Secret Key Generation Rates over Frequency Selective Channels

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

***Abstract*—The emergence of Internet of things (IoT) applications brings the challenge of finding lightweight schemes for secret key distribution. A promising solution coming from the physical layer is the so called secret key generation (SKG) from shared randomness. SKG allows two communicating parties to extract the shared randomness already present in wireless channels. This work investigates the achievable SKG rates using received signal strength (RSS) as a key generation parameter. A multi-path scenario is considered and the probability density function of the RSS in different channel conditions is evaluated. Next, through a numerical evaluation, the SKG rates are derived in the form of mutual information (MI) estimates, using a 3GPP standard channel model. The simulations are performed for different values of bandwidth and delay spread. We demonstrate that while both parameters have similar impact on how multi-path components are resolved, their effect on the MI is opposite.**

***Index Terms*—IoT, multi-path, mutual information, physical layer security, secret key generation.**

I. Introduction

A major limitation of 5G networks, especially in the context Internet of things (IoT), is related to security [1]. Standard cryptographic solutions rely on computationally intensive modulo arithmetic operations which makes them unsuitable for power constraint devices. This issue has been addressed in the report on approved lightweight cryptographic primitives from the national institute of standards and technology (NIST), where well-known public key encryption (PKE) algorithms, such as Diffie-Hellman and Rivest-Shamir-Adleman (RSA), are excluded. The report includes mainly symmetric key block ciphers, such as the advanced encryption standard (AES), which achieve quantum supremacy with appropriate key lengths for IoT scenarios [2]. In this sense, lightweight alternatives to PKE for secret key distribution are sought.

A viable solution, considered for deployment in the sixth generation (6G) mobile networks [3], is provided by the physical layer security (PLS) paradigm. The PLS-based secret key generation (SKG), first proposed in [4] and [5], allows two wireless nodes to extract shared randomness leveraging the reciprocity of wireless channels, within the coherence time of the channel. A major advantage of the SKG process is that it can be employed using any PHY waveform (as long as the transmitted power spectral density is known) and it does not require matched receivers [6], [7]. However, the key generation rates depend strongly on the stochastic behaviour of the wireless medium. While other studies suggest that the achievable rates of received signal strength (RSS) based SKG are independent from the number of multi-path components (MPCs) [8], here we demonstrate that this might not be true in general.

Motivated by the above, in this work we evaluate the achievable RSS-based SKG rates over frequency selective channels. Our system model consists of two legitimate parties, namely Alice and Bob, and an eavesdropper, Eve. To generate a shared random bit sequence, Alice and Bob simply use their observations of the measured power, i.e., entropy is extracted without additional computational overhead. The dependence of the SKG rates with respect to (w.r.t.) different channel and system parameters, such as delay spread and signal bandwidth is evaluated. The performance of the key generation is verified through numerical evaluation using 3GPP-based tapped delay line (TDL) channel model [9].

The focus of this work is to evaluate the SKG rates using the RSS parameter. First, the RSS distribution is evaluated, showing its dependency on the receivers' bandwidth. Next, through numerical analysis we evaluate the number of key bits that can be extracted using different channel and system parameters. Overall, the analysis shows that no single solution can be applied for all cases. The rest of the paper is organized as follows: Section II introduces the SKG concept, Section III presents the system model and gives the derivation of the distribution of the measured power at Alice and Bob, Section IV shows our numerical analysis for the achievable SKG rates, and, Section V concludes this paper.

II. Secret key generation

The standard SKG process consists of three phases:
1) Advantage distillation: In this phase Alice and Bob sequentially exchange probe signals, with constant power [10], and obtain estimates of their reciprocal channel state information (CSI). Due to the presence of noise and imperfect CSI estimation, the observations at both sides, will differ. Eve, a third party who acts as an eavesdropper, could also obtain channel estimates during this phase. CSI can be obtained with different granularity levels, however, due to its availability in off-the-shelf chipsets, typical parameter is the RSS.
2) Information reconciliation: At the beginning of this phase, Alice and Bob quantize their observations to binary vectors. Next, errors due to imperfect CSI estimation are corrected through a public exchange of reconciliation information. Numerous information reconciliation algorithms have been proposed in the literature [11].

