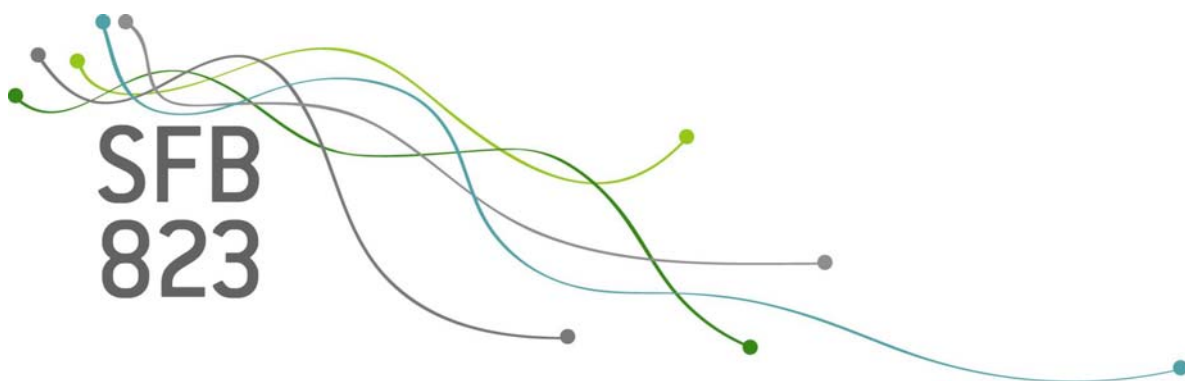# Predictive, finite-sample model choice for time series under stationarity and non-stationarity

Tobias Kley, Philip Preuß,
Piotr Fryzlewicz

Predictive, finite-sample model choice for time series under stationarity and non-stationarity

Tobias KLEY[a], Philip PREUß[b], Piotr FRYZLEWICZ[a],

[a]London School of Economics and Political Science

[b]Ruhr-Universität Bochum

November 14, 2016

Abstract

In statistical research there usually exists a choice between structurally simpler or more complex models. We argue that, even if a more complex, locally stationary time series model were true, then a simple, stationary time series model may be advantageous to work with under parameter uncertainty. We present a new model choice methodology, where one of two competing approaches is chosen based on its empirical finite-sample performance with respect to prediction. A rigorous, theoretical analysis of the procedure is provided. As an important side result we prove, for possibly diverging model order, that the localised Yule-Walker estimator is strongly, uniformly consistent under local stationarity. An R package, **forecastSNSTS**, is provided and used to apply the methodology to financial and meteorological data in empirical examples. We further provide an extensive simulation study and discuss when it is preferable to base forecasts on the more volatile time-varying estimates and when it is advantageous to forecast as if the data were from a stationary process, even though they might not be.

Keywords: forecasting, Yule-Walker estimate, local stationarity, covariance stationarity.

1 Introduction

A well-trodden path in applied statistical research is to propose a model believed to be a good approximation to the data-generating process, and then to estimate the model parameters with a view to performing a specific task, for example prediction. However, even

if the analyst were 'lucky' and chose the right model family, thereby reducing modelling bias, the resulting parameter estimators could be so variable that the selected model might well be sub-optimal from the point of view of the task in question. Choosing a slightly wrong model but with less variable parameter estimates may well lead to superior performance in, for example, prediction. This paper explores how this unsurprising but interesting phenomenon could and should affect model choice in time series analysis.

Choosing between stationary and non-stationary modelling is, typically, an important step in the analysis of time series data. Stationarity, which assumes that certain probabilistic properties of the time series model do not evolve over time, is a key assumption in time series analysis, and several excellent monographs focus on stationary modelling; see, e.g., Brillinger (1975), Brockwell and Davis (1991) or Priestley (1981). However, in practice, many time series are deemed to be better-suited for non-stationary modelling; this judgement can be based on diverse factors, such as, for example, visual inspection, formal tests against stationarity, or the observation that the data have been collected in a time-evolving environment and therefore are unlikely to have come from a stationary model.

One popular framework for the rigorous description of non-stationary time series models is that of local stationarity (Dahlhaus, 1997, 2012), in which the data are modelled locally as approximately stationary. We illustrate the main idea of the paper using a simple example of a locally stationary time series model, the time-varying autoregressive model (of order 1)

$$X_{t,T} = a(t/T)X_{t-1,T} + Z_t, \quad t = 1, ..., T,$$

with T denoting the sample size, $a : [0,1] \to (-1,1)$ being some suitable function and Z_t being an i.i.d. sequence with mean zero and variance one. Typically, to forecast future observations, one would require an estimate of $a(1)$, see e.g. Chen et al. (2010). Before constructing a suitable estimator, some analysts would wish to test if a was indeed time-varying, and there exist a vast amount of techniques to validate the assumption of a constant second-order structure in this framework; see von Sachs and Neumann (2000), Paparoditis (2009), Dwivedi and Subba Rao (2010), Paparoditis (2010), Dette et al. (2011), Nason (2013), Preuß et al. (2013) or Vogt and Dette (2015). If the process was found to be non-stationary, it would be tempting to estimate $a(1)$ by a localised estimate based on the most recent observations of $X_{t,T}$. This localisation would most likely reduce the bias of the estimator if the true dependency structure was indeed time-varying, but also increase its variance. However, if, for example, the function a was varying only slowly over time, this estimation procedure might result in sub-optimal estimation from the point of view of the mean squared prediction error, yielding inferior forecasts compared to the classical stationary AR(1) model. This would be particularly likely if the test of stationarity employed at the start was not constructed with the same performance measure in mind (i.e., mean squared prediction error) and was therefore 'detached' from the task in question (i.e., prediction). One of the findings of this paper is that even if the function a varied over time, one should in some cases treat it as constant in order to obtain smaller prediction errors, or in other words, 'prefer the wrong model' from the point of view of prediction.

The main aim of this paper is to propose an alternative model choice methodology in time series analysis that avoids the pitfalls of the above-mentioned process of testing followed by model choice. More precisely, our work has the following objectives:

- To propose a generic procedure for *finite-sample* model choice which avoids the path of hypothesis testing but instead chooses the model that offers better empirical finite-sample performance in terms of prediction on a validation set, with associated performance guarantees for the test set of yet unobserved data. Although the procedure is proposed and analysed theoretically in the framework of choice between stationarity and local stationarity and in the context of prediction, the procedure is applicable more generally whenever a decision needs to be made between two competing approaches, and can therefore be viewed as model- and problem-free.

- To suggest 'rules of thumb' indicating when the (wrong) stationary model may be preferred in a time-varying, locally stationary situation from the point of view of forecasting; and when a time-varying model should be preferred.

Our procedure validates and puts on a solid footing the possibly counter-intuitive observation that it is sometimes beneficial to choose the 'wrong' (but possibly simpler) model in time series analysis, if that model relies on more reliable estimators of its parameters than the right (but possibly more complex) model. While we stop short of conveying the message that simplicity in time series should always be preferred, part of our aim is to draw time series analysts' attention to the fact that particularly complex time series models may well appear attractive on first glance as they have the potential to capture features of the data well, but on the other hand can be so hard to estimate that this makes them inferior to simple and easy-to-estimate alternative models, even if the latter are wrong.

We now briefly describe related recent literature. The work of Xia and Tong (2011), who, while discussing time series prediction, select the model based on the minimisation of up to m-step ahead prediction errors (rather than the usual 1-step ahead ones) also appears to carry the general message that different models may be preferred for the same dataset depending on the task in question, or, in the language of the authors, on the 'features to be matched'. Besides similarities in this general outlook, our model-fitting methodology and the context in which it is proposed are entirely different. Forecasting in the presence of structural changes is a widely studied topic in the econometrics literature, see e.g. the comprehensive review by Rossi (2013) and the references therein. In particular, Giraitis et al. (2013) also use the minimisation of the 1-step ahead prediction error as a basis for model choice under non-stationarity, but, unlike us, do not consider the question of how this may lead to the preference for the 'wrong' model in finite samples. Instead of pursuing the cross-validation approach as in our work and in Giraitis et al. (2013), McDonald et al. (2016) evaluate the upper bound on the generalisation error in time series forecasting, and use its heuristically estimated version to guide model choice. We note, however, that this approach requires the estimation of some possibly difficult to estimate parameters, unlike cross-validation-based approaches.

The empirical mean squared prediction error (MSPE) which we will employ in our method is closely related to the population MSPE under parameter uncertainty. The strand of literature discussing this population quantity includes Baillie (1979) and Reinsel (1980), where approximating expressions were derived for stationary VAR time series. Further, Lewis and Reinsel (1988) compare the MSPEs of the optimal h-step predictor obtained in general r-dimensional linear processes and a wrongly assumed VARMA(p,q) model. They also determine an asymptotic approximation for the MSPE in a VAR(p) model when the parameters are estimated from the data of a general r-dimensional linear process. Note that the above references' results are restricted to stationary time series. For locally stationary tvMA(∞) processes, Palma et al. (2013) discuss optimal h-step ahead forecasting, in terms of the true model characteristics. Yet, they do not take parameter uncertainty into account.

While the main question we are concerned with is whether a stationary or a time-varying autoregressive model should be used for prediction, a nested question is what order the stationary or non-stationary model should have. Traditionally, order selection is done via minimisation of an information criterion, see, e.g., Brockwell and Davis (1991), p. 301. For stationary AR models Shibata (1980) and Bhansali (1996) discuss how to select the model order efficiently, which here means that asymptotically a lower bound is achieved for the MSPE. Zhang and Koreisha (2015) develop an adaptive criterion for model selection based on predictive risk. Peña and Sánchez (2007) derive and compare MSPE for univariate and multivariate predictors when the parameters are known. They then define and estimate a criterion (a measure of predictability) to choose between these two prediction options. Their approach is similar to ours in spirit, but, firstly, it chooses between univariate and multivariate models while we consider stationary and non-stationary models and, secondly, their methodology works with the population MSPE (which moves the focus away from the observed data to the postulated model), while we work with the corresponding empirical quantity directly. This difference in approaching the problem also holds for another, more general class of special-purpose-criteria: the focused information criteria (FIC), which were introduced in Claeskens and Hjort (2003). The FIC methodology with the focus on choosing the model best suited for prediction was then applied in the field of time series analysis in Claeskens et al. (2007), where the best AR(p) model for prediction is chosen, in Rohan and Ramanathan (2011), where the best ARMA(p,q) model for this purpose is chosen, and in Brownlees and Gallo (2008), where models for volatility forecasting are chosen. The idea of the FIC is that the model which minimises the *asymptotic* MSPE is the best one and the FIC is then based on an estimator of that asymptotic MSPE. Contrary to this, our approach is based on the *empirical* MSPE directly, which we believe to be the more relevant quantity in many applications. Contrary to the FIC which is based on the large-sample theory of the estimators involved, we provide finite-sample exponential bounds that imply a performance guarantee for our method. This approach can be advantageous, when it is preferred that the model choice also depends on the size of the sample, which in our view should be a natural requirement.

Our paper is organised as follows. In Section 2 we provide a simple motivating example. In Sections 3.1 and 3.2 we introduce and comment on our new time series model choice methodology. The statistical properties of our procedure are discussed in Section 3.3,

where also the performance guarantee (Theorem 3.1) is provided. The results of a simulation study and the analysis of three empirical examples can be found in Sections 4 and 5. In Section 6 we discuss statistical properties of the local Yule-Walker estimator and prove its strong uniform consistency under local stationarity (Corollary 6.2). We conclude with a summary in Section 7. In Appendix A technical conditions for the results from Sections 3.3 and 6 are listed and the main proofs are collected. Technical details of the proofs and tables and figures from the simulations section are gathered in an Online Supplement (Appendices B–I).

2 Motivating example

We consider the time-varying autoregressive (tvAR) model of order 2:

$$X_{t,T} = a_1(t/T)X_{t-1,T} + a_2(t/T)X_{t-2,T} + Z_t, \quad t = 1, ..., T,$$

where $a_1(u) := 0.15 + 0.15u$, $a_2(u) := 0.25 - 0.15u$, and Z_t is Gaussian white noise. $X_{t,T}$ is a non-stationary process which lies in the locally stationary class of Dahlhaus (1997). We will now compare different forecasting procedures for $X_{0.9T,T}$, where $T \in \{50, 500, 5000\}$. The predictor that minimises the mean squared prediction error is given by

$$\hat{X}_{0.9T,T}^{\text{oracle}} = 0.285 X_{0.9T-1,T} + 0.115 X_{0.9T-2,T}.$$

Yet, since in practice the underlying model is unknown, the analyst needs to

(1) make assumptions regarding the model, and

(2) estimate the assumed model's parameters.

For the purpose of this illustration, we discuss four possible models. In the first two models we falsely assume that the data were stationary and model $X_{t,T}$ to satisfy a traditional, autoregressive (AR) equation.

- In the first of the two cases we assume an AR(1) model and

- in the second case we assume the model to be an AR(2) model.

We further, discuss cases three and four, where the correct class of models (tvAR) is assumed. Yet,

- in case three, we falsely assume a tvAR(1) model, before

- in case four, we correctly assume the model to be a tvAR(2) model.

Note that the true model, the tvAR(2) model, is the most complex one of the four choices. In each of the models we estimate the parameters by solving the empirical Yule-Walker equations. In the case of the tvAR models we localise by using the segment
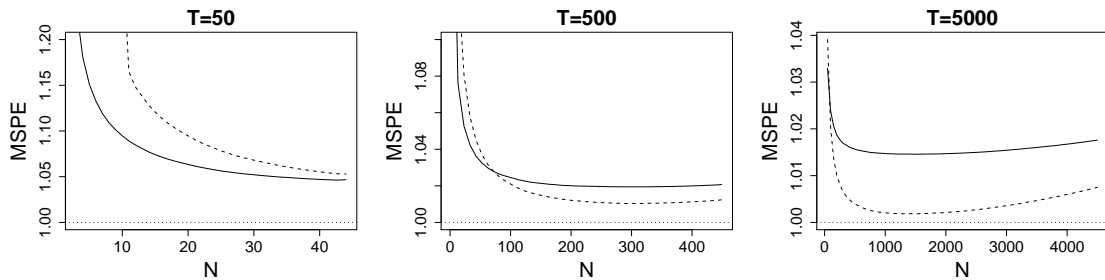
Figure 1: Mean squared prediction errors (MSPEs) for forecasting $X_{0.9T,T}$ with predictors $\hat{X}^{1,N}_{0.9T,T}$ and $\hat{X}^{2,N}_{0.9T,T}$ associated with tvAR(1) and tvAR(2) modelling of the data, where N varies. Left, middle and right column correspond to $T = 50$, $T = 500$ or $T = 5000$, respectively. The solid lines corresponds to $\mathbb{E}(\hat{X}^{1,N}_{0.9T,T} - X_{0.9T,T})^2$, the dashed line corresponds to $\mathbb{E}(\hat{X}^{2,N}_{0.9T,T} - X_{0.9T,T})^2$. The endpoints of each line indicate the MSPEs of the predictor associated with the stationary AR(1) and AR(2) models. The dotted lines at level 1.00 indicate the MSPE of the oracle predictor.

$X_{0.9T-N,T}, \ldots X_{0.9T-1,T}$. In the case of the traditional, stationary AR models we use all available observations $X_{1,T}, \ldots X_{0.9T-1,T}$. Details on the estimation are deferred to Section 3.1.

Denoting the localised Yule-Walker estimates of order 1 by $\hat{a}^{(1)}_{N,T}(0.9T - 1)$ and the ones of order 2 by $\hat{a}^{(2)}_{1;N,T}(0.9T - 1)$ and $\hat{a}^{(2)}_{2;N,T}(0.9T - 1)$ we obtain the predictors

$$\hat{X}^{1,N}_{0.9T,T} := \hat{a}^{(1)}_{N,T}(0.9T - 1)X_{0.9T-1,T},$$
$$\hat{X}^{2,N}_{0.9T,T} := \hat{a}^{(2)}_{1;N,T}(0.9T - 1)X_{0.9T-1,T} + \hat{a}^{(2)}_{2;N,T}(0.9T - 1)X_{0.9T-2,T},$$

where $\hat{X}^{1,N}_{0.9T,T}$ corresponds to the models of order 1 and $\hat{X}^{2,N}_{0.9T,T}$ corresponds to the models of order 2. The segment length N will be chosen as $0.9T - 1$ in the AR models and strictly smaller than this in the tvAR models.

In Figure 1, we observe that the predictors associated with the simpler, stationary AR model perform better or similarly well if $T = 50$ or $T = 500$. If $T = 5000$ the predictor associated with the locally stationary tvAR model performs visibly better in terms of its MSPE when the segment size N is chosen appropriately. In conclusion, this example illustrates how it can be advantageous to assume a wrong, but structurally simpler model when only a short time series is available. In particular, the model chosen should depend on the task at hand (here: prediction) and on the amount of data available. For $T = 50$ the best result is obtained by assuming the AR(1) model which is the simplest of the four candidates. When $T = 500$ the more complex AR(2) model becomes advantageous. Note that this model is more complex than the AR(1) model and thus provides a better approximation to the true tvAR(2) mechanism, but is still simplifying, because it does not take the time-varying characteristics into account at all. Only when even more data (here: $T = 5000$) are available, the variability of the parameter estimates of the tvAR(2) model is small enough not to overshadow the modelling bias, which in this example is rather small.

6

The important observation to be made here, thus, is that finding the 'right' model may not always be a suitable way of proceeding when it comes to the prediction of future observations. We point out that this observation was made in other contexts of time series analysis. For example, basic exponential smoothing is a widely used forecasting and trend extrapolation technique, and although it is well-known that it corresponds to standard Box-Jenkins forecasting in the $\mathrm{ARIMA}(0,1,1)$ model, it is also frequently used for data that does not follow it.

This paper investigates the question of what is the best model in terms of forecasting performance in the context of the choice between stationarity and non-stationarity. To ask this question explicitly instead of applying a test for stationarity is important since the smallest sample size T needed to reject the null hypothesis of stationarity may be smaller than the sample size needed to obtain improvement in terms of our task of interest, namely forecasting. In the following section, we will elaborate more on this question. Further, in Section 4, we see, as results of a simulation study, under which conditions using the true model is advantageous and when it can become disadvantageous.

3 When (not) to use locally stationary models under local stationarity: the new model choice methodology

3.1 Precise description of the procedure

We work in the framework of general locally stationary time series (a rigorous definition is deferred to Appendix A), in which the available data is a finite stretch $X_{1,T}, \ldots, X_{T,T}$ from an array $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ of random variables with mean zero and finite variances. Our aim is to determine a linear predictor for the unobserved $X_{T+h,T}$ from the observed $X_{1,T}, \ldots, X_{T,T}$.

Our proposal is to compare candidate h-step ahead predictors in terms of their empirical mean squared prediction error and choose the predictor with the best forecasting performance. To this end, we proceed as follows:

Step 1. Separate the final $2m$ observations from the T available observations. The observations with indices $M_0 := \{1, \ldots, T - 2m\}$, $M_1 := \{T - 2m + 1, \ldots, T - m\}$ and $M_2 := \{T - m + 1, \ldots, T\}$ will be referred to as the *(main part of the) training set*, the *final part of the training set* and the *validation set*, respectively. The set of unobserved data with the indices $M_3 := \{T + 1, \ldots, T + m\}$ will be referred to as the *test set*. The size m of the separated sets will be small in comparison to the sample size T (and hence also to the training set).

Step 2. Compute the linear 1-step ahead prediction coefficients

$$\hat{a}_{N,T}^{(p)}(t) := \big(\hat{\Gamma}_{N,T}^{(p)}(t)\big)^{-1} \hat{\gamma}_{N,T}^{(p)}(t) = \big(\hat{a}_{1;N,T}^{(p)}(t), \ldots, \hat{a}_{p;N,T}^{(p)}(t)\big)', \tag{1}$$

(a' denotes the transposed vector a) for $t + h \in M_1 \cup M_2$, $p = 1, \ldots, p_{\max}$, and $N \in \mathcal{N}$,

$$\hat{\Gamma}_{N,T}^{(p)}(t) := \big[\hat{\gamma}_{i-j;N,T}(t)\big]_{i,j=1,\ldots,p}, \quad \hat{\gamma}_{N,T}^{(p)}(t) := \big(\hat{\gamma}_{1;N,T}(t), \ldots, \hat{\gamma}_{p;N,T}(t)\big)' \tag{2}$$

7

and

$$\hat{\gamma}_{k;N,T}(t) := \frac{1}{N-|k|} \sum_{\ell=t-N+|k|+1}^{t} X_{\ell-|k|,T} X_{\ell,T}, \quad k = 0, \ldots, p_{\max}. \tag{3}$$

The parameters $p_{\max} \geq 1$ and $\mathcal{N} \subset \{p_{\max} + 1, \ldots, T - m - h\}$ are to be specified by the user. Comments on how they are to be chosen are deferred to Section 3.2.

Step 3. Compute $\hat{v}_{N,T}^{(p,h)}(t) := \big(\hat{v}_{1;N,T}^{(p,h)}(t), \ldots, \hat{v}_{p;N,T}^{(p,h)}(t)\big)$, the linear h-step ahead prediction coefficients, recursively,

$$
\begin{aligned}
\hat{v}_{i;N,T}^{(p,1)}(t) &:= \hat{a}_{i;N,T}^{(p)}(t), \quad i = 1, \ldots, p, \\
\hat{v}_{i;N,T}^{(p,\eta)}(t) &:= \hat{a}_{i;N,T}^{(p)}(t)\hat{v}_{1;N,T}^{(p,\eta-1)}(t) + \hat{v}_{i+1;N,T}^{(p,\eta-1)}(t)I\{i \leq p-1\}, \quad \eta = 2, 3, \ldots, h.
\end{aligned}
\tag{4}
$$

Step 4. Define $f_{t,h;0,N}^{\text{ls}}(X_{t,T}, \ldots, X_{1,T}) := 0$ and, for $p = 1, \ldots, p_{\max}$ and $N \in \mathcal{N}$, compute

$$f_{t,h;p,N}^{\text{ls}}(X_{t,T}, \ldots, X_{1,T}) := \sum_{i=1}^{p} \hat{v}_{i;N,T}^{(p,h)}(t)X_{t-i+1,T}, \quad p \geq 1, \tag{5}$$

the h-step ahead, plug-in forecast for $X_{t+h,T}$ typically used for $X_{t,T}$ that follow a tvAR(p) model. Further, define $f_{t,h;0}^{\text{s}}(X_{t,T}, \ldots, X_{1,T}) := 0$ and, for $p = 1, \ldots, p_{\max}$, compute

$$f_{t,h;p}^{\text{s}}(X_{t,T}, \ldots, X_{1,T}) := \sum_{i=1}^{p} \hat{v}_{i;t,T}^{(p,h)}(t)X_{t-i+1,T}, \quad p \geq 1, \tag{6}$$

the h-step ahead, plug-in forecast for $X_{t+h,T}$ typically used for $X_{t,T}$ that follow an AR(p) model.

Step 5. Amongst predictors (5) select $f_{t,h}^{\text{ls}} := f_{t,h;\hat{p}_{\text{ls}},\hat{N}}^{\text{ls}}$, with

$$(\hat{p}_{\text{ls}}, \hat{N}) := \arg \min_{\substack{p=0,\ldots,p_{\max} \\ N \in \mathcal{N}}} \sum_{t+h \in M_1} \big(X_{t+h,T} - f_{t,h;p,N}^{\text{ls}}(X_{t,T}, \ldots, X_{1,T})\big)^2,$$

and, amongst predictors (6) select $f_{t,h}^{\text{s}} := f_{t,h;\hat{p}_{\text{s}}}^{\text{s}}$, with

$$\hat{p}_{\text{s}} := \arg \min_{p=0,\ldots,p_{\max}} \sum_{t+h \in M_1} \big(X_{t+h,T} - f_{t,h;p}^{\text{s}}(X_{t,T}, \ldots, X_{1,T})\big)^2.$$

Note that $f_{t,h}^{\text{ls}}$ and $f_{t,h}^{\text{s}}$ are the forecasts of type (5) and (6) that minimise the empirical MSPE (on M_1) within the classes of tvAR and AR models of orders $p = 0, 1, \ldots, p_{\max}$, respectively.

Step 6. Use $f_{t,h}^{\text{ls}}$ as h-step ahead forecast of X_{t+h}, with $t + h > T$, if

$$\hat{R}_{T,j}^{\text{s,ls}}(h) := \frac{\text{MSPE}_{T,j}^{\text{s}}(h)}{\text{MSPE}_{T,j}^{\text{ls}}(h)} \geq 1 + \delta \tag{7}$$

holds for $j = 2$, and $f_{t,h}^{\text{s}}$ otherwise, where

$$\text{MSPE}_{T,j}^{*}(h) := \frac{1}{m} \sum_{t+h \in M_j} (X_{t+h,T} - f_{t,h}^{*})^2, \tag{8}$$

8

with $*$ indicating the corresponding model (we write 'ls' for the locally stationary approach and 's' for the stationary model) and $\delta \geq 0$ is a parameter by which the user of the procedure specifies which degree of superiority of the more complex procedure is required before it is preferred over the simpler alternative (cf. the end of Section 3.2).

By Theorem 3.1 we have that, with an appropriately chosen δ, the decision rule of type (7) will, with high probability, prefer the same models on the validation and on the test sets.

3.2 Remarks on the procedure

Some further explanations regarding the procedure are in order now. Our comments are organised according to the steps of the previous section.

Step 2. The coefficients (1) are estimates for the coefficient functions $a_1(t/T), \ldots, a_p(t/T)$ in the tvAR(p) model

$$X_{t,T} = \sum_{j=1}^{p} a_j(t/T) X_{t-j,T} + \sigma(t/T) Z_t, \quad t = 1, ..., T, \tag{9}$$

(see, for example, Dahlhaus and Giraitis (1998)). Recall that Z_t is usually assumed to be white noise and that $X_{t,T}$ is non-stationary if at least one of the functions a_j, $j = 1, \ldots, p$, or σ is non-constant.

We are interested in *linear forecasts* that will perform well for time series possessing a *general dependency structure*. The tvAR(p) model (9) is a natural choice to approximate the linear dynamics of the observed, non-stationary time series. This is the case, because the coefficient functions at time t/T coincide with the 1-step prediction coefficients (of order p) which define the best linear predictor. In Section 6, we show that $\hat{a}_{N,T}^{(p)}(t)$ from Step 2 are estimates for the 1-step linear prediction coefficients

$$\tilde{a}_T^{(p)}(t) := \arg \min_{a = (a_1, ..., a_p)' \in \mathbb{R}^p} \mathbb{E}\left[\left(X_{t,T} - \sum_{j=1}^{p} a_j X_{t-j,T}\right)^2\right],$$

also when the observations do not satisfy (9). A forecasting procedure derived within the tvAR(p) model can therefore be expected to behave reasonably, irrespective of whether the tvAR(p) model is true or 'just' an approximation to the truth. Note that we use the tvAR(p) model to approximate the dynamic structure of the data in Section 3.2 and most of our examples in Section 4 are of this kind, but we do *not* assume that the data actually satisfies it.

Step 3. Linear h-step ahead predictors can either be obtained by iterating the model equation (9) or by using a separate model for each h in which the indices of the sum on the right hand side run from $j = h, \ldots, p + h - 1$. These approaches have been referred to as the *plug-in method* and the *direct method*, respectively. A comparison of the two approaches can, for example, be found in Bhansali (1996), where results for a class of linear, stationary processes were derived. We employ the plug-in method.

Step 4. From the previous comments it can be seen how the predictors $f_{t,h;p,N}^{\text{ls}}$ and $f_{t,h;p}^{\text{s}}$ relate to the choice of modelling the time series' dynamics by a tvAR(p) or AR(p) model, respectively. In each of these model classes, increasing the order p will give a better approximation of the dynamics, but increase the complexity of the model, and make it more difficult to deal with under parameter uncertainty.

The parameters $p_{\max}$ and $\mathcal{N} \subset \{p_{\min} + 1, \ldots, T - m - h\}$ are to be chosen by the user and should depend on T. $N \in \mathcal{N}$ determines the degree of locality in the estimation of the coefficients. The parameters $p = 1, \ldots, p_{\max}$ and $N \in \mathcal{N}$ will influence the degree of bias and variance of the predictor. Our selection mechanism will balance them implicitly. The p corresponds to the order of the tvAR(p) model used to approximate the dynamics. A larger p will give a better approximation, but it will also be more difficult to estimate. A larger N will most likely reduce the variance, but, apart from exceptional cases, inflict a bias if non-stationarity is present.

Traditional choice of N. It is obvious that the variance of the estimator can decrease when a larger segment is used, but that the non-stationarity will potentially inflict an additional bias that increases with N. Under the condition that $N/T + T/N^2 = o(1)$, Dahlhaus and Giraitis (1998) derive asymptotic expansions for the local Yule-Walker estimator's bias and variance. They conclude that N should be chosen at the order of $T^{4/5}$, with the constant depending on the second derivatives of the true model quantities, which are unknown and difficult to estimate. The choice of $\mathcal{N}$ should thus, ideally, be such that $N \asymp T^{4/5}$, for all $N \in \mathcal{N}$. In practice, since the true model parameters are unknown, this rate provides very little guidance to the user of the method. We recommend, though, to adhere to two facts: the upper and lower bound of $\mathcal{N}$ should be bounded away from 0 and T. In other words, we recommend to choose $\mathcal{N}$ with $\min \mathcal{N}$ 'large enough', for the performance guarantee to be valid (cf. Theorem 3.1) and $\max \mathcal{N}$ being 'visibly smaller' than T, to ensure that there is a clear boundary between the locally stationary and the stationary approach.

Traditional choice of p. As described in the beginning of this section we use the tvAR(p)-model to *approximate* the dynamic structure of the data. Intuitively, we have that the larger the order p the better the approximation to the true dynamic structure. In opposition to the previously discussed question of how to choose the segment length N, we here have that a smaller p will inflict a modelling bias, while a larger p will typically be accompanied by an inflation of the variance of the estimation, because it implies that more parameters need to be estimated. Traditionally, the model order is chosen by minimizing information criteria as for example AIC or BIC. Claeskens et al. (2007) propose to use a version of the focused information criterion (FIC, see Claeskens and Hjort (2003)) to select the model order of a stationary AR(p) model optimal with respect to forecasting when the true model is known to be AR(∞). However, as mentioned in the introduction, the FIC-based methods employs an estimator of the asymptotic MSPE, while our approach is based on the empirical MSPE, which facilitates our focus on the finite sample performance.

Step 5 and 6. Our procedure performs two stages of selections via a training step followed by a validation step, as is common in machine learning. Firstly (in Step 5), it selects the model order p and, for the locally stationary approach, the segment length N by comparing predictors within each class of models under consideration (i.e., time-varying or non-time-varying autoregressive models). The parameters p and N are chosen such that the empirical MSPE (predicting M_1) is minimised. Secondly (in Step 6), a final competition of the winners is performed to select among the two classes of models. The procedure that minimises the empirical MSPE (predicting M_2) is selected and used for forecasting of the test set (M_3).

In our theoretical analysis of the next section (see, in particular, Theorem 3.1) we show that the proposed procedure will, with high probability, choose the same class of models on the validation as on the test set, implying that the procedure with the best empirical performance will be selected.

The parameter δ. We choose the predictor derived from the more complex, locally stationary model only if it performs at least $\delta \cdot 100\%$ better in forecasting the observations from the validation set than the simpler, stationary model. The user-selected parameter $\delta \geq 0$ specifies how much, in percentage terms, the user wishes to gain before deciding to use the more involved locally stationary approach. $\delta = 0$ corresponds to the decision rule whereby the time-varying model is chosen if it performs better or equally well on the validation set. If we choose δ such that $1 + \delta$ is larger (or smaller) than the true superiority of the locally stationary procedure over the stationary approach in terms of MSPE, then our procedure will, with high probabilty, choose the predictor with the best empirical performance on the test set (cf. Section 3.3).

3.3 Theoretical results

In this section, we establish theoretical results that will facilitate our analysis of the model choice suggested by decision rule (7). We show that the probability of choosing different models on the validation and the test set decays to zero at an exponential rate, which can be viewed as a "performance guarantee" of our model choice methodology.

Our results hold under reasonable assumptions on the observed data. We now provide the reader with an idea of the generality of the result by commenting briefly on each of the five main assumptions, the rigorous statement of which and further, more detailed comments are deferred to Appendix A so as not to impair the flow of the paper.

Heuristics of the main assumptions. By Assumption 1 the data are required to be locally stationary. The notion of local stationarity we impose goes beyond that of locally stationary linear processes and, in particular, we do not require the data to be tvAR. The constant $C \geq 0$ in Assumption 1 is associated with the degree of non-stationarity and will vanish for second order stationary processes. Assumption 2 is that the data be α-mixing at an exponential rate, which we require to prove the result of the probability in our result to be vanishing at an exponential rate as well. In Assumptions 3 and 4 we require that the local spectral density $f(u, \lambda)$ be bounded from below, by $m_f > 0$, and from above,

by $M_f < \infty$, and that $f(\cdot, \lambda)$ be continuously differentiable with $|\partial/\partial u f(u, \lambda)|$ being bounded by $M_f' \geq 0$, respectively. Finally, in Assumption 5, we require the moments to satisfy a Bernstein-type condition involving the constants c and d. The assumptions are non-restrictive in the sense that many popular and widely used time series models (tvARMA, tvGARCH, etc.) satisfy the full set of assumptions.

We are now ready to state the main result that guarantees that our procedure will, with high probability, choose the predictor that achieves the best empirical performance on the test set.

Theorem 3.1 *Let* $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ *satisfy Assumptions 1–5 (Appendix A) and* $\mathbb{E} X_{t,T} = 0$. *Let* $\delta \geq 0$, $m, p_{\max}, h \in \mathbb{N}^*$, *and* $\mathcal{N} \subset \{p_{\max} + 1, \dots, T - m - h\}$ *be such that*

$$\min_{\substack{p_1, p_2 = 0, \dots, p_{\max} \\ N \in \mathcal{N}}} \left| \text{MSPE}_{s_1/T, m/T}^{(p_1, h)}\left(\frac{s_1}{T}\right) - (1 + \delta) \cdot \text{MSPE}_{N/T, m/T}^{(p_2, h)}\left(\frac{s_1}{T}\right) \right| =: f(\delta) > 0, \quad (10)$$

where $s_1 := T - m - h + 1$ *and* $\text{MSPE}_{\Delta_1, \Delta_2}^{(p,h)}(u)$ *is defined in (48), as a function of the local moments of the process* $(X_{t,T})$. *Assume that*

$$T \geq \max\left\{ 3h2^h C_1 p_{\max}^2, 60f(\delta)^{-1}(1 + \delta)h2^{h+4}\left(C_0\right)^{2h+1} p_{\max}^2 (2\pi M_f' + C), \right.$$
$$\left. 40f(\delta)^{-1}(1 + \delta)m\left(2h + 1\right)\left(C_0\right)^{2h+1} M_f' \right\}, \quad (11)$$

where C_0, C_1 *are defined in (43) and* C *and* M_f' *are from Assumptions 1 and 4, respectively. Then, with* $\hat{R}_{T,i}^{\text{s,ls}}(h)$, $i = 1, 2$, *defined in (7), we have*

$$\mathbb{P}\left((\hat{R}_{T,2}^{\text{s,ls}}(h) \geq 1 + \delta \text{ and } \hat{R}_{T,3}^{\text{s,ls}}(h) \geq 1 + \delta) \text{ or } (\hat{R}_{T,2}^{\text{s,ls}}(h) < 1 + \delta \text{ and } \hat{R}_{T,3}^{\text{s,ls}}(h) < 1 + \delta) \right)$$
$$\geq 1 - 6p_{\max}^2 |\mathcal{N}| P_{m, \min \mathcal{N}}^{(p_{\max}, h)}\left(\frac{f(\delta)}{20(1 + \delta)} \right),$$

where $P_{m,N}^{(p,h)}$ *is defined in Lemma A.2 of Appendix A.*

Remark 3.2 *In a typical application the bound* $P_{m,N}^{(p,h)}(\varepsilon)$ *will be small. For example, if the parameters* $p = p(T)$, $h = h(T)$ *and* $\varepsilon = \varepsilon(T)$ *were bounded sequences with* $\varepsilon \geq \varepsilon_0 > 0$, *then*

$$P_{m,N}^{(p,h)}(\varepsilon) \leq K_1\left(\exp(-m^{1/(2+4d)} K_2) + m \exp(-N^{1/(2+2d)} K_3) \right),$$

where K_1, K_2, K_3 *are constants that do not depend on* m *or* N *and* d *is the constant from Assumption 5 (e.g., for sub-Gaussian processes:* $d = 1/2$*). Thus, if* $m, N \to \infty$ *(faster than logarithmically) and* $\log(m) = o(N^{1/(2+2d)})$, *as* $T \to \infty$, *then* $P_{m,N}^{(p,h)}(\varepsilon)$ *will vanish (at an exponential rate).*

In condition (11), we require the size T to be above a threshold. This requirement is for technical reasons in our proof and, in fact, our simulation results in Section 4 suggest that the probability bounded in Theorem 3.1 will be large also for T smaller than the threshold, as long as δ is chosen appropriately. More precisely, we require that $f(\delta)$ will be non zero and the larger it is, the larger the probability will be. The quantity $f(\delta)$ is constructed to measure the difference between the MSPEs of the stationary predictors for

different p_1 and the MSPEs of the locally stationary predictors for different (p_2, N) scaled by a factor of $1+\delta$. It is obvious that a condition like this is required for consistency of the procedure, because if there is no difference in performance either approach may equally likely be chosen. Note that in the situation where both approaches perform equally well we do not need the selection to be consistent. Further, note that the condition in our result is slightly stronger than necessary, as we do not only require the best performing procedures to perform differently for the p_1 and (p_2, N) that yield the best result, but we require it for any combination. This is due to our method of proof.

The quantity $f(\delta)$ depends on the model under consideration and, as $p_{\max}$ and $|\mathcal{N}|$ get larger, may potentially tend to zero. Thus, to employ Theorem 3.1 in practice, one has to analyse $p_{\max}^2 |\mathcal{N}| P_{m,\min \mathcal{N}}^{(p_{\max}, h)}\left(\frac{f(\delta)}{1+\delta}\right)$, whether it will vanish and if so, at what rate this happens. We expect that, if $p_{\max}$ is kept fix, $\min \mathcal{N} \to \infty$ and $\max \mathcal{N} = o(T)$ that, in many cases, $f(\delta)$ can be bounded away from zero. Numerical illustrations which suggest this are provided in Section 4.

To illustrate the usefulness of Theorem 3.1, we derive two corollaries of the result. The corollaries consider a simplified version of our procedure in which the model order is always chosen as $p = 1$, for both the stationary and the locally stationary approach. Restricting ourselves to $p = 1$ allows us to formulate the equivalent of Theorem 3.1 that rephrases assumption (10) in more explicit terms. To make the modified procedure precise: we now consider the procedure where, after Step 1–4, as described in Section 3.1, we continue with Steps 5' and 6', as follows:

Step 5'. Amongst the predictors (5) select $\tilde{f}_{t,1}^{\mathrm{ls}} := f_{t,1;\hat{p}_{\mathrm{ls}},\hat{N}}^{\mathrm{ls}}$, with $\hat{p}_{\mathrm{ls}} := 1$ and

$$\hat{N} := \arg \min_{N \in \mathcal{N}} \sum_{t+1 \in M_1} \left(X_{t+1,T} - f_{t,1;1,N}^{\mathrm{ls}}(X_{t,T}, \ldots, X_{1,T})\right)^2,$$

and, amongst the predictors (6) select $\tilde{f}_{t,1}^{\mathrm{s}} := f_{t,1;\hat{p}_{\mathrm{s}}}^{\mathrm{s}}$, with $\hat{p}_{\mathrm{s}} := 1$. Note that $\tilde{f}_{t,1}^{\mathrm{ls}}$ and $\tilde{f}_{t,1}^{\mathrm{s}}$ are the forecasts of type (5) and (6) that minimise the empirical MSPE (on M_1) within the classes of tvAR(1) and AR(1), respectively.

Step 6'. Use $\tilde{f}_{t,1}^{\mathrm{ls}}$ as 1-step ahead forecast of X_{t+1}, with $t + 1 > T$, if

$$\tilde{R}_{T,j}^{\mathrm{s,ls}}(1) := \frac{\widetilde{\mathrm{MSPE}}_{T,j}^{\mathrm{s}}(1)}{\widetilde{\mathrm{MSPE}}_{T,j}^{\mathrm{ls}}(1)} \geq 1 + \delta \tag{12}$$

holds for $j = 2$, and $\tilde{f}_{t,1}^{\mathrm{s}}$ otherwise, where

$$\widetilde{\mathrm{MSPE}}_{T,j}^{*}(1) := \frac{1}{m} \sum_{t+1 \in M_j} (X_{t+1,T} - \tilde{f}_{t,1}^{*})^2. \tag{13}$$

The modified procedure, based on steps 1–4 and 5'–6', is then consistent (a) if δ is chosen large enough or (b) in case the true model is non-stationary, if δ is chosen small enough. To make the non-stationarity requirement in the second case precise, we define

$$D_{\sup} := \sup_{\frac{T-m-h+1}{T} \leq u \leq \frac{T-h+1}{T}} \left| \frac{\int_0^1 \gamma_1\left(u + \frac{s_1}{T}(x-1)\right) \mathrm{d}x}{\int_0^1 \gamma_0\left(u + \frac{s_1}{T}(x-1)\right) \mathrm{d}x} - \frac{\gamma_1(u)}{\gamma_0(u)} \right|, \tag{14}$$

and

$$D_{\inf} := \inf_{\frac{T-m-h+1}{T} \leq u \leq \frac{T-h+1}{T}} \left| \frac{\int_0^1 \gamma_1\big(u + \frac{s_1}{T}(x-1)\big)\mathrm{d}x}{\int_0^1 \gamma_0\big(u + \frac{s_1}{T}(x-1)\big)\mathrm{d}x} - \frac{\gamma_1(u)}{\gamma_0(u)} \right|, \tag{15}$$

the maximum and minimum difference between the ratio of local auto-covariances $\gamma_0(u)$, $\gamma_1(u)$, from Assumption 1, at any point u of the validation set, and the ratio of averaged auto-covariances (averaged across the training set). Note that $D_{\sup} \leq 2$ and that $D_{\inf}$ is a measure for the non-stationarity of the training set. In particular, it will vanish if the data stems from a stationary process. We will further require that the local autocorrelations of order 1 for the validation set are bounded and to this end define

It is important to observe that while the quantity $f(\delta)$ in Theorem 3.1 depends on $\mathcal{N}$ we have that the D's only depend on m and T and the first and second order local moments (γ_1 and γ_0). The definition does not require that the data actually stems from a tvAR(1) process, but if the true model were indeed a tvAR(1) with coefficient function a, then we have $\gamma_k(u) = a(u)^k/(1 - a(u)^2)$. In this important class of models the integrals in the first term of the D's thus are explicit in terms of the function a, as

$$\int_0^1 \gamma_k(u + \Delta(x-1))\mathrm{d}x = \Delta^{-1} \int_{u-\Delta}^u \left(\frac{a(w)^k}{1 - a(w)^2} \right) \mathrm{d}w.$$

We now state two results about the modified procedure of 1-step ahead forecasting. The first result illustrates that the procedure will always be consistent if δ is chosen large enough:

Corollary 3.3 *Let $(X_{t,T})_{t\in\mathbb{Z}, T\in\mathbb{N}^*}$ satisfy Assumptions 1–5 (Appendix A), $\mathbb{E}X_{t,T} = 0$, and*

$$\rho := \sup_{\frac{T-m-h+1}{T} \leq u \leq \frac{T-h+1}{T}} \left| \frac{\gamma_1(u)}{\gamma_0(u)} \right| < 1.$$

Let $m \in \mathbb{N}^*$, $\mathcal{N} \subset \{2, \ldots, T - m - 1\}$ and $\delta \geq 2D_{\sup}^2/(1 - \rho^2)$, with $D_{\sup}$ defined in (14), $\gamma_k(u)$, $k = 0,1$, from Assumption 1, m_f from Assumption 3 and M_f' from Assumption 4. Then, for T large enough to satisfy condition (11) with $p_{\max} = h = 1$ and $f(\delta) := \delta\pi m_f(1 - \rho^2)$, with $\tilde{R}_{T,i}^{s,ls}(1)$, $i = 1,2$, defined in (12), we have

$$\mathbb{P}\Big((\tilde{R}_{T,2}^{s,ls}(1) \geq 1 + \delta \ \text{and} \ \tilde{R}_{T,3}^{s,ls}(1) \geq 1 + \delta) \ \text{or} \ (\tilde{R}_{T,2}^{s,ls}(1) < 1 + \delta \ \text{and} \ \tilde{R}_{T,3}^{s,ls}(1) < 1 + \delta) \Big)$$

$$\geq 1 - 6|\mathcal{N}|P_{m,\min\mathcal{N}}^{(1,1)}\Big(\frac{\delta}{(1+\delta)} \frac{\pi m_f}{20}(1 - \rho^2) \Big),$$

where $P_{m,N}^{(p,h)}$ is defined in Lemma A.2 of Appendix A.

Further more, we have as a second result that if the true model is non-stationary in the sense that the quantity $D_{\inf}$ is large compared to the N/T for all $N \in \mathcal{N}$, then we also have consistency for δ's that are small enough:

Corollary 3.4 *Let $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ satisfy Assumptions 1–5 (Appendix A), $\mathbb{E}X_{t,T} = 0$, and $D_{\inf} > 0$, with $D_{\inf}$ defined in (15). Let $m \in \mathbb{N}^*$, and $\mathcal{N} \subset \{2,\ldots,T-m-1\}$, and $\delta \leq \frac{1}{8}D_{\inf}^2$, Then, with m_f from Assumption 3 and M_f' from Assumption 4, if*

$$D_{\inf}^2 \geq 2\Big(\frac{M_f'}{m_f}\frac{\max\mathcal{N}}{T}\Big)^2, \tag{16}$$

and T is large enough to satisfy condition (11) with $p_{\max} = h = 1$ and $f(\delta) := \pi D_{\inf}^2 m_f/2$, with $\tilde{R}_{T,i}^{s,ls}(h)$, $i = 1,2$, defined in (12), we have

$$\mathbb{P}\Big((\tilde{R}_{T,2}^{s,ls}(1) \geq 1+\delta \ \text{and} \ \tilde{R}_{T,3}^{s,ls}(1) \geq 1+\delta) \ \text{or} \ (\tilde{R}_{T,2}^{s,ls}(1) < 1+\delta \ \text{and} \ \tilde{R}_{T,3}^{s,ls}(1) < 1+\delta)\Big)$$

$$\geq 1 - 6|\mathcal{N}|P_{m,\min\mathcal{N}}^{(1,1)}\Big(\frac{\pi m_f D_{\inf}^2}{40(1+\delta)}\Big),$$

where $P_{m,N}^{(p,h)}$ is defined in Lemma A.2 of Appendix A.

Note that, in Corollaries 3.3 and 3.4, to show that ε of $P_{m,\min\mathcal{N}}^{(1,1)}(\varepsilon)$ is bounded away from zero, it suffices to bound $D_{\inf}$ and Remark 3.2 then fully applies, yielding the exponential bound.

4 Simulations

In this section we discuss finite sample properties of the estimates $\hat{R}_{T,i}^{s,ls}(h)$, defined in (7), and their population counterparts $R_{T,j}^{s,ls}(h) := (\mathbb{E}(\text{MSPE}_{T,j}^s(h)))/(\mathbb{E}(\text{MSPE}_{T,j}^{ls}(h)))$. The simulation study was conducted with the R package forecastSNSTS (R Core Team, 2016; Kley et al., 2016), available from The Comprehensive R Archive Network (CRAN). In particular, we investigate the performance of decision rule (7). To this end, we have considered 15 different tvAR models. Three of the models are stationary, the other 12 are non-stationary. Amongst the non-stationary processes we have some where the covariance structure changes quickly and some where the covariances change slowly. Further, we will have examples where the processes given by the parameters at some local time u are almost unit root and some where they are not.

For each of the models we proceed as follows. We simulate sequences of length $T + m = n \in \{100, 200, 500, 1000, 2000, 4000, 6000, 8000, 10000\}$. The $T + m$ observations, with T and m as in Section 3, contain the training, validation and test set. We separate the test and validation set of length $m := \lfloor n^{.85}/4\rfloor$. Thus, $n_i := n - (3-i)\lfloor n^{.85}/4\rfloor$, $i = 0,...,3$, mark the end indices of the main part of the training set, the final part of the training set, the validation set and the test set, respectively. We have chosen m as a function of n in such a way that $m = o(n)$ and $m \to \infty$, as $n \to \infty$. The sizes of the three sets therefore are $12, 22, 49, 88, 159, 288, 406, 519$, and 627 for the different sequence sizes, respectively.

As described in Section 3.1 we then, for any $h = 1,\ldots,H := 10$, determine linear h-step ahead predictions for $X_{t+h,T}$ with $t + h \in \{n_0 + 1,\ldots,n_1\}$. We determine the 'stationary predictions', with coefficients estimated for a given $p = 0,\ldots,p_{\max} := 7$, from

$X_{1,T}, \ldots, X_{t,T}$ by $\hat{v}_{t,T}^{(p,h)}(t)$ from Step 3 of the procedure. For simplicity, we have chosen the same $p_{\max}$ for every T. We further determine 'locally stationary predictions' where the coefficients $\hat{v}_{N,T}^{(p,h)}(t)$ are used for $p = 0, \ldots, p_{\max}$ and

$$\mathcal{N} = \{N := N_{\min} + i\lceil (N_{\max} - N_{\min})/25\rceil \; : \; i \in \mathbb{N}, N \leq N_{\max}\},$$

where $N_{\min} := \lfloor (n/2)^{4/5} \rfloor$ and $N_{\max} := \lfloor n^{4/5} \rfloor$. Note that $\mathcal{N}$ always contains 25 (or fewer) equally spaced elements, the smallest being $N_{\min}$. We limit the number of elements in $\mathcal{N}$ to reduce computation time and point out that results did not change significantly when a larger number of elements was used. We then compare the predictors with respect to their empirical mean squared prediction error (MSPE) on the final part of the training set and, according to Step 5 of the procedure, choose the stationary predictor with $\hat{p}_s$ that minimises the MSPE on M_1 amongst all stationary predictors and the locally stationary predictor with $(\hat{p}_{ls}, \hat{N}_{ls})$ which minimizes the MSPE on M_1 amongst all the locally stationary predictors.

For those two predictors we then determine the empirical mean squared prediction errors $\mathrm{MSPE}_{T,2}^*$ and $\mathrm{MSPE}_{T,3}^*$, defined in (8), on the validation and test set, respectively. We record seven pieces of information: $\hat{p}_s$, $\hat{p}_{ls}$, $\hat{N}_{ls}$, $\mathrm{MSPE}_{T,2}^s$, $\mathrm{MSPE}_{T,2}^{ls}$, $\mathrm{MSPE}_{T,3}^s$, and $\mathrm{MSPE}_{T,3}^{ls}$. We replicate the experiment 10000 times.

Now we define the models. The first two tvAR(1) models (the innovations Z_t are i. i. d Gaussian white noise, in all 15 models) are defined by two periodic coefficient functions, namely the models are

$$X_{t,T} = (0.8 + 0.19\sin(4\pi\frac{t}{T}))X_{t-1,T} + Z_t, \tag{17}$$

$$X_{t,T} = (0.3 + 0.19\sin(4\pi\frac{t}{T}))X_{t-1,T} + Z_t. \tag{18}$$

We then look at six tvAR(1) models where the AR coefficient increases linearly, namely

$$X_{t,T} = (0.5 + 0.19\frac{t}{T})X_{t-1,T} + Z_t, \quad (19) \qquad X_{t,T} = (0.9 + 0.09\frac{t}{T})X_{t-1,T} + Z_t, \quad (22)$$

$$X_{t,T} = (0.5 + 0.09\frac{t}{T})X_{t-1,T} + Z_t, \quad (20) \qquad X_{t,T} = (0.5 + 0.49\frac{t}{T})X_{t-1,T} + Z_t, \quad (23)$$

$$X_{t,T} = (0.8 + 0.19\frac{t}{T})X_{t-1,T} + Z_t, \quad (21) \qquad X_{t,T} = (0.5 + 0.4\frac{t}{T})X_{t-1,T} + Z_t. \quad (24)$$

Finally, we consider the stationary AR(1) model with coefficient -0.6 and two models with independent observations, one with and one without heteroscedasticity, namely

$$X_{t,T} = Z_t, \tag{26}$$

$$X_{t,T} = -0.6X_{t-1,T} + Z_t, \qquad (25) \qquad X_{t,T} = \left(5 - 16|\frac{t}{T} - 0.5|^2\right)Z_t, \qquad (27)$$

We further consider the following tvAR(2) model from Dahlhaus (1997)

$$X_{t,T} = 1.8\cos(1.5 - \cos(4\pi\frac{t}{T}))X_{t-1,T} - 0.81X_{t-2,T} + Z_t \tag{28}$$

and one of its tangent processes (cf. the comment after Assumption 1)

$$X_{t,T} = X_{t-1,T} - 0.81X_{t-2,T} + Z_t. \tag{29}$$

Further, we consider two tvAR(1) models where the coefficient decreases linearly:

$$X_{t,T} = (0.99 - 0.49t/T)X_{t-1,T} + Z_t \qquad (30)$$

$$X_{t,T} = (0.5 - t/T)X_{t-1,T} + Z_t. \qquad (31)$$

We now discuss our analysis of process (17), the first of the models, in detail and comment briefly on the outcome for the other processes with tables and figures for the other models being deferred to Appendix I (Online Supplement).

In Figure 2, note that, since in the numerator we have the MSPE for the best stationary predictor and in the denominator the MSPE for the best locally stationary predictor, a ratio above 1 corresponds to the situation where the best locally stationary predictor outperforms the best stationary predictor. It can be seen whether this happens on average, while in Table 2 we can see the proportion of simulated cases in which this has happens. In Figure 2, we thus observe that, for $n = 100$, the stationary approach performs better on average across all values of h on both the test and validation set. For $n = 200$ the locally stationary approach performs better for $3 \leq h \leq 6$ on the test set, while the stationary approach still excels for all h on the validation set. For $n \geq 500$ the locally stationary approach is better across all values of h on the test set and for $2 \leq h \leq 4$ it outperforms the stationary approach on the validation set. For $n \geq 1000$ the locally stationary approach is always as least as good as the stationary approach for all h. It is striking that, for this particular model and for the larger n's we see that as h gets larger the two approaches (stationary and locally stationary) perform almost equally well on average, which can be seen from the lines in Figure 2 being close to one. Another important observation is that, as n gets larger and m/n gets smaller, we see the lines for the validation and test set converging, which is in line with what Theorem 3.1 suggests should happen.

We now, briefly, compare the outcome of model (17) to that of model (18); details are shown in Appendix I (Online Supplement). Note that in model (17) the coefficient function ranges from 0.61 to 0.99, placing some of its tangent processes close to the unit root. In model (18) the coefficient function ranges from 0.11 to 0.49. Thus, the two models have the same variation of the coefficient function, but in model (18) the tangent processes are further away from the unit root. In Figure 9, it can be seen that the stationary approach is preferred over the locally stationary approach for sequences up to length $n = 1000$. Further, we observe that the advantage of using the locally stationary approach for sequences of length $n \geq 4000$ is minuscule and visible only for 1-step ahead forecasting. For the other models we can make similar observations:

Rules of Thumb. The locally stationary approach outperforms the stationary approach only if either the sequence is long, or the coefficient function varies a lot, or the tangent processes are close to the unit root. In any other case the stationary approach can be chosen without (a large) loss.

