Guaranteed Private Communication with Secret Block Structure

Maxime Ferreira Da Costa, Jianxiu Li, Urbashi Mitra

Guaranteed Private Communication with Secret Block Structure

Maxime Ferreira Da Costa, Jianxiu Li, and Urbashi Mitra

***Abstract*—A novel private communication framework is proposed where privacy is induced by transmitting over channel instances of linear inverse problems that are identifiable to the legitimate receiver, but unidentifiable to an eavesdropper. The gap in identifiability is created in the framework by leveraging secret knowledge between the transmitter and the legitimate receiver. Specifically, the case where the legitimate receiver harnesses a secret block structure to decode a transmitted block-sparse message from underdetermined linear measurements in conditions where classical compressed sensing would provably fail is examined. The applicability of the proposed scheme to practical multiple access wireless communication systems is discussed. The protocol's privacy is studied under a single transmission, and under multiple transmissions without refreshing the secret block structure. It is shown that, under a specific scaling of the channel dimensions and transmission parameters, the eavesdropper can attempt to overhear the block structure from the fourth-order moments of the channel output. Computation of a statistical lower bound, suggests that the proposed fourth-order moment secret block estimation strategy is near optimal. The performance of a spectral clustering algorithm is studied to that end, defining scaling laws on the lifespan of the secret key before the communication is compromised. Finally, numerical experiments corroborating the theoretical findings are conducted.**

***Index Terms*—Private communication, inverse problems, structured compressed sensing, moment method.**

I. Introduction

While communication privacy is often ensured at higher network layers [1]–[3], and can be achieved via cryptographic means, there are new methods in *physical layer security* [4], which can leverage the structural properties of a communication channel to generate privacy. Physical layer privacy can strengthen security in modern data exchange protocols, such as next-generation wireless systems, the Internet of Things, and satellite constellations. Physical layer security offers numerous complementary guarantees to usual cryptography: It can protect users' identities, physical locations, or even conceal the existence of a communication to an eavesdropper; and can be implemented opportunistically over wireless channels with no or little computational overhead. There is interest in realizing the theoretical promises of physical layer security in realistic systems [5].

Traditional physical layer privacy strategies involve the use of artificial noise [6]–[9]. The noise can be either injected into the nullspace of channel state information (CSI) and mitigated by exploiting CSI, or directly injected noise into the transmitted message and resolved by the legitimate receiver side by exploiting a secret key [10], [11]. Other privacy schemes involve random and adversarial beamforming design [12], [13], or the injection of fake paths over geometric channels to diminish the capability of an eavesdropper to distinguish between true and fake paths and challenge the estimation of CSI [14] by an eavesdropper.

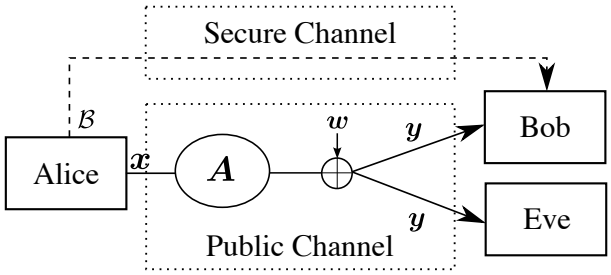

Figure 1. Communication model with secure channel.

The previously mentioned physical layer security schemes induce privacy by performing a linear action on the transmitted message that is statistically hard to invert without additional knowledge. In a related fashion, the compressed sensing framework [15] assumes a non-linear prior on the input message and has been exploited as a means to ensure privacy [16]. If the sensing matrix is kept secret to an eavesdropper, perfect secrecy can be guaranteed in the information-theoretic sense [17] under restrictive conditions [18]. Typical sensing matrices are functions of the CSI. The computational secrecy of this approach has also been investigated [19], [20], restricting Eve's ability to recover the encoded message via a polynomial time algorithm.

Motivated by applications to multiple access wireless systems, we focus here, instead, on a novel model where the sensing matrix (*e.g.* the channel matrix) is imposed by the environment and is *not* under the control of the transmitter. Privacy is achieved by sharing an additional structure on said message with the legitimate receiver, easing the decoding of the message [21]. From the eavesdropper's perspective, the decoding amounts to solving a bilinear inverse problem, which is known to demand much more stringent assumptions to be identifiable [22]–[27]. Thus, statistical hardness is exploited to provide privacy.

A. Linear Inverse Problem Based Privacy

We consider the classical secret communication problem with side information: A transmitter (Alice) wishes to privately

M. Ferreira Da Costa is with the Laboratory of Signals and Systems (L2S) at CentraleSupélec, Université Paris–Saclay. The work of M. Ferreira Da Costa is supported in part by ANR-20-IDEES-0002. Email: maxime.ferreira@centralesupelec.fr.

J. Li and U. Mitra are with the Viterbi School of Engineering, University of Southern California. The work of J. Li and U. Mitra is supported in part by the USC + Amazon Center on Secure and Trusted Machine Learning, NSF CCF-1817200, DOE DE-SC0021417, Swedish Research Council 2018-04359, NSF CCF-2008927, NSF CCF-2200221, ONR 503400-78050, and ONR N00014-15-1-2550. Emails: jianxiul@usc.edu, ubli@usc.edu.

transmit a vector $\boldsymbol{x} \in \mathbb{R}^n$ to a legitimate receiver (Bob) over a public channel. The noisy channel output $\boldsymbol{y} = f(\boldsymbol{x}) + \boldsymbol{w}$, with noise $\boldsymbol{w}$ being received by Bob and the eavesdropper (Eve). To achieve privacy and prevent Eve from recovering the message $\boldsymbol{x}$, Alice and Bob may communicate a low information rate signal over a secure channel inaccessible to Eve.

In the proposed setting, the effect of the channel is assumed to be linear and modeled by a "fat" matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, with $m < n$ so that $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w}$, where $\boldsymbol{w}$ is white Gaussian noise $\boldsymbol{w} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_m)$. The matrix $\boldsymbol{A}$ is imposed by the *environment* and is assumed to be known by Bob and Eve. For the purposes of our analysis, we will assume $\boldsymbol{A}$ to satisfy certain incoherence properties, which are detailed in the sequel. Finally, we assume that Eve is aware of the communication protocol established by Alice. The overall communication model is depicted in Fig. 1.

To ensure privacy, Alice, who designs the message $\boldsymbol{x}$ and the side information, must ensure two properties. First, Bob must be able to provably recover $\boldsymbol{x}$ from the observation $\boldsymbol{y}$ via the side information from the secure channel. Second, Eve cannot provably recover $\boldsymbol{x}$ without knowing the side information. Thus, Alice is left to design an inverse problem that is *identifiable* to Bob, but *unidentifiable* to Eve. These goals can typically be jointly achieved by imposing an additional structure on $\boldsymbol{x}$, and by privately sharing this structure over the secure channel. For practicality, this structure must be comprised of a small number of bits, and reusable over multiple transmissions. Building onto our prior work [28], we propose that Alice shares with Bob a secret block structure, and encodes her message in the form of a block-sparse signal whose support follows this secret structure. Harnessing a block-sparse prior to recover signals through underdetermined linear measurements has been extensively shown to allow exact recovery in conditions where classical compressed sensing would provably fail [21], [29]–[31]. We exploit this result to propose a novel private communication framework where secrecy is achieved by transmitting instances of an unidentifiable compressed sensing problem over a public channel. Furthermore, if the secret block structure is reused for multiple transmissions, we show that Eve can eavesdrop on the block structure by a spectral clustering technique applied on fourth-order empirical moments of the probability distribution at the channel output, and the trade-off between key-reuse and secrecy is discussed. Spectral clustering has been considered with significant success as a fast and robust method to recover low-dimensional structures in high-dimensional datasets. It has been applied, for instance, to recovering partitions and cliques in high-dimensional graphs [32], and in the context of supervised classification in machine learning [33].

The proposed signaling scheme is motivated by its applicability to modern multi-user wireless communication protocols. As an example, we assume an uplink scenario with r many transmitters sending within a symbol interval a message $\boldsymbol{u}_q \in \mathbb{R}^d$ using a precoding scheme $\boldsymbol{S}_q \in \mathbb{R}^{n \times d}$ through a linear channel $\boldsymbol{H}_q \in \mathbb{R}^{n \times n}$ that is imposed by the environment. The received message at the base station classically reads $\boldsymbol{y} = \sum_{q=1}^{r} \boldsymbol{H}_q \boldsymbol{S}_q \boldsymbol{u}_q + \boldsymbol{w}$. When the channel users parsimoniously transmit at a given symbol interval, *i.e.* when a random fraction of users remain inactive, the channel input can be modeled with a group-sparse prior. If this prior is only known by the legitimate base station (Bob), the relative identifiability of block-sparse signals versus unstructured sparse signals can be exploited to induce privacy against an eavesdropper. We note two practical schemes where this framework is applicable:

1) In *overloaded CDMA communications*, the transmitters rely on unique sequences $\{\boldsymbol{S}_1, \cdots, \boldsymbol{S}_r\}$, known to the base station, to spread the messages onto a larger dimension space before transmission [34], [35]. Sparse coded multiple access schemes have been considered to improve user detection when the system is overloaded [36], and adapted coding sequences are proposed in [37], [38]. However, the privacy benefits of overloading have not yet been considered in that context.
2) In *massive MIMO communications*, the number of identifiable spatial streams is equal to the number of receive antennas. If the transmitter has more antennas than the receiver, she had intermittently activate sub-groups of antennas according to a pattern shared with the receiver and transmits on the active sub-groups at each symbol interval, at the price of a reduced bit-rate. Such MIMO systems have been considered to reduce implementation cost [39] or improve spectral efficacy [40], [41].

B. Contributions and Paper Organization

We build upon our prior work [28] and present an improved eavesdropping scheme based on fourth moments with full proofs and numerical simulations. Computation of a statistical lower bound suggests that the improved eavesdropping scheme is asymptotically near-optimal. In Section II, we propose a novel communication protocol that leverages the advantageous recoverability of block-sparse signals to ensure privacy. We provide the encoding and decoding strategies of Alice and Bob, respectively. In our design, Alice transmits secretly to Bob a block structure and uses this structure to encode her message, which can be done at a very low transmission rate, while the channel matrix $\boldsymbol{A}$ *cannot* be designed by Alice and is provided by nature. To the authors' knowledge, the proposed protocol is the first linear inverse problem-based privacy method that does not require the matrix $\boldsymbol{A}$ to be secretly shared. Furthermore, Corollary 3 guarantees that Alice can adjust the block length and the sparsity level of the message she transmits so that the transmission is provably identifiable for Bob and unidentifiable to Eve as the signal length increases. In Section III, we consider the possibility of Eve recovering the secret block structure from the observation of *multiple* snapshots of the observation $\{\boldsymbol{y}_\ell\}$ that Alice has generated with the same block structure $\mathcal{B}$. We show in Proposition 9 that, depending on Alice's choice of the block length and sparsity level, it is possible to extract $\mathcal{B}$ from the fourth-order moments of the observation and propose an eavesdropping algorithm to that end. We investigate the case of a finite number of snapshots and derive an upper bound on the rate at which Alice must generate a new $\mathcal{B}$ to prevent Eve from deciphering Bob's messages.

We present numerical results that validate our theoretical findings in Section IV. In Section V, a conclusion is drawn, and further research directions are discussed.

C. Notations

Vectors of $\mathbb{R}^n$ and matrices of $\mathbb{R}^{n_1 \times n_2}$ are denoted by boldface letters $\boldsymbol{a}$ and capital boldface letters $\boldsymbol{A}$, respectively. The entry (i,j) of a matrix $\boldsymbol{A}$ is written as $a_{i,j}$. The matrix norms $\|\boldsymbol{M}\|_2$, $\|\boldsymbol{M}\|_{\mathsf{F}}$, and $\|\boldsymbol{M}\|_{\max}$ refer to the spectral norm, the Frobenius norm, and the maximal absolute value of the entries in $\boldsymbol{M}$, respectively. Given a positive semi-definite matrix $\boldsymbol{M}$, we write $\lambda_{\min}(\boldsymbol{M})$ and $\lambda_r(\boldsymbol{M})$ as its smallest eigenvalue, and rth-largest eigenvalue (with multiplicity), respectively. The Hadamard product between two matrices $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ is denoted as $\boldsymbol{M}_1 \odot \boldsymbol{M}_2$. We write by $\boldsymbol{I}_n$ the identity matrix and by $\boldsymbol{J}_n$ the all-one matrix in dimension $n \times n$. Given a random vector $\boldsymbol{z} \in \mathbb{R}^n$, we denote by $\boldsymbol{\Sigma}_{\boldsymbol{z}} \in \mathbb{R}^{n \times n}$ its covariance matrix. A block structure over $\mathbb{R}^n$ into r blocks is described by a mapping $\mathcal{B} : [1, \dots, n] \to [1, \dots, r]$, and is associated with the indicator matrix of $\boldsymbol{B} \in \mathbb{R}^{n \times n}$ defined by

$$b_{i,j} = \begin{cases} 1 & \text{if } \mathcal{B}(i) = \mathcal{B}(j); \\ 0 & \text{if } \mathcal{B}(i) \neq \mathcal{B}(j). \end{cases} \tag{1}$$

We denote by $\boldsymbol{x}[q]$ the subvector of $\boldsymbol{x}$ with entries x_i ensuring $\mathcal{B}(i) = q$. The "block-ℓ_0-norm" of a vector $\boldsymbol{x}$ is defined as $\|\boldsymbol{x}\|_{\mathcal{B},0} = \sum_{q=1}^{r} \mathbf{1}_{\boldsymbol{x}[q] \neq \mathbf{0}}$ and counts the number blocks in $\boldsymbol{x}$ that are not exactly equal to $\mathbf{0}$. For two functions f and g, we use the Landau notation $f = o(g)$ to denote that the ratio $\frac{f(t)}{g(t)}$ tends to 0 as $t \to \infty$.

