A Theoretically Novel Trade-off for Sparse Secret-key Generation

Makan Zamanipour, Member, IEEE

Abstract—We in this paper theoretically go over a rate-distortion based sparse dictionary learning problem. We show that the Degrees-of-Freedom (DoF) interested to be calculated – standing for the minimal set that guarantees our rate-distortion trade-off – are basically accessible through a Langevin equation. We indeed explore that the relative time evolution of DoF, i.e., the transition jumps is the essential issue for a relaxation over the relative optimisation problem. We subsequently prove the aforementioned relaxation through the Graphon principle w.r.t. a stochastic Chordal Schramm-Loewner evolution, etc. via a minimisation over a distortion between the relative realisation times of two given graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ as $\min_{\mathcal{G}_1, \mathcal{G}_2} \mathcal{D}(\mathcal{G}_1, \mathcal{G}_2)$. We also extend our scenario to the eavesdropping case. We finally prove the efficiency of our proposed scheme via simulations.

Index Terms—NP-hard, Alice, Bochner’s theorem, Bob, Chordal Schramm-Loewner evolution, dictionary learning, distortion function, Eve, harmonic functions, Herglotz-Riesz theorem, holomorphic functions, probabilistic constraint, Riemannian gradients, side-information, sparse coding, sparse patterns, tangent space.

I. INTRODUCTION

Any machine learning and modern signal processing applications — such as bio-metric authentication/identification and recommending systems — follow sparse signal processing techniques [1], [2], [3], [4], [5], [6]. The sparse synthesis model focuses on those data sets that can be approximated using a linear combination of only a small number of cells of a dictionary. The applications of sparse coding based dictionary learning go chiefly over either or both extraction and estimation of local features. Typically, this kind of discipline is controlled via a prior decomposition of the original signal into overlapping blocks — call side-information. In relation to this strategy, there mainly exist two demerits: (i) each point in the signal is estimated multiple times — something that shows a redundancy; and (ii) a shifted version of the features interested to be learnt are captured since the correlations/side-information among neighboring data sets are not fully taken into account — something that results in the inevitable fact that some portions of the learnt dictionaries may be practically of a partially useless nature.

A quick overview over the literature in terms of key-generation and authentication is strongly widely provided here in details. In [7] and for cloud-assisted autonomous vehicles, Q. Jiang et al. technically proposed a bio-metric privacy preserving three-factor authentication and key agreement. In [8] and through vehicular crowd-sourcing, F. Song et al. technically proposed a privacy-preserving task matching with threshold similarity search. In [9] and for mobile Internet-of-Things (IoT), X. Zeng, et al. technically proposed a highly effective anonymous user authentication protocol. In [10] and for fog-assisted IoT, J. Zhang et al. technically proposed a revocable and privacy-preserving decentralized data sharing scheme. In [11] and for resource-constrained IoT devices, Q. Hu et al. technically proposed a half-duplex mode based secure key generation method. In [12] and for mobile crowdsensing under un-trusted platform, Z. Wang et al. theoretically explored privacy-driven truthful incentives. In [13] and for IoT-enabled probabilistic smart contracts, N. S. Patel et al. theoretically explored blockchain-envisioned trusted random estimators. In [14] and for IoT, R. Chaudhary et al. theoretically explored Lattice-based public key crypto-mechanisms. In [15] and for vehicle to vehicle communication, V. Hassija et al. theoretically explored a scheme with the aid of directed acyclic graph and game theory. In [16] and for vehicular social networks, J. Sun et al. theoretically explored a secure flexible and tampering-resistant data sharing mechanism. In [17] and for mission-critical IoT applications, H. Wang et al. theoretically explored a secure short-packet communications. In [18] and for edge-assisted IoT, P. Gope et al. theoretically proposed a highly effective privacy-preserving authenticated key agreement framework.

Overall, many works have been fulfilled in the context of secret key-generation, however, sparse coding and an optimised problem in this area has not been introduced so far. In addition, although the literature has tried its best to propose novel frameworks [7], [9], [10], [13], [16], [17], [18] as well as the optimisation techniques [8], [11], [12], [14], [15] and the relative relaxations, the population of the research and work done in this context still essentially lacks and requires to be further enhanced.

A. Motivations and contributions

In this paper, we are interested in responding to the following question: How can we guarantee highly adequate relaxations over a rate-distortion based sparse dictionary learning? How is Alice able to efficiently utilise totally disparate resources to guarantee her private message to be concealed from Eve? With regard to the non-complete version of the literature, the expressed question strongly motivate us to find an interesting solution, according to which our contributions are fundamentally described as follows.
PROBLEM & SOLUTION:
To find the minimal set which guarantees a rate-distortion trade-off
DEGREES-of-FREEDOM:
The minimal set expressed above which theoretically talk about the transition jumps
TRADE-OFF:
Rate-distortion problem
REASON of SIDE-INFORMATION:
To control the sparse coding
SPARSE SECRET-KEY GENERATION:
To guarantee a correlated minimal set available between Bob and Eve
TIME EVOLUTION of DoF:
Via Langevin equation
RELAXATION of the RATE-DISTORTION PROBLEM:
Via the Graphon principle w.r.t. a Chordal Schramm-Loewner evolution
LOGIC of TIME in OPTIMISATION:
Instantaneous optimisation

Fig. 1: Essential details of our problem-and solution and its flow.

- (i) We initially write a rate-distortion based sparse coding optimisation problem. Subsequently, we try to add some extra constraints which are of a purely meaningful nature in this context.
- (ii) We show that we can get access to the time evolution of the Degrees-of-Freedom (DoF) in which we are interested. We do this through a Langevin equation.
- (iii) We then add an additional constraint in relation to the side-information.
- (iv) Pen-ultimately, we make trials to theoretically relax the resultant optimisation problem w.r.t. the facts of Graphon and a stochastic Chordal Schramm-Loewner evolution etc. – something that can be chiefly elaborated according to some theoretical principles such as Riemannian gradients in Riemannian geometry and the Kroner’s graph entropy.
- (v) Finally, we also extend our scenario to the eavesdropping case. We in fact technically define some probabilistic constraints to be added to the main optimisation problem.

