ADAPTIVE SPECTRAL DENSITY ESTIMATION BY MODEL SELECTION UNDER LOCAL DIFFERENTIAL PRIVACY

MARTIN KROLL

Abstract. We study spectral density estimation under local differential privacy. Anonymization is achieved through truncation followed by Laplace perturbation. We select our estimator from a set of candidate estimators by a penalized contrast criterion. This estimator is shown to attain nearly the same rate of convergence as the best estimator from the candidate set. A key ingredient of the proof are recent results on concentration of quadratic forms in terms of sub-exponential random variables obtained in [GSS19]. We illustrate our findings in a small simulation study.

1. Introduction

Let $X = (X_t)_{t\in\mathbb{Z}}$ be a stationary time series with autocorrelation function $\gamma\colon \mathbb{Z} \to \mathbb{R}$ defined via $\gamma_k = \gamma(k) = \mathrm{Cov}(X_t, X_{t+k})$ for any $t \in \mathbb{Z}$. Then, under the assumption that the series $(\gamma_k)_{k\in\mathbb{Z}}$ is absolutely convergent, its spectral density f can be represented by the inversion formula as

$$f(\omega) = \frac{1}{2\pi}\sum_{k\in\mathbb{Z}} \gamma_k e^{-\mathrm{i}\omega k}, \qquad \omega \in [-\pi, \pi]. \tag{1.1}$$

The present paper treats the nonparametric estimation of f from a finite snippet $X_{1:n} = (X_1, \ldots, X_n)$ of length n under privacy constraints. More precisely, we assume that each X_i, $i = 1, \ldots, n$ belongs to another data holder who is willing to publish an anonymized version Z_i of the actual value X_i only. Then, the complete snippet $X_{1:n}$ is not accessible to the statistician, and estimation can be done based on the privatized snippet $Z_{1:n} = (Z_1, \ldots, Z_n)$ solely. Such a situation might, for instance, be of relevance for so-called random walk survey designs [Ben+91, LR85] where people are successively interviewed and the next person to be interviewed is determined by a precisely defined random walk strategy. Hence, one might suppose that data of consecutively interviewed persons should incorporate some amount of dependency since they live nearby each other. Since survey interviews often aim at collecting data concerning sensitive social and biological data concerning health [Fly+13] there is certainly need for anonymization.

As our mathematical setup for privacy we use the notion of local α-differential privacy that has gained increasing popularity in the statistics community in recent years. Until now, theoretical research in this framework has focused on models with independent observations and estimation tasks like density estimation [DJW18, But+20], estimation of functionals [RS20, BRS20], testing [BB20, LLL20, BRS20], and classification [BB19]. To the best of our knowledge, this work is the first one that applies the concept of differential privacy to time series data and the task of estimating the dependency structure of a process under privacy restrictions.

Of course, in the classical scenario without any privacy restriction there exists an overwhelming amount of literature on spectral density estimation from a snippet of finite length in both parametric and nonparametric models [Com01, Dah89, Dav73, Efr98, FT86,

Date: October 8, 2020.

2010 *Mathematics Subject Classification.* 62G05 (primary), and 62M10 (secondary).

Key words and phrases. Differential privacy, spectral density estimation, orthonormal series estimator, adaptive estimation, model selection.

Gol93, Neu96, Sou00, Tan87]. From a methodological point of view we explicitly point out the paper [Com01] that is our point of origin and uses the penalized contrast approach (that we will also use later in Section 4) in the non-private framework. For this reason, we give in the following a recap of this technique.

The non-private model selection device from [Com01] in a nutshell. The so-called periodogram is the point of origin of many procedures for spectral density estimation but it is has to be smoothed in order to obtain consistency or even rate optimal estimators. The *centred periodogram*, based on n consecutive observations of the time series, is defined via

$$I_n^X(\omega) = \frac{1}{2\pi n}\left|\sum_{t=1}^n (X_t - \bar{X}_n)e^{-\mathrm{i}t\omega}\right|^2, \tag{1.2}$$

where $\bar{X}_n$ is the sample mean of the observed snippet $X_{1:n}$.

One possibility to obtain a smoothed version of the periodogram is projection of I_n^X to a finite-dimensional subspace $S_\mathfrak{m}$ of $L^2([-\pi,\pi])$, say $S_\mathfrak{m} = \operatorname{span}(\varphi_i)_{i\in\mathcal{I}_\mathfrak{m}}$, where $(\varphi_i)_{i\in\mathcal{I}_\mathfrak{m}}$ is an orthonormal basis of $L^2([-\pi,\pi])$ and $\mathcal{I}_\mathfrak{m}$ a finite subset. We denote $D_\mathfrak{m} = \dim(S_\mathfrak{m}) = |\mathcal{I}_\mathfrak{m}|$. The choice of a subspace $S_\mathfrak{m}$ might be interpreted as the choice of a finite-dimensional model $\mathfrak{m}$ (which explains the choice of the letter $\mathfrak{m}$ here). Potential models include spaces generated by trigonometric functions, regular piecewise polynomials, general piecewise polynomials, regular compactly supported periodic wavelets, and general compactly supported periodic wavelets; see Section 2.2 in [Com01] for a detailed description of all of these models. With such a model we associate the estimator

$$\widehat{f}_\mathfrak{m}^X = \sum_{i\in\mathcal{I}_\mathfrak{m}} \widehat{a}_i^X \varphi_i$$

with coefficients

$$\widehat{a}_i^X = \int_{-\pi}^{\pi} I_n^X(\omega)\varphi_i(\omega)\mathrm{d}\omega.$$

Note that so defined estimator minimizes the quantity

$$\int_{-\pi}^{\pi} (I_n(\omega) - t(\omega))^2 \mathrm{d}\omega,$$

or, equivalently, the contrast $\Upsilon_n(t) = \int_{-\pi}^{\pi} t^2(\omega)\mathrm{d}\omega - 2\int_{-\pi}^{\pi} t(\omega)I_n^X(\omega)\mathrm{d}\omega$, over all $t \in S_\mathfrak{m}$.

An upper risk bound for the estimator $\widehat{f}_\mathfrak{m}^X$ can be derived from the following decomposition (defining $f_\mathfrak{m}$ as the projection of f onto the space $S_\mathfrak{m}$):

$$\begin{aligned}\mathbf{E}\|f - \widehat{f}_\mathfrak{m}^X\|^2 &= \|f - f_\mathfrak{m}\|^2 + \mathbf{E}\|\widehat{f}_\mathfrak{m}^X - f_\mathfrak{m}\|^2 \\ &\leq \|f - f_\mathfrak{m}\|^2 + C(f)\cdot\frac{D_\mathfrak{m}}{n},\end{aligned} \tag{1.3}$$

where the inequality is taken from Equation (5) in [Com01]. It is based on the following assumption (Assumption 2 in [Com01]) that we will adopt for this work.

Assumption 1.1. The autocovariance function γ of the time series X is such that $\sum_{k\in\mathbb{Z}}|\gamma_k| = M < +\infty$ and $\sum_{k\in\mathbb{Z}}|k\gamma_k^2| = M_1 < +\infty$.

From the results in [Com01] it is easy to see that the optimal rate that can be achieved for smooth spectral density functions belonging to a Sobolev space with smoothness parameter s is $n^{-2s/(2s+1)}$. However, as often in nonparametric statistics, the optimal model from a set of potential models that has to be selected to reach this rate can be chosen directly only if a priori knowledge concerning the smoothness is available. Since such knowledge is usually not given, one has to find a method for model selection that is completely data-driven. A by now classical method for this purpose is model selection [BM97, BBM99, Mas07]. This general toolbox has been used by F. Comte in [Com01] for adaptive spectral density estimation in the non-private case where $X_{1:n}$ is observable. Her method consists in choosing a model as the minimizer $\widehat{\mathfrak{m}}$ of a penalized contrast criterion over a set $\mathcal{M}_n$ of potential models, that is,

$$\widehat{\mathfrak{m}} = \operatorname*{argmin}_{\mathfrak{m}\in\mathcal{M}_n} \Upsilon_n(\widehat{f}^X_{\mathfrak{m}}) + \mathrm{pen}(\mathfrak{m}).$$

Here Υ_n is a contrast function (for instance, the one defined above) and pen: $\mathcal{M}_n \to [0,\infty)$ a penalty function that penalizes too complex potential models. Usually, pen is a monotone function in the dimension $D_{\mathfrak{m}}$ of the space $S_{\mathfrak{m}}$. In [Com01] it has been shown that, under sufficiently mild assumptions, the estimator $\widehat{f}^X_{\widehat{\mathfrak{m}}}$ behaves nearly as well as the oracle given by the optimal model from the collection:

$$\mathbf{E}\|\widehat{f}^X_{\widehat{\mathfrak{m}}} - f\|^2 \leq C_1 \inf_{\mathfrak{m}\in\mathcal{M}_n} \{\|f - f_{\mathfrak{m}}\|^2 + \mathrm{pen}(\mathfrak{m})\} + \frac{C_2}{n}. \tag{1.4}$$

Here, the constant C_1 is purely numerical whereas C_2 might depend on f through its sup-norm, and additionally on quantities related to the class $\mathcal{M}_n$ of potential models (but, of course, not on n). If the penalty term can be chosen of the same order as the variance term $D_{\mathfrak{m}}/n$ in (1.3) (maybe up to logarithmic factors), then the adaptive estimator attains the same rate as the best possible estimator over all potential models (up to logarithmic factors).

Contributions of the paper. The principal purpose of this work is to derive an oracle inequality in the spirit of (1.4) when only anonymized data are available. The main difficulty in this scenario is that the periodogram I^X_n is not directly available. Hence, one approach would be to define differentially private Z_i in a way such that a suitable substitute I^Z_n for I^X_n can be defined in terms of the Z_i only. We introduce a procedure to define such $Z_{1:n}$ in the framework of α-differential privacy by a combination of truncation and Laplace perturbation. Using the privatized version of the periodogram, one can then apply the general toolbox as in the non-private case. We first consider upper bounds in the spirit of (1.3) for projections of I^Z_n to finite-dimensional spaces $S_{\mathfrak{m}}$ for fixed models $\mathfrak{m}$. In the specific case where the privacy level α is fixed and interpreted as a constant whereas n tends to $+\infty$, the rate of convergence over Sobolev ellipsoids that we obtain is the same as in the non-private setup up to logarithmic factors. Complementary to these upper bounds, we also state some first lower bound results that show that in some cases there might be a loss in the rate caused by the privacy level α when it is allowed to vary with n. The main theoretical result of this paper is an oracle inequality in the spirit of (1.4) for private data. For our completely data-driven estimator $\widetilde{f} = \widehat{f}^Z_{\widehat{\mathfrak{m}}}$ with the model $\widehat{\mathfrak{m}}$ determined via a model selection device (with a penalty that is adapted to the privacy framework), we derive

$$\mathbf{E}\|\widetilde{f} - f\|^2 \leq C_1 \inf_{\mathfrak{m}\in\mathcal{M}_n} [\|f - f_{\mathfrak{m}}\|^2 + \mathrm{pen}(\mathfrak{m})] + C_2 \max\left\{\frac{1}{n}, \frac{\log^2(n)}{n^3\alpha^4}\right\}$$

where α is the privacy parameter (see Section 2 for the significance of this parameter). In contrast to (1.4), also the penalty depends on the privacy parameter α in our case. However, as in the non-private setup, the adaptive estimator suffers at most from an additional loss in extra logarithmic terms in contrast to the optimal possible estimator taken from the considered collection of models. Remarkably, recent results on the concentration of quadratic forms in sub-exponential random variables [GSS19] turn out to be useful for our theoretical analysis. From a methodological point of view the present work might be of interest since it is, to the best of the author's knowledge, the first paper where model selection has been used to perform adaptive estimation under privacy constraints ([But+20] uses wavelet estimators to achieve adaptation in the privacy setup).

Notation. For real numbers a, b we set $[\![a,b]\!] = [a,b] \cap \mathbb{Z}$. By $\mathcal{L}(b)$ we denote the Laplace distribution with parameter b, by $\mathcal{N}(\mu, \Sigma)$ the normal distribution with mean μ and covariance matrix Σ. With $P_H v$ we denote the projection of a vector v to some subspace H. With E_n we denote the $n \times n$-identity matrix, with $\mathbf{0}_n$ the $n \times n$-zero matrix, and with $\vec{c}$ the $n \times 1$-vector containing only the value $c \in \mathbb{R}$ as entry. $\rho(A)$ denotes the spectral radius of a matrix A.

For any real-valued random variable X and $\beta > 0$ define the (quasi-)norm

$$\|X\|_{\psi_\beta} := \inf \left\{ t > 0 : \mathbf{E} \exp \left(\frac{|X|^\beta}{t^\beta} \right) \leq 2 \right\}$$

(as usual, one puts $\inf \emptyset = +\infty$). The (quasi-)norms $\|\cdot\|_{\psi_\beta}$ are called *exponential Orlicz norms*. By $\|\cdot\|$ we denote the usual L^2-norm, by $\|\cdot\|_{\mathrm{op}}$ the operator norm of a matrix.

We write $a_n \lesssim b_n$ if $a_n \leq C b_n$ for some purely numerical constant C and all sufficiently large n. Throughout the paper, C denotes a generic constant whose value might change with every appearance. By writing $C(\dots)$ we indicate the dependence of a numerical constant on one or several parameters that are listed within the brackets.

Organization of the paper. The paper is organized as follows. Section 2 introduces the notion of α-differential privacy and we introduce our algorithm to anonymize time series data. Section 3 is devoted to the derivation of upper risk bounds for fixed models $\mathfrak{m}$, and we also give some lower bound results. In the main Section 4 we state the oracle inequality for privatized time series data. A small sample simulation study is presented in Section 5 followed by a summary in Section 6 where we also indicate directions for further research.

2. Privacy

The notion of local α-differential privacy. Let us denote by $X_1, \dots, X_n$ the unanonymized random variables, that is $X_{1:n} = (X_1, \dots, X_n)$ is a snippet from the stationary time series X whose spectral density is the quantity of interest. We assume that each X_i belongs to a certain data holder who does not want to publish the value X_i but only an anonymized version of it, which will be denoted with Z_i. A theoretical framework for formalizing the vague catchwords *anonymization* and *privacy* is α-differential privacy which originally goes back to [Dwo06] and has obtained increasing interest in the statistics community within the last decade. There is a distinction between *global* differential privacy (for instance, considered in [HRW13]) where a trusted curator is given access to the complete data (that is, in our case, the snippet $X_{1:n}$) and a privatized version of standard estimators can be published, and *local* differential privacy where the original data are anonymized directly by the data holders and estimation has to be performed using the resulting private snippet $Z_{1:n}$. We will stick to this latter framework of local differential privacy in this paper. Under local privacy, data are successively obtained applying appropriate Markov kernels. More precisely, given $X_i = x_i$ and $Z_j = z_j$ for $j = 1, \dots, i-1$, the i-th privatized output Z_i is drawn as

$$Z_i \sim Q_i(\cdot \mid X_i = x_i, Z_1 = z_1, \dots, Z_{i-1} = z_{i-1}) \tag{2.1}$$

for Markov kernels $Q_i : \mathscr{Z} \times (\mathcal{X} \times \mathcal{Z}^{i-1}) \to [0,1]$ with $(\mathcal{X}, \mathscr{X})$, $(\mathcal{Z}, \mathscr{Z})$ denoting the measure spaces of non-private and private data, respectively (cf. Figure 2 in [DJW18] for a representation of this sampling scheme as a graphical model). In this paper, we propose a *non-interactive* algorithm where the random value Z_i depends on X_i only: thus, there is no dependence on previously generated Z_i's on the right-hand side of Equation (2.1). We also dispense with the dependence of Q_i on i, that is, we consider procedures with

$$Z_i \sim Q(\cdot \mid X_i = x_i)$$

for all $i \in [\![1, \dots, n]\!]$.

The quantification of privacy is achieved via the notion of α-differential privacy. In our context, this notion means that the estimate

$$\sup_{A\in\mathscr{Z}} \frac{Q_i(A\,|\,X_i=x)}{Q_i(A\,|\,X_i=x')} \leq \exp(\alpha) \tag{2.2}$$

is supposed to hold for all $x,x'\in\mathcal{X}$. If there exist densities $q(z\,|\,X=x)$ for the Markov kernel for all $x\in\mathcal{X}$ it is easy to verify that condition (2.2) is equivalent to

$$\sup_{z\in\mathcal{Z}} \frac{q(z\,|\,X_i=x)}{q(z\,|\,X_i=x')} \leq \exp(\alpha) \tag{2.3}$$

for all $x,x'\in\mathcal{X}$.

Anonymization procedure. It well-known that adding centred Laplace distributed noise on bounded random variables with sufficiently large variance guarantees α-differential privacy [DJW18, RS20]. We will use this general technique but have to transform the X_i in a first step because we do not want to impose a boundedness assumption on the X_i in general since this is obviously not satisfied in the most important case of Gaussian time series. This transformation consists in a truncation of the original X_i. More precisely, we put

$$\widetilde{X}_i = (X_i \wedge \tau_n) \vee (-\tau_n), \qquad i\in[\![1,n]\!] \tag{2.4}$$

where $\tau_n>0$. Note that the truncation can be performed locally by the data holders once all of them have agreed on the value τ_n. Our estimators are quite sensitive to the choice of the threshold τ_n. On the one hand, we want τ_n to tend to $+\infty$ in order to bound the probability that truncation occurs for at least one variable X_i by the rate of convergence that we aim at. On the other hand, τ_n arises in the rates of convergence and should be as small as possible. For our purposes, a logarithmically increasing (in terms of the snippet length n) sequence τ_n will turn out to be convenient.

By construction, we trivially have $\widetilde{X}_i\in[-\tau_n,\tau_n]$, and the above mentioned Laplace technique can be applied on the transformed data.

Lemma 2.1. *The random variables*

$$Z_i = \widetilde{X}_i + \xi_i \tag{2.5}$$

with ξ_i i.i.d. $\sim\mathcal{L}(2\tau_n/\alpha)$ are α-differentially private views of the original X_i.

Proof. We only have to check (2.3). Recall that the density of a centred Laplace distributed random variable with scale parameter $b>0$ is given by $1/(2b)\exp(-|x|/b)$. Put $\widetilde{x}=(x\wedge\tau_n)\vee(-\tau_n)$ and $\widetilde{x}'=(x'\wedge\tau_n)\vee(-\tau_n)$. By the reverse triangle inequality, we have

$$\begin{aligned}\sup_{z\in\mathcal{Z}} \frac{q(z\,|\,X=x)}{q(z\,|\,X=x')} &= \sup_{z\in\mathcal{Z}} \exp\left(-\alpha\cdot\frac{|z-\widetilde{x}|}{2\tau_n} + \alpha\cdot\frac{|z-\widetilde{x}'|}{2\tau_n}\right)\\ &\leq \exp\left(\alpha\cdot\frac{|\widetilde{x}-\widetilde{x}'|}{2\tau_n}\right)\\ &\leq \exp(\alpha),\end{aligned}$$

and (2.3) holds. □

Remark 2.2. Let use mention that the privacy mechanism defining the Z_i in (2.5) is convenient for our purposes in this paper but not optimal in other scenarios. For instance, imagine that the X_i are i.i.d. and the statistician wants to estimate the underlying probability density function f. Then, apart from the additional threshold, (2.5) defines a convolution model with Laplace distributed error density. Convolution models are well studied and it is known that the rate of convergence for s-smooth functions based on observations Z_i is at least $n^{-2s/(2s+3)}$ [Fan91]. However, the optimal rate under local differential privacy (considering α as a fixed constant), that can only be attained using privacy mechanisms different from (2.5)), is $n^{-s/(s+1)}$ as has been shown in [DJW18] and [But+20]. This emphasizes the fact that the privacy mechanisms to be used should not only depend on the available data but also on the statistical problem at hand.

