IEEE Communications Magazine, Volume 59, Issue 5, May 2021, pp. 47-53
DOI:10.1109/MCOM.001.2000957



# Signal Detection for Molecular Communication: Model-Based vs. Data-Driven Methods

Yu Huang, Fei Ji, Zhuangkun Wei, Miaowen Wen, and Weisi Guo

***Abstract*—Multi-scale molecular communication (MC) employs the characteristics of information molecules for information exchange. The received signal in MC inevitably encounters severe inter-symbol interference and signal-dependent noise due to the stochastic diffusion mechanism. Focusing on the critical signal detection in MC, first this article reviews the commonly used model-based detectors, and exposes their limitations in practical implementation. Then, the emerging data-driven detectors that can make up for some deficiencies of the model-based detectors are presented. Despite the black-box nature of the data-driven detectors, the explainable artificial intelligence can be further investigated for the performance improvement of transparency and trust.**

## I. Introduction

Molecular communication (MC) is an emerging multi-scale communication paradigm that conveys messages via the patterns of molecules [1]. MC has been regarded as an alternative for conventional communication techniques in virtue of its unique transmission mechanism.

For the nano-scale and micro-scale channels *in vivo*, the implementation of the conventional communication techniques is hindered for the following reasons. The electromagnetic (EM) wave based wireless communication has its constraint in wave generation. As the wavelength in operation is proportional to the antenna size, nanomachine requires the Tera-hertz band for communications, which is, however, still in its research infancy, let alone the micro-scale implementation. Besides, optical communication is not an appropriate choice due to complex cell/tissue obstacles in the channel. Finally, acoustic communication has an energy issue on such an extremely tiny scale. In addition to the implementation restrictions, the bio-compatibility issue of the above schemes has not been verified in the open literature. As for the macro-scale channels, the EM wave based communication has the propagation problem in metal and aqueous environments. Inspired by the biological systems, MC can remedy the flaws in the existing communication systems. For micro-scale channels, MC via diffusion (MCvD) is an energy-efficient signal transmission mechanism that uses information molecules from the communication environments, and no extra energy is required for the carrier propagation thanks to the Brownian motion. During the coronavirus disease pandemic, MC has been recognized as a potential model for the viral aerosol transmission and detection process [2], providing a deeper understanding of the virus study.

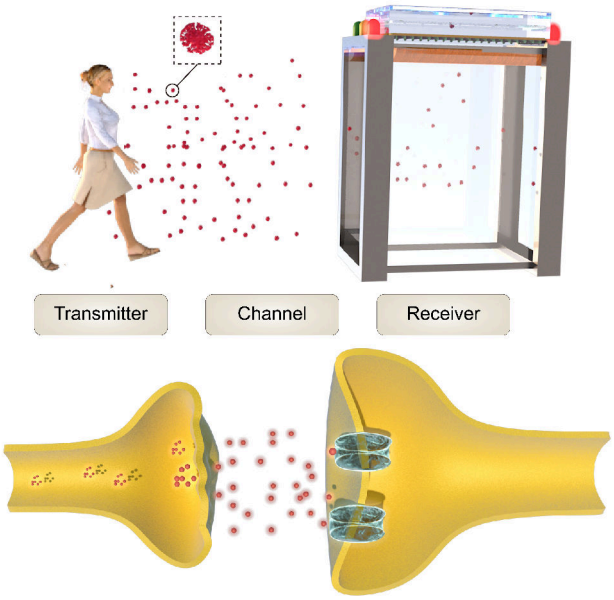


Fig. 1. MC block diagram under different scales.

## II. System Model and Signal Characteristics

Typically, an MC system model consists of three fundamental blocks, i.e., transmitter, channel, and receiver. The transmitter first modulates the digital signal into molecular forms. The information molecules then enter the channel, undergoing the mechanisms of diffusion, advection, or reaction, etc. The receiver senses molecular concentration or particle number as the received signal. Figure 1 depicts both the macro-scale and micro-scale MC systems, where the environments can be nonlinear and turbulent in practice. The macro-scale virus-laden aerosol detection is shown in the upper part of Fig. 1, where the infected/uninfected human is treated as the transmitter, and the breath, cough, and sneeze processes can emit virus into the air (channel) through the aerosol transmission mechanism. A specific receiver with viral detection capability is placed within a certain distance to decide whether the source has the infection or not based on the received signal [2]. Alternatively, the lower part of Fig. 1 demonstrates the micro-scale neurotransmission mechanism at the synapse. In this process, the axon terminal acts as the transmitter, releasing the neurotransmitters into the synaptic cleft channel. After a given

Y. Huang is with the School of Electronics and Information Engineering, South China University of Technology, Guangzhou 510641, China, and also with the School of Aerospace, Transport and Manufacturing, Cranfield University, Bedford MK43 0AL, U.K. (e-mail: ee06yuhuang@mail.scut.edu.cn).

