IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 January 2023, accepted 7 February 2023, date of publication 9 February 2023, date of current version 22 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3243797

SURVEY

Coding in Diffusion-Based Molecular Nanonetworks: A Comprehensive Survey

PIT HOFMANN[1], (Graduate Student Member, IEEE), JUAN A. CABRERA[1], RICCARDO BASSOLI[1,2], (Member, IEEE), MARTIN REISSLEIN[3], (Fellow, IEEE), AND FRANK H. P. FITZEK[1,2], (Senior Member, IEEE)

[1]Deutsche Telekom Chair of Communication Networks, Faculty of Electrical and Computer Engineering, Institute of Communication Technology, Technische Universität Dresden, 01062 Dresden, Germany
[2]Centre for Tactile Internet with Human-in-the-Loop (CeTI), Cluster of Excellence, 01062 Dresden, Germany
[3]School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287, USA

Corresponding author: Martin Reisslein (reisslein@asu.edu)

This work was supported by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) as part of Germany's Excellence Strategy–EXC 2050/1–Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI) of Technische Universität Dresden under Project ID 390696704, as well as by the Federal Ministry of Education and Research of Germany through the program "Souverän. Digital. Vernetzt." and by the Federal Ministry of Education and Research of Germany through the Joint project 6G-life, project id number: 16KISK001K.

ABSTRACT Diffusion-based molecular nanonetworks exploit the diffusion of molecules, e.g., in free space or in blood vessels, for the purpose of communication. This article comprehensively surveys coding approaches for communication in diffusion-based molecular nanonetworks. In particular, all three main purposes of coding for communication, namely source coding, channel coding, and network coding, are covered. We organize the survey of the channel coding approaches according to the different categories of channel codes, including linear block codes, convolutional codes, and inter-symbol interference (ISI) mitigation codes. The network coding studies are categorized into duplex network coding, physical-layer network coding, multi-hop nanonetwork coding, performance improvements of network-coded nanosystems, and network coding in mobile nanonetworks. We also present a comprehensive set of future research directions for the still nascent area of coding for diffusion-based molecular nanonetworks; specifically, we outline research imperatives for each of the three main coding purposes, i.e., for source, channel, and network coding, as well as for overarching research goals.

INDEX TERMS Channel coding, diffusion, Internet of Bio-Nano Things (IoBNT), molecular communications, network coding, source coding.

I. INTRODUCTION

A. MOTIVATION FOR DIFFUSION-BASED MOLECULAR NANONETWORKS

In a lecture in 1959, Richard Feynman presented ideas on how technology could work at the microscopic level, entitled 'There's Plenty of Room at the Bottom' [1], [2]. His work presents approaches to data storage, the need for better electron microscopes, the advantages of miniature computers, and the design of microscopic machines, among other topics. Later, his proposals became the origin of nanotechnology—the starting point of nanonetworks.

The associate editor coordinating the review of this manuscript and approving it for publication was Marco Martalo.

With the later IEEE P1906.1, there is a working group on nanonetworking for nanoscale and molecular communication. The IEEE P1906.1 working group [3] defines a nanonetwork as a human-designed system communicating with or at nanoscale, using physical principles that are suited to nanoscale systems. Human-designed means a system occurring as a result of conscious human intervention. These systems may include naturally occurring components in an arrangement or for a purpose that is otherwise not naturally occurring. ISO/TS 80004-2 [4] defines nanoscale dimensions as one to 100 nanometers. Nanonetworks are communications networks existing entirely or mostly at the nanometer scale; whereby, typically, thousands of nanometers separate physical nodes. The nodes are assumed to be mostly

self-powered and mobile or rapidly deployable [5]. Besides interfacing among the nanometer scale [6], [7], the interoperability of nanoscale and macroscale networks poses challenging problems.

In the past years, there has been a drive to make devices and networks smaller—-the trend towards nanoscience. Future applications arise in the context of health monitoring, biohybrid implants, human augmentation, and environmental monitoring, as well as nuclear, biological and chemical attack detection.

Diffusion-Based Molecular Communication (DBMC) is an emerging area of research in communications engineering [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18]. DBMC uses chemical signals to convey information from a transmitter to a receiver via diffusion [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36]. Compared to traditional communication systems using Electromagnetic (EM) signals, DBMC can be employed in areas where EM signals are not feasible or fail. DBMC can be used for propagating pheromones [37], [38], biomedical drug delivery [39], [40], [41], [42], [43], [44], [45], [46], or for monitoring in confined environments [47], e.g., in metallic pipes, where wireless signals will fail. Compared to EM signals, another essential benefit of DBMC is energy efficiency [48]. The data rates in DBMC are lower than with EM signals, but DBMC expends less energy per transmitted bit [49], [50]. In particular, DBMC is suitable for applications with ultra-low power levels that do not require high data rates [51].

B. RELATED SURVEYS

Several surveys have been published to date on nanonetworks, whereby all of them are complementary to our survey. Terahertz nanocommunication and networking, which exploit electromagnetic signals in the terahertz range, have been surveyed in [52] and [53]. Medium access control and routing mechanisms for terahertz-based nanonetworks have been surveyed in [54] and [55], respectively. In contrast, we focus on diffusion-based nanonetworks that transmit signals via the diffusion of molecules.

Both terahertz and molecular diffusion-based communication in the context of body-centric nanonetworks [44], [56], [57], [58] are covered in [59], whereby Sec. VI-B.5 in [59] gives a brief two-paragraph selective overview of channel coding techniques for DBMC. The review articles [60], [61] on reliability and performance enhancing techniques for DBMC include each a one-page overview of channel coding. We also note that an overview of channel coding in DBMC is provided by Damrath in his dissertation [62]. On the other hand, the present survey covers source, channel, and network coding for DBMC and does so comprehensively and in-depth.

The overview article [63] focuses on DBMC, in particular, on the physical design aspects of the transmitters and receivers [64]. As part of this overview, [63, Sec. IV-B., Table 3] gives a two-page overview of channel coding, covering ten selected studies up to 2018. Similarly, the four-page conference paper [65] gives a brief overview of channel coding techniques up to 2018. In contrast, we provide a comprehensive up-to-date in-depth survey of channel coding approaches for DBMC; specifically, our survey includes 39 channel coding studies; whereby 17 of the channel coding studies have not been included in any previous survey; most of these 17 studies appeared after 2018.

We note for completeness that several other surveys have covered complementary aspects of DBMC nanonetworks, including synthetic biological building blocks [66], channel modeling [67], modulation techniques [51], detection techniques [68], estimation techniques [69], retroactivity aspects [70], [71], interfaces [72], [73], and security aspects [74].

C. CONTRIBUTION AND STRUCTURE

This article contributes the first comprehensive in-depth survey of all three types of coding for diffusion-based molecular communication (DBMC) nanonetworks. As illustrated in Fig. 1, we cover source coding, channel coding, and network coding centered around DBMC.

Section II describes the required background to understand the operational principles of DBMC nanonetworks, including microscopic and macroscopic diffusion theory, and the different coding types, namely the source, channel, and network coding types. The subsequent sections are the core survey sections that are dedicated to the coding types of source, channel, and network coding; specifically, Section III surveys source coding approaches (mechanisms, methods, techniques), Sections IV–VIII cover the different subtypes (categories) of channel coding approaches, and Section IX surveys network coding approaches. For each coding type, the corresponding section gives an overview of the historical development of the respective coding techniques in the field of DBMC nanonetworks from the very beginning to the latest studies by thoroughly surveying the existing literature. Throughout, the survey covers the development of the different coding mechanisms for DBMC nanonetworks, including the coding processes and the achieved coding performance tradeoffs.

We summarize and discuss the surveyed coding approaches in Section X to draw inferences and identify the shortcomings and limitations of the existing DBMC coding approaches. Building on the identified shortcomings and limitations of the existing DBMC coding approaches, Section XI outlines the future research directions for each type of coding for DBMC as well as the overarching research imperatives that cut across the individual coding types. The article concludes in Section XII with a brief summary.

II. BACKGROUND

This section first guides the reader through the microscopic and macroscopic theories of diffusion-based molecular communication (DBMC) nanonetworking. Subsequently, the

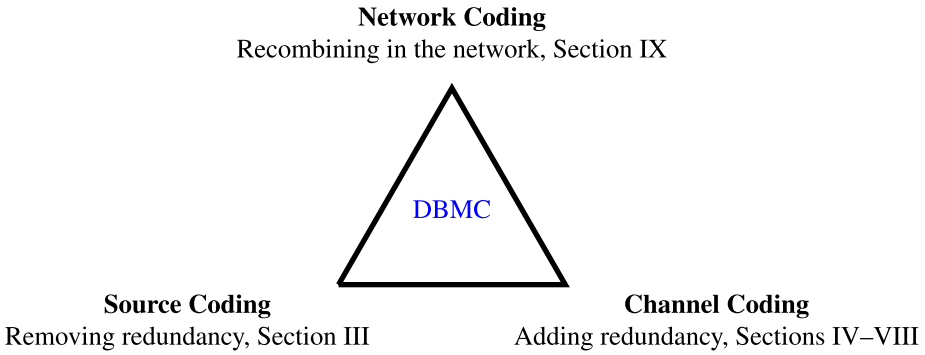


FIGURE 1. Coding triangle for DBMC nanonetworks with the three different types (purposes) of coding, namely source, channel, and network coding.

principles of source, channel, and network coding are explained.

In order to provide the reader with a detailed overview of coding in DBMC systems in the further course of this survey, we first describe the flow of a message from the sender to the receiver, which is illustrated in Fig. 2. The message, an input bit stream of an information source, is encoded in the first step by the source encoder, i.e., redundancy is removed, the length of the message is minimized (independent of the selected communication type). The channel encoder then adds redundancy to mitigate errors or the occurrence of Intersymbol Interference (ISI) after the transmission by the source and propagation through the channel. This is followed by modulation, which depends on the communication type, e.g., Molecular Communication (MC) or EM communication, and the propagation mechanism of the channel. After the propagation of the modulated signal through the channel, the steps follow in reverse order, i.e., demodulation, channel decoding, and source decoding.

A. DIFFUSION-BASED MOLECULAR NANONETWORKING

The process of DBMC consists of transmitters modulating information by expelling particles into a medium. Modulation in Molecular Communications (MC) systems uses different techniques. Thereby, the modulation techniques of MC can be divided into concentration-based, type-based, timing-based, spatial, and hybrid modulation techniques [51]. Generally, the MC coding approaches, especially the MC channel coding approaches, must be adapted to the used modulation technique.

The physical process of diffusion distributes the expelled molecules, and the receivers detect the concentration of these particles to extract the modulated information. In brief, the thermal movement of particles, such as atoms, molecules, or ions, causes Brownian motion, i.e., random non-directional particle movements. In the case of a non-uniform distribution, over time, Brownian motion moves statistically more particles from areas of high concentration (particle density) to areas of low concentration (particle density) [75]. This movement of particles results in a net macroscopic mass transport, which is defined as diffusion [76].

In the field of MC, researchers mimic natural techniques in order to develop alternative communication approaches for cooperating nanomachines. MC, especially DBMC, is energy-efficient. Three different application areas of DBMC can be identified: free-space, in-vessel, and underwater applications [77]. Free-space diffusion can, for instance, occur in the open air [78]. The strengths of free-space diffusion over EM communications are eavesdropping robustness and high energy efficiency. Free-space diffusion channels imitate nature, e.g., colony communications via pheromones.

Vessel-bounded channels are also based on diffusion, but reinforced by an additional drift [79]. Possible applications are inside veins, e.g., for nanobot health monitoring or drug delivery. Diffusion with drift is a mimesis, e.g., of the natural process of organs controlling a system via hormones.

Finally, underwater channels mainly rely on gravity and buoyancy. Only a small proportion of the movement driving force (especially in the horizontal direction) is due to diffusion; rather, current forces are typically mainly responsible for the movement in the horizontal direction, at least in open waters. Under laboratory conditions, diffusion is the strongest force. Underwater diffusion communication often uses a model based on the mimesis of the olfactory sense, e.g., of sharks. Potential applications arise in abyssal sensing, as well as leakage detection [80].

The following subsections provide the mathematical principles necessary to understand the results and developments of DBMC nanonetworks.

1) MICROSCOPIC THEORY

Arising from motion due to thermal energy, diffusion describes the random migration of nanoscale particles or molecules. A particle with mass m and initial velocity v_x has a kinetic energy for one-dimensional motion of

$$\mathcal{E}\left[\frac{mv^2}{2}\right] = \frac{kT}{2}, \tag{1}$$

where $\mathcal{E}[\cdot]$ denotes the operative mean value (expectation) as a functional operator, k denotes the Boltzmann's constant, and T denotes the absolute temperature. Consequently, the

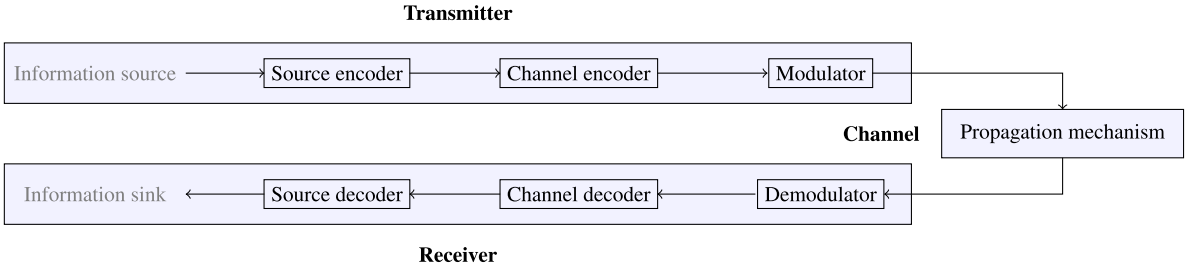


FIGURE 2. Information flow in communication systems from information source in transmitter via channel to information sink in receiver.

root-mean-square velocity becomes:

$$\sqrt{\mathcal{E}[v^2]} = \left(\frac{kT}{m}\right)^{\frac{1}{2}}. \tag{2}$$

Further, consider a random walk of the particles with starting position of molecule i at $x_i(t_0) = 0$. Every time step of duration τ seconds, each particle moves to the left or to the right with velocity $\pm v_x$. Depending on the aqueous medium, the temperature, and the particle size, a particle moves the distance $\delta = \pm v_x \cdot \tau$ in a time step. For each time step, the probability of moving to the left or moving to the right is equal to $1/2$. Furthermore, each particle moves independently of all other particles. Hence, the successive movements of a given particle are statistically independent.

As a first consequence, the particles go nowhere on the average over all particles, e.g., if a particle moves the distance δ to the right, there is also a particle (with statistical certainty) that moves the distance δ to the left. From a time step to the next one, the mean position of the particles does not change. If all particles start at the origin, where the mean position is zero, the expected value of the position at a later time is also equal to zero, i.e., $\mathcal{E}[x(n)] = \mathcal{E}[x(n-1)] = 0$, where n denotes the time step index [81].

Next, let's continue on to the second consequence, the root-mean-square displacement of N particles based on $\mathcal{E}[x^2(n)] = \mathcal{E}[x^2(n-1)] + \delta^2$ by considering the square of the particle displacement, i.e., δ^2. Different from $\mathcal{E}[x(n)]$, now δ^2 is not cancelled out when evaluating $\mathcal{E}[x^2(n)]$ due to the squaring of the δ. Hence:

$$\mathcal{E}[x^2(n)] = n\delta^2, \tag{3}$$

where n denotes the number of time steps. Subsequently, with $n = t/\tau$:

$$\mathcal{E}[x^2(t)] = \frac{t\delta^2}{\tau} = 2Dt \tag{4}$$

$$\mathcal{E}[x^2(t)]^{\frac{1}{2}} = \sqrt{2Dt}, \tag{5}$$

where D denotes the diffusion coefficient.

