

# Chapter 3

## New Challenges in Covariance Estimation: Multiple Structures and Coarse Quantization



Johannes Maly, Tianyu Yang, Sjoerd Dirksen, Holger Rauhut, and Giuseppe Caire

### 3.1 Introduction

The key objective in covariance estimation is simple to state: given  $n \in \mathbb{N}$  i.i.d. samples  $\mathbf{X}^1, \dots, \mathbf{X}^n \stackrel{d}{\sim} \mathbf{X}$  of a random vector  $\mathbf{X} \in \mathbb{R}^p$ , compute a reliable estimate of the covariance matrix  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] = \mathbf{\Sigma} \in \mathbb{R}^{p \times p}$  (without loss of generality, we restrict ourselves here to mean-zero distributions, i.e.,  $\mathbb{E}[\mathbf{X}] = \mathbf{0}$ ). For this purpose, a natural estimator is the *sample covariance matrix*

$$\hat{\mathbf{\Sigma}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{X}^k (\mathbf{X}^k)^\top \quad (3.1)$$

as it converges to  $\mathbf{\Sigma}$ , for  $n \rightarrow \infty$ , by the law of large numbers. Nevertheless, an asymptotic result is of limited use from practical perspective. Given  $n \in \mathbb{N}$ , it provides no information on the reconstruction error  $\|\hat{\mathbf{\Sigma}}_n - \mathbf{\Sigma}\|$  measured in the operator norm  $\|\cdot\|$ . (Although other norms or error metrics might be considered

---

J. Maly (✉)  
Catholic University of Eichstaett-Ingolstadt, Eichstätt, Germany  
e-mail: [johannes.maly@ku.de](mailto:johannes.maly@ku.de)

T. Yang · G. Caire  
Technical University of Berlin, Berlin, Germany  
e-mail: [tianyu.yang@tu-berlin.de](mailto:tianyu.yang@tu-berlin.de); [caire@tu-berlin.de](mailto:caire@tu-berlin.de)

S. Dirksen  
Utrecht University, Utrecht, Netherlands  
e-mail: [s.dirksen@uu.nl](mailto:s.dirksen@uu.nl)

H. Rauhut  
RWTH Aachen University, Aachen, Germany  
e-mail: [rauhut@mathc.rwth-aachen.de](mailto:rauhut@mathc.rwth-aachen.de)

as well, e.g., the Frobenius norm, we mainly restrict ourselves in this chapter on operator norm bounds as the most common representative.)

In the last two decades, numerous works on non-asymptotic analysis of covariance estimation showed that reliable approximation of  $\Sigma$  by  $\hat{\Sigma}_n$  becomes feasible for sub-Gaussian distributions if  $n \gtrsim p$ , where  $a \lesssim b$  denotes  $a \leq Cb$  for some absolute constant  $C > 0$ . For instance, if  $\mathbf{X}$  has a Gaussian distribution, then it is well known [61] that with probability at least  $1 - 2e^{-t}$

$$\|\hat{\Sigma}_n - \Sigma\| \lesssim \|\Sigma\| \left( \sqrt{\frac{p+t}{n}} + \frac{p+t}{n} \right). \quad (3.2)$$

This classical result exhibits various weaknesses. For instance, it requires strong concentration of the distribution of  $\mathbf{X}$  around its mean. The estimator in (3.1) is sensitive to outliers and not reliable if concentration fails [12, 34]. Furthermore, in applications the ambient dimension can easily exceed the number of accessible samples such that even if concentration may be assumed, the estimate in (3.2) is void.

### 3.1.1 Outline and Notation

In Sect. 3.2, we briefly discuss massive MIMO as one specific modern application of covariance estimation. The massive MIMO setting originates from wireless communications research and will serve as a motivation for investigating multiple structures and quantized samples in a mathematical framework. Section 3.3 then surveys recent theoretical advances on estimation of structured covariance matrices, and Sect. 3.4 shows the impact of coarse sample quantization on estimation guarantees. Having the theoretical results from Sects. 3.3 and 3.4 in mind, in Sect. 3.5, we finally return to the details of massive MIMO and present our recent approach in engineering literature. We conclude in Sect. 3.6 by discussing the gap between existing theoretical guarantees and practical solutions. Some technical details of Sect. 3.3 are deferred to the Appendix.

We denote  $[n] = \{1, \dots, n\}$ . For any absolute constant  $C > 0$ , we abbreviate  $a \leq Cb$  (resp.,  $\geq$ ) as  $a \lesssim b$  (resp.,  $\gtrsim$ ). We furthermore write  $a \lesssim_L b$  (resp.,  $\gtrsim_L$ ) if  $C$  only depends on the quantity  $L$ . Whenever we use absolute constants  $c, C > 0$ , their values may vary from line to line. Scalar-valued functions act component-wise on vectors and matrices. For a set  $S$ , the *indicator function*  $\chi_S$  is 1 on  $S$  and 0 on its complement  $S^c$ . We denote the all ones-matrix by  $\mathbf{1} \in \mathbb{R}^{p \times p}$  and the identity by  $\mathbf{I} \in \mathbb{R}^{p \times p}$ . In particular,

$$[\text{sign}(\mathbf{x})]_i = \begin{cases} 1 & \text{if } x_i \geq 0 \\ -1 & \text{if } x_i < 0, \end{cases}$$

for all  $\mathbf{x} \in \mathbb{R}^p$  and  $i \in [p]$ . For  $\mathbf{Z} \in \mathbb{R}^{p \times p}$ , we denote the *operator norm* (the maximum singular value) by  $\|\mathbf{Z}\| = \sup_{\mathbf{u} \in \mathbb{S}^{p-1}} \|\mathbf{Z}\mathbf{u}\|_2$ , the *nuclear norm* (the sum of singular values) by  $\|\mathbf{Z}\|_* = \text{tr}(\sqrt{\mathbf{Z}^\top \mathbf{Z}})$ , the *Frobenius norm* (trace norm) by  $\|\mathbf{Z}\|_F^2 = \text{tr}(\mathbf{Z}^\top \mathbf{Z}) = \sum_{i,j=1}^p Z_{i,j}^2$ , the *max norm* by  $\|\mathbf{Z}\|_\infty = \max_{i,j} |Z_{i,j}|$ , and the *maximum column norm*  $\|\mathbf{Z}\|_{1 \rightarrow 2} = \max_{j \in [p]} \|\mathbf{z}_j\|_2$ , where  $\mathbf{z}_j$  denotes the  $j$ -th column of  $\mathbf{Z}$ . We use  $\odot$  for the Hadamard (i.e., entry-wise) product of two matrices. The *uniform distribution* on a set  $S$  is denoted by  $\text{Unif}(S)$ . The *multivariate Gaussian distribution* with mean  $\boldsymbol{\mu} \in \mathbb{R}^p$  and covariance matrix  $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$  is denoted by  $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . The *sub-Gaussian* ( $\psi_2$ -) and *subexponential* ( $\psi_1$ -) norms of a random variable  $X$  are defined by

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

A mean-zero random vector  $\mathbf{X}$  on  $\mathbb{R}^n$  is called *K-sub-Gaussian* if

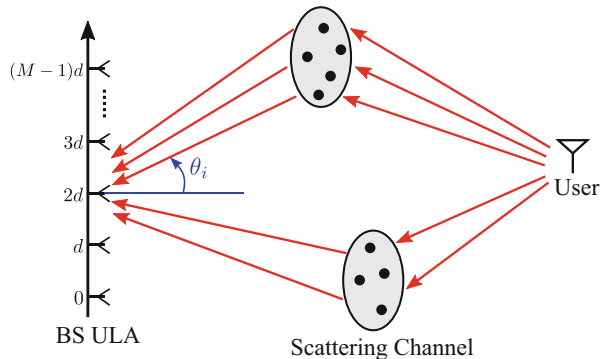
$$\|\langle \mathbf{X}, \mathbf{x} \rangle\|_{\psi_2} \leq K \mathbb{E}[\langle \mathbf{X}, \mathbf{x} \rangle^2]^{1/2} \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

### 3.2 Motivation: Massive MIMO

*Multiple-input multiple-output (MIMO)* is a method in wireless communication to enhance the capacity of a radio link by using multiple transmission and multiple receiving antennas. It has become an essential element of wireless communication standards for Wi-Fi and mobile devices [24, 50]. *Massive MIMO* equips the *base station (BS)* with a large number of antennas to further increase bandwidth and potential number of users [44, 45].

In a classical massive MIMO communication system, the BS is equipped with a *uniform linear array (ULA)* of  $M$  antennas and communicates with multiple users through a scattering channel, e.g., wave reflection on buildings or objects. See Fig. 3.1 for an exemplary setup. During *uplink (UL)*, the BS receives user pilots and aims at estimating the respective channel covariance matrices, which characterize

**Fig. 3.1** An exemplary multipath propagation channel, where the user signal is received at the BS through two scattering clusters



each transmission channel. By assuming mutual orthogonality of all UL pilots, it suffices to focus on a single user channel. We denote the corresponding UL channel vector at time–frequency resource  $s$  by  $\mathbf{h}(s) \in \mathbb{C}^M$  (standard block-fading model, e.g., [59]). Furthermore, we assume that the user transmits a single pilot per channel coherence block such that the channel vectors  $\mathbf{h}(s)$  are i.i.d. complex Gaussian vectors, for  $s \in [N]$  [27, 28].

The received pilot signal at the BS at resource block  $s$  is then given as

$$\mathbf{y}(s) = \mathbf{h}(s)x(s) + \mathbf{z}(s), \quad (3.3)$$

for  $s \in [N]$ , where  $x(s) \in \mathbb{C}$  is the known pilot symbol and  $\mathbf{z}(s) \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}) = \mathcal{N}(\mathbf{0}, \frac{N_0}{2}\mathbf{I}) + j\mathcal{N}(\mathbf{0}, \frac{N_0}{2}\mathbf{I})$  models *additive white Gaussian noise (AWGN)*. Without loss of generality, one may assume that the pilot symbols are normalized, i.e.,  $x(s) = 1$ . The core problem of massive MIMO channel estimation is now to estimate the channel covariance matrix

$$\mathbf{\Sigma}_{\mathbf{h}} = \mathbb{E}[\mathbf{h}(s)\mathbf{h}(s)^H] \quad (3.4)$$

from  $N$  noisy samples  $\mathbf{y}(s)$ ,  $s \in [N]$ . Since the number of samples  $N$  is limited due to time constraints of the UL phase, one expects for massive MIMO that  $N \approx M$ . Translating this into our initial theoretical setting, i.e., identifying the ambient dimension  $p$  with the number of antennas  $M$ , the number of samples  $n$  with the number of independent time–frequency resources  $N$ , and the sample vectors  $\mathbf{X}^k$  with the channel vectors  $\mathbf{h}(s)$ , we see that the sample covariance matrix will not provide a reliable estimate of  $\mathbf{\Sigma}_{\mathbf{h}}$  in this case, cf. Eq. (3.2) for  $n \approx p$ . Nevertheless, a closer look into the channel model reveals that  $\mathbf{\Sigma}_{\mathbf{h}}$  naturally exhibits intrinsic structures such as low-rankness and Toeplitz structure, cf. Sect. 3.5.

