

Received 13 November 2024, accepted 13 December 2024, date of publication 17 December 2024,  
date of current version 26 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3519310

## RESEARCH ARTICLE

# Synchronization and Detection in Molecular Communication Using a Deep-Learning-Based Approach

DUARTE CASALEIRO<sup>1</sup>, NUNO M. B. SOUTO<sup>2,3</sup>, (Senior Member, IEEE), AND JOÃO C. SILVA<sup>3</sup>

<sup>1</sup>Telecommunications and Computer Engineering, Iscte–University Institute of Lisbon, 1649-026 Lisbon, Portugal

<sup>2</sup>Instituto de Telecomunicações, 1049-001 Lisbon, Portugal

<sup>3</sup>Department of Information Science and Technology, Iscte–University Institute of Lisbon, 1649-026 Lisbon, Portugal

Corresponding author: Duarte Casaleiro (duarte\_casaleiro@iscte-iul.pt)

This work was supported by the Fundação para a Ciência e Tecnologia (FCT), I.P. by project reference UIDB/50008/2020, and DOI identifier <https://doi.org/10.54499/UIDB/50008/2020>.

**ABSTRACT** The concept of Internet of Bio-Nano Things (IoBNT) has emerged due to its revolutionary possibilities that transcend traditional wireless communication systems. Molecular Communication (MC) arises as a potential centrepiece for this paradigm, enabling applications in challenging environments. However, this type of communication, which often relies on molecular diffusion, suffers from a high inter-symbol interference (ISI), which deteriorates the reliability of the transmission. To cope with the strong ISI as well as the typical short coherence time of the MC channel, this work considers the adoption of a data-driven approach to accomplish non-coherent based detection at the receiver. In particular, we investigate the performance of a low complexity one-dimensional Convolutional Neural Network (1-D CNN) based in dilated causal convolutional layers and of a Gated Recurrent Unit based Recurrent Neural Network (GRU-RNN) aimed at the tasks of symbol detection and synchronisation, comparing the results with a conventional non-coherent detection. Initially, we study the performance of the proposed Neural Networks (NNs) based detectors assuming prior synchronisation between the transmitter and the receiver and, afterwards, we extend the approach for scenarios without prior synchronisation. Furthermore, we also investigate the robustness of the proposed NNs schemes against unknown variations in the distance between the transmitter and the receiver as well as in the diffusion coefficient. Finally, the results presented in this work lead to the conclusion that the implementation of NNs for both synchronisation and non-coherent detection can be a very effective approach for the challenging MC channel, ensuring more robustness than conventional model-based approaches.

**INDEX TERMS** 6G, future wireless networks, molecular communications, neural networks.

## I. INTRODUCTION

Throughout history, the Telecommunications domain has been improving in order to enhance our knowledge, our society and our industries. Currently, the research work regarding to the sixth generation (6G) wireless systems is expected to enable communications in environments previously considered unfeasible [1]. Therefore, to fulfil these challenges, the Internet of Bio-Nano Things (IoBNT) paradigm emerges due to its ability to enable the formation

of networks between bio-devices [2]. Given the extreme environments in which IoBNT is envisioned to be deployed, it is imperative that the devices communicate in a way that does not deteriorate the surrounding environment, so-called biocompatibility.

With this in mind, Electromagnetic Communication (EM) at nanoscale has restrictions since the antenna size is proportional to the signal wavelength [3], [4]. This implies operating in the Terahertz (THz) range [4], [5]. Still, operating at these higher frequency bands, not only increases the attenuation but also emits high-frequency radiation, making it unfeasible for the considered applications in IoBNT [2]. On the other hand,

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Fadda<sup>1</sup>.

Molecular Communication (MC) has the potential to be at the heart of this paradigm since the information is encoded in the physical properties of the information molecules, which act as carriers [2], [4], [6], [7], [8], [9], preventing any emission of radiation, thus ensuring biocompatibility. Additionally, MC is also energy efficient [4] since the propagation of the information molecules follows the random Brownian motion [3], [7], thus not requiring extra energy for the information to propagate through the molecular channel [4]. Furthermore, given the lack of limitations and regulations regarding to MC [4], it is possible to occupy the totality of the bandwidth. This enables a wide range of applications such as targeted drug delivery, health monitoring, pipeline communications for leak detection and swarm control for search-and-rescue missions [6], [10].

Despite the advantages of MC mentioned above, there are still challenges to overcome regarding these communication systems. One of the main challenges is the fact that the molecular channel has memory [3], [8], [9], which means that molecules are slowly absorbed by the molecular receivers [3], [8], remaining in the channel for a random period of time. This causes molecules from previous symbols to interfere with molecules from the current symbol, leading to a high inter-symbol interference (ISI) [4], [8], which restricts the transmission rates. Additionally, the scale and the environments in which MC communications systems are envisioned to be deployed require low complexity modulation, coding and detection schemes, while still ensuring robustness in the communication.

Similarly to traditional communication systems, MC systems are composed by a transmitter, a receiver and a channel through which the information is conveyed [4]. Relatively to the transmission process, it can be characterised as passive [10], where the information molecules propagate through the molecular channel via diffusion, or as propulsive [10], where the propagation of the molecules is induced by external forces, requiring extra energy. Regarding the propagation in the molecular channel via diffusion, it is one of the most studied methods due to its simplicity. In this type of MC, the information molecules propagate through the molecular channel, introducing a certain level of randomness into the channel itself. This implies that the molecular channel is constantly changing, making it difficult to estimate [11]. As for molecular receivers, they can be passive [9], where molecules are observed in an indirect form which does not change the concentration of molecules, or absorbing [9], where the molecules are absorbed and removed from the molecular channel. In order to encode the information in the physical properties of the molecules, there are several modulation schemes available. However, concentration shift keying (CSK) has been the most studied given its low complexity, since the information is encoded in the concentration of the information molecules [3].

Given the complexity constraints of MC, a non-coherent detection can be a solution for the detection process due to its simplicity [4] since this type of approach does not

require the estimation of channel state information (CSI) of the molecular channel [11]. Additionally, in MC, data-driven detectors can potentially have a better performance than model-driven ones [4]. Against this background, non-coherent detection based on a data-driven approach using a low-complexity neural network (NN) can be beneficial when the characteristics of the molecular channel are unknown [4].

Moreover, the implementation of neural networks in the demodulation process is proving to be very successful [12], [13]. Both in [12] and [14] it was shown the efficiency of convolutional neuronal networks (CNNs) in MC systems in order to decode the information sent through the inherently noisy molecular channel. These types of NNs are formed by convolutional layers, each one generating a higher level of abstraction, in order to extract essential information from the input [15]. For this to be possible, a filter of a given size sweeps across the input performing several convolutions. As already mentioned, the molecular channel has memory. Therefore, recurrent neural networks (RNNs) have the potential to surpass the accuracy of CNNs [12], [14]. This is due to the fact that RNNs use the feedback from past inputs to predict the current output [4], [16], thus taking into account the temporal variations of the molecular channel.

Based on the work and concepts mentioned above, in this paper we investigate the implementation of low complexity NNs in MC receivers, which can accomplish synchronisation and non-coherent detection of the MC signal. In order to demonstrate the effectiveness of the proposed NN architectures, we rely on extensive Monte Carlo simulations to evaluate the bit error rate (BER) of the molecular transmission. The main objective is to achieve a NN with reduced implementation complexity, while ensuring robustness in the communication itself. The main contributions presented in this research paper are the following:

- 1) We propose both a low complexity one-dimensional convolutional neural network (1-D CNN) based on dilated causal convolutional layers and a gated recurrent unit based recurrent neural network (GRU-RNN) architecture, which can perform non-coherent molecular detection, assuming prior synchronisation between the transmitter and the receiver.
- 2) We extend the design of the proposed 1-D CNN and GRU-RNN based approaches for scenarios where there is no previous synchronisation between the transmitter and the receiver. Two different methods are presented. One relies on two separate NNs dedicated to the synchronisation and detection tasks, whereas the second adopts a single NN for accomplishing simultaneous synchronisation and detection.
- 3) We train the proposed 1-D CNN and GRU-RNN based synchronisation and detection schemes in order to enhance their robustness against unknown variations in the distance between the transmitter and the receiver as well as in the diffusion coefficient. It is shown that both NN architectures achieve better performance than conventional non data-aided approaches.

At last, regarding its structure, this paper is organised as follows:

- In section II we introduce the considered system model.
- In section III, we present a description of the proposed architectures adopted for the neural networks, as well as the methods considered for synchronisation.
- Section IV describes the numerical simulations carried out for evaluating the performance of the different approaches.
- Section V presents a summary of the key conclusions obtained with this investigation.

## II. SYSTEM MODEL

### A. TRANSMITTED SIGNAL

In MC, the modulation process consists of encoding the information in the physical properties of the information molecules. As described in [9], there are several types of modulation techniques such as concentration-based, type-based, timing-based and spatial-based. Additionally, it is also possible to have hybrid techniques that rely on the combination of two or more modulation schemes.