The diffusion coefficient D describes the migration of a given type of particles at a given temperature in a given medium [81]. In analogy to the one-dimensional random walk, for the two- and three-dimensional random walk:

$$\mathcal{E}[r^2] = \mathcal{E}[x^2 + y^2] = 4Dt \tag{6}$$

$$\mathcal{E}[r^2] = \mathcal{E}[x^2 + y^2 + z^2] = 6Dt, \tag{7}$$

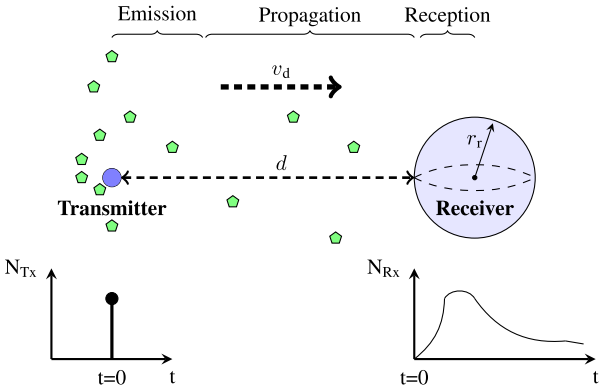


FIGURE 3. System model of molecular communication via diffusion with number N_{Tx} of emitted molecules, number N_{Rx} of received molecules, time t, distance d between the transmitter (point source) and the receiver (spherical receiver), velocity v_d of the drift, and radius r_r of the spherical receiver, adapted from [82].

whereby, the movements in each direction are independent of each other.

From a probabilistic perspective, suppose that a particle moves to the right with probability p and to the left with probability $(1-p)$. The number K of movements to the right in n time steps is binomially distributed, i.e., $K \sim \mathcal{B}(n, p)$. For large n and np, $\mathcal{B}(n, p)$ can be approximated by a Gaussian (Normal) $\mathcal{N}(\mu, \sigma^2)$ or Poisson $\mathcal{P}(\lambda)$ distribution.

In [81], [82], and [83], the approximations of Gaussian and Poisson of the binomial model are analyzed for a point transmitter and a spherical receiver in an aqueous medium, as illustrated in Fig. 3. Based on MC with drift and considering the parameters distance, drift velocity, and number of emitted molecules, the study [82] investigates the regions where either the Gaussian or Poisson model fits better in terms of the Root-Mean-Squared Error (RMSE) of the Cumulative Distribution Function (CDF). The given system, consisting of a source, information molecules, a receiver, and the aqueous medium, was analytically modeled, and the transmission process was simulated. The messenger molecules represent the information particles at the nanoscale.

2) MACROSCOPIC THEORY

Transitioning from the microscopic to the macroscopic theory of diffusion, we suppose that a known number $N(x)$ of particles is at position x (on the one-dimensional x axis) at

time t [81], [84]. By the time $t + \tau$ seconds, half of the particles at x, i.e., $N(x)/2$ particles, cross an imaginary line to the right and half of the particles at $x + \delta$ (to the right of the imaginary line) cross the imaginary line toward the left. The net number of particles crossing the imaginary line to the right for a given time interval duration τ and area A normal to the x-axis leads to the net flux

$$J_x = -\frac{\frac{1}{2}[N(x+\delta) - N(x)]}{A\tau}. \tag{8}$$

The microscopic theory defines the diffusion coefficient as $D = \delta^2/(2\tau)$, see Eq. (4). With the number $C(x) = N(x)/(A\delta)$ of particles per unit volume,

$$J_x = -D\frac{1}{\delta}[C(x+\delta) - C(x)]. \tag{9}$$

For the limit $\delta \to 0$, Eq. (9) becomes a partial derivative, giving Fick's first equation [81]:

$$J_x = -D\frac{\partial C}{\partial x}. \tag{10}$$

Summarized, the slope of the concentration function is proportional to the net flux. Thereby, $-D$ becomes the coefficient of proportionality. A slope equal to 0 and, accordingly, $J_x = 0$, uniformly distributes the particles [81].

Using Fick's first law, the diffusion can be determined for a given time t, but the concentrations also change over time. The time dependence induces a different magnitude of the flux. Justified at a reference volume, Fick's second law, with dependency on time and place, follows as

$$\frac{\partial C}{\partial t} = -\frac{\partial J_x}{\partial x} = D\frac{\partial^2 C}{\partial x^2}. \tag{11}$$

For a radial flux of a spherical symmetric problem, this equation becomes [81]:

$$\frac{\partial C}{\partial t} = D\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial C}{\partial r}\right). \tag{12}$$

Given an initial distribution of the particles and other boundary conditions, this partial differential equation can be solved [81].

For the analytical solution of Fick's second law, we assume a concentration at time $t = 0$ at point $x = 0$ of $c(x, t = 0) = c_0\delta(x)$, where c_0 and δ denote the concentration at time $t = 0$ and the Dirac delta function. Using the Fourier transformation [85], the solution follows [81]:

$$1D : c(t, x) = \frac{c_0}{2\sqrt{\pi Dt}}\exp\left(\frac{x^2}{4Dt}\right) \tag{13}$$

$$3D : c(t, r) = \frac{c_0}{(4\pi Dt)^{\frac{3}{2}}}\exp\left(-\frac{r^2}{4Dt}\right). \tag{14}$$

This is one of the most fundamental equations for diffusion-based nanonetworking. One-dimensional (1D) means, that the concentration $c(t, x)$ as a function of time t and location x can only be determined directionally, in one dimension. Three-dimensional (3D) describes the uniform spreading of the molecules in all spatial directions due to diffusion. The underlying model is based on a point source, whereby r denotes the distance from the source and can be regarded as a point in space.

The following three subsections introduce the fundamentals of the three coding types (purposes of coding), namely source, network, and channel coding.

B. BASICS OF SOURCE CODING

1) SHORT HISTORICAL INTRODUCTION TO SOURCE CODING

Claude Shannon laid the foundation for source coding in [86]. Shannon formulated the key concept of information entropy, along with other fundamental notions in information theory [87] such as the bit as a unit of information. Entropy describes the average information content per symbol of a random source. Shannon proved in what is known as Shannon's first source coding theorem that random sources can be encoded, noiselessly, to the entropy rate of the random variable associated with the source [87]. According to Shannon's first source coding theorem [86], the entropy rate of a stationary ergodic finite-alphabetic Markov source is the optimal rate at which a source can be encoded with block codes of a fixed rate or with noise-free block codes of variable rate. Furthermore, Shannon proposed fixed-rate block codes; the corresponding theoretical result is nowadays called the Shannon-McMillan Theorem [87]. In addition, he constructed a noiseless variable-rate block code known as Shannon-Fano code. This block code encodes a block into a code word of the length within one bit of $-\log_2 p$ [87], whereby p denotes the probability of the occurrence of the block.

However, Shannon left the problem of a noiseless prefix code for a random variable open [87]. David Huffman published in 1952 the algorithm, named after him, that determines at least one optimal code [88]. Since 1952, numerous publications have been built on this algorithm. Probably the most important modification is the work of Robert Gallager on adaptive Huffman coding in 1978 [89].

Other seminal publications have further contributed fundamental advances: establishment of the Shannon-McMillan theorem in 1953 [90], discovery of the Lloyd algorithm in 1957 [91] and its modification [92], and the systematization of the rate distortion theory in 1959 [93]. Additionally, the commencement of multiterminal source coding theory by David Slepian and Jack Wolf in 1973 [94], the universal source coding theory also in 1973 [95], and first practical arithmetic coding schemes in 1976 [96], [97] link to the historic context. For additional background, the reader is referred to [87] and [92] for a more complete historical survey of source coding.

2) DIFFUSION-BASED MOLECULAR COMMUNICATIONS AND SOURCE CODING: A NEW STORY

The history of source coding in combination with molecular nanocommunication systems via diffusion is still nascent.

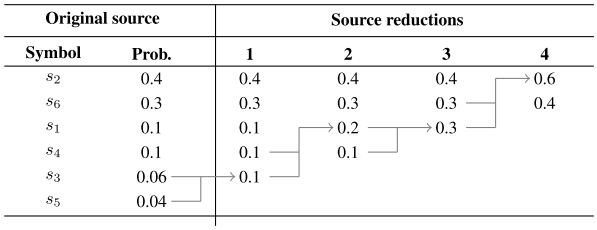


FIGURE 4. Example of Huffman encoding algorithm for source reduction, adapted from [102].

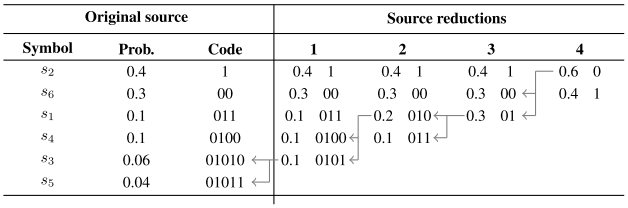


FIGURE 5. Example of Huffman encoding algorithm for assigning the code words, adapted from [102].

We briefly sketch the main historical development steps of source coding for diffusion-based molecular communications here, and conduct a detailed survey of these source coding approaches in Sectino III. Dhayabaran et al. [98] first published a source coding technique called Modified Inverse Source Coding (MISC). This technique allows the control of the *a priori* probability of the transmitted data. In 2021, another study by Dhayabaran et al. presented two low-complexity receivers for MISC DBMC [99].

3) MECHANISMS OF SOURCE CODING

In conventional source coding techniques, code words with a shorter length are assigned to symbols with a higher probability of occurrence; and, conversely, code words with a longer length are assigned to symbols with a lower probability of occurrence. By reducing the length of the code words, better spectral efficiency can be achieved [100], [101]. Essential to the representation of the source coding techniques for MISC DBMC are Huffman coding [100] and inverse Huffman coding [101].

The Huffman code is a popular entropy-based encoding technique [102]. In the Huffman code, a prefix code is created from a given message that minimizes the length of the encoded message [103]. A constant probability distribution of the individual characters is exploited. To uniquely recognize characters with a variable bit length from a bit string, the encoding must have the following property: no bit encoding of a character may occur as a prefix in the encoding of another character [103]. As illustrated in Fig. 4, the first step of the Huffman algorithm is to create a set of source reductions by sorting the probabilities of each symbol. The two least likely symbols are combined into a single symbol, which is used in the next stage of the source reduction [102].

The original source symbols are in the left column of Fig. 4, ordered from bottom to top in ascending order of probability of occurrence. In the first reduction, the symbols with the lowest probabilities are combined, i.e., $P(s_5) = 0.04$ and $P(s_3) = 0.06$. The probability of the composite symbol is $P(s_5) + P(s_3) = 0.04 + 0.06 = 0.1$. Then, the composite symbol is moved to the next column of the source reduction, and the arrangement follows again in ascending order from bottom to top. The process is repeated until only a source reduced to two symbols remains. In the example in Fig. 4, this criterion is reached after four reductions [102].

The first step of the Huffman algorithm is to encode each reduced source symbol, starting on the right side of Fig. 5 and moving towards the original source symbols. For a binary source, the most compact code is represented by the symbols 0 and 1. The 0 and 1 are assigned to the two symbols furthest to the right, according to the common convention (for conventional EM communication) that (depending on a tree structure) the edge of the upper (left) arm is a 0 and the edge of the lower (right) arm is a 1. Symbols created from two other symbols during source reduction are assigned a 0 or a 1 in the next smaller reduction column. The assignment is made according to the same convention—the lower symbol is assigned a 1. This process is repeated for each reduced source symbol until the set of original source symbols is reached, see Fig. 5.

The resulting code is given in the third column (from the left) of Fig. 5. The average length L_{avg} of the code words is:

$$\begin{aligned} L_{avg} &= 0.4 \times 1 + 0.3 \times 2 + 0.1 \times 3 + 0.1 \times 4 \\ &\quad + 0.06 \times 5 + 0.04 \times 5 \\ &= 2.2 \text{ bits/symbol}. \end{aligned} \tag{15}$$

The Huffman algorithm produces the shortest possible average code word length for a given source [102]. Huffman coding is often referred to as "instantaneous, uniquely decodable block coding" because each symbol of the source is encoded into a fixed bit sequence (block) [102]. Each code word can be decoded immediately, i.e., without reference to previous or subsequent symbols. In addition, there is only one way to decode a given sequence of 1s and 0s. Consequently, no code word is allowed to be the prefix of another valid code word. For encoding large numbers of source symbols, the original source reduction algorithm becomes impractical as well as computationally inefficient [102]. Due to this limitation of the original Huffman coding algorithm, it has been modified to reduce the computational complexity (to the detriment of the coding performance). The new modified algorithm is named in the literature as modified, truncated, or adaptive Huffman coding. Detailed descriptions of some of these variants are available in [104].

C. BASICS OF CHANNEL CODING

1) SHORT HISTORICAL INTRODUCTION TO CHANNEL CODING

After the first successful transmissions, the demand for high data rates and high reliability increased from year to year.

Researchers were increasingly confronted with the question of how to transmit data reliably over a noisy channel [105]. Noise is thereby the greatest challenge and limitation in communication systems. The communication channel is susceptible to different noise effects, e.g., spurious molecules in MC, interference, distortion due to hardware imperfections, or physical limitations [105]. All these effects can lead to errors. Consequently, the need arises for error control encoding to reliably recover the information [105]. The idea of recovering the erroneous or incomplete data represents the birth of error control coding or channel coding.

Claude Shannon is considered the father of information theory, and as early as 1948 he showed that even in a noisy channel there is a possibility of encoding the data in such a way that there is a probability arbitrarily close to one of transmitting data reliably and decoding it correctly [86], [105]. This was the theoretical limit, but no practical coding technique was presented [105]. The first practical code was found by Richard Hamming in 1950, namely the Hamming codes, see illustration in Fig. 6. Hamming added additional parity bits (the so-called redundancy) to the information bits, allowing the correction of up to one bit error in a transmission [106].

Research on channel coding approaches for conventional (EM) communication can be classified into the primary categories of block codes, convolutional codes, and concatenated codes [105], see Fig. 6. Block codes, such as the Hamming code, implement an algebraic structure in the code so that computationally efficient algorithms can solve systems of equations by using boolean polynomials. Convolutional codes encode trellis-based algorithms, such as the Viterbi algorithm [107], generating parity symbols. Concatenated codes are designed from at least two codes (inner and outer code) to achieve a good performance, e.g., classic concatenated codes [108]. In particular, the channel coding research was accelerated by the implementation of turbo iterative decoding [109], which strongly approaches Shannon's limit in terms of performance. The Low Density Parity Check (LDPC) codes developed by Gallager [110], and rediscovered by MacKay and Neal towards the end of the 1990s [111], [112], performed well and attracted attention. Erdal Arikan brought further progress in 2008 with the introduction of polar codes [113]. Among other properties, he showed that the channel capacity can be achieved asymptotically with a simple, recursive decoding algorithm [113].

2) DIFFUSION-BASED MOLECULAR COMMUNICATIONS AND CHANNEL CODING

Schneider published in [126] the first explorations of the channel capacity of molecular machines. He explored how Shannon's theorem can be implemented in molecular biology. In the early 2000s, further research in channel coding and wireless sensor networks with LDPC codes followed [127]. After that, for the first time, in 2012, Ko et al. presented a molecular coding distance function [128], which marked the beginning of channel coding in MC. In the following two years, channel codes for mitigating ISI and for enhancing the reliability in MC systems were researched, see Section VII. Based on [129] and [130], this was followed in 2014 by the first-ever analysis of Hamming codes examined by simulations with respect to MC systems via diffusion [131]. Between 2014 and 2016, additional ISI mitigation techniques and ISI-free modulations have been presented and verified against simulation results [132], [133], [134].