The proportions shown in Table 3 provide information on the consistency of the procedure, as we see the proportion of cases in which the same procedure (stationary or locally stationary) is chosen on both the test and validation set. This validates Theorem 3.1 for
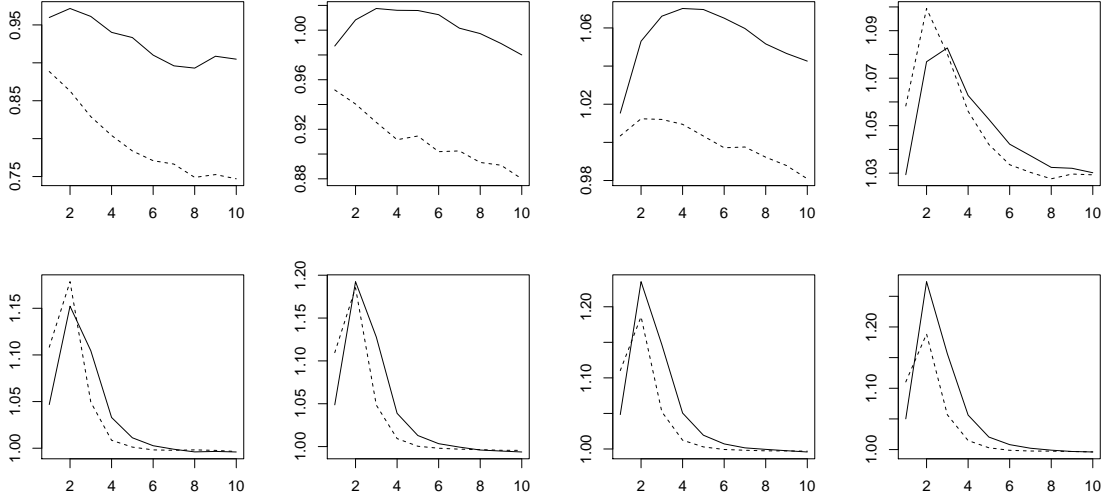
Figure 2: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (17) and different values of n (from left to right: n=100, n=200, n=500, n=1000 [first row], n=4000, n=6000, n=8000, n=10000 [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.1273	0.2765	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.0994	0.225
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1369	0.4593	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.1042	0.5714

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	7e-04	0.0056	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0.07	0.0671
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	0.0053	0.9884	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	0.0652	0.7977

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.9925	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.7577	0.2045
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.0075	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.0252	0.0126

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0.4961	4e-04	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0.0018	0.1697
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0.5033	2e-04	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0.0025	0.826

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0.0034	8e-04	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	1e-04	0.012
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0.7029	0.2929	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	1e-04	0.9878

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.15	< 1.15		≥ 1.15	< 1.15
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.15$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.15$	0	0.001
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.15$	0.0185	0.9815	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.15$	2e-04	0.9988

Table 1: *Proportions of the individual events in (32) for the process (17) and selected combinations of n and δ.*

	(32) holds for $h=1$, $i=2$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.2934	0.3364	0.5013	0.8595	0.9959	0.9999	1	1	1
0.01	0.2642	0.2971	0.4316	0.814	0.9924	0.9999	1	1	1
0.05	0.1814	0.1713	0.2182	0.5585	0.9058	0.9851	0.9958	0.9985	0.9994
0.1	0.1209	0.084	0.0747	0.2146	0.4628	0.6176	0.6691	0.6989	0.7063
0.15	0.082	0.0437	0.0193	0.0431	0.0673	0.0606	0.0432	0.0299	0.0185
0.2	0.0558	0.0234	0.0053	0.006	0.0042	6e-04	0	0	0
0.4	0.016	0.0031	8e-04	2e-04	4e-04	1e-04	0	0	0
0.6	0.0074	0.0015	5e-04	2e-04	3e-04	0	0	0	0
	(32) holds for $h=1$, $i=3$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.4382	0.477	0.6085	0.7049	0.8148	0.9318	0.9698	0.9831	0.9925
0.01	0.4038	0.4303	0.5525	0.6496	0.7542	0.877	0.9303	0.9574	0.9759
0.05	0.2914	0.2723	0.3305	0.4001	0.4206	0.4639	0.4843	0.4778	0.4965
0.1	0.1987	0.1485	0.1359	0.1455	0.0957	0.0462	0.0196	0.0112	0.0042
0.15	0.134	0.079	0.0472	0.0324	0.0094	3e-04	2e-04	0	0
0.2	0.09	0.0412	0.017	0.0063	5e-04	1e-04	0	0	0
0.4	0.0223	0.0054	0.0011	4e-04	4e-04	1e-04	0	0	0
0.6	0.0085	0.0022	8e-04	3e-04	4e-04	1e-04	0	0	0
	(32) holds for $h=5$, $i=2$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.2145	0.2731	0.3195	0.3082	0.1759	0.1895	0.2479	0.3013	0.3577
0.01	0.2036	0.2534	0.2972	0.3012	0.1592	0.1376	0.1625	0.1838	0.2089
0.05	0.1666	0.1858	0.2181	0.2671	0.1191	0.0305	0.0122	0.0072	0.0043
0.1	0.1288	0.1279	0.1464	0.2261	0.0975	0.0131	0.0013	2e-04	2e-04
0.15	0.1049	0.0898	0.0989	0.18	0.0834	0.0117	0.0011	0	2e-04
0.2	0.088	0.0635	0.0684	0.1352	0.0685	0.01	9e-04	0	2e-04
0.4	0.0477	0.0199	0.0134	0.0252	0.01	0.0013	1e-04	0	0
0.6	0.0284	0.0064	0.0041	0.0033	4e-04	0	0	0	0
	(32) holds for $h=5$, $i=3$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.3404	0.4694	0.5116	0.3117	0.2059	0.2918	0.3921	0.4706	0.537
0.01	0.3244	0.4467	0.4921	0.3031	0.1855	0.2516	0.3438	0.4222	0.4909
0.05	0.2684	0.3653	0.4142	0.2677	0.127	0.0998	0.1161	0.148	0.1715
0.1	0.2165	0.2784	0.322	0.2233	0.0849	0.0242	0.0154	0.0119	0.0121
0.15	0.1733	0.2161	0.2432	0.1781	0.0622	0.0098	0.0023	7e-04	0.001
0.2	0.1393	0.1598	0.1779	0.1371	0.0448	0.0065	8e-04	0	0
0.4	0.0646	0.0536	0.0483	0.0387	0.0081	5e-04	0	0	0
0.6	0.0334	0.0174	0.0131	0.0075	4e-04	0	0	0	0

Table 2: *Proportion of (7) being fulfilled for the process (17) and different values of h, δ and n.*

	(32) holds for $h = 1$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.5718	0.5422	0.4948	0.638	0.8119	0.9317	0.9698	0.9831	0.9925
0.01	0.5866	0.561	0.4897	0.5884	0.7502	0.8769	0.9303	0.9574	0.9759
0.05	0.667	0.6744	0.6125	0.5054	0.441	0.466	0.4845	0.4783	0.4963
0.1	0.7592	0.8079	0.8212	0.7173	0.5409	0.3984	0.3391	0.3057	0.2963
0.15	0.828	0.8929	0.9395	0.9295	0.9253	0.9395	0.9566	0.9701	0.9815
0.2	0.8814	0.9428	0.9803	0.9891	0.9961	0.9995	1	1	1
0.4	0.9703	0.9941	0.9995	0.9998	1	1	1	1	1
0.6	0.9887	0.9983	0.9993	0.9999	0.9999	0.9999	1	1	1
	(32) holds for $h = 5$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.6617	0.6033	0.6639	0.8889	0.8874	0.8015	0.774	0.7699	0.7703
0.01	0.6708	0.6131	0.6601	0.8847	0.8895	0.788	0.7345	0.6874	0.655
0.05	0.7074	0.6531	0.6669	0.871	0.9287	0.9011	0.8805	0.849	0.8278
0.1	0.7457	0.7207	0.7104	0.8598	0.9546	0.9825	0.9853	0.9879	0.9879
0.15	0.7872	0.7753	0.7655	0.8593	0.9552	0.9939	0.998	0.9993	0.9988
0.2	0.8195	0.8243	0.8193	0.8677	0.9553	0.9945	0.9991	1	0.9998
0.4	0.9029	0.9333	0.9475	0.9477	0.9875	0.9984	0.9999	1	1
0.6	0.9442	0.9774	0.9844	0.9896	0.9992	1	1	1	1

Table 3: *Proportion of (32) being fulfilled for the process (17) and different values of h, δ and n.*

	Value of $f(\delta)$, defined in (10), for $h = 1$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0	0	0	0	0	0	0	0	0
0.01	5.3e-05	2.4e-06	0.019	0.02	0.021	0.021	0.021	0.022	0.022
0.05	5.1e-05	8.3e-05	2.4e-05	0.0078	0.0095	0.0047	0.0018	1.8e-06	1.9e-05
0.1	0.0023	0.00059	1.4e-07	0.0056	0.01	0.016	0.018	0.02	0.021
0.15	0.053	0.048	0.05	0.056	0.061	0.067	0.069	0.07	0.071
0.2	0.1	0.098	0.1	0.11	0.11	0.12	0.12	0.12	0.12
0.4	0.31	0.3	0.3	0.31	0.31	0.32	0.32	0.32	0.32
0.6	0.13	0.18	0.32	0.41	0.46	0.5	0.52	0.53	0.53
	Value of $f(\delta)$, defined in (10), for $h = 5$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0	0	0	0	0	0	0	0	0
0.01	9e-05	0.00037	0.00045	0.0017	0.021	0.021	0.021	0.022	0.022
0.05	5.8e-05	0.00034	0.047	0.037	0.029	0.022	0.017	0.014	0.011
0.1	0.00074	0.00058	0.073	0.13	0.13	0.12	0.12	0.12	0.11
0.15	0.00018	0.00025	0.0015	0.17	0.14	0.12	0.11	0.098	0.091
0.2	0.00032	0.00013	0.00031	0.069	0.04	0.014	0.00014	0.0096	0.0011
0.4	0.016	0.081	0.12	0.16	0.19	0.22	0.23	0.24	0.25
0.6	0.35	0.44	0.49	0.55	0.59	0.62	0.64	0.66	0.66

Table 4: *Values of $f(\delta)$, defined in (10), for the process (17) and different values of h, δ and n.*

the example. It is interesting to compare the observed proportions with the corresponding value of $f(\delta)$, which we provide in Table 4. We see that a larger proportion typically goes along with a larger value of $f(\delta)$ indicating the relevance of condition (10). To make it more precise: the tables are concerned with the proportion for which the decision rule (7) yields the same result no matter if we take $i = 2$ or $i = 3$, i.e. we count what proportion of runs satisfies

$$(\hat{R}_{T,2}^{s,ls}(h) \geq 1 + \delta \text{ and } \hat{R}_{T,3}^{s,ls}(h) \geq 1 + \delta) \text{ or } (\hat{R}_{T,2}^{s,ls}(h) < 1 + \delta \text{ and } \hat{R}_{T,3}^{s,ls}(h) < 1 + \delta). \quad (32)$$

We see that if δ is chosen large enough the probability for the event (32) approaches 1, as T and m increase. More precisely, this is the case, if δ is chosen smaller than the ratio of MSPEs depicted in Figure 2 on both the validation and test set or larger than both those ratios. This is as expected from Corollary 3.3 and 3.4. A more detailed analysis is possible, employing the information provided in Table 1. In the third row of tables we see, for example, that for $n = 10000$ and $\delta = 0$ the procedure will consistently choose the locally stationary approach on both the test and validation set for 1-step ahead forecasting. For $n = 10000$ and $\delta = 0.05$, on the other hand, we see that the procedure almost consistently chooses the locally stationary approach on the validation set while it is rather undecided (50%-50%) on the test set. For $\delta = 0.1$ the procedure almost consistently chooses the stationary approach on the test set and is to some degree undecided (70%-30%) on the validation set. Finally, if $\delta = 0.15$, we see that the stationary approach gets chosen almost consistently on both validation and test sets. This is just what we would expect, as a smaller δ must lead to the locally stationary approach being preferred, as the more complex locally stationary approach only gets selected if the empirical MSPE of the stationary approach is at least $(1 + \delta)$-times of the empirical MSPE of the locally stationary approach.

We now briefly summarise our findings for the other models and refer to Appendix I (Online Supplement) for further details.

For the different models we observe that for $n = 100$ the best stationary predictor usually (in all models, except the strongly non-stationary (30) and (31)) outperforms the locally stationary approach, which can be seen from the figures where the lines are below 1 and also in the proportions from the tables. In 11 of the 15 models the 1-step ahead stationary predictor yields the better performance in more than 65% of the cases. Even in the highly non-stationary models (17), (23), (30), and (31) where the tangent processes for the observations to be predicted are close to the unit root it is still more than 47% in which the stationary approach yields the better performance. As a general conclusion we thus see that locally stationary modelling might not be advantageous for forecasting short time series. Further, for those models, when $n = 1000$, we can see that the locally stationary approach will result in a significantly reduced MSPE as the sample size increases. For example, with model (23) the 1-step ahead locally stationary approach outperforms the stationary approach by about 20% on average and in almost 92% of the cases.

We further observe that in all of the figures the plots in the bottom row, corresponding to $n = 4000, 6000, 8000$ and 10000 are of a generally similar shape (in (28) we see this only for $n \geq 6000$). For an explanation note that, as n grows larger, m/n tends to zero and

consequently the data in the validation and test set behave more similarly. For $n = 100$, for example, the validation and test set make up 24% of the data, while for $n = 10000$ it is only 12.54%. We observe two different kinds of shapes: (a) for some of the models, the locally stationary and the stationary approach appear to be performing similarly well as h increases, and (b) in other models we find that the locally stationary approach outperforms the stationary approach even more as h increases. More precisely, for models (17)–(20), (28), (30), and (31), we see that the locally stationary approach excels for small h and both approaches are almost equally good as h gets larger, while for models (25)–(27) the stationary approach is advantageous for smaller h and the two approaches are of similar performance for the larger h's. In model (29) the stationary procedure is always better (uniformly with respect to h), but as T gets larger the advantage is less pronounced. Note that models (25), (26), and (29) are stationary and (26) and (27) are the independent observations. It is thus not very surprising that the stationary approach outperforms the locally stationary one. Finally, in models (21)–(24), where the tangent processes for the observations from the validation and test set are close to the unit root, we see that as h gets larger the ratio of MSPEs first increases and then (slowly) decreases. These observations corroborate the rules of thumbs we have given before.

5 Data examples

5.1 London Housing Prices

We analyse average housing prices from the UK House Price Index (HPI). The HPI is updated monthly with data from the Land Registry, the Registers of Scotland, and the Land and Property Services Northern Ireland. The data is combined by the Office of National Statistics using hedonic regression.[1] The sequence we used for the analysis contains 252 monthly index values from 1995 to 2015.[2] The data is depicted in the left panel of Figure 3. For the analysis we consider $T + m = 251$ monthly changes (in percent). The prices are centred by subtracting the arithmetic mean prior to the analysis. We clearly see autocorrelation at lags less or equal than 4 and at lag 12 in the right panel of Figure 3.

We then compute the 1-step to 6-step prediction coefficients, defined in (4), with which we can predict an observation X_{t+h} from $X_t, \ldots, X_{t-p+1}$, where X_{t+h} is an observation made either in 2013, 2014 or 2015, respectively. We choose $p = 0, 1, \ldots, 18$, where $p = 0$ shall mean that we are predicting with 0. Note that the maximum p was chosen larger than 12, as we are dealing with monthly data and dependence at lag 12 can be seen from the autocorrelation function. We consider the stationary predictors as well as locally stationary predictors with $N = 35, 36, \ldots, 84 = \lceil 251^{4/5} \rceil$.

Interestingly, in Figure 4, we observe that the MSPE of the locally stationary forecasts

---

[1]See http://landregistry.data.gov.uk/app/ukhpi.

[2]In the 'customise your search' part of the 'search the UK house price index' form we have selected the 'English region' London, the period from 01-1995 to 12-2015, and then obtained the 'average price' for 'all property types'.
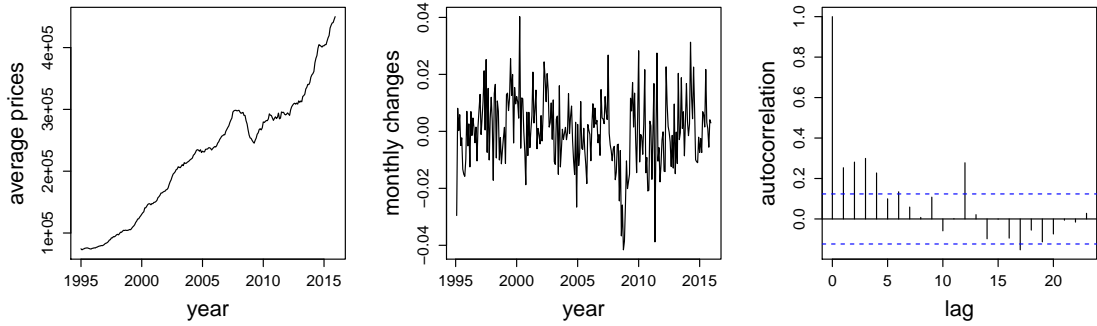
Figure 3: Data for London from the UK House Price Index. Left: average monthly housing prices. Middle: monthly changes of average housing prices in percent, demeaned by subtracting arithmetic mean. Right: autocorrelation function, computed from the sequence in the middle.

are typically larger than corresponding ones of the stationary forecasts. Further, note that amongst the stationary forecasts the mean squared prediction errors for both the 1-step and 6-step ahead predictor with $p = 0$ (shown as a grey solid horizontal line) is smaller than those for $0 < p < 12$, but larger than those with $12 \leq p \leq 15$. Amongst the locally stationary estimators this observation is not visible.

As described in our procedure we now determine the $\hat{p}_\mathrm{s}$, $\hat{p}_\mathrm{ls}$, and $\hat{N}$ that minimise the MSPE within each class of predictors. For 1-step ahead prediction we find $\hat{p}_\mathrm{s} = 13$, $\hat{p}_\mathrm{ls} = 16$, and $\hat{N} = 56$. For 6-step ahead prediction we find $\hat{p}_\mathrm{s} = 14$, $\hat{p}_\mathrm{ls} = 13$, and $\hat{N} = 51$. The numbers are summarised in Table 5.

We then determine the MSPE for forecasting the observations from the validation set (here: the year 2014) using these predictors. For 1-step ahead prediction we find that $\mathrm{MSPE}^\mathrm{s}_{239,2}(1) = 0.000116$ and $\mathrm{MSPE}^\mathrm{ls}_{239,2}(1) = 0.000132$, with $\mathrm{MSPE}^*_{T,j}(h)$ defined in (8). For 6-step ahead prediction we find that $\mathrm{MSPE}^\mathrm{s}_{239,2}(6) = 0.000131$ and $\mathrm{MSPE}^\mathrm{ls}_{239,2}(6) = 0.000132$. Consequently we decide to use the stationary approach for 1-step and 6-step ahead forecasting of the observations made in 2015.

The MSPEs computed from 1-step ahead forecasting the observations from the test set (here: the year 2015) are $\mathrm{MSPE}^\mathrm{s}_{239,3}(1) = 3.11 \cdot 10^{-5}$ and $\mathrm{MSPE}^\mathrm{ls}_{239,3}(1) = 4.72 \cdot 10^{-5}$. The MSPEs computed from 6-step ahead forecasting the observations from 2015 are $\mathrm{MSPE}^\mathrm{s}_{239,3}(6) = 4.70 \cdot 10^{-5}$ and $\mathrm{MSPE}^\mathrm{ls}_{239,3}(6) = 4.93 \cdot 10^{-5}$. We have thus chosen the better performing procedure for 1-step and 6-step ahead forecasting.

In conclusion, our analysis has revealed that, from the point of view of prediction of the 2015 observations, treating the data as stationary does not have a negative effect. We were able to see that using the estimates from the (stationary) AR(13) or AR(14) models (for month and half-year ahead forecasts, respectively) gave us better predictions than using the (locally stationary) estimates of segments of roughly 4.5 years (56 and 51 months, respectively). In particular, the impact of, for example, the 2008-2009 financial crisis on the stationary estimates, as one might naively expect, is not profound enough to worsen the predictors' performance.

23

Figure 4: Empirical mean squared prediction errors (MSPEs) computed on the end of the training set (predicting the 12 observations from 2013). Top panel shows MSPEs for 1-step ahead prediction. Bottom panel shows 6-step ahead prediction. The colours indicate which p was used. The colour code is described in the plot's legend. The solid lines correspond to the MSPEs for different N when the locally stationary approach is used. The dashed lines show the MSPE when the stationary approach is used. The horizontal grey line indicates the MSPE for the trivial forecasts ($f^{\text{ls}}_{t,h;0,N}$ and $f^{\text{s}}_{t,h;0}$). The MSPE in this case is 0.000129.

h	$\hat{p}_{\mathrm{s}}(1)$	$\mathrm{MSPE}^{\mathrm{s}}_{T,1}(h)$	$\hat{p}_{\mathrm{ls}}(1)$	$\hat{N}_{\mathrm{ls}}(1)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,1}(h)$
1	13	1.085033e-04	16	56	1.117308e-04
2	13	9.367297e-05	16	55	8.999408e-05
3	13	1.028582e-04	16	63	9.535843e-05
4	16	8.752610e-05	16	68	7.943293e-05
5	14	1.035185e-04	13	67	1.138279e-04
6	14	1.115081e-04	13	51	1.075563e-04

h	$\mathrm{MSPE}^{\mathrm{s}}_{T,2}(h)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,2}(h)$	$\hat{R}^{\mathrm{s,ls}}_{T,2}(h)$	$\mathrm{MSPE}^{\mathrm{s}}_{T,3}(h)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,3}(h)$	$\hat{R}^{\mathrm{s,ls}}_{T,3}(h)$
1	0.0001156469	0.0001317383	0.878	3.110699e-05	4.716983e-05	0.659
2	0.0001316892	0.000136116	0.967	3.078684e-05	5.139516e-05	0.599
3	0.0001460473	0.0001457031	1.002	3.531227e-05	3.129435e-05	1.128
4	0.0001120983	0.0001143668	0.980	4.226533e-05	3.240333e-05	1.304
5	0.0001367077	0.000160234	0.853	4.424542e-05	3.762859e-05	1.176
6	0.0001307965	0.0001322581	0.989	4.6993e-05	4.928723e-05	0.953

Table 5: *Minimum empirical mean squared prediction errors (MSPEs) for h-step ahead prediction, $h = 1, \ldots, 6$, of the housing price data. Top table shows values computed on the end of the training set. Bottom table shows values computed on the end of the validation and test set.*

5.2 Temperatures Hohenpeißenberg

In this example, we analyse seasonally adjusted, daily temperature data collected at the meteorological observatory in Hohenpeißenberg (Germany). More precisely, we use $n = T + m = 11315$ observations of daily mean temperatures that were recorded between 1985 and 2015.[3] The data are shown in the left panel of Figure 5. To eliminate the clearly visible trend and seasonality, we have fitted a harmonic linear regression model of the form

$$y_t = c + \alpha t + \sum_{i=1}^{4} \beta_i \sin(2\pi t i/365) + \gamma_i \cos(2\pi t i/365),$$

to capture the trend and annual variation. The red curve in the left panel of Figure 5 is the prediction of the fitted model. We then consider the residuals of this model which are shown in the middle panel of Figure 5. The right panel of Figure 5 shows the auto-correlation function, which clearly indicates that serial dependence is present. Birr et al. (2016) analyse the same data set and fit a stationary ARMA(3,1) model to capture the serial dependence.

In Figure 6, the MSPE are presented in the same manner as in Section 5.1. In this example we have chosen $p_{\max} = 10$ and $\mathcal{N} := \{\lceil n^{4/5}/2 \rceil, \ldots, \lceil n^{4/5} \rceil\} = \{875, 876, \ldots, 1748, 1749\}$ and $m := 365$. The MSPE corresponding to $p = 0$ is 116 in this example and therefore not visible in the plot. Note further, that for $p = 1$ the 1-step ahead forecasts are performing far worse than those for $p \geq 2$.

By minimising the empirical MSPE on the end of the training set the procedure chooses, for the stationary approach $\hat{p}_{\mathrm{s}} = 5$ and $\hat{p}_{\mathrm{s}} = 6$ for $h = 1$ and $h = 2$, respectively. For

---

[3]The data was obtained from `http://www.dwd.de/DE/klimaumwelt/cdc/cdc_node.html`.
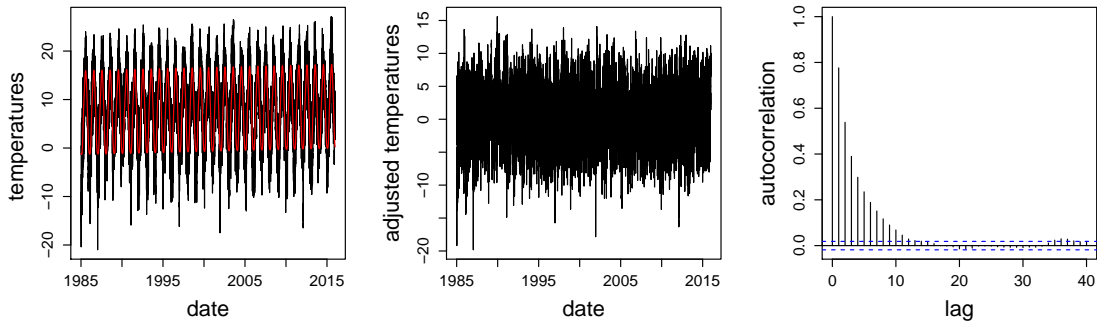
Figure 5: Temperature data Hohenpeißenberg. Left: daily temperatures and fitted harmonic regression model. Middle: adjusted data (demeaned and detrended). Right: autocorrelation function, computed from the sequence in the middle.

h	$\hat{p}_{\mathrm{s}}(1)$	$\mathrm{MSPE}^{\mathrm{s}}_{T,1}(h)$	$\hat{p}_{\mathrm{ls}}(1)$	$\hat{N}_{\mathrm{ls}}(1)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,1}(h)$
1	5	7.613696	5	1255	7.534377
2	6	15.153486	5	1235	15.050641
3	6	18.940623	5	1234	18.912431
4	6	20.800483	5	1234	20.882383
5	1	21.516749	1	1512	21.508112

h	$\mathrm{MSPE}^{\mathrm{s}}_{T,2}(h)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,2}(h)$	$\hat{R}^{\mathrm{s,ls}}_{T,2}(h)$	$\mathrm{MSPE}^{\mathrm{s}}_{T,3}(h)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,3}(h)$	$\hat{R}^{\mathrm{s,ls}}_{T,3}(h)$
1	7.211691	7.3375	0.983	8.019931	8.092247	0.991
2	13.01974	13.26302	0.982	14.77994	14.92948	0.990
3	15.64904	15.88363	0.985	17.6721	17.81343	0.992
4	16.924	17.1055	0.989	19.56872	19.74487	0.991
5	17.7535	17.8397	0.995	20.66451	20.61552	1.002

Table 6: *Minimum empirical mean squared prediction errors (MSPEs) for h-step ahead prediction, $h = 1, 2, 3, 4, 5$, of the temperature data Hohenpeißenberg. Top table shows values computed on the end of the training set. Bottom table shows values computed on the end of the validation and test set.*

the locally stationary approach the procedure chooses $(\hat{p}_{\mathrm{ls}}, \hat{N}) = (5, 1255)$ and $(\hat{p}_{\mathrm{ls}}, \hat{N}) = (5, 1235)$ for $h = 1$ and $h = 2$, respectively. Empirical MSPEs for other values of p and N are shown in Figure 6. The numbers are summarised in Table 6.

For 1-step ahead forecasting and on the validation set this yields, given the $\hat{p}_{\mathrm{s}}$ chosen by the procedure, that $\mathrm{MSPE}^{\mathrm{s}}_{10950,2}(1) = 7.21$ and, given the $\hat{p}_{\mathrm{ls}}$ and N chosen by the procedure, that $\mathrm{MSPE}^{\mathrm{ls}}_{10950,2}(1) = 7.34$. Similarly, for 2-step ahead forecasting, we have $\mathrm{MSPE}^{\mathrm{s}}_{10950,2}(2) = 13.02$ and $\mathrm{MSPE}^{\mathrm{ls}}_{10950,2}(2) = 13.26$. Because the respective ratios are both smaller than 1, the procedure chooses the stationary approach over the locally stationary approach. Obviously, this superiority will continue to hold if δ is chosen strictly larger than one by the user.

On the test set we have $\mathrm{MSPE}^{\mathrm{s}}_{10950,3}(1) = 8.02$ and $\mathrm{MSPE}^{\mathrm{ls}}_{10950,3}(1) = 8.09$ for 1-step ahead forecasting. Likewise, for 2-step ahead forecasting, we have $\mathrm{MSPE}^{\mathrm{s}}_{10950,3}(2) = 14.77$ and $\mathrm{MSPE}^{\mathrm{ls}}_{10950,3}(2) = 14.93$. Thus, again, the stationary approach is superior for 1-step
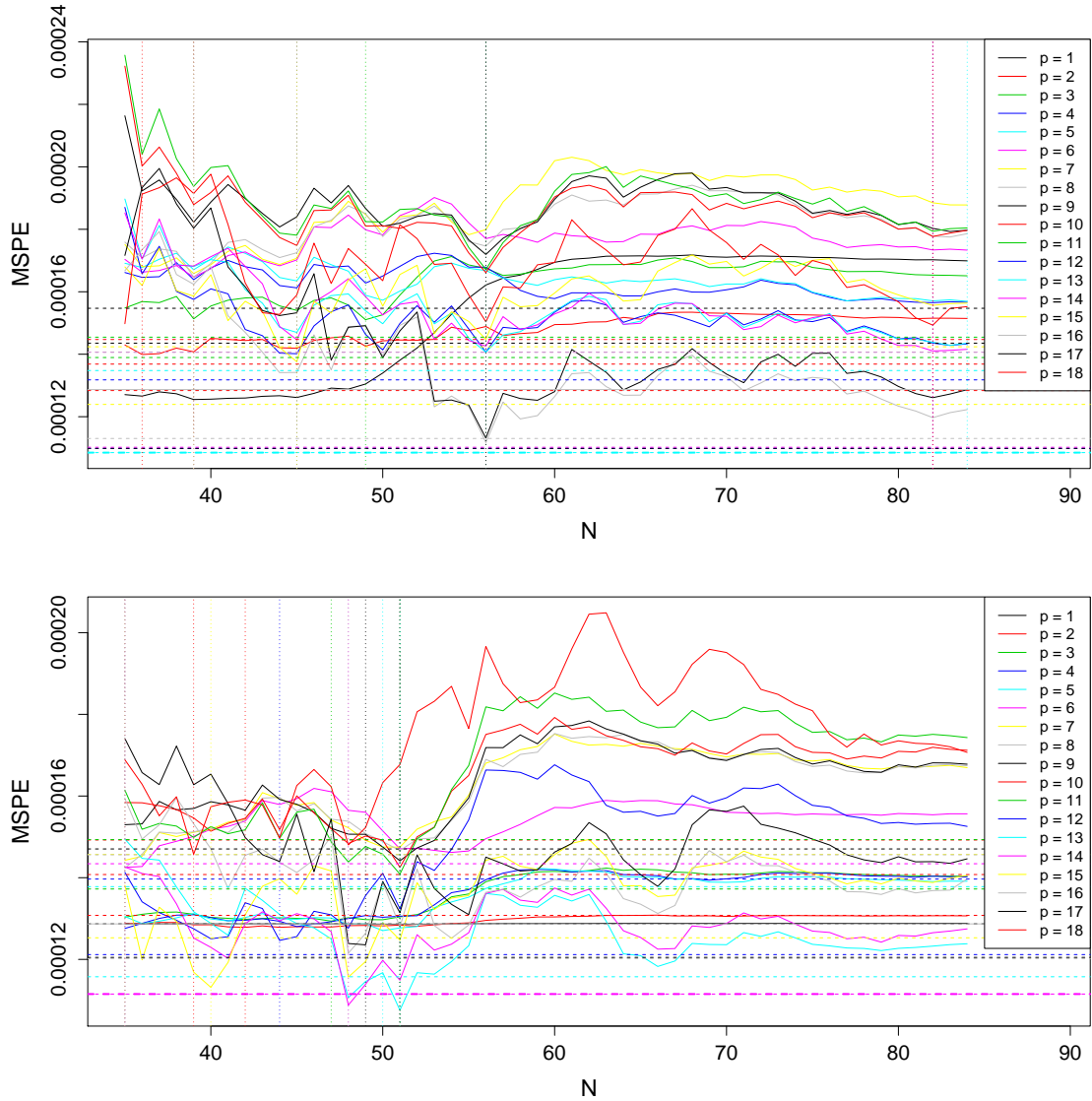
26

Figure 6: Empirical mean squared prediction errors (MSPEs) computed on the end of the training set (predicting the 365 observations from 2013) of the temperature data. Top panel shows MSPEs for 1-step ahead prediction. Bottom panel shows 2-step ahead prediction. The colours indicate which p was used. The colour code is described in the plot's legend. The solid lines correspond to the MSPEs for different N when the locally stationary approach is used. The dashed lines show the MSPE when the stationary approach is used.
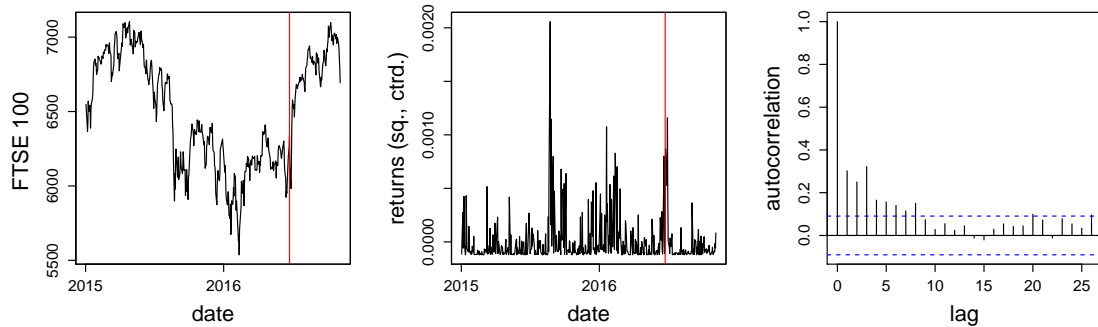
Figure 7: Volatility of the FTSE 100 Index, for 2 January 2015 to 4 November 2016. Left: FTSE 100 closing price. Middle: squared and centred returns. Right: autocorrelation function, computed from the sequence in the middle. Red vertical line in the left and middle plot marks 23/06/2016, the day of the EU referendum in the UK.

and 2-step ahead forecasting and we see that the approach that our procedure chose, the one performing best on the validation set, is also performing best on the test set.

In conclusion, in this example, we have provided clear evidence that the temperature data collected in the Hohenpeißenberg observatory, from the point of view of prediction, should be treated as if they were stationary. We see that using the estimates related to a AR(5), AR(6), or AR(1) model (for $h = 1$, $h = 2, 3, 4$, or $h = 5$, respectively) yielded the better forecasts on the validation set in all cases. In particular, using the estimates localised to the segment suggested by the procedure (using the past roughly 3.5–4 years; 1234–1512 days) the performance was always worse. This observation is remarkable, in the sense that, in 30 years of data an analyst might typically expect non-stationarity (e. g., changes due to global warming). Further, our procedure consistently chose the approach with the better performance on the test set, but for the case where $h = 5$, in which both approaches perform almost equally well. In this case ($h = 5$), choosing the stationary approach, though it simplifies the analysis, results in a performance-loss of only 0.2%.

5.3 Volatility around the time of the EU referendum in the UK, 2016

This example is about forecasting volatility of the FTSE 100 stock index. More precisely, we consider a sequence of $n = T + m = 468$ (daily) opening prices p_{open} and closing prices p_{close}, dated from 2 January 2015 to 4 November 2016.[4] The analysis is then based on the sequence $((p_{\text{close}} - p_{\text{open}})/p_{\text{close}})^2$, centered by subtracting the arithmetic mean of this sequence. The data are shown in Figure 7.

We separate the final 45 observations of the data as test, validation and end of the training set (used for determining $\hat{p}_{\text{s}}$, $\hat{p}_{\text{ls}}$ and $\hat{N}$), respectively. Thus, each of the sets is of size $m :=$ 15. In this example we have further chosen $p_{\max} = 8$ and $\mathcal{N} := \{30, 31, \ldots, \lceil n^{4/5} \rceil\} = \{30, 31, \ldots, 136, 137\}$. For the first step, in Figure 8, we consider the empirical MSPEs

---

[4]The data was obtained from `http://www.finanzen.net/index/FTSE_100/Historisch`.

h	$\hat{p}_\text{s}(1)$	$\text{MSPE}^\text{s}_{T,1}(h)$	$\hat{p}_\text{ls}(1)$	$\hat{N}_\text{ls}(1)$	$\text{MSPE}^\text{ls}_{T,1}(h)$
1	6	9.353206e-09	6	35	6.397327e-09
2	6	8.545269e-09	7	47	5.135330e-09
3	6	8.651203e-09	7	46	5.079011e-09
4	8	7.067758e-09	6	43	4.141545e-09
5	8	6.348194e-09	6	30	3.511903e-09

h	$\text{MSPE}^\text{s}_{T,2}(h)$	$\text{MSPE}^\text{ls}_{T,2}(h)$	$\hat{R}^\text{s,ls}_{T,2}(h)$	$\text{MSPE}^\text{s}_{T,3}(h)$	$\text{MSPE}^\text{ls}_{T,3}(h)$	$\hat{R}^\text{s,ls}_{T,3}(h)$
1	4.250396e-09	3.77037e-09	1.127	3.103276e-09	3.318341e-09	0.935
2	4.352724e-09	3.55879e-09	1.223	3.638043e-09	3.310187e-09	1.099
3	4.588659e-09	3.463196e-09	1.325	4.203515e-09	3.548548e-09	1.185
4	4.913305e-09	3.626292e-09	1.355	4.556755e-09	3.637315e-09	1.253
5	4.862813e-09	3.989628e-09	1.219	4.829544e-09	3.54644e-09	1.362

Table 7: *Minimum empirical mean squared prediction errors (MSPEs) for h-step ahead prediction, $h = 1, 2, 3, 4, 5$, of the squared and centred FTSE 100 data. Analysis performed with $m := 15$ and $p_\text{max} = 8$. Top table shows values computed on the end of the training set. Bottom table shows values computed on the end of the validation and test set.*

of forecasting the 15 observations from the end of the training set. We can see, for the 1-step, 2-step and 3-step ahead forecasts, that the lines have a characteristic shape: as N increases from 30 to about 45–50 the MSPE is constant or slightly decreasing (for each p at a different level), from 45–50 to about 55–60 it rises, sharply and 'in turmoil', to a significantly higher level and then it slowly decreases from around 65 onwards.

Note that observations 1 through to 373 were recorded from 2 January 2015 to 23 June 2016 (the day of the EU referendum) and observations 374 through to 468 were recorded from 24 June 2016 to 4 November 2016. This implies that the final 95 observations were recorded after the EU referendum, meaning that there are 50 observations between the EU referendum and the observations to be forecast in the first step. Thus, the increase of the lines when N is roughly between 50 and 60 corresponds to an increased bias of the Yule Walker estimator due to non-stationarity when pre-referendum data is starting to be used for the estimation of the prediction coefficients. Another important observation is that in the post-referendum part of the diagram ($25 \leq N \leq 45$) the data does not show a great degree of non-stationarity, for the purpose of forecasting. More clearly, we cannot distinguish it from data that were stationary in terms of the MSPE observed. This can be seen from the fact that the MSPE has more of a constant or decreasing shape than it is 'U-shaped'. For all h the minimum empirical MSPE for forecasting the data from the end of the estimation set is obtained somewhere between around 30 and 50. More precisely, we choose $(p_\text{ls}, \hat{N}_\text{ls})$ as $(6, 35)$, $(7, 47)$ and $(7, 46)$, for $h = 1, 2, 3$, respectively. For $h = 1$ we see 6.40×10^{-9}, for $h = 2$ we see 5.14×10^{-9}, for $h = 3$ we see 5.08×10^{-9}). When using the stationary approach we choose $\hat{p}_\text{s}$ as 6, when $h = 1, 2, 3$, respectively. For $h = 1$ we see 9.35×10^{-9}, for $h = 2$ we see 8.55×10^{-9}, for $h = 3$ we see 8.65×10^{-9}). The numbers are summarised in Table 7.

Using these predictors to forecast the 15 observations from the validation set we see that, for $h = 1$ the MSPE of the stationary approach is 4.25×10^{-9} and the MSPE of the locally stationary approach is 3.77×10^{-9}. For $h = 2$, we see that the MSPE of the stationary
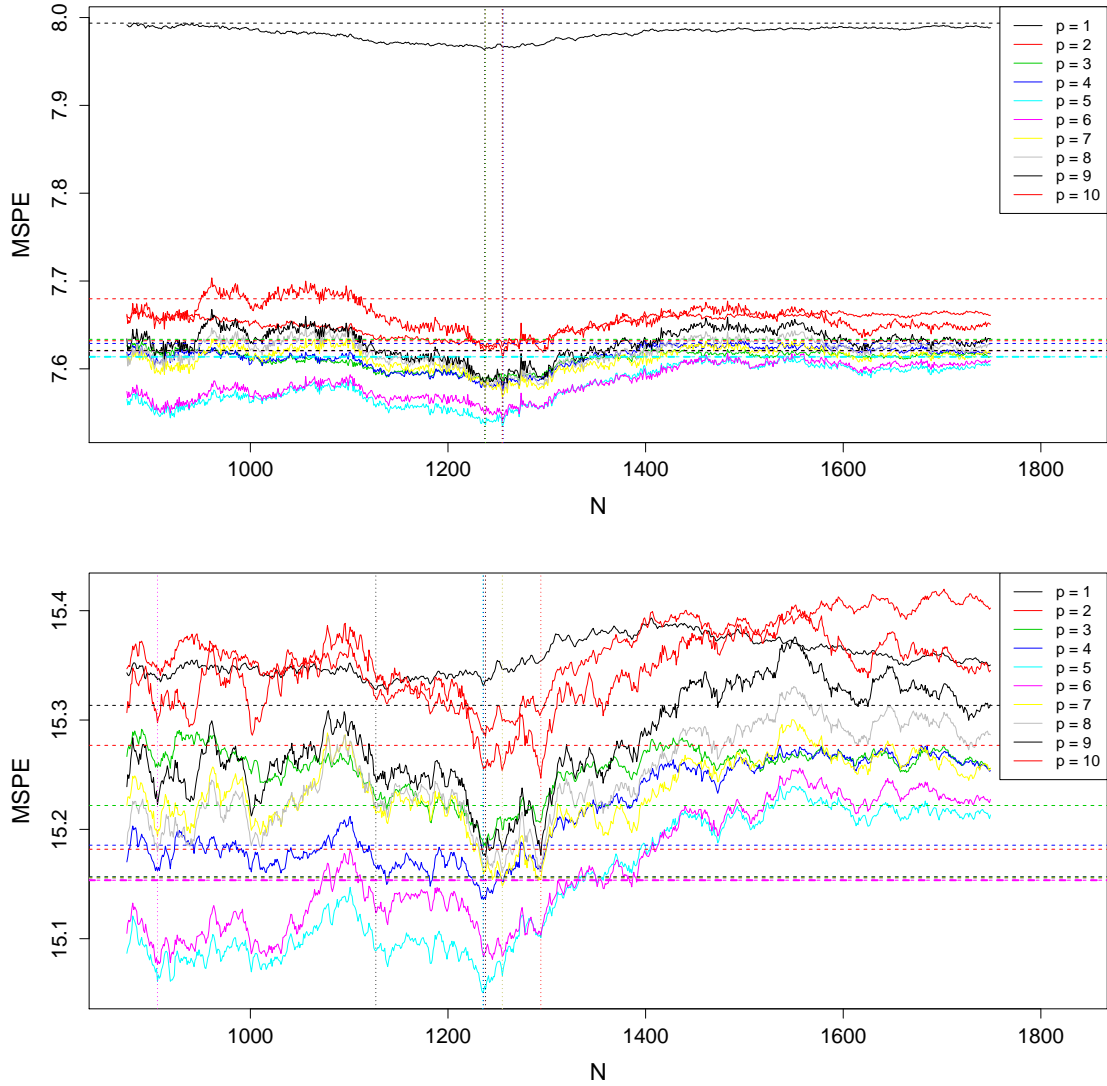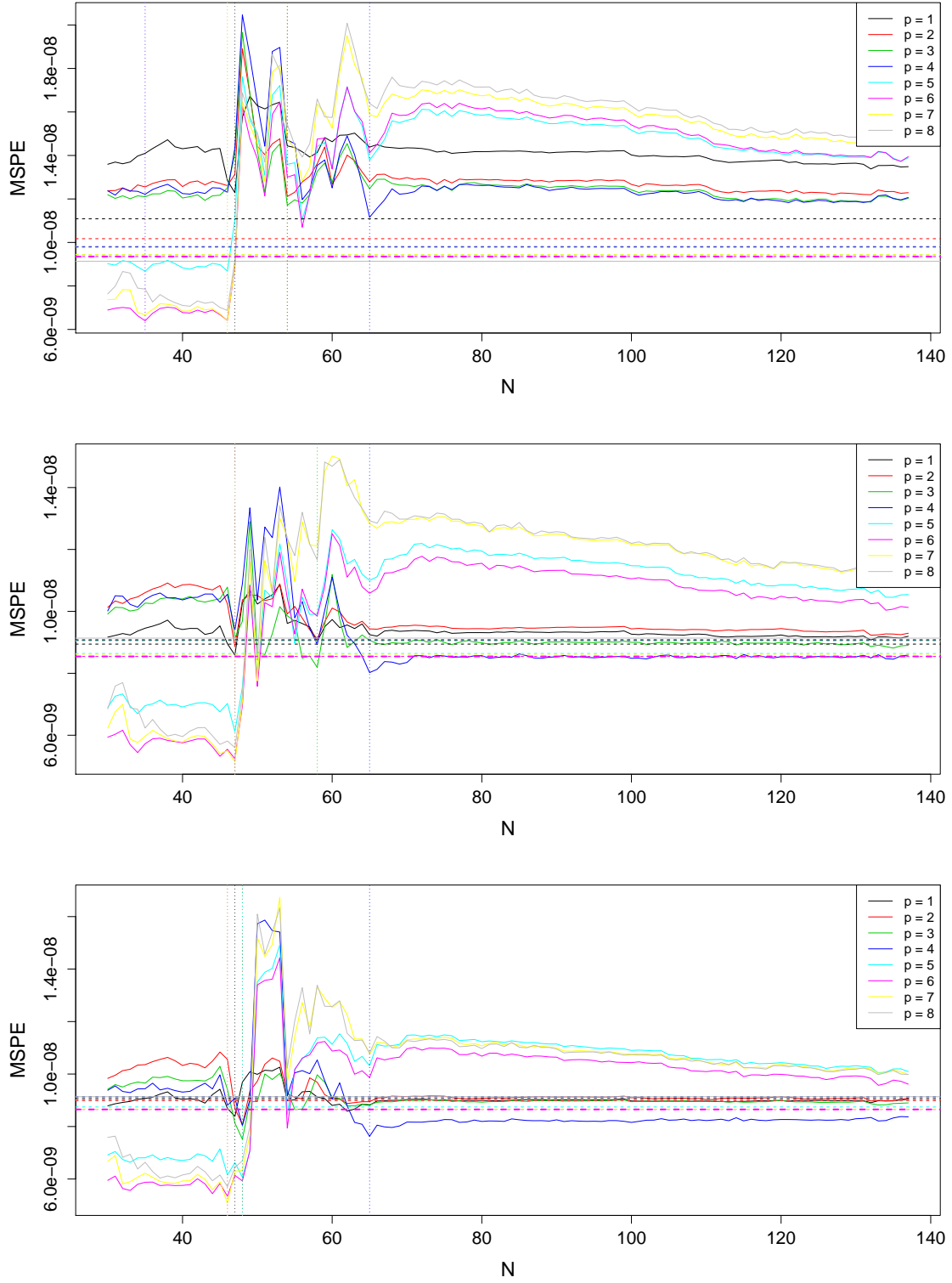
Figure 8: Empirical mean squared prediction errors (MSPEs) computed on the end of the training set (predicting the observations 424 to 438) of the squared and centred FTSE returns. Top, middle and bottom panel show 1, 2 and 3-step ahead prediction, respectively. The colours indicate which p was used. The colour code is described in the plot's legend. The solid lines correspond to the MSPEs for different N when the locally stationary approach is used. The dashed lines show the MSPE when the stationary approach is used. The horizontal grey line indicates the MSPE for the trivial forecasts ($f_{t,h;0,N}^{\mathrm{ls}}$ and $f_{t,h;0}^{\mathrm{s}}$). The MSPE in this case is 9.13×10^{-9}.

30

approach is 4.35×10^{-9} and the MSPE of the locally stationary approach is 3.56×10^{-9}. For $h = 3$, we see that the MSPE of the stationary approach is 4.59×10^{-9} and the MSPE of the locally stationary approach is 3.46×10^{-9}. Thus, for all three $h = 1, 2, 3$ we choose the locally stationary approach, as it performs better on the validation set.

Now, forecasting the final 15 observations (the ones from the test set) we see that, for $h = 1$ the MSPE of the stationary approach is 3.10×10^{-9} and the MSPE of the locally stationary approach is 3.32×10^{-9}. For $h = 2$, we see that the MSPE of the stationary approach is 3.64×10^{-9} and the MSPE of the locally stationary approach is 3.31×10^{-9}. For $h = 3$, we see that the MSPE of the stationary approach is 4.20×10^{-9} and the MSPE of the locally stationary approach is 3.55×10^{-9}. Thus, choosing the stationary approach would have yield a slightly better result in forecasting the observations from the test set, but for $h = 2, 3$ we chose the approach that has the better performance.

In this example, we see that, for the case of $h = 2, \ldots, 5$, our procedure choose the procedure consistently, while for $h = 1$ it did not. We see that the 'mistakenly' chosen locally stationary approach, for $h = 1$, yielded only an 6.5% worse performance (in terms of the MSPE on the test set), while for $h = 2, 3, 4, 5$ it chose the procedure that is substantially better (9.9% to more than 36% improvement). An evaluation of the application of the method should also take into account that the recent changes are ongoing and that less than 5 month of observations after the EU referendum in the UK were available, which requires that m was chosen rather small. For this example, we further conducted a sensitivity analysis, by varying the parameters m and $p_{\max}$. Selected results, in which we see how results change are shown in Appendix H (Online Supplement).

6   Analysis of the localised Yule-Walker estimator under general conditions and local stationarity

In this section we discuss the probabilistic properties of the localised Yule-Walker estimator $\hat{a}_{N,T}^{(p)}(t)$ defined in (1). We believe the results to be of independent interest and therefore present them in this separate section. They are also key results for the proofs of the result in Section 3.3. Our results will hold under the five main assumptions discussed in Section 3.3 and rigorously stated in Appendix A. The assumptions are not restrictive and, in particular, the concentration result in this section will hold for a broad class of locally stationary processes and, in particular, does not require that the data come from a tvAR(p) model. Further, we allow for any $1 + p \leq N \leq T$ and, in particular, allow for a diverging model order p, as $T \to \infty$. We do not, as do for example Dahlhaus and Giraitis (1998), require that $N = o(T)$.

Our main result (Theorem 6.1) provides a non-asymptotic bound for the distance of $\hat{a}_{N,T}^{(p)}(t)$ to the following population quantity:

$$\bar{a}_{N,T}^{(p)}(t) := \big(\mathbb{E}\hat{\Gamma}_{N,T}^{(p)}(t)\big)^{-1}\big(\mathbb{E}\hat{\gamma}_{N,T}^{(p)}(t)\big) = \big(\bar{a}_{1,N,T}^{(p)}(t), \ldots, \bar{a}_{p,N,T}^{(p)}(t)\big)'. \tag{33}$$

The Yule-Walker estimator is widely used in practice and $\hat{a}_{N,T}^{(p)}(t)$ and its properties have

been studied in detail under various conditions. Bercu et al. (1997, 2000) and Bercu (2001) derive large deviation principles for Gaussian AR processes when the model order is 1. A simple exponential inequality, also for model order 1, is given in Section 5.2 of Bercu and Touati (2008). Yu and Si (2009) prove a large deviation principle for general, but fixed, model order. The cited results all require that the underlying process is stationary. Dahlhaus and Giraitis (1998) analyse the bias and variance of the localised Yule-Walker estimator in the framework of local stationarity. They do not, however, provide an exponential inequality, and, as far as we are aware, no result as the one we provide below is available at present.

The exponential inequality in Theorem 6.1, which we now state, is explicit in terms of all parameters and constants. We make use of the explicitness to derive Corollary 6.2, by which the localised Yule-Walker estimator is strongly, uniformly consistent, even when the model order is diverging as the sample size grows.

Theorem 6.1 *Let $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ satisfy Assumptions 1–5 (Appendix A) and $\mathbb{E}X_{t,T} = 0$. Then, for every $T \geq 2C_1p^2$, $N \geq 1 + p \geq 2$ and $\varepsilon > 0$, we have:*

$$\mathbb{P}\left(\|\hat{a}_{N,T}^{(p)}(t) - \bar{a}_{N,T}^{(p)}(t)\| > \varepsilon\right)$$

$$\leq 3p\exp\left(-\frac{\left(\frac{m_f}{4p}\min\left\{1,\varepsilon\frac{1}{8C_0}\right\}\right)^2}{2\left(C_{1,1}\frac{p}{N-p} + \left(\frac{m_f}{4p}\min\left\{1,\varepsilon\frac{1}{8C_0}\right\}\right)^{(3+4d)/(2+2d)}\left(C_{2,1}\frac{p}{N-p}\right)^{1/(2+2d)}\right)}\right)$$

$$\leq \begin{cases} 3p\exp\left(-\dfrac{m_f^2}{32C_{1,1}\frac{p^3}{N-p} + m_f^{(3+4d)/(2+2d)}\left(32C_{2,1}\frac{p^2}{N-p}\right)^{1/(2+2d)}}\right) & \varepsilon \geq 1/(8C_0) \\[2ex] 3p\exp\left(-\varepsilon^2\dfrac{m_f^2}{2^{12}C_{1,1}}\left(C_0^2\dfrac{p^3}{N-p}\right)^{-1}\right) & \varepsilon \leq \min\{U_{p,N}, \frac{1}{8C_0}\} \\[2ex] 3p\exp\left(-\varepsilon^{1/(2+2d)}\left(\dfrac{m_f}{2^{9+4d}C_{2,1}}\right)^{1/(2+2d)}\left(C_0\dfrac{p^2}{N-p}\right)^{-1/(2+2d)}\right) & \frac{1}{8C_0} > \varepsilon \geq \min\{U_{p,N}, \frac{1}{8C_0}\} \end{cases}$$

where $\hat{a}_{N,T}^{(p)}(t)$ is defined in (1), $\bar{a}_{N,T}^{(p)}(t)$ is defined in (33),

$$U_{p,N} := \frac{32C_0}{m_f}\left(\frac{C_{1,1}^{2+2d}}{C_{2,1}}\right)^{1/(3+4d)}\left(\frac{p^{(4+6d)/(3+4d)}}{(N-p)^{(1+2d)/(3+4d)}}\right),$$

and C_0, C_1 and $C_{1,1}$, $C_{2,1}$ are defined in (43) and (85), respectively.