CSK is a concentration-based modulation technique where the information is conveyed in the concentration of the information molecules. For this to be possible, the transmitter must change the concentration of molecules over each discrete time slot. In this paper, we consider a Binary CSK (BCSK) modulation technique that relies upon the fact that one symbol is represented by a single bit. In its simplest form, bit '1' can correspond to the release of a specific amount of molecules and bit '0' can correspond to not releasing any molecules. This specific modulation form is called On-Off Keying (OOK) and is represented in figure 1.

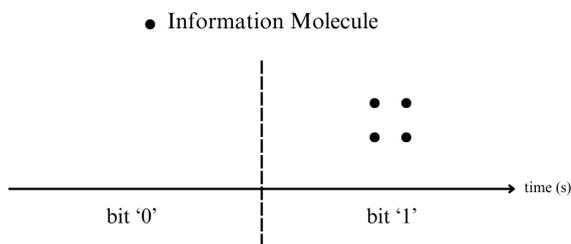


FIGURE 1. OOK modulation scheme, adapted from [3].

In communication systems, synchronisation is required for establishing a link between the transmitter and the receiver. In this work, we assume that a short synchronisation code is employed to ensure that the transmitter and receiver are synchronised. This code consists of a block of known symbols appended to the beginning of the information sequence, in order for the receiver to accurately detect the start of the frame sent by the transmitter. In this work, we consider the adoption of Barker Codes due to their good autocorrelation properties. In addition, we can append some '0' bits to the considered Barker Code to act as a waiting delay and reduce interference between the synchronisation

code and the information frame, thus resulting in a modified Barker Code. Figure 2 illustrates the frame structure where  $N_{BC}$  represents the length of the modified Barker Code and  $N_s$  denotes the number of information symbols.

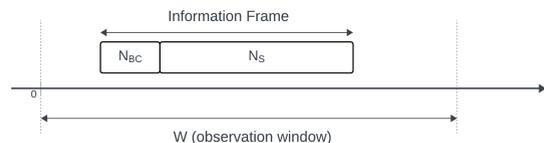


FIGURE 2. Frame structure.

### B. CHANNEL MODEL

In MC, information molecules act as information carriers. However, the propagation of the molecules is induced by vibrations due to their thermal energy and collisions with other particles, making their movement purely random [17]. Therefore, the propagation of molecules in any medium can be described by the Brownian motion [7], [9], [17] which, in a three-dimensional space, can be modelled by

$$\begin{aligned} (x_{i+1}, y_{i+1}, z_{i+1}) = & ([x_i + \mathcal{N}(0, 2D\Delta t)], \\ & [y_i + \mathcal{N}(0, 2D\Delta t)], \\ & [z_i + \mathcal{N}(0, 2D\Delta t)]). \end{aligned} \quad (1)$$

Hence, the position of an information molecule after a time interval ( $\Delta t$ ) is given by its initial position ( $x_i, y_i, z_i$ ), adding a spatial step  $\mathcal{N}(0, 2D\Delta t)$  following a standard Gaussian distribution with mean 0 and variance  $2D\Delta t$ .

As shown in (1), the diffusion coefficient ( $D$ ) has a direct impact on the movement of the information molecules. The mathematical model for this parameter can be described by the Stokes-Einstein relation as

$$D = \frac{k_b T}{\delta \pi \mu R} \quad [\text{m}^2/\text{s}] \quad (2)$$

where  $k_b$  is the Boltzmann constant ( $k_b = 1.38 \times 10^{-23} \text{J/K}$ ),  $T$  is the temperature of the environment in Kelvin,  $\mu$  is the viscosity of the fluid and  $R$  is the radius of the particle. Additionally, it is important to mention that  $\delta$  is determined by the relation between the information molecules and the molecules in the fluid [9]. If the information molecules are bigger than the fluid particles, this parameter is equal to 6, if not,  $\delta$  has a value of 4. For the investigation proposed in this paper, we consider a scenario where  $\delta = 6$ .

To fulfil the objectives proposed in this paper, we assume a transmission from a point transmitter to a fully-absorbing receiver [17]. The point transmitter considers the transmitter as a one-dimensional point where  $N_{molec}$  molecules are generated instantaneously, thus not taking into account either the geometry or the effects of the release of the molecules. Moreover, the model of a fully-absorbing receiver assumes that  $N_{rx}$  molecules are absorbed by the receiver's membrane during an observation window  $[t_u, t_l]$ , taking the geometry of the receiver into account. With this in mind, the probability

of an information molecule released at time  $t = 0$  being absorbed by the molecular receiver at time  $t$  during the observation window can be obtained as [17]

$$F_{hit}(t_u - t_l) = \frac{a_{rx}}{d_0} \left[ \operatorname{erfc} \left( \frac{d_0 - a_{rx}}{\sqrt{4Dt_u}} \right) - \operatorname{erfc} \left( \frac{d_0 - a_{rx}}{\sqrt{4Dt_l}} \right) \right] \quad (3)$$

where  $\operatorname{erfc}()$  is the complementary error function,  $a_{rx}$  represents the radius of the considered receiver and  $d_0$  is the distance between the centre of the transmitter and the centre of the receiver.

Assuming a time-invariant channel, we can define  $\bar{h}[t, \tau]$  as the expected number of molecules absorbed at time  $t$  by the receiver, after a release of  $N_{molec}$  molecules at time  $\tau$ .

$$\bar{h}[t, \tau] = N_{molec} F_{hit}(t, \tau). \quad (4)$$

Considering the diffusion of the molecules released by the transmitter and assuming the independency of the observations of the different molecules at the receiver, we adopt a Poisson distribution model which is commonly used in MC literature [11], [17], [18] as an accurate approximation to the binomial distribution that models the number of observed molecules. Therefore, to determine the concentration of the information molecules released at time  $\tau$  and absorbed by the receiver at time interval  $t$  we use

$$h[t, \tau] \sim \text{Poisson}(\bar{h}[t, \tau]). \quad (5)$$

Furthermore, as already mentioned in section I, the molecular channel suffers from a high level ISI. Hence, taking this effect into account, it is possible to define the concentration of molecules observed at the  $k$ -th symbol as

$$y[k] = \sum_{l=0}^{L-1} h[l, k] x[k-l] + n[k] \quad (6)$$

where  $L$  is the channel memory length and  $x[k]$  represents the  $k$ -th symbol modulated with binary CSK, thus  $x[k] \in [A_0, A_1]$  where  $A_0$  and  $A_1$  represent the level for the bit '0' and '1', respectively. In addition,  $n[k]$  represents the concentration of interfering molecules. These molecules are of the same type of the information carrying ones, even though they result from an independent process. As discussed in [11], they can be modelled using the same statistical distribution considered in (5), namely  $n[k] \sim \text{Poisson}(\bar{n}[k])$ .

At last, it is possible to define the concentration of information molecules observed by the receiver as

$$y[k] = \underbrace{\bar{h}[0, k] x[k]}_{\text{signal}} + \underbrace{n[k]}_{\text{environmental noise}} + \underbrace{v[k]}_{\text{diffusion noise}} + \underbrace{I[k]}_{\text{ISI}} \quad (7)$$

where the diffusion noise is modelled as a Poisson random variable whose mean has been subtracted. Hence, this variable is defined as  $v[k] \sim \text{Poisson}_0(\bar{h}[0, k] x[k])$ . Additionally,  $I[k]$  represents the ISI effect and is given by  $I[k] = \sum_{l=1}^{L-1} h[l, k] x[k-l]$ .

### III. SIGNAL DETECTION

The main goal of this work is to develop and study NN-based approaches for non-coherent detection and compare the performance against a direct non-coherent detection. The diagrams in figure 3 and 4 illustrate the main steps that are performed inside the receiver in order to obtain the final estimates of the information bits, using the two mentioned approaches.

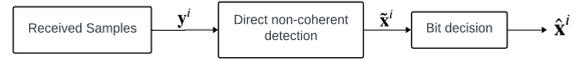


FIGURE 3. Demodulation process for the direct non-coherent detection.



FIGURE 4. Demodulation process for the NN-based non-coherent detection.

#### A. DIRECT NON-COHERENT DETECTION

In this work, we consider the adoption of non-coherent detection, which means that the detection process is carried out without any CSI estimation. In the case of a conventional non data-driven approach as illustrated in figure 3, we apply a direct non-coherent detection to the received samples and, afterwards, we perform a hard bit decision to the soft estimates.