With the increasing attention to MC research, channel coding approaches for MC systems have continued to be in focus over the past few years [135], [136]. Especially in the early 2020s, a trend towards low-complexity MC channel codes can be identified [137].

3) CHANNEL CODING TECHNIQUES

In principle, channel coding can be divided into the class of block and convolutional codes [138]. Section II-C3.a introduces block codes, especially linear block codes. Then, Section II-C3.b briefly explains convolutional codes. Forward Error Correction (FEC) describes the approach when the receiver equipment does most of the error correction work [138].

Fig. 7 shows schematically the channel block coding procedure using MC as an example. A message vector $m = m_1, m_2, \ldots, m_k$, consisting of k message elements of a defined set, is converted by a block code into a longer vector, e.g., a code word $U = u_1, u_2, \ldots, u_n$, consisting of n coding elements of the same defined set of message elements [138]. The elements of the defined set have a one-to-one correspondence with drawn elements of a finite field [138]. Finite fields go back to the French mathematician Evariste Galois, who published the Galois fields [139]. A Galois field has the notation $GF(q)$, where q describes the number of elements contained in the field. The simplest finite field is $GF(2)$, which contains the binary elements $\{1, 0\}$, which have a logical connection to bits [138]. A block code is characterized by the number n of output elements (code bits) and number k of input elements (information bits), so that the notation (n, k) follows for a block code. Often the extension by the variable t follows, which describes in the code (n, k, t) the number t of correctable errors in the n-element code word [138]. In the encoding process, G describes the generator matrix for transforming the message vector m into the code word U.

Various modulation techniques can be used for the transmission of the code bits U. The most common application is On-Off Keying (OOK), where the values $\{1, 0\}$ are represented by the release of a prescribed number of molecules and the release of no or few molecules, respectively. The modulation depending on the time t is denoted as $s(t)$. Noise, interference, and channel distortion mechanisms [140], collectively denoted by $n(t)$, are responsible for the conversion of the transmitted form $s(t)$ into the corrupted received form $r(t) = s(t) + n(t)$. A widely used model for $n(t)$ is the additive white Gaussian noise process [141]. The corrupted form is received by the demodulator and fed into the detector [138].

	Block Codes	Convolutional Codes	Concatenated Codes
		Shannon Capacity Theorem	
1950	Hamming Code [106]	Classic Convolutional Codes [114]	
1960	Berlekamp-Massey algorithm [119], [120]	LDPC (Gallager) [110] Viterbi algorithm [107]	Classic Concatenated Codes (Forney) [108]
1970	Chase algorithm [121]		
	Linear Block Codes (Wolf) [122]	Decoding algorithm (Bahl et al.) [115]	
1980		Classic Trellis-Coded Modulation (Ungerboeck) [116]	
1990	SISO[1] Chase algo. (Pyndiah) [123]	Max-Log MAP[2] algorithm (Koch) [117] Space Time Trellis Code (Tarokh) [118]	Turbo Codes (Berrou et al.) [109]
2000	Space Time Block Coding (Alamouti) [124]	LDPC Codes revisited [111], [112]	Turbo Trellis-Coded Modulation (Robertson) [125]
2010	Polar Code [113]	LDPC Codes as a powerful code [111], [112]	
2020			

[1]Soft-Input Soft-Output [2]Maximum A-Posteriori

FIGURE 6. Historical view of channel coding approaches for conventional (electromagnetic) communication, following [105].

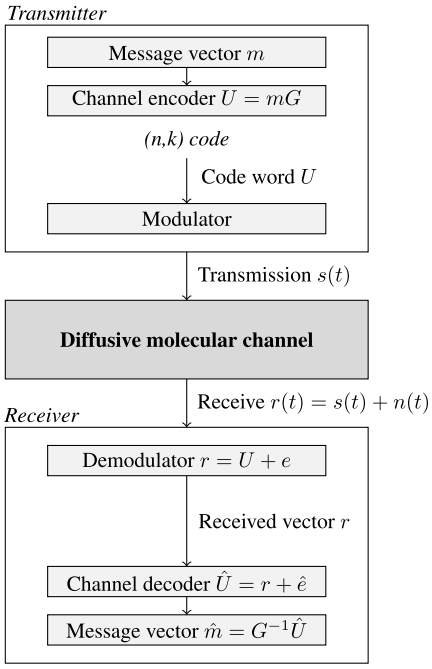


FIGURE 7. Technical scheme of channel coding applied to MC, adapted from [138].

Consequently, the detector does not reproduce the exact transmitted code word U, but rather a corrupted version $r = U+e$. Whereby, $r = r_1, r_2, \ldots, r_n$ and $e = e_1, e_2, \ldots, e_n$ denote the received block of n elements and the corruption.

a: LINEAR BLOCK CODES

This section provides a brief overview of linear block codes. We refer to [138] for more details. First, let's turn our attention to the Simple Parity Codes (SPC). On the transmitter side, the encoder adds redundancy. A number of boundary conditions apply, which must be fulfilled by the set of all code words. If a received vector does not satisfy the conditions, then error detection occurs. The simplest approach to error detection adds an extra bit, called a parity bit, to the binary sequence. For instance, if the boundary condition requires that the code word must contain an even number of ones, then the procedure is called even parity [138]. The same applies vice versa, i.e., for an odd number of ones, which corresponds to odd parity [138]. For even parity, the parity bit p is formed as the modulo-2 sum of the message bits to satisfy the even parity constraint:

$$p = m_1 \oplus m_2 \oplus m_3 \oplus \cdots \oplus m_k, \tag{16}$$

whereby, $\oplus$ denotes the modulo-2 addition. The parity check performed at the receiver side checks whether the modulo-2 sum of the parity and message bits in the received sequence is zero for an even parity [138]. An error is detected when the complementary counter value, i.e., a one, occurs. The test result is denoted by S [138]:

$$S = r_1 \oplus r_2 \oplus \cdots \oplus r_k \oplus r_{k+1}. \tag{17}$$

Single parity bits are only used in error detection [138]. For error correction, additional information is required for error

TABLE 1. Two-dimensional parity check in rows and columns.

m_1	m_2	m_3	m_4	p_1
m_5	m_6	m_7	m_8	p_2
p_3	p_4	p_5	p_6	

localization. A simple example is provided by [138], in which the message elements are transformed into a two-dimensional array; whereby, parity can be determined for each row and each column in the array, see Table 1.

Secondly, we briefly review the cyclic codes. Cyclic codes belong to the linear block codes, based on an algebraic structure. Cyclic codes have good error correction capabilities and consequently exhibit computationally efficient encoding and decoding [138]. Examples are the Bose, Chaudhuri, and Hocquenghem (BCH) codes and the Reed-Solomon codes. A code word U consisting of n elements can be described as

$$U = u_0, u_1, u_2, \ldots, u_{n-1}. \tag{18}$$

A single cyclic shift of the elements, then also represents a code word,

$$U^{(1)} = u_{n-1}, u_0, u_1, \ldots, u_{n-2}. \tag{19}$$

A number of properties, as well as structures, can be derived from Eq. (19) [138]. Data sequences and code words can thus be mapped to polynomials in the indeterminate X. When polynomials are used, the counting of elements starts with the index 0, see Eq. (18). Based on the mapping, a code word of the form shown in Eq. (18) can be naturally mapped to its corresponding polynomial [138]:

$$U(X) = u_0 + u_1 X + u_2 X^2 + \cdots + u_{n-1} X^{n-1}. \tag{20}$$

The elements of the sequence correspond to the coefficients of the polynomial. For the indefinite variable X, only positive integer exponents are used. As per Eq. (20), it is possible to represent all code words by polynomials of the maximum degree $n-1$ [138]. The coefficients are taken from the binary field. For non-binary codes, non-binary coefficients follow the definition in [138]. In accordance with Eq. (19),

$$U^{(1)}(X) = u_{n-1} + u_0 X + u_1 X^2 + \cdots + u_{n-2} X^{n-1}. \tag{21}$$

Consequently, the cyclically shifted form can be obtained by using X as a delay operator, followed by trivial manipulations [138]. In the next step, the polynomial $XU(X)$ is formed to linearly shift the elements. The terms u_{n-1} and u_{n-1} can then be added. Adding the terms u_{n-1} and u_{n-1} does not change the shifted polynomial, because they sum up to zero in the binary field [138]. Hence,

$$\begin{aligned} XU(X) &= u_0 X + u_1 X^2 + \cdots + u_{n-2} X^{n-1} + u_{n-1} X^n \\ &= u_{n-1} + u_0 X + u_1 X^2 + \cdots \\ &\quad + u_{n-2} X^{n-1} + u_{n-1} X^n + u_{n-1} \\ &= u_{n-1} + u_0 X + u_1 X^2 + \cdots \\ &\quad + u_{n-2} X^{n-1} + u_{n-1}(X^n + 1). \end{aligned} \tag{22}$$

Dividing the right side of Eq. (22) by the expression (X^n+1), discarding the quotient polynomial, and keeping the remainder polynomial is called a remainder reduction operation [138]. If the cyclic shift is repeated using the remainder reduction operation, the polynomial that equals all cyclic shifts of a code word in the subspace is obtained [138].

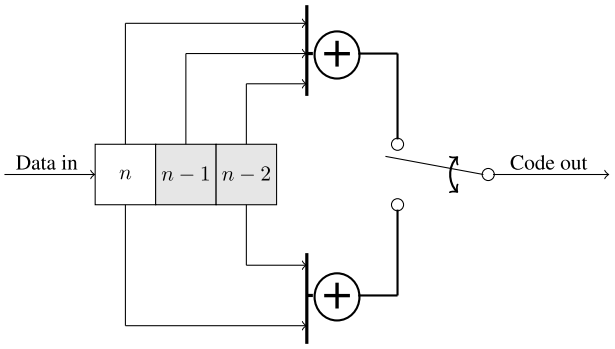


FIGURE 8. Technical scheme of a convolutional encoder with shift register of length $K = 3$ and $r = 2$ summation blocks, adapted from [146].

Finally, we turn our attention to the well-known Hamming and BCH channel codes. Hamming channel codes are single-error correcting codes which exist for $(n, k, 1)$. They are described by [138]:

$$(n, k, 1) = (2^m - 1,\ 2^m - 1 - m,\ 1), \tag{23}$$

whereby m denotes the number of parity bits in the code. For the operation of Hamming codes, refer to [142]. Generalized Hamming codes lead to the BCH codes [143], [144], [145]. BCH codes are a class of cyclic codes that have a large variance of block lengths, different code rates, and error correction strengths. A t-error-correcting BCH code is described with the (n, k, t) notation by [138]:

$$(n, k, t) = (2^m - 1,\ 2^m - 1 - mt,\ t). \tag{24}$$

The generating polynomial, as well as the number of parity bits in the code, is less than or equal to mt. Using the code following Eq. (24), t binary errors in a code word of length $2m - 1$ bits can be corrected [138].

b: CONVOLUTIONAL CODES

In addition to block codes, convolutional codes are a widely used channel codes [146]. In contrast to blocks consisting of bits; with convolutional codes, the output bits are coded by logical operations with the current bit in the data stream and a few previous bits, see Fig. 8. In the encoder, the data bits are placed in a shift register of length K (also called constraint length). If a bit enters at the left position, the remaining bits are shifted to the right. The oldest bit, the right-most bit, is removed [146]. Binary summation operations (at least two), the number of summation operations denoted as r (and $R = 1/r$ denoting the code rate), create the code bits that can be output in one period of the data flow. Consequently, the code bit rate is $1/r$ times the data rate. Thereby, the code rate characterizes the number of information symbols (here 1) transmitted per code word related to the length of the code words (here r). The encoder is called a convolutional encoder, with parameters K and $1/r$. In addition, the encoder's connections between the individual stages in the shift register and the r summation blocks are required. The connections are called generator vectors, consisting of a series of K binary digits.

Parity bits are generated at the outputs of the binary adders. Thereby, connections with an even number of ones lead to the output of zero and vice versa.

Fig. 8 visualizes an example with parameters $K = 3$, $r = 2$, and generator vectors [1, 1, 1] and [1, 0, 1], whereby n denotes the discrete sampling time points. The data stream enters from the left; the bit at the current time n, the most recent bit $(n - 1)$, and the preceding bit $(n - 2)$ take the places in the register. Two parity bits from the upper and lower adder are executed in the interval between n and $n - 1$. When the next bit arrives, a shift to the right follows. The $(K - 1) = 2$ bits shown in gray in Fig. 8 define the state of the encoder. In total, there are 2^{K-1} states [146]. For each state, two possibilities can be used for the output of the code bits, depending on the input symbol zero or one. The time sequence of the different states is therefore a function of the data stream [146]. For more details, see [146].

D. BASICS OF NETWORK CODING

1) SHORT HISTORICAL INTRODUCTION TO NETWORK CODING

In 1948, Claude E. Shannon published a paper titled "A Mathematical Theory of Communication" [86]—the era of digital communications had started, and information theory was born. The seminal work of Shannon led to the source and channel coding principles reviewed in Sections II-B and II-C. The initial problems of information theory were related mainly to point-to-point systems, i.e., systems with one source (or multiple sources) and a destination, and without intermediate nodes. In the case of source coding, the question was how to reduce the redundancy of information sources to transmit the messages such that a destination could recover them error-free. In the case of channel coding, the question was how to add redundancy to the communication such that errors introduced in the communication link (channel) could be repaired at the destination. Therefore, initially, the information sent through a network was interpreted as an exchange of commodities (the so-called commodity flow). In this commodity flow context, the information bits are only stored and forwarded by the intermediate network nodes.

Later on, in 2000, Ahlswede et al. [147] introduced the concept of network information flow, which added a new question to the coding problems of information theory. With source and channel coding, we learned what operations the sources and the destination should carry out to compress the sources and to achieve the channel capacity. Ahlswede et al. questioned what the intermediate network nodes that interconnect the sources and destinations should do, and what the maximum capacity of such networks is. The main idea in [147] was to allow intermediate nodes to combine the incoming bits at the network nodes. The maximum capacity of an arbitrary multicast network was proven to be determined by the min-cut max-flow theorem [147]. Furthermore, [147] proved that linear combinations of the information bits at the intermediate nodes can achieve the maximum capacity of a network beyond the limit reached by commodity flow routing.

This extension gave birth of a new research area called network coding. Network coding can be seen as an addition to the known source coding (compressing the information at the source to increase the efficiency of transmission) and channel coding (introducing redundant bits into the information sequence to correct potential errors introduced by a noisy channel), as we illustrated in Figure 1. Network coding opened the way for another coding operation, which allows the nodes in the network to perform coding operations over the received information bits and recombining them. Thus, a node can transmit functions of the messages previously received on the incoming edges to the outgoing ones. Over the last decades, there has been an increasing interest in network coding theory [148], [149], [150] and its applications [151], [152], [153], [154], [155], [156], [157], [158], [159].

2) DIFFUSION-BASED MOLECULAR COMMUNICATIONS AND NETWORK CODING: THE ANTECEDENTS

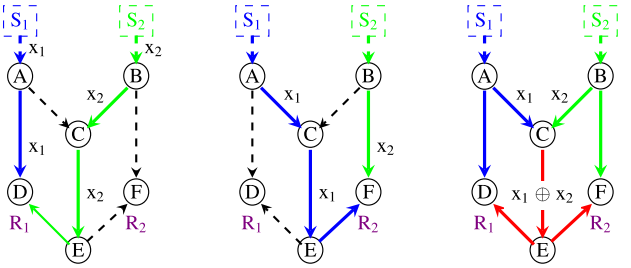
After the publication of [147] in 2000, significant research efforts began in the area of network coding theory. While network coding studies in the context of DBMC nanonetworks are surveyed in detail in Section IX, we briefly sketch here the historical developments. The beginning of nanonetworks was represented in particular by the so-called wireless sensor networks [160]—with nodes becoming smaller and smaller. In this context, collections of highly distributed, small, and lightweight wireless sensor nodes monitoring the environment or systems based on physical measurements are referred to as a wireless sensor network [161], [162]. During the early 2000s, the first network coding approaches began to be implemented in sensor networks [163], [164], [165], including in underwater sensor networks [166].