3. Risk bounds for fixed models

In this section, we propose an estimator of the spectral density f based only on observations of the privatized data Z_i as defined in (2.5). In this case, we derive an upper risk bound similar to (1.3) for any fixed model $\mathfrak{m}$. As a consequence we obtain that, regarding the privacy parameter α as a fixed numerical constant, the proposed estimator attains the nearly same rate of convergence in terms of the snippet length n as in the non-private setup up to an additional logarithmic factor. Our estimator is based on the *privatized periodogram*

$$I_n^Z(\omega) = \frac{1}{2\pi n} \left| \sum_{t=1}^n (Z_t - \bar{Z}_n) e^{-\mathrm{i}t\omega} \right|^2 .$$

The function I_n^Z formally resembles the definition of the periodogram in (1.2) with X_i being replaced with Z_i. Put $Z_i' = X_i + \xi_i$. Then $Z_i' = Z_i$ holds whenever $X_i = \widetilde{X}_i$, that is, the value X_i is not modified in the truncation step (2.4). It is intuitively clear that in this 'nice' case one can hope to extract much more information from the dependency structure of the time series than in the case where truncation leads to a value $\widetilde{X}_i$ different from X_i. This 'nice' event is formalized in the proofs of Theorems 3.2 and 4.4 below via the event $A = \{X_i = \widetilde{X}_i \text{ for all } i \in [\![1, n]\!]\}$. For $i, j \in [\![1, n]\!]$, the covariance between Z_i' and Z_j' can be calculated as

$$\begin{aligned} \operatorname{Cov}(Z_i', Z_j') &= \operatorname{Cov}(X_i + \xi_i, X_j + \xi_j) \\ &= \operatorname{Cov}(X_i, X_j) + \operatorname{Cov}(X_i, \xi_j) + \operatorname{Cov}(\xi_i, X_j) + \operatorname{Cov}(\xi_i, \xi_j) \\ &= \gamma_{|i-j|} + \frac{8\tau_n^2}{\alpha^2}\,\delta_{ij}, \end{aligned}$$

where δ_{ij} is the Kronecker delta. Thus, by the inversion formula (1.1), we have

$$f^{Z'}(\omega) = f(\omega) + \frac{8\tau_n^2}{\alpha^2} \tag{3.1}$$

with $f^{Z'}$ denoting the spectral density of the stationary time series $(Z_t')_{t\in\mathbb{Z}}$. There is only hope to be able to estimate this quantity if we can observe the Z_i' for a significant amount of i. This is the more likely the larger the threshold τ_n is chosen. Under our technical assumptions that will be introduced below, a logarithmically increasing sequence τ_n guarantees that $Z_i = Z_i'$ for all $i \in [\![1, n]\!]$ with sufficiently high probability. In this scenario, it then turns out convenient to *define*

$$\widehat{I}_n(\omega) = I_n^Z(\omega) - \frac{8\tau_n^2}{\alpha^2},$$

which can be seen as an substitute of the quantity I_n^X. Based on the definition of $\widehat{I}_n$ we can now proceed as in the non-private case. For a fixed model $\mathfrak{m}$, we put

$$\widehat{f}_{\mathfrak{m}} = \sum_{i\in S_{\mathfrak{m}}} \widehat{a}_i \varphi_i, \tag{3.2}$$

where

$$\widehat{a}_i = \int_{-\pi}^{\pi} \widehat{I}_n(\omega)\varphi_i(\omega)\mathrm{d}\omega. \tag{3.3}$$

The following assumption is used in the proof of Theorem 3.2 to bound the probability of the event that $Z_i \neq Z_i'$ for at least one index i.

Assumption 3.1. [Sub-Gaussianity, see Section 2.3 in [BLM13]] Let μ denote the (unknown) mean of the time series X. The marginals $X_t - \mu$ of the stationary time series $(X_t - \mu)_{t\in\mathbb{Z}}$ are sub-Gaussian with variance factor $\nu > 0$, that is,

$$\psi_{X_t-\mu}(\lambda) \leq \frac{\lambda^2\nu}{2} \qquad \forall\lambda \in \mathbb{R},$$

where $\psi_{X_t-\mu}(\lambda) = \log \mathbf{E}e^{\lambda(X_t-\mu)}$ denotes the logarithmic moment-generating function of the random variable $X_t - \mu$.

Note that we do not assume the mean μ to be known for our analysis. A direct consequence of Assumption 3.1 is the bound

$$\mathbf{P}(|X_t - \mu| > x) \leq 2e^{-x^2/(2\nu)} \qquad \text{for all } x > 0, \tag{3.4}$$

see, for instance, [BLM13], Theorem 2.1. We will only need this bound for our further results.

Theorem 3.2 (Upper bound)**.** *Let Assumptions 1.1 and 3.1 hold. Further assume that the model* $\mathfrak{m}$ *is given by a subspace* $S_\mathfrak{m}$ *of symmetric functions that satisfies* $\|\varphi_i\|_\infty \leq C\sqrt{n}$ *for all* $i \in \mathcal{I}_\mathfrak{m}$*. Let* Z_i *be defined as in (2.5) with* $\tau_n^2 = 56\nu\log(n)$*. Consider the estimator* $\widehat{f}_\mathfrak{m}$ *defined through Equations (3.2) and (3.3). Then,*

$$\mathbf{E}\|\widehat{f}_\mathfrak{m} - f\|^2 \leq \|f - f_\mathfrak{m}\|^2 + CD_\mathfrak{m}(1 + \log(n))\left[\frac{1}{n} \vee \frac{\tau_n^4}{n\alpha^4}\right] \tag{3.5}$$

where $f_\mathfrak{m}$ *denotes the projection of* f *on the space* $S_\mathfrak{m}$*.*

Remark 3.3. Of course, if the time series X is known to be bounded, say $|X_t| \leq K$ for all $t \in \mathbb{Z}$, the quantity τ_n in this section can be replaced with K which removes at least some of the logarithmic factors (the ones arising via τ_n) in the upper bound.

Remark 3.4. In the proof of Theorem 3.2, Assumption 3.1 is only needed to bound the probability of the event $\{\exists i : X_i \neq \widetilde{X}_i\}$. For this purpose, the assumption of sub-Gaussianity might be replaced with assuming subexponential tails for the marginals. This would lead to a slightly different (but still logarithmic in terms of n) definition of the truncation threshold τ_n. However, for Theorem 4.4 we will have to impose Gaussian marginals.

Remark 3.5. The quantity ν in Assumption 3.1 is usually not given to the statistician but can be easily replaced by taking an estimator for this upper variance bound instead.

Example 3.6 (Sobolev ellipsoids and analytic functions)**.** In order to illustrate the upper bound (3.5), we consider the case where each model can be identified with a natural number: we have $\mathfrak{m} \in \mathbb{N}_0$, set $\mathcal{I}_\mathfrak{m} = [\![-\mathfrak{m}, \mathfrak{m}]\!]$, and $S_\mathfrak{m} = \operatorname{span}(\mathbf{e}_j)_{j\in\mathcal{I}_\mathfrak{m}}$ with $\mathbf{e}_j(\omega) = \exp(-\mathrm{i}j\omega)$ denoting the (complex) Fourier basis functions. In terms of these basis functions, smoothness may be expressed by assuming membership of f to an ellipsoid

$$\mathcal{F}(\beta, L) = \left\{ f = \sum_{j\in\mathbb{Z}} f_j\mathbf{e}_j : f \geq 0 \text{ and } \sum_{j\in\mathbb{Z}} f_j^2\beta_j^2 \leq L^2 \right\}$$

where $L > 0$ and $\beta = (\beta_j)_{j\in\mathbb{Z}}$ is a strictly positive symmetric sequence such that $\beta_0 = 1$ and $(\beta_n)_{n\in\mathbb{N}_0}$ is non-decreasing. Typical examples of sequences include the cases where $\beta_j \asymp |j|^s$ (Sobolev ellipsoids) and $\beta_j \asymp \exp(p|j|)$ for some $p \geq 0$ (class of analytic functions). Under our assumptions, the squared bias in the upper bound (3.5) may be bounded as

$$\|f_\mathfrak{m} - f\|^2 = \sum_{|j|>\mathfrak{m}} f_j^2 \leq \beta_\mathfrak{m}^{-2} \sum_{|j|>\mathfrak{m}} f_j^2\beta_j^2 \leq L^2\beta_\mathfrak{m}^{-2}.$$

Thus, the trade-off between squared bias and variance is equivalent to the best compromise between $\beta_{\mathfrak{m}}^{-2}$ and $(2\mathfrak{m}+1)\cdot(1+\log(n))\left[1/n \vee \tau_n^4/(n\alpha^4)\right]$. In the polynomial case $\beta_j = |j|^s$, the best compromise is realized by choosing $\mathfrak{m}^* \asymp \left[(1+\log(n))\left(1/n \vee \tau_n^4/(n\alpha^4)\right)\right]^{-1/(2s+1)}$ leading to the rate $\left[(1+\log(n))\left(1/n \vee \tau_n^4/(n\alpha^4)\right)\right]^{2s/(2s+1)}$. It is remarkable in the setup of spectral density estimation that also the part of the rate in terms of the privacy parameter α does not suffer from a loss in the exponent whereas in the setup of density estimation the optimal non-private rate $n^{-2s/(2s+1)}$ deteriorates to $n^{-2s/(2s+1)} \vee (n\alpha^2)^{-s/(s+1)}$ under differential privacy. In the case where $\beta_j = \exp(p|j|)$, we take $\mathfrak{m}^* \asymp \log(n) + \log(\alpha)$ to obtain the rate $(\log(n)+\log(\alpha))\cdot(1+\log(n))\left[1/n \vee \tau_n^4/(n\alpha^4)\right]$.

Lower bounds. In this subsection, we derive minimax lower bounds for function classes that can be written as ellipsoids in terms of the Fourier coefficients of the function, that is, the classes $\mathcal{F}(\beta, L)$ introduced in Example 3.6. As discussed above, this general approach includes Sobolev ellipsoids and classes of analytic functions. We determine both a non-private and and private lower bound, the former one valid already in the framework where a snippet from the original time series X can be observed, the second one being special to the considered privacy scenario with observation $Z_{1:n}$.

Theorem 3.7 (Lower bound)**.** *Assume that the time series X is Gaussian, and consider the class $\mathcal{F}(\beta, L)$ of potential spectral densities introduced in Example 3.6. Further assume that anonymized data $Z_{1:n}$ are generated via a (potentially interactive) channel Q ensuring local differential privacy.*

a) Assume that $B := \sum_{j\in\mathbb{Z}} \beta_j^{-2} < \infty$. Define k_n^* and Ψ_n via

$$k_n^* = \operatorname*{argmin}_{k\in\mathbb{N}} \left[\max\left(\beta_k^{-2}, \frac{2k+1}{n}\right)\right],$$

$$\Psi_n = \max\left(\beta_{k_n^*}^{-2}, \frac{2k_n^*+1}{n}\right),$$

and assume that there is a positive constant η such that

$$0 < \eta^{-1} \le \Psi_n^{-1} \min\left\{\beta_{k_n^*}^{-2}, \frac{2k_n^*+1}{n}\right\}.$$

Then,

$$\inf_{\widetilde{f}} \sup_{f\in\mathcal{F}(\beta,L)} \mathbf{E}\|\widetilde{f} - f\|^2 \gtrsim \frac{2k_n^*+1}{n}$$

holds where the infimum is taken over all estimators $\widetilde{f}$ of f based on the privatized sample $Z_{1:n}$.

b) It holds

$$\inf_{\widetilde{f}} \sup_{f\in\mathcal{F}(\beta,L)} \mathbf{E}\|\widetilde{f} - f\|^2 \gtrsim \min\left\{\frac{\pi}{n(e^\alpha-1)^2}, \frac{L^2}{4}\right\},$$

where the infimum is taken over all estimator $\widetilde{f}$ of f based on the privatized sample $Z_{1:n}$.

Remark 3.8. The proof of statement a) of Theorem 3.7 is based on a reduction to estimators in terms of the original sample $X_{1:n}$. Indeed, any lower bound valid for estimators in terms of the original sample stays valid in the privacy case since working with differentially private data can equivalently be interpreted as restricting the set of potentially available estimators from the set of all measurable functions in terms of $X_{1:n}$ to the set of functions of the form $\widetilde{f}\circ Q$ where Q is a channel that yields differential privacy and $\widetilde{f}$ any measurable function in terms of $Z_{1:n}$. In the appendix, we give the complete proof since we were not able to find a good reference in the existing literature (the articles [Ben85] and [Efr98] consider different function classes).

By combining the non-private and the private lower bound we directly obtain the following corollary.

Corollary 3.9. *Under the Assumptions of Theorem 3.7 we have*

$$\inf_{\widetilde{f}} \sup_{f \in \mathcal{F}(\beta, L)} \mathbf{E}\|\widetilde{f} - f\|^2 \gtrsim \max\left\{\Psi_n^2, \min\left\{1, \frac{1}{n(e^\alpha - 1)^2}\right\}\right\}.$$

Remark 3.10. Up to logarithmic factors the lower bounds determined coincide with the given upper bounds. However, our results here do not give a complete answer concerning the exact dependence of the optimal convergence rate in terms of the privacy parameter α. Intuitively, part b) states only the loss that can be explained from the constant basis function when the spectral density is written in terms of the trigonometric basis. It can already be seen here that a deterioration of the usual rate (given by part a)) is unavoidable if α is too small. In this case, one can obtain a lower bound by comparing distributions characterized by two different but constant spectral densities (see the proof of part b) in the appendix). Then, there is no dependence between the X_t, that is, we have access to an i.i.d. sample and the well-known information theoretic inequalities for differential privacy from the paper [DJW18] are available. These data processing inequalities do not longer hold for dependent X_t. Developing tools in this direction that help to understand the exact scaling behaviour represent an interesting point of departure for further investigations. Note also that even for the Fourier coefficient associated with the constant basis function we do not have coincidence for the scaling in terms of α: we have a term $1/(n(e^\alpha - 1)^2)$ in the lower bound (which behaves as $1/(n\alpha^2)$ for small α) but a term of order $1/(n\alpha^4)$ in the upper bound (plus extra logarithmic factors). This last issue might be tackled by publishing an anonymized version of X_t^2 in addition to the privatized version of X_t since for computation of the empirical correlation coefficient associated with the constant basis function no interaction between the data holders is necessary.

4. Risk bound for the adaptive estimator

In Section 3 we have derived the upper risk bound (3.5) for fixed models $\mathfrak{m}$. The near optimality for the class of Sobolev ellipsoids was equally illustrated in Example 3.6 and the accompanying lower bounds established in Theorem 3.7. The performance of the rate optimal estimators hinges on the choice of a suitable approximating model $\mathfrak{m}$ the choice of which depends on both the sample size n and the regularity of the functions in the considered function class. Since such regularity assumptions are usually not realistic to be fulfilled, there is need to obtain a suitable model in completely data-driven way.

In order to define the adaptive estimator, first introduce the contrast

$$\Upsilon_n(t) = \int_{-\pi}^{\pi} t^2(\omega)\mathrm{d}\omega - 2\int_{-\pi}^{\pi} t(\omega)\widehat{I}_n(\omega)\mathrm{d}\omega.$$

Note that, the estimator $\widehat{f}_{\mathfrak{m}}$ associated with the fixed model $\mathfrak{m}$ in Section 3 satisfies

$$\Upsilon_n(\widehat{f}_{\mathfrak{m}}) = \min_{t \in \mathcal{S}_{\mathfrak{m}}} \Upsilon_n(t).$$

The model selection step is performed by putting

$$\widehat{\mathfrak{m}} = \operatorname*{argmin}_{\mathfrak{m} \in \mathcal{M}_n} \{\Upsilon_n(\widehat{f}_{\mathfrak{m}}) + \mathrm{pen}(\mathfrak{m})\} \tag{4.1}$$

where $\mathcal{M}_n$ is some set of potential models, and $\mathrm{pen}\colon \mathcal{M}_n \to [0, \infty)$ a penalty function is given by

$$\mathrm{pen}(\mathfrak{m}) = CD_{\mathfrak{m}} \max\left\{\frac{1}{n}, \frac{\tau_n^4}{n\alpha^4}\right\} \cdot (L_{\mathfrak{m}}^4 + L_{\mathfrak{m}} + \log(n)) \cdot (1 + \|f\|_\infty)^2 \tag{4.2}$$

for some constant $C > 0$ that has to be chosen large enough. Finally, the adaptive estimator of the spectral density f is defined as

$$\widetilde{f} = \widehat{f}_{\widehat{\mathfrak{m}}}.$$

Before we can state our main result, we have to impose the following assumptions on the collection $\mathcal{M}_n$ of models. These assumptions are already present in the work of [Com01], and no extra assumptions on the models are needed in the privacy framework.

Assumption 4.1. Each $\mathcal{S}_\mathfrak{m}$ is a linear finite-dimensional subspace of $L^2([-\pi,\pi])$ containing symmetric functions with dimension $\dim(S_\mathfrak{m}) = D_\mathfrak{m} \geq 1$. Moreover, $D_n := \max_{\mathfrak{m}\in\mathcal{M}_n} D_\mathfrak{m} \leq n$.

Assumption 4.2. Let $(\varphi_i)_{i\in\mathcal{I}_\mathfrak{m}}$ be an orthonormal basis of $S_\mathfrak{m}$, and $\beta = (\beta_i)_{i\in\mathcal{I}_\mathfrak{m}} \in \mathbb{R}^{D_\mathfrak{m}}$. Set $|\beta|_\infty = \sup_{i\in\mathcal{I}_\mathfrak{m}} |\beta_i|$. Then, for all $\mathfrak{m} \in \mathcal{M}_n$,

$$\bar{r}_\mathfrak{m} := \frac{1}{\sqrt{D_\mathfrak{m}}} \sup_{\beta\neq 0} \frac{\|\sum_{i\in\mathcal{I}_\mathfrak{m}} \beta_i\varphi_i\|_\infty}{|\beta|_\infty} \leq C_{\bar{r}}\sqrt{\frac{n}{D_\mathfrak{m}}}.$$

Assumption 4.3. $\sum_{\mathfrak{m}\in\mathcal{M}_n} e^{-L_\mathfrak{m} D_\mathfrak{m}} \leq C_L < \infty$ for some positive weights $L_\mathfrak{m}$.

Remarks 2.3–2.6 from [Com01] show that Assumptions 4.2 and 4.3 are satisfied for the models mentioned in the introduction for suitable values $C_{\bar{r}}$ and $L_\mathfrak{m}$.