For computing soft symbol estimates in the direct non-coherent detector, we follow our previous work in [19]. First we define the number of received molecules in each position of the  $N_s$  bits that compose the information block  $i$  as  $\mathbf{y}^i = (y^i[1], \dots, y^i[N_s])$ . Additionally, considering a BCSK modulation scheme, the  $k$ -th symbol of information block  $i$  is represented by  $x^i[k] \in [A_0, A_1]$ ,  $k = 1, \dots, N_s$ , where  $A_0$  and  $A_1$  denote the levels for bit '0' and '1', respectively, then, we define the probabilities  $p_{1,k} = P(x^i[k] = A_1 | \mathbf{y}^i)$  and  $p_{0,k} = P(x^i[k] = A_0 | \mathbf{y}^i)$  assuming an uniform distribution with bounds  $y_{min}^i$  and  $y_{max}^i$ , where  $y_{min}^i = \min(\mathbf{y}^i)$  and  $y_{max}^i = \max(\mathbf{y}^i)$ . These probabilities can be computed as

$$p_{1,k} = \frac{y^i[k] - y_{min}^i}{y_{max}^i - y_{min}^i} \quad (8)$$

and

$$p_{0,k} = \frac{y_{max}^i - y^i[k]}{y_{max}^i - y_{min}^i}. \quad (9)$$

Using these probabilities it is possible to obtain the log-likelihood ratio (LLR) for each bit as

$$\lambda^i[k] = \log \left( \frac{p_{1,k}}{p_{0,k}} \right) \quad (10)$$

which allow us to write

$$p_{1,k} = \frac{1}{1 + e^{-\lambda^i[k]}}. \quad (11)$$

Bearing in mind that soft values are the conditional expected value of the received symbols, these can be calculated as

$$\tilde{x}^i[k] = E(x^i[k]|y^i) = (A_1 - A_0)p_{1,k} + A_0. \quad (12)$$

Finally, using (11), we can rewrite the expression for the computation of soft estimates as

$$\tilde{x}^i[k] = \frac{(A_1 + A_0)}{1 + e^{-\lambda^i[k]}} + A_0. \quad (13)$$

For OOK, we have  $A_0 = 0$  and  $A_1 = 1$  and (13) can be reduced to

$$\tilde{x}^i[k] = \frac{1}{2} \left( \tanh \left( \frac{\lambda^i[k]}{2} \right) + 1 \right). \quad (14)$$

Although we do not consider here, these soft symbol estimates can be used as input for a channel decoder when channel coding is employed. Since in this work we assume the transmission of uncoded information, we apply hard decision to the soft estimates coming from the direct non-coherent detector. Therefore, the  $k$ -th estimated bit of information block  $i$ ,  $\hat{x}^i[k]$ , can be obtained as

$$\hat{x}^i[k] = \begin{cases} 0, & \tilde{x}^i[k] \leq \phi \\ 1, & \tilde{x}^i[k] > \phi, \end{cases} \quad (15)$$

where  $\phi$  represents the threshold value and is given by  $\phi = (A_0 + A_1)/2$ . In the case of OOK, the threshold value becomes  $\phi = 0.5$ .

### B. NEURAL NETWORK BASED DETECTION

As an alternative to the direct non-coherent detection, described above in section III-A, we propose a data-driven approach based on the implementation of NNs at the receiver. As it can be observed in figure 4, in this approach we apply a NN to the received samples in order to predict the information that was sent through the molecular channel. Similarly to the direct non-coherent detection, the adopted NN architectures output ‘‘soft’’ symbol estimates. Therefore, we apply hard bit decision to map these estimates into bits ‘0’ or ‘1’ using (15). Given that the molecular channel is very noisy and can be constantly changing, the received signal can vary considerably. Therefore, to address this challenge, in this work we adopted regression networks in order to obtain ‘‘soft’’ bit estimates, which are relevant for a channel decoder when the transmitted information is encoded. However, it is important to note that the architectures adopted can be easily adapted to classification based neural networks.

Finally, it is important to note that, given the constraints regarding MC, the proposed NNs were designed in order to achieve a good trade-off between the complexity of the NN and reliability, in terms of BER values achieved, for the molecular transmission. Two different architectures were considered: a 1-D CNN based on dilated causal convolutional layers and a GRU-RNN, as detailed next.

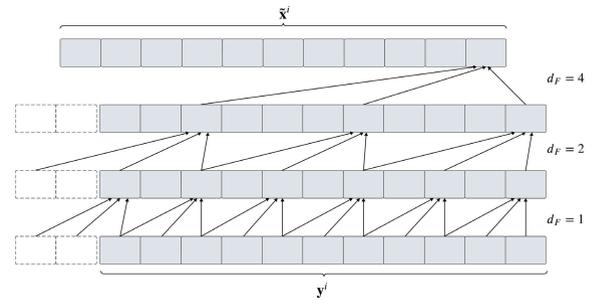


FIGURE 5. Architecture of a 1-D CNN based in dilated causal convolutional layers, adapted from [21].

#### 1) 1-D CONVOLUTIONAL NEURAL NETWORK

The first NN-based approach proposed for accomplishing non-coherent detection of the information that passes through a molecular channel, corresponds to a low complexity 1-D CNN based on dilated causal convolutional layers architecture. The main feature of this CNN architecture is that it has the ability to memorise past inputs [20]. Therefore, bearing in mind that the molecular channel has memory, with this specific NN architecture, it is possible to take into account the temporal variations of the molecular channel.

The proposed NN is a variation of traditional CNNs. Specifically, the considered 1-D CNN is formed by dilated causal convolutional layers, which operate over the time steps of each sequence. In addition, one of the main features of these specific layers is that they are causal, meaning that each current output ( $\tilde{x}^i[k]$ ) does not depend on future inputs [21]. The basic architecture of this NN is illustrated in figure 5. This NN is formed by convolutional layers with dilation, in order to increase the receptive field of each layer, without increasing the number of parameters. The dilation factor ( $d_F$ ) determines the step size for sampling the input ( $y^i$ ). Consequently, as the dilation factor increases, the distance between two consecutive filter taps also increases. The receptive field ( $R$ ) is equivalent to the time steps of the input sequence used by the network for each estimate, which is calculated as

$$R = (f - 1)(2^j - 1) + 1 \quad (16)$$

for a stride of 1 and a dilation factor of  $2^{j-1}$  at layer  $j$ , where  $f$  denotes the filter size and  $J$  is the number of convolutional layers in the NN.

In particular, we propose a deep NN comprising three one-dimension convolutional layers with a rectified linear unit (ReLU) activation function and a batch normalization layer between each of these layers. For the input layer, which inputs sequence data to the 1-D CNN, a data normalization is applied so that the input is rescaled to be in a range between 0 and 1. This network also includes a fully-connected layer as the output layer. Furthermore, each 1-D convolutional layer considers a stride of 1, a filter size of  $f = 4$  and a causal method to determine the padding size, which is given by  $p = (f - 1)d_F$ , to ensure consistency in the size of the

frames along the network layers. Moreover, while the first 1-D convolutional layer has 2 filters, the other two have 4 filters. At last, to ensure a larger receptive field, the dilation factor increases exponentially in each convolutional layer, ranging from  $d_F = 1$  to  $d_F = 4$ . In table 1, presented below, it is possible to observe the detailed architecture of this NN, as well as the parameters of each layer.

In this paper, all of the data used for training and validation of the NN was generated using Monte Carlo simulations. This dataset is independent from the data utilized for the find testing simulations. Therefore, we used a training and a validation dataset ( $N_{seq}$ ) of 19200 sequences with a 99%/1% proportion using a length of  $N_s = 20$  bits per frame. Given the large quantity of data, the dataset was divided into batches of 32 sequences to reduce overfitting and also to increase the performance and the accuracy of the network itself. Each sequence in the dataset is generated by first modulating a random sequence of 20 bits using binary CSK as described in section III-B, which represents the information that was transmitted. Afterwards, this information sequence passes through a time-invariant molecular channel modelled as (6), and onto which additive stationary noise is added. Therefore, each training pair that will feed the Neural Network will comprise this generated received sequence as well as the corresponding original information sequence in order to be used in the mean squared error loss function (17). The sequences used in the testing simulations are generated using a similar approach in order to be used in the loss function adopted for the NN.

Furthermore, we adopted the mean squared error loss function ( $L_{MSE}(\theta)$ ), which is described as

$$L_{MSE}(\theta) = \frac{1}{N_{seq}} \sum_{i=1}^{N_{seq}} (f_{\theta}(\mathbf{y}^i) - \mathbf{x}^i)^2 \quad (17)$$

considering the NN as a function  $f_{\theta}$ , where  $\theta$  represents the learnable parameters. Additionally,  $\mathbf{y}^i$  and  $\mathbf{x}^i$  represent the  $i$ -th block of received molecules and information symbols, respectively. Moreover, the Adaptive Moment Estimation (Adam) method [22] is used as the optimisation algorithm, with an initial learning rate of 0.001. A maximum of 80 epochs were considered, where the training data is shuffled before each epoch to prevent overfitting. At last, the NN has a validation patience of 5, meaning that if after five attempts the loss function has not decreased its value, the training progress is stopped.

## 2) GATED RECURRENT UNIT BASED RECURRENT NEURAL NETWORK

As an alternative approach to the previous 1-D CNN, we propose a low complexity GRU-RNN to also perform non-coherent detection of the information sent through a molecular channel. Specifically, we adopt a three-layer GRU-RNN with a sequence input layer, two gated recurrent unit (GRU) layers and a fully-connected layer as the output layer.