Then, in the early 2010s, MC via diffusion gained attention. First, communication approaches were determined analytically and compared with numerical simulations [167], [168]. Further, in the early 2010s, a trend towards network coding in combination with DBMC finally developed.

3) NETWORK CODING MECHANISMS

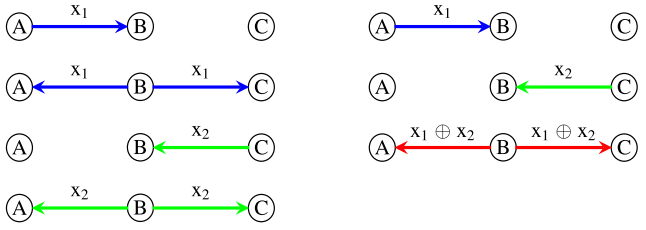
We briefly review the basics of network coding. Fundamentally, network coding can be based on the so-called eXclusive OR (XOR). The output of an XOR is one if the inputs A and B are unequal. Equivalently, the XOR operation (symbol: $\oplus$) is addition over the binary field.

With the advent of network coding, the simple and important observation was made that in communication networks we are not only able to forward, but also to process the incoming information streams. For example, at the network layer, binary addition of independent bit streams can be performed between nodes, e.g., at the physical layer of optical networks, intermediate nodes can overlay incoming optical signals. In other words, data streams that are produced



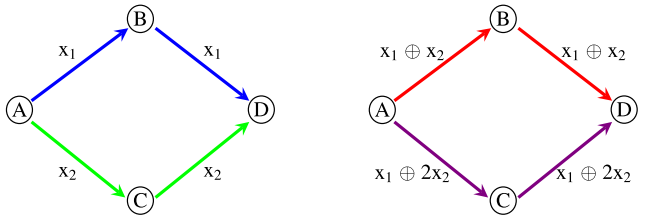
(a) Routing to R_1 *(b) Routing to R_2* *(c) Network Coding*

FIGURE 9. Illustration of the butterfly network: With network coding, sources S_1 and S_2 superimpose (XOR) their information bits on the link $\{CE\}$ to simultaneously deliver both bit streams to both receivers R_1 and R_2 [169].



(a) Without Network Coding *(b) With Network Coding*

FIGURE 10. Nodes A and C need four time slots to communicate with each other without network coding (a), and only three time slots with networking coding (b) [169].



(a) Without Network Coding *(b) With Network Coding*

FIGURE 11. Transmission of linear combinations in a communication network to protect against one eavesdropping intermediate node [169].

and consumed independently do not necessarily have to be transported separately. There are ways to combine independent information streams and extract them later. Combining independent data streams allows for a better adaptation of the information flow to the network environment and the requirements of the network. Network coding offers benefits in various dimensions of communication networks, such as wireless resources, complexity, throughput, security, and resilience [169]. To illustrate the advantages of network coding, we give three examples. The examples are intended to generate a basic understanding. For the theoretical background, please refer to [148], [150], [169], [170], and [171].

The first benefit enhances throughput. Fig. 9 shows a communication network represented as a directed graph where vertices correspond to end devices and arrows correspond to channels. This example is known in the network coding literature as a butterfly network. Assume that time is slotted and that we can send one bit per time slot over each channel. The network consists of two sources S_1, S_2 and two receivers R_1, R_2. Each source generates one bit per time slot, which we denote as x_1 and x_2, respectively.

If the receiver R_1 uses all network resources by itself, it could receive both sources. For instance, the bit x_1 can be routed via the channel $\{AD\}$, while the bit x_2 can be routed via the path $\{BC, CE, ED\}$, as illustrated in Fig. 9(a). Analogously, if the receiver R_2 claims all network resources for itself, then the bit x_2 can be routed via the channel $\{BF\}$, while the bit x_1 can be routed via path $\{AC, CE, EF\}$, see Fig. 9(b).

If both receivers want to simultaneously receive the information from both sources, there is an emerging problem in the channel $\{CE\}$, arising from the fact, that only one bit per second can pass through this channel (edge). Sending the bit x_1 from the source S_1 via $\{AD\}$ to R_1, and from S_1 via $\{AC, CE, EF\}$ to R_2, only receiver R_1 will receive the bit x_1; analogously, sending x_2 from the source S_2 via $\{BF\}$ to R_1, and from S_2 via $\{BC, CE, ED\}$ to R_2, only receiver R_2 will receive the bit x_2 due to the bottleneck $\{CE\}$. The simple but highly important observation in [147] is that intermediate network nodes should process the incoming information streams, instead of merely forwarding them. The node C can take bits x_1 and x_2, and XOR them to generate a third bit $x_1 \oplus x_2$, which can be sent over the edge $\{CE\}$, see Fig. 9(c). R_1 receives $\{x_1,\ x_1 \oplus x_2\}$ and can solve this system of equations to obtain x_1 and x_2. Similarly, R_2 receives $\{x_2,\ x_1 \oplus x_2\}$, and can solve this system of equations to obtain x_1 and x_2 [169].

In terms of throughput, network coding can bring another benefit for communication networks. Especially for DBMC nanonetworks, throughput is a very important metric for improving communication because of the low MC bit rate in comparison to EM communication. Consider the wireless ad hoc network in Fig. 10 in which the devices A and C exchange the binary files x_1 and x_2, via the relay device B. We further consider slotted time and a device either sending or receiving a file during a time slot (half-duplex communication). Without network coding, nodes A and C send their files to relay B, which forwards each file to the corresponding destination within four time slots, as illustrated in Fig. 10(a).

The network coding approach leverages the natural capability of wireless channels to broadcast transmissions. In particular, node B receives the two files x_1 and x_2 and concatenates (XORs) them bitwise. Then, node B sends the XORed file to both receivers (A and C) via one common broadcast transmission. Node A has knowledge of x_1 and thus can decode x_2; analogously, node C has knowledge of x_2 and can decode x_1. This approach enhances the energy efficiency (node B only needs one transmission), reduces the delay (the complete file exchange is completed after three time slots instead of four), improves the wireless bandwidth utilization (the wireless channel is occupied for shorter time durations), and reduces the interference (in case of other wireless nodes in the vicinity trying to communicate) [169].

Finally, sending linear combinations of packets instead of uncoded data exploits the network multipath diversity to protect against eavesdropping attacks. Thus, systems that only need protection against simple attacks can operate without additional security mechanisms.

In Fig. 11, node A sends information to node D via two paths: $\{ABD\}$ and $\{ACD\}$. We also assume that an attacker eavesdrops on a single path, but not on the complementary path. If the independent symbols x_1 and x_2 are sent unencrypted, then the attacker can intercept one of them. If instead linear combinations (over a finite field) of the symbols are sent, then the attacker cannot decrypt a single piece of data [169, Section 7.].

Further detailed overviews of the theoretical areas of network coding can be found in [148], [169], [172], [173], [174], [175], [176], [177], and [178]. General application-oriented network coding (not specific to DBMC) is surveyed in [179] and [180].

E. DIFFERENCES BETWEEN CODING IN EM COMMUNICATIONS AND MC

This section briefly outlines main differences between signal propagation in conventional EM channels versus MC channels [34], [181] which uniquely influence the design and performance of all types of coding mechanisms (source, channel, and network coding) for MC.

1) SOURCE CODING

Source coding for MC is fundamentally different from source coding for EM communication in that source coding for MC strives to reduce the occurrence of errors during propagation on the channel by reducing the code word weight, i.e., the number of bit-1s to be transmitted, at the expense of potentially increasing the average code word length. For example, the MISC source coding approach for MC controls the percentage of 0s and 1s in the transmitted data [98]. In contrast, source coding for EM communication focuses on the removal of redundancy, i.e., the shortening of the code words. Specifically, conventional source coding methods assign shorter code words to more probable information symbols and longer code words to less probable information symbols. In contrast, MISC assigns code words with longer length and lower code word weight (i.e., fewer bit-1s) to more probable symbols [98].

2) CHANNEL CODING

A major challenge for channel coding in MC is the random delay of the transmitted molecules due to the channel propagation via diffusion, which is commonly considered in MC. In conventional EM communications, the underlying propagation mechanism is commonly the propagation of EM waves in an unguided (open air or water) or guided medium (e.g., cable or fiber), which results in relatively small to moderate delay spreads [182]. In contrast, the diffusion-based propagation in MC results in relatively large delay spreads [183], [184]. If the molecules are encoded in a time-slotted system, then delayed molecule arrivals can lead to ISI, which in turn increases the Bit Error Rate (BER).

Among the numerous error models for conventional EM communication [185], the Binary Symmetric Channel (BSC) model is a basic common model [186], [187], [188]. In the BSC model, the transmitter sends a bit (bit-0 or bit-1) and the receiver will receive a bit (bit-0 or bit-1). The received bit is flipped (an error) with a probability p; whereas the received bit is correct with probability $1-p$. With X and Y denoting the input and output of the channel, respectively, the channel crossovers for a symmetric channel can be characterized as (following [101]):

$$\begin{aligned} P(Y=0|X=0) &= 1-p \\ P(Y=0|X=1) &= p \\ P(Y=1|X=0) &= p \\ P(Y=1|X=1) &= 1-p. \end{aligned} \tag{25}$$

In contrast, MC channels can be asymmetric [189], [190], [191]. For instance, considering an MC system with the common OOK modulation technique and without ISI (whereby the ISI can be eliminated with sufficient time spacing between successive bit transmissions [192]), Einolghozati and Fekri [190] analytically model the asymmetric channel error probabilities: A bit-0 (absence of released molecules) is received correctly with probability one (for an ISI free channel without residual molecules from preceding bit-1 transmissions). In contrast, a bit-1 transmission (release of molecules) is received incorrectly with probability p (which depends on the diffusion characteristics and the molecular binding at the receiver); and correctly with probability $1-p$, i.e., Einolghozati and Fekri [190] model a so-called z-channel [101]:

$$\begin{aligned} P(Y=0|X=0) &= 1 \\ P(Y=0|X=1) &= p \\ P(Y=1|X=0) &= 0 \\ P(Y=1|X=1) &= 1-p. \end{aligned} \tag{26}$$

For MC with ISI, McGuiness et al. [191] examine the effects of the residual molecules that remain from preceding bit-1 transmissions. These residual molecules lead to a $P(Y=1|X=0) > 0$ in a future slot. That is, despite a bit-0 being transmitted, the residual molecules can result in a bit-1 at the receiver, and this probability of a transmitted bit-0 turning into a received bit-1 increases as more residual molecules accumulate, resulting in a binary asymmetric channel model [191].

3) NETWORK CODING

In classical network coding in wireless (EM) communication, the received signal can be a superposition at the EM wave level, i.e., superposition of EM waves of the transmitted signals from multiple network nodes, resulting in physical-layer network coding [193], [194]. On the other hand, chemical

TABLE 2. Overview of source coding studies in DBMC nanonetworks.

Coding Approach	Reference	Modulation	Channel Model	Complexity
Mod. Inverse Source Coding (MISC)	[98], [99]	BCSK/OOK	Diffusion/3D/No background noise	Low

TABLE 3. Huffman coding for desired percentage $p = 0.2$ of 1s and desired percentage $q = 0.8$ of 0s in source code words, following [98].

Symbol/Length	Probability	Code word/Length
00 / 2	0.64	1 / 1
01 / 2	0.16	00 / 2
10 / 2	0.16	011 / 3
11 / 2	0.04	010 / 3

TABLE 4. Inverse Huffman coding for $p = 0.2$ and $q = 0.8$, following [98].

Symbol/Length	Probability	Code word/Length
010 / 3	0.125	1 / 1
011 / 3	0.125	01 / 2
00 / 2	0.25	001 / 3
1 / 1	0.5	000 / 3

reactions can be used to design physical-layer network coding approaches in DBMC systems. In physical-layer DBMC network coding, two molecules can react with each other through an irreversible chemical reaction. Thereby, chemical reactions represent a simple implementation method for the XOR operator in DBMC networks.

III. SOURCE CODING IN DIFFUSION-BASED MOLECULAR NANONETWORKS

This section comprehensively surveys the existing studies on source coding in molecular nanocommunication systems. The section is divided into Section III-A Modified Inverese Source Coding (MISC), Section III-B Low-Complexity Receiver Design for MISC, and ends with a summary and discussion in Section III-C. Table 2 gives an overview of the source coding studies for DBMC systems.

We note that general source coding (such as Huffman coding) does not account for the idiosyncrasies of MC. Therefore, efficient MC source coding needs to modify the conventional source coding approaches of conventional EM communication; specifically, MISC, which is the only currently existing MC source coding approach, controls the percentage of 1s and 0s in the transmitted message. Thereby, the MISC source coding approach builds on the fundamental idea of source coding of wireless communication and the Huffman coding principles, but adapts these Huffman source coding principles to the specific characteristics of MC.

A. MODIFIED INVERSE SOURCE CODING (MISC)

MISC was originally introduced in [98] in 2020 for DBMC, and allows control of the *a priori* probability of the transmitted data. MISC was inspired by the Inverse Source Coding (ISC) technique, which was originally examined for controlling the dimming percentage in visible light communication [195].

The model in [98] follows a three-dimensional DBMC system consisting of a fully absorbing spherical receiver with a radius r_r and a point transmitter at a distance of d, see Fig. 3. The data transmission is based on the OOK modulation scheme: Bit-1 is represented by the release of N_{Tx} molecules, while no molecules are released for the modulation of bit-0. The OOK modulation is a special case of the more general Binary Concentration-Based Shift Keying (BCSK) modulation which adapts the concentration for a bit-0 and a bit-1.

MISC controls the percentage of 1s and 0s in the transmitted data. Conventional source coding methods follow the principle of assigning shorter-length code words to the more probable information symbols and longer-length code words to the less probable ones, according to Section II-B. The average code word length can thus be reduced, and better spectral efficiency can be achieved [100]. MISC assigns code words with longer length and lower code word weight to the more probable symbols, and vice versa [98]. As a result, MISC can reduce the ISI and the BER in OOK modulated DBMC systems [98]. Adversely, MISC affects the channel capacity of the molecular communication system: By effectively decompressing the data to be transmitted, MISC decreases the capacity compared to non-decompressed data transmissions [98].

More specifically, Dhayabaran et al. [98] use Huffman coding (cf. Section II-B) to generate inverse Huffman codes. It is assumed that the desired percentage of 1s and 0s is $p = 20\%$ and $q = 80\%$, respectively [98]. In inverse Huffman coding, the symbols with high probability of occurrence are assigned codes with the longest length and the symbols with low probability of occurrence are assigned codes with the shortest length [98]. MISC assigns 0s instead of 1s to the likely symbols, contrary to a possible Huffman coding conventions that assign the 1s to one particular branch (e.g., the right branch or edge) or to the more likely symbols. The Huffman code and a corresponding inverse Huffman code for $p = 0.2$ and $q = 0.8$ are given in Table 3 and Table 4.

