A FRAMEWORK FOR PRIVATE COMMUNICATION WITH SECRET BLOCK STRUCTURE

Maxime Ferreira Da Costa and Urbashi Mitra

Ming Hsieh Department of Electrical Engineering, Viterbi School of Engineering
University of Southern California, Los Angeles, CA, USA

ABSTRACT

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. We exploit this result to propose a novel private communication framework where the secrecy is achieved by transmitting instances of an unidentifiable compressed sensing problem over a public channel. The legitimate receiver can attempt to overcome this ill-posedness by leveraging secret knowledge of a block structure that was used to encode the transmitter's message. We study the privacy guarantees of this communication protocol to a single transmission, and to multiple transmissions without refreshing the shared secret. Additionally, we propose an algorithm for an eavesdropper to learn the block structure via the method of moments and highlight the privacy benefits of this framework through numerical experiments.

Index Terms— Private communication, inverse problems, structured compressed sensing, block compressed sensing.

1. INTRODUCTION

While security is commonly considered over the *transport layer* and achieved by the mean of cryptographic algorithms, recent advances in *physical layer* security [1], which aims at exploiting the physical properties of a communication channel to discriminate in favor of legitimate parties, have allowed a leap forward for private communication with modern applications to next-generation wireless systems [2]. To that end, the compressed sensing framework [3] has been extensively considers as a mean to ensure privacy [4]. If the sensing matrix is kept secret to an eavesdropper, perfect secrecy can be guaranteed in the information theoretic sense [5] under restrictive conditions [6]. The computational secrecy of this approach have also been discussed [7], [8], restricting Eve's ability to recover the encoded message via a polynomial time algorithm.

Motivated by applications to MIMO systems, we focus here instead on a novel model where the sensing matrix (*e.g.* the channel matrix) is imposed by the environment and not left to design by the transmitter. Privacy is rather achieved
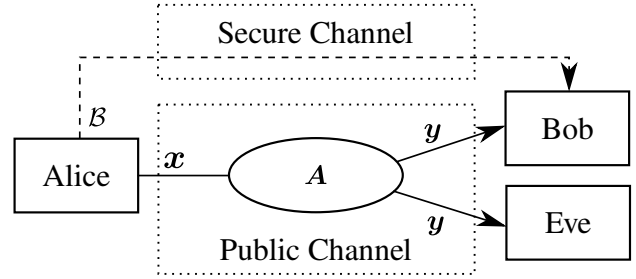
Fig. 1. Communication Model with Secure Channel

by sharing an additional structure on the message with the legitimate receiver, easing the decoding of the message [9]. On the eavesdropper perspective, the decoding amounts to solving a bilinear inverse problem, which are known to demand much stringent assumptions to be identifiable [10]–[13].

1.1. Linear Inverse Problem Based Privacy

We consider the classical secret communication problem with side information: A transmitter (Alice) wishes to privately transmit a vector $\boldsymbol{x} \in \mathbb{R}^N$ to a legitimate receiver (Bob) over a public channel. The channel output $\boldsymbol{y} = f(\boldsymbol{x})$ can be overheard by both Bob and an 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 that is inaccessible by Eve.

In the proposed setting, the effect of the channel is assumed to be linear and noiseless and modelled by a fat matrix $\boldsymbol{A} \in \mathbb{R}^{M \times N}$, with $M < N$ so that $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$. The matrix $\boldsymbol{A}$ is imposed by the environment and is assumed to be known by Bob and Eve. For the purposes of the analysis, we suppose $\boldsymbol{A}$ to be drawn at random with i.i.d. Gaussian entries $a_{i,j} \sim \mathcal{N}\left(0, \frac{1}{M}\right)$. Finally, we assume that Eve is aware of the communication protocol established by Alice. The overall communication model is depicted in Fig. 1.

In order to ensuring privacy, Alice, who designs the message $\boldsymbol{x}$ and the side information, needs to certify two things. 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 achieved by imposing an additional

structure on 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.