A. Flow of the problem-and-solution
Initially speaking, take a quick look at Fig. 1. It basically says:
- Problem-and-Solution is to find the minimal set which guarantees a rate-distortion trade-off;
- DoF stands fundamentally theoretically for the minimal set expressed above which essentially talk about the transition jumps;
- Trade-off is the rate-distortion problem;
- Reason of side-information and the logic behind of using that is to control the sparse coding
- Sparse secret-key generation and the logic behind of using that is to guarantee a correlated minimal set available between Bob and Eve;
- Time evolution of DoF is explored via a Langevin equation;
- Relaxation of the rate-distortion problem is interpreted through the Graphon principle and w.r.t. a stochastic Chordal Schramm-Loewner evolution; and
- Logic of time in optimisation is to guarantee an instantaneous optimisation.

B. General notation
The notations widely used throughout the paper is given in Table 1.

C. Organisation
The rest of the paper is organised as follows. The system set-up and our main results are given in Sections II and III. Subsequently, the evaluation of the framework and conclusions are given in Sections IV and V.

II. SYSTEM MODEL AND PROBLEM FORMULATION
In this section, we describe the system model, subsequently, we formulate the basis of our problem.

1Something that is traditional as mentioned in the paper in the next parts.
2According to the Kerckchoffs’s principle in cryptography [1].
Table I: List of notations.

| Notation | Definition | Notation | Definition |
|----------|-----------|----------|-----------|
| $\mathcal{P}_1$ | Problem 1 | $\mathcal{P}_2$ | Problem 2 |
| $\mathcal{P}_3$ | Problem 3 | $\mathcal{P}_3^{(i)}$ | New Version of Problem 3 |
| $\theta$ | Perturbation | $E$ | Expected-value |
| $t$ | Time Instance | $\mathcal{A}$ | Side-information |
| $\sim$ | Distributed | $\mathcal{H}, i \in \{1, 2\}$ | Random-walk |
| $\xi$ | Specific Time Instance | $\lambda_i, i \in \{1, 2, \ldots, 7\}$ | Arbitrary Threshold |
| $\mathcal{D}(\cdot)$ | Distortion Function | $\mathcal{I}(\cdot)$ | Mutual Information |
| $\mathcal{G}_i, i \in \{1, 2\}$ | Graph | $\mathcal{R}_i^{(1)}, i \in \{1, 2\}$ | Brownian Motion |

- Server – Step 1 as an indexing for the sender: The server marks some indexes on the received sparse-codes in a database.
- Client: Via the shared sparsifying transform, Bob as the client generates a sparse representation from his query data and subsequently sends it to the server.
- Server – Step 2 as a matching test for the client: The server searches to find the most similar sparse-code set, and subsequently replies to the client’s request.

In fact, Alice has a secret message while she decides to send it to Bob in the presence of the passive – potential – or/and active – actual – Eve(s).

C. Main problem

A sparse dictionary learning problem can be mathematically formulated as [1], [2], [3], [4], [5], [6]

$$
\mathcal{P}_1: \min_{\mathcal{A}, \Phi_0(\cdot)} \mathbb{E}_{\theta \sim \Theta} \mathcal{D}\left(\mathcal{Y}, \mathcal{A}\Phi_0(\mathcal{X}; \theta)\right),
$$

$$
\text{s.t.} \quad |\mathcal{A}| \leq \lambda_0,
$$

$$
|\Phi_0(\mathcal{X})| \leq \lambda_1,
$$

with regard to the thresholds $\lambda_0$ and $\lambda_1$, where $\mathcal{Y}$ denotes the observed data, $\mathcal{A}$ represents the unknown dictionary, $\mathcal{X}$ stands for the original data set w.r.t. the data instance $x \in \mathbb{R}^n$, $\Phi_0(\mathcal{X})$ refers to the sparse representation coefficient set, $\Theta$ is a distribution over appropriate perturbations $\theta \sim \Theta$ [6], while we have the pre-trained models $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\Phi_0 \subset \Phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$.

**Definition 1**: $\alpha$–secret key. A random variable $\mathcal{K}$ with the finite range $\mathcal{K}$ information-theoretically represents an $\alpha$–secret key for Alice and Bob, achievable through the public communication $\beta$, if there exist two functions $\mathcal{F}_a$ and $\mathcal{F}_b$ such that for any $\alpha > 0$ the following three conditions hold:

(i) $\mathcal{P}_1(\mathcal{K} = \mathcal{F}_a(\mathcal{X}, \beta) = \mathcal{F}_b(\mathcal{Y}, \beta)) \geq 1 - \alpha$;

(ii) $\mathcal{F}(\mathcal{X}, \beta) \leq \alpha$; and

(iii) $\mathcal{H}(\mathcal{K}) \geq \log|\mathcal{K}| - \alpha$.

These three conditions respectively ensure that (i) Alice and Bob generate the same secret key with high probability as much as possible; the generated secret key is efficiently kept hide from any user except for Bob who gets access to the public communication $\beta$; and (iii) the secret key is sub-uniformly distributed.

This definition is based fully upon a passive Eve which can be extended to an active scenario – although we in the next parts, use the active one.

III. MAIN RESULTS

In this section, our main results are theoretically provided in details.

A. Time evolution of the sparsity

**Proposition 1**: The time evolution of the sparsity values, that is, as DoF relating to the transition jumps for the rate-distortion based sparse coding optimisation problem $\mathcal{P}_1$ expressed in Eqn. (1) can be mathematically attainable for the distortion term $\mathcal{D}\left(\Phi_1(\mathcal{X}^{(t)}), \Phi_0(\mathcal{X}^{(t)})\right)$ while $\Phi_1(\mathcal{X}^{(t)}) := \Phi(\Phi_0(\mathcal{X}^{(t)}); \theta^{(t)}), t$ for the instant $t$.