Theorem 4.4. *Let Assumption 1.1 hold. Let Z_i be defined as in (2.5) with $\tau_n^2 = 56\nu\log(n)$. Consider the estimator $\widetilde{f} = \widehat{f}_{\widehat{\mathfrak{m}}}$ defined through Equations (3.2), (3.3), and (4.1) where the penalty function is defined in (4.2). Let Assumptions 4.1–4.3 hold. Then,*

$$\mathbf{E}\|\widetilde{f} - f\|^2 \lesssim \inf_{\mathfrak{m}\in\mathcal{M}_n} [\|f - f_\mathfrak{m}\|^2 + \mathrm{pen}(\mathfrak{m})] + C(C_{\bar{r}}, \|f\|_\infty) \max\left\{\frac{1}{n}, \frac{\tau_n^4}{n^3\alpha^4}\right\}.$$

Remark 4.5. Unfortunately, the definition of the penalty function introduced above depends on the unknown value $\|f\|_\infty$. In practise, one can replace this quantity by an appropriate estimator. Theoretical results can be proved for this more realistic estimator as well. We do not realize this here, and refer the interested reader to the papers [Com01] and [Kro19] where this idea has been put into practise. The resulting fully-adaptive estimator can be shown to satisfy an oracle inequality as in the case of known $\|f\|_\infty$ under mild assumptions.

5. Numerical study

In this section, we illustrate our findings by a small simulation study. The code that can be used to (re)produce the results is available under

https://gitlab.com/kroll.martin/adaptive-private-spectral-density-estimation.

We consider the same time series model as [Neu96] and [Com01], that is, we consider the time series $(X_t)_{t\in\mathbb{Z}}$ defined as

$$X_t = X_t^{\mathrm{ARMA}} + \sigma X_t^{\mathrm{WN}}$$

where X_t^{ARMA} is an ARMA(2,2)-process,

$$X_t^{\mathrm{ARMA}} + a_1 X_{t-1}^{\mathrm{ARMA}} + a_2 X_{t-2}^{\mathrm{ARMA}} = b_0\varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2},$$

and $(\varepsilon_t)_{t\in\mathbb{Z}}$ and $(X_t^{\mathrm{WN}})_{t\in\mathbb{Z}}$ are independent Gaussian white noise processes with unit variance. From the cited papers we also adopt the choices of the parameters ($a_1 = 0.2$, $a_2 = 0.9$, $b_0 = 1$, $b_1 = 0$, $b_2 = 1$, and $\sigma = 0.5$). We consider time series snippets of length $n \in \{10000, 20000\}$ and simulate $T = 100$ replications of each setup. In contrast to the mentioned papers, which consider a non-private framework, our principal aim is to illustrate the effect of the privacy level α. For this purpose, we consider $\alpha \in \{+\infty, 5, 2.5\}$ where formally putting $\alpha = +\infty$ corresponds to the case without any privacy constraint.

Note that these choices of the privacy parameter are very conservative and provide only a moderate anonymization of the data (see, for instance, Figure 3 in [DJW18] where the link between the privacy parameter and a hypothesis testing problem is illustrated).

For each parameter setup, we computed the mean L^2-risk over the $T = 100$ replications, its standard deviation $\widehat{v}$, and the $\pm 95\%$ confidence intervals defined as $1.96\widehat{v}/\sqrt{T}$ (see [Com01, Neu96]). We slightly modified the method considered in Sections 3 and 4 for the theoretical analysis in order to perform our simulation experiments. Instead of a logarithmically increasing sequence τ_n (which was principally introduced to control the probability of the event A^{c} introduced in the analysis in the appendix), we took $\tau_n = 4$ after some calibrations. As [Com01], we restricted ourselves to histogram estimators of the spectral density. For a given dimension $D_{\mathfrak{m}} = d$, the orthonormal basis functions are defined as

$$\varphi_j^{(d)} = \sqrt{\frac{d}{\pi}} \mathbf{1}_{[\pi j/d, \pi(j+1/d)}, \qquad j \in [\![0, d-1]\!]$$

(we define the basis functions only on $[0, \pi)$ and extend the final estimator on the interval $[-\pi, \pi]$ by exploiting the symmetry of the target spectral density). For the model $\mathfrak{m}$, the estimator $\widehat{f}_{\mathfrak{m}}$ is then

$$\widehat{f}_{\mathfrak{m}} = \sum_{j=0}^{d-1} \widehat{a}_j^{(d)} \varphi_j^{(d)}$$

where the estimated coefficients are calculated via the formula

$$\widehat{a}_j^{(d)} = \sqrt{\frac{d}{\pi}} \left[\frac{c_0}{2d} + \frac{1}{\pi} \sum_{r=1}^{n-1} \frac{c_r}{r} \left(\sin\left(\frac{\pi(j+1)r}{d} \right) - \sin\left(\frac{\pi j r}{d} \right) \right) \right]$$

for $j \in [\![0, d-1]\!]$ where $c_r = c_{r,n}$, $r \in [\![0, n-1]\!]$ are the empirical covariances of the masked data $Z_{1:n}$, that is,

$$c_{r,n} = \frac{1}{n} \sum_{k=1}^{n-r} (Z_k - \bar{Z}_n)(Z_{k+r} - \bar{Z}_n)$$

(the value $c_{0,n}$ has to be modified by subtracting $8\tau_n^2/\alpha^2$ afterwards). Ignoring logarithmic factors and constants in the theoretical penalty in Section 4, this leads to the following form of the penalized contrast criterion:

$$-\sum_{j=0}^{d-1} (\widehat{a}_j^{(d)})^2 + \frac{\kappa d}{n} \max\left\{1, \frac{\tau_n^4}{\alpha^4}\right\}$$

(we took $\kappa = 1$). We minimized this criterion over potential dimensions $d \in [\![1, 50]\!]$. The results of our simulation study are summarized in Table 1 and illustrated in Figures 1, 2, and 3 (note the different scaling of the y-axes in the plots). A profound loss of performance is encountered for decreasing values of α which can be compensated with taking considerably longer snippets from the time series only. This might make inference from privatized data difficult in applications where only samples of moderate size can be collected.

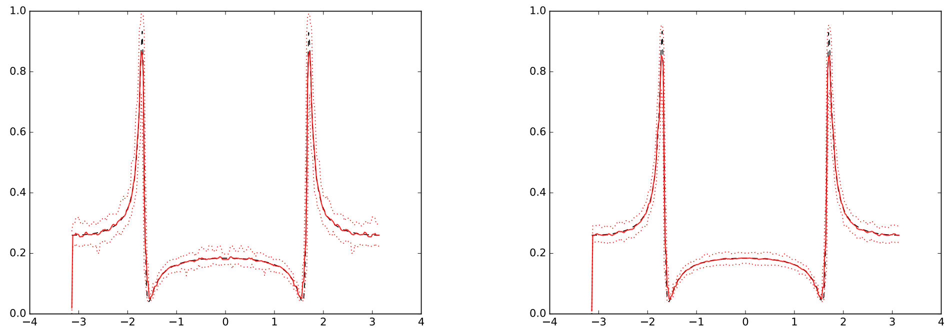
6. Summary and outlook

In this paper, we have extended the model selection approach for adaptive nonparametric spectral density estimation to the framework of local α-differential privacy. We were able to derive an oracle inequality similar to the one in the non-private setup. Since the proposed adaptive procedure is limited to Gaussian time series it might also be of interest to study whether known adaptive estimators that work in non-Gaussian frameworks (for instance, the wavelet estimator considered in [Neu96]) can also be transferred to the

Table 1. Results of the simulation study. The table contains the mean of the L^2-risk over $T = 100$ replications of the experiment, and the $\pm 95\%$ intervals computed as in [Neu96] as $1.96\widehat{v}/\sqrt{T}$ where $\widehat{v}$ is the standard deviation.

	n = 10000			n = 20000		
α	$+\infty$	5.0	2.5	$+\infty$	5.0	2.5
L^2-risk	0.00216	0.01316	0.13629	0.00159	0.00734	0.07126
$\pm$ 95% CI	0.00012	0.00048	0.00464	0.00007	0.00022	0.00243

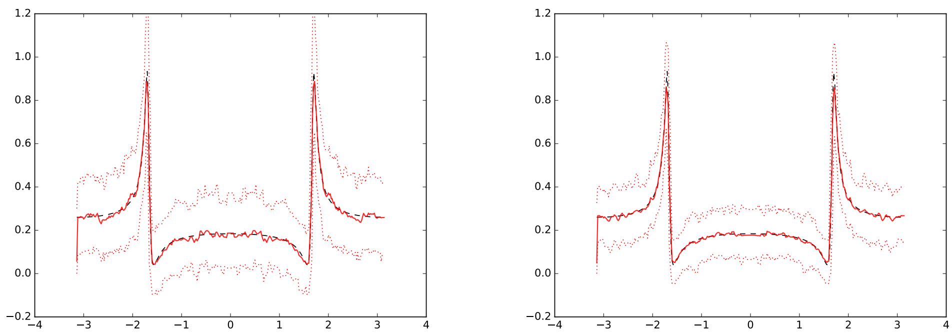
$\boldsymbol{\alpha = +\infty}$



a) Snippet length $n = 10000$ **b)** Snippet length $n = 20000$

Figure 1. The figures show for the two considered snippet sizes $n \in \{10000, 20000\}$ the mean of the estimator (red solid line) and both the 0.95 and 0.05 pointwise quantile (red dotted lines) over $T = 100$ replications for the case $\alpha = +\infty$ (this corresponds to the case without privacy constraints). The true spectral density is represented as a black dashed line.

$\boldsymbol{\alpha = 5.0}$

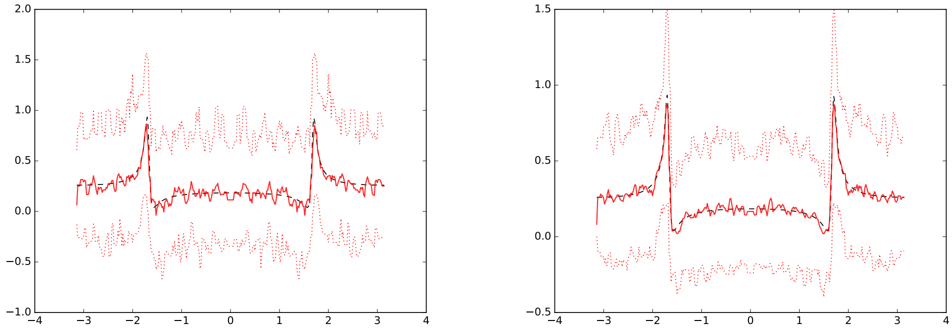


a) Snippet length $n = 10000$ **b)** Snippet length $n = 20000$

Figure 2. The figures show for the two considered snippet sizes $n \in \{10000, 20000\}$ the mean of the estimator (red solid line) and both the 0.95 and 0.05 pointwise quantile (red dotted lines) over $T = 100$ replications for the case $\alpha = 5.0$. The true spectral density is represented as a black dashed line.

framework of the present paper. The exact dependence of minimax rates of convergence on the privacy parameter as well as the unclear necessity of logarithmic factors in these rates is a remaining open problem that hopefully stimulates the development of further

$\alpha = 2.5$



a) Snippet length $n = 10000$ **b)** Snippet length $n = 20000$

FIGURE 3. The figures show for the two considered snippet sizes $n \in \{10000, 20000\}$ the mean of the estimator (red solid line) and both the 0.95 and 0.05 pointwise quantile (red dotted lines) over $T = 100$ replications for the case $\alpha = 2.5$. The true spectral density is represented as a black dashed line.

theoretical results. In addition, a more detailed series of simulation experiments seems to be necessary in order to calibrate an estimator that produces reliable results in practise.

APPENDIX A. PROOFS OF SECTION 3

The following result (which is valid without any distributional assumptions on the stationary time series) has been proven in [Com01].

Proposition A.1 ([Com01], Proposition 1)**.** *Let X be a stationary sequence with autocovariance function satisfying Assumption 1.1. Then*

$$\int_{-\pi}^{\pi} (f(\omega) - \mathbf{E}(I_n(\omega)))^2 \mathrm{d}\omega \le \frac{M_1 + 39M^2}{2\pi n} =: \frac{M_2}{n}.$$

This result can also be applied to the time series Z'. Then the constant M_1 does not change but for the constant M we have $M^{Z'} = M^X + 8\tau_n^2/\alpha^2$.

A.1. **Proof of Theorem 3.2 (Upper bound for fixed model $\mathfrak{m}$).**
Let us introduce the event A and its complement defined as follows:

$$A = \bigcap_{i=1}^{n} \{\widetilde{X}_i = X_i\}, \qquad A^{\mathsf{c}} = \bigcup_{i=1}^{n} \{\widetilde{X}_i \neq X_i\}.$$

As above, let us denote with $f_{\mathfrak{m}}$ the projection of f on the space $S_{\mathfrak{m}}$. We have the decomposition

$$\begin{aligned} \mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f\|^2 &= \|f_{\mathfrak{m}} - f\|^2 + \mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2 \\ &= \|f_{\mathfrak{m}} - f\|^2 + \mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2 \mathbf{1}_A + \mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2 \mathbf{1}_{A^{\mathsf{c}}}. \end{aligned} \tag{A.1}$$

The first (pure bias) term on the right-hand side is already in the form of the statement of the theorem, and we have to study the terms including $\mathbf{1}_A$ and $\mathbf{1}_{A^{\mathsf{c}}}$ only.

Bound for $\mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2 \mathbf{1}_A$: By the very definition of A we have $X_i = \widetilde{X}_i$ on A, and hence $Z_i = Z_i' = X_i + \xi_i$ for $\xi_i \sim \mathcal{L}(2\tau_n/\alpha)$. Hence, on the event A the identity

$$I_n^Z(\omega) = I_n^{Z'}(\omega)$$

holds (with $I_n^{Z'}$ defined exactly as I_n^Z with Z replaced with Z'), and we have

$$\begin{aligned}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2 \mathbf{1}_A &= \sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle f - \widehat{I}_n, \varphi_i \rangle|^2 \mathbf{1}_A \\ &= \sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle f - (I_n^Z(\omega) - \frac{8\tau_n^2}{\alpha^2}), \varphi_i \rangle|^2 \mathbf{1}_A \\ &= \sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle f - (I_n^{Z'}(\omega) - \frac{8\tau_n^2}{\alpha^2}), \varphi_i \rangle|^2 \mathbf{1}_A \\ &\leq \sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle f - (I_n^{Z'}(\omega) - \frac{8\tau_n^2}{\alpha^2}), \varphi_i \rangle|^2 \\ &= \sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle f^{Z'} - I_n^{Z'}, \varphi_i \rangle|^2.\end{aligned}$$

where the last identity is established in (3.1). From this we get

$$\begin{aligned}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2 \mathbf{1}_A &\leq 2 \sum_{i \in \mathcal{I}_{\mathfrak{m}}} (|\langle f^{Z'} - \mathbf{E} I_n^{Z'}, \varphi_i \rangle|^2 + |\langle \mathbf{E} I_n^{Z'} - I_n^{Z'}, \varphi_i \rangle|^2) \\ &= 2\|(f^{Z'} - \mathbf{E} I_n^{Z'})_{\mathfrak{m}}\|^2 + 2 \sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle \mathbf{E} I_n^{Z'} - I_n^{Z'}, \varphi_i \rangle|^2 \\ &\leq 2\|f^{Z'} - \mathbf{E} I_n^{Z'}\|^2 + 2 \sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle \mathbf{E} I_n^{Z'} - I_n^{Z'}, \varphi_i \rangle|^2\end{aligned}$$

in order to bound the term $\|f^{Z'} - \mathbf{E} I_n^{Z'}\|^2$, we use Proposition A.1 in order to obtain:

$$\|f^{Z'} - \mathbf{E} I_n^{Z'}\|^2 = \int_{-\pi}^{\pi} (f^{Z'} - \mathbf{E} I_n(\omega))^2 \mathrm{d}\omega \leq \max\left(\frac{M_1}{\pi n}, \frac{39(M^{Z'})^2}{\pi n}\right).$$

Note that Assumption 1.1 can also be applied to the time series Z' instead of X with $M_1 = M_1^{Z'} = M_1^X$ and with $M = M^X$ replaced with $M^{Z'} = M^X + \frac{8\tau_n^2}{\alpha^2}$. Hence,

$$\|f^{Z'} - \mathbf{E} I_n^{Z'}\|^2 \lesssim \max\left(\frac{\tau_n^4}{n\alpha^4}, \frac{1}{n}\right). \tag{A.2}$$

Let us now consider the expression $\mathbf{E} \sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle \mathbf{E} I_n^{Z'} - I_n^{Z'}, \varphi_i \rangle|^2$. We write

$$I_n^{Z'}(\omega) = I_n^X + I_n^\xi + \widetilde{I}_n,$$

where

$$\begin{aligned} I_n^X &= \frac{1}{2\pi n} \left| \sum_{t=1}^n (X_t - \bar{X}_n) e^{-\mathrm{i}t\omega} \right|^2, \\ I_n^\xi &= \frac{1}{2\pi n} \left| \sum_{t=1}^n (\xi_t - \bar{\xi}_n) e^{-\mathrm{i}t\omega} \right|^2, \qquad \text{and} \\ \widetilde{I}_n &= \frac{1}{2\pi n} \left(\sum_{t=1}^n (X_t - \bar{X}_n) e^{-\mathrm{i}t\omega} \right) \left(\sum_{t=1}^n (\xi_t - \bar{\xi}_n) e^{\mathrm{i}t\omega} \right) + \\ &\quad + \frac{1}{2\pi n} \left(\sum_{t=1}^n (X_t - \bar{X}_n) e^{\mathrm{i}t\omega} \right) \left(\sum_{t=1}^n (\xi_t - \bar{\xi}_n) e^{-\mathrm{i}t\omega} \right).\end{aligned}$$

Hence, by exploiting that $\mathbf{E}\widetilde{I}_n = 0$, we obtain

$$\sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle \mathbf{E} I_n^{Z'} - I_n^{Z'}, \varphi_i \rangle|^2 \leq \sum_{i \in \mathcal{I}_{\mathfrak{m}}} |\langle \mathbf{E} I_n^X - I_n^X, \varphi_i \rangle|^2$$

$$+\sum_{i\in\mathcal{I}_\mathfrak{m}}|\langle \mathbf{E}I_n^\xi - I_n^\xi,\varphi_i\rangle|^2 + \sum_{i\in\mathcal{I}_\mathfrak{m}}|\langle \mathbf{E}\widetilde{I}_n - \widetilde{I}_n,\varphi_i\rangle|^2.$$

Put

$$G_{i,X}(\mathfrak{m}) = \sup_{\zeta\in\{\pm1\}}\langle I_n^X - \mathbf{E}I_n^X,\zeta\varphi_i\rangle,$$
$$G_{i,\xi}(\mathfrak{m}) = \sup_{\zeta\in\{\pm1\}}\langle I_n^\xi - \mathbf{E}I_n^\xi,\zeta\varphi_i\rangle,$$
$$\widetilde{G}_i(\mathfrak{m}) = \sup_{\zeta\in\{\pm1\}}\langle \widetilde{I}_n,\zeta\varphi_i\rangle.$$

Then, for any constant $\kappa_X>0$, we have

$$\mathbf{E}\langle \mathbf{E}I_n^X - I_n^X,\varphi_i\rangle^2 \le \mathbf{E}\left[\left((G_{i,X}(\mathfrak{m}))^2 - \frac{4\kappa_X\|f\|_\infty^2(1+C_{\bar r}^2)}{n}\right)_+\right] + \frac{4\kappa_X\|f\|_\infty^2(1+C_{\bar r}^2)}{n}.$$