3) Privacy amplification: To generate a secret key, the reconciled information at each of the legitimate users is compressed using a one-way collision-resistant function, e.g., hash function. This last phase ensures that the generated key sequence is uniformly distributed and unpredictable by an adversary [12].

As noted above, during the "advantage distillation" phase Alice and Bob extract entropy from the wireless channel, i.e., the nature and the stochastic properties of the link between them strongly impact the SKG rates. In this sense, "channel awareness" is a vital ingredient for the success of the SKG process. It would allow the adaptation of SKG parameters and provide upper layers with information on the available key bits. Hence, it is important to understand: *How the channel properties affect the number of available secret bits and how many can be extracted in different channel conditions?*

To answer the above question, we focus on the estimation of the achievable SKG rates during the "advantage distillation" phase. The investigation of the second and third phases of the process falls out of the scope of this paper and will be considered as a future work.

III. System Model

A. Channel model

In this work, we consider a multi-path channel model with additive white Gaussian noise (AWGN). The analyses assumes complex baseband signal and a low-pass filter applied by each party. Based on that, the observations at Alice and Bob are:

$$y_A(t) = f(t) * [x(t) * h(t) + w_A(t)]\,, \tag{1}$$

$$y_B(t) = f(t) * [x(t) * h(t) + w_B(t)]\,, \tag{2}$$

where $f(t)$ is the filter impulse response, $x(t)$ represents the transmitted signal, $h(t)$ denotes the channel response and $w_A(t) \sim \mathcal{N}(0, \sigma_A^2)$, $w_B(t) \sim \mathcal{N}(0, \sigma_B^2)$ are the independent noise variables at Alice and Bob, respectively. The channel impulse response (CIR) is defined using the 3GPP TDL channel model [9]. The TDL model has 5 different modes, (from A to E), all with the same root mean square delay spread of 1 ns, but with different power delay profiles. One advantage of this model is that, it allows the delays to be scaled to obtain different delay spreads, i.e., by making $\tau_i = D\tau_{i,\text{TDL}}$, where $\tau_{i,\text{TDL}}$ is the propagation delay for the i-th MPC, as defined in [9] and D is the desired delay spread in ns. Based on that the CIR is modelled in the time domain as:

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

where L is the number of MPCs, $\delta(t)$ is the Dirac delta function and α_i is the complex amplitude for the i-th component, respectively. Furthermore, as noted in [9], α_i, $i = 0, \ldots, L-1$ are circularly symmetric complex Gaussian variables. Their magnitudes $|\alpha_i|$ could follow either Rayleigh or Rician distribution. This paper focuses on the Rayleigh scenario, and assumes that MPCs are independently distributed.

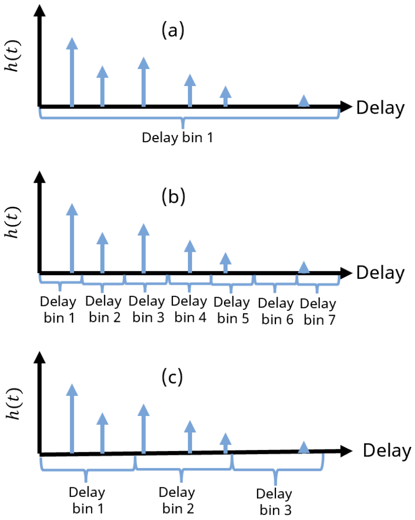


Fig. 1: Different cases for MPCs addition: a) all MPCs add non-coherently; b) all MPCs add coherently; c) general case.