**Structure and Quantization** Let us highlight two crucial points. First, whereas engineers are successful in boosting the sample covariance matrix by using special features of their problem setting, cf. Sect. 3.5, it might simplify existing approaches if alternatives to the sample covariance matrix are used that automatically leverage intrinsic structure(s) of the covariance matrix. As Sect. 3.3 will show, the last decade substantially improved our theoretical understanding in this regard. Second, if the above methods are used in real applications, one has to take into account that the sample vectors  $\mathbf{y}(s)$  have to be quantized to finite alphabets before digital processing. Especially, in massive MIMO, the information loss due to quantization can be significant since fine quantization at a multitude of antennas leads to enormous energy consumption. The results presented in Sect. 3.4 can be seen as a first theoretical step into understanding the non-asymptotic behavior of covariance estimators under coarse quantization of the samples. Since we concentrate on memoryless quantization schemes (each vector entry is quantized independently of all others), our model should be applicable to massive MIMO in a straightforward way.

### 3.3 Estimation of Structured Covariance Matrices and Robustness Against Outliers

As we have already seen in Sect. 3.2, there are several structures of interest that  $\Sigma$  might exhibit in applications. We concentrate here on three important instances—sparsity, low-rankness, and Toeplitz structure—that naturally emerge in engineering, biology, and data science, e.g., [42, 53]. Parts of the results we review below are not restricted to Gaussian random vectors but allow to treat heavy-tailed distributions that only satisfy assumptions on their lower moments. Techniques for robust covariance estimation include median of means [31, 49], element- and spectrum-wise truncation [12, 47], and  $M$ -estimators [47, 48]. The recent work [46] even constructs a “sub-Gaussian” estimator that only requires a finite kurtosis assumption ( $L_4$ – $L_2$ -norm equivalence). In this context, an estimator is called sub-Gaussian if it performs on non-Gaussian distributions as well as the sample covariance matrix applied to Gaussian distributions, for further discussion see [46]. Although the proposed construction is computationally intractable, it illustrates the potential of robust estimation. For further information on early and recent approaches to robust covariance estimation, we refer the reader to [29, 34].

#### 3.3.1 Sparse Covariance Matrices

We begin with the assumption that  $\Sigma$  is a *sparse* matrix, i.e., only few entries of  $\Sigma$  are relevant and hence non-zero. If  $\mathbf{X}$  models ordered variables, the non-zero entries of  $\Sigma$ , for instance, might cluster around the diagonal such that  $\Sigma$  is a banded or tapered matrix. A straightforward way to estimate such covariance matrices is to band/taper the sample covariance matrix  $\hat{\Sigma}_n$  [6, 11, 23]. If the variables are not ordered and the non-zero entries of  $\Sigma$  do not cluster, thresholding of  $\hat{\Sigma}_n$  is a viable alternative [5, 19]. As remarked in [40], the aforementioned approaches can be treated in a unified way by introducing a *mask*  $\mathbf{M} \in [0, 1]^{p \times p}$  and considering the *masked sample covariance matrix*  $\mathbf{M} \odot \hat{\Sigma}_n$ . The masked formulation allows to decompose the estimation error

$$\|\mathbf{M} \odot \hat{\Sigma}_n - \Sigma\| \leq \|\mathbf{M} \odot \hat{\Sigma}_n - \mathbf{M} \odot \Sigma\| + \|\mathbf{M} \odot \Sigma - \Sigma\|$$

into a variance term that behaves well if  $\mathbf{M}$  is (close to) sparse and a bias term that is small whenever  $\mathbf{M}$  encodes the support of  $\Sigma$ . The bias term is deterministic and solely depends on a proper choice of  $\mathbf{M}$ . For understanding the influence of sparsity on the required sample size, it thus suffices to control the variance term. The corresponding state-of-the-art result can be found in [13] which extends [40] from Gaussian distributions to general distributions of finite fourth moment and strengthens [40] if applied to Gaussian distributions. To facilitate the comparison with (3.2), we present the result only in the Gaussian case.

**Theorem 3.1** ([13, Theorem 1.1]) *Let  $\mathbf{M} \in [0, 1]^{p \times p}$ , for  $p \geq 3$ , be fixed and  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$  with  $\Sigma \in \mathbb{R}^{p \times p}$ . Then,*

$$\begin{aligned} & \left( \mathbb{E} \|\mathbf{M} \odot \hat{\Sigma}_n - \mathbf{M} \odot \Sigma\|^2 \right)^{\frac{1}{2}} \\ & \lesssim \|\Sigma\| \left( \sqrt{\frac{\|\Sigma\|_\infty}{\|\Sigma\|} \cdot \frac{\|\mathbf{M}\|_{1 \rightarrow 2}^2 \log(p)}{n}} + \frac{\|\Sigma\|_\infty}{\|\Sigma\|} \cdot \frac{\|\mathbf{M}\| \log(p) \log(np)}{n} \right). \end{aligned}$$

Theorem 3.1 only bounds the second moment of the variance term, which yields high-probability estimates via Markov's inequality. However, the same proof techniques apply to higher moments of the variance term as well such that exponential tail bounds can be achieved for Gaussian  $\mathbf{X}$ , cf. [13, Section 3.3].

Let us compare Theorem 3.1 with (3.2). For general covariance estimation, i.e.,  $\mathbf{M} = \mathbf{1}$ , we have  $\|\mathbf{M}\|_{1 \rightarrow 2}^2 = \|\mathbf{M}\| = p$ , which implies that up to log-factors both results are of the same order  $O(\sqrt{\frac{p}{n} + \frac{p}{n}})$ . If  $\mathbf{M}$  encodes sparsity, however, meaning that only up to  $s \ll p$  columns and rows are non-zero and  $\|\mathbf{M}\|_{1 \rightarrow 2}^2 = \|\mathbf{M}\| = s$ , the estimation error is considerably reduced when applying Theorem 3.1. A similar error reduction occurs if  $\mathbf{M} \odot \hat{\Sigma}_n$  is a banded estimator of bandwidth  $B$ .

**Estimation via Thresholding** While the masked framework provides a unified understanding of the intrinsic complexity of sparse covariance estimation, in practice the mask  $\mathbf{M}$  is unknown. A more realistic approach to the problem is hence thresholding procedures as, e.g., [5]. To allow for non-ordered covariance matrices, i.e., general sparsity and not only limited bandwidth of the matrix, the authors of [5] introduce the set of bounded and (effectively) sparse covariance matrices

$$\mathcal{U}(q, s, M) := \left\{ \Sigma : \Sigma_{i,i} \leq M \text{ and } \sum_{j=1}^p |\Sigma_{i,j}|^q \leq s, \text{ for all } i \in [p] \right\},$$

for  $q \in [0, 1)$  and  $s, M > 0$ . If  $q = 0$ , the matrices in  $\mathcal{U}(q, s, M)$  have at most  $s$  non-zero entries per row; if  $q > 0$ , the rows are close to  $s$ -sparse vectors. To estimate  $\Sigma \in \mathcal{U}(q, s, M)$ , the thresholded estimator  $\mathbb{T}_\tau(\hat{\Sigma}_n)$  is considered, where

$$[\mathbb{T}_\tau(\mathbf{A})]_{i,j} = \begin{cases} A_{i,j} & \text{if } |A_{i,j}| \geq \tau, \\ 0 & \text{else,} \end{cases} \quad (3.5)$$

for any  $\tau > 0$  and  $\mathbf{A} \in \mathbb{R}^{p \times p}$ .

**Theorem 3.2** ([5, Theorem 1]) *Let  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ , for  $\Sigma \in \mathcal{U}(q, s, M)$ , and  $M' > 0$  be sufficiently large (depending on  $M$ ). If*

$$\tau = M' \sqrt{\frac{\log(p)}{n}},$$

for  $n \gtrsim \log(p)$ , then with probability at least  $1 - e^{-cn\tau^2}$

$$\|\mathbb{T}_\tau(\hat{\Sigma}_n) - \Sigma\| = O\left(s \left(\frac{\log(p)}{n}\right)^{\frac{1-q}{2}}\right).$$

Theorem 3.2 does not require knowledge on the support of  $\Sigma$  and respects sparsity defects. However, if we once more consider the case  $q = 0$ , we see that the estimate in Theorem 3.2 is suboptimal since the error behaves (up to log-factors) like  $O(\sqrt{\frac{s^2}{n}})$  and not like  $O(\sqrt{\frac{s}{n}})$  as one would expect.

### 3.3.2 Low-Rank Covariance Matrices

When working with high-dimensional random vectors, another commonly considered structural prior is to assume that the distribution concentrates around a low-dimensional manifold. This may manifest itself in  $\Sigma$  being a low-rank matrix. Interestingly, the sample covariance matrix in (3.1) intrinsically leverages low-rankness of  $\Sigma$ . To understand this phenomenon, we consider the *effective rank* of  $\Sigma$  defined as

$$\mathbf{r}(\Sigma) = \frac{\|\Sigma\|_*}{\|\Sigma\|}.$$

It is straightforward to verify that  $1 \leq \mathbf{r}(\Sigma) \leq \text{rank}(\Sigma)$ . In contrast to the rank of  $\Sigma$ , the quantity  $\mathbf{r}(\Sigma)$  is small even if  $\Sigma$  is only close to a low-rank matrix, e.g., consider  $\Sigma$  to be a full rank matrix with exponentially decaying spectrum.

**Theorem 3.3 ([37, Corollary 2])** *Let  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ , for  $\Sigma \in \mathbb{R}^{p \times p}$ , and  $n \gtrsim \mathbf{r}(\Sigma)$ . Then with probability at least  $1 - e^{-t}$  the sample covariance matrix satisfies*

$$\|\hat{\Sigma}_n - \Sigma\| \lesssim \|\Sigma\| \left( \sqrt{\frac{\mathbf{r}(\Sigma)}{n}} + \frac{\mathbf{r}(\Sigma)}{n} + \sqrt{\frac{t}{n}} + \frac{t}{n} \right).$$

The authors of [37] further show that the bound in Theorem 3.3 is tight up to constants. If we compare the result to (3.2), we see that both estimates agree for (effectively) full rank matrices like  $\Sigma = \mathbf{I}$ . If  $\Sigma$  is of low rank, however, Theorem 3.3 controls the estimation error even in the case  $n < p$ .

**Low-Rank Estimators** We could stop at this point since  $\hat{\Sigma}_n$  apparently meets our requirements. Nevertheless, two questions remain. First, if one assumes  $\Sigma$  to be

low rank, one would wish for an estimator that is low rank itself, and, second, Theorem 3.3 fails if  $\mathbf{X}$  does not exhibit strong concentration around its mean. The first point can be addressed by using the LASSO estimator

$$\hat{\Sigma}_n^\lambda = \arg \min_{\mathbf{S} \succeq \mathbf{0}} \|\mathbf{S} - \hat{\Sigma}_n\|_F^2 + \lambda \|\mathbf{S}\|_* , \quad (3.6)$$

where  $\lambda > 0$  is a tunable parameter. Initially introduced in [43] to estimate covariance matrices from incomplete observations, the result reads in our setting as follows.

**Theorem 3.4 ([43, Corollary 1])** *Let  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ , for  $\Sigma \in \mathbb{R}^{p \times p}$ , and  $n \gtrsim \mathbf{r}(\Sigma) \log(2p + n)^2$ . If*

$$\lambda = C \sqrt{\text{tr}(\hat{\Sigma}_n) \|\hat{\Sigma}_n\|} \sqrt{\frac{\log(2p)}{n}},$$

*for a sufficiently large absolute constant  $C > 0$ , then with probability at least  $1 - \frac{1}{2p}$  the estimator in (3.6) satisfies*

$$\|\hat{\Sigma}_n^\lambda - \Sigma\| \lesssim \|\Sigma\| \sqrt{\frac{\mathbf{r}(\Sigma) \log(2p)}{n}}.$$

The nuclear norm regularization in (3.6) induces (effective) low-rankness on  $\hat{\Sigma}_n^\lambda$  [21, 51] and the order of estimation error reflects up to log-factors the one in Theorem 3.3. Furthermore, the construction of  $\hat{\Sigma}_n^\lambda$  can easily be adapted to heavy-tailed distributions by replacing  $\hat{\Sigma}_n$  with an appropriate robust counterpart, e.g., the spectrum-wise truncated sample covariance matrix [34]. A corresponding version of Theorem 3.4 that is not restricted to (sub)-Gaussian distributions is [34, Theorem 5.2].

