# Variational Secret Common Randomness Extraction

Xinyang Li , Graduate Student Member, IEEE, Vlad C. Andrei , Graduate Student Member, IEEE, Peter J. Gu , Graduate Student Member, IEEE, Yiqi Chen , Member, IEEE, Ullrich J. Mönich , Senior Member, IEEE, and Holger Boche , Fellow, IEEE

Abstract—This paper studies the problem of extracting common randomness (CR) or secret keys from correlated random sources observed by two legitimate parties, Alice and Bob, through public discussion in the presence of an eavesdropper, Eve. We propose a practical two-stage CR extraction framework. In the first stage, the variational probabilistic quantization (VPQ) step is introduced, where Alice and Bob employ probabilistic neural network (NN) encoders to map their observations into discrete, nearly uniform random variables (RVs) with high agreement probability while minimizing information leakage to Eve. This is realized through a variational learning objective combined with adversarial training. In the second stage, a secure sketch using code-offset construction reconciles the encoder outputs into identical secret keys, whose secrecy is guaranteed by the VPQ objective. As a representative application, we study physical layer key (PLK) generation. Beyond the traditional methods, which rely on the channel reciprocity principle and require two-way channel probing, thus suffering from large protocol overhead and being unsuitable in high mobility scenarios, we propose a sensing-based PLK generation method for integrated sensing and communications (ISAC) systems, where paired range-angle (RA) maps measured at Alice and Bob serve as correlated sources. The idea is verified through both end-to-end simulations and real-world software-defined radio (SDR) measurements, including scenarios where Eve has partial knowledge about Bob's position. The results demonstrate the feasibility and convincing performance of both the proposed CR extraction framework and sensing-based PLK generation method.

*Index Terms*—Common randomness, variational learning, physical layer security, integrated sensing and communications, secret key generation.

#### I. INTRODUCTION

#### A. Background and Related Works

Common randomness (CR) [1], [2] plays an essential role in information theory, referring to the generation of identical random variables (RVs) by two parties, Alice and Bob, from correlated observations. When secrecy is required, the generated RVs must remain statistically independent of any side information available to an external observer, Eve. These concepts have been extensively studied in applications such as secure communications [3], identification codes [4], and quantum cryptography [5]. Prior works mainly focus on characterizing the maximum entropy of CR, known as the CR capacity, for two correlated sources under different settings. For example, [6] shows that the CR capacity without public discussion is equal to the entropy of the Gács-Körner-Witsenhausen (GKW) common components of the two random sources, which becomes zero if they have an indecomposable

The authors are with the Department of Electrical and Computer Engineering, Technical University of Munich, Munich, 80333 Germany (e-mail: {xinyang.li, ylad.andrei, peter.gu, yiqi.chen, moenich, boche}@tum.de).

Code will be available at https://github.com/xinyanglii/vcr after acceptance.

joint distribution. Furthermore, [1], [2] derived the CR capacity if both parties are allowed to communicate publicly, subject to a rate constraint, both with and without the secrecy requirement. Achievable and upper bounds of the CR capacity have also been studied in extended scenarios, including multi-way communications [7] and in the presence of a helper [8].

Despite the instructive meaning of the theoretical foundation, practical approaches to extracting CR remain largely unexplored. Most existing works focus narrowly on the application of physical layer key (PLK) generation, where Alice and Bob safeguard their wireless communication link by deriving secret keys from channel measurements like received signal strength or channel state information [9]–[12], These schemes typically follow a pipeline consisting of channel probing, quantization, information reconciliation, and privacy amplification, and the effectiveness often relies on channel reciprocity and temporal variation. Moreover, many practical implementations assume that the generated keys remain unknown to Eve due to spatial decorrelation and thus omit the secrecy requirement.

In this work, we address CR extraction from a more general information-theoretic perspective and propose a practical two-stage framework. In the first stage, termed variational probabilistic quantization (VPQ), Alice and Bob each employ probabilistic neural network (NN) encoders to transform their observations into discrete, nearly uniform RVs with high agreement probability and low leakage. The design objective jointly optimizes these properties through a variational formulation, and to further suppress leakage, we integrate adversarial training based on mutual information bounds [13], [14]. In the second stage, one-way public communication is used for secret key reconciliation via a secure sketch, implemented through a code-offset construction [15], and the resulting secret keys remain information-theoretically secure provided that the VPO objectives are met. Compared to the conventional PLK generation scheme, this two-stage design eliminates the need for explicit privacy amplification, since secrecy is already embedded in the VPQ stage. Additionally, unlike the traditional quantization rules, which are often tailored to specific sources [16]-[18], such as received signal strength or channel phase, and thus lack flexibility, VPQ is a learning-based and data-driven method that can, in principle, be applied to arbitrary data types. Under the proposed CR extraction framework, we will demonstrate, using an example of fading channels, that the extracted PLKs not only achieve a uniform distribution and a high key agreement rate but also are robust to Eve's correlated observation, owing to the adversarial

The traditional PLK generation methods are often limited

by low key generation rates due to the scarcity of randomness sources and non-ideal channel reciprocity. The channel probing step requires multi-way communications between Alice and Bob, making it unsuitable for high-mobility scenarios. Recent advances in integrated sensing and communications (ISAC) provide existing wireless networks with sensing capability to simultaneously communicate and sense the environment [19], such as detecting targets and estimating range and velocity, offering new opportunities to enhance the physical layer security (PLS) in ISAC systems. Existing works mainly leverage sensing for waveform design, such as artificial noise injection [20] or interference management [21], [22], to impair wiretap channels [23]. In these approaches, sensing is primarily used to detect potential eavesdroppers or adversaries [24]. While effective, these wiretap coding methods fail to ensure security when the wiretap channel is stronger than the legitimate one [7] or the location of the eavesdropper is unavailable.

To address these limitations, we propose a novel PLK generation framework in ISAC systems that directly utilizes the sensing data collected by the legitimate users. When Alice and Bob sense their shared propagation environment, the resulting measurements inherently contain CR that can serve as a source for PLK generation. As a case study, we focus on the relative distance and angle between Alice and Bob. In the presence of line of sight (LoS) path and a detectable echo signal reflected from Bob to Alice, the measured range-angle (RA) information at both parties becomes highly correlated. Under high mobility conditions, Bob's position varies rapidly and the measured RA maps can be treated as an independent random variable when its coherence time is shorter than the PLK update interval.

To validate this concept, we conduct an end-to-end system simulation that involves all necessary signal processing steps and channel effects using the NR physical data shared channel (PDSCH) signal. After receiver processing, the resulting RA maps at Alice and Bob are then used as inputs to the proposed learning-based CR extraction framework for PLK generation. Unlike conventional reciprocity-based approaches, the proposed method does not require Bob to perform active channel probing, thereby significantly reducing communication overhead. To further examine robustness, we also consider cases where Eve has partial knowledge of Bob's relative position to Alice. Beyond simulations, we also apply the software-defined radio (SDR) technique to collect the real-world RA map data in both the lab room and the anechoic chamber environments. The models pretrained on the synthesized dataset are then fine-tuned on the real-world RA maps with the backbone NN frozen, demonstrating both the generalizability of the pretrained models and the effectiveness of the proposed CR extraction and sensing-based PLK generation framework.

## B. Contributions

The main contributions of this work are summarized as follows:

 We propose a practical two-stage CR extraction framework by combining a learning-based VPQ method with a secure sketch. VPQ employs probabilistic NN encoders to map correlated observations into nearly uniform RVs with low mismatch probability and minimal leakage to Eve. To achieve this, we derive variational lower and upper bounds on the leakage rate and introduce an adversarial training strategy. In the second stage, reconciliation is performed via a code-offset construction, and we prove that the secrecy of the resulting secret keys is ensured by the learning objective established in the VPQ stage.

- We apply the proposed framework to synthesized correlated Gaussian RVs, representing a typical PLK generation scenario from wireless fading channels. We investigate cases without Eve, with uncorrelated observations at Eve, and with correlated observations at Eve. Unlike conventional PLK schemes, which often assume spatial decorrelation of Eve's channel, our learning-based approach adapts to more general and challenging scenarios.
- We propose a novel PLK generation approach by exploiting the correlated sensing information at Alice and Bob in ISAC systems. We treat the RA maps simultaneously estimated at Alice and Bob as the CR source, which is highly correlated if a LoS link exists. To validate the idea, we perform both end-to-end 5G NR simulations and real-world measurements using SDR devices. To bridge simulation and practice, the NN models trained on large synthesized datasets are fine-tuned on measured data with a frozen backbone, demonstrating convincing performance of both the CR extraction and sensing-based PLK generation scheme, even when Eve has partial knowledge of Bob's location.

#### II. SECRET COMMON RANDOMNESS

In many problems in information theory, CR refers to RVs generated by two parties, Alice and Bob, from a pair of correlated random sources  $(X,Y) \sim p(x,y)$  with the aid of public discussion [1], [2]. Specifically, let Alice observes a sequence  $X^n = (X_1, \ldots, X_n)$ , while Bob observes  $Y^n = (Y_1, \ldots, Y_n)$ . We consider the one-way communication setting, where Alice sends a public message  $M = \Phi(X^n) \in \mathcal{M} = \{1, \ldots, |\mathcal{M}|\}$  and both parties map their observations into RVs

$$K = f(X^n), \quad L = g(Y^n, M), \tag{1}$$

with  $K, L \in \mathcal{K} = \{1, \dots, |\mathcal{K}|\}$ , such that K = L with high probability. The mappings  $f, g, \Phi$  may be either deterministic or stochastic.