**Proof**: See Appendix A. ■

B. Relaxation of the rate-distortion problem

**Proposition 2**: While the degrees of freedom of interest so far is $\Phi_1(\mathcal{X}^{(t)})$, following the previous result, i.e., Proposition 1 and its relative proof given in Appendix A, the main problem over the distortion term $\mathcal{D}\left(\Phi_1(\mathcal{X}^{(t)}), \Phi_0(\mathcal{X}^{(t)})\right)$ is NP-hard. However, this problem can be theoretically relaxable through a new definition over the given side-information $\mathcal{A}$ as a new distortion constraint $\mathcal{D}\left(\Phi_0(\mathcal{X}^{(t)}); \mathcal{A}^{(t)}\right)$.

**Proof**: See Appendix B. ■

**Algorithm 1**: A greedy algorithm to the Problem $\mathcal{P}_2$.

**Initialisation**.

while TRUE do

(i) Solve $\mathcal{P}_2$ given in Eqn. (6) and (ii) Update.

endwhile

**Output**.

end

\[3\text{See e.g. [19], [20].}\]
C. Positive definity

Proposition 3: DoF, if they are non-zero, they are conditionally positive definite.

Proof: See Appendix C. □

D. Extension to the eavesdropping scenario

Assume$^4$ that we have two users, without loss of generality one of which is Bob as the legitimate user, and another one is an eavesdropper – call Eve – who can get access to the information in relation to the private message arose between Alice and Bob.

Proposition 4: For the $t$–th instant, call the sparsity patterns $S_{ab}^{(t)} \in \{0, 1\}$ and $S_{ae}^{(t)} \in \{0, 1\}$ respectively for Bob and Eve, that is, in relation to the channel between Alice and Bob, and the one between Alice and Eve, respectively. For the $t$–th instant and $j$–th sub-channel, define $\mathcal{P} = \{j \in S_{ae}^{(t)} : j \in S_{ab}^{(t)} = \psi(t), \forall j \in \mathbb{N}\}$ The Problem $\mathcal{P}_2$ expressed in Eqn. (6) can be re-written as the Problem $\mathcal{P}_3$ expressed in Eqn. (7), or the Problem $\mathcal{P}_3'$ expressed in Eqn. (8) or even the Problem $\mathcal{P}_5'$ expressed in Eqn. (9).

Proof: See Appendix D. □

IV. Numerical results

We have done our simulations w.r.t. the Bernoulli-distributed data-sets using GNU Octave of version 4.2.2 on Ubuntu 16.04.

Fig. 2 shows the optimum value earned from the Problem $\mathcal{P}_2$ given in Eqn. (6) while $\lambda_2$ and $\lambda_3$ are $\ll \infty$ divided by the one when $\lambda_2 = \lambda_3 = \infty$, i.e., the case in which we do not constrain our optimisation problem w.r.t. the relative constraints. As totally, obvious, the more constraints we intelligently and purposefully provide, the significantly more favourable response we can undoubtedly guarantee with regard to our proposed scheme.

Fig. 3 shows the instantaneous key-rate $\mathcal{R}_key^{(t)}$ earned from the Problem $\mathcal{P}_2''$ given in Eqn. (9) while $\lambda_2$ and $\lambda_3$ are $\ll \infty$ divided by the one when $\lambda_2 = \lambda_3 = \infty$, i.e., the case in which we do not constrain our optimisation problem w.r.t. the relative constraints. The same as the previous trial, the more constraints we smartly provide, the considerably more adequate response we consequently assure in relation to our proposed scheme.

Fig. 4 shows the instantaneous key-rate $\mathcal{R}_key^{(t)}$ against the iteration regime while it is proorted version of the with case to the without one of the first constraint given in Eqn. (4), i.e., relating to the threshold $\lambda^{(1)}$.

Finally, the outage probability $1 - \lambda_0$ versus the normalised version of the Key-Rate $\mathcal{R}_key^{(t)}$ in Eqn. (9) is shown in Fig. 5 while it is proorted of $||\Delta \xi||$ = 15% to $||\Delta \xi||$ = 5%, where $|| \cdot ||$ stands for the Frobenius norm.

$^4$See e.g. [21], [22].

$^5$See e.g. [21], [22].

V. Conclusion

We have theoretically shown a novel solution to an $NP$–hard rate-distortion based sparse dictionary learning in the context of a novel method of relaxation. We theoretically showed the accessibility of DoF in which er are interested via a Langevin equation. In fact, the relative time evolution of DoF, i.e., the transition jumps for a relaxation over the relative optimisation problem was calculated in the context of a tractable fashion. The aforementioned relaxation was proven with the aid of the Graphon principle w.r.t. a stochastic Chordal Schramm-Loewner evolution etc. Our proposed method and solutions were extended to the eavesdropping case. The effectiveness of our proposed framework was ultimately shown by simulations.
The problem $\mathcal{P}_1$ written in Eqn. (1) can be theoretically re-expressed as

$$\min_{\Phi(\mathcal{X})} \mathbb{E}_{\theta \sim \mathcal{D}} \left\{ \mathcal{D}\left( \Phi\left(\mathcal{X}; \theta, \Phi(\mathcal{X})\right), \Phi(\mathcal{X})\right) \right\},$$

(2)

s.t. $|\Phi(\mathcal{X})| \leq \lambda_1$, while $\Phi_1 \subset \Phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ for which $\Phi_1 \supset \Phi_0$ holds, or correspondingly in the context of an instantaneous\(^6\) optimisation

\(^6\)See e.g. [23], [24] to understand what an instantaneous expected-value it means.

problem

$$\min_{\Phi(\mathcal{X})} \mathbb{E}_{\theta \sim \mathcal{D}} \left\{ \mathcal{D}\left( \Phi\left(\mathcal{X}^{(i)}; \theta^{(i)}, \Phi_0(\mathcal{X}^{(i)})\right), \Phi_1(\mathcal{X}^{(i)})\right) \right\},$$

(3)

s.t. $|\Phi_0(\mathcal{X}^{(i)})| \leq \lambda_1^{(i)}$.

