physiological trembling. The muscle activity is recorded at a sampling rate of 1000 Hz by EMG. A short epoch of rectified EMG data is depicted in Fig. 1a. Of particular interest in tremor is the progression of the disease. It is, e.g., still under debate whether tremor becomes more severe with age. One measure of severity is the amplitude of the trembling, which is typically measured as the variance of the signal in a given frequency band. To be able to study the potential long term changes in amplitude, a reliable approximation of the confidence intervals for the variance are needed. In Peifer et al. (2005), block-bootstrapping was suggested to approximate the statistics of the variance.

For EMG signals, the patho-physiological basis of tremor investigations, the approach suggested in Peifer et al. (2005) is not directly applicable. The envelope of the autocorrelation function is not just an exponential decay (Fig. 2a). Since the EMG is modulated noise, there is a discontinuity at time lag zero. In the following, we will suggest a block-bootstrapping approach that is not influenced by such a discontinuity. Before discussing the novel approach to block-bootstrapping, we introduce a rough model, which we will utilise in the following to substantiate our numerical results by analytic derivations.

Interestingly, the nonlinear rectified EMG (Jachan et al., 2009) can be roughly modelled by a so-called autoregressive process of order p, AR[p],

$$x_t = \sum_{i=1}^{p} a_i x_{t-i} + \epsilon_t. \tag{1}$$

The order p, coefficients a_i and dynamical Gaussian white noise term ϵ_t with zero-mean and variance σ_t^2 can be chosen such that the frequency and second order properties, i.e. autocorrelation properties, model the second order properties of the EMG except for the discontinuity at lag 0.

In the example of Fig. 1a, the tremor frequency was 5.59 Hz. The parameters of a low order autoregressive process ($p = 2$) captures the main oscillation of the EMG tremor signal for the process parameters $a_1 = 1.9975$ and $a_2 = -0.9987$, and intrinsic noise variance $\sigma^2 = 0.05^2$. The EMG is modulated noise such that the noise strength is roughly of the same order as the signal. In the model system (Eq. (1)) the modulated noise of the EMG can be modelled as

$$y_t = x_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \Sigma^2) \tag{2}$$

by adding approximately 100% measurement noise with variance $\Sigma^2 = 30^2$ to the AR[2] model data x_t derived from Eq. (1). This corresponds to a 1:1 noise to signal ratio defined as the variance of the measurement noise divided by the variance of the signal. The corresponding autocorrelation functions are shown in Fig. 2 for both the EMG (Fig. 2a) and the described AR[2] model (Fig. 2b). The qualitative agreement warrants usage of this simplified model for the theoretical considerations.

2.2. Block-bootstrapping in the presence of noise

Block-bootstrapping is based on segmenting the data into independent blocks of optimal length. The temporal correlation in a stationary process is quantified by the autocorrelation function,

$$\gamma_\tau = \frac{\mathrm{E}[x_t x_{t+\tau}]}{\mathrm{Var}[x]}, \tag{3}$$

for time lags τ, where Var[·] denotes the variance and E[·] the expected value of the respective zero-mean process. An estimate of the autocorrelation function is

$$\hat{\gamma}_\tau = \frac{1}{\widehat{\mathrm{Var}}[x]} \frac{1}{N} \sum_{t=1}^{N-\tau} x_t x_{t+\tau}. \tag{4}$$

For exponentially mixing processes, as typically present in applications, the envelope of the autocorrelation function, Eq. (3), is of the form ϕ^τ. We define the decay rate $\phi := \exp(-1/\theta) < 1$ where θ is the relaxation time or time constant of an exponential decay. For a correct statistics of the variance estimation, the optimal block length (Peifer et al., 2005),

$$l_{\mathrm{opt}} = \left[N \frac{\left(\sum_{\tau=-\infty}^{\infty} |\tau| \gamma_\tau\right)^2}{\left(\sum_{\tau=-\infty}^{\infty} \gamma_\tau\right)^2} \right]^{1/3}, \tag{5}$$

is a function of the decay rate ϕ, leading to

$$l_{\mathrm{opt}} = \left[4N \frac{\left(\phi/(1-\phi)^2\right)^2}{\left(1 + 2(\phi/(1-\phi))\right)^2} \right]^{1/3} \tag{6}$$

for non-overlapping blocks (Peifer et al., 2005).