If an eavesdropper Eve observes another correlated sequence  $Z^n$  that is jointly distributed with  $(X^n,Y^n)$ , it is additionally desirable that the extracted  $\operatorname{CR} K$  (or L) remains unpredictable from  $(Z^n,M)$ . This leads to the requirement that the averaged mutual information  $\frac{1}{n}I(K;Z^n,M)$  is arbitrarily small, and K tends to be uniformly distributed. The generated RVs K or L are also referred to as secret keys. The entropy rate  $\frac{1}{n}H(K)$  is called an achievable  $\operatorname{CR}$  rate and the supremum over all such achievable rates defines the  $\operatorname{CR}$  capacity. A schematic illustration of such a process is given in Fig. 1. More formally, we have the following definition.

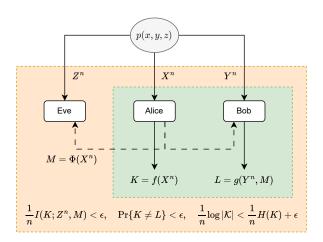


Fig. 1: Common randomness (CR) with public discussion.

Definition 1. Let Alice, Bob and Eve observe random sequences  $\{(X_i, Y_i, Z_i)\}_{i=1}^n$  generated independent and identically distributed (i.i.d.) from the joint distribution p(x, y, z), with  $Z=\varnothing$  if Eve is absent. A function  $\Phi$  at Alice maps  $X^n$ into a public message  $M = \Phi(X^n)$  and a pair of functions f, g extract RVs  $K = f(X^n), L = g(Y^n, M)$  at Alice and Bob, respectively. K or L is called CR or secret key if the following conditions hold:

$$\Pr\left\{K \neq L\right\} < \epsilon,\tag{2}$$

$$\frac{1}{n}\log|\mathcal{K}| < \frac{1}{n}H(K) + \epsilon, \tag{3}$$

$$\frac{1}{n}I(K; \mathbb{Z}^n, M) < \epsilon. \tag{4}$$

$$\frac{1}{n}I(K;Z^n,M) < \epsilon. \tag{4}$$

for every  $\epsilon > 0$  and sufficiently large n. The supremum of achievable entropy rates  $\frac{1}{n}H(K)$  as  $n \to \infty$  defines the CR or secret key capacity.

For the scenario without public discussion, i.e.,  $\Phi = \emptyset$ , it has been proved that the CR capacity is given by  $H(X_0|Z)$ where  $X_0$  is the GKW common component of X and Y[6], [25]. If X, Y have an indecomposable joint distribution such as a joint Gaussian, the CR capacity is zero. If Eve is absent, the CR capacity becomes I(X;Y) which is achieved by transmitting the compression of  $X^n$  at rate H(X|Y) such that Bob can decode  $X^n$  losslessly with the side information  $Y^n$  according to the Slepian-Wolf theorem [26]. In the general one-way communication case, the CR capacity with present Eve is given by  $\max I(T; Y|U) - I(T; Z|U)$  where the maximum is taken over all possible auxiliary RVs (U,T) such that the Markov chain U - T - X - (Y, Z) holds [1].

Although the theoretical properties of CR have been well established, practical extraction methods remain scarce, especially when the joint distribution p(x, y, z) is inaccessible and complicated. To this end, this work develops a twostage CR extraction framework that combines a variational learning approach and the secure sketch-based information reconciliation.

#### III. PROPOSED METHOD

The proposed CR extraction framework consists of two stages. In the first stage, Alice and Bob independently map their respective observations to sequences of discrete RVs that are (i) nearly uniform, (ii) closed to each other (low mismatch probability), and (iii) unpredictable from Eve's observations. To achieve this, we introduce a VPQ scheme, where probabilistic NN encoders are trained under a variational adversarial objective. In the second stage, Alice applies a secure sketch based on the code-offset construction to assist Bob in correcting the disagreement between the VPQ output pair and consequently recovering the secret keys.

Specifically, Alice and Bob quantize each observed pair (X,Y) into discrete RVs (W,V) by learning two probabilistic NN encoders  $p_{\theta}(w|x), p_{\phi}(v|y)$  with learnable parameters  $\theta, \phi$ . W and V take value from a finite alphabet W. Hence, the last two layers of  $p_{\theta}$ ,  $p_{\phi}$  are typically a linear layer followed by a softmax layer with output dimension  $|\mathcal{W}|$ . If X, Y have the same data structure, Alice and Bob may also share the same encoder parameters. The distributions of W and V are expected to be uniform such that their entropy H(W) and H(V) are maximized toward  $\log |\mathcal{W}|$ , and the mismatch rate  $\Pr\{W \neq V\}$  is minimized. If Eve is present and observes Z, another predictor  $p_{\psi}(w|z)$  for Eve is designed and trained with the encoders  $p_{\theta}$ ,  $p_{\phi}$  in an adversarial manner to minimize the mutual information I(W; Z). In the reconciliation stage, given the quantized sequence  $W^n$  transformed from  $X^n$ , Alice samples uniformly a codeword C from an error-correcting code  $\mathcal{C}$  and computes the code offset  $S = W^n - C$  on the corresponding finite field. The offset S is sent to Bob as the secure sketch. Bob computes  $C' = V^n + S$  and then decodes  $\hat{C} \in \mathcal{C}$  from C'. Finally, both parties use K = C and  $L = \hat{C}$ as the resulting shared secret key.

In this section, we first present the design of the learning objective and training strategy for VPQ, including the adversarial predictor for Eve. We then describe the implementation of the secure sketch based on code-offset construction and prove that the VPQ training objective ensures the secrecy of the reconciled keys. An overview of the proposed framework is illustrated in Fig. 2.

## A. Mismatch Rate

The first training target of VPQ is to minimize the mismatch rate between the encoder outputs of Alice and Bob, i.e.,  $\Pr\{W \neq V\}$ . To convert it to a differentiable function that can be used to train the NNs, we note that

$$\Pr\{W \neq V\} = \mathbb{E}_{p(x,y)}[\Pr\{W \neq V | X, Y\}]$$

$$= 1 - \mathbb{E}_{p(x,y)}[\mathbb{E}_{p_{\theta}(w|x), p_{\phi}(v|y)}[\mathbb{1}\{W = V\}]]$$
(6)

with  $\mathbb{1}\{\cdot\}$  the indicator function. Denoting w and v as the  $|\mathcal{W}|$ dimensional one-hot vector of W.V. respectively, we have

$$\mathbb{E}_{p_{\theta}(w|x),p_{\phi}(v|y)}[\mathbb{1}\{W=V\}] = \mathbb{E}_{p_{\theta}(w|x),p_{\phi}(v|y)}[\boldsymbol{w}^{\top}\boldsymbol{v}]$$
(7)
$$= \mathbb{E}_{p_{\theta}(w|x)}[\boldsymbol{w}]^{\top}\mathbb{E}_{p_{\phi}(v|y)}[\boldsymbol{v}]$$
(8)
$$= \sum_{w'\in\mathcal{W}} p_{\theta}(w'|x)p_{\phi}(w'|y),$$
(9)

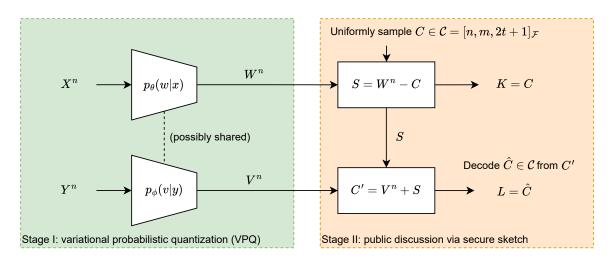


Fig. 2: Overview of the proposed two-stage CR extraction framework.

where the second equality follows from the independence between w and v conditioned on (x, y). Hence, given a data batch  $\{(x_i, y_i)\}_{i=1}^B$ , the mismatch-rate loss is defined as

$$\mathcal{L}_{MR} = -\frac{1}{B} \sum_{i=1}^{B} \sum_{w' \in \mathcal{W}} p_{\theta}(w'|x_i) p_{\phi}(w'|y_i).$$
 (10)

It turns out that  $\mathcal{L}_{MR}$  will force  $p_{\theta}(w|x)$  and  $p_{\phi}(v|y)$  to the same one-hot vector for each paired input, corresponding to deterministic mappings at both Alice and Bob.

# B. Uniformity

The second objective requires the generated (W,V) to be as close to a uniform distribution as possible. From a security perspective, uniformity guarantees unpredictability, while from a learning perspective it prevents mode collapse, where both VPQ encoders always output the same one-hot vector due to the mismatch-rate loss.

