A Low-Complexity Deep Neural Network for Signal-to-Interference-Plus-Noise Ratio Estimation

Roberto M. M. Kagami [1], **Luciano L. Mendes** [1]

[1]Instituto Nacional de Telecomunicações (Inatel)
CEP 37540-000 - Santa Rita do Sapucaí - MG - Brazil

Abstract. *Mobile network technology has been driven by a huge demand for throughput and reliability to support new emerging services. The quality of service is based on measurements of indicators with a high level of precision. Accurate controlling of parameters to fulfil the quality requirements will be essential for future applications. In LTE and 5G standards, the Channel Quality Indicator can be calculated using different algorithms. It is key to determine the best coding and modulation as well as the power control. Thus, it depends on the exact signal-to-noise ratio estimation. MSE based on hard-decision has a very low computational cost, however, it can insert non-linearities. This paper proposes a neural network to estimate an SINR from a modified MSE function.*

1. Introduction

The future wireless communications systems will provide an extensive variety of applications and use cases. Many of them have distinct requirements, which can lead to divergent decisions to be addressed accordingly to the same indicator. The Modulation Coding Scheme (MCS) index, for instance, must be lower than an error-free limit condition when the requirement is reliability. On the other hand, if the service demands throughput, the MCS index has to be as close as possible to the highest spectral efficiency point that still achieves the maximum target Bit Error Rate (BER). But in both cases, the indicator needs a very precise Signal-to-Interference-Plus-Noise Ratio (SINR) estimation.

In OFDM systems, this measurement can be performed using pilot carriers [Khan et al. 2017], cyclic prefix [Baumgartner et al. 2014] or preambles [Malik et al. 2011]. Basically, the benefit of such methods is the use of pre-known transmitted information to estimate the SINR on the receiver side. However, only small portions of the frame are taken into account. Statistically, the larger is the amount of data used for estimation, the better is the accuracy and predictability of results. The computation of all sub-carriers provides superior estimation about the data region. This approach can also consider inaccuracies in respect of interference and channel estimation processes. Thus, the adoption of classical Mean Square Error (MSE) can be a viable solution. However, this method has some drawbacks. To achieve an acceptable estimation of the SINR over real channel conditions, it is necessary to increase the performance of the estimator and Deep Neural Network (DNN) is an interesting approach to accomplish this goal.

The most recent mobile communication standards employ robust channel coding schemes, such as Polar Code, Turbo Code and Low Density Check Parity Code (LDPC), which allows the system to operate with very low SINR, including negative values in logarithmic scale [Bin Li et al. 2002]. When estimation is calculated using grid-based and hard decisions, another problem emerges. If a link condition worsens quickly, the current modulation order can lead to very high BER, which also means that the estimation based on Euclidean distance on the receiver side is poor. This paper proposes a modified MSE calculation to mitigate this inaccuracy. Based on that, it is possible to fit an even better solution using a small neural network. The remaining of this paper is organized as follows: Section 2 presents the system model used to generate data. Section 3 describes the

proposed modified MSE function. Section 4 presents nonlinear regression issues. Section 5 presents the proposed DNN-based estimator and Section 6 brings the final conclusions of this paper.

2. System Model

The mobile network system considered in this paper is based on the PHY layer description provided by 3GPP for the Fifth Generation of Mobile Networks (5G) in Release 15 [3GPP TS 38.211 version 15.2.0 Release 15 2018]. Using Orthogonal Frequency Division Multiplexing (OFDM) technique, the data Resource Blocks (RBs) are transmitted employing M-Quadrature Amplitude Modulation (QAM) with $M \in \{4, 16, 64, 256\}$.The data symbols are mapped into the OFDM subcarriers and the inverse Fast Fourier Transform (iFFT) is used to generate the time-domain OFDM block. Cyclic Prefix (CP) is inserted to protect the data from the time dispersive channels. On the receiver side, the CP is removed after the time-domain synchronization. The Fast Fourier Transform (FFT) is used to obtain the received symbols in the frequency-domain.

Since the estimator only employs the M-QAM symbols to evaluate the MSE, this estimator can also be employed in Single Carrier (SC) schemes. The setup will consider a system with perfect synchronization, carrier recovery and equalization in order to generate data and evaluate the function behaviour with pure Additive White Gaussian Noise (AWGN) and without interference. The system model considered in this paper assumes the M-QAM symbols in the frequency domain after the equalization, then the scheme depicted in Figure 1 can be properly used. The sweep control block selects the gain values of the AWGN source and provides the specified Signal-to-Noise Ratio (SNR) to the data set block (η_k). A random bits generator provides input data to the M-QAM block. This block, in turn, produces the symbols to be added to the noise. Subsequently, the MSE function to be evaluated process this signal, calculates the average and generates the output (λ_k) to the data set block. The number of symbols used to calculate the average value is specified further. The adjusted range is from -40dB to +40dB with steps of 0.1dB.
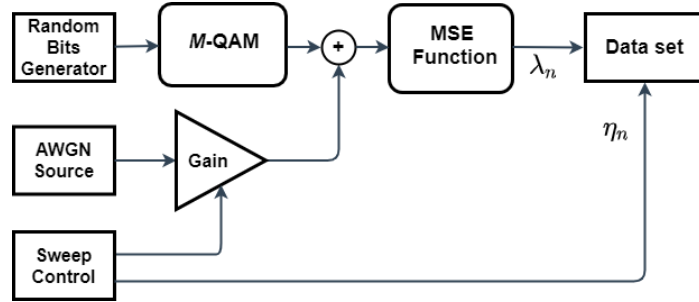