Theorem 6.1 is a key ingredient to the proof of Lemma A.2 which is essential to the proof of the performance-guarantee-result (Theorem 3.1) of our procedure. Further, it implies

Corollary 6.2 *Let $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ satisfy Assumptions 1–5 (Appendix A), $\mathbb{E}X_{t,T} = 0$ and let $P = P_T$ and $N = N_T$ be sequences of integers that satisfy $2 \leq 1 + P \leq N \leq T$. Assume that $P = o(N^{(1+2d)/(4+6d)})$ and $N \to \infty$, as $T \to \infty$. Further, assume that there exists a sequence R_T with $0 \leq R_T \to \infty$ and $R_T\log(T) = o((N/P)^{1/(3+4d)})$, as $T \to \infty$, where d is the constant from Assumption 5. Then, we have*

$$\sup_{p=1,\ldots,P}\sup_{t=N,\ldots,T}\|\hat{a}_{N,T}^{(p)}(t) - \bar{a}_{N,T}^{(p)}(t)\| = O\left(P^{3/2}\left(\frac{\log(T)}{N}\right)^{1/2}\right), \text{ almost surely, as } T \to \infty.$$

Remark 6.3 *For any stationary $AR(p)$ model we have that $\bar{a}_{N,T}^{(p)}(u)$ corresponds to the vector of coefficients. This can be seen from Lemma 6.5, below, and the fact that $C_1 = 0$ if the model is stationary. Thus, choosing $N_T = T$ and $P_T = p$, our result yields the same rate as Theorem 1 in Lai and Wei (1982), by which the (least squares) estimator is strongly consistent with rate $(\log(T)/T)^{1/2}$.*

We now discuss the quantity $\bar{a}_{N,T}^{(p)}(t)$ defined in (33). It is related to $a_\Delta^{(p)}(u)$ defined in (51) in the sense that $a_0^{(p)}(t/T)$ and $a_{N/T}^{(p)}(t/T)$ approximate $\bar{a}_{N,T}^{(p)}(t)$. We further have that $a_0^{(p)}(t/T)$ approximates the *time-varying 1-step linear prediction coefficients* which, for $p \in \mathbb{N}^*$ and $t = 1, \ldots, T$, are defined as

$$\tilde{a}_T^{(p)}(t) := \arg \min_{a=(a_1,\ldots,a_p)' \in \mathbb{R}^p} \mathbb{E}\left[\left(X_{t,T} - \sum_{j=1}^p a_j X_{t-j,T} \right)^2 \right] = (\tilde{\Gamma}_T^{(p)}(t))^{-1} \tilde{\gamma}_T^{(p)}(t),$$

where

$$\begin{aligned} \tilde{\gamma}_T^{(p)}(t) &:= \left(\text{Cov}\left(X_{t,T}, X_{t-1,T} \right), \ldots, \text{Cov}\left(X_{t,T}, X_{t-p,T} \right) \right)', \\ \tilde{\Gamma}_T^{(p)}(t) &:= \left(\text{Cov}\left(X_{t-i,T}, X_{t-j,T} \right); i,j = 1, \ldots, d \right). \end{aligned} \tag{34}$$

The following two lemmas discuss approximation properties of $\bar{a}_{N,T}^{(p)}(t)$ and $\tilde{a}_T^{(p)}(t)$:

Lemma 6.4 *Let $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ satisfy Assumptions 1, 3, 4 (Appendix A), $\mathbb{E}X_{t,T} = 0$. Define C_0 and C_1 as in (43). Then, if $T \geq 2p^2 C_1$, we have*

$$\|\tilde{a}_T^{(p)}(t) - a_0^{(p)}(t/T)\| \leq \frac{1}{T}\left(5 C_0 C_1\, p^2 \right).$$

Lemma 6.5 *Let $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ satisfy Assumptions 1, 3, 4 (Appendix A), $\mathbb{E}X_{t,T} = 0$. Define C_0 and C_1 as in (43). Then,*

(i) if $T \geq 4pNC_1$, then $\|\bar{a}_{N,T}^{(p)}(t) - a_0^{(p)}(t/T)\| \leq \frac{N}{T}\left(9 C_0 C_1\, p \right)$.

(ii) if $T \geq 2p^2 C_1$, then $\|\bar{a}_{N,T}^{(p)}(t) - a_{N/T}^{(p)}(t/T)\| \leq \frac{1}{T}\left(5 C_0 C_1\, p^2 \right)$.

The proofs of Lemmas 6.4 and 6.5 can be found in Appendix D (Online Supplement). Note that, if $p^2 = o(T)$, as $T \to \infty$, then we have, by Lemma 6.4, that $\tilde{a}_T^{(p)}(t)$ and $a_0^{(p)}(t/T)$ are asymptotically equivalent in the sense that the Euclidean norm of the difference vanishes asymptotically. For $Np = o(T)$ we have, by Lemma 6.5, that $\bar{a}_{N,T}^{(p)}(t)$ and $a_0^{(p)}(t/T)$ are asymptotically equivalent, too. Therefore, since $0 \leq p^2 \leq Np$, we have: if $Np = o(T)$, then $\bar{a}_{N,T}^{(p)}(t)$ and $\tilde{a}_T^{(p)}(t)$ are asymptotically equivalent. Note further, that in the case of a tvAR(p) model, the quantity $\tilde{a}_T^{(p)}(t)$ coincides with the vector of coefficients $(a_1(t/T), \ldots, a_p(t/T))$, as is evident from the Yule-Walker equations.

It is worth mentioning that in case of a stationary process, where $C_1 = 0$, Lemmas 6.4 and 6.5 imply that the time-varying versions of the quantities and their local approximation coincide for all T (as they should).

Roueff and Sanchez-Perez (2016) prove a similar bound (Lemma 3):

$$\|\tilde{a}_T^{(p)}(t) - a_0^{(p)}(t/T)\| \leq \frac{D_1}{T}, \quad D_1 := \frac{Cp^{1/2}(p2^p + 1)}{\pi m_f},$$

for $T \geq T_0 := \frac{Cp^{3/2}}{\pi m_f}$. Note that (for larger p) their constant D_1 can be substantially larger than the constant in Lemma 6.4, which is largely due to a different representations of $\tilde{a}_T^{(p)}(t) - a_0^{(p)}(t/T)$ in their proof.

We will also need the following result that bounds the norm of $a_\Delta^{(p)}(u)$:

Lemma 6.6 *Let $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ satisfy Assumptions 1, 3, and 4, from Appendix A, and $\mathbb{E}X_{t,T} = 0$. Then, for $u \in \mathbb{R}$ and $\Delta \geq 0$, we have*

$$\|a_\Delta^{(p)}(u)\| \leq (2\pi)^{1/2}M_f/m_f =: C_0.$$

By Lemma 2 in Roueff and Sanchez-Perez (2016) we have $\|a_0^{(p)}(u)\| \leq 2^p$. Their proof adapts arguments from Lemma 4.2 in Dahlhaus and Giraitis (1998) where $\|\hat{a}_0^{(p)}(u)\| \leq 2^p$ almost surely is proven. We choose to work with the bound from Lemma 6.6, because it has the advantage that it does not depend on p. Further, note that neither of the bounds is sharp, as by Cauchy-Schwarz inequality we clearly have $\|a_0^{(1)}(u)\| \leq 1$.

In Lemmas C.1(i) and(ii) (Online Supplement) we show similar bounds for the approximation of $\bar{v}_{N,T}^{(p,h)}(t)$ with $v_0^{(p,h)}(t/T)$ or $v_{N/T}^{(p,h)}(t/T)$.

7 Conclusion

In this paper, we have presented a method to choose between different forecasting procedures, based on the empirical mean squared prediction errors the procedures achieve. Using the empirical rather than the asymptotic mean squared prediction error, our procedure automatically takes into account that different models should be preferred depending on the amount of data available, which is an important difference to the Focused Information Criterion by Claeskens and Hjort (2003). Working in the general framework of locally stationary time series we choose from two classes of forecasts that were motivated by approximating the serial dependence of the time series by time-varying or traditional autoregressive models. The procedure implicitly balances the modelling bias (which is lower if the model is more complex) and the variance of estimation (which increases for more complex models). Our two step procedure automatically chooses the number of forecasting coefficients to be used and the segment size from which the forecasting coefficients are estimated.

In a comprehensive simulation study we have illustrated that it is often advisable to use a forecasting procedure derived from a simpler model when not a vast amount of data is available. In particular, in the tvAR-models of our simulations, if the variation over time is not very pronounced and when the tangent processes are not close to being unit

34

root it is advisable to work with the simpler stationary model, even when the data are non-stationary.

As an important side result of our rigorous theoretical analysis of the method, we have shown that the localised Yule-Walker estimator is strongly, uniformly consistent under local stationarity.

References

Baillie, R. T. (1979). Asymptotic prediction mean squared error for vector autoregressive models. *Biometrika*, 66(3):675–8.

Bercu, B. (2001). On large deviations in the gaussian autoregressive process: stable, unstable and explosive cases. *Bernoulli*, 7(2):299–316.

Bercu, B., Gamboa, F., and Lavielle, M. (2000). Sharp large deviations for gaussian quadratic forms with applications. *ESAIM: Probability and Statistics*, 4:1–24.

Bercu, B., Gamboa, F., and Rouault, A. (1997). Large deviations for quadratic forms of stationary gaussian processes. *Stochastic Processes and their Applications*, 71(1):75 – 90.

Bercu, B. and Touati, A. (2008). Exponential inequalities for self-normalized martingales with applications. *The Annals of Applied Probability*, 18(5):1848–1869.

Bhansali, R. J. (1996). Asymptotically efficient autoregressive model selection for multi-step prediction. *Annals of the Institute of Statistical Mathematics*, 48:577–602.

Birr, S., Volgushev, S., Kley, T., Dette, H., and Hallin, M. (2016). Quantile spectral analysis for locally stationary time series. *ArXiv:1404.4605*.

Bosq, D. (1998). *Nonparametric statistics for stochastic processes: estimation and prediction*. Springer, New York, 2 edition.

Brillinger, D. R. (1975). *Time Series: Data Analysis and Theory*. Holt, Rinehart and Winston, Inc., New York.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer, New York.

Brownlees, C. T. and Gallo, G. M. (2008). On variable selection for volatility forecasting: The role of focused selection criteria. *Journal of Financial Econometrics*, 6:513–539.

Chen, Y., Härdle, W., and Pigorsch, U. (2010). Localized realized volatility modeling. *Journal of the American Statistical Association*, 105(492):1376–1393.

Claeskens, G., Croux, C., and Van Kerckhoven, J. (2007). Prediction-focused model selection for autoregressive models. *Australian & New Zealand Journal of Statistics*, 49(4):359–379.

Claeskens, G. and Hjort, N. L. (2003). The focused information criterion [with discussion]. *Journal of the American Statistical Association*, 98:900–916.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Annals of Statistics*, 25(1):1–37.

Dahlhaus, R. (2012). Locally stationary processes. In Rao, T. S., Rao, S. S., and Rao, C., editors, *Time Series Analysis: Methods and Applications*, volume 30 of *Handbook of Statistics*, pages 351 – 413. Elsevier.

Dahlhaus, R. and Giraitis, L. (1998). On the optimal segment length for parameter estimates for locally stationary time series. *Journal of Time Series Analysis*, 19(6):629–655.

Demmel, J. W. (1997). *Applied Numerical Linear Algebra*. siam, Philadelphia.

Dette, H., Preuß, P., and Vetter, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *Journal of the American Statistical Association*, 106(495):1113–1124.

Doukhan, P. (1994). *Mixing: Properties and Examples*. Springer, New York.

Dwivedi, J. and Subba Rao, S. (2010). A test for second order stationarity based on the discrete fourier transform. *Journal of Time Series Analysis*, 32:68–91.

Fryzlewicz, P. and Subba Rao, S. (2011). Mixing properties of ARCH and time-varying ARCH processes. *Bernoulli*, 17(1):320–346.

Giraitis, L., Kapetanios, G., and Price, S. (2013). Adaptive forecasting in the presence of recent and ongoing structural change. *Journal of Econometrics*, 177(2):153–170.

Goldenshluger, A. and Zeevi, A. (2001). Nonasymptotic bounds for autoregressive time series modeling. *Annals of Statistics*, 29(2):417–444.

Gray, R. M. (2009). *Toeplitz and Circulant Matrices: A review*. now publishers, Boston.

Kley, T., Fryzlewicz, P., and Preuss, P. (2016). *forecastSNSTS: Forecasting for Stationary and Non-Stationary Time Series*. R package version 1.0.0-0.

Lai, T. L. and Wei, C. Z. (1982). Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *Annals of Statistics*, 10(1):154–166.

Lewis, R. A. and Reinsel, G. C. (1988). Prediction error of multivariate time series with mis-specified models. *Journal of Time Series Analysis*, 9(1):43–57.

Magnus, J. R. and Neudecker, H. (1988). *Matrix differential calculus with applications in statistics and econometrics*. Wiley.

McDonald, D. J., Shalizi, C. R., and Schervish, M. (2016). Nonparametric risk bounds for time-series forecasting. *ArXiv:1212.0463*.

Nason, G. (2013). A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *Journal of the Royal Statistical Society: Series B*, 75:879–904.

Palma, W., Olea, R., and Ferreira, G. (2013). Estimation and forecasting of locally stationary processes. *Journal of Forecasting*, 32:86–96.

Paparoditis, E. (2009). Testing temporal constancy of the spectral structure of a time series. *Bernoulli*, 15:1190–1221.

Paparoditis, E. (2010). Validating stationarity assumptions in time series analysis by rolling local periodograms. *Journal of the American Statistical Association*, 105:839–851.

Peña, D. and Sánchez, I. (2007). Measuring the advantages of multivariate vs. univariate forecasts. *Journal of Time Series Analysis*, 28(6):886–909.

Preuß, P., Vetter, M., and Dette, H. (2013). A test for stationarity based on empirical processes. *Bernoulli*, 19(5B):2715–2749.

Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Reinsel, G. (1980). Asymptotic properties of prediction errors for the multivariate autoregressive model using estimated parameters. *Journal of the Royal Statistical Society: Series B*, 42(3):328–333.

Rohan, N. and Ramanathan, T. V. (2011). Order selection in arma models using the focused information criterion. *Australian & New Zealand Journal of Statistics*, 53(2):217—-231.

Rossi, B. (2013). Advances in forecasting under instability. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2B, pages 1203–1324. Elsevier.

Roueff, F. and Sanchez-Perez, A. (2016). Locally stationary processes prediction by autoregression. *ArXiv:1602.01942*.

Saulis, L. and Statulevičus, V. A. (1991). *Limit Theorems for Large Deviations*. Kluwer, Dordrecht.

Shibata, R. (1980). Asymptotically efficient selection of the order of the model for estimating parameters of a linear process. *Annals of Statistics*, 8:147–164.

Stroock, D. W. (2011). *Probability Theory: An Analytic View*. Cambridge University Press, New York, 2 edition.

Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Annals of Statistics*, 40(5):2601–2633.

Vogt, M. and Dette, H. (2015). Detecting smooth changes in locally stationary processes. *Annals of Statistics*, 43(2):713–740.

von Sachs, R. and Neumann, M. H. (2000). A wavelet-based test for stationarity. *Journal of Time Series Analysis*, 21:597–613.

Xia, Y. and Tong, H. (2011). Feature matching in time series modeling. *Statistical Science*, 26(1):21–46.

Yu, M. and Si, S. (2009). Moderate deviation principle for autoregressive processes. *Journal of Multivariate Analysis*, 100(9):1952 – 1961.

Zhang, Y. and Koreisha, S. (2015). Adaptive order determination for constructing time series forecasting models. *Communications in Statistics - Theory and Methods*, 44(22):4826–4847.

A Assumptions, technical details and proofs

In this section we provide proofs of the main results in the paper. We keep the proofs as short as possible, by deferring technical results to the Online Supplement.

We use the usual symbols to denote sets of numbers: $\mathbb{N} := \{0, 1, 2, \ldots\}$, $\mathbb{N}^* := \{1, 2, \ldots\}$, etc. Further, for a vector $x \in \mathbb{R}^d$ we denote by $\|x\|$ its Euclidean norm and by $\|x\|_\infty$ its sup-norm. For a $m \times n$ matrix $A = (a_{i,j})_{i=1,\ldots,m,\ j=1,\ldots,n} \in \mathbb{R}^{m \times n}$ we denote by $\|A\| := \sup\{\|Ax\|/\|x\| : x \neq 0\}$ its spectral norm, by $\|A\|_1 := \max_{j=1,\ldots,n} \sum_{i=1}^{m} |a_{i,j}|$ and $\|A\|_\infty := \max_{i=1,\ldots,m} \sum_{j=1}^{n} |a_{i,j}|$ the matrix norms induced by the corresponding vector norms, and by $\|A\|_F := (\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{i,j}|^2)^{1/2} = \|\text{vec} A\|$ the Frobenius norm.

In some of the technical lemmas we denote by $\|\cdot\|_M$ or $\|\cdot\|_v$ an arbitrary matrix or vector norm, respectively. Special properties we require include *submultiplicativity* of a matrix norm, and *compatibility* of a matrix norm with a vector norm. A matrix norm which satisfies $\|AB\|_M \leq \|A\|_M \|B\|_M$ for all square matrices ($m = n$), is said to be submultiplicative. A matrix norm $\|\cdot\|_M$ and vector norm $\|\cdot\|_v$ are said to be compatible if $\|Ax\|_v \leq \|A\|_M \|x\|_v$ for all square matrices A and vectors x (of sizes that allow for the matrix product).

To rigorously prove the results, some technical assumptions are in order. We assume the time series $X_{1,T}, \ldots, X_{T,T}$ to be from a locally stationary process as described in Roueff and Sanchez-Perez (2016); see Assumption 1, below. More precisely, we will work with the doubly indexed process $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ of random variables with finite variances and use the following definitions:

The *local time-varying covariance function* is defined for all $t \in \mathbb{Z}, T \in \mathbb{N}^*$ and $k \in \mathbb{Z}$ as

$$\tilde{\gamma}_{k,T}(t) = \text{Cov}\left(X_{t,T}, X_{t-k,T}\right). \tag{35}$$

A *local spectral density* f is a $\mathbb{R}^2 \to \mathbb{R}_+$ function, (2π)-periodic and locally integrable with respect to the second variable. The *local covariance function* γ associated with the time varying spectral density f is defined on $\mathbb{R} \times \mathbb{Z}$ by

$$\gamma_k(u) = \int_{-\pi}^{\pi} \exp\left(ik\lambda\right) f(u, \lambda) d\lambda. \tag{36}$$

Assumption 1 (Locally stationary processes) $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ *is an array of random variables with finite variances. We say that* $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ *is locally stationary with local spectral density* f *if the time varying covariance function of* $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ *and the local covariance function associated with* f *satisfy*

$$\left|\tilde{\gamma}_{k,T}\left(t\right) - \gamma_k\left(\frac{t}{T}\right)\right| \leq \frac{C}{T}, \tag{37}$$

where $C \geq 0$ *is a constant.*

Many processes considered in the literature fulfil this definition. Examples include any second order stationary process (then we have $C = 0$), the general (linear) locally stationary process first considered by Dahlhaus (1997), but also non-linear processes as elaborated by Roueff and Sanchez-Perez (2016) [see their Example 3].

In the tvAR-model that we used to motivate our prediction approach in Section 3.2, see (9), and as examples in Section 4, the local covariances are naturally those of the stationary AR-process when the parameter u of the coefficient functions is chosen as t/T. We will refer to these AR-processes as the *tangent processes* of the tvAR-process.

We recall, from Section 3.2, that the tvAR(p)model is used to approximate the linear dynamic structure of the data, but that we do *not* assume that the data actually satisfies it. Thus our results are more general.

We will further make an assumption with respect to the process's range of dependence in terms of its strong mixing coefficient:

$$\alpha(k) := \sup_{T\in\mathbb{N}^*} \sup_{t\in\mathbb{Z}} \alpha_T(s, t+k),$$

where

$$\alpha_T(s, t+k) := \sup_{A\in\sigma(X_{s,T}:s\leq t)} \sup_{B\in\sigma(X_{s,T}:s\geq t+k)} \left|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)\right|.$$

More specifically, we will assume that the underlying process satisfies

Assumption 2 (geometrically α-mixing) *There exist constants* $K > 0$ *and* $\rho > 1$ *such that*

$$\alpha(n) \leq K\rho^{-n}. \tag{38}$$

Assumption 2 is satisfied for a broad class of locally stationary time series models (linear and non-linear); see, for example Fryzlewicz and Subba Rao (2011) or Vogt (2012).

It is further common to make assumptions with respect to the local spectral density.

Assumption 3 *The local spectral density f is bounded from above and below:*

$$0 < m_f \le |f(u, \lambda)| \le M_f. \tag{39}$$

As an immediate consequence to (35), (36), and (37) we have that

$$\left| \text{cum}(X_{t-k,T}, X_{t,T}) - \int_{-\pi}^{\pi} f(t/T, \lambda) e^{ik\lambda} d\lambda \right| \le C/T.$$

Taking (39) into account we have (choosing $k = 0$) that

$$\pi m_f < \sigma_{t,T}^2 := \text{Var}(X_{t,T}) < 3\pi M_f, \tag{40}$$

for all $T \ge \frac{C}{\pi m_f}$.

We further assume that $f(u, \lambda)$ is smooth with respect to the local time u:

Assumption 4 *The local spectral density f is continuously differentiable with respect to the first argument and the partial derivative is uniformly bounded. More precisely, assume the existence of $M_f' \ge 0$ such that*

$$\left| \frac{\partial}{\partial u} f(u, \lambda) \right| \le M_f'. \tag{41}$$

Note that by assuming the existence of a continuous partial derivative we have, by Leibniz's integral rule, that $\gamma_k'(u)$ exists and has the following form

$$\gamma_k'(u) := \frac{\partial}{\partial u} \gamma_k(u) = \frac{\partial}{\partial u} \int_{-\pi}^{\pi} \exp(i\ell\lambda) f(u, \lambda) d\lambda = \int_{-\pi}^{\pi} \exp(i\ell\lambda) \frac{\partial}{\partial u} f(u, \lambda) d\lambda, \tag{42}$$

which, in particular, implies that $|\gamma_k'(u)| \le 2\pi M_f'$.

Finally, to prove the exponential inequalities in Lemma (A.2), we make assumptions regarding moments of the marginal distributions. More precisely, we will assume the following, generalized version of Bernstein's condition (cf. Saulis and Statulevičus (1991))

Assumption 5 (Bernstein-type condition) *Assume that there exists $c > 0$ and $d \ge 1/2$, such that*

$$\mathbb{E}|X_{t,T}|^k \le c^{k-2} (k!)^d \sigma_{t,T}^2 \quad t = 1, \dots, T; \ k = 2, 3, \dots$$

By the following lemma we have that many Gaussian processes will satisfy Assumption 5 with $c := 6\pi M_f$ and $d := 1/2$.

Lemma A.1 *Let $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ satisfy Assumptions 1 and 3. Assume that $T \ge \frac{C}{\pi m_f}$ and that the standardized variables $X_{t,T}/\sigma_{t,T}$ are b-sub-Gaussian ($b > 0$); i.e.,*

$$\mathbb{E}\left(\exp\left(\xi X_{t,T}/\sigma_{t,T} \right) \right) \le \exp\left(\frac{b^2 |\xi|^2}{2} \right), \quad \xi \in \mathbb{R}.$$

Then, the process $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ satisfies Assumption 5 with $c := 6\pi b M_f$ and $d := 1/2$.

Proof. Deferred to Appendix B (Online Supplement).

We now introduce two constants that combine constants from the assumptions. Stating the results in terms of these constants will help to better interpret the bounds and significantly shorten otherwise complicated expressions. To this end, we define

$$C_0 := (2\pi)^{1/2} M_f / m_f, \quad C_1 := (2\pi M_f' + C) m_f^{-1}. \tag{43}$$

The constant C_0 can be interpreted in terms of the strength of serial correlation. Note that C_0 will be smaller if there is little variation (uniform in local time) of the spectral density with respect to frequency. In particular, it will be minimal if the spectral density is constant. This would corresponds to the case of white noise. The constant C_1 can be interpreted as divergence from stationarity. In particular, note the meaning of the two summands of the first factor. The constant M_f' corresponds to the rapidity of changes in stationarity and will vanish in case of stationarity. The constant C corresponds to the quality of locally approximating the correlation structure with a stationary processes correlation structure. It, also, vanishes if the underlying process is stationary.

Before we state the proof of Theorem 3.1, we derive auxiliary results that are concerned with a general mean squared prediction error defined as

$$\text{MSPE}_{s,m,N,T}^{(p,h)} := \frac{1}{m} \sum_{t=s+1}^{s+m} \left(X_{t+h,T} - \sum_{i=1}^{p} \hat{v}_{i;N,T}^{(p,h)}(t) X_{t-i+1,T} \right)^2, \tag{44}$$

with $\hat{v}_{i;N,T}^{(p,h)}(t)$ defined in (4) and $\hat{v}_{i;0,T}^{(p,h)}(t) := \hat{v}_{i;t,T}^{(p,h)}(t)$.

The first auxiliary result (Lemma A.2) entails that the quantity defined in (44) is, with high probability, close to

$$\overline{\text{MSPE}}_{s,m,N,T}^{(p,h)} := \frac{1}{n} \sum_{t=s+1}^{s+m} \mathbb{E} \left(X_{t+h,T} - \sum_{i=1}^{p} \bar{v}_{i,N,T}^{(p,h)}(t) X_{t-i+1,T} \right)^2, \tag{45}$$

with

$$\begin{aligned} \bar{v}_{N,T}^{(p,h)}(t) &:= \left(\bar{v}_{1;N,T}^{(p,h)}(t), \bar{v}_{2;N,T}^{(p,h)}(t), \cdots, \bar{v}_{p;N,T}^{(p,h)}(t) \right) \\ &:= e_1' \left(\bar{A}_{N,T}^{(p)}(t) \right)^h := e_1' \left(e_1 \left(\bar{a}_{N,T}^{(p)}(t) \right)' + H \right)^h, \end{aligned} \tag{46}$$

where $\bar{a}_{N,T}^{(p)}(t)$ is defined in (33), e_1 denotes the first canonical unity vector of dimension p and H denotes a $p \times p$ Jordan block with all eigenvalues equal to zero:

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad H := \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}. \tag{47}$$

41

By our second auxiliary result (Lemma A.3) we have that $\overline{\mathrm{MSPE}}^{(p,h)}_{s,m,N,T}$ in turn can be approximated by $\mathrm{MSPE}^{(p,h)}_{N/T,m/T}(s/T)$, where the function $\mathrm{MSPE}^{(p,h)}_{\Delta_1,\Delta_2}(u)$ is defined, for $u \in \mathbb{R}$, $\Delta_1, \Delta_2 \geq 0$, in terms of the second order properties of the process:

$$\mathrm{MSPE}^{(p,h)}_{\Delta_1,\Delta_2}(u) := \int_0^1 g^{(p,h)}_{\Delta_1}\Big(u + \Delta_2(1-x)\Big)\mathrm{d}x, \tag{48}$$

with $g^{(0,h)}_{\Delta}(u) := \gamma_0(u)$ and, for $p \in \mathbb{N}^*$,

$$g^{(p,h)}_{\Delta}(u) := \gamma_0(u) - 2\big(v^{(p,h)}_{\Delta}(u)\big)'\gamma^{(p,h)}_0(u) + \big(v^{(p,h)}_{\Delta}(u)\big)'\Gamma^{(p)}_0(u)v^{(p,h)}_{\Delta}(u) \tag{49}$$

$$\gamma^{(p,h)}_0(u) := \big(\gamma_h(u), \ldots, \gamma_{h+p-1}(u)\big)',$$

where

$$v^{(p,h)}_{\Delta}(u) := \big(v^{(p,h)}_{1;\Delta}(u), v^{(p,h)}_{2;\Delta}(u), \cdots, v^{(p,h)}_{p;\Delta}(u)\big)$$
$$:= e_1'\big(A^{(p)}_{\Delta}(u)\big)^h := e_1'\big(e_1\big(a^{(p)}_{\Delta}(t)\big)' + H\big)^h, \tag{50}$$

with e_1 and H defined in (47) and, for every $u \in \mathbb{R}$ and $\Delta \geq 0$,

$$a^{(p)}_{\Delta}(u) := \Gamma^{(p)}_{\Delta}(u)^{-1}\gamma^{(p)}_{\Delta}(u), \tag{51}$$

where

$$\gamma^{(p)}_{\Delta}(u) := \int_0^1 \gamma^{(p)}(u + \Delta(x-1))\mathrm{d}x, \quad \gamma^{(p)}(u) := [\gamma_1(u) \ \ldots \ \gamma_p(u)]',$$
$$\Gamma^{(p)}_{\Delta}(u) := \int_0^1 \Gamma^{(p)}(u + \Delta(x-1))\mathrm{d}x, \quad \Gamma^{(p)}(u) := (\gamma_{i-j}(u); i,j = 1, \ldots, p). \tag{52}$$

Some comments on the quantities defined in (45), (48), and (51) are in order: Note that $\overline{\mathrm{MSPE}}^{(p,h)}_{s,m,N,T}$ is defined as the expectation of a modified version of $\mathrm{MSPE}^{(p,h)}_{s,m,N,T}$, the modification being that $\hat{v}^{(p,h)}_{N,T}(t)$ is exchanged by $\bar{v}^{(p,h)}_{N,T}(t)$. As before, we will denote $\bar{v}^{(p,h)}_{0,T}(t) := \bar{v}^{(p,h)}_{t,T}(t)$.

Note that the definition of $a^{(p)}_{\Delta}(u)$ in (51) resembles the solution to the Yule-Walker equations. In fact, for $\Delta = 0$ we have the solution to the Yule-Walker equations with the locally approximating stationary covariance structure, at local time u. We allow for $\Delta > 0$ to capture the evolving second moments of the processes with indices from $[u - \Delta, u]$. This averaging is necessary to take the bias from the changing second moments into account, as we do not require $N/T = o(1)$ for our results. The quantity $g^{(p,h)}_{N/T}(t/T)$ approximates the MSPE of $f^{\mathrm{ls}}_{t,h;p,N}$ defined in (5). Note that expectation of the empirical mean squared prediction error (44) we are considering is naturally an average of these quantities:

$$\mathbb{E}[\mathrm{MSPE}^{(p,h)}_{s,m,N,T}] = \frac{1}{m}\sum_{t+h \in M_j}\mathbb{E}[(X_{t+h,T} - f^{\mathrm{ls}}_{t,h;p,N})^2].$$

We now state the results that the quantities defined in (44) and (45) are close, with high probability.

Lemma A.2 *Let $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ satisfy Assumptions 1–5 and $\mathbb{E}X_{t,T} = 0$. Then, for every $m, h \in \mathbb{N}^*$, $p \in \mathbb{N}$, $N \geq p+1$, $\varepsilon > 0$ and $T \geq 5C_1 p^2$, with $\mathrm{MSPE}_{s,m,N,T}^{(p,h)}$ defined in (44) and $\overline{\mathrm{MSPE}}_{s,m,N,T}^{(p,h)}$ defined in (45), we have that*

$$\mathbb{P}\left(\left| \mathrm{MSPE}_{s,m,N,T}^{(p,h)} - \overline{\mathrm{MSPE}}_{s,m,N,T}^{(p,h)} \right| > \varepsilon \right) \leq P_{m,N}^{(p,h)}(\varepsilon)$$

and

$$\mathbb{P}\left(\left| \mathrm{MSPE}_{s,m,0,T}^{(p,h)} - \overline{\mathrm{MSPE}}_{s,m,0,T}^{(p,h)} \right| > \varepsilon \right) \leq P_{m,s}^{(p,h)}(\varepsilon)$$

with

$$P_{m,N}^{(p,h)}(\varepsilon) := (1 + 4p + 2p^2)$$

$$\cdot \exp\left(- \frac{\frac{\varepsilon^2}{(p+1)^4}}{8\left((2C_0+1)^{4h \frac{\bar{C}_{1,2}(h+p-1)}{m}} + (\frac{\varepsilon}{2(p+1)^2})^{(3+8d)/(2+4d)} \left((2C_0+1)^{2h \frac{\bar{C}_{2,2}(h+p-1)}{m}} \right)^{1/(2+4d)} \right)} \right)$$

$$+ 6mp^2(p+1) \exp\left(- \frac{\eta^2}{2\left(\bar{C}_{1,1} \frac{p}{N-p} + \eta^{(3+4d)/(2+2d)} \left(\bar{C}_{2,1} \frac{p}{N-p} \right)^{1/(2+2d)} \right)} \right),$$

where

$$\eta := \frac{m_f}{4p} \min\left\{ 1, \bar{\mu}/(8C_0) \right\}, \qquad \bar{\mu} := 2^{1-h} \frac{\mu}{\mu + h(2C_0)^{h-1}},$$

$$\mu := \frac{\bar{\varepsilon}}{2\left((2C_0+1)^{2h} + \bar{\varepsilon} \right)^{1/2}}, \qquad \bar{\varepsilon} := \frac{\varepsilon/(p+1)^2}{2\left((6\pi M_f c^2 24^d)^2 + \varepsilon^2/(p+1)^4 \right)^{1/4}},$$

and the constants C_0, C_1, and $\bar{C}_{1,1}$, $\bar{C}_{1,2}$, $\bar{C}_{2,1}$, $\bar{C}_{2,2}$, and m_f, M_f, and c, d are defined in (43), (73), and Assumptions 3 and 5, respectively.

As detailed in Remark 3.2 of the main text, $P_{m,N}^{(p,h)}(\varepsilon) \to 0$, as $T \to \infty$, at an exponential rate if p, h, ε are fixed and $m, N \to \infty$ faster than logarithmically.

We now state the result that the quantities defined in (45) and (48) are close. The quality of the approximation depends on the parameters T, p and h, but is uniform with respect to s, m and N:

Lemma A.3 *Let $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ satisfy Assumptions 1–5 and $\mathbb{E}X_{t,T} = 0$. Then, for every $m, h \in \mathbb{N}^*$, $p \in \mathbb{N}$, $T \geq 3h2^h C_1 p^2$, and $N \geq 1+p$, with $\overline{\mathrm{MSPE}}_{s,m,N,T}^{(p,h)}$ defined in (45) and $\mathrm{MSPE}_{\Delta_1,\Delta_2}^{(p,h)}(u)$ defined in (48), we have*

$$\left| \overline{\mathrm{MSPE}}_{s,m,N,T}^{(p,h)} - \mathrm{MSPE}_{N/T,m/T}^{(p,h)}(s/T) \right| \leq 3h2^{h+4}(C_0)^{2h+1} \frac{p^2}{T}(2\pi M_f' + C)$$

and

$$\left| \overline{\mathrm{MSPE}}_{s,m,0,T}^{(p,h)} - \mathrm{MSPE}_{s/T,m/T}^{(p,h)}(s/T) \right| \leq 3h2^{h+4}(C_0)^{2h+1} \frac{p^2}{T}(2\pi M_f' + C).$$

The proofs of Lemmas A.2 and A.3 are long and technical. We therefore defer them to Appendix B (Online Supplement).

A few comments about Lemma A.3 are in order. Note that the approximation error is zero in case of a stationary time series, as then $2\pi M_f' + C = 0$. Note further, that the approximation will be good, if h and p are small compared to T. More precisely, if $h(2C_0^2)^h p^2 = o(T)$, then the difference will vanish asymptotically. In particular, if $h = O(1)$, then it would suffice to assume that $p = o(T^{1/2})$, for the approximation error to vanish.

We now provide the proofs of the results from Section 3.3 and conclude with the proofs of the results from Section 6.

Proof of Theorem 3.1.

We compact notation: denote $s_2 := T - h + 1$, $\text{MSPE}_{s_{i+1},m,N,T}^{(p_1,h)}$ by X_i and $\text{MSPE}_{s_{i+1},m,0,T}^{(p_2,h)}$ by Y_i. Further, denote $\text{MSPE}_{N/T,m/T}^{(p_1,h)}(\frac{s_{i+1}}{T})$ and $\text{MSPE}_{s_1/T,m/T}^{(p_2,h)}(\frac{s_{i+1}}{T})$ by $\bar{Y}_i$ and $\bar{X}_i$, respectively. Further, we abbreviate $A := Y_1 - X_1(1+\delta)$ and $B := Y_1 - Y_2 + (X_2 - X_1)(1+\delta)$. In Appendix B (Online Supplement) we prove that

$$\mathbb{P}\Big((\hat{R}_{T,2}^{s,ls}(h) \geq 1 + \delta \text{ and } \hat{R}_{T,3}^{s,ls}(h) \geq 1 + \delta) \text{ or } (\hat{R}_{T,2}^{s,ls}(h) < 1 + \delta \text{ and } \hat{R}_{T,3}^{s,ls}(h) < 1 + \delta)\Big)$$

$$\geq 1 - \sum_{p_1,p_2=0}^{p_{\max}} \sum_{N \in \mathcal{N}} \Big(\mathbb{P}\big(|A| \leq f(\delta)/2\big) + \mathbb{P}\big(|B| > f(\delta)/2\big)\Big). \tag{53}$$

We now bound the part of the right hand side of (53) that involves the quantity A. Using the fact that

$$|\bar{Y}_1 - \bar{X}_1(1+\delta)| = |Y_1 + \bar{Y}_1 - Y_1 - X_1(1+\delta) + (X_1 - \bar{X}_1)(1+\delta)|$$
$$\leq |Y_1 - X_1(1+\delta)| + |\bar{Y}_1 - Y_1| + |X_1 - \bar{X}_1|(1+\delta),$$

we have the first inequality of

$$\mathbb{P}\Big(|A| \leq f(\delta)/2\Big) = \mathbb{P}\Big(|Y_1 - X_1(1+\delta)| \leq f(\delta)/2\Big)$$

$$\leq \mathbb{P}\Big(|Y_1 - \bar{Y}_1| + |X_1 - \bar{X}_1|(1+\delta) \geq |\bar{Y}_1 - \bar{X}_1(1+\delta)| - f(\delta)/2\Big)$$

$$\leq \mathbb{P}\Big(|Y_1 - \bar{Y}_1| \geq \frac{1}{2}(|\bar{Y}_1 - \bar{X}_1(1+\delta)| - f(\delta)/2)\Big)$$

$$\quad + \mathbb{P}\Big(|X_1 - \bar{X}_1| \geq \frac{1}{2(1+\delta)}(|\bar{Y}_1 - \bar{X}_1(1+\delta)| - f(\delta)/2)\Big)$$

$$\leq \mathbb{P}\Big(|Y_1 - \bar{Y}_1| > f(\delta)/10\Big) + \mathbb{P}\Big(|X_1 - \bar{X}_1| > \frac{f(\delta)}{10(1+\delta)})\Big) \tag{54}$$

$$\leq P_{m,T-m}^{(p_2,h)}\Big(\frac{f(\delta)}{20}\Big) + P_{m,N}^{(p_1,h)}\Big(\frac{f(\delta)}{20(1+\delta)}\Big) \leq 2P_{m,N_{\min}}^{(p_{\max},h)}\Big(\frac{f(\delta)}{20(1+\delta)}\Big). \tag{55}$$

For the inequality in (54) we have used the definition of $f(\delta)$ and $1/4 > 1/10$. For the first inequality in (55) we have used that, we have

$$T \geq 3h2^{h+4}\big(C_0\big)^{2h+1} p_{\max}^2 (2\pi M_f' + C)/\varepsilon,$$

with $\varepsilon := f(\delta)/(20(1+\delta))$, by assumption, and thus, by Lemmas A.2 and A.3, we have

$$\mathbb{P}\Big(\Big|\text{MSPE}_{s,m,N,T}^{(p,h)} - \text{MSPE}_{N/T,m/T}^{(p,h)}(s/T)\Big| > 2\varepsilon\Big) \leq P_{m,N}^{(p,h)}(\varepsilon) \tag{56}$$

and

$$\mathbb{P}\left(\left|\mathrm{MSPE}_{s,m,0,T}^{(p,h)} - \mathrm{MSPE}_{s/T,m/T}^{(p,h)}(s/T)\right| > 2\varepsilon\right) \leq P_{m,s}^{(p,h)}(\varepsilon). \tag{57}$$

For the second inequality in (55) we have used that

$$p_1 \leq p_2 \Rightarrow P_{m,N}^{(p_1,h)}(\varepsilon) \leq P_{m,N}^{(p_2,h)}(\varepsilon), \quad N_1 \leq N_2 \Rightarrow P_{m,N_1}^{(p,h)}(\varepsilon) \geq P_{m,N_2}^{(p,h)}(\varepsilon),$$

and $\varepsilon_1 \leq \varepsilon_2 \Rightarrow P_{m,N}^{(p,h)}(\varepsilon_1) \geq P_{m,N}^{(p,h)}(\varepsilon_2)$.

We now bound the part of the right hand side of (53) that involves the quantity B. We have

$$\mathbb{P}\left(|B| > f(\delta)/2\right) = \mathbb{P}\left(|Y_1 - Y_2 + (X_2 - X_1)(1 + \delta)| > f(\delta)/2\right)$$
$$\leq \mathbb{P}\left(|Y_1 - Y_2| > f(\delta)/4\right) + \mathbb{P}\left(|X_2 - X_1| > \frac{f(\delta)}{4(1 + \delta)}\right)$$
$$\leq 2P_{m,T-m}^{(p_2,h)}\left(\frac{f(\delta)}{20}\right) + 2P_{m,N}^{(p_1,h)}\left(\frac{f(\delta)}{20(1 + \delta)}\right) \leq 4P_{m,N_{\min}}^{(p_{\max},h)}\left(\frac{f(\delta)}{20(1 + \delta)}\right).$$

Note that we have

$$\mathbb{P}\left(\left|\mathrm{MSPE}_{s_1,m,N,T}^{(p,h)} - \mathrm{MSPE}_{s_2,m,N,T}^{(p,h)} > 5\varepsilon\right)\right.$$
$$\leq \mathbb{P}\left(\left|\mathrm{MSPE}_{s_1,m,N,T}^{(p,h)} - \mathrm{MSPE}_{N/T,m/T}^{(p,h)}(s_1/T)\right| > 2\varepsilon\right)$$
$$+ \mathbb{P}\left(\left|\mathrm{MSPE}_{s_2,m,N,T}^{(p,h)} - \mathrm{MSPE}_{N/T,m/T}^{(p,h)}(s_2/T)\right| > 2\varepsilon\right)$$
$$+ I\left\{\left|\mathrm{MSPE}_{N/T,m/T}^{(p,h)}(s_1/T) - \mathrm{MSPE}_{N/T,m/T}^{(p,h)}(s_2/T)\right| > \varepsilon\right)\right\},$$

where the first two terms can be bound by an application of (56) and the indicator function vanishes for all T satisfying the condition of the Theorem, because

$$\left|\mathrm{MSPE}_{N/T,m/T}^{(p,h)}(s_1/T) - \mathrm{MSPE}_{N/T,m/T}^{(p,h)}(s_2/T)\right| \leq 4(2h + 1)(C_0)^{2h+1}M_f'\left|\frac{s_1 - s_2}{T}\right|,$$

where Lemma C.4(iv) from the Online Supplement was employed for the last inequality. This completes the proof of the theorem. $\square$

Remark A.4 *Equations (56)–(57), which are immediate consequences of Lemmas A.2 and A.3, can be used to derive the almost sure convergence of*

$$\left|\mathrm{MSPE}_{s,m,N,T}^{(p,h)} - \mathrm{MSPE}_{N/T,m/T}^{(p,h)}(s/T)\right| \text{ and } \left|\mathrm{MSPE}_{s,m,0,T}^{(p,h)} - \mathrm{MSPE}_{s/T,m/T}^{(p,h)}(s/T)\right|,$$

under appropriate conditions, using a classical Borel-Cantelli argument.

This asymptotic view of $\mathrm{MSPE}_{s,m,N,T}^{(p,h)}$ and $\mathrm{MSPE}_{s,m,0,T}^{(p,h)}$, in particular, implies that we may interpret $\mathrm{MSPE}_{\Delta_1,\Delta_2}^{(p,h)}(u)$ as an approximation of the expectation of the empirical MSPE for an h-step ahead linear forecast of order p, where observations up to (local) time u have been made. The Δ_1 and Δ_2 are (localised) length which are related to the segment length of observations used for the estimation of the forecasting coefficients and the segment from which the observations $X_{t+h,T}$ that are being forecasted are taken, respectively.

Proofs of Corollary 3.3 and 3.4.

Following the proof of Theorem 3.1 we have that Corollary 3.3 holds if we show that

$$\min_{N \in \mathcal{N}} \left| \text{MSPE}_{s_1/T, m/T}^{(1,1)}\left(\frac{s_1}{T}\right) - (1+\delta) \cdot \text{MSPE}_{N/T, m/T}^{(1,1)}\left(\frac{s_1}{T}\right) \right| \geq \delta \pi m_f (1 - \rho^2).$$

Likewise, to show Corollary 3.4, we bound the left hand side inequality with $\pi m_f D_{\inf}^2/2$ on the right hand side.

Denoting

$$\gamma_k(u, \Delta) := \int_0^1 \gamma_k(u + \Delta(x - 1))\mathrm{d}x = \Delta^{-1} \int_{u-\Delta}^u \gamma_k(v)\mathrm{d}v$$

we have, by definition (48), that

$$\text{MSPE}_{\Delta_1,\Delta_2}^{(1,1)}(u) = \int_0^1 g_{\Delta_1}^{(1,1)}\big(u + \Delta_2(1-x)\big)\mathrm{d}x$$

$$= \Delta_2^{-1} \int_u^{u+\Delta_2} \left(\gamma_0(w) - 2\frac{\gamma_1(w; \Delta_1)}{\gamma_0(w; \Delta_1)}\gamma_1(w) + \left(\frac{\gamma_1(w; \Delta_1)}{\gamma_0(w; \Delta_1)}\right)^2 \gamma_0(w) \right)\mathrm{d}w.$$

To find the lower bound we want, it therefore suffices to proof lower bounds, for every $w \in [s_1/T, (s_1 + m)/T]$, of the following difference

$$\begin{aligned} &\left(\left(\gamma_0(w) - 2\frac{\gamma_1(w; s_1/T)}{\gamma_0(w; s_1/T)}\gamma_1(w) + \left(\frac{\gamma_1(w; s_1/T)}{\gamma_0(w; s_1/T)}\right)^2 \gamma_0(w)\right) \right. \\ &\left. - (1+\delta)\left(\gamma_0(w) - 2\frac{\gamma_1(w; N/T)}{\gamma_0(w; N/T)}\gamma_1(w) + \left(\frac{\gamma_1(w; N/T)}{\gamma_0(w; N/T)}\right)^2 \gamma_0(w)\right) \right), \end{aligned} \tag{58}$$

or for $-1 \times (58)$. For notational convenience we omit the w's and denote

$$E := \frac{\gamma_1(w, N/T)}{\gamma_0(w, N/T)}, \text{ and } F := \frac{\gamma_1(w, s_1/T)}{\gamma_0(w, s_1/T)}.$$

In the Appendix B (Online Supplement) we show that

$$(58) = \gamma_0 \left(\left(F - \frac{\gamma_1}{\gamma_0}\right)^2 - \left(\frac{\gamma_1}{\gamma_0} - E\right)^2 - \delta\left(1 - \left(\frac{\gamma_1}{\gamma_0}\right)^2 + \left(\frac{\gamma_1}{\gamma_0} - E\right)^2\right) \right). \tag{59}$$

By (15), we have $|F - \frac{\gamma_1}{\gamma_0}| \geq D_{\inf}$ and by (14), we have $|F - \frac{\gamma_1}{\gamma_0}| \leq D_{\sup}$. Further, we have that $\frac{\gamma_1}{\gamma_0} - E = O(N/T)$, uniformly with respect to ω, which can be seen as follows: first, note that

$$\left| \gamma_k(w, N/T) - \gamma_k(w, 0) \right| \leq \int_0^1 \left| \gamma_k(w) - \gamma_k(w - \frac{N}{T}(1-x)) \right|\mathrm{d}x$$

$$\leq 2\pi M_f' \int_0^1 \frac{N}{T}(1-x)\mathrm{d}x = \pi M_f' \frac{N}{T}$$

Further, note that we have $\frac{x}{y} - \frac{x_0}{y_0} = \frac{1}{y_0}\left(\frac{x}{y}(y_0 - y) + (x - x_0)\right)$ and thus

$$\left| \frac{\gamma_1}{\gamma_0} - E \right| \leq \frac{1}{\gamma_0(w; N/T)}\left(\frac{|\gamma_1|}{\gamma_0} + 1\right)\pi M_f' \frac{N}{T} \leq \frac{M_f'}{m_f}\frac{N}{T}, \tag{60}$$

46

where we have used that $2\pi m_f \leq \gamma_0(w; \Delta) := \int_0^1 \gamma_0(w + \Delta(x-1))\mathrm{d}x$ and $|\gamma_1|/\gamma_0 \leq 1$. Employing (59), we have now brought the tools together to prove Corollary 3.3:

$$-1 \times (58) \geq 2\pi m_f\Big(1 - \rho^2\Big)\Big(\delta/2 + \delta/2 - D_{\sup}^2/\big(1 - \rho^2\big)\Big)\Big)\Big) \geq \pi m_f \delta(1 - \rho^2).$$

For the first inequality we have used the fact that $(\gamma_1/\gamma_0 - E)^2 \geq 0$ and the definitions of ρ and $D_{\sup}$. For the second inequality we have used the condition imposed on δ.

Finally, employing (59) again, we prove Corollary 3.4:

$$(58) \geq m_f\Big(\Big(F - \frac{\gamma_1}{\gamma_0}\Big)^2 - \Big(\frac{\gamma_1}{\gamma_0} - E\Big)^2 - 2\delta\Big)$$

$$\geq 2\pi m_f\Big(D_{\inf}^2 - \Big(\frac{M_f' N}{m_f T}\Big)^2 - 2\delta\Big) \geq 2\pi m_f\Big(D_{\inf}^2/2 - 2\delta\Big) \geq \pi m_f D_{\inf}^2/2,$$

where in the first inequality we have used

$$\Big(\frac{\gamma_1}{\gamma_0} - E\Big)^2 \leq \Big(\frac{M_f' N}{m_f T}\Big)^2 \leq 1,$$

as we have $D_{\inf} \leq 2$ and thus $\max \mathcal{N} \leq (m_f/M_f')T$ follows from (16). For the second inequality we have used the definition of $D_{\inf}$ and assumption (60) of the Corollary, by which we have $D_{\inf}^2/2 \geq \big(M_f' N/(m_f T)\big)^2$. Finally, for the third inequality we have used that by assumption in the Corollary $2\delta \leq D_{\inf}^2/4$. $\qquad\square$

We now proceed with the proofs of the results from Section 6.

Proof of Theorem 6.1.

Let $M := \hat{\Gamma}_{N,T}^{(p)}(t)$, $M_0 := \mathbb{E}M$, $v := \hat{\gamma}_{N,T}^{(p)}(t)$, and $v_0 := \mathbb{E}v$. By Lemma E.3(ii-c) we deduce that M_0 is invertible for $T \geq 2p^2 C_1$, because it is positive definite with smallest eigenvalue larger or equal to $m_f/2$. An application of Lemma F.6, with the spectral norm as the matrix norm and the Euclidean norm as the vector norm yields

$$\mathbb{P}\big(\|\hat{a}_{N,T}^{(p)}(t) - \bar{a}_{N,T}^{(p)}(t)\| > \varepsilon\big) = \mathbb{P}\big(\|M^{-1}v - M_0^{-1}v_0\| > \varepsilon\big)$$

$$\leq \mathbb{P}\Big(\|M - M_0\| > \frac{1}{2\|M_0^{-1}\|}\Big) + \mathbb{P}\Big(\|v - v_0\| > \frac{\varepsilon}{4}\frac{1}{\|M_0^{-1}\|}\Big)$$

$$+ \mathbb{P}\Big(\|M - M_0\| > \frac{\varepsilon}{4}\frac{1}{(\|M_0^{-1}\|)^2\|v_0\|}\Big)I\{\|v_0\| \neq 0\}$$

$$\leq \mathbb{P}\Big(\max_{k=0,\ldots,p-1}|\hat{\gamma}_{k;N,T}(t) - \mathbb{E}\hat{\gamma}_{k;N,T}(t)| > \frac{1}{4p}m_f\Big)$$

$$+ \mathbb{P}\Big(\max_{k=1,\ldots,p}|\hat{\gamma}_{k;N,T}(t) - \mathbb{E}\hat{\gamma}_{k;N,T}(t)| > \frac{\varepsilon}{8p^{1/2}}m_f\Big)$$

$$+ \mathbb{P}\Big(\max_{k=0,\ldots,p-1}|\hat{\gamma}_{k;N,T}(t) - \mathbb{E}\hat{\gamma}_{k;N,T}(t)| > \frac{\varepsilon}{32(2\pi)^{1/2}M_f p}m_f^2\Big)$$

$$\leq 3p \max_{k=0,\ldots,p}\mathbb{P}\Big(|\hat{\gamma}_{k;N,T}(t) - \mathbb{E}\hat{\gamma}_{k;N,T}(t)| > \frac{m_f}{4p}\min\Big\{1, \frac{\varepsilon p^{1/2}}{2}, \frac{\varepsilon}{8C_0}\Big\}\Big),$$

$$= 3p \max_{k=0,\ldots,p}\mathbb{P}\Big(|\hat{\gamma}_{k;N,T}(t) - \mathbb{E}\hat{\gamma}_{k;N,T}(t)| > \frac{m_f}{4p}\min\Big\{1, \frac{\varepsilon}{8C_0}\Big\}\Big),$$

where we have use Lemma E.3(ii-c) again to bound $1/\|M_0^{-1}\|$. In the last step we employed that $\frac{p^{1/2}}{2} \geq \frac{1}{8C_0}$. Further, we have used that $M - M_0$ satisfies

$$\|M - M_0\|_1 = \|M - M_0\|_\infty = \max_{1 \leq \ell \leq p} \sum_{h=1}^{p} |\hat{\gamma}_{h-\ell;N,T}(t) - \mathbb{E}\hat{\gamma}_{h-\ell;N,T}(t)|$$

$$\leq p \max_{k=0,\dots,p-1} |\hat{\gamma}_{k;N,T}(t) - \mathbb{E}\hat{\gamma}_{k;N,T}(t)|.$$

Thus, by Hölder's inequality

$$\|M - M_0\| \leq \left(\|M - M_0\|_1 \|M - M_0\|_\infty\right)^{1/2} \leq p \max_{k=0,\dots,p-1} |\hat{\gamma}_{k;N,T}(t) - \mathbb{E}\hat{\gamma}_{k;N,T}(t)|.$$

For the Euclidean norm we have used

$$\|v - v_0\| \leq p^{1/2}\|v - v_0\|_\infty = p^{1/2} \max_{k=1,\dots,p} |\hat{\gamma}_{k;N,T}(t) - \mathbb{E}\hat{\gamma}_{k;N,T}(t)|.$$

Finally, by Corollary E.2(iii) and Lemma E.3(i-b), we have

$$\|v_0\| = \|\mathbb{E}\hat{\gamma}_{N,T}^{(p)}(t/T)\| \leq \|\gamma_{N/T}^{(p)}(t/T)\| + \|\mathbb{E}\hat{\gamma}_{N,T}^{(p)}(t/T) - \gamma_{N/T}^{(p)}(t/T)\|$$
$$\leq (2\pi)^{1/2}M_f + 2T^{-1}p^{3/2}C_1m_f \leq 2(2\pi)^{1/2}M_f,$$

where the second inequality holds for $T \geq 2\frac{p^{3/2}C_1m_f}{(2\pi)^{1/2}M_f} = 2C_1p^{3/2}/C_0$, which is the case, as $T \geq 2C_1p^2$ is assumed. Applying Lemma E.4 yields the assertion, because

$$\mathbb{P}\left(\|\hat{a}_{N,T}^{(p)}(t) - a_{N,T}^{(p)}(t)\| > \varepsilon\right)$$

$$\leq 3p \max_{h=0,\dots,p} \exp\left(-\frac{\eta^2}{2\left(C_{1,1}\frac{h_*}{N-|h|} + \eta^{(3+4d)/(2+2d)}\left(C_{2,1}\frac{h_*}{N-|h|}\right)^{1/(2+2d)}\right)}\right)$$

$$= 3p \exp\left(-\frac{\eta^2}{2\left(C_{1,1}\frac{p}{N-p} + \eta^{(3+4d)/(2+2d)}\left(C_{2,1}\frac{p}{N-p}\right)^{1/(2+2d)}\right)}\right)$$

where $\eta := \frac{m_f}{4p} \min\left\{1, \varepsilon\frac{1}{8C_0}\right\}$, and the third line follows from the fact, for any two integers N and p with $N \geq 1 + p \geq 2$ we have that $(\frac{h_*}{N-|h|})_{h=0,1,\dots,p}$ is an increasing sequence. This is easy to see: $\frac{1}{N-0} \leq \frac{1}{N-1} \leq \dots \leq \frac{p-1}{N-p+1} \leq \frac{p}{N-p}$. Note that $T \geq 2pC_1 \geq C/(\pi m_f)$, such that this condition of Lemma E.4 is met. $\qquad\square$

Proof of Corollary 6.2.