The uniformity of (W,V) is quantified by their respective information entropy H(W) and H(V), which are to be maximized. In practice, the marginals are estimated empirically by first averaging encoder outputs over a training data batch:

$$p(w) = \frac{1}{B} \sum_{i=1}^{B} p_{\theta}(w|x_i), \ q(v) = \frac{1}{B} \sum_{i=1}^{B} p_{\phi}(v|y_i).$$
 (11)

If the output dimension  $|\mathcal{W}|$  is too large compared to the batch size B, such that the above estimation over a single batch is inaccurate, one can perform the exponentially moving average (EMA) approach to marginalizing the probabilities over multiple batches:

$$p_t(w) = \alpha p_{t-1}(w) + \frac{1-\alpha}{B} \sum_{i=1}^{B} p_{\theta}(w|x_i), \qquad (12)$$

$$q_t(v) = \alpha q_{t-1}(v) + \frac{1-\alpha}{B} \sum_{i=1}^{B} p_{\phi}(v|y_i), \qquad (13)$$

where  $0 \le \alpha < 1$  and t denotes the training step. At training step t,  $p_{t-1}(w)$  and  $q_{t-1}(v)$  are detached from the gradient

computation graph as they do not depend on the current encoder outputs.  $p_t(w)$  and  $q_t(v)$  are then used to compute the empirical entropy values:

$$\hat{H}(W) = -\sum_{w \in \mathcal{W}} p_t(w) \log p_t(w), \tag{14}$$

$$\hat{H}(V) = -\sum_{v \in \mathcal{W}} q_t(v) \log q_t(v). \tag{15}$$

The uniformity loss is thus given by

$$\mathcal{L}_{ENT} = -\frac{1}{2(1-\alpha)}(\hat{H}(W) + \hat{H}(V)),$$
 (16)

where we divide the entropy by  $(1-\alpha)$  to compensate for the downscaled gradient caused by EMA.

For validation and testing, the marginal probabilities p(w) and q(v) are computed over the entire dataset to obtain a more accurate entropy estimate.

# C. Leakage Rate

Besides the objective of lower mismatch rate and uniformity of (W,V), it is also desired that I(W;Z) approaches 0 such that there is no leakage of the encoder output to Eve. This requirement can ensure the unpredictability of the final secret keys in the second stage, which will be shown later. If Eve is absent, or observes uncorrelated information such that p(x,y,z) = p(x,y)p(z), it is satisfied automatically  $I(W;Z) \leq I(X;Z) = 0$  due to the data processing inequality. In this case, the combination of (10) and (16) as the total loss function

$$\mathcal{L}_{AB} = \mathcal{L}_{ENT} + \lambda_1 \mathcal{L}_{MB} \tag{17}$$

with  $\lambda_1>0$  is sufficient to train the encoders  $p_\theta,p_\phi$ . In contrast, when Z is correlated to (X,Y), one should design another loss function to suppress the leakage rate I(W;Z). However, computing mutual information without knowing the underlying distributions is difficult. To this end, we introduce both variational lower and upper bounds for I(W;Z) and propose to train the encoders and another predictor at Eve

in an adversarial manner. We will show this procedure to be equivalent to jointly estimating and minimizing I(W; Z).

We start with the variational lower bound of I(W; Z) [13], [27]. By noting that  $W \sim p_{\theta}(w|x)$  is independent of Z conditioned on X, we have

$$I(W; Z) = \mathbb{E}_{p(w,z)} \left[ \log \frac{p(w,z)}{p(w)p(z)} \right]$$
(18)

$$= \mathbb{E}_{p(w,x,z)} \left[ \log \frac{p(w|z)}{p(w)} \right] \tag{19}$$

$$= \mathbb{E}_{p_{\theta}(w|x)p(x,z)} \left[ \log \frac{p(w|z)}{p(w)} \right]$$
 (20)

$$= \mathbb{E}_{p_{\theta}(w|x)p(x,z)} \left[ \log \frac{p(w|z)p_{\psi}(w|z)}{p_{\psi}(w|z)p(w)} \right]$$
 (21)

$$= D(p(w|z)||p_{\psi}(w|z))$$

$$+ \mathbb{E}_{p_{\theta}(w|x)p(x,z)}[\log p_{\psi}(w|z)] + H(W) \quad (22)$$

$$\geq \mathbb{E}_{p_{\theta}(w|x)p(x,z)} \left[ \log p_{\psi}(w|z) \right] + H(W) \tag{23}$$

$$\triangleq I_{\text{VLB}}(W; Z),$$
 (24)

where we introduce a conditional probability  $p_{\psi}(w|z)$  parameterized by a NN with parameter  $\psi$ . Because the Kullback–Leibler divergence (KLD) term  $D(p(w|z)\|p_{\psi}(w|z))$  is always nonnegative,  $I_{\text{VLB}}(W;Z)$  provides a variational lower bound for I(W;Z). By fixing the encoder  $p_{\theta}(w|x)$  and thus I(W;Z), one may maximize the lower bound to estimate I(W;Z), and at the optimum  $p_{\psi}(w|z)$  is equal to the true p(w|z), the KLD term becomes 0 and  $I_{\text{VLB}}(W;Z)$  equals I(W;Z). By replacing the expectation over p(x,z) by the empirical mean, the variational lower bound objective is given by

$$\mathcal{I}_{\text{VLB}} = \frac{1}{B} \sum_{i=1}^{B} \sum_{w \in \mathcal{W}} p_{\theta}(w|x_i) \log p_{\psi}(w|z_i), \qquad (25)$$

where H(W) is omitted as it is not affected by  $\psi$ . In fact, (25) is the negative cross entropy between  $p_{\theta}(w|x)$  and  $p_{\psi}(w|z)$ . Intuitively, maximizing (25) to estimate I(W;Z) while fixing  $p_{\theta}$  is equivalent to training a predictor  $p_{\psi}$  at Eve to infer the encoder output from the correlated observation Z.

With the optimal  $p_{\psi}(w|z)$  and the estimated I(W;Z), the goal of Alice and Bob is to minimize it as much as possible. One simple idea is to directly minimize  $I_{\text{VLB}}(W;Z)$  with respect to  $p_{\theta}$  by fixing  $p_{\psi}$ . However, this could lead to two main issues. The first issue is that minimizing  $I_{\text{VLB}}(W;Z)$  conflicts with maximizing the entropy H(W) in the previous section. On the other hand, even if one can only minimize  $\mathcal{I}_{\text{VLB}}$  while omitting the term H(W), updating  $p_{\theta}$  will also change p(w|z) implicitly such that the fixed  $p_{\psi}$  is no more optimal and  $I_{\text{VLB}}(W;Z)$  becomes again an untight lower bound, whose reduction can not ensure the decreasing of I(W;Z).

To this end, we shall consider a variational upper bound for I(W;Z) [14]. By assuming  $p_{\psi}$  to be optimal, I(W;Z) is given by

$$\mathbb{E}_{p_{\theta}(w|x)p(x,z)} \left[ \log p_{\psi}(w|z) \right] - \mathbb{E}_{p(w)} \left[ \log p(w) \right]$$

$$= \mathbb{E}_{p_{\theta}(w|x)p(x,z)} \left[ \log p_{\psi}(w|z) \right] - \mathbb{E}_{p(w)} \left[ \log \mathbb{E}_{p(z)} [p_{\psi}(w|z)] \right]$$
(27)

$$\leq \mathbb{E}_{p_{\theta}(w|x)p(x,z)} \left[ \log p_{\psi}(w|z) \right] - \mathbb{E}_{p(w)p(z)} \left[ \log p_{\psi}(w|z) \right]$$
(28)

$$\triangleq I_{\text{VUB}}(W; Z) \tag{29}$$

where we adopt Jensen's inequality. In fact,  $I_{\text{VUB}}(W; Z)$  is not always a valid upper bound of I(W; Z) as we use  $p_{\psi}(w|z)$  to approximate the true p(w|z). Nonetheless, by Theorem 3.2 in [14],  $I_{\text{VUB}}(W; Z) \geq I(W; Z)$  holds true if

$$D(p(w|z)p(z)||p_{\psi}(w|z)p(z)) \le D(p(w)p(z)||p_{\psi}(w|z)p(z)).$$
(30)

When  $p_{\psi}(w|z)$  is optimal, such that the left-hand side is zero, although updating  $p_{\theta}(w|z)$  will change p(w|z), the upper bound  $I_{\text{VUB}}(W;Z)$  is still valid as long as the change of p(w|z) doesn't violate the condition (30).

The variational upper bound  $I_{\mathrm{VUB}}(W;Z)$  can also be computed empirically as

 $\mathcal{I}_{\mathrm{VUB}}$ 

$$= \frac{1}{B^2} \sum_{i,j=1}^{B} \sum_{w \in \mathcal{W}} p_{\theta}(w|x_i) \left[ \log p_{\psi}(w|z_i) - \log p_{\psi}(w|z_j) \right]$$
(31)

$$= \mathcal{I}_{VLB} - \frac{1}{B^2} \sum_{i,j=1}^{B} \sum_{w \in \mathcal{W}} p_{\theta}(w|x_i) \log p_{\psi}(w|z_j).$$
 (32)

Therefore, minimizing  $\mathcal{I}_{VUB}$  will not only reduce the variational lower bound  $\mathcal{I}_{VLB}$  to decrease the prediction accuracy at Eve, but also force the encoder output to be predicted by Eve more likely from uncorrelated observations. The update of  $\mathcal{I}_{VLB}$  and  $\mathcal{I}_{VUB}$  are thus performed alternately in an adversarial way. That is, while fixing  $p_{\theta}$ ,  $p_{\psi}$  is trained to maximize  $\mathcal{I}_{VLB}$ , and while  $p_{\psi}$  is frozen,  $p_{\theta}$  is learned to decrease  $\mathcal{I}_{VUB}$ .