GRU layers were proposed in [23] as a low complexity alternative to long short-term memory (LSTM) layers firstly introduced in 1997 by Hochreiter and Schmidhuber in [24]. Both these types of layers have the ability to learn long-term dependencies between data sequences. However, GRU layers have fewer parameters and a simpler architecture, making them more efficient. Figure 6 illustrates the structure of a GRU unit as adopted in this work. This unit relies in two different gate units that depend on the previous hidden state [23]: a reset gate ( $\Gamma_r$ ) and an update gate ( $\Gamma_z$ ). As the name suggests, the reset gate enables the unit to forget its previous state and the update gate determines how much the unit will be updated [25]. Both these two gates are computed with a sigmoid function ( $\sigma$ ), which is applied to the current input ( $a_t$ ) and to the previous value of the memory cell ( $b_{t-1}$ ). Moreover,  $\tilde{b}_t$  is the candidate activation, which is updated at every timestep, and  $b_t$  is the output of the GRU unit at time  $t$ .

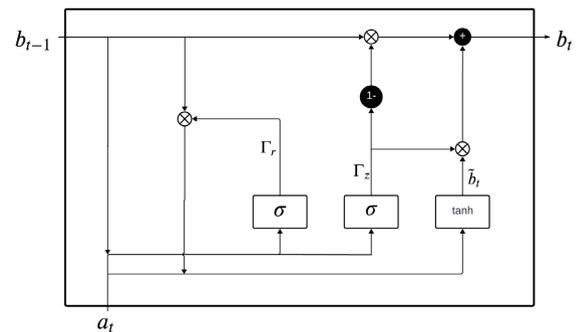


FIGURE 6. Structure of a GRU unit.

In the adopted GRU-RNN, the two GRU layers are implemented with 4 hidden units each. This was found to result in a good compromise between complexity and performance. The detailed architecture of this GRU-RNN is presented, in table 2.

Regarding the training parameters, this NN was trained with the same settings as the 1-D CNN described previously. Hence, a training and validation dataset was used, comprising 19200 sequences, each one with  $N_s = 20$  bits per frame, divided into batches of 32 sequences to prevent overfitting. Adam was also used as the optimisation algorithm, with an initial learning rate of 0.001, and the mean squared error loss function, already described in (17). Finally, a validation patience of 5 and a maximum of 80 epochs was considered, where the training data is shuffled before each epoch.

It is important to highlight that the previously described NNs architectures were initially aimed at non-coherent molecular detection. However, as explained in the next section, they were also extended for scenarios where there is no prior synchronisation between the transmitter and the receiver. Hence, the proposed NN-based approaches are able to accomplish both synchronisation and non-coherent detection.

TABLE 1. 3 layer CNN architecture.

Layer	Type	Filter size	Filters	Dilation Factor	Activation	Learnable Parameters
1	Sequence input	–	–	–	–	–
2	1-D Convolutional	4	4	1	ReLu	20
3	Batch Normalization	–	–	–	–	8
4	1-D Convolutional	4	2	2	ReLu	34
5	Batch Normalization	–	–	–	–	4
6	1-D Convolutional	4	2	4	ReLu	18
7	Batch Normalization	–	–	–	–	4
8	Fully-connected	–	–	–	–	3

TABLE 2. GRU-RNN architecture.

Layer	Type	Hidden Units	Learnable Parameters
1	Sequence input	–	–
2	GRU	4	72
3	GRU	4	108
4	Fully-connected	–	5

### C. FRAME SYNCHRONISATION

For the scenario where there is no prior synchronisation between the transmitter and the receiver, we propose three distinct methods for performing the synchronisation as well as the detection of the information sent from the transmitter to the receiver:

- **Method 1** - In this first method, a NN-based detection is implemented after a conventional synchronisation with the considered modified Barker Codes.
- **Method 2** - The second method relies on the implementation of a NN over the entire observation window of the receiver. Hence, in this case, we accomplish joint synchronisation and detection using a single NN.
- **Method 3** - With this method we propose the implementation of two different NNs, one for the synchronisation and another for the detection process.

In **method 1**, the NNs proposed in section III-B are only implemented for the detection process. Therefore, we apply conventional synchronisation, which relies on correlations between the received signal and the modified Barker Code along the whole observation window. The location of the maximum value for all the performed correlations is assumed to be the position of the synchronisation code in the observation window. Therefore, the position of the beginning of the information frame ( $\hat{k}_0$ ) can be computed as

$$\hat{k}_0 = \operatorname{argmax}(\mathbf{c}^T \mathbf{y}[k : k + N_{BC} - 1]), \quad 0 \leq k \leq W - N_{BC} - N_s. \quad (18)$$

In this case, the bit sequence and the length of the modified Barker Code are represented by  $\mathbf{c} = [c_0, \dots, c_{N_{BC}-1}]^T$  and  $N_{BC}$ , respectively. Moreover,  $N_s$  defines the total number of information symbols and  $W$  represents the size of the observation window. After obtaining the position of the synchronisation code, a NN is applied to the following 20 symbol positions in order to estimate the information that was received.

As for **method 2**, we follow an approach where a NN similar to the ones proposed in section III-B is applied to the entire observation window for accomplishing joint synchronisation and detection. As it can be observed in figure 7, we start by applying a NN to the whole observation window as is done in the case of the NN-based detection, thus obtaining soft bit estimates. Then, we perform a hard bit decision to the result in order to map the soft estimates onto bits '0' or '1', i.e., using (15). Afterwards, to find the beginning of the information frame, we apply a simple correlation between the estimated sequence and the Barker Code, similarly to (18). Finally, knowing the position of the synchronisation code in the observation window, we segment the frame by extracting the following  $N_s$  bits, which are considered to be the information that was sent by the transmitter.

Finally, for **method 3**, we propose the implementation of two different NNs also similar to the ones presented in section III-B: one for the synchronisation and another one for the detection process, as can be observed in figure 8. For this to be possible, we start by following an approach similar to method 2. However, after the correlation process to detect the beginning of the information frame, we segment the bit sequence  $\mathbf{y}^i$ . Then, we apply a NN-based detection to the resulting  $N_s$  samples in order to obtain soft bit estimates. Finally, using (15), we apply a hard decision to the soft estimates to map them into bits '0' or '1' to detect the information that was sent through the molecular channel.

## IV. RESULTS

### A. SIMULATION PARAMETERS

In this section, we present the results of the simulations conducted in this investigation, in order to evaluate the proposed NN-based approaches for MC.

Table 3 presents the main parameters considered for the simulations performed in this work. For this investigation, we consider the transmission of blocks of  $N_s = 20$  bits using OOK as a modulation scheme with  $A_0 = 0$ ,  $A_1 = 1$  and  $N_{molec} = 500$ . Moreover, we consider a deterministic model of a point transmitter ( $tx$ ), a spherical absorbing receiver ( $rx$ ) and a Poisson distribution as a stochastic model for the channel. Furthermore, all simulations considered in this paper assume the transmission of uncoded unipolar non-return to zero (NRZ) signals where each sample corresponds to a symbol.

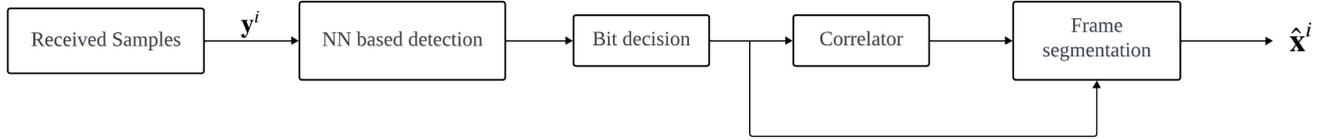


FIGURE 7. Demodulation process for method 2.

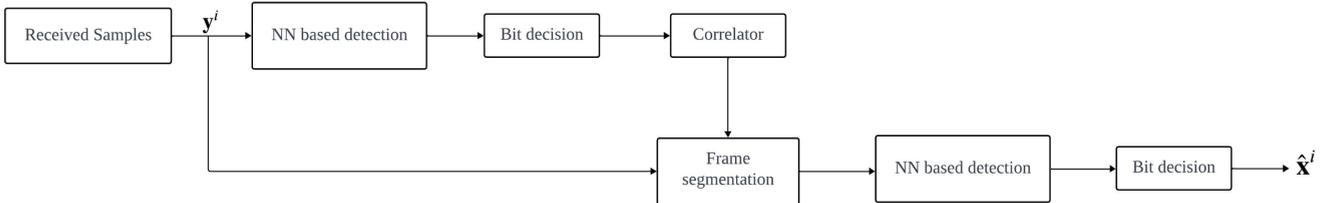


FIGURE 8. Demodulation process for method 3.

TABLE 3. Main parameters considered for the Monte Carlo simulations.

Parameters	Values
Molecules transmitted	500
Deterministic model	Point transmitter to Spherical absorbing receiver
Stochastic model	Poisson
Signal pulse type	Unipolar NRZ
Symbol duration ( $T_s$ )	$\alpha \cdot \frac{d_{tx-rx}^2}{6D}$ [s]
Modulation Scheme	OOK

Moreover, for the symbol duration ( $T_s$ ), which has a significant impact on the transmission process, we used

$$T_s = \alpha \cdot \frac{d_{tx-rx}^2}{6D}. \tag{19}$$

It is important to note that, for all simulations presented in this paper, we have defined the scaling factor ( $\alpha$ ) as 1.