Note the fact that, if $R_n \geq 0$ is a sequence with $R_n \to \infty$, as $n \to \infty$, then

$$b_n = O(1) \Leftrightarrow b_n = o(r_n), \forall 0 \leq r_n \leq R_n, \text{ with } r_n \to \infty, \text{ as } n \to \infty.$$

Thus, employing the Borel-Cantelli lemma, it suffices to show that, for any given $\varepsilon > 0$ and sequence $0 \leq r_T \leq R_T^{1/2}$ with $r_T \to \infty$, we have

$$\sum_{T=1}^{\infty} \mathbb{P}\left(\sup_{p=1,\dots,P} \sup_{t=N,\dots,T} \|\hat{a}_{N,T}^{(p)}(t) - \bar{a}_{N,T}^{(p)}(t)\| > \varepsilon P^{3/2}\left(\frac{\log(T)}{N}\right)^{1/2} r_T\right) < \infty.$$

48

This follows, since we have

$$\mathbb{P}\Big(\sup_{p=1,\ldots,P}\sup_{t=N,\ldots,T}\|\hat{a}^{(p)}_{N,T}(t) - a^{(p)}_{N,T}(t)\| > \varepsilon P^{3/2}\Big(\frac{\log(T)}{N}\Big)^{1/2} r_T\Big)$$

$$\leq P \cdot T \cdot \sup_{p=1,\ldots,P}\sup_{t=N,\ldots,T}\mathbb{P}\Big(\|\hat{a}^{(p)}_{N,T}(t) - a^{(p)}_{N,T}(t)\| > \varepsilon P^{3/2}\Big(\frac{\log(T)}{N}\Big)^{1/2} r_T\Big)$$

$$\leq P \cdot T \cdot \sup_{p=1,\ldots,P}\sup_{t=N,\ldots,T}\mathbb{P}\Big(\|\hat{a}^{(p)}_{N,T}(t) - a^{(p)}_{N,T}(t)\| > \varepsilon p^{3/2}\tilde{C}^{1/2}\Big(\frac{\log(T)}{N-p}\Big)^{1/2} r_T\Big)$$

$$\leq P \cdot T \cdot \sup_{p=1,\ldots,P}\sup_{t=N,\ldots,T} 3p\exp\Big(-\varepsilon^2\frac{p^3\log(T)}{N-p}\tilde{C}r_T^2\frac{m_f^2}{2^{12}C_{1,1}}\Big(C_0^2\frac{p^3}{N-p}\Big)^{-1}\Big)\Big)$$

$$= 3T^3\exp\Big(-\varepsilon^2\log(T)\tilde{C}r_T^2\frac{m_f^2}{2^{12}C_{1,1}}C_0^{-2}\Big)\Big) \leq 3T^{-2},$$

for T large enough. In the second inequality we have used the fact that, due to $P = o(N)$, there exists a $\tilde{C} > 0$ such that $1/N \geq \tilde{C}/(N-P)$, for T large enough. Note that we have $P = o(T^{1/2})$, from $N \leq T$, $P = o(N^{(1+2d)/(4+6d)})$ and $d \geq 1/2$, such that, in the third inequality, Theorem 6.1 can be applied, where we have also used the fact that, under the assumptions made

$$p^{3/2}\Big(\frac{\log(T)}{N-p}\Big)^{1/2} R_T^{1/2} = o\Big(\frac{P^{(4+6d)(3+4d)}}{N^{(1+2d)/(3+4d)}}\Big),$$

implying that, for T large enough, we have

$$\varepsilon p^{3/2}\Big(\frac{\log(T)}{N-p}\Big)^{1/2} r_T \leq \min\{U_{p,N}, 1/(8C_0)\} = U_{p,N}.$$

This completes the proof. $\qquad\square$

Online Supplement: Predictive, finite-sample model choice for time series under stationarity and non-stationarity

Tobias KLEY,[a] Philip PREUß,[b] Piotr FRYZLEWICZ,[a]

[a]London School of Economics and Political Science

[b]Ruhr-Universität Bochum

November 14, 2016

B Proofs of Lemmas A.1, A.2, A.3, (53) and (59)

Proof of Lemma A.1. It is well know that (see, for example, Exercise 2.3.25(i) in Stroock (2011)), if X is a b-sub-Gaussian random variable, then

$$\mathbb{E}|X|^p \leq 2^{p/2+1}\Gamma(p/2+1)b^p,$$

for every $p \in (0, \infty)$. Thus, we have

$$\mathbb{E}|X_{t,T}|^k = \sigma_{t,T}^{k-2}\mathbb{E}\left|\frac{X_{t,T}}{\sigma_{t,T}}\right|^k \sigma_{t,T}^2 \leq (3\pi M_f)^{(k-2)/2}2^{k/2+1}\Gamma(k/2+1)b^k\sigma_{t,T}^2$$

$$\leq (3\pi M_f)^k 2^k (k!)^{1/2}b^k\sigma_{t,T}^2 = (6\pi b M_f)^k (k!)^{1/2}\sigma_{t,T}^2,$$

where we have used (40) for the first inequality and for the second inequality assumed that without loss of generality $3\pi M_f \geq 1$ and we have used that, for $p \in \mathbb{N}^*$, we have

$$\Gamma(p/2+1) = \begin{cases} (p/2)! & \text{if } p \text{ is even,} \\ \sqrt{\pi}\prod_{j=0}^{\lfloor p/2\rfloor}\frac{p-2j}{2} & \text{if } p \text{ is odd,} \end{cases}$$

which implies $\Gamma(p/2+1)^2 \leq p!$, for $p \geq 3$, which finishes the proof, as $\mathbb{E}X_{t,T} = 0$. $\square$

Proof of Lemma A.2. We prove the first equation in full detail and comment on how the proof needs to be adapted for the second inequality in the end. First, we consider the case when $p \geq 1$. In the end of the proof we will comment on the (easier) case $p = 0$. Denote $\hat{v}_{N,T}^{(p,h)}(t) =: (\hat{v}_{1,t}, \ldots, \hat{v}_{p,t})'$ and note that, omitting the second index of $X_{t,T}$ for the sake of brevity, we have

$$
\begin{aligned}
\text{MSPE}_{s,n,N,T}^{(p,h)} \\
= n^{-1} \sum_{t=s+1}^{s+n} \left(X_{t+h}^2 - 2X_{t+h} \sum_{i=1}^{p} \hat{v}_{i,t} X_{t-i+1} + \sum_{i_1,i_2=1}^{p} \hat{v}_{i_1,t} \hat{v}_{i_2,t} X_{t-i_1+1} X_{t-i_2+1} \right)
\end{aligned}
$$

and an analogous equation for $\overline{\text{MSPE}}_{s,n,N,T}^{(p,h)}$, with $\bar{v}_{N,T}^{(p,h)}(t) =: (\bar{v}_{1,t}, \ldots, \bar{v}_{p,t})'$.

With this notation we have

$$
\begin{aligned}
&\mathbb{P}\left(\left| \text{MSPE}_{s,n,N,T}^{(p,h)} - \overline{\text{MSPE}}_{s,n,N,T}^{(p,h)} \right| > \varepsilon \right) \\
&= \mathbb{P}\left(\left| n^{-1} \sum_{t=s+1}^{s+n} \left(X_{t+h}^2 - \mathbb{E}(X_{t+h}^2) \right) \right. \right. \\
&\quad - 2 \sum_{i=1}^{p} \left(n^{-1} \sum_{t=s+1}^{s+n} \left(\hat{v}_{i,t} X_{t+h} X_{t-i+1} - \bar{v}_{i,t} \mathbb{E}(X_{t+h} X_{t-i+1}) \right) \right) \\
&\quad \left. \left. + \sum_{i_1,i_2=1}^{p} n^{-1} \sum_{t=s+1}^{s+n} \left(\hat{v}_{i_1,t} \hat{v}_{i_2,t} X_{t-i_1+1} X_{t-i_2+1} - \bar{v}_{i_1,t} \bar{v}_{i_2,t} \mathbb{E}(X_{t-i_1+1} X_{t-i_2+1}) \right) \right| > \varepsilon \right),
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
&\mathbb{P}\left(\left| \text{MSPE}_{s,n,N,T}^{(p,h)} - \overline{\text{MSPE}}_{s,n,N,T}^{(p,h)} \right| > \varepsilon \right) \\
&\leq \mathbb{P}\left(\left| n^{-1} \sum_{t=s+1}^{s+n} \left(X_{t+h}^2 - \mathbb{E}(X_{t+h}^2) \right) \right| > \varepsilon/(p+1)^2 \right) \tag{61}
\end{aligned}
$$

$$
+ 2 \sum_{i=1}^{p} \mathbb{P}\left(\left| n^{-1} \sum_{t=s+1}^{s+n} \left(\hat{v}_{i,t} X_{t+h} X_{t-i+1} - \bar{v}_{i,t} \mathbb{E}(X_{t+h} X_{t-i+1}) \right) \right| > \varepsilon/(p+1)^2 \right) \tag{62}
$$

$$
\begin{aligned}
+ \sum_{i_1,i_2=1}^{p} \mathbb{P}\left(\left| n^{-1} \sum_{t=s+1}^{s+n} \left(\hat{v}_{i1,t} \hat{v}_{i2,t} X_{t-i_1+1} X_{t-i_2+1} \right. \right. \right. \\
\left. \left. \left. - \bar{v}_{i1,t} \bar{v}_{i2,t} \mathbb{E}(X_{t-i_1+1} X_{t-i_2+1}) \right) \right| > \varepsilon/(p+1)^2 \right). \tag{63}
\end{aligned}
$$

We now use Lemma F.5 to disentangle (62) and (63). More precisely, for the ith addend

$(i = 1, \ldots, p)$ in (62) we have

$$\mathbb{P}\left(\left|n^{-1}\sum_{t=s+1}^{s+n}\left(\hat{v}_{i,t}X_{t+h}X_{t-i+1} - \bar{v}_{i,t}\mathbb{E}(X_{t+h}X_{t-i+1})\right)\right| > \varepsilon/(p+1)^2\right)$$

$$\leq n \sup_{t=s+1,\ldots,s+n}\mathbb{P}\left(|\hat{v}_{i,t} - \bar{v}_{i,t}| > \frac{\varepsilon/(p+1)^2}{2\left((6\pi M_f c^2 24^d)^2 + \varepsilon^2/(p+1)^4\right)^{1/4}}\right) \tag{64}$$

$$+ \mathbb{P}\left(\left|\sum_{t=s+1}^{s+n}(X_{t+h}^2 X_{t-i+1}^2 - \mathbb{E}(X_{t+h}^2 X_{t-i+1}^2))\right| > \frac{n\varepsilon}{2(p+1)^2}\right) \tag{65}$$

$$+ \mathbb{P}\left(\left|\sum_{t=s+1}^{s+n}(\bar{v}_{i,t})(X_{t+h}X_{t-i+1} - \mathbb{E}(X_{t+h}X_{t-i+1}))\right| > \frac{n\varepsilon}{2(p+1)^2}\right), \tag{66}$$

where we have used the Cauchy-Schwarz inequality, Assumption 5 and (40), which holds as $T \geq \frac{C}{\pi m_f}$, to (uniformly) bound the second moments

$$\mathbb{E}(X_{t+h}^2 X_{t-i+1}^2) \leq \left(\mathbb{E}(X_{t+h}^4)\mathbb{E}(X_{t-i+1}^4)\right)^{1/2} \leq \left(c^4(4!)^{2d}\sigma_{t+h,T}^2\sigma_{t-i+1,T}^2)\right)^{1/2} \leq 3\pi M_f c^2 24^d$$

Note that we have used $m_2^2 := 3\pi M_f c^2 24^d$ in the application of Lemma F.5. For the (i_1, i_2)th addend $(i_1, i_2 = 1, \ldots, p)$ in (63) we analogously have

$$\mathbb{P}\left(\left|n^{-1}\sum_{t=s+1}^{s+n}\left(\hat{v}_{i1,t}\hat{v}_{i2,t}X_{t-i_1+1}X_{t-i_2+1} - \bar{v}_{i1,t}\bar{v}_{i2,t}\mathbb{E}(X_{t-i_1+1}X_{t-i_2+1})\right)\right| > \varepsilon/(p+1)^2\right)$$

$$\leq n \sup_{t=s+1,\ldots,s+n}\mathbb{P}\left(|\hat{v}_{i1,t}\hat{v}_{i2,t} - \bar{v}_{i1,t}\bar{v}_{i2,t}| > \frac{\varepsilon/(p+1)^2}{2\left((6\pi M_f c^2 24^d)^2 + \varepsilon^2/(p+1)^4\right)^{1/4}}\right) \tag{67}$$

$$+ \mathbb{P}\left(\left|\sum_{t=s+1}^{s+n}(X_{t-i_1+1}^2 X_{t-i_2+1}^2 - \mathbb{E}(X_{t-i_1+1}^2 X_{t-i_2+1}^2))\right| > \frac{n\varepsilon}{2(p+1)^2}\right) \tag{68}$$

$$+ \mathbb{P}\left(\left|\sum_{t=s+1}^{s+n}\left(\bar{v}_{i1,t}\bar{v}_{i2,t}\right)(X_{t-i_1+1}X_{t-i_2+1} - \mathbb{E}(X_{t-i_1+1}X_{t-i_2+1}))\right| > \frac{n\varepsilon}{2(p+1)^2}\right) \tag{69}$$

where we bounded $\sup_{t=s,\ldots,s+n-1}\mathbb{E}(X_{t-i_1+1}^2 X_{t-i_2+1}^2)$ by m_2^2 as before.

We now use Lemma F.4 to bound (67) by a sum of two quantities resembling those from (64). Applying Lemmas F.4 and C.3(i), for the second we have required that $T \geq 5C_1p^2$, yields that (67) can be bounded by

$$2n\sup_{i\in\{i_1,i_2\}}\sup_{t=s+1,\ldots,s+n}\mathbb{P}\left(|\hat{v}_{i,t} - \bar{v}_{i,t}| > \mu\right). \tag{70}$$

In conclusion, we have shown

$$\mathbb{P}\left(\left|\mathrm{MSPE}_{s,n,N,T}^{(p,h)} - \overline{\mathrm{MSPE}}_{s,n,N,T}^{(p,h)}\right| > \varepsilon\right)$$

$$\leq (61) + 2p\sup_{i=1,\ldots,p}\left((65)_i + (66)_i\right) + p^2\sup_{i_1,i_2=1,\ldots,p}\left((68)_{i_1,i_2} + (69)_{i_1,i_2}\right) \tag{71}$$

$$+ 2p\sup_{i=1,\ldots,p}(64)_i + p^2\sup_{i_1,i_2=1,\ldots,p}(70)_{i_1,i_2}, \tag{72}$$

	b	a_t	α	h	ε
(61)	$s + h + 1$	1	2	0	$\varepsilon/(p+1)^2$
(65)	$s - i + 2$	1	2	$h + i - 1$	$\frac{1}{2}\varepsilon/(p+1)^2$
(66)	$s - i + 2$	$\bar{v}_{i,t}$	1	$h + i - 1$	$\frac{1}{2}\varepsilon/(p+1)^2$
(68)	$s - \max\{i_1, i_2\} + 2$	1	2	$|i_1 - i_2|$	$\frac{1}{2}\varepsilon/(p+1)^2$
(69)	$s - \max\{i_1, i_2\} + 2$	$\bar{v}_{i_1,t}\bar{v}_{i_2,t}$	1	$|i_1 - i_2|$	$\frac{1}{2}\varepsilon/(p+1)^2$

Table 8: Parameters for the application of Lemma E.5.

where the subscript at the equation numbers indicate that the respective expressions depend on i or i_1, i_2. We now discuss how to bound (71) and (72).

To bound (71) we note that (61), (65), (66), (68), and (69) are all of the type to which Lemma E.5 can be applied. We do so with the quantities b, a_t, α, h, ε of Lemma E.5 chosen as in Table 8.

Note that for the discussion of (66) and (69), we have, by Lemma C.3(i) that

$$|a_{t,(69)}| \leq (2C_0 + 1)^h =: A_{(66)}, \text{ and } |a_{t,(69)}| \leq (2C_0 + 1)^{2h} =: A_{(69)},$$

for (66) and (69), respectively. (We wrote $a_{t,(\#)}$ for the sequence of equation (#).)

We now enlarge the bound that we get from Lemma E.5 by enlarging the constants $C_{1,\alpha}$ and $C_{2,\alpha}$ to

$$\begin{aligned}
\bar{C}_{1,\alpha} &:= 12 \cdot 2^{8\alpha d + 9}(\alpha!)^{4d}\big(\max\{1, c^2, 3\pi M_f\}\big)^{2\alpha}\mathrm{e}\Big(1 + \frac{1}{\log \rho}\Big)\Big(1 + K^{1/2}\Big), \\
\bar{C}_{2,\alpha} &:= 12 \cdot 2^{3(\alpha d + 1)}(\alpha!)^{2d}\big(\max\{1, c^2, 3\pi M_f\}\big)^{\alpha}\mathrm{e}\Big(1 + \frac{1}{\log \rho}\Big).
\end{aligned} \tag{73}$$

Note that (i) the bound from Lemma E.5 is independent of b, that (ii) increasing A, α or h_* will increase the bound and (iii) decreasing ε will also increase the bound. Note that comment (ii), with respect to α, holds, because we have enlarged the constants $C_{1,\alpha}$ and $C_{2,\alpha}$ to $\bar{C}_{1,\alpha}$ and $\bar{C}_{2,\alpha}$.

Thus, (71) can be bounded by

$$(1 + 4p + 2p^2)$$

$$\cdot \exp\left(-\frac{\frac{\varepsilon^2}{(p+1)^4}}{8\Big((2C_0 + 1)^{4h}\frac{\bar{C}_{1,2}(h+p-1)}{n} + (\frac{\varepsilon}{2(p+1)^2})^{(3+8d)/(2+4d)}\Big((2C_0 + 1)^{2h}\frac{\bar{C}_{2,2}(h+p-1)}{n}\Big)^{1/(2+4d)}\Big)}\right).$$

It remains to bound (72). Note that

$$(72) \leq (2pn + 2p^2 n) \sup_{i=1,\ldots,p} \sup_{t=s+1,\ldots,s+n} \mathbb{P}\Big(|\hat{v}_{i,t} - \bar{v}_{i,t}| > \mu\Big), \tag{74}$$

where we have used the fact that $\mu \leq \bar{\varepsilon}$, because of $(2C_0 + 1)^{2h} \geq 1$.

Thus, to bound (72) we can employ Lemma F.7, which yields that

$$
\begin{aligned}
\mathbb{P}\Big(|\hat{v}_{i,t} - \bar{v}_{i,t}| > \mu\Big) &\leq \mathbb{P}\Big(\|\hat{v}_{N,T}^{(p,h)}(t) - \bar{v}_{N,T}^{(p,h)}(t)\| > \mu\Big) \\
&\leq \mathbb{P}\Big(\|\hat{a}_{N,T}^{(p)}(t) - \bar{a}_{N,T}^{(p)}(t)\| > 2^{1-h}\frac{\mu}{\mu + h(2C_0)^{h-1}}\Big)
\end{aligned}
\tag{75}
$$

where we have used that $\max\{\|\bar{a}_{N,T}^{(p)}(t)\|, 1\} \leq \max\{2C_0, 1\} = 2C_0$ by Lemma C.3(i) and the fact that $C_0 \geq (2\pi)^{1/2}$.

Thus, employing (74), (75) and Theorem 6.1 (note: $5C_1 p^2 \geq 2C_1 p^2$), we have that (72) is bounded by

$$
\begin{aligned}
&2np(p+1)\mathbb{P}\Big(\|\hat{a}_{i,N,T}^{(p)}(t) - \bar{a}_{i,N,T}^{(p)}(t)\| > 2^{1-h}\frac{\mu}{\mu + h(2C_0)^{h-1}}\Big) \\
&\leq 6np^2(p+1)\exp\Big(-\frac{\eta^2}{2\big(\bar{C}_{1,2}\frac{p}{N-p} + \eta^{(3+4d)/(2+2d)}\big(\bar{C}_{2,2}\frac{p}{N-p}\big)^{1/(2+2d)}\big)}\Big).
\end{aligned}
\tag{76}
$$

This completes the proof of the first inequality of Lemma A.2. For the second inequality denote $\hat{v}_{t,T}^{(p,h)}(t) =: (\hat{v}_{1,t}, \ldots, \hat{v}_{p,t})'$ and apply the same stream of arguments, but set $N = t$ and instead of (76) use

$$
\begin{aligned}
&2np(p+1)\mathbb{P}\Big(\|\hat{a}_{i,t,T}^{(p)}(t) - \bar{a}_{i,t,T}^{(p)}(t)\| > 2^{1-h}\frac{\mu}{\mu + h(2C_0)^{h-1}}\Big) \\
&\leq 6np^2(p+1)\exp\Big(-\frac{\eta^2}{2\big(\bar{C}_{1,2}\frac{p}{t-p} + \eta^{(3+4d)/(2+2d)}\big(\bar{C}_{2,2}\frac{p}{t-p}\big)^{1/(2+2d)}\big)}\Big) \\
&\leq 6np^2(p+1)\exp\Big(-\frac{\eta^2}{2\big(\bar{C}_{1,2}\frac{p}{s-p} + \eta^{(3+4d)/(2+2d)}\big(\bar{C}_{2,2}\frac{p}{s-p}\big)^{1/(2+2d)}\big)}\Big),
\end{aligned}
$$

as $s \leq t$. The second inequality follows.

Lastly, for the case where $p = 0$, note that it suffices to bound (61).

$\square$

Proof of Lemma A.3. Note the two representations

$$
\begin{aligned}
\overline{\mathrm{MSPE}}_{s,n,N,T}^{(p,h)} &= \frac{1}{n}\sum_{t=s+1}^{s+n}\mathbb{E}\Big(X_{t+h,T} - \sum_{i=1}^{p}\bar{v}_{i;N,T}^{(p,h)}(t)X_{t-i+1,T}\Big)^2 \\
&= \frac{1}{n}\sum_{t=s+1}^{s+n}\Big(g_{h,T}^{(1)}(t) + g_{p,h,N,T}^{(2)}(t) + g_{p,h,N,T}^{(3)}(t)\Big)
\end{aligned}
$$

where

$$
\begin{aligned}
g_{h,T}^{(1)}(t) &:= \mathbb{E}\big(X_{t+h,T}^2\big) \\
g_{p,h,N,T}^{(2)}(t) &:= -2\sum_{i=1}^{p}\bar{v}_{i;N,T}^{(p,h)}(t)\mathbb{E}\big(X_{t-i+1,T}X_{t+h,T}\big) \\
g_{p,h,N,T}^{(3)}(t) &:= \sum_{i_1=1}^{p}\sum_{i_2=1}^{p}\bar{v}_{i_1;N,T}^{(p,h)}(t)\bar{v}_{i_2;N,T}^{(p,h)}(t)\mathbb{E}\big(X_{t-i_1+1,T}X_{t-i_2+1,T}\big),
\end{aligned}
$$

54

and

$$\mathrm{MSPE}_{N/T,n/T}^{(p,h)}(s/T) = \int_0^1 g_{N/T}^{(p,h)}\Big(\frac{s}{T} + \frac{n}{T}x\Big)\mathrm{d}x$$

$$= \frac{1}{n}\sum_{t=s+1}^{s+n}\Big(g_h^{(1)}(t/T) + g_{p,h;N/T}^{(2)}(t/T) + g_{p,h;N/T}^{(3)}(t/T)\Big) + R_{p,h,N,T}(t)$$

where $g_\Delta^{(p,h)}(u) := g_h^{(1)}(u) + g_{p,h;\Delta}^{(2)}(u) + g_{p,h;\Delta}^{(3)}(u)$, and

$$g_h^{(1)}(u) := \gamma_0(u)$$

$$g_{p,h;\Delta}^{(2)}(u) := -2\sum_{i=1}^p v_{i;\Delta}^{(p,h)}(u)\gamma_{i+h-1}(u)$$

$$= -2\big(v_\Delta^{(p,h)}(u)\big)'\gamma^{(p,h)}(u)$$

$$g_{p,h;\Delta}^{(3)}(u) := \sum_{i_1=1}^p\sum_{i_2=1}^p v_{i_1;\Delta}^{(p,h)}(u)v_{i_2;\Delta}^{(p,h)}(u)\gamma_{i_2-i_1}(u)$$

$$= \big(v_\Delta^{(p,h)}(u)\big)'\Gamma^{(p)}(u)v_\Delta^{(p,h)}(u).$$

Note that we have

$$\overline{\mathrm{MSPE}}_{s,n,N,T}^{(p,h)} - \mathrm{MSPE}_{N/T,n/T}^{(p,h)}(s/T)$$

$$= R_{p,h,N,T}(t) + \frac{1}{n}\sum_{t=s+1}^{s+n}\Big(r_{h,T}^{(1)}(t) + r_{p,h,N,T}^{(2a)}(t) + r_{p,h,N,T}^{(3a)}(t)\Big), \tag{77}$$

and

$$\overline{\mathrm{MSPE}}_{s,n,0,T}^{(p,h)} - \mathrm{MSPE}_{s/T,n/T}^{(p,h)}(s/T)$$

$$= R_{p,h,N,T}(t) + \frac{1}{n}\sum_{t=s+1}^{s+n}\Big(r_{h,T}^{(1)}(t) + r_{p,h,t,T}^{(2)}(t) + r_{p,h,t,T}^{(3)}(t)\Big),$$

with

$$r_{h,T}^{(1)}(t) := g_{h,T}^{(1)}(t) - g_h^{(1)}(t/T),$$

$$r_{p,h,N,T}^{(2)}(t) := g_{p,h,N,T}^{(2)}(t) - g_{p,h;N/T}^{(2)}(t/T),$$

$$r_{p,h,N,T}^{(3)}(t) := g_{p,h,N,T}^{(3)}(t) - g_{p,h;N/T}^{(3)}(t/T),$$

Because, by Lemma F.8 and Lemma C.4(ii) we have

$$|R_{p,h,N,T}(t)| \le \frac{1}{T}\sup_{\frac{s}{T}<u<\frac{s+n}{T}}|\frac{\partial}{\partial u}g_{N/T}^{(p,h)}(u)| \le \frac{4}{T}(2h+1)\big(C_0\big)^{2h+1}M_f'$$

it now suffices to prove bounds (uniform with respect to t) for $r_{h,T}^{(1)}(t)$, $r_{p,h,N,T}^{(2)}(t)$, and $r_{p,h,N,T}^{(3)}(t)$. For $r_{h,T}^{(1)}(t)$ we have

$$g_{h,T}^{(1)}(t) := \mathbb{E}\big(X_{t+h,T}^2\big) = \tilde{\gamma}_{0,T}(t+h)$$

$$= \gamma_0(\frac{t}{T}) + \gamma_0(\frac{t+h}{T}) - \gamma_0(\frac{t}{T}) + \tilde{\gamma}_{0,T}(t+h) - \gamma_0(\frac{t+h}{T})$$

$$=: g_h^{(1)}(t/T) + r_{h,T}^{(1)}(t),$$

55

where $|r_{h,T}^{(1)}(t)| \leq (2\pi M_f' h + C)/T$, by (41) and (37).

For $r_{p,h,N,T}^{(2)}(t)$ we have

$$|r_{p,h,N,T}^{(2)}(t)| := \Big| -2\sum_{i=1}^{p} \Big(\bar{v}_{i;N,T}^{(p,h)}(t)\mathbb{E}\big(X_{t-i+1,T}X_{t+h,T}\big)$$

$$- v_{i;N/T}^{(p,h)}(t/T)\gamma_{i+h-1}(\frac{t}{T})\Big)\Big|$$

$$= 2\Big|\sum_{i=1}^{p} \Big(\big(\bar{v}_{i;N,T}^{(p,h)}(t) - v_{i;N/T}^{(p,h)}(t/T)\big)\gamma_{i+h-1}(\frac{t}{T})\Big)$$

$$+ \sum_{i=1}^{p} \Big(\bar{v}_{i;N,T}^{(p,h)}(t)\Big(\mathbb{E}\big(X_{t-i+1,T}X_{t+h,T}\big) - \gamma_{i+h-1}(\frac{t+h}{T})$$

$$+ \gamma_{i+h-1}(\frac{t+h}{T}) - \gamma_{i+h-1}(\frac{t}{T})\Big)\Big)\Big)\Big|$$

$$\leq 2\Big(\|\bar{v}_{N,T}^{(p,h)}(t) - v_{N/T}^{(p,h)}(t/T)\| \cdot \|\gamma^{(p,h)}(\frac{t}{T})\| + p\|\bar{v}_{N,T}^{(p,h)}(t)\|_\infty(C + 2\pi h M_f')/T\Big)$$

$$\leq 2\Big(3h\big(2C_0\big)^h C_1\frac{p^2}{T} \cdot (2\pi)^{1/2}M_f + p2\big(C_0\big)^h \cdot (C + 2\pi M_f')h/T\Big)$$

$$\leq 2\big(C_0\big)^h\frac{1}{T}\Big(h2^h C_1\big(3p^2\big)(2\pi)^{1/2}M_f + 2hp(C + 2\pi M_f')\Big)$$

$$\leq 2\big(C_0\big)^h\frac{1}{T}\Big(h2^h\big(3p^2\big)C_0(2\pi M_f' + C) + 2p^2 C_0(2\pi M_f' + C)\Big)$$

$$\leq 2\big(C_0\big)^{h+1}(3h2^h + 2h)\frac{1}{T}p^2(2\pi M_f' + C) \leq 5(2\pi M_f' + C)h\big(2C_0\big)^{h+1}\frac{1}{T}p^2$$

where the second inequality holds by employing Lemmas C.1(ii) and C.3, both of which can be applied, since we assumed $T \geq 3h2^h C_1 p^2$, and $3h2^h C_1 p^2 \geq 5C_1 p$, as $h \geq 1$. Also note that $\|\gamma^{(p,h)}(\frac{t+h}{T})\| \leq \|\gamma_0^{(p+h)}(\frac{t+h}{T})\|$ to which we apply Lemma E.3(i-b). We also employed the bound on $\gamma_{i+h-1}'(u)$ from below (42) and used the mean value theorem.

For $r_{p,h,N,T}^{(3)}(t)$, employing $abc - a_0 b_0 c_0 = b_0 c(a - a_0) + ac(b - b_0) + a_0 b_0(c - c_0)$, we have

$$|r_{p,h,N,T}^{(3)}(t)| := \Big| \sum_{i_1=1}^{p} \sum_{i_2=1}^{p} \Big(\bar{v}_{i_2;N,T}^{(p,h)}(t) \gamma_{i_2-i_1}(\frac{t}{T}) \big(v_{i_1;N/T}^{(p,h)}(t/T) - \bar{v}_{i_1;N,T}^{(p,h)}(t) \big)$$

$$+ v_{i_1;N/T}^{(p,h)}(t/T) \gamma_{i_2-i_1}(\frac{t}{T}) \big(v_{i_2;N/T}^{(p,h)}(t) - \bar{v}_{i_2;N,T}^{(p,h)}(t/T) \big)$$

$$+ \bar{v}_{i_1;N,T}^{(p,h)}(t) \bar{v}_{i_2;N,T}^{(p,h)}(t) \big(\gamma_{i_2-i_1}(\frac{t}{T}) - \mathbb{E}\big(X_{t-i_1+1,T} X_{t-i_2+1,T}\big) \big) \Big) \Big|$$

$$\leq \Big| (\bar{v}_{N,T}^{(p,h)}(t))' \Gamma_0^{(p)} \big(v_{N/T}^{(p,h)}(t/T) - \bar{v}_{N,T}^{(p,h)}(t) \big)$$

$$+ (v_{N/T}^{(p,h)}(t/T))' \Gamma_0^{(p)}(t/T) \big(v_{N/T}^{(p,h)}(t/T) - \bar{v}_{N,T}^{(p,h)}(t) \big) \Big|$$

$$+ \|\bar{v}_{N,T}^{(p,h)}(t)\|_1^2 (2\pi M_f'(p-1) + C)/T$$

$$\leq \big(\|\bar{v}_{N,T}^{(p,h)}(t)\| + \|v_{N/T}^{(p,h)}(t/T)\| \big) \|\Gamma_0^{(p)}(t/T)\| \cdot \|v_{N/T}^{(p,h)}(t/T) - \bar{v}_{N,T}^{(p,h)}(t)\|$$

$$+ p \|\bar{v}_{N,T}^{(p,h)}(t)\|^2 (2\pi M_f'(p-1) + C)/T$$

$$\leq \big(2(C_0)^h + (C_0)^h \big) M_f \cdot 3h(2C_0)^h C_1 \frac{p^2}{T} + p \big(2(C_0)^h \big)^2 (2\pi M_f'(p-1) + C)/T$$

$$= T^{-1}(C_0)^{2h} \big(9h2^h M_f C_1 p^2 + 4p(2\pi M_f'(p-1) + C) \big)$$

$$\leq T^{-1}(C_0)^{2h} \big(9h2^h M_f C_1 p^2 + 4p^2(2\pi M_f' + C) \big)$$

$$= T^{-1}(C_0)^{2h} p^2 (2\pi M_f' + C) \big(9(2\pi)^{-1/2} h2^h C_0 + 4 \big) \leq 8(2\pi M_f' + C) h2^h (C_0)^{2h+1} T^{-1} p^2$$

where the last inequality is due to Lemma C.3(ii), C.2(i), E.3(i-a), C.1(ii), and C.3(ii) to the respective terms. Note that Lemma C.3(ii) can be applied as we have assumed $T \geq 3h2^h C_1 p^2$, which implies the condition for Lemma C.1(ii), as $h \geq 1$.

For the first inequality, we have used the fact that

$$|\gamma_{i_2-i_1}(\frac{t}{T}) - \mathbb{E}\big(X_{t-i_1+1,T} X_{t-i_2+1,T}\big)|$$

$$\leq |\gamma_{i_2-i_1}(\frac{t}{T}) - \gamma_{i_2-i_1}(\frac{t-i_1+1}{T})| + |\gamma_{i_2-i_1}(\frac{t-i_1+1}{T}) - \mathbb{E}\big(X_{t-i_1+1,T} X_{t-i_2+1,T}\big)|$$

For the very last inequality we used $9/(2\pi)^{1/2} + 4 \leq 8$.

Substituting these four results into (77), we have

$$\Big| \overline{\text{MSPE}}_{s,n,N,T}^{(p,h)} - \text{MSPE}_{N/T,n/T}^{(p,h)}(s/T) \Big|$$

$$\leq \frac{4}{T}(2h+1)(C_0)^{2h+1} M_f' + (2\pi M_f' h + C)/T$$

$$+ 5(2\pi M_f' + C)h(2C_0)^{h+1} \frac{1}{T} p^2 + 8(2\pi M_f' + C)h2^h(C_0)^{2h+1} T^{-1} p^2$$

$$\leq 4 \cdot 12(2\pi M_f' + C)h2^h(C_0)^{2h+1} \frac{p^2}{T},$$

which completes this proof. $\qquad \square$

Proof of (53). With the notation from the proof of Theorem 3.1, we have

$$\mathbb{P}\Big((\hat{R}_{T,2}^{\mathrm{s,ls}}(h) \geq 1+\delta \text{ and } \hat{R}_{T,3}^{\mathrm{s,ls}}(h) \geq 1+\delta)$$

$$\text{or } (\hat{R}_{T,2}^{\mathrm{s,ls}}(h) < 1+\delta \text{ and } \hat{R}_{T,3}^{\mathrm{s,ls}}(h) < 1+\delta)\Big)$$

$$= 1 - \mathbb{P}\Big(\bigcup_{p_1,p_2=0}^{p_{\max}} \bigcup_{N \in \mathcal{N}} \{(\hat{R}_{T,2}^{\mathrm{s,ls}}(h) < 1+\delta \text{ or } \hat{R}_{T,3}^{\mathrm{s,ls}}(h) < 1+\delta)$$

$$\text{and } (\hat{R}_{T,2}^{\mathrm{s,ls}}(h) \geq 1+\delta \text{ or } \hat{R}_{T,3}^{\mathrm{s,ls}}(h) \geq 1+\delta),\ \hat{p}_{\mathrm{ls}} = p_1,\ \hat{N}_{\mathrm{ls}} = N,\ \hat{p}_{\mathrm{s}} = p_2\}\Big)$$

$$\geq 1 - \sum_{p_1,p_2=0}^{p_{\max}} \sum_{N \in \mathcal{N}} \mathbb{P}\Big((Y_1 < X_1(1+\delta) \text{ or } Y_2 < X_2(1+\delta))$$

$$\text{and } \big(Y_1 \geq X_1(1+\delta) \text{ or } Y_2 \geq X_2(1+\delta)\big)\Big)$$

$$= 1 - \sum_{p_1,p_2=0}^{p_{\max}} \sum_{N \in \mathcal{N}} \mathbb{P}\Big((Y_1 < X_1(1+\delta) \text{ or } Y_1 < X_1(1+\delta) + Y_1 - Y_2 + (X_2 - X_1)(1+\delta))$$

$$\text{and } \big(Y_1 \geq X_1(1+\delta) \text{ or } Y_1 \geq X_1(1+\delta) + Y_1 - Y_2 + (X_2 - X_1)(1+\delta)\big)\Big)$$

$$= 1 - \sum_{p_1,p_2=0}^{p_{\max}} \sum_{N \in \mathcal{N}} \mathbb{P}\Big((A < 0 \text{ or } A < B) \text{ and } (A \geq 0 \text{ or } A \geq B)\Big)$$

$$= 1 - \sum_{p_1,p_2=0}^{p_{\max}} \sum_{N \in \mathcal{N}} \mathbb{P}\Big(A < \max\{0,B\},\ A \geq \min\{0,B\}\Big)$$

$$= 1 - \sum_{p_1,p_2=0}^{p_{\max}} \sum_{N \in \mathcal{N}} \Big(\mathbb{P}\Big(A < \max\{0,B\},\ A \geq \min\{0,B\},\ |B| > \varepsilon\Big)$$

$$+ \mathbb{P}\Big(A < \max\{0,B\},\ A \geq \min\{0,B\},\ -\varepsilon \leq B < 0\Big)$$

$$+ \mathbb{P}\Big(A < \max\{0,B\},\ A \geq \min\{0,B\},\ 0 \leq B \leq \varepsilon\Big)\Big)$$

$$= 1 - \sum_{p_1,p_2=0}^{p_{\max}} \sum_{N \in \mathcal{N}} \Big(\mathbb{P}\Big(A < \max\{0,B\},\ A \geq \min\{0,B\},\ |B| > \varepsilon\Big)$$

$$+ \mathbb{P}\Big(A < 0,\ A \geq B,\ -\varepsilon \leq B < 0\Big)$$

$$+ \mathbb{P}\Big(A < B,\ A \geq 0,\ 0 \leq B \leq \varepsilon\Big)\Big)$$

$$\geq 1 - \sum_{p_1,p_2=0}^{p_{\max}} \sum_{N \in \mathcal{N}} \Big(\mathbb{P}\Big(|B| > \varepsilon\Big) + \mathbb{P}\Big(-\varepsilon \leq A < 0\Big) + \mathbb{P}\Big(0 \leq A < \varepsilon\Big)\Big)$$

$$\geq 1 - \sum_{p_1,p_2=0}^{p_{\max}} \sum_{N \in \mathcal{N}} \Big(\mathbb{P}\Big(|B| > \varepsilon\Big) + \mathbb{P}\Big(|A| \leq \varepsilon\Big)\Big).$$

Note that in (53) we use $\varepsilon = f(\delta)/2$. $\qquad\square$

Proof of (59). With the notation from the proof of Corollary 3.3, we have

$$
\begin{aligned}
(58) &= \Big(-\delta\gamma_0 - 2\gamma_1\big(F - (1+\delta)E\big) + \gamma_0\big(F^2 - (1+\delta)E^2\big)\Big) \\
&= -\gamma_0\Big(\delta + 2\frac{\gamma_1}{\gamma_0}\big(F - (1+\delta)E\big) - \big(F^2 - (1+\delta)E^2\big)\Big) \\
&= -\gamma_0\Big(\delta - 2\delta E\Big(\frac{\gamma_1}{\gamma_0} - E/2\Big) + 2\frac{\gamma_1}{\gamma_0}(F - E) - \big(F^2 - E^2\big)\Big) \\
&= -\gamma_0\Big(\delta\Big(1 - 2E\Big(\frac{\gamma_1}{\gamma_0} - E/2\Big)\Big) + 2\frac{\gamma_1}{\gamma_0}(F - E) - (F + E)(F - E)\Big) \\
&= -\gamma_0\Big(\delta\Big(1 - 2E\Big(\frac{\gamma_1}{\gamma_0} - E/2\Big)\Big) + \Big(2\frac{\gamma_1}{\gamma_0} - (F + E)\Big)(F - E)\Big) \\
&= -\gamma_0\Big(\delta\Big(1 - 2E\Big(\frac{\gamma_1}{\gamma_0} - E/2\Big)\Big) + \Big(\frac{\gamma_1}{\gamma_0} - E - \Big(F - \frac{\gamma_1}{\gamma_0}\Big)\Big)\Big(F - \frac{\gamma_1}{\gamma_0} + \frac{\gamma_1}{\gamma_0} - E\Big)\Big) \\
&= -\gamma_0\Big(\delta\Big(1 - E^2 - 2E\Big(\frac{\gamma_1}{\gamma_0} - E\Big)\Big) + \Big(\frac{\gamma_1}{\gamma_0} - E\Big)^2 - \Big(F - \frac{\gamma_1}{\gamma_0}\Big)^2\Big) \\
&= \gamma_0\Big(\Big(F - \frac{\gamma_1}{\gamma_0}\Big)^2 - \Big(\frac{\gamma_1}{\gamma_0} - E\Big)^2\Big) - \delta\Big(1 - E^2 - 2E\Big(\frac{\gamma_1}{\gamma_0} - E\Big)\Big)\Big),
\end{aligned}
$$

where in the last step the (reverse) triangle inequality was used. The inequality then follows, as

$$
\begin{aligned}
&1 - E^2 - 2E\Big(\frac{\gamma_1}{\gamma_0} - E\Big) \\
&= 1 - \Big[\Big(E - \frac{\gamma_1}{\gamma_0}\Big)^2 + 2\frac{\gamma_1}{\gamma_0}\Big(E - \frac{\gamma_1}{\gamma_0}\Big) + \Big(\frac{\gamma_1}{\gamma_0}\Big)^2\Big] - 2E\Big(\frac{\gamma_1}{\gamma_0} - E\Big) \\
&= 1 - \Big(E - \frac{\gamma_1}{\gamma_0}\Big)^2 + 2\frac{\gamma_1}{\gamma_0}\Big(\frac{\gamma_1}{\gamma_0} - E\Big) - \Big(\frac{\gamma_1}{\gamma_0}\Big)^2 - 2E\Big(\frac{\gamma_1}{\gamma_0} - E\Big) \\
&= 1 - \Big(E - \frac{\gamma_1}{\gamma_0}\Big)^2 + 2\Big(\frac{\gamma_1}{\gamma_0} - E\Big)^2 - \Big(\frac{\gamma_1}{\gamma_0}\Big)^2 = 1 + \Big(E - \frac{\gamma_1}{\gamma_0}\Big)^2 - \Big(\frac{\gamma_1}{\gamma_0}\Big)^2. \qquad \square
\end{aligned}
$$

C   Lemmas regarding v, g and MSPE

For the proofs we will employ (Lemma C.1(i) and(ii)) that $\bar{v}_{N,T}^{(p,h)}(t)$ can be approximated by $v_0^{(p,h)}(t/T)$ or $v_{N/T}^{(p,h)}(t/T)$, where $\bar{v}_{\Delta}^{(p,h)}(u)$ was defined in (46).

More precisely, the following lemma that is constructed analogously to Lemma 6.5, but for the h-step ahead coefficients, holds:

Lemma C.1 *Let* $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ *satisfy Assumptions 1, 3, and 4 (Appendix A) and* $\mathbb{E}X_{t,T} = 0$. *Define* C_0 *and* C_1 *as in (43). Then, we have, for* $\bar{v}_{N,T}^{(p,h)}(t)$ *defined in (46),*

(i) if $T \geq 9C_1pN$, *with* $v_0^{(p,h)}(t/T)$ *defined in (50), that*

$$
\|\bar{v}_{N,T}^{(p,h)}(t) - v_0^{(p,h)}(t/T)\| \leq 5h\big(2C_0\big)^h C_1 \frac{pN}{T}.
$$

(ii) if $T \geq 5C_1p$, *with* $v_{N/T}^{(p,h)}(t/T)$ *defined in (50), that*

$$
\|\bar{v}_{N,T}^{(p,h)}(t) - v_{N/T}^{(p,h)}(t/T)\| \leq 3h\big(2C_0\big)^h C_1 \frac{p^2}{T},
$$

59

Further, for the norms of $u \mapsto v_\Delta^{(p,h)}(u)$ and $\Delta \mapsto v_\Delta^{(p,h)}(u)$ and the derivatives, we have the following bounds:

Lemma C.2 *Let $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ satisfy Assumptions 1, 3, and 4 (Appendix A) and $\mathbb{E}X_{t,T} = 0$. C_0 as in (43) and m_f, M_f, M_f' from the assumptions. Then, with $v_\Delta^{(p,h)}(u)$ defined in (50), we have*

(i) $\|v_\Delta^{(p,h)}(u)\| \leq (C_0)^h$,

(ii) $v_\Delta^{(p,h)}(\cdot)$ *is continuously differentiable, with*

$$\left\| \frac{\partial}{\partial u} v_\Delta^{(p,h)}(u) \right\| \leq h(C_0)^h M_f'(m_f^{-1} + M_f^{-1}),$$

(iii) $\Delta \mapsto v_\Delta^{(p,h)}(u)$, $\Delta > 0$, *is continuously differentiable, with*

$$\left\| \frac{\partial}{\partial \Delta} v_\Delta^{(p,h)}(u) \right\| \leq 2h(C_0)^h M_f'(m_f^{-1} + M_f^{-1})/\Delta.$$

Lemma C.2 also holds for $h = 1$. Part (i) thus extends Lemma 6.6.

Finally, we use Lemmas 6.4, 6.5 and 6.6 to bound the norm of $\bar{a}_{N,T}^{(p)}(t)$ and $\bar{v}_{N,T}^{(p)}(t,h)$.

Lemma C.3 *Let $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ satisfy Assumptions 1, 3, and 4 (Appendix A) and $\mathbb{E}X_{t,T} = 0$. Define C_0 and C_1 as in (43). Then,*

(i) for $T \geq 5C_1 p^2$ we have, for $\bar{a}_{N,T}^{(p,h)}(t)$ defined in (33),

$$\|\bar{a}_{N,T}^{(p)}(t)\| \leq 2C_0, \ \ and \ \|\bar{v}_{N,T}^{(p,h)}(t)\|_\infty \leq \left(2C_0 + 1\right)^h.$$

Further,

(ii) for $T \geq 3h2^h C_1 p^2$ we have, for $\bar{v}_{N,T}^{(p,h)}(t)$ defined in (46),

$$\|\bar{v}_{N,T}^{(p,h)}(t)\| \leq 2(C_0)^h.$$

Note that Lemma C.3(i) implies that we have $\|\bar{v}_{N,T}^{(p,h)}(t)\| \leq p^{1/2}\left(2C_0 + 1\right)^h$. The bound in Lemma C.3(ii) is tighter, but requires T to be larger.

An important observation is that, as a function of u, $\text{MSPE}_{N/T,n/T}^{(p,h)}(u)$ is differentiable with bounded derivative

Lemma C.4 *Let* $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ *satisfy Assumptions 1, 3, and 4 (Appendix A) and* $\mathbb{E}X_{t,T} = 0$. C_0 *as in (43) and the other constants from the assumptions. Then, the function* $g_\Delta^{(p,h)}$, *defined in (49), is continuously differentiable and the derivatives are bounded. More precisely, We have*

(i)
$$\left|g_\Delta^{(p,h)}(u)\right| \le 4M_f\big(C_0\big)^{2h},$$

(ii)
$$\left|\frac{\partial}{\partial u}g_\Delta^{(p,h)}(u)\right| \le 4\big(2h+1\big)\big(C_0\big)^{2h+1}M_f',$$

(iii)
$$\left|\frac{\partial}{\partial\Delta}g_\Delta^{(p,h)}(u)\right| \le 8\big(2h+1\big)\big(C_0\big)^{2h+1}M_f'/\Delta.$$

In particular, this implies, or $\text{MSPE}_{\Delta_1,\Delta_2}^{(p,h)}(u)$ *defined in (48), that*

(iv)
$$\left|\frac{\partial}{\partial u}\text{MSPE}_{\Delta_1,\Delta_2}^{(p,h)}(u)\right| \le 4\big(2h+1\big)\big(C_0\big)^{2h+1}M_f'.$$

(v)
$$\left|\frac{\partial}{\partial\Delta_1}\text{MSPE}_{\Delta_1,\Delta_2}^{(p,h)}(u)\right| \le 8\big(2h+1\big)\big(C_0\big)^{2h+1}M_f'/\Delta_1.$$

(vi)
$$\left|\frac{\partial}{\partial\Delta_2}\text{MSPE}_{\Delta_1,\Delta_2}^{(p,h)}(u)\right| \le 8M_f\big(C_0\big)^{2h}/\Delta_2.$$

Proof of Lemma C.1. By the definitions of $\bar{A}_{N,T}^{(p)}(t)$ and $A_\Delta^{(p)}(t/T)$, which we now denote by A and A_0, respectively. We have

$$\begin{aligned}
\|\bar{v}_{N,T}^{(p,h)}(t) - v_0^{(p,h)}(t/T)\| &= \left\|(1,0,\ldots,0)\big(A^h - A_0^h\big)\right\| \\
&\le \left\|A^h - A_0^h\right\| \le h\|A - A_0\|\big(\|A - A_0\| + \|A_0\|\big)^{h-1} \\
&\le h\|\bar{a}_{N,T}^{(p)}(t) - a_0^{(p)}(t/T)\|\big(\|\bar{a}_{N,T}^{(p)}(t) - a_0^{(p)}(t/T)\| + \|a_0^{(p)}(t/T)\|\big)^{h-1} \\
&\le h\big(9\frac{N}{T}C_0C_1p\big)\big(9\frac{N}{T}C_0C_1p + C_0\big)^{h-1} \\
&\le h\big(9\frac{N}{T}C_0C_1p\big)\big(2C_0\big)^{h-1} \le 5C_1hp\frac{N}{T}\big(2C_0\big)^h,
\end{aligned}$$

where the first inequality is due to the submultiplicativity of the spectral norm, the second inequality is due to Lemma F.3, the third inequality is due to the fact that $\|A - A_0\| \le$

$\|A - A_0\|_F$ and $\|A_0\| = \|e_1 a^{(p)}(t/T)\|$, because adding/subtracting the Jordan block doesn't change the eigenvalues (see also the proof of Lemma C.2).

The fourth inequality is due to Lemma 6.6 and Lemma 6.5(i). The fifth inequality holds if $T \geq T_{0,1} := N 9 C_0 C_1 p C_0^{-1}$.

For (ii) we have, by a similar argument as above that

$$
\begin{aligned}
\|\bar{v}_{N,T}^{(p)}(t) - v_{N/T}^{(p)}(t/T)\| &\leq h \big(C_1 \tfrac{1}{T}(5C_0 p^2)\big)\big(C_1 \tfrac{1}{T}(5C_0 p^2) + C_0\big)^{h-1} \\
&\leq h C_1 \tfrac{1}{T}(5C_0 p^2)\big(2C_0\big)^{h-1} = 3h C_1 \frac{p^2}{T}\big(2C_0\big)^h
\end{aligned}
$$

where the second inequality holds if $T \geq T_{0,1} := 5 C_0 C_1 p^2 C_0^{-1}$. Note that $T \geq T_{0,1}$ implies $T \geq 2C_1 p^2$, which is the condition required to apply Lemma 6.5(ii). $\qquad\square$

Proof of Lemma C.2. Let e_1 and H be defined as in (47). We derive a compact expression for $\mathrm{d}v$ where $v(u)$ is short for

$$
v_\Delta^{(p,h)}(u) := e_1'\big(e_1(a(u))' + H\big)^h =: e_1' X(u)^h,
$$

with $a(u)$ being short for $a_\Delta^{(p)}(u) := \Gamma_\Delta^{(p)}(u)^{-1}\gamma_\Delta^{(p)}(u)$. We will further abbreviate $\Gamma(u) := \Gamma_\Delta^{(p)}(u)$ and $\gamma(u) := \gamma_\Delta^{(p)}(u)$.

By Weyl's inequality we have that the eigenvalues of $\mu_1 \leq \ldots \leq \mu_p$ of $e_1 a(u)' + H$ fulfil $\lambda_1 + \rho_i \leq \mu_i \leq \rho_i + \lambda_p$, where $\rho_1 \leq \ldots \leq \rho_p$ denote the eigenvalues of $e_1 a(u)'$ and $\lambda_1 \leq \ldots \leq \lambda_p$ denote the eigenvalues of H. Note further, that the eigenvalues of H are $\lambda_1 = \ldots = \lambda_p = 0$ (it is a Jordan block). Therefore: $\mu_i = \rho_i$. In particular: $\|e_1 a(u)' + H\| = \|e_1 a(u)'\| \leq \|a(u)\| \leq C_0$, where the last inequality follows from Lemma 6.6. This observation, obviously, implies that $\|v\| = \|e_1' X^h\| \leq \|X\|^h \leq \|a\|^h$, which yields (i).