### 3.3.3 Toeplitz Covariance Matrices and Combined Structures

The third structure we discuss here in detail naturally arises in various engineering problems. If the entries of  $\mathbf{X}$  resemble measurements on a temporal or spatial grid whose covariances only depend on the distances of measurements (in time or space) but not their location,  $\Sigma$  is a *symmetric Toeplitz matrix*, i.e.,



$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1 & \sigma_2 & \cdots & \sigma_p \\ \sigma_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sigma_2 \\ \sigma_p & \cdots & \sigma_2 & \sigma_1 \end{pmatrix},$$

and the first column  $\boldsymbol{\sigma} \in \mathbb{R}^p$  determines  $\boldsymbol{\Sigma}$  via  $\Sigma_{i,j} = \sigma_{|i-j|+1}$ . (For simplicity, we identify Toeplitz matrices with their first column in the following.) Such a structure appears, for instance, in Direction-Of-Arrival (DOA) estimation [38] and medical/radar imaging processing [9, 56]. For further examples, we refer the reader to [53]. Since Toeplitz structure reduces the degrees of freedom in  $\boldsymbol{\Sigma}$  from  $p^2$  to  $p$ , leveraging this structure can lead to a notable reduction in sample complexity.

The authors of [10] propose to average, the sample covariance matrix along its diagonals to obtain the Toeplitz estimator  $\hat{\boldsymbol{\Sigma}}_n^{\text{Toep}}$  defined as

$$[\hat{\boldsymbol{\sigma}}_n^{\text{Toep}}]_r = \frac{1}{(p+1)-r} \sum_{i-j=r-1} [\hat{\boldsymbol{\Sigma}}_n]_{i,j}, \quad \text{for } r \in [p]. \quad (3.7)$$

They derive error estimates for Gaussian distributions with banded Toeplitz covariance matrices.

The more recent work [33] extends these results to non-Gaussian distributions and general masks as introduced in Sect. 3.3.1. To be more precise, the authors of [33] assume that the distribution of  $\mathbf{X}$  has the so-called convex concentration property.

**Definition 3.1** A random vector  $\mathbf{X} \in \mathbb{R}^p$  has the *convex concentration property* with constant  $K$  if for any 1-Lipschitz function  $\phi: \mathbb{R}^p \rightarrow \mathbb{R}$ , one has  $\mathbb{E}[\phi(\mathbf{X})] < \infty$  and

$$\Pr[|\phi(\mathbf{X}) - \mathbb{E}[\phi(\mathbf{X})]| \geq t] \leq 2e^{-\frac{t^2}{K^2}}, \quad \text{for all } t > 0.$$

By setting  $\phi(\cdot) = \langle \cdot, \mathbf{x} \rangle$ , for  $\mathbf{x} \in \mathbb{R}^p$ , one easily sees that all distributions that have the convex concentration property are sub-Gaussian. For the sake of consistency, we therefore restrict ourselves here to Gaussian distributions as their most prominent representative. For a symmetric Toeplitz mask  $\mathbf{M} \in [0, 1]^{p \times p}$  characterized by its first column  $\mathbf{m} \in [0, 1]^p$ , we furthermore define the weighted  $\ell_1$ - and  $\ell_2$ -norms of  $\mathbf{m}$  as

$$\|\mathbf{m}\|_{1,*} = \sum_{r=1}^p \frac{m_r}{(p+1)-r} \quad \text{and} \quad \|\mathbf{m}\|_{2,*} = \left( \sum_{r=1}^p \frac{m_r^2}{(p+1)-r} \right)^{\frac{1}{2}}.$$

**Theorem 3.5** ([33, Theorem 3]) *Let  $\mathbf{M} \in [0, 1]^{p \times p}$  be a symmetric Toeplitz mask and  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$  with  $\Sigma \in \mathbb{R}^{p \times p}$  symmetric and Toeplitz. Then,*

$$\mathbb{E} \|\mathbf{M} \odot \hat{\Sigma}_n^{\text{Toep}} - \mathbf{M} \odot \Sigma\| \lesssim \|\Sigma\| \left( \sqrt{\frac{\|\mathbf{m}\|_{2,*} \log(p)}{n}} + \frac{\|\mathbf{m}\|_{1,*} \log(p)}{n} \right).$$

As Theorem 3.1, the result is not restricted to an estimate of the expected error but includes respective high probability bounds with exponential tail decay. Let us compare Theorem 3.5 to Theorem 3.1. If we ignore log-factors and assume that  $\mathbf{M}$  is a banding or tapering mask with support bandwidth  $B \leq \frac{p}{2}$ , i.e., only the  $B$  innermost diagonals of  $\mathbf{M}$  are non-zero, Theorem 3.5 guarantees an estimation error of order  $O(\sqrt{\frac{B}{pn}} + \frac{B}{pn})$ , cf. [33, Corollary 2], which improves the estimate  $O(\sqrt{\frac{B}{n}} + \frac{B}{n})$  of Theorem 3.1 by a factor  $p$ . This improvement corresponds to the reduction in degrees of freedom when comparing Toeplitz to general matrices. Note, however, that the additional assumption  $B \leq \alpha p$ , for  $\alpha \in (0, 1)$ , is required for such a reduction since estimation of the outermost diagonals of  $\Sigma$  is hardly enhanced by averaging over the Toeplitz structure. This is expressed by Theorem 3.5 since  $\|\mathbf{m}\|_{1,*}$  and  $\|\mathbf{m}\|_{2,*}$  are  $O(1)$  and not  $O(\frac{1}{p})$  if the tail entries of  $\mathbf{m}$  are not of vanishing magnitude.

**Estimation via Thresholding** Theorem 3.5 differs from the previously discussed results in the sense that it allows to simultaneously leverage two structures of  $\Sigma$ , sparsity and Toeplitz structure. Nevertheless, as in Sect. 3.3.1, the masked framework leaves open the question of how to choose  $\mathbf{M}$  in practice. By combining the thresholded approach in Theorem 3.2 with the techniques of Theorem 3.5, one can obtain a thresholded Toeplitz estimator which profits from both structural priors. To state a corresponding estimate, let us define the set of bounded Toeplitz covariance matrices with (effectively) sparse first column  $\sigma$  by

$$\mathcal{U}^{\text{Toep}}(q, s, M) := \left\{ \Sigma : \Sigma_{i,j} = \sigma_{|i-j|+1} \leq M, \text{ for } \sigma \in \mathbb{R}^p \text{ with } \sum_{r=1}^p |\sigma_r|^q \leq s \right\}.$$

We furthermore denote by  $\mathbb{B}_{\alpha p}(\Sigma)$  the matrix  $\Sigma$  restricted to bandwidth  $\alpha p$ , i.e.,  $[\mathbb{B}_{\alpha p}(\Sigma)]_{i,j} = \Sigma_{i,j}$  if  $|i-j|+1 \leq \alpha p$  and  $[\mathbb{B}_{\alpha p}(\Sigma)]_{i,j} = 0$  else.

**Theorem 3.6** *There exists an absolute constant  $C > 0$  such that the following holds. Let  $\mathbf{X}$  have the convex concentration property with constant  $K$ . Let  $\mathbb{E}[\mathbf{X}] = \mathbf{0}$  and  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] = \Sigma$ , for  $\Sigma \in \mathcal{U}^{\text{Toep}}(q, s, M)$ . For all  $\alpha \in (0, 1)$  and  $c > 1$ , we have with probability at least  $1 - (2\alpha p)^{-(c-1)}$  that if*

$$\tau = \sqrt{\frac{2c}{(1-\alpha)}} \max\{CK^2, \sqrt{CK}\} \sqrt{\frac{\log(p)}{np}}, \quad (3.8)$$

then

$$\left\| \mathbb{T}_\tau(\mathbb{B}_{\alpha p}(\hat{\Sigma}_n^{\text{Toep}})) - \Sigma \right\| \lesssim s \left( \max\{C^2 K^4, C K^2\} \frac{c}{1-\alpha} \frac{\log(p)}{np} \right)^{\frac{1-q}{2}} + \|\mathbb{B}_{\alpha p}(\Sigma) - \Sigma\|,$$

where  $\mathbb{T}_\tau$  is the thresholding operator from (3.5).

Two comments are in order here. To gain from the Toeplitz structure, Theorem 3.6 requires  $\Sigma$  to be close to a banded matrix. This is as in Theorem 3.5 before and has been discussed previously. Moreover, by adapting the proof strategy of Theorem 3.2, the result inherits the slightly suboptimal error decay in the sparsity level  $s$ , cf. the discussion of Theorem 3.2 for the case  $q = 0$ . The proof, which combines ideas from [5] and [33], can be found in the Appendix.

**Combining Toeplitz Structure and Low-Rankness** Sparsity is not the only structure that can be imposed on Toeplitz matrices. For instance, in massive MIMO, see Sect. 3.2, low-rankness of  $\Sigma$  may naturally be assumed in addition to Toeplitz structure [28]. The recent works [20, 39] propose several algorithms to estimate low-rank Toeplitz covariance matrices from partial observations by a technique called “sparse ruler.” In particular, the authors can show that the sufficient number of samples to approximate  $\Sigma$  scales (up to log-factors) polynomial in the (effective) rank of  $\Sigma$ .

*Remark 3.1* Before closing this section, let us briefly comment on the three types of structures discussed above and their mutual relation:

*Sparsity:* The concept of sparsity is the maybe most fundamental way of theoretically describing intrinsic “low-complexity” of points in a vector space. Whereas we only introduced sparsity of vectors in  $\mathbb{R}^n$  with respect to the canonical basis, it is straightforward to generalize the definition to arbitrary vector spaces and other bases (or even frames). Note, however, that sparsity strongly depends on the chosen representation of objects in space, i.e., a point that is sparse in one basis need not be sparse in another.

*Low-rankness:* One can view low-rankness as a special case of sparsity since a matrix is low rank if and only if the vector of its singular values is sparse in the canonical basis. Stated differently, a matrix is low rank if its induced linear mapping only acts on low-dimensional subspaces of the ambient input and output space. This second characterization shows that, in contrast to sparsity, low-rankness is not representation dependent. Furthermore, one can generalize the concept to higher dimensional linear operators as well, e.g., tensor spaces.

*Toeplitz structure:* Just as low-rankness, Toeplitz structure is a special type of sparsity that requires matrix structure of the points in space. Its low-dimensional structure lies in the fact that only  $(2n - 1)$  parameters are necessary to characterize an  $\mathbb{R}^{n \times n}$  Toeplitz matrix. In contrast to low-rankness, Toeplitz structure is representation dependent. Nevertheless, Toeplitz matrices naturally appear as covariance matrices of stationary random processes, i.e., if the covariance of two

events does not depend on their localization in time but only their distance in time.

For further discussion and literature on the subject, we refer the interested reader to [22].

