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 considered as a mean to ensure privacy [3]. 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 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 \( x \in \mathbb{R}^N \) to a legitimate receiver (Bob) over a public channel. The channel output \( y = f(x) \) can be overheard by both Bob and an eavesdropper (Eve). To achieve privacy and prevent Eve from recovering the message \( x \), 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 \( A \in \mathbb{R}^{M \times N} \), with \( M < N \) so that \( y = Ax \). The matrix \( A \) is imposed by the environment and is assumed to be known by Bob and Eve. For the purposes of the analysis, we suppose \( A \) be drawn at random with i.i.d. Gaussian entries \( a_{i,j} \sim \mathcal{N}(0, \frac{1}{M}) \). 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 \( x \) and the side information, needs to certify two things. First, Bob must be able to provably recover \( x \) from the observation \( y \) via the side information from the secure channel. Second, Eve cannot provably recover \( x \) 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
2. PRIVACY WITH BLOCK SPARSITY

2.1. Alice’s encoding

Alice constructs her message $x$ has follows. Given the knowledge of the channel matrix, Alice initializes the communication by drawing a block structure $B : [1, \ldots, N] \rightarrow [1, \ldots, R]$ and sends this structure to Bob over the secret channel. We highlight that this exchange only requires $R \log_2(N)$ bits of information which 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 Bernoulli-Gaussian distribution such that

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

where $z[r] \sim N(0, I_d)$ 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 $B$ that was securely sent by Alice to recover the ground truth message $x$. To do so, Bob formulates the block compressed sensing problem

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

Harnessing a block-sparse prior in compressed sensing has been extensively shown in the literature to enable 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 $\|x\|_{B,0}$. Thus, Bob computes instead an estimate of $\hat{x}_{B,\ell}$ 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 of 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

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

in the limit where $N \rightarrow \infty$, then Bob can recover $x$ asymptotically almost surely.
2.3. Privacy Guarantees under a Single Snapshot

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

\[
\hat{x}_E = \arg \min_{x \in \mathbb{R}^N} \|x\|_0 \text{ such that } y = Ax. \tag{5}
\]

The identifiability condition \( x = \hat{x}_E \) of (5) is well-understood to be related to the Restricted Isometry Property (RIP) of the measurement operator \( A \). In the case of a Gaussian matrix \( A \), 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 \( A \) is a matrix with i.i.d. random Gaussian entries. Then if

\[
\frac{M}{p N} = \alpha \left( - \log(p)^{-1} \right) \tag{6}
\]

holds in the limit where \( N \to \infty \), then the solution of algorithm (5) is different from \( x \) with overwhelming probability.

Altogether, Propositions 1 and 2 suggest that, given the dimensions \( M \) and \( N \) of \( A \), 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 determinantal ratio \( \alpha \triangleq \frac{M}{p N} > 1 \) with \( d \geq \frac{M}{N} \), 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 \( x \) via the Block-BP and BP algorithms respectively, for different values of the ratio \( \alpha \).

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 \( B \) in transmitting several independent signals \( \{x_1, \ldots, x_L\} \). In that scenario, if Eve can acquire multiple snapshots of observation \( \{y_1, \ldots, y_L\} \) given by \( y_\ell = Ax_\ell, \ell = 1, \ldots, L, \) and under the knowledge of the prior distribution (2) of \( x \), she can attempt to gain statistical information about \( B \) without having to reconstruct the messages by studying the posterior distribution of \( y \). In particular, if the expectancy \( E[x] = 0_N \) and covariance \( \Sigma_x = p I_N \) of \( x \) carry no information about the block structure \( B \), the even forth order moments of \( x \) given by

\[
E[x_\ell^2] = \begin{cases} 
3p & \text{if } \ell = \ell' \\
0 & \text{if } B(\ell) = B(\ell') \text{ and } \ell \neq \ell' \\
p^2 & \text{if } B(\ell) \neq B(\ell'),
\end{cases} \tag{7}
\]
do encode information about \( B \). Additionally, as the odd fourth order moments of \( x \) are equal to zero, Eve can restrict herself to the study of the covariance \( \Sigma_v \) of the vector \( v = (A^\ell y) \cap (A^\ell y) \). The following proposition indicates that the indicator matrix \( B \) of the block structure \( B \) can be approximately inferred from a proper translation of \( \Sigma_v \).

Proposition 4 (Structure of the 4th order moments). Denote by \( \Sigma_v \) the covariance matrix of the vector \( v \). Moreover let by \( H = \beta_1 I_N + \beta_2 J_N \) the matrix given by the coefficients

\[
\beta_1 = 8p - 2p^2, \quad \beta_2 = 2p^2, \tag{8}
\]

there exists \( C > 0 \) such that the inequality

\[
\left\| \frac{1}{3p(1-p)} \left( \Sigma_v - H \right) - B \right\|_{\text{max}} \leq \frac{d N \log^2(N)}{M^2} \tag{9}
\]

holds with pr. greater than \( 1 - CN^{-1} \) when \( N \to \infty \).

