Filterbank Secret Key Generation Rates in Multipath Channels

Miroslav Mitev, André N. Barreto, Thuy M. Pham, Maximilian Matthé, Gerhard Fettweis
Barkhausen Institut, Dresden, Germany
{miroslav.mitev, andre.nollbarreto, minhthuy.pham, maximilian.matthe, gerhard.fettweis}@barkhauseninstitut.org

***Abstract*—The sixth generation of wireless networks (6G) is expected to support the deployment of Internet of things (IoT) devices in massive scales. Finding lightweight and decentralized secret-key distribution primitives is therefore a challenge. Secret-key generation (SKG) from wireless channel coefficients is seen as a possible solution. It allows the extraction of secret keys using the channel randomness observed at the physical layer, without a centralized key distribution server. In this work an SKG approach suitable for wideband IoT devices is proposed. We investigate a filterbank-based SKG method, in which secret bits are generated through power measurements over different frequencies. To minimize dependencies and correlation among frequencies the quantile and the Karhunen-Loève transforms are used. Finally, we perform a numerical evaluation of the achievable SKG rates, in the form of mutual-information (MI) estimates, using 3GPP channel models. Our numerical evaluation shows that the achievable SKG rate depends on the channel statistics, and, hence, to optimally harvest the information, devices need to be channel-aware.**

***Index Terms*—IoT, frequency selective channels, mutual information, physical layer security, secret key generation.**

I. Introduction

The current security framework of connected wireless systems relies on public infrastructure and complex modulo arithmetic operations. However, todays' security algorithms may not satisfy the needs of future Internet of things (IoT) devices. In fact, public-key encryption (PKE) protocols do not satisfy the scalability and low-energy requirements of massive IoT deployment [1]. Furthermore, quantum computing may pose a threat to asymmetric cryptography mechanisms, which play a vital role in standard security solutions, unless key sizes increase to impractical lengths [2]. Therefore, finding lightweight alternatives to PKE is of great interest.

A promising and quantum-secure key distribution technique, considered for 6G [3], is the physical-layer security (PLS) based secret key generation (SKG). In SKG devices extract a shared secret using the reciprocity of wireless channels. However, the achievable key generation rates strongly depend on the statistical properties of the channel, and, depending on these, channel response in neighbouring frequencies may be correlated and dependent. This impacts the security of the generated key, because correlation and dependence between key bits can greatly reduce the search space for brute-force attacks [4]. To overcome this problem different techniques have been proposed [4], [5]. [4] considers an orthogonal frequency division multiplexing (OFDM) system, where keys are generated using a set of subcarriers. As these are correlated, [4] proposes to transform the channel measurements using principal component analysis (PCA), which successfully eliminates the correlation among the measurements, as shown through numerical evaluation. A similar approach is taken in [5], where an OFDM system with multiple antennas is considered. To remove the correlation between measurements from different antennas and subcarriers, the Karhunen-Loève transform (KLT) is applied. Through an experimental setup it is shown that the KLT can remove the correlation between measurements and provide stable key generation rates. Both works show promising results, but uncorrelated observations do not imply that they are also independent, and, hence, further investigation is required to achieve both.

This work proposes a transformation technique that successfully minimizes both correlation and dependence. We focus on SKG for wideband communication, in which the channel randomness is harvested using a filterbank. The system model includes two legitimate nodes, Alice and Bob, who perform SKG, and a passive adversary, Eve. This work is a continuation of our recent findings in [6], [7], with the concept of filterbank SKG being first presented in [6]. In this method, a set of filters are used to divide the communication bandwidth into multiple sub-bands and, then, keys are generated from power measurements for each sub-band. That is, unlike the OFDM approach from [4], [5], we do not use the phase information. The advantages of our proposal are: i) it does not require full CSI information and can be employed without pilot symbols; ii) it does not require matched receivers, as long as the power spectral density of the transmitted signal is known; iii) it eliminates the need for fine synchronization in time and frequency. The distribution of the power measurements and the achievable SKG rate when Alice and Bob apply only a single filter were evaluated in [7]. Our findings show that the SKG rates strongly depend on channel parameters such as bandwidth, delay spread and number of resolvable multipath components (MPCs). Therefore, bringing SKG into practice would require devices to be channel-aware and able to adaptively optimize their parameters depending on the channel conditions.