II. Privacy with Block Sparsity

A. Alice's encoding

In the proposed protocol, Alice constructs her message $\boldsymbol{x}$ as follows. Given the knowledge of the channel dimension, Alice initializes the communication by randomly selecting a block structure $\mathcal{B} : [1, \dots, n] \to [1, \dots, r]$. Alice sends this structure to Bob over the secret channel. We highlight that this exchange only requires $n \log_2(r)$ bits of information which is significantly less than schemes relying on exchanging the entire matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ (mn infinite precision numbers). Although not required in practice, we assume for simplicity that the r blocks have equal block size d, *i.e.* $n = r \cdot d$. Next, Alice selects a probability of block activation $p \in \left[0, \frac{1}{2}\right]$, where $p \leq \frac{1}{2}$ is assumed for convenience in the analysis, and encodes her message in a block-sparse vector $\boldsymbol{x}$. In the sequel, we assume that $\boldsymbol{x}$ is distributed according to a block Bernouilli–Gaussian distribution such that

$$\boldsymbol{x}[q] = \begin{cases} \mathbf{0}_d & \text{w.p. } 1-p \\ \boldsymbol{z}[q] & \text{w.p. } p, \end{cases} \tag{2}$$

where $\boldsymbol{z}[q] \sim \mathcal{N}(\mathbf{0}_d, \boldsymbol{I}_d)$ is a random i.i.d. standard Gaussian vector of dimension d. A visualization of the block sparsity encoding is provided in Figure 2.

B. Bob's decoding

At the public channel output, Bob receives a vector $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w}$, and leverages $\mathcal{B}$ that Alice securely sent to recover the ground truth message $\boldsymbol{x}$. To do so, Bob formulates the block-compressed sensing problem:

$$\widehat{\boldsymbol{x}}_B = \underset{\boldsymbol{x} \in \mathbb{R}^n}{\arg\min} \|\boldsymbol{x}\|_{\mathcal{B},0} \text{ such that } \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2 \leq \epsilon, \tag{3}$$

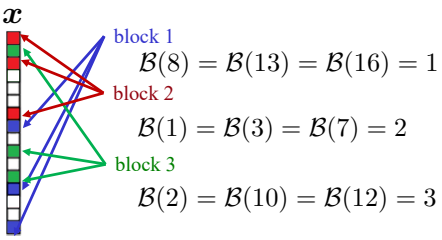

Figure 2. Example of block sparse encoding.

where $\epsilon > 0$ is a parameter that scales proportionally with the standard deviation of the noise $\|\boldsymbol{w}\|_2$. Harnessing a block-sparse prior in compressed sensing has been extensively shown in the literature to enhance the identifiability of (3) and to allow an exact reconstruction of the message with much fewer measurements than classical compressed sensing [29], [31]. However, directly solving (3) remains NP-hard in the general case, due to the combinatorics inherent to the minimization of $\|\boldsymbol{x}\|_{\mathcal{B},0}$. Thus, Bob computes, instead, an estimate of $\widehat{\boldsymbol{x}}_B$ using a polynomial time algorithm of his choice. Among the many addressed algorithms proposed in the literature, Block Matching Pursuit (Block MP) [42], Block Iterative Harding Thresholding (Block IHT), Block Basis Pursuit (Block BP) [43] or block-based CoSaMP [21], have been shown to have provable performance guarantees.

In the sequel, we denote $\beta = \frac{m}{np}$ as the *redundancy parameter*, defined as the ratio between the number of measurements at the channel output and the *expected* number of non-zero entries in the block-sparse input vector $\boldsymbol{x}$. We remark that $\beta \geq 1$ is obviously needed for successful decoding of the message. In fact, asymptotic phase transitions for the success of greedy algorithms to recover the block-sparse ground truth have been studied in the literature [21]. Proposition 1 reinterprets this result in terms of the parameter β, the block-length d, and the transmission parameter p in the asymptotics $n \to \infty$.

Proposition 1 (Success of Bob's decoding)**.** *Suppose that $\boldsymbol{A}$ is a matrix with i.i.d. random Gaussian entries and assume a noise-free environment $\boldsymbol{w} = \mathbf{0}$. If*

$$\log\left(\frac{1}{p}\right) = o\left(\frac{d}{\log(d)}\right) \qquad \textit{and} \qquad \beta \to \infty \tag{4}$$

in the limit where $n \to \infty$, then Bob can stably recover $\boldsymbol{x}$ asymptotically almost surely.

Additionally, denoising bounds on the estimate of the input vector $\boldsymbol{x}$ are provided in the presence of noise [21].

C. Privacy Guarantees under a Single Snapshot

If only one snapshot $\boldsymbol{y}$ is observed, it is impossible for Eve to reliably infer $\mathcal{B}$, which remains ambiguous even with perfect knowledge of $\boldsymbol{x}$. Therefore, from her perspective, the best possible approach consists of attempting to recover $\boldsymbol{x}$ without leveraging the existence of a latent block structure in the message. This amounts to solving a *classical* compressed sensing program

$$\widehat{\boldsymbol{x}}_E = \underset{\boldsymbol{x} \in \mathbb{R}^n}{\arg\min} \|\boldsymbol{x}\|_0 \text{ such that } \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2 \leq \epsilon. \tag{5}$$

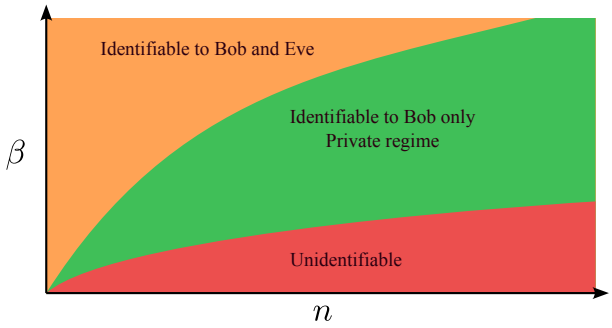

Figure 3. Regions of (non)identifiability for Eve and Bob in the single snapshot case for a block-length $d = n\log^{-\delta}(n)$ with $\delta > 0$.

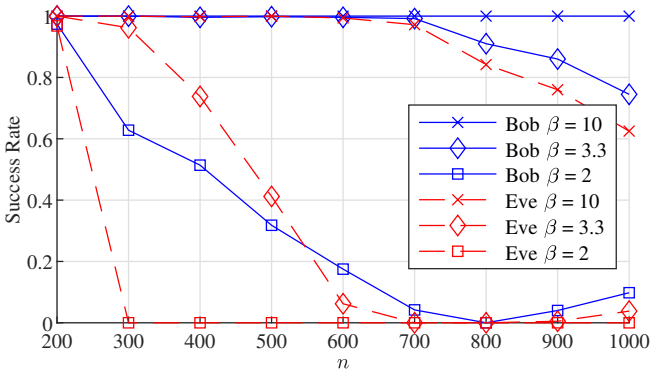

Figure 4. Success Rate of Bob and Eve to recover $\boldsymbol{x}$ for different values of β in absence of noise. The parameters are set to $m = 200$, $r = 20$. The results are averaged over 500 trials.

The identifiability condition $\boldsymbol{x} = \widehat{\boldsymbol{x}}_E$ of (5) is well-understood to be related to the Restricted Isometry Property (RIP) of the measurement operator [44]. In the case of a Gaussian matrix $\boldsymbol{A}$, the following proposition links the asymptotic failure of (5) to a function of the model's parameters, translating results in [45] to our context.

Proposition 2 (Failure of Eve's decoding [45])**.** *Suppose that $\boldsymbol{A}$ is a matrix with i.i.d. random Gaussian entries and assume a noise-free environment $\boldsymbol{w} = \boldsymbol{0}$. Then if*

$$\beta = o\left(\log\left(\frac{1}{p}\right)\right) \tag{6}$$

holds in the limit where $n \to \infty$, then the solution $\widehat{\boldsymbol{x}}_E$ of (5) is different from $\boldsymbol{x}$ with overwhelming probability.

Altogether, Propositions 1 and 2 suggest that, given the dimensions m and n of $\boldsymbol{A}$, Alice can select the parameters β and d so that (4) and (6) are jointly satisfied, which is summarised in the sequel,

Corollary 3 (Single snapshot privacy)**.** *If Alice selects a diverging redundancy parameter $\beta \to \infty$ with $\beta = o\left(\log\left(\frac{1}{p}\right)\right)$ and $\log\left(\frac{1}{p}\right) = o\left(\frac{d}{\log(d)}\right)$, then the protocol is asymptotically private to the exchange of a single message in the limit $n \to +\infty$.*

As an example, if Alice allows the block length to grow with the channel input n at a rate $d \sim n\log^{-\delta}(n)$ for some $\delta > 0$ and scales p as $\log\left(\frac{1}{p}\right) = \beta\log(n)$, then the region $\log(n) \ll \beta \ll n\log^{-\delta-1}(n)$ will be asymptotically private. Fig. 3 depicts the three different communication regions under this block-length assumption. This result suggests that parameter intervals for the private regime are increasing with the channel length. This highlights that the proposed communication protocol benefits from larger channel dimensions. Larger channel dimensions can be realized in practice by selecting longer spreading sequences in CDMA systems or increasing the number of antennae in MIMO systems.

Fig. 4 shows the success rate of Bob and Eve to recover $\boldsymbol{x}$ via the Block-BP and BP algorithms, respectively, for different values of the ratio β. We see that as β gets small, the success rates for both Bob and Eve diminish. This is intuitive as $\beta = \frac{m}{pn}$ measures the number of observations relative to the number of active components. The lower the activity level, the fewer non-zero signals that are sent. However, it is also clear that given $m = 200$, there is a sweet spot at $\beta = 3.3$, where Bob achieves good performance while Eve does not.

III. Eavesdropping via Higher Order Moments

A. Structure of the Moments

In effect, our results above are for a single-pad key, *i.e.* a new block structure is created for each message to be sent [18]. To reduce the usage of the secure channel, we want to understand the reusability of $\mathcal{B}$ in transmitting several independent signals $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\}$. In this scenario, if Eve can acquire multiple snapshots of observation $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L\}$ given by $\boldsymbol{y}_\ell = \boldsymbol{A}\boldsymbol{x}_\ell + \boldsymbol{w}_\ell$, $\ell = 1, \ldots, L$, and under the knowledge of the prior distribution (2) of $\boldsymbol{x}$, she can attempt to gain statistical information about $\mathcal{B}$ without having to reconstruct the messages by studying the posterior distribution of $\boldsymbol{y}$. In particular, we observe that given our block signaling, the mean $\mathrm{E}[\boldsymbol{x}] = \boldsymbol{0}_N$ and covariance $\boldsymbol{\Sigma}_{\boldsymbol{x}} = p\boldsymbol{I}_n$ of $\boldsymbol{x}$ carry no information about the block structure $\mathcal{B}$. However, the even fourth-order moments of $\boldsymbol{x}$ **do** provide information about the block structure, $\mathcal{B}$, as seen below:

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{x}\odot\boldsymbol{x}}(i,j) &= \mathrm{E}[x_i^2 x_j^2] - \mathrm{E}[x_i^2]\mathrm{E}[x_j^2] \\ &= \begin{cases} 3p - p^2 & \text{if } i = j \\ p - p^2 & \text{if } \mathcal{B}(i) = \mathcal{B}(j) \text{ and } i \neq j \\ 0 & \text{if } \mathcal{B}(i) \neq \mathcal{B}(j), \end{cases}\end{aligned} \tag{7}$$

Additionally, as the odd fourth-order moments of $\boldsymbol{x}$ equal zero, the terms in (7) are the moments of smallest order containing information about the block structure $\mathcal{B}$. As the number of samples that is necessary to estimate moments increases with their order, Eve can restrict herself to the study of the covariance $\boldsymbol{\Sigma}_{\boldsymbol{z}}$ of the vector $\boldsymbol{z} = \left(\boldsymbol{A}^\mathsf{T}\boldsymbol{y}\right) \odot \left(\boldsymbol{A}^\mathsf{T}\boldsymbol{y}\right)$ in attempt to eavesdrop $\mathcal{B}$ from the observation of the channel output. Given this observation, understanding the reusability of the block structure is equivalent to understanding Eve's capability to learn these fourth moments.