Now, as we are interested in theoretically exploring the sparsity, as DoF in relation to the transition jumps, one can apply the Langevin equation according to which the time evolution of the aforementioned DoF can be accessible. Thus, we have

$$\min_{\Phi(\mathcal{X})} \mathbb{E}_{\theta \sim \mathcal{D}} \left\{ \mathcal{D}\left( \Phi\left(\mathcal{X}^{(i)}; \theta^{(i)}, \Phi_0(\mathcal{X}^{(i)})\right), \Phi_1(\mathcal{X}^{(i)})\right) \right\},$$

(4)

s.t. $|\Phi_0(\mathcal{X}^{(i)})| \leq \lambda_1^{(i)}$, $d\left(\Phi_1(\mathcal{X}^{(i)})\right) = -\gamma_0^{(i)} \Phi_1(\mathcal{X}^{(i)}) + \mathbb{W}_1^{(i)}$, while $\gamma_0^{(i)}$ is an arbitrary parameter, and the minus sign principally shows a backward orientation in the solution to the problem, as well as the theoretical fact that $\mathbb{W}_1^{(i)}$ is a random-walk.

The proof is now completed. ■

**APPENDIX B**

**PROOF OF PROPOSITION 2**

The proof is provided here in terms of the following multi-step solution.

**STEP 1.** It is strongly conventional [25], [26] to define a constraint over a distortion between the side-information and the sparse data set. Therefore, we do this here as well according to which our optimisation problem is conveniently re-written as

$$\min_{\Phi_0(\mathcal{X})} \mathbb{E}_{\theta \sim \mathcal{D}} \left\{ \mathcal{D}\left( \Phi\left(\mathcal{X}^{(i)}; \theta^{(i)}, \Phi_0(\mathcal{X}^{(i)})\right), \Phi_0(\mathcal{X}^{(i)})\right) \right\},$$

(5)

w.r.t. the arbitrary threshold $\lambda_2^{(i)}$.

**STEP 2.** One can here say that DoF cited above can be in terms of the possibly available ways via which one can calculate a specific amount of information. Accordingly, the possibly available pieces of a specific amount of information calculable such as the relative side-information set can be DoF intersted here to be calculated.

**STEP 3.** Now, assume that the possibly available pieces of the specific amount of information calculable talked above are
in the context of probabilistic energy\textsuperscript{7} quanta. Subsequently, we assume a given Ball $\mathcal{B}_0$ in which the number of energy quanta $\epsilon$ should require
\[ \mathbb{P}(\epsilon \leq \epsilon_0), \]
while $\epsilon$ is in relation to $e$, and we have a $2 - d$ probability density function (pdf) $f(\epsilon, p, t)$ while $e$, $p$ and $t$ stand for the energy, momentum and time, respectively. Subsequently, we assign
\[ d\mathcal{B}_0 = d^2 e d^2 p, \]
while $d^2$ is the second derivative in the polar axis. Thus, defining
\[ d\Phi_1(\mathcal{X}^{(t)}) := f(\epsilon, p, t) d\mathcal{B}_0, \]
the relative energy quanta numbers is calculated as
\[ \Phi_1(\mathcal{X}^{(t)}) := \int_{\mathcal{B}_0} f(\epsilon, p, t) d\mathcal{B}_0 = \int_p \int_\theta f(\epsilon, p, t, p, \theta) \cdots \]
while $\epsilon$ is an arbitary threshold to constraint the amount of information calculable in the previous step, for each relative DoF as the users.

**Remark 1.** In our game, the population physically has a positive growth rate. This is because of the fact that the principle of Diffusion-And-Reaction hold here, that is, a diffusive flow exists from the side of a higher concentration to a less one, and reciprocally, a reactive movement exists from the less concentration towards the higher one.

**Step 5.** Now, let us mathematically define an initial condition for the Langevin based constraint. We have $\Phi_1^{(0)}(\mathcal{X}^{(t_0)})$. Now, call the Picard-Lindelof Theorem\textsuperscript{8} which can be of a purely useful nature to say that: the Langevin equation
\[ \frac{d\Phi_1(\mathcal{X}^{(t)})}{dt} = -\gamma_0^{(t)} \Phi_1(\mathcal{X}^{(t)}) + \mathcal{W}_1^{(t)}, \]
w.r.t. the initial condition $\Phi_1^{(0)}(\mathcal{X}^{(t_0)})$ can have a unique solution $\Phi_1^{(t)}(\mathcal{X}^{(t)})$ w.r.t. a surely existing $\xi > 0$ over the zone $[t_0 - \xi, t_0 + \xi]$; while $\Phi_1(\mathcal{X}^{(t)})$ is uniformly Lipschitz continuous in $\Phi_0(\mathcal{X}^{(t)})$ and continuous in $t$. Thus, we indeed branch and localise the optimisation problem into its localities, without loss of generality and optimality. This is because of the NP-hard structure of our main optimisation problem since that is not convex, so, the strong duality does not theoretically hold [31].

**Step 6.** Re-call Remark 1 for the fact that the population size of the relative users has a growth over the time zone, according to which we can theoretically consider a Graphon\textsuperscript{9} in between. Indeed, we can prove that there exists a duality
\[ \text{between the main optimisation problem and its localised version created in Step 5. Now, w.r.t. the fact that there would exist a stochastic Chordal\textsuperscript{10} Schramm-Loewner evolution}\textsuperscript{11} which shows a Brownian motion over the surface of the relative evolving Riemannian-manifold relating to the graph
\[ \lim_{\Delta t \to 0} \frac{\Phi_0(\mathcal{X}^{(t, \xi + \Delta t)}) - \Phi_0(\mathcal{X}^{(t, \xi)})}{\Delta t} = \mathcal{F}(\Phi_0(\mathcal{X}^{(t, \xi)}), \mathcal{W}_2^{(t, \xi)}), \]
w.r.t. the arbitrary function $\mathcal{F}(\cdot)$ and the random-walk $\mathcal{W}_2^{(t, \xi)}$. On other hand, since a graphon – as a density measure – is the number of times $\mathcal{G}$ occurs as a sub-graph of the graph $\mathcal{G}$, and w.r.t. the fact that two graphs, do they have the final approximation in terms of a same graphon, they actualise the same phenomenon [34], and regarding the fact that the positive growth of the population size literally defined above highly requires an adjacent matrix for the evolving Riemannian-manifold and the graph in between, a proof for a the duality we are looking for can be proven. Now, the duality basically holds iff and only iff $\mathcal{G}_1 \in \mathcal{G}$ and $\mathcal{G}_2 \in \mathcal{G}$ conclude a common phenomenon, while $\mathcal{G}_2$ fundamentally is a stochastic Schramm-Loewner evolutioned version of $\mathcal{G}_1$ in the nearest physical vicinity of the Reimannian sub-manifold, relating to the surface of the relative evolving Riemannian-manifold, that is
\[ \mathcal{D}(t(\mathcal{G}_1, \mathcal{G}), t(\mathcal{G}_2, \mathcal{G})) \to \nu_t^+, \]
while $\nu_t^+$ is a sufficiently small non-zero value and $t(\mathcal{G}_1, \mathcal{G})$ and $t(\mathcal{G}_2, \mathcal{G})$ respectively show the final realisation of the relative graphs $\mathcal{G}_1$ and $\mathcal{G}_2$. This is equivalent to\textsuperscript{12}
\[ \min_{\mathcal{G}_1, \mathcal{G}_2} \mathcal{D}(t(\mathcal{G}_1, \mathcal{G}), t(\mathcal{G}_2, \mathcal{G})). \]