We extend here the findings from [6], [7], by evaluating the SKG rates considering the filterbank SKG method. Unfortunately, power measurements taken at neighbouring frequencies may be strongly correlated to each other, and as discussed above, would produce insecure keys after quantization. To overcome this, we propose to use two transformation techniques, namely, quantile transform and KLT, in a sequential manner. KLT is a well-known linear transform, which is used here to decompose the power measurements into a set of uncorrelated

components. Note that if the KLT input has a jointly Gaussian distribution, then the components at its output will be both uncorrelated and independent, but independence cannot be guaranteed for other distributions. Therefore, we propose that before performing the KLT, measurements should go through a quantile transform. This is a technique that transforms numerical measurements to Gaussian variables. Even though these may be not jointly Gaussian, we show that the dependence between the components after both transforms is successfully minimized. Both are invertible, and, hence, do not impact the MI from initial measurements [8].

Another motivation for these transformations is that measuring MI for multi-dimensional (multi-filter in our scenario) systems is a fundamental problem [9]. Minimizing the correlation and dependence between the dimensions, by using the proposed transforms, allows us to convert the complex system to a set of independent and tractable one-dimensional systems. Through numerical evaluation we estimate the MI between Alice and Bob using a 3GPP tapped delay line (TDL) channel model [10] and show that the proposed approach successfully overcomes this fundamental limitation of the MI estimation.

The rest of the paper is organized as follows: Section II introduces the system model and our filterbank approach. It also shows that the proposed transformations successfully minimize the dependency among channel measurements in different frequencies. Next, a numerical analysis of the SKG rates is discussed in Section III, and Section IV concludes this paper.

II. Filterbank secret key generation

A. Secret key generation

Typically, the SKG process is based on three steps: i) advantage distillation, during which Alice and Bob estimate the reciprocal channel between them; ii) information reconciliation, in which channel estimates are quantized to bits and bit mismatch errors are corrected through a public exchange of helper data [11]; iii) privacy amplification, in which it is ensured that the publicly exchanged data does not reveal information about the key, and for that purpose Alice and Bob apply a one-way compression function, such as a universal hash function [12].

For the successful deployment of SKG it is of utmost importance to answer the following question: *How many secret bits can be extracted from different channels?* In this work we estimate the achievable SKG rates during advantage distillation. This information would allow receivers to flexibly adapt their parameters and report the number of available secret bits to upper layers. Investigation of other steps in the SKG process is out of the scope for this paper. Finally, it is assumed that the eavesdropper Eve is at least a few wavelengths away from the legitimate users and, due to spatial decorrelation in wireless channels, we can assume that her measurements will be uncorrelated with measurements at Alice and Bob [1]. Hence, Eve is excluded from the analysis in the next sections.

B. System model

The channel model considered in this work is a multi-path channel with additive white Gaussian noise (AWGN). For the analysis, we assume that Alice and Bob exchange passband modulated signals, which upon reception are first converted to baseband and then passed through a filterbank. Note that the described signals are all complex baseband signals. Their observations are consequently denoted as:

$$y_{A,n}(t) = g_n(t) * \left[x(t) * h(t) + w_A(t)\right], \tag{1}$$

$$y_{B,n}(t) = g_n(t) * \left[x(t) * h(t) + w_B(t)\right], \tag{2}$$

where $x(t)$ is the transmitted baseband signal with bandwidth $B/2$, $h(t)$ represents a complex time-invariant channel impulse response (CIR)[1], $w_A(t) \sim \mathcal{N}(0, \sigma_A^2)$, $w_B(t) \sim \mathcal{N}(0, \sigma_B^2)$ are AWGN variables observed at Alice and Bob, respectively, and $g_n(t)$ denotes the impulse response of the n-th filter within the filterbank, with $n = 1, \ldots, N$. Each individual filter can be represented as $g_n(t) = g(t)e^{(j2\pi f_n(t))}$, where the center frequency of each filter is calculated as $f_n = -\frac{B(N-2n+1)}{2N}$, and $g(t)$ is a prototyping filter.

The CIR considered for this study is based on the 3GPP TDL channel model [10]. There are 5 different TDL channel models, (from A to E), all having a root-mean-square delay spread of 1 ns, but each one with a different power delay profile. A nice property of these models is that they allow one to easily scale the delay spread to any desired value, i.e., the propagation delay of the l-th MPC can be obtained as $\tau_l = D\tau_{l,\text{TDL}}$, with $\tau_{l,\text{TDL}}$ the corresponding propagation delay of the reference TDL power delay profile, and D the value of the desired delay spread in ns [10]. Given that, the CIR in the time domain is given as:

$$h(t) = \sum_{l=0}^{L-1} \alpha_l \delta(t - \tau_{l,}), \tag{3}$$

where L denotes the number of MPCs, $\delta(t)$ is the Dirac delta function, and α_l is the complex amplitude for the l-th MPC. As per [10], the magnitude of each MPC, $|\alpha_l|$, could follow either Rayleigh or Ricean distribution. For this work we assume that $|\alpha_l|$ are independently distributed Rayleigh variables.