As motivated by the EMG and the AR[2] example (Fig. 2), noise leads to a discontinuity of the autocorrelation function at $\tau = 0$. To briefly discuss the reason for this discontinuity, consider an autoregressive process of order one, AR[1],

$$x_t = a x_{t-1} + \epsilon_t. \tag{7}$$

The autocorrelation function of this process with parameter $|a| < 1$, for stationarity, and Gaussian white noise $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$ with zero-mean and variance σ^2 is

$$\gamma_\tau = \frac{\mathrm{E}[x_t x_{t-\tau}]}{\mathrm{Var}[x]} = \frac{\mathrm{E}[a^\tau x_{t-\tau} x_{t-\tau}]}{\mathrm{Var}[x]} = a^\tau. \tag{8}$$

Note that this process is not only exponentially mixing but there is also an immediate link of the decay rate to the process parameter, $\phi = a$.

Adding Gaussian measurement noise $\eta_t \sim \mathcal{N}(0, \Sigma^2)$ to the process in Eq. (7) according to Eq. (2) the autocorrelation function becomes

$$\frac{\mathrm{E}[y_t y_{t-\tau}]}{\mathrm{Var}[y]} \overset{(2)}{=} \begin{cases} \frac{\mathrm{Var}[x]}{\mathrm{Var}[y]} + \frac{\mathrm{Var}[\eta]}{\mathrm{Var}[y]} & \text{if } \tau = 0 \\ \frac{\mathrm{E}[x_t x_{t-\tau}]}{\mathrm{Var}[y]} & \text{else} \end{cases} \tag{9}$$

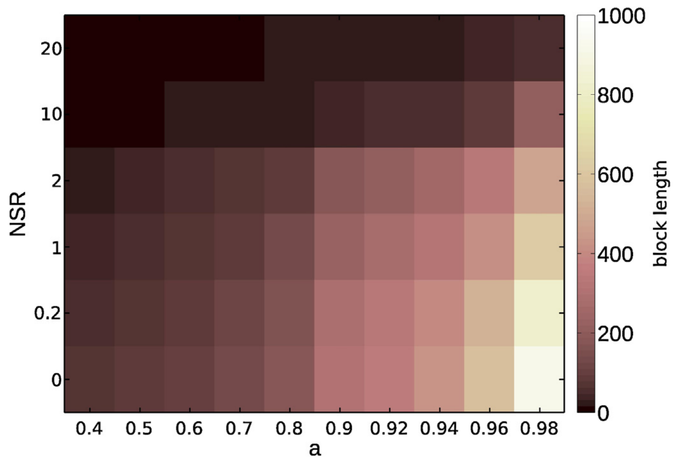

Fig. 3. Block length simulation for AR[1] with varying parameters a and noise to signal ratios NSR. The colour encodes the block length, which was estimated fitting the model ϕ^τ to the autocorrelation function. Block lengths are highly dependent on the noise to signal ratio, no matter which parameter, block lengths become small for high noise to signal ratios. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