B. Power distribution

As mentioned above, the source of common randomness between Alice and Bob is the RSS, which can be expressed as the average of the squared magnitudes of (1), (2) over a frame interval, i.e., $\hat{P}_A = \overline{|y_A|^2}$ and $\hat{P}_B = \overline{|y_B|^2}$, where $\overline{(\cdot)}$ and $\hat{(\cdot)}$ denote the time average and estimate of a given value, respectively. In this work, we assume that $f(t)$, the filter applied at both receivers, is an ideal low-pass filter with unity gain and cut-off frequency $B/2$. The transmitted signal is assumed to be a complex chirp with constant modulation. The chirp waveform has a nearly-flat and well-contained power spectral density which allows us to better illustrate the changes in the power distribution, i.e., the distribution will depend only on the magnitude variation of the channel in the frequency domain and on the noise within the considered bandwidth. Next, due to the independence between channel and noise realizations we first treat both processes independently and then we consider their combined effect.

1) Channel power distribution: As noted earlier, it is assumed that the magnitude of the MPCs are independent Rayleigh distributed variables. Depending on the considered bandwidth B and delay spread D, three cases can be differentiated: i) all MPCs are non-resolvable and they add non-coherently (i.e., complex sum); ii) all MPCs are resolvable and they add coherently (i.e., power sum); iii) a general case where we have a mixture of non-coherent and coherent addition, which depends on the number of MPCs that fall in each delay bin. The three cases are illustrated in Fig. 1. The width of the delay bins T_s is proportional to the filter's bandwidth at the receiver and is given as $T_s = 1/f_s$, where the sampling frequency is given as $f_s = B$. The two extremes, illustrated in Fig. 1a) and b), represent cases i) and ii) where B is small and large, respectively. The channel distribution in these two cases is known; if all L MPCs add up non-coherently, the resulting distribution is exponential; if all MPCs add up coherently the resulting power sum has non-standard chi-square with degrees of freedom equal to $\kappa = 2L$ [8]. It is important to note that both exponential and chi-square distributions are particular cases of the gamma distribution and are related as follows: $\sum_{i=0}^{L} \mathrm{Exp}(v) \sim \Gamma(L, v)$ where v is scale parameter, and $\chi^2_\kappa \sim \Gamma(\frac{\kappa}{2}, 2)$, considering gamma distribution with shape and scale parameterization. Therefore, in the case where different number of MPCs may fall in each delay bin, Fig. 1c, the distribution of the squared magnitude of the channel can be represented as a sum of gamma-distributed random variables:

$$|h(t)|^2 = \left|\sum_{i_1=0}^{L_1-1} \alpha_{i_1}\right|^2 + \left|\sum_{i_2=1}^{L_2-1} \alpha_{i_2}\right|^2 \cdots + \left|\sum_{i_M=1}^{L_M-1} \alpha_{i_M}\right|^2, \quad (4)$$

where the indices $i_1, \ldots, i_M$, correspond to a particular delay bin, i.e., $M = \lceil \tau_{\text{max}}/T_s \rceil$, where $\lceil \cdot \rceil$ is the ceiling operator and τ_{max} the maximum delay spread. As noted earlier, the distribution of (4) can be represented as a sum of gamma-distributed random variables, where each of the inner sums follows $\Gamma(k_{i_m}, \theta_{i_m})$, $m = 1, \ldots, M$. To evaluate the resulting probability density function (PDF) we use the Welch–Satterthwaite approximation [13], which states that the sum of gamma variables is also a gamma variable and (4) follows a gamma distribution with parameters as:

$$|h(t)|^2 \sim \Gamma\left(k = \frac{(\sum_{j=1}^{M} k_j \theta_j)^2}{\sum_{j=1}^{M} k_j \theta_j^2}, \theta = \frac{\sum_{j=1}^{M} k_j \theta_j}{k}\right). \quad (5)$$