1.2. Contributions and Paper Organization

In Section 2, we propose a novel communication protocol that leverages the advantageous recoverability of block-sparse signals in order to ensuring privacy. We provide the encoding and decoding strategies of Alice and Bob, respectively. In our design Alice transmits secretly to Bob a block structure, which can be done at a very low transmission rate, and uses it to encode her message, while the channel matrix A is *not* left for Alice to design. To the authors knowledge, the showcased protocol is the first linear inverse problem based privacy method that does not require A to be secretly shared. Furthermore, Corollary 3 guarantee 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 in Eve's perspective as the signal length increases.

In Section 3, we consider the possibility for Eve to recover the secret block structure from the observation of *multiple* snapshots of the observation $\{y_\ell\}$ that have been generated by Alice with the same block structure $\mathcal{B}$. We show in Corollary 5 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. Finally, we investigate the case of a finite number of snapshots and derive an upper bound of the rate at which Alice must recycle $\mathcal{B}$ to prevent Eve from deciphering Bob's messages. We present numerical results that validate our theoretical findings.

1.3. Notation

Vectors of $\mathbb{R}^n$ and matrices of $\mathbb{R}^{n_1 \times n_2}$ are denoted by boldface letters a and capital boldface letters A respectively. The matrix norms $\|M\|_2$ and $\|M\|_{\max}$ refer to the spectral norm and the maximal absolute value of the entries in M, respectively. The Hadamard product between two matrices M_1 and M_2 is denoted $M_1 \odot M_2$. We write by I_n the identity matrix and by J_n the all-one matrix in appropriate dimension $n \times n$. A block structure over $\mathbb{R}^N$ into R blocks is described by a mapping $\mathcal{B} : [1, \ldots, N] \to [1, \ldots, R]$, and is associated with the indicator matrix of $B \in \mathbb{R}^{N \times N}$ defined by

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

We denote by $x[r]$ the subvector of x with entries x_ℓ verifying $\mathcal{B}(\ell) = r$. The "block-ℓ_0-norm" of a vector x is defined as $\|x\|_{\mathcal{B},0} = \sum_{r=1}^R \mathbf{1}_{x[r] \neq \mathbf{0}}$ and counts the number blocks in x that are not exactly equal to $\mathbf{0}$. For two functions f and g, we use the Landau notations $f = \Omega(g)$, $f = \mathcal{O}(g)$ and $f = o(g)$ to denote that the ratio $\frac{f}{g}$ is asymptotically greater than some constant C, smaller than C, and tends to 0, respectively.

2. PRIVACY WITH BLOCK SPARSITY

2.1. Alice's Encoding

Alice constructs her message x as follows. Given the knowledge of the channel matrix, Alice initializes the communication by drawing a block structure $\mathcal{B} : [1, \ldots, N] \to [1, \ldots, R]$ and sends this structure to Bob over the secret channel. We highlight that this exchange requires less than $R \log_2(N)$ bits of information which is significantly less than schemes relying on exchanging the matrix A. For simplicity, we assume that the R blocks have equal block size d, *i.e.* $N = Rd$. Next, Alice selects a probability of block activation $p \in [0, 1]$, and encodes her message in a block-sparse vector x. In the sequel, we assume that x is distributed according to a block Bernouilli-Gaussian distribution such that

$$x[r] = \begin{cases} \mathbf{0}_d & \text{w.p. } 1 - p \\ z[r] & \text{w.p. } p, \end{cases} \quad (2)$$

where $z[r] \sim \mathcal{N}(\mathbf{0}_d, I_d)$ is a random i.i.d. standard Gaussian vector of dimension d.

2.2. Bob's Decoding

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

$$\widehat{x}_B = \underset{x \in \mathbb{R}^N}{\arg\min} \|x\|_{\mathcal{B},0} \text{ such that } y = Ax. \quad (3)$$

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 [9], [14], [15]. However, directly solving (3) remains NP-hard in the generic case, due to the combinatorics inherent to the minimization of $\|x\|_{\mathcal{B},0}$. Thus, Bob computes instead an estimate of $\widehat{x}_B$ using a polynomial time algorithm of his choice. Among the many addressed in the literature, Block Matching Pursuit (Block MP) [16], Block Basis Pursuit (Block BP) [17] or Block Iterative Harding Thresholding (Block IHT) [9], have been proposed with provable performance guarantees. Proposition 1 recalls that if the block length is large enough and the number of observation per non-zero elements remains constant then Bob can provably recover x.