#### D. VPQ Training Strategy

By combining the loss functions associated with the three objectives, the overall VPQ loss function is defined as

$$\mathcal{L} = \mathcal{L}_{AB} + \lambda_2 \mathcal{I}_{VUB} \tag{33}$$

$$= \mathcal{L}_{ENT} + \lambda_1 \mathcal{L}_{MR} + \lambda_2 \mathcal{I}_{VUB}, \tag{34}$$

with  $\lambda_2 \geq 0$  a weight factor. In our experiments, if Eve is present,  $\lambda_2$  is either fixed or updated adaptively according to

$$\lambda_2 = \frac{\|\nabla_{\theta_L} \mathcal{L}_{AB}\|_2}{\|\nabla_{\theta_L} \mathcal{I}_{VUB}\|_2 + \delta}$$
 (35)

following the same scaling strategy as VQ-GAN [28], where  $\nabla_{\theta_L}$  denotes the gradient with respect to the last layer before softmax of the encoder  $p_{\theta}$ , and  $\delta = 10^{-7}$  is used for numerical stability. This choice guarantees the gradient norms of  $\mathcal{L}_{AB}$  and  $\mathcal{I}_{VUB}$  remain comparable, preventing one objective from dominating the update.

The overall training pseudocode is given in Algorithm 1. In each iteration, Alice and Bob generate encoder outputs and compute  $\mathcal{L}_{\mathrm{MR}}$  and  $\mathcal{L}_{\mathrm{ENT}}$ . If Eve is present, her predictor  $p_{\psi}$  is first updated to maximize the variational lower bound  $\mathcal{I}_{\mathrm{VLB}}$ . Then, with  $\psi$  fixed, Alice and Bob update their encoders to minimize the combined loss  $\mathcal{L}$ . This alternating optimization

## Algorithm 1 VPQ Training Algorithm

```
for each training step t = 1, 2, \ldots do
     Sample a data batch \{(x_i, y_i, z_i)\}_{i=1}^B
                                                   \triangleright z_i = \emptyset if Eve absent
     Alice and Bob outputs p_{\theta}(w|x_i), p_{\phi}(v|y_i) for all i
     Compute \mathcal{L}_{MR} according to (10)
     Compute p_t(w) and q_t(v) via EMA
     Compute \mathcal{L}_{ENT} according to (16)
     \mathcal{L}_{AB} \leftarrow \mathcal{L}_{ENT} + \lambda_1 \mathcal{L}_{MR}
     if Eve is present then
          if update \psi then
               Eve output p_{\psi}(w|z_i) for all i
               Compute \mathcal{I}_{VLB} according to (25)
                Update p_{\psi} by maximizing \mathcal{I}_{VLB}
          end if
          if update \theta, \phi then
               Eve output p_{\psi}(w|z_i) for all i
               Compute \mathcal{I}_{VUB} according to (31)
               \mathcal{L} \leftarrow \mathcal{L}_{AB} + \lambda_2 \mathcal{I}_{VUB}
               Update p_{\theta}, p_{\phi} by minimizing \mathcal{L}
          end if
     else
          Update p_{\theta}, p_{\phi} by minimizing \mathcal{L}_{AB}
     end if
end for
```

implements the adversarial training strategy: Eve learns to infer Alice's output as accurately as possible, while Alice and Bob adjust their encoders to minimize the information leaked to Eve.

## E. Secret Key Reconciliation

In the VPQ stage, Alice and Bob extract quantized sequences  $(W^n,V^n)$  from their observations  $(X^n,Y^n)$  without exchanging information, which are expected to be uniformly distributed and remain unpredictable by Eve. Usually, the mismatch rate  $\Pr\{W \neq V\}$  is a nonzero value, and thus the agreement rate between  $W^n$  and  $V^n$  decays exponentially with n. Consequently, a public discussion step is required to reconcile the sequences into a common secret key.

We adopt the secure sketch technique [15] for one-way reconciliation, ensuring that the public message remains independent of the final key. Specifically, based on the code-offset construction in [15], we consider the finite field  $\mathcal{F}=\mathrm{GF}(|\mathcal{W}|)$  and a  $[n,m,2t+1]_{\mathcal{F}}$  error-correcting code  $\mathcal{C}$  that can correct up to t symbol errors under Hamming distance. Alice uniformly samples a codeword C from  $\mathcal{C}$ , computes the offset

$$S = W^n - C \tag{36}$$

and transmits S publicly. Bob computes

$$C' = V^n + S \tag{37}$$

decodes it to  $\hat{C} \in \mathcal{C}$ . If the error between C and C' is within the correction capability, Bob can recover  $\hat{C} = C$ , and the final secret keys are set as K = C and  $L = \hat{C}$ . Note that

the subtraction and addition are defined over the finite field  $\mathcal{F}$  [29].

The resulting secret key rate is

$$\frac{1}{n}H(K) = \frac{m}{n}\log|\mathcal{W}|\tag{38}$$

because the codebook size is  $|\mathcal{C}| = |\mathcal{W}|^m$ . Thus, there exists a trade-off between the key rate and key agreement rate. In other words, increasing m improves the key entropy but reduces the error-correcting capability of  $\mathcal{C}$  and vice versa. highlights the importance of minimizing the mismatch probability in the VPO stage.

To analyze security, we have

$$I(K; Z^n, S) (39)$$

$$=H(Z^n,S)-H(Z^n,S|C)$$
(40)

$$= H(Z^n) + H(S|Z^n) - H(Z^n|C) - H(S|Z^n, C)$$
 (41)

$$= H(S|Z^{n}) - H(S, W^{n}|Z^{n}, C)$$
(42)

$$\leq n \log |\mathcal{W}| - H(W^n | Z^n, C) \tag{43}$$

$$= n\log|\mathcal{W}| - nH(W|Z) \tag{44}$$

$$= n\log|\mathcal{W}| - nH(W) + nI(W; Z), \tag{45}$$

where (42) holds because  $Z^n$  is independent of C and  $W^n$  is a function of S and C, (43) follows that condition doesn't increase entropy and  $S \in \mathcal{W}^n$ , (44) uses the fact that C is sampled independently of  $W^n, Z^n$  and the sequence  $\{(W_i, Z_i)\}_{i=1}^n$  is i.i.d.. Consequently, if W follows uniform distribution and I(W; Z) is arbitrarily small, the resulting key leakage rate  $\frac{1}{n}I(K; Z^n)$  is upper bounded by an arbitrarily small value. This confirms that the VPQ objective directly guarantees the secrecy of the reconciled keys.

**Remark 1.** Unlike the conventional usage of secure sketch in the PLK generation [18], [30], where Bob reconstructs  $W^n$  and both parties adopt a further privacy amplification step to extract secret keys to remove the leaked information contained in the public message, our proposed key reconciliation method uses the randomly sampled codeword as the final secret keys without any additional steps. The security performance of the generated keys is guaranteed by the VPQ step, and no further privacy amplification is necessary, as proved above.

# F. Case Study: PLK Generation from Fading Channels

PLK generation is one of the key enablers for PLS [9], where both legitimate parties, Alice and Bob, aim to extract the common secret keys from their wireless channel measurement [10], [11]. Traditional methods leverage the channel reciprocity property, meaning that the wireless channel from Alice to Bob is highly correlated with that from Bob to Alice within the channel coherence time. Most practical methods assume spatial decorrelation of Eve to Alice and Bob, thus omitting the secrecy requirement. However, the spatial decorrelation does not always hold true [17], leaving them vulnerable to key leakage. By contrast, our proposed learning-based CR extraction framework directly ensures secrecy in the quantization stage and is therefore well-suited for PLK generation.

We study the case of fading channels, where the estimated wireless channels at Alice and Bob are modeled by two correlated Gaussian random variables:

$$X = H + W_1, \quad Y = H + W_2,$$
 (46)

where  $H \sim \mathcal{N}(0,P)$  is the true channel between Alice and Bob, and  $W_1 \sim \mathcal{N}(0,N_1)$ ,  $W_2 \sim \mathcal{N}(0,N_2)$  are independent additive white Gaussian noise (AWGN). In our experiments, we set P=0 dBm,  $N_1=N_2=-20$  dBm. Algorithm 1 is first applied to learn the encoders at both parties, and the Reed-Solomon codes are then adopted to realize the proposed secret key reconciliation to extract final PLKs.