From (19), we can conclude that the diffusion coefficient ( $D$ ) and the distance between the transmitter and the receiver ( $d_{tx-rx}$ ) have a direct effect on the assumed symbol duration, which has a major influence on the ISI effect in the communication. Depending on the applications in which molecular communications are envisioned to be deployed, it is possible that the surrounding environment will be constantly changing. This implies that the transmitter and the receiver may move closer or further apart from each other, resulting in variations in the distance between these two devices. Furthermore, the diffusion coefficient, previously described in (2), influences the propagation speed of the information molecules through the molecular channel [17] and directly depends on the temperature and the viscosity of the environment. Thus, a variation in these parameters may change the value of the diffusion coefficient, which also has a significant impact on the transmission process. Since it is assumed that the transmitter and receiver have no knowledge of these variations,  $T_s$  is set with the expected distance and diffusion values.

In this investigation, we start always by evaluating the performance of the proposed NNs in a scenario where there are no variations in the distance and in the diffusion coefficient. Hence, we consider a distance between the transmitter and the receiver ( $d_{tx-rx}$ ) of  $10 \mu\text{m}$  and a value for the diffusion coefficient ( $D$ ) of  $79.4 \mu\text{m}^2/\text{s}$ . Then, bearing in mind the influence that parameters such as the distance between the transmitter and the receiver and the diffusion coefficient have in the transmission process, we also present the results for a scenario considering variations of these parameters.

It is important to note that, following some experiments, it was decided to train both neural networks considering 100 dB in terms of signal to noise ratio (SNR), which is defined as  $N_{molec}/\sigma_n$  where  $\sigma_n^2$  denotes the noise variance.

### B. NEURAL NETWORK BASED DETECTION WITH PERFECT SYNCHRONISATION

In this section, we simulate a molecular transmission using the 1-D CNN and the GRU-RNN based detectors, which were previously described in section III-B, assuming perfect synchronisation between the transmitter and the receiver. For the analysis of the results, all simulations are compared with a direct non-coherent detection, using the approach described in section III-A.

The presented test results are obtained from simulations using a set of transmissions with 10000 frames each with  $N_s = 20$  bits. Given that the quality of the communication directly depends on the channel conditions, the main objective is to evaluate the BER of the conventional molecular communication in comparison against a molecular communication where a NN-based approach is implemented for non-coherent detection.

As described in section III-B1 and III-B2, the training and validation datasets are composed by 19200 sequences, each with  $N_s$  bits per frame. In order to help verify the suitability of

the selected size for the training and validation datasets, the proposed 1-D CNN was trained with different dataset sizes, namely with 3840, 19200 and 38400 sequences. The results obtained for the different sets are illustrated in figure 9.

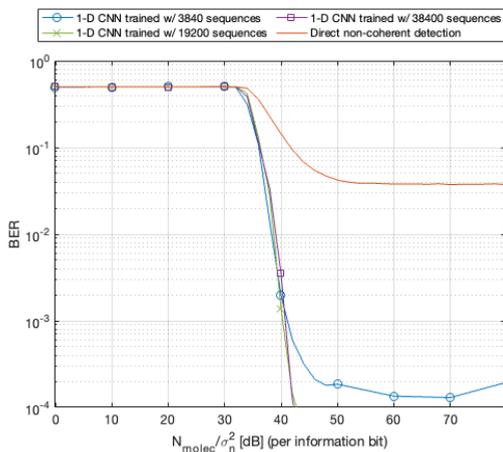


FIGURE 9. BER comparison between the NNs trained with different dataset sizes.

As expected, in figure 9 it is possible to observe that the NNs trained with 19200 and 38400 sequences exhibited better performance in comparison with the NN trained with only 3840 sequences. However, there is no apparent improvement in the performance between the NN trained with 19200 sequences and the one trained with 38400 sequences. Consequently, given the complexity constraints regarding MC, the training dataset with 19200 sequences revealed to be the most suitable option.

The decision to implement the specific 1-D CNN configuration, previously described in table 1, is justified by the superior complexity-performance trade-off exhibited by this specific NN architecture in comparison with other NNs with a greater number of layers. In order to help assess the advantage of the adopted three-layer NN architecture configuration, we compared it against other 1-D CNNs with different depths, namely five and seven one-dimensional convolutional layers. All NNs have a rectified linear unit (ReLU) activation function and a batch normalization layer between each one-dimensional convolutional layer. For the input layer, which inputs sequence data to each 1-D CNN, a data normalization is applied so that the input is rescaled to be in a range between 0 and 1. These networks also include a fully-connected layer as the output layer. In tables 4 and 5, it is possible to observe the detailed architecture of these additional 1-D CNNs, as well as the parameters of each layer. The comparison between the performance of these two NNs and of the NN from table 1 is illustrated in figure 10. It is important to note that due to the larger number of learnable parameters, a larger dataset was used for training these two deeper NNs.

From figure 10, it is possible to observe that there is no significant improvement in the performance of the NNs

comprising more convolutional layers in comparison with the three-layer NN. In addition, the reduced number of layers in the three-layer 1-D CNN results in a lower complexity NN when compared with the other considered 1-D CNN, which have a higher number of learnable parameters. Consequently, based on the presented results, we concluded that the three-layer 1-D CNN aligns effectively with the complexity constraints of MC due to its low-complexity, therefore being more suitable for the intended application.

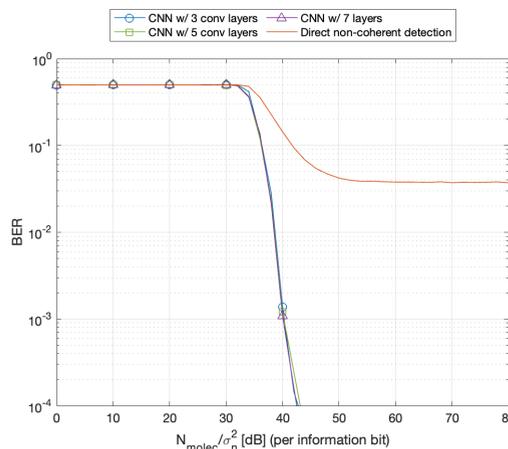


FIGURE 10. BER comparison between the NNs with different depths.

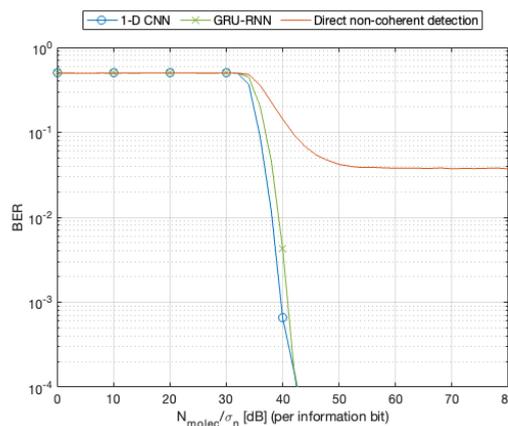


FIGURE 11. BER comparison between the NNs and the direct non-coherent detection.

Figure 11 illustrates a comparison between the BER of the conventional molecular communication and the NN-based non-coherent detection. As it can be seen in this figure, for SNR values above 42 dB, the proposed NNs can achieve BER values lower than  $10^{-4}$ . This is a difference of three orders of magnitude when compared with the BER of the direct non-coherent detection, which has a BER floor value of  $10^{-1}$ . In addition, it is possible to observe that, for both NNs, the value of the BER decreases for a SNR value higher than 30 dB, matching the behaviour of the direct non-coherent detection. At last, we can observe that both the proposed NN-based detectors have a similar performance.

TABLE 4. 5 layer CNN architecture.

Layer	Type	Filter size	Filters	Dilation Factor	Activation	Learnable Parameters
1	Sequence input	–	–	–	–	–
2	1-D Convolutional	2	8	1	ReLu	24
3	Batch Normalization	–	–	–	–	16
4	1-D Convolutional	2	4	1	ReLu	68
5	Batch Normalization	–	–	–	–	8
6	1-D Convolutional	4	4	1	ReLu	68
7	Batch Normalization	–	–	–	–	8
8	1-D Convolutional	4	2	2	ReLu	34
9	Batch Normalization	–	–	–	–	4
10	1-D Convolutional	4	2	4	ReLu	18
11	Batch Normalization	–	–	–	–	4
12	Fully-connected	–	–	–	–	3

TABLE 5. 7 layer CNN architecture.