Proposition 1 (Success of Bob's decoding [9]). *Suppose that A is a matrix with i.i.d. random Gaussian entries. If*

$$\log(p^{-1}) = o(\log(d)); \qquad \frac{M}{pN} = \Omega(1) \quad (4)$$

in the limit where $N \to \infty$, then Bob can recover x asymptotically almost surely.
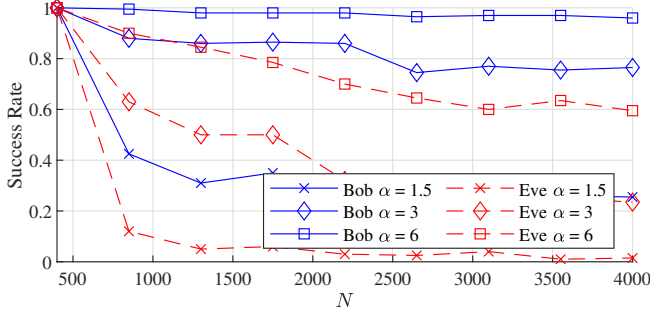
Fig. 2. Success Rate of Bob and Eve to recover $\boldsymbol{x}$ for different values of the ratio $\alpha = \frac{M}{pN}$. Here the parameters are set to $M = 200$, $d = 50$. The results are averaged over 200 trials.

2.3. 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 the perfect knowledge of $\boldsymbol{x}$. Therefore, under her perspective, the best possible approach consists in attempting to recover $\boldsymbol{x}$ without leveraging the existence of a latent block structure in the message. This amounts to solving the classical compressed sensing program

$$\widehat{\boldsymbol{x}}_E = \arg\min_{\boldsymbol{x} \in \mathbb{R}^N} \|\boldsymbol{x}\|_0 \text{ such that } \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}. \qquad (5)$$

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 [18]. In the case of a Gaussian matrix $\boldsymbol{A}$, the following proposition, links the asymptotic failure of (5) as a function of the parameters of the model.

Proposition 2 (Failure of Eve's decoding [19]). *Suppose that $\boldsymbol{A}$ is a matrix with i.i.d. random Gaussian entries. Then if*

$$\frac{M}{pN} = o\left(-\log(p)^{-1}\right) \qquad (6)$$

holds in the limit where $N \to \infty$, then the solution of algorithm (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 p and d so that (4) and (6) are jointly satisfied.

Corollary 3 (Single snapshot privacy). *If Alice selects the determining ratio[1] $\alpha \triangleq \frac{M}{pN} > 1$ with $d \geq \frac{M}{\alpha N}$, then the protocol is asymptotically private to the exchange of a single message in the limit $N \to +\infty$.*

Fig. 2 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 α.

---
[1] The parameter α must be greater than 1 for the problem to be identifiable.

3. EAVESDROPPING VIA HIGHER ORDER MOMENTS

3.1. Structure of the Moments

In order 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 that scenario, if Eve can acquires multiple snapshots of observation $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L\}$ given by $\boldsymbol{y}_\ell = \boldsymbol{A}\boldsymbol{x}_\ell$, $\ell = 1, \ldots, L$, 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, if the expectancy $\mathrm{E}[\boldsymbol{x}] = \boldsymbol{0}_N$ and covariance $\boldsymbol{\Sigma}_{\boldsymbol{x}} = p\boldsymbol{I}_N$ of $\boldsymbol{x}$ carry no information about the block structure $\mathcal{B}$, the even fourth order moments of $\boldsymbol{x}$ given by

$$\mathrm{E}[x_\ell^2 x_{\ell'}^2] = \begin{cases} 3p & \text{if } \ell = \ell' \\ p & \text{if } \mathcal{B}(\ell) = \mathcal{B}(\ell') \text{ and } \ell \neq \ell' \\ p^2 & \text{if } \mathcal{B}(\ell) \neq \mathcal{B}(\ell'), \end{cases} \qquad (7)$$