F. Ji and M. Wen are with the School of Electronics and Information Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: eefeiji@scut.edu.cn; eemwwen@scut.edu.cn).

Z. Wei is with the School of Engineering, University of Warwick, Coventry CV4 7AL, U.K. (e-mail: zhuangkun.wei@warwick.ac.uk).

W. Guo is with the School of Aerospace, Transport, and Manufacturing, Cranfield University, Milton Keynes MK43 0AL, U.K., and also with the Alan Turing Institute, London NW1 2DB, U.K. (e-mail: weisi.guo@cranfield.ac.uk).

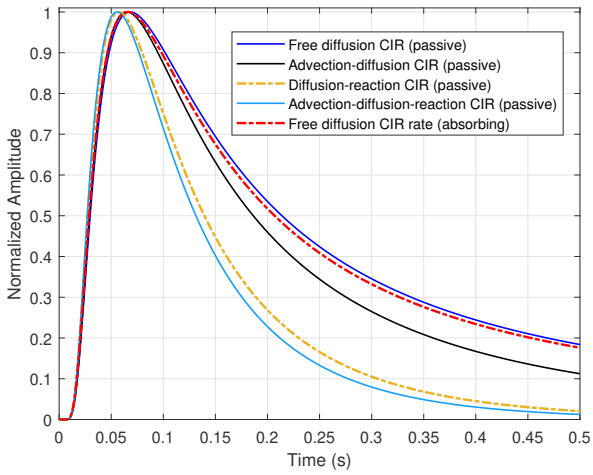


Fig. 2. CIRs in response to different conditions of the diffusive channels, where the propagation distance, spherical receiver radius, and diffusion coefficient are 30μm, 0.5μm, and $2.2 \times 10^{-9}$m/s$^2$, respectively. The flow velocity is 100μm/s, and the reaction rate constant is 5 s$^{-1}$.

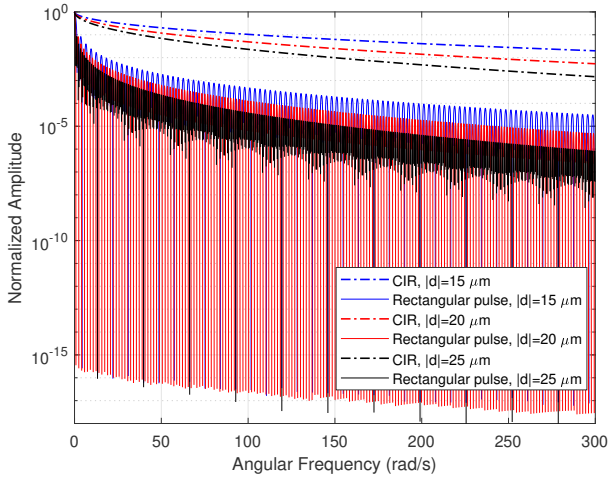


Fig. 3. Frequency response of the one-shot modulated signal with the impulsive release (CIR) and rectangular pulse in the free diffusion MC channels, where $D = 2.2 \times 10^{-9}$m/s$^2$.

period, some of the neurotransmitters bind to and activate the receptors of the postsynaptic membrane.

### A. Time-Domain Signal

Different effects in the channel are reflected by the variation of the channel gains and the noise model [1]. Moreover, the corresponding channel impulse response (CIR) shares common transient feature [3]. Figure 2 demonstrates the CIRs and the CIR rate with normalized amplitude associated with various channel/receiver models. Intuitively, they all have an increasing trend initially, and a following the downturn with different rates after passing the peak value. Therefore, for the consecutive transmission with a given symbol duration, the received signals in response to different transmit symbols can overlap with each other, known as the inter-symbol interference (ISI) phenomenon [1]. This article focuses on the MC via diffusion (MCvD) system with the three-dimensional channel, which is typically considered in the literature. From the observations of Fig. 2, the received signal in the free diffusion channel is more sensitive to ISI compared with the other channel conditions. Besides, the received signal in the free diffusion channel encounters the signal-dependent counting noise, which is different from the conventional communication systems, where the signal and noise are usually independent [1].