Layer	Type	Filter size	Filters	Dilation Factor	Activation	Learnable Parameters
1	Sequence input	–	–	–	–	–
2	1-D Convolutional	2	16	1	ReLu	48
3	Batch Normalization	–	–	–	–	32
4	1-D Convolutional	2	8	1	ReLu	264
5	Batch Normalization	–	–	–	–	16
6	1-D Convolutional	3	8	1	ReLu	200
7	Batch Normalization	–	–	–	–	16
8	1-D Convolutional	3	4	1	ReLu	100
9	Batch Normalization	–	–	–	–	8
10	1-D Convolutional	4	4	1	ReLu	68
11	Batch Normalization	–	–	–	–	8
12	1-D Convolutional	4	2	2	ReLu	34
13	Batch Normalization	–	–	–	–	4
14	1-D Convolutional	4	2	4	ReLu	18
15	Batch Normalization	–	–	–	–	4
16	Fully-connected	–	–	–	–	3

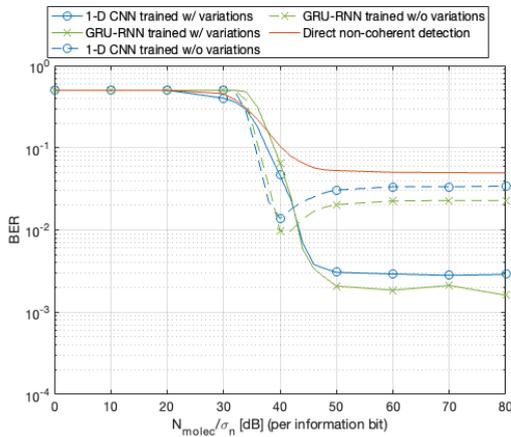
Taking into account the previous results, in this testbed we also wanted to evaluate the impact of the variability of the channel conditions. Therefore, we simulated the performance of the proposed NNs, trained with a varying range of distances between the transmitter and the receiver and diffusion coefficient errors. For both the validation and the training process, we considered a range of distances between  $5 \mu\text{m}$  and  $15 \mu\text{m}$  and diffusion coefficient errors randomly changing between 5% and 15%. In addition, to cope with the expected variations in the molecular channel, which are assumed unknown at the receptor, the symbol duration parameter was modified in order to reduce the ISI effect. Hence, considering (19),  $d_{Tx-Rx}$  was set with a value of  $15 \mu\text{m}$  whereas a maximum diffusion coefficient error of 15% was added to the standard value for the diffusion coefficient, resulting  $D = 91.31 \mu\text{m}^2/\text{s}$ .

In order to study the behaviour of the NNs in this scenario, in figure 12 we present the results for the original NNs trained both with and without the variations in the molecular channel. Comparing these results with the ones obtained in figure 11, we observe an increase in the BER for both NNs, regardless of the training method employed. This can be attributed to the more challenging nature of this scenario, as the considered variations in the distance between the transmitter and the receiver and in the diffusion coefficient, which are unknown at the receiver, lead to a significant deterioration of the

molecular transmission, thus increasing the BER. However, training both NNs under the effect of these variations improves the robustness, leading to a reduction in the overall BER of the molecular transmission for this scenario. We can observe a difference of approximately one order of magnitude when compared with a scenario where no NN is adopted or where a NN trained without considering variations in these parameters is used. Moreover, it is also noticeable that, independently of the training method, the BER for the GRU-RNN is slightly lower when compared with the BER achieved by the 1-D CNN.

### C. NEURAL NETWORKS FOR THE DETECTION AND SYNCHRONISATION PROCESS

In this section we study the performance of the proposed NNs for a scenario where there is no previous synchronisation between the transmitter and the receiver. As already mentioned in section II-A, for this scenario we chose to employ a modified Barker Code for the synchronisation process. Hence, we consider a Barker Code of length 7, appending three additional '0' bits. Therefore, we consider the following synchronisation binary code:  $\mathbf{c} = [1110010000]$ . It is important to note that this specific synchronisation code was chosen after several simulations considering Barker Codes with different lengths. The considered synchronisation code



**FIGURE 12. BER comparison between the NNs and the direct non-coherent detection, considering variations in the distance and in the diffusion coefficient.**

is a trade-off between the complexity of the synchronisation process and the correct synchronism rate achieved.

Regarding the observation window, since it is assumed that initially the considered molecular receiver only detects an increase in the concentration of the information molecules, it will not be able to determine whether it detected the end or the beginning of the information frame. Hence, in this work, we consider that the receiver works with an observation window three times larger than the transmitted frame. Therefore, for a frame of  $N_s = 20$  bits and a modified Barker Code with a total of  $N_{BC} = 10$  bits, the receiver has an observation window of  $W = 90$  bits, plus the maximum channel memory length which is considered to be of  $L = 7$  symbols.

Finally, it is important to mention that all the simulations presented in this section are compared with a direct non-coherent detection, where no NN is applied.

To evaluate the performance of the NNs for each of the proposed methods, we consider the same parameters and values described in section IV-A. Moreover, we present the results of the BER and the synchronisation rate as a function of  $N_{molec}/\sigma_n$ , in order to observe the evolution of the accuracy of each method. Additionally, it must be noted that the BER results only consider the cases where the synchronism was correct.

From the analysis of figure 13.a) we can conclude that, for each method, both of the proposed NNs have a similar behaviour. Finally, it is important to mention that, since the detection process is the same for method 1 and method 3, the results for these methods are illustrated in the same plot line.

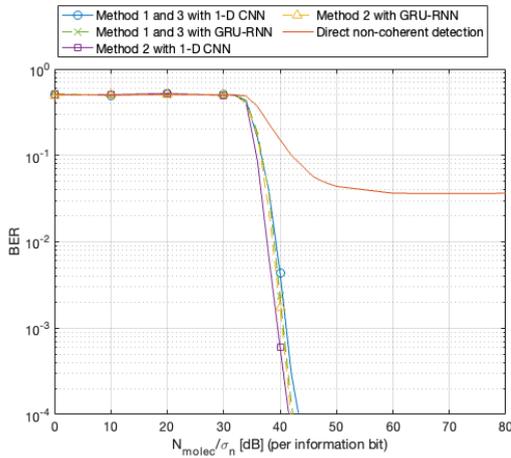
As it can be seen in figure 13.b), since both method 1 and the direct non-coherent detection apply a conventional correlation for the synchronisation process, the results for these two scenarios are illustrated in the same plot line. Moreover, we can observe that, in this case, there is an increase in the rate of correct synchronisms for SNR

values between 10 dB and 40 dB, achieving a successful synchronisation rate above 90% for values over 40 dB. Although method 1 shows better synchronisation rates for SNRs below 35 dB, it corresponds to a scenario in which the BER is very high, approaching 50%, therefore making it less relevant. In contrast, for method 2 we can observe synchronisation rates of 100% for values of 40 dB and above for both NNs. However, it is important to note that, for this case, the synchronisation rate only starts to increase at SNRs around 35 dB but it quickly reaches its maximum value. Finally, for method 3, we can observe that the synchronisation rate is identical for both of the proposed NNs. In addition, this method has a similar performance to method 2. However, since this method relies in the implementation of two NNs, the synchronisation and detection process incurs in a higher complexity when compared with method 2.

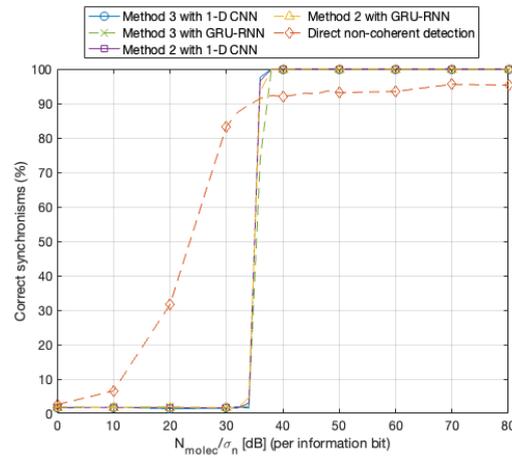
Bearing in mind the previous results, in the following testbed we evaluate the performance of the proposed NNs for each of the previously mentioned methods of synchronism, considering a scenario with variations in the distance between the transmitter and the receiver as well as the diffusion coefficient. For this to be possible, we consider a range of distances between  $5 \mu\text{m}$  and  $15 \mu\text{m}$  and diffusion coefficient errors randomly varying between 5% and 15% for both the training and the validation process.

As it can be seen in figure 14.a), for each of the considered methods, the NNs that were trained considering variations in the distance and in the diffusion coefficient have a better performance, achieving lower BER values when compared with the other scenarios. In addition, we observe that, for all methods, the proposed GRU-RNN has a lower BER floor value when compared with the 1-D CNN. Moreover, it is noticeable that, for method 1 and 3, the BER value of the 1-D CNN trained with variations decreases for a SNR value higher than 20 dB, surpassing the behaviour of the direct non-coherent detection.