Towards such end, we re-express our optimisation problem as Eqn. (6) on the top of the next page in which the term $\xi + \Delta \xi$ indicates the stochastic Schramm-Loewner evolution expressed above and $\lambda_3^{(t)}$ is an arbitrary threshold to constraint the amount of the Schramm-Loewner evolution. Now, let us continue by the following definition.

**Definition 2\textsuperscript{13}: The $d$-Order Binomial Extension.** Consider the mapping $\mathcal{V} : \mathcal{A} \to \mathcal{B}$ such that $B = \mathcal{V}(\mathcal{A})$ holds – and a $d$ which is a discrete random variable that takes values on the non-zero negative integers $\mathbb{Z} \setminus \{0\}$. The mapping $\mathcal{V}^{\oplus d} : \mathcal{A} \to \mathcal{B}^d$ such that $\mathcal{B}^d = \mathcal{V}^{\oplus d}(\mathcal{A})$ holds – is theoretically called the $d$-order
\[ \text{mapping.} \]

\textsuperscript{7}In order to see the relevance between energy and information see e.g. [27], [28], [29].
\textsuperscript{8}See e.g. [30] to see what it is.
\textsuperscript{9}See e.g. [6].
\textsuperscript{10}Which maps a boundary point to another one, which is in contrast to the Radial Schramm-Loewner evolution which maps a boundary point to an interior one; See e.g. [32], [33] to see what they are.
\textsuperscript{11}See e.g. [32], [33] to see what it is.
\textsuperscript{12}This kind of optimisation problem can be solved by the alternating direction method of multipliers; See e.g. [27], [28], [29].
\textsuperscript{13}See e.g. [35].
since the terms \( t \)

\[
\mathcal{P}_2 : \begin{cases}
\min_{\Phi_0 \in \mathcal{X}(t, \xi + \Delta \xi)} \mathbb{E}_{\theta(t, \xi + \Delta \xi) \in \Theta} \left\{ \Delta \left( \Phi_1 \left( \mathcal{X}(t, \xi + \Delta \xi) \right), \Phi_0 \left( \mathcal{X}(t, \xi + \Delta \xi) \right) \right) \right\},
\end{cases}
\]

s.t. 6.a : \( |\Phi_0(\mathcal{X}(t, \xi + \Delta \xi))| \leq \lambda_1^{(t, \xi + \Delta \xi)}, \)

6.b : \( \frac{d}{dt} \Phi_0(\mathcal{X}(t, \xi + \Delta \xi)) = -\gamma_0^{(t, \xi + \Delta \xi)} \Phi_1(\mathcal{X}(t, \xi + \Delta \xi)) + \gamma_1^{(t, \xi + \Delta \xi)} \Phi_1^{(0)}(\mathcal{X}(t, \xi + \Delta \xi)), \)

6.c : \( \mathcal{D}(\Phi_0(\mathcal{X}(t, \xi + \Delta \xi)), \delta(\mathcal{X}(t, \xi + \Delta \xi))) \leq \lambda_2^{(t, \xi + \Delta \xi)}, \)

6.d : \( t \in [t - \xi - \Delta \xi, t + \xi + \Delta \xi], \exists \xi > 0, \exists \Delta \xi > 0, \)

6.e : \( \mathcal{D}(\Phi_0(\mathcal{X}(t, \xi)), \Phi_0(\mathcal{X}(t, \xi + \Delta \xi))) \leq \lambda_3^{(t)}. \)

\[
\mathcal{P}_3 : \begin{cases}
\min_{\Phi_0 \in \mathcal{X}(t, \xi + \Delta \xi)} \mathbb{E}_{\theta(t, \xi + \Delta \xi) \in \Theta} \left\{ \Delta \left( \Phi_1 \left( \mathcal{X}(t, \xi + \Delta \xi) \right), \Phi_0 \left( \mathcal{X}(t, \xi + \Delta \xi) \right) \right) \right\},
\end{cases}
\]

s.t. 6.a - 6.e,

7.f : \( \mathcal{P}_r \left( \Phi(t, \xi + \Delta \xi) \left( \mathcal{X}(t, \xi + \Delta \xi) \right) \right) \geq \lambda_4^{(t, \xi + \Delta \xi)} \)

\[
\mathcal{P}_3' : \begin{cases}
\min_{\Phi_0 \in \mathcal{X}(t, \xi + \Delta \xi)} \mathbb{E}_{\theta(t, \xi + \Delta \xi) \in \Theta} \left\{ \Delta \left( \Phi_1 \left( \mathcal{X}(t, \xi + \Delta \xi) \right), \Phi_0 \left( \mathcal{X}(t, \xi + \Delta \xi) \right) \right) \right\},
\end{cases}
\]

s.t. 6.a - 6.e, 7.f,

8.h : \( \mathcal{P}_r \left( \omega(t, \xi + \Delta \xi) \left( \mathcal{X}(t, \xi + \Delta \xi) \right) \right) \geq \lambda_6^{(t, \xi + \Delta \xi)} \)