### B. Frequency-Domain Analysis

As the time-domain study provides an intuitive expression, the research of MC had focused on the time-domain signal processing, while the frequency-domain analysis was neglected, which can bring new insights into the MC system design. The CIR in the frequency domain manifests its low-pass nature as the amplitude components monotonically decrease along the frequency axis. However, just like the time-domain CIR that lasts for infinite time, the frequency-domain CIR occupies the whole frequency range. The frequency response for the modulated signals in the free diffusion channel is shown in Fig. 3, where the receiver radius is assumed to be far less than the transmission distance. Explicitly, the amplitude of the frequency response monotonically decays with the increasing angular frequency. Besides, the longer transmission distance leads to a more rapid decline rate of the amplitude component.

## III. Model-Based Detection Schemes

Signal detection plays an important role in MC for information recovery. This section reviews some fundamental model-based detection schemes. Generally, the transmitter is modeled by a point source that can emit the information molecules impulsively, and a passive spherical receiver can perfectly sense the received signal. Also, perfect channel estimation is assumed to be available in coherent detection schemes. More details about some of the mentioned signal detection schemes are provided by a recent survey for MC in [4].

### A. Maximum Likelihood Sequence Detection

Due to the existence of ISI, the sequence-wise detector outperforms its symbol-wise counterparts as the information of previously transmitted symbols are taken into account. The MLSD can achieve desirable detection performance, while the complexity of its brute-force search grows exponentially with the sequence length. The Viterbi can reduce search complexity by discarding the unlikely sequences, while complexity is still exponentially dependent on the number of channel taps, which is unrealistic when the transmission rate is high [5]. The MLSD metric in MC can be different when the signal-dependent counting noise is considered. Despite its impracticality in implementation, MLSD can be used as the benchmark for evaluating the existing detection schemes in terms of the error performance bound. Alternatively, the reduced-state sequence estimation (RSSE) for MC was mentioned in [6] by means of numerical simulation. When the perfect knowledge of channel

gains is unavailable at the receiver, the RSSE with the hard decision of the previously transmitted symbol can counter-intuitively outperform the MLSD.

### B. Time-Domain and Frequency-Domain Equalizers

Time-domain equalization (TDE) shows lower complexity and attains a sub-optimal detection performance compared with MLSD, while it has the following limitations. On the one hand, TDE is inherently infeasible for the high transmission rate scenario as its computational complexity of TDE is proportional to the ISI length. On the other hand, the non-stationary noise further complicates the TDE with the minimum mean square error (MMSE) criterion. First, the tap coefficients vary symbol by symbol due to the signal-dependent noise feature [5], and such computational complexity is prohibitively large for the nanomachine receiver. Second, the calculation of tap coefficients depends on the previously detected bits. Hence, the error propagation problem may exist, deteriorating the error performance. For the consistency issue, the current symbol to be detected is unknown at the receiver. However, as the noise power of the current symbol is required, the current symbol may be estimated according to the statistical distribution of the input symbol. Analogously, the noise power of the received signal after the symbol to be detected is unavailable, hindering the collection of its multipath copies.

Inspired by the previous frequency-domain analysis, the signal detection, pulse shaping, and channel estimation techniques can be studied from a new perspective. Against this background, frequency-domain equalization (FDE), recognized for its low-complexity implementation, can be deployed in MC to solve the implementation complexity problem of some time-domain detectors when the number of channel taps is large. MC can fully use the whole bandwidth without any division, as there is no limitation or regulation. Against this background, the transmission of MC in the frequency domain is the single-carrier system. Correspondingly, the FDE in MC is inherently a single-carrier FDE that has been widely studied in conventional wireless communication systems. The principle of FDE in MC is illustrated as follows. The transmitter side first packs the symbols for transmission in one block and inserts redundant guard interval with a proper length at the end of the block. The new data block can eliminate the inter-block interference (IBI), and it is converted into the molecular form according to the modulation scheme. The receiver senses the continuous-time molecular signal, sampled as the form of the discrete-time digital signal. After removing the guard interval signal, the receiver deploys the discrete Fourier transform (DFT) operation, converting the time-domain signal to the frequency domain. Then, the FDE, dealing with the complex signal, is implemented to combat the ISI effect based on specific criteria. In response to the FDE output, the inverse DFT operation transforms it back to the time domain. Finally, the detection techniques recover the data stream according to the equalized signal, which preserves its original property, i.e., being real and positive.