do encode information about $\mathcal{B}$. Additionally, as the odd fourth order moments of $\boldsymbol{x}$ are equal to zero, Eve can restrict herself to the study of the covariance $\boldsymbol{\Sigma}_{\boldsymbol{v}}$ of the vector $\boldsymbol{v} = (\boldsymbol{A}^\mathsf{T}\boldsymbol{y}) \odot (\boldsymbol{A}^\mathsf{T}\boldsymbol{y})$. The following proposition indicates that the indicator matrix $\boldsymbol{B}$ of the block structure $\mathcal{B}$ can be approximately inferred from a proper translation of $\boldsymbol{\Sigma}_{\boldsymbol{v}}$.

Proposition 4 (Structure of the 4th order moments). *Denote by $\boldsymbol{\Sigma}_{\boldsymbol{v}}$ the covariance matrix of the vector $\boldsymbol{v}$. Moreover let by $\boldsymbol{H} = \beta_1 \boldsymbol{I}_N + \beta_2 \boldsymbol{J}_N$ the matrix given by the coefficients*

$$\beta_1 = 8p - 2p^2, \qquad \beta_2 = 2p^2, \qquad (8)$$

there exists $C > 0$ such that the inequality

$$\left\| \frac{1}{3p(1-p)} (\boldsymbol{\Sigma}_{\boldsymbol{v}} - \boldsymbol{H}) - \boldsymbol{B} \right\|_{\max} \leq \frac{dN\log^2(N)}{M^2} \qquad (9)$$

holds with probability greater than $1 - CN^{-1}$ when $N \to \infty$.

A sketch proof of Proposition 4 is presented in Appendix A. If the parameter p selected by Alice was known to Eve, she could compute $\boldsymbol{H}$ and use the matrix $\widetilde{\boldsymbol{B}} = (3p(1-p))^{-1} (\boldsymbol{\Sigma}_{\boldsymbol{v}} - \boldsymbol{H})$ as a first estimate of $\boldsymbol{B}$. Additionally, since $\boldsymbol{B}$ takes binary values, the matrix $\widehat{\boldsymbol{B}}$ obtained by rounding each entry of $\widetilde{\boldsymbol{B}}$ to its closest value in $\{0, 1\}$ is equal to $\boldsymbol{B}$ whenever the right hand side of (9) is smaller than $1/2$. The messages $\{\boldsymbol{x}_\ell\}$ can subsequently be recovered by Eve, who solves (3) with its estimate $\widehat{\mathcal{B}}$.

Corollary 5 (Asymptotic vulnerability). *If $d = \mathcal{O}\left(\frac{N\log^2(N)}{M^2}\right)$, then the structure $\mathcal{B}$ and the messages $\{\boldsymbol{x}_\ell\}$ are asymptotically identifiable from the fourth order moment $\boldsymbol{\Sigma}_{\boldsymbol{v}}$.*