For the proof of (ii) recall that the notation reflects that v, a, Γ, and γ are functions of the variable u (for a fixed $\Delta \geq 0$). By elementary rule (15), p. 148 and Theorem 3, p. 151 from Magnus and Neudecker (1988), we have, that

$$
\mathrm{d}a = -\Gamma^{-1}(\mathrm{d}\Gamma)\Gamma^{-1}\gamma + \Gamma^{-1}\mathrm{d}\gamma \tag{78}
$$

Iterating elementary rule (15), p. 148, we get that, for every square matrix function X and $h = 1, 2, \ldots$ that

$$
\mathrm{d}(X^h) = \sum_{k=1}^{h} X^{k-1}(\mathrm{d}X)X^{h-k}.
$$

Obviously we have $\mathrm{d}X = e_1(\mathrm{d}a)'$, which yields

$$
\begin{aligned}
\mathrm{d}v &= e_1' \sum_{k=1}^{h} X^{k-1}(\mathrm{d}X)X^{h-k} \\
&= e_1' \sum_{k=1}^{h} \big(e_1 a' + H\big)^{k-1}(e_1(\mathrm{d}a)')\big(e_1 a' + H\big)^{h-k}
\end{aligned} \tag{79}
$$

For $\|\mathrm{d}a\|$ note that, by (78) and employing Lemma E.3(i), we have

$$
\begin{aligned}
\|\mathrm{d}a\| &\le \|\Gamma^{-1}\|^2 \|\mathrm{d}\Gamma\| \|\gamma\| + \|\Gamma^{-1}\| \|\mathrm{d}\gamma\| \\
&\le m_f^{-2} M_f' ((2\pi)^{1/2} M_f) + m_f^{-1} ((2\pi)^{1/2} M_f') \\
&= (2\pi)^{1/2} m_f^{-1} M_f' (M_f/m_f + 1).
\end{aligned}
$$

Here we have used that, by the assumed continuity of $\frac{\partial}{\partial u} f(u, \lambda)$, we have

$$
(\mathrm{d}\Gamma_\Delta^{(p)})_{i,j}(u) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}(i-j)\lambda} \Big(\int_0^1 \frac{\partial}{\partial u} f\big(u + \Delta(x-1), \lambda\big) \mathrm{d}x \Big) \mathrm{d}\lambda,
$$

such that an application of Lemma 4.1 from Gray (2009) (cf. the proof of Lemma E.3(i)) yields that $\|\mathrm{d}\Gamma\| \le M_f'$. Further, since

$$
(\mathrm{d}\gamma_\Delta^{(p)})_j(u) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}j\lambda} \Big(\int_0^1 \frac{\partial}{\partial u} f\big(u + \Delta(x-1), \lambda\big) \mathrm{d}x \Big) \mathrm{d}\lambda,
$$

we have by Bessel's inequality (see, again, the proof of Lemma E.3(i)) that $\|\mathrm{d}\gamma\| \le (2\pi)^{1/2} M_f'$. In conclusion, [cf. (79)] we have shown

$$
\|\mathrm{d}v\| \le h\|a\|^{h-1} \|\mathrm{d}a\| \le h(C_0)^{h-1} (2\pi)^{1/2} m_f^{-1} M_f' (M_f/m_f + 1).
$$

For the proof of (iii) consider Δ to be the argument of the functions (instead of u). Let this be reflected by changing the notation to $v(\Delta) := v_\Delta^{(p,h)}(u)$, $a(\Delta) := a_\Delta^{(p)}(u)$, $\Gamma(\Delta) := \Gamma_\Delta^{(p)}(u)$, and $\gamma(\Delta) := \gamma_\Delta^{(p)}(u)$ (for a fixed $u \in \mathbb{R}$). Note that all previous arguments remain the same, but for the derivation of $\mathrm{d}\Gamma$ and $\mathrm{d}\gamma$. To this end, denote $f_\lambda(u) := f(\lambda, u)$ and note that we have

$$
\begin{aligned}
\frac{\partial}{\partial \Delta}(\gamma_\Delta^{(p)}(u))_j &= \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}j\lambda} \Big(\int_0^1 \frac{\partial}{\partial \Delta} f\big(u + \Delta(x-1), \lambda\big) \mathrm{d}x \Big) \mathrm{d}\lambda \\
&= \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}j\lambda} \Big(\int_0^1 f_\lambda'(u + \Delta(x-1)) x \, \mathrm{d}x \Big) \mathrm{d}\lambda \\
&= \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}j\lambda} \Big(\Big[\frac{f_\lambda(u + \Delta(x-1))}{\Delta} x \Big]_0^1 - \int_0^1 \frac{f_\lambda(u + \Delta(x-1))}{\Delta} \, \mathrm{d}x \Big) \mathrm{d}\lambda \\
&= \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}j\lambda} \Big(\frac{1}{\Delta} \int_0^1 \big(f_\lambda(u) - f_\lambda(u + \Delta(x-1)) \big) \mathrm{d}x \Big) \mathrm{d}\lambda.
\end{aligned}
$$

Thus,

$$
\Big\| \frac{\partial}{\partial \Delta} \gamma_\Delta^{(p)}(u) \Big\| \le 2(2\pi)^{1/2} M_f/\Delta, \text{ and } \Big\| \frac{\partial}{\partial \Delta} \Gamma_\Delta^{(p)}(u) \Big\| \le 2 M_f/\Delta.
$$

$\square$

Proof of Lemma C.3. This follows, because

$$
\begin{aligned}
\|\bar{a}_{N,T}^{(p)}(t)\| &= \|a_{N/T}^{(p)}(t/T) + \bar{a}_{N,T}^{(p)}(t) - a_{N/T}^{(p)}(t/T)\| \\
&\le \|a_{N/T}^{(p)}(t/T)\| + \|\bar{a}_{N,T}^{(p)}(t) - a_{N/T}^{(p)}(t/T)\| \\
&\le C_0 + C_1 \frac{1}{T} \big(5 C_0 p^2 \big) \le 2 C_0,
\end{aligned}
$$

63

The second inequality follows by application of Lemmas 6.6 and 6.5(ii). Note that Lemma 6.5(ii) can be applied since $T \geq 2C_1 p^2$ by assumption. The third inequality follows, if $T \geq 5C_1 p^2$, which is what we assumed.

For the second inequality note that, similar to the definition of $\hat{v}_{N,T}^{(p,h)}(t)$ in Step 3 of Section 3.1 and (4), we have the following recursive relationship

$$\bar{v}_{N,T}^{(p,h)}(t) = e_1' \big(e_1(\bar{a}_{N,T}^{(p)}(t))' + H\big)^h = \bar{v}_{N,T}^{(p,h-1)}(t)\big(e_1(\bar{a}_{N,T}^{(p)}(t))' + H\big)$$
$$= \bar{v}_{N,T}^{(p,h-1)}(t)e_1(\bar{a}_{N,T}^{(p)}(t))' + \bar{v}_{N,T}^{(p,h-1)}(t)H.$$

In other terms, we have
$$\bar{v}_{i,N,T}^{(p,1)}(t) = \bar{a}_{i,N,T}^{(p)}(t)$$
and, for every h = 2, 3, ..., we have

$$\bar{v}_{i,N,T}^{(p)}(t,h) = \bar{a}_{i,N,T}^{(p)}(t)\bar{v}_{1,N,T}^{(p,h-1)}(t) + \bar{v}_{i+1,N,T}^{(p,h-1)}(t)I\{i \leq p-1\},$$

which implies that

$$\|\bar{v}_{N,T}^{(p,h)}(t)\|_\infty \leq (\|\bar{a}_{N,T}^{(p)}(t)\|_\infty + 1)\|\bar{v}_{N,T}^{(p,h-1)}(t)\|_\infty \leq \big(\|\bar{a}_{N,T}^{(p)}(t)\|_\infty + 1\big)^{h-1}\|\bar{a}_{N,T}^{(p)}(t)\|_\infty.$$

Employing the first part of this lemma we thus have

$$\|\bar{v}_{N,T}^{(p,h)}(t)\|_\infty \leq \big(2C_0 + 1\big)^{h-1}2C_0 \leq \big(2C_0 + 1\big)^h.$$

For the proof of (ii), as in the proof of the bound for $\|\bar{a}_{N,T}^{(p)}(t)\|$ in (i), we have

$$\|\bar{v}_{N,T}^{(p,h)}(t)\| \leq \|v_{N/T}^{(p,h)}(t/T)\| + \|\bar{v}_{N,T}^{(p,h)}(t) - v_{N/T}^{(p,h)}(t/T)\|$$
$$\leq \big(C_0\big)^h + 3h\big(2C_0\big)^h C_1 \frac{p^2}{T} \leq 2\big(C_0\big)^h,$$

where, for the second inequality, we have employed Lemma C.2(i) and C.1(ii). The third inequality holds, if

$$T \geq 3h\big(2C_0\big)^h\big(C_0\big)^{-h}C_1 p^2 = 3h2^h C_1 p^2,$$

which, for all $h = 1, 2, \ldots$, implies the condition required to apply Lemma C.1(ii). $\qquad\square$

Proof of Lemma C.4. For the sake of brevity, we omit sub and superscripts that are not needed for this particular argument. More precisely, in this section we denote $g_\Delta^{(p,h)}(u)$ by g, $v_\Delta^{(p,h)}(u)$ by v, $\gamma_0(u)$ by γ_0, $\gamma_0^{(p,h)}(u)$ by γ, and $\Gamma_0^{(p)}(u)$ by Γ, respectively.

For (i) note that we have, by (36) together with Assumption 3, by Lemma C.2(i), and by Lemma E.3(i),

$$\|g\| \leq \|\gamma_0\| + 2\|v'\| \cdot \|\gamma\| + \|v'\| \cdot \|\Gamma\| \cdot \|v\|$$
$$\leq 2\pi M_f + 2(2\pi)^{1/2}M_f\big(C_0\big)^h + \big(C_0\big)^{2h}M_f$$
$$= \big((2\pi)^{1/2} + \big(C_0\big)^h\big)^2 M_f \leq 4M_f\big(C_0\big)^{2h}$$

We now compute the derivatives of $g_\Delta^{(p,h)}(u)$. We have:

$$
\begin{aligned}
\mathrm{d}g &= \mathrm{d}\big(\gamma_0 - 2v'\gamma + v'\Gamma v\big) \\
&= \mathrm{d}\gamma_0 - 2\big((\mathrm{d}v')\gamma + v'\mathrm{d}\gamma\big) \\
&\quad + (\mathrm{d}v')\Gamma v + v'(\mathrm{d}\Gamma)v + v'\Gamma(\mathrm{d}v) \\
&= \mathrm{d}\gamma_0 + 2(\mathrm{d}v')(\Gamma v - \gamma) + v'\big((\mathrm{d}\Gamma)v - 2\mathrm{d}\gamma\big).
\end{aligned}
$$

Note that, for $\Delta = 0$ and $h = 1$, we have $v = a$ and by the definition of a, it follows that $\Gamma v - \gamma = 0$, such that the second term of the derivative will only show if $h \geq 2$.

In consequence, for (i) we bound $\mathrm{d}g_{p,h}$ as follows:

$$
\begin{aligned}
&|\mathrm{d}g_{p,h}| \\
&= \|\mathrm{d}\gamma_0 + 2(\mathrm{d}v')(\Gamma v - \gamma) + v'\big((\mathrm{d}\Gamma)v - 2\mathrm{d}\gamma\big)\| \\
&\leq \|\mathrm{d}\gamma_0\| + 2\|\mathrm{d}v'\|(\|\Gamma\|\|v\| + \|\gamma\|) + \|v'\|(\|\mathrm{d}\Gamma\|\|v\| + 2\|\mathrm{d}\gamma\|) \\
&\leq 2\pi M_f' + 2h(C_0)^h M_f'(m_f^{-1} + M_f^{-1})\Big(M_f(C_0)^h + (2\pi)^{1/2}M_f\Big) \\
&\quad + (C_0)^h\Big(M_f'(C_0)^h + 2(2\pi)^{1/2}M_f'\Big) \\
&= M_f'\Big[2\pi + 2h(C_0)^h(M_f/m_f + 1)\big((C_0)^h + (2\pi)^{1/2}\big) + (C_0)^h\big((C_0)^h + 2(2\pi)^{1/2}\big)\Big] \\
&\leq M_f'\Big[(C_0)^{2h} + 2h(C_0)^h(M_f/m_f + 1)\big((C_0)^h + (C_0)^h\big) + (C_0)^h\big((C_0)^h + 2(C_0)^h\big)\Big] \\
&= M_f'\Big[(C_0)^{2h} + (C_0)^{2h}\big(4h(M_f/m_f + 1) + 3\big)\Big] \leq 4M_f'(C_0)^{2h}\Big(h(M_f/m_f + 1) + 1\Big) \\
&\leq 4M_f'(C_0)^{2h}\Big(h(C_0 + C_0) + C_0\Big) \leq 4M_f'(C_0)^{2h+1}(2h + 1).
\end{aligned}
$$

In the above, for the second inequality, we have used (42) together with Assumption 4, Lemma C.2(i)+(ii), and Lemma E.3(i).

For the proof of (ii) use the partial derivatives of γ and Γ with respect to Δ that were derived in the proof of Lemma C.2(iii). For the proof of (iii) use Lemma C.2(iii) instead of Lemma C.2(ii) and use the bounds for $\|\mathrm{d}\Gamma\|$ and $\|\mathrm{d}\gamma\|$ that were derived in the end of the proof of Lemma C.2(iii). For the proof of (iv) note that the partial derivative of g is continuous by (i) and employ Leibniz's integral rule. (v) follows analogously from (ii). For the proof of (vi) we note that, by the argument in the proof of Lemma C.2(iii), we have

$$
\frac{\partial}{\partial \Delta_2} \mathrm{MSPE}_{\Delta_1,\Delta_2}^{(p,h)}(u) = \frac{1}{\Delta_2} \int_0^1 \big(g_{\Delta_1}^{(p,h)}(u) - g_{\Delta_1}^{(p,h)}(u + \Delta_2(x - 1))\big)\, \mathrm{d}x.
$$

(vi) then follows from (i) of this lemma. $\qquad \square$

D Proof of auxiliary results from Section 6

Proof of Lemma 6.6. Note that

$$
\|a_\Delta^{(p)}(u)\| = \|\Gamma_\Delta^{(p)}(u)^{-1}\gamma_\Delta^{(p)}(u)\| \leq \|\Gamma_\Delta^{(p)}(u)^{-1}\|\|\gamma_\Delta^{(p)}(u)\|.
$$

The assertion follows from Lemma E.3(i). □

Outline for the proof of Lemma 6.4 and Lemma 6.5.

Note that the norm to be bounded is, in all cases, of a difference of quantities of the form $a_1 := \Gamma_1^{-1}\gamma_1$ and $a_2 := \Gamma_2^{-1}\gamma_2$, where the norms of the components of a_2 and norms of the differences of the components can be bounded using results from Section E.1. We denote these bounds by $\|\Gamma_2^{-1}\| \le K$, $\|\gamma_2\| \le \kappa$, $\|\Gamma_1 - \Gamma_2\| \le D$, and $\|\gamma_1 - \gamma_2\| \le \delta$.

For the bound of the norm of interest note that

$$a_1 - a_2 = \Gamma_1^{-1}\gamma_1 - \Gamma_2^{-1}\gamma_2 = (\Gamma_1^{-1} - \Gamma_2^{-1})\gamma_1 + \Gamma_2^{-1}(\gamma_1 - \gamma_2)$$

Thus, we have $\|a_1 - a_2\| \le \|\Gamma_1^{-1} - \Gamma_2^{-1}\| \cdot \|\gamma_1\| + \|\Gamma_2^{-1}\| \cdot \|\gamma_1 - \gamma_2\|$. For each of the inequalities to be proven, we will provide a $T_{0,1}$ such that

$$\|\Gamma_2 - \Gamma_1\| \cdot \|\Gamma_2^{-1}\| \le D \cdot K \le 1/2, \text{ for all } T \ge T_{0,1}. \tag{80}$$

Under this condition, we have, by Lemma F.2, that

$$\|\Gamma_1^{-1} - \Gamma_2^{-1}\| = \|(\Gamma_2 + \Gamma_1 - \Gamma_2)^{-1} - \Gamma_2^{-1}\| \le 2\|\Gamma_1 - \Gamma_2\| \cdot \|\Gamma_2^{-1}\|^2,$$

which yields

$$\|a_1 - a_2\| \le 2\|\Gamma_1 - \Gamma_2\| \cdot \|\Gamma_2^{-1}\|^2 \cdot \|\gamma_1\| + \|\Gamma_2^{-1}\| \cdot \|\gamma_1 - \gamma_2\|. \tag{81}$$

Finally, we will provide a $T_{0,2}$ such that

$$\kappa - \delta \ge 0, \text{ for all } T \ge T_{0,2}. \tag{82}$$

Under this condition, we have $\|\gamma_1\| \le \kappa + \delta \le 2\kappa$, by the triangle inequality. It is implied that, for $T \ge \max\{T_{0,1}, T_{0,2}\}$, the norm $\|a_1 - a_2\|$ does not exceed

$$4 \, K^2 \, \kappa \, D + K \, \delta. \tag{83}$$

For the proof of the individual cases it thus suffices to refer to Section E.1 where appropriate bounds (i.e., K, κ, D, and δ) can be found, from which $T_{0,1}$, $T_{0,2}$, and (83) are obtained.

Proof of Lemma 6.4.

We proceed as outlined and choose $\tilde{a}_T^{(p)}(t)$ to be a_1 and $a_0^{(p)}(t/T)$ as a_2. Lemma E.3(i) and Corollary E.2(i) provide the needed bounds. We have

$$\|\tilde{\Gamma}_T^{(p)}(t) - \Gamma_0^{(p)}(t/T)\| \cdot \|\Gamma_0^{(p)}(u)^{-1}\| \le T^{-1}p^2(2\pi M_f' + C)m_f^{-1}$$

Thus, (80) holds if $T \ge T_{0,1} := 2p^2(2\pi M_f' + C)m_f^{-1}$. Further, we have

$$\|\tilde{\gamma}_T^{(p)}(t) - \gamma_0^{(p)}(t/T)\| \le T^{-1}p^{1/2}C, \quad \|\gamma_0^{(p)}(t/T)\| \le (2\pi)^{1/2}M_f,$$

such that (82) holds if $T \geq T_{0,2} := \frac{C}{M_f}(\frac{p}{2\pi})^{1/2}$. Consequently, by (83), we have the following bound

$$\|\tilde{a}_T^{(p)}(t) - a_0^{(p)}(t/T)\| \leq 4T^{-1}p^2(2\pi M_f' + C)m_f^{-2}(2\pi)^{1/2}M_f + m_f^{-1}T^{-1}p^{1/2}C$$
$$= \frac{1}{T}\left(4C_0C_1\,p^2 + p^{1/2}C/m_f\right),$$

which holds, if $T \geq \max\{2p^2C_1, p^{1/2}C_0^{-1}C/m_f\}$. The assertion follows, as we have $p^{1/2}C/m_f \leq C_1p^2$ and $C_0^{-1} < 1$. $\qquad\square$

Proof of Lemma 6.5. We proceed as outlined. For (i) we choose $\bar{a}_{N,T}^{(p)}(t)$ to be a_1 and $a_0^{(p)}(t/T)$ as a_2. Lemma E.3(i) and Corollary E.2(ii) provide the needed bounds. We have

$$\|\mathbb{E}\hat{\Gamma}_{N,T}^{(p)}(t) - \Gamma_0^{(p)}(t/T)\| \cdot \|\Gamma_0^{(p)}(u)^{-1}\| \leq T^{-1}p(2\pi M_f'(N+1) + C)m_f^{-1}.$$

Thus, (80) holds if $T \geq T_{0,1} := 2p(2\pi(N+1)M_f' + C)m_f^{-1}$. Further, we have

$$\|\mathbb{E}\hat{\gamma}_{N,T}^{(p)}(t) - \gamma_0^{(p)}(t/T)\| \leq T^{-1}p^{1/2}(2\pi M_f'N + C), \quad \|\gamma_0^{(p)}(t/T)\| \leq (2\pi)^{1/2}M_f,$$

such that (82) holds if $T \geq T_{0,2} := \frac{2\pi M_f'N+C}{M_f}(\frac{p}{2\pi})^{1/2}$. Consequently, by (83), we have the following bound

$$\|\bar{a}_{N,T}^{(p)}(t) - a^{(p)}(t/T)\| \leq 4T^{-1}p(2\pi M_f'(N+1) + C)m_f^{-2}(2\pi)^{1/2}M_f$$
$$+ m_f^{-1}T^{-1}p^{1/2}(2\pi M_f'N + C)$$
$$\leq 4T^{-1}p2N(2\pi M_f' + C)m_f^{-2}(2\pi)^{1/2}M_f$$
$$+ m_f^{-1}T^{-1}p^{1/2}N(2\pi M_f' + C)$$
$$= C_1\frac{N}{T}\left(8C_0p + p^{1/2}\right)$$

for $T \geq \max\{2p(2\pi M_f'(N+1) + C)m_f^{-1}, \frac{2\pi M_f'N+C}{M_f}(\frac{p}{2\pi})^{1/2}\}$. We chose to state the result to hold for $T \geq NC_1\max\{4p, p^{1/2}C_0^{-1}\}$, which is more restrictive, but allows for the more compact expression, as $C_0 > 1$.

For the proof of (ii) we choose $\bar{a}_{N,T}^{(p)}(t)$ to be a_1 and $a_{N/T}^{(p)}(t/T)$ as a_2. Lemma E.3(i) and Corollary E.2(iii) provide the needed bounds. We have

$$\|\mathbb{E}\hat{\Gamma}_{N,T}^{(p)}(t) - \Gamma_{N/T}^{(p)}(t/T)\| \cdot \|\Gamma_{N/T}^{(p)}(t/T)^{-1}\| \leq T^{-1}p^2(2\pi M_f' + C)m_f^{-1}$$

Thus, (80) holds if $T \geq T_{0,1} := 2p^2(2\pi M_f' + C)m_f^{-1}$. Further, we have

$$\|\mathbb{E}\hat{\gamma}_{N,T}^{(p)}(t) - \gamma_{N/T}^{(p)}(t/T)\| \leq 2T^{-1}p^{3/2}(2\pi M_f' + C), \quad \|\gamma_{N/T}^{(p)}(t/T)\| \leq (2\pi)^{1/2}M_f,$$

such that (82) holds if $T \geq T_{0,2} := \frac{2p^{3/2}(2\pi M_f'+C)}{(2\pi)^{1/2}M_f}$. Consequently, by (83), we have the following bound

$$\|\bar{a}_{N,T}^{(p)}(t) - a_{N/T}^{(p)}(t/T)\| \leq 4m_f^{-2}(2\pi)^{1/2}M_fT^{-1}p^2(2\pi M_f' + C)$$
$$+ 2m_f^{-1}T^{-1}p^{3/2}(2\pi M_f' + C)$$
$$= C_1\frac{1}{T}\left(4C_0p^2 + 2p^{3/2}\right),$$

for $T \geq C_1\max\{2p^2, 2p^{3/2}C_0^{-1}\}$. (ii) follows from $C_0^{-1} < 1$ and $p^{3/2} < C_0p^2$. $\qquad\square$

E Properties of empirical localised auto-covariances

E.1 Approximations for moments

Lemma E.1 *Let $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ satisfy Assumptions 1 and 4 (Appendix A), $\mathbb{E}X_{t,T}=0$. Then, with $\hat{\gamma}_{k;N,T}(t)$ defined in (3), $f(u,\lambda)$ and C from Assumption 1, and M_f' from Assumption 4, we have: (i)*

$$\left|\mathbb{E}\hat{\gamma}_{k;N,T}(t) - \int_{-\pi}^{\pi}\left[\int_0^1 f\big(\frac{t-N+|k|}{T} + \frac{N-|k|}{T}u,\lambda\big)\mathrm{d}u\right]\mathrm{e}^{\mathrm{i}|k|\lambda}\mathrm{d}\lambda\right| \le \frac{2\pi M_f' + C}{T}$$

and (ii)

$$\left|\mathbb{E}\hat{\gamma}_{k;N,T}(t) - \int_{-\pi}^{\pi}\left[\int_0^1 f\big(\frac{t-N}{T} + \frac{N}{T}u,\lambda\big)\mathrm{d}u\right]\mathrm{e}^{\mathrm{i}|k|\lambda}\mathrm{d}\lambda\right| \le \frac{2\pi(|k|+1)M_f' + C}{T}$$

and (iii)

$$\left|\mathbb{E}\hat{\gamma}_{k;N,T}(t) - \gamma_k(t/T)\right| \le \frac{2\pi M_f'(N-|k|+1) + C}{T}.$$

Proof of Lemma E.1. We have

$$\mathbb{E}\hat{\gamma}_{k;N,T}(t) = \frac{1}{N-|k|}\sum_{\ell=t-N+|k|+1}^{t}\mathbb{E}(X_{\ell-|k|,T}X_{\ell,T})$$

$$= \frac{1}{N-|k|}\sum_{\ell=t-N+|k|+1}^{t}\big(\int_{-\pi}^{\pi}f(\ell/T,\lambda)\mathrm{e}^{\mathrm{i}|k|\lambda}\mathrm{d}\lambda + R^{(1)}_{\ell,|k|,T}\big),\ |R^{(1)}_{\ell,|k|,T}| \le C/T$$

$$= \int_{-\pi}^{\pi}\left[\frac{1}{N-|k|}\sum_{\ell=t-N+|k|+1}^{t}f(\ell/T,\lambda)\right]\mathrm{e}^{\mathrm{i}|k|\lambda}\mathrm{d}\lambda + R^{(2)}_{t,N,|k|,T},\ |R^{(2)}_{t,N,|k|,T}| \le C/T$$

$$= \int_{-\pi}^{\pi}\left[\int_0^1 f\big(\frac{t-N+|k|}{T} + \frac{N-|k|}{T}u,\lambda\big)\mathrm{d}u\right.$$

$$\left. + R^{(3)}_{t-N+|k|,t,T}(\lambda)\right]\mathrm{e}^{\mathrm{i}|k|\lambda}\mathrm{d}\lambda + R^{(2)}_{t,N,|k|,T}$$

$$= \int_{-\pi}^{\pi}\left[\int_0^1 f\big(\frac{t-N+|k|}{T} + \frac{N-|k|}{T}u,\lambda\big)\mathrm{d}u\right]\mathrm{e}^{\mathrm{i}|k|\lambda}\mathrm{d}\lambda + R^{(4)}_{t,N,|k|,T},$$

$$|R^{(4)}_{t,N,|k|,T}| \le \frac{2\pi M_f' + C}{T}$$

where we have

$$\left|R^{(3)}_{A,B,T}(\lambda)\right| \le \sup_{(t-N+|k|)/T<u<t/T}\left|\frac{\partial}{\partial u}f(u,\lambda)\right|\frac{1}{T} \le M_f'/T,$$

by Lemma F.8, where the differentiability (with respect to u) is guaranteed by Assumption 4.

For the second equation in the lemma, note that (for fixed t, N, k, u) we have

$$
\begin{aligned}
f\Big(&\frac{t - N + |k|}{T} + \frac{N - |k|}{T}u, \lambda\Big) \\
&= f\Big(\frac{t - N}{T} + \frac{N}{T}u, \lambda\Big) \\
&\quad + f\Big(\frac{t - N + |k|}{T} + \frac{N - |k|}{T}u, \lambda\Big) - f\Big(\frac{t - N}{T} + \frac{N}{T}u, \lambda\Big) \\
&= f\Big(\frac{t - N}{T} + \frac{N}{T}u, \lambda\Big) + \frac{\partial}{\partial x} f(x, \lambda)\Big|_{x = \xi}\Big(\frac{|k|}{T}(1 - u)\Big),
\end{aligned}
\tag{84}
$$

where $\xi \in \Big(\frac{t-N}{T} + \frac{N}{T}u, \frac{t-N+|k|}{T} + \frac{N-|k|}{T}u\Big)$.

Thus, by Assumption 4 and because $u \in [0, 1]$, we have

$$
\Big|f\Big(\frac{t - N + |k|}{T} + \frac{N - |k|}{T}u, \lambda\Big) - f\Big(\frac{t - N}{T} + \frac{N}{T}u, \lambda\Big)\Big| \leq M'_f\Big(\frac{|k|}{T}\Big).
$$

Thus,

$$
\begin{aligned}
\Big|\int_{-\pi}^{\pi} &\Big[\int_0^1 \Big(f\Big(\frac{t - N + |k|}{T} + \frac{N - |k|}{T}u, \lambda\Big) - f\Big(\frac{t - N}{T} + \frac{N}{T}u, \lambda\Big)\Big)\mathrm{d}u\Big]\mathrm{e}^{\mathrm{i}|k|\lambda}\mathrm{d}\lambda\Big| \\
&\leq \int_{-\pi}^{\pi}\int_0^1 \Big|f\Big(\frac{t - N + |k|}{T} + \frac{N - |k|}{T}u, \lambda\Big) - f\Big(\frac{t - N}{T} + \frac{N}{T}u, \lambda\Big)\Big|\mathrm{d}u\mathrm{d}\lambda \leq 2\pi M'_f\Big(\frac{|k|}{T}\Big).
\end{aligned}
$$

For the proof of the third part of the Lemma, note that (for fixed t, N, k, u) we have, in analogy to (84)

$$
f\Big(\frac{t - N + |k|}{T} + \frac{N - |k|}{T}u, \lambda\Big) = f\Big(\frac{t}{T}, \lambda\Big) + \frac{\partial}{\partial x} f(x, \lambda)\Big|_{x = \xi}\Big(\frac{|k| - N}{T}(1 - u)\Big),
$$

where $\xi \in \Big(\frac{t-N}{T} + \frac{N}{T}u, \frac{t}{T}\Big)$.

Thus, by Assumption 4 and because $u \in [0, 1]$, we have

$$
\Big|f\Big(\frac{t - N + |k|}{T} + \frac{N - |k|}{T}u, \lambda\Big) - f\Big(\frac{t}{T}, \lambda\Big)\Big| \leq M'_f\Big(\frac{N - |k|}{T}\Big).
$$

Thus,

$$
\begin{aligned}
\Big|\int_{-\pi}^{\pi} &\Big[\int_0^1 \Big(f\Big(\frac{t - N + |k|}{T} + \frac{N - |k|}{T}u, \lambda\Big) - f\Big(\frac{t}{T}, \lambda\Big)\Big)\mathrm{d}u\Big]\mathrm{e}^{\mathrm{i}|k|\lambda}\mathrm{d}\lambda\Big| \\
&\leq \int_{-\pi}^{\pi}\int_0^1 \Big|f\Big(\frac{t - N + |k|}{T} + \frac{N - |k|}{T}u, \lambda\Big) - f\Big(\frac{t}{T}, \lambda\Big)\Big|\mathrm{d}u\mathrm{d}\lambda \leq 2\pi M'_f\Big(\frac{N - |k|}{T}\Big),
\end{aligned}
$$

which completes the proof of the lemma. $\qquad\square$

Corollary E.2 *Under the conditions of Lemma E.1, with $\tilde{\Gamma}_T^{(p)}(t)$ and $\tilde{\gamma}_T^{(p)}(t)$ defined in (34), $\Gamma_\Delta^{(p)}(u)$ and $\gamma_\Delta^{(p)}(u)$ defined in (52), $\hat{\Gamma}_{N,T}^{(p)}(t)$ and $\hat{\gamma}_{N,T}^{(p)}(t)$ defined in (2), we have:*
(i)

$$\|\tilde{\Gamma}_T^{(p)}(t) - \Gamma_0^{(p)}(t/T)\| \leq \frac{p^2}{T}(2\pi M_f' + C) \qquad \|\tilde{\gamma}_T^{(p)}(t) - \gamma_0^{(p)}(t/T)\| \leq \frac{p^{1/2}}{T}C$$

and (ii)

$$\|\mathbb{E}\hat{\Gamma}_{N,T}^{(p)}(t) - \Gamma_0^{(p)}(t/T)\| \leq \frac{p}{T}(2\pi M_f'(N+1) + C)$$

$$\|\mathbb{E}\hat{\gamma}_{N,T}^{(p)}(t) - \gamma_0^{(p)}(t/T)\| \leq \frac{p^{1/2}}{T}(2\pi M_f'N + C)$$

and (iii)

$$\|\mathbb{E}\hat{\Gamma}_{N,T}^{(p)}(t) - \Gamma_{N/T}^{(p)}(t/T)\| \leq \frac{p^2}{T}(2\pi M_f' + C)$$

$$\|\mathbb{E}\hat{\gamma}_{N,T}^{(p)}(t) - \gamma_{N/T}^{(p)}(t/T)\| \leq 2\frac{p^{3/2}}{T}\Big(2\pi M_f' + C\Big)$$

Proof of Corollary E.2.

For (i), note that, letting

$$\tilde{\Gamma}_T^{(p)}(t) - \Gamma_0^{(p)}(t/T) =: \big(\delta_{j,k}(t,T)\big)_{j,k=1,\ldots,p}$$

$$\tilde{\gamma}_T^{(p)}(t) - \gamma_0^{(p)}(t/T) =: \big(\eta_j(t,T)\big)_{j=1,\ldots,p}$$

we have

$$|\delta_{j,k}(t,T)| = \Big|\tilde{\gamma}_{k-j,T}(t-j) - \gamma_{j-k}\big(\frac{t-j}{T}\big) + \gamma_{j-k}\big(\frac{t-j}{T}\big) - \gamma_{j-k}\big(\frac{t}{T}\big)\Big|$$

$$\leq C/T + \Big|\int_{-\pi}^{\pi} \exp\big(\mathrm{i}(j-k)\lambda\big)\Big[f\big(\frac{t-j}{T},\lambda\big) - f\big(\frac{t}{T},\lambda\big)\Big]\mathrm{d}\lambda\Big|$$

$$\leq (C + 2\pi M_f'j)/T$$

where the first part of the first inequality is due to (37) and the second part is due to (41) and the mean value theorem.

We bound the norms of $\Delta := \big(\delta_{j,k}(t,T)\big)_{j,k=1,\ldots,p}$

$$\|\Delta\|_1 = \max_{1\leq k\leq p}\sum_{j=1}^{p}|\delta_{j,k}| \leq \max_{1\leq k\leq p}\sum_{j=1}^{p}\frac{C + 2\pi M_f'j}{T} = T^{-1}p(C + \pi M_f'(p+1))$$

$$\|\Delta\|_\infty = \max_{1\leq j\leq p}\sum_{k=1}^{p}|\delta_{j,k}| \leq \max_{1\leq j\leq p}\sum_{k=1}^{p}\frac{C + 2\pi M_f'j}{T} = p\frac{C + 2\pi M_f'p}{T}$$

Then, we have, by Hölder's inequality

$$\|\Delta\|_2 \leq \Big(\|\Delta\|_1\|\Delta\|_\infty\Big)^{1/2}$$

$$\leq T^{-1}p\Big((C + \pi M_f'(p+1))(C + 2\pi M_f'p)\Big)^{1/2}$$

$$\leq T^{-1}p(C + 2\pi M_f'p) \leq T^{-1}p^2(C + 2\pi M_f')$$

Further, note that

$$|\eta_j(t,T)| = \left| \tilde{\gamma}_{j,T}(t) - \gamma_j\left(\frac{t}{T}\right) \right| \le C/T$$

and we get the second part of (i) by bounding the Euclidean norm by $p^{1/2}$ times the maximum norm.

For (ii) note that, by Lemma E.1(iii), we have

$$\mathbb{E}\hat{\Gamma}_{N,T}^{(p)}(t) - \Gamma_0^{(p)}(t/T) =: \left(\delta_{j-k}(t,T)\right)_{j,k=1,\dots,p} =: \Delta$$

$$\mathbb{E}\hat{\gamma}_{N,T}^{(p)}(t) - \gamma_0^{(p)}(t/T) =: \left(\delta_j(t,T)\right)_{j=0,\dots,p-1} =: \delta$$

with

$$|\delta_k(t,T)| \le \left| \frac{2\pi M_f'(N - |k| + 1) + C}{T} \right|.$$

We bound the spectral norm of Δ by its Frobenius norm:

$$\|\Delta\|_2^2 \le \|\Delta\|_F^2 = \sum_{i=1}^{p}\sum_{j=1}^{p} |\delta_{i-j}|^2 \le \left| \frac{2\pi M_f'(N+1) + C}{T} \right|^2 p^2$$

For the second part of (ii) note, that

$$\left\| \mathbb{E}\hat{\gamma}_{N,T}^{(p)}(t) - \gamma_0^{(p)}(t/T) \right\| = \left(\sum_{j=1}^{p} |\delta_j(t,T)| \right)^{1/2}$$

$$\le \left(\sum_{j=1}^{p} \left| \frac{2\pi M_f'N + C}{T} \right|^2 \right)^{1/2} \le p^{1/2} \left| \frac{2\pi M_f'N + C}{T} \right|$$

For the proof of (iii) note that, letting

$$\mathbb{E}\Gamma_{N,T}^{(p)}(t) - \Gamma_{N/T}^{(p)}(t/T) =: \left(\delta_{j,k}(t,T)\right)_{j,k=1,\dots,p}$$

$$\mathbb{E}\gamma_{N,T}^{(p)}(t) - \gamma_{N/T}^{(p)}(t/T) =: \left(\eta_j(t,T)\right)_{j=1,\dots,p}$$

we have

$$\|\Delta\|_1 = \|\Delta\|_\infty = \max_{1 \le \ell \le p} \sum_{h=1}^{p} |\delta_{h-\ell}| \le \sum_{j=0}^{p-1} \frac{2\pi(j+1)M_f' + C}{T}$$

$$= \frac{p(C + \pi M_f'(p+1))}{T} \le \frac{p^2}{T}(C + 2\pi M_f'),$$

where the first inequality is due to Lemma E.1(ii) and for the second inequality we have used $\pi(p+1) \le 2\pi p$ and $Ap^2 + Bp \le (A+B)p^2$ which holds as $A, B \ge 0$, and $p \ge 1$. By Hölder's inequality this implies

$$\|\Delta\|_2 \le \left(\|\Delta\|_1 \|\Delta\|_\infty \right)^{1/2} \le \frac{p^2}{T}(C + 2\pi M_f').$$

71

For the other bound note

$$\Big(\sum_{j=1}^{p}(\frac{2\pi(j+1)M'_f+C}{T})^2\Big)^{1/2} \le \frac{1}{T}\Big(2\pi M'_f(\sum_{j=1}^{p}(j+1)^2)^{1/2}+p^{1/2}C\Big)$$

$$\le \frac{1}{T}\Big(2\pi M'_f(p(p+1)^2)^{1/2}+p^{1/2}C\Big) \le \frac{1}{T}\Big(4\pi M'_f p^{3/2}+p^{1/2}C\Big).$$

□

Lemma E.3 *Let* $(X_{t,T})_{t\in\mathbb{Z},T\in\mathbb{N}^*}$ *satisfy Assumptions 1, 3, and 4 (Appendix A) and* $\mathbb{E}X_{t,T}=0$. *Then, we have:*

(i-a) the matrices $\Gamma^{(p)}(u)$ *and* $\Gamma_{\Delta}^{(p)}(u)$ *are positive definite, hence invertible, for* $u\in\mathbb{R}$ *and* $\Delta\ge 0$, *with their eigenvalues between* m_f *and* M_f. *In other words, the norms of the matrices and their inverses are uniformly bounded:*

$$m_f \le 1/\|\Gamma_{\Delta}^{(p)}(u)^{-1}\| \le \|\Gamma_{\Delta}^{(p)}(u)\| \le M_f.$$

(i-b) the norms of the respective vectors are uniformly bounded:

$$\|\gamma_{\Delta}^{(p)}(u)\| \le (2\pi)^{1/2}M_f.$$

(ii-a) The largest eigenvalue of $\mathbb{E}\hat\Gamma_{N,T}^{(p)}(t)$ *satisfies the following bound:*

$$\|\mathbb{E}\hat\Gamma_{N,T}^{(p)}(t)\| \le M_f + \frac{p^2}{T}(2\pi M'_f+C).$$

(ii-b) if $T > m_f^{-1}p^2(2\pi M'_f+C)$, *then the matrix* $\mathbb{E}\hat\Gamma_{N,T}^{(p)}(t)$ *is positive definite, and the smallest eigenvalue satisfies the following bound:*

$$m_f - \frac{p^2}{T}(2\pi M'_f+C) \le \frac{1}{\|\big(\mathbb{E}\hat\Gamma_{N,T}^{(p)}(t)\big)^{-1}\|}$$

(ii-c) in particular, if $T \ge 2m_f^{-1}p^2(2\pi M'_f+C)$ *we thus have*

$$\frac{1}{2}m_f \le \frac{1}{\|\big(\mathbb{E}\hat\Gamma_{N,T}^{(p)}(t)\big)^{-1}\|} \le \|\mathbb{E}\hat\Gamma_{N,T}^{(p)}(t)\| \le \frac{3}{2}M_f.$$

Proof of Lemma E.3. We use the notation from Gray (2009). For every function f defined on $[0,2\pi]$, with a Fourier series that has absolutely summable Fourier coefficients and for every $n\in\mathbb{N}^*$, he defines, in (4.8), the $p\times p$ Toeplitz matrix

$$T_n(p) := \left[\int_{-\pi}^{\pi}f(\lambda)e^{-i(k-j)\lambda}d\lambda \;:\; k,j=0,1,\ldots,p-1\right].$$

We will apply Lemma 4.1 from Gray (2009) according to which, if $T_p(f)$ is Hermitian (which is the case if and only f is real valued), its eigenvalues lie between the essential infimum and the essential supremum of f.

Letting $g_\Delta(u,\lambda) := \int_0^1 f\big(u - \Delta(x-1), \lambda\big)\mathrm{d}x$, we have $\Gamma_\Delta^{(p)}(u) = T_p\big(g_\Delta(u,\cdot)\big)$. The proof of (i-a) is thus completed, as $g_\Delta(u, \mathbb{R}) \subset [m_f, M_f]$, by Assumption 3. (i-b) follows from Bessel's inequality, as we have

$$\|\gamma_\Delta^{(p)}(u)\|^2 = \sum_{\ell=1}^p \Big| \int_{-\pi}^\pi \mathrm{e}^{\mathrm{i}\ell\lambda} g_\Delta(u,\lambda)\mathrm{d}\lambda \Big|^2$$

$$\leq \int_{-\pi}^\pi \big|g_\Delta(u,\lambda)\mathrm{d}\lambda\big|^2 \leq \int_{-\pi}^\pi \big|f(u,\lambda)\mathrm{d}\lambda\big|^2 \leq 2\pi M_f^2.$$

For the proof of (ii-a) apply the triangle inequality:

$$\|\mathbb{E}\hat{\Gamma}_{N,T}^{(p)}(t)\| \leq \|\Gamma_{N/T}^{(p)}(t/T)\| + \|\mathbb{E}\hat{\Gamma}_{N,T}^{(p)}(t) - \Gamma_{N/T}^{(p)}(t/T)\|$$

and note that the first term on the right hand side can be bounded by (i-a) of this lemma and the second part can be bounded by Corollary E.2(iii).

For the proof of (ii-b) we employ Lemma F.1. Denote $A := \Gamma_{N/T}^{(p)}(t/T)$ and note that, by Corollary E.2(iii) we have that $\Delta := \mathbb{E}\Gamma_{N,T}^{(p)}(t) - \Gamma_{N/T}^{(p)}(t/T)$ satisfies $\|\Delta\| \leq \frac{p^2}{T}(C + 2\pi M_f')$. Thus, by assuming $T > m_f^{-1} p^2 (C + 2\pi M_f')$ we have

$$\|A^{-1}\|\|\Delta\| \leq m_f^{-1} \frac{p^2}{T}(C + 2\pi M_f') < 1.$$

Under this condition Lemma F.1 asserts that $\mathbb{E}\Gamma_{N,T}^{(p)}(t)$ is invertible with

$$\frac{1}{\|\mathbb{E}\Gamma_{N,T}^{(p)}(t)^{-1}\|} \geq \frac{1 - \|A^{-1}\| \cdot \|\Delta\|}{\|A^{-1}\|} = \frac{1}{\|A^{-1}\|} - \|\Delta\| \geq m_f - \frac{p^2}{T}(2\pi M_f' + C).$$

Finally, because $2m_f^{-1} p^2 (2\pi M_f' + C) \geq 2M_f^{-1} p^2 (2\pi M_f' + C)$ we see that (ii-c), the "in particular", holds. $\qquad\square$

E.2 Exponential inequalities for empirical covariances

We now state an exponential inequalities for the empirical covariances:

Lemma E.4 *Let $(X_{t,T})_{t\in\mathbb{Z}, T\in\mathbb{N}^*}$ satisfy Assumptions 1, 2, 3, and 5 (Appendix A) and $\mathbb{E}X_{t,T} = 0$. Then, for $T \geq C/(\pi m_f)$, $n \in \mathbb{N}^*$, $h \in \mathbb{N}$ and $\varepsilon > 0$, we have*

$$\mathbb{P}\Big(\big|\hat{\gamma}_{h;N,T}(t) - \mathbb{E}\hat{\gamma}_{h;N,T}(t)\big| \geq \varepsilon\Big) \leq \exp\Bigg(- \frac{\varepsilon^2}{2\Big(\frac{C_{1,1}h_*}{N-|h|} + \varepsilon^{(3+4d)/(2+2d)}\Big(\frac{C_{2,1}h_*}{N-|h|}\Big)^{1/(2+2d)}\Big)}\Bigg)$$

$$\leq \begin{cases} \exp\Big(-\frac{\varepsilon^2}{4}\frac{N-|h|}{C_{1,1}h_*}\Big) & \varepsilon \leq \Big(\frac{h_*}{N-|h|}\Big)^{(1+2d)/(3+4d)} \Big(\frac{C_{1,1}^{2+2d}}{C_{2,1}}\Big)^{1/(3+4d)} \\ \exp\Big(-\frac{1}{4}\big(\varepsilon\frac{N-|h|}{C_{2,1}h_*}\big)^{1/(2+2d)}\Big) & \varepsilon \geq \Big(\frac{h_*}{N-|h|}\Big)^{(1+2d)/(3+4d)} \Big(\frac{C_{1,1}^{2+2d}}{C_{2,1}}\Big)^{1/(3+4d)} \end{cases},$$

where $h_* := |h| + I\{h = 0\}$, $\hat{\gamma}_{h;N,T}(t)$ is defined in (3), and the constants $C_{1,1}$ and $C_{2,1}$ are defined in (85).

Note that the right hand side does not depend on t.

Proof of Lemma E.4. We state and prove a more general version, which entails Lemma E.4 as a special case (let $a_t := 1$, $\alpha = 1$, $b := t - N + 1$, and $n := N - |h|$).

Lemma E.5 *Let $(X_{t,T})_{t \in \mathbb{Z}, T \in \mathbb{N}^*}$ satisfy Assumptions 1, 2, 3, and 5 (Appendix A) and $\mathbb{E} X_{t,T} = 0$. Let a_t, $t = b, \ldots, b + n - 1$ be a bounded sequence of numbers; i. e., $|a_t| \leq A$. Then, for $\alpha \in \mathbb{N}^*$, $T \geq C/(\pi m_f)$, $n \in \mathbb{N}^*$, $b \in \mathbb{Z}$, $h \in \mathbb{N}$ and $\varepsilon > 0$, we have*

$$
\mathbb{P}\Big(\Big| n^{-1} \sum_{t=b}^{b+n-1} a_t (X_{t,T}^\alpha X_{t+h,T}^\alpha - \mathbb{E}(X_{t,T}^\alpha X_{t+h,T}^\alpha)) \Big| > \varepsilon \Big)
$$

$$
\leq \exp\left(- \frac{\varepsilon^2}{2\Big(\frac{C_{1,\alpha} A^2 h_*}{n} + \varepsilon^{(3+4\alpha d)/(2+2\alpha d)} \Big(\frac{C_{2,\alpha} A h_*}{n} \Big)^{1/(2+2\alpha d)} \Big)} \right)
$$

$$
\leq \begin{cases} \exp\Big(- \dfrac{(\varepsilon/A)^2}{4} \dfrac{n}{C_{1,\alpha} h_*} \Big) & \varepsilon \leq A\Big(\dfrac{C_{1,\alpha}^{2+2\alpha d}}{C_{2,\alpha}} \Big)^{1/(3+4\alpha d)} \Big(\dfrac{h_*}{n} \Big)^{(1+2\alpha d)/(3+4\alpha d)} \\[3mm] \exp\Big(- \dfrac{1}{4}\Big(\dfrac{\varepsilon}{A} \dfrac{n}{C_{2,\alpha} h_*} \Big)^{1/(2+2\alpha d)} \Big) & \varepsilon \geq A\Big(\dfrac{C_{1,\alpha}^{2+2\alpha d}}{C_{2,\alpha}} \Big)^{1/(3+4\alpha d)} \Big(\dfrac{h_*}{n} \Big)^{(1+2\alpha d)/(3+4\alpha d)} \end{cases},
$$

where $h_* := |h| + I\{h = 0\}$ and the constants $C_{1,\alpha}$ and $C_{2,\alpha}$ are defined in (85) in the proof [depending only on α, d, C, M_f, ρ, and K].

Note the important fact that the bounds in the inequality do not depend on b.

Proof of Lemma E.5.

We proceed along the lines of the proof of Lemma 2 in Goldenshluger and Zeevi (2001) and apply the two results from Saulis and Statulevičus (1991) which are cite in Section G. As Goldenshluger and Zeevi (2001) pointed out, several other exponential-type inequalities are available in the literature; e. g., Doukhan (1994) and Bosq (1998). These result could be used to derive the type of bound we need, but under our conditions the results from Saulis and Statulevičus (1991) seem most suitable here.

The assertion of Lemma E.5 follows from Lemma G.1 if we show that (93) holds with $\gamma := 1 + 2\alpha d$,

$$
\xi := \sum_{t=1}^{n} n^{-1} a_t \big(X_{t+b-1,T}^\alpha X_{t+b-1+h,T}^\alpha - \mathbb{E}(X_{t+b-1,T}^\alpha X_{t+b-1+h}^\alpha) \big),
$$

and

$$
H := \frac{C_{1,\alpha} A^2 h_*}{n}, \quad \bar{\Delta} := \frac{n}{C_{2,\alpha} A h_*},
$$

where $h_* := |h| + I\{h = 0\}$, and

$$
\begin{aligned}
C_{1,\alpha} &:= 12 \cdot 2^{8\alpha d + 9} (\alpha!)^{4d} \big(\max\{c^2, 3\pi M_f\} \big)^{2\alpha} \mathrm{e}\Big(1 + \frac{1}{\log \rho} \Big)\Big(1 + K^{1/2} \Big), \\
C_{2,\alpha} &:= 12 \cdot 2^{3(\alpha d + 1)} (\alpha!)^{2d} \big(\max\{c^2, 3\pi M_f\} \big)^{\alpha} \mathrm{e}\Big(1 + \frac{1}{\log \rho} \Big).
\end{aligned} \tag{85}
$$

We show this by employing Lemma G.2 with

$$Y_{\ell,T} := n^{-1} a_t \big(X^\alpha_{\ell+b-1,T} X^\alpha_{\ell+b-1+h,T} - \mathbb{E}(X^\alpha_{\ell+b-1,T} X^\alpha_{\ell+b-1+h,T}) \big),$$

$\ell = 1, \ldots, n$. The first step in the application of Lemma G.2 is to show that (94) holds with appropriately chosen H_1 and γ_1. The second step will be to show that the bound we get from Lemma G.2 can again be bounded by the right hand side in (93) with H, $\bar{\Delta}$, and γ as defined before.

For the first step, observe that for any $\ell = 1, \ldots, n$ and $k = 2, 3, \ldots$ we have

$$
\begin{aligned}
&n^k \mathbb{E}|Y_{\ell,T}|^k \\
&= |a_t|^k \mathbb{E}\big| X^\alpha_{\ell+b-1,T} X^\alpha_{\ell+b-1+h,T} - \mathbb{E}(X^\alpha_{\ell+b-1,T} X^\alpha_{\ell+b-1+h,T}) \big|^k \\
&\leq A^k 2^{k-1} \mathbb{E}\Big(|X^\alpha_{\ell+b-1,T} X^\alpha_{\ell+b-1+h,T}|^k + \big|(\mathbb{E}X^{2\alpha}_{\ell+b-1,T})(\mathbb{E}X^{2\alpha}_{\ell+b-1+h,T})\big|^{k/2} \Big) \\
&\leq A^k 2^k (k!)^{2\alpha d} \big(((2\alpha)!)^d \max\{c^{2\alpha}, (3\pi M_f)^\alpha\} \big)^k.
\end{aligned}
\tag{86}
$$

For the first part of (86) we have used that

$$
\begin{aligned}
\mathbb{E}|X^\alpha_{\ell+b-1,T} X^\alpha_{t-s-1+h,T}|^k &= \mathbb{E}(|X_{\ell+b-1,T}|^{\alpha k} |X_{t-s-1+h,T}|^{\alpha k}) \\
&\leq (\mathbb{E}|X_{\ell+b-1,T}|^{2\alpha k})^{1/2} (\mathbb{E}|X_{\ell+b-1+h,T}|^{2\alpha k})^{1/2} \\
&\leq \big(c^{2\alpha k-2} ((2\alpha k)!)^d \sigma^2_{\ell+b-1,T} c^{2\alpha k-2}((2\alpha k)!)^d \sigma^2_{\ell+b-1+h,T} \big)^{1/2} \\
&\leq c^{2\alpha k-2}((2\alpha)!)^k)^d (k!)^{2\alpha d}(3\pi M_f) \leq (k!)^{2\alpha d}\big(((2\alpha)!)^d (\max\{c,(3\pi M_f)^{1/2}\})^{2\alpha}\big)^k.
\end{aligned}
$$

In the above we have used Cauchy-Schwarz inequality, Assumption 5, Assumption 3, which, if $T \geq C/(\pi m_f)$, implies (40), and the elementary inequality $(Lk)! \leq (L!)^k (k!)^L$, which holds for $L, k \in \mathbb{N}^*$. For the second part of (86) we have used that

$$
\begin{aligned}
\big|(\mathbb{E}X^{2\alpha}_{\ell+b-1})(\mathbb{E}X^{2\alpha}_{t-s-1+h,T})\big|^{k/2} &\leq \big| c^{2\alpha-2}((2\alpha)!)^d \sigma^2_{\ell+b-1,T} c^{2\alpha-2}((2\alpha)!)^d \sigma^2_{\ell+b-1+h,T} \big|^{k/2} \\
&\leq \big(c^{2\alpha-2}((2\alpha)!)^d (3\pi M_f) \big)^k \leq (k!)^{2\alpha d}\big(((2\alpha)!)^d (\max\{c,(3\pi M_f)^{1/2}\})^{2\alpha}\big)^k,
\end{aligned}
$$

where we have employed arguments as before and $1 \leq (k!)^{2\alpha d}$.

Thus we have shown that (94) holds, with $H_1 := 2A((2\alpha)!)^d \max\{c^{2\alpha}, (3\pi M_f)^\alpha\}/n$ and $\gamma_1 := 2\alpha d - 1$.

For the second step we turn our attention to $\Delta_n(\nu)$ defined in Lemma G.2. Note that $\sigma(Y_{\ell,T} : 1 \leq \ell \leq s) \subseteq \sigma(X_{\ell,T} : \ell \leq s - b)$ and $\sigma(Y_{\ell,T} : \ell \geq t) \subseteq \sigma(X_{\ell,T} : \ell \geq t - b - |h|)$, which implies that

$$
\begin{aligned}
\alpha^Y(s,t) &:= \sup_{A \in \sigma(Y_{\ell,T}:1\leq\ell\leq s)} \sup_{B \in \sigma(Y_{ell,T}:u\geq t)} \big| \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \big| \\
&\leq \sup_{A \in \sigma(X_{\ell,T}:\ell\leq s-b)} \sup_{B \in \sigma(X_{\ell,T}:\ell\geq t-b-|h|)} \big| \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \big| \\
&=: \alpha_T(s-b, t-b-|h|) \leq \alpha(t-s-|h|),
\end{aligned}
$$

where α_T and α are the functions from Assumption 2. Further, note that the trivial inequality $\alpha(t-s-|h|) \leq 1$ holds for $t \leq s + |h|$. Thus, adopting the argument from

step 3 in the proof of Lemma 2 in Goldenshluger and Zeevi (2001), we have

$$\sum_{t=s}^{n}\alpha^Y(s,t)^{1/\nu} \le \sum_{t=s}^{n}\alpha(t-s-|h|)^{1/\nu} \le \sum_{t=s}^{s+|h|}1+\sum_{t=s+|h|+1}^{n}\alpha(t-s-|h|)^{1/\nu}$$

$$\le |h|+\sum_{t=1}^{n-s-|h|}K^{1/\nu}\rho^{-t/\nu} \le |h|+K^{1/\nu}\sum_{t=0}^{\infty}\rho^{-t/\nu}$$

$$= |h|+K^{1/\nu}\frac{\rho^{1/\nu}}{\rho^{1/\nu}-1} \le |h|+K^{1/\nu}\Big(1+\frac{\nu}{\log\rho}\Big),$$

where the last inequality follows from $\exp(x)-1\ge x$, for $x\ge 0$.

Thus, for $\delta=1$ and $k=2,3,\ldots$, we have

$$\Delta_n((1+1/\delta)(k-1))^{k-1} \le \Big(|h|+K^{\frac{1}{2(k-1)}}\Big(1+\frac{2(k-1)}{\log\rho}\Big)\Big)^{k-1}$$

$$\le 2^{k-2}\Big(|h|^{k-1}+K^{1/2}\Big(1+\frac{2(k-1)}{\log\rho}\Big)^{k-1}\Big)$$

$$\le 2^{k-2}\Big(h_*^{k-1}+K^{1/2}\Big(1+\frac{2}{\log\rho}\Big)^{k-1}(k-1)^{k-1}\Big)$$

$$\le 2^{k-2}\Big(h_*^{k-1}+K^{1/2}\Big(1+\frac{2}{\log\rho}\Big)^{k-1}(k-1)!e^{k-1}\Big)$$

$$\le 2^{k-2}\Big(h_*^{k-1}\Big(1+\frac{2}{\log\rho}\Big)^{k-1}(k-1)!e^{k-1}\Big)\Big(1+K^{1/2}\Big)$$

$$\le \Big(4\,e\,h_*\Big(1+\frac{1}{\log\rho}\Big)\Big)^{k-1}k!\Big(1+K^{1/2}\Big).$$

where we have used the fact that $p^p\le p!\,e^p$, for $p\in\mathbb{N}$.