In the MC over free diffusion channels, zero-forcing (ZF) FDE has the same form as in previous research as the noise components are neglected. Unlike the ZF-FDE, the design of MMSE-FDE depends on the power of counting noise, while its signal-dependent feature indicates the variation of noise power in each block. In this case, the noise power estimation becomes a critical issue in MMSE-FDE. A statistical approach based on the distribution of input symbols can easily estimate the noise power. Furthermore, the output signal in response to this statistical estimation can be used as either the output directly or the feedback for advanced noise power estimation in an iterative manner for detection performance enhancement. In light of this, the MMSE-FDE can determine the iteration times concerning the application requirement and the device capability. Despite its advantages, FDE requires additional calculation for the DFT and its inverse operation. Besides, the guard interval also decreases the transmission efficiency. Finally, the whole sequence is decoded after the block transmission in FDE due to its block-wise detection procedure, hindering the real-time applications.

### C. Non-Coherent Detectors

The fixed threshold detector is one of the simplest detection methods in MC, and it is effective only when a sufficiently long symbol duration is provided. For high transmission rate, the severe ISI exacerbates the detection. Alternatively, the detector with the optimal threshold can enhance the error performance, while it requires the knowledge of previously transmitted symbols, which can be impractical due to the error propagation phenomenon [7].

However, whether the threshold is fixed or optimal, the CIR knowledge is indispensable for both schemes. Alternatively, some model-based detectors avoid this requirement by using the features of the received signal. Note that they are referred to as the non-coherent detectors, dispensing with the CIR knowledge. An adaptive threshold detector was proposed in [6], which exploits the ISI effect for signal detection, and determines the threshold with previously sensed signal. Thus, the decision rule is simply the comparison between the currently sensed signal and the threshold. The characteristics of the MC signal can be further elaborated as follows. When the symbol duration is properly set in MCvD (greater than the peak concentration time), the received signal in response to the bit-1 transmission leads to a conspicuous rising edge, and its sampled mean is greater than the expectation of the previously transmitted signal. On the contrary, the bit-0 transmission shows the opposite trend. In this case, the channel estimation process can be avoided, which enhances the communication efficiency. Except for the rising edge, the MC signal also exhibits features like the inflexion and energy difference. These metrics constitute a high-dimensional form of the received signal, which converts the original signal into the new domain insensitive to the ISI. Nevertheless, such a high-dimensional signal detector requires multiple samplings in each symbol duration for signal construction [8].

### D. Comparison between Coherent and Non-Coherent Detectors

Table I summarizes the features of the aforementioned detection techniques. The performance comparison between

TABLE I
COMPARISON ANALYSIS BETWEEN DIFFERENT DETECTORS

| Detector Type | Detection Performance | Complexity | Memory Requirement | Guard Interval/Transformation | Channel Information |
|---|---|---|---|---|---|
| Adaptive Threshold | Low | Low | Only One | Not Required | Not Required |
| TDE | Medium | Relatively High | ISI Length | Not Required | Required |
| FDE | Medium | Medium | Block Length | Required | Required |
| MLSD with Viterbi algorithm | High | High | ISI Length | Not Required | Required |

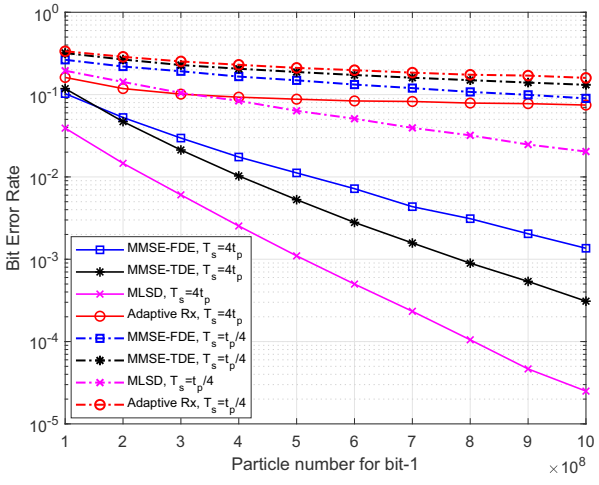


Fig. 4. Comparison of BER performance between the time-domain and frequency-domain detectors, where $T_b$ denotes the symbol duration and $t_p$ is the arrival time of the peak concentration signal. The propagation distance, spherical receiver radius, and diffusion coefficient are 20μm, 0.1μm, and $2.2\times 10^{-9}$m/s$^2$, respectively.