We first consider three cases of Eve: absent, uncorrelated, and correlated. That is, Eve observes  $Z = \emptyset$ , some independent random Gaussian noise, or  $Z = H + W_3$  for  $W_3 \sim \mathcal{N}(0, N_3)$  with  $N_3 = 0$  dBm. Alice and Bob share the same encoder  $p_{\theta}$ , implemented as a 4-layer fullyconnected network (FCN) with 1024 neurons per layer, batch normalization, and ReLU activation function. The input to the FCN is vectors of length 8, each component being an independent sample of X or Y. We set the batch size to B=2048, EMA factor  $\alpha=0.6$ . The training runs with Adam optimizer with learning rate  $3 \times 10^{-5}$  for a maximum 60,000 steps. To carry out the training with uncorrelated or correlated Eve, we build a larger 8-layer FCN with 2048 neurons per layer as the predictor  $p_{\psi}$ , updated once per training step with the same optimizer setting as the encoder, and in the last 10,000 steps, only the predictor is updated while freezing the encoder to obtain a tighter mutual information estimate  $\mathcal{I}_{VLB}$ . We evaluate  $|\mathcal{W}| \in \{16, 32, 64, 128\}$  with  $\lambda_1 = 1.0$  except  $\lambda_1 = 4.0$  for  $|\mathcal{W}| = 128$  user correlated Eve.  $\lambda_2$  is adaptively updated according to (35) during training.

After training, the encoder  $p_{\theta}$  and predictor  $p_{\psi}$  are tested with 81,920 data points. The test results are shown in Fig. 3, where "No Eve" denotes the training without an adversarial predictor, "U. Eve" indicates training with an adversarial predictor, but Eve observes independent random Gaussian noise, and "C. Eve" means that Eve observes correlated Z. In all cases,  $p_{\theta}$  achieves almost the maximal entropy H(W), implying that the outputs approach uniformity. The agreement rate  $\Pr\{W=V\}$  decreases with the dimension  $|\mathcal{W}|$ , and training with correlated Eve further reduces agreement due to the added unpredictability constraint. The test results of  $\mathcal{I}_{\text{VLB}}$  and  $\mathcal{I}_{\text{VUB}}$  reflect that both the variational lower and upper bounds of I(W;Z) are close to zero, indicating negligible leakage to Eve.

We then test the proposed secret key reconciliation step on the outputs of the trained encoder  $p_{\theta}$ . The Reed-Solomon codes  $RS(|\mathcal{W}|-1,m)$  are adopted with different choices of m, such that the secret key rate is given by

$$\frac{1}{n}H(K) = \frac{m}{n}\log|\mathcal{W}| = \frac{m}{|\mathcal{W}| - 1}\log|\mathcal{W}|. \tag{47}$$

The test results are illustrated in Fig. 4, demonstrating the relationship between the resulting key rate and key mismatch rate. It shows that the key mismatch rate  $\Pr\{K \neq L\}$  increases with the key rate because a larger m leads to lower error-correcting performance of the Reed-Solomon codes. Overall,

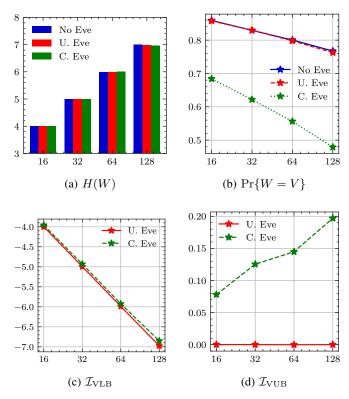


Fig. 3: Test results of the extracted sequence in VPQ stage for PLK generation from fading channels example. The x-axis is  $|\mathcal{W}|$ .

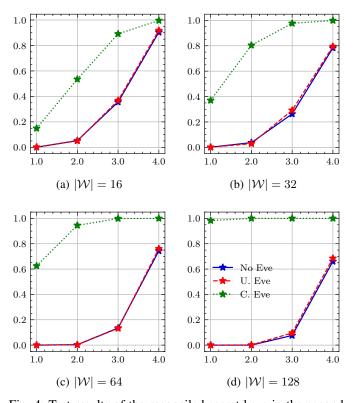


Fig. 4: Test results of the reconciled secret keys in the second stage for the fading channel example: key mismatch rate  $\Pr\{K \neq L\}$  vs. key rate  $\frac{1}{n}H(K)$  in bits.

these results demonstrate that the proposed framework can extract uniform, high-entropy, and secure keys from fading channels, even in the presence of a correlated Eve.

## IV. SENSING-BASED PHYSICAL LAYER KEY GENERATION

The traditional PLK generation based on channel reciprocity requires two-way channel probing between Alice and Bob, leading to large protocol overhead, especially in high-mobility scenarios, where the channel coherence time is too short to accommodate probing. To address this limitation, we propose a sensing-based PLK generation method enabled by emerging ISAC technology [19], which unifies communication and sensing within a common waveform. In our approach, only Alice transmits signals, while both Alice and Bob perform sensing. This design reduces the PLK update interval, thereby improving practicality under fast-varying conditions. Because Alice and Bob share the same propagation environment, their sensing outputs are expected to contain CR that can be exploited for PLK extraction. As a case study, we consider the scenario where Alice can detect an echo signal reflected from Bob in the presence of LoS path, and the measured range-angle (RA) information at both parties becomes highly correlated. Under high mobility conditions, Bob's position varies rapidly and can be modeled as an independent RV when its coherence time is short enough. To validate this concept, we conduct an end-to-end system simulation incorporating all relevant signal processing steps and channel effects. In addition, we develop a real-world SDR testbed to collect measurement data. The proposed method is first evaluated on synthesized data, after which the trained NN models are fine-tuned on the measured dataset to demonstrate generalizability and robustness.

# A. Channel Model and RA Map

We consider an ISAC system consisting of a static transmitter Alice and a dynamic receiver Bob. Alice is equipped with two co-located beamformers, enabling her to simultaneously transmit and receive signals. Bob is equipped with a receiving beamformer. For the purpose of simplified notations, the following channel model only considers 2D beamforming, and its generalization to 3D scenarios is straightforward.

In particular, Alice maps a set of data symbols and pilots to an orthogonal frequency-division multiplexing (OFDM) grid spanning  $N_{\rm sc}$  subcarriers and  $N_{\rm sym}$  OFDM symbols with subcarrier spacing  $\Delta f$  and symbol duration  $T_0$ , which are then modulated into the time domain as s(t) and sent out through her transmitting beamformer toward the angle  $\varphi$ . Both parties apply beamforming to receive the signal at the same time. We assume that a LoS path exists between Alice and Bob, and their antenna arrays are parallel. Let  $\theta_{\rm A}$  and  $\theta_{\rm B}$  be the receiving beamforming angle of Alice and Bob,  $y_{\rm A}(t)$  and  $y_{\rm B}(t)$  be the respective received signals, we have [31]

$$\begin{aligned} y_{\mathrm{A}}(t) &= \boldsymbol{a}_{\mathrm{Rx,A}}^{\mathsf{H}}(\theta_{\mathrm{A}}) \sum_{l=0}^{L_{\mathrm{A}}} \boldsymbol{H}_{\mathrm{A},l}(t) \boldsymbol{a}_{\mathrm{Tx}}(\varphi) s(t-\tau_{\mathrm{A},l}) + n_{\mathrm{A}}(t), \\ y_{\mathrm{B}}(t) &= \boldsymbol{a}_{\mathrm{Rx,B}}^{\mathsf{H}}(\theta_{\mathrm{B}}) \sum_{l=0}^{L_{\mathrm{B}}} \boldsymbol{H}_{\mathrm{B},l}(t) \boldsymbol{a}_{\mathrm{Tx}}(\varphi) s(t-\tau_{\mathrm{B},l}) + n_{\mathrm{B}}(t), \end{aligned}$$

where for  $* \in \{A, B\}^1$ ,  $n_*(t)$  is AWGN,

$$H_{*,l}(t) = b_{*,l} a_{\text{Rx},*}(\theta_{*,l}) a_{\text{Tr}}^{\text{H}}(\varphi_{*,l}) e^{j2\pi f_{D,*,l}t},$$
 (48)

and  $b_{*,l}$ ,  $\tau_{*,l}$ ,  $f_{D,*,l}$ ,  $\varphi_{*,l}$ ,  $\theta_{*,l}$  are the path attenuation factor, path delay, Doppler shift, angle of departure (AoD) and angle of arrival (AoA) of the l-th path, respectively.  $a_{\rm Tx}(\varphi)$  and  $a_{\rm Rx,*}(\theta)$  are the steering vectors of the transmitting and receiving beamformers in the direction  $\varphi$  and  $\theta$ . We denote path 0 as the LoS path such that  $\tau_{\rm A,0}=2\tau_{\rm B,0}$ ,  $\theta_{\rm A,0}=\theta_{\rm B,0}=\varphi_{\rm A,0}=\varphi_{\rm B,0}$  and  $|b_{\rm A,0}|\propto \sqrt{\sigma_{\rm RCS}}|b_{\rm B,0}|/2$  because the free space path loss is proportional to the square of distance and  $\sigma_{\rm RCS}$  is the radar cross section (RCS) of Bob. The maximum path delay is assumed to be less than the cyclic prefix (CP) length to guarantee the subcarrier orthogonality.