Hence, by Lemma C.3[1] we get

$$\mathbf{E}\langle \mathbf{E}I_n^X - I_n^X,\varphi_i\rangle^2 \le \frac{C(C_{\bar r},\|f\|_\infty)}{n}$$

provided that κ_X is sufficiently large. Analogously, for the terms incorporating I_n^ξ and $\widetilde{I}_n$, we obtain with sufficiently large constants $\kappa_\xi,\widetilde{\kappa}>0$ by using Lemmata C.5 and C.7

$$\mathbf{E}\langle \mathbf{E}I_n^\xi - I_n^\xi,\varphi_i\rangle^2 \le \kappa_\xi\frac{\tau_n^4(1+\log(n))}{n\alpha^4} + \frac{C(C_{\bar r})\tau_n^4}{n^3\alpha^4}$$

and

$$\mathbf{E}\langle \mathbf{E}\widetilde{I}_n - \widetilde{I}_n,\varphi_i\rangle^2 \le \widetilde{\kappa}(3+4\tau_n/\alpha)^4(1+\|f\|_\infty^2)(1+\log(n))\frac{1}{n} + \frac{C(C_{\bar r},\|f\|_\infty)(3+4\tau_n/\alpha)^4}{n^3},$$

respectively. Putting the obtained estimates together, we get

$$\mathbf{E}\sum_{i\in\mathcal{I}_\mathfrak{m}}|\langle \mathbf{E}I_n^{Z'} - I_n^{Z'},\varphi_i\rangle|^2 \le D_\mathfrak{m}C(C_{\bar r},\|f\|_\infty)(1+\log(n))\left[\frac{1}{n}\vee\frac{\tau_n^4}{n\alpha^4}\right]. \tag{A.3}$$

Combining (A.2) and (A.3), we obtain

$$\mathbf{E}\|\widehat{f}_\mathfrak{m} - f_\mathfrak{m}\|^2\mathbf{1}_A \lesssim D_\mathfrak{m}(1+\log(n))\left[\frac{1}{n}\vee\frac{\tau_n^4}{n\alpha^4}\right].$$

Bound for $\mathbf{E}\|\widehat{f}_\mathfrak{m} - f_\mathfrak{m}\|^2\mathbf{1}_{A^c}$: By the Cauchy-Schwarz inequality, we have

$$\mathbf{E}\|\widehat{f}_\mathfrak{m} - f_\mathfrak{m}\|^2\mathbf{1}_{A^c} \le (\mathbf{E}\|\widehat{f}_\mathfrak{m} - f_\mathfrak{m}\|^4)^{1/2}\cdot(\mathbf{P}(A^c))^{1/2}, \tag{A.4}$$

and we analyse the two factors on the right-hand side separately. First,

$$\mathbf{E}\|\widehat{f}_\mathfrak{m} - f_\mathfrak{m}\|^4 = \mathbf{E}\left[\left(\sum_{i\in\mathcal{I}_\mathfrak{m}}|\langle f-\widehat{I}_n,\varphi_i\rangle|^2\right)^2\right]$$

---

[1] Admittedly, using Lemmata C.3, C.5, and C.7 here is like using a sledgehammer to crack a nut. At least for the term containing I_n^X we can directly refer to p. 294 in [Com01] for an alternative reasoning. For the other terms, one could perform in the same manner with some tedious calculations but we do currently not see how one could establish an upper bound without any logarithmic terms and a better dependence on α than in our current estimate. Note that instead of Assumption 4.2 we only need to assume that $\|\varphi_i\|_\infty \le C\sqrt{n}$ for the an orthonormal basis $(\varphi_i)_{i\in\mathcal{I}_\mathfrak{m}}$ of the considered model. In addition, we can also put $L_\mathfrak{m}=1$ here since in contrast to the proof of Theorem 4.4 no summation over all potenial models is performed.

$$
\begin{aligned}
&= \mathbf{E}\left[\left(\sum_{i\in\mathcal{I}_{\mathfrak{m}}} |\langle f + \frac{8\tau_n^2}{\alpha^2} - I_n^Z, \varphi_i\rangle|^2 \right)^2 \right] \\
&\leq \mathbf{E}\left[\left(\sum_{i\in\mathcal{I}_{\mathfrak{m}}} \| f + \frac{8\tau_n^2}{\alpha^2} - I_n^Z \|^2 \right)^2 \right] \\
&= \mathbf{E}\left[D_{\mathfrak{m}}^2 \cdot \| f + \frac{8\tau_n}{\alpha^2} - I_n^Z\|^4 \right].
\end{aligned}
$$

Now,

$$
\begin{aligned}
\| f + \frac{8\tau_n}{\alpha^2} - I_n^Z\|^4 &\leq 4\pi^2 \cdot \| f + \frac{8\tau_n}{\alpha^2} - I_n^Z\|_\infty^4 \\
&\leq 32\pi^2 \cdot \| f + \frac{8\tau_n^2}{\alpha^2}\|_\infty^4 + 32\pi^2 \|I_n^Z\|_\infty^4 \\
&\leq 32\pi^2 \left(\frac{1}{2\pi} \sum_{k\in\mathbb{Z}} |\gamma(k)| + \frac{8\tau_n^2}{\alpha^2} \right)^4 + 32\pi^2 \|I_n^Z\|_\infty^4.
\end{aligned}
$$

Furthermore, using $|\widetilde{X}_t| \leq \tau_n$,

$$
\begin{aligned}
\mathbf{E}\left[\|I_n^Z\|_\infty^4\right] &= \mathbf{E}\left[\frac{1}{(2\pi n)^4} \|\sum_{t=1}^n (\widetilde{X}_t - \bar{\widetilde{X}}_n) e^{-\mathrm{i}t\omega} + \sum_{t=1}^n (\xi_t - \bar{\xi}_n) e^{-\mathrm{i}t\omega}\|_\infty^8 \right] \\
&\leq \frac{2^7}{(2\pi n)^4} \cdot \mathbf{E}\left[\|\sum_{t=1}^n (\widetilde{X}_t - \bar{\widetilde{X}}_n) e^{-\mathrm{i}t\omega}\|_\infty^8 \right] + \frac{2^7}{(2\pi n)^4} \mathbf{E}\left[\|\sum_{t=1}^n (\xi_t - \bar{\xi}_n) e^{-\mathrm{i}t\omega}\|_\infty^8 \right] \\
&\leq \frac{2^7}{(2\pi n)^4} \cdot \mathbf{E}\left[\left(\sum_{t=1}^n |\widetilde{X}_t - \bar{\widetilde{X}}_n| \right)^8 \right] + \frac{2^7}{(2\pi n)^4} \cdot \mathbf{E}\left[\left(\sum_{t=1}^n |\xi_t| + n|\bar{\xi}_n| \right)^8 \right] \\
&\leq \frac{2^{15} n^8 \tau_n^8}{(2\pi n)^4} + \frac{2^7}{(2\pi n)^4} \cdot \mathbf{E}\left[\left(2\sum_{t=1}^n |\xi_t| \right)^8 \right] \\
&\leq \frac{2^{11} n^4 \tau_n^8}{\pi^4} + \frac{2^{15}}{(2\pi n)^4} \cdot \mathbf{E}\left[\left(\sum_{t=1}^n |\xi_t| \right)^8 \right] \\
&= \frac{2^{11} n^4 \tau_n^8}{\pi^4} + \frac{2^{19} \tau_n^8 (n+7)\cdot(n+6)\cdot \ldots \cdot n}{\pi^4 n^4 \alpha^8} \\
&\lesssim \frac{n^4 \tau_n^8}{1 \wedge \alpha^8}
\end{aligned}
$$

where we have also used that $\sum_{t=1}^n |\xi_t| \sim \Gamma(n, \alpha/(2\tau_n))$ together with the fact that the k-th moment of a $\Gamma(n,\beta)$-distributed random variable is equal to $(n+k-1)\cdot \ldots \cdot n/\beta^k$. Thus,

$$
\mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^4 \lesssim D_{\mathfrak{m}}^2 \cdot \left[\left(\sum_{k\in\mathbb{Z}} |\gamma(k)| + \frac{\tau_n^2}{\alpha^2} \right)^4 + \frac{n^4 \tau_n^8}{1\wedge \alpha^8} \right]. \tag{A.5}
$$

Putting this bound into (A.4), we note that it is sufficient to show that $\mathbf{P}(A^{\mathsf{c}}) \lesssim n^{-6}$ to obtain a bound that is bounded from above by the rate obtained for the term $\mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2 \mathbf{1}_A$ above. We will derive such a bound in the following by means of Assumption 3.1. For n sufficiently large (namely $\tau_n > 2\mu$ has to hold) we have by 3.4

$$
\mathbf{P}(A^{\mathsf{c}}) = \mathbf{P}(\exists i : X_i \neq \widetilde{X}_i)
$$

$$
\begin{aligned}
&\leq \sum_{i=1}^{n} \mathbf{P}(X_i \neq \widetilde{X}_i)\\
&= \sum_{i=1}^{n} \mathbf{P}(|X_i| > \tau_n)\\
&\leq \sum_{i=1}^{n} \mathbf{P}(|X_i - \mu| > \tau_n/2)\\
&\leq 2\sum_{i=1}^{n} e^{-\frac{\tau_n^2}{8\nu}}\\
&\leq 2ne^{-\frac{\tau_n^2}{8\nu}}.
\end{aligned}
$$

With $\tau_n^2 = 56\nu\log(n)$ (our definition), we obtain $\mathbf{P}(A^{\mathsf{c}}) \lesssim n^{-6}$. Combining this estimate with (A.4) and (A.5), we obtain desired bound for $\mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2\mathbf{1}_{A^{\mathsf{c}}}$. Putting the obtained bounds for the terms $\mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2\mathbf{1}_{A}$ and $\mathbf{E}\|\widehat{f}_{\mathfrak{m}} - f_{\mathfrak{m}}\|^2\mathbf{1}_{A^{\mathsf{c}}}$ into the right-hand side of (A.1) yields the claim of the theorem.

A.2. **Proof of Theorem 3.7 (Lower bounds).**

Proof of statement a). First, note that the minimax risk based on the sample $Z_{1:n}$ can be bounded from below by the one based on the sample $X_{1:n}$:

$$
\begin{aligned}
\inf_{\substack{\widetilde{f}\\ \widetilde{f}=\widetilde{f}(Z_{1:n})}} \sup_{f\in\mathcal{F}(\beta,L)} \mathbf{E}\|\widetilde{f} - f\|^2 &= \inf_{\substack{\widetilde{f}\\ \widetilde{f}=\widetilde{f}(Q(X_{1:n}))}} \sup_{f\in\mathcal{F}(\beta,L)} \mathbf{E}\|\widetilde{f} - f\|^2\\
&\geq \inf_{\substack{\widetilde{f}\\ \widetilde{f}=\widetilde{f}(X_{1:n})}} \sup_{f\in\mathcal{F}(\beta,L)} \mathbf{E}\|\widetilde{f} - f\|^2,
\end{aligned}
$$

because the original infimum on the right-hand side is taken over a smaller set of potential estimators.

Put $\zeta = \min\{1/(B\eta), 1/(2\eta), \pi/2\}$. For any $\theta = (\theta_j)_{0\leq j\leq k_n^*} \in \{\pm 1\}^{k_n^*+1}$, we consider the function f^θ defined through

$$
\begin{aligned}
f^\theta &= \frac{2L}{3} + \theta_0\left(\frac{L^2\zeta}{9n}\right)^{1/2} + \left(\frac{L^2\zeta}{9n}\right)^{1/2} \sum_{1\leq|j|\leq k_n^*} \theta_{|j|}\mathbf{e}_j\\
&= \frac{2L}{3} + \left(\frac{L^2\zeta}{9n}\right)^{1/2} \sum_{0\leq|j|\leq k_n^*} \theta_{|j|}\mathbf{e}_j.
\end{aligned}
$$

Let us first check whether the functions f^θ belong to the set $\mathcal{F}(\beta, L)$ of admissible functions for any $\theta \in \{\pm 1\}^{k_n^*+1}$. First, f^θ is a real-valued function since $f_j^\theta = f_{-j}^\theta$ holds for all j and all θ by construction. Second, f^θ is non-negative since

$$
\begin{aligned}
\left\| \left(\frac{L^2\zeta}{9n}\right)^{1/2} \sum_{0\leq|j|\leq k_n^*} \theta_{|j|}\mathbf{e}_j \right\|_\infty &\leq \left(\frac{L^2\zeta}{9n}\right)^{1/2} \sum_{0\leq|j|\leq k_n^*} 1\\
&= \left(\frac{L^2\zeta}{9}\right)^{1/2} \left(\sum_{0\leq|j|\leq k_n^*} \beta_j^{-2}\right)^{1/2} \cdot \left(\sum_{0\leq|j|\leq k_n^*} \frac{\beta_j^2}{n}\right)^{1/2}\\
&\leq \left(\frac{L^2\zeta B}{9}\right)^{1/2} \cdot \left(\beta_{k_n^*}^2 \cdot \frac{2k_n^*+1}{n}\right)^{1/2}
\end{aligned}
$$

$$\leq \left(\frac{L^2 \zeta B \eta}{9} \right)^{1/2}$$
$$\leq \frac{L}{3},$$

and hence we even have $f^\theta \geq L/3 \geq 0$ (the fact that the functions f^θ are uniformly bounded from below will be exploited later).

Third, $\sum_{j \in \mathbb{Z}} |f_j^\theta|^2 \beta_j^2 \leq L^2$ for any $\theta \in \{\pm 1\}^{k_n^* + 1}$ thanks to the estimate

$$\begin{aligned}
\sum_{j \in \mathbb{Z}} |f_j^\theta|^2 \beta_j^2 &= \sum_{0 \leq |j| \leq k_n^*} |f_j^\theta|^2 \beta_j^2 \\
&= \left[\frac{2L}{3} + \theta_0 \left(\frac{L^2 \zeta}{9n} \right)^{1/2} \right]^2 + \frac{L^2 \zeta}{9} \sum_{1 \leq |j| \leq k_n^*} \frac{\beta_j^2}{n} \\
&\leq \frac{8L^2}{9} + \frac{2L^2 \zeta}{9n} + \frac{L^2 \zeta}{9} \cdot \beta_{k_n^*}^2 \cdot \frac{2k_n^*}{n} \\
&\leq \frac{8L^2}{9} + \frac{2L^2 \zeta}{9} \cdot \beta_{k_n^*}^2 \cdot \frac{2k_n^* + 1}{n} \\
&\leq L^2.
\end{aligned}$$

Combining the three derived properties ensures $f^\theta \in \mathcal{F}(\beta, L)$. Denote with $\mathbf{P}_\theta$ the law of the snippet $X_{1:n}$ when $(X_t)_{t \in \mathbb{Z}}$ is a stationary time series with zero mean and spectral density f^θ. Now, let $\widetilde{f}$ be an arbitrary estimator defined in terms of the snippet $X_{1:n}$. Its maximal risk can be bounded from below by reduction to a finite set of hypotheses as follows:

$$\begin{aligned}
\sup_{f \in \mathcal{F}(\beta, L)} \mathbf{E} \|\widetilde{f} - f\|^2 &\geq \sup_{\theta \in \{\pm 1\}^{k_n^*+1}} \mathbf{E}_\theta \|\widetilde{f} - f^\theta\|^2 \\
&\geq \frac{1}{2^{k_n^*+1}} \sum_{\theta \in \{\pm 1\}^{k_n^*+1}} \mathbf{E}_\theta \|\widetilde{f} - f^\theta\|^2 \\
&\geq \frac{1}{2^{k_n^*+1}} \sum_{\theta \in \{\pm 1\}^{k_n^*+1}} \sum_{0 \leq |j| \leq k_n^*} \mathbf{E}_\theta [|\widetilde{f}_j - f_j^\theta|^2] \\
&= \frac{1}{2^{k_n^*+1}} \sum_{0 \leq |j| \leq k_n^*} \sum_{\theta \in \{\pm 1\}^{k_n^*+1}} \frac{1}{2} [\mathbf{E}_\theta |\widetilde{f}_j - f_j^\theta|^2 + \mathbf{E}_{\theta^{|j|}} |\widetilde{f}_j - f_j^{\theta^{|j|}}|^2], \quad \text{(A.6)}
\end{aligned}$$

where for $\theta \in \{\pm 1\}^{k_n^*+1}$ and $j \in [\![-k_n^*, k_n^*]\!]$ the element $\theta^{|j|}$ is defined by $\theta_k^{|j|} = \theta_k$ for $k \neq |j|$ and $\theta_{|j|}^{|j|} = -\theta_{|j|}$ ('flip in the j-th coordinate'). Recall the notion of Hellinger affinity which is defined via $\rho(\mathbf{P}_\theta, \mathbf{P}_{\theta^{|j|}}) = \int \sqrt{\mathrm{d}\mathbf{P}_\theta \mathrm{d}\mathbf{P}_{\theta^{|j|}}}$. For any estimator $\widetilde{f}$, we have

$$\begin{aligned}
\rho(\mathbf{P}_\theta, \mathbf{P}_{\theta^{|j|}}) &\leq \int \frac{|\widetilde{f}_j - f_j^\theta|}{|f_j^\theta - f_j^{\theta^{|j|}}|} \sqrt{\mathrm{d}\mathbf{P}_\theta \mathrm{d}\mathbf{P}_{\theta^{|j|}}} + \int \frac{|\widetilde{f}_j - f_j^{\theta^{|j|}}|}{|f_j^\theta - f_j^{\theta^{|j|}}|} \sqrt{\mathrm{d}\mathbf{P}_\theta \mathrm{d}\mathbf{P}_{\theta^{|j|}}} \\
&\leq \left(\int \frac{|\widetilde{f}_j - f_j^\theta|^2}{|f_j^\theta - f_j^{\theta^{|j|}}|^2} \mathrm{d}\mathbf{P}_\theta \right)^{1/2} + \left(\int \frac{|\widetilde{f}_j - f_j^{\theta^{|j|}}|^2}{|f_j^\theta - f_j^{\theta^{|j|}}|^2} \mathrm{d}\mathbf{P}_\theta^{|j|} \right)^{1/2},
\end{aligned}$$

from which we obtain using the elementary estimate $(a+b)^2 \leq 2a^2 + 2b^2$

$$\frac{1}{2} |f_j^\theta - f_j^{\theta^{|j|}}|^2 \rho^2(\mathbf{P}_\theta, \mathbf{P}_{\theta^{|j|}}) \leq \mathbf{E}_\theta |\widetilde{f}_j - f_j^\theta|^2 + \mathbf{E}_{\theta^{|j|}} |\widetilde{f}_j - f_j^{\theta^{|j|}}|^2. \quad \text{(A.7)}$$

For the squared Hellinger distance between the laws $\mathbf{P}_\theta$ and $\mathbf{P}_{\theta|j|}$ we obtain

$$\begin{aligned} H^2(\mathbf{P}_\theta, \mathbf{P}_{\theta|j|}) &\leq K(\mathbf{P}_\theta, \mathbf{P}_{\theta|j|}) \\ &\leq \left| \mathbf{E}_\theta \log \frac{\mathrm{d}\mathbf{P}_\theta}{\mathrm{d}\mathbf{P}_{\theta|j|}} \right| \\ &\leq \frac{n}{4\pi(\min_\theta \inf_\omega f^\theta(\omega))^2} \cdot \| f^\theta - f^{\theta|j|} \|^2 \\ &\leq \frac{9n}{4\pi L^2} \cdot \left[|f_j^\theta - f_j^{\theta|j|}|^2 + |f_{-j}^\theta - f_{-j}^{\theta|j|}|^2 \right] \end{aligned}$$

by using Equation (2.21) from [Tsy04], Lemma 3.4 from [Ben85], and the fact that $f^\theta \geq L/3$ (the latter was *en passant* established above). Thus, by the very definition of ζ

$$H^2(\mathbf{P}_\theta, \mathbf{P}_{\theta|j|}) \leq \frac{18n}{\pi L^2} \cdot \left(\frac{L^2 \zeta}{9n} \right) \leq 1,$$

and consequently $\rho(\mathbf{P}_\theta, \mathbf{P}_{\theta|j|}) \geq 1/2$. Putting this estimate into (A.7) and combining it with (A.6) yields

$$\sup_{f \in \mathcal{F}(\beta, L)} \mathbf{E} \| \widetilde{f} - f \|^2 \geq \frac{1}{4} \sum_{0 \leq |j| \leq k_n^*} \frac{L^2 \zeta}{9n} = \frac{L^2 \zeta}{36} \cdot \frac{2k_n^* + 1}{n}$$

which is the claim assertion.