Recall that $\gamma:=1+2\alpha d$ and $\gamma_1:=2\alpha d-1$. Thus, by Lemma G.1, we have that

$$\Big|\text{cum}_k\Big(\sum_{t=1}^{n}Y_t\Big)\Big|$$

$$\le 2(k!)^{2+2\alpha d-1}12^{k-1}\big(H_1\big)^k 2^{2\alpha dk}\Big(4\,e\,h_*\Big(1+\frac{1}{\log\rho}\Big)\Big)^{k-1}k!\Big(1+K^{1/2}\Big)n$$

$$= 12^{k-1}2^{1+2\alpha dk+2+2\alpha d+2(k-1)}(k!/2)^{2+2\alpha d}\big(nH_1\big)^k\Big(e\,h_*\Big(1+\frac{1}{\log\rho}\Big)/n\Big)^{k-1}\Big(1+K^{1/2}\Big)$$

$$= 12\cdot 12^{k-2}2^{2(\alpha d+1)k+2\alpha d+1}(k!/2)^{2+2\alpha d}\big(nH_1\big)^k\Big(e\,h_*\Big(1+\frac{1}{\log\rho}\Big)/n\Big)^{k-1}\Big(1+K^{1/2}\Big)$$

$$= \Big(\frac{k!}{2}\Big)^{1+\gamma}\frac{12\cdot 2^{4(\alpha d+1)+2\alpha d+1}\big(nH_1\big)^2 e\,h_*\Big(1+\frac{1}{\log\rho}\Big)\Big(1+K^{1/2}\Big)/n}{\Big(n12^{-1}2^{-2(\alpha d+1)}\big(nH_1\big)^{-1}\Big(e\,h_*\Big(1+\frac{1}{\log\rho}\Big)\Big)^{-1}\Big)^{k-2}} \le \Big(\frac{k!}{2}\Big)^{1+\gamma}\frac{H}{\overline{\Delta}^{k-2}},$$

where we have used the fact that $nH_1\le 2^{\alpha d+1}A(\alpha!)^{2d}\max\{c^{2\alpha},(3\pi M_f)^{\alpha}\}$. Applying Lemma 2.4 from Saulis and Statulevičus (1991) yields the expression in the lemma. $\square$

F   Technical Results

In the proofs we use the following general results, which are not restricted to locally stationary processes.

Lemma F.1 *Let $A \in \mathbb{R}^{p \times p}$ an invertible matrix and $\Delta \in \mathbb{R}^{p \times p}$ be a matrix with $\|A^{-1}\|_M \cdot \|\Delta\|_M < 1$ for a submultiplicative matrix norm $\|\cdot\|_M$. Then, the matrix $A + \Delta$ is invertible and we have*

$$\|(A + \Delta)^{-1}\|_M \le \frac{\|A^{-1}\|_M}{1 - \|A^{-1}\|_M \cdot \|\Delta\|_M}$$

Proof of Lemma F.1. Let $X := I - A^{-1}(A + \Delta) = -A^{-1}\Delta$. Then, because $\|X\|_M \le \|A^{-1}\|_M \|\Delta\|_M < 1$, Lemma 2.1 in Demmel (1997) can be applied, which asserts that $I - X = A^{-1}(A + \Delta)$ is invertible and that

$$\|(I - X)^{-1}\|_M \le \frac{1}{1 - \|X\|_M}.$$

We deduce that $A + \Delta$ is invertible, because we have

$$\det(A + \Delta) = \det(A)\det(I - X) \ne 0,$$

where we have used that A is invertible by assumption and $I - X$ is invertible by Lemma 2.1 in Demmel (1997). Finally, the bound on the matrix norm of the inverse follows from the assumed submultiplicativity

$$\|(A + \Delta)^{-1}\|_M = \|(I - X)^{-1}A^{-1}\|_M \le \|(I - X)^{-1}\|_M \cdot \|A^{-1}\|_M$$
$$\le \frac{\|A^{-1}\|_M}{1 - \|A^{-1}\Delta\|_M} \le \frac{\|A^{-1}\|_M}{1 - \|A^{-1}\|_M \|\Delta\|_M}.$$

This finishes the proof of the lemma. $\qquad\square$

An important corollary to the above lemma is the following:

Lemma F.2 *Let $A \in \mathbb{R}^{p \times p}$ be an invertible matrix and $\Delta \in \mathbb{R}^{p \times p}$ be a matrix with $\|A^{-1}\|_M \cdot \|\Delta\|_M \le c < 1$ for a submultiplicative matrix norm $\|\cdot\|_M$. Then, the matrix $A + \Delta$ is invertible and we have*

$$\|(A + \Delta)^{-1} - A^{-1}\|_M \le \|\Delta\|_M \frac{\|A^{-1}\|_M^2}{1 - c}.$$

Proof of Lemma F.2. Note that Lemma F.1 can be applied, which yields that $A + \Delta$ is invertible as an immediate consequence. Further, note that

$$(A + \Delta)^{-1} - A^{-1} = (A + \Delta)^{-1}\big(A - (A + \Delta)\big)A^{-1}.$$

Employing the submultiplicativity of the norm and the inequality from Lemma F.1 yields

$$\|(A + \Delta)^{-1} - A^{-1}\|_M \le \frac{\|A^{-1}\|_M}{1 - \|A^{-1}\|_M \cdot \|\Delta\|_M} \cdot \| - \Delta\|_M \cdot \|A^{-1}\|_M.$$

The assertion then follows, because we have $\| - \Delta\|_M = \|\Delta\|_M$ and, by assumption, $\|A^{-1}\|_M \cdot \|\Delta\|_M \le c$ holds. $\qquad\square$

Lemma F.3 *Let A and A_0 be two square matrices and $\|\cdot\|_M$ be a submultiplicative matrix norm. Then, for any $h \in \mathbb{N}$,*

$$\|A^h - A_0^h\|_M \leq h\|A - A_0\|_M \big(\|A - A_0\|_M + \|A_0\|_M\big)^{h-1}.$$

Proof of Lemma F.3. The statement for $h \in \{0,1\}$ is obvious. For $h = 2, 3, \dots$ note that by the binomial theorem we have

$$A^h = \big((A - A_0) + A_0\big)^h = \sum_{\substack{(i_1, \dots, i_h) \\ \in \{0,1\}^h}} \prod_{\ell=1}^{h} (A - A_0)^{1 - i_\ell} A_0^{i_\ell}.$$

This obviously implies

$$
\begin{aligned}
\|A^h - A_0^h\|_M &= \Big\| \sum_{\substack{(i_1, \dots, i_h) \\ \in \{0,1\}^h \setminus \{(1, \dots, 1)\}}} \prod_{\ell=1}^{h} (A - A_0)^{1 - i_\ell} A_0^{i_\ell} \Big\|_M \\
&\leq \sum_{j=0}^{h-1} \binom{h}{j+1} \|A - A_0\|_M^{j+1} \|A_0\|_M^{h-j-1} \\
&= \sum_{j=0}^{h-1} \frac{h}{j+1} \frac{(h-1)!}{j!(h-1-j)!} \|A - A_0\|_M^{j+1} \|A_0\|_M^{h-j-1} \\
&\leq h\|A - A_0\|_M \sum_{j=0}^{h-1} \binom{h-1}{j} \|A - A_0\|_M^{j} \|A_0\|_M^{h-1-j} \\
&= h\|A - A_0\|_M \big(\|A - A_0\|_M + \|A_0\|_M\big)^{h-1}.
\end{aligned}
$$

Lemma F.4 *Let u and v be two real-valued random variables. Further, let u_0 and v_0 be two real numbers. Then, for all $\varepsilon > 0$*

$$\mathbb{P}(|uv - u_0 v_0| > \varepsilon) \leq \mathbb{P}\Big(|u - u_0| > \frac{1}{2} \frac{\varepsilon}{(|v_0|^2 + \varepsilon)^{1/2}}\Big) + \mathbb{P}\Big(|v - v_0| > \frac{1}{2} \frac{\varepsilon}{(|u_0|^2 + \varepsilon)^{1/2}}\Big).$$

Proof of Lemma F.4. Note that

$$
\begin{aligned}
\mathbb{P}(|uv - u_0 v_0| > \varepsilon) &= \mathbb{P}(|(u - u_0)(v - v_0) + u_0(v - v_0) + (u - u_0)v_0| > \varepsilon) \\
&\leq \mathbb{P}(|u - u_0||v - v_0| + |u_0||v - v_0| + |u - u_0||v_0| > \varepsilon) \\
&\leq \mathbb{P}(\tfrac{1}{2}|u - u_0|^2 + \tfrac{1}{2}|v - v_0|^2 + |u_0||v - v_0| + |v_0||u - u_0| > \varepsilon) \\
&\leq \mathbb{P}(|u - u_0|^2 + 2|v_0||u - u_0| > \varepsilon) + \mathbb{P}(|v - v_0|^2 + 2|u_0||v - v_0| > \varepsilon),
\end{aligned}
$$

where the first inequality is due to the triangle inequality and the second is due to Young's inequality. We now bound the first term and note that the second term can be handled analogously. Note that

$$\mathbb{P}(|u - u_0|^2 + 2|v_0||u - u_0| > \varepsilon) = \mathbb{P}(|u - u_0| > \sqrt{\varepsilon + |v_0|^2} - |v_0|) \tag{87}$$

The assertion then follows, because

$$(|v_0|^2 + \varepsilon)^{1/2} - (|v_0|^2)^{1/2} = \frac{1}{2}\xi^{-1/2}\varepsilon \geq \frac{1}{2}(|v_0|^2 + \varepsilon)^{-1/2}\varepsilon. \tag{88}$$

The equality in the above holds due to the mean value theorem, for some $\xi \in [|v_0|^2, |v_0|^2 + \varepsilon]$. The inequality is due to $\xi \leq |v_0|^2 + \varepsilon$. $\qquad\square$

For the proof in the main part we need the following lemma:

Lemma F.5 *Let X_t and $\hat{a}_t$, $t = 1, \ldots, n$, be two sequences of random variables, and α_t, $t = 1, \ldots, n$ be a sequence of numbers. Assume that there exists a constant $m_2^2 > 0$ such that $\max_{t=1,\ldots,n} \mathbb{E}X_t^2 \leq m_2^2 < \infty$. Then, for any $\varepsilon > 0$, we have*

$$\mathbb{P}\Big(\Big|\sum_{t=1}^n \big(\hat{a}_t X_t - \alpha_t \mathbb{E}(X_t)\big)\Big| > n\varepsilon\Big)$$

$$\leq \mathbb{P}\Big(\sup_{t=1,\ldots,n} |\hat{a}_t - \alpha_t| > \frac{\varepsilon}{2\big((2m_2^2)^2 + \varepsilon^2\big)^{1/4}}\Big)$$

$$+ \mathbb{P}\Big(\Big|\sum_{t=1}^n (X_t^2 - \mathbb{E}X_t^2)\Big| > n\varepsilon/2\Big) + \mathbb{P}\Big(\Big|\sum_{t=1}^n \alpha_t(X_t - \mathbb{E}X_t)\Big| > n\varepsilon/2\Big).$$

Proof of Lemma F.5. Note that, denoting $\mu_t := \mathbb{E}X_t$, we have

$$\mathbb{P}\Big(\Big|\sum_{t=1}^n \big(\hat{a}_t X_t - \alpha_t \mathbb{E}(X_t)\big)\Big| > n\varepsilon\Big) = \mathbb{P}\Big(\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)X_t + \sum_{t=1}^n \alpha_t(X_t - \mu_t)\Big| > n\varepsilon\Big)$$

$$\leq \mathbb{P}\Big(\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)X_t\Big| + \Big|\sum_{t=1}^n \alpha_t(X_t - \mu_t)\Big| > n\varepsilon\Big)$$

$$\leq \mathbb{P}\Big(\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big|^{1/2} \Big|\sum_{t=1}^n X_t^2\Big|^{1/2} + \Big|\sum_{t=1}^n \alpha_t(X_t - \mu_t)\Big| > n\varepsilon\Big)$$

$$\leq \mathbb{P}\Big(\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big|\Big|\sum_{t=1}^n (X_t^2 - \mathbb{E}X_t^2 + \mathbb{E}X_t^2)\Big| > \big(\frac{n\varepsilon}{2}\big)^2\Big) + \mathbb{P}\Big(\Big|\sum_{t=1}^n \alpha_t(X_t - \mu_t)\Big| > \frac{n\varepsilon}{2}\Big)$$

Further, we have

$$\mathbb{P}\Big(\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big| \cdot \Big|\sum_{t=1}^n (X_t^2 - \mathbb{E}X_t^2 + \mathbb{E}X_t^2)\Big| > (n\varepsilon/2)^2\Big)$$

$$\leq \mathbb{P}\Big(\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big| \cdot \Big|\sum_{t=1}^n (X_t^2 - \mathbb{E}X_t^2)\Big| + \Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big| \cdot \Big|\mathbb{E}\sum_{t=1}^n X_t^2\Big| > (n\varepsilon/2)^2\Big)$$

$$\leq \mathbb{P}\Big(\frac{1}{2}\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big|^2 + \frac{1}{2}\Big|\sum_{t=1}^n (X_t^2 - \mathbb{E}X_t^2)\Big|^2 + \Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big| nm_2^2 > (n\varepsilon/2)^2\Big)$$

$$\leq \mathbb{P}\Big(\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big|^2 + 2nm_2^2\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big| > \frac{(n\varepsilon)^2}{4}\Big) + \mathbb{P}\Big(\Big|\sum_{t=1}^n (X_t^2 - \mathbb{E}X_t^2)\Big| > \frac{n\varepsilon}{2}\Big)$$

$$\leq \mathbb{P}\Big(\Big|\sum_{t=1}^n (\hat{a}_t - \alpha_t)^2\Big| > \sqrt{(n\varepsilon/2)^2 + (nm_2^2)^2} - nm_2^2\Big) + \mathbb{P}\Big(\Big|\sum_{t=1}^n (X_t^2 - \mathbb{E}X_t^2)\Big| > \frac{n\varepsilon}{2}\Big)$$

79

where for the second inequality we employed Young's inequality and in the fourth inequality we have used the argument in (87) from the proof of Lemma F.4. The assertion then follows, because using (88) from the proof of Lemma F.4, we have

$$\sqrt{(n\varepsilon/2)^2 + (nm_2^2)^2} - nm_2^2 \geq \frac{1}{2}(n\varepsilon/2)^2\big((nm_2^2)^2 + (n\varepsilon/2)^2\big)^{-1/2} = \frac{n}{2}2(\varepsilon/2)^2\big((2m_2^2)^2 + \varepsilon^2\big)^{-1/2},$$

which finishes the proof. $\qquad\square$

We will further use the following lemmas:

Lemma F.6 *Let $M \in \mathbb{R}^{p\times p}$ be a random $p \times p$ matrix with existing expectation $M_0 := \mathbb{E}M$, which is assumed to be invertible. Further, let v be a $\mathbb{R}^p$-valued random vector v with existing expectation $\mathbb{E}v := v_0$. Then, for every submultiplicative matrix norm $\|\cdot\|_M$ that is compatible with the (vector) norm $\|\cdot\|_v$, we have: for every $\varepsilon > 0$*

$$\mathbb{P}\Big(\big\|M^{-1}v - M_0^{-1}v_0\big\|_v > \varepsilon\Big)$$
$$\leq \mathbb{P}\Big(\|M - M_0\|_M > \frac{1}{2\|M_0^{-1}\|_M}\Big) + \mathbb{P}\Big(\|v - v_0\|_v > \frac{\varepsilon}{4}\frac{1}{\|M_0^{-1}\|_M}\Big)$$
$$+ \mathbb{P}\Big(\|M - M_0\|_M > \frac{\varepsilon}{4}\frac{1}{(\|M_0^{-1}\|_M)^2\|v_0\|_v}\Big)I\{\|v_0\|_v \neq 0\}.$$

Proof of Lemma F.6. For every $\delta \in (0, 1/\|M_0^{-1}\|_M)$, we have

$$\mathbb{P}\Big(\big\|M^{-1}v - M_0^{-1}v_0\big\|_v > \varepsilon\Big)$$
$$= \mathbb{P}\Big(\big\|M^{-1}v - M_0^{-1}v_0\big\|_v > \varepsilon \,\Big|\, \|M - M_0\|_M > \delta\Big)\mathbb{P}(\|M - M_0\|_M > \delta)$$
$$+ \mathbb{P}\Big(\big\|M^{-1}v - M_0^{-1}v_0\big\|_v > \varepsilon, \ \|M - M_0\|_M \leq \delta\Big)$$
$$\overset{(*)}{\leq} \mathbb{P}(\|M - M_0\|_M > \delta)$$
$$+ \mathbb{P}\Big(\|v - v_0\|_v > \frac{\varepsilon}{2}\frac{1 - \|M_0^{-1}\|_M\,\delta}{\|M_0^{-1}\|_M}\Big)$$
$$+ \mathbb{P}\Big(\|M - M_0\|_M\|v_0\|_v > \frac{\varepsilon}{2}\frac{1 - \|M_0^{-1}\|_M\,\delta}{(\|M_0^{-1}\|_M)^2}\Big),$$

where for $(*)$ we have used the fact that any conditional probability is ≤ 1 for the first part and the following argument for the second part:

$$\mathbb{P}\Big(\big\|M^{-1}v - M_0^{-1}v_0\big\|_v > \varepsilon, \ \|M - M_0\|_M \leq \delta\Big)$$
$$= \mathbb{P}\Big(\big\|M^{-1}v - M^{-1}v_0 + M^{-1}v_0 - M_0^{-1}v_0\big\|_v > \varepsilon, \ \|M - M_0\|_M \leq \delta\Big)$$
$$\leq \mathbb{P}\Big(\|M^{-1}\|_M\|v - v_0\|_v + \|M^{-1}\|_M\|M_0 - M\|_M\|M_0^{-1}\|_M\|v_0\|_v > \varepsilon, \ \|M - M_0\|_M \leq \delta\Big)$$
$$\leq \mathbb{P}\Big(\frac{\|M_0^{-1}\|_M}{1 - \|M_0^{-1}\|_M\,\delta}\|v - v_0\|_v + \frac{(\|M_0^{-1}\|_M)^2}{1 - \|M_0^{-1}\|_M\,\delta}\|M - M_0\|_M\|v_0\|_v > \varepsilon, \ \|M - M_0\|_M \leq \delta\Big)$$
$$\leq \mathbb{P}\Big(\|v - v_0\|_v > \frac{\varepsilon}{2}\frac{1 - \|M_0^{-1}\|_M\,\delta}{\|M_0^{-1}\|_M}, \ \|M - M_0\|_M \leq \delta\Big)$$
$$+ \mathbb{P}\Big(\|M - M_0\|_M\|v_0\|_v > \frac{\varepsilon}{2}\frac{1 - \|M_0^{-1}\|_M\,\delta}{(\|M_0^{-1}\|_M)^2}, \ \|M - M_0\|_M \leq \delta\Big)$$

In the above inequalities we have used Lemma F.1, by which we have:

$$\|M^{-1}\|_M = \|(M_0 + M - M_0)^{-1}\|_M \leq \frac{\|M_0^{-1}\|_M}{1 - \|M_0^{-1}\|_M \|M - M_0\|_M} \leq \frac{\|M_0^{-1}\|_M}{1 - \|M_0^{-1}\|_M \delta}$$

The assertion then follows by choosing $\delta = \frac{1}{2\|M_0^{-1}\|_M}$. $\hfill\square$

Lemma F.7 *Let $x = (x_1, \ldots, x_p)$ be a random vector and $x_0 = (x_{0,1}, \ldots, x_{0,p})$ be a deterministic vector. Define two $p \times p$ matrices*

$$A := \begin{pmatrix} x_1 & x_2 & \cdots & x_{p-1} & x_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \quad and \quad A_0 := \begin{pmatrix} x_{0,1} & x_{0,2} & \cdots & x_{0,p-1} & x_{0,p} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

For any $h = 1, 2, \ldots$ define $v := (1, 0, \ldots, 0)A^h$ and $v_0 := (1, 0, \ldots, 0)A_0^h$. Then, for every $\varepsilon > 0$,

$$\mathbb{P}(\|v - v_0\| > \varepsilon) \leq \mathbb{P}\left(\|x - x_0\| > 2^{1-h} \frac{\varepsilon}{\varepsilon + h(\max\{\|x_0\|, 1\})^{h-1}}\right).$$

Proof of Lemma F.7. Note that, for $h = 1$ we have $v = x$ and $v_0 = x_0$. Thus, the assertion holds as an equality. For $h = 2, 3, \ldots$, we have

$$\|v - v_0\| = \|(1, 0, \ldots, 0)(A^h - A_0^h)\| \leq \|A^h - A_0^h\|,$$

Thus, with this and by Lemma F.3, we have

$$\begin{aligned} \mathbb{P}(\|v - v_0\| > \varepsilon) &\leq \mathbb{P}(\|A^h - A_0^h\| > \varepsilon) \\ &\leq \mathbb{P}(\|A - A_0\|(\|A - A_0\| + \|A_0\|)^{h-1} > \varepsilon/h) \\ &\leq \mathbb{P}(\|x - x_0\|(\|x - x_0\| + \|x_0\|)^{h-1} > \varepsilon/h) =: (*), \end{aligned}$$

where the second inequality uses $\|A - A_0\| \leq \|A - A_0\|_F = \|x - x_0\|$ and the fact that $\|A_0\| = \|e_1 x_0'\| \leq \|x_0\|$. For $\|A_0\| = \|e_1 x_0'\|$ note that subtracting the Jordan block doesn't change the eigenvalues (see also the argument regarding the norm of X in the proof of Lemma C.2). For the rest of the derivation denote $y := \|x - x_0\|^{1/(h-1)}$. Then, we have

$$(*) = \mathbb{P}(y^h + \|x_0\|y - (\varepsilon/h)^{1/(h-1)} > 0) \leq \mathbb{P}\left(y > y_{\min}\right),$$

where $y_{\min} := \min\{|y_j| : j = 1, \ldots, h\}$ is the minimum of the absolute values of the (complex) roots $y_1, \ldots, y_h$ of the polynomial function $p(y) := y^h + \|x_0\|y - (\varepsilon/h)^{1/(h-1)}$. The inequality in the above is due to the fact that, for all $y \geq 0$, we have

$$p(y) > 0 \Rightarrow y > y_{\min}, \tag{89}$$

which we now prove. Note that p is continuous and $p(0) < 0$. For a proof by contradiction, we assume that $y \leq y_{\min}$. We find that $p(y) \leq 0$ for $y = y_{\min}$, because (a) it is either

81

a root of the polynomial and then we have $p(y) = 0$, or (b) we have $p(y) \neq 0$, in which case it has to be negative: we cannot have $p(y) > 0$, because then the intermediate value theorem would implies that $p(\xi) = 0$, for some $\xi \in [0, y_{\min})$, which is not possible, because it would have $|\xi| < y_{\min}$. In conclusion, we have shown that $p(y) \leq 0$ for all $|y| \leq y_{\min}$, which finishes the proof of (89).

By Rouché's theorem, we have

$$y_{\min} \geq \frac{(\varepsilon/h)^{1/(h-1)}}{(\varepsilon/h)^{1/(h-1)} + \max\{\|x_0\|, 1\}}.$$

This implies

$$
\begin{aligned}
\mathbb{P}(\|v - v_0\| > \varepsilon) &\leq \mathbb{P}\Big(\|x - x_0\|^{1/(h-1)} > \frac{(\varepsilon/h)^{1/(h-1)}}{(\varepsilon/h)^{1/(h-1)} + \max\{\|x_0\|, 1\}}\Big) \\
&\leq \mathbb{P}\Big(\|x - x_0\| > \frac{\varepsilon/h}{\big((\varepsilon/h)^{1/(h-1)} + \max\{\|x_0\|, 1\}\big)^{h-1}}\Big) \\
&= \mathbb{P}\Big(\|x - x_0\| > \frac{\varepsilon}{\big(\varepsilon^{1/(h-1)} + h^{1/(h-1)}\max\{\|x_0\|, 1\}\big)^{h-1}}\Big) \\
&\leq \mathbb{P}\Big(\|x - x_0\| > \frac{\varepsilon}{\max\{2^{h-2}, 1\}\big(\varepsilon + h\max\{\|x_0\|, 1\}^{h-1}\big)}\Big)
\end{aligned}
$$

Thus,

$$
\mathbb{P}(\|v - v_0\| > \varepsilon) \leq \begin{cases} \mathbb{P}\Big(\|x - x_0\| > \varepsilon\Big) & \text{if } h = 1, \\ \mathbb{P}\Big(\|x - x_0\| > 2^{2-h}\dfrac{\varepsilon}{\varepsilon + h\max\{\|x_0\|, 1\}^{h-1}}\Big) & \text{if } h = 2, 3, \ldots. \end{cases}
$$

We report a bound that is larger, to have an expression that is valid for all h. Note further, that for $\varepsilon \leq h\max\{\|x_0\|, 1\}^{h-1}$, we have

$$\mathbb{P}(\|v - v_0\| > \varepsilon) \leq \mathbb{P}\Big(\|x - x_0\| > \frac{\varepsilon}{h\max\{2\|x_0\|, 2\}^{h-1}}\Big).$$

$\square$

Lemma F.8 *Let* $f : [0, 1] \to \mathbb{R}$ *be continuous and differentiable on* $(0, 1)$. *Then, for every* $A, B = 0, \ldots, T$, $T \in \mathbb{N}_*$, $A < B$, *we have*

$$\Big|\frac{1}{B-A}\sum_{\ell=A+1}^{B} f(\ell/T) - \int_0^1 f\Big(\frac{A}{T} + \frac{B-A}{T}u\Big)\mathrm{d}u\Big| \leq \frac{1}{T}\sup_{A/T \leq u \leq B/T}|f'(u)|.$$

Proof of Lemma F.8.

By the mean value theorem and the assumed conditions we have that for every $\ell \in \mathbb{Z}$ there exists $\xi_\ell \in [(\ell-1)/T, \ell/T]$ and $\zeta_\ell \in (\xi_\ell, \ell/T)$ such that

$$
\begin{aligned}
\int_{(\ell-1)/T}^{\ell/T} f(u)\mathrm{d}u &= f(\xi_\ell)(\frac{\ell}{T} - \frac{\ell-1}{T}) = f(\ell/T)\frac{1}{T} + (f(\xi_\ell) - f(\ell/T))\frac{1}{T} \\
&= f(\ell/T)\frac{1}{T} + f'(\zeta_\ell)(\ell/T - \xi_\ell)\frac{1}{T}
\end{aligned}
$$

82

Thus, for $A, B = 0, \ldots, T$ we have

$$\int_{A/T}^{B/T} f(u)\mathrm{d}u = \sum_{\ell=A+1}^{B} \int_{(\ell-1)/T}^{\ell/T} f(u)\mathrm{d}u = \frac{1}{T} \sum_{\ell=A+1}^{B} f(\ell/T) + R_T(A, B) \tag{90}$$

where

$$|R_T(A, B)| = \Big| \sum_{\ell=A+1}^{B} f'(\zeta_\ell)(\ell/T - \xi_\ell)\frac{1}{T} \Big| \leq \frac{B - A}{T^2} \sup_{A/T < u < B/T} |f'(u)|. \tag{91}$$

In the above we have used that $|\ell/T - \xi_\ell| \leq 1/T$, if $\xi_\ell \in [(\ell-1)/T, \ell/T]$.

Integration by substitution yields that

$$\int_{A/T}^{B/T} f(u)\mathrm{d}u = \frac{B - A}{T} \int_0^1 f\Big(\frac{A}{T} + \frac{B - A}{T}u\Big). \tag{92}$$

Substitute (92) into (90) and multiply by $T/(B - A)$, then the assertion follows from (91).
$\square$

G Results from Saulis and Statulevičus (1991)

In this section we cite two results of Saulis and Statulevičus (1991) that are needed for the proofs in the previous sections.

Lemma G.1 (Lemma 2.4, Saulis and Statulevičus (1991)) *For an arbitrary random variable ξ with $\mathbb{E}\xi = 0$, let there exist $\gamma \geq 0$, $H > 0$ and $\bar{\Delta} > 0$ such that*

$$|\mathrm{cum}_k(\xi)| \leq \left(\frac{k!}{2}\right)^{1+\gamma} \frac{H}{\bar{\Delta}^{k-2}}, \quad k = 2, 3, \ldots. \tag{93}$$

Then, for all $x \geq 0$,[5]

$$\mathbb{P}(|\xi| \geq x) \leq \exp\left(-\frac{x^2}{2(H + (x/\bar{\Delta}^{1/(1+2\gamma)})^{(1+2\gamma)/(1+\gamma)})}\right)$$

$$\leq \begin{cases} \exp\left(-\frac{x^2}{4H}\right) & 0 \leq x \leq (H^{1+\gamma}\bar{\Delta})^{1/(1+2\gamma)} \\ \exp\left(-\frac{1}{4}(x\bar{\Delta})^{1/(1+\gamma)}\right) & x \geq (H^{1+\gamma}\bar{\Delta})^{1/(1+2\gamma)} \end{cases}$$

Lemma G.2 (Theorem 4.17, Saulis and Statulevičus (1991)) *Let X_t, $t = 1, 2, \ldots$ be a random process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and define the strong mixing coefficients by*

$$\alpha^Y(s, t) := \sup_{A \in \sigma(X_u : 1 \leq u \leq s)} \sup_{B \in \sigma(X_u : u \geq t+k)} \left|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)\right|.$$

[notation taken from Saulis and Statulevičus (1991), p. 60] If for some $\gamma_1 \geq 0$ and $H_1 > 0$

$$\mathbb{E}|Y_t|^k \leq (k!)^{1+\gamma_1} H_1^k, \quad t = 1, \ldots, n, \ k = 2, 3, \ldots, \tag{94}$$

then for all $\delta > 0$

$$\left|\mathrm{cum}_k\left(\sum_{t=1}^n Y_t\right)\right| \leq 2(k!)^{2+\gamma_1} 12^{k-1} H_1^k (1+\delta)^{(1+\gamma_1)k} \Delta_n((1 + 1/\delta)(k-1))^{k-1} n,$$

where

$$\Delta_n(\nu) := \max\left\{1, \max_{1 \leq s \leq n} \sum_{t=s}^n \alpha^Y(s, t)^{1/\nu}\right\}.$$

---

[5]Note that in Saulis and Statulevičus (1991) there is a typo!

H Further output for the analysis in Section 4

h	$\hat{p}_{\mathrm{s}}(1)$	$\mathrm{MSPE}^{\mathrm{s}}_{T,1}(h)$	$\hat{p}_{\mathrm{ls}}(1)$	$\hat{N}_{\mathrm{ls}}(1)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,1}(h)$
1	6	7.694097e-09	6	51	4.439184e-09
2	6	7.173126e-09	5	51	4.154400e-09
3	6	7.250332e-09	5	50	4.066610e-09
4	7	7.801423e-09	5	49	4.859895e-09
5	6	6.925372e-09	5	32	4.588102e-09

h	$\mathrm{MSPE}^{\mathrm{s}}_{T,2}(h)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,2}(h)$	$\hat{R}^{\mathrm{s,ls}}_{T,2}(h)$	$\mathrm{MSPE}^{\mathrm{s}}_{T,3}(h)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,3}(h)$	$\hat{R}^{\mathrm{s,ls}}_{T,3}(h)$
1	2.833838e-09	3.133783e-09	0.904	3.308841e-09	2.83092e-09	1.169
2	3.470928e-09	3.232187e-09	1.074	3.895253e-09	2.914949e-09	1.336
3	3.519066e-09	2.95315e-09	1.192	4.503667e-09	3.228895e-09	1.395
4	4.440684e-09	3.330911e-09	1.333	5.120855e-09	3.315358e-09	1.545
5	4.490788e-09	3.643417e-09	1.233	5.66247e-09	4.524367e-09	1.252

Table 9: *Minimum empirical mean squared prediction errors (MSPEs) for h-step ahead prediction, $h = 1, 2, 3, 4, 5$, of the squared and centred FTSE 100 data. Analysis performed with $m := 14$ and $p_{\max} = 8$. Top table shows values computed on the end of the training set. Bottom table shows values computed on the end of the validation and test set.*

h	$\hat{p}_{\mathrm{s}}(1)$	$\mathrm{MSPE}^{\mathrm{s}}_{T,1}(h)$	$\hat{p}_{\mathrm{ls}}(1)$	$\hat{N}_{\mathrm{ls}}(1)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,1}(h)$
1	2	1.914662e-08	6	46	1.682087e-08
2	1	1.710659e-08	7	47	1.424715e-08
3	1	1.704640e-08	7	46	1.363948e-08
4	6	1.616856e-08	7	45	1.283952e-08
5	6	1.555465e-08	3	44	1.280425e-08

h	$\mathrm{MSPE}^{\mathrm{s}}_{T,2}(h)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,2}(h)$	$\hat{R}^{\mathrm{s,ls}}_{T,2}(h)$	$\mathrm{MSPE}^{\mathrm{s}}_{T,3}(h)$	$\mathrm{MSPE}^{\mathrm{ls}}_{T,3}(h)$	$\hat{R}^{\mathrm{s,ls}}_{T,3}(h)$
1	5.157787e-09	3.711193e-09	1.390	3.684765e-09	3.144637e-09	1.172
2	6.006821e-09	3.807446e-09	1.578	6.820476e-09	3.129964e-09	2.179
3	6.400765e-09	3.620512e-09	1.768	7.657029e-09	3.374028e-09	2.269
4	5.518483e-09	3.797911e-09	1.453	4.960166e-09	3.461703e-09	1.432
5	5.19019e-09	4.835555e-09	1.073	5.309015e-09	5.584192e-09	0.951

Table 10: *Minimum empirical mean squared prediction errors (MSPEs) for h-step ahead prediction, $h = 1, 2, 3, 4, 5$, of the squared and centred FTSE 100 data. Analysis performed with $m := 16$ and $p_{\max} = 8$. Top table shows values computed on the end of the training set. Bottom table shows values computed on the end of the validation and test set.*

h	$\hat{p}_\text{s}(1)$	$\text{MSPE}^\text{s}_{T,1}(h)$	$\hat{p}_\text{ls}(1)$	$\hat{N}_\text{ls}(1)$	$\text{MSPE}^\text{ls}_{T,1}(h)$
1	6	9.353206e-09	6	35	6.397327e-09
2	6	8.545269e-09	6	47	5.237836e-09
3	6	8.651203e-09	6	46	5.342035e-09
4	6	7.271229e-09	6	43	4.141545e-09
5	6	6.542987e-09	6	30	3.511903e-09

h	$\text{MSPE}^\text{s}_{T,2}(h)$	$\text{MSPE}^\text{ls}_{T,2}(h)$	$\hat{R}^{\text{s,ls}}_{T,2}(h)$	$\text{MSPE}^\text{s}_{T,3}(h)$	$\text{MSPE}^\text{ls}_{T,3}(h)$	$\hat{R}^{\text{s,ls}}_{T,3}(h)$
1	4.250396e-09	3.77037e-09	1.127	3.103276e-09	3.318341e-09	0.935
2	4.352724e-09	3.650413e-09	1.192	3.638043e-09	3.24847e-09	1.120
3	4.588659e-09	3.55735e-09	1.290	4.203515e-09	3.529553e-09	1.191
4	5.226132e-09	3.626292e-09	1.441	4.943403e-09	3.637315e-09	1.359
5	5.203778e-09	3.989628e-09	1.304	5.298157e-09	3.54644e-09	1.494

Table 11: *Minimum empirical mean squared prediction errors (MSPEs) for h-step ahead prediction, $h = 1, 2, 3, 4, 5$, of the squared and centred FTSE 100 data. Analysis performed with $m := 15$ and $p_\text{max} = 6$. Top table shows values computed on the end of the training set. Bottom table shows values computed on the end of the validation and test set.*

h	$\hat{p}_\text{s}(1)$	$\text{MSPE}^\text{s}_{T,1}(h)$	$\hat{p}_\text{ls}(1)$	$\hat{N}_\text{ls}(1)$	$\text{MSPE}^\text{ls}_{T,1}(h)$
1	6	9.353206e-09	9	46	6.125191e-09
2	6	8.545269e-09	9	45	5.006806e-09
3	6	8.651203e-09	7	46	5.079011e-09
4	8	7.067758e-09	6	43	4.141545e-09
5	8	6.348194e-09	6	30	3.511903e-09

h	$\text{MSPE}^\text{s}_{T,2}(h)$	$\text{MSPE}^\text{ls}_{T,2}(h)$	$\hat{R}^{\text{s,ls}}_{T,2}(h)$	$\text{MSPE}^\text{s}_{T,3}(h)$	$\text{MSPE}^\text{ls}_{T,3}(h)$	$\hat{R}^{\text{s,ls}}_{T,3}(h)$
1	4.250396e-09	3.877702e-09	1.096	3.103276e-09	3.442897e-09	0.901
2	4.352724e-09	3.935614e-09	1.106	3.638043e-09	3.282647e-09	1.108
3	4.588659e-09	3.463196e-09	1.325	4.203515e-09	3.548548e-09	1.185
4	4.913305e-09	3.626292e-09	1.355	4.556755e-09	3.637315e-09	1.253
5	4.862813e-09	3.989628e-09	1.219	4.829544e-09	3.54644e-09	1.362

Table 12: *Minimum empirical mean squared prediction errors (MSPEs) for h-step ahead prediction, $h = 1, 2, 3, 4, 5$, of the squared and centred FTSE 100 data. Analysis performed with $m := 15$ and $p_\text{max} = 10$. Top table shows values computed on the end of the training set. Bottom table shows values computed on the end of the validation and test set.*

I Further simulation results for Section 5.3

Note the organisation of the tables and figures. For each model we have three pages with one figure, in which eight plots are displayed that showing the ratio of median performances, and four tables.

Figure 9: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (18) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.079	0.204	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.0707	0.1458
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1519	0.5651	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.136	0.6475
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	0	2e-04	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0	1e-04
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	5e-04	0.9993	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	1e-04	0.9998
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.8504	0.0697	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.4758	0.1442
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.0797	2e-04	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1318	0.2482
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0.0015	0.2485	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0.0208	0.7292	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0	1
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0	0.0024	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0	0.9976	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0	1
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.15	< 1.15		≥ 1.15	< 1.15
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.15$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.15$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.15$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.15$	0	1

Table 13: *Proportions of the individual events in (32) for the process (18) and selected combinations of n and δ.*

	(32) holds for $h=1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.625	0.6414	0.6217	0.608	0.65	0.6972	0.7478	0.8018	0.8506
0.01	0.6441	0.6751	0.675	0.6398	0.6165	0.5728	0.5825	0.6012	0.6304
0.05	0.7245	0.7967	0.8551	0.866	0.86	0.8402	0.8063	0.7608	0.7307
0.1	0.8035	0.8919	0.9569	0.9752	0.9878	0.995	0.9959	0.9965	0.9976
0.15	0.8654	0.9432	0.9875	0.9946	0.9994	0.9999	1	0.9999	1
0.2	0.9042	0.9695	0.9952	0.9993	1	1	1	1	1
0.4	0.9745	0.997	0.9999	1	1	1	1	1	1
0.6	0.9914	0.9996	1	1	1	1	1	1	1
	(32) holds for $h=5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.6852	0.6876	0.6846	0.6903	0.71	0.7095	0.7173	0.723	0.724
0.01	0.7182	0.7422	0.7978	0.8505	0.9128	0.9621	0.9744	0.9845	0.99
0.05	0.7964	0.8651	0.948	0.979	0.9964	0.9998	1	1	1
0.1	0.8687	0.9349	0.9833	0.9969	0.9999	1	1	1	1
0.15	0.9069	0.962	0.993	0.9991	1	1	1	1	1
0.2	0.9327	0.977	0.9963	0.9998	1	1	1	1	1
0.4	0.9745	0.9941	0.9998	1	1	1	1	1	1
0.6	0.9857	0.9981	1	1	1	1	1	1	1

Table 14: *Proportion of* (32) *being fulfilled for the process* (18) *and different values of h, δ and n.*

	Value of $f(\delta)$, defined in (10), for $h=1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	8.3e-05	7.2e-05	3.6e-05	8.7e-05	0.002	0.005	0.0059	0.0063	0.0066
0.05	0.021	0.016	0.012	0.0082	0.005	0.0033	0.0013	7.9e-05	6.1e-05
0.1	0.071	0.066	0.062	0.058	0.055	0.054	0.052	0.05	0.049
0.15	0.12	0.12	0.11	0.11	0.11	0.1	0.1	0.1	0.099
0.2	0.17	0.17	0.16	0.16	0.16	0.15	0.15	0.15	0.15
0.4	0.37	0.37	0.36	0.36	0.36	0.36	0.35	0.35	0.35
0.6	0.58	0.57	0.56	0.56	0.56	0.56	0.55	0.55	0.55
	Value of $f(\delta)$, defined in (10), for $h=5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.05	0.051	0.051	0.052	0.052	0.052	0.053	0.053	0.053	0.053
0.1	0.1	0.1	0.1	0.1	0.1	0.11	0.11	0.11	0.11
0.15	0.15	0.15	0.16	0.16	0.16	0.16	0.16	0.16	0.16
0.2	0.2	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21
0.4	0.41	0.41	0.41	0.42	0.42	0.42	0.42	0.42	0.42
0.6	0.61	0.62	0.62	0.63	0.63	0.63	0.63	0.63	0.63

Table 15: *Values of $f(\delta)$, defined in* (10), *for the process* (18) *and different values of h, δ and n.*

	(32) holds for $h=1$, $i=2$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.2541	0.2364	0.2686	0.4015	0.5398	0.5853	0.598	0.6493	0.6992
0.01	0.2309	0.2002	0.206	0.3103	0.4495	0.4948	0.4825	0.5055	0.529
0.05	0.1553	0.1043	0.0737	0.092	0.1061	0.0889	0.0567	0.0385	0.0223
0.1	0.1001	0.0505	0.0243	0.0168	0.0079	7e-04	2e-04	1e-04	0
0.15	0.0644	0.025	0.0075	0.0036	4e-04	0	0	0	0
0.2	0.0432	0.0134	0.0031	5e-04	0	0	0	0	0
0.4	0.0115	0.0015	1e-04	0	0	0	0	0	0
0.6	0.0041	2e-04	0	0	0	0	0	0	0
	(32) holds for $h=1$, $i=3$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.3091	0.3082	0.3423	0.3547	0.3704	0.4677	0.5596	0.6263	0.6892
0.01	0.283	0.2643	0.2668	0.2645	0.2682	0.353	0.4654	0.5491	0.6298
0.05	0.2	0.1472	0.0928	0.0632	0.0445	0.0781	0.141	0.2045	0.25
0.1	0.1336	0.0706	0.0218	0.0098	0.0043	0.0043	0.0039	0.0034	0.0024
0.15	0.0856	0.0362	0.0054	0.0018	2e-04	1e-04	0	1e-04	0
0.2	0.0588	0.0191	0.0017	2e-04	0	0	0	0	0
0.4	0.0144	0.0015	0	0	0	0	0	0	0
0.6	0.0045	2e-04	0	0	0	0	0	0	0
	(32) holds for $h=5$, $i=2$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.248	0.2496	0.2553	0.2552	0.2366	0.227	0.2309	0.2233	0.2195
0.01	0.2067	0.1719	0.1272	0.083	0.0439	0.0174	0.0109	0.0053	0.0034
0.05	0.1232	0.0746	0.0251	0.0087	5e-04	0	0	0	0
0.1	0.0723	0.0335	0.007	0.001	0	0	0	0	0
0.15	0.0487	0.0186	0.0022	1e-04	0	0	0	0	0
0.2	0.035	0.0105	7e-04	1e-04	0	0	0	0	0
0.4	0.0134	0.0022	0	0	0	0	0	0	0
0.6	0.0075	4e-04	0	0	0	0	0	0	0
	(32) holds for $h=5$, $i=3$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.259	0.272	0.2709	0.2697	0.248	0.2413	0.233	0.2305	0.2319
0.01	0.2165	0.1853	0.134	0.0923	0.0525	0.0227	0.0151	0.0104	0.0066
0.05	0.134	0.0829	0.0303	0.0125	0.0031	2e-04	0	0	0
0.1	0.079	0.0368	0.0099	0.0021	1e-04	0	0	0	0
0.15	0.0532	0.0202	0.0048	8e-04	0	0	0	0	0
0.2	0.0381	0.0127	0.003	1e-04	0	0	0	0	0
0.4	0.0133	0.0037	2e-04	0	0	0	0	0	0
0.6	0.0068	0.0015	0	0	0	0	0	0	0

Table 16: *Proportion of* (7) *being fulfilled for the process* (18) *and different values of h, δ and n.*
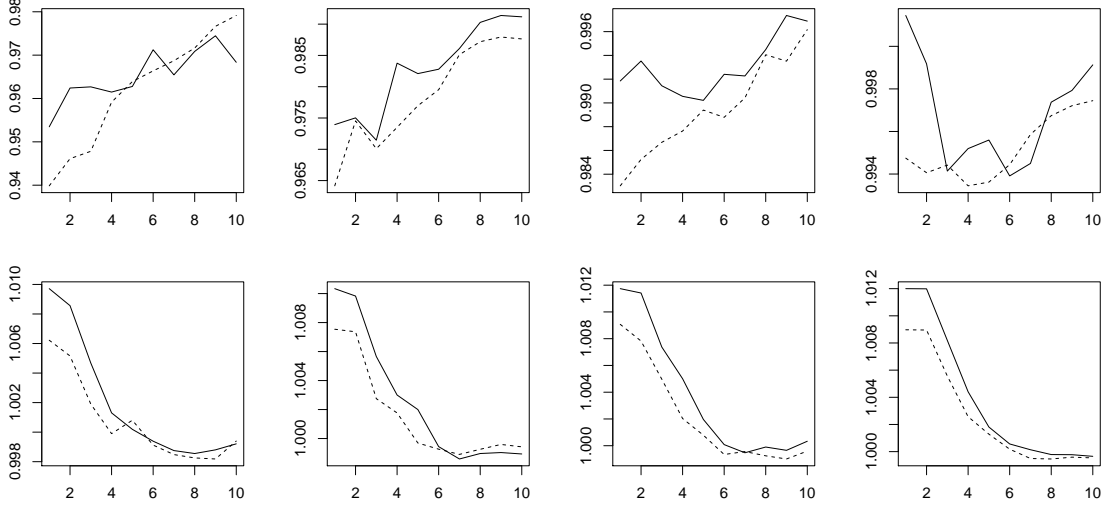
Figure 10: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (19) and different values of n (from left to right: n=100, n=200, n=500, n=1000 [first row], n=4000, n=6000, n=8000, n=10000 [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.1082	0.1936	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.0684	0.1481
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1745	0.5237	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.1436	0.6399
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	3e-04	2e-04	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0	0.0031
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	0	0.9995	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	0.002	0.9949
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.7681	0.132	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.4061	0.2357
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.0835	0.0164	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1828	0.1754
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.2385	0.2906	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.0326	0.1073
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1973	0.2736	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.0794	0.7807
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0	0.0017	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	1e-04	0.0025
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	9e-04	0.9974	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0.0025	0.9949
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0	1

Table 17: *Proportions of the individual events in (32) for the process (19) and selected combinations of n and δ.*

	(32) holds for $h=1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.6103	0.5896	0.5583	0.5519	0.5532	0.6044	0.6691	0.7362	0.7845
0.01	0.6319	0.6273	0.6189	0.6139	0.5877	0.5505	0.5309	0.5238	0.5121
0.05	0.7188	0.7656	0.8554	0.8987	0.9466	0.9802	0.9907	0.9955	0.9974
0.1	0.7986	0.8794	0.964	0.9895	0.9987	0.9999	1	0.9999	1
0.15	0.8578	0.9361	0.99	0.9991	1	1	1	1	1
0.2	0.8986	0.9649	0.9972	0.9998	1	1	1	1	1
0.4	0.962	0.9911	0.9994	0.9999	1	1	1	1	1
0.6	0.977	0.9924	0.9991	0.9997	1	1	1	1	1
	(32) holds for $h=5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.6959	0.6963	0.6908	0.6695	0.6437	0.6218	0.6141	0.5958	0.5815
0.01	0.7083	0.7244	0.7409	0.7473	0.7607	0.7789	0.8001	0.7854	0.8133
0.05	0.7754	0.8255	0.8856	0.9211	0.9536	0.9784	0.9871	0.9914	0.995
0.1	0.8297	0.8941	0.9495	0.9721	0.9893	0.9982	0.9991	0.9997	1
0.15	0.8702	0.9308	0.9737	0.9886	0.9973	0.9999	0.9999	1	1
0.2	0.8965	0.95	0.9854	0.9949	0.9991	1	1	1	1
0.4	0.9528	0.9838	0.9973	0.9997	1	1	1	1	1
0.6	0.9747	0.9936	0.9989	1	1	1	1	1	1

Table 18: *Proportion of* (32) *being fulfilled for the process* (19) *and different values of h, δ and n.*

	Value of $f(\delta)$, defined in (10), for $h=1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	7.7e-06	8.8e-07	0.00082	0.0014	0.0019	0.0023	0.0025	0.0026	0.0027
0.05	0.039	0.039	0.038	0.038	0.037	0.037	0.037	0.037	0.037
0.1	0.089	0.089	0.088	0.088	0.087	0.087	0.087	0.087	0.087
0.15	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
0.2	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
0.4	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39
0.6	0.25	0.26	0.26	0.27	0.27	0.27	0.28	0.28	0.28
	Value of $f(\delta)$, defined in (10), for $h=5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.012	0.011	0.01	0.0097	0.0094	0.0091	0.0089	0.0088	0.0088
0.05	0.056	0.054	0.052	0.052	0.051	0.05	0.05	0.05	0.05
0.1	0.15	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
0.15	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
0.2	0.32	0.32	0.33	0.33	0.33	0.33	0.33	0.33	0.33
0.4	0.68	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
0.6	1	1	1.1	1.1	1.1	1.1	1.1	1.1	1.1

Table 19: *Values of $f(\delta)$, defined in (10), for the process* (19) *and different values of h, δ and n.*

δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
	(32) holds for $h=1$, $i=2$								
0	0.3129	0.329	0.3657	0.4304	0.5329	0.6673	0.7452	0.8055	0.8516
0.01	0.2827	0.2748	0.2692	0.2834	0.3135	0.3528	0.384	0.4149	0.4358
0.05	0.1927	0.1502	0.0781	0.0482	0.0246	0.0077	0.0026	0.0018	9e-04
0.1	0.1249	0.0775	0.0182	0.0046	3e-04	1e-04	0	0	0
0.15	0.0875	0.0442	0.0048	5e-04	0	0	0	0	0
0.2	0.0617	0.0278	0.0024	3e-04	0	0	0	0	0
0.4	0.0226	0.0099	0.002	2e-04	0	0	0	0	0
0.6	0.0119	0.0054	0.0013	0	0	0	0	0	0
	(32) holds for $h=1$, $i=3$								
0	0.333	0.3492	0.4062	0.4959	0.5991	0.7385	0.8141	0.8735	0.9001
0.01	0.3018	0.2983	0.3085	0.3467	0.373	0.4393	0.4795	0.5205	0.5291
0.05	0.2157	0.1676	0.0969	0.0661	0.0324	0.0125	0.0067	0.0027	0.0017
0.1	0.1429	0.0831	0.024	0.0075	0.001	0	0	1e-04	0
0.15	0.1031	0.0461	0.0094	0.001	0	0	0	0	0
0.2	0.0761	0.0281	0.0042	5e-04	0	0	0	0	0
0.4	0.0306	0.0108	0.0016	3e-04	0	0	0	0	0
0.6	0.0183	0.0074	0.001	3e-04	0	0	0	0	0
	(32) holds for $h=5$, $i=2$								
0	0.2408	0.2295	0.2407	0.2613	0.2881	0.3542	0.3949	0.4391	0.4716
0.01	0.212	0.1869	0.1718	0.1667	0.1456	0.139	0.1219	0.126	0.112
0.05	0.143	0.1011	0.0606	0.0414	0.0217	0.0119	0.0067	0.0049	0.0026
0.1	0.0944	0.0549	0.0247	0.0135	0.0046	8e-04	3e-04	1e-04	0
0.15	0.0669	0.0333	0.0115	0.0045	0.0011	1e-04	0	0	0
0.2	0.0507	0.0222	0.0069	0.002	5e-04	0	0	0	0
0.4	0.0211	0.0056	0.001	3e-04	0	0	0	0	0
0.6	0.012	0.0023	4e-04	0	0	0	0	0	0
	(32) holds for $h=5$, $i=3$								
0	0.2411	0.2434	0.2487	0.2758	0.318	0.3742	0.4258	0.4899	0.5245
0.01	0.2165	0.2025	0.1827	0.1794	0.1719	0.1583	0.1504	0.1606	0.1399
0.05	0.1532	0.114	0.0714	0.0483	0.0291	0.0123	0.0086	0.0045	0.0026
0.1	0.1087	0.0676	0.03	0.0158	0.0065	0.001	6e-04	2e-04	0
0.15	0.0821	0.0437	0.016	0.0073	0.0018	0	1e-04	0	0
0.2	0.0654	0.031	0.0087	0.0031	4e-04	0	0	0	0
0.4	0.0299	0.0108	0.0017	0	0	0	0	0	0
0.6	0.0151	0.0041	7e-04	0	0	0	0	0	0

Table 20: *Proportion of* (7) *being fulfilled for the process* (19) *and different values of h, δ and n.*

Figure 11: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (20) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.0979	0.1914	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.0696	0.1506
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1699	0.5408	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.1363	0.6435
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	5e-04	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0	9e-04
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	1e-04	0.9994	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	5e-04	0.9986
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.2773	0.2654	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.361	0.1876
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.2072	0.2501	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1629	0.2885
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.0037	0.0441	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.003	0.0244
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.0304	0.9218	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.0212	0.9514
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0	5e-04
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	1e-04	0.9994
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0	1

Table 21: *Proportions of the individual events in (32) for the process (20) and selected combinations of n and δ.*

δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
				(32) holds for $h=1$					
0	0.616	0.6082	0.5902	0.591	0.5702	0.5459	0.5435	0.532	0.5274
0.01	0.6387	0.6543	0.6866	0.7303	0.7861	0.853	0.8849	0.8997	0.9255
0.05	0.7346	0.8017	0.9196	0.9742	0.9951	0.9999	0.9998	0.9999	1
0.1	0.813	0.9005	0.9851	0.9984	1	1	1	1	1
0.15	0.8673	0.9475	0.9962	0.9999	1	1	1	1	1
0.2	0.9016	0.969	0.9976	0.9999	1	1	1	1	1
0.4	0.9639	0.9881	0.9967	0.9995	0.9999	1	1	1	1
0.6	0.9809	0.9914	0.9979	0.9997	1	1	1	1	1
				(32) holds for $h=5$					
0	0.6976	0.6961	0.6942	0.6876	0.68	0.6796	0.6736	0.6638	0.6495
0.01	0.7131	0.7258	0.7667	0.7906	0.8404	0.8981	0.9282	0.9375	0.9544
0.05	0.7845	0.8393	0.9115	0.9519	0.9797	0.996	0.9987	0.9994	0.9994
0.1	0.8404	0.904	0.9616	0.9867	0.997	0.9999	1	1	1
0.15	0.8822	0.9379	0.9817	0.9958	0.9996	1	1	1	1
0.2	0.9103	0.9587	0.9905	0.9986	0.9999	1	1	1	1
0.4	0.9602	0.9869	0.999	0.9999	1	1	1	1	1
0.6	0.9781	0.9942	0.9998	1	1	1	1	1	1

Table 22: *Proportion of* (32) *being fulfilled for the process* (20) *and different values of h, δ and n.*

δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
				Value of $f(\delta)$, defined in (10), for $h=1$					
0	0	0	0	0	0	0	0	0	0
0.01	0.0079	0.0078	0.0077	0.0076	0.0075	0.0075	0.0075	0.0074	0.0074
0.05	0.048	0.048	0.048	0.048	0.048	0.047	0.047	0.047	0.047
0.1	0.098	0.098	0.098	0.098	0.098	0.097	0.097	0.097	0.097
0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
0.4	0.12	0.12	0.12	0.12	0.12	0.12	0.13	0.13	0.13
0.6	0.081	0.08	0.078	0.077	0.076	0.075	0.074	0.074	0.074
				Value of $f(\delta)$, defined in (10), for $h=5$					
0	0	0	0	0	0	0	0	0	0
0.01	0.0084	0.0082	0.008	0.0079	0.0079	0.0078	0.0078	0.0078	0.0078
0.05	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069
0.1	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
0.15	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22
0.2	0.29	0.29	0.3	0.3	0.3	0.3	0.3	0.3	0.3
0.4	0.59	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
0.6	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9

Table 23: *Values of $f(\delta)$, defined in* (10), *for the process* (20) *and different values of h, δ and n.*