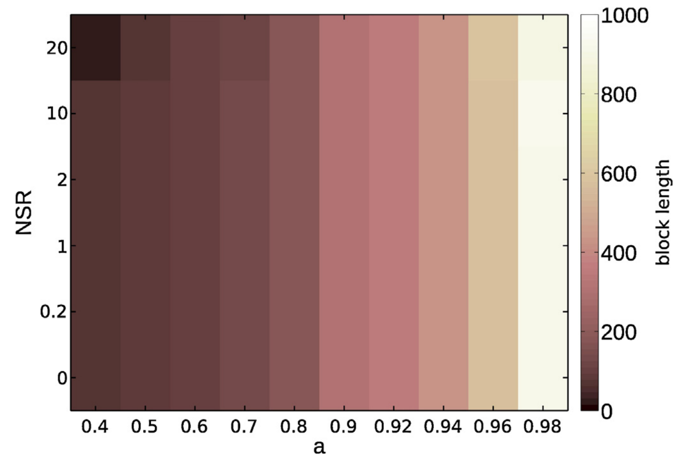

Fig. 4. Block length simulation for AR[1] with varying parameters a and noise to signal ratios as in Fig. 3. The colour encodes the block length, which was estimated fitting the model $A\phi^\tau$ to the autocorrelation function. Block lengths are not dependent on the noise to signal ratio anymore. They only depend on the dynamical parameter a. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$= \begin{cases} 1 & \text{if} \quad \tau = 0 \\ \dfrac{\text{Var}[x]}{\text{Var}[y]} \dfrac{\text{E}[x_t x_{t-\tau}]}{\text{Var}[x]} = Aa^\tau & \text{else} \end{cases} \tag{10}$$

with $A = [1 + (\text{Var}[\eta]/\text{Var}[x])]^{-1}$. Eqs. (9) and (10) make use of the independence of measurement noise η_t and the dynamics x_t, such that $\text{Var}[y] = \text{Var}[x] + \text{Var}[\eta]$. The measurement noise contributes to the autocorrelation function of the process at $\tau = 0$, but not at any $\tau > 0$. Therefore, the autocorrelation function drops from time lag $\tau = 0$ to $\tau = 1$ by a factor A, cf. Eq. (10), which depends on the noise to signal ratio, NSR=$\text{Var}[\eta]/\text{Var}[x]$. Consequently, the factor A has to be introduced to model the decay properties, notably the discontinuity at time lag $\tau = 0$. Such a discontinuity appears also in the autocorrelation function of the EMG and the AR[2] with measurement noise (Fig. 2).

When fitting an exponential decay ϕ^τ to the autocorrelation function in Eq. (10), the decay rate is estimated inadequately such that a systematic error occurs in the block length selection, cf. Eq. (6). In Fig. 3 the impact of the measurement noise on the block length estimation when assuming the autocorrelation function to be $\gamma_\tau = \phi^\tau$ in Eq. (5) instead of $A\phi^\tau$, as proposed by Eq. (10), is shown for different process parameters a and different noise to signal ratios $\frac{\text{Var}[\eta]}{\text{Var}[x]}$. The optimal block length of the underlying process x with respective parameters $a = \phi$ is given for NSR = 0. Block lengths decay with increasing noise to signal ratio for all parameter values a. With measurement noise the optimal block length is underestimated.

To prevent this, the assumptions of the exponential decay have to be modified, such that in the optimal block length equation the autocorrelation function is modelled by $A\phi^\tau$ instead of ϕ^τ. Analogously to the optimal block length for non-overlapping blocks derived in Peifer et al. (2005), the optimal block length is derived from Eq. (5) by inserting $A\phi^\tau$ instead of ϕ^τ. Using $\gamma_\tau = \gamma_{-\tau}$ and $\phi < 1$ in the geometric series $\sum_{\tau=0}^{\infty} \phi^\tau = 1/(1-\phi)$ and its derivative $\sum_{\tau=0}^{\infty} \tau\phi^{\tau-1} = 1/(1-\phi)^2$ leads to Eq. (6) as before. Note that the term for $\tau = 0$ enters the geometric series. Therefore the optimal block length is estimated from the decay rate ϕ only. The difference is that the autocorrelation function is modelled incorporating the constant A such that the decay property ϕ for $\tau > 0$ can be fit correctly. The factor A itself eliminates.

Fig. 4 shows the block lengths estimated based on this modification. For all parameters a the obtained block lengths are, other than for the conventional method (Fig. 3), independent of the noise influence. This independence of the block length from the noise to signal ratio shows that the modification renders the block length selection robust to noise influences.

In Figs. 3 and 4 the block lengths are colour coded based on the noise to signal ratio and the parameter values of the processes. Both figures show the median optimal block length of three realisations, in order to remove outliers for robustness. Since a fit is not guaranteed to converge to the global optimum, but may end in a local one, non-optimal estimates of the envelope of the autocorrelation function may occur. They result in outliers of the optimal block length. Therefore, the median has been chosen to calculate the final optimal block length. We want to stress that this is not an issue inherent to the method proposed, but to every fit algorithm.