Alice and Bob then perform OFDM demodulation and channel estimation. If they are clock synchronized and demodulation is time aligned to transmission , the estimated OFDM channels at subcarrier  $n_{\rm sc}$  and OFDM symbol  $n_{\rm sym}$  allocated for pilots are obtained as

$$\hat{\boldsymbol{H}}_{*}(\theta_{*})[n_{\mathrm{sc}}, n_{\mathrm{sym}}] = \sum_{l=0}^{L_{*}} b_{*,l} \boldsymbol{a}_{\mathrm{Rx},*}^{\mathsf{H}}(\theta_{*}) \boldsymbol{a}_{\mathrm{Rx},*}(\theta_{*,l})$$

$$\cdot \boldsymbol{a}_{\mathrm{Tx}}^{\mathsf{H}}(\varphi_{*,l}) \boldsymbol{a}_{\mathrm{Tx}}(\varphi) e^{-j2\pi(n_{\mathrm{sc}}\tau_{*,l}\Delta f - n_{\mathrm{sym}}f_{D,*,l}T_{0})}$$

$$+ \boldsymbol{W}_{*}[n_{\mathrm{sc}}, n_{\mathrm{sym}}],$$

where  $W_*[n_{\rm sc}, n_{\rm sym}]$  is complex AWGN if pilots are phase shift keying (PSK) modulated [32].  $\hat{H}_*(\theta_*)[n_{\rm sc}, n_{\rm sym}]$  are then interpolated to obtain the channel estimates for the whole OFDM grid. Note that Alice can further apply clutter suppression techniques and use data symbols for channel estimation to improve the performance.

We take the channel estimate of a certain OFDM symbol, e.g.,  $n_{\rm sym}=0$ , to eliminate the impact of Doppler shifts. Let  $\hat{\boldsymbol{h}}_*(\theta_*)[n_{\rm sc}]=\hat{\boldsymbol{H}}_*(\theta_*)[n_{\rm sc},0]$ . For fixed  $\varphi$  and  $\theta_*$ , Alice and Bob apply inverse fast Fourier transform (IFFT) of length  $N_{\rm IFFT}$  to the zero padded channel estimates to calculate the channel range profile

$$r_*(\theta_*)[n] = \frac{1}{N_{\text{IFFT}}} \left| \sum_{n'=0}^{N_{\text{IFFT}}-1} \hat{h}_*(\theta_*)[n'] e^{j2\pi \frac{nn'}{N_{\text{IFFT}}}} \right|^2, \quad (49)$$

where  $\hat{h}_*(\theta_*)[n'] = 0$  if  $n' \ge N_{\rm sc}$ . A peak at index  $\hat{n}_*$  of  $r_*(\theta_*)$  corresponds to a target or an environmental scatter at distance

$$\hat{d}_{\mathrm{B}} = \frac{\hat{n}_{\mathrm{B}}c_{0}}{\Delta f N_{\mathrm{IFFT}}}, \quad \hat{d}_{\mathrm{A}} = \frac{\hat{n}_{\mathrm{A}}c_{0}}{2\Delta f N_{\mathrm{IFFT}}} \tag{50}$$

with the speed of light  $c_0$ , and the factor  $\frac{1}{2}$  in  $\hat{d}_A$  resulting from the round-trip propagation [32]. By repeating the above steps at different  $\theta_*$  and stacking the resulting range profiles, Alice and Bob construct their respective RA maps as the PLK generation source. It is also shown that the corresponding angle of a peak in the RA map indicates the AoA of a path, which is denoted as  $\hat{\theta}_A$  and  $\hat{\theta}_B$  at Alice and Bob, respectively. By assuming the existence of LoS path and that Alice performs

<sup>&</sup>lt;sup>1</sup>In the following, we use the subscript \* to denote both {A,B}, if there is no confusion

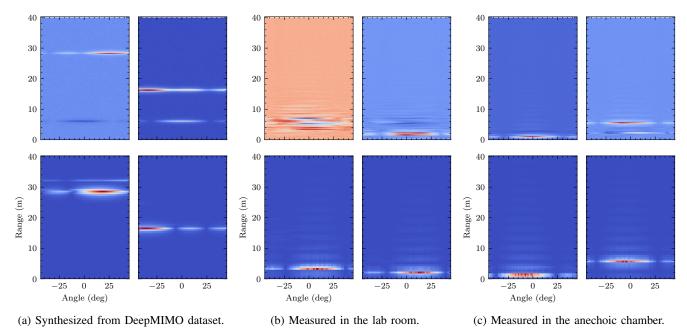


Fig. 5: Examples of synthesized and measured RA maps of Alice (above) and Bob (bottom).

clutter suppression, which subtracts the environment RA map from the RA maps containing Bob, the RA estimation of the strongest peak  $(\hat{d}_A, \hat{\theta}_A)$  and  $(\hat{d}_B, \hat{\theta}_B)$  should most likely coincide with each other, as shown in Fig. 5.

**Remark 2.** The RA maps can also be estimated using the MIMO technique, such that one time transmission is sufficient. The angle information is then extracted using spatial matched filters or superresolution methods such as MUSIC. In this case, the sensing overhead will be further reduced, leading to faster PLK generation. However, to keep the simulation setup consistent with our testbed, as detailed later, we apply the beam sweeping method in this work.

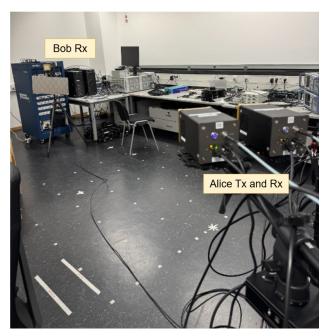
Remark 3. In practical scenarios of asynchronous bistatic sensing, Alice and Bob may not be perfectly synchronized in either time or angle domain, due to clock offsets or nonparallel beamformers. To mitigate this misalignment, the two parties can first establish a reference point at the beginning of PLK generation. Specifically, Alice transmits a probing signal to Bob, and both record the received timing and angle as their local references. These references are then used to align subsequent RA map measurements. In practice, the NNs can either (i) be trained on RA maps that have been pre-calibrated using the reference point, or (ii) take the raw RA maps together with the reference measurement as additional input features. This mechanism ensures that meaningful common randomness can still be extracted despite asynchrony. More general scenarios with practical imperfections, such as residual timing errors or beam misalignment, will be considered in future work.

#### B. Synthesized Dataset

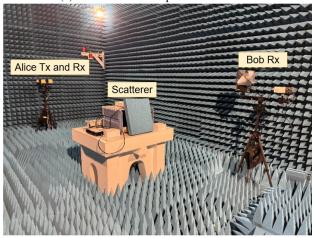
We first generate RA maps based on the DeepMIMO dataset [33], which contains the propagation path parameters of dif-

ferent scenarios synthesized by the 3D ray tracing software Remcom Wireless InSite [34]. Specifically, we extract the LoS path parameters of the DeepMIMO O1\_28 scenario and treat base stations as Alice, users as Bob, and each beamformer is specified as a  $4 \times 4$  uniform rectangular antenna array by the MATLAB phased array Toolbox [35]. Alice's beamformers are assumed to be 30° down tilt and horizontally directed to one of four directions  $\{-90^{\circ}, 0^{\circ}, 90^{\circ}, 180^{\circ}\}\$ , each covering a  $90^{\circ}$ azimuth sector, depending on its relative direction to Bob. On the other hand, Bob's beamformer is set to be 30° up tilt and parallel to Alice. As for the echo channel parameters, we add the reflection path between Alice and Bob to the Alice-to-Alice channel paths (the self-interference channel of base stations in the DeepMIMO dataset), where we modify the Alice-to-Bob LoS path by doubling the path delay and reducing the path gain by 6 dB with an additional randomly sampled  $\sigma_{\rm RCS}$ to simulate the echo path. The path parameters are then fed into the 3GPP clustered delay line (CDL) channel model [36] to construct channel objects in MATLAB using the 5G NR Toolbox.

The transmit signal is carried by the 5G PDSCH occupying 275 resource blocks with a  $120\,\mathrm{kHz}$  subcarrier spacing, corresponding to the full  $400\,\mathrm{MHz}$  bandwidth at the frequency band n257. The transmit beamforming direction is fixed to  $0^\circ$  in both azimuth and elevation angles, and processed by the constructed channel objects with AWGN added. With the received signals, Alice and Bob build their RA maps following the aforementioned steps by sweeping their receiving beams over 64 uniformly spaced azimuth angles in  $[-45^\circ, 45^\circ]$ . The range axis of the RA maps is then truncated to the maximum value allowed by the CP length.



(a) Measurement setup in the lab room.



(b) Measurement setup in the anechoic chamber.

Fig. 6: Hardware setup for RA map measurement.