\[
\mathcal{P}_3'' : \begin{cases}
\min_{\Phi_0 \in \mathcal{X}(t, \xi + \Delta \xi)} \mathbb{E}_{\theta(t, \xi + \Delta \xi) \in \Theta} \left\{ \Delta \left( \Phi_1 \left( \mathcal{X}(t, \xi + \Delta \xi) \right), \Phi_0 \left( \mathcal{X}(t, \xi + \Delta \xi) \right) \right) \right\},
\end{cases}
\]

s.t. 6.a - 6.e,

9.f : \( \mathcal{P}_r \left( \mathcal{P}_{key} \left( \mathcal{X}(t, \xi + \Delta \xi) \right) \right) \geq \lambda_8^{(t, \xi + \Delta \xi)} \)

binomial extension of \( V \) if 

\[ V^{b_d} [\delta^d | \alpha] = \prod_{i=0}^{d-1} V[\delta_i | \alpha], \forall \alpha \in \mathcal{A}, \delta^d \in \mathcal{B}^d. \]

In the subsequence of Definition 2, one can re-consider the optimisation

\[
\min_{\mathcal{G}_1, \mathcal{G}_2} \mathcal{D}(\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_2),
\]

as

\[
\min_{\mathcal{G}_1, \mathcal{P}(d)} \mathcal{D}(\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_2 = V^{b_d}(\mathcal{G}_1), \mathcal{G}_2),
\]

while \( \mathcal{P}(d) \) basically is the probability mass function (pmf) of \( d \). It is also importantly interesting to be noticed that the term \( (\cdot)^{b_d} \) comes principally from the convolutional nature of the Graph \( \mathcal{G}_2 \) associated fundamentally with \( \mathcal{G}_1 \) since the terms \( \mathcal{D}(\mathcal{G}_1, \mathcal{G}_2) \) and \( \mathcal{D}(\mathcal{G}_2, \mathcal{G}_2) \) may be technically of a Poisson-randomly-distributed nature. Meanwhile, we may interchangeably use \( \mathcal{G}_1^{b_d} \) with a slight abuse of notation here, without loss of generality.

Now, consider the graph \( \mathcal{G} \) and the subgraphs \( \mathcal{G}_1 \) and \( \mathcal{G}_2 \) with respectively random vertex sets \( \mathcal{X} \), \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \). Let \( \mathcal{P} \), \( \mathcal{P}_1 \) and \( \mathcal{P}_2 \) be the probability density distributions respectively on \( \mathcal{X} \), \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \) whose marginal on \( \mathcal{X} \) are equal to \( \mathcal{P} \). Now the
later optimisation can be re-written as

$$\min_{\mathcal{V}, \mathcal{P}(d)} D\left(t(\mathcal{F}_1, \mathcal{V}), t(\mathcal{F}_2 = \mathcal{V}^{Bd}(\mathcal{F}_1), \mathcal{V})\right),$$

s.t. $$\min_{\mathcal{P}^i} \mathcal{J}(X, X_i) = f_1(\mathcal{P}(d)), \min_{\mathcal{P}^2} \mathcal{J}(X, X_2) = f_2(\mathcal{P}(d)),$$

$$D\left(f_1(\mathcal{P}(d)), f_2(\mathcal{P}(d))\right) \leq V^2, \lambda_2 = \mathcal{V}^{Bd}(X_1) \approx \Lambda^{Bd}, \lambda_2 = \mathcal{V}^{Bd}(f_1) = f_1^{Bd},$$

where $$\mathcal{P} := \mathcal{P}(d)$$ holds, w.r.t. the non-zero non-negative threshold $$V^2$$ as well as arbitrary functions $$f_1$$ and $$f_2$$, while the first two additional constraints are thoroughly justified according to the Korner’s graph entropy\textsuperscript{14} – something that can be strongly followed up via Riemannian gradients in Riemannian geometry as described below. Indeed, the first two constraints have come from the fact that

$$\min_{\mathcal{V}, \mathcal{P}(d)} D\left(t(\mathcal{F}_1, \mathcal{V}), t(\mathcal{F}_2 = \mathcal{V}^{Bd}(\mathcal{F}_1), \mathcal{V})\right),$$

is equivalent to

$$\min \mathcal{J}(X, X_1|X_1) - \mathcal{J}(X, X_2|X_1) = \min \mathcal{J}(X|X_1, X_2) - \mathcal{J}(X|X_1) + \mathcal{J}(X|X_2, X_1)$$

$$\equiv \min \mathcal{J}(X, X_2) - \mathcal{J}(X|X_1) = \mathcal{J}(X, X_2) - \mathcal{J}(X, X_1) + \mathcal{J}(X_1)$$

$$\equiv \min \mathcal{J}(X, X_2) - \mathcal{J}(X, X_1).$$

The strongly theoretical logic behind of our interpretation and consideration over the evolution expressed above can be physically justified as follows as well, according to the Riemannian geometry and Riemannian optimisation principles\textsuperscript{15}. Let us initially define the Riemannian gradients\textsuperscript{16} in relation to the sub-manifolds as the orthogonal projection of the distortion function

$$D\left(t(\mathcal{F}_1, \mathcal{V}), t(\mathcal{F}_2, \mathcal{V})\right),$$

onto the associated tangent space as

$$\text{Grad}_{Bce} := \text{Proj}_{\mathcal{F}_1}(\nabla \mathcal{F}_1),$$

which can be widely re-written in terms of the Riemannian logarithmic map operator

$$\text{Grad}_{Bce} = \Sigma_1 \lambda \log(\Sigma_2 + 1 \Sigma_2^{-1} + 1),$$

while $$\Sigma_1$$ and $$\Sigma_2$$ theoretically are the eigen-value matrices associated with the graphs $$\mathcal{F}_1$$ and $$\mathcal{F}_2$$, respectively. Now, defining translation functions $$\mathcal{T}_1$$ and $$\mathcal{T}_2$$ respectively for the real-parts of $$\Sigma_1$$ and $$\Sigma_2$$, and rotation functions $$\mathcal{R}_1^{(\rho)}$$ and $$\mathcal{R}_2^{(\rho)}$$ for the relative imaginary-parts, while the following conditions satisfy