### 3.4 Estimation from Quantized Samples

All results stated above assume that the sample vectors  $\mathbf{X}^k$  are real-valued, i.e., one has access to infinite precision representations of the samples. In applications, this assumption is not always fulfilled. Especially in signal processing, samples are collected via sensors and, hence, need to be quantized to finitely many bits before they can be digitally transmitted and processed. Engineers have been examining the influence of coarse quantization on correlation and covariance estimation for decades, e.g., [2, 14, 30, 41, 54]. However, in contrast to classical covariance estimation from unquantized samples, so far only asymptotic estimation guarantees have been derived in the quantized setting. To improve our understanding on the effect of quantization on covariance estimation, we analyzed two memoryless one-bit quantization schemes in our recent work [16]. We call a quantizer memoryless if it quantizes each entry of  $\mathbf{X}^k$  independently of all remaining entries. This is fundamentally different from feedback systems, e.g.,  $\Sigma\Delta$ -quantization [4, 55], and of particular interest for large-scale applications like massive MIMO where the entries of  $\mathbf{X}^k$  correspond to inputs from different antennas, cf. Sect. 3.2. We conclude by providing a detailed discussion of the models and results in [16].

#### 3.4.1 Sign Quantization

In the first setting, we assume to receive one-bit quantized samples

$$\text{sign}(\mathbf{X}^k) \in \{-1, 1\}^p, \quad (3.9)$$

for  $k \in [n]$ , instead of  $\mathbf{X}^k$  itself. (Recall that we apply scalar functions like  $\text{sign}$  entry-wise to vectors and matrices.) Since the quantizer  $\text{sign}$  is scale-invariant, i.e.,  $\text{sign}(\mathbf{z}) = \text{sign}(\mathbf{D}\mathbf{z})$  for any diagonal matrix  $\mathbf{D} \in \mathbb{R}^{p \times p}$  with strictly positive entries and  $\mathbf{z} \in \mathbb{R}^p$ , we can only hope to recover the correlation matrix of the distribution, i.e., a normalized version of  $\Sigma$  with entries  $\left[ \frac{\Sigma_{i,j}}{\sqrt{\Sigma_{i,i}}\sqrt{\Sigma_{j,j}}} \right]_{i,j}$ . We thus assume that  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ , where  $\Sigma$  has ones on its diagonal.

It has been known for decades that

$$\tilde{\Sigma}_n = \sin \left( \frac{\pi}{2n} \sum_{k=1}^n \text{sign}(\mathbf{X}^k) \text{sign}(\mathbf{X}^k)^\top \right) \quad (3.10)$$

is well suited to approximate  $\Sigma$  from the quantized samples, cf. [30]. Note that the specific form of  $\tilde{\Sigma}_n$  is motivated by Grothendieck's identity (see, e.g., [61, Lemma 3.6.6]), also known as ‘‘arcsin-law’’ in the engineering literature [30, 60], which implies that

$$\Gamma := \mathbb{E}[\text{sign}(\mathbf{X}^k) \text{sign}(\mathbf{X}^k)^\top] = \frac{2}{\pi} \arcsin(\Sigma) \quad (3.11)$$

if  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ . Applying the strong law of large numbers and the continuity of the sine function to (3.10), one easily obtains with (3.11) that  $\tilde{\Sigma}_n$  is a consistent estimator of  $\Sigma$ .

The two key quantities for understanding the non-asymptotic performance of  $\tilde{\Sigma}_n$  are  $\Gamma$  and

$$\mathbf{A} := \cos(\arcsin(\Sigma)) = \cos\left(\frac{\pi}{2}\Gamma\right).$$

Furthermore, we define

$$\sigma(\mathbf{Z})^2 := \mathbf{Z}^2 \odot \Gamma - (\mathbf{Z} \odot \Gamma)^2 = \frac{2}{\pi} \mathbf{Z}^2 \odot \arcsin(\Sigma) - \frac{4}{\pi^2} (\mathbf{Z} \odot \arcsin(\Sigma))^2,$$

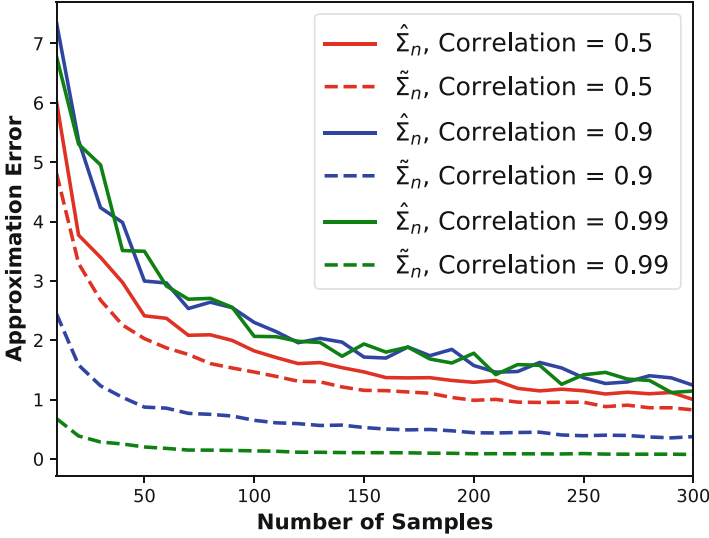
for symmetric  $\mathbf{Z} \in \mathbb{R}^{p \times p}$ .

**Theorem 3.7** ([16, Theorem 1]) *There exist constants  $c_1, c_2 > 0$  such that the following holds. Let  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$  with  $\Sigma_{i,i} = 1$ , for  $i \in [p]$ , and  $\mathbf{X}^1, \dots, \mathbf{X}^n \stackrel{\text{i.i.d.}}{\sim} \mathbf{X}$  be i.i.d. samples of  $\mathbf{X}$ . Let  $\mathbf{M} \in [0, 1]^{p \times p}$  be a fixed symmetric mask. Then, for all  $t \geq 0$  with  $n \geq c_1 \log^2(p)(\log(p) + t)$ , the biased sign estimator  $\tilde{\Sigma}_n$  fulfills with probability at least  $1 - 2e^{-c_2 t}$*

$$\begin{aligned} \|\mathbf{M} \odot \tilde{\Sigma}_n - \mathbf{M} \odot \Sigma\| &\lesssim \|\sigma(\mathbf{M} \odot \mathbf{A})\| \sqrt{\frac{\log(p) + t}{n}} \\ &\quad + (\max\{\|\mathbf{M} \odot \mathbf{A}\|, \|\mathbf{M} \odot \Sigma\|\}) \frac{\log(p) + t}{n}. \end{aligned} \quad (3.12)$$

The right-hand side in Theorem 3.7 (for convenience, we only consider the case  $\mathbf{M} = \mathbf{1}$  here) can be trivially estimated to get

$$\|\tilde{\Sigma}_n - \Sigma\| \lesssim \max\{\|\cos(\arcsin(\Sigma))\|, \|\Sigma\|\} \left( \sqrt{\frac{\log(p) + t}{n}} + \frac{\log(p) + t}{n} \right),$$



**Fig. 3.2** The experiment from [16] depicts average estimation error of  $\hat{\Sigma}_n$  and  $\tilde{\Sigma}_n$  in operator norm, for  $p = 20$ ,  $n$  varying from 10 to 300 and three different choices of the ground truth  $\Sigma$  with ones on the diagonal and off-diagonal entries equal to  $c = 0.5$ ,  $c = 0.9$ , and  $c = 0.99$

which is up to the additional dependence on  $\cos(\arcsin(\Sigma))$  comparable to the error bound in (3.2) for  $\hat{\Sigma}_n$ . This is remarkable since  $\tilde{\Sigma}_n$  accesses considerably less information on the samples than  $\hat{\Sigma}_n$ .

Theorem 3.7 even suggests that for strongly correlated distributions of  $\mathbf{X}$ , i.e.,  $\Sigma \approx \mathbf{1}$ , the dominant first term on the right-hand side of (3.12) vanishes. In other words, the bound in (3.12) predicts  $\tilde{\Sigma}_n$  to outperform  $\hat{\Sigma}_n$  if the entries of  $\mathbf{X}$  strongly correlate. Numerical experiments from [16] confirm this counter-intuitive fact, cf. Fig. 3.2. A possible explanation is that by construction,  $\tilde{\Sigma}_n$  implicitly uses the assumption that  $\Sigma$  has ones on its diagonal which is not provided to  $\hat{\Sigma}_n$ .

Furthermore, a corresponding lower bound on the second moment of the estimation error shows that the unconventional term  $\|\sigma(\mathbf{M} \odot \mathbf{A})\|$  is factual and not an artifact of the proof, cf. [16, Proposition 14].

### 3.4.2 Dithered Quantization

The results of Sect. 3.4.1 are restricted to the estimation of correlation matrices of Gaussian distributions. Both limitations stem from the chosen quantization model: first, (3.9) is blind to the rescaling of variances and, second, Grothendieck's identity only holds for Gaussian distributions. Nevertheless, by introducing a *dither* in the one-bit quantizer in (3.9), we can fully estimate the covariance matrix of general sub-Gaussian distributions. Dithering means adding artificial random noise (with

a suitably chosen distribution) to the samples before quantizing them to improve reconstruction from quantized observations, cf. [25, 26, 52]. In the context of one-bit compressed sensing, the effect of dithering was recently rigorously analyzed in [3, 17, 18, 32, 36], see also the survey [15].

To be precise, we require two bits per entry of each sample vector where each bit is dithered by an independent uniformly distributed dither, i.e., we are given

$$\text{sign}(\mathbf{X}^k + \boldsymbol{\tau}^k), \text{sign}(\mathbf{X}^k + \bar{\boldsymbol{\tau}}^k)^\top, \quad k = 1, \dots, n, \quad (3.13)$$

where the dithering vectors  $\boldsymbol{\tau}^1, \bar{\boldsymbol{\tau}}^1, \dots, \boldsymbol{\tau}^n, \bar{\boldsymbol{\tau}}^n$  are independent and uniformly distributed in  $[-\lambda, \lambda]^p$ , with  $\lambda > 0$  to be specified later. From the quantized observations in (3.13), we construct the estimator

$$\tilde{\boldsymbol{\Sigma}}_n^{\text{dith}} = \frac{1}{2} \tilde{\boldsymbol{\Sigma}}_n' + \frac{1}{2} (\tilde{\boldsymbol{\Sigma}}_n')^\top, \quad (3.14)$$

where

$$\tilde{\boldsymbol{\Sigma}}_n' = \frac{\lambda^2}{n} \sum_{k=1}^n \text{sign}(\mathbf{X}^k + \boldsymbol{\tau}^k) \text{sign}(\mathbf{X}^k + \bar{\boldsymbol{\tau}}^k)^\top. \quad (3.15)$$

**Theorem 3.8 ([16, Theorem 3])** *Let  $\mathbf{X}$  be a mean-zero,  $K$ -sub-Gaussian vector with covariance matrix  $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] = \boldsymbol{\Sigma}$ . Let  $\mathbf{X}^1, \dots, \mathbf{X}^n \stackrel{\text{d}}{\sim} \mathbf{X}$  be i.i.d. samples of  $\mathbf{X}$ . Let  $\mathbf{M} \in [0, 1]^{p \times p}$  be a fixed symmetric mask. If  $\lambda^2 \gtrsim_K \log(n) \|\boldsymbol{\Sigma}\|_\infty$ , then with probability at least  $1 - e^{-t}$ ,*

$$\begin{aligned} & \|\mathbf{M} \odot \tilde{\boldsymbol{\Sigma}}_n^{\text{dith}} - \mathbf{M} \odot \boldsymbol{\Sigma}\| \\ & \lesssim_K \|\mathbf{M}\|_{1 \rightarrow 2} (\lambda \|\boldsymbol{\Sigma}\|^{1/2} + \lambda^2) \sqrt{\frac{\log(p) + t}{n}} + \lambda^2 \|\mathbf{M}\| \frac{\log(p) + t}{n}. \end{aligned}$$