2) Noise distribution: Similarly to the channel, the distribution of the measured noise power is strongly dependent on the number of samples, which depends on the frame length and bandwidth. For Alice, it is given as $\hat{P}_N = ||\mathbf{w}_A||^2/N$, where $\mathbf{w}_A$ is a the measurement noise vector composed of N elements and $|| \cdot ||$ is the norm operator. Note that the analysis for Bob is identical and is, therefore, omitted here. It can be seen that the distribution of the measured noise power varies with the number of noise samples within the considered bandwidth, i.e., as the legitimate users sample at the Nyquist rate a wider bandwidth will correspond to a higher number of noise samples. We note that the distribution of $||\mathbf{w}_A||^2$ can be derived as $\chi^2_2 s^2$, where s^2 denotes the sample variance. To derive the distribution of the sample variance, we use Cochran's theorem [14], which states that the quantity $\frac{(N-1)s^2}{\sigma_A^2}$ follows χ^2_{N-1} distribution. Based on that, it can be concluded that the sample variance is distributed as: $s^2 \sim \frac{\chi^2_{N-1}\sigma_A^2}{N-1} \sim \Gamma(\frac{N-1}{2}, \frac{2\sigma_A^2}{N-1})$. Finally, we can estimate the distribution of $\hat{P}_N$ as follows:

$$\hat{P}_N = \frac{||\mathbf{w}_A||^2}{N} \sim \Gamma\left(\frac{N-1}{2}, \frac{4\sigma_A^2}{N-1}\right). \quad (6)$$

It is important to note that the variance of the measured noise power is inversely proportional to the number of samples:

$$\mathrm{Var}\left(\hat{P}_N\right) = \frac{N-1}{2}\left(\frac{4\sigma_A^2}{N-1}\right)^2 = \frac{8\sigma_A^4}{N-1}. \quad (7)$$

3) Received power distribution: Next, the distribution of the measured power is evaluated. As noted earlier, the distributions of P_A and P_B depend on the channel and noise realizations, on the measurement period (i.e., number of observed samples) and on the considered bandwidth. Due to the fact that the three variables $h(t), w_A(t), w_B(t)$ are independent and zero mean and that noise always adds non-coherently, the measured power at Alice and Bob can be represented as $|h(t)|^2 + |w_A|^2$ and $|h(t)|^2 + |w_B|^2$, respectively, (this can be applied only if multiple samples are considered, i.e., the measured power is obtained through averaging). Using the results in (5) and (6), once again, we can apply the Welch–Satterthwaite approximation to obtain:

$$P_A \sim \Gamma\left(k = \frac{(\sum_{j=1}^{M} k_j \theta_j + 2\sigma_A^2)^2}{\sum_{j=1}^{M} k_j \theta_j^2 + \frac{8\sigma_A^4}{N-1}}, \theta = \frac{\sum_{j=1}^{M} k_j \theta_j + 2\sigma_A^2}{k}\right), \quad (8)$$

The distribution for P_B is derived similarly. The findings above are demonstrated through numerical simulations in Fig. 2. The simulations assume a fixed number of MPCs, $L = 30$, all with equal power, and received SNR approximately equal to 10 dB. Each MPC is assumed to be a circularly symmetric Gaussian random variable. All figures assume the same channel realizations, however, depending on the considered bandwidth, different number of MPCs will be resolved, i.e., different values and distributions are observed for the measured power. For each scenario the noise variance is set in accordance with the SNR.