C. Extracting channel randomness

To extract the randomness from the channel, Alice and Bob measure the power at the output of each filter within their filterbank. The power measurements are taken in the digital domain with the assumption that the legitimate users want to also demodulate the signals, hence, they sample at the Nyquist rate. Their power measurements over a set of multiple frames, $j = 1, \ldots, J$, are then denoted as:

$$\hat{\mathbf{P}}_A = \begin{bmatrix} \hat{p}_{A_{1,1}} \cdots \hat{p}_{A_{1,N}} \\ \vdots \\ \hat{p}_{A_{J,1}} \cdots \hat{p}_{A_{J,N}} \end{bmatrix}. \tag{4}$$

Assuming measurements are taken over long periods, the power of the sampled version of the filtered signal can be approximated as the average of the continuous-time signals, i.e., $\hat{p}_{A_{1,n}} = \overline{|y_{A,n}(t)|^2}$ with $\overline{(\cdot)}$ denoting time-averaging operator. We assume that consecutive channel measurements

[1]To simplify the notation, the CIR is assumed constant during a frame.

$\hat{p}_{A_{1,n}}, \hat{p}_{A_{2,n}}, \ldots, \hat{p}_{A_{J,n}}$ are taken at intervals greater than the channel coherence time and are, therefore, considered to be independent. This is a realistic assumption, as for example in the sub-6 GHz frequency bands, (used in many IoT scenarios, including wideband applications), the coherence time of the channel could be as low as 2 ms for low-mobility applications [13].

In our earlier work [7], the probability distribution of a single power measurement was evaluated. In fact, it is a sum of Gamma-distributed variables $\sum_{m=1}^{M} \Gamma(\kappa_m, \theta_m)$, each parameterized by shape κ_m and scale θ_m values. The number of these Gamma variables is equal to the number of resolvable MPCs, denoted here by M, with $M \leq L$. The final distribution for a single power measurement, i.e., $\hat{p}_{A_{j,n}}$ was evaluated as:

$$\Gamma\left(\kappa=\frac{(\sum_{m=1}^{M}\kappa_m\theta_m+2\sigma_A^2)^2}{\sum_{m=1}^{M}\kappa_m\theta_m^2+\frac{8\sigma_A^4}{K-1}}, \theta=\frac{\sum_{m=1}^{M}\kappa_m\theta_m+2\sigma_A^2}{\kappa}\right), \quad (5)$$

where K is the number of samples falling within the filter bandwidth. As mentioned above, Alice and Bob sample at the Nyquist rate, hence a wider bandwidth will correspond to a higher number of noise observations. Two conclusions were drawn from these findings: i) by increasing the delay spread, the number of resolvable MPCs, $M \in \{0, L\}$, increases, forming a less random distribution of the power, which leads to a decrease in the MI between Alice and Bob; ii) similarly, by increasing the bandwidth, the number of resolvable MPCs, $M \in \{0, L\}$, increases, forming a less random distribution. However, as discussed above, increasing the bandwidth results in more observations K in a given time frame, and, as it can be seen in (5), increasing K minimizes the impact of the noise variance σ_A^2. Interestingly, it was demonstrated that, due to this effect, an increase in the bandwidth leads to an increase in the MI between the legitimate users.

Here, we extend our work to a filterbank approach, in which correlated power measurements from different filters are used for key generation. We have considered the prototyping filter $g(t)$ to be a raised-cosine filter for a symbol rate B/N with roll-off factor $\rho = 0.25$. Figure 1 shows an example of an arbitrary frame considering a filterbank with 10 filter elements. It illustrates the estimation of the power spectral densities (PSDs) of the received signals at Alice and Bob as well as the measured power at the filters' outputs. Note that this is not the real PSD but the estimate obtained using Welch's method for a measurement period $T \approx 4$ μs. The transmit signal is a 500 MHz chirp with constant modulation, which is passed through a TDL-A channel with 100 ns delay spread and signal to noise ration (SNR) of 5 dB. As correlation reduces the entropy of the observed vector, an ideal scenario would be with uncorrelated and independent power measurements from different filters. However, this is not a realistic assumption and power measurement at Alice, $\hat{p}_{A_{j,1}}, \ldots, \hat{p}_{A_{j,N}}$ are correlated to each other. On the other hand, thanks to the reciprocity of the wireless channel, Alice's measurements are also coupled to those at Bob's side $\hat{p}_{B_{j,1}}, \ldots, \hat{p}_{B_{j,N}}$. While the correlation between Alice's and Bob's measurements allows them to obtain