the aforementioned detectors is demonstrated in Fig. 4, where low and high transmission rates are considered according to the relationship between the symbol duration $T_s$ and the peak concentration time $t_p$. Note that all the detectors have the perfect channel knowledge at the receiver side. In both cases, the adaptive threshold detector, whose performance is comparable to that of the fixed threshold detector [6], demonstrates the worst performance due to its low complexity implementation, while the MLSD prevails among all the schemes with its involved search complexity, which grows exponentially with the number of the channel taps. In the low transmission rate scenario, the TDE has a better performance than the FDE, while the FDE, in turn, outperforms TDE when the transmission rate is high. Therefore, one may choose between them according to the transmission rate for attaining a satisfactory detection performance.

### *E. Problems in Model-Based Detectors*

On the one hand, the coherent model-based detectors highly depend on the CIR knowledge. Though MLSD reaches the near-optimal performance with perfect CIR knowledge, a slight estimation error can lead to severe error performance degradation, which is demonstrated in [9]. Besides, both FDE and TDE face the same problem. The CIR can be theoretically calculated only within some specific conditions with the known system parameters to derive the corresponding analytical performance; for practical implementations, the channel model may not be analytically available, whereas the time-varying property further complicates the CIR calculation issue. Thus, the channel estimation procedure is required to track and provide the estimation CIR, typically the channel gains, regardless of the channel models or the system parameters. For instance, the training sequence known by the transceivers can be used for channel estimation, which occupies the extra overhead, while such estimated CIR is usually imperfect, deteriorating the detection performance [10].

On the other hand, the non-coherent model-based detectors have undesirable error performance, especially in the high transmission rate case, which lays the foundation for useful MC applications. They become unreliable as either the difference between the consecutive transmission diminishes [6], or the detection metrics are lost [8]. Besides, the non-coherent schemes generally have worse performance than the coherent detection schemes.

## IV. DATA-DRIVEN DETECTORS

Obviously, the model-based detectors have either the CIR knowledge or the detection performance problems. Against this background, the data-driven detectors have been presented to overcome the limitations in the model-based detector. Typically, the machine learning based algorithms are adopted by the data-driven detectors. Depending on the learning methods, machine learning has three fundamental categories, namely, supervised learning, unsupervised learning, and reinforcement learning.

### *A. Channel Model Issues*

The channel model is critical for both analytical and practical MC systems. The analytical performance evaluation provides insights for the model-based detectors in MC, which usually considers perfect CIR knowledge with simple initial/boundary conditions, such as the impulsive emission, open space, or perfectly transparent/absorbing behaviors. These assumptions are unique in MC. The complex boundaries and environments in realistic MC channels make the analytical models intractable. For instance, the multiple-input multiple-output (MIMO) MC is one example that exploits the spatial domain to achieve diversity or multiplexing, while its channel model becomes elusive when multiple absorbing receivers are deployed at the receiver side. To obtain a channel model in MIMO-MC, the analytical channel model in the single-input single-output (SISO) MC is treated as the empirical form of the MIMO-MC channel model, whereas certain parameters are further introduced for the adjustment purpose. Based on the potential model knowledge, the artificial neural network (NN)

is trained to find the input-output relationship, revealing the approximated form of the corresponding channel model [11].

Alternatively, the channel of realistic MC prototype is generally far from that in the analytical study. On the one hand, the channel changes due to the change of channel parameters (e.g., temperature fluctuation), transceiver movement, or the effect of stochastic turbulence. On the other hand, the quantification of the boundary conditions can be difficult for certain channels, such as the in-vessel environment. The data-driven detectors are useful to deal with the MC system with unknown CIR as they can learn to detect the received signals regardless of the channel model and its parameters.

### *B. Comparison with Model-Based Schemes*

Two data-driven detectors were proposed for the vessel-like prototype in [12] with unknown channel model. First, the support vector machine performs the symbol-by-symbol detection, which is trained with current received signal and transmitted symbols, regardless of the ISI. Second, the recurrent NN (RNN) takes the ISI into account, acting like the sequence detection. Additionally, inspired by the signal feature from human observation, the model-based slope detector has been proposed, which possesses the non-coherent property. From the bit error rate (BER) of the experimental results, the RNN prevails among these detectors, while the model-based slope detector exhibits the worst performance. The RNN is inherently feed-forward, which means that it only uses the currently and previously received signal for sequence detection. Nevertheless, the received signal after the detection symbol slot may carry its information in the form of ISI. In light of this, the bidirectional RNN (BRNN) introduces the feed-backward architecture, considering more useful signals to improve the detection performance. The BRNN, however, suffers from the IBI that affects the sequence detection performance. While the insertion of guard interval accounts for the extra overhead, a sliding window in the BRNN can circumvent this problem, enhancing the detection performance through a soft decision manner [9]. Though the sliding BRNN is inferior to the MLSD with perfect CIR knowledge, it outperforms the MLSD with imperfect knowledge (2.5% and 5% deviation) in terms of the error performance.