$$\mathcal{T}_i = \mathcal{T}_i + \mathcal{R}_i^{(\rho)} + \mathcal{R}_i^{(\rho)}, i \in \{1, 2\},$$

$$\mathcal{R}_i^{(\rho)} = \mathcal{R}_i^{(\rho)} + \mathcal{R}_i^{(\rho)}, i \in \{1, 2\},$$

while ($$\mathcal{T}$$) principally indicates the average value and $$\mathcal{R}_i^{(\rho)}, i \in \{1, 2\}$$ and $$\mathcal{R}_i^{(\rho)}, i \in \{1, 2\}$$ are Brownian motions. Now, our relaxation we are focusing on is equivalent to the following optimisation

$$\min_{\mathcal{T}_1, \mathcal{T}_2, \mathcal{R}_1^{(\rho)}, \mathcal{R}_2^{(\rho)}} D\left(D\left(\mathcal{T}_1, \mathcal{T}_2\right), D\left(\mathcal{R}_1^{(\rho)}, \mathcal{R}_2^{(\rho)}\right)\right) + D\left(D\left(\mathcal{R}_1^{(\rho)}, \mathcal{R}_2^{(\rho)}\right), D\left(\mathcal{R}_1^{(\rho)}, \mathcal{R}_2^{(\rho)}\right)\right).$$

**Lemma 1:** For the equation derived above, the necessary and sufficient condition to be relaxable fundamentally is the fact that there should exist $$\eta^f$$ and $$\eta^{R^{(\rho)}}$$ such that the following satisfies

$$D\left(f(\mathcal{T}_1), f(\mathcal{T}_2)\right) = \eta^f D\left(\mathcal{T}_1, \mathcal{T}_2\right),$$

$$D\left(f(\mathcal{R}_1^{(\rho)}), f(\mathcal{R}_2^{(\rho)}\right) = \eta^{R^{(\rho)}} D\left(\mathcal{R}_1^{(\rho)}, \mathcal{R}_2^{(\rho)}\right).$$

**Proof:** The proof is omitted for brevity. \(\blacksquare\)

**Step 7.** Finally, the following lemma can help us in concluding our interpretation.

**Lemma 2:** The constraint

$$D\left(\Phi_0(\mathcal{X}(\mathcal{E})), \Phi_0(\mathcal{X}(\mathcal{E} + \Delta \mathcal{E}))\right) \leq \lambda_3^{(t)},$$

is interpreted according to the Bochner’s theorem.

**Proof:** The Bochner’s theorem\textsuperscript{17} says that for a locally compact abelian group\textsuperscript{18} $$\mathcal{G}$$, with dual group $$\tilde{\mathcal{G}}$$, there exists a unique probability measure $$\kappa$$ such that

$$f(\tilde{\mathcal{G}}) = \int_{\mathcal{G}} x(\tilde{\mathcal{G}}) d\kappa(x),$$

holds. This completes the proof of Lemma 1 here.

The proof is now completed. \(\blacksquare\)

**APPENDIX C**

**Proof of Proposition 3**

The Herglotz-Riesz representation theorem for harmonic functions\textsuperscript{19} can be useful here for us saying that DoF we are interested in, if they are non-zero, have positive real-parts iff and only iff there is a probability measure $$\mu$$ on the unit circle such that

$$\min_{\mu^{(\circ)}} D\left(\mathcal{F}_0(\rho e^{i\theta}), \int_{0}^{2\pi} 1 - \rho^2 |1 - 2\rho \cos(\theta - \theta') + \rho^2| d\mu(\theta')\right),$$

holds, while

$$\mathcal{F}_0 := d\left(\Phi_1(\mathcal{X}(\mathcal{E}))\right) := f(e, \rho, i) d\mathcal{B}_0,$$

\textsuperscript{17}see e.g. [43] to see what it is.

\textsuperscript{18}A group while each pair of group members are interchangeable.

\textsuperscript{19}see e.g. [44] to see what it is.
since it is proven that $\mathcal{F}_0$ is harmonic, that is
$$\nabla^2 \mathcal{F}_0 = 0,$$
according to the Newton’s second law of motion, roughly speaking.

The minimisation defined above can be re-written as
$$\min_{\mu(t)} D \left( \mathcal{F}_0(\rho e^{i \theta} \mu), \int_0^{2\pi} \frac{1 + z(\theta)e^{-i \theta'}}{1 - z(\theta)e^{-i \theta'}} d\mu(\theta') \right),$$
for the case where $\mathcal{F}_0$ is holomorphic.

**Remark 3.** The term $\frac{1 + z(\theta)e^{-i \theta'}}{1 - z(\theta)e^{-i \theta'}}$ can be re-expressed as
$$\frac{1 + z(\theta)e^{-i \theta'}}{1 - z(\theta)e^{-i \theta'}} = e^{i \frac{\varphi}{2}e^{i \varphi}e^{\frac{\varphi}{2}e^{-i \varphi} + e^{-i \frac{\varphi}{2}e^{-i \varphi}}},$$
while $z(\theta) = e^{i \theta}$. In this term, the term $\frac{\varphi}{2}e^{i \varphi}e^{-i \varphi}$ shows the story in relation to the tangent space expressed in Step 6 in Appendix B.

The proof is now completed. ■

**Appendix D**

**Proof of Proposition 4**

The achievable secret-key rate per dimension $\mathcal{R}_{\text{key}}$ is $^{20}$
$$\mathcal{R}_{\text{key}} = \mathbb{E} \left\{ (1 - \psi(t))\sigma(t)\mathcal{J}_1(t) \left( \frac{\omega(t)}{\sigma(t)} \right) + \ldots \right\}$$
$$\psi(t)\sigma(t)\left( \mathcal{J}_1(t) \left( \frac{\omega(t)}{\sigma(t)} \right) - \mathcal{J}_2(t) \left( \frac{\omega(t)}{\sigma(t)} \right) \right),$$
while
$$|S_{ab}(t)| = |S_{ae}(t)| = \sigma(t),$$
holds, and $\omega(t)$ is the instantaneous signal-to-noise (SNR), as well as the fact that
$$\mathcal{J}_1(t) \left( \frac{\omega(t)}{\sigma(t)} \right) = I(\mathcal{X}(t); \mathcal{Y}(t)|S_{ab}(t)),$$
and
$$\mathcal{J}_2(t) \left( \frac{\omega(t)}{\sigma(t)} \right) = I(\mathcal{X}(t); \mathcal{Z}(t)|S_{ab}(t)),$$
hold$^{21}$, while $\mathcal{L}$ is the observed data by Eve.