*In particular, if  $\lambda^2 \approx_K \log(n) \|\boldsymbol{\Sigma}\|_\infty$ , then*

$$\begin{aligned} & \|\mathbf{M} \odot \tilde{\boldsymbol{\Sigma}}_n^{\text{dith}} - \mathbf{M} \odot \boldsymbol{\Sigma}\| \\ & \lesssim_K \log(n) \|\mathbf{M}\|_{1 \rightarrow 2} \sqrt{\frac{\|\boldsymbol{\Sigma}\| \|\boldsymbol{\Sigma}\|_\infty (\log(p) + t)}{n}} + \log(n) \|\mathbf{M}\| \|\boldsymbol{\Sigma}\|_\infty \frac{\log(p) + t}{n}. \end{aligned} \quad (3.16)$$

The error bound (3.16) coincides (up to different logarithmic factors) with the best known estimate for the masked sample covariance matrix in Theorem 3.1, even though the sample covariance matrix requires direct access to the samples  $\mathbf{X}^k$ . This performance, however, heavily depends on the choice of  $\lambda$ , cf. [16]. Furthermore, it should be mentioned that there are cases where the performance of the dithered

estimator is significantly worse than the performance of the sample covariance matrix. Let us consider for simplicity the case  $\mathbf{M} = \mathbf{1}$ . If the samples  $\mathbf{X}^k$  are Gaussian, then [37] shows that

$$\mathbb{E}\|\hat{\Sigma}_n - \Sigma\| \simeq \sqrt{\frac{\|\Sigma\|\text{Tr}(\Sigma)}{n}} + \frac{\text{Tr}(\Sigma)}{n},$$

whereas (3.16) yields

$$\mathbb{E}\|\tilde{\Sigma}_n^{\text{dith}} - \Sigma\| \lesssim \log(n) \sqrt{\frac{p\|\Sigma\| \|\Sigma\|_{\infty} \log(p)}{n}} + \log(n) \frac{p\|\Sigma\|_{\infty} \log(p)}{n}$$

via tail integration. Since  $\text{Tr}(\Sigma) \leq p\|\Sigma\|_{\infty}$ , the second estimate is worse in general. Numerical experiments in [16] have shown that this difference is not an artifact of proof. Simply put,  $\hat{\Sigma}_n$  and  $\tilde{\Sigma}_n^{\text{dith}}$  perform similarly if  $\Sigma$  has a constant diagonal, whereas  $\hat{\Sigma}_n$  performs significantly better whenever  $\text{Tr}(\Sigma) \ll p\|\Sigma\|_{\infty}$ .

Theorem 3.8 can be extended to heavier-tailed random vectors. This, however, requires a larger choice of  $\lambda$  and thus more samples to reach the same error. For a sub-exponential random vector  $\mathbf{X}$ , one would already need  $\lambda^2 \gtrsim \log(n)^2 \cdot \max_{i \in [p]} \|X_i\|_{\psi_1}^2$ . The dependence of  $\lambda$  on  $n$ , both in the latter statement and in Theorem 3.8, can be observed in numerical experiments [16] as well.

Let us finally mention that the quantized estimators in (3.10) and (3.14) are not necessarily positive semi-definite as one expects from covariance matrices. In applications, one would thus replace both estimators by their projection onto the cone of positive semi-definite matrices, which is efficiently computed via the singular value decomposition [8, Section 8.1.1]. The obtained estimates also apply to the projected estimators since convex projections are 1-Lipschitz.

### 3.5 The Underlying Structures of Massive MIMO Covariance Estimation

Having the just surveyed theoretical insights on covariance matrix estimation in mind, let us return to the massive MIMO setup of Sect. 3.2. To understand the intrinsic structure of  $\Sigma_{\mathbf{h}}$  in (3.4) and consequent approaches in engineering literature to leverage it, we have to dive deeper into the underlying model and its physical interpretation. We thus follow the notational conventions of engineering literature in this section. Recall that the number of antennas  $M$  can be identified with the ambient dimension  $p$ , the number of time–frequency resources  $N$  with the number of samples  $n$ , and the channel vectors  $\mathbf{h}(s)$  correspond to samples  $\mathbf{X}^k$ .

Under the assumptions stated in the beginning of Sect. 3.2, i.e., the BS is equipped with  $M$  antennas in a ULA, the channel vector  $\mathbf{h}(s)$  can be written explicitly as



$$\mathbf{h}(s) = \int_{-1}^1 \rho(\xi, s) \mathbf{a}(\xi) d\xi,$$

for  $s \in [N]$ . Here,  $\xi = \frac{\sin(\theta)}{\sin(\theta_{\max})}$  are the normalized *angles of arrival (AoA)* with  $\theta_{\max} \in [0, \frac{\pi}{2}]$  being the maximum array angular aperture (cf. Fig. 3.1), the vectors  $\mathbf{a}(\xi) \in \mathbb{C}^M$  denote the respective array response at the BS, and the *channel gain*  $\rho(\xi, s)$  is a complex Gaussian process with zero mean. By assuming the antenna spacing to be  $d = \frac{\lambda}{2}$ , where  $\lambda = \frac{c_0}{f_0}$  denotes the wavelength with  $c_0$  being the speed of light and  $f_0$  the carrier frequency, we obtain

$$\mathbf{a}(\xi) = (1, e^{j\pi\xi}, \dots, e^{j\pi(M-1)\xi})^\top,$$

where  $j$  denotes the imaginary unit. With the additional assumption of *wide sense stationary uncorrelated scattering (WSSUS)*, the second-order statistics of the Gaussian process  $\rho(\xi, s)$  are time invariant and uncorrelated across AoAs so that

$$\mathbb{E}[\rho(\xi, s)\rho^*(\xi', s)] = \gamma(\xi) \delta(\xi - \xi'),$$

where  $\gamma: [-1, 1] \rightarrow \mathbb{R}_{\geq 0}$  is the real and non-negative measure that represents the *angular scattering function (ASF)* and  $\delta$  is the Dirac delta function. In particular, this implies that

$$\Sigma_{\mathbf{h}} = \mathbb{E}[\mathbf{h}(s)\mathbf{h}(s)^H] = \int_{-1}^1 \gamma(\xi) \mathbf{a}(\xi)\mathbf{a}(\xi)^H d\xi. \quad (3.17)$$

Building upon this explicit representation of  $\mathbf{h}(s)$  and structural assumptions on  $\gamma$ , one can refine the estimate obtained from the sample covariance matrix of  $\mathbf{y}$  defined in (3.3).

**A Hands-on Approach** In [35], we choose the following approach. First note that by (3.17) the channel covariance matrix belongs to the set

$$\mathcal{M} = \left\{ \int_{-1}^1 \gamma(\xi) \mathbf{a}(\xi)\mathbf{a}(\xi)^H d\xi : \gamma \in \mathcal{A} \right\},$$

where  $\mathcal{A}$  denotes the class of typical ASFs in wireless propagation. If one assumes sparse scattering propagation, the set  $\mathcal{A}$  consists of sparse ASFs. In particular, we assume that  $\gamma(\xi)$  can be decomposed as the sum of a discrete spike component  $\gamma_d$  (modeling the power received from *line-of-sight (LOS)* paths and narrow scatterers) and a continuous component  $\gamma_c$  (modeling the power received from wide scatterers). Mathematically, we can write

$$\gamma(\xi) = \gamma_d(\xi) + \gamma_c(\xi) = \sum_{k=1}^r c_k \delta(\xi - \xi_k) + \gamma_c(\xi), \quad (3.18)$$

where  $\gamma_d$  consists of  $r \ll M$  Dirac deltas with AoAs  $\xi_1, \dots, \xi_r$  and strengths  $c_1, \dots, c_r > 0$  corresponding to  $r$  specular propagation elements. Furthermore, by sparsity assumptions on  $\gamma$ , we have that  $\text{meas}(\gamma_c) \ll \text{meas}([-1, 1])$ , where  $\text{meas}(\gamma_c)$  denotes here the measure of the support of  $\gamma_c$ . Combining (3.17) and (3.18), we decompose the channel covariance matrix as

$$\mathbf{\Sigma}_{\mathbf{h}} = \mathbf{\Sigma}_{\mathbf{h}}^d + \mathbf{\Sigma}_{\mathbf{h}}^c = \sum_{k=1}^r c_k \mathbf{a}(\xi_k) \mathbf{a}(\xi_k)^H + \int_{-1}^1 \gamma_c(\xi) \mathbf{a}(\xi) \mathbf{a}(\xi)^H d\xi, \quad (3.19)$$

where  $\mathbf{\Sigma}_{\mathbf{h}}^d$  is rank- $r$  and positive semi-definite and  $\mathbf{\Sigma}_{\mathbf{h}}^c$  is full rank and positive semi-definite with few dominant singular values. We can approximate  $\mathbf{\Sigma}_{\mathbf{h}}$  now in three consecutive steps:

- (i) *Spike Location Estimation for  $\gamma_d$* : Applying the *MUltiple Signal Classification (MUSIC)* algorithm [58], we estimate the AoAs  $\xi_k$  of the spike component  $\gamma_d$  from the noisy samples  $\mathbf{y}(1), \dots, \mathbf{y}(N)$ , cf. [35, Theorem 1]. Since this step is fairly standard, we do not discuss the details here but refer the interested reader to [35]. Let us only mention that the number of spikes is estimated by the number of dominant eigenvalues of  $\mathbf{\Sigma}_{\mathbf{y}} := \mathbb{E}[\mathbf{y}(s)\mathbf{y}(s)^H]$  (where one can naturally assume a corresponding gap in the spectrum since the power received via LOS paths in  $\gamma_d$  dominates the power received from wide scatterers in  $\gamma_c$ ). As a result, we obtain estimated spike locations  $\hat{\xi}_k$ , for  $k \in [\hat{r}]$ , and define an approximation of  $\gamma_d$

$$\tilde{\gamma}_d(\xi) = \sum_{k=1}^{\hat{r}} \tilde{c}_k \delta(\xi - \hat{\xi}_k),$$

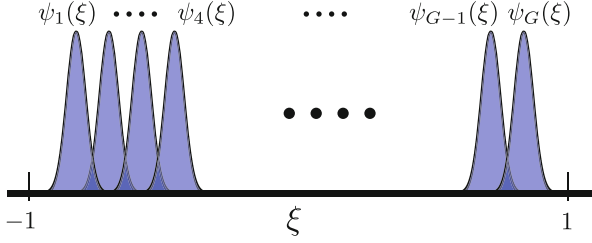
where the coefficients  $\tilde{c}_1, \dots, \tilde{c}_{\hat{r}} \geq 0$  still need to be estimated.

- (ii) *Sparse Dictionary-Based Method*: We approximate the continuous component  $\gamma_c$  over a finite dictionary of densities  $\mathcal{G}_c := \{\psi_i : [-1, 1] \rightarrow \mathbb{R}, i \in [G]\}$  that are suitably chosen, e.g., Gaussian, Laplacian, or rectangular kernels, cf. Fig. 3.3. We hence define

$$\tilde{\gamma}_c(\xi) = \sum_{i=1}^G \tilde{b}_i \psi_i(\xi),$$

where only the coefficients  $\tilde{b}_1, \dots, \tilde{b}_G \geq 0$  need to be estimated.