The comparative studies of the error performance between the model-based detectors and the data-driven detectors have been also presented concerning the analytical channel models, where different learning strategies are used. First, an Artificial NN architecture is trained by the supervised learning method with labeled inputs and outputs. After the training process, it acts as the receiver with BER minimization regardless of the CIR knowledge. On the contrary, the model-based detector finds the optimal threshold with perfect CIR knowledge. Both detectors exploit the prior estimated symbol, and one can observe a good agreement between them [7]. Another research uses the unsupervised method to classify different transmitted symbols according to the characteristics of the received signals [13]. Note that it surpasses conventional model-based schemes in terms of error performance, except for the MLSD with perfect CIR knowledge. Besides, it requires much lower computational complexity than the MLSD, and neither the CIR knowledge nor the labeling is needed. While the previous studies merely focus on the detector design at the receiver side, the transmitter can also be considered for the joint transceiver optimization. In this case, the transmitter and receiver can be trained by the deep reinforcement learning and supervised learning, respectively [14]. Finally, the results based on the joint optimization display a better BER performance than the model-based threshold detection.

The training phase is essential and vital for the data-driven detectors with supervised learning, which determines the performance of the NN. Despite its similarity to the pilot sequence, the training phase can be done in an offline manner, which has three remarkable advantages below [9]. First, it avoids the periodical channel estimation process, which accounts for the extra overhead. Second, when the big data, i.e., the input-output training sequences over many channel scenarios, is provided, the trained receiver acquires the resistance capability to the time-varying channels. Finally, the macro-scale devices can be used to facilitate the complex offline training, and then transplant it into the micro-scale MC systems with less computing power.

### *C. A Pathway Towards Explainable Artificial Intelligence*

Unlike the model-based detectors that have intuitive blocks, the data-driven detector typically lacks the transparency and trust due to its black-box nature, which is obscure for human. This loss of explainability suggests an increasing vulnerability to noisy and even malicious inputs. To overcome this, several definitions and methods to measure and improve the explainability are proposed, known as the explainable artificial intelligence (XAI). These include but are not limited to intuitive visualization, symbolic representations, and Neural Network compression.

**Visualization** is referred to as an intuitive evaluation (watch) of a neuron's features. In the context of MC (as is shown by Fig. 5), we highlight one neuron's output that resembles the rising-edge shape of the molecular signal (when corresponds to a '1' bit), and this indicates an intuitive explainability of this trained neuron. One difficulty lies in that with the increasing depth of the hidden layers, the non-linearity increases and will prevent the intuitive features from being noticed.

**Symbolic representation** is to find a rigorous equation for the representation of the Deep NN (or locally one neuron). Current approaches focused on using the general hyper-geometric function (e.g., the Meijer-G functions and Gaussian process approximations). In the example of Fig. 5, one neuron's output is verified by the finding of a corresponding hyper-geometric function. As such, the outputs are to some extent guaranteed, as no malicious outputs will be generated out of the range of the representation function.

An **NN compression** can be pursued by removing the neurons with trivial weights and in-explainable results (shown by Fig. 5). This will undoubtedly scarify the accuracy, but help derive a more trustworthy and energy-efficient NN, which is important for the nano-scale system with limited energy