Proof of statement b). Set

$$\Psi_{n,\alpha}^2 = \frac{1}{4} \min \left\{ \frac{\pi}{n(e^\alpha - 1)^2}, \frac{L^2}{4} \right\}.$$

Grant to the general reduction principle for the proof of minimax lower bounds (see Chapitre 2.2 in [Tsy04]) it is suffcient to find two candidate functions f^0, f^1 such that

(i) $f^0, f^1 \in \mathcal{F}(\beta, L)$,
(ii) $\| f^0 - f^1 \|_2^2 \gtrsim 4\Psi_{n,\alpha}^2$, and
(iii) $\mathrm{KL}(\mathbf{P}_f^Z, \mathbf{P}_g^Z) \leq C$ for some constant $C < \infty$ depending neither on α nor n.

Then, for any estimator $\widetilde{f}$

$$\begin{aligned} \sup_{f \in \mathcal{F}(\beta, L)} \mathbf{E} \| \widetilde{f} - f \|_2^2 &\geq \Psi_{n,\alpha}^2 \sup_{f \in \mathcal{F}(\beta, L)} \mathbf{P}(\| \widetilde{f} - f \| \geq \Psi_{n,\alpha}) \\ &\geq \Psi_{n,\alpha}^2 \sup_{\theta \in \{0,1\}} \mathbf{P}(\| \widetilde{f} - f^\theta \| \geq \Psi_{n,\alpha}) \\ &\geq \Psi_{n,\alpha}^2 \inf_T \max_{\theta \in \{0,1\}} \mathbf{P}_\theta(T \neq \theta) \end{aligned}$$

where the last infimum runs over all tests T with values in $\{0, 1\}$ and $\mathbf{P}_\theta$ denotes the distribution of Z when the true spectral density is f^θ.

Let us define the functions f^θ for $\theta \in \{0, 1\}$ via

$$\begin{aligned} &f^0 \equiv L, \\ &f^1 \equiv f^0 - \min \left\{ L - \sqrt{\frac{\pi}{n(e^\alpha - 1)^2}}, L/2 \right\} = L - \min \left\{ \sqrt{\frac{\pi}{n(e^\alpha - 1)^2}}, L/2 \right\}, \end{aligned}$$

and we need to verify the conditions (i)–(iii). Condition (i) is trivially satisfied and (ii) follows from the identity

$$\| f^0 - f^1 \|_2^2 = (f_0^0 - f_0^1)^2 = \min \left\{ \frac{\pi}{n(e^\alpha - 1)^2}, \frac{L^2}{4} \right\} = 4\Psi_{n,\alpha}^2.$$

It remains to prove (iii). First note that the fact that both candidate spectral densites f^θ are constant ensures, by Gaussianity, that the random variables $X_1, \ldots, X_n$ are independent. Thus, we can apply Corollary 1 from [DJW18] together with Lemma 3.4 from [Ben85] and the bound $\mathrm{TV}^2 \leq \mathrm{KL}$ (see (2.21) in [Tsy04], for instance) in order to obtain

$$\begin{aligned}
\mathrm{KL}(\mathbf{P}_0^Z, \mathbf{P}_1^Z) &\leq 4(e^\alpha - 1)^2 \sum_{i=1}^n \mathrm{TV}^2(\mathbf{P}_0^{X_i}, \mathbf{P}_1^{X_i})\\
&\leq 4(e^\alpha - 1)^2 \sum_{i=1}^n \mathrm{KL}(\mathbf{P}_0^{X_i}, \mathbf{P}_1^{X_i})\\
&= 4(e^\alpha - 1)^2 \mathrm{KL}(\mathbf{P}_0^{X}, \mathbf{P}_1^{X})\\
&\leq \frac{(e^\alpha - 1)^2 n}{\pi(\min_{\theta=0,1} \inf_\omega f^\theta(\omega))^2} \cdot \|f^0 - f^1\|_2^2\\
&= \frac{4(e^\alpha - 1)^2 n}{\pi L^2} \cdot (f_0 - g_0)^2\\
&= \frac{4(e^\alpha - 1)^2 n}{\pi L^2} \cdot \min\left\{ \sqrt{\frac{\pi}{n(e^\alpha - 1)^2}}, L/2 \right\}^2\\
&\leq \frac{4(e^\alpha - 1)^2 n}{\pi L^2} \cdot \frac{\pi}{n(e^\alpha - 1)^2}\\
&= 4/L^2.
\end{aligned}$$

Now, application of Théorème 2.2., (iii) from [Tsy04] yields the bound

$$\inf_T \max_{\theta \in \{0,1\}} \mathbf{P}_\theta(T \neq \theta) \geq \max\left\{ \frac{1}{4} e^{-4/L^2}, \frac{1 - \sqrt{2}/L}{2} \right\}$$

which finishes the proof.

Appendix B. Proofs of Section 4

We define the event A (and its complement) exactly as in the proof of Theorem 3.2, namely

$$A = \bigcap_{i=1}^n \{X_i = \widetilde{X}_i\},$$

and consider the decomposition

$$\mathbf{E}\|\widetilde{f} - f\|^2 = \mathbf{E}\|\widetilde{f} - f\|^2 \mathbf{1}_A + \mathbf{E}\|\widetilde{f} - f\|^2 \mathbf{1}_{A^c}.$$

Upper bound for $\mathbf{E}\|\widetilde{f} - f\|^2 \mathbf{1}_A$: We can write the contrast as

$$\Upsilon_n(t) = \|t\|^2 - 2\langle \widehat{I}_n, t\rangle = \|t - f\|^2 - 2\langle \widehat{I}_n - f, t\rangle - \|f\|^2.$$

By the definitions of $\widetilde{f}$ and $\widehat{\mathfrak{m}}$combined with the fact that $\widehat{f}_{\mathfrak{m}}$ minimizes the contrast over the space $S_{\mathfrak{m}}$ the estimate

$$\Upsilon_n(\widetilde{f}) + \mathrm{pen}(\widehat{\mathfrak{m}}) \leq \Upsilon_n(f_{\mathfrak{m}}) + \mathrm{pen}(\mathfrak{m})$$

holds for all $\mathfrak{m} \in \mathcal{M}_n$, we obtain

$$\|f - \widetilde{f}\|^2 - 2\langle \widehat{I}_n - f, \widetilde{f}\rangle + \mathrm{pen}(\widehat{\mathfrak{m}}) \leq \|f - f_{\mathfrak{m}}\|^2 - 2\langle \widehat{I}_n - f, f_{\mathfrak{m}}\rangle + \mathrm{pen}(\mathfrak{m}).$$

Then, by elementary algebraic manipulations,

$$\begin{aligned}
\|f - \widetilde{f}\|^2 &\leq \|f - f_{\mathfrak{m}}\|^2 + 2\langle \widehat{I}_n - f, \widetilde{f} - f_{\mathfrak{m}}\rangle + \mathrm{pen}(\mathfrak{m}) - \mathrm{pen}(\widehat{\mathfrak{m}})\\
&\leq \|f - f_{\mathfrak{m}}\|^2 + 2\langle f - \mathbf{E}\widehat{I}_n, f_{\mathfrak{m}} - \widetilde{f}\rangle + 2\langle \widehat{I}_n - \mathbf{E}\widehat{I}_n, \widetilde{f} - f_{\mathfrak{m}}\rangle + \mathrm{pen}(\mathfrak{m}) - \mathrm{pen}(\widehat{\mathfrak{m}}).
\end{aligned}$$

On the event A, we have $Z'_{1:n} = Z_{1:n}$ and $\widehat{I}_n = I_n^Z - \frac{8\tau_n^2}{\alpha^2} = I_n^{Z'} - \frac{8\tau_n^2}{\alpha^2}$. Hence, on A the identity

$$\langle \widehat{I}_n - \mathbf{E}\widehat{I}_n, \widetilde{f} - f_{\mathfrak{m}} \rangle = \langle I_n^{Z'} - \mathbf{E} I_n^{Z'}, \widetilde{f} - f_{\mathfrak{m}} \rangle$$

holds. By definition of $I_n^{Z'}$, we have

$$\begin{aligned}
I_n^{Z'}(\omega) &= \frac{1}{2\pi n} \left| \sum_{t=1}^n (Z'_t - \overline{Z}'_n) e^{-\mathrm{i}t\omega} \right|^2 \\
&= \frac{1}{2\pi n} \left| \sum_{t=1}^n (X_t - \bar{X}_n) e^{-\mathrm{i}t\omega} + \sum_{t=1}^n (\xi_t - \bar{\xi}_n) e^{-\mathrm{i}t\omega} \right|^2 \\
&= \frac{1}{2\pi n} \left| \sum_{t=1}^n (X_t - \bar{X}_n) e^{-\mathrm{i}t\omega} \right|^2 + \frac{1}{2\pi n} \left(\sum_{t=1}^n (X_t - \bar{X}_n) e^{-\mathrm{i}t\omega} \right) \left(\sum_{t=1}^n (\xi_t - \bar{\xi}_n) e^{\mathrm{i}t\omega} \right) \\
&\quad + \frac{1}{2\pi n} \left(\sum_{t=1}^n (X_t - \bar{X}_n) e^{\mathrm{i}t\omega} \right) \left(\sum_{t=1}^n (\xi_t - \bar{\xi}_n) e^{-\mathrm{i}t\omega} \right) + \frac{1}{2\pi n} \left| \sum_{t=1}^n (\xi_t - \bar{\xi}_n) e^{-\mathrm{i}t\omega} \right|^2 \\
&=: I_n^X + \widetilde{I}_n + I_n^\xi
\end{aligned}$$

(as above $\widetilde{I}_n$ is defined as the sum of the two 'mixed' terms). For $\mathfrak{m}, \mathfrak{m}' \in \mathcal{M}_n$, set

$$\begin{aligned}
G_X(\mathfrak{m}, \mathfrak{m}') &= \sup_{u \in \mathcal{B}_{\mathfrak{m},\mathfrak{m}'}} \langle I_n^X - \mathbf{E} I_n^X, u \rangle, \\
G_\xi(\mathfrak{m}, \mathfrak{m}') &= \sup_{u \in \mathcal{B}_{\mathfrak{m},\mathfrak{m}'}} \langle I_n^\xi - \mathbf{E} I_n^\xi, u \rangle, \\
\widetilde{G}(\mathfrak{m}, \mathfrak{m}') &= \sup_{u \in \mathcal{B}_{\mathfrak{m},\mathfrak{m}'}} \langle \widetilde{I}_n, u \rangle,
\end{aligned}$$

where $\mathcal{B}_{\mathfrak{m},\mathfrak{m}'}$ denotes the unit ball in $S_{\mathfrak{m}} + S_{\widehat{\mathfrak{m}}}$, and we write $G_X(\mathfrak{m})$, $G_\xi(\mathfrak{m})$, and $\widetilde{G}(\mathfrak{m})$ when $\mathfrak{m} = \mathfrak{m}'$. We have $G_X(\mathfrak{m}, \mathfrak{m}') \leq G_X(\mathfrak{m}) + G_X(\mathfrak{m}')$, and the same type of bound holds for G_ξ and $\widetilde{G}$. As a consequence, using the estimate $2xy \leq \tau x^2 + \tau^{-1} y^2$ for $\tau = 16$ we have

$$\begin{aligned}
\|f - \widetilde{f}\|^2 \mathbf{1}_A &\leq \left(\|f - f_{\mathfrak{m}}\|^2 + 2\langle f - \mathbf{E} I_n^{Z'}, f_{\mathfrak{m}} - \widetilde{f} \rangle + 2\langle I_n^{Z'} - \mathbf{E} I_n^{Z'}, \widetilde{f} - f_{\mathfrak{m}} \rangle + \mathrm{pen}(\mathfrak{m}) - \mathrm{pen}(\widehat{\mathfrak{m}}) \right) \mathbf{1}_A \\
&= \Big(\|f - f_{\mathfrak{m}}\|^2 + 2\langle f - \mathbf{E} I_n^{Z'}, f_{\mathfrak{m}} - \widetilde{f} \rangle + 2\langle I_n^X - \mathbf{E} I_n^X, \widetilde{f} - f_{\mathfrak{m}} \rangle + 2\langle I_n^\xi - \mathbf{E} I_n^\xi, \widetilde{f} - f_{\mathfrak{m}} \rangle \\
&\qquad + 2\langle \widetilde{I}_n, \widetilde{f} - f_{\mathfrak{m}} \rangle + \mathrm{pen}(\mathfrak{m}) - \mathrm{pen}(\widehat{\mathfrak{m}}) \Big) \mathbf{1}_A \\
&\leq \Big(\|f - f_{\mathfrak{m}}\|^2 + 2\langle f - \mathbf{E} I_n^{Z'}, f_{\mathfrak{m}} - \widetilde{f} \rangle + 2\|\widetilde{f} - f_{\mathfrak{m}}\| G_X(\mathfrak{m}, \widehat{\mathfrak{m}}) + 2\|\widetilde{f} - f_{\mathfrak{m}}\| G_\xi(\mathfrak{m}, \widehat{\mathfrak{m}}) \\
&\qquad + 2\|\widetilde{f} - f_{\mathfrak{m}}\| \widetilde{G}(\mathfrak{m}, \widehat{\mathfrak{m}}) + \mathrm{pen}(\mathfrak{m}) - \mathrm{pen}(\widehat{\mathfrak{m}}) \Big) \mathbf{1}_A \\
&= (\|f - f_{\mathfrak{m}}\|^2 + \tau \|f - \mathbf{E} I_n^{Z'}\|^2 + 4\tau^{-1} \|f_{\mathfrak{m}} - \widetilde{f}\|^2 + \tau G_X^2(\mathfrak{m}, \widehat{\mathfrak{m}}) + \tau G_\xi^2(\mathfrak{m}, \widehat{\mathfrak{m}}) \\
&\qquad + \tau \widetilde{G}^2(\mathfrak{m}, \widehat{\mathfrak{m}}) + \mathrm{pen}(\mathfrak{m}) - \mathrm{pen}(\widehat{\mathfrak{m}})) \mathbf{1}_A \\
&= (\|f - f_{\mathfrak{m}}\|^2 + 16 \|f - \mathbf{E} I_n^{Z'}\|^2 + \frac{1}{4} \|f_{\mathfrak{m}} - \widetilde{f}\|^2 + 32 G_X^2(\widehat{\mathfrak{m}}) + 32 G_\xi^2(\widehat{\mathfrak{m}}) \\
&\qquad + 32 \widetilde{G}^2(\widehat{\mathfrak{m}}) + 32 G_X^2(\mathfrak{m}) + 32 G_\xi^2(\mathfrak{m}) + 32 \widetilde{G}^2(\mathfrak{m}) + \mathrm{pen}(\mathfrak{m}) - \mathrm{pen}(\widehat{\mathfrak{m}})) \mathbf{1}_A \\
&\leq (3\|f - f_{\mathfrak{m}}\|^2/2 + 16 \|f - \mathbf{E} I_n^{Z'}\|^2 + \frac{1}{2} \mathbf{E} \|f - \widetilde{f}\|^2 + 32 G_X^2(\widehat{\mathfrak{m}}) + 32 G_\xi^2(\widehat{\mathfrak{m}}) \\
&\qquad + 32 \widetilde{G}^2(\widehat{\mathfrak{m}}) + 32 G_X^2(\mathfrak{m}) + 32 G_\xi^2(\mathfrak{m}) + 32 \widetilde{G}^2(\mathfrak{m}) + \mathrm{pen}(\mathfrak{m}) - \mathrm{pen}(\widehat{\mathfrak{m}})) \mathbf{1}_A.
\end{aligned}$$

Hence,

$$\|f - \widetilde{f}\|^2 \mathbf{1}_A \leq (3\|f - f_{\mathfrak{m}}\|^2 + 32 \|f - \mathbf{E} I_n^{Z'}\|^2 + 64 G_X^2(\widehat{\mathfrak{m}}) + 64 G_\xi^2(\widehat{\mathfrak{m}})$$

$$+ 64\widetilde{G}^2(\widehat{\mathfrak{m}}) + 64G_X^2(\mathfrak{m}) + 64G_\xi^2(\mathfrak{m}) + 64\widetilde{G}^2(\mathfrak{m}) + 2\,\mathrm{pen}(\mathfrak{m}) - 2\,\mathrm{pen}(\widehat{\mathfrak{m}}))\mathbf{1}_A.$$

If the numerical constant in the definition of the penalty is large enought, we can write $\mathrm{pen}(\mathfrak{m}) = \mathrm{pen}_X(\mathfrak{m}) + \mathrm{pen}_\xi(\mathfrak{m}) + \widetilde{\mathrm{pen}}(\mathfrak{m})$ such that

$$\begin{aligned} \mathrm{pen}_X(\mathfrak{m}) &\geq 32\kappa_X \|f\|_\infty (1 + C_{\bar{r}}^2) \frac{D_\mathfrak{m}(1+L_\mathfrak{m})^2}{n}, \\ \mathrm{pen}_\xi(\mathfrak{m}) &\geq 32\kappa_\xi \frac{\tau_n^4 D_\mathfrak{m}(L_\mathfrak{m}^4 + L_\mathfrak{m} + \log(n))}{n\alpha^4}, \qquad \text{and} \\ \widetilde{\mathrm{pen}}(\mathfrak{m}) &\geq 32\widetilde{\kappa} M^4 (1 + \|f\|_\infty)^2 (L_\mathfrak{m}^4 + L_\mathfrak{m} + \log(n)) \frac{D_\mathfrak{m}}{n} \end{aligned}$$

holds for any model $\mathfrak{m} \in \mathcal{M}_n$. Summing over all potential models and taking expectations implies

$$\begin{aligned} \mathbf{E}\|f - \widetilde{f}\|^2 \mathbf{1}_A \leq 3\|f - f_\mathfrak{m}\|^2 &+ 32\|f - \mathbf{E} I_n^{Z'}\|^2 + 4\,\mathrm{pen}(\mathfrak{m}) \\ &+ 128 \sum_{\mathfrak{m} \in \mathcal{M}_n} \mathbf{E}\left[\left(G_X^2(\mathfrak{m}) - \mathrm{pen}_X(\mathfrak{m})/32\right)_+\right] \\ &+ 128 \sum_{\mathfrak{m} \in \mathcal{M}_n} \mathbf{E}\left[\left(G_\xi^2(\mathfrak{m}) - \mathrm{pen}_\xi(\mathfrak{m})/32\right)_+\right] \\ &+ 128 \sum_{\mathfrak{m} \in \mathcal{M}_n} \mathbf{E}\left[\left(\widetilde{G}^2(\mathfrak{m}) - \widetilde{\mathrm{pen}}(\mathfrak{m})/32\right)_+\right]. \end{aligned}$$