- (iii) *Non-Negative Least Square (NNLS) estimator*: Collecting the coefficients in a single vector  $\mathbf{u} = (\tilde{b}_1, \dots, \tilde{b}_G, \tilde{c}_1, \dots, \tilde{c}_{\hat{r}})^T \in \mathbb{R}_{\geq 0}^{G+\hat{r}}$  and recalling (3.19), we define our coefficient-dependent estimate of the channel covariance



**Fig. 3.3** Example of a Gaussian dictionary that might be used to express  $\gamma_c$

$$\Sigma_{\mathbf{h}}(\mathbf{u}) = \sum_{k=1}^{\hat{r}} \tilde{c}_k \mathbf{a}(\hat{\xi}_k) \mathbf{a}(\hat{\xi}_k)^H + \sum_{i=1}^G \tilde{b}_i \int_{-1}^1 \psi_i(\xi) \mathbf{a}(\xi) \mathbf{a}(\xi)^H d\xi =: \sum_{i=1}^{G+\hat{r}} u_i \mathbf{S}_i, \quad (3.20)$$

where

$$\mathbf{S}_i = \begin{cases} \int_{-1}^1 \psi_i(\xi) \mathbf{a}(\xi) \mathbf{a}(\xi)^H d\xi & \text{if } 1 \leq i \leq G \\ \mathbf{a}(\hat{\xi}_{i-G}) \mathbf{a}(\hat{\xi}_{i-G})^H & \text{if } G < i \leq G + \hat{r}. \end{cases}$$

All that remains is to determine the coefficient vector  $\mathbf{u}$ . Since  $\Sigma_{\mathbf{y}} = \Sigma_{\mathbf{h}} + N_0 \mathbf{I}$ , we can do so by fitting (3.20) to the sample covariance matrix  $\hat{\Sigma}_{\mathbf{y}}$  of  $\mathbf{y}(1), \dots, \mathbf{y}(N)$ , i.e.,

$$\mathbf{u}^* = \arg \min_{\mathbf{u} \geq \mathbf{0}} \left\| \hat{\Sigma}_{\mathbf{y}} - \sum_{i=1}^{G+\hat{r}} u_i \mathbf{S}_i - N_0 \mathbf{I} \right\|_F^2. \quad (3.21)$$

Since  $\Sigma_{\mathbf{h}}$  is Hermitian Toeplitz, one can incorporate the structure in (3.21) by replacing  $\hat{\Sigma}_{\mathbf{h}} = \hat{\Sigma}_{\mathbf{y}} - N_0 \mathbf{I}$  with its projection  $\tilde{\Sigma}_{\mathbf{h}}$  onto the space of Hermitian Toeplitz matrices (which can be done by averaging the diagonals as in (3.7)). Denoting the first column of  $\tilde{\Sigma}_{\mathbf{h}}$  by  $\tilde{\sigma} \in \mathbb{C}^M$  and collecting the first columns of the matrices  $\mathbf{S}_i$  in a matrix  $\tilde{\mathbf{S}} \in \mathbb{C}^{M \times (G+\hat{r})}$ , we may instead solve

$$\mathbf{u}^* = \arg \min_{\mathbf{u} \geq \mathbf{0}} \left\| \mathbf{W}(\tilde{\mathbf{S}}\mathbf{u} - \tilde{\sigma}) \right\|_F^2, \quad (3.22)$$

where  $\mathbf{W} = \text{diag}((\sqrt{M}, \sqrt{2(M-1)}, \sqrt{2(M-2)}, \dots, \sqrt{2})^\top)$  is a weight matrix compensating the averaging process.

**A Hands-on Approach: Empirical Evaluation** Let us empirically compare the NNLS estimator to the sample covariance matrix right away. We consider a ULA with  $M = 128$  antennas, where the spacing between two consecutive antenna

elements is set to  $d = \frac{\lambda}{2}$ . We produce random ASFs in the following general format:

$$\begin{aligned} \gamma(\xi) &= \gamma_d(\xi) + \gamma_c(\xi) \\ &= \frac{\alpha}{r} \sum_{i=1}^r \delta(\xi - \xi_i) + \frac{1-\alpha}{Z} \left( \sum_{j=1}^{n_r} \text{rect}_{\mu_j, \sigma_j}(\xi) + \sum_{k=1}^{n_g} \text{Gaussian}_{\mu_k, \sigma_k}(\xi) \right), \end{aligned} \quad (3.23)$$

where we set the number of delta, rectangular, and Gaussian functions to be  $r := 2$ ,  $n_r := 2$ , and  $n_g := 2$ , respectively. The spike locations are chosen uniformly at random from  $[-1, 1]$ , i.e.,  $\xi_i \sim \text{Unif}([-1, 1])$  for  $i \in [2]$ . The rectangular functions are defined as

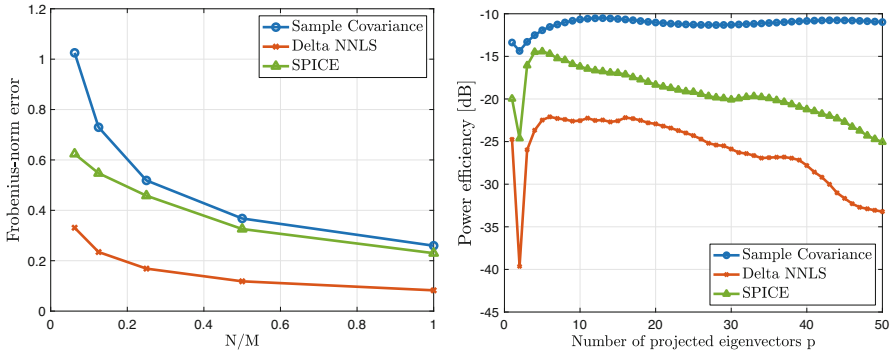
$$\text{rect}_{\mu_j, \sigma_j}(\xi) = \chi_{\left[\mu_j - \frac{\sigma_j}{2}, \mu_j + \frac{\sigma_j}{2}\right]}(\xi),$$

where  $\mu_1 \sim \text{Unif}([-1, 0])$ ,  $\mu_2 \sim \text{Unif}([0, 1])$ , and  $\sigma_j \sim \text{Unif}([0.1, 0.3])$ , for  $j \in [2]$ . The Gaussian functions  $\text{Gaussian}_{\mu_k, \sigma_k}$  are densities of  $\mathcal{N}(\mu_k, \sigma_k)$ , where  $\mu_k \sim \text{Unif}([-0.7, 0.7])$  and  $\sigma_k \sim \text{Unif}([0.03, 0.04])$ , for  $k \in [2]$ . Moreover,  $\alpha := 0.5$  is set to present the power contribution of discrete spikes. The constant  $Z = \int_{-1}^1 \gamma_c(\xi) d\xi$  normalizes  $\gamma_c$  in measure. The signal-to-noise ratio (SNR) is set to 10 dB.

In addition to the sample covariance, we compare our NNLS estimator to sparse iterative covariance-based estimation (SPICE) [57]. This method also exploits the ASF domain to fit a covariance matrix. Note that SPICE can only be applied with Dirac delta dictionaries and that it does not include a step of spike support detection as in our method.

Denoting a generic covariance estimate as  $\bar{\Sigma}$ , we consider two metrics to evaluate the estimation quality. The first metric, the *normalized Frobenius-norm error*, is defined as  $E_{\text{NF}} = \frac{\|\Sigma_{\mathbf{h}} - \bar{\Sigma}\|_{\text{F}}}{\|\Sigma_{\mathbf{h}}\|_{\text{F}}}$ . The second metric, the *power efficiency*, evaluates the similarity of dominant subspaces between the estimated and true matrices, which is an important factor in various applications of massive MIMO such as user grouping and group-based beamforming. Specifically, let  $d \in [M]$  denote a subspace dimension parameter, and let  $\mathbf{U}_d \in \mathbb{C}^{M \times d}$  and  $\bar{\mathbf{U}}_d \in \mathbb{C}^{M \times d}$  be the  $d$  dominant eigenvectors of  $\Sigma_{\mathbf{h}}$  and  $\bar{\Sigma}$  corresponding to their largest  $d$  eigenvalues, respectively. Then, the power efficiency based on  $d$  is defined as  $E_{\text{PE}}(d) = 1 - \frac{\langle \Sigma_{\mathbf{h}}, \bar{\mathbf{U}}_d \bar{\mathbf{U}}_d^{\text{H}} \rangle}{\langle \Sigma_{\mathbf{h}}, \mathbf{U}_d \mathbf{U}_d^{\text{H}} \rangle}$ . Note that  $E_{\text{PE}}(d) \in [0, 1]$  where a value closer to 0 means that more power is captured by the estimated  $d$ -dominant subspace.

SPICE and the proposed NNLS estimators are applied with  $G = 2M$  Dirac delta dictionaries for the continuous part  $\mathcal{G}_c$ . The resulting Frobenius-norm error and power efficiency are depicted in Fig. 3.4. All results are averaged over 20 random ASFs and 200 random channel realizations for each ASF. The proposed NNLS method outperforms the sample covariance matrix and SPICE for both metrics.



**Fig. 3.4** Frobenius-norm error (left) and power efficiency with  $\frac{N}{M} = 0.5$  (right)

Finally, one can observe a similar outcome for smaller sample sizes as well, e.g.,  $N/M = 0.125$ , which occur naturally in massive MIMO.

### 3.6 Conclusion

The present chapter shows that in the last decade good progress has been made on understanding the influence of intrinsic structure of covariance matrices on the non-asymptotic performance of suitably designed estimators. As we have seen, such estimators with strong guarantees are available for sparse, low-rank, and Toeplitz covariance matrices. At the same time, the chapter illustrates that practitioners still continue to tweak the basic sample covariance matrix using their specific knowledge of the application at hand—seemingly unaware of the progress in theory. We hope that this essay helps mathematicians and practitioners alike to gain an overview of recent theoretical developments on structural and quantized covariance estimation and that it motivates mathematicians to look deeper into the underlying physical models of concrete applications to better understand the structures of interest. Furthermore, our recent theoretical progress on quantized covariance estimation suggests that reliable reconstruction of the covariance matrix is possible even under heavy loss of information during sampling. The use of coarse quantization might thus lead to a considerable increase in capacity in massive MIMO systems and related applications.

**Acknowledgments** All authors acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the project CoCoMIMO funded within the priority program SPP 1798 Compressed Sensing in Information Processing (COSIP).

## Appendix: Proof of Theorem 3.6

To prove Theorem 3.6, we need two technical lemmas. In the remaining section,  $\sigma$  always refers to the first column of  $\Sigma$  and  $\hat{\sigma}$  to the first column of  $\hat{\Sigma}_n^{\text{Toep}}$ .

**Lemma 3.1** *Under the assumptions of Theorem 3.6, we have for  $\alpha \in (0, 1)$  and  $0 < u < 1$  that*

$$\Pr \left[ \max_{r \leq \alpha p} |\hat{\sigma}_r - \sigma_r| \geq \sqrt{u} \right] \leq 2\alpha p e^{-(1-\alpha) \min\left\{\frac{1}{CK^4}, \frac{1}{CK^2}\right\} n p u},$$

where  $C > 0$  is an absolute constant.