Figure 1. Data set generation diagram.

3. Modified MSE Function

This section presents the assumptions and adjustments used to improve a MSE calculation using a grid-based decision. By considering the k-th received symbol $\mathbf{x}_k$, correspondent to the transmitted symbol $\mathbf{s}_k$,

$$\|\mathbf{e}_k\|^2 = \|\mathbf{x}_k - \mathbf{s}_k\|^2, \tag{1}$$

where $\mathbf{s}_k$ is a symbol pertained to the sample space of a defined constellation. If the receiver applies a grid-based hard-decision for a M-QAM scheme to estimate the transmitted symbols, where $d_{\min}$ is the minimum distance between adjacent symbols, the absolute

values of coordinates of estimated transmitted symbols $\hat{\mathbf{s}}_k$ can be determined as

$$|\Re(\hat{\mathbf{s}}_k)| = d_{\min} \left\lfloor \frac{|\Re(\mathbf{x}_k)|}{d_{\min}} \right\rfloor + 1, \tag{2}$$

$$|\Im(\hat{\mathbf{s}}_k)| = d_{\min} \left\lfloor \frac{|\Im(\mathbf{x}_k)|}{d_{\min}} \right\rfloor + 1. \tag{3}$$

However, depending on the constellation, the maximum value of a $\|\hat{\mathbf{s}}\|$ coordinate is

$$S_{\max} = \frac{d_{\min}}{2}(\sqrt{M} - 1). \tag{4}$$

If an absolute value of a received symbol coordinate crosses the most external decision limit, i.e $S_{\max} + d_{\min}/2$, we can assume that the most probable absolute coordinate is $S_{\max}$. This occurrence is a detectable portion of all grid crossings. Moreover, it is the key event for the following considerations:

i) If there is no events, the calculated errors using (2) and (3) are correct.

ii) If the value $S_{\max}$ is attributed to the respective coordinate of $\|\hat{\mathbf{s}}\|$ when an event occurs, the error is closer to the correct value than those calculated using (2) or (3).

iii) On condition that the symbols are equiprobable and equidistant, the frequency of this detected event is statistically the same as undetected events for this coordinate and for each one of the internal symbols ($N/2$, where $N = \sqrt{M}$ and M is the modulation order of a square constellation).

By considering a first level event as the detected component that does not surpass $S_{\max} + d_{\min}$ and if $d_{\min}=2$, a similar undetected error δ_k inserts an underestimation calculated as

$$\xi_k = (1 + \delta_k) - (1 - \delta_k) = 2\delta_k. \tag{5}$$

A second level event is a detected component that exceeds $S_{\max} + d_{\min}$ but does not surpass the next hard decision limit. The undetected inserted error, in this case, is given by

$$\xi_m = (3 + \delta_m) - (1 - \delta_m) = 2(1 + \delta_m). \tag{6}$$

As the SNR decreases, higher levels of hard decision limits are overstepped and higher values of errors are inserted. Compensation factors could be used to adjust each level of outdistancing based on its statistical rate and error amplitude. But this estimation would comprise the computation of a complex arrangement.

Taking into account only the first level event and the mentioned considerations, a better function can be rewritten. If an α factor is used to represent the rate of undetected events and $\beta = 1 + \alpha$, the modified function for a k-th received symbol

$$\|\mathbf{e}_k\|^2 = a_k + b_k, \tag{7}$$

where

$$a_k = \begin{cases} (|\Re(x_k)| - |\Re(\hat{s}_k)|)^2, & \text{if } |\Re(x_k)| \le S_{\max} \\ \beta(|\Re(x_k)| - S_{\max})^2, & \text{otherwise,} \end{cases} \tag{8}$$

$$b_k = \begin{cases} (|\Im(x_k)| - |\Im(\hat{s}_k)|)^2, & \text{if } |\Im(x_k)| \le S_{\max} \\ \beta(|\Im(x_k)| - S_{\max})^2, & \text{otherwise.} \end{cases} \tag{9}$$