Through MISC, the percentage of 1s, can be obtained from Table 4, as a fraction with the numerator representing the average number of 1s in the code words, and the denominator representing the average code word length:

$$Per = \frac{1 \times 0.125 + 1 \times 0.125 + 1 \times 0.25}{3 \times 0.75 + 2 \times 0.125 + 1 \times 0.125} \approx 0.190 \backsimeq 20\%. \quad (27)$$

The simulation results in [98] demonstrate that the MISC DBMC system with a larger range achieves a lower average BER than the uncoded system with a smaller range up to a

certain value for the released molecules N_{Tx}. The reason is a reduction of 1s in the transmitted code words compared to the non-coded system. By decompressing the data in the MISC algorithm, the data rate of the MC system is effectively reduced. Specifically, compared to an uncoded DBMC system, the simulation results indicate that the channel capacity of the MISC DBMC system is lower as a function of the decompression rate of the associated MISC scheme [98]. The study [98] also demonstrated that MISC systems exhibit lower BER compared to uncoded systems even at a greater distance between the transmitter and receiver, thus enabling reliable communication over a longer distance.

B. LOW-COMPLEXITY RECEIVER DESIGN FOR MISC

Building on the findings from [98], Dhayabaran et al. [99] develop two low-complexity receivers for MISC DBMC systems. The first low-complexity receiver, the Reduced State Sequence Detector (RSSD), is based on the Viterbi algorithm [99]. If the atypical states generated by the MISC technique are removed from the trellis structure of the conventional based sequence detectors, the RSSD receiver can be designed. The second algorithm is the Decision Feedback Based Maximum A-Posteriori (DF-MAP) threshold detector, which requires less computing power [99]. The DF-MAP receiver relies on feeding back previous decisions to a simple MAP-based symbol detector, rather than a more complex trellis-based receiver. Fig. 12 shows the placement of both the MISC encoder and the detector in the MC system. The following sections briefly explain each of the low-complexity receivers.

1) RSSD

The diffusive molecular transmission channel, affected by ISI, can be modeled as a system of finite-state machines. Whereby, the different states correspond to the possible combinations of ISI symbols [196]. The sequence detectors based on the use of a trellis provide the optimal performance for this type of channels [197]. In this context, a computationally efficient way to implement the detection using the trellis is provided by the Viterbi algorithm [198]. ISI symbols with a larger number of ones have a lower probability in MISC DBMC systems [99], which the MISC scheme exploits to reduce the computational complexity of the Viterbi algorithm. The states with lower probability are removed from the trellis structure, see Fig. 13. Since this property is unique to MISC, the same technique cannot be applied in uncoded DBMC systems [99]. Nevertheless, for each removed state, a new connection is realized from the current state to the next one. Despite the formation of a new connection, the number of connections from the current state to the next state does not change. Thus, the reduction in computational complexity only depends on the number of removed states [99].

2) DF-MAP

Dhayabaran et al. [99] develop a decision-feedback receiver using MAP criteria that conform with the MISC perspective.

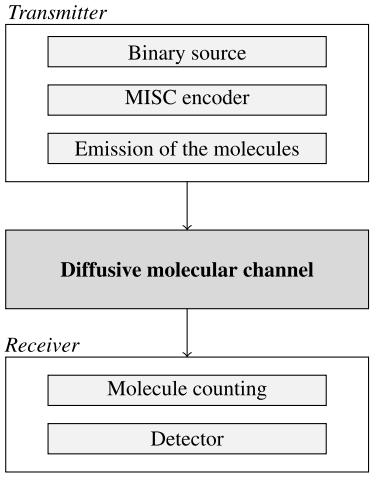
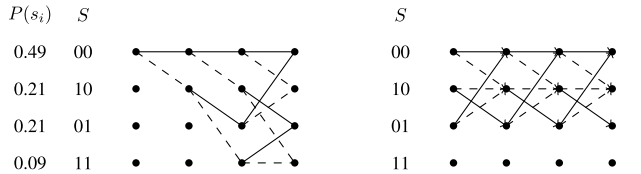


FIGURE 12. Diffusion-based molecular communication system block scheme, including the MISC encoder and detector proposed in [99].



(a) Trellis for conventional Full State Maximum Likelihood Sequence Detector

(b) Trellis in three states for RSSD

FIGURE 13. MISC DBMC system trellis diagram for $p = 0.3$ (ISI length $I = 2$), adapted from [99].

The receiver is particularly well suited for MISC DBMC systems since the encoding imposes unequal *a priori* probabilities on the source symbols [99]. The BER of the DF-MAP receiver is evaluated for Poisson and Gaussian approximations of the received signal and for different error propagation scenarios of the decision feedback, e.g., one bit error (slight) and 50% (severe) decision feedback error propagation [99]. The computational complexity of the DF-MAP receiver is similar to the conventional MAP receiver [99]. The main difference between the DF-MAP receiver and the conventional MAP receiver is the additional memory, which takes into account the decisions made in the previous time slots.

The simulations in [99] demonstrate that the performance of the RSSD of a MISC DBMC system is very close to the performance of the trellis based Full State Maximum Likelihood Sequence Detector (FS-MLSD) in terms of the average BER. Additionally, the RSSD can reduce the computing complexity by up to 50% compared to an uncoded DBMC system with a FS-MLSD receiver [99]. In MISC DBMC system simulations, the DF-MAP receiver exhibits similar BER performance as a threshold receiver and a RSSD receiver [99]. Ideally, in the case that the ISI can be completely cancelled, the DF-MAP is significantly better than the MLSD receiver [99]. Furthermore, the coded systems (RSSD and DF-MAP) achieve lower BER than the uncoded (threshold detector) system [99].

In summary, the MISC technique, also in combination with low-complexity receivers (RSSD and DF-MAP), minimizes

the ISI effect and makes an MC communication link more reliable.

C. SUMMARY AND DISCUSSION

In [98] and [99], Dhayabaran et al. derived approximations for the lower and upper bounds of the number of molecules arriving at the receiver. These bounds are applied to determine the dependence of the BER of a DBMC system on the *a priori* probability of the information to be transmitted and the length of the ISI. Dhayabaran et al. also demonstrate that the BER can be significantly reduced by lowering the fraction of bit-1s in the code word. Thus, Dhayabaran et al. provide a crucial starting point for future source coding schemes. The MISC technique with the two low-complexity receivers RSSD and DF-MAP assigns code words of numerous bits, i.e., long code word length, but few bit-1s, i.e., a low code word weight, to symbols with a high probability of occurrence, and vice versa [98], [99]. This assignment strategy reduces the ISI effects that are caused by the molecules that are transmitted for bit-1s and take an excessively long time to propagate through the channel. The MISC technique can mitigate ISI and consequently reduce the BER [98], [99]. Furthermore, the presented MISC scheme offers the possibility to construct new low-complexity receivers, providing a crucial advantage over other channel coding algorithms [99].

Limitations arise in RSSD because the ISI cannot be compensated equivalently compared to conventional detectors [99]. In addition, DBMC systems are very susceptible to a change in symbol duration. Also, all approaches in [98] and [99] have only been examined for the OOK modulation technique. Future research should address this limitation and consider a wide range of modulation techniques. Ideally, future research should develop new DBMC source coding approaches that are not limited to a particular modulation technique.

The existing MISC source coding approaches minimize the proportion of bit-1s in the code word, which can make the decoding on the receiver side more challenging due to the relatively small number of transmitted molecules (that represent the bit-1s in the commonly considered OOK modulation). Future source coding schemes should address the reduction of the bit-1s by using fixed (e.g., [199]) or adaptive demodulation thresholds (e.g., [200], [201]) on the receiver side of the DBMC system. In particular, adaptive thresholds on the receiver side can adapt to a small number of transmitted molecules, so that a reduced number of bit-1s can be demodulated successfully.

Research in source coding in DBMC has focused on a hybrid of efficiently representing the source and avoiding or mitigating channel errors (by reducing or controlling the number of transmitted bit-1s). Thus, the existing DBMC source coding approaches can be considered hybrid source-channel coding approaches. Future research should further explore hybrid source-channel coding for DBMC nanonetworks so as to achieve a controllable balance between short code words (reducing the overall transmission duration and energy) and low code word weights (reducing the ISI due to few transmitted molecules or few time slots with molecule transmissions).

IV. CHANNEL CODING IN DIFFUSION-BASED MOLECULAR NANONETWORKS: AN OVERVIEW

Sections V–VIII comprehensively survey the field of channel coding in DBMC nanonetworks. Table 5 gives an overview of all channel coding studies covered in Sections V–VIII, which we have organized into the four categories of linear block codes, convolutional codes, ISI-mitigation codes, and miscellaneous codes. Table 5 encompasses all channel coding studies covered in [63] and [65], as well as 17 additional studies. The additional studies differ from the studies covered in [63] and [65] primarily in their timeliness (appeared after 2018) and partly the topicality (cover additional topic aspects that were not covered in [63] and [65]). Throughout Sections V–VIII, we present comprehensive surveys of the various channel coding techniques in DBMC nanonetworks treating the studies covered in [63] and [65], and the 17 additional studies in an integrated manner so as to give a complete cohesive survey of the field of DBMC channel coding.

V. CHANNEL CODING: LINEAR BLOCK CODES

This section is based on the fundamentals from Section II-C and comprehensively surveys the existing research on linear block channel coding for DBMC nanonetworks. In order to provide a comprehensive survey, we survey both the studies that had previously been covered in [63] and [65] as well as the additional studies that had not previously been covered in a cohesive integrated manner.

A. HAMMING CODES

This section first surveys the Hamming code based channel coding studies that had been included in [63] and [65], i.e., the studies [128], [131], [202], [203], [204], [205]. Subsequently, we survey the additional studies [206], [207] that appeared recently.

In 2012, Leeson and Higgins in [202] and [203] were the first to propose error correction channel codes in DBMC systems using Hamming codes. They simulated simple block codes at the nanoscale to evaluate the performance of the coded communication between nanomachines. Leeson and Higgins found that a block of size 11 achieves a coding gain of approximately 1.6 dB [65]. In addition, Leeson and Higgins [202], [203] present for the first time an energy analysis of error correction codes in DBMC systems. Especially in nanoscale MC, error correction plays an important role in energy utilization [65]. In [202], Leeson and Higgins note that for future research, more powerful channel codes are needed if the range of energy effectiveness for a high BER is to be enhanced [203]. According to [202], the limited amount of energy of the nanomachines must be taken into account, so that powerful iterative decoding methods probably cannot

TABLE 5. Overview of channel coding studies in DBMC nanonetworks, further developed from [65]. In terms of complexity, the attribute "Low" describes the ease of implementation, application to molecular modulation techniques, and understanding the coding scheme. The "High" attribute, on the other hand, requires a more sophisticated implementation and more computing power. "Medium" is in between.

Category	Coding Approach	Reference	Error Correct. Cap.	Complexity
Linear Block Codes (Section V)	Hamming (Sec. V-A)	[128], [131], [202]–[207]	Low	Low
	Reed-Muller (Sec. V-B)	[133], [208]	Medium	Medium
	Reed-Solomon (Sec. V-C)	[209]–[211]	Medium	Medium
Convolutional Codes (Section VI)	Pulse Amplitude Modulation (Sec. VI-A)	[212]	Low	Medium
	Self-Orthogonal Convol. (Sec. VI-B)	[213], [214]	Medium	Medium
	Distance Metric Functions (Secs. VI-C + VI-D)	[215], [216]	Low	Medium
ISI-Mitigation Codes (Section VII)	ISI-Free (Sec. VII-A)	[134], [217]–[221]	–	Low
	ISI-Aware (Sec. VII-B)	[132], [135], [136], [222], [223]	Medium	Medium
	TCH (Sec. VII-C)	[137]	High	High
Miscellaneous Codes (Sec. VIII)	Error-Free Constant-Weight Code (Sec. VIII-A)	[190], [224]–[226]	High	Low
	Misc. Channel Error Mitigating Codes (Sec.VIII-B)			
	Single Parity Check (Sec. VIII-B1)	[227]	Low	Medium
	EGLDPC (Sec. VIII-B2)	[133]	Low	Low
	Parity Check Erasure (Sec. VIII-B3)	[228]	Low	Low
	Bin. Non-Lin. – Fair Weak Flip (Sec. VIII-B4)	[229]	High	–
	Spatial Code – Repetition MIMO (Sec.VIII-C)	[219], [230], [231]	–	Low
	Minimum Energy (Sec. VIII-D)	[204]	Low	Medium

be used. In [203], Leeson and Higgins found, that ∼1.7 dB coding gain can be achieved over a transmission distance of 1 μm. Leeson and Higgins also demonstrated, that the critical distance, i.e., the distance at which the energy gain of the code overcomes the operational costs of the code, is $\gg$30 μm [203], i.e., the coding is sensible for relatively large distances of more than 30 μm.

Also in 2012, Ko et al. [128] introduced a molecular coding distance function instead of using the already known Hamming codes. The molecular distance function from a code word x to a code word y is defined by [128]:

$$d_N^{(S)}(x, y) = -\log\left\{\sum_{\{z\in Z_N : x_z = y\}} P(x, z)\right\}, \tag{28}$$

whereby N denotes the level of crossover, S denotes the signal space, z is the shift index, $P(x, z)$ is the probability for shifting x by z, and Z_N denotes the set of all shift indices with a crossover level less than or equal to N. The shift index describes the crossover effect starting from the original code word. Furthermore, x_z denotes a shifted code, i.e., x shifted by z. The molecular coding distance function performs significantly better than the Hamming codes for channel coding in time based systems [128]. However, the molecular coding distance function approach in [128] is limited in that it can only be applied in DBMC systems with timing-based modulation.

The work of Leeson and Higgins [202], [203] was extended by further research of Hamming codes in DBMC systems that was conducted by Lu et al. [131]. Lu et al. compare non-coded and coded channels in [131], using Hamming codes up to a size of 10 [65]. Lu et al. demonstrate that Hamming codes can achieve a coding gain down to a BER of 10^{-9} compared with an uncoded system and that optimal Hamming code parameters exist [131]. Specifically, Lu et al. [65], [131] found that (32, 26) Hamming codes have the highest efficiency in terms of energy usage, while (63, 57) Hamming codes have the highest coding gain. Following Lu et al., the energy loss or saving ΔE for a coded MC system can be defined as [131]:

$$\Delta E = 2450 \times (N_u - N_c) - E_{encode} - E_{decode}, \tag{29}$$

whereby the constant 2450 characterizes the energy cost for synthesizing a molecule [49], N_u denotes the required number of molecules for achieving a prescribed BER with an uncoded MC system, N_c denotes the corresponding required number of molecules for achieving a prescribed BER with a coded MC system, while E_{encode} and E_{decode} denote the required energy for encoding and decoding.

In the following research, existing channel coding schemes in wireless communications were compared with Hamming codes [65]. Bai et al. [204] found that Minimum Energy Codes (MEC) [232] (see Section VIII-D) achieve lower BER and lower energy usage than Hamming codes. In [204], Bai et al. demonstrated that both Hamming codes and MEC achieve coding gains on the order of a few dB. Thereby, MEC offer a lower BER and a lower energy consumption compared to Hamming codes, but require longer code words [204]. However, the longer code words of the MEC are disadvantageous for DBMC systems due to the corresponding required longer transmission time.

Marcone et al. [205] present a design of a biological encoding and decoding system. In particular, they proposed a biologically modulated parity-check encoder in a cell based on molecular information [205]. For the implementation, a biological XOR is applied between the input bits. The output of the system is a block code word of the length 3 bits, modulated by OOK. With the help of SimBiology, a software tool for simulating, modeling, and analyzing biological systems, and based on synthetic biology, the proposed encoder is able to encode molecular information. The encoding is based on the serialization of a naturally parallel information [205]. Marcone et al. proposed a methodology for the design of a molecular encoder that could be replicated in future research.

Building on this system, more complex codes for encoding can be constructed [65], also for the use in future devices for the Internet of Bio-Nano Things [205].