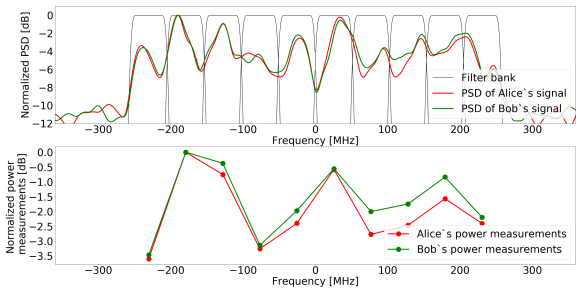

Fig. 1: PSD of received signals, filterbank constructed of 10 filters and power measurements at the output of the filterbank. All values are normalized to 0dB.

a shared secret, the correlation between measurements taken at different frequencies results in correlated bits in the SKG key. To overcome this, in the next section, we propose a technique that minimizes the correlation and dependency among measurements taken at different frequencies, without reducing the MI between both parties. Another advantage of this approach is that it allows the evaluation of the achievable SKG rate between Alice and Bob, which as mentioned in Sec. II-A, is of vital importance for the successful deployment of SKG in practical systems.

D. Measurement transformation

In this section, we describe the proposed transformation techniques. The goal is to minimize dependency and correlation between adjacent filter measurements taken at Alice and Bob, while keeping the MI between the two parties unaffected. To achieve this, two transformations are performed in sequential manner. First, the measurements are passed through a quantile transformation. This is a widely used transformation for data normalization. It transforms all data features into normal random variables. It does not guarantee that they are jointly normal, though. In this work, data features correspond to the power measurements taken at the output of each filter. Once the features are transformed into normal variables, we propose to apply a second transform, known as KLT, which is a technique used for data compression, by redistributing the information into a set of uncorrelated components through a linear transformation.

As mentioned earlier, both transforms are invertible, and, hence, preserve the information from the initial measurements [8]. The reasoning for using both transforms is as follows: if the KLT inputs were jointly normal, then its output components would not only be uncorrelated but also independent. Unfortunately, quantile transform does not ensure joint normality. However, as shown later in Section II-E, it still allows the KLT to minimize the dependencies to a negligible level.

1) Quantile transform: Quantile transform is a preprocessing technique used to reduce the impact of marginal outliers by transforming data features to a normal distribution [14]. In the scenario considered here, Alice and Bob measure the power at the output of their filterbanks. Assuming measurements are taken using N filters, the source for their

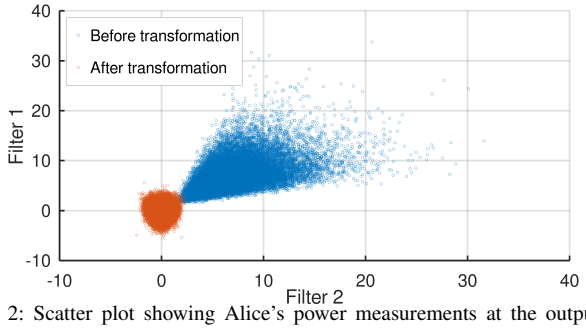

Fig. 2: Scatter plot showing Alice's power measurements at the output of 2 filters before and after transformation.

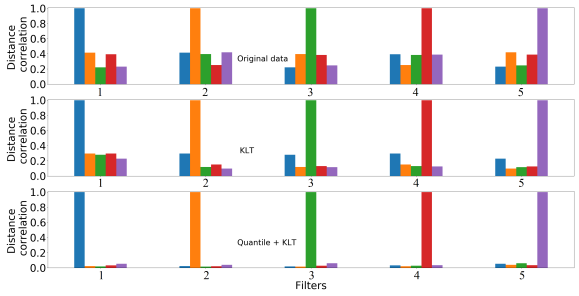

Fig. 3: Distance correlation measure of power measurements before and after transformation. The considered SNR of the measurements is $\gamma = 10$ dB.

key generation are the matrices $\hat{\mathbf{P}}_A$ and $\hat{\mathbf{P}}_B$, as defined in (4). Quantile transformation is then performed filter-wise as:

$$\dot{\mathbf{P}}_A = \left[Q\left(F_1 \begin{bmatrix} \hat{p}_{A_{1,1}} \\ \vdots \\ \hat{p}_{A_{J,1}} \end{bmatrix} \right), \cdots, Q\left(F_N \begin{bmatrix} \hat{p}_{A_{1,N}} \\ \vdots \\ \hat{p}_{A_{J,1}} \end{bmatrix} \right) \right], \tag{6}$$

where F_i, with $i = 1, \ldots, N$ is the cumulative distribution function (CDF) of the power measurements corresponding to taken at the output of the i-th filter, and Q is the quantile function of the normal distribution, which is equal to its inverse CDF [14]. Similarly, Bob obtains $\dot{\mathbf{P}}_B$.