Regarding the synchronisation rate, illustrated in figure 14.b), we can observe a significant reduction in the synchronisation rate for the direct non-coherent detection (and for method 1 since this also implements a conventional synchronisation). Unlike the 90% synchronisation rate achieved previously in figure 13.b), in this case it is only possible to achieve a maximum synchronisation rate around 84%. As for method 2, unlike what was observed in figure 13.b), for both NNs trained considering variations in the distance and in the diffusion coefficient, it is only possible to achieve a synchronisation rate of 100% for  $N_{molec}/\sigma_n$  values of approximately 45 dB and above. For method 3, it is possible to observe that the synchronisation rate for the 1-D CNN trained with variations starts to increase from SNR values of 20 dB and above. At last, it is noticeable that, for the GRU-RNN, it is possible to achieve a synchronisation rate of 100% for  $N_{molec}/\sigma_n$  values of approximately 45 dB and above. However, for the 1-D CNN, it is only possible to achieve a maximum synchronisation rate of approximately 96%. In contrast, the same does not occur when the proposed

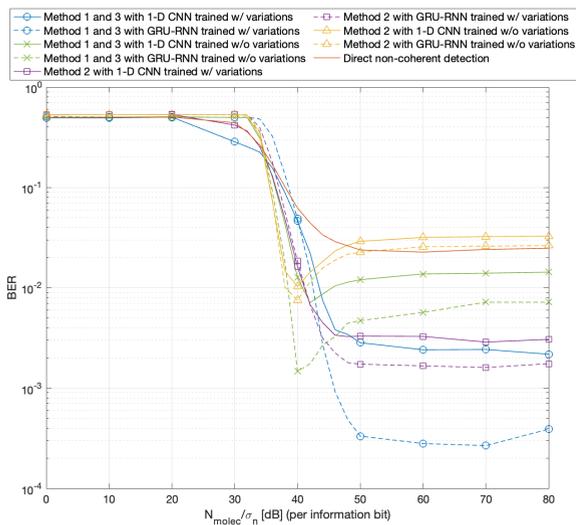


a) BER comparison between the NNs and the direct non-coherent detection for each method.

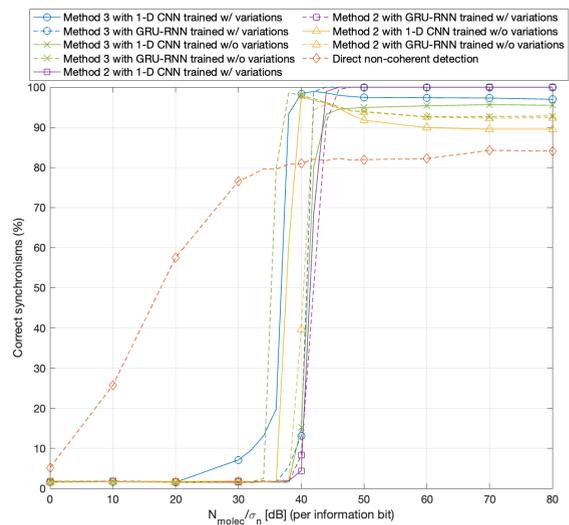


b) Synchronisation rate comparison between the NNs and the direct non-coherent detection for each method.

FIGURE 13. NNs applied to each synchronisation and detection method.



a) BER comparison between the NNs and the direct non-coherent detection for each method, considering variations in the distance and in the diffusion coefficient.



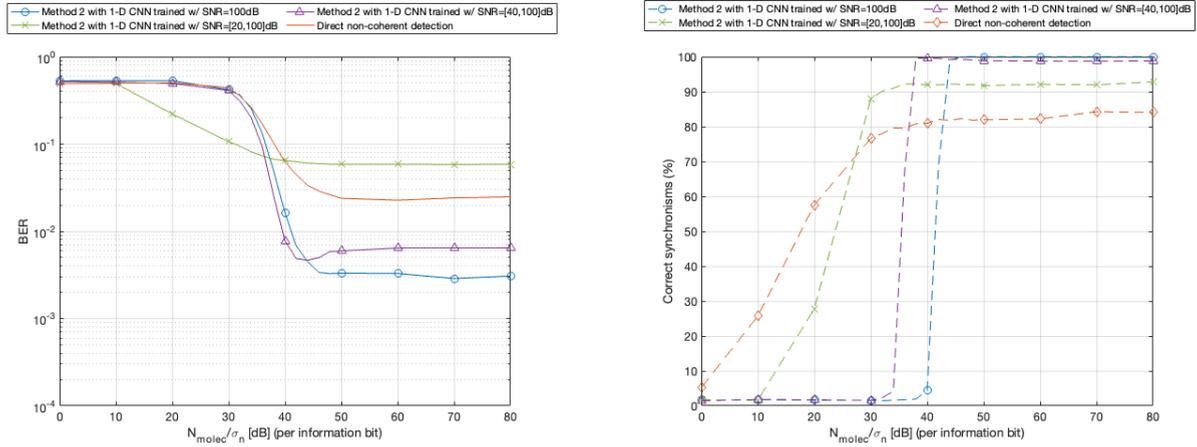
b) Synchronisation rate comparison between the NNs and the direct non-coherent detection for each method, considering variations in the distance and in the diffusion coefficient.

FIGURE 14. NNs applied to each synchronisation and detection method, considering variations in the distance and in the diffusion coefficient.

NNs are trained without considering variations in the distance and in the diffusion coefficient.

To provide further analysis regarding the behaviour of the proposed NNs under noisy conditions, in the following results we evaluate the impact of the type of data used for training, namely how the level of noise present on the training sequences influences the performance of the proposed NN based synchronisation and detection schemes. Therefore, figure 15 presents the results of the proposed 1-D CNN trained with datasets covering different ranges of SNR values. For this evaluation, and considering the previous results we selected method 2 since, while it exhibits slightly worse performance in terms of BER, it presents a higher correct

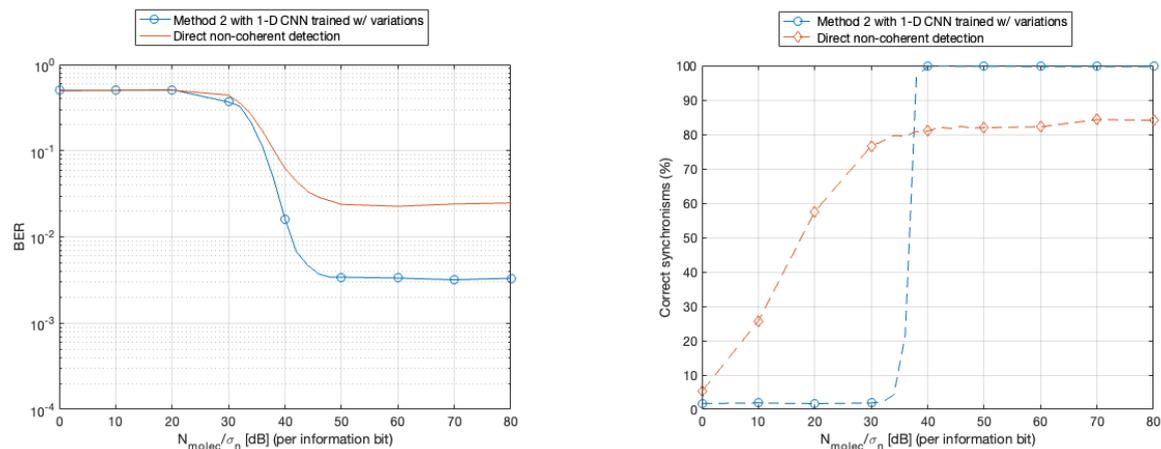
synchronisation rate when compared to method 3. Moreover, method 2 relies on a single NN for joint synchronisation and detection, making it less computationally complex than method 3, which relies on two separate NNs. In figure 15.a), we can observe that the NN trained with a SNR value of 100 dB exhibits better performance, achieving lower BER values than the other considered NNs. In comparison, the NNs trained with a large interval of SNR values achieve lower BER values for lower  $N_{molec}/\sigma_n$  values. However, for higher  $N_{molec}/\sigma_n$  values, the BER floor value for these NNs becomes higher in comparison with the NN trained with a single SNR value of 100 dB. Moreover, as it can be observed in figure 15.b), the NN trained with a SNR value



a) BER comparison between the NNs trained with different SNR values and the direct non-coherent detection, considering variations in the distance and in the diffusion coefficient.

b) Synchronisation rate comparison between the NNs trained with different SNR values and the direct non-coherent detection, considering variations in the distance and in the diffusion coefficient.

FIGURE 15. 1-D CNN trained with different SNR values applied to method 2, considering variations in the distance and in the diffusion coefficient.



a) BER comparison between the NN and the direct non-coherent detection, considering variations in the distance and in the diffusion coefficient in a time-variant channel.

b) Synchronisation rate comparison between the NN and the direct non-coherent detection, considering variations in the distance and in the diffusion coefficient in a time-variant channel.

FIGURE 16. 1-D CNN applied to method 2, considering variations in the distance and in the diffusion coefficient in a time-variant channel.

of 100 dB achieves a correct synchronisation rate of 100% for  $N_{molec}/\sigma_n$  values of 45 dB and above. In comparison, the NNs trained with an interval of SNR values achieve higher synchronization rates for lower  $N_{molec}/\sigma_n$  values. However, for higher  $N_{molec}/\sigma_n$  values, these NNs achieve lower synchronization rates than that achieved by the NN trained with a SNR value of 100 dB. These results seem to suggest that while training the NN using some sequences corrupted by higher levels of noise may provide some additional robustness when operating in worse conditions, it can severely limit the performance that can be achieved in more favourable scenarios.