#### C. Real-World Measurements

In addition to the synthesized dataset, we perform realworld measurements using the SDR technique in both a lab room and an anechoic chamber at the Advanced Communication Systems and Embedded Security (ACES) Lab of Technical University of Munich (TUM). As described above and shown in Fig. 6, Alice is equipped with an up/downconverter (TMYTEK UDBox) and two mmWave beamformers (TMYTEK BBox), one as the transmitter and the other one as the receiver, while Bob has one UDBox and one BBox as the receiver. Both Alice's UDBox and Bob's UDBox are connected to the same universal software radio peripheral (USRP) X410 for timing-synchronized transmission and reception. For transmission at Alice, we load the generated baseband PDSCH signal to the USRP X410, which is first converted to the intermediate frequency (IF) at 3.3 GHz, and then further upconverted by the UDBox to 28 GHz and sent out through the BBox. Simultaneously, the received signals at Alice and Bob are acquired by the USRP and saved to the host PC. During the measurement, Alice's transmitting beam is fixed toward  $0^{\circ}$ , while the receiving BBox of Alice and Bob performs beam sweeping in azimuth from  $-45^{\circ}$  to  $45^{\circ}$  with 64 beams in total. The received signals from all beams are processed using the method described above to obtain the RA maps. Therefore, the real-world measurement setup resembles the simulation, allowing us to train the NNs first on the synthesized large dataset and then fine-tune them on the measured dataset.

For both the lab room and anechoic chamber environments, we conduct measurements by placing Bob at different locations while keeping Alice fixed. After completing measurements at all locations, we remove Bob and let Alice perform another measurement, which can be used for clutter suppression to eliminate environmental scattering. Examples of synthesized and measured RA maps are shown in Fig. 5. These results demonstrate that the synthesized data is consistent with the measured data to some extent, but cannot always reflect the complex environment of the real world. The measurements in the lab room are also much noisier than those in the anechoic chamber due to the presence of more scatterers. Furthermore, both synthesized and real-world data validate our idea that the sensing information at Alice and Bob is expected to contain CR and thus can be used as the PLK source.

#### D. Experiments

We first apply Algorithm 1 to the synthesized dataset. The Transformer [37] is selected as the NN architecture for learning the individual encoder  $p_{\theta}$  and  $p_{\phi}$  from the paired RA maps. The Transformer is originally designed for sequence modeling, leveraging the self-attention mechanism to capture long-range dependencies in data. Since the RA maps are also naturally sequential along the range axis, the Transformer is suitable to extract low-dimensional but representative features from them. Nevertheless, other types of NNs, such as CNN or RNN, may also be chosen in place of the Transformer as the encoder. Given a RA map  $X \in \mathbb{R}^{N_r \times N_a}$ , we first apply the positional encoding along its range axis by adding X with a learnable vector of length  $N_r$ . The position-encoded matrix X' is fed into a multi-head self-attention block. Each selfattention layer projects X' linearly into query  $Q \in \mathbb{R}^{N_r \times d_k}$ , key  $K \in \mathbb{R}^{N_r \times d_k}$  and value  $V \in \mathbb{R}^{N_r \times d_v}$ , and performs the attention operation

$$\operatorname{Attention}(oldsymbol{Q}, oldsymbol{K}, oldsymbol{V}) = \operatorname{Softmax}\left(rac{oldsymbol{Q} oldsymbol{K}^ op}{\sqrt{d_k}}
ight) oldsymbol{V} \in \mathbb{R}^{N_r imes d_v},$$

where Softmax is applied along each row of the input matrix. The multi-head self-attention block comprises multiple independent self-attention layers and concatenates their outputs, which are then transformed linearly to the size  $N_r \times N_a$ . Subsequently, the block output passes through a FCN, followed by a residual connection and layer normalization. The combination of the above operations constitutes a Transformer layer. By stacking multiple such layers, one constructs the Transformer encoder. The Transformer encoder output has the same size as

Case	$\Delta D$ (meter)	$\Delta\Theta$ (degree)	H(W)	$\Pr\left\{W=V\right\}$	$\mathcal{I}_{ ext{VLB}}$	$\mathcal{I}_{ ext{VUB}}$
No Eve	-	-	3.958	0.946	-	-
Uncorrelated Eve	-	-	3.978	0.949	-3.963	0.003
	10	15	3.984	0.957	-3.576	1.107
Correlated Eve		10	3.985	0.935	-3.763	0.552
		5	3.992	0.955	-3.800	0.457
	5	15	3.979	0.901	-3.747	0.497
		10	3.983	0.927	-3.755	0.489
		5	3.988	0.934	-3.802	0.414
	3	15	3.988	0.879	-3.674	0.636
		10	3.988	0.885	-3.695	0.594
		5	3.991	0.874	-3.707	0.572
	1	15	3.993	0.799	-3.533	0.767
		10	3.990	0.803	-3.514	0.780
		5	3.996	0.803	-3.495	0.828

TABLE I: Test results of VPQ models trained on the synthesized dataset.

the input, which is then truncated to obtain the logits before the final softmax layer.

We set  $|\mathcal{W}|=16$  throughout the experiments. Alice and Bob use independent Transformer encoders as their RA maps have different range resolutions and patterns. For the case of absent Eve, we set  $\lambda_1=1.0$ , each encoder is trained with an individual AdamW optimizer [38] with the same setting of learning rate  $10^{-4}$  and weight decay of  $10^{-4}$ . Then, we assume that Eve has an estimation of Bob's relative position to Alice with different levels of uncertainty. Let  $(d,\theta)$  be the true relative range (meter) and angle (degree) of Bob to Alice, then Eve's estimation is

$$\hat{d}_{\rm E} = d + \Delta d, \ \hat{\theta}_{\rm E} = \theta + \Delta \theta,$$
 (51)

with  $\Delta d, \Delta \theta$  uniformly distributed within  $[-\frac{\Delta D}{2}, \frac{\Delta D}{2}]$  and  $[-\frac{\Delta \Theta}{2}, \frac{\Delta \Theta}{2}]$ , respectively. We set  $\Delta D \in \{1,3,5,10\}$  in meter and  $\Delta \Theta \in \{5,10,15\}$  in degree in the experiment. For all the experiments trained with adversarial strategy, i.e., with Eve, we set  $\lambda_1=1.0, \ \lambda_2=2.0,$  and the AdamW optimizer with learning rate  $5\times 10^{-5}$  and weight decay  $1\times 10^{-4}$  is used for each encoder and predictor, and in the last training 50 epochs only Eve's predictor  $p_{\psi}$  is trained with Alice's and Bob's encoders frozen, as in the fading channel case, to obtain a tighter lower bound  $\mathcal{I}_{\text{VLB}}.$  As a comparison, we also consider the case where Eve's estimation  $(\hat{d}_{\text{E}}, \hat{\theta}_{\text{E}})$  is totally random, i.e., uncorrelated to the RA maps at Alice and Bob. We expect that the uncorrelated case should lead to the same result as the case without Eve.

The experimental results of the VPQ stage for the synthesized RA map dataset are summarized in Table I. The proposed learning framework extracts RVs (W,V) that are nearly uniform, with H(W) approaching  $\log |\mathcal{W}| = 4$  across all scenarios. The agreement rate between W and V exceeds 90% when Eve is absent, uncorrelated, or has relatively large uncertainty in her position estimates. As Eve's estimation

accuracy improves, corresponding lower  $\Delta D$  and  $\Delta \Theta$ , the encoder performance degrades as expected, reflected either by lower agreement rate or larger  $\mathcal{I}_{\text{VLB}}$  and  $\mathcal{I}_{\text{VUB}}$ . Nonetheless, (W,V) remain highly unpredictable in all cases, meaning that the CR source for sensing-based PLK generation does not solely come from Bob's location information but also arises from shared scattering and channel fluctuations.

The models trained on the synthesized dataset are then fine-tuned on the measured data. Since the dataset size of each measurement environment contains fewer than 200 data points, training from scratch or fine-tuning the entire pretrained models easily leads to overfitting. To mitigate this, we replace the output linear layer of both Alice's and Bob's pretrained encoders with a new randomly initialized linear layer and freeze the remaining parameters. Only the output linear layers are trained on the measurement dataset to avoid overfitting. If Eve is present, all her predictor parameters are fine-tuned. Since all Bob's locations are closed to Alice in both real-world datasets, we only consider the extreme case with  $\Delta D = 1 \text{m}$ ,  $\Delta \Theta = 5^{\circ}$ for the correlated Eve's observations. The test results with and without the fine-tuning strategy are reported in Table II and Table III for the lab room and anechoic chamber environment, respectively. Fine-tuning significantly improves both entropy H(W) and agreement rate  $Pr\{W = V\}$ , demonstrating that pretrained models capture meaningful low-dimensional features from the synthesized dataset. Interestingly, the anechoic chamber yields higher performance than the lab room when Eve is absent or uncorrelated, whereas the lab room produces closer encoder outputs under correlated Eve. This can be attributed to the richer scattering and reflections in the lab environment, which provide additional CR sources beyond pure location information. This observation is also reflected in the reconciled secret key results with RS(15, m) codes, as shown in Fig. 7.