	(32) holds for $h=1$, $i=2$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2962	0.2991	0.3003	0.3075	0.3319	0.3746	0.4137	0.4525	0.4845
0.01	0.2678	0.2456	0.2084	0.1631	0.1232	0.0791	0.0551	0.0468	0.0341
0.05	0.1781	0.1285	0.0475	0.0131	0.0021	1e-04	0	0	0
0.1	0.1173	0.0648	0.0123	0.0011	1e-04	0	0	0	0
0.15	0.0807	0.0376	0.0064	6e-04	1e-04	0	0	0	0
0.2	0.0565	0.0249	0.0057	6e-04	1e-04	0	0	0	0
0.4	0.0207	0.0065	0.0034	3e-04	0	0	0	0	0
0.6	0.0093	0.0029	0.0015	3e-04	0	0	0	0	0
	(32) holds for $h=1$, $i=3$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.32	0.3125	0.3241	0.3443	0.3641	0.4241	0.4728	0.5189	0.5427
0.01	0.2893	0.2613	0.2216	0.189	0.1345	0.0949	0.0752	0.0607	0.0478
0.05	0.2007	0.1396	0.0509	0.0163	0.003	0	2e-04	1e-04	0
0.1	0.1341	0.0707	0.0146	0.0017	1e-04	0	0	0	0
0.15	0.094	0.0399	0.0074	5e-04	1e-04	0	0	0	0
0.2	0.0703	0.0253	0.0055	5e-04	1e-04	0	0	0	0
0.4	0.0258	0.0122	0.0027	4e-04	1e-04	0	0	0	0
0.6	0.0142	0.0079	0.0012	2e-04	0	0	0	0	0
	(32) holds for $h=5$, $i=2$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2412	0.2309	0.2329	0.2401	0.2344	0.2429	0.2546	0.2697	0.2717
0.01	0.2059	0.183	0.1474	0.1286	0.0895	0.0542	0.04	0.0341	0.0242
0.05	0.1326	0.094	0.0462	0.0247	0.0086	0.0017	5e-04	3e-04	1e-04
0.1	0.0866	0.0485	0.0177	0.0057	0.001	0	0	0	0
0.15	0.0591	0.0296	0.0077	0.0016	0	0	0	0	0
0.2	0.0444	0.0191	0.004	5e-04	0	0	0	0	0
0.4	0.0182	0.0046	4e-04	1e-04	0	0	0	0	0
0.6	0.0099	0.0018	1e-04	0	0	0	0	0	0
	(32) holds for $h=5$, $i=3$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2474	0.2428	0.2401	0.2433	0.2544	0.2647	0.2722	0.2901	0.2964
0.01	0.2202	0.197	0.1585	0.1326	0.1021	0.0623	0.0424	0.0392	0.0274
0.05	0.1497	0.1027	0.0529	0.026	0.0119	0.0023	8e-04	3e-04	5e-04
0.1	0.102	0.0589	0.0221	0.0078	0.002	1e-04	0	0	0
0.15	0.0755	0.0375	0.011	0.0026	4e-04	0	0	0	0
0.2	0.0561	0.0256	0.0055	9e-04	1e-04	0	0	0	0
0.4	0.0236	0.0087	6e-04	0	0	0	0	0	0
0.6	0.0128	0.004	1e-04	0	0	0	0	0	0

Table 24: *Proportion of* (7) *being fulfilled for the process* (20) *and different values of h, δ and n.*
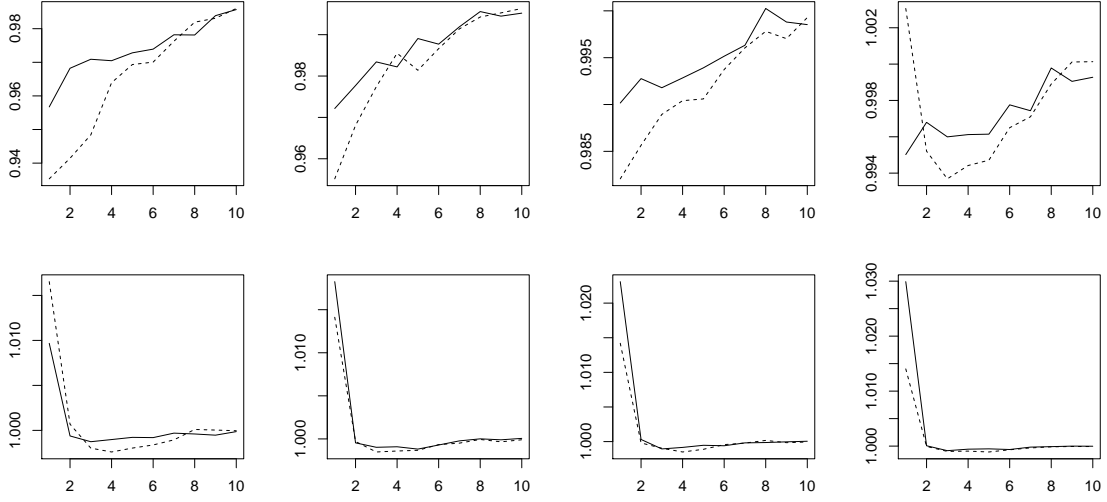
Figure 12: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (21) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.1489	0.1747	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.1004	0.1456
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.17	0.5064	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.1465	0.6075

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	0.0011	0.0572	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0.0617	0.2354
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	0.0134	0.9283	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	0.0967	0.6062

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.9901	0.007	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.9842	0.0108
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.0029	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.005	0

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0.4201	0.367	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0.884	0.0857
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0.1251	0.0878	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0.0291	0.0012

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	1e-04	0.0303	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0.2004	0.4467
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	3e-04	0.9693	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	0.127	0.2259

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.4	< 1.4		≥ 1.4	< 1.4
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.4$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.4$	0.0033	0.1876
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.4$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.4$	0.0171	0.792

Table 25: *Proportions of the individual events in (32) for the process (21) and selected combinations of n and δ.*

	(32) holds for $h = 1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.6404	0.6018	0.5998	0.6319	0.75	0.9035	0.9609	0.9807	0.9901
0.01	0.6553	0.624	0.5893	0.5753	0.6486	0.802	0.8862	0.9291	0.9569
0.05	0.7193	0.7135	0.6695	0.5726	0.5011	0.4696	0.4664	0.4805	0.5079
0.1	0.781	0.7978	0.7899	0.73	0.6828	0.6286	0.6121	0.5904	0.5965
0.15	0.833	0.8633	0.8774	0.8575	0.8447	0.8413	0.8564	0.8553	0.8727
0.2	0.8702	0.9122	0.9264	0.9294	0.9309	0.943	0.9581	0.9597	0.9694
0.4	0.9508	0.9843	0.9929	0.9967	0.9986	0.9995	1	0.9999	1
0.6	0.9805	0.9977	0.9994	1	1	1	1	1	1
	(32) holds for $h = 5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.7044	0.6508	0.5888	0.6245	0.7502	0.8941	0.952	0.9722	0.9842
0.01	0.7079	0.655	0.5826	0.5983	0.7085	0.8614	0.9279	0.9576	0.9752
0.05	0.7369	0.7019	0.6134	0.5471	0.5787	0.6946	0.7873	0.8476	0.8852
0.1	0.7741	0.7531	0.6622	0.556	0.5112	0.5341	0.5841	0.628	0.6849
0.15	0.8029	0.7953	0.7173	0.6059	0.5162	0.4616	0.4711	0.4731	0.4988
0.2	0.8302	0.8303	0.7661	0.6679	0.5691	0.4898	0.4523	0.437	0.4263
0.4	0.9004	0.9111	0.8993	0.8688	0.8295	0.8072	0.8015	0.7899	0.7953
0.6	0.9331	0.949	0.954	0.9577	0.9498	0.9496	0.9581	0.9584	0.963

Table 26: *Proportion of* (32) *being fulfilled for the process* (21) *and different values of h, δ and n.*

	Value of $f(\delta)$, defined in (10), for $h = 1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.04	0.047	0.055	0.06	0.065	0.07	0.072	0.074	0.075
0.05	7.4e-05	0.006	0.014	0.02	0.025	0.029	0.032	0.033	0.035
0.1	0.028	0.023	0.017	0.012	0.0078	0.0037	0.0012	7.1e-07	2.3e-06
0.15	0.079	0.074	0.068	0.063	0.058	0.054	0.051	0.05	0.048
0.2	0.13	0.12	0.12	0.11	0.11	0.1	0.1	0.1	0.099
0.4	0.33	0.33	0.32	0.31	0.31	0.3	0.3	0.3	0.3
0.6	0.53	0.53	0.52	0.51	0.51	0.5	0.5	0.5	0.5
	Value of $f(\delta)$, defined in (10), for $h = 5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.18	0.22	0.26	0.28	0.3	0.31	0.32	0.33	0.33
0.05	0.24	0.42	0.62	0.75	0.86	0.95	1	1	1.1
0.1	0.0055	0.18	0.38	0.51	0.61	0.71	0.76	0.79	0.81
0.15	7.5e-05	0.00048	0.14	0.27	0.37	0.47	0.52	0.55	0.57
0.2	0.082	0.00098	0.00016	0.025	0.13	0.23	0.28	0.31	0.33
0.4	1	0.79	0.58	0.44	0.33	0.23	0.18	0.14	0.12
0.6	1.9	1.7	1.5	1.4	1.3	1.2	1.1	1.1	1.1

Table 27: *Values of $f(\delta)$, defined in (10), for the process* (21) *and different values of h, δ and n.*

					(32) holds for $h=1$, $i=2$				
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.3496	0.3822	0.4943	0.6506	0.8054	0.9362	0.9727	0.9861	0.993
0.01	0.3189	0.3328	0.4169	0.5524	0.7087	0.8591	0.9179	0.9465	0.9683
0.05	0.233	0.2039	0.2176	0.2647	0.3387	0.4203	0.4784	0.5122	0.5452
0.1	0.165	0.1237	0.1024	0.1086	0.1118	0.1119	0.1141	0.1074	0.0987
0.15	0.1211	0.0764	0.0456	0.0393	0.0335	0.0258	0.0183	0.0167	0.0096
0.2	0.0872	0.0462	0.0225	0.0145	0.0097	0.0044	0.0018	0.0014	4e-04
0.4	0.0318	0.0086	0.0014	3e-04	1e-04	0	0	0	0
0.6	0.0157	0.003	2e-04	0	0	0	0	0	0
					(32) holds for $h=1$, $i=3$				
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.352	0.4106	0.5507	0.7357	0.8898	0.9615	0.9874	0.9946	0.9971
0.01	0.3236	0.3684	0.4852	0.6567	0.8169	0.9179	0.9627	0.9818	0.9886
0.05	0.2435	0.244	0.2975	0.4125	0.5162	0.6429	0.709	0.7563	0.7871
0.1	0.1772	0.1555	0.1665	0.2192	0.263	0.3149	0.333	0.3634	0.3644
0.15	0.1285	0.0987	0.0948	0.1144	0.127	0.1359	0.1277	0.1302	0.1189
0.2	0.1014	0.0638	0.0579	0.0583	0.0598	0.0528	0.0401	0.0389	0.0304
0.4	0.0404	0.0127	0.0057	0.003	0.0013	5e-04	0	1e-04	0
0.6	0.019	0.0037	4e-04	0	0	0	0	0	0
					(32) holds for $h=5$, $i=2$				
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.2616	0.3316	0.4864	0.6481	0.8103	0.9344	0.97	0.9809	0.9892
0.01	0.2469	0.3117	0.451	0.6039	0.7711	0.9103	0.9517	0.9707	0.9828
0.05	0.2024	0.2398	0.3383	0.4548	0.6174	0.7682	0.8468	0.8864	0.9131
0.1	0.1605	0.1828	0.2358	0.3219	0.4419	0.5687	0.649	0.6947	0.7399
0.15	0.1319	0.1407	0.168	0.2271	0.3006	0.3872	0.4515	0.4906	0.5244
0.2	0.1092	0.1116	0.1227	0.1584	0.2038	0.2556	0.2953	0.3099	0.3274
0.4	0.0604	0.0532	0.0425	0.0366	0.0381	0.0338	0.0277	0.0243	0.0204
0.6	0.04	0.0314	0.0172	0.0085	0.006	0.0035	0.0019	0.0011	3e-04
					(32) holds for $h=5$, $i=3$				
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.2564	0.3404	0.5404	0.7278	0.8817	0.9549	0.9816	0.9911	0.995
0.01	0.246	0.3233	0.5108	0.6954	0.854	0.9419	0.9748	0.9863	0.9924
0.05	0.2075	0.2627	0.4125	0.5835	0.7497	0.8708	0.9251	0.9544	0.9697
0.1	0.1708	0.2093	0.3232	0.4697	0.6179	0.7556	0.8255	0.8699	0.8972
0.15	0.1442	0.1682	0.2525	0.3734	0.5006	0.6326	0.694	0.7477	0.7852
0.2	0.1202	0.1369	0.203	0.2971	0.3995	0.5066	0.564	0.6141	0.6471
0.4	0.0658	0.0687	0.0852	0.108	0.146	0.1698	0.1782	0.1934	0.1909
0.6	0.0423	0.0398	0.0412	0.0378	0.045	0.0469	0.0404	0.0405	0.0367

Table 28: *Proportion of* (7) *being fulfilled for the process* (21) *and different values of h, δ and n.*
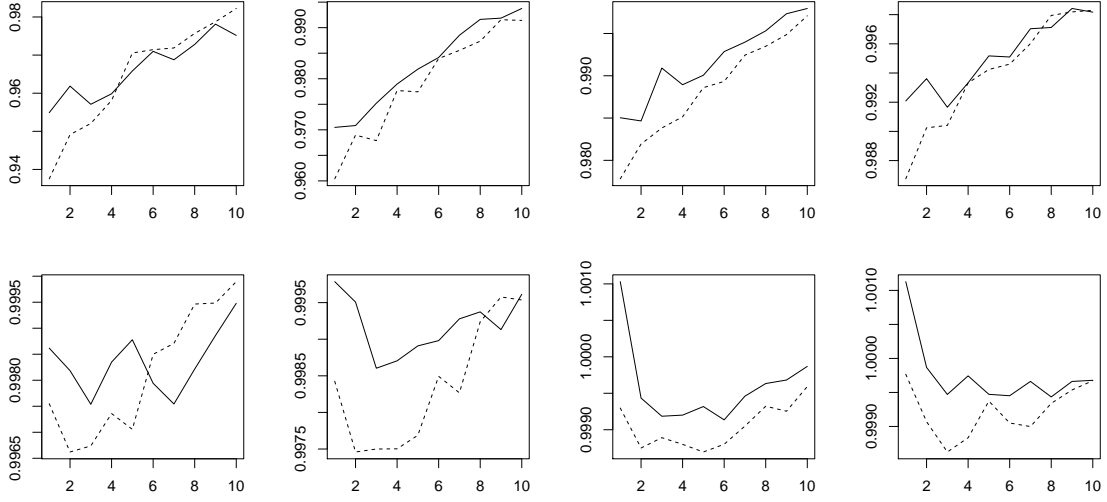
Figure 13: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (22) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.1398	0.1678	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.1077	0.134
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1714	0.521	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.1543	0.604
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	0	0.0058	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0.0158	0.0997
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	0.0026	0.9916	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	0.0551	0.8294
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.8702	0.0826	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.8849	0.0707
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.0436	0.0036	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.0419	0.0025
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0.0098	0.1703	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0.4632	0.2837
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0.053	0.7669	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0.1531	0.1
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0	0.0129	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0.1556	0.3354
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	8e-04	0.9863	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0.1588	0.3502
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0.0074	0.1566
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	0.0441	0.7919

Table 29: *Proportions of the individual events in (32) for the process (22) and selected combinations of n and δ.*

δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
	(32) holds for $h=1$								
0	0.6493	0.6142	0.6138	0.584	0.5953	0.673	0.7638	0.8233	0.8738
0.01	0.6608	0.6464	0.6539	0.6156	0.5676	0.5291	0.5587	0.5705	0.6219
0.05	0.7315	0.7667	0.8127	0.8154	0.798	0.776	0.7756	0.7658	0.7767
0.1	0.8067	0.8605	0.921	0.9372	0.946	0.9617	0.9789	0.9802	0.9863
0.15	0.8497	0.9172	0.9659	0.9787	0.9882	0.9957	0.9984	0.9986	0.9995
0.2	0.8858	0.9516	0.9844	0.9916	0.9976	0.9998	0.9998	1	1
0.4	0.9602	0.9938	0.9995	1	1	1	1	1	1
0.6	0.9839	0.9992	1	1	1	1	1	1	1
	(32) holds for $h=5$								
0	0.7065	0.6485	0.5748	0.546	0.582	0.6998	0.7896	0.8482	0.8874
0.01	0.7117	0.6596	0.5918	0.554	0.5566	0.6298	0.71	0.7695	0.8197
0.05	0.7331	0.7045	0.678	0.6252	0.5561	0.5115	0.5187	0.5228	0.5632
0.1	0.7689	0.7632	0.7627	0.7114	0.639	0.5747	0.5308	0.5026	0.5058
0.15	0.7959	0.8095	0.8196	0.7828	0.7415	0.6918	0.6657	0.6468	0.6501
0.2	0.8224	0.845	0.8631	0.8452	0.8178	0.7994	0.8002	0.788	0.7993
0.4	0.8906	0.9195	0.9505	0.9608	0.9669	0.9757	0.9853	0.9863	0.9894
0.6	0.9303	0.9545	0.9775	0.9896	0.9943	0.9983	0.9992	0.9993	0.9998

Table 30: *Proportion of (32) being fulfilled for the process (22) and different values of h, δ and n.*

δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
	Value of $f(\delta)$, defined in (10), for $h=1$								
0	0	0	0	0	0	0	0	0	0
0.01	0.0089	0.011	0.014	0.015	0.017	0.018	0.019	0.02	0.02
0.05	0.025	0.024	0.022	0.021	0.02	0.018	0.018	0.017	0.017
0.1	0.076	0.074	0.072	0.071	0.07	0.068	0.068	0.067	0.067
0.15	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
0.2	0.18	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
0.4	0.38	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
0.6	0.58	0.58	0.57	0.57	0.57	0.57	0.57	0.57	0.57
	Value of $f(\delta)$, defined in (10), for $h=5$								
0	0	0	0	0	0	0	0	0	0
0.01	0.21	0.29	0.34	0.36	0.37	0.39	0.39	0.39	0.4
0.05	0.023	0.095	0.17	0.22	0.26	0.29	0.31	0.32	0.32
0.1	0.095	0.022	7.7e-05	4.6e-05	0.018	0.05	0.066	0.076	0.085
0.15	0.33	0.26	0.19	0.15	0.11	0.087	0.072	0.063	0.055
0.2	0.56	0.49	0.43	0.38	0.35	0.33	0.31	0.3	0.29
0.4	1.5	1.4	1.4	1.3	1.3	1.3	1.3	1.3	1.2
0.6	2.4	2.4	2.3	2.3	2.3	2.2	2.2	2.2	2.2

Table 31: *Values of $f(\delta)$, defined in (10), for the process (22) and different values of h, δ and n.*

		(32) holds for $h=1$, $i=2$							
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.3392	0.3501	0.3702	0.4495	0.57	0.7341	0.8276	0.8758	0.9138
0.01	0.3112	0.2987	0.2791	0.3177	0.3918	0.5016	0.5931	0.641	0.6887
0.05	0.2228	0.1684	0.1134	0.0952	0.0908	0.0794	0.0777	0.0737	0.0628
0.1	0.1549	0.091	0.0383	0.024	0.0153	0.0076	0.0033	0.0025	8e-04
0.15	0.1131	0.0533	0.0126	0.007	0.0027	3e-04	0	0	0
0.2	0.0851	0.0317	0.0053	0.0026	7e-04	0	0	0	0
0.4	0.0302	0.0056	2e-04	0	0	0	0	0	0
0.6	0.0139	0.0028	0	0	0	0	0	0	0
		(32) holds for $h=1$, $i=3$							
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.3373	0.3487	0.3862	0.4969	0.6467	0.8087	0.8906	0.9321	0.9528
0.01	0.3076	0.2977	0.3044	0.3773	0.4782	0.6285	0.7074	0.7671	0.8016
0.05	0.2187	0.1765	0.1331	0.1382	0.1474	0.1712	0.1721	0.1833	0.1801
0.1	0.1506	0.0965	0.0553	0.0436	0.0399	0.0307	0.018	0.0173	0.0129
0.15	0.111	0.0537	0.0241	0.0149	0.0091	0.004	0.0016	0.0014	5e-04
0.2	0.0795	0.0307	0.0111	0.0058	0.0017	2e-04	2e-04	0	0
0.4	0.0286	0.0062	3e-04	0	0	0	0	0	0
0.6	0.0144	0.003	0	0	0	0	0	0	0
		(32) holds for $h=5$, $i=2$							
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.2727	0.3242	0.3954	0.4996	0.6266	0.7847	0.8613	0.9018	0.9268
0.01	0.262	0.3061	0.358	0.4399	0.5548	0.7093	0.7924	0.842	0.876
0.05	0.2225	0.2367	0.2462	0.28	0.3524	0.4535	0.5311	0.5675	0.6163
0.1	0.1801	0.1761	0.1593	0.17	0.205	0.246	0.284	0.2932	0.3144
0.15	0.1493	0.1378	0.1133	0.1098	0.1185	0.1293	0.1408	0.1422	0.135
0.2	0.128	0.1075	0.0822	0.0709	0.0728	0.0669	0.0618	0.0611	0.0515
0.4	0.0735	0.0498	0.0259	0.0147	0.008	0.0053	0.0019	0.001	5e-04
0.6	0.0468	0.0263	0.0137	0.0049	0.0015	2e-04	0	0	0
		(32) holds for $h=5$, $i=3$							
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.2524	0.3177	0.4114	0.5402	0.6984	0.8359	0.8971	0.9352	0.9556
0.01	0.2417	0.2979	0.3768	0.4899	0.6414	0.7851	0.8566	0.8987	0.9269
0.05	0.2062	0.2394	0.2712	0.3486	0.4597	0.581	0.653	0.7093	0.7469
0.1	0.1704	0.1817	0.1916	0.2384	0.3098	0.3881	0.4264	0.4708	0.491
0.15	0.1426	0.1423	0.1407	0.1656	0.2048	0.2469	0.2647	0.2856	0.2915
0.2	0.1204	0.1143	0.1049	0.1155	0.1362	0.1549	0.1558	0.1677	0.164
0.4	0.0687	0.0585	0.041	0.0315	0.0271	0.019	0.013	0.0127	0.0101
0.6	0.0429	0.0352	0.0204	0.0097	0.0052	0.0015	8e-04	7e-04	2e-04

Table 32: *Proportion of* (7) *being fulfilled for the process* (22) *and different values of h, δ and n.*
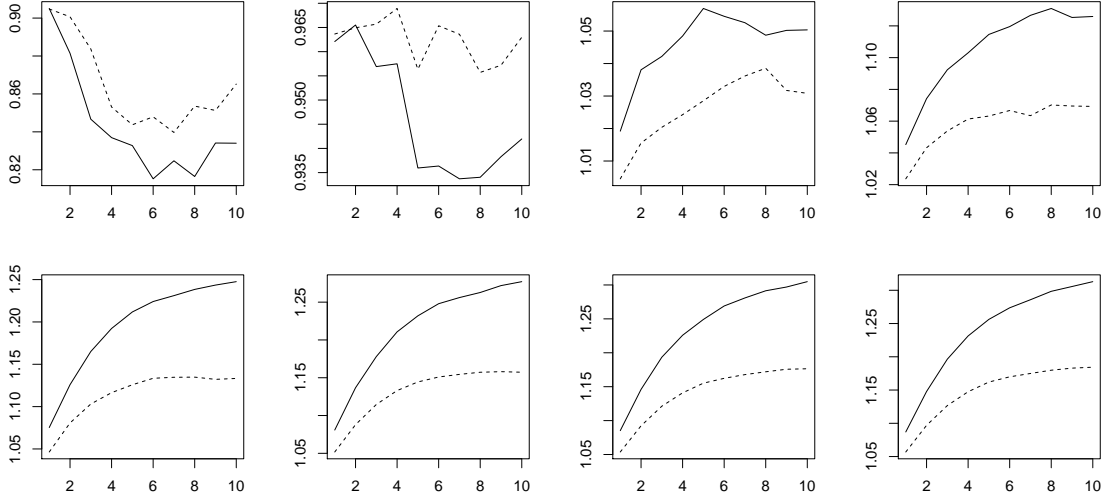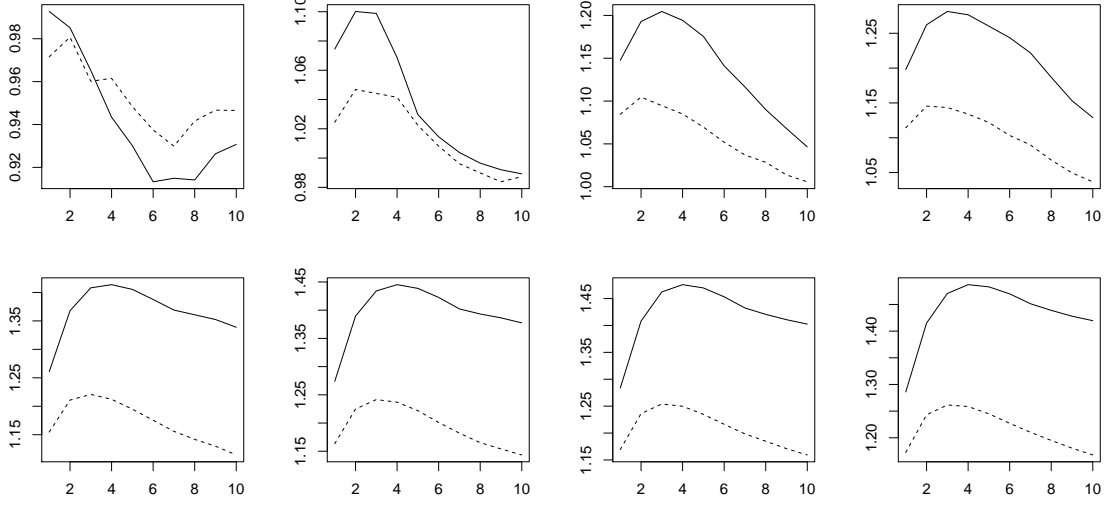
Figure 14: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (23) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.2401	0.2135	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.1045	0.1435
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1712	0.3752	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.1276	0.6244
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	0.0922	0.38	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0.1509	0.3644
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	0.1079	0.4199	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	0.0879	0.3968
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	1	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.9998	2e-04
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0	0
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	0.2392	0.5791	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0.6439	0.3225
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	0.0699	0.1118	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	0.0254	0.0082
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.4	< 1.4		≥ 1.4	< 1.4
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.4$	2e-04	0.1541	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.4$	0.0428	0.6228
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.4$	7e-04	0.845	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.4$	0.0282	0.3062
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.6	< 1.6		≥ 1.6	< 1.6
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.6$	0	0.0084	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.6$	3e-04	0.2679
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.6$	0	0.9916	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.6$	0.0011	0.7307

Table 33: *Proportions of the individual events in (32) for the process (23) and selected combinations of n and δ.*

	(32) holds for $h=1$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.6116	0.5964	0.7286	0.892	0.977	0.9985	0.9999	1	1
0.01	0.6153	0.5848	0.6851	0.8555	0.9608	0.9963	0.9995	0.9998	1
0.05	0.6363	0.5731	0.5608	0.6818	0.8296	0.951	0.9823	0.9932	0.9965
0.1	0.6772	0.6075	0.4997	0.5094	0.6009	0.7432	0.8304	0.8787	0.9183
0.15	0.724	0.6635	0.534	0.4612	0.4395	0.4783	0.5363	0.5852	0.6224
0.2	0.7679	0.72	0.6055	0.5121	0.4366	0.3819	0.3629	0.3464	0.351
0.4	0.8962	0.9043	0.872	0.8406	0.8242	0.8245	0.8343	0.8311	0.8452
0.6	0.954	0.9713	0.9691	0.9634	0.966	0.976	0.9853	0.9876	0.9916
	(32) holds for $h=5$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.7218	0.686	0.7029	0.794	0.9175	0.9829	0.9965	0.9978	0.9998
0.01	0.7289	0.6839	0.6846	0.7635	0.8969	0.9756	0.9944	0.9973	0.9993
0.05	0.759	0.7026	0.6366	0.6501	0.7906	0.9264	0.9705	0.986	0.9937
0.1	0.795	0.7412	0.6237	0.5603	0.6299	0.791	0.8865	0.9261	0.9552
0.15	0.8263	0.7804	0.649	0.5295	0.5088	0.6172	0.7215	0.7856	0.839
0.2	0.8513	0.8117	0.6862	0.5477	0.4523	0.4613	0.5319	0.5915	0.6521
0.4	0.91	0.8971	0.8355	0.7496	0.63	0.5083	0.4426	0.3829	0.349
0.6	0.942	0.9363	0.9131	0.8852	0.8411	0.8022	0.7743	0.7462	0.731

Table 34: *Proportion of* (32) *being fulfilled for the process* (23) *and different values of h, δ and n.*

	Value of $f(\delta)$, defined in (10), for $h=1$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.14	0.16	0.19	0.2	0.21	0.22	0.22	0.23	0.23
0.05	0.099	0.12	0.15	0.17	0.19	0.21	0.22	0.22	0.23
0.1	0.045	0.068	0.098	0.12	0.14	0.15	0.16	0.17	0.18
0.15	3e-05	0.015	0.046	0.067	0.086	0.1	0.11	0.12	0.12
0.2	2.5e-05	7.1e-05	0.00011	0.015	0.034	0.052	0.062	0.068	0.073
0.4	0.14	0.12	0.088	0.064	0.042	0.021	0.0083	1.9e-05	4.9e-05
0.6	0.34	0.32	0.29	0.27	0.24	0.22	0.21	0.2	0.19
	Value of $f(\delta)$, defined in (10), for $h=5$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.096	0.12	0.16	0.18	0.2	0.21	0.22	0.22	0.23
0.05	0.24	0.53	0.79	0.9	0.98	1.1	1.1	1.1	1.1
0.1	0.011	0.29	0.71	1	1.3	1.5	1.6	1.7	1.8
0.15	0.00098	0.048	0.46	0.76	1	1.3	1.4	1.5	1.6
0.2	7.4e-07	0.00072	0.22	0.51	0.77	1	1.2	1.2	1.3
0.4	0.18	0.00052	0.00087	1.7e-06	0.00022	0.039	0.18	0.26	0.34
0.6	1	0.47	0.0013	0.0021	0.0012	0.00027	0.0011	0.00071	0.00074

Table 35: *Values of $f(\delta)$, defined in (10), for the process* (23) *and different values of h, δ and n.*

δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
				(32) holds for $h=1$, $i=2$					
0	0.4369	0.5535	0.7806	0.9193	0.9832	0.9987	0.9999	1	1
0.01	0.4113	0.5197	0.7403	0.8897	0.9712	0.9968	0.9997	0.9999	1
0.05	0.3177	0.3939	0.572	0.7397	0.8674	0.9611	0.9847	0.994	0.9966
0.1	0.2389	0.2776	0.3916	0.5189	0.6519	0.7839	0.8521	0.8914	0.9262
0.15	0.1833	0.1984	0.2572	0.3333	0.4123	0.5032	0.5698	0.6169	0.6522
0.2	0.1407	0.1384	0.163	0.2001	0.2314	0.2602	0.2862	0.2939	0.3091
0.4	0.0528	0.0301	0.0215	0.0151	0.0121	0.0061	0.0031	0.0021	9e-04
0.6	0.024	0.0081	0.0021	0.0013	4e-04	1e-04	0	0	0
				(32) holds for $h=1$, $i=3$					
0	0.4745	0.6387	0.8644	0.9647	0.9936	0.9998	1	1	1
0.01	0.4536	0.6067	0.8374	0.9512	0.989	0.9995	0.9998	0.9999	1
0.05	0.377	0.5004	0.7268	0.8699	0.9526	0.9897	0.9976	0.9992	0.9999
0.1	0.3061	0.3925	0.5873	0.7415	0.859	0.9397	0.9725	0.9859	0.9921
0.15	0.2515	0.3055	0.4646	0.6049	0.7268	0.8339	0.8909	0.9295	0.9456
0.2	0.2012	0.2402	0.3589	0.4722	0.586	0.6833	0.7453	0.7957	0.8183
0.4	0.088	0.0806	0.1119	0.1487	0.1659	0.1714	0.1634	0.1672	0.1543
0.6	0.0424	0.026	0.029	0.0355	0.0336	0.0239	0.0147	0.0124	0.0084
				(32) holds for $h=5$, $i=2$					
0	0.2489	0.336	0.549	0.7606	0.918	0.9864	0.9971	0.9983	0.9998
0.01	0.2321	0.3114	0.5185	0.732	0.902	0.9805	0.9956	0.9978	0.9994
0.05	0.1757	0.2382	0.4066	0.6059	0.8069	0.9394	0.9742	0.9876	0.9941
0.1	0.1307	0.1738	0.2988	0.4535	0.6476	0.8196	0.8988	0.932	0.9582
0.15	0.1015	0.1318	0.2216	0.3372	0.4848	0.6468	0.7489	0.8026	0.8485
0.2	0.0812	0.1004	0.1624	0.2388	0.3406	0.4621	0.5577	0.6143	0.6693
0.4	0.0395	0.0447	0.0623	0.0634	0.0623	0.0641	0.0671	0.0696	0.071
0.6	0.026	0.028	0.0323	0.0233	0.0126	0.0075	0.0029	0.0031	0.0014
				(32) holds for $h=5$, $i=3$					
0	0.2597	0.3846	0.6393	0.8392	0.9611	0.9955	0.9992	0.9995	1
0.01	0.248	0.3679	0.6213	0.8279	0.9555	0.9939	0.9986	0.9995	0.9999
0.05	0.2043	0.3092	0.5476	0.7684	0.9283	0.984	0.9961	0.9984	0.9996
0.1	0.1627	0.2454	0.4627	0.6874	0.8675	0.9578	0.9863	0.9935	0.9966
0.15	0.1348	0.201	0.3924	0.6009	0.7886	0.9138	0.9602	0.98	0.9889
0.2	0.1137	0.1663	0.3302	0.5153	0.7031	0.8534	0.9132	0.9498	0.9664
0.4	0.0653	0.0856	0.1626	0.2506	0.3649	0.4864	0.5589	0.6247	0.6656
0.6	0.041	0.0491	0.0842	0.1183	0.1577	0.1955	0.2246	0.2525	0.2682

Table 36: *Proportion of* (7) *being fulfilled for the process* (23) *and different values of h, δ and n.*
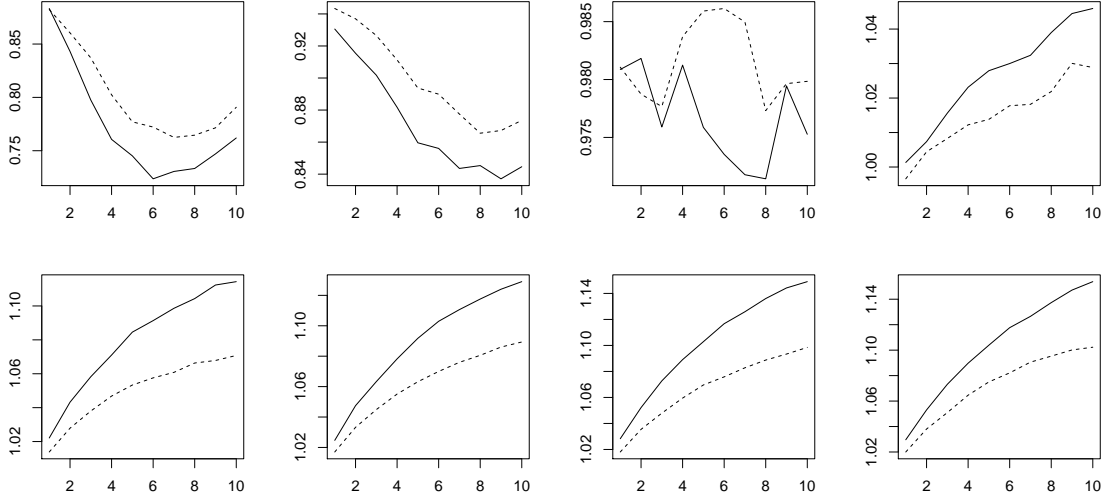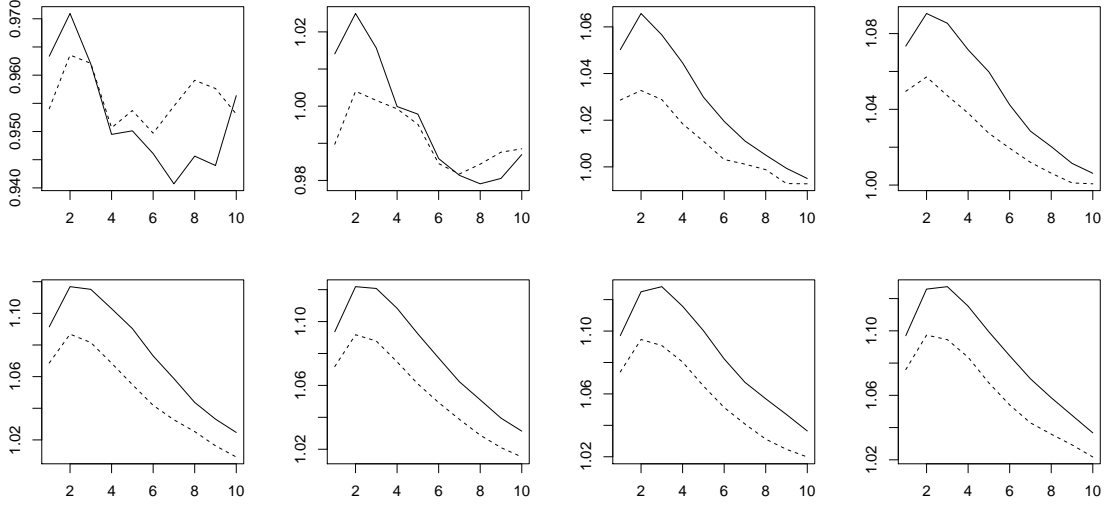
Figure 15: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (24) and different values of n (from left to right: n=100, n=200, n=500, n=1000 [first row], n=4000, n=6000, n=8000, n=10000 [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.1703	0.2128	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.0795	0.1501
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1778	0.4391	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.1372	0.6332
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0.2776	0.3122	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0.1577	0.2497
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0.1697	0.2405	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0.1284	0.4642
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.9997	3e-04	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.9741	0.0185
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.0071	3e-04
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0.7686	0.168	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0.574	0.2864
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0.0541	0.0093	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0.0947	0.0449
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0.0835	0.3575	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0.0988	0.375
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0.1111	0.4479	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0.1015	0.4247
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.15	< 1.15		≥ 1.15	< 1.15
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.15$	9e-04	0.074	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.15$	0.0057	0.1586
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.15$	0.0122	0.9129	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.15$	0.0247	0.811

Table 37: *Proportions of the individual events in (32) for the process (24) and selected combinations of n and δ.*

	(32) holds for $h=1$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.5993	0.5685	0.6091	0.7451	0.8887	0.9782	0.9944	0.9984	0.9997
0.01	0.6094	0.5714	0.5797	0.6737	0.8231	0.9471	0.9815	0.9928	0.9977
0.05	0.6678	0.6201	0.5476	0.5181	0.5292	0.6068	0.6807	0.7397	0.7779
0.1	0.7327	0.7093	0.6683	0.6175	0.5923	0.5553	0.5478	0.5274	0.5314
0.15	0.7916	0.7963	0.799	0.7919	0.8198	0.8547	0.8788	0.8967	0.9138
0.2	0.8406	0.8675	0.8878	0.9059	0.9364	0.9679	0.9828	0.9878	0.9938
0.4	0.945	0.9754	0.9919	0.9968	0.9996	1	1	1	1
0.6	0.9739	0.9935	0.9993	0.9999	1	1	1	1	1
	(32) holds for $h=5$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0.7058	0.6763	0.6509	0.6454	0.7256	0.8519	0.9299	0.9556	0.9744
0.01	0.7127	0.6872	0.6521	0.6188	0.6617	0.7821	0.8714	0.9121	0.9471
0.05	0.7577	0.7481	0.6969	0.6219	0.5453	0.5217	0.5467	0.5883	0.6189
0.1	0.8059	0.8105	0.7818	0.731	0.6543	0.582	0.5597	0.5269	0.5235
0.15	0.8414	0.8571	0.8472	0.8225	0.8081	0.7965	0.8058	0.8052	0.8167
0.2	0.872	0.8848	0.8895	0.8869	0.9007	0.9231	0.9406	0.9455	0.9569
0.4	0.9293	0.948	0.967	0.9738	0.986	0.997	0.9996	0.9998	1
0.6	0.9539	0.9698	0.9867	0.9912	0.9949	0.9994	0.9999	0.9999	1

Table 38: *Proportion of (32) being fulfilled for the process (24) and different values of h, δ and n.*

	Value of $f(\delta)$, defined in (10), for $h=1$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.044	0.044	0.045	0.045	0.046	0.047	0.047	0.047	0.047
0.05	0.01	0.017	0.024	0.029	0.034	0.037	0.039	0.04	0.041
0.1	0.018	0.014	0.0092	0.0053	0.002	1.2e-05	3.5e-05	0.00055	0.00022
0.15	0.069	0.065	0.059	0.056	0.052	0.049	0.047	0.046	0.046
0.2	0.12	0.11	0.11	0.11	0.1	0.099	0.098	0.096	0.096
0.4	0.32	0.32	0.31	0.31	0.3	0.3	0.3	0.3	0.3
0.6	0.52	0.52	0.51	0.51	0.5	0.5	0.5	0.5	0.5
	Value of $f(\delta)$, defined in (10), for $h=5$								
δ	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=4000$	$n=6000$	$n=8000$	$n=10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.039	0.042	0.044	0.045	0.046	0.046	0.047	0.047	0.047
0.05	3.6e-06	0.0051	0.058	0.089	0.12	0.14	0.15	0.16	0.16
0.1	0.039	9.4e-05	7.8e-05	7.3e-05	2.1e-05	5.6e-05	1.6e-06	0.00011	4.1e-05
0.15	0.19	0.14	0.086	0.055	0.029	0.0082	7.9e-05	6.8e-05	5.6e-05
0.2	0.16	0.29	0.25	0.22	0.19	0.17	0.16	0.15	0.15
0.4	0.37	0.23	0.11	0.038	8.7e-05	0.02	0.048	0.065	0.08
0.6	0.98	0.86	0.75	0.68	0.64	0.6	0.57	0.56	0.55

Table 39: *Values of $f(\delta)$, defined in (10), for the process (24) and different values of h, δ and n.*

	(32) holds for $h=1$, $i=2$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.3792	0.4583	0.6423	0.8055	0.9225	0.9854	0.9958	0.9988	0.9997
0.01	0.3481	0.4165	0.5806	0.7373	0.8727	0.9634	0.9863	0.9944	0.9985
0.05	0.2595	0.2806	0.3544	0.4473	0.5494	0.6719	0.7423	0.7895	0.8227
0.1	0.1854	0.1727	0.1714	0.1889	0.1976	0.2084	0.2057	0.2026	0.1946
0.15	0.1325	0.1073	0.081	0.0732	0.0551	0.0335	0.0222	0.0189	0.0131
0.2	0.0982	0.0656	0.0374	0.0254	0.0139	0.0051	0.0013	0.0015	1e-04
0.4	0.0329	0.0112	0.001	3e-04	0	0	0	0	0
0.6	0.0165	0.0048	1e-04	0	0	0	0	0	0
	(32) holds for $h=1$, $i=3$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.4091	0.526	0.7302	0.881	0.9578	0.9928	0.9986	0.9996	1
0.01	0.3831	0.4873	0.6761	0.83	0.9308	0.9829	0.9952	0.9984	0.9992
0.05	0.2931	0.3517	0.4704	0.5898	0.7142	0.8187	0.8794	0.9182	0.9366
0.1	0.2157	0.2282	0.2747	0.3308	0.3663	0.4019	0.4199	0.4518	0.441
0.15	0.1651	0.1522	0.154	0.1631	0.1437	0.1202	0.104	0.0874	0.0749
0.2	0.1224	0.0951	0.0862	0.0747	0.0505	0.027	0.0159	0.0107	0.0061
0.4	0.0463	0.0212	0.0073	0.0029	4e-04	0	0	0	0
0.6	0.0258	0.0073	8e-04	1e-04	0	0	0	0	0
	(32) holds for $h=5$, $i=2$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.2381	0.2744	0.3839	0.5387	0.7317	0.8854	0.9445	0.9652	0.9808
0.01	0.2167	0.244	0.3406	0.4772	0.6636	0.8256	0.9008	0.9304	0.9597
0.05	0.1611	0.1626	0.211	0.2861	0.3856	0.5006	0.5796	0.634	0.6687
0.1	0.1135	0.1081	0.1244	0.1488	0.166	0.181	0.1887	0.1997	0.2003
0.15	0.085	0.0758	0.0817	0.0811	0.0738	0.0573	0.0436	0.0372	0.0304
0.2	0.0646	0.0561	0.053	0.0489	0.0355	0.0166	0.0087	0.0051	0.0043
0.4	0.0305	0.0232	0.0167	0.0116	0.0061	0.0015	4e-04	1e-04	1e-04
0.6	0.0192	0.0138	0.0065	0.0038	0.0019	3e-04	0	0	0
	(32) holds for $h=5$, $i=3$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.2453	0.3009	0.4392	0.6167	0.8003	0.9287	0.9732	0.9868	0.9922
0.01	0.2296	0.2754	0.4013	0.5706	0.7581	0.9005	0.9548	0.9773	0.986
0.05	0.178	0.2019	0.2831	0.4074	0.5485	0.7025	0.7807	0.8379	0.8604
0.1	0.1368	0.1408	0.1876	0.2518	0.3241	0.3914	0.4224	0.467	0.4738
0.15	0.1094	0.1045	0.1251	0.1582	0.1731	0.1774	0.1704	0.1724	0.1643
0.2	0.088	0.0831	0.0881	0.1006	0.0912	0.0703	0.0551	0.0504	0.0402
0.4	0.0466	0.0354	0.0245	0.0222	0.0111	0.0033	4e-04	3e-04	1e-04
0.6	0.0301	0.0194	0.0098	0.0068	0.0034	5e-04	1e-04	1e-04	0

Table 40: *Proportion of (7) being fulfilled for the process (24) and different values of h, δ and n.*

108
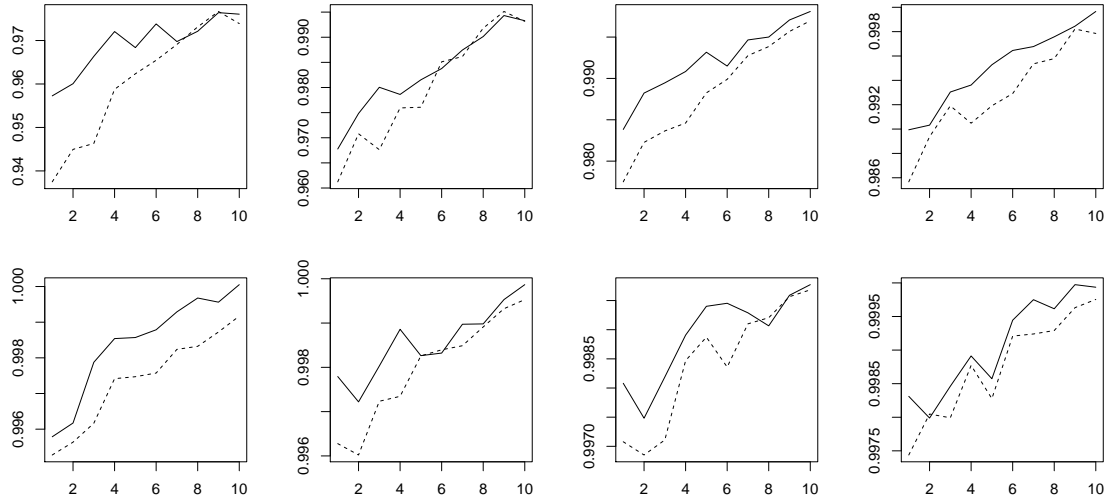
Figure 16: *Plot of $h \mapsto R_{T,i}^{s,ls}(h)$ for model (25) and different values of n (from left to right: n=100, n=200, n=500, n=1000 [first row], n=4000, n=6000, n=8000, n=10000 [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{s,ls}(1)$			$R_{T,2}^{s,ls}(5)$	
		≥ 1	< 1		≥ 1	< 1
100	$R_{T,3}^{s,ls}(1) \geq 1$	0.2098	0.2057	$R_{T,3}^{s,ls}(5) \geq 1$	0.4545	0.1547
	$R_{T,3}^{s,ls}(1) < 1$	0.1876	0.3969	$R_{T,3}^{s,ls}(5) < 1$	0.1567	0.2341
n		$R_{T,2}^{s,ls}(1)$			$R_{T,2}^{s,ls}(5)$	
		≥ 1	< 1		≥ 1	< 1
1000	$R_{T,3}^{s,ls}(1) \geq 1$	0.0869	0.2148	$R_{T,3}^{s,ls}(5) \geq 1$	0.4039	0.1789
	$R_{T,3}^{s,ls}(1) < 1$	0.187	0.5113	$R_{T,3}^{s,ls}(5) < 1$	0.1505	0.2667
n		$R_{T,2}^{s,ls}(1)$			$R_{T,2}^{s,ls}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{s,ls}(1) \geq 1$	0.0709	0.2013	$R_{T,3}^{s,ls}(5) \geq 1$	0.3138	0.183
	$R_{T,3}^{s,ls}(1) < 1$	0.1821	0.5457	$R_{T,3}^{s,ls}(5) < 1$	0.1728	0.3304
n		$R_{T,2}^{s,ls}(1)$			$R_{T,2}^{s,ls}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
10000	$R_{T,3}^{s,ls}(1) \geq 1.01$	0	0.0021	$R_{T,3}^{s,ls}(5) \geq 1.01$	0.0029	0.021
	$R_{T,3}^{s,ls}(1) < 1.01$	0.0012	0.9967	$R_{T,3}^{s,ls}(5) < 1.01$	0.0145	0.9616
n		$R_{T,2}^{s,ls}(1)$			$R_{T,2}^{s,ls}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{s,ls}(1) \geq 1.05$	0	0	$R_{T,3}^{s,ls}(5) \geq 1.05$	0	2e-04
	$R_{T,3}^{s,ls}(1) < 1.05$	0	1	$R_{T,3}^{s,ls}(5) < 1.05$	1e-04	0.9997
n		$R_{T,2}^{s,ls}(1)$			$R_{T,2}^{s,ls}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{s,ls}(1) \geq 1.1$	0	0	$R_{T,3}^{s,ls}(5) \geq 1.1$	0	0
	$R_{T,3}^{s,ls}(1) < 1.1$	0	1	$R_{T,3}^{s,ls}(5) < 1.1$	0	1

Table 41: *Proportions of the individual events in (32) for the process (25) and selected combinations of n and δ.*

	(32) holds for $h = 1$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.6067	0.6027	0.6065	0.5982	0.6032	0.609	0.6252	0.6077	0.6166
0.01	0.6271	0.6528	0.7084	0.7653	0.8584	0.9527	0.9814	0.9913	0.9967
0.05	0.7177	0.8075	0.93	0.985	0.999	1	1	1	1
0.1	0.803	0.9049	0.9884	0.9996	1	1	1	1	1
0.15	0.8607	0.9521	0.9969	1	1	1	1	1	1
0.2	0.8973	0.9719	0.9976	0.9999	1	1	1	1	1
0.4	0.9599	0.9858	0.996	0.9997	1	1	1	1	1
0.6	0.9776	0.9902	0.9969	0.9997	1	1	1	1	1
	(32) holds for $h = 5$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.6886	0.6837	0.6885	0.6706	0.6791	0.6623	0.667	0.6594	0.6442
0.01	0.7008	0.7124	0.75	0.7844	0.8366	0.8966	0.927	0.9466	0.9645
0.05	0.766	0.8187	0.8988	0.9462	0.9812	0.9939	0.9979	0.999	0.9997
0.1	0.8313	0.8914	0.9573	0.9854	0.9972	0.9995	1	0.9999	1
0.15	0.871	0.9308	0.9804	0.9946	0.9993	0.9999	1	1	1
0.2	0.9013	0.9505	0.9896	0.998	1	1	1	1	1
0.4	0.9556	0.9825	0.9989	0.9998	1	1	1	1	1
0.6	0.9758	0.9926	0.9998	1	1	1	1	1	1

Table 42: *Proportion of* (32) *being fulfilled for the process* (25) *and different values of h,* δ *and n.*

	Value of $f(\delta)$, defined in (10), for $h = 1$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
0.4	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
0.6	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037
	Value of $f(\delta)$, defined in (10), for $h = 5$								
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.0061	0.0061	0.0061	0.0061	0.0061	0.0061	0.0061	0.0061	0.0061
0.05	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068
0.1	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
0.15	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22
0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
0.4	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61
0.6	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92

Table 43: *Values of* $f(\delta)$, *defined in* (10), *for the process* (25) *and different values of h,* δ *and n.*

		(32) holds for $h=1$, $i=2$							
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.3048	0.2943	0.2827	0.2727	0.2694	0.2563	0.2522	0.2552	0.253
0.01	0.2752	0.2439	0.1868	0.1315	0.0713	0.0242	0.0091	0.0034	0.0012
0.05	0.1872	0.1228	0.0442	0.0085	4e-04	0	0	0	0
0.1	0.1182	0.0639	0.0128	0.0011	0	0	0	0	0
0.15	0.0805	0.0377	0.0085	8e-04	0	0	0	0	0
0.2	0.0596	0.0251	0.0075	8e-04	0	0	0	0	0
0.4	0.0214	0.0113	0.0037	3e-04	0	0	0	0	0
0.6	0.0116	0.0072	0.0026	2e-04	0	0	0	0	0
		(32) holds for $h=1$, $i=3$							
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.3229	0.3234	0.3056	0.3005	0.2894	0.2853	0.2744	0.2819	0.2722
0.01	0.2929	0.2647	0.1996	0.1482	0.0833	0.0251	0.0097	0.0053	0.0021
0.05	0.2053	0.1385	0.0458	0.0087	6e-04	0	0	0	0
0.1	0.1378	0.0658	0.0138	9e-04	0	0	0	0	0
0.15	0.099	0.036	0.0086	8e-04	0	0	0	0	0
0.2	0.0725	0.0238	0.0075	7e-04	0	0	0	0	0
0.4	0.0303	0.0115	0.0051	6e-04	0	0	0	0	0
0.6	0.0164	0.0064	0.0027	5e-04	0	0	0	0	0
		(32) holds for $h=5$, $i=2$							
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2527	0.2462	0.2297	0.2329	0.2382	0.2386	0.2556	0.2588	0.2714
0.01	0.2218	0.1999	0.1535	0.1237	0.0894	0.0496	0.0373	0.0248	0.0174
0.05	0.1468	0.1023	0.0502	0.024	0.0086	0.002	8e-04	5e-04	1e-04
0.1	0.0962	0.0585	0.02	0.0063	0.0015	2e-04	0	1e-04	0
0.15	0.0679	0.0357	0.0078	0.002	2e-04	0	0	0	0
0.2	0.0493	0.0239	0.0035	7e-04	0	0	0	0	0
0.4	0.0205	0.0084	4e-04	0	0	0	0	0	0
0.6	0.0107	0.0033	0	0	0	0	0	0	0
		(32) holds for $h=5$, $i=3$							
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2507	0.2529	0.2478	0.2613	0.2593	0.2717	0.2758	0.2782	0.2816
0.01	0.222	0.2097	0.1781	0.1437	0.1036	0.0676	0.0457	0.0352	0.0239
0.05	0.1532	0.1132	0.0658	0.0344	0.0108	0.0041	0.0013	5e-04	2e-04
0.1	0.1033	0.0635	0.0259	0.0091	0.0013	3e-04	0	0	0
0.15	0.0755	0.0391	0.0124	0.0034	5e-04	1e-04	0	0	0
0.2	0.0576	0.0274	0.0069	0.0013	0	0	0	0	0
0.4	0.0259	0.0093	7e-04	2e-04	0	0	0	0	0
0.6	0.0143	0.0043	2e-04	0	0	0	0	0	0