The expectations are bounded by Lemmata C.3, C.5, and C.7, combined with Assumption 4.3 in order to obtain

$$\begin{aligned} \mathbf{E}\|f - \widetilde{f}\|^2 \mathbf{1}_A \leq 3\|f - f_\mathfrak{m}\|^2 &+ 32\|f - \mathbf{E} I_n^{Z'}\|^2 + 4\,\mathrm{pen}(\mathfrak{m}) \\ &+ C(C_{\bar{r}}, \|f\|_\infty) \max\left\{\frac{1}{n}, \frac{\tau_n^4}{n^3\alpha^4}\right\}. \end{aligned}$$

Finally, by Proposition A.1 we get (using the same argument as in the proof of Theorem 3.2)

$$\begin{aligned} \mathbf{E}\|f - \widetilde{f}\|^2 \mathbf{1}_A \lesssim \|f - f_\mathfrak{m}\|^2 &+ \max\left\{\frac{\tau_n^4}{n\alpha^4}, \frac{1}{n}\right\} + \mathrm{pen}(\mathfrak{m}) \\ &+ C(C_{\bar{r}}, \|f\|_\infty) \max\left\{\frac{1}{n}, \frac{\tau_n^4}{n^3\alpha^4}\right\}. \end{aligned}$$

Since, this estimate holds for any fixed model $\mathfrak{m}$, we can take the infimum over all potential models which yields

$$\begin{aligned} \mathbf{E}\|f - \widetilde{f}\|^2 \mathbf{1}_A \lesssim \inf_{\mathfrak{m} \in \mathcal{M}_n} \left[\|f - f_\mathfrak{m}\|^2, \mathrm{pen}(\mathfrak{m})\right] &+ \max\left\{\frac{\tau_n^4}{n\alpha^4}, \frac{1}{n}\right\} \\ &+ C(C_{\bar{r}}, \|f\|_\infty) \max\left\{\frac{1}{n}, \frac{\tau_n^4}{n^3\alpha^4}\right\}. \end{aligned}$$

Upper bound for $\mathbf{E}\|\widetilde{f} - f\|^2 \mathbf{1}_{A^c}$: This term can be bounded exactly as in the upper bound for any fixed model (the only property of the model that we have exploited in that proof was the fact that $D_\mathfrak{m} \leq n$ which holds true also for the randomly selected model $\widehat{\mathfrak{m}}$):

$$\mathbf{E}\|\widehat{f}_\mathfrak{m} - f_\mathfrak{m}\|^2 \mathbf{1}_{A^c} \lesssim \frac{1}{n}.$$

Appendix C. Concentration results for the proof of Theorem 4.4

C.1. **A general chaining argument.** Let $\overline{S}$ be a finite dimensional subspace of $L^2 \cap L^\infty$ spanned by some orthonormal basis $(\varphi_i)_{i\in\mathcal{I}}$. We denote the dimension $|\mathcal{I}|$ of $\overline{S}$ with D, and define the quantity

$$\bar{r}_\varphi = \frac{1}{\sqrt{D}} \sup_{\beta\in\mathbb{R}^D, \beta\neq 0} \frac{\|\sum_{i\in\mathcal{I}} \beta_i\varphi_i\|_\infty}{|\beta|_\infty}.$$

In addition, we define $\bar{r}$ as the infimum of $\bar{r}_\varphi$ taken over all possible orthonormal bases of $\overline{S}$.

Proposition C.1 (Proposition 1 from [BM98])**.** *Let $\overline{S}$ be a D-dimensional linear subspace of $L^2 \cap L^\infty$ with its index $\bar{r}$ defined as above. Let $\mathcal{B}$ be any ball of radius σ in $\overline{S}$ and $0 < \delta < \sigma/5$. Then there exists a finite set $T \subset \mathcal{B}$ which is simultaneously a δ-net for $\mathcal{B}$ with respect to the L^2-norm and an $\bar{r}\delta$-net with respect to the L^∞-norm and such that $|T| \leq (6\sigma/\delta)^D$.*

We will apply Proposition C.1 with $\sigma = 1$ which reduces the choice of δ to $\delta < \frac{1}{5}$.

In the sequel, we will use the following chaining argument. For $0 < \delta_0 < 1/5$ and any $k \in \mathbb{N}$, we set $\delta_k = 2^{-k}\delta_0$ and consider a sequence of δ_k-nets $(T_k)_{k\in\mathbb{N}}$ with $T_k = T_{\delta_k}$. Then, for any $u \in \mathcal{B}_\mathfrak{m}$ ($\mathcal{B}_\mathfrak{m}$ is defined in the proof of Theorem 4.4 as the unit ball in the space $S_\mathfrak{m}$), we are able to find a sequence $(u_k)_{k\geq 0}$ with $u_k \in T_k$ such that $\|u - u_k\|^2 \leq \delta_k^2$ and $\|u - u_k\|_\infty \leq \bar{r}_\mathfrak{m}\delta_k$. Moreover, one can achieve $|T_k| \leq (6/\delta_k)^{D_\mathfrak{m}}$. We have the following decomposition:

$$u = u_0 + \sum_{k=1}^{\infty}(u_k - u_{k-1}). \tag{C.1}$$

From the above properties it follows that $\|u_0\| \leq \delta_0$, $\|u_0\|_\infty \leq \bar{r}_\mathfrak{m}\delta_0$, and, for $k \geq 1$, $\|u_k - u_{k-1}\|^2 \leq 2(\delta_k^2 + \delta_{k-1}^2) = 5\delta_{k-1}^2/2$ and $\|u_k - u_{k-1}\|_\infty \leq 3\bar{r}_\mathfrak{m}\delta_{k-1}/2$. These estimates will be used below without further reference.

Let us finally note that we will work with different definitions of δ_0 below. For the purely Gaussian terms in Subsection C.3 it will turn out convenient to choose $0 < \delta_0 < 1/5$ as a numerical constant independent of n whereas for the analysis of the Laplace term in Subsection C.4 and the mixed term C.5 we will need to choose $\delta_0 \asymp n^{-1}$ in order to get better rates (at the cost of slightly worse logarithmic terms). We put $H_k = \log(|T_k|)$. Then

$$H_k \leq D_\mathfrak{m}\log(6/\delta_k) = D_\mathfrak{m}[\log(6/\delta_0) + k\log 2]$$

which will be used below without further reference.

C.2. **The Toeplitz matrix $T_n(u)$.** In the following three Subsections C.3–C.5 we will consider the following Toeplitz matrix $T_n(u)$ associated with the function u that is given by the entries

$$[T_n(u)]_{j,k} = \int_{-\pi}^{\pi} u(\omega)\mathrm{e}^{\mathrm{i}\omega(j-k)}\mathrm{d}\omega, \qquad 1 \leq j,k \leq n.$$

The matrix $T_n(u)$ is always Hermitian but since we consider only symmetric u, the same holds true for $T_n(u)$ (which is then real-valued).

C.3. **Gaussian terms.**

Proposition C.2. *Put $\Xi_n^X(u) = \langle I_n^X - \mathbf{E}I_n^X, u\rangle$. For any symmetric function u,*

$$\mathbf{P}(\Xi_n^X(u) \geq t) \leq 2\exp\left[-c\min\left(\frac{4\pi^2nt^2}{9\|f\|_\infty^2\|u\|^2}, \frac{2\pi nt}{3\|f\|_\infty\|u\|_\infty}\right)\right].$$

Proof. Denote $X = (X_1, \dots, X_n)^\top$. First, we can write

$$\Xi_n^X(u) = \frac{1}{2\pi n}[(X - \bar{X}_n\vec{1})^\top T_n(u)(X - \bar{X}_n\vec{1}) - \mathbf{E}(X - \bar{X}_n\vec{1})^\top T_n(u)(X - \bar{X}_n\vec{1})].$$

Let H be the hyperplane orthogonal to the linear subspace generated by the vector $\vec{1}$ in $\mathbb{R}^n$. Note that $X - \bar{X}_n\vec{1} = P_H X = P_H \Sigma_X^{1/2} Y$ where $Y \sim \mathcal{N}(\vec{0}, E_n)$ and Σ_X is the covariance matrix of $X_{1:n}$. Now, we the Hanson-Wright inequality (Proposition D.1) with $A = (\Sigma_X^{1/2})^\top P_H^\top T_n(u) P_H \Sigma_X^{1/2}$. Since the Y_i are i.i.d. $\sim \mathcal{N}(0,1)$, we have $\|Y_i\|_{\psi_2} \leq \sqrt{8/3} \leq \sqrt{3} = K$. For the given choice of A, we need to bound the quantities $\|A\|_{\mathrm{HS}}$ and $\|A\|_{\mathrm{op}}$ appearing on the right-hand side of the Hanson-Wright inequality. First,

$$\begin{aligned}
\|A\|_{\mathrm{HS}}^2 &= \mathrm{tr}(A^\top A) = \mathrm{tr}(\Sigma_X^{1/2} P_H^\top T_n(u) P_H \Sigma_X P_H^\top T_n(u) P_H \Sigma_X^{1/2}) \\
&= \mathrm{tr}(P_H \Sigma_X P_H^\top T_n(u) P_H \Sigma_X P_H^\top T_n(u)) \\
&\leq \|f\|_\infty^2 \cdot \mathrm{tr}(T_n(u)^2) \\
&\leq n\|f\|_\infty^2 \|u\|^2,
\end{aligned}$$

where we have used the bound $\mathrm{tr}((AB)^2) \leq \rho(A)^2 \mathrm{tr}(B^2)$, and the fact that $\mathrm{tr}(T_n(u)^2) \leq n\|u\|^2$ from p. 284 in [Com01]. Second,

$$\begin{aligned}
\|A\|_{\mathrm{op}} &= \|\Sigma_X^{1/2} P_H^\top T_n(u) P_H \Sigma_X^{1/2}\|_{\mathrm{op}} \\
&\leq \|\Sigma_X^{1/2}\|_{\mathrm{op}} \cdot \|T_n(u)\|_{\mathrm{op}} \cdot \|\Sigma_X^{1/2}\|_{\mathrm{op}} \\
&= \|\Sigma_X\|_{\mathrm{op}} \cdot \|T_n(u)\|_{\mathrm{op}} \\
&= \rho(\Sigma_X) \cdot \rho(T_n(u)) \\
&\leq \|f\|_\infty \cdot \|u\|_\infty.
\end{aligned}$$

Using these estimates, application of the Hanson-Wright inequality (Proposition D.1) yields

$$\mathbf{P}(\Xi^X(u) \geq t) \leq 2\exp\left[-c\min\left(\frac{4\pi^2 nt^2}{9\|f\|_\infty^2\|u\|^2}, \frac{2\pi nt}{3\|f\|_\infty\|u\|_\infty}\right)\right].$$

□

Lemma C.3. *For any fixed model $\mathfrak{m} \in \mathcal{M}_n$ and a sufficiently large constant $\kappa_X > 0$, we have*

$$\mathbf{E}\left[\left((G^X(\mathfrak{m}))^2 - \kappa_X\|f\|_\infty^2(1 + C_{\bar{r}}^2)\frac{D_\mathfrak{m}(1 + L_\mathfrak{m})^2}{n}\right)_+\right] \lesssim e^{-L_\mathfrak{m} D_\mathfrak{m}} \cdot \frac{C(C_{\bar{r}}, \|f\|_\infty)}{n}.$$

Proof. We consider a sequence $(\eta_k)_{k\geq 0}$ of positive numbers and $\eta \geq \sum_{k\geq 0}\eta_k$ (these quantities will be specified later on). Then, using the decomposition (C.1),

$$\begin{aligned}
\mathbf{P}(\sup_{u\in B_\mathfrak{m}} \Xi_n^X(u) > \eta) &= \mathbf{P}\left[\exists (u_k)_{k\geq 0} \in \prod_{k\geq 0} T_k : \Xi^X(u_0) + \sum_{k\geq 1}\Xi_n^X(u_k - u_{k-1}) > \eta_0 + \sum_{k\geq 1}\eta_k\right] \\
&\leq P_1 + P_2,
\end{aligned}$$

where

$$\begin{aligned}
P_1 &= \sum_{u_0\in T_0} \mathbf{P}(\Xi_n^X(u_0) > \eta_0), \\
P_2 &= \sum_{k\geq 1} \sum_{\substack{u_{k-1}\in T_{k-1} \\ u_k \in T_k}} \mathbf{P}(\Xi_n^X(u_k - u_{k-1}) > \eta_k).
\end{aligned}$$

For any $u_0 \in T_0$, we obtain from Proposition C.2 that

$$\mathbf{P}(\Xi_n^X(u_0) > \eta_0) \leq 2\exp\left(-c\min\left(\frac{4\pi^2 n\eta_0^2}{9\|f\|_\infty^2\delta_0^2}, \frac{2\pi n\eta_0}{3\|f\|_\infty \bar{r}_{\mathfrak{m}}\delta_0}\right)\right),$$

and hence

$$P_1 \leq 2\exp(H_0)\exp\left(-c\min\left(\frac{4\pi^2 n\eta_0^2}{9\|f\|_\infty^2\delta_0^2}, \frac{2\pi n\eta_0}{3\|f\|_\infty \bar{r}_{\mathfrak{m}}\delta_0}\right)\right).$$

For $\lambda > 0$, we consider η_0 such that

$$c\min\left(\frac{n\eta_0^2}{9\|f\|_\infty^2\delta_0^2}, \frac{n\eta_0}{3\|f\|_\infty \bar{r}_{\mathfrak{m}}\delta_0}\right) \geq H_0 + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda,$$

that is,

$$\eta_0 = C\|f\|_\infty\delta_0 \cdot \max\left(\sqrt{\frac{H_0 + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda}{n}}, \frac{\bar{r}_{\mathfrak{m}}(H_0 + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda)}{n}\right).$$

for some sufficiently large constant $C > 0$. For any $k \geq 1$, we get from Proposition C.2 with $u_{k-1} \in T_{k-1}$ and $u_k \in T_k$

$$\mathbf{P}(\Xi_n^X(u_k - u_{k-1}) > \eta_k) \leq 2\exp\left(-c\min\left(\frac{8\pi^2 n\eta_k^2}{45\|f\|_\infty^2\delta_{k-1}^2}, \frac{4\pi n\eta_k}{9\|f\|_\infty \bar{r}_{\mathfrak{m}}\delta_{k-1}}\right)\right).$$

Here, for $\lambda \geq 0$, we choose the η_k such that

$$c\min\left(\frac{8\pi^2 n\eta_k^2}{45\|f\|_\infty^2\delta_{k-1}^2}, \frac{4\pi n\eta_k}{9\|f\|_\infty \bar{r}_{\mathfrak{m}}\delta_{k-1}}\right) \geq H_{k-1} + H_k + kD_{\mathfrak{m}} + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda$$

which in turn is satisfied whenever

$$\eta_k = C\|f\|_\infty\delta_{k-1}\max\left(\sqrt{\frac{H_{k-1} + H_k + kD_{\mathfrak{m}} + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda}{n}}, \frac{\bar{r}_{\mathfrak{m}}(H_{k-1} + H_k + kD_{\mathfrak{m}} + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda)}{n}\right)$$

for some sufficiently large constant $C > 0$. Under this choice of $(\eta_k)_{k\geq 0}$, we obtain for $\eta \geq \sum \eta_k$ (using the assumption that $D_{\mathfrak{m}} \geq 1$)

$$\begin{aligned}\mathbf{P}(\sup_{u\in B_{\mathfrak{m}}} \Xi_n^X(u) > \eta) &\leq P_1 + P_2 \\ &\leq 2\exp(-L_{\mathfrak{m}}D_{\mathfrak{m}} - \lambda) + 2\sum_{k\geq 1}\exp(-kD_{\mathfrak{m}} - L_{\mathfrak{m}}D_{\mathfrak{m}} - \lambda) \\ &= 2\exp(-L_{\mathfrak{m}}D_{\mathfrak{m}} - \lambda)\left[1 + \sum_{k\geq 1} e^{-kD_{\mathfrak{m}}}\right] \\ &\leq 3.2\exp(-L_{\mathfrak{m}}D_{\mathfrak{m}} - \lambda).\end{aligned}$$

We compute a bound for $\sum_{k\geq 0}\eta_k$, and take $0 < \delta_0 < 1/5$ as a purely numerical constant from now on.

$$\begin{aligned}\left(\sum_{k\geq 0}\eta_k\right)^2 &\leq C\|f\|_\infty^2\left(\delta_0\sqrt{\frac{H_0 + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda}{n}} + \sum_{k\geq 1}\delta_{k-1}\sqrt{\frac{H_{k-1} + H_k + kD_{\mathfrak{m}} + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda}{n}}\right. \\ &\quad \left. + \delta_0\frac{\bar{r}_{\mathfrak{m}}(H_0 + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda)}{n} + \sum_{k\geq 1}\delta_{k-1}\frac{\bar{r}_{\mathfrak{m}}(H_{k-1} + H_k + kD_{\mathfrak{m}} + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda)}{n}\right)^2 \\ &\leq C\|f\|_\infty^2\left(\frac{1}{n}(\delta_0 + \sum_{k\geq 1}\delta_{k-1})(\delta_0(H_0 + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda) + \sum_{k\geq 1}\delta_{k-1}(H_{k-1} + H_k + kD_{\mathfrak{m}} + L_{\mathfrak{m}}D_{\mathfrak{m}} + \lambda))\right.\end{aligned}$$

$$+\frac{\bar{r}_{\mathfrak{m}}^2}{n^2}\left(\delta_0(H_0+L_{\mathfrak{m}}D_{\mathfrak{m}}+\lambda)+\sum_{k\geq 1}\delta_{k-1}H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}}D_{\mathfrak{m}}+\lambda\right)^2\Bigg)$$
$$\leq C\|f\|_\infty^2\left[\left(\frac{D_{\mathfrak{m}}+D_{\mathfrak{m}}L_{\mathfrak{m}}+\lambda}{n}\right)+\frac{\bar{r}_{\mathfrak{m}}^2}{n^2}(D_{\mathfrak{m}}^2+D_{\mathfrak{m}}^2L_{\mathfrak{m}}^2+\lambda^2)\right]$$
$$\leq C\|f\|_\infty^2\left[\frac{D_{\mathfrak{m}}(1+L_{\mathfrak{m}})}{n}+\frac{\lambda}{n}+\frac{C_{\bar{r}}^2D_{\mathfrak{m}}(1+L_{\mathfrak{m}}^2)}{n}+\frac{\bar{r}_{\mathfrak{m}}^2\lambda^2}{n^2}\right]$$
$$\leq \kappa_X\|f\|_\infty^2(1+C_{\bar{r}}^2)\frac{D_{\mathfrak{m}}(1+L_{\mathfrak{m}})^2}{n}+2\left[\frac{\lambda}{n}\vee\frac{\bar{r}_{\mathfrak{m}}^2\lambda^2}{n^2}\right]$$

for some numerical constant κ_X. Then,

$$\mathbf{E}\left[\left((G^X(\mathfrak{m}))^2-\kappa_X\|f\|_\infty^2(1+C_{\bar{r}}^2)\frac{D_{\mathfrak{m}}(1+L_{\mathfrak{m}})^2}{n}\right)_+\right]$$
$$=\int_0^\infty \mathbf{P}\left((G^X(\mathfrak{m}))^2>\kappa_X\|f\|_\infty^2(1+C_{\bar{r}}^2)\frac{D_{\mathfrak{m}}(1+L_{\mathfrak{m}})^2}{n}+u\right)\mathrm{d}u$$
$$\leq e^{-L_{\mathfrak{m}}D_{\mathfrak{m}}}\left(\int_{2\kappa_X\|f\|_\infty^2/\bar{r}_{\mathfrak{m}}^2}^\infty e^{-nu/(2\kappa_X\|f\|_\infty^2)}\mathrm{d}u+\int_0^{2\kappa_X\|f\|_\infty^2/\bar{r}_{\mathfrak{m}}^2}e^{-n\sqrt{u}/(2\sqrt{\kappa_X}\bar{r}_{\mathfrak{m}}\|f\|_\infty)}\mathrm{d}u\right)$$
$$\leq e^{-L_{\mathfrak{m}}D_{\mathfrak{m}}}\cdot\frac{2\kappa_X\|f\|_\infty^2}{n}\left(\int_0^\infty e^{-v}\mathrm{d}v+\frac{2\bar{r}_{\mathfrak{m}}}{n}\int_0^\infty e^{-\sqrt{v}}\mathrm{d}v\right)$$
$$\lesssim e^{-L_{\mathfrak{m}}D_{\mathfrak{m}}}\frac{C(C_{\bar{r}},\|f\|_\infty)}{n}$$

which is the claim. □

C.4. **Subexponential terms.**

Proposition C.4. *Let $\Xi_n^\xi(u)=\langle I_n^\xi-\mathbf{E}I_n^\xi,u\rangle$. For any symmetric function u,*

$$\mathbf{P}(\Xi_n^\xi(u)\geq t)\leq 2\exp\left(-\frac{1}{C}\min\left(\frac{\pi^2nt^2\alpha^4}{64\tau_n^4\|u\|^2},\frac{\sqrt{2\pi nt}\alpha}{4\tau_n\|u\|_\infty^{1/2}}\right)\right).$$