Algorithm 1 Eavesdropping of the Block Structure

1: **function** ESTIMATEBLOCKS($\boldsymbol{Y}, \boldsymbol{A}, p, d$)
2: $\quad \boldsymbol{V} \leftarrow \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{Y}\right) \odot \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{Y}\right)$
3: $\quad \mathrm{E}[\boldsymbol{v}] \leftarrow p \operatorname{diag}((\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A})^2)$
4: $\quad \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{v}} \leftarrow \frac{1}{L} \sum_{\ell=1}^{L} \left(\boldsymbol{v}_\ell - \mathrm{E}[\boldsymbol{v}]\right)\left(\boldsymbol{v}_\ell - \mathrm{E}[\boldsymbol{v}]\right)^{\mathsf{T}}$
5: $\quad \boldsymbol{H} \leftarrow \beta_1 \boldsymbol{I}_N + \beta_2 \boldsymbol{J}_N \quad \triangleright$ With (β_1, β_2) as given in (8)
6: $\quad \widetilde{\boldsymbol{B}} \leftarrow p^{-1}(1-p)^{-1}\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{v}} - \boldsymbol{H}\right)$
7: $\quad \widehat{\boldsymbol{B}} \leftarrow \mathbf{1}_{\widetilde{\boldsymbol{B}} > \frac{1}{2}} \qquad\qquad \triangleright$ Rounding operation
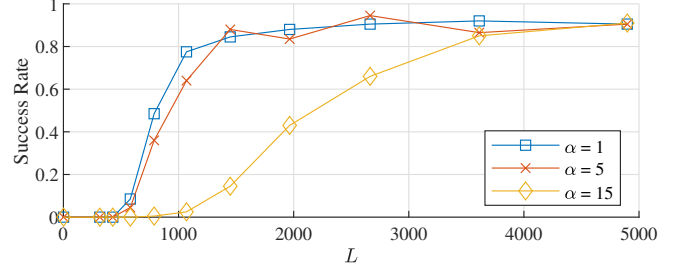8: $\quad$ **return** $\widehat{\boldsymbol{B}}$
9: **end function**

Fig. 3. Success rates of Algorithm 1 to recover the block structure $\mathcal{B}$ as a function of the number of snapshots L and different ratios $\alpha = \frac{M}{pN}$. Herein, $N = 1000$, $M = 400$ and $d = 20$. Results are averaged over 200 trials.

3.2. Estimation with a Finite Number of Snapshots

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

As for the conclusion drawn in Subsection 3.1, Eve can compute the estimator $\widehat{\boldsymbol{B}}$ by rounding the matrix $\widehat{\boldsymbol{B}} = (3p(1-p))^{-1}\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{v}} - \boldsymbol{H}\right)$. This complete procedure is summarized in Algorithm 1. However, the quality of the estimate $\widetilde{\boldsymbol{B}}$ will be worsened by the estimation error on $\boldsymbol{\Sigma}_{\boldsymbol{v}}$. The next lemma provide a high probability bound on this error via the matrix Bernstein inequality (e.g. [20]).

Lemma 6 (Concentration of $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{v}}$). *There exists two constants* C_1, C_2 *such that the inequality*

$$\left\|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{v}} - \boldsymbol{\Sigma}_{\boldsymbol{v}}\right\|_2 \leq \frac{pN^2 \log^2(N) \log(L)}{M^2 \sqrt{d}\sqrt{L}} \qquad (10)$$

holds with probability greater than $1 - C_1 N^{-1} - C_2 L^{-1}$.

Equations (9) and (10) with the bound $\|\cdot\|_{\max} \leq \|\cdot\|_2$ yields an upper bound on the number of snapshots that are necessary of Eve to compromise the protocol.

Corollary 7 (Vulnerability under a finite number of snapshots). *Let* $L_{\mathrm{crit}} \triangleq \frac{N^4 \log^4(N)}{M^4 d}$. *If* $d = \mathcal{O}\left(\frac{N \log^2(N)}{M^2}\right)$ *and* $L = \Omega\left(L_{\mathrm{crit}}\right)$ *then Algorithm 1 recovers the truth structure* $\mathcal{B}$ *with overwhelming probability when* $N \to \infty$.

If Alice follows the scaling $d \sim \frac{M}{\alpha N}$ proposed in Subsection 2.3 the lifespan of $\mathcal{B}$ will be $L_{\max} \sim \frac{\alpha N^5 \log^4(N)}{M^5}$ which suggests that: large channel are more robust to statistical eavesdropping; and that higher values of the determining parameter α increases the lifespan of $\mathcal{B}$. This last observation is corroborated by Fig. 3. However increasing α amounts to decreasing the amount of information transmitted in each message $\boldsymbol{x}_\ell$.

This last observation suggests the existence of a trade-off between the achievable communication rate and the robustness of the privacy of the protocol which is left for future studies.