Table 44: *Proportion of* (7) *being fulfilled for the process* (25) *and different values of h, δ and n.*

Figure 17: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (26) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.3309	0.1776	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.4733	0.1481
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.1477	0.3438	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1473	0.2313

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.3363	0.1584	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.488	0.1436
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.1351	0.3702	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.129	0.2394

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.3218	0.1509	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.4882	0.1404
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.141	0.3863	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1213	0.2501

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0	0.0038	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0	0.0052
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.0032	0.993	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	9e-04	0.9939

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0	1

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0	1

Table 45: *Proportions of the individual events in (32) for the process (26) and selected combinations of n and δ.*

	(32) holds for $h = 1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.6747	0.6888	0.6969	0.7065	0.7094	0.7101	0.7111	0.7097	0.7081
0.01	0.6926	0.7188	0.761	0.8106	0.8702	0.9482	0.9766	0.9888	0.993
0.05	0.7647	0.8258	0.9256	0.9776	0.9972	0.9998	1	1	1
0.1	0.8329	0.9127	0.9858	0.9983	1	1	1	1	1
0.15	0.8838	0.9542	0.9966	0.9996	1	1	1	1	1
0.2	0.9169	0.9779	0.9986	1	1	1	1	1	1
0.4	0.9768	0.998	1	1	1	1	1	1	1
0.6	0.9908	0.9998	1	1	1	1	1	1	1
	(32) holds for $h = 5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.7046	0.7173	0.7211	0.7274	0.7276	0.7358	0.7337	0.7394	0.7383
0.01	0.7349	0.7713	0.8265	0.8783	0.9293	0.9678	0.9831	0.9883	0.9939
0.05	0.8168	0.8839	0.9527	0.9837	0.9962	0.9997	1	1	1
0.1	0.8805	0.9441	0.9858	0.9973	0.9998	1	1	1	1
0.15	0.9143	0.9675	0.9947	0.9998	1	1	1	1	1
0.2	0.9367	0.9795	0.9983	0.9999	1	1	1	1	1
0.4	0.9754	0.9957	0.9999	1	1	1	1	1	1
0.6	0.9867	0.9987	1	1	1	1	1	1	1

Table 46: *Proportion of* (32) *being fulfilled for the process* (26) *and different values of h, δ and n.*

	Value of $f(\delta)$, defined in (10), for $h = 1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
	Value of $f(\delta)$, defined in (10), for $h = 5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

Table 47: *Values of $f(\delta)$, defined in (10), for the process* (26) *and different values of h, δ and n.*

δ	(32) holds for $h=1$, $i=2$								
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2183	0.2192	0.208	0.2034	0.2065	0.1965	0.197	0.2044	0.2022
0.01	0.1981	0.1857	0.1456	0.1121	0.0679	0.0253	0.0104	0.0049	0.0032
0.05	0.1394	0.099	0.0353	0.0107	7e-04	1e-04	0	0	0
0.1	0.0889	0.0456	0.0061	4e-04	0	0	0	0	0
0.15	0.0581	0.0219	0.0017	1e-04	0	0	0	0	0
0.2	0.0401	0.0115	5e-04	0	0	0	0	0	0
0.4	0.0109	8e-04	0	0	0	0	0	0	0
0.6	0.0048	0	0	0	0	0	0	0	0

δ	(32) holds for $h=1$, $i=3$								
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2482	0.235	0.2291	0.2267	0.2217	0.2186	0.2169	0.2201	0.2121
0.01	0.2263	0.2003	0.1628	0.1233	0.0745	0.0287	0.0134	0.0065	0.0038
0.05	0.1563	0.107	0.0425	0.0117	0.0021	1e-04	0	0	0
0.1	0.1046	0.0501	0.0081	0.0013	0	0	0	0	0
0.15	0.0689	0.0257	0.0017	3e-04	0	0	0	0	0
0.2	0.048	0.0112	9e-04	0	0	0	0	0	0
0.4	0.0125	0.0012	0	0	0	0	0	0	0
0.6	0.0044	2e-04	0	0	0	0	0	0	0

δ	(32) holds for $h=5$, $i=2$								
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2426	0.2257	0.222	0.2123	0.2154	0.1964	0.2049	0.1979	0.1966
0.01	0.1946	0.1534	0.105	0.0663	0.0362	0.0144	0.0064	0.0044	9e-04
0.05	0.1136	0.0624	0.0207	0.0061	0.0013	0	0	0	0
0.1	0.0668	0.0266	0.005	5e-04	1e-04	0	0	0	0
0.15	0.0427	0.0131	0.0012	0	0	0	0	0	0
0.2	0.0302	0.0075	2e-04	0	0	0	0	0	0
0.4	0.011	9e-04	0	0	0	0	0	0	0
0.6	0.0061	2e-04	0	0	0	0	0	0	0

δ	(32) holds for $h=5$, $i=3$								
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2434	0.2348	0.2309	0.2269	0.2296	0.2216	0.2254	0.2193	0.2157
0.01	0.1969	0.1597	0.1157	0.0742	0.0399	0.0196	0.0109	0.0075	0.0052
0.05	0.1168	0.0707	0.0298	0.0104	0.0025	3e-04	0	0	0
0.1	0.0745	0.0331	0.0094	0.0022	1e-04	0	0	0	0
0.15	0.0512	0.0202	0.0041	2e-04	0	0	0	0	0
0.2	0.0367	0.0132	0.0015	1e-04	0	0	0	0	0
0.4	0.014	0.0034	1e-04	0	0	0	0	0	0
0.6	0.0072	0.0011	0	0	0	0	0	0	0

Table 48: *Proportion of* (7) *being fulfilled for the process* (26) *and different values of h, δ and n.*
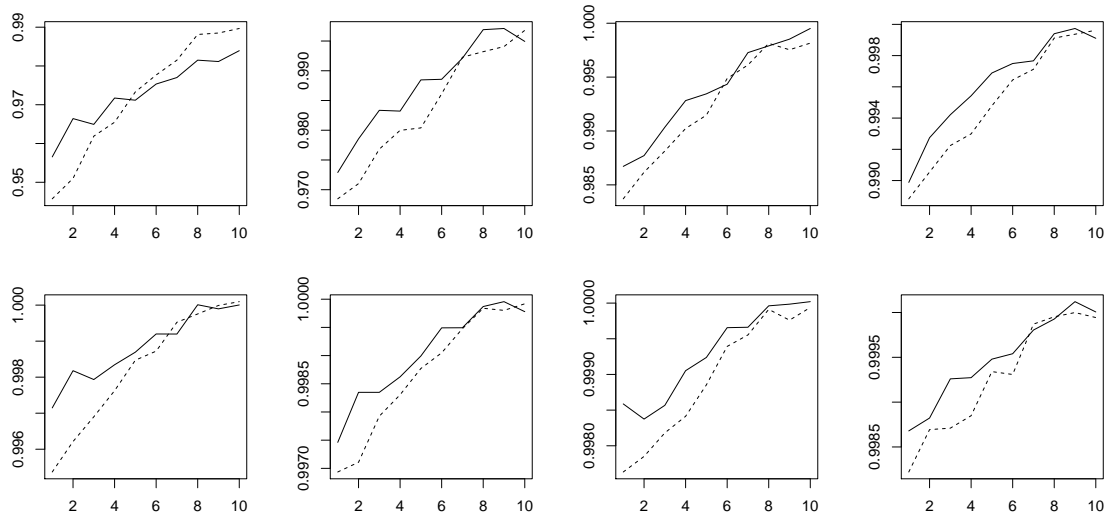
Figure 18: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (27) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i=3$ (test set), dashed line: $i=2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.3101	0.1865	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.3966	0.1642
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.1585	0.3449	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1613	0.2779
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.2802	0.1909	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.3757	0.178
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.1619	0.367	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1608	0.2855
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.249	0.1897	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.3532	0.1798
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.1715	0.3898	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1587	0.3083
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0	0.009	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.0026	0.0212
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.0048	0.9862	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.0116	0.9646
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0	1
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0	1	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0	1

Table 49: *Proportions of the individual events in (32) for the process (27) and selected combinations of n and δ.*

	(32) holds for $h=1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.655	0.6537	0.6518	0.6472	0.6483	0.6417	0.6419	0.6454	0.6388
0.01	0.6741	0.6829	0.7162	0.7536	0.8268	0.9198	0.9593	0.9762	0.9862
0.05	0.7415	0.7965	0.889	0.962	0.9935	0.9992	1	1	1
0.1	0.8066	0.879	0.9707	0.9952	0.9997	1	1	1	1
0.15	0.8544	0.9308	0.9892	0.9989	1	1	1	1	1
0.2	0.8879	0.959	0.9959	0.9998	1	1	1	1	1
0.4	0.9571	0.9923	0.9997	1	1	1	1	1	1
0.6	0.9784	0.9977	1	1	1	1	1	1	1
	(32) holds for $h=5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.6745	0.6759	0.671	0.6612	0.663	0.6552	0.6586	0.6672	0.6615
0.01	0.7072	0.7284	0.7882	0.8344	0.8929	0.9372	0.9527	0.9602	0.9672
0.05	0.7873	0.8565	0.9289	0.9563	0.9764	0.9928	0.9989	0.9997	1
0.1	0.8514	0.9209	0.9557	0.9757	0.9962	0.9995	1	1	1
0.15	0.887	0.9451	0.9668	0.9887	0.999	1	1	1	1
0.2	0.9063	0.9538	0.9759	0.9956	0.9999	1	1	1	1
0.4	0.9471	0.9742	0.9947	0.9999	1	1	1	1	1
0.6	0.9606	0.9851	0.9988	1	1	1	1	1	1

Table 50: *Proportion of* (32) *being fulfilled for the process* (27) *and different values of h, δ and n.*

	Value of $f(\delta)$, defined in (10), for $h=1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.038	0.035	0.032	0.029	0.027	0.025	0.024	0.023	0.023
0.05	0.19	0.18	0.16	0.15	0.13	0.12	0.12	0.12	0.11
0.1	0.38	0.35	0.32	0.29	0.27	0.25	0.24	0.23	0.23
0.15	0.57	0.53	0.48	0.44	0.4	0.37	0.36	0.35	0.34
0.2	0.76	0.7	0.63	0.58	0.54	0.5	0.48	0.46	0.45
0.4	1.5	1.4	1.3	1.2	1.1	1	0.95	0.93	0.91
0.6	2.3	2.1	1.9	1.7	1.6	1.5	1.4	1.4	1.4
	Value of $f(\delta)$, defined in (10), for $h=5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.061	0.046	0.036	0.031	0.028	0.025	0.024	0.023	0.023
0.05	0.3	0.23	0.18	0.15	0.14	0.13	0.12	0.12	0.11
0.1	0.61	0.46	0.36	0.31	0.28	0.25	0.24	0.23	0.23
0.15	0.91	0.69	0.54	0.46	0.42	0.38	0.36	0.35	0.34
0.2	1.2	0.91	0.71	0.62	0.56	0.51	0.48	0.47	0.46
0.4	2.4	1.8	1.4	1.2	1.1	1	0.97	0.94	0.91
0.6	3.6	2.7	2.1	1.9	1.7	1.5	1.4	1.4	1.4

Table 51: *Values of $f(\delta)$, defined in (10), for the process* (27) *and different values of h, δ and n.*

δ				(32) holds for $h = 1$, $i = 2$					
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2398	0.2396	0.2411	0.2471	0.2493	0.2372	0.239	0.2495	0.2467
0.01	0.2158	0.2023	0.1708	0.1392	0.0876	0.034	0.0152	0.0092	0.0048
0.05	0.1429	0.1051	0.0462	0.0131	0.0019	0	0	0	0
0.1	0.0951	0.0504	0.0102	6e-04	0	0	0	0	0
0.15	0.0633	0.0254	0.0026	1e-04	0	0	0	0	0
0.2	0.0437	0.0126	6e-04	0	0	0	0	0	0
0.4	0.013	0.002	0	0	0	0	0	0	0
0.6	0.0055	2e-04	0	0	0	0	0	0	0
δ				(32) holds for $h = 1$, $i = 3$					
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2678	0.2717	0.2745	0.2761	0.27	0.2741	0.2677	0.2637	0.2649
0.01	0.2487	0.2394	0.2042	0.1654	0.1066	0.0522	0.0263	0.015	0.009
0.05	0.1878	0.1452	0.0738	0.0261	0.0046	8e-04	0	0	0
0.1	0.1357	0.0832	0.0199	0.0042	3e-04	0	0	0	0
0.15	0.0999	0.0496	0.0082	0.001	0	0	0	0	0
0.2	0.0782	0.0304	0.0035	2e-04	0	0	0	0	0
0.4	0.0307	0.0057	3e-04	0	0	0	0	0	0
0.6	0.0161	0.0021	0	0	0	0	0	0	0
δ				(32) holds for $h = 5$, $i = 2$					
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2945	0.2998	0.2999	0.2976	0.2932	0.283	0.2831	0.2885	0.2851
0.01	0.2398	0.2101	0.1549	0.1175	0.074	0.0396	0.0262	0.0232	0.0142
0.05	0.1488	0.0977	0.0467	0.0258	0.0085	0.0011	1e-04	0	0
0.1	0.0956	0.0532	0.0214	0.0074	4e-04	0	0	0	0
0.15	0.0669	0.0346	0.0109	0.0026	0	0	0	0	0
0.2	0.05	0.0262	0.006	6e-04	0	0	0	0	0
0.4	0.0236	0.01	6e-04	0	0	0	0	0	0
0.6	0.0144	0.004	1e-04	0	0	0	0	0	0
δ				(32) holds for $h = 5$, $i = 3$					
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2974	0.3045	0.3175	0.3148	0.3178	0.3138	0.3097	0.3027	0.3062
0.01	0.251	0.2365	0.1893	0.1383	0.0905	0.0524	0.0363	0.028	0.0238
0.05	0.1791	0.1272	0.0696	0.0369	0.0169	0.0061	0.001	3e-04	0
0.1	0.1274	0.0765	0.0391	0.0183	0.0034	5e-04	0	0	0
0.15	0.0999	0.0537	0.0261	0.0087	0.001	0	0	0	0
0.2	0.0833	0.0414	0.0189	0.0038	1e-04	0	0	0	0
0.4	0.0453	0.02	0.0047	1e-04	0	0	0	0	0
0.6	0.0308	0.0115	0.0011	0	0	0	0	0	0

Table 52: *Proportion of* (7) *being fulfilled for the process* (27) *and different values of h, δ and n.*
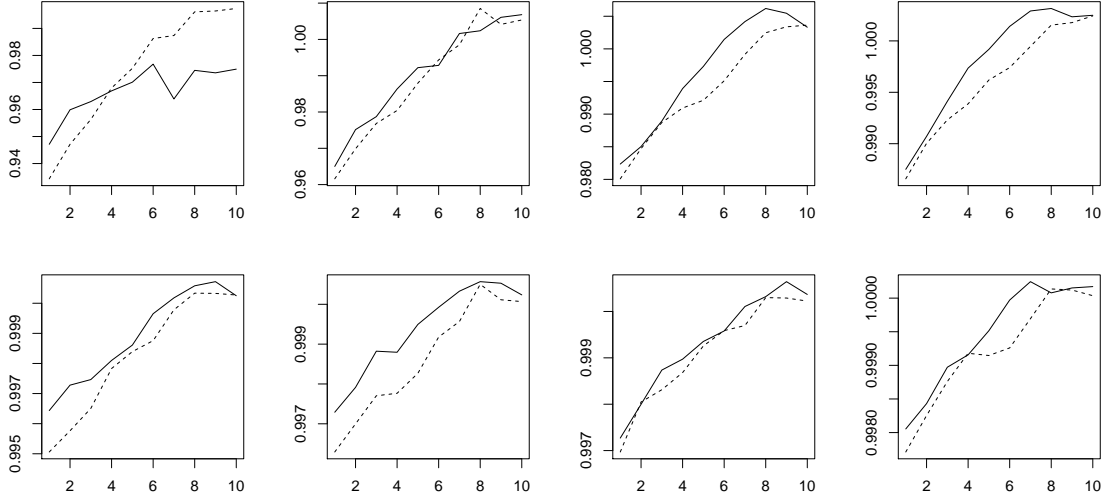
Figure 19: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (28) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.5407	0.1704	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.6052	0.0926
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.1376	0.1513	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1406	0.1616
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.4616	0.2526	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.0404	0.333
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0.2332	0.0526	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.0821	0.5445
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	0.9965	0.0035	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	4e-04	5e-04
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.0675	0.9316
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.9952	0.0048	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0	2e-04
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.0474	0.9524
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0.9856	0.0144	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0.0159	0.9841
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0.9478	0.0522	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0.0024	0.9976

Table 53: *Proportions of the individual events in (32) for the process (28) and selected combinations of n and δ.*

δ	(32) holds for $h = 1$								
	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.692	0.5882	0.431	0.5142	0.6672	0.7326	0.7851	0.9674	0.9965
0.01	0.6915	0.5867	0.4201	0.4993	0.6555	0.7347	0.7704	0.9566	0.9952
0.05	0.6989	0.5877	0.3784	0.445	0.6058	0.7411	0.7059	0.9041	0.9856
0.1	0.7097	0.5944	0.3426	0.396	0.5393	0.7359	0.6099	0.802	0.9478
0.15	0.7276	0.6114	0.3313	0.3711	0.4838	0.7244	0.5148	0.6668	0.8715
0.2	0.7474	0.6284	0.3386	0.383	0.4417	0.7095	0.4284	0.5193	0.7488
0.4	0.8273	0.7147	0.4739	0.5773	0.4935	0.6665	0.1986	0.1108	0.2059
0.6	0.8893	0.7857	0.6429	0.7783	0.7191	0.669	0.1489	0.0105	0.0237
δ	(32) holds for $h = 5$								
	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.7668	0.7467	0.6302	0.5849	0.6015	0.4949	0.7791	0.9101	0.932
0.01	0.7796	0.7551	0.6582	0.624	0.6449	0.6759	0.8643	0.9389	0.9524
0.05	0.8177	0.8129	0.7557	0.7536	0.781	0.8864	0.9597	0.9842	0.9841
0.1	0.8617	0.8608	0.8339	0.8542	0.8701	0.9537	0.9931	0.9987	0.9976
0.15	0.89	0.8914	0.8881	0.9155	0.9217	0.9796	0.9988	0.9999	0.9992
0.2	0.9111	0.9138	0.9225	0.9486	0.9532	0.9906	0.9999	1	0.9997
0.4	0.9549	0.9614	0.9829	0.9916	0.9928	0.9997	1	1	1
0.6	0.9714	0.9788	0.995	0.998	0.9992	1	1	1	1

Table 54: *Proportion of* (32) *being fulfilled for the process* (28) *and different values of h, δ and n.*

δ	Value of $f(\delta)$, defined in (10), for $h = 1$								
	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.0051	0.071	0.006	0.00037	0.00086	0.00064	0.00027	0.027	0.094
0.05	0.27	0.35	0.0097	0.0012	0.0042	0.0052	0.0014	0.019	0.01
0.1	0.61	0.71	0.003	0.00032	0.0025	0.0052	0.011	0.0039	0.041
0.15	0.95	1.1	0.033	0.00082	0.0044	0.00078	0.0011	0.0021	0.063
0.2	1.3	1.4	0.0033	0.0079	0.0014	0.0056	0.0026	0.00093	0.034
0.4	2.6	2.8	0.93	0.02	0.011	0.00081	0.001	0.0066	0.00021
0.6	4	4.3	2.1	0.023	0.007	0.00064	0.00019	0.0015	0.0093
δ	Value of $f(\delta)$, defined in (10), for $h = 5$								
	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	5.1e-06	1e-04	0.00052	0.00035	4e-04	9.9e-05	0.00099	0.00096	0.00025
0.05	0.0041	0.13	0.0058	3.2e-05	0.00017	0.01	0.0074	0.0013	0.0024
0.1	0.21	0.4	0.0035	0.006	0.00045	0.0033	0.0013	0.00065	0.0015
0.15	0.41	0.67	0.11	0.0039	0.018	0.0014	0.0028	0.00034	0.0019
0.2	0.62	0.94	0.42	0.19	0.13	0.066	0.012	0.0063	0.00083
0.4	1.4	2	1.6	1.5	1.5	1.5	1.5	1.5	1.5
0.6	2.3	3.1	2.8	2.8	2.9	3	3	3	3

Table 55: *Values of $f(\delta)$, defined in (10), for the process* (28) *and different values of h, δ and n.*

		(32) holds for $h = 1$, $i = 2$							
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.2255	0.4266	0.7888	0.6948	0.6751	0.3301	0.7315	0.9674	0.9965
0.01	0.2201	0.4204	0.7799	0.6819	0.6593	0.3222	0.7157	0.9566	0.9952
0.05	0.2024	0.4017	0.7479	0.6292	0.5972	0.2853	0.6466	0.9041	0.9856
0.1	0.1808	0.3773	0.7036	0.5623	0.5149	0.2413	0.5456	0.802	0.9478
0.15	0.1598	0.3518	0.6548	0.4917	0.4361	0.2007	0.4446	0.6668	0.8715
0.2	0.1449	0.3291	0.6111	0.4289	0.3612	0.1611	0.3523	0.5193	0.7488
0.4	0.0983	0.2497	0.4403	0.2234	0.147	0.048	0.087	0.1108	0.2059
0.6	0.0656	0.1895	0.2914	0.1023	0.0467	0.0107	0.008	0.0105	0.0237
		(32) holds for $h = 1$, $i = 3$							
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.2583	0.2752	0.5358	0.7142	0.8767	0.5429	0.9462	1	1
0.01	0.2514	0.2663	0.5242	0.7026	0.87	0.5361	0.9451	1	1
0.05	0.2295	0.2382	0.4709	0.6492	0.8428	0.5114	0.9405	1	1
0.1	0.2015	0.2037	0.4104	0.5785	0.8038	0.4868	0.9357	1	1
0.15	0.1774	0.1728	0.3539	0.5098	0.7627	0.4645	0.9298	1	1
0.2	0.1541	0.1481	0.3013	0.4451	0.7103	0.4436	0.9239	1	1
0.4	0.088	0.0724	0.1574	0.2373	0.4631	0.3809	0.8884	1	1
0.6	0.0489	0.0354	0.0805	0.1234	0.2422	0.3377	0.8591	1	1
		(32) holds for $h = 5$, $i = 2$							
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.2128	0.1875	0.173	0.099	0.1259	0.2291	0.1444	0.0796	0.0675
0.01	0.1916	0.1647	0.1534	0.0894	0.1126	0.1843	0.1129	0.0589	0.0474
0.05	0.1425	0.1161	0.106	0.0624	0.0711	0.0819	0.0401	0.0157	0.0159
0.1	0.1068	0.0838	0.0735	0.0418	0.0427	0.0303	0.0071	0.0013	0.0024
0.15	0.0846	0.0618	0.0504	0.0264	0.0249	0.0126	0.0013	1e-04	8e-04
0.2	0.0693	0.0479	0.0355	0.0158	0.0142	0.0054	1e-04	0	3e-04
0.4	0.0353	0.0187	0.0071	0.003	0.0026	2e-04	0	0	0
0.6	0.0226	0.0079	0.002	9e-04	3e-04	0	0	0	0
		(32) holds for $h = 5$, $i = 3$							
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.1648	0.2186	0.3212	0.3499	0.3042	0.3432	0.0843	0.0107	5e-04
0.01	0.1422	0.1916	0.2806	0.3112	0.2575	0.1704	0.0264	0.0022	2e-04
0.05	0.0972	0.1306	0.1799	0.1926	0.1499	0.0357	0.001	1e-04	0
0.1	0.0623	0.088	0.1102	0.1064	0.0874	0.0162	2e-04	0	0
0.15	0.0432	0.0636	0.0699	0.0591	0.0534	0.0078	1e-04	0	0
0.2	0.0314	0.0471	0.046	0.0356	0.0326	0.004	0	0	0
0.4	0.0124	0.0213	0.01	0.0054	0.0046	1e-04	0	0	0
0.6	0.0072	0.0133	0.003	0.0011	5e-04	0	0	0	0

Table 56: *Proportion of* (7) *being fulfilled for the process* (28) *and different values of h, δ and n.*
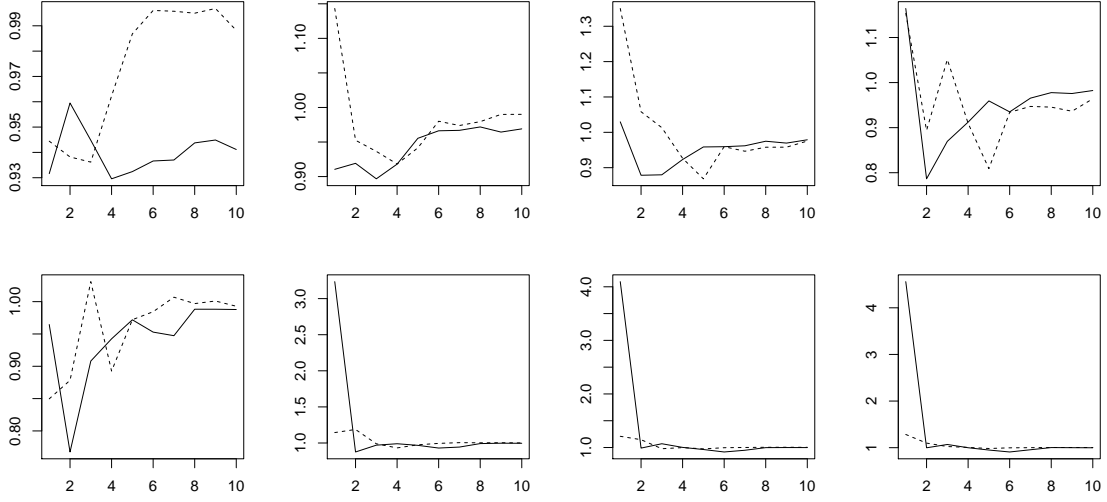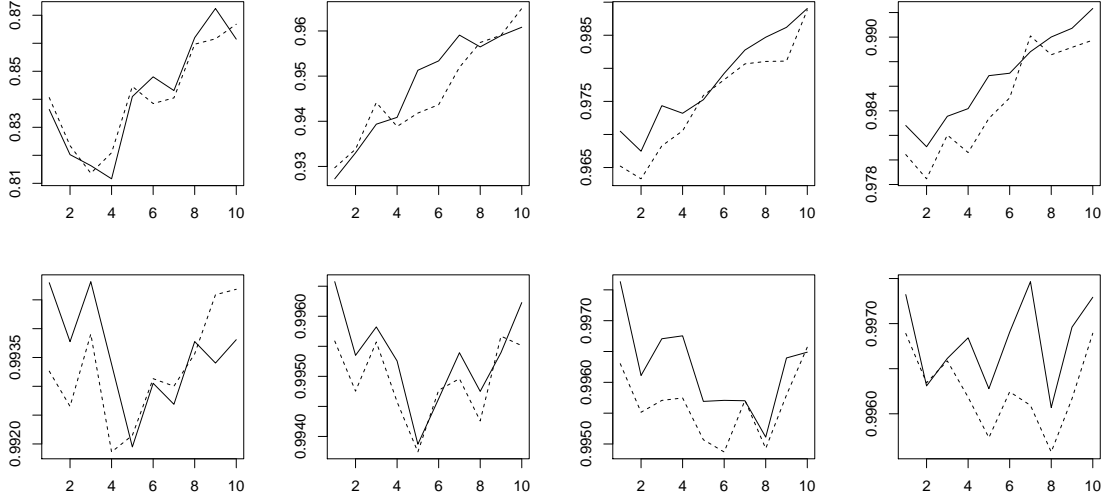
Figure 20: *Plot of $h \mapsto R^{s,ls}_{T,i}(h)$ for model (29) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R^{s,ls}_{T,2}(1)$			$R^{s,ls}_{T,2}(5)$	
		≥ 1	< 1		≥ 1	< 1
100	$R^{s,ls}_{T,3}(1) \geq 1$	0.1001	0.1732	$R^{s,ls}_{T,3}(5) \geq 1$	0.3487	0.1514
	$R^{s,ls}_{T,3}(1) < 1$	0.1709	0.5558	$R^{s,ls}_{T,3}(5) < 1$	0.1464	0.3535
n		$R^{s,ls}_{T,2}(1)$			$R^{s,ls}_{T,2}(5)$	
		≥ 1	< 1		≥ 1	< 1
1000	$R^{s,ls}_{T,3}(1) \geq 1$	0.0695	0.2001	$R^{s,ls}_{T,3}(5) \geq 1$	0.1521	0.2034
	$R^{s,ls}_{T,3}(1) < 1$	0.1783	0.5521	$R^{s,ls}_{T,3}(5) < 1$	0.1923	0.4522
n		$R^{s,ls}_{T,2}(1)$			$R^{s,ls}_{T,2}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R^{s,ls}_{T,3}(1) \geq 1$	0.051	0.1883	$R^{s,ls}_{T,3}(5) \geq 1$	0.0632	0.2063
	$R^{s,ls}_{T,3}(1) < 1$	0.1737	0.587	$R^{s,ls}_{T,3}(5) < 1$	0.1903	0.5402
n		$R^{s,ls}_{T,2}(1)$			$R^{s,ls}_{T,2}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
10000	$R^{s,ls}_{T,3}(1) \geq 1.01$	0	0.0027	$R^{s,ls}_{T,3}(5) \geq 1.01$	0.001	0.0284
	$R^{s,ls}_{T,3}(1) < 1.01$	0.0022	0.9951	$R^{s,ls}_{T,3}(5) < 1.01$	0.0261	0.9445
n		$R^{s,ls}_{T,2}(1)$			$R^{s,ls}_{T,2}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R^{s,ls}_{T,3}(1) \geq 1.05$	0	0	$R^{s,ls}_{T,3}(5) \geq 1.05$	0	0
	$R^{s,ls}_{T,3}(1) < 1.05$	0	1	$R^{s,ls}_{T,3}(5) < 1.05$	0	1
n		$R^{s,ls}_{T,2}(1)$			$R^{s,ls}_{T,2}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R^{s,ls}_{T,3}(1) \geq 1.1$	0	0	$R^{s,ls}_{T,3}(5) \geq 1.1$	0	0
	$R^{s,ls}_{T,3}(1) < 1.1$	0	1	$R^{s,ls}_{T,3}(5) < 1.1$	0	1

Table 57: *Proportions of the individual events in (32) for the process (29) and selected combinations of n and δ.*

δ	(32) holds for $h=1$								
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.6559	0.6135	0.6187	0.6216	0.6259	0.6336	0.6395	0.6358	0.638
0.01	0.6742	0.6485	0.6997	0.7682	0.863	0.9455	0.9772	0.9906	0.9951
0.05	0.7426	0.7852	0.9182	0.9826	0.9988	0.9999	1	1	1
0.1	0.8101	0.8967	0.9862	0.9992	1	1	1	1	1
0.15	0.8629	0.9501	0.998	1	1	1	1	1	1
0.2	0.8993	0.9766	0.9996	1	1	1	1	1	1
0.4	0.9648	0.9956	0.9999	1	1	1	1	1	1
0.6	0.9786	0.996	0.9999	1	1	1	1	1	1
δ	(32) holds for $h=5$								
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.7022	0.6495	0.6215	0.6043	0.5847	0.594	0.5964	0.5975	0.6034
0.01	0.7158	0.6753	0.6773	0.6956	0.7471	0.8326	0.8898	0.9226	0.9455
0.05	0.7716	0.7571	0.8269	0.9101	0.9727	0.9969	0.9997	1	1
0.1	0.8097	0.8235	0.9194	0.9769	0.996	0.9999	0.9999	1	1
0.15	0.8418	0.8762	0.9571	0.9863	0.9968	0.9997	0.9999	1	1
0.2	0.867	0.906	0.9728	0.9868	0.9959	0.9995	0.9998	1	1
0.4	0.9319	0.9633	0.9845	0.9897	0.9953	0.9995	0.9995	1	1
0.6	0.9588	0.9767	0.9903	0.9952	0.9975	0.9996	1	1	1

Table 58: *Proportion of* (32) *being fulfilled for the process* (29) *and different values of h, δ and n.*

δ	Value of $f(\delta)$, defined in (10), for $h=1$								
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
0.4	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
0.6	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47
δ	Value of $f(\delta)$, defined in (10), for $h=5$								
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021
0.05	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
0.1	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
0.15	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45
0.2	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54
0.4	0.042	0.042	0.042	0.042	0.042	0.042	0.042	0.042	0.042
0.6	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65

Table 59: *Values of $f(\delta)$, defined in (10), for the process* (29) *and different values of h, δ and n.*

δ				(32) holds for $h=1$, $i=2$				
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2636	0.2748	0.2544	0.2478	0.2406	0.2338	0.2213	0.2262	0.2247
0.01	0.2421	0.2372	0.1809	0.1332	0.0737	0.0264	0.0105	0.0039	0.0022
0.05	0.1807	0.128	0.0417	0.009	7e-04	1e-04	0	0	0
0.1	0.1268	0.062	0.0075	2e-04	0	0	0	0	0
0.15	0.0935	0.0329	0.0018	0	0	0	0	0	0
0.2	0.0704	0.0181	9e-04	0	0	0	0	0	0
0.4	0.0268	0.008	7e-04	0	0	0	0	0	0
0.6	0.0169	0.0063	5e-04	0	0	0	0	0	0

δ				(32) holds for $h=1$, $i=3$				
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2659	0.2753	0.2671	0.2696	0.2517	0.2456	0.2464	0.2336	0.2393
0.01	0.2459	0.2381	0.1908	0.14	0.0751	0.0303	0.0127	0.0057	0.0027
0.05	0.1789	0.132	0.0465	0.0092	5e-04	0	0	0	0
0.1	0.1253	0.0651	0.0079	6e-04	0	0	0	0	0
0.15	0.0898	0.033	0.0016	0	0	0	0	0	0
0.2	0.0661	0.0201	9e-04	0	0	0	0	0	0
0.4	0.0282	0.0082	6e-04	0	0	0	0	0	0
0.6	0.0177	0.0067	6e-04	0	0	0	0	0	0

δ				(32) holds for $h=5$, $i=2$				
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2516	0.301	0.3079	0.307	0.28	0.2633	0.2552	0.2586	0.2534
0.01	0.2123	0.2464	0.2364	0.2124	0.1598	0.0896	0.0575	0.0407	0.0271
0.05	0.1601	0.1721	0.1198	0.07	0.0231	0.0031	6e-04	0	0
0.1	0.1246	0.1139	0.0597	0.0299	0.0102	0.0016	6e-04	0	0
0.15	0.1003	0.08	0.036	0.0201	0.0082	0.0015	6e-04	0	0
0.2	0.0821	0.0579	0.0255	0.0164	0.0069	0.0014	5e-04	0	0
0.4	0.0392	0.0203	0.0116	0.0066	0.0027	1e-04	4e-04	0	0
0.6	0.0221	0.0113	0.0053	0.0021	0.0011	0	0	0	0

δ				(32) holds for $h=5$, $i=3$				
	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.2566	0.3073	0.3164	0.3181	0.2931	0.2689	0.2706	0.2631	0.2694
0.01	0.2245	0.2519	0.2423	0.2218	0.1639	0.1002	0.0645	0.0425	0.0294
0.05	0.1631	0.1718	0.1251	0.0711	0.0234	0.0034	9e-04	0	0
0.1	0.1275	0.1214	0.0653	0.0304	0.0106	0.0017	5e-04	0	0
0.15	0.1025	0.0848	0.0397	0.0214	0.0088	0.0016	5e-04	0	0
0.2	0.0843	0.0647	0.0281	0.018	0.0074	0.0015	5e-04	0	0
0.4	0.0393	0.0258	0.0135	0.0089	0.0038	6e-04	1e-04	0	0
0.6	0.0235	0.0162	0.0064	0.0035	0.0018	4e-04	0	0	0

Table 60: *Proportion of* (7) *being fulfilled for the process* (29) *and different values of h, δ and n.*
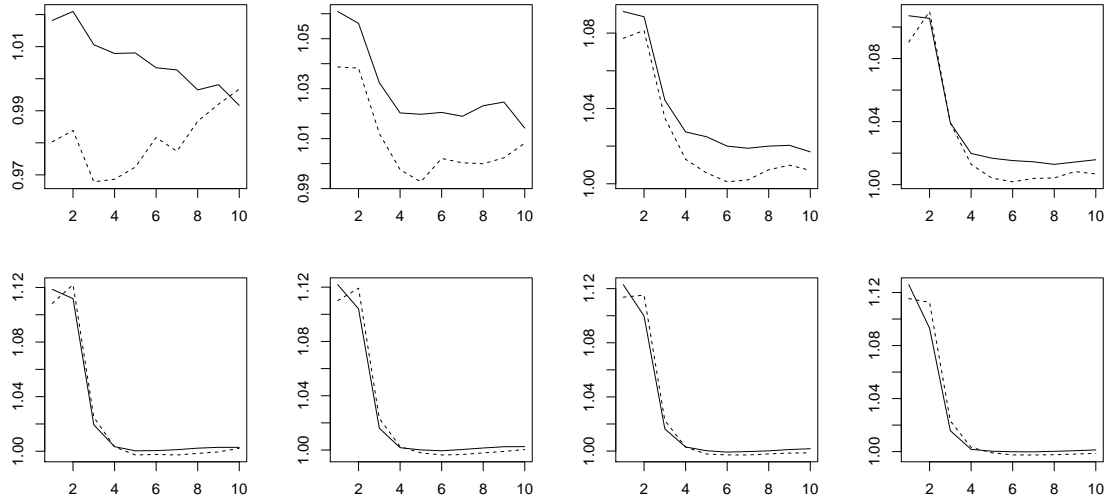
Figure 21: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (30) and different values of n (from left to right: n=100, n=200, n=500, n=1000 [first row], n=4000, n=6000, n=8000, n=10000 [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.2522	0.2389	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.1592	0.188
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1737	0.3352	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.147	0.5058
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0.2615	0.2754	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0.0627	0.0688
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0.1585	0.3046	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0.0467	0.8218
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	1	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.3928	0.2349
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.2009	0.1714
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	1	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.0069	0.1104
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.0899	0.7928
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.05	< 1.05		≥ 1.05	< 1.05
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.05$	0.9995	3e-04	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.05$	7e-04	8e-04
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.05$	2e-04	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.05$	0	0.9985
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0.7138	0.1872	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	5e-04	1e-04
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	0.0675	0.0315	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	1e-04	0.9993

Table 61: *Proportions of the individual events in (32) for the process (30) and selected combinations of n and δ.*

					(32) holds for $h=1$				
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.5878	0.6068	0.771	0.923	0.991	0.9997	1	1	1
0.01	0.5874	0.5899	0.7366	0.8916	0.9834	0.9991	1	1	1
0.05	0.596	0.5754	0.6068	0.7068	0.8602	0.9705	0.9931	0.9982	0.9995
0.1	0.6471	0.6304	0.5927	0.5661	0.5585	0.6014	0.6465	0.6889	0.7453
0.15	0.7089	0.725	0.7456	0.7626	0.7842	0.8116	0.8357	0.8553	0.8588
0.2	0.7665	0.8122	0.8819	0.929	0.9742	0.9934	0.9978	0.9989	0.9995
0.4	0.9111	0.9492	0.9717	0.9856	0.9955	0.9993	0.9999	1	1
0.6	0.959	0.9703	0.9838	0.9925	0.9988	1	1	1	1
					(32) holds for $h=5$				
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0.665	0.6758	0.6887	0.6723	0.6344	0.5987	0.5841	0.5736	0.5642
0.01	0.665	0.6735	0.6919	0.6897	0.6904	0.7027	0.7427	0.7656	0.7997
0.05	0.6821	0.6916	0.7375	0.8033	0.9097	0.9803	0.9942	0.9979	0.9992
0.1	0.7173	0.7381	0.8039	0.8845	0.9626	0.9932	0.9988	0.9998	0.9998
0.15	0.7546	0.7837	0.8528	0.9252	0.9694	0.9926	0.9978	0.9995	0.9996
0.2	0.7839	0.8221	0.8931	0.9513	0.9779	0.9954	0.9992	0.9998	0.9999
0.4	0.8752	0.9308	0.9798	0.9949	0.9993	1	1	1	1
0.6	0.9265	0.9735	0.9966	0.9997	1	1	1	1	1

Table 62: *Proportion of* (32) *being fulfilled for the process* (30) *and different values of h, δ and n.*

					Value of $f(\delta)$, defined in (10), for $h=1$				
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014
0.05	0.034	0.041	0.048	0.052	0.056	0.059	0.061	0.062	0.063
0.1	2.9e-05	3.9e-05	3.7e-05	0.0018	0.0055	0.0088	0.011	0.012	0.013
0.15	0.037	0.032	0.026	0.022	0.018	0.014	0.012	0.011	0.0099
0.2	0.087	0.082	0.077	0.072	0.068	0.064	0.062	0.061	0.06
0.4	0.015	0.019	0.024	0.028	0.032	0.035	0.037	0.038	0.039
0.6	0.22	0.22	0.22	0.23	0.23	0.24	0.24	0.24	0.24
					Value of $f(\delta)$, defined in (10), for $h=5$				
δ	$n{=}100$	$n{=}200$	$n{=}500$	$n{=}1000$	$n{=}2000$	$n{=}4000$	$n{=}6000$	$n{=}8000$	$n{=}10000$
0	0	0	0	0	0	0	0	0	0
0.01	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
0.05	0.069	0.068	0.067	0.067	0.067	0.067	0.066	0.066	0.066
0.1	0.02	0.026	0.034	0.039	0.045	0.05	0.053	0.056	0.057
0.15	3.1e-05	1.7e-05	0.0017	1.2e-05	0.0057	0.012	0.013	0.011	0.0089
0.2	0.071	0.063	0.054	0.046	0.039	0.032	0.028	0.026	0.023
0.4	0.36	0.34	0.33	0.32	0.31	0.31	0.3	0.3	0.3
0.6	0.64	0.63	0.61	0.6	0.59	0.58	0.57	0.57	0.57

Table 63: *Values of $f(\delta)$, defined in* (10), *for the process* (30) *and different values of h, δ and n.*

δ									
	(32) holds for $h=1$, $i=2$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.4513	0.6159	0.828	0.9465	0.9928	0.9998	1	1	1
0.01	0.4259	0.5832	0.7952	0.9247	0.9879	0.9993	1	1	1
0.05	0.3443	0.4493	0.6197	0.7622	0.8987	0.978	0.9947	0.9984	0.9997
0.1	0.2554	0.3047	0.3576	0.42	0.503	0.5985	0.6725	0.7259	0.7813
0.15	0.1909	0.1961	0.1773	0.1438	0.104	0.0739	0.057	0.0484	0.0411
0.2	0.1427	0.1282	0.0868	0.0483	0.0146	0.0022	4e-04	1e-04	0
0.4	0.0524	0.037	0.0259	0.0154	0.0052	8e-04	1e-04	0	0
0.6	0.0245	0.0186	0.012	0.0058	0.001	0	0	0	0
	(32) holds for $h=1$, $i=3$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.5137	0.6557	0.8692	0.9689	0.9982	0.9999	1	1	1
0.01	0.4911	0.6249	0.8424	0.9553	0.9955	0.9998	1	1	1
0.05	0.4027	0.5055	0.6997	0.8462	0.9475	0.9923	0.9984	0.9998	0.9998
0.1	0.3137	0.3615	0.4507	0.5369	0.6407	0.7551	0.8208	0.8656	0.901
0.15	0.236	0.2409	0.2347	0.219	0.1894	0.1605	0.1417	0.1189	0.1213
0.2	0.1764	0.158	0.1097	0.0691	0.027	0.0072	0.002	0.001	5e-04
0.4	0.0609	0.0392	0.0194	0.0102	0.0027	5e-04	0	0	0
0.6	0.0253	0.0177	0.008	0.0025	2e-04	0	0	0	0
	(32) holds for $h=5$, $i=2$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.3231	0.3527	0.3637	0.3386	0.3099	0.2916	0.2962	0.3114	0.3249
0.01	0.3062	0.3324	0.3378	0.3012	0.238	0.1757	0.1388	0.1189	0.0968
0.05	0.2549	0.2673	0.2475	0.1874	0.0954	0.0266	0.0075	0.0016	7e-04
0.1	0.2016	0.2004	0.1604	0.1094	0.0545	0.0158	0.0044	0.0011	6e-04
0.15	0.1626	0.1469	0.1032	0.0627	0.0323	0.0095	0.0027	5e-04	1e-04
0.2	0.1287	0.1076	0.0661	0.0351	0.0173	0.0037	9e-04	1e-04	0
0.4	0.0596	0.0292	0.0087	0.0015	3e-04	0	0	0	0
0.6	0.0301	0.0076	0.0013	0	0	0	0	0	0
	(32) holds for $h=5$, $i=3$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.3621	0.3897	0.4066	0.3879	0.3509	0.3347	0.3327	0.3584	0.3589
0.01	0.3472	0.3719	0.3819	0.3501	0.2814	0.2024	0.1619	0.1411	0.1173
0.05	0.2986	0.3057	0.2926	0.2263	0.1229	0.0321	0.0079	0.0025	0.0015
0.1	0.2455	0.2361	0.2041	0.1315	0.0641	0.0162	0.0044	9e-04	6e-04
0.15	0.1994	0.1808	0.1336	0.0771	0.0367	0.0099	0.0021	6e-04	3e-04
0.2	0.1636	0.1413	0.0876	0.0442	0.0204	0.0049	7e-04	1e-04	1e-04
0.4	0.0874	0.0496	0.0137	0.0042	4e-04	0	0	0	0
0.6	0.0504	0.0193	0.0021	3e-04	0	0	0	0	0

Table 64: *Proportion of* (7) *being fulfilled for the process* (30) *and different values of h, δ and n.*
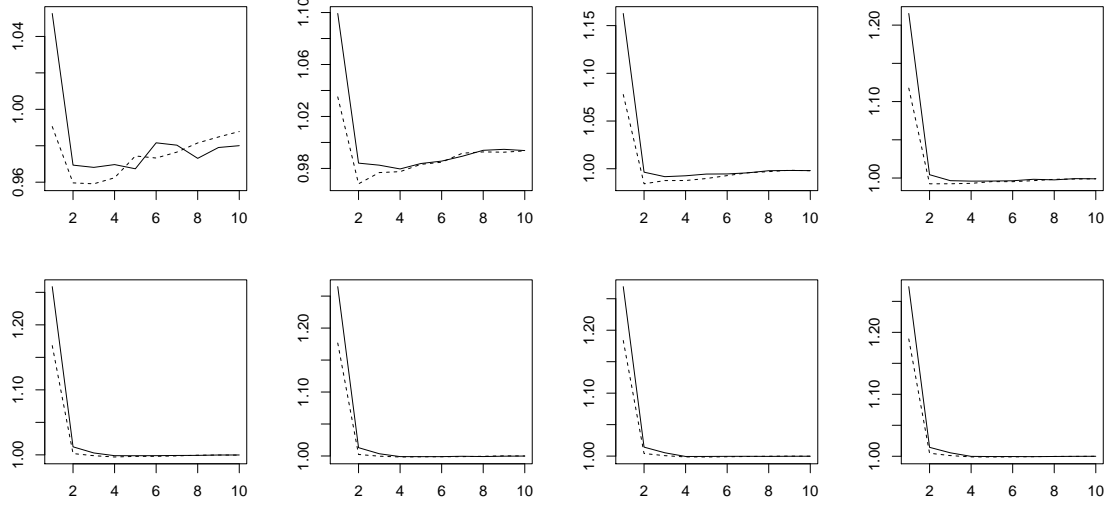
Figure 22: *Plot of $h \mapsto R_{T,i}^{\mathrm{s,ls}}(h)$ for model (31) and different values of n (from left to right: $n=100$, $n=200$, $n=500$, $n=1000$ [first row], $n=4000$, $n=6000$, $n=8000$, $n=10000$ [second row]). Solid line: $i = 3$ (test set), dashed line: $i = 2$ (validation set).*

n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
100	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	0.2168	0.2322	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	0.0612	0.1345
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0.1404	0.4106	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.1364	0.6679
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.15	< 1.15		≥ 1.15	< 1.15
1000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.15$	0.247	0.4444	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.15$	0	8e-04
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.15$	0.0875	0.2211	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.15$	1e-04	0.9991
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1	< 1		≥ 1	< 1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1$	1	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1$	0.4834	0.1485
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1$	0.1227	0.2454
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.01	< 1.01		≥ 1.01	< 1.01
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.01$	1	0	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.01$	1e-04	0.0113
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.01$	0	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.01$	0.0055	0.9831
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.1	< 1.1		≥ 1.1	< 1.1
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.1$	0.994	0.0059	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.1$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.1$	1e-04	0	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.1$	0	1
n		$R_{T,2}^{\mathrm{s,ls}}(1)$			$R_{T,2}^{\mathrm{s,ls}}(5)$	
		≥ 1.2	< 1.2		≥ 1.2	< 1.2
10000	$R_{T,3}^{\mathrm{s,ls}}(1) \geq 1.2$	0.3571	0.5775	$R_{T,3}^{\mathrm{s,ls}}(5) \geq 1.2$	0	0
	$R_{T,3}^{\mathrm{s,ls}}(1) < 1.2$	0.0234	0.042	$R_{T,3}^{\mathrm{s,ls}}(5) < 1.2$	0	1

Table 65: *Proportions of the individual events in (32) for the process (31) and selected combinations of n and δ.*

	(32) holds for $h = 1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.6294	0.6616	0.8011	0.936	0.9933	1	1	1	1
0.01	0.6274	0.6509	0.7712	0.9171	0.9887	0.9999	1	1	1
0.05	0.6308	0.6162	0.6522	0.7891	0.934	0.9943	0.9996	0.9999	0.9999
0.1	0.6576	0.6199	0.5535	0.5771	0.7349	0.9	0.9638	0.984	0.994
0.15	0.688	0.655	0.5516	0.4681	0.4815	0.6058	0.7128	0.7846	0.8487
0.2	0.7303	0.7043	0.6193	0.5033	0.4267	0.3676	0.3584	0.3792	0.3991
0.4	0.8666	0.8924	0.908	0.9155	0.9359	0.964	0.9774	0.986	0.991
0.6	0.9378	0.9648	0.9852	0.9948	0.999	1	1	1	1
	(32) holds for $h = 5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.7055	0.7197	0.7289	0.7271	0.7287	0.7232	0.7325	0.7302	0.7288
0.01	0.7291	0.7654	0.8161	0.8581	0.9061	0.9433	0.9624	0.9754	0.9832
0.05	0.8145	0.8797	0.9437	0.9765	0.9933	0.9986	0.9998	1	1
0.1	0.8786	0.9346	0.9807	0.996	0.9991	1	1	1	1
0.15	0.916	0.9602	0.9927	0.9991	0.9999	1	1	1	1
0.2	0.937	0.9757	0.9969	0.9998	0.9999	1	1	1	1
0.4	0.9742	0.9929	0.9994	1	1	1	1	1	1
0.6	0.9862	0.9979	0.9999	1	1	1	1	1	1

Table 66: *Proportion of* (32) *being fulfilled for the process* (31) *and different values of h, δ and n.*

	Value of $f(\delta)$, defined in (10), for $h = 1$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.012	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.013
0.05	0.062	0.063	0.063	0.063	0.064	0.064	0.064	0.064	0.064
0.1	0.065	0.081	0.099	0.11	0.12	0.13	0.13	0.13	0.13
0.15	0.013	0.029	0.048	0.06	0.069	0.078	0.082	0.085	0.087
0.2	0.00034	0.00046	2e-04	0.0087	0.019	0.027	0.031	0.034	0.036
0.4	0.18	0.17	0.15	0.14	0.14	0.13	0.13	0.12	0.12
0.6	0.38	0.37	0.35	0.35	0.34	0.33	0.33	0.32	0.32
	Value of $f(\delta)$, defined in (10), for $h = 5$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0	0	0	0	0	0	0	0	0
0.01	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
0.05	0.06	0.061	0.062	0.063	0.063	0.063	0.063	0.063	0.063
0.1	0.12	0.12	0.12	0.13	0.13	0.13	0.13	0.13	0.13
0.15	0.18	0.18	0.19	0.19	0.19	0.19	0.19	0.19	0.19
0.2	0.24	0.24	0.25	0.25	0.25	0.25	0.25	0.26	0.26
0.4	0.48	0.49	0.5	0.5	0.51	0.51	0.51	0.51	0.51
0.6	0.72	0.73	0.75	0.75	0.76	0.76	0.77	0.77	0.77

Table 67: *Values of $f(\delta)$, defined in (10), for the process* (31) *and different values of h, δ and n.*

	(32) holds for $h = 1$, $i = 2$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.374	0.4813	0.7151	0.9084	0.9875	1	1	1	1
0.01	0.3572	0.46	0.6885	0.8914	0.9834	0.9999	1	1	1
0.05	0.2976	0.3746	0.5613	0.7766	0.9337	0.9943	0.9996	0.9999	0.9999
0.1	0.231	0.2756	0.3915	0.5557	0.7498	0.9035	0.9644	0.9841	0.9941
0.15	0.1751	0.1965	0.2517	0.3345	0.4587	0.6126	0.7202	0.7887	0.8503
0.2	0.1354	0.1382	0.1509	0.1734	0.2172	0.2686	0.3097	0.3434	0.3805
0.4	0.0472	0.0285	0.0133	0.0055	0.0024	1e-04	1e-04	0	0
0.6	0.0185	0.0067	0.0012	0	0	0	0	0	0
	(32) holds for $h = 1$, $i = 3$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.4616	0.5921	0.8148	0.9466	0.993	1	1	1	1
0.01	0.449	0.5763	0.8031	0.9433	0.9925	1	1	1	1
0.05	0.399	0.508	0.7449	0.9137	0.9873	1	1	1	1
0.1	0.3344	0.4217	0.6334	0.8292	0.9525	0.9957	0.9994	0.9999	0.9999
0.15	0.2813	0.3391	0.5017	0.6914	0.8528	0.9556	0.9822	0.9923	0.9968
0.2	0.2275	0.2659	0.3762	0.5271	0.6697	0.8066	0.8697	0.9082	0.9346
0.4	0.1016	0.0901	0.0835	0.0806	0.0625	0.0359	0.0225	0.014	0.009
0.6	0.0475	0.0299	0.014	0.0052	0.001	0	0	0	0
	(32) holds for $h = 5$, $i = 2$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.239	0.2112	0.2084	0.2041	0.1924	0.193	0.198	0.1953	0.2055
0.01	0.1976	0.1464	0.107	0.0734	0.0434	0.021	0.0148	0.008	0.0056
0.05	0.1148	0.0592	0.0238	0.0083	0.0018	1e-04	0	0	0
0.1	0.0686	0.0288	0.0067	0.0013	1e-04	0	0	0	0
0.15	0.0422	0.0166	0.0025	1e-04	0	0	0	0	0
0.2	0.0297	0.0091	9e-04	0	0	0	0	0	0
0.4	0.0106	0.0022	1e-04	0	0	0	0	0	0
0.6	0.0048	6e-04	1e-04	0	0	0	0	0	0
	(32) holds for $h = 5$, $i = 3$								
δ	n=100	n=200	n=500	n=1000	n=2000	n=4000	n=6000	n=8000	n=10000
0	0.2337	0.2309	0.2119	0.21	0.2199	0.223	0.2185	0.2281	0.2313
0.01	0.1957	0.1688	0.1229	0.0939	0.0607	0.0387	0.0248	0.017	0.0114
0.05	0.1195	0.0789	0.0367	0.0156	0.0049	0.0013	2e-04	0	0
0.1	0.0736	0.0414	0.0132	0.0027	8e-04	0	0	0	0
0.15	0.0516	0.0244	0.005	8e-04	1e-04	0	0	0	0
0.2	0.0369	0.0154	0.0022	2e-04	1e-04	0	0	0	0
0.4	0.0162	0.0049	5e-04	0	0	0	0	0	0
0.6	0.0094	0.0015	0	0	0	0	0	0	0

Table 68: *Proportion of* (7) *being fulfilled for the process* (31) *and different values of h, δ and n.*