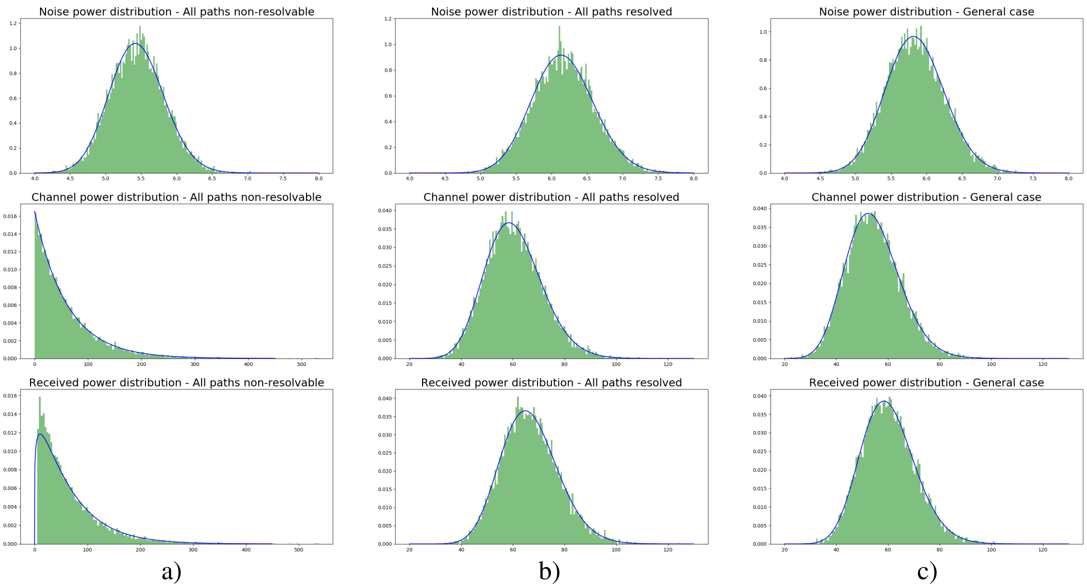
C. Theoretical limits

The upper bound on the SKG rate has been derived in [4] as $I(P_A; P_B|P_E)$, where I denotes mutual information (MI) and P_E denotes the observations at Eve. In this work, we assume that Eve is few wavelengths away from both Alice and Bob. Therefore, due to the decorelative properties of the wireless channel, in rich multi-path fading scenarios, it can be assumed that Eve's observations are independent from the observations at Alice and Bob. Hence, the upper bound for the SKG rate reduces to $I(P_A; P_B)$, which is calculated as:

$$\int_{\mathcal{P}_A}\int_{\mathcal{P}_B} p(P_A, P_B) \log \frac{p(P_A, P_B)}{p(P_A)p(P_B)} dP_A dP_B, \quad (9)$$

where $p(P_A, P_B)$ is the joint PDF of P_A and P_B and $p(P_A), p(P_B)$ are the marginal PDFs of P_A and P_B, respectively. As showed in the previous section, $p(P_A), p(P_B)$ are gamma PDFs with parameters as in (8). Regarding the joint PDF $p(P_A, P_B)$, [15] demonstrated that if two gamma variables have the same scale or shape parameter their joint PDF is also gamma. These parameters could be equal only if the noise observations at Alice and Bob are the same. For the general case, a closed-form solution for $p(P_A, P_B)$ is not yet found. However, even for the case when $p(P_A, P_B)$ is known, the integral will be hard to solve and its calculation is left for future work.

Given the above, to estimate the MI between Alice's and Bob's measurements we use two numerical MI estimators,



a) b) c)

Fig. 2: Measurement noise, channel and signal powers (green histogram) and the derived theoretical PDFs (blue curves) for three cases: a) all MPCs fall into the same delay bin; b) all MPCs fall into different delay bins; c) different number of MPCs fall into each delay bin. The parameters for all figures are: total number of MPCs $L = 30$, received SNR$\approx$ 10dB

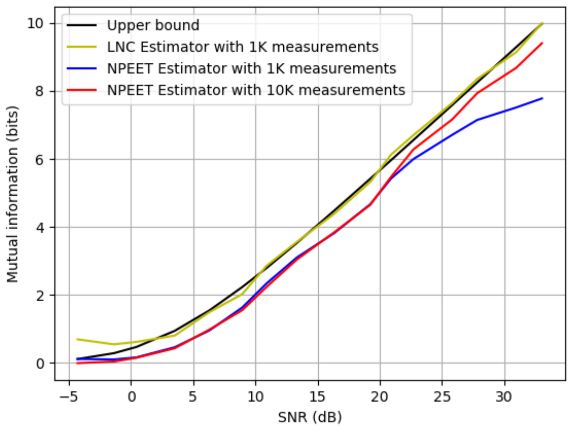


Fig. 3: Comparison of LNC and NPEET estimators against the MI upper bound given in Eq. (10) for Rayleigh fading channels.

non-parametric entropy estimation toolbox (NPEET) [16] and mutual information estimation with local non-uniformity correction (LNC) [17]. The estimators allow the estimation of the MI between P_A and P_B without knowledge of their joint PDF. First, the performance of both estimators is tested in a simple scenario, Rayleigh-faded AWGN channel for narrow-band communication systems (illustrated in Figs. 1a, 2a). The theoretical upper bound for this scenario is [11]:

$$I\left(P_A; P_B\right) = -\log_2\left[1 - \left(\frac{\text{SNR}}{1+\text{SNR}}\right)^2\right], \quad \frac{\text{bits}}{\text{sample}}. \tag{10}$$