Two methods are conceivable in order to obtain the optimal block length based on the decay properties of the autocorrelation function. Either the decay rate ϕ is fit to the envelope of the autocorrelation function according to the model $A\phi^\tau$ as shown in Fig. 5 for the tremor data. Alternatively, an autoregressive process of order two (AR[2]) in the state space model (Honerkamp and Stüttgen, 2012),

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \epsilon_t \quad t = 1, \ldots, N, \tag{11}$$

$$y_t = x_t + \eta_t \quad t = 1, \ldots, N, \tag{12}$$

is fit to the data. Eq. (11) is the dynamical equation, which models the underlying process x_t as an autoregressive process with parameters a_1 and a_2, and dynamical Gaussian white noise $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$ with variance σ^2. Eq. (12) is the observation equation. It models the

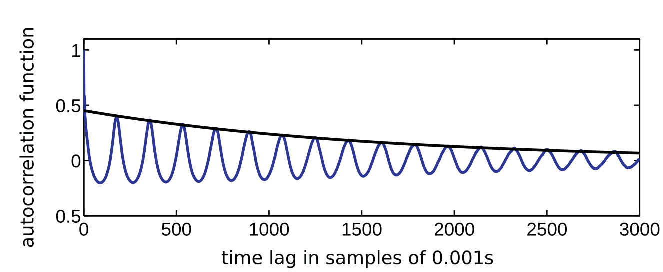

Fig. 5. Estimated autocorrelation function of EMG data of a trembling patient (blue) and the fit $A\phi^\tau$ of its envelope (black). It yields an optimal block length of $l_{\text{opt}} = 9143$.

measurement y_t by the underlying process x_t and additive measurement noise $\eta_t \sim \mathcal{N}(0, \Sigma^2)$. Once the model (Eqs. (11) and (12)) is fit, the decay rate can be derived from the parameters, as follows.

An AR[2] is linked to a damped harmonic oscillator with driving noise, when setting the parameters (Honerkamp and Stüttgen, 2012)

$$a_1 = 2e^{-1/\theta} \cos\left(\frac{2\pi}{T}\right) \tag{13}$$

$$a_2 = -e^{-2/\theta}, \tag{14}$$

where θ is the relaxation time and T is the period of the oscillator. From the relaxation time the decay rate $\phi = \exp(-1/\theta)$ can be determined. Given the parameter a_2 of an AR[2] process, the decay rate can be obtained from

$$\phi = \sqrt{-a_2}. \tag{15}$$

Once the decay rate is estimated, either by fitting $A\phi^\tau$ to the autocorrelation function or by fitting an AR[2] and determining the decay rate from the parameter a_2, cf. Eq. (15), the bootstrap is conduced as follows:

1. Compute the optimal block length l_{opt} for your statistic, e.g. for the variance according to Eq. (6).
2. Cut data into non-overlapping blocks of length l_{opt}.
3. Estimate the variable of interest, here the variance, from each block separately, and from the whole set of data under investigation.
4. Randomly draw the estimated variable with replacement from the ensemble of the blocks.
5. Sample the distribution of the variable of interest from the random draws.
6. Compare the variable of interest derived from the whole set of data, see 3., to the distribution sampled by the blocks.

Here, we sample the distribution by estimating the standard deviation of blockwise variances from the random draws. From the standard deviation, confidence intervals are derived. If the quantiles of the distribution are needed, the distribution can be sampled from the bootstrap by building the histogram from the independent draws of the bootstrap. Quantiles then can be estimated from the histogram.

3. Results

In order to illustrate the effect of noise on the choice of the optimal block length, we simulated $N = 308{,}960$ samples of an AR[2], as in Eqs. (11) and (12). The number of data points N and the parameters a_i were chosen such that the autoregressive process modelled the EMG data of the tremor patient according to frequency, decay properties and noise to signal ratio (Appendix A). This yielded a true decay rate of the AR[2], $\phi = 0.999364$, and a respective optimal block length, $l_{opt} = 9143$ data points according to Eq. (6).