2) Karhunen Loève transform: KLT is a linear decomposition technique that can transform stochastic processes to pairwise uncorrelated random variables. The transformation is applied over $\dot{\mathbf{P}}_A$ and $\dot{\mathbf{P}}_B$ in two steps. First, the covariance matrix is calculated. Since the transformed power measurements $\dot{\mathbf{P}}_A$ (and $\dot{\mathbf{P}}_B$) are zero-mean processes, the empirical covariance matrix is equal to the dot product $\mathbf{R}_A = \dot{\mathbf{P}}_A \dot{\mathbf{P}}_A^T$. Next, the transformation is applied as $\ddot{\mathbf{P}}_A = \mathbf{A}^T \dot{\mathbf{P}}_A$, where $\mathbf{A} \in \mathbb{R}^{N \times N}$ contains the eigenvectors of $\mathbf{R}_A$ sorted in a descending order according to their eigenvalues. Note that the total energy of the input is preserved, $||\dot{\mathbf{P}}_A||^2 = ||\ddot{\mathbf{P}}_A||^2$,with $||\cdot||$ the Frobenius norm operator, but it will be distributed along the output features in a descending order (due to the descending order of the eigenvectors). Therefore, the first feature has the highest energy and will bring the most information, and the last brings the least information and would typically be dominated by noise.

Figure 2 illustrates an example of $\hat{\mathbf{P}}_A$ and $\ddot{\mathbf{P}}_A$ when $N = 2$ filters. The axes represent the power measurements taken at the output of the two filters. It can be seen that before transformation the measurements are highly correlated, while after transformation the points are centered around zero with approximately equal variance in all directions. Not only correlation but also dependencies between different filters at Alice (and Bob) are successfully minimized, which will be shown in the following section, where we evaluate their distance correlation (d_{Cor}).

E. Distance correlation

d_{Cor} measures both linear and non-linear statistical dependencies between random variables [15]. The value of d_{Cor} varies between 0 and 1, with 0 denoting fully independent variables and 1 fully dependent variables. d_{Cor} successfully captures statistical dependencies, as opposed to linear measures such as Pearson correlation [15]. Its empirical form is given as:

$$d_{\text{Cor}} = \begin{cases} \frac{{d_{\text{Cov}}}^2(X,Y)}{\sqrt{d_{\text{Var}}(X) d_{\text{Var}}(Y)}}, & \text{if } d_{\text{Var}}(X) d_{\text{Var}}(Y) > 0 \\ 0, & \text{if } d_{\text{Var}}(X) d_{\text{Var}}(Y) = 0, \end{cases} \tag{7}$$

where the definition of distance covariance, d_{Cov}, and distance variance d_{Var} can be found in [15].

To show the advantage of using both transformation techniques described in the previous section, we evaluate the distance correlation for the power measurements at Alice when $N = 5$ filters. This is illustrated in Figure 3, which shows the distance correlation of power measurements taken at Alice's filterbank output before and after transformation. Before transformation we see that there is strong dependence between the filters and the value of d_{Cor} reaches 0.4. Next, the figure shows the dependence between filters when we apply only the KLT. As it successfully removes correlation between measurements, the KLT has been previously used in different SKG methods [5]. However, as seen here, even if the correlation between filters is removed, the dependence remains strong. Finally, we evaluate d_{Cor} when using both quantile transform and KLT. As seen in the figure, by performing both transforms, in a sequential manner, the dependence between filters' outputs is reduced to values below 0.1, which can be considered as negligible [15].