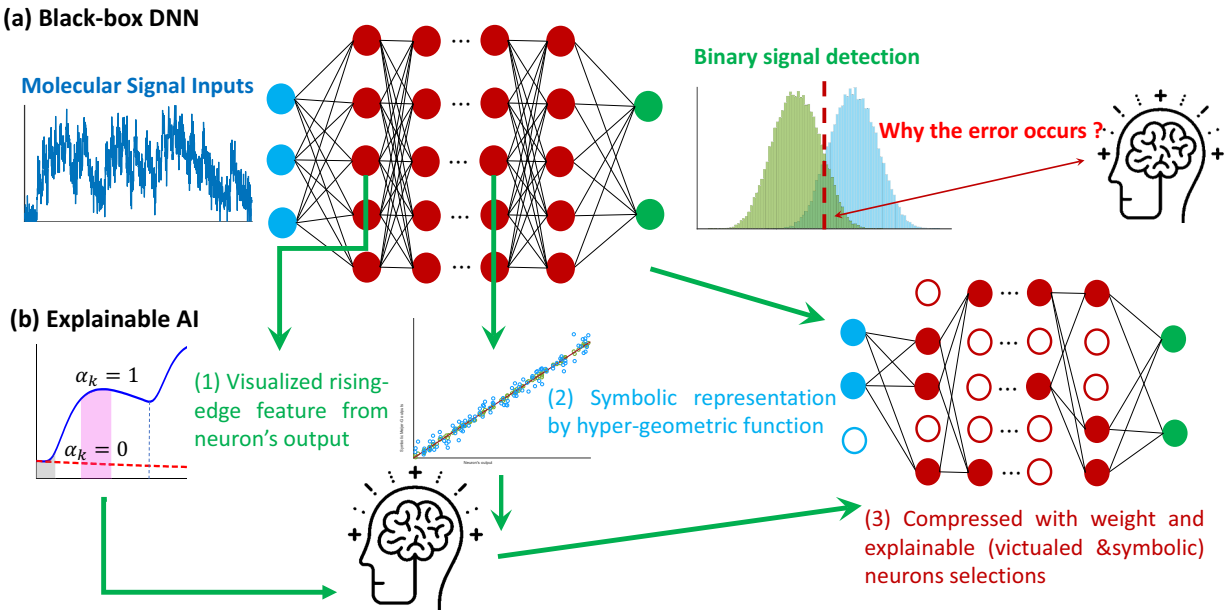


Fig. 5. XAI for the interpretation and refinement of black-box DNN for molecular signal detection.

and resources. Besides, the resulting decisions can be more likely in the way that a human expert can understand, which is vital for the MC applications *in vivo*, such as the target drug delivery.

## V. Open Challenges and Future Directions

There are still some open challenges for signal detection of the MC receiver, which are discussed below. Meanwhile, some future research directions are also elaborated in this section.

### A. Signal Detection with Complex Channel Conditions

The model-based detectors usually avoid the channel variation issue by assuming perfect channel estimation or using the static channel model. Nevertheless, the release manner can be spatio-temporal, the channel can be highly dynamic, and the reception mechanism can be nonlinear in practice. Against this background, the feasibility of the model-based detectors for practical MC systems requires further validation. For instance, in the mobile environments, the training symbols may occupy a longer time than the channel coherence time due to its slowly diffusive manner. In this circumstance, the detection scheme should be robust to the inaccuracy of the estimated channel gain. The non-coherent detection scheme without channel estimation is more appropriate [15], where the signal feature is captured for threshold update.

Besides, the data-driven detectors can overcome this issue through the offline training with sufficient data. However, the access to big data remains challenging. Alternatively, current data-driven detectors in MC are in the static state as once being trained, the architecture and weights stay fixed during the detection process. However, it cannot deal with the situation when undergoing the new channel beyond the training sets. Therefore, inspired by the channel estimation in the model-based methods, the data-driven detectors should evolve with an update mechanism.

### B. Inspiration From Existing Communication Systems

MC has both similarities and differences with the conventional communication techniques. Despite its macro-scale applications that can resort to the equipment with high computing power like other communication systems. The nanomachine is an exception in MC, which is deployed in the cellular biological environments. Due to the size constraint, some communication techniques, such as data-driven detection algorithms, may not be implanted into the nanomachine. Nevertheless, some biological mechanism may kindle the detection scheme for artificial MC systems. As a consequence, the nano-network mimics the cell signaling network to realize the nano-scale or micro-scale communication. The model-based adaptive threshold detection [6] is one of the examples inspired by the short-term synaptic plasticity of the neurotransmission. Besides, the techniques developed in the mature communication system also facilitate the design of signal detection schemes in MC, where the model-based MLSD, TDE, and FDE are inspired by those in wireless communications. Except for this, some optical communication techniques may be easily transplanted into the MC system as the signals in both systems are real and positive [9].

## VI. Conclusions

MC has versatile applications in multiple scales that require specific demands. Therefore, multiple types of model-based detectors were proposed with different error performance and computational complexity. The data-driven detectors were presented to overcome the high dependency on the CIR knowledge of the coherent model-based detectors and the poor error performance of the non-coherent model-based detectors. The ideology of XAI was envisioned to combat the transparency and trust issues in the data-driven detectors.