As described in Section II-C, Hamming codes are linear block codes that specify individual errors in each block and detect up to two errors. Rai et al. [206] presented for the first time a Very-Large-Scale Integration (VLSI) implementation of the Hamming Code for MC. An encoder using Complementary Metal-Oxid-Semiconductor (CMOS) logic circuits is implemented for different values of the parity check bits. Rai et al. develop the encoder for three and five Hamming code parity check bits, which are encoded by shift registers and a two-input XOR gate. A set-reset flip-flop is used in each shift register, which can be built in CMOS logic [206]. The two-input XOR gate, which calculates the parity bits depending on the generator polynomial, can also be implemented in CMOS logic [206]. The encoder was designed and simulated in [206] using a 250 nm technology in the Tanner Electronic Design Automation (EDA) tool. The study [206] fills a gap in the MC field, as the performance of MC systems can be significantly improved, especially in terms of power consumption and delay of the circuits.

Cheong [207] introduces an improved Hamming decoder based on a so-called soft value. The simulation environment is based on a three-dimensional MC channel. Assuming the Binary Concentration-Based Shift Keying (BCSK) modulation technique, where many molecules are released for symbol 1 and a few molecules are released for symbol 0, the molecules injected by the transmitter propagate through the MC channel to the receiver via diffusion, following Fig. 3. The receiver determines the number of molecules absorbed by the receiver during the reception time slot. If a prescribed threshold is exceeded, then a bit-1 follows; if the threshold is not reached, then a bit-0 follows. If many molecules are released (in order to transmit a bit-1) in a time slot, but fewer molecules are absorbed than required to exceed the threshold, then an error occurs [207]. After successful transformation and evaluation of the number of molecules received in zeros and ones, the number of molecules in a given time slot is discarded in the hard value approach. In the soft value approach [207], the number of received molecules after symbol determination is not discarded, but used as a soft value to improve the decoding performance of the encoder. Simulations of a (15,11) and a (7,4) Hamming code in [207] demonstrate that the presented Hamming code using soft values achieves a significantly lower BER than a hard value Hamming code.

B. REED-MULLER CODES

We first survey the study [133], which was covered in [63], [65]. In [133], for the first time, cyclic Reed-Muller codes were implemented in a DBMC system and the performance compared to Hamming codes in terms of the coding gain and energy requirements. The simulations by Lu et al. [133] found a coding gain of 4.46 dB for C-RM(2,5), compared to a coding gain of 2.77 dB for $m = 5$ Hamming codes and 7.26 dB for $s = 4$ LDPC codes at a BER of 10^{-9}. For nanomachine-to-macromachine communication, C-RM(1,4), (2,4), (2,5), and (3,5) were found to be beneficial for BER levels less than 10^{-4} for all propagation distances [133]. For a propagation distance of at least 6.1 μm and a BER level higher than 10^{-4}, C-RM(3,5) can be used for the design of DBMC systems [133]. Especially for large distances between transmitter and receiver, Hamming codes achieve lower BER than Reed-Muller codes [133].

Continuing with the survey of the additional study [208] that was not covered in [63] and [65], we note that Rai et al. [208] presented for the first time a VLSI implementation of Cyclic Reed-Muller (C-RM) codes. Rai et al. [208] proposed a non-systematic C-RM encoder, which is based on shift registers and XOR gates with two inputs [208]. Thereby, non-systematic encoding means, that the information sequence is not directly embedded in the encoded sequence [233]. A CMOS logic circuit forms the shift registers and the XOR gates. The parity bit can be calculated using a two-input XOR gate. The two-input XOR gate depends on the generator polynomial of each code [208]. The generator polynomial for the C-RM (1, 3) code is $g(1,3) = x^3 + x + 1$ [208].

The C-RM (1, 3) code is decoded with a simple shift register and the two-level maximum-likelihood method [208]. A four-input XOR gate is used as the majority logic method. Rai et al. form a two-input Majority Logic Gate (MLG) using two two-input Not AND (NAND) gates. XOR gates with four inputs, for use in the two-step decoding method of majority logic, are formed by merging several XOR gates [208]. The test polynomial for C-RM is $h(1,3) = x^4 + x^2 + x + 1$. If two NAND gates are combined, the result is an MLG. The MLG can make the following, logical decision: If the majority of the inputs is a 1, then the output is also a 1, otherwise it is a 0 [208].

The encoder and decoder in [208] were simulated using 250 nm technology in the Tanner EDA tool. All circuits used CMOS logic. The approaches fill thereby the gap of VLSI implementation of C-RM systems in MC to improve the performance [208].

C. REED-SOLOMON CODES

Dissanayake et al. implement Reed-Solomon codes in [209] for error correction in MC systems. The implemented Reed-Solomon (RS) codes were compared in simulations with uncoded systems. The simulation results demonstrate that RS coded systems can significantly reduce the BER of DBMC [209]. Building on this research in [209], Dissanayake et al. in [210] demonstrate through Monte Carlo simulations that RS codes can reduce the BER in a multiuser DBMC system. Moreover, by appropriately selecting the code length n and information bit length k of the (n, k) RS code, the BER performance can be tailored to system requirements and constraints [210].

In a three-dimensional molecular multiuser communication system via diffusion, there are several transmitters

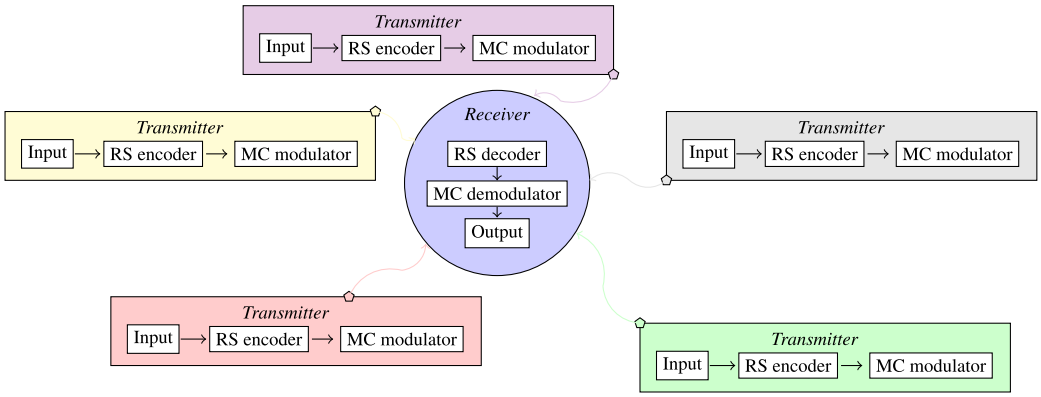


FIGURE 14. System model of partially absorbing receiver and swarm located transmitters, adapted from [211].

located in a swarm in [211] in addition to the partially absorbing receiver. The transmitters are arranged in a stochastic geometry, as illustrated in Fig. 14. Dissanayake et al. [211] propose RS error correction coding to mitigate the Multi-User Interference (MUI) and ISI, due to the high effectiveness of RS codes in combating burst and random errors (RS codes have the capability to correct more than one bit in a single code word).

RS codes are non-binary BCH codes, which are used for error correction coding for a three-dimensional MC system with a swarm of interfering transmitters in [211]. RS codes embed the redundant data into the input message to facilitate subsequent error recovery at the receiver [211]. A Galois field, analogous to Section II-C, defines $RS(n, k)$. The (n, k) encoder of the RS code expands the input message M of length k into a code word C of length n by adding redundant parity bits P of length $n - k$. In this process, the parity bits represent a multiple of the RS code generator polynomial. The RS code decoder multiplies the syndrome S of the received code word C' by a predefined verification polynomial [234]. The received code word is called erroneous if the syndrome calculation results in a non-zero value. The identified erroneous code word can be corrected by determining the polynomial of the symbol error locator in the first step using the Berlekamp-Massey algorithm [119], [120]. In the next step, the roots of the error locator polynomial are determined using the Chien search algorithm [235] and finally the error magnitudes are estimated using the Forney algorithm [108] for the overall correction processing [211]. After successful estimation of the error location and the error value, a correction can be attempted. Thereby, RS codes can correct up to $(n - k)/2$ symbol errors. Consequently, the higher the redundancy, the more errors can be corrected. For further information on RS codes, see [234] and [236].

Dissanayanake et al. [211] implement RS codes in an MC system and validate the results using Monte Carlo simulations. They demonstrate good agreement between the equations established analytically in [211] and the simulation results. The simulations reveal that systems with $RS(16, 2)$ and $RS(10, 2)$ achieve significantly lower BER than systems without channel coding [211]. On the other hand, systems with $RS(8, 4)$ exhibit higher BER. The reason is that the error correction capacity $(n - k)/2$ of $RS(8, 4)$ codes cannot satisfy the conditions of the MC system [211]. Under these conditions, channel coding degrades the BER performance of the system due to the additional redundant information. Overall, the simulation results demonstrate that RS codes are capable of reducing the BER; however, the code length n and message length k of the $RS(n, k)$ code must be matched carefully to the MC system settings and the required BER.

In summary, RS codes provide flexibility in the capacity of error correction, due to matching the different parameters of the RS coder to the MC system requirements [211].

D. SUMMARY AND DISCUSSION

In 2012, Leeson and Higgins were the first to propose error correction channel codes in DBMC systems using Hamming codes [202], [203]; these seminal studies were the starting point for Hamming codes in DBMC systems. In further research, comparing Hamming codes, cyclic Reed Muller codes [237], and Euclidean Geometry Low Density Parity Check (EGLDPC) codes [110], Lu et al. [133] found that Hamming codes are well suited for DBMC systems with long distances [65]. Cheong demonstrated in 2020 that Hamming codes using soft values achieve a significantly lower BER than hard-valued Hamming codes [207].

For Reed-Muller codes, Lu et al. [133] indicated that in comparison, Hamming codes achieve lower BER than the Reed-Muller codes in DBMC systems.

With the proposal of RS coded systems for DBMC, Dissanayake et al. demonstrated that RS coded systems can significantly lower the BER of DBMC systems [209]. Moreover, by appropriately selecting the code length n and information bit length k of the (n, k) RS code, the BER performance can be tailored to MC system characteristics [210]. The greater flexibility of RS codes in error correction capacity allows the RS encoder to be adapted to MC system requirements [211].

Generally, linear block codes reduce the BER of DBMC systems compared to uncoded DBMC systems, and can often be better adapted to the system requirements of an MC

system than convolutional codes (see Section VI) to further reduce the BER. Overall, linear block codes thus represent a promising starting point for channel coding in DBMC systems. Table 5 indicates that for linear block channel codes in DBMC systems, the error correction capability is generally proportional to the complexity of the linear block code: Hamming codes have low complexity (which is desirable), but also achieve only low error correction capability (which is undesirable). On the other hand, RM and RS codes achieve medium error correction capabilities, but incur also a medium level of complexity.

The three linear block codes that have been evaluated in simulations of DBMC systems (Hamming, Reed-Muller, and Reed-Solomon codes) are channel codes known from wireless EM communication, which have been transferred to the modulation techniques of DBMC systems. Linear block codes are often inadequate for MC systems. For example, due to the linear nature of the Hamming codes, only one error can be detected. However, in DBMC systems, multiple bit flips often occur due to the occurrence of ISI (similarly multiple bit flips are common in EM communication due to the EM propagation characteristics, e.g., fading). Therefore, Hamming codes are not widely used in wireless systems, and have similarly only limited usefulness in DBMC. Generally, linear block codes are challenged by the error characteristics in wireless EM systems and in DBMC systems, and more advanced codes are needed for achieving better performance.

VI. CHANNEL CODING: CONVOLUTIONAL CODES

This section first surveys the studies [212], [213], [214] (which were included in [63] and [65]), including Self Orthogonal Concolutional Codes (SOCCs). Subsequently, the new convolutional channel coding studies (not covered in [63] and [65]) are surveyed in detail, classified into crossover-distance metric for DBMC systems (Section VI-C) and asymmetric-distance metric for DBMC systems (Section VI-D).

A. PULSE AMPLITUDE MODULATION (PAM)

To evaluate the performance of channel coding in diffusive MC systems, Pulse Amplitude Modulation (PAM) for concentration-encoded convolutional channel coding was introduced by Mahfuz et al. [212] using pulse amplitude levels $M = 1, 2, 4$. Different parameters, e.g., the transmission rate or the communication range, are varied. For short- and medium-range communication, simulation results indicate that convolutional codes with $M = 1$, i.e., OOK, outperform uncoded transmission; for long-range communication, convolutional codes are essential to achieve practically acceptable BER levels [212]. Mahfuz et al. [212] also highlight that OOK is more suitable for DBMC systems than PAM.

Furthermore, the simulation results in [212] demonstrate, that the range of communication can be increased by implementing convolutional codes on the transmitter side of the DBMC system. Increasing signal strength and communication range are desired due to the major challenge of signal attenuation in MC systems. Even for short communication ranges, Mahfuz et al. showed that using a signalling scheme with multiple levels (M-ary) increases the BER. In addition, convolutional code designs with a larger alphabet of signalling symbols/constellations are not recommended for DBMC systems because of the increased ISI.

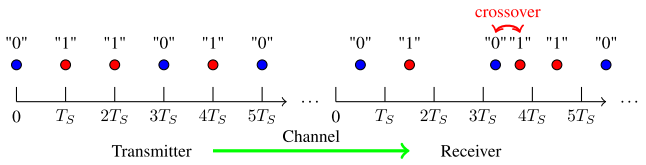


FIGURE 15. Binary Molecule Shift Keying (MoSK) modulation for molecular communication system via diffusion, adapted from [215]. The transmitter encodes the bit sequence [0,1,1,0,1,0] into two types of molecules (blue/red). The receiver decodes the sequence [0,1,0,1,1,0] due to a level-1 crossover at the third bit.

B. SELF-ORTHOGONAL CONVOLUTIONAL CODES

Self Orthogonal Concolutional Codes (SOCC)s with a ''majority-logic decoding scheme'' were first introduced for DBMC systems by Lu et al. [213]. SOCCs are a subcategory of convolutional codes that can be used simply by implementing a logic decoding scheme. The implementation was evaluated in comparison to cyclic Hamming codes [234]. In general, simulation results demonstrate that SOCC have significantly better energy efficiency (for a low BER requirement) [213]. Lu et al. [213] demonstrated a coding gain of 2.28 dB for a (3, 2, 13) SOCC at a BER level of 10^{-9}, compared to a coding gain of 1.76 dB for $m = 5$ Hamming codes. Additionally, Lu et al. [213] suggest using a (3, 2, 13) SOCC for a transmission distance greater than 28 μm, also due to the easy implementation.

In [214], Lu et al. expand their research: If a low BER is required, the (2, 1, 17) SOCC has a higher energy efficiency than Hamming codes [214]. Lu et al. [214] also evaluate the critical distance as a performance metric of a DBMC system. The higher the tolerable operating BER, the longer the critical distance [214]. Lu et al. also indicate that for nanomachine-to-macromachine communication (and vice versa for macromachine-to-nanomachine communication), the critical distance decreases compared to nanomachine-to-nanomachine communication [214].

C. CROSSOVER-DISTANCE METRIC FOR DBMC

As discussed in Section VII, ISI is a primary factor in the BER degradation of a DBMC system. In [215], a crossover-distance was introduced to determine the distance between the received bit sequence and the probably transmitted bit sequence. Based on this crossover-distance, convolutional codes using the introduced crossover-distance were developed by Li and Li [215] and evaluated for an MC system using a transmitter and receiver pair [215]. The used modulation technique is binary MoSK [238], [239]. MoSK transmits one molecule type for a bit-1 and a different molecule type for a bit-0, see Fig. 15, whereby T_S denotes the time interval

for the release of the molecules for a given bit. Each particle diffuses independently in the one-dimensional channel, and is absorbed upon reaching the receiver. The receiver detects the type of each particle and the particle is subsequently removed from the system.