Eq. (7) can be used to compute the average MSE for the lth user as

$$\bar{\varepsilon}_l = \frac{\sum_{k \in \mathcal{L}} \|\mathbf{e}_k\|^2}{L}, \tag{10}$$

where $\mathcal{L}$ is the set of symbols received by user U_l and L is the number of symbols of $\mathcal{L}$. Figure 2 shows the MSE obtained by applying (10) and setting $L = 1000$, using 16-QAM and 256-QAM as examples. The parameter $\beta = 1 + N/2$, where $N = \sqrt{M}$, is the factor based on the first level event. The difference caused by the omission of other levels will be further estimated by the proposed neural network. The results for $\beta = 1$ and assuming the classical grid-based method for evaluating the MSE are also presented as benchmarks.
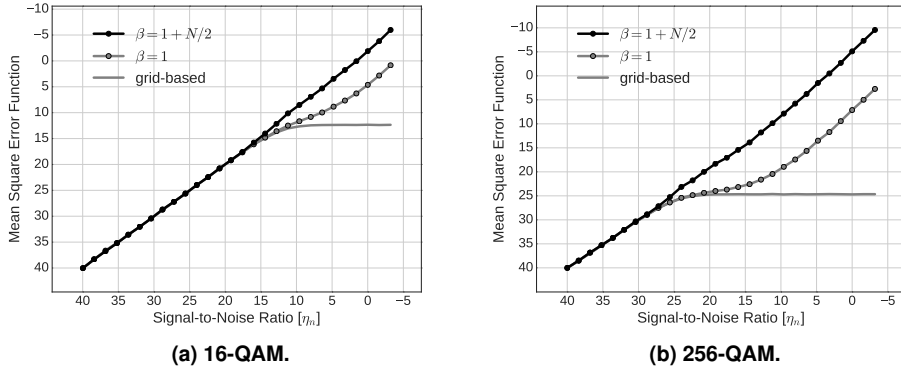

(a) 16-QAM. **(b) 256-QAM.**

Figure 2. Modified MSE vs SNR - AWGN channel.

Other constellation formats can be considered using the same approach. The main difference will be the employed proportion of detectable grid crossings (β factor).

4. Nonlinear regression

The samples provided by (10) are compared to the target (η_n) in order to produce data for the regression function. From the difference vector, a function can be created via curve fitting techniques. Polynomial regression is an option. However, in this case, the result is not satisfactory. A high order polynomial is necessary to fit the curve. Moreover, the curve fitting diverges from the observed data samples for values at the edge of the range.

A deep learning neural network, constituted of two hidden layers, can provide a suitable function. Figure 3 shows the achievement of a trained neural network. The chosen modulation option is 256-QAM. The values of predictions close to the edge are also consistent.
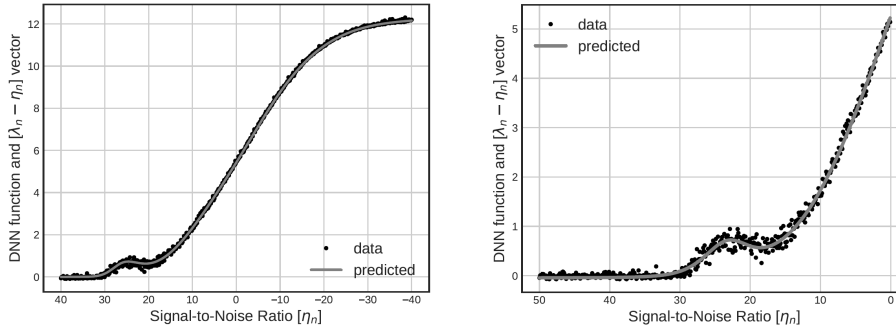

Figure 3. Nonlinear Regression using DNN.

As the DNN precisely fulfils the nonlinear regression, it can be used for the inference of imprecision given by the modified MSE function output, as shown in Figure 4.
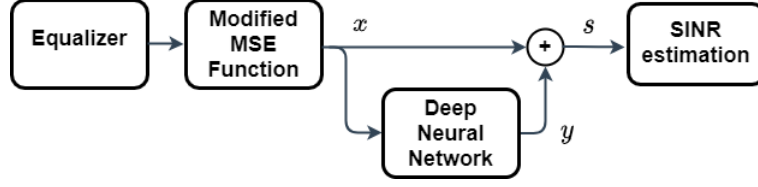

Figure 4. Modified MSE plus DNN scheme.