All the results presented previously consider a time-invariant molecular channel. However, to assess the

robustness of the proposed approach, we extend this investigation for a scenario considering a time-variant molecular channel with unknown variations in the distance between the transmitter and the receiver and in the diffusion coefficient. In this case, for the time-variant channel it is assumed that the expected number of molecules absorbed at each time  $t$  by the receiver, as given by (3) and (4), will change for each release time  $\tau$ , due to random variations on the diffusion coefficient. This provides a further analysis of the impact of stochastic variations in the channel on the performance of the proposed NN-based non-coherent detection and synchronisation schemes. As already mentioned, given that method 2 is less complex than method 3, we simulate the results for the considered

1-D CNN applied to method 2, properly trained considering a time-variant molecular channel with unknown variations in the distance and in the diffusion coefficient. As can be seen in figure 16.a), the considered NN still exhibits lower BER values when compared with the direct non-coherent detection. Moreover, as illustrated in figure 16.b), we can observe that the proposed 1-D CNN applied to method 2 achieves a synchronisation rate of 100% for  $N_{molec}/\sigma_n$  values of approximately 40 dB and above, therefore providing more robustness than the conventional direct approach.

## V. CONCLUSION

Molecular Communication has the potential to be at the heart of IoBNT due to its revolutionary capabilities that transcend traditional communication systems. However, the unpredictable nature of the molecular channel induces environment noise and causes high ISI interference, making these two of the main challenges to overcome regarding this type of communication. To help address these problems, in this paper we propose a low complexity 1-D CNN comprising dilated causal convolutional layers and a GRU-RNN based approach, both aimed at achieving non-coherent detection and synchronisation in MC receivers, in order to increase the robustness of the molecular transmission.

Initially, we described these data-aided NNs-based approaches for accomplishing detection, assuming that the transmitter and the receiver were previously synchronised. For this scenario, it was observed that both of the proposed NNs perform significantly better when compared with a direct non-coherent detection, with the GRU-RNN having lower complexity and a slightly better performance when compared with the 1-D CNN.

Subsequently, we have extended the two NN architectures for scenarios where there is no prior synchronisation between the transmitter and the receiver, proposing two different synchronisation methods. In this part of our study, we have concluded that the implementation of the NNs for the synchronisation process increases the correct synchronisation rate, thus adding greater robustness to the molecular transmission. Moreover, comparing the use of a single NN for joint synchronisation and detection against the use of two separate NNs for synchronisation and detection, it was observed that the later achieves a better performance in terms of BER. However, this is not the case when considering the synchronisation rate, which is very similar for both methods.

Finally, in this work, we also investigated the behaviour of the proposed NNs for scenarios where there are unknown variations in the distance and in the diffusion coefficient. Given this more challenging setting, it was observed that the proposed NN-based non-coherent detection and synchronisation schemes can provide more robustness than a conventional non-data aided approach.

## REFERENCES

- [1] W. Guo, M. Abbaszadeh, L. Lin, J. Charnet, P. Thomas, Z. Wei, B. Li, and C. Zhao, "Molecular physical layer for 6G in wave-denied environments," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 33–39, May 2021.
- [2] I. F. Akyildiz, M. Pierobon, S. Balasubramaniam, and Y. Koucheryavy, "The Internet of Bio-Nano Things," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 32–40, Mar. 2015.
- [3] N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo, "A comprehensive survey of recent advancements in molecular communication," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1887–1919, 3rd Quart., 2016.
- [4] Y. Huang, F. Ji, Z. Wei, M. Wen, and W. Guo, "Signal detection for molecular communication: Model-based vs. data-driven methods," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 47–53, May 2021.
- [5] J. M. Jornet and A. Sangwan, "Nanonetworking in the terahertz band and beyond," *IEEE Nanotechnol. Mag.*, vol. 17, no. 3, pp. 21–31, Apr. 2023.
- [6] T. Nakano, M. J. Moore, F. Wei, A. V. Vasilakos, and J. Shuai, "Molecular communication and networking: Opportunities and challenges," *IEEE Trans. Nanobiosci.*, vol. 11, no. 2, pp. 135–148, Jun. 2012.
- [7] L. Chouhan and M.-S. Alouini, "Interfacing of molecular communication system with various communication systems over Internet of Every Nano Things," *IEEE Internet Things J.*, vol. 10, no. 16, pp. 14552–14568, Aug. 2023.
- [8] H. Zhai, Q. Liu, A. V. Vasilakos, and K. Yang, "Anti-ISI demodulation scheme and its experiment-based evaluation for diffusion-based molecular communication," *IEEE Trans. Nanobiosci.*, vol. 17, no. 2, pp. 126–133, Apr. 2018.
- [9] M. S. Kuran, H. B. Yilmaz, I. Demirkol, N. Farsad, and A. Goldsmith, "A survey on modulation techniques in molecular communication via diffusion," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 7–28, 1st Quart., 2021.
- [10] S. Lotter, L. Brand, V. Jamali, M. Schäfer, H. M. Loos, H. Unterwieser, S. Greiner, J. Kirchner, C. Alexiou, D. Drummer, G. Fischer, A. Buettner, and R. Schober, "Experimental research in synthetic molecular communications—Part I," *IEEE Nanotechnol. Mag.*, no. 3, 2023.
- [11] V. Jamali, N. Farsad, R. Schober, and A. Goldsmith, "Non-coherent detection for diffusive molecular communication systems," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2515–2531, Jun. 2018.
- [12] M. Bartunik, J. Kirchner, and O. Keszocze, "Artificial intelligence for molecular communication," *Inf. Technol.*, vol. 65, nos. 4–5, pp. 155–163, Aug. 2023, doi: 10.1515/itit-2023-0029.
- [13] N. Farsad and A. Goldsmith, "Neural network detection of data sequences in communication systems," *IEEE Trans. Signal Process.*, vol. 66, no. 21, pp. 5663–5678, Nov. 2018.
- [14] M. Bartunik, O. Keszocze, B. Schiller, and J. Kirchner, "Using deep learning to demodulate transmissions in molecular communication," in *Proc. IEEE 16th Int. Symp. Med. Inf. Commun. Technol. (ISMICT)*, May 2022, pp. 1–6.
- [15] S. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," *Proc. IEEE*, vol. 105, no. 12, pp. 2295–2329, Dec. 2017.
- [16] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444.
- [17] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel, and R. Schober, "Channel modeling for diffusive molecular communication—A tutorial review," *Proc. IEEE*, vol. 107, no. 7, pp. 1256–1301, Jul. 2019.
- [18] M. Kuscü, E. Dinc, B. A. Bilgin, H. Ramezani, and O. B. Akan, "Transmitter and receiver architectures for molecular communications: A survey on physical design with modulation, coding, and detection techniques," *Proc. IEEE*, vol. 107, no. 7, pp. 1302–1341, Jul. 2019.
- [19] S. Figueiredo, N. Souto, and F. Cercas, "Low-complexity channel codes for reliable molecular communication via diffusion," *Sensors*, vol. 22, no. 1, pp. 41–55, Dec. 2021. [Online]. Available: <https://search.ebscohost.com/login.aspx?direct=true&db=edspsc&AN=000741422400001&site=eds-live>
- [20] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1003–1012.
- [21] S. Bai, J. Zico Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," 2018, *arXiv:1803.01271*.
- [22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [23] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder–decoder approaches," 2014, *arXiv:1409.1259*.

- [24] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997. [Online]. Available: <https://search.ebscohost.com/login.aspx?direct=true&db=cmedm&AN=9377276&site=eds-live>
- [25] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014, *arXiv:1412.3555*.



**NUNO M. B. SOUTO** (Senior Member, IEEE) received the Graduate, Ph.D., and Habilitation degrees in aerospace engineering from the Avionics Branch, Instituto Superior Técnico, Lisbon, Portugal, in 2000, 2006, and 2021, respectively. From November 2000 to January 2002, he was a Researcher in the field of automatic speech recognition with the Instituto de Engenharia e Sistemas de Computadores, Lisbon. He joined the Iscte–University Institute of Lisbon, as an Assistant Professor, in 2006. He has been a Researcher with the Instituto de Telecomunicações (IT), Portugal, since 2002, and has been involved in several international research projects and many national projects. His research interests include wireless networks, signal processing for communications, multicarrier communications, frequency domain equalization, channel coding, modulation, channel estimation, synchronization, MIMO schemes, wireless sensor networks, unmanned aerial vehicles, mmWave/terahertz communications, machine learning for wireless communications, and molecular communications. He is a member of the IEEE Communications Society, the IEEE Signal Processing Society, and the IEEE Instrumentation and Measurement Society.



**JOÃO C. SILVA** received the B.S. degree in aerospace engineering from the Instituto Superior Técnico (IST), Lisbon Technical University, in 2000, and the M.B.A. degree from ISEG, in 2015. He finished his Ph.D. thesis at IST, in 2006, focusing on spread spectrum techniques, multiuser detection schemes, and MIMO systems. From 2000 to 2002, he was a Business Consultant with McKinsey & Company. Since 2006, he has been working on computer networks as an Assistant Professor with ISCTE-IUL.

...



**DUARTE CASALEIRO** received the degree in telecommunications and computer engineering from the Iscte–Instituto Universitário de Lisboa, Lisbon, Portugal, in 2022, where he is currently pursuing the M.S. degree in telecommunications and computer engineering, with specialization in the Internet of Things. His research interests include molecular communication and machine learning for wireless communications.