1) CROSSOVER TRANSITION PROBABILITY

The phenomenon of the order of the incoming particles differing from the order of transmitted particles (and the order of the decoded bit sequence consequently differing from the original bit sequence), is called crossover [215]. A level-l crossover denotes the crossover of a bit, i.e., the delayed arrival of a bit, after l bits. For example, Fig. 15 illustrates a level-1 crossover, while the example [010001] $\rightarrow$ [000101] has a level-2 crossover for the second bit in the input bit stream (first bit-1 in the input bit stream) [215]. No background noise is assumed in this case in [215]. Generally, in coding theory, the distance between two code words is an important concept which can be used to evaluate the performance of the code, to generate code words or to develop efficient decoding algorithms. The crossover transition probability $P_c(l)$, i.e., a level-l crossover at the i^{th} bit, is [215]:

$$P_c(l) = Pr[t_i > lT_S + t_{i+l},\ t_i < (l+1)T_S + t_{i+l+1},\ t_i < (l+2)T_S + t_{i+l+2}, \ldots], \quad (30)$$

whereby t_i denotes the travel time of the i^{th} particle, and T_S the fixed duration of the time interval of an adjacent transmission [215]. The bracket expression means that the i^{th} particle arrives at the receiver side after the $(i+l)^{th}$ particle, but before the $(i+l+1)^{th}$, $(i+l+2)^{th}$, ... particles [215]. Complementarily, level-0 describes the absence of crossover; accordingly,

$$P_c(0) = 1 - \sum_{l=1}^{\infty} P_c(l). \quad (31)$$

2) CROSSOVER-DISTANCE

Li and Li [215] measure the distance between the incoming and outgoing binary sequences of the MC system with the crossover vector v_i, which captures the state of crossover between two binary sequences. Let $x = [x_1, x_2, \ldots, x_N]$ and $y = [y_1, y_2, \ldots, y_N]$ represent the input and output bit sequences of an MC channel, respectively. The crossover vector $v = [v_1, v_2, \ldots, v_N]$ is defined as [215]:

$$v_i = \begin{cases} l & \text{if for bit } x_i \text{ a level-}l \text{ crossover occurred} \\ 0 & \text{if no crossover for bit } x_i \text{ occurs.} \end{cases} \quad (32)$$

For example, for the input sequence $x = [0, 1, 0, 0, 0, 1]$ and the output sequence $y = [0, 0, 1, 0, 1, 0]$, the crossover vector is $v = [0, 1, 0, 0, 1, 0]$, with two level-1 crossovers occurring at the second and fifth bits of x. For a given crossover vector v of the input sequence x and the output sequence y of the channel, the distance is described based on the probability of occurrence of crossover in the crossover vector [215]. The

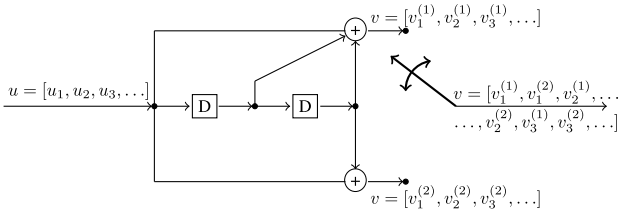


FIGURE 16. Convolutional encoding structure, adapted from [215] with constraint length $K = 3$ and code rate $R = 1/2$. For the technical scheme of a convolutional encoder (basic general convolutional encoding principle), the reader is referred to Fig. 8.

probability that x is transformed to y, with the occurrence of $v = [l_1, l_2, \ldots, l_N]$ is [215]:

$$P_v = Pr[v : x \rightarrow y] = \prod_{k=1}^{N} P_c(l_k). \quad (33)$$

Li and Li [215] define the crossover-distance D_v as:

$$D_v(x, y) = -\log P_v + N \log P_c(0) \quad (34)$$

$$= -\sum_{k=1}^{N} \log \frac{P_c(l_k)}{P_c(0)} \quad (35)$$

$$= \sum_{k=1}^{N} W_c(l_k), \quad (36)$$

whereby $W_c(l_k)$ denotes the distance contribution for level-l_k crossover. Li and Li define $W_c(l_k)$ as [215]:

$$W_c(l_k) = \log P_c(0) - \log P_c(l). \quad (37)$$

For $l \geq 1$, Li and Li assume that $P_c(0) > P_c(l)$ such that $D_v(x, y) \geq 0$. The crossover-distance $D_v(x, y) \geq 0$ has the following properties [215]:

- If no crossover effect occurs, then $v = 0$, $y = x$, and $D_v(x, x) = 0$.
- If the input x_1 is transmitted with a higher probability than x_2, then $D(x_1, y) < D(x_2, y)$ for the given channel output y.
- If the output y_1 is received with a higher probability than y_2, then $D(x, y_1) < D(x, y_2)$ for the given channel input x.
- If $D_{v_1}(x, y) < D_{v_2}(x, y)$, then $P_{v_1} < P_{v_2}$. The minimum crossover-distance between x and y is estimated using the minimum crossover vector v_0 between x and y.

3) CODING SCHEME FOR CROSSOVER-DISTANCE

The crossover-distance presented in Sections VI-C1 and VI-C2 can be used to develop a new metric for coding convolutional codes, especially decoding. First, the encoding scheme is similar to the scheme for classical convolutional codes [215]. Fig. 16 shows a simple encoding structure of (7, 5) convolutional codes (constraints of length $K = 3$ and code rate $R = 1/2$).

Typically, the classical decoding algorithm for convolutional codes is a Viterbi algorithm. The Viterbi algorithm uses the Hamming distance to determine the distance between

the transmitted and the received sequence for hard decision decoding and the Euclidean distance for soft decision in the additive white Gaussian noise channel [215]. In the decoding algorithm presented in [215], the crossover-distance is used as a metric for the Viterbi algorithm. By deliberately minimizing the crossover-distance between the received sequence and the probably transmitted sequence, the maximum-likelihood estimation of the transmitted bits can be determined.

The decoding scheme in [215] is compared in terms of BER with known channel coding schemes, such as the Hamming code, ISI-free codes, and convolutional codes based on the Hamming distance. To compare the performance of different channel codes on the same basis, the BER of each code has the same information bit throughput. The simulation results demonstrate that the BERs of the convolutional codes based on the Hamming distance, including Hamming codes, are not significantly lower than the uncoded systems [215]. The redundant bits in the transmitted sequence are not able to correct the ISI errors. On the other hand, the convolutional codes based on the crossover-distance achieve the same or slightly lower BER than the ISI codes, for the same throughput [215].

D. ASYMMETRIC-DISTANCE METRIC FOR DBMC

In [216], Li partially builds on the findings from [215] to develop an asymmetric-distance decoding scheme for DBMC systems. Time-based release of molecules is used as the modulation technique for encoding [238]. The molecules representing a bit-1 and a bit-0 are released at different times. The model representation of the modulation technique is illustrated in Fig. 17. Each information bit corresponds to exactly one molecule. The molecules diffuse through the three-dimensional environment.

1) MODULATION AND DEMODULATION SCHEME

A maximum-likelihood estimator (from [240]) is used in [216] to obtain the information from the times of the received molecules. The transmission time of the molecules is represented by the variable X, where $X \in \{0, t_e\}$ within a given time slot (symbol interval) of duration T_S. If a molecule is released at the time instant $X = 0$, it is a bit-0; at time $X = t_e$, a bit-1 [216]. If several information symbols are transmitted via the channel, the molecule representing a bit-0 is released at the beginning of the respective time slot, i.e., at time instant 0, or T_S, or $2T_S$, and so on. The molecule for a bit-1, on the other hand, is released at time instant t_e, or T_S+t_e, or $2T_S + t_e$, and so on.

On the receiver side, the arrival time of the information bits b_i is marked as y_i and the corresponding vector $y = [y_0, y_1, \ldots, y_{N-1}]$, respectively. If $X_i \in \{iT_S, iT_S + t_e\}$, then the following distinction can be made [216]:

$$\hat{X}_{i,ML} = \begin{cases} iT_S & \text{if } y_i < \theta_i \\ iT_S + t_e & \text{if } y_i > \theta_i. \end{cases} \tag{38}$$

However, the vector $y = [y_0, y_1, \ldots, y_{N-1}]$ cannot be determined directly due to ISI, but only indirectly through the sorted vector z.

$$z = [z_0, z_1, \ldots, z_{N-1}] = sort([y_0, y_1, \ldots, y_{N-1}]). \tag{39}$$

The *sort(·)* function sorts the elements of the vector in ascending order. At the receiver, the vector z is used to approximate y [216]:

$$\hat{X}_{i,ML} = \begin{cases} iT_S & \text{if } z_i < \theta_i \\ iT_S + t_e & \text{if } z_i > \theta_i. \end{cases} \tag{40}$$

2) ASYMMETRIC-DISTANCE

For multiple uses of an MC channel, $P(1|0) \neq P(0|1)$, where $P(1|0)$ is the probability of having decoded a bit-1 but sent a bit-0, and vice versa for $P(0|1)$ [216]. This distinguishes the MC channel from convolutional channels, such as the additive white Gaussian noise channel [216]. Li and Li [216] establish an asymmetric metric that determines the distance from the transmitted code word to the received code word in a timing molecular channel. The asymmetric-distance is denoted from a transmitted bit-0 to a received bit-1 as $D(0, 1)$, and from a transmitted bit-1 to a received bit-0 as $D(1, 0)$, following

$$D(0, 1) = -\log[P(1|0)], \quad D(1, 0) = -\log[P(1|0)]. \tag{41}$$

Based on $P(0|0) = 1 - P(1|0)$ and $P(1|1) = 1 - P(0|1)$, the asymmetric-distance from a transmitted bit-0 to a received bit-0, denoted as $D(0|0)$, and from a transmitted bit-1 to a received bit-1, denoted as $D(1|1)$, can be determined to be [216]:

$$D(0, 0) = -\log[1 - P(1|0)], \quad D(1, 1) = -\log[1 - P(1|0)]. \tag{42}$$

Li and Li additionally define the operation $x.y$ and $\bar{x}$ for the code words $x = [x_1, x_2, \ldots, x_N]$ and $y = [y_1, y_2, \ldots, y_N]$ as follows [216]:

$$x.y = [x_1 y_1, x_2 y_2, \ldots, x_N y_N] \tag{43}$$

$$\bar{x} = [1 + x_1, 1 + x_2, \ldots, 1 + x_N]. \tag{44}$$

The '+' operation is the *mod 2 operation*. The weight $\omega(x)$ of a code word x equals the number of bit-1s in x. The asymmetric-distance from a transmitted code word x to a received code word y is then defined according to [216] as:

$$D(x, y) = -\log p(y|x) \tag{45}$$

$$\begin{aligned} D(x, y) = {} & -\omega(\bar{x}.\bar{y}) \log[1 - P(1|0)] \\ & -\omega(\bar{x}.y) \log[1 - P(1|0)] \\ & -\omega(x.\bar{y}) \log[1 - P(0|1)] \\ & -\omega(x.y) \log[1 - P(0|1)]. \end{aligned} \tag{46}$$

In a timing molecular channel, the asymmetric-distance satisfies $D(x, y) \neq D(y, x)$ [216]; in contrast, the Hamming distance satisfies $D(x, y) = D(y, x)$.

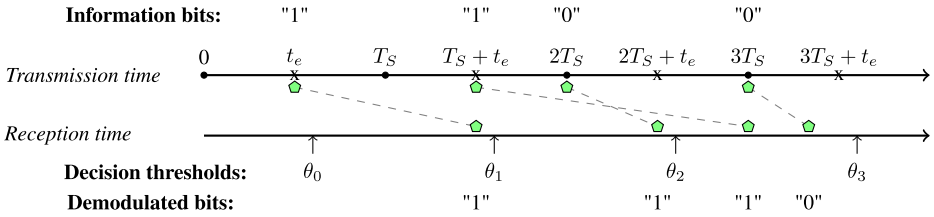


FIGURE 17. Timing diagram for time modulation of the molecules, adapted from [216]. Thereby, T_S denotes the symbol interval and θ_0, θ_1, θ_2, and θ_3 are the decision thresholds. The channel is called a timing molecular channel [216], which assumes that exactly one molecule is released in each symbol interval (of duration T_S), either at time 0 for a bit-0 or at time t_e for a bit-1. If the molecule released in a given symbol interval is received before the corresponding decision threshold time θ (at the receiver), then a bit-0 is decoded; for a later received molecule, a bit-1 is decoded. Thus, the third released molecule (at time $2T_S$) is interpreted as arriving after the threshold θ_1 and thus as a "1" (specifically, as the second transmitted bit, an example of the ISI phenomenon); while the second released molecule (at $T_S + t_e$) is interpreted as arriving after θ_2, i.e., as a "1" (third transmitted bit). The fourth transmitted bit (at $3T_S$) is correctly interpreted as arriving before threshold θ_3 as a "0" (fourth transmitted bit).

3) CODING SCHEME FOR ASYMMETRIC-DISTANCE METRICS

Based on the asymmetric-distance, Li [216] develops a decoding scheme with hard decision decoding. More specifically, the asymmetric distance is used as a metric for the Viterbi algorithm. By minimizing the distance from the likely transmitted sequences to the received sequence, the maximum-likelihood estimation of the transmitted bits can be achieved. Furthermore, when the asymmetric distance is used as a metric, the timing molecular channel is a binary-input-binary-output channel, which is similar to the Hamming distance, so that a comparison is possible [216].

The simulation results in [216] indicate that the BER of the convolutional code based on the asymmetric-distance is significantly lower than the BER of an uncoded system when the BER magnitude is lower than 10^{-2}. For reliable communication, a magnitude of 10^{-5} is typically required. The coding scheme of the convolutional codes from [216], following Fig. 16, can therefore significantly increase the information throughput [216]. Convolutional codes with asymmetric-distance also use a lower constraint length, which reduces the decoding complexity [216]. Consequently, the asymmetric distance in a Viterbi algorithm for convolutional codes significantly increases the reliability of a DBMC nanonetwork [216]. In future research, further modulation techniques should be adapted to the asymmetric-distance metric.

E. SUMMARY AND DISCUSSION

Convolutional channel codes in DBMC systems can be divided into techniques based on the crossover-distance and techniques based on the asymmetric-distance metric. Thereby, the simulation results of Li and Li [215] demonstrate that convolutional codes based on the Hamming distance (including Hamming codes) with constraint lengths $K = 3$, 5, 7 achieve about the same (or slightly higher) BER as uncoded systems. The redundant bits of the transmitted bit sequence are often unable to correct the errors that occur through ISI. In contrast, the crossover-distance-based convolutional codes presented in [215] achieve similar or even lower BER than other ISI codes in simulations.

Based on the results from [215], Li [216] develops an asymmetric distance on which a modulation and decoding scheme is built. A special feature is the time-based release of the molecules as the used modulation technique, i.e., the molecules representing a bit-0 or a bit-1 are released at different times. Thus, this approach differs from the BCSK and MoSK schemes often used in the literature (e.g., [132], [215], [238], [239], see Section VII and Section VII-B). The simulations demonstrate that the convolutional code based on the asymmetric-distance significantly lowers the BER compared to an uncoded system [216], and therefore significantly increases the information throughput [216]. Hybrid modulation techniques [51], e.g., modulation techniques that combine timing-based modulation with molecule type-based modulation, should be adapted to the asymmetric distance metric in future research.

For an MC system with the asymmetric distance metric based timing, see Fig. 17, using in addition to time-slotting two different molecule types could enhance the system performance. More specifically, the release of two molecule types, e.g., molecule type A at time iT_S and molecule type B at time $i(T_S + t_e)$, can reduce the noise since molecules of type A remaining in the channel do not interfere with the reception of molecules of type B in the following time slot. Moreover, future research could explore quadrature (2×2) modulation based on the time slots (iT_S and $i(T_S + t_e)$) and molecule types (A and B).