For notational convenience, let $\boldsymbol{M} = \boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}$ and $\boldsymbol{P} = \boldsymbol{M} \odot \boldsymbol{M}$. Moreover, we define the matrices $\boldsymbol{E}_{\mathcal{B}}$, $\boldsymbol{F}$, and $\boldsymbol{G}$ where each component is given, respectively, by,

$$\boldsymbol{E}_{\mathcal{B}}(i,j) = \sum_{k} \sum_{\substack{k' \neq k \\ \mathcal{B}(k')=\mathcal{B}(k)}} m_{i,k} m_{i,k'} m_{j,k} m_{j,k'}, \tag{8a}$$

$$\boldsymbol{F}(i,j) = \sum_{k} \sum_{k' \neq k} m_{i,k} m_{i,k'} m_{j,k} m_{j,k'}, \tag{8b}$$

$$\boldsymbol{G}(i,j) = \sum_{k} \sum_{k' \neq k} a_{k,i} a_{k,j} a_{k',i} a_{k',j}. \tag{8c}$$

The next proposition, whose proof is presented in Appendix A, gives an expression of the covariance $\boldsymbol{\Sigma}_{\boldsymbol{v}}$ as a function of the matrices $\boldsymbol{E}_{\mathcal{B}}$, $\boldsymbol{F}$, and of the block structure matrix $\boldsymbol{B}$.

Proposition 4. *Let* $\boldsymbol{z} = \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{y}\right) \odot \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{y}\right)$. *If* $\boldsymbol{x}$ *is drawn according to* (2) *then the covariance* $\boldsymbol{\Sigma}_{\boldsymbol{z}}$ *of* $\boldsymbol{z}$ *is given by*

$$\boldsymbol{\Sigma}_{\boldsymbol{z}} = p(1-p)\boldsymbol{P}\boldsymbol{B}\boldsymbol{P} + 2p\boldsymbol{E}_{\mathcal{B}} + \boldsymbol{C} \tag{9}$$

where the matrix $\boldsymbol{C}$ *is given by*

$$\boldsymbol{C} = 2p\boldsymbol{P}^2 + 2p^2\boldsymbol{F} + 2p\sigma^2\boldsymbol{M}^2 \odot \boldsymbol{M} + 2\sigma^4\left((\boldsymbol{A}\odot\boldsymbol{A})^{\mathsf{T}}(\boldsymbol{A}\odot\boldsymbol{A}) + \boldsymbol{G}\right) \tag{10}$$

Proposition 4 proposes an decomposition of the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{z}}$ into two main terms:

1) The term $p(1-p)\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}+2p\boldsymbol{E}_{\mathcal{B}}$, which captures properties of the block structure $\mathcal{B}$.
2) The term $\boldsymbol{C}$ which only depends on the block activation probability p, on the channel $\boldsymbol{A}$, and on the noise power σ^2.

In the sequel, we propose a strategy by which to exploit this structure to learn $\mathcal{B}$.

B. Reconstruction via Spectral Clustering

In this section, we propose a provable spectral clustering-based algorithm for Eve to infer the block structure $\mathcal{B}$ from observing a finite number of snapshots L. In our setting, the block structure matrix $\boldsymbol{B}$ that Eve aims to recover has a rank equal to the number of blocks r, which is assumed to be much smaller than the ambient signal dimension n. As a result, spectral clustering's reliability can be anticipated for inferring the low-dimensional block structure.

We first review Algorithm 1. This is a straightforward algorithm that employs the matrix $\boldsymbol{Y}$, whose columns $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L\}$ are sampled from the channel output, to determine an estimate $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}}$ of the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{z}}$ in Equation (9). This equation is consecutively "inverted", yielding an estimator $\widetilde{\boldsymbol{B}}$ of the indicator matrix $\boldsymbol{B}$. As the r-leading eigenvectors of the indicator matrix $\boldsymbol{B}$ identify exactly the block structure $\mathcal{B}$, an estimate $\widehat{\mathcal{B}}$ of the true block structure $\mathcal{B}$ is constructed by clustering the rows of the r leading eigenvectors of the matrix $\widetilde{\boldsymbol{B}}$, following a K-means-type procedure described by Algorithm 2.

The rest of this section is dedicated to the theoretical analysis of the estimation procedure proposed by Algorithm 1. Under incoherence assumptions on the channel matrix $\boldsymbol{A}$, we first assess Eve's capability to eavesdrop $\mathcal{B}$ using Algorithm 2 when she has access to infinitely many channel outputs $\boldsymbol{y}_\ell$, and thus to the ground truth covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{z}}$. Then, we consider the case where Eve observes a finite number of channel outputs.

Algorithm 1 Eavesdropping by Spectral Clustering

function MOMENTMETHOD($\boldsymbol{Y} \in \mathbb{R}^{m\times L}, \boldsymbol{A} \in \mathbb{R}^{m\times n}, p, r$)
  $\boldsymbol{Z} \leftarrow \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{Y}\right) \odot \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{Y}\right)$
  $\overline{\boldsymbol{z}} \leftarrow p\,\mathrm{diag}\left(\left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}\right)^2\right) + \sigma^2\,\mathrm{diag}\left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}\right)$
  $\gamma \leftarrow \frac{2(d-1)}{m} + \frac{(n-2)(d-1)}{m^2}$ with $d = \frac{n}{r}$
  $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} \leftarrow \frac{1}{L}\sum_{\ell=1}^{L}(\boldsymbol{z}_\ell - \overline{\boldsymbol{z}})(\boldsymbol{z}_\ell - \overline{\boldsymbol{z}})^{\mathsf{T}}$
  $\boldsymbol{K} \leftarrow \boldsymbol{C} + 2p\gamma\boldsymbol{I}_n$ ▷ With $\boldsymbol{C}$ as in (10)
  $\widetilde{\boldsymbol{B}} \leftarrow (p(1-p))^{-1}\boldsymbol{P}^{-1}(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \boldsymbol{K})\boldsymbol{P}^{-1}$
  $\widetilde{\boldsymbol{U}} \leftarrow$ the r dominant eigenvectors of $\widetilde{\boldsymbol{B}}$
  $\widehat{\mathcal{B}} \leftarrow$ GREEDYKMEANS $(\widetilde{\boldsymbol{U}}, r)$
  return $\widehat{\mathcal{B}}$
end function

Algorithm 2 Greedy implementation of K-means

function GREEDYKMEANS($\widetilde{\boldsymbol{U}} \in \mathbb{R}^{n\times r}$)
  $\widetilde{r} \leftarrow 0$
  for $j = 1 \ldots n$ **do**
    if $\min_{q\in\{1,\ldots,\widetilde{r}\}} \|\widetilde{\boldsymbol{c}}_q - \widetilde{\boldsymbol{u}}_j\|_2 < \frac{1}{\sqrt{2d}}$ **then**
      $\widehat{\mathcal{B}}(j) \leftarrow \arg\min_q \|\widetilde{\boldsymbol{c}}_q - \widetilde{\boldsymbol{u}}_j\|_2$
      ▷ Assign jth entry to cluster with closest centroid
      $\widetilde{\boldsymbol{c}}_q \leftarrow \mathrm{mean}\left\{\widetilde{\boldsymbol{u}}_q;\ q \le j \text{ and } \widehat{\mathcal{B}}(q) = \widehat{\mathcal{B}}(j)\right\}$
      ▷ Update the centroid
    else
      $\widetilde{r} \leftarrow \widetilde{r} + 1$ ▷ Create a new cluster
      $\widetilde{\boldsymbol{c}}_{\widetilde{r}} \leftarrow \widetilde{\boldsymbol{u}}_j$ ▷ Assign jth entry to new cluster
    end if
  end for
  return $\widehat{\mathcal{B}}$
end function

C. Conditions for exact clustering

Eve's ability to estimate $\widetilde{\boldsymbol{B}}$ sufficiently close to the indicator matrix $\boldsymbol{B}$ is a determining factor in her attempt to recover $\mathcal{B}$. When the spectral distance $\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2$ is small enough, the eigenvectors of $\widetilde{\boldsymbol{B}}$ will align with those of $\boldsymbol{B}$, and the block structure will become identifiable by spectral clustering. We start the theoretical derivations by finding in Proposition 5, a sufficient condition on $\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2$ under which the K-means clustering procedure described by Algorithm 2 returns the exactly the secret block structure $\mathcal{B}$.

Proposition 5 (Exact clustering)**.** *Assume* $\boldsymbol{B} \in \{0,1\}^{n\times n}$ *is the indicator matrix of a block structure* $\mathcal{B}$ *with* $d \ge 2$. *Then, for any* $\widetilde{\boldsymbol{B}} \in \mathbb{R}^{n\times n}$ *with* $\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2 < \frac{\sqrt{2d}}{8}$, *the output of Algorithm 1 applied the matrix* $\widetilde{\boldsymbol{U}} \in \mathbb{R}^{n\times r}$ *that is composed by the* r *leading eigenvector of* $\widetilde{\boldsymbol{B}}$ *exactly recovers the block structure, i.e.* $\widehat{\mathcal{B}} = \mathcal{B}$.

Proof. First, it is easy to confirm from Equation (1) that $\|\boldsymbol{B}\|_2 = d$. As both $\boldsymbol{B}$ and $\widetilde{\boldsymbol{B}}$ are Hermitian matrices, they have

orthogonal bases of eigenvectors. We write $\boldsymbol{U}, \widetilde{\boldsymbol{U}} \in \mathbb{R}^{n\times r}$ the matrices whose columns are the eigenvectors corresponding to the r leading eigenvalues of $\boldsymbol{B}$ and $\widetilde{\boldsymbol{B}}$, respectively. By the Davis-Kahan eigenvector perturbation theorem [46], there exists an orthogonal matrix $\boldsymbol{O} \in \mathbb{R}^{r\times r}$ such that

$$\begin{aligned}\left\|\widetilde{\boldsymbol{U}} - \boldsymbol{U}\boldsymbol{O}\right\|_{\mathsf{F}} \leq & \frac{\sqrt{2}\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2}{\lambda_r(\boldsymbol{B}) - \left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2} \\ = & \frac{\sqrt{2}\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2}{d - \left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2} \\ \leq & \frac{2}{d}\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2 < \frac{\sqrt{2}}{4\sqrt{d}}, \end{aligned} \tag{11}$$

where we used $\frac{t}{1-t} \leq \sqrt{2}t$ when $0 \leq t \leq \frac{1}{4}$ in the second inequality. Next, we denote by $\boldsymbol{u}_j$ and $\widetilde{\boldsymbol{u}}_j$ the jth columns of the matrices $\boldsymbol{U}^{\mathsf{T}}$ and $\widetilde{\boldsymbol{U}}^{\mathsf{T}}$, respectively. From the expression (1) of $\boldsymbol{B}$, the vector $\boldsymbol{u}_j$ indicates the block in which the jth element belongs, more precisely we have

$$\boldsymbol{u}_j(q) = \begin{cases} \frac{1}{\sqrt{d}} & \text{if } \mathcal{B}(j) = q \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Suppose that $\mathcal{B}(j) = q$ and let $\boldsymbol{c}_q = \boldsymbol{c}_{\mathcal{B}(j)} = \boldsymbol{O}^{\mathsf{T}}\boldsymbol{u}_j$, which represent the rotated true centroid of the qth block. Equation (11) implies that $\|\widetilde{\boldsymbol{u}}_j - \boldsymbol{c}_q\|_2 < \frac{\sqrt{2}}{4\sqrt{d}}$. Therefore, this also implies that the estimated centroid of the qth block $\widetilde{\boldsymbol{c}}_q$ satisfies $\|\widetilde{\boldsymbol{c}}_q - \boldsymbol{c}_q\|_2 < \frac{\sqrt{2}}{4\sqrt{d}}$ at each step of the algorithm. From the triangle inequality, we have,

$$\|\widetilde{\boldsymbol{u}}_j - \boldsymbol{c}_q\|_2 \leq \|\widetilde{\boldsymbol{u}}_j - \widetilde{\boldsymbol{c}}_q\|_2 + \|\widetilde{\boldsymbol{c}}_q - \boldsymbol{c}_q\|_2 < \frac{1}{\sqrt{2d}}. \tag{13}$$

By orthogonality of the eigenvectors $\boldsymbol{u}_q$ and $\boldsymbol{u}_{q'}$, we also have that $\|\boldsymbol{c}_{q'} - \boldsymbol{c}_q\|_2 = \left\|\boldsymbol{O}^{\mathsf{T}}(\boldsymbol{u}_{q'} - \boldsymbol{u}_q)\right\|_2 = \sqrt{\frac{2}{d}}$ for any $q \neq q'$. Hence if $q' \neq \mathcal{B}(j)$ we may write

$$\begin{aligned}\|\widetilde{\boldsymbol{u}}_j - \widetilde{\boldsymbol{c}}_{q'}\|_2 &= \|\widetilde{\boldsymbol{u}}_j - \boldsymbol{c}_q + \boldsymbol{c}_q - \widetilde{\boldsymbol{c}}_{q'}\|_2 \\ &\geq \|\boldsymbol{c}_q - \widetilde{\boldsymbol{c}}_{q'}\|_2 - \|\widetilde{\boldsymbol{u}}_j - \boldsymbol{c}_q\|_2 \\ &> \sqrt{\frac{2}{d}} - \frac{\sqrt{2}}{2\sqrt{d}} = \frac{1}{\sqrt{2d}}. \end{aligned} \tag{14}$$

Hence $\|\widetilde{\boldsymbol{u}}_j - \widetilde{\boldsymbol{c}}_q\|_2 < \|\widetilde{\boldsymbol{u}}_j - \widetilde{\boldsymbol{c}}_{q'}\|_2$ for any $q \neq q'$, and we conclude with (13) and (14) that at the jth iteration, Algorithm 2 associate $\widehat{\mathcal{B}}(j) = \mathcal{B}(j) = q$ if there was an element in $\{1, \ldots, j-1\}$ that is in the qth cluster, otherwise associate j to a new cluster q. This results in $\widehat{\mathcal{B}} = \mathcal{B}$ at the algorithm's output. □