TABLE II: Test results of VPQ stage for RA maps measured in lab room with / without fine-tuning

Case	H(W)	$\Pr\left\{W=V\right\}$	$\mathcal{I}_{ ext{VLB}}$	$\mathcal{I}_{ ext{VUB}}$
No Eve	3.675 / 2.768	0.747 / 0.214	-	-
Uncorrelated Eve	3.780 / 2.662	$0.662\ /\ 0.144$	$-4.050 \ / \ -4.077$	$-0.005 \ / \ 0.002$
Correlated Eve ( $\Delta D=1, \Delta\Theta=5$ )	3.882 / 2.715	$0.605\ /\ 0.004$	$-3.497 \ / \ -4.917$	$0.741 \ / \ -0.063$

TABLE III: Test results of VPQ stage for RA maps measured in anechoic chamber with / without fine-tuning

Case	H(W)	$\Pr\left\{W=V\right\}$	$\mathcal{I}_{ ext{VLB}}$	$\mathcal{I}_{ ext{VUB}}$
No Eve	3.744 / 2.775	0.879 / 0.403	-	-
Uncorrelated Eve	3.707 / 3.191	$0.762 \ / \ 0.333$	-3.996 / -3.947	$-0.031 \ / \ -0.003$
Correlated Eve ( $\Delta D = 1, \Delta \Theta = 5$ )	3.953 / 3.162	$0.490 \ / \ 0.060$	$-3.468 \ / \ -5.416$	$0.820 \ / \ -0.220$

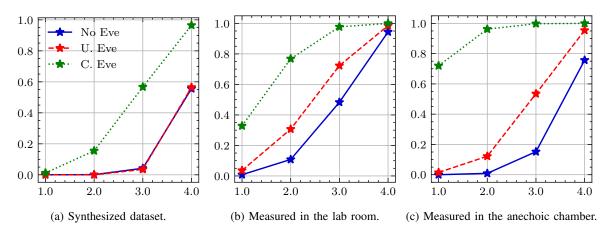


Fig. 7: Key mismatch rate vs. key rate (in bits) for the secret keys generated from synthesized dataset, measured dataset from the lab room and the anechoic chamber. C. Eve corresponds to the case of  $\Delta D = 1$ m,  $\Delta \Theta = 5^{\circ}$ .

#### V. CONCLUSION

In this work, we introduced a variational CR extraction framework, consisting of two stages. In the first stage, VPQ learns probabilistic encoders that quantize correlated observations at Alice and Bob into nearly uniform and highly correlated RVs while suppressing information leakage to Eve via an adversarial mutual information objective. In the second stage, a secure sketch based on the code-offset construction reconciles the quantized outputs into identical secret keys with theoretical secrecy guaranteed.

The proposed framework was validated extensively. We first demonstrated its effectiveness on fading channel models, showing that it can achieve near-maximal entropy, high agreement rates, and negligible leakage even in the presence of correlated eavesdroppers. We then applied the framework to sensing-based PLK generation in ISAC systems, where RA maps serve as the source of CR. Both end-to-end 5G NR simulations and real-world SDR measurements confirmed that the framework can reliably extract secure keys from sensing information, while transfer learning enables pretrained models to generalize effectively across environments. Compared with conventional PLK schemes that rely on reciprocity and require two-way channel probing, our method reduces protocol overhead, supports high-mobility scenarios, and naturally in-

tegrates secrecy without a separate privacy amplification step.

Looking forward, an analysis of the gap between the proposed learning-based CR framework and the information-theoretic CR capacity is necessary. Additionally, a theoretical characterization of sensing-based PLK generation, including fundamental limits of achievable key rates under sensing constraints, is of particular interest. Moreover, extending the framework to multi-user and distributed deployments, as well as validating its performance on larger-scale real-time testbeds, could further broaden its applicability.

#### REFERENCES

- R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, 2002.
- [2] —, "Common randomness in information theory and cryptography. ii. CR capacity," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 225–240, 2002.
- [3] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE transactions on information theory*, vol. 39, no. 3, pp. 733–742, 2002.
- [4] R. Ahlswede and G. Dueck, "Identification via channels," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 15–29, 2002.
- [5] C. Portmann and R. Renner, "Security in quantum cryptography," Reviews of Modern Physics, vol. 94, no. 2, p. 025008, 2022.
- [6] P. Gács, J. Körner et al., "Common information is far less than mutual information." Problems of Control and Information Theory, vol. 2, pp. 149–162, 1973.

- [7] A. El Gamal and Y.-H. Kim, Network information theory. Cambridge university press, 2011.
- [8] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 344–366, 2002.
- [9] V.-L. Nguyen, P.-C. Lin, B.-C. Cheng, R.-H. Hwang, and Y.-D. Lin, "Security and privacy for 6g: A survey on prospective technologies and challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2384–2428, 2021.
- [10] K. Ren, H. Su, and Q. Wang, "Secret key generation exploiting channel characteristics in wireless communications," *IEEE Wireless Communications*, vol. 18, no. 4, pp. 6–12, 2011.
- [11] K. Zeng, "Physical layer key generation in wireless networks: challenges and opportunities," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 33–39, 2015.
- [12] D. S. Bhatti, H. Choi, and H.-N. Lee, "Beyond traditional security: A review on information-theoretic secret key generation at wireless physical layer," *Authorea Preprints*, 2024.
- [13] B. Poole, S. Ozair, A. Van Den Oord, A. Alemi, and G. Tucker, "On variational bounds of mutual information," in *International conference* on machine learning. PMLR, 2019, pp. 5171–5180.
- [14] P. Cheng, W. Hao, S. Dai, J. Liu, Z. Gan, and L. Carin, "Club: A contrastive log-ratio upper bound of mutual information," in *International conference on machine learning*. PMLR, 2020, pp. 1779–1788.
- [15] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," in *International* conference on the theory and applications of cryptographic techniques. Springer, 2004, pp. 523–540.
- [16] T. Aono, K. Higuchi, T. Ohira, B. Komiyama, and H. Sasaoka, "Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 11, pp. 3776–3784, 2005.
- [17] J. Zhang, R. Woods, T. Q. Duong, A. Marshall, Y. Ding, Y. Huang, and Q. Xu, "Experimental study on key generation for physical layer security in wireless communications," *IEEE Access*, vol. 4, pp. 4464–4477, 2016.
- [18] Q. Wang, H. Su, K. Ren, and K. Kim, "Fast and scalable secret key generation exploiting channel phase randomness in wireless networks," in 2011 Proceedings IEEE INFOCOM. IEEE, 2011, pp. 1422–1430.
- [19] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6g and beyond," *IEEE journal on selected areas in communications*, vol. 40, no. 6, pp. 1728–1767, 2022.
- [20] N. Su, F. Liu, and C. Masouros, "Secure radar-communication systems with malicious targets: Integrating radar, communications and jamming functionalities," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 83–95, 2020.
- [21] N. Su, F. Liu, Z. Wei, Y.-F. Liu, and C. Masouros, "Secure dual-functional radar-communication transmission: Exploiting interference for resilience against target eavesdropping," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 7238–7252, 2022.
- [22] X. Wang, Z. Fei, P. Liu, J. A. Zhang, Q. Wu, and N. Wu, "Sensing aided covert communications: Turning interference into allies," *IEEE Transactions on Wireless Communications*, 2024.
- [23] A. D. Wyner, "The wire-tap channel," *Bell system technical journal*, vol. 54, no. 8, pp. 1355–1387, 1975.
- [24] N. Su, F. Liu, J. Zou, C. Masouros, G. C. Alexandropoulos, A. Mourad, J. L. Hernando, Q. Zhang, and T.-T. Chan, "Integrating sensing and communications in 6G? not until it is secure to do so," arXiv preprint arXiv:2503.15243, 2025.
- [25] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," SIAM Journal on Applied Mathematics, vol. 28, no. 1, pp. 100–113, 1975.
- [26] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on information Theory*, vol. 19, no. 4, pp. 471–480, 1973.
- [27] D. Barber and F. Agakov, "The im algorithm: a variational approach to information maximization," *Advances in neural information processing* systems, vol. 16, no. 320, p. 201, 2004.
- [28] P. Esser, R. Rombach, and B. Ommer, "Taming transformers for high-resolution image synthesis," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 12873–12883.
- [29] R. Roth, Introduction to coding theory. Cambridge University Press, 2006.
- [30] J. Zhang, T. Q. Duong, A. Marshall, and R. Woods, "Key generation from wireless channels: A review," *IEEE Access*, vol. 4, pp. 614–626, 2016.

- [31] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, "An overview of signal processing techniques for joint communication and radar sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1295–1315, 2021.
- [32] K. M. Braun, "Ofdm radar algorithms in mobile communication networks," Ph.D. dissertation, Karlsruhe, Karlsruher Institut für Technologie (KIT), Diss., 2014, 2014.
- [33] A. Alkhateeb, "Deepmimo: A generic deep learning dataset for millimeter wave and massive mimo applications," arXiv preprint arXiv:1902.06435, 2019.
- [34] Remcom, "Wireless InSite," http://www.remcom.com/wireless-insite.
- [35] T. M. Inc., "Matlab version: 23.2 (r2023b)," Natick, Massachusetts, United States, 2023. [Online]. Available: https://www.mathworks.com
- [36] 3GPP TR. 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project; Technical Specification Group Radio Access Network, Technical Report, 2020.
- [37] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.
- [38] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," arXiv preprint arXiv:1711.05101, 2017.