## REFERENCES


[1] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel, and R. Schober, "Channel modeling for diffusive molecular communicationA tutorial review," *Proc. IEEE*, vol. 107, no. 7, pp. 1256–1301, July 2019.

[2] M. Khalid, O. Amin, S. Ahmed, B. Shihada, and M. Alouini, "Modeling of viral aerosol transmission and detection," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4859–4873, Aug. 2020.

[3] B. Li, W. Guo, X. Wang, Y. Deng, Y. Lan, C. Zhao, and A. Nallanathan, "CSI-independent non-linear signal detection in molecular communications," *IEEE Trans. Signal Process.*, vol. 68, pp. 97–112, Jan. 2020.

[4] M. Kuscu, 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, July 2019.

[5] D. Kilinc and O. B. Akan, "Receiver design for molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 705–714, Dec. 2013.

[6] M. Damrath and P. A. Hoeher, "Low-complexity adaptive threshold detection for molecular communication," *IEEE Trans. Nanobiosci.*, vol. 15, no. 3, pp. 200–208, Apr. 2016.

[7] X. Qian, M. Di Renzo, and A. Eckford, "Molecular communications: Model-based and data-driven receiver design and optimization," *IEEE Access*, vol. 7, pp. 53 555–53 565, 2019.

[8] Z. Wei, W. Guo, B. Li, J. Charmet, and C. Zhao, "High-dimensional metric combining for non-coherent molecular signal detection," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1479–1493, Mar. 2020.

[9] 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.

[10] V. Jamali, A. Ahmadzadeh, C. Jardin, H. Sticht, and R. Schober, "Channel estimation for diffusive molecular communications," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4238–4252, Oct. 2016.

[11] C. Lee, H. B. Yilmaz, C. Chae, N. Farsad, and A. Goldsmith, "Machine learning based channel modeling for molecular MIMO communications," in *Proc. IEEE Int. Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, Sapporo, Japan, June 2017, pp. 1–5.

[12] N. Farsad, D. Pan, and A. Goldsmith, "A novel experimental platform for in-vessel multi-chemical molecular communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Singapore, Dec. 2017, pp. 1–6.

[13] S. Liu, Z. Wei, B. Li, and C. Zhao, "Unsupervised clustering-based non-coherent detection for molecular communications," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1687–1690, Aug. 2020.

[14] S. Mohamed, J. Dong, A. R. Junejo, and D. C. Zuo, "Model-based: End-to-End molecular communication system through deep reinforcement learning auto encoder," *IEEE Access*, vol. 7, pp. 70 279–70 286, 2019.

[15] X. Mu, H. Yan, B. Li, M. Liu, R. Zheng, Y. Li, and L. Lin, "Low-complexity adaptive signal detection for mobile molecular communication," *IEEE Trans. Nanobiosci.*, vol. 19, no. 2, pp. 237–248, Apr. 2020.


**Yu Huang** is a Ph. D. student at South China University of Technology. His main research interests include molecular communications, deep learning for communication systems.

**Fei Ji** received her Ph.D. degree from South China University of Technology in 1998. She is currently a full-time professor with South China University of Technology. Her research focuses on wireless communication systems and networking.

**Zhuangkun Wei** is a Ph.D. student at School of Engineering, University of Warwick. His research interests cover graph signal processing, molecular communications, and X-AI.

**Miaowen Wen** received the Ph.D. degree from Peking University, China, in 2014. He is currently an Associate Professor with South China University of Technology, China. His research interests include a variety of topics in the areas of wireless and molecular communications. He is currently serving as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE COMMUNICATIONS LETTERS.

**Weisi Guo** received his M.Eng., M.A., and Ph.D. degrees from the University of Cambridge. He is Chair Professor of Human Machine Intelligence at Cranfield University. He has published over 150 papers and is PI on over 12 research projects from EPSRC, Royal Society, EC H2020, and InnovateUK. He is a Turing Fellow at the Alan Turing Institute and a Fellow of the Royal Statistical Society.

2021-06-03

# Signal detection for molecular communication: model-based vs. data-driven methods

Huang, Yu

IEEE

Huang Y, Ji F, Wei Z, Wen M, Guo W. (2021) Signal detection for molecular communication: model-based vs. data-driven methods. IEEE Communications Magazine, Volume 59, Issue 5, pp. 47-53

https://doi.org/10.1109/MCOM.001.2000957