A comparison of the estimated MI with the upper bound in (10) is given in Fig. 3. As it is observed, the NPEET estimator requires more samples to converge to the true value. This is mostly noticeable for highly correlated measurements, i.e., at high SNR. However, even at low SNR the LNC estimator shows better performance. Based on this result, in the rest of this paper only the LNC estimator is used.

IV. Numerical results

This section presents a numerical evaluation of the MI between P_A and P_B for different system and channel parameters. The considered radio channel is the 3GPP TDL-A model [9]. The TDL-A model assumes non-line-of-sight communication with a fixed number of MPCs, $L = 23$. Large-scale fading (i.e., path-loss and shadowing) is not considered. The communication waveform $x(t)$ is a complex chirp signal with constant modulation and bandwidth of 500 MHz. The power observation at Alice P_A is obtained by convolving $x(t)$ with the TDL-A CIR $h(t)$ and adding AWGN w_A, then passing the resulting signal through a low-pass filter $f(t)$ with cut-off frequency $B/2$, and, finally, the power is measured through averaging the square magnitude of the filter's output.

First, we evaluate the MI w.r.t. different values of the SNR and B, while $D = 100$ ns. The resulting estimates of the MI are illustrated in Fig. 4. Recalling that by increasing B more and more MPCs are getting resolved, this leads to a change in the distribution of the measured power, as shown in Fig. 2. As observed in Fig. 2, resolving more MPCs leads to a decrease in the variance of the observations, which has a negative effect on the MI. However, the opposite is observed here. This is a consequence of (7). Increasing B increases the number of noise samples N during the measurement period and the noise variance decreases. Effectively, this allows Alice

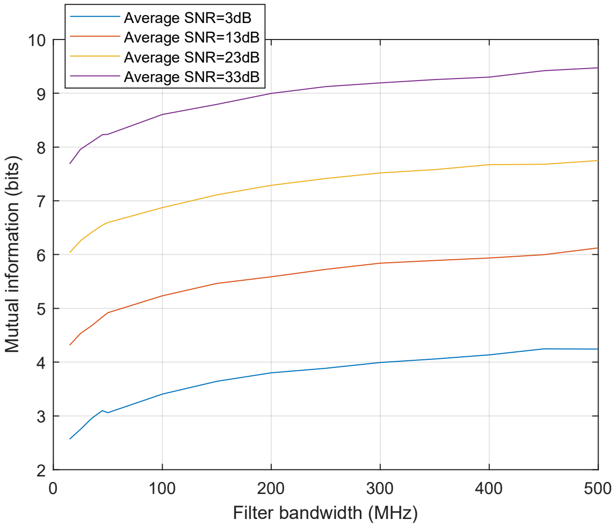


Fig. 4: MI estimation for different B and SNR for fixed $D = 100$ ns.

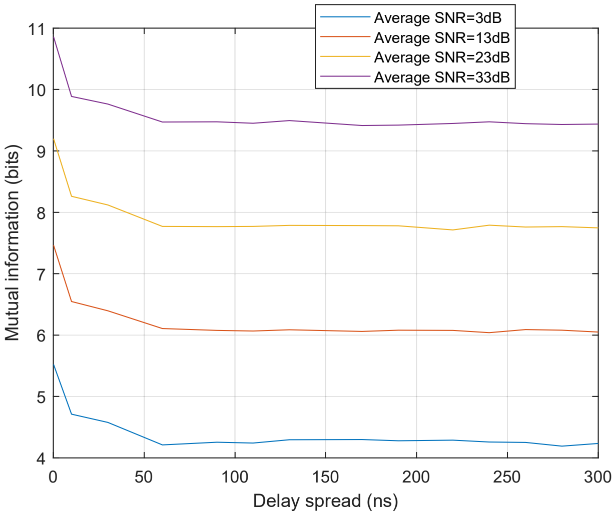


Fig. 5: MI estimation for different SNR and D, and fixed $B = 500$ MHz.

and Bob to control the measured SNR, which as illustrated in Fig. 4, directly impacts the MI.