F. Mutual information

Next, we look at the achievable SKG rate. The upper bound on the number of secret bits that can be extracted between Alice and Bob has been derived in [16] as $I\left(\hat{\mathbf{P}}_A; \hat{\mathbf{P}}_B\right)$, with I denoting MI, and is calculated by:

$$\int_{\hat{\mathcal{P}}_A} \int_{\hat{\mathcal{P}}_B} p(\hat{\mathbf{P}}_A, \hat{\mathbf{P}}_B) \log \frac{p(\hat{\mathbf{P}}_A, \hat{\mathbf{P}}_B)}{p(\hat{\mathbf{P}}_A) p(\hat{\mathbf{P}}_B)} d\hat{\mathbf{P}}_A d\hat{\mathbf{P}}_B, \tag{8}$$

where $p(\hat{\mathbf{P}}_A), p(\hat{\mathbf{P}}_B)$ denote the marginal probability density functions (PDFs) of $\hat{\mathbf{P}}_A$ and $\hat{\mathbf{P}}_B$, respectively, and $p(\hat{\mathbf{P}}_A, \hat{\mathbf{P}}_B)$ is their joint PDF. Unfortunately, a closed-form solution for the PDFs is hard to find and, even if they are known, solving the integral in (8) analytically might not be possible.

To overcome the problem above and to estimate the MI between Alice's and Bob's measurements we use a numerical MI estimator from the non-parametric entropy estimation toolbox

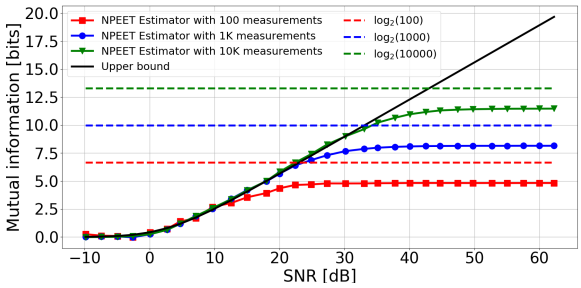

Fig. 4: Mutual information estimation using different number of samples.

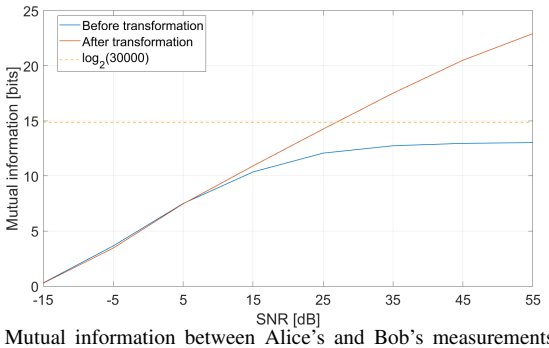

Fig. 5: Mutual information between Alice's and Bob's measurements before and after transformation for $N = 2$ and the number of samples is 30000.

(NPEET) [17]. An advantage of the NPEET estimator is that it can be used to estimate the MI between $\hat{\mathbf{P}}_A$ and $\hat{\mathbf{P}}_B$ without any assumptions on their PDFs. However, a disadvantage is that the reliability of its estimate is strongly dependent on the number of samples, N_{Samples}. It was proven that its estimate is upper bounded by $\log_2(N_{\text{Samples}})$, if considering MI in bits [9]. Due to this limitation the estimator is unable to provide a good estimate for high MI values. An example is illustrated in Figure 4. The figure illustrates a simple scenario with known upper bound on the SKG rate. In this case, we consider a one-dimensional Rayleigh-faded AWGN channel. The theoretical limit for this scenario is [11], $-\log_2\left[1-\left(\gamma/(1+\gamma)\right)^2\right]$, with γ the SNR. The figure compares the upper bound with the NPEET estimation for different N_{Samples}. It shows that the estimation is bounded by both the theoretical limit and $\log_2(N_{\text{Samples}})$ [9], hence, reliable estimation of high MI values requires big number of samples.

As discussed in Section I, increasing the number of samples might help to find the MI for one-dimensional measurements. For higher dimensions, as in our filterbank approach, the MI is expected to reach values whose estimation would require an impractical number of samples. Fortunately, as the proposed transformation technique minimizes dependencies and correlation among measurements from different filters, we can treat them as independent one-dimensional variables. Then, for a single frame $j = 1, \ldots, J$, the MI can be approximated by:

$$I\left(\ddot{\mathbf{p}}_{A_j}; \ddot{\mathbf{p}}_{B_j}\right) \approx \sum_{n=1}^{N} I\left(\ddot{p}_{A_{j,n}}; \ddot{p}_{B_{j,n}}\right), \tag{9}$$

where $\ddot{\mathbf{p}}_{A_j} = \left[\ddot{p}_{A_{j,1}}, \ldots, \ddot{p}_{A_{j,N}}\right]$ and $\ddot{\mathbf{p}}_B = \left[\ddot{p}_{B_{j,1}}, \ldots, \ddot{p}_{B_{j,N}}\right]$ are the transformed power measurements taken over N filters at Alice and Bob, respectively. Based on that, in the next section we evaluate the achievable SKG rates for different parameters.