To compare the performance of the proposed modification to the conventional method, we compared the decay properties and the resulting block lengths. From the two approaches proposed in Section 2.2, we chose the first. Thus, instead of fitting an AR[2] to the data as the second approach suggests, we fit an exponential decay to the envelope of the autocorrelation function. The latter was estimated from the simulated AR[2] both without and with measurement noise, i.e. x_t and y_t in Eqs. (11) and (12). For the results of the conventional method the envelope of the estimated autocorrelation function was fit by ϕ^τ. For the modification, the decay was fit by $A\phi^\tau$. The decay rates were used to estimate the optimal block lengths according to Eq. (6). Different settings, summarised

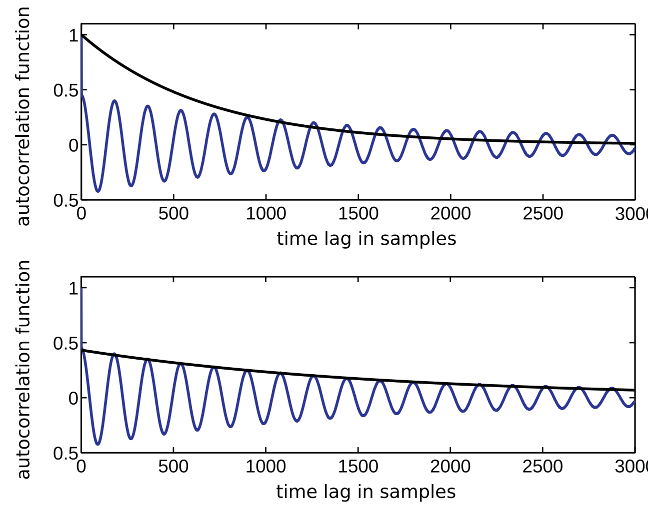

Fig. 6. Decay of estimated autocorrelation function (blue) and the estimate of its envelope (black) without (top) and with (bottom) the proposed modification.

in Table 1, were investigated to compare the conventional method to the proposed modification.

First, the optimal block length was estimated from the adapted fit of the estimated autocorrelation function of the simulated AR[2] without measurement noise, i.e. x_t in Eq. (11). The parameters of the AR[2] were chosen as described in Appendix A, such that the optimal block length was known. This yielded a relative deviation $(|l_{opt} - \hat{l}_{opt}|)/l_{opt} = 4.2\%$ of the true optimal block length $l_{opt} = 9143$ data points to the estimated optimal block length $\hat{l}_{opt} = 9527$ data points. The relative deviation is a measure of accuracy for the adapted fit.

Second, the optimal block length was estimated from the measurement y_t (Eq. (12)), instead of x_t. Adapted fitting yields a relative deviation $|l_{opt} - \hat{l}_{opt}|/l_{opt} = 2.5\%$ of the estimated block length $\hat{l}_{opt} = 9357$ data points to l_{opt}. However, when ignoring the modification, the relative deviation of the estimated optimal block length, $\hat{l}_{opt} = 5235$ data points, to l_{opt} is 42.7%. In Fig. 6 the estimated autocorrelation function (blue) of the simulated autoregressive process y_t with measurement noise as well as both the unadapted (ϕ^τ, Fig. 6a, black) and the adapted ($A\phi^\tau$, Fig. 6b, black) fits are shown.

To summarise, the relative deviation of estimated and true optimal block length of the AR[2] contaminated with measurement noise is one order of magnitude higher for the conventional than for the adapted method. In contrast, the relative deviations obtained from the adapted fit of the estimated autocorrelation functions of the measured signal y_t and the one of the underlying process x_t are comparable.

Once optimal block lengths are derived, they can be used to determine the confidence bounds of the estimated variance $\widehat{V}$ of a signal. For this, the variance needs to be estimated on each block, yielding $\widehat{V}_{block}$. The standard deviation $\widehat{\Delta V}$ of the variance $\widehat{V}$ is estimated by the standard deviation of randomly chosen blockwise estimated variances $\widehat{V}_{block}$.