Proof. Let H be the hyperplane orthogonal to the space generated by the vector $\vec{1}$ in $\mathbb{R}^n$. Then, for $\xi=(\xi_1,\dots,\xi_n)^\top$, $\xi-\bar{\xi}_n\vec{1}=P_H\xi$. We have

$$\Xi_n^\xi(u)=\frac{1}{2\pi n}[(P_H\xi)^\top T_n(u)P_H\xi-\mathbf{E}(P_H\xi)^\top T_n(u)P_H\xi].$$

We will now use Proposition D.2 from Appendix D which is taken from [GSS19]. More precisely, we would like to apply this result with our ξ_i playing the role of the X_i, with $A=P_H^\top T_n(u)P_H$, and $\beta=1$. We have $\mathbf{E}\xi_i^2=\sigma_i^2=8\tau_n^2/\alpha^2$ for all $i\in[\![1,n]\!]$. Moreover $\|\xi_i\|_{\Psi_1}\leq 4\tau_n/\alpha$ which will play the role of M. The last estimate is easily derived using the fact that $|\xi_i|$ obeys an exponential distribution with parameter $\lambda=\alpha/(2\tau_n)$ and then considering the moment generating function for the exponential distribution. It remains to bound the quantities $\|A\|_{\mathrm{HS}}$ and $\|A\|_{\mathrm{op}}$. First,

$$\begin{aligned}\|A\|_{\mathrm{HS}}^2=\mathrm{tr}(A^\top A)&=\mathrm{tr}(P_H^\top T_n(u)P_HP_H^\top T_n(u)P_H)\\&=\mathrm{tr}((P_HP_H^\top T_n(u))^2)\qquad\text{[cyclic property]}\\&=\rho(P_HP_H^\top)^2\cdot\mathrm{tr}(T_n(u)^2)\qquad[\text{since }\mathrm{tr}((MN)^2)\leq\rho(M)^2\mathrm{tr}(N^2)]\\&\leq\mathrm{tr}(T_n(u)^2).\end{aligned}$$

Using the same argument as on p. 284 in [Com01], we have $\mathrm{tr}(T_n(u)^2)\leq n\|u\|^2$, and hence

$$\|A\|_{\mathrm{HS}}^2\leq n\|u\|^2.$$

Second, for $\|A\|_{\mathrm{op}}$ have the bound

$$\|A\|_{\mathrm{op}} = \rho(A) \le \rho(T_n(u)) \le \|u\|_\infty.$$

Thus, we finally obtain

$$\mathbf{P}(\Xi_n^\xi(u) \ge t) \le 2\exp\left(-\frac{1}{C}\min\left(\frac{\pi^2 n t^2\alpha^4}{64\tau_n^4\|u\|^2}, \frac{\sqrt{2\pi n t}\alpha}{4\tau_n\|u\|_\infty^{1/2}}\right)\right)$$

which is the claim assertion. □

Lemma C.5. *For any fixed model $\mathfrak{m} \in \mathcal{M}_n$ and a sufficiently large constant $\kappa_\xi > 0$, we have*

$$\mathbf{E}\left[\left((G^\xi(\mathfrak{m}))^2 - \kappa_\xi\frac{\tau_n^4 D_\mathfrak{m}(L_\mathfrak{m}^4 + L_\mathfrak{m} + \log(n))}{n\alpha^4}\right)_+\right] \lesssim e^{-L_\mathfrak{m} D_\mathfrak{m}}\frac{C(C_{\bar r})\tau_n^4}{n^3\alpha^4}.$$

Proof. As in the proof of Lemma C.3 we consider

$$\begin{aligned}\mathbf{P}(\sup_{u\in\mathcal{B}_\mathfrak{m}} \Xi_n^\xi(u) > \eta) &= \mathbf{P}\left[\exists(u_k) \in \prod_{k\ge 0} T_k : \Xi_n^\xi(u_0) + \sum_{k\ge 1}\Xi_n^\xi(u_k - u_{k-1}) > \eta_0 + \sum_{k\ge 1}\eta_k\right]\\ &\le P_1 + P_2\end{aligned}$$

with

$$\begin{aligned}P_1 &= \sum_{u_0\in T_0}\mathbf{P}(\Xi_n^\xi(u_0) > \eta_0),\\ P_2 &= \sum_{k\ge 1}\sum_{\substack{u_{k-1}\in T_{k-1}\\ u_k\in T_k}}\mathbf{P}(\Xi_n^\xi(u_k - u_{k-1}) > \eta_k).\end{aligned}$$

Now, for any $u_0 \in T_0$, we obtain from Proposition C.4 that

$$\begin{aligned}\mathbf{P}(\Xi_n^\xi(u_0) > \eta_0) &\le 2\exp\left(-\frac{1}{C}\min\left(\frac{\pi^2 n\eta_0^2\alpha^4}{64\tau_n^4\|u_0\|^2}, \frac{\sqrt{2\pi n\eta_0}\alpha}{4\tau_n\|u_0\|_\infty}\right)\right)\\ &\le 2\exp\left(-\frac{1}{C}\min\left(\frac{\pi^2 n\eta_0^2\alpha^4}{64\tau_n^4\delta_0^2}, \frac{\sqrt{2\pi n\eta_0}\alpha}{4\tau_n\sqrt{\bar r_\mathfrak{m}}\delta_0}\right)\right)\end{aligned}$$

and hence

$$P_1 \le 2\exp(H_0)\exp\left(-\frac{1}{C}\min\left(\frac{\pi^2 n\eta_0^2\alpha^4}{64\tau_n^4\delta_0^2}, \frac{\sqrt{2\pi n\eta_0}\alpha}{4\tau_n\sqrt{\bar r_\mathfrak{m}}\delta_0}\right)\right)$$

We choose η_0 such that

$$\min\left(\frac{\pi^2 n\eta_0^2\alpha^4}{64\tau_n^4\delta_0^2}, \frac{\sqrt{2\pi n\eta_0}\alpha}{4\tau_n\sqrt{\bar r}\delta_0}\right) \ge H_0 + L_\mathfrak{m} D_\mathfrak{m} + \lambda$$

which in turn is satisfied whenever

$$\eta_0 \ge C\cdot\frac{\tau_n^2\delta_0}{\alpha^2}\max\left\{\sqrt{\frac{H_0 + L_\mathfrak{m} D_\mathfrak{m} + \lambda}{n}}, \frac{\bar r_\mathfrak{m}}{n}(H_0 + L_\mathfrak{m} D_\mathfrak{m} + \lambda)^2\right\}$$

for some sufficiently large constant $C > 0$. By Proposition C.4 for any choice of u_{k-1} and u_k

$$\begin{aligned}\mathbf{P}(\Xi_n^\xi(u_k - u_{k-1}) > \eta_k) &\le 2\exp\left(-\frac{1}{C}\min\left(\frac{\pi^2 n\eta_k^2\alpha^4}{64\tau_n^4\|u_k - u_{k-1}\|^2}, \frac{\sqrt{2\pi n\eta_k}\alpha}{4\tau_n\|u_k - u_{k-1}\|_\infty^{1/2}}\right)\right)\\ &\le 2\exp\left(-\frac{1}{C}\min\left(\frac{\pi^2 n\eta_k^2\alpha^4}{160\tau_n^4\delta_{k-1}^2}, \frac{\sqrt{2\pi n\eta_k}\alpha}{4\sqrt{3/2}\tau_n\sqrt{\bar r_\mathfrak{m}}\delta_{k-1}}\right)\right).\end{aligned}$$

Thus,

$$P_2 \leq 2 \sum_{k \geq 1} \exp(H_{k-1}) \exp(H_k) \exp\left(-\frac{1}{C} \min\left(\frac{\pi^2 n \eta_k^2 \alpha^4}{160 \tau_n^4 \delta_{k-1}^2}, \frac{\sqrt{2\pi n \eta_k} \alpha}{4\sqrt{3/2} \tau_n \sqrt{\bar{r}_{\mathfrak{m}}} \delta_{k-1}} \right) \right).$$

Here we choose the η_k such that

$$\min\left(\frac{\pi^2 n \eta_k^2 \alpha^4}{160 \tau_n^4 \delta_{k-1}^2}, \frac{\sqrt{2\pi n \eta_k} \alpha}{4\sqrt{3/2} \tau_n \sqrt{\bar{r}_{\mathfrak{m}}} \delta_{k-1}} \right) \geq H_{k-1} + H_k + kD_{\mathfrak{m}} + L_{\mathfrak{m}} D_{\mathfrak{m}} + \lambda$$

which in turn is satisfied whenever

$$\eta_k \geq C \frac{\tau_n^2 \delta_{k-1}}{\alpha^2} \max\left\{ \sqrt{\frac{H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda}{n}}, \frac{\bar{r}_{\mathfrak{m}}}{n} (H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda)^2 \right\}$$

for some sufficiently large constant $C > 0$. Under this choice of $(\eta_k)_{k \geq 0}$, we obtain for $\eta \geq \sum \eta_k$ (under the assumption that $D_{\mathfrak{m}} \geq 1$)

$$\begin{aligned}
\mathbf{P}(\sup_{u \in \mathcal{B}_{\mathfrak{m}}} \Xi_n^{\xi}(u) > \eta) &\leq P_1 + P_2 \\
&\leq 2 \exp(-L_{\mathfrak{m}} D_{\mathfrak{m}} - \lambda) + 2 \sum_k \exp(-kD_{\mathfrak{m}} - L_{\mathfrak{m}} D_{\mathfrak{m}} - \lambda) \\
&= 2 \exp(-L_{\mathfrak{m}} D_{\mathfrak{m}} - \lambda) \left[1 + \sum_k e^{-kD_{\mathfrak{m}}} \right] \\
&\leq 3.2 \exp(-L_{\mathfrak{m}} D_{\mathfrak{m}} - \lambda).
\end{aligned}$$

Let us now find a bound for $\sum_{k \geq 0} \eta_k$. We have

$$\begin{aligned}
\left(\textstyle\sum_{k \geq 0} \eta_k \right)^2 &\lesssim \Bigg(\frac{\tau_n^2 \delta_0}{\alpha^2} \left[\sqrt{\frac{H_0+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda}{n}} + \frac{\bar{r}_{\mathfrak{m}}}{n} (H_0+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda)^2 \right] \\
&\quad + \frac{\tau_n^2}{\alpha^2} \textstyle\sum_k \delta_{k-1} \left[\sqrt{\frac{H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda}{n}} + \frac{\bar{r}_{\mathfrak{m}} (H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda)^2}{n} \right])^2 \\
&= \frac{\tau_n^4}{\alpha^4} \Bigg[\left(\delta_0 \sqrt{\frac{H_0+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda}{n}} + \textstyle\sum_{k \geq 1} \delta_{k-1} \sqrt{\frac{H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda}{n}} \right) \\
&\quad + \frac{\bar{r}_{\mathfrak{m}}}{n} \left(\delta_0 (H_0+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda)^2 + \textstyle\sum_{k \geq 1} \delta_{k-1} (H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda)^2 \right) \Bigg]^2 \\
&\lesssim \frac{\tau_n^4}{\alpha^4} \left(\delta_0 \sqrt{\frac{H_0+L_{\mathfrak{m}} D_{\mathfrak{m}}+\tau}{n}} + \textstyle\sum_{k \geq 1} \delta_{k-1} \sqrt{\frac{H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda}{n}} \right)^2 \\
&\quad + \frac{\tau_n^4}{\alpha^4} \cdot \frac{\bar{r}_{\mathfrak{m}}^2}{n^2} \cdot \left(\delta_0 (H_0+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda)^2 + \textstyle\sum_{k \geq 1} \delta_{k-1} (H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda)^2 \right)^2 \\
&\lesssim \frac{\tau_n^4}{n\alpha^4} \left(\delta_0 + \textstyle\sum_k \delta_{k-1} \right) \left(\delta_0 (H_0+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda) + \textstyle\sum_{k \geq 1} \delta_{k-1} (H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda) \right) \\
&\quad + \frac{\tau_n^4 \bar{r}_{\mathfrak{m}}^2}{\alpha^4 n^2} \delta_0^2 \left[H_0^2 + (L_{\mathfrak{m}} D_{\mathfrak{m}})^2 + \lambda^2 + \textstyle\sum 2^{-(k-1)} (H_{k-1}^2 + H_k^2 + k^2 D_{\mathfrak{m}}^2 + L_{\mathfrak{m}}^2 D_{\mathfrak{m}}^2 + \lambda^2) \right]^2 \\
&\lesssim \frac{\tau_n^4}{n\alpha^4} \cdot \delta_0^2 (H_0+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda+\textstyle\sum_{k \geq 1} 2^{-(k-1)} (H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda)) \\
&\quad + \frac{\tau_n^4 \bar{r}_{\mathfrak{m}}^2}{\alpha^4 n^2} \delta_0^2 \left[D_{\mathfrak{m}}^2 \log^2(1/\delta_0) + L_{\mathfrak{m}}^2 D_{\mathfrak{m}}^2 + \lambda^2 + L_{\mathfrak{m}}^2 D_{\mathfrak{m}}^2 + \lambda^2 \right]^2 \\
&\lesssim \frac{\tau_n^4}{n\alpha^4} \cdot \delta_0^2 (H_0+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda+D_{\mathfrak{m}} \log(1/\delta_0)+L_{\mathfrak{m}} D_{\mathfrak{m}}+\lambda) \\
&\quad + \frac{\tau_n^4 \bar{r}_{\mathfrak{m}}^2}{n^2 \alpha^4} \delta_0^2 \left[L_{\mathfrak{m}}^2 D_{\mathfrak{m}}^2 + D_{\mathfrak{m}}^2 \log^2(1/\delta_0) + \lambda^2 \right]^2 \\
&\lesssim \frac{\tau_n^4}{n\alpha^4} \cdot \delta_0^2 [L_{\mathfrak{m}} D_{\mathfrak{m}} + D_{\mathfrak{m}} \log(1/\delta_0) + \lambda] \\
&\quad + \frac{\tau_n^4 C_{\bar{r}}^2}{n\alpha^4} \delta_0^2 L_{\mathfrak{m}}^4 D_{\mathfrak{m}}^3 + \frac{\tau_n^4 C_{\bar{r}}^2}{n\alpha^4} \delta_0^2 \log^4(1/\delta_0) D_{\mathfrak{m}}^3 + \frac{\lambda^4 \tau_n^4 \bar{r}_{\mathfrak{m}}^2}{\alpha^4 n^2} \delta_0^2.
\end{aligned}$$

Now, taking $\delta_0 = c/n$ for some numerical constant $0 < c < 1/5$, we obtain (note that we assume $D_{\mathfrak{m}} \le n$ for all $\mathfrak{m} \in \mathcal{M}_n$)

$$\left(\sum_{k\ge 1} \eta_k\right)^2 \le \kappa_\xi \left\{ \frac{\tau_n^4}{n\alpha^4}(L_{\mathfrak{m}} + L_{\mathfrak{m}}^4 + \log^4(n))D_{\mathfrak{m}} + \frac{\tau_n^4}{n^3\alpha^4}\left[\lambda \vee \frac{\lambda^4 \bar{r}_{\mathfrak{m}}^2}{n}\right]\right\}$$

for a sufficiently large constant $\kappa_\xi = \kappa_\xi(C_{\bar{r}})$. Finally,

$$\begin{aligned}
&\mathbf{E}\left[\left((G^\xi(\mathfrak{m}))^2 - \kappa_\xi \frac{\tau_n^4 D_{\mathfrak{m}}(L_{\mathfrak{m}}^4 + L_{\mathfrak{m}} + \log^4(n))}{n\alpha^4}\right)_+\right] \\
&\le \int_0^\infty \mathbf{P}\left((G^\xi(\mathfrak{m}))^2 > \kappa_\xi \frac{\tau_n^4 D_{\mathfrak{m}}(L_{\mathfrak{m}}^4 + L_{\mathfrak{m}} + \log^4(n))}{n\alpha^4} + u\right) \mathrm{d}u \\
&\le e^{-L_{\mathfrak{m}}D_{\mathfrak{m}}}\left(\int_{(n/\bar{r}_{\mathfrak{m}}^2)^{1/3}}^\infty e^{-n\alpha/(\tau_n\sqrt{\bar{r}_{\mathfrak{m}}})\cdot(u/(2\kappa_\xi))^{1/4}}\mathrm{d}u + \int_0^{(n/\bar{r}_{\mathfrak{m}}^2)^{1/3}} e^{-un^3\alpha^4/(2\kappa_\xi\tau_n^4)}\mathrm{d}u\right) \\
&\le e^{-L_{\mathfrak{m}}D_{\mathfrak{m}}} \cdot \left(\frac{2\kappa_\xi\tau_n^4\bar{r}_{\mathfrak{m}}^2}{n^4\alpha^4} + \frac{2\kappa_\xi\tau_n^4}{n^3\alpha^4}\right) \\
&\lesssim e^{-L_{\mathfrak{m}}D_{\mathfrak{m}}}\frac{\tau_n^4}{n^3\alpha^4}
\end{aligned}$$

which finishes the proof. □

C.5. **Mixed terms.**

Proposition C.6. *For any symmetric function u,*

$$\mathbf{P}\left(\tilde{\Xi}_n(u) \ge t\right) \le 2\exp\left(-\frac{1}{C}\min\left(\frac{t^2 n}{2M^4\|u\|^2\cdot\|f\|_\infty}, \left(\frac{nt}{M^2\|u\|_\infty\cdot\|f\|_\infty^{1/2}}\right)^{1/2}\right)\right)$$

where $M = 3 + 4\tau_n/\alpha$.