Now, one can assert that the Problem $\mathcal{P}_2$ expressed in Eqn. (6) can be re-written as the Problem $\mathcal{P}_3$ expressed in Eqn. (7) or the Problem $\mathcal{P}_3'$ expressed in Eqn. (8), while for the $t$-th instant and $j$-th sub-channel, we have
$$\Pr \left( j \in S_{ae}(t)| \right. \left. j \in S_{ab}(t) \right) = \psi(t), \forall j \in \mathbb{N},$$
and we principally define the following probabilistic constraint
$$\Pr \left( \psi(t) \left( \mathcal{X}(t) \right) \geq \lambda_4(t) \right) \leq \lambda_5(t),$$
which can be in parallel with
$$\Pr \left( \omega(t) \left( \mathcal{X}(t) \right) \geq \lambda_6(t) \right) \geq \lambda_7(t),$$
\[^{20}\text{see e.g. [21], [22].}\]
\[^{21}\text{see e.g. [21], [22].}\]

w.r.t. the arbitrary thresholds $\lambda_4(t)$, $\lambda_5(t)$, $\lambda_6(t)$ and $\lambda_7(t)$, roughly speaking.

The two probabilistic constraints theoretically introduced here stand respectively for the following facts: (i) the upper-bound of the probability value $\psi(t)$ should be constrained through a threshold; (ii) and the lower-bound of the instantaneous SNR $\omega(t)$ should also be constrained through a threshold as well.

Instead of the two constraints given above, one can constrain the secret-key rate as
$$\Pr \left( \mathcal{R}_{\text{key}}(\mathcal{X}(t)) \geq \lambda_4(t) \right) \geq \lambda_5(t),$$
\[^{22}\text{See e.g. [27], [28], [29] for further discussion.}\]

w.r.t. the arbitrary thresholds $\lambda_4(t)$ and $\lambda_5(t)$,

**Remark 4.** It is absolutely interesting to note that although perturbations are considered in the literature mainly due to attacks, we do not ignore the term $\xi + \Delta$ in Eqns. (7-8-9) in order to preserve the optimality and generality of the problem.

**Remark 5.** The three probabilistic constraints newly declared above are relaxable according to the principle of the Concentration-of-Measure – something that theoretically proves an exponential bound for the two aforementioned constraints$^{22}$.

The proof is now completed. ■

**References**

[1] B. Razeghi, S. Ferdowsi, D. Kostadinov, F. P. Calmon, S. Voloshynovskiy, "Privacy-Preserving Near Neighbor Search via Sparse Coding with Ambiguation," IEEE Trans. Automatic Control, https://arxiv.org/abs/2102.04274, 2021.

[2] C. Cheng, W. Dai, "Dictionary Learning with Convex Update (ROMD)," https://arxiv.org/abs/2110.06641, 2021.

[3] L. Shi, Y. Chi, "Manifold Gradient Descent Solves Multi-Channel Sparse Blind Deconvolution Provably and Efficiently," https://arxiv.org/abs/1911.11167, 2021.

[4] F. G. Veshki, S. A. Vorobyov, "Efficient ADMM-based Algorithms for Convolutional Sparse Coding," https://arxiv.org/abs/2109.02969, 2021.

[5] P. Mayo, O. Karakus, R. Holmes, A. Achim, "Representation Learning via Cauchy Convolutional Sparse Coding," https://arxiv.org/abs/2008.03473, 2021.

[6] S. Kolek, D. A. Nguyen, R. Levie, J. Bruna, G. Kutyniok, "A Rate-Distortion Framework for Explaining Black-box Model Decisions," https://arxiv.org/abs/2110.08252, 2021.

[7] Q. Jiang, N. Zhang, J. Ni, J. Ma, X. K. Kwang, "Unified Biometric Privacy Preserving Three-Factor Authentication and Key Agreement for Cloud-Assisted Autonomous Vehicles," IEEE Trans. Vehicular Technol., Vol. 69, no. 9, pp. 9390-9401, 2020.

[8] F. Song, Z. Qin, D. Liu, J. Zhang, X. Lin, "Privacy-Preserving Task Matching With Threshold Similarity Search via Vehicular Crowdsourcing," IEEE Trans. Vehicular Technol., Vol. 70, no. 7, pp. 7161-7175, 2021.

[9] X. Zeng, G. Xu, X. Zheng, Y. Xian, W. Zhou, "E-AUA: An Efficient Anonymous User Authentication Protocol for Mobile IoT," IEEE Internet Thing. J., Vol. 6, no. 2, pp. 1506-1519, 2019.

[10] J. Zhang, J. Ma, Y. Yang, X. Liu, N. N. Xiong, "Recollectable and Privacy-Preserving Decentralized Data Sharing Framework for Fog-Assisted Internet of Things," IEEE Internet Thing. J., Vol. pp, no. 99, pp. 1-1, 2021.

[11] Q. Hu, J. Zhang, G. Hancke, Y. Hu, W. Li, H. Jiang, "Half-duplex mode based secure key generation method for resource-constrained IoT devices," IEEE Internet Thing. J., Vol. pp, no. 99, pp. 1-1, 2021.

[12] Z. Wang, J. Li, J. Hu, J. Ren, Q. Wang, Z. Li, Y. Li, "Towards Privacy-Driven Truthful Incentives for Mobile Crowdsensing Under Untrusted Platform," IEEE Trans. Mob. Comput., Vol. pp, no. 99, pp. 1-1, 2021.