A. PROOF OF PROPOSITION 3

Herein, we provide a proof sketch of Proposition 4. First, let $\boldsymbol{M} = \boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}$, and denote by $\boldsymbol{Z}$:

$$\boldsymbol{Z} = 2 \operatorname{diag}(\boldsymbol{M}) \operatorname{diag}(\boldsymbol{M})^{\mathsf{T}} \cdot$$
$$\left(2p\boldsymbol{I}_N + p(1-p)\boldsymbol{B} + p^2 \boldsymbol{J}_N\right) \cdot \operatorname{diag}(\boldsymbol{M}) \operatorname{diag}(\boldsymbol{M})^{\mathsf{T}} \qquad (11)$$

with expected value $\mathrm{E}_{\boldsymbol{M}}[\boldsymbol{Z}] = 4p\boldsymbol{I}_N + 2p(1-p)\boldsymbol{B} + 2p^2 \boldsymbol{J}_N$. A direct calculation yields the expression of the covariance $\boldsymbol{\Sigma}_{\boldsymbol{v}}$,

$$\boldsymbol{\Sigma}_{\boldsymbol{v}} = p(1-p)\left(\boldsymbol{M} \odot \boldsymbol{M}\right)\left(2\boldsymbol{I}_N + \boldsymbol{B}\right)\left(\boldsymbol{M} \odot \boldsymbol{M}\right) + \boldsymbol{Z}. \quad (12)$$

Next, as the entries of $\boldsymbol{A}$ are i.i.d Gaussian with $a_{i,j} \sim \mathcal{N}\left(0, \frac{1}{M}\right)$ the event $\max_{i,j} \left|\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle^2 - \delta_{i,j} - \frac{1}{M}\right| \leq \frac{\log(N)}{M}$ holds with probability greater than $1 - C_0 N^{-1}$ for some $C_0 > 0$, which leads after some algebra to

$$\left\|\left(\boldsymbol{M} \odot \boldsymbol{M}\right) \boldsymbol{I}_N \left(\boldsymbol{M} \odot \boldsymbol{M}\right) - \frac{N}{M^2}\boldsymbol{J}_N\right\|_{\max} \leq \frac{N \log^2(N)}{M^2}, \qquad (13a)$$

$$\left\|\left(\boldsymbol{M} \odot \boldsymbol{M}\right) \boldsymbol{B} \left(\boldsymbol{M} \odot \boldsymbol{M}\right) - \frac{dN}{M^2}\boldsymbol{J}_N\right\|_{\max} \leq \frac{dN \log^2(N)}{M^2}. \qquad (13b)$$

The next lemma uses the Hanson-Wright inequality [21], [22] to provide a bound on each entries of the quantity $\boldsymbol{W} = \boldsymbol{Z} - \mathrm{E}_{\boldsymbol{M}}[\boldsymbol{Z}]$.

Lemma 8 (Uniform Hanson-Wright type inequality). *There exists a constant* $C > 0$ *such that the inequality* $\|\boldsymbol{W}\|_{\max} \leq \frac{p\sqrt{dN}\log^6(N)}{M^2}$ *holds with probability greater than* $1 - CN^{-1}$.

One concludes on the desired statement with Lemma 8 by applying the triangle inequality on (12), (13) for a sufficiently large N. $\qquad\square$

References

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

[2] 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.

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

[4] 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.

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

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] S. Choudhary and U. Mitra, "On identifiability in bilinear inverse problems", in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2013, pp. 4325–4329.

[11] 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.

[12] 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.

[13] Y. Li, K. Lee, and Y. Bresler, "A unified framework for identifiability analysis in bilinear inverse problems with applications to subspace and sparsity models", *arXiv preprint arXiv:1501.06120*, 2015.

[14] 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.

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

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

[17] 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.

[18] 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.

[19] 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.

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

[21] M. Rudelson and R. Vershynin, "Hanson-Wright inequality and sub-Gaussian concentration", *Electronic Communications in Probability*, vol. 18, pp. 1–9, 2013.

[22] Y. Klochkov and N. Zhivotovskiy, "Uniform Hanson-Wright type concentration inequalities for unbounded entries via the entropy method", *Electronic Journal of Probability*, vol. 25, pp. 1–30, 2020.