III. Numerical Evaluation

In this section, we present a numerical evaluation for the achievable SKG rates between Alice and Bob for different channel and system parameters. The radio channel used for the evaluation is the 3GPP TDL-A model [10]. This model assumes non-line-of-sight (NLOS) communication with 23 MPCs. Here we consider 10 ns delay spread. Large-scale fading phenomena, such as path-loss and shadowing, are not included. The transmit signal, $x(t)$, is a 500 MHz complex chirp with constant modulation. Finally, the SKG rate is evaluated through MI estimates between Alice's and Bob's power measurements.

First, (9) is validated through a simple example, where $N = 2$. This is shown in Figure 5, where we measure the MI using 30 000 power measurements. The limitation of the estimator, i.e., the numerical estimate is bounded by $\log_2(N_{\text{Samples}})$ is clearly visible. First, we evaluate the MI between the measurements before transformation. While the estimate rises linearly for low SNR, it saturates when gets close to $\log_2(30\,000) \approx 14.87$ and remains almost constant. Next, we estimate the MI as in the right-hand side of (9). Instead of measuring the total MI at once, we can calculate the the MI between Alice and Bob for each feature of the KLT output. The final result is obtained by summing all values. This is illustrated by the red curve in Figure 5. The curve shows that if we distribute the MI among several independent features, and measure for each one separately, the limitation of the MI estimator can be overcome, hence see the expected linear MI growth with the increase of the SNR.

Next, we look at the achievable SKG rates when considering different number of filters. This is given in Figure 6, which compares the achievable SKG rates for different SNR values. The MI is lowest when Alice and Bob apply only a single filter (this scenario is equivalent to RSS-based SKG methods), i.e., a single power measurement cannot capture all the information contained in the channel. Next, we see that depending on the SNR different number of secret bits can be extracted. Focusing on SNR, $\gamma = 10$ dB, it can be observed that every 10 new filters added to the system bring less and less information. The reason behind this is two-fold: i) the total information that can be harvested from the considered channel is limited by bandwidth, SNR, MPCs, etc.; ii) more filters result in smaller bandwidth for each one and therefore, each filter will contribute less information. For this particular scenario, it can be seen that already at 60 filters the total information contained in the channel is extracted. Hence, increasing the number of filters above this value might not be beneficial, as it increases the complexity of system and bring only negligible amount of information. Next, looking at $\gamma = -10$ dB it is observed that the total information that can be extracted from the channel drastically decreases. Due to that, already at $N = 10$ Alice and Bob can capture most of the available information. We note that the range for the SNR was chosen to account for the considered scenario, i.e., low-power wideband devices [13].

Finally, in Figure 7, we look at the achievable SKG rates

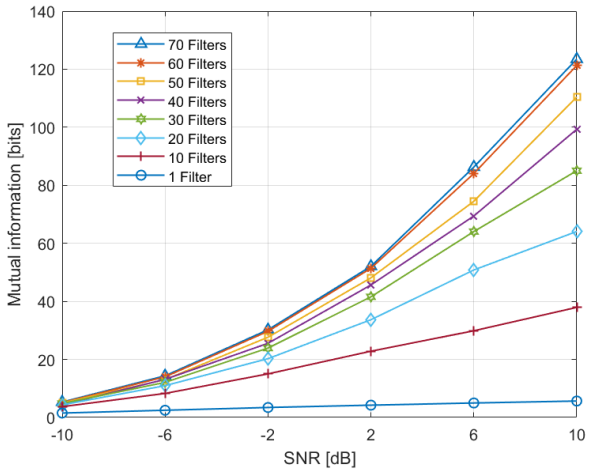

Fig. 6: Mutual information between Alice's and Bob's measurements for different number of filters and SNR values. The delay spread is 10 ns and the number of samples is 30000.