**Proof** We proceed similar as in [33]. First note that, for all  $k \in [n]$ ,  $r \in [\alpha p]$ , we can write

$$\begin{aligned} |\hat{\sigma}_r - \sigma_r| &= \frac{1}{(p+1)-r} \left| \sum_{j-i=r-1} (X_i^k X_j^k - \sigma_r) \right| \\ &= \left| \langle \mathbf{M}_r \mathbf{X}^k, \mathbf{X}^k \rangle - \mathbb{E}[\langle \mathbf{M}_r \mathbf{X}^k, \mathbf{X}^k \rangle] \right|, \end{aligned} \quad (3.24)$$

where the mask  $\mathbf{M}_r$  is defined by  $[M_r]_{i,j} = \frac{1}{(p+1)-r}$  if  $j-i=r-1$  and  $[M_r]_{i,j} = 0$  else, i.e., only the  $r$ -th co-diagonal of  $\mathbf{M}$  is non-zero. By using a version of the Hanson–Wright inequality for random vectors with the convex concentration property [1], we get that

$$\Pr \left[ \underbrace{|\langle \mathbf{M}_r \mathbf{X}^k, \mathbf{X}^k \rangle - \mathbb{E}[\langle \mathbf{M}_r \mathbf{X}^k, \mathbf{X}^k \rangle]}_{=: Z_i^r} \geq u \right] \leq 2e^{-\min\left\{\frac{u^2}{CK^4 \|\mathbf{M}_r\|_F^2}, \frac{u}{CK^2 \|\mathbf{M}_r\|}\right\}},$$

which, by integration, leads to

$$\begin{aligned} \mathbb{E}[|Z_i^r|^{2q}] &\leq 2q(2CK^4 \|\mathbf{M}_r\|_F^2)^q \Gamma(q) + 4q(CK^2 \|\mathbf{M}_r\|)^{2q} \Gamma(2q) \\ &\leq q!(4CK^4 \|\mathbf{M}_r\|_F^2)^q + (2q)!(2CK^2 \|\mathbf{M}_r\|)^{2q}, \end{aligned}$$

for any  $q \geq 1$ . The random variables  $Z_i^r$  are thus sub-gamma with variance parameter  $\nu = 16K^4(C\|\mathbf{M}_r\|_F^2 + C^2\|\mathbf{M}_r\|^2) \leq CK^4\|\mathbf{M}_r\|_F^2$  and scale parameter  $\gamma = 4CK^2\|\mathbf{M}_r\|^2$  [7, Theorem 2.3]. By independence, we get for all  $0 < \mu < \frac{1}{\gamma}$

$$\mathbb{E} \left[ e^{\mu \sum_{i=1}^n Z_i^r} \right] = \prod_{i=1}^n \mathbb{E}[e^{\mu Z_i^r}] \leq e^{\frac{\mu^2 n \nu}{2(1-\gamma\mu)}}$$

(and the same holds for  $-Z_i^r$ ) such that  $\sum_{i=1}^n Z_i^r$  is sub-gamma with variance parameter  $\nu n$  and scale parameter  $\gamma$  [7, Chapter 2.4]. Consequently,

$$\begin{aligned} \Pr \left[ \left| \frac{1}{n} \sum_{i=1}^n Z_i^r \right| \geq CK^2 \left( \|\mathbf{M}_r\|_F \sqrt{\frac{u}{n}} + \|\mathbf{M}_r\| \frac{u}{n} \right) \right] \\ \leq \Pr \left[ \left| \sum_{i=1}^n Z_i^r \right| \geq \sqrt{2\nu nu} + \gamma u \right] \leq 2e^{-u}, \end{aligned}$$

for any  $u > 0$  [7, Chapter 2.4]. Recalling (3.24) and noting that  $\|\mathbf{M}_r\|_F^2 = \|\mathbf{M}_r\| = \frac{1}{(p+1)-r}$  yield with the choice  $u = \min \left\{ \frac{1}{C^2K^4}, \frac{1}{CK^2} \right\} ((p+1)-r)n\tilde{u}$  that

$$\begin{aligned} \Pr \left[ |\tilde{\sigma}_r - \sigma_r| \geq 2\sqrt{\tilde{u}} \right] &\leq \Pr \left[ |\tilde{\sigma}_r - \sigma_r| \geq \sqrt{\tilde{u}} + \tilde{u} \right] \\ &\leq 2e^{-\min \left\{ \frac{1}{C^2K^4}, \frac{1}{CK^2} \right\} ((p+1)-r)n\tilde{u}}. \end{aligned}$$

A union bound over  $r \in [\alpha p]$  and the bound  $r \leq \alpha p$  conclude the proof.  $\square$

The second lemma follows along the lines of [5, Theorem 1].

**Lemma 3.2** *Under the assumptions of Theorem 3.6, assume in addition that  $\Sigma$  has a bandwidth of at most  $\alpha p$ , i.e.,  $\mathbb{B}_{\alpha p}(\Sigma) = \Sigma$  and thus  $\text{supp}(\sigma) \subset [\alpha p]$  and that*

$$\max_{i,j \in [p]} |\mathbb{B}_{\alpha p}(\hat{\Sigma}_n^{\text{Toep}})_{i,j} - \Sigma_{i,j}| = \max_{r \leq \alpha p} |\hat{\sigma}_r - \sigma_r| \leq (1-\gamma)\tau, \quad (3.25)$$

for some  $\gamma \in (0, 1)$ . Then,

$$\|\mathbb{T}_\tau(\mathbb{B}_{\alpha p}(\hat{\Sigma}_n^{\text{Toep}})) - \Sigma\| \lesssim_\gamma \tau^{1-q} s.$$

**Proof** For convenience, let us abbreviate  $\tilde{\Sigma} := \mathbb{B}_{\alpha p}(\hat{\Sigma}_n^{\text{Toep}})$  and denote its first column by  $\tilde{\sigma}$ . We write

$$\|\mathbb{T}_\tau(\tilde{\Sigma}) - \Sigma\| \leq \|\mathbb{T}_\tau(\Sigma) - \Sigma\| + \|\mathbb{T}_\tau(\tilde{\Sigma}) - \mathbb{T}_\tau(\Sigma)\|.$$

Since  $\Sigma \in \mathcal{U}^{\text{Toep}}(q, s, M)$ ,

$$\sum_{j=1}^p |\Sigma_{i,j}| \chi_{\{|\Sigma_{i,j}| \leq \tau\}} = \sum_{j=1}^p |\Sigma_{i,j}|^q |\Sigma_{i,j}|^{1-q} \chi_{\{|\Sigma_{i,j}| \leq \tau\}} \leq \tau^{1-q} \sum_{j=1}^p |\Sigma_{i,j}|^q$$

and Gershgorin's disc theorem imply

$$\|\mathbb{T}_\tau(\boldsymbol{\Sigma}) - \boldsymbol{\Sigma}\| \leq \max_i \sum_{j=1}^p |\Sigma_{i,j}| \chi_{\{|\Sigma_{i,j}| \leq \tau\}} \leq \tau^{1-q} s. \quad (3.26)$$

Moreover,

$$\begin{aligned} \|\mathbb{T}_\tau(\tilde{\boldsymbol{\Sigma}}) - \mathbb{T}_\tau(\boldsymbol{\Sigma})\| &\leq \max_i \sum_{j=1}^p |\tilde{\Sigma}_{i,j}| \chi_{\{|\tilde{\Sigma}_{i,j}| \geq \tau, |\Sigma_{i,j}| < \tau\}} \\ &\quad + \max_i \sum_{j=1}^p |\Sigma_{i,j}| \chi_{\{|\tilde{\Sigma}_{i,j}| < \tau, |\Sigma_{i,j}| \geq \tau\}} \\ &\quad + \max_i \sum_{j=1}^p |\tilde{\Sigma}_{i,j} - \Sigma_{i,j}| \chi_{\{|\tilde{\Sigma}_{i,j}| \geq \tau, |\Sigma_{i,j}| \geq \tau\}} \\ &= (I) + (II) + (III). \end{aligned}$$

First recall that by assumption  $\text{supp}(\boldsymbol{\sigma}) \subset [\alpha p]$  and  $\text{supp}(\tilde{\boldsymbol{\sigma}}) \subset [\alpha p]$ . Hence, using the observation that  $\tilde{\sigma}_r = \hat{\sigma}_r$ , for  $r \leq \alpha p$ , and

$$\sum_{j=1}^p \chi_{\{|\Sigma_{i,j}| \geq \tau\}} = \sum_{j=1}^p \tau^q \tau^{-q} \chi_{\{|\Sigma_{i,j}| \geq \tau\}} \leq \sum_{j=1}^p |\Sigma_{i,j}|^q \tau^{-q}, \quad (3.27)$$

we may estimate with (3.25) and  $\boldsymbol{\Sigma} \in \mathcal{U}^{\text{Toep}}(q, s, M)$

$$(III) \leq \max_{r \leq \alpha p} |\hat{\sigma}_r - \sigma_r| \cdot \max_i \sum_{j=1}^p |\Sigma_{i,j}|^q \tau^{-q} \leq s \tau^{1-q}.$$

Furthermore,

$$\begin{aligned} (I) &\leq \max_i \sum_{j=1}^p |\tilde{\Sigma}_{i,j} - \Sigma_{i,j}| \chi_{\{|\tilde{\Sigma}_{i,j}| \geq \tau, |\Sigma_{i,j}| < \tau\}} + \max_i \sum_{j=1}^p |\Sigma_{i,j}| \chi_{\{|\tilde{\Sigma}_{i,j}| \geq \tau, |\Sigma_{i,j}| < \tau\}} \\ &= (IV) + (V). \end{aligned}$$

By (3.26), we know that (V)  $\leq \tau^{1-q} s$ . Furthermore, we get that

$$(IV) \leq \max_i \sum_{j=1}^p |\tilde{\Sigma}_{i,j} - \Sigma_{i,j}| \chi_{\{|\tilde{\Sigma}_{i,j}| \geq \tau, |\Sigma_{i,j}| < \gamma \tau\}}$$



$$\begin{aligned}
& + \max_i \sum_{j=1}^p |\tilde{\Sigma}_{i,j} - \Sigma_{i,j}| \chi_{\{|\tilde{\Sigma}_{i,j}| \geq \tau, \gamma\tau \leq |\Sigma_{i,j}| < \tau\}} \\
& \leq \max_{r \leq \alpha p} |\tilde{\sigma}_r - \sigma_r| \cdot \max_i N_i(1 - \gamma) + s(\gamma\tau)^{-q} \tau,
\end{aligned}$$

where we defined  $N_i(1 - \gamma) := \sum_{j=1}^p \chi_{\{|\tilde{\Sigma}_{i,j} - \Sigma_{i,j}| > (1 - \gamma)\tau\}}$  and reused the bound on (III) for the second term (replacing  $\tau$  with  $\gamma\tau$  in the summation). Since we have by (3.25) that  $N_i(1 - \gamma) = 0$ , for  $i \in [p]$ , we get that (IV)  $\lesssim_{\gamma} s\tau^{1-q}$ . Hence,

$$(I) \lesssim_{\gamma} s\tau^{1-q}.$$

Finally, note that by (3.27) and  $\Sigma \in \mathcal{U}^{\text{Toep}}(q, s, M)$ ,

$$\begin{aligned}
(II) & \leq \max_i \sum_{j=1}^p (|\tilde{\Sigma}_{i,j} - \Sigma_{i,j}| + |\tilde{\Sigma}_{i,j}|) \chi_{\{|\tilde{\Sigma}_{i,j}| < \tau, |\Sigma_{i,j}| \geq \tau\}} \\
& \leq \max_{r \leq \alpha p} |\tilde{\sigma}_r - \sigma_r| \cdot \max_i \sum_{j=1}^p \chi_{\{|\Sigma_{i,j}| \geq \tau\}} + \tau \max_i \sum_{j=1}^p \chi_{\{|\Sigma_{i,j}| \geq \tau\}} \\
& \leq s\tau^{1-q} + s\tau^{1-q}.
\end{aligned}$$

Combining the bounds for (I), (II), and (III) yields the claim.  $\square$

**Proof of Theorem 3.6** Note that

$$\|\mathbb{T}_{\tau}(\mathbb{B}_{\alpha p}(\hat{\Sigma}_n^{\text{Toep}})) - \Sigma\| \leq \|\mathbb{T}_{\tau}(\mathbb{B}_{\alpha p}(\hat{\Sigma}_n^{\text{Toep}})) - \mathbb{B}_{\alpha p}(\Sigma)\| + \|\mathbb{B}_{\alpha p}(\Sigma) - \Sigma\|. \quad (3.28)$$