In our simulation study the estimated variance of the AR[2] is $\widehat{V} = 1638$ (in a.u.). For the bootstrap, 20 independent choices of the blockwise variances were drawn based on the optimal block lengths estimated as above. Block lengths were estimated from the underlying process (Table 1, row 1), from the measured process with (Table 1, row 2) and without (Table 1, row 3) adaption of the fit. The estimated standard deviation based on the optimal block length according to the adapted fit was $\widehat{\Delta V} = 230$ (in a.u.). The relative deviation of the halfwidth of the 95% confidence interval

Table 1
Simulation results of optimal block length selection based on a fit of the envelope of the estimated autocorrelation function of the underlying process x_t (underlying, 1st row), adapted fitting of the autocorrelation function of the measured data y_t (adapted, 2nd row) as proposed here, and the unadapted fitting results from the measured data y_t (unadapted, 3rd row). Simulation results are the estimated block length $\hat{l}_{opt}$ from the autocorrelation function fit, relative deviation $(l_{opt} - \hat{l}_{opt})/l_{opt}$ from the true optimal block length $l_{opt} = 9143$, bootstrapped standard deviation of the variance $\widehat{\Delta \mathrm{Var}}$, and relative 1.96-fold standard deviation of variance to its absolute estimated variance $1.96\Delta\hat{\mathrm{V}}\mathrm{ar}/\widehat{\mathrm{Var}}$. For comparability, the variance $\widehat{Var}$ is the variance of the process with measurement noise, even for the row of the underlying process. For the latter the known measurement variance is added to the estimated underlying process variance. The 1.96-fold of the standard deviation constitutes the 95%-confidence interval about the estimated variance. The last column indicates (Eqs. (11) and (12)), whether the optimal block length was estimated from the autocorrelation function of the process without measurement noise x_t or with measurement noise y_t.

$l_{opt} = 9143$	$\hat{l}_{opt}$	$(l_{opt} - \hat{l}_{opt})/l_{opt}$ (%)	$\widehat{\mathrm{Var}}$	$\widehat{\Delta\,\mathrm{Var}}$	$(1.96 \cdot \widehat{\Delta\,\mathrm{Var}})/\widehat{\mathrm{Var}}$ (%)	Process
Underlying	9527	4.2	1642	230.1	27.5	x_t
Adapted	9357	2.5	1638	217.9	26.1	y_t
Unadapted	5235	42.7	1638	348.1	41.7	y_t

and the estimated variance was $1.96\widehat{\Delta V}/\widehat{V} = 27.5\%$. The estimated standard deviation based on the optimal block length which was derived from adapted fitting is in the same range: $\widehat{\Delta V} = 218$ (in a.u.), with a similar relative deviation $1.96\widehat{\Delta V}/\widehat{V} = 26.1\%$ as for the true optimal block length. However, the unadapted block length selection yields a 50% larger estimate of the standard deviation $\widehat{\Delta V} = 348$ (in a.u.) with a relative deviation $1.96\widehat{\Delta V}/\widehat{V} = 41.7\%$.

To test the power of the confidence bounds obtained by block-bootstrapping, we compared them to the true variance. For an AR[2], the latter can be computed from the process parameters a_1, a_2, and the variances σ^2 and Σ^2 (Appendix B).

All confidence intervals obtained by block-bootstrapping include the true value of the variance $V = 1704$ (in a.u.), but the modification of the block length selection renders the confidence interval as small as possible.

As an example for the applicability of this method, the strength of tremor is quantified by the variance of the EMG recording. The optimal block length for the EMG data was determined from the adapted fit, see Fig. 5. For tremor patients, the variance of the EMG is used as a measure of tremor strength. As for the simulated data, the variance $\widehat{V} = 921$ (in a.u.) was estimated from the data. The standard deviation $\widehat{\Delta V} = 142$ (in a.u.) was estimated from block-bootstrapping using the modified block length selection framework proposed. This yielded an optimal block length of 9143 data points. The relative halfwidth of a 95%-confidence interval to the estimated variance is $1.96\widehat{\Delta V}/\widehat{V} = 30.1\%$.

4. Discussion and conclusion

We proposed a crucial modification of the optimal block length estimation in the framework of block-bootstrap when noise is present. Guided by an example of electromyographic (EMG) recordings of a trembling tremor patient, we simulated the impact of noise onto the statistics by autoregressive processes. This enabled us to compare the true optimal block length to the estimated (a) based on the unadapted block length selection and (b) based on our proposed modification. The adapted version is comparable to the true optimal block length while the unadapted block length selection deviates from the true block length by almost 50% in our simulation. As a second step we compared the performances of the different block length selection algorithms with respect to the statistics. The adapted block length selection method performs just as good as the true block length does. The unadapted method is imprecise and leads to a 50% larger confidence interval.