D. Asymptotic vulnerability

In this subsection, we assume that Eve can sample infinitely many channel output $\{\boldsymbol{y}_\ell\}$ that have been produced with *the same* secret block structure $\mathcal{B}$, and wish to understand Eve's capability to recover $\mathcal{B}$ from Algorithm 1. Of particular interest, Eve knows in this setting the probability distribution $\boldsymbol{y}$, and consequently has access to the ground truth covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{z}}$ given in (9). In the additional pessimistic hypothesis where Eve knows the activation probability [1] p, the block length d, the channel matrix $\boldsymbol{A}$, and the statistics of the noise $\boldsymbol{w}$, she can compute the matrices $\boldsymbol{P}$ and $\boldsymbol{C}$ in Proposition 4, and the constant γ defined in the fourth step of Algorithm 1. Hence, she can formulate the estimate $\widetilde{\boldsymbol{B}}$ of the block structure $\boldsymbol{B}$ as

$$\begin{aligned}\widetilde{\boldsymbol{B}} &= \frac{1}{p(1-p)}\boldsymbol{P}^{-1}\left(\boldsymbol{\Sigma}_{\boldsymbol{z}} - 2p\gamma\boldsymbol{I}_n - \boldsymbol{C}\right)\boldsymbol{P}^{-1} \\ &= \boldsymbol{B} + 2(1-p)^{-1}\boldsymbol{P}^{-1}\left(\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n\right)\boldsymbol{P}^{-1}, \end{aligned} \tag{15}$$

and achieves a spectral distance to the ground truth indicator matrix

$$\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2 = 2(1-p)^{-1}\left\|\boldsymbol{P}^{-1}\left(\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n\right)\boldsymbol{P}^{-1}\right\|_2. \tag{16}$$

The crux is to understand whenever (16) matches the sufficiency criteria of Proposition 5 to access Eve's perfect recovery $\mathcal{B}$, and the vulnerability of the proposed scheme.

To that end, we must note that the matrices $\boldsymbol{E}_{\mathcal{B}}, \boldsymbol{F}$, and $\boldsymbol{G}$ introduced in (8) are summations of fourth-order moments of the matrices $\boldsymbol{M}$ and $\boldsymbol{A}$. Furthermore, even if the entries of the matrix $\boldsymbol{A}$ are assumed to be drawn *i.i.d.*, the products considered in (8) are coupled, and the summations are over dependent terms. As a result, additional statistical assumptions on the distribution of the channel matrix $\boldsymbol{A}$ are needed to control the estimate of the block structure $\widetilde{\boldsymbol{B}}$. For that reason, we provide Definition 6, which introduces a new notion of coherence relevant to our spectral clustering context.

Definition 6 (Coherence)**.** *For an $m \times n$ matrix $\boldsymbol{A}$, we let $\boldsymbol{M} = \boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}$ and $\boldsymbol{P} = \boldsymbol{M} \odot \boldsymbol{M}$. Given two positive numbers $\mu > 0$ and $\nu > 0$, a matrix $\boldsymbol{A}$ is said to be (μ, ν)-*coherent *if and only if the following bounds holds:*

1) <u>First-order bounds:</u>

$$\|\boldsymbol{A}\|_2 \leq \sqrt{\frac{n}{m}}\mu \tag{17a}$$

$$\|\boldsymbol{A}\|_{\max} \leq \sqrt{\frac{n\log(n)}{m}}\mu \tag{17b}$$

2) <u>Second-order bound:</u>

$$\|\boldsymbol{M}\|_{\max} \leq \log(n)\mu^2 \tag{17c}$$

3) <u>Fourth-order bounds:</u> For any block structure $\mathcal{B}$ over n element with maximal block length d, and for $(i,j) \in \{1, \ldots, n\}^2$, the fourth order matrix $\boldsymbol{E}_{\mathcal{B}}$ satisfies

$$\|\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n\|_2 \leq \max\left\{\frac{1}{m^2}, \frac{n}{m^4}\right\} d\sqrt{n}\log(n)\mu^8 \tag{17d}$$

where $\gamma = \frac{2(d-1)}{m^2} + \frac{(n-2)(d-1)}{m^4}$, and the fourth order matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ satisfy

$$\|\boldsymbol{F}\|_2 \leq \frac{n^2}{m^2}\log^2(n)\mu^8 \tag{17e}$$

$$\|\boldsymbol{G}\|_2 \leq \frac{n}{m}\log(n)\mu^4; \tag{17f}$$

The matrix $\boldsymbol{P}$ is invertible and

$$\lambda_{\min}(\boldsymbol{P}) \geq \nu^{-1}. \tag{17g}$$

[1] In more practical considerations, the transmission parameter p can be estimated by Eve from the covariance $\boldsymbol{\Sigma}_{\boldsymbol{y}}$ of the channel output as $\boldsymbol{\Sigma}_{\boldsymbol{y}} = p\boldsymbol{M}$.

The parameter μ is raised to different exponents in (17) to maintain homogeneity across the different matrix norms. Understanding when a matrix $\boldsymbol{A}$ is (μ,ν)-coherent is crucial to apply our theoretical analysis of Algorithm 1. However, finding coherence parameters when assuming the entries $\{a_{i,j}\}$ of $\boldsymbol{A}$ to be drawn *i.i.d.* from a know prior distribution can be particularly challenging as the quantities defined in (8) are summations of fourth and eighth-order terms in the matrix $\boldsymbol{A}$. As a result, the terms in those summations are *dependent*, and the usual incoherence bounds for matrix sensing [47], [48] cannot be directly applied.

Nonetheless, an interesting class of matrix $\boldsymbol{A}$ to consider is the one whose columns are drawn *i.i.d.* according to an *unitary* and *isotropic* distribution. In that case, we have

$$\mathrm{E}[\boldsymbol{P}] = \boldsymbol{I}_n + \frac{1}{m}\left(\boldsymbol{J}_n - \boldsymbol{I}_n\right). \tag{18}$$

Under the additional assumption that the columns of $\boldsymbol{A}$ have a *bounded inner product*, *i.e.* if there exists a small enough $\varepsilon > 0$ such that

$$p_{i,j} = |\langle \boldsymbol{a}_i, \boldsymbol{a}_j\rangle|^2 \leq \begin{cases} (1+\varepsilon) & \text{if } i = j \\ \frac{1}{m}(1+\varepsilon) & \text{if } i \neq j \end{cases} \tag{19}$$

for all (i,j), then we can show that (μ,ν)-coherence holds with high probability. Indeed, (17a) holds because of the unitary isotropic assumption on $\boldsymbol{A}$, (17b) is induced by the bounded, hence sub-Gaussian concentration of the matrix $\boldsymbol{A}$ (see *e.g.* [49]), (17c) is immediate from (19), and (17g) occurs from $\|\boldsymbol{P} - \mathrm{E}[\boldsymbol{P}]\|_2 < \lambda_{\min}(\mathrm{E}[\boldsymbol{P}]) = 1 - m^{-1}$ given a small enough ε. Finally, Lemma 7 validates (17d) with high probability, and its proof is provided in Appendix C-A. The proofs of the two later bounds (17e) and (17f) are omitted for brevity and can be re-derived by following analogous reasoning.

Lemma 7 (Concentration of $\boldsymbol{E}_{\mathcal{B}}$)**.** *Suppose that the columns of $\boldsymbol{A}$ are drawn* i.i.d. *according to a unitary isotropic random distribution and that* (19) *holds for some $\varepsilon > 0$, then there exists a constant $\mu > 0$ such that* (17d) *is satisfied with probability greater than $1 - 2n^{-1}$.*

The (μ,ν)-coherence assumption on the matrix $\boldsymbol{A}$ can be exploited with $p \leq \frac{1}{2}$ to control the spectral distance (16) as

$$\begin{aligned}\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2 &= 2(1-p)^{-1}\left\|\boldsymbol{P}^{-1}\left(\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n\right)\boldsymbol{P}^{-1}\right\|_2 \\ &\leq 4\left\|\boldsymbol{P}^{-1}\right\|_2^2\left\|\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n\right\|_2 \\ &\leq 4\max\left\{\frac{1}{m^2}, \frac{n}{m^4}\right\} d\sqrt{n}\log(n)\nu^2\mu^8. \end{aligned} \tag{20}$$

A direct application of Proposition 5 with (20) yields the following characterization of the asymptotic vulnerability of the communication protocol proposed in Section II from an eavesdropper attempting to learn the secret block structure $\mathcal{B}$ via Algorithm 2.

Corollary 8 (Asymptotic vulnerability)**.** *Suppose that $\boldsymbol{A}$ is (μ,ν)-coherent, then if*

$$\delta := \frac{\sqrt{2}}{16}\nu^{-2}\mu^{-8} - 2\max\left\{\frac{1}{m^2}, \frac{n}{m^4}\right\}\sqrt{nd} \geq 0 \tag{21}$$

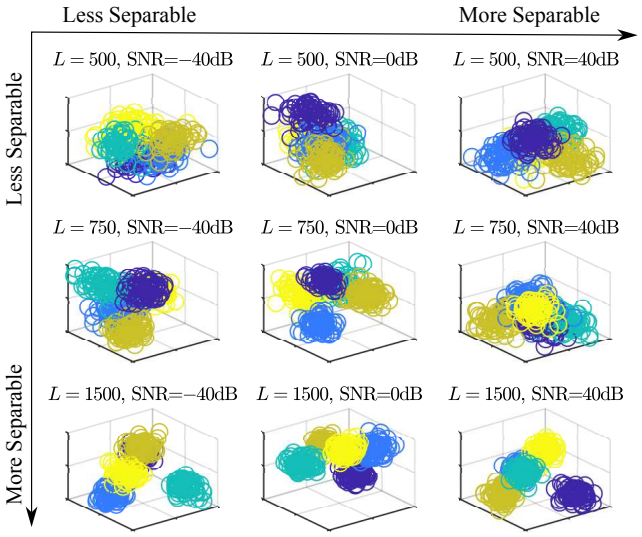

Figure 5. Projections of the clusters estimated by Algorithm 2 unto $\mathbb{R}^3$ for different numbers of snapshots and SNRs. Rows (from top to bottom): $L = 500$, $L = 750$, $L = 1500$. Columns (from left to right): SNR $= -40$dB, SNR $= 0$dB, SNR $= 40$dB. Other system parameters are $n = 400$, $m = 200$, $\beta = 2.5$ and $r = 5$.

Eve can recover the block structure $\mathcal{B}$ by applying Algorithm 1 provided access to infinitely many samples of the channel outputs $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots\}$.

This result suggests that the communication protocol between Alice and Bob proposed in Section II is compromised from the knowledge of the ground truth covariance and to a channel of large enough output dimension m with constant reuse of the secret key. We call this regime *asymptotic vulnerability*.

E. Estimation with a Finite Number of Snapshots

In practice, Eve can access a limited number of snapshots L before Alice terminates the communication or refreshes the structure $\mathcal{B}$. Consequently, the true covariance $\boldsymbol{\Sigma}_{\boldsymbol{z}}$ always remains unknown to Eve. Instead, she can attempt to estimate $\mathcal{B}$ from the empirical estimator of the covariance given by $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} = \frac{1}{L}\sum_{\ell=1}^{L}\left(\boldsymbol{z}_\ell - \mathrm{E}[\boldsymbol{z}]\right)\left(\boldsymbol{z}_\ell - \mathrm{E}[\boldsymbol{z}]\right)^{\mathsf{T}}$, where $\mathrm{E}[\boldsymbol{z}] = p\,\mathrm{diag}\left(\boldsymbol{M}^2\right)$ and $\mathrm{diag}(\cdot)$ is the operator that stacks the diagonal elements of a $n \times n$ matrix into an n-dimensional vector. Proposition 9 provides recovery guarantees for Eve under the proviso she accesses a large enough number of snapshots L.

Proposition 9 (Estimation with Finite Numbers of Snapshots)**.** *Let the quantity δ be as defined in* (21) *and suppose that $\delta > 0$, then there exist a constants $C > 0$ such that if L satisfies*

$$\begin{aligned}\frac{\sqrt{L}}{\log(L)} \geq \delta^{-1}\frac{n^2}{m^2}\log^2(n)\sqrt{d} \\ \cdot\left(1 + \frac{2}{7\log(n)}\frac{\sigma^2}{\mu^2} + \frac{4\beta n}{7m\log(n)}\frac{\sigma^4}{\mu^4}\right),\end{aligned} \tag{22}$$

then the output $\widehat{\mathcal{B}}$ of Algorithm 1 satisfies $\widehat{\mathcal{B}} = \mathcal{B}$ with probability greater than $1 - CL^{-1}$.

The proof of Proposition 9 is presented in Appendix B.