Next, we compute the MI w.r.t. different values for SNR and delay spread, D, while the filter bandwidth is fixed to $B = 500$ MHz. The results are given in Fig 5. In contrast to Fig. 4, the effect of resolving different number of MPCs is clearly visible here. For small values of D, when all MPCs fall into the same delay bin, the MI takes its highest value. Next, increasing D increases the number of resolvable MPCs and, as discussed earlier, decreases the randomness. From Fig. 5, we see that shortly after $D = 50$ ns all MPCs are resolved and the MI remains constant. This is expected, as the number of MPCs in the 3GPP TDL-A model is fixed to $N = 23$, i.e., once all are resolved the sum in (4) will have 23 elements and will remain constant regardless of D. On the other hand, the measured noise power given in (6) is independent from the delay spread and does not changes w.r.t. D.

Overall, we see that the MI between Alice and Bob strongly depends on the scenario. All variables, such as SNR, number and power of MPCs, delay spread and bandwidth affect the MI. This leads to a single conclusion: *In order to successfully incorporate SKG in practice, devices need to be channel-aware and able to adaptively adjust their parameters.*

V. Conclusions

In this work, we provided an estimation of the achievable SKG rates using RSS in a 3GPP TDL-A channel model. First, the PDFs of the observations at Alice and Bob were derived. It was shown that by adjusting the bandwidth Alice and Bob can manipulate the distribution of both channel and noise power, which directly impacts the MI. Furthermore, while delay spread affects the number of resolvable MPCs, its increment after a certain value will not have impact on the MI. Based on these results, as a future work, the authors intend to improve the system model by adding more degrees of freedom, such as multiple antennas and multiple filters (in the form of a filterbank) and investigate possible adaptive solutions to harvest the channel entropy.

References

[1] "3GPP TR 33.825 V0.3.0, Study on the Security for 5G URLLC (Release 16)," 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects.

[2] K. McKay, L. Bassham, M. Turan, and N. Mouha, "Report on lightweight cryptography," *NIST Interagency/Internal Report (NISTIR) - 8114*, Mar 2017.

[3] M. Latvaaho and K. Leppänen, "Key drivers and research challenges for 6G ubiquitous wireless intelligence," October 2019, published online by the University of Oulu.

[4] 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.

[5] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.

[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 Journal on Wireless Communications and Networking*, 2020.

[7] M. Zoli, M. Mitev, A. N. Barreto, and G. Fettweis, "Estimation of the secret key rate in wideband wireless physical-layer-security," in *ISWCS*, 2021, pp. 1–6.

[8] Y. Liu, S. C. Draper, and A. M. Sayeed, "Exploiting channel diversity in secret key generation from multipath fading randomness," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1484–1497, 2012.

[9] 3GPP, "Release 16, TR 38.901, Study on channel model for frequencies from 0.5 to 100 GHz."

[10] M. Mitev, A. Chorti, E. V. Belmega, and H. V. Poor, "Protecting physical layer secret key generation from active attacks," *Entropy*, vol. 23, 2021.

[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 Journal on Wireless Communications and Networking*, 2020.

[13] N. W. F. Spellman, *Handbook of Mathematics and Statistics for the Environment*. CRC Press, 2013.

[14] W. G. Cochran, "The distribution of quadratic forms in a normal system, with applications to the analysis of covariance," *Math. Proc. of the Cambridge Philosophical Society*, vol. 30, no. 2, 1934.

[15] S. Nadarajah and A. K. Gupta, "Some bivariate gamma distributions," *Applied Mathematics Letters*, vol. 19, no. 8, pp. 767–774, 2006.

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

[17] S. Gao, G. V. Steeg, and A. Galstyan, "Efficient estimation of mutual information for strongly dependent variables." [Online]. Available: https://github.com/BiuBiuBiLL/NPEET_LNC