The results are useful whenever the block length selection is determined by the decay properties of the autocorrelation function. This is commonly the case except for processes like fractional Gaussian noise.

The method can be applied for processes contaminated with any noise that is independent of the process – even for coloured noise, as e.g. in-band noise. Moreover, noise which is correlated to the process can be treated, if the time-scale on which the autocorrelation function of the noise decays is sufficiently small. For an oscillating process, e.g., this time scale must be smaller than the period of the process.

We propose two approaches to derive the decay properties: (1) The envelope of the autocorrelation function is fit by an exponential decay $A\phi^{\tau}$, with $\phi < 1$, and $\tau > 0$ the time lags. (2) The data is fit by an AR[2] in the state space model. From the parameters of the process, the decay rate ϕ can be determined. For our results we used the first approach. However, we used the idea of the second approach to compare the results from our fit to the true parameters derived from the process and its parameters.

From our simulations we conclude that the modification improves the statistics tremendously and should be used whenever the decay of the autocorrelation function is used for the block length selection.

In neuroscientific applications, often nonstationary processes are investigated. Typically local stationarity is assumed, when analysing these processes. The duration for which the process is treated as stationary can be related to the typical time scale of the process. Based on this time scale, i.e. the decay constant of the autocorrelation function, the optimal block length is determined. If this block length is estimated to be in the order of the duration during which the process is considered stationary, the block-bootstrap based statistics cannot be meaningfully applied. While other statistical tests, such as the t-test (Student, 1908) would provide a potentially wrong result, the proposed method of block-bootstrapping makes a reliable statement upon its applicability.

Acknowledgement

We thank Bernhardt Hellwig and Florian Amtage for the tremor data and the fruitful discussions.

Appendix A. Choice of simulation parameters

Autoregressive processes of order 2 (AR[2]) are the discrete analogue to damped stochastic driven harmonic oscillators. Rewriting the parameter Eqs. (13) and (14) to

$$a_1 = \phi \cos(2\pi f) \quad \text{(A.1)}$$

$$a_2 = -\phi^2 \quad \text{(A.2)}$$

the parameters are linked to the frequency $f = 1/T$ as the inverse of the period, and the decay rate $\phi = e^{1/\theta}$ as a function of the relaxation constant θ. For the first parameter a_1, we chose the frequency f from the tremor application. The frequency of the tremor was 5.59 Hz, which corresponds to 0.01118 of the Nyquist frequency 500 Hz. The time scale of autoregressive processes is sample points. In order to define a frequency, a reasonable assumption is to define one sample as one second. This yields a sampling rate of 1 Hz and thus a Nyquist frequency of 0.5 Hz. According to the tremor data, we chose $2\pi f = 0.01118\pi$ in Eq. (A.1).

To obtain the decay rate ϕ, used for both the first and the second autoregressive parameter, a_1 (Eq. (A.1)) and a_2 (Eq. (A.2)), we fit $A\phi^{\tau}$ to the estimated autocorrelation function (Eq. (4)) of the EMG-recording. This yielded $\phi = 0.999364$. These parameters were used as true parameters of the simulation studies based on autoregressive processes of order 2.

Finally, the variances of the intrinsic noise, σ^2, and measurement noise, Σ^2, respectively were chosen such that the noise to signal ratio matched the one observed from the EMG. This yielded the variances $\sigma^2 = 0.05^2$ and $\Sigma^2 = 30^2$.