By Lemma 3.1, we get with probability at least  $1 - (2\alpha p)^{-(c-1)}$  that

$$\max_{r \leq \alpha p} |\hat{\sigma}_r - \sigma_r| \leq \sqrt{\frac{c}{1 - \alpha}} \max\{CK^2, \sqrt{C}K\} \sqrt{\frac{\log(p)}{np}} = (1 - \gamma)\tau, \quad (3.29)$$

where  $c > 1$  and  $1 - \gamma = \frac{1}{\sqrt{2}}$ . The claim now follows by applying Lemma 3.2 to the first term on the right-hand side of (3.28).  $\square$

## References

1. Adamczak, R.: A note on the Hanson-Wright inequality for random vectors with dependencies. *Electron. Commun. Probab.* **20** (2015)

2. Bar-Shalom, O., Weiss, A.J.: DOA estimation using one-bit quantized measurements. *IEEE Trans. Aerospace Electron. Syst.* **38**(3), 868–884 (2002)
3. Baraniuk, R.G., Foucart, S., Needell, D., Plan, Y., Wootters, M.: Exponential decay of reconstruction error from binary measurements of sparse signals. *IEEE Trans. Inform. Theory* **63**(6), 3368–3385 (2017)
4. Benedetto, J.J., Powell, A.M., Yilmaz, O.: Sigma-delta quantization and finite frames. *IEEE Trans. Inform. Theory* **52**(5), 1990–2005 (2006)
5. Bickel, P.J., Levina, E.: Covariance regularization by thresholding. *Ann. Stat.* **36**(6), 2577–2604 (2008)
6. Bickel, P.J., Levina, E.: Regularized estimation of large covariance matrices. *Ann. Stat.* **36**(1), 199–227 (2008)
7. Boucheron, S., Lugosi, G., Massart, P.: Concentration inequalities: a nonasymptotic theory of independence. Oxford University Press, Oxford (2013)
8. Boyd, S., Vandenberghe, L.: *Convex Optimization*. Cambridge University Press, Cambridge (2004)
9. Brookes, M.J., Vrba, J., Robinson, S.E., Stevenson, C.M., Peters, A.M., Barnes, G.R., Hillebrand, A., Morris, P.G.: Optimising experimental design for MEG beamformer imaging. *Neuroimage* **39**(4), 1788–1802 (2008)
10. Cai, T.T., Ren, Z., Zhou, H.H.: Optimal rates of convergence for estimating Toeplitz covariance matrices. *Probab. Theory Related Fields* **156**(1-2), 101–143 (2013)
11. Cai, T.T., Zhang, C.H., Zhou, H.H.: Optimal rates of convergence for covariance matrix estimation. *Ann. Stat.* **38**(4), 2118–2144 (2010)
12. Catoni, O.: Challenging the empirical mean and empirical variance: a deviation study. In: *Annales de l’IHP Probabilités et statistiques*, vol. 48, pp. 1148–1185 (2012)
13. Chen, R.Y., Gittens, A., Tropp, J.A.: The masked sample covariance estimator: an analysis using matrix concentration inequalities. *Inform. Inference J. IMA* **1**(1), 2–20 (2012)
14. Choi, J., Mo, J., Heath, R.W.: Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs. *IEEE Trans. Commun.* **64**(5), 2005–2018 (2016)
15. Dirksen, S.: Quantized compressed sensing: a survey. In: *Compressed Sensing and Its Applications: Third International MATHEON Conference 2017*, pp. 67–95. Applied and Numerical Harmonic Analysis. Birkhäuser, Cham (2019)
16. Dirksen, S., Maly, J., Rauhut, H.: Covariance estimation under one-bit quantization. *arXiv preprint arXiv:2104.01280* (2021)
17. Dirksen, S., Mendelson, S.: Robust one-bit compressed sensing with partial circulant matrices. *ArXiv:1812.06719* (2018)
18. Dirksen, S., Mendelson, S.: Non-gaussian hyperplane tessellations and robust one-bit compressed sensing. *J. Eur. Math. Soc. Arxiv: 1805.09409* (2021)
19. El Karoui, N.: Operator norm consistent estimation of large-dimensional sparse covariance matrices. *Ann. Stat.* **36**(6), 2717–2756 (2008)
20. Eldar, Y.C., Li, J., Musco, C., Musco, C.: Sample efficient Toeplitz covariance estimation. In: *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 378–397. SIAM (2020)
21. Fazel, M.: Matrix rank minimization with applications. Ph.D. Thesis, Stanford University (2002)
22. Foucart, S., Rauhut, H.: *A Mathematical Introduction to Compressive Sensing*. Applied and Numerical Harmonic Analysis. Birkhäuser, Basel (2013)
23. Furrer, R., Bengtsson, T.: Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *J. Multivariate Anal.* **98**(2), 227–255 (2007)
24. Goldsmith, A., Jafar, S.A., Jindal, N., Vishwanath, S.: Capacity limits of MIMO channels. *IEEE J. Selected Areas Commun.* **21**(5), 684–702 (2003)
25. Gray, R.M., Neuhoff, D.L.: Quantization. *IEEE Trans. Inform. Theory* **44**(6), 2325–2383 (1998)

26. Gray, R.M., Stockham, T.G.: Dithered quantizers. *IEEE Trans. Inform. Theory* **39**(3), 805–812 (1993)
27. Haghhighatshoar, S., Caire, G.: Massive MIMO channel subspace estimation from low-dimensional projections. *IEEE Trans. Signal Process.* **65**(2), 303–318 (2016)
28. Haghhighatshoar, S., Caire, G.: Low-complexity massive MIMO subspace estimation and tracking from low-dimensional projections. *IEEE Trans. Signal Process.* **66**(7), 1832–1844 (2018)
29. Hubert, M., Rousseeuw, P.J., Van Aelst, S.: High-breakdown robust multivariate methods. *Statistical Science*, pp. 92–119 (2008)
30. Jacovitti, G., Neri, A.: Estimation of the autocorrelation function of complex Gaussian stationary processes by amplitude clipped signals. *IEEE Trans. Inform. Theory* **40**(1), 239–245 (1994)
31. Jerrum, M.R., Valiant, L.G., Vazirani, V.V.: Random generation of combinatorial structures from a uniform distribution. *Theor. Comput. Sci.* **43**, 169–188 (1986)
32. Jung, H.C., Maly, J., Palzer, L., Stollenwerk, A.: Quantized compressed sensing by rectified linear units. *IEEE Trans. Inform. Theory* **67**(6), 4125–4149 (2021)
33. Kabanava, M., Rauhut, H.: Masked Toeplitz covariance estimation. [ArXiv:1709.09377](https://arxiv.org/abs/1709.09377) (2017)
34. Ke, Y., Minsker, S., Ren, Z., Sun, Q., Zhou, W.X.: User-friendly covariance estimation for heavy-tailed distributions. *Stat. Sci.* **34**(3), 454–471 (2019)
35. Khalilsarai, M.B., Yang, T., Haghhighatshoar, S., Caire, G.: Structured channel covariance estimation from limited samples in massive MIMO. In: *IEEE International Conference on Communications (ICC)*, pp. 1–7 (2020)
36. Knudson, K., Saab, R., Ward, R.: One-bit compressive sensing with norm estimation. *IEEE Trans. Inform. Theory* **62**(5), 2748–2758 (2016)
37. Koltchinskii, V., Lounici, K.: Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli* **23**(1), 110–133 (2017)
38. Krim, H., Viberg, M.: Two decades of array signal processing research: the parametric approach. *IEEE Signal Process. Mag.* **13**(4), 67–94 (1996)
39. Lawrence, H., Li, J., Musco, C., Musco, C.: Low-rank Toeplitz matrix estimation via random ultra-sparse rulers. In: *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4796–4800. *IEEE* (2020)
40. Levina, E., Vershynin, R.: Partial estimation of covariance matrices. *Probab. Theory Related Fields* **153**(3–4), 405–419 (2012)
41. Li, Y., Tao, C., Seco-Granados, G., Mezghani, A., Swindlehurst, A.L., Liu, L.: Channel estimation and performance analysis of one-bit massive MIMO systems. *IEEE Trans. Signal Process.* **65**(15), 4075–4089 (2017)
42. Liu, L., Hawkins, D.M., Ghosh, S., Young, S.S.: Robust singular value decomposition analysis of microarray data. *Proc. Nat. Acad. Sci.* **100**(23), 13167–13172 (2003)
43. Lounici, K.: High-dimensional covariance matrix estimation with missing observations. *Bernoulli* **20**(3), 1029–1058 (2014)
44. Lu, L., Li, G.Y., Swindlehurst, A.L., Ashikhmin, A., Zhang, R.: An overview of massive MIMO: Benefits and challenges. *IEEE J. Selected Topics Signal Process.* **8**(5), 742–758 (2014)
45. Marzetta, T.L., Ngo, H.Q.: *Fundamentals of massive MIMO*. Cambridge University Press, Cambridge (2016)
46. Mendelson, S., Zhivotovskiy, N.: Robust covariance estimation under L4-L2 norm equivalence. *Ann. Stat.* **48**(3), 1648–1664 (2020)
47. Minsker, S.: Sub-Gaussian estimators of the mean of a random matrix with heavy-tailed entries. *Ann. Stat.* **46**(6A), 2871–2903 (2018)
48. Minsker, S., Wei, X.: Robust modifications of U-statistics and applications to covariance estimation problems. *Bernoulli* **26**(1), 694–727 (2020)
49. Nemirovskij, A.S., Yudin, D.B.: *Problem complexity and method efficiency in optimization* (1983)
50. Paulraj, A.J., Gore, D.A., Nabar, R.U., Bolcskei, H.: An overview of MIMO communications—a key to gigabit wireless. *Proc. IEEE* **92**(2), 198–218 (2004)

51. Recht, B., Fazel, M., Parrilo, P.A.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.* **52**(3), 471–501 (2010)
52. Roberts, L.: Picture coding using pseudo-random noise. *IRE Trans. Inform. Theory* **8**(2), 145–154 (1962)
53. Romero, D., Ariananda, D.D., Tian, Z., Leus, G.: Compressive covariance sensing: Structure-based compressive sensing beyond sparsity. *IEEE Signal Process. Mag.* **33**(1), 78–93 (2016)
54. Roth, K., Munir, J., Mezghani, A., Nosssek, J.A.: Covariance based signal parameter estimation of coarse quantized signals. In: 2015 IEEE International Conference on Digital Signal Processing (DSP), pp. 19–23. IEEE (2015)
55. Schreier, R., Temes, G.C., Norsworthy, S.R.: *Delta-Sigma Data Converters: Theory, Design, and Simulation*. IEEE Press (1996)
56. Snyder, D.L., O’Sullivan, J.A., Miller, M.I.: The use of maximum likelihood estimation for forming images of diffuse radar targets from delay-doppler data. *IEEE Trans. Inform. Theory* **35**(3), 536–548 (1989)
57. Stoica, P., Babu, P., Li, J.: SPICE: a sparse covariance-based estimation method for array processing. *IEEE Trans. Signal Process.* **59**(2), 629–638 (2011)
58. Stoica, P., Moses, R.L.: *Spectral analysis of signals* (2005)
59. Tse, D., Viswanath, P.: *Fundamentals of Wireless Communication*. Cambridge University Press, Cambridge (2005)
60. Van Vleck, J.H., Middleton, D.: The spectrum of clipped noise. *Proc. IEEE* **54**(1), 2–19 (1966)
61. Vershynin, R.: *High-Dimensional Probability: An Introduction with Applications in Data Science*, vol. 47. Cambridge University Press, Cambridge (2018)