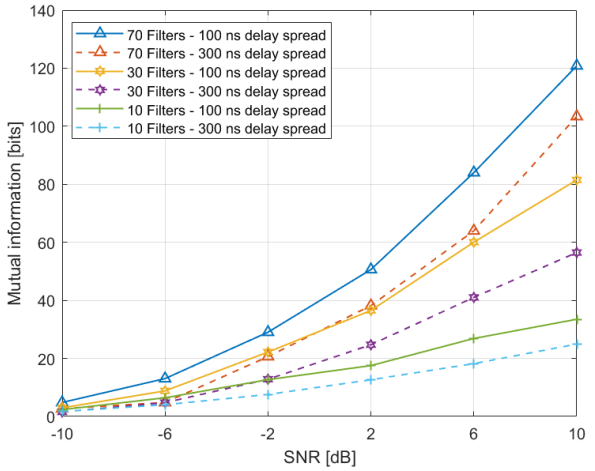

Fig. 7: Mutual information between Alice and Bob for different number of filters, SNR and delay spread. The number of samples is 30000.

considering the delay spread, which determines how many MPCs are resolved at the receiver. A higher delay spread allows more MPCs to be resolved. Given our earlier findings in [7], the results are expected. In [7], it was seen that resolving more MPCs decreases the randomness of their power measurements at each filter (see (5)). As shown here, this is also the case for the filterbank approach, i.e., a lower delay spread increases the randomness of the measurements, leading to higher MI values. Overall, the results in this section show that no single solution would be optimal for all cases, and, therefore, bringing SKG to practical systems would require nodes to be channel-aware.

IV. Conclusions

In this work, we evaluated the achievable SKG rates for a filterbank approach considering wideband communication and a 3GPP TDL-A channel model. First, we proposed a transformation technique that successfully minimizes the dependency and correlation between different frequencies. The transformation provides the following benefits: i) it allows for secure key generation, ii) it converts the complex multidimensional system to several tractable one dimensional systems which allows us to overcome the fundamental limitation of MI estimation. Our evaluation shows that the information available for SKG varies greatly with the channel characteristics, hence, system parameters must be chosen accordingly. As a future work, the authors plan to further improve the system model by considering multiple antenna scenarios and work on adaptive solution through investigation of the full SKG generation protocol.

Acknowledgement

This work is financed by the Saxon State government out of the State budget approved by the Saxon State Parliament.

References

[1] J. Zhang, G. Li, A. Marshall, A. Hu, and L. Hanzo, "A new frontier for IoT security emerging from three decades of key generation relying on wireless channels," *IEEE Access*, vol. 8, 2020.

[2] M. Mosca, "Cybersecurity in an Era with Quantum Computers: Will We Be Ready?" *IEEE Security & Privacy*, vol. 16, no. 5, Sep. 2018.

[3] M. Latvaaho and K. Leppänen, "Key drivers and research challenges for 6G ubiquitous wireless intelligence," Oct 2019, university of Oulu.

[4] G. Li, A. Hu, J. Zhang, L. Peng, C. Sun, and D. Cao, "High-agreement uncorrelated secret key generation based on principal component analysis preprocessing," *IEEE Trans. on Comm.*, vol. 66, no. 7, 2018.

[5] G. Li, H. Yang, J. Zhang, and H. Liu, "Fast and secure key generation with channel obfuscation in slowly varying environments," 2021.

[6] M. Zoli, A. N. Barreto, S. Köpsell, P. Sen, and G. Fettweis, "Physical-layer-security box: a concept for time-frequency channel-reciprocity key generation," *Eurasip J. Wirel. Commun. Netw.*, 2020.

[7] M. Mitev, A. N. Barreto, T. M. Pham, and G. Fettweis, "Secret key generation rates over frequency selective channels," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–5.

[8] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E*, vol. 69, p. 066138, Jun 2004.

[9] D. McAllester and K. Stratos, "Formal limitations on the measurement of mutual information," in *Proc. 23 Int. Conf. on AI and Stat.*, vol. 108. PMLR, 26–28 Aug 2020, pp. 875–884.

[10] 3GPP, "Release 16, TR 38.901, Study on channel model for frequencies from 0.5 to 100 GHz."

[11] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. B. Mandayam, "Information-theoretically secret key generation for fading wireless channels," *IEEE Trans. Inf. Forensics Security*, vol. 5, 2010.

[12] M. Mitev, A. Chorti, M. Reed, and L. Musavian, "Authenticated secret key generation in delay-constrained wireless systems," *Eurasip J. Wirel. Commun. Netw.*, 2020.

[13] G. Llano, J. C. Cuellar, and A. Navarro, "Frequency UWB channel," in *Ultra Wideband Communications*. IntechOpen, 2011, ch. 4.

[14] S. Hazarika, A. Biswas, and H.-W. Shen, "Uncertainty visualization using copula-based analysis in mixed distribution models," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, 2018.

[15] R. Wang, A.-H. Karimi, and A. Ghodsi, "Distance correlation autoencoder," in *2018 IJCNN*, 2018, pp. 1–8.

[16] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[17] G. V. Steeg, "Non-parametric entropy estimation toolbox." [Online]. Available: https://github.com/gregversteeg/NPEET