Appendix B. Variance of AR[2]

The variance of an autoregressive process of order two (AR[2]) is

$$\mathrm{Var}\,[x] = a_1^2 \mathrm{E}[x_{j-1}^2] + a_2 \mathrm{E}[x_{j-2}^2] + 2a_1 a_2 \mathrm{E}[x_{j-1}x_{j-2}] + \mathrm{E}[\epsilon_j^2] \quad (\mathrm{B.1})$$

$$= a_1^2 \mathrm{Var}[x] + a_2 \mathrm{Var}[x] + 2a_1 a_2 \mathrm{E}[x_{j-1}x_{j-2}] + \sigma^2, \quad (\mathrm{B.2})$$

where a_k, $k = 1, 2$ are the process parameters, σ^2 the noise variance, and

$$\mathrm{E}[x_{j-1}x_{j-2}] = a_1 \mathrm{E}[x_{j-2}^2] + a_2 \mathrm{E}[x_{j-3}x_{j-2}] \quad (\mathrm{B.3})$$

$$= a_1 \mathrm{Var}[x] + a_2 \mathrm{E}[x_{j-1}x_{j-2}] \quad (\mathrm{B.4})$$

$$\Rightarrow \mathrm{E}[x_{j-1}x_{j-2}] = \frac{a_1}{1-a_2} \mathrm{Var}[x], \quad (\mathrm{B.5})$$

the autocovariance function of lag $\tau = 1$. With Eq. (B.2), this yields

$$\mathrm{Var}[x] - a_1^2 \mathrm{Var}[x] - a_2^2 \mathrm{Var}[x] - 2a_1 a_2 \frac{a_1}{1-a_2} \mathrm{Var}[x] = \sigma^2 \quad (\mathrm{B.6})$$

and after sorting,

$$\mathrm{Var}[x] = \frac{\sigma^2}{1 - a_1^2 - a_2^2 - 2(a_1^2 a_2 / 1 - a_2)}. \quad (\mathrm{B.7})$$

References

Arndt S, Cizadlo T, Andreasen NC, Heckel D, Gold S, O'Leary DS. Tests for comparing images based on randomization and permutation methods. Journal of Cerebral Blood Flow & Metabolism 1996];16(6):1271–9.

Bellec P, Rosa-Neto P, Lyttelton OC, Benali H, Evans AC. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. NeuroImage 2010];51(3):1126–39.

Carlstein E. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. Annals of Statistics 1986];14(3):1171–9.

Davison A, Hinkley D. Bootstrap methods and their application. Cambridge: Cambridge University Press; 1997].

Efron B. Bootstrap methods: Another look at the jackknife. Annals of Statistics 1979];7(1):1–26.

Efron B, Gong G. A leisurely look at the bootstrap, the jackknife, and cross-validation. American Statistician 1983];37(1):36–48.

Foster D, Bischof W. Bootstrap variance estimators for the parameters of small-sample sensory-performance functions. Biological Cybernetics 1987];57(4/5):341–7.

Gotman J. Automatic recognition of epileptic seizures in the EEG. Electroencephalography and Clinical Neurophysiology 1982];54(5):530–40.

Hall P, Horowitz J, Jing BY. On blocking rules for the bootstrap with dependent data. Biometrika 1995];82(3):561–74.

Hentschke H, Stüttgen MC. Computation of measures of effect size for neuroscience data sets. European Journal of Neuroscience 2011];34(12):1887–94.

Honerkamp J. Statistical physics: an advanced approach with applications. 3rd ed. Berlin: Springer; 2012].

Jachan M, Henschel K, Nawrath J, Schad A, Timmer J, Schelter B. Inferring direct directed-information flow from multivariate nonlinear time series. Physical Review E 2009];80(1 Pt 1):11138.

Künsch HR. The jackknife and the bootstrap for general stationary observations. Annals of Statistics 1989];17(3):1217–41.

Mammen E. When does bootstrap work? Asymptotic results and simulations. No. 77 in Lect Notes Stat. New York: Springer; 1992].

Peifer M, Schelter B, Guschlbauer B, Hellwig B, Lücking C, et al. On studentising and blocklength selection for the bootstrap on time series. Biometrical Journal 2005];47(3):346–57.

Percival D, Walden A. Spectral analysis for physical applications: multitaper and conventional univariate techniques. Cambridge: Cambridge University Press; 1993].

Schelter B, Winterhalder M, Timmer J, Peifer M. Testing for phase synchronization. Physics Letters A 2007];366(4/5):382–90.

Student. The probable error of a mean. Biometrika 1908];6(1):1–25.

Zoubir AM, Boashash B. The bootstrap and its application in signal processing. IEEE Signal Processing Magazine 1998];15(1):56–76.