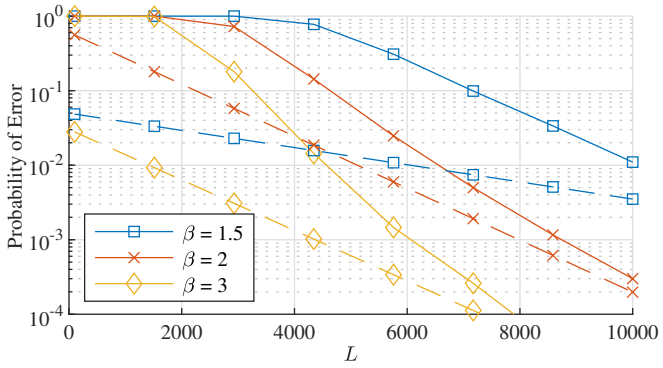

Figure 6. Probability of failure of Algorithm 1 as a function of the number of snapshots L for different communications rates β. Dashed lines represent Hoeffding's error rates p_{Hoeff} detailed in Section IV for the corresponding values of β. Herein, we set $n = 200$, $m = 100$, $r = 5$, and SNR $= 0$dB. Experiments are averaged over 10^5 trials.

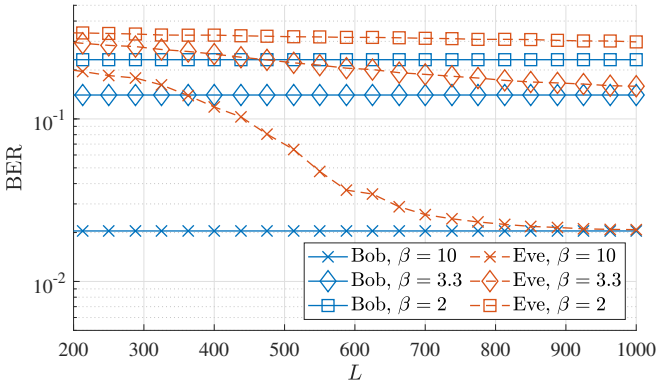

Figure 7. BER as a function of the number of snapshots L for different communication rates β. Herein, we set $n = 400$, $m = 200$, $r = 20$, and SNR $= 0$dB. Experiments are averaged over 10^5 trials.

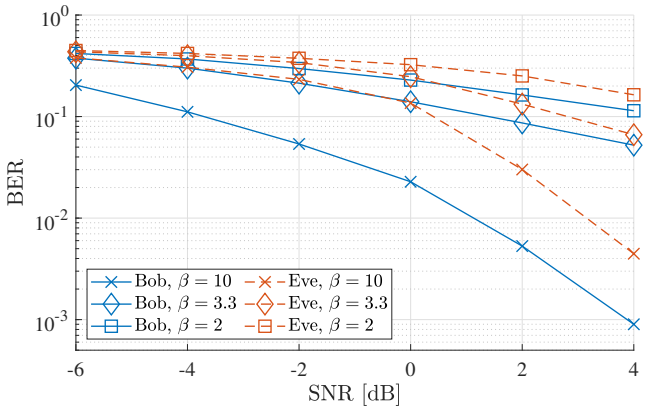

Figure 8. BER as a function of SNR for different communication rates β. Herein, we set $n = 400$, $m = 200$, $r = 20$, and $L = 400$. Experiments are averaged over 10^5 trials.

IV. Numerical simulations

In this section, we validate the theoretical findings presented in Section II-C and Section III through numerical simulations. Herein, the block compressed sensing problem (3) and compressed sensing problem (5) are solved using the block-basis pursuit (Block-BP) and basis pursuit (BP) convex relaxation with MATLAB and the SPGL1 package [50]. For a unitary and isometric matrix $\boldsymbol{A}$, the signal-to-noise ratio (SNR) at the channel output is defined as $\text{SNR} \triangleq \frac{\mathrm{E}\left[\|\boldsymbol{A}\boldsymbol{x}\|_2^2\right]}{\mathrm{E}\left[\|\boldsymbol{w}\|_2^2\right]} = \frac{pn^2}{\sigma^2 m^2}$. We subsequently select $\boldsymbol{A}$ at random with independent Gaussian entries $a_{i,j} \sim \mathcal{N}\left(0, \frac{1}{m}\right)$.

We start by considering the clustering capabilities of Algorithm 1. Figure 5 shows the clusters returned by the subroutine Algorithm 2 for different numbers of snapshots and different SNRs, for the case where $n = 400, m = 200$ and $\beta = 2.5$ and $r = 5$; that is due to the block structure, we have 5 clusters. It is clear that the value of L (number of snapshots) impacts whether we can identify the clusters and, thus, the block structure. Additionally, high SNR values result in better identifiability of the clusters, especially under a limited number of snapshots, when the signal and the noise empirical covariances are not yet decoupled.

Next, we evaluate the probability for Eve to recover the correct block structure $\mathcal{B}$ from the output of Algorithm 1 as a function of the number of observed snapshots, L, that she has acquired without a refresh of the block structure. We evaluate the empirical error rate of Algorithm 1, defined by the fraction random problem instances where $\widehat{\mathcal{B}} \neq \mathcal{B}$. To assess the secrecy of the proposed protocol, we compare this empirical error rate with the error rate of a Hoeffding test between the probability distribution $\mathcal{Y}$ of the channel output produced by the true block structure $\mathcal{B}$, and the probability distribution $\mathcal{Y}'$ produced by another block structure $\mathcal{B}'$. Given the Kullback–Leibler divergence $\mathrm{KL}(\mathcal{Y}, \mathcal{Y}')$ between those two distributions, Hoeffding's error rate is given by $p_{\text{Hoeff}} = C \exp\left(-L \min_{\mathcal{Y}'} \left\{\mathrm{KL}(\mathcal{Y}, \mathcal{Y}')^2\right\}\right)$ for some $C > 0$, where the minimum is taken over all possible block structures $\mathcal{B}'$ of r-blocks of length d that is not equal to $\mathcal{Y}$. Hoeffding's error rate is an asymptotic statistical lower bound on the error probability for hypothesis testing [51]. As $\mathcal{Y}$ and $\mathcal{Y}'$ are Gaussian mixtures in dimension m with 2^r classes, calculating the KL-divergence by a Monte Carlo method is computationally prohibitive, and we evaluate instead its variational approximation [52]. The findings are shown in Figure 6 suggest that larger values of β increase Eve's learning rate of the secret block structure, which corroborates with the theoretical results of Proposition 9 as $\sigma^2 \propto \frac{pn^2}{m^2}$ in fixed SNR settings. Additionally, for larger values of β, we observe that Algorithm 1 achieves an error exponent close to Hoeffding's rate, indicating the near-optimally of the proposed moment method to eavesdrop the block structure in the asymptotic $L \to \infty$.

Finally, motivated by communication applications, we consider the downlink of a massive MIMO system. We assume Alice transmits parsimoniously messages encoded on a block-sparse BPSK constellation to Bob, meaning that $\boldsymbol{x}$ is drawn according to a block-Bernoulli probability distribution, *i.e.* within an active block $x_i = \pm 1$ with independent and equal probability $\frac{1}{2}$, and $x_i = 0$ within a non-active block. We define the bit-error-rate (BER) as the ratio of entries that are in the active support of Alice's message ($x_i \neq 0$), and that are incorrectly decoded by the receiver, *i.e.* $\widehat{x}_i \neq x_i$. Assuming

that Eve relies on its estimate $\widehat{\mathcal{B}}$ of the block structure obtained from the output of Algorithm 1, we empirically evaluate Bob's and Eve's BERs as a function of the number of snapshots in Figure 7, and as a function of the SNR in Figure 8. The figures suggest that larger values of β ease both Bob's and Eve's decoding. Eve can achieve the same BER as Bob if the secret structure is reused sufficiently many times. Additionally, for a fixed number of snapshots, Eve's decoding is more impeded by the noise than Bob's, and the BER margin between Bob and Eve increases with the redundancy parameter β. Hence, for fixed channel dimensions, if Alice reduces her communication rate with Bob by selecting a smaller block activation probability p, she can harden Eve's decoding. This observation shows the trade-off between the communication rate Alice can achieve and the secrecy against an eavesdropper the protocol can induce. For the example considered, it is clear that for $\beta = 10$ and $L < 500$ Eve cannot decode while Bob can.

V. Conclusions and Future Work

In this article, we introduced a novel communication protocol with provable privacy guarantees. The proposed method harnesses a secret block-sparse prior to recovering the initial message from underdetermined linear measurements gathered at the output of a fat channel matrix. As block sparsity allows exact recovery in conditions where classical compressed sensing would provably fail, we established the existence of a secure transmission regime to a single snapshot between Alice and Bob. We studied the privacy guarantees of this communication protocol to multiple transmissions without refreshing the shared secret and proposed an algorithm for an eavesdropper to learn the block structure via the method of moments. The proposed block structure estimator appears to be asymptotically near-optimal. We validated the privacy benefits of this framework through numerical experiments.

Possible extensions of this work include a comprehensive study of the trade-off between the communication rate that Alice and Bob can achieve and the lifespan of the secret block structure. Additionally, the proposed scheme paves the way for further linear inverse problem-based implementation of private communication protocols over the physical layer.

Appendix A
Proof of Proposition 4

Let $\boldsymbol{u} = \boldsymbol{M}\boldsymbol{x} = \boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{x}$ and $\widetilde{\boldsymbol{w}} = \boldsymbol{A}^{\mathsf{T}}\boldsymbol{w}$ for convenience purposes. Moreover let $\boldsymbol{z} = \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{y}\right) \odot \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{y}\right)$. We have that

$$\begin{aligned}\boldsymbol{z} &= (\boldsymbol{A}^{\mathsf{T}}\boldsymbol{y}) \odot (\boldsymbol{A}^{\mathsf{T}}\boldsymbol{y}) = (\boldsymbol{u} + \widetilde{\boldsymbol{w}}) \odot (\boldsymbol{u} + \widetilde{\boldsymbol{w}}) \\ &= \boldsymbol{u} \odot \boldsymbol{u} + 2\boldsymbol{u} \odot \widetilde{\boldsymbol{w}} + \widetilde{\boldsymbol{w}} \odot \widetilde{\boldsymbol{w}}. \end{aligned} \tag{23}$$

We aim to derive the expression of the covariance of $\boldsymbol{z}$. First, the independence between $\boldsymbol{x}$ and $\boldsymbol{w}$ implies the independence between $\boldsymbol{u}$ and $\widetilde{\boldsymbol{w}}$. Additionally, the assumptions $\mathrm{E}[\boldsymbol{x}] = \boldsymbol{0}$ and $\mathrm{E}[\boldsymbol{w}] = \boldsymbol{0}$ imply that $\mathrm{E}[\boldsymbol{u}] = 0$ and $\mathrm{E}[\widetilde{\boldsymbol{w}}] = 0$. This yields

$$\mathrm{Cov}(\boldsymbol{u} \odot \boldsymbol{u}, \widetilde{\boldsymbol{w}} \odot \widetilde{\boldsymbol{w}}) = \boldsymbol{0}, \tag{24a}$$

$$\mathrm{Cov}(\boldsymbol{u} \odot \boldsymbol{u}, \boldsymbol{u} \odot \widetilde{\boldsymbol{w}}) = \boldsymbol{0}, \tag{24b}$$

$$\mathrm{Cov}(\boldsymbol{u} \odot \widetilde{\boldsymbol{w}}, \widetilde{\boldsymbol{w}} \odot \widetilde{\boldsymbol{w}}) = \boldsymbol{0}. \tag{24c}$$

Hence the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{z}} = \mathrm{Cov}(\boldsymbol{z}, \boldsymbol{z})$ of the random vector $\boldsymbol{z}$ reduces to

$$\boldsymbol{\Sigma}_{\boldsymbol{z}} = \boldsymbol{\Sigma}_{\boldsymbol{u}\odot\boldsymbol{u}} + 2\boldsymbol{\Sigma}_{\boldsymbol{u}\odot\widetilde{\boldsymbol{w}}} + \boldsymbol{\Sigma}_{\widetilde{\boldsymbol{w}}\odot\widetilde{\boldsymbol{w}}}. \tag{25}$$

We derive in the sequel the expression of each of the three matrices on the right-hand side of (25).

a) Expression of $\boldsymbol{\Sigma}_{\boldsymbol{u}\odot\boldsymbol{u}}$: By definition of the vector $\boldsymbol{u}$, we have for any $i = 1, \ldots, n$ that

$$u_i = \sum_{k=1}^{n} x_k \langle \boldsymbol{a}_i, \boldsymbol{a}_k \rangle = \sum_{k=1}^{n} x_k m_{i,k}. \tag{26}$$

Thus, denoting $p_{i,k}$ the (i,k)-th term of the matrix $\boldsymbol{P} = \boldsymbol{M} \odot \boldsymbol{M}$, the expected value $\mathrm{E}[u_i^2]$ of the random variable u_i^2 is given by

$$\begin{aligned}\mathrm{E}[u_i^2] &= \mathrm{E}\left[\left(\sum_{k=1}^{n} x_k m_{i,k}\right)\left(\sum_{k'=1}^{n} x_{k'} m_{i,k'}\right)\right] \\ &= \sum_{k=1}^{n}\sum_{k'=1}^{n} \mathrm{E}[x_k x_{k'}] m_{i,k} m_{i,k'} \\ &= \sum_{k=1}^{n} \mathrm{E}[x_k^2] m_{i,k}^2 = \sum_{k=1}^{n} \mathrm{E}[x_k^2] p_{i,k}. \end{aligned} \tag{27}$$