5. Deep learning network description

Keras platform was applied for implementing the neural network. The learning process considered an SINR=-10dB as the worst scenario. The configured DNN topology is shown in Figure 5. It has two hidden layers with four neurons per layer and a single input. A sigmoid activation function is applied in the second hidden layer. All the other layers employ linear activation functions. The biases are labelled as b_1, b_2 and b_3. More neurons can be added to the input layer with a specific time gap in a future implementation as an attempt to anticipate the estimation. Other approaches were already used to predict, as proposed in [Ngo et al. 2020]. Predictive strategies are interesting when the link condition indicates a fast-changing situation. However, this paper is focused on the deployment of an accurate and instantaneous indicator. But it can also be connected to other neural network topologies for prediction purposes.
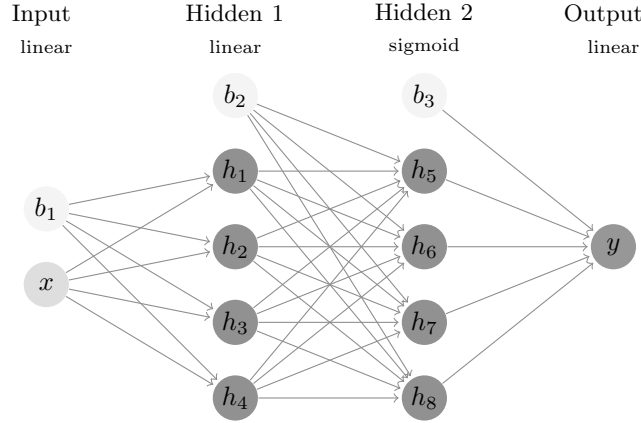

Figure 5. Implemented Deep Learning Network topology.

Figure 6 shows the achieved results for the proposed scheme. The chosen modulation is 256-QAM since it is the most complex case. As can be seen, the estimated values of SNR are very close to the ideal. A mean square error of 400 measures, made by the proposed function, was calculated. By considering ρ_k as the k-th measure of this vector and η_k as the respective ideal target, the value is given by

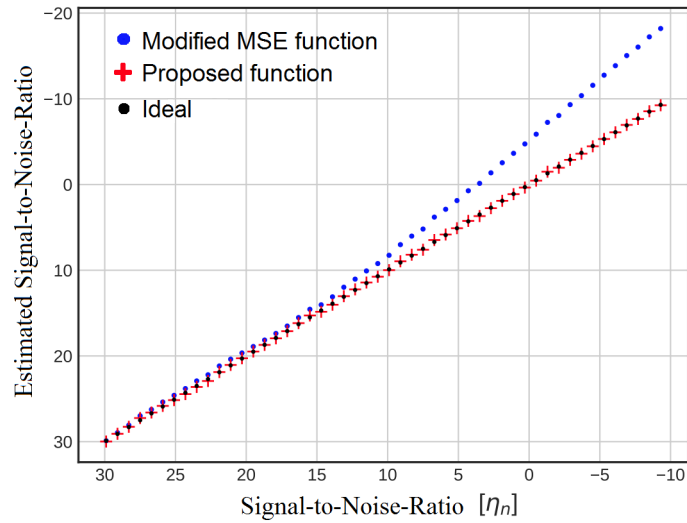$$M = \frac{1}{400} \sum_{k=1}^{400} (\rho_k - \eta_k)^2 = 0.00608495. \tag{11}$$

Figure 6. Proposed function response.

6. Conclusions

A precise estimation of the Signal-to-Interference-Plus-Noise Ratio was efficiently executed adopting the proposed modified Mean Square Error function plus a low-complexity Deep Learning Network. The DNN can be used in place of complex statistical computation. The adoption of machine learning as support for deterministic algorithms is an interesting approach. It can provide a lower complexity solution compared to an end-to-end neural network system. From this point on, different strategies can be executed to perform a better accuracy and prediction. Even a simplified implementation of lookup tables needs a good scheme to set the values. In this case, this proposed method can be also utilized. The same neural network topology can be applied to all modulation options described in the system model section using different predefined biases and weights.

References

3GPP TS 38.211 version 15.2.0 Release 15 (2018). *Physical channels and modulation*.

Baumgartner, S., Hirtz, G., and Baumgartner, A. (2014). A modified maximum likelihood method for SNR estimation in OFDM based systems. In *2014 IEEE International Conference on Consumer Electronics (ICCE)*, pages 155–158.

Bin Li, DiFazio, R., and Zeira, A. (2002). A low bias algorithm to estimate negative SNRs in an AWGN channel. *IEEE Communications Letters*, 6(11):469–471.

Khan, A. M., Jeoti, V., Rehman, M., and Jilani, M. (2017). Noise power estimation for broadcasting OFDM systems. In *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–6.

Malik, S., Portugal, S., Seo, C., Kim, C., and Hwang, I. (2011). Proposal and performance analysis of a novel preamble-based SNR estimation algorithm. In *2011 34th International Conference on Telecommunications and Signal Processing*, pages 100–103.

Ngo, T., Kelley, B., and Rad, P. (2020). Deep learning based prediction of signal-to-noise ratio for LTE and 5G systems. In *2020 8th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, pages 1–6.