Proof. In order to deal with the mixed term, we first write

$$\begin{pmatrix} X \\ \xi \end{pmatrix} = \begin{pmatrix} \sqrt{\Sigma_{\mathbf{X}}} & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{E}_n \end{pmatrix}\begin{pmatrix} Y \\ \xi \end{pmatrix}$$

where $Y = (Y_1, \ldots, Y_n)^\top$ is a vector of i.i.d. standard Gaussian random variables. Then, the term of interest can be written as

$$\begin{aligned}
\begin{pmatrix} X^\top & \xi^\top \end{pmatrix}\begin{pmatrix} \mathbf{0}_n & T_n(u) \\ T_n(u) & \mathbf{0}_n \end{pmatrix}\begin{pmatrix} X \\ \xi \end{pmatrix} &= \begin{pmatrix} Y^\top & \xi^\top \end{pmatrix}\begin{pmatrix} \sqrt{\Sigma_{\mathbf{X}}} & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{E}_n \end{pmatrix}\begin{pmatrix} \mathbf{0}_n & T_n(u) \\ T_n(u) & \mathbf{0}_n \end{pmatrix}\begin{pmatrix} \sqrt{\Sigma_{\mathbf{X}}} & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{E}_n \end{pmatrix}\begin{pmatrix} Y \\ \xi \end{pmatrix} \\
&= \begin{pmatrix} Y^\top & \xi^\top \end{pmatrix}\begin{pmatrix} \mathbf{0}_n & \sqrt{\Sigma_{\mathbf{X}}}T_n(u) \\ T_n(u)\sqrt{\Sigma_{\mathbf{X}}} & \mathbf{0}_n \end{pmatrix}\begin{pmatrix} Y \\ \xi \end{pmatrix} \\
&=: \begin{pmatrix} Y^\top & \xi \end{pmatrix} A \begin{pmatrix} Y \\ \xi \end{pmatrix}.
\end{aligned}$$

Since all components of the vector $(Y^\top \xi^\top)$ are independent, and the matrix A is symmetric, we can apply Proposition D.2 again with $\beta = 1$ as in the proof of Proposition C.4. We have $\mathbf{E}Y_i^2 = 1$, $\mathbf{E}\xi_i^2 = 8\tau_n^2/\alpha^2$. As seen above $\|\xi_i\|_{\psi_1} \le 4\tau_n/\alpha$ and moreover $\|Y_i\|_{\psi_1} \le \|1\|_{\psi_2} \cdot \|Y_i\|_{\psi_2} \le (\log 2)^{-1/2} \cdot \sqrt{3} \le 3$. Hence, we can take $M = 3 + 4\tau_n/\alpha$. Application of Proposition [GSS19] yields

$$\mathbf{P}\left(\tilde{\Xi}_n(u) \ge t\right) \le 2\exp\left(-\frac{1}{C}\min\left(\frac{4\pi^2 t^2 n^2}{M^4\|A\|_{\mathrm{HS}}^2}, \left(\frac{2\pi nt}{M^2\|A\|_{\mathrm{op}}}\right)^{1/2}\right)\right),$$

and we have to find appropriate bounds for the quantities $\|A\|_{\mathrm{HS}}$ and $\|A\|_{\mathrm{op}}$. Now, using the estimate H.1.g in Section II.9 from[MOA11], p. 341, we have

$$\begin{aligned}\|A\|_{\mathrm{HS}}^2 &= \operatorname{tr}(A^\top A) = \operatorname{tr}\begin{pmatrix}\sqrt{\Sigma_{\mathbf{X}}}T_n(u)^2\sqrt{\Sigma_{\mathbf{X}}} & \mathbf{0}_n\\ \mathbf{0}_n & T_n(u)\Sigma_{\mathbf{X}}T_n(u)\end{pmatrix}\\ &= \operatorname{tr}(\sqrt{\Sigma_{\mathbf{X}}}T_n(u)^2\sqrt{\Sigma_{\mathbf{X}}}) + \operatorname{tr}(T_n(u)\Sigma_{\mathbf{X}}T_n(u))\\ &= 2\operatorname{tr}(\Sigma_X T_n(u)^2)\\ &\leq 2n\|u\|^2\cdot\|f\|_\infty.\end{aligned}$$

Finally, in order to bound $\|A\|_{\mathrm{op}}$, note that

$$\begin{aligned}\|A\|_{\mathrm{op}} &\leq \|T_n(u)\|_{\mathrm{op}}\cdot\|\sqrt{\Sigma_X}\|_{\mathrm{op}}\\ &\leq \|u\|_\infty\cdot\|f\|_\infty^{1/2}.\end{aligned}$$

□

Lemma C.7. *For any fixed model $\mathfrak{m}\in\mathcal{M}_n$ and a sufficiently large constant $\widetilde{\kappa}>0$, we have*

$$\mathbf{E}\left[\left((\widetilde{G}(\mathfrak{m}))^2 - \widetilde{\kappa}M^4(1+\|f\|_\infty)^2(L_{\mathfrak{m}}^4 + L_{\mathfrak{m}} + \log(n))\frac{D_{\mathfrak{m}}}{n}\right)_+\right] \lesssim e^{-L_{\mathfrak{m}}D_{\mathfrak{m}}}\frac{C(C_{\bar{r}},\|f\|_\infty)M^4}{n^3}$$

where $M = 3+4\tau_n/\alpha$.

Proof. We define P_1 and P_2 in analogy to the definition in the proof of Lemma C.5, and using Proposition C.6 we obtain

$$\begin{aligned}\mathbf{P}(\widetilde{\Xi}_n(u_0)>\eta_0) &\leq 2\exp\left(-\frac{1}{C}\min\left(\frac{2\pi^2 n\eta_0^2}{\|f\|_\infty M^4\|u_0\|^2}, \frac{\sqrt{2\pi n\eta_0}}{M\|f\|_\infty^{1/2}\|u_0\|_\infty}\right)\right)\\ &\leq 2\exp\left(-\frac{1}{C}\min\left(\frac{2\pi^2 n\eta_0^2}{\|f\|_\infty M^4\delta_0^2}, \frac{\sqrt{2\pi n\eta_0}}{M\|f\|_\infty^{1/2}\sqrt{\bar{r}_{\mathfrak{m}}}\delta_0}\right)\right),\end{aligned}$$

and hence

$$P_1 \leq 2\exp(H_0)\exp\left(-\frac{1}{C}\min\left(\frac{2\pi^2 n\eta_0^2}{\|f\|_\infty M^4\delta_0^2}, \frac{2\pi\sqrt{n\eta_0}}{M\|f\|_\infty^{1/2}\sqrt{\bar{r}_{\mathfrak{m}}}\delta_0}\right)\right).$$

We choose η_0 such that

$$\frac{1}{C}\min\left(\frac{n\eta_0^2}{2\|f\|_\infty M^4\delta_0^2}, \frac{\sqrt{n\eta_0}}{M\|f\|_\infty^{1/2}\sqrt{\bar{r}_{\mathfrak{m}}}\delta_0}\right) \geq H_0 + L_{\mathfrak{m}}D_{\mathfrak{m}} + \delta$$

which in turn is satisfied whenever

$$\eta_0 \geq CM^2\delta_0(1+\|f\|_\infty)\max\left\{\sqrt{\frac{H_0+L_{\mathfrak{m}}D_{\mathfrak{m}}+\lambda}{n}}, \frac{\bar{r}_{\mathfrak{m}}(H_0+L_{\mathfrak{m}}D_{\mathfrak{m}}+\lambda)^2}{n}\right\}$$

for some sufficiently large constant $C>0$. From Proposition C.6 we obtain for any choice of u_{k-1} and u_k that

$$\begin{aligned}\mathbf{P}(\widetilde{\Xi}_n(u_k-u_{k-1})>\eta_k) &\leq 2\exp\left(-\frac{1}{C}\min\left(\frac{2\pi^2 n\eta_k^2}{\|f\|_\infty M^4\|u_k-u_{k-1}\|^2}, \frac{\sqrt{2\pi n\eta_k}}{M\|f\|_\infty^{1/2}\|u_k-u_{k-1}\|_\infty^{1/2}}\right)\right)\\ &\leq 2\exp\left(-\frac{1}{C}\min\left(\frac{4\pi^2 n\eta_k^2}{5\|f\|_\infty M^4\delta_{k-1}^2}, \frac{2\sqrt{\pi n\eta_k}}{M\|f\|_\infty^{1/2}\sqrt{3\bar{r}_{\mathfrak{m}}\delta_{k-1}}}\right)\right).\end{aligned}$$

As a consequence,

$$P_2 \le 2\sum_{k\ge 1}\exp(H_{k-1})\exp(H_k)\exp\left(-\frac{1}{C}\min\left(\frac{n\eta_k^2}{5\|f\|_\infty M^4\delta_{k-1}^2}, \frac{2\sqrt{\pi n\eta_k}}{M\|f\|_\infty^{1/2}\sqrt{3\bar{r}_{\mathfrak{m}}}\delta_{k-1}}\right)\right).$$

Here we choose the η_k such that

$$\frac{1}{C}\min\left(\frac{4\pi^2 n\eta_k^2}{5\|f\|_\infty M^4\delta_{k-1}^2}, \frac{2\sqrt{\pi n\eta_k}}{M\|f\|_\infty^{1/2}\sqrt{3\bar{r}_{\mathfrak{m}}}\delta_{k-1}}\right) \ge H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}}D_{\mathfrak{m}}+\lambda,$$

which in turn is satisfied whenever

$$\eta_k \ge CM^2\delta_{k-1}(1+\|f\|_\infty)\max\left\{\sqrt{\frac{H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}}D_{\mathfrak{m}}+\lambda}{n}}, \frac{\bar{r}_{\mathfrak{m}}(H_{k-1}+H_k+kD_{\mathfrak{m}}+L_{\mathfrak{m}}D_{\mathfrak{m}}+\lambda)^2}{n}\right\}.$$

Apart from the dependence of the leading numerical constant on $\|f\|_\infty$ and the different dependence in terms of α (which is hidden in the quantity M), the obtained expressions for η_k, $k\ge 0$ are the same as in the proof of Lemma C.5. Taking $\delta_0 = c/n$ for some numerical constant $0<c<1/5$ again, we obtain

$$\left(\sum_{k\ge 1}\eta_k\right)^2 \le \widetilde{\kappa}M^4(1+\|f\|_\infty)^2\left\{\frac{D_{\mathfrak{m}}}{n}(L_{\mathfrak{m}}+L_{\mathfrak{m}}^4+\log(n)) + +\frac{1}{n^3}\left[\lambda\vee\frac{\lambda^4\bar{r}_{\mathfrak{m}}^2}{n}\right]\right\}.$$

A calculation similar to the one in the proof of Lemma C.5 yields

$$\mathbf{E}\left[\left((\widetilde{G}(\mathfrak{m}))^2 - \widetilde{\kappa}M^4(1+\|f\|_\infty)^2(L_{\mathfrak{m}}+L_{\mathfrak{m}}^4+\log(n))\frac{D_{\mathfrak{m}}}{n}\right)_+\right] \le e^{-L_{\mathfrak{m}}D_{\mathfrak{m}}}\frac{C(C_{\bar{r}},\|f\|_\infty)M^4}{n^3}.$$

□

Appendix D. Auxiliary results

Proposition D.1 (Hanson-Wright inequality, [RV13], Theorem 1.1)**.** *Let $X=(X_1,\dots,X_n)\in\mathbb{R}^n$ be a random vector with independent components X_i which satisfy $\mathbf{E}X_i=0$ and $\|X_i\|_{\psi_2}\le K$. Let A be an $n\times n$-matrix. Then, for every $t\ge 0$,*

$$\mathbf{P}\left(|X^\top AX - \mathbf{E}X^\top AX| > t\right) \le 2\exp\left[-c\min\left(\frac{t^2}{K^4\|A\|_{\mathrm{HS}}^2}, \frac{t}{K^2\|A\|_{\mathrm{op}}}\right)\right].$$

The following result generalizes Proposition D.1 because it can also deal with other exponential Orlicz norms than $\|\cdot\|_{\psi_2}$. This permits to apply the result to subexponential random variables as the Laplace noise used for our anonymization algorithm.

Proposition D.2 ([GSS19], Proposition 1.1)**.** *Let $X_1,\dots,X_n$ be independent random variables satisfying $\mathbf{E}X_i=0$, $\mathbf{E}X_i^2=\sigma_i^2$, $\|X_i\|_{\psi_\beta}\le M$ for some $\beta\in(0,1]\cup\{2\}$, and A be a symmetric $n\times n$ matrix. For any $t>0$ we have*

$$\mathbf{P}\left(|\sum_{i,j}a_{ij}X_iX_j - \sum_{i=1}^n\sigma_i^2 a_{ii}| \ge t\right) \le 2\exp\left(-\frac{1}{C}\min\left(\frac{t^2}{M^4\|A\|_{\mathrm{HS}}^2}, \left(\frac{t}{M^2\|A\|_{\mathrm{op}}}\right)^{\beta/2}\right)\right).$$

Acknowledgements

This research has been supported in part by the research grant DFG DE 502/27-1 of the German Research Foundation (DFG).

References

[BB19] Berrett, T. and Butucea, C. Classification under local differential privacy. *Preprint* (2019). http://arxiv.org/abs/1912.04629.

[BB20] Berrett, T. B. and Butucea, C. Locally private non-asymptotic testing of discrete distributions is faster using interactive mechanisms. *Preprint* (2020). http://arxiv.org/abs/2005.12601.

[BBM99] Barron, A., Birgé, L., and Massart, P. Risk bounds for model selection via penalization. *Probab. Theory Related Fields* 113 (1999), 301–413.

[Ben+91] Bennett, S. et al. A simplified general method for cluster-sample surveys of health in developing countries. *World Health Statistics Quarterly* 44 (1991), 98–106.

[Ben85] Bentkus, R. Y. The rate of uniform convergence of statistical estimates of a spectral density in spaces of differentiable functions. *Litovsk. Mat. Sb.* 25 (1985), 17–31.

[BLM13] Boucheron, S., Lugosi, G., and Massart, P. *Concentration inequalities.* A nonasymptotic theory of independence. Oxford University Press, Oxford, 2013.

[BM97] Birgé, L. and Massart, P. *From model selection to adaptive estimation.* In: *Festschrift for Lucien Le Cam.* Springer, New York, 1997, 55–87.

[BM98] Birgé, L. and Massart, P. Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli* 4 (1998), 329–375.

[BRS20] Butucea, C., Rohde, A., and Steinberger, L. Interactive versus non-interactive locally differentially private estimation: Two elbows for the quadratic functional. *Preprint* (2020). http://arxiv.org/abs/2003.04773.

[But+20] Butucea, C. et al. Local differential privacy: elbow effect in optimal density estimation and adaptation over Besov ellipsoids. *Bernoulli* 26 (2020), 1727–1764.

[Com01] Comte, F. Adaptive estimation of the spectrum of a stationary Gaussian sequence. *Bernoulli* 7 (2001), 267–298.

[Dah89] Dahlhaus, R. Efficient parameter estimation for self-similar processes. *Ann. Statist.* 17 (1989), 1749–1766.

[Dav73] Davies, R. B. Asymptotic inference in stationary Gaussian time-series. *Advances in Appl. Probability* 5 (1973), 469–497.

[DJW18] Duchi, J. C., Jordan, M. I., and Wainwright, M. J. Minimax optimal procedures for locally private estimation. *J. Amer. Statist. Assoc.* 113 (2018), 182–201.

[Dwo06] Dwork, C. *Differential privacy.* In: *Automata, languages and programming. Part II.* Vol. 4052. Lecture Notes in Comput. Sci. Springer, Berlin, 2006, 1–12.

[Efr98] Efromovich, S. Data-driven efficient estimation of the spectral density. *J. Amer. Statist. Assoc.* 93 (1998), 762–769.

[Fan91] Fan, J. On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.* 19 (1991), 1257–1272.

[Fly+13] Flynn, A. et al. A modified random walk door-to-door recruitment strategy for collecting social and biological data relating to mental health, substance use, addiction, and violence problems in a Canadian community. *The International Journal of Alcohol and Drug Research* 2(2) (2013), 7–16.

[FT86] Fox, R. and Taqqu, M. S. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* 14 (1986), 517–532.

[Gol93] Golubev, G. K. Nonparametric estimation of smooth spectral densities of Gaussian stationary sequences. *Teor. Veroyatnost. i Primenen.* 38 (1993), 775–786.

[GSS19] Götze, F., Sambale, H., and Sinulis, A. Concentration inequalities for polynomials in α-sub-exponential random variables. *Preprint* (2019). http://arxiv.org/abs/1903.05964.

[HRW13] Hall, R., Rinaldo, A., and Wasserman, L. Differential privacy for functions and functional data. *Journal of Machine Learning Research* 14 (2013), 703–727.

[Kro19] Kroll, M. Non-parametric Poisson regression from independent and weakly dependent observations by model selection. *J. Statist. Plann. Inference* 199 (2019), 249–270.

[LLL20] Lam-Weil, J., Laurent, B., and Loubes, J.-M. Minimax optimal goodness-of-fit testing for densities under a local differential privacy constraint. *Preprint* (2020). http://arxiv.org/abs/2002.04254.

[LR85] Lemeshow, S. and Robinson, D. Surveys to measure programme coverage and impact: a review of the methodology used by the expanded programme on immunization. *World Health Statistics Quarterly* 38 (1985), 65–75.

[Mas07] Massart, P. *Concentration inequalities and model selection.* Vol. 1896. Lecture Notes in Mathematics. Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003. Springer, Berlin, 2007, pp. xiv+337.

[MOA11] Marshall, A. W., Olkin, I., and Arnold, B. C. *Inequalities: theory of majorization and its applications.* Springer Series in Statistics. Springer, New York, 2011, pp. xxviii+909.

[Neu96] Neumann, M. H. Spectral density estimation via nonlinear wavelet methods for stationary non-Gaussian time series. *J. Time Ser. Anal.* 17 (1996), 601–633.

[RS20] Rohde, A. and Steinberger, L. Geometrizing rates of convergence under local differential privacy constraints. *Ann. Statist.* 48 (2020), 2646–2670.

[RV13] Rudelson, M. and Vershynin, R. Hanson-Wright inequality and sub-Gaussian concentration. *Electron. Commun. Probab.* 18 (2013), paper no. 82, 9.

[Sou00] Soulier, P. Estimation adaptative de la densité spectrale d'un processus gaussien faiblement ou fortement dépendant. *C. R. Acad. Sci. Paris Sér. I Math.* 330 (2000), 733–736.

[Tan87] Taniguchi, M. Minimum contrast estimation for spectral densities of stationary processes. *J. Roy. Statist. Soc. Ser. B* 49 (1987), 315–325.

[Tsy04] Tsybakov, A. B. *Introduction à l'estimation non-paramétrique.* Springer-Verlag, Berlin, 2004, pp. x+175.

Ruhr-Universität Bochum, Fakultät für Mathematik, D-44780 Bochum
E-mail address: martin.kroll-k9x@rub.de