For readability, we drop the summation interval from 1 to n in the following. A direct calculation of the expected value $\mathrm{E}[u_i^2 u_j^2]$ of the product $u_i^2 u_j^2$ yields

$$\begin{aligned}&\mathrm{E}[u_i^2 u_j^2] \\ &= \sum_{k_1}\sum_{k_1'}\sum_{k_2}\sum_{k_2'} \mathrm{E}[x_{k_1} x_{k_1'} x_{k_2} x_{k_2'}] m_{i,k_1} m_{i,k_1'} m_{j,k_2} m_{j,k_2'} \\ &= \sum_{k}\sum_{k'} \mathrm{E}[x_k^2 x_{k'}^2] p_{i,k} p_{j,k'} \\ &\qquad + 2\sum_{k}\sum_{k'\neq k} \mathrm{E}[x_k^2 x_{k'}^2] m_{i,k} m_{i,k'} m_{j,k} m_{j,k'}. \end{aligned} \tag{28}$$

Equations (27) and (28) immediately lead to an expression of the generic term of the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{u}\odot\boldsymbol{u}}$ of the form

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{u}\odot\boldsymbol{u}}(i,j) &= \mathrm{E}[u_i^2 u_j^2] - \mathrm{E}[u_i^2]\mathrm{E}[u_j^2] \\ &= \sum_{k}\sum_{k'} \left(\mathrm{E}[x_k^2 x_{k'}^2] - \mathrm{E}[x_k^2]\mathrm{E}[x_{k'}^2]\right) p_{i,k} p_{j,k'} \\ &\qquad + 2\sum_{k}\sum_{k'\neq k} \mathrm{E}[x_k^2 x_{k'}^2] m_{i,k} m_{i,k'} m_{j,k} m_{j,k'} \\ &= \boldsymbol{p}_i^{\mathsf{T}} \boldsymbol{\Sigma}_{\boldsymbol{x}\odot\boldsymbol{x}} \boldsymbol{p}_j \\ &\qquad + 2\sum_{k}\sum_{k'\neq k} \mathrm{E}[x_k^2 x_{k'}^2] m_{i,k} m_{i,k'} m_{j,k} m_{j,k'}. \end{aligned} \tag{29}$$

Next, the second sum on the right-hand side of expression (29) can be reformulated as

$$\begin{aligned}&\sum_{k}\sum_{k'\neq k} \mathrm{E}[x_k^2 x_{k'}^2] m_{i,k} m_{i,k'} m_{j,k} m_{j,k'} \\ &\quad = \sum_{k} \sum_{\substack{k'\neq k \\ \mathcal{B}(k')=\mathcal{B}(k)}} \mathrm{E}[x_k^2 x_{k'}^2] m_{i,k} m_{i,k'} m_{j,k} m_{j,k'} \\ &\qquad + \sum_{k} \sum_{\substack{k'\neq k \\ \mathcal{B}(k')\neq\mathcal{B}(k)}} \mathrm{E}[x_k^2 x_{k'}^2] m_{i,k} m_{i,k'} m_{j,k} m_{j,k'} \end{aligned}$$

$$= p \sum_{k} \sum_{\substack{k' \neq k \\ \mathcal{B}(k')=\mathcal{B}(k)}} m_{i,k} m_{i,k'} m_{j,k} m_{j,k'} + p^2 \sum_{k} \sum_{\substack{k' \neq k \\ \mathcal{B}(k') \neq \mathcal{B}(k)}} m_{i,k} m_{i,k'} m_{j,k} m_{j,k'}$$
$$= p \boldsymbol{E}_{\mathcal{B}}(i,j) + p^2 \boldsymbol{F}(i,j). \quad (30)$$

Given the fourth-order statistics (7) on the block Bernouilli-Gaussian distribution of $\boldsymbol{x}$, we have the matrix expression $\boldsymbol{\Sigma}_{\boldsymbol{x}\odot\boldsymbol{x}} = 2p\boldsymbol{I}_n + p(1-p)\boldsymbol{B}$. Substituting the previous with (30) into (29) yields the expression

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{u}\odot\boldsymbol{u}} &= \boldsymbol{P}\boldsymbol{\Sigma}_{\boldsymbol{x}\odot\boldsymbol{x}}\boldsymbol{P} + 2p\boldsymbol{E}_{\mathcal{B}} + 2p^2\boldsymbol{F} \\ &= p(1-p)\boldsymbol{P}\boldsymbol{B}\boldsymbol{P} + 2p\boldsymbol{E}_{\mathcal{B}} + 2p\boldsymbol{P}^2 + 2p^2\boldsymbol{F}. \quad (31)\end{aligned}$$

b) Expression of $\boldsymbol{\Sigma}_{\boldsymbol{u}\odot\widetilde{\boldsymbol{w}}}$: From the independence assumption between $\boldsymbol{x}$ and $\boldsymbol{w}$, the expectation reduces to

$$\begin{aligned}\mathrm{E}[(\boldsymbol{u}\odot\widetilde{\boldsymbol{w}})_i] &= \mathrm{E}\left[\left(\sum_k m_{i,k}x_k\right)\left(\sum_{k'} a_{k',i}w_{k'}\right)\right] \\ &= \sum_k \sum_{k'} \mathrm{E}[x_k]\mathrm{E}[w_{k'}] m_{i,k}a_{k',i} = 0, \quad (32)\end{aligned}$$

and the correlation term is

$$\begin{aligned}&\mathrm{E}[(\boldsymbol{u}\odot\widetilde{\boldsymbol{w}})_i(\boldsymbol{u}\odot\widetilde{\boldsymbol{w}})_j] \\ &= \mathrm{E}\left[\left(\sum_{\ell_1} m_{i,\ell_1}x_{\ell_1}\right)\left(\sum_{k_1'} a_{k_1',i}w_{\ell'_1}\right)\right. \\ &\qquad \left.\left(\sum_{\ell_2} m_{j,\ell_2}x_{\ell_2}\right)\left(\sum_{\ell_2'} a_{\ell_2',j}w_{\ell_2'}\right)\right] \\ &= \left(\sum_k \mathrm{E}[x_k^2] m_{i,k}m_{j,k}\right)\left(\sum_{k'} \mathrm{E}[w_{k'}^2]a_{k',i}a_{k',j}\right) \\ &= p\sigma^2\left(\sum_k m_{i,k}m_{j,k}\right)\left(\sum_{k'} a_{k',i}a_{k',j}\right) \\ &= p\sigma^2\left(\sum_k m_{i,k}m_{j,k}\right) m_{i,j} \quad (33)\end{aligned}$$

and we use (32) and (33) to get

$$\boldsymbol{\Sigma}_{\boldsymbol{u}\odot\widetilde{\boldsymbol{w}}} = p\sigma^2 \boldsymbol{M}^2 \odot \boldsymbol{M}. \quad (34)$$

c) Expression of $\boldsymbol{\Sigma}_{\widetilde{\boldsymbol{w}}\odot\widetilde{\boldsymbol{w}}}$: We follow analogous reasoning than for the previous term. First, the expectation is given by

$$\begin{aligned}\mathrm{E}[(\widetilde{\boldsymbol{w}}\odot\widetilde{\boldsymbol{w}})_i] &= \mathrm{E}\left[\left(\sum_k a_{k,i}w_k\right)\left(\sum_{k'} a_{k',i}w_{k'}\right)\right] \\ &= \sum_k \sum_{k'} \mathrm{E}[w_k w_{k'}] a_{k,i}a_{k',i} \\ &= \sum_k \mathrm{E}[w_k^2]a_{k,i}^2 = \sigma^2 \sum_k a_{k,i}^2. \quad (35)\end{aligned}$$

Hence, the generic covariance term is

$$\begin{aligned}&\mathrm{E}[(\widetilde{\boldsymbol{w}}\odot\widetilde{\boldsymbol{w}})_i(\widetilde{\boldsymbol{w}}\odot\widetilde{\boldsymbol{w}})_j] - \mathrm{E}[(\widetilde{\boldsymbol{w}}\odot\widetilde{\boldsymbol{w}})_i]\mathrm{E}[(\widetilde{\boldsymbol{w}}\odot\widetilde{\boldsymbol{w}})_j] \\ &= \mathrm{E}\left[\left(\sum_{k_1} a_{k_1,i}w_{k_1}\right)\left(\sum_{k_1'} a_{k_1',i}w_{k_1'}\right)\right. \\ &\qquad \left.\left(\sum_{k_2} a_{k_2,j}w_{k_2}\right)\left(\sum_{k_2'} a_{k_2',j}w_{k_2'}\right)\right] \\ &= \sum_k \sum_{k'} \left(\mathrm{E}[w_k^2 w_{k'}^2] - \mathrm{E}[w_k^2]\mathrm{E}[w_k^2]\right) a_{k,i}^2 a_{k',j}^2 \\ &\quad + 2\sum_k \sum_{k'\neq k} \mathrm{E}[w_k^2 w_{k'}^2] a_{k,i}a_{k,j}a_{k',i}a_{k',j} \\ &= \sum_k \sum_{k'} \left(\mathrm{E}[w_k^2 w_{k'}^2] - \mathrm{E}[w_k^2]\mathrm{E}[w_k^2]\right) a_{k,i}^2 a_{k',j}^2 \\ &\quad + 2\sigma^4 \sum_k \sum_{k'\neq k} a_{k,i}a_{k,j}a_{k',i}a_{k',j}. \quad (36)\end{aligned}$$

Furthermore as $\boldsymbol{w}$ is an *i.i.d.* white Gaussian random vector with variance σ^2, we have that $\boldsymbol{\Sigma}_{\boldsymbol{w}\odot\boldsymbol{w}} = 2\sigma^4\boldsymbol{I}_m$. Hence, we may write

$$\begin{aligned}\boldsymbol{\Sigma}_{\widetilde{\boldsymbol{w}}\odot\widetilde{\boldsymbol{w}}} &= (\boldsymbol{A}\odot\boldsymbol{A})^{\mathsf{T}} \boldsymbol{\Sigma}_{\boldsymbol{w}\odot\boldsymbol{w}} (\boldsymbol{A}\odot\boldsymbol{A}) + 2\sigma^4\boldsymbol{G} \\ &= 2\sigma^4\left((\boldsymbol{A}\odot\boldsymbol{A})^{\mathsf{T}}(\boldsymbol{A}\odot\boldsymbol{A}) + \boldsymbol{G}\right). \quad (37)\end{aligned}$$

We achieve the desired statement by substituting (31), (34) and (37) into (25). □

Appendix B
Proof of Proposition 9

We start the proof by noticing that from Proposition 4, the indicator matrix $\boldsymbol{B}$ of the block structure matrix $\mathcal{B}$ is given by

$$\boldsymbol{B} = \frac{1}{p(1-p)} \boldsymbol{P}^{-1}\left(\boldsymbol{\Sigma}_{\boldsymbol{z}} - 2p\boldsymbol{E}_{\mathcal{B}} - \boldsymbol{C}\right)\boldsymbol{P}^{-1}, \quad (38)$$

where $\boldsymbol{C}$ is defined in (10) and is independent of $\mathcal{B}$. The spectral distance $\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2$ can be bounded with the triangle inequality, the (μ,ν)-incoherence of the matrix $\boldsymbol{A}$, and the assumption $p \leq \frac{1}{2}$ as follows

$$\begin{aligned}&\left\|\widetilde{\boldsymbol{B}} - \boldsymbol{B}\right\|_2 \\ &= \frac{1}{p(1-p)}\left\|\boldsymbol{P}^{-1}\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \boldsymbol{\Sigma}_{\boldsymbol{z}} + 2p(\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n)\right)\boldsymbol{P}^{-1}\right\|_2 \\ &\leq \left\|\boldsymbol{P}^{-1}\right\|_2^2\left(2p^{-1}\left\|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \boldsymbol{\Sigma}_{\boldsymbol{z}}\right\|_2 + 4\left\|\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n\right\|_2\right) \\ &\leq \nu^2\left(2p^{-1}\left\|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \boldsymbol{\Sigma}_{\boldsymbol{z}}\right\|_2\right. \\ &\qquad \left. + 4\max\left\{\frac{1}{m^2}, \frac{n}{m^4}\right\} d\sqrt{n}\log(n)\mu^8\right). \quad (39)\end{aligned}$$

We obtain from (39) and Proposition 5 that Algorithm 1 outputs the true block structure if

$$\begin{aligned}&\left\|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \boldsymbol{\Sigma}_{\boldsymbol{z}}\right\|_2 \\ &\quad \leq p\sqrt{d}\mu^8\left(\frac{\sqrt{2}}{16}\nu^{-2}\mu^{-8} - 2\max\left\{\frac{1}{m^2}, \frac{n}{m^4}\right\} d\sqrt{n}\right) \\ &\quad \leq p\sqrt{d}\mu^8\delta \quad (40)\end{aligned}$$

where the right-hand side of (40) is non-negative by the assumption $\delta > 0$. The estimated covariance error on the left-hand side of (40) can be made arbitrarily small for a sufficiently large number of snapshots L. Lemma 10 provides a

high-probability bound on the error on the estimated covariance error as $L \to \infty$ in terms of the problem parameters.