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

Corollary 5 (Asymptotic vulnerability). If \( d = O \left( \frac{N \log^2(N)}{M^2} \right) \), then the structure \( B \) and the messages \( \{x_\ell\} \) are asymptotically identifiable from the fourth order moment \( \Sigma_v \).

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 $B$. Consequently, the true covariance $\Sigma_v$ always remains unknown to Eve. Instead, she can attempt to estimate $B$ from the empirical estimator of the covariance given by $\hat{\Sigma}_v = \frac{1}{T} \sum_{t=1}^{T} (v_t - \mathbb{E}[v])(v_t - \mathbb{E}[v])^T$, where $\mathbb{E}[v] = p \text{diag} \left( (A^T A)^2 \right)$ and diag(·) 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 $\hat{B}$ by rounding the matrix $\hat{B} = (3p(1-p))^{-1} \left( \hat{\Sigma}_v - H \right)$. This complete procedure is summarized in Algorithm 1. However, the quality of the estimate $\hat{B}$ will be worsened by the estimation error on $\Sigma_v$. The next lemma provide a high probability bound on this error via the matrix Bernstein inequality (e.g. [20]).

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

$$\| \hat{\Sigma}_v - \Sigma_v \|_2 \leq \frac{pN^2 \log^2(N) \log(L)}{M^2d/N}$$

holds with probability greater than $1 - C_1 N^{-1} - C_2 L^{-1}$. Equations (3) 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 $L_{\text{crit}} \triangleq N^4 \log^4(N)/M^4d$. If $d = O \left( N \log^2(N)/M^2 \right)$ and $L = \Omega \left( L_{\text{crit}} \right)$ then Algorithm 1 recovers the truth structure $B$ 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 $B$ will be $L_{\max} \sim \frac{\alpha N^3 \log^4(N)}{M^2}$ which suggests that: large channel are more robust to statistical eavesdropping; and that higher values of the determinantal parameter $\alpha$ increases the lifespan of $B$. This last observation is corroborated by Fig. 3. However increasing $\alpha$ amounts to decreasing the amount of information transmitted in each message $x_I$. This last observation suggests the existence of a trade-off between the achievable communication rate and the robustness of the privacy of the proposed protocol which is proposed for future study.

---

**Algorithm 1** Eavesdropping of the Block Structure

1. **function** \textsc{EstimateBlocks}(Y, A, p, d)
2. \hspace{1em} $V \leftarrow (A^T Y) \circ (A^T Y)$
3. \hspace{1em} $E[v] \leftarrow p \text{diag}((A^T A)^2)$
4. \hspace{1em} $\hat{\Sigma}_v \leftarrow \frac{1}{T} \sum_{t=1}^{L} (v_t - \mathbb{E}[v])(v_t - \mathbb{E}[v])^T$
5. \hspace{1em} $H \leftarrow \beta_1 I_N + \beta_2 J_N$ \hspace{1em} $\triangleright$ With $(\beta_1, \beta_2)$ as given in [8]
6. \hspace{1em} $B \leftarrow p^{-1}(1-p)^{-1} \left( \hat{\Sigma}_v - H \right)$
7. \hspace{1em} $\hat{B} \leftarrow \hat{1}_{B > \frac{\ell}{2}}$ \hspace{1em} $\triangleright$ Rounding operation
8. \hspace{1em} return $\hat{B}$

---

**A. PROOF OF PROPOSITION 3**

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

$$Z = 2 \text{diag}(M) \text{diag}(M)^T.$$  

\begin{equation}
(2pI_N + p(1-p)B + p^2J_N) \cdot \text{diag}(M) \text{diag}(M)^T
\end{equation}

with expected value $E_M[Z] = 4pI_N + 2p(1-p)B + 2p^2J_N$. A direct calculation yields the expression of the covariance $\Sigma_v$,

$$\Sigma_v = p(1-p) (M \circ M) (2I_N + B) (M \circ M) + Z.$$  

Next, as the entries of $A$ are i.i.d Gaussian with $a_{i,j} \sim N(0, \frac{1}{M})$ there exists the event $\max_{i,j} \left| \frac{a_{i,j}}{\sqrt{M}} \right| = \frac{\log(N)}{\sqrt{M}}$ holds with probability greater than $1 - C_0 N^{-1}$ for some $C_0 > 0$, which leads after some algebra to

\begin{equation}
\left\| (M \circ M) I_N (M \circ M) \right\|_{\max} \leq \frac{N \log^2(N)}{M^2},
\end{equation}

\begin{equation}
\left\| (M \circ M) B (M \circ M) \right\|_{\max} \leq \frac{dN \log^2(N)}{M^2}.
\end{equation}

The next Lemma, obtained uniformly controlling each entries using Hanson-Wright inequality [21], [22], proposes a bound on the quantity $W = Z - E_M[Z]$.

**Lemma 8** (Uniform Hanson-Wright type inequality). There exists a constant $C > 0$ such that the inequality $\| W \|_{\max} \leq \frac{\sqrt{d} \log^2(N)}{M^2}$ 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 $N$. 

---

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