Lemma 10 (Covariance estimation)**.** *Under the hypothesis of Proposition 9, there exist a constant $C > 0$ such that the event*

$$\left\|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \boldsymbol{\Sigma}_{\boldsymbol{z}}\right\|_2 \leq \frac{\log(L)}{\sqrt{L}} \beta^{-1} \frac{n}{m} \log^2(n) d\mu^8 \cdot \left(1 + \frac{2}{7\log(n)} \frac{\sigma^2}{\mu^2} + \frac{4\beta n}{7m\log(n)} \frac{\sigma^4}{\mu^4}\right) \quad (41)$$

holds with probability greater than $1 - CL^{-1}$.

The proof of Lemma 10 is deferred to Appendix C-B for readability. It suffices to replace the left-hand side of inequation (40) with the high probability bound given Lemma 10 to yield the desired statement. □

Appendix C
Proof of the Technical Lemmas

A. Proof of Lemma 7

We start by studying the expected value of the diagonal terms of $\boldsymbol{E}_{\mathcal{B}}$. We have that

$$\boldsymbol{E}_{\mathcal{B}}(i,i) = \sum_k \sum_{\substack{k' \neq k \\ \mathcal{B}(k) = \mathcal{B}(k')}} p_{i,k} p_{i,k'}, \quad (42)$$

and we write for each $i = \{1, \ldots, n\}$

$$\gamma_i = \sum_k \sum_{\substack{k' \neq k \\ \mathcal{B}(k) = \mathcal{B}(k')}} \mathrm{E}[p_{i,k}]\mathrm{E}[p_{i,k'}]. \quad (43)$$

By the isotropy assumption on the matrix $\boldsymbol{A}$, γ_i is constant for different values of i, and we may write $\gamma \triangleq \gamma_1 = \cdots = \gamma_n$. Moreover, by (18), the right hand side of (43) is a summation over $n(d-1)$ elements yielding

$$\mathrm{E}[p_{i,k}]\mathrm{E}[p_{i,k'}] = \begin{cases} \frac{1}{m} & \text{if } k = i \text{ or } k' = i \\ \frac{1}{m^2} & \text{otherwise.} \end{cases} \quad (44)$$

Counting the number of occurrences in each case, we have $\gamma = \frac{2(d-1)}{m} + \frac{(n-2)(d-1)}{m^2}$. Additionally, under the lemma's conditions, (42) and (43) imply

$$\begin{aligned} |\boldsymbol{E}_{\mathcal{B}}(i,i) - \gamma| &= |\boldsymbol{E}_{\mathcal{B}}(i,i) - \gamma_i| \\ &= \left| \boldsymbol{E}_{\mathcal{B}}(i,i) - \sum_k \sum_{\substack{k' \neq k \\ \mathcal{B}(k) = \mathcal{B}(k')}} \mathrm{E}[p_{i,k}]\mathrm{E}[p_{i,k'}] \right| \\ &\leq \sum_k \sum_{\substack{k' \neq k \\ \mathcal{B}(k) = \mathcal{B}(k')}} |p_{i,k}p_{i,k'} - \mathrm{E}[p_{i,k}]\mathrm{E}[p_{i,k'}]| \\ &\leq \sum_k \sum_{\substack{k' \neq k \\ \mathcal{B}(k) = \mathcal{B}(k')}} |p_{i,k} - \mathrm{E}[p_{i,k}]|\, \mathrm{E}[p_{i,k'}]\, (1+\varepsilon) \\ &= (1+\varepsilon) \sum_k |p_{i,k} - \mathrm{E}[p_{i,k}]| \left(\sum_{\substack{k' \neq k \\ \mathcal{B}(k) = \mathcal{B}(k')}} \mathrm{E}[p_{i,k'}] \right) \end{aligned} \quad (45)$$

where we used in the second inequality the assumption $|p_{i,k'} - \mathrm{E}[p_{i,k'}]| \leq \mathrm{E}[p_{i,k'}]\varepsilon$. As a result, by the isometry assumption, the terms of the summation in the right-hand side of (45) are *independent* and *bounded* by $(1+\varepsilon)$ and $\frac{(1+\varepsilon)}{m^2}$ when $k = i$ and $k \neq i$, respectively. Hence, the Chernoff bound can be applied [53], and we have

$$\mathbb{P}\left\{ |\boldsymbol{E}_{\mathcal{B}}(i,i) - \gamma| \leq \left(1 + \frac{d-1}{m^2}\sqrt{2n}\log(n)\right)(1+\varepsilon)^2 \right\} \geq 1 - 2n^{-2}. \quad (46)$$

On the off-diagonal, because of the isotropy assumption, the random variable $m_{i,k}m_{i,k'}m_{j,k}m_{j,k'}$ with $i \neq j$ has an even distribution for all k and k' whenever $i \neq j$. Therefore its expected value is null, that is $\mathrm{E}[m_{i,k}m_{i,k'}m_{j,k}m_{j,k'}] = 0$. Denote by $\overline{\boldsymbol{E}}_{\mathcal{B}}$ the matrix with off-diagonal terms equal to $\boldsymbol{E}_{\mathcal{B}}$ with diagonal entries $\overline{\boldsymbol{E}}_{\mathcal{B}}(i,i) = 0$ for all $i \in \{1, \ldots, n\}$. Relying on the symmetrization principle, we introduce the Rademacher random variable $\rho_{i,j} = \mathrm{sgn}(\boldsymbol{E}_{\mathcal{B}}(i,j))$, where $\mathrm{sgn}(\cdot)$ denotes the signum function. We note that $\{\rho_{i,j}\}_{i;j \geq i+1}$ are pair-wise independent. Furthermore, we can decompose the matrix $\overline{\boldsymbol{E}}_{\mathcal{B}}$ as the sum

$$\overline{\boldsymbol{E}}_{\mathcal{B}} = \sum_{i=1}^n \sum_{j=i+1}^n \rho_{i,j} \left|\boldsymbol{E}_{\mathcal{B}}(i,j)\right| \left(\boldsymbol{e}_i^{\mathsf{T}} \boldsymbol{e}_j + \boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{e}_i\right). \quad (47)$$

Next, we recall in Proposition 11 (see *e.g.* [49, Theorem 4.1.1]) a matrix norm concentration inequality for matrices with Rademacher entries.

Proposition 11 (Sum of symmetric Rademacher matrix series)**.** *Consider a fixed symmetric matrix $\boldsymbol{B}$ of dimension n. Let $\{\rho_{i,j}\}_{i;j \geq i+1}$ be a finite sequence of independent Rachemacher variables, and introduce the matrix Rademacher series*

$$\boldsymbol{Z} = \sum_{i=1}^n \sum_{j=i+1}^n \rho_{i,j} b_{i,j} \left(\boldsymbol{e}_i^{\mathsf{T}} \boldsymbol{e}_j + \boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{e}_i\right). \quad (48)$$

Let v be the matrix variance statistic of the Rademacher sum defined as $v(\boldsymbol{Z}) = \max_j \{\|\boldsymbol{b}_j\|_2^2\}$ then for all $t > 0$ we have

$$\mathbb{P}\{\|\boldsymbol{Z}\|_2 \geq t\} \leq 2n \exp\left(\frac{-t^2}{2v(\boldsymbol{Z})}\right). \quad (49)$$

To bound $\left\|\overline{\boldsymbol{E}}_B\right\|_2$ using Proposition 11, we evaluate the matrix variance $v(\overline{\boldsymbol{E}}_{\mathcal{B}})$ from the decomposition (47). It yields

$$\begin{aligned} v(\overline{\boldsymbol{E}}_{\mathcal{B}}) &= \max_j \{\left\|\overline{\boldsymbol{E}}_{\mathcal{B},j}\right\|_2^2\} \\ &= \max_j \left\{ \sum_{i \neq j} \left(\sum_{\substack{k' \neq k \\ \mathcal{B}(k) = \mathcal{B}(k')}} |m_{i,k}m_{i,k'}m_{j,k}m_{j,k'}| \right)^2 \right\} \\ &= \max_j \left\{ \sum_{i \neq j} \left(\sum_{\substack{k' \neq k \\ \mathcal{B}(k) = \mathcal{B}(k')}} \sqrt{p_{i,k}p_{i,k'}p_{j,k}p_{j,k'}} \right)^2 \right\} \\ &\leq \max_j \left\{ \sum_{i \neq j} \left(\sum_{\substack{k' \neq k \\ \mathcal{B}(k) = \mathcal{B}(k')}} \sqrt{\mathrm{E}[p_{i,k}]\mathrm{E}[p_{i,k'}]\mathrm{E}[p_{j,k}]\mathrm{E}[p_{j,k'}]} \right)^2 \right\} \\ &\quad \cdot (1+\varepsilon)^4 \end{aligned} \quad (50)$$

where the quantity to maximize in the last inequality is constant across different values for j and can be evaluated for $j = 1$ without loss of generality. The inner summation in (50) is taken over $n(d-1)$ terms, which are equal to $\frac{1}{m^2}$ when $k \neq i$ and $k' \neq i$, and equal to $\frac{1}{m}$ when $k = i$ or $k' = i$. After counting the occurrences, we may reduce (50) to

$$v(\overline{\boldsymbol{E}}_{\mathcal{B}}) \leq 2nd^2 \left(\frac{1}{m} + \frac{n}{m^2}\right)^2 (1+\varepsilon)^4. \tag{51}$$

Applying the matrix concentration inequality of Proposition 11 with $t = 2\sqrt{n\log(n)}d\left(\frac{1}{m^2} + \frac{n}{m^4}\right)(1+\varepsilon)^2$ induces

$$\mathbb{P}\left\{\left\|\overline{\boldsymbol{E}}_{\mathcal{B}}\right\|_2 \leq 2\sqrt{2n\log(n)}d\left(\frac{1}{m^2} + \frac{n}{m^4}\right)(1+\varepsilon)^2\right\} \geq 1 - 2n^{-1}. \tag{52}$$

We are now ready to achieve the desired statement. First, by the triangle inequality, we have

$$\begin{aligned}\left\|\boldsymbol{E}_{\mathcal{B}} - \gamma \boldsymbol{I}_n\right\|_2 &\leq \left\|\overline{\boldsymbol{E}}_{\mathcal{B}}\right\|_2 + \left\|\boldsymbol{E}_{\mathcal{B}} - \overline{\boldsymbol{E}}_{\mathcal{B}} - \gamma \boldsymbol{I}_n\right\|_2 \\ &= \left\|\overline{\boldsymbol{E}}_{\mathcal{B}}\right\|_2 + \max_i \left|\boldsymbol{E}_{\mathcal{B}}(i,i) - \gamma\right|. \end{aligned} \tag{53}$$

It suffices to substitute the probability bounds (46) and (52) into (53) with the union bound to yield

$$\left\|\boldsymbol{E}_{\mathcal{B}} - \boldsymbol{I}_n\right\|_2 \leq 6\sqrt{2n}\log(n)d\max\left\{\frac{1}{m^2}, \frac{n}{m^4}\right\}(1+\varepsilon)^2 \tag{54}$$

with probability greater than $1 - 4n^{-1}$. The statement of Lemma 7 follows by selecting the incoherence parameter $\mu = \left(6\sqrt{2}\,(1+\varepsilon)^2\right)^{\frac{1}{8}}$. □

B. Proof of Lemma 10

We seek to upper bound the quantity $\left\|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \boldsymbol{\Sigma}_{\boldsymbol{z}}\right\|_2$ with overwhelming probability. We start the proof by recalling in Proposition 12 the matrix Bernstein concentration inequality in the case of covariance estimation [49].

Proposition 12 (Matrix Bernstein for covariance estimation)**.** *Assume that there exist a constant C such that $\|\boldsymbol{z}_k - \mathrm{E}[\boldsymbol{z}_k]\|_2 \leq C\log(L)\|\boldsymbol{\Sigma}_{\boldsymbol{z}}\|_2$ for all $k = 1, \ldots, L$, we have that*

$$\begin{aligned}&\mathbb{P}\left(\left\|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \boldsymbol{\Sigma}_{\boldsymbol{z}}\right\|_2 \geq t\right) \\ &\quad\leq 2n\exp\left(\frac{-Lt^2/2}{C\log(L)\left(\|\boldsymbol{\Sigma}_{\boldsymbol{z}}\|_2^2 + \frac{2}{3}\|\boldsymbol{\Sigma}_{\boldsymbol{z}}\|_2 t\right)}\right). \end{aligned} \tag{55}$$

Hence, it is sufficient to provide a high probability bound on $\|\boldsymbol{\Sigma}\|_2$ to prove the desired statement. To that end, we apply the triangle inequality on (9). This yields

$$\begin{aligned}\|\boldsymbol{\Sigma}_{\boldsymbol{z}}\|_2 &= \|p(1-p)\boldsymbol{PBP} + 2p\boldsymbol{E}_{\mathcal{B}} + \boldsymbol{C}\|_2 \\ &\leq p(1-p)\|\boldsymbol{PBP}\|_2 + 2p\|\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n\|_2 \\ &\quad + \|\boldsymbol{C}\|_2 + \|2p\gamma\boldsymbol{I}_n\|_2 \\ &\leq p\|\boldsymbol{P}\|_2^2\|\boldsymbol{B}\|_2 + 2p\|\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n\|_2 + \|\boldsymbol{C}\|_2 + 2p\gamma. \end{aligned} \tag{56}$$

Now, we bound each of the elements on the right-hand side of (56) individually. We recall $\|\boldsymbol{B}\|_2 = d$, $\|\boldsymbol{E}_{\mathcal{B}} - \gamma\boldsymbol{I}_n\|_2$ is controlled by the (μ,ν)-coherence assumption on $\boldsymbol{A}$. Furthermore, we recall that for any Hermitian matrices $\boldsymbol{X}, \boldsymbol{Y}$ of same dimension, we have $\|\boldsymbol{X}\odot\boldsymbol{Y}\|_2 \leq \|\boldsymbol{X}\|_2\|\boldsymbol{Y}\|_{\max}$ (see *e.g.* [54, p. 113]). This implies that

$$\begin{aligned}\|\boldsymbol{P}\|_2 &= \|\boldsymbol{M}\odot\boldsymbol{M}\|_2 \leq \|\boldsymbol{M}\|_2\|\boldsymbol{M}\|_{\max} \\ &= \|\boldsymbol{A}\|_2^2\|\boldsymbol{M}\|_{\max} \end{aligned} \tag{57a}$$

$$\begin{aligned}\left\|\boldsymbol{M}^2\odot\boldsymbol{M}\right\|_2 &\leq \left\|\boldsymbol{M}^2\right\|_2\|\boldsymbol{M}\|_{\max} \\ &= \|\boldsymbol{A}\|_2^4\|\boldsymbol{M}\|_{\max} \end{aligned} \tag{57b}$$

$$\|\boldsymbol{A}\odot\boldsymbol{A}\|_2 \leq \|\boldsymbol{A}\|_2\|\boldsymbol{A}\|_{\max}. \tag{57c}$$

We are now ready to bound $\|\boldsymbol{C}\|_2$ to derive an upper bound on $\|\boldsymbol{\Sigma}_{\boldsymbol{z}}\|_2$. Applying the triangle inequality on the expression of $\boldsymbol{C}$ given in (10) gives

$$\begin{aligned}\|\boldsymbol{C}\|_2 \leq{}& 2p\|\boldsymbol{P}\|_2^2 + 2p^2\|\boldsymbol{F}\|_2 + 2\sigma^4\|\boldsymbol{G}\|_2 \\ &+ 2p\sigma^2\left\|\boldsymbol{M}^2\odot\boldsymbol{M}\right\|_2 + 2\sigma^4\|\boldsymbol{A}\odot\boldsymbol{A}\|_2^2 \end{aligned} \tag{58}$$

Substituting (57) into (58) and leveraging the (μ,ν)-coherence assumption on the matrix $\boldsymbol{A}$ yield

$$\begin{aligned}\|\boldsymbol{C}\|_2 &\leq 2p\|\boldsymbol{A}\|_2^4\|\boldsymbol{M}\|_{\max}^2 + 2p^2\|\boldsymbol{F}\|_2 + 2\sigma^4\|\boldsymbol{G}\|_2 \\ &\quad + 2p\sigma^2\|\boldsymbol{A}\|_2^4\|\boldsymbol{M}\|_{\max} + 2\sigma^4\|\boldsymbol{A}\|_2^2\|\boldsymbol{A}\|_{\max}^2 \\ &\leq 4p\frac{n^2}{m^2}\log^2(n)\mu^8 + 2p\frac{n^2}{m^2}\log(n)\mu^6\sigma^2 \\ &\quad + 4\frac{n^2}{m^2}\log(n)\mu^4\sigma^4. \end{aligned} \tag{59}$$

Finally, we can substitute (59) into (56) to obtain

$$\begin{aligned}\|\boldsymbol{\Sigma}_{\boldsymbol{z}}\|_2 &\leq 7pd\frac{n^2}{m^2}\log^2(n)\mu^8 + 2p\frac{n^2}{m^2}\log(n)\mu^6\sigma^2 \\ &\quad + 4\frac{n}{m}\log(n)\mu^4\sigma^4 \\ &\leq 7pd\frac{n^2}{m^2}\log^2(n)\mu^8 \\ &\quad \cdot\left(1 + \frac{2}{7\log(n)}\frac{\sigma^2}{\mu^2} + \frac{4}{7p\log(n)}\frac{\sigma^4}{\mu^4}\right) \end{aligned} \tag{60}$$

We achieve the desired statement with $p = \frac{m}{\beta n}$ and by letting $t = \frac{\log(L)}{7\sqrt{L}}\|\boldsymbol{\Sigma}_{\boldsymbol{z}}\|_2$ in the matrix Bernstein bound (55). □.

REFERENCES

[1] R. Yu, Z. Bai, L. Yang, P. Wang, O. A. Move, and Y. Liu, "A Location Cloaking Algorithm Based on Combinatorial Optimization for Location-Based Services in 5G Networks", *IEEE Access*, vol. 4, pp. 6515–6527, 2016.

[2] S. Tomasin and J. G. L. Hidalgo, "Virtual private mobile network with multiple gateways for b5g location privacy", in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, IEEE, 2021, pp. 1–6.

[3] P. Schmitt and B. Raghavan, "Pretty Good Phone Privacy", in *30th USENIX Security Symposium (USENIX Security 21)*, USENIX Association, 2021, pp. 1737–1754.

[4] M. Bloch and J. Barros, *Physical-layer security: from information theory to security engineering*. Cambridge University Press, 2011.

[5] H. V. Poor and R. F. Schaefer, "Wireless physical layer security", *Proceedings of the National Academy of Sciences*, vol. 114, no. 1, pp. 19–26, 2017.

[6] S. Goel and R. Negi, "Guaranteeing Secrecy Using Artificial Noise", *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[7] S. Tomasin, "Beamforming and Artificial Noise for Cross-Layer Location Privacy of E-Health Cellular Devices", in *2022 IEEE International Conference on Communications Workshops (ICC Workshops)*, May 2022, pp. 568–573.

[8] M. Rajiv and U. Mitra, "Securing BMOCZ Signaling: A Two Layer Artificial Noise Injection Scheme", in *2022 IEEE 23rd International Workshop on Signal Processing Advances in Wireless Communication (SPAWC)*, 2022, pp. 1–5.

[9] M. Krunz and P. Siyari, "Secure Linear Precoding in Overloaded MU-MIMO Wireless Networks", *IEEE Transactions on Communications*, 2023.

[10] Q. E. Zhang, M. Bakshi, and S. Jaggi, "Covert Communication over Adversarially Jammed Channels", in *2018 IEEE Information Theory Workshop (ITW)*, Nov. 2018, pp. 1–5.

[11] R. F. Schaefer, A. Khisti, and H. V. Poor, "Secure Broadcasting Using Independent Secret Keys", *IEEE Transactions on Communications*, vol. 66, no. 2, pp. 644–661, Feb. 2018.

[12] R. Ayyalasomayajula, A. Arun, W. Sun, and D. Bharadia, "Users are Closer than they Appear: Protecting User Location from WiFi APs", in *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications*, 2023, pp. 124–130.

[13] J. J. Checa and S. Tomasin, "Location-Privacy-Preserving Technique for 5G mmWave Devices", *IEEE Communications Letters*, vol. 24, no. 12, pp. 2692–2695, 2020.

[14] J. Li and U. Mitra. "Channel State Information-Free Location-Privacy Enhancement: Fake Path Injection". arXiv: 2307.05442 [eess]. (Jul. 11, 2023), preprint.

[15] D. L. Donoho, "Compressed sensing", *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[16] Y. Zhang, L. Y. Zhang, J. Zhou, L. Liu, F. Chen, and X. He, "A review of compressive sensing in information security field", *IEEE access*, vol. 4, pp. 2507–2519, 2016.

[17] Y. Liang, H. V. Poor, and S. Shamai, *Information theoretic security*. Now Publishers Inc, 2009.

[18] T. Bianchi, V. Bioglio, and E. Magli, "Analysis of one-time random projections for privacy preserving compressed sensing", *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 313–327, 2015.

[19] A. Orsdemir, H. O. Altun, G. Sharma, and M. F. Bocko, "On the security and robustness of encryption via compressed sensing", in *MILCOM 2008-2008 IEEE Military Communications Conference*, IEEE, 2008, pp. 1–7.

[20] Y. Rachlin and D. Baron, "The secrecy of compressed sensing measurements", in *2008 46th Annual Allerton conference on communication, control, and computing*, IEEE, 2008, pp. 813–817.

[21] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing", *IEEE Transactions on information theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[22] S. Choudhary and U. Mitra, "On the Properties of the Rank-Two Null Space of Nonsparse and Canonical-Sparse Blind Deconvolution", *IEEE Transactions on Signal Processing*, vol. 66, no. 14, pp. 3696–3709, 2018.

[23] S. Choudhary and U. Mitra, "Identifiability bounds for bilinear inverse problems", in *2013 Asilomar Conference on Signals, Systems and Computers*, IEEE, 2013, pp. 1677–1681.

[24] M. Ferreira Da Costa and Y. Chi, "Self-calibrated super resolution", in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, IEEE, 2019, pp. 230–234.

[25] Y. Li, K. Lee, and Y. Bresler, "Identifiability in bilinear inverse problems with applications to subspace or sparsity-constrained blind gain and phase calibration", *IEEE Transactions on Information Theory*, vol. 63, no. 2, pp. 822–842, 2016.

[26] K. Lee, N. Tian, and J. Romberg, "Fast and guaranteed blind multichannel deconvolution under a bilinear system model", *IEEE Transactions on Information Theory*, vol. 64, no. 7, pp. 4792–4818, 2018.

[27] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming", *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1711–1732, 2013.

[28] M. Ferreira Da Costa and U. Mitra, "A Framework for Private Communication with Secret Block Structure", in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 5657–5661.

[29] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces", *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.

[30] Y. C. Eldar and H. Bolcskei, "Block-sparsity: Coherence and efficient recovery", in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2009, pp. 2885–2888.

[31] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases", *IEEE transactions on Information theory*, vol. 49, no. 12, pp. 3320–3325, 2003.

[32] K. Rohe, S. Chatterjee, and B. Yu, "Spectral Clustering and the High-Dimensional Stochastic Blockmodel", *The Annals of Statistics*, vol. 39, no. 4, pp. 1878–1915, Aug. 2011.

[33] U. Von Luxburg, "A tutorial on spectral clustering", *Statistics and computing*, vol. 17, pp. 395–416, 2007.

[34] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading", *IEEE Transactions on Information theory*, vol. 45, no. 2, pp. 622–640, 1999.

[35] H.-H. Chen, J.-F. Yeh, and N. Suehiro, "A multicarrier CDMA architecture based on orthogonal complementary codes for new generations of wideband wireless communications", *IEEE communications Magazine*, vol. 39, no. 10, pp. 126–135, 2001.

[36] M. Alam and Q. Zhang, "Non-orthogonal multiple access with sequence block compressed sensing multiuser detection for 5G", *IEEE Access*, vol. 6, pp. 63 058–63 070, 2018.

[37] X. Liu, H.-H. Chen, M. Peng, and F. Yang, "Identical Code Cyclic Shift Multiple Access—A Bridge Between CDMA and NOMA", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 2878–2890, Mar. 2020.

[38] H. Zhu and G. B. Giannakis, "Exploiting Sparse User Activity in Multiuser Detection", *IEEE Transactions on Communications*, vol. 59, no. 2, pp. 454–465, Feb. 2011.

[39] W. Ni and X. Dong, "Hybrid Block Diagonalization for Massive Multiuser MIMO Systems", *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 201–211, Jan. 2016.

[40] T. Wang, L. Shi, K. Cai, L. Tian, and S. Zhang, "Non-coherent NOMA with massive MIMO", *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 134–138, 2019.

[41] L. Liu, C. Yuen, Y. L. Guan, Y. Li, and C. Huang, "Gaussian message passing for overloaded massive MIMO-NOMA", *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 210–226, 2018.

[42] F. R. Bach, "Consistency of the group LASSO and multiple kernel learning.", *Journal of Machine Learning Research*, vol. 9, no. 6, 2008.

[43] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery", *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[44] E. J. Candès, "The restricted isometry property and its implications for compressed sensing", *Comptes rendus mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[45] J. D. Blanchard, C. Cartis, and J. Tanner, "Compressed sensing: How sharp is the restricted isometry property?", *SIAM review*, vol. 53, no. 1, pp. 105–125, 2011.

[46] C. Davis and W. M. Kahan, "The rotation of eigenvectors by a perturbation. III", *SIAM Journal on Numerical Analysis*, vol. 7, no. 1, pp. 1–46, Mar. 1970, Publisher: Society for Industrial and Applied Mathematics.

[47] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information", *IEEE Transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.

[48] M. A. Davenport and J. Romberg, "An overview of low-rank matrix recovery from incomplete observations", *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 4, pp. 608–622, 2016.

[49] J. A. Tropp, "An introduction to matrix concentration inequalities", *Foundations and Trends® in Machine Learning*, vol. 8, no. 1-2, pp. 1–230, 2015.

[50] E. Van Den Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions", *Siam journal on scientific computing*, vol. 31, no. 2, pp. 890–912, 2009.

[51] W. Hoeffding, "Asymptotically optimal tests for multinomial distributions", *The Annals of Mathematical Statistics*, pp. 369–401, 1965.

[52] J. R. Hershey and P. A. Olsen, "Approximating the Kullback Leibler divergence between Gaussian mixture models", in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, IEEE, vol. 4, 2007, pp. IV–317.

[53] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, Feb. 2013.

[54] C. R. Johnson, *Matrix theory and applications*. American Mathematical Soc., 1990, vol. 40.