Overall, as Table 5 summarizes the convolutional coding approaches for DBMC systems are of medium complexity, but differ in their error correction capabilities, since in particular in [215] the focus is on distance metric functions. Moreover, Table 5 indicates that relatively few studies to date have examined convolutional codes for MC systems; thus, overall, convolutional codes are understudied for DBMC systems to date.

VII. CHANNEL CODING: ISI-MITIGATION

This section first surveys the ISI-mitigation channel coding studies [217], [218], [219], [220], [221] that were covered in [63] and [65]. Then, we survey in detail the studies that were not included in [63] and [65]. Specifically, we survey in detail the ISI-free codes in Section VII-A, the ISI-aware

codes in Section VII-B, and the Tomlinson, Cercas, Hughes (TCH) codes in Section VII-C.

A. ISI-FREE CODES

This subsection introduces ISI-free codes in DBMC and explains how they work. In [219], Yeh et al. establish a new molecular distance metric and evaluate its operation in MC systems. Based on the presented metric, new channel coding techniques and schemes can be constructed. For this purpose, Yeh et al. distinguish two types: ISI-free code construction and ISI cancellation [219]. The ISI cancellation approach can significantly reduce the BER [219]. The proposed ISI-free code has the lowest BER in most cases compared to traditional convolutional codes [219]. The designed channel codes, based on the minimum distance criterion with molecular distance metric, achieve a significant gain over using the Hamming distance in MC [219].

Using MoSK modulation [238], [239], Shih et al. [217] proposed an ISI-free coding scheme for diffusion channels. Shih et al. designed a (4, 2, 1) ISI-free code, where (n, k, l) denotes a block code of length n, information sequence length k, and a crossover of $l = 1$ (level-1-permutation). Shih et al. counteract the crossovers among successive code words by ending the preceding code word and starting the next code word such that the successive code words are connected either with $2l = 2$ zeros or ones. Crossovers within a code word are countered by ensuring that all code word permutations with $l = 1$ crossovers are distinct, see [217] for details. Furthermore, Shih et al. revisit the repetition codes and evaluate repetition codes in a diffusive channel [217]. Repetition codes repeat a message over a noisy channel several times. Repetition codes are based on the assumption that the channel corrupts only a few of these repetitions. The receiver detects errors as changing received sequences; the receiver can also correct errors, since the most frequent sequence is most likely the transmitted sequence. Simulation results in [217] demonstrate that for transmitter-receiver distances of 100 μm and 500 μm, ISI-free codes achieve lower BER than repetition codes. For a transmitter-receiver distance of 20 μm, ISI-free codes give higher BER than repetition codes [217]. Simulations also demonstrate that (4, 2, 1) ISI-free codes achieve lower BER than repetition codes with a high code rate (1/3), but give higher BER than repetition codes with a small code rate (1/5).

In [218], Shih et al. extend their approaches to (n, k, l) ISI-free codes for crossovers up to level 5 ($l = 2, 3, 4, 5$). In addition, Shih et al. [218] present the (n, k, l, s) ISI-free codes in which the code words have at least s initial and l final identical bits, instead of the symmetric l bits at both ends of the ISI-free code. The simulations in [218] demonstrate that the (n, k, l, s) ISI-free codes are computationally simpler than the (n, k, l) ISI-free codes. Furthermore, through mathematical analysis of the approximated crossover probability functions and the exponentially-decaying character of these functions, Shih et al. demonstrate that the ISI-free index is the dominant factor that governs the BER performance [218]. The ISI-free index is defined as $(l + 1)R$, whereby R denotes the code rate, i.e., the ratio between the information sequence length k and the block code length n, i.e., $R = k/n$ [218]. For example, for the ISI-free (4, 2, 1) code, the ISI-free index is $(l + 1)R = (1 + 1)2/4 = 1$. For future research, Shih et al. [218] suggest to modify the ISI-free codes to more than two different molecule types.

Finally, Qiu et al. [220] compare ISI-free codes with Hamming codes and cyclic Reed-Muller codes. The results indicate a significantly lower BER of the ISI-free codes with a simultaneously lower coding rate [65], [220]. In addition, Qiu et al. [220] demonstrated on a real testbed with mobile robots equipped with molecular transceivers that the transposition of bits is more damaging than the additive noise from random molecule arrivals at the receiver side. Based on the experimental testbed results, Qiu et al. consider positional-distance codes. Specifically, Qiu et al. [220] review state-of-the-art positional-distance codes and implement positional-distance codes on the robot. Thereby, Qiu et al. used a step-by-step approach of Shih et al. [218] to construct a code for reducing bit transposition errors. The constructed (n, k, l) ISI-free code is a proof-of-concept positional-distance code with block length n, message length k, and parameter l denoting the level of transposition that can be corrected. Whereby, the ISI-free codes primarily target the correction of the transposition; the removal of the ISI is a side effect [218]. The considered (4, 2, 1) ISI-free code uses Hamming weight features (distinct Hamming weight [241]), decreasing the effect of transposition errors in a code word. The level-l bit transposition errors across code words can be corrected by ensuring that for two given consecutive code words, the first l bits of the latter code word are the same as the last l bits of the first code word [218]. Simulation results demonstrate that the considered positional-distance codes significantly improve the performance in comparison to Hamming distance codes [220]. Based on real testbed results, Qiu et al. demonstrate that the BER is independent of the starting point of the robots [220].

Arjmandi et al. [134] consider ISI-free time-slotted modulation schemes, which consider modulations with small delay and thus a short block length. In principle, to prevent ISI, the transmitter has to ensure that the transmissions of molecules of the same type are separated by sufficiently long time intervals. Then, the transmissions do not lead to any superposition, i.e., avoid ISI.

The study [134] considers the following setup: Time is divided into fixed-duration time slots. At the beginning of each time slot, the transmitter either releases molecules or does not release molecules. In [134], the information is encoded only in the respective molecule type, but not in the number of released molecules. However, an extension to amplitude modulation would be conceivable for future research [134]. To prevent ISI at the receiver side, there must be a sufficiently large delay between two molecules of the same type. The amount of the delay depends on the communication medium [134]. The receiver also has receptors

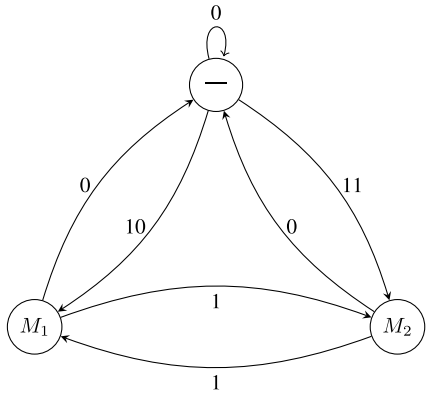


FIGURE 18. ISI-free sequences Moore machine state diagram (with states corresponding to output, i.e., released molecules) of the symbol set $M = \{-, M_1, M_2\}$ and modulation strategy, adapted from [134]. Assuming we are in state '–', then '–' is sent when a 0 is to be transmitted as the next bit. If the next two bits in the bit sequence to be transmitted are 10, then M_1 is sent for the first bit-1; in contrast, for 11, M_2 is sent for the first bit-1 and M_1 for the second bit-1. Generally, in a sequence of successive bit-1s, M_2 is transmitted for the bit-1s in the odd time slots, and M_1 for the bit-1s in the even time slots. An exception is a bit-1 surrounded by 0s.

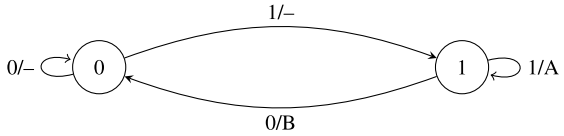


FIGURE 19. Molecular Transition Shift Keying (MTSK) Mealy machine state diagram (with labels on state transition arrows indicating "input/output") for the encoder, adapted from [132]. The release of the molecules depends on the succession of bits in the input bit stream. If a bit-0 (currently in state "0") is followed by a bit-1 or again by a bit-0, then no molecules are released. If a bit-1 (state "1") is followed by a bit-1, type-A molecules are released; if a bit-1 is followed by a bit-0, type-B molecules are released.

that can detect and identify the different types of molecules over time.

An example of ISI-free modulation is the Molecular Concentration Shift Keying (MCSK) modulation scheme [242]. MCSK applies the symbol set $M = \{-, M_1, M_2\}$. Thus, as illustrated in the state diagram in Fig. 18, in binary MCSK, information bit '0' is mapped to the symbol '–'; whereas, bit-1 is mapped to the symbol M_1 in even time slots of the successive bit-1, and to M_2 in odd time slots of the successive bit-1. An exception is an isolated 1, surrounded by 0s, i.e., the bit sequence '010', which is mapped to $\{-, M_1, -\}$. In contrast, the sequence '0110' is mapped to $\{-, M_2, M_1, -\}$; whereby the first 1 is, within the bit sequence of the 1s, in the first, i.e., odd position, mapped to the release of M_2, followed by M_1 for the 1 in the second (even) position. Analogously, the bit sequence '011110' is mapped to $\{-, M_2, M_1, M_2, M_1, -\}$. Thereby, quaternary and higher modulation approaches can be applied in future research if different concentrations (levels) of the different molecules are used [134].

Building on the BCSK modulation scheme, Arjmandi et al. [134] developed ISI-free codes with limited delay and decoder buffer size. If there is no limit of the delay between the release of the molecules, the ISI-free capacity is approximately as large as the constrained coding capacity [134]. The constrained coding capacity involves block lengths approaching infinity and complex coding schemes. However, block lengths approaching infinity and complex coding schemes are unsuitable for nanomachines. Rather, in practical DBMC systems, it is necessary to assign ISI-free sequences of the symbols with a limited length to the input bits, i.e., to use codes with a short block length. The reasons are the limitations of memory resources and that long delays are not practical in decoding for some applications [134]. According to the definition of Arjmandi et al., the delay constraint d_c is the delayed decoding of the i^{th} bit of the binary input string b_i by up to d_c channel uses. Consequently, no code word with a length greater than d_c should be used [134]. In addition to the transmission of the bit b_i, the code word is decoded within d_c time slots. An ISI-free code of length n and block length constraints consists of [134]:

- An encoder function $\mathfrak{E}$ that maps the binary input sequence $b_i = \{b_1, \ldots, b_n\}$ to the ISI-free sequence $s_i = \{s_1, \ldots, s_n\}$. The number n of output symbols depends on the input sequence.
- A decoder function $\mathfrak{D}$ for time step j, which takes over the decoding and maps the sequence into the symbols $s_i = \{s_1, s_2, \ldots, s_{\tau_{i+1}}\}$ and finally into the bits $b_j = \{b_{l_{i-1}+1}, \ldots, b_{l_i}\}$, whereby l_i denotes the number of the input bits which were decoded in time step i, and τ_i denotes the number of symbols that are transmitted during this step. More specifically, given that the input bits $(b_1, \ldots, b_{l_i})$ have already been encoded into the symbol string $(s_1, \ldots, s_{\tau_i})$, the encoder $\mathfrak{E}_i$ takes in new bits $(b_{l_i+1}, \ldots, b_{l_{i+1}})$ and maps them to the output symbol string $(s_{\tau_i+1}, \ldots, s_{\tau_{i+1}})$ with a length that is less than d_c, e.g., the difference between τ_{i+1} and τ_i is less than or equal to d_c.

The algorithm for constructing the codes is based on the approach of using a Moore machine state diagram of maximum depth d_c to represent a practical ISI-free modulation family, analogously to Fig. 18, see [134] for details. Furthermore, the modulator develops the ISI-free sequence by transferring the modified graph to the binary input sequence and transfers the associated symbols [134]. The demodulator decodes the transmitted code word by removing the same modified graph from the received symbol sequence [134]. The results demonstrate that the presented scheme is close to the optimum, and closely approximates the constrained coding capacity [134].

B. ISI-AWARE CODES

1) PHYSICAL-LAYER ISI-AWARE CODES: MOLECULAR TRANSITION SHIFT KEYING (MTSK)

Both binary MoSK and BCSK are inefficient in terms of energy consumption and prevention of ISI at high data rates [132]. Tepekule et al. address exactly this limitation by developing a transmitter-based modulation technique for DBMC systems, called MTSK, that extends a modulation technique towards a coding technique in DBMC systems.

MTSK can increase the data rate without additional occurrence of ISI. The communication model follows Fig. 3, i.e., it consists of a point source and a spherical receiver in a three-dimensional environment.

MTSK is a transmitter-based ISI-mitigation technique, which is energy efficient and mitigates the ISI effects by using two different molecule types [132]. For a progressive transmission of a bit sequence, difficulties in detecting a bit-0 after many consecutive bit-1s occur in MC systems, due to the ISI and using OOK or BCSK. The bit sequence $b_i^5 = [1, 1, 1, 1, 0]$ serves as an example. The receiver may not be able to detect $b_5^5 = 0$ due to the ISI caused by $b_4^5 = 1$ (and the preceding bit-1s) [132]. In MTSK, bit-0s are represented by the absence of molecules, i.e., no molecules are released. The bit-1s, on the other hand, are encoded by two different types of molecules, namely types A and B. Thereby, a constant number $N_A = N_B = N$ of molecules is released. The choice of the corresponding molecule type depends on the following symbol of the bit sequence [132].

As illustrated in the MTSK encoder state diagram in Fig. 19, MTSK operates as follows: For a bit-0, no molecules are released. For a bit-1, molecules of type A are released, unless the corresponding bit is the only or the last bit in a subsequence from bit-1s [132]. For example, following Fig. 19, starting in state 0, the sequence $b_i^7 = [0, 1, 1, 1, 0, 1, 0]$ is transmitted as the encoded sequence $m_i^7 = [-, -, A, A, B, -, B]$. The release of the molecules depends on the following input bits. If a bit-0 is followed by a bit-1 or again by a bit-0, then no molecules are released. If a bit-1 is followed by a bit-1, then type-A molecules are released; finally, if a bit-1 is followed by a bit-0, then type-B molecules are released. Monte Carlo simulations demonstrate that the modulation technique presented in [132] significantly reduces the BER, and outperforms both the BCSK and binary MoSK modulation techniques.

In addition, Tepekule et al. [132] present an analytical technique to determine the optimal threshold of a DBMC system using BCSK or MoSK. Until that time (i.e., until 2014), thresholds were determined empirically, i.e., by a long sequence of pilot symbols. However, if subsequently the system parameters change, such as the temperature, the diffusion coefficient, or the distance between transmitter and receiver, then the empirical process has to be repeated [132]. Tepekule et al. [132] propose to avoid this repetition of the empirical process by implementing a least mean squares regression for the first 25 threshold values; whereby, the least mean squares regression is essentially a function of the dynamics of the diffusion channel. The reason for using the regression for only the first 25 threshold values is the increasing computational power needed for sequences of increasing length. By considering only the first 25 threshold values, the optimal thresholds can be determined in a computationally efficient manner, irrespective of the sequence length. Monte Carlo simulations in [132] indicate a reduced BER for the regression approach compared to the empirically determined threshold values.

Finally, energy efficiency is a major limitation for nanoscale communications due to the very small scale of the devices [132]. Tepekule et al. [132] address the energy efficiency problem by developing a receiver-based solution, namely the decision feedback filter, which is more energy efficient than comparable solutions, such as the minimum mean squared error equalizer [197] or the decision feedback equalizer [132], [197], [243]. The decision feedback filter computes the optimal threshold for the corresponding symbol and continuously renews the threshold. In the simulations, the decision feedback filter exhibits a higher memory requirement than the minimum mean squared error equalizer, but the decision feedback filter has a higher energy efficiency [132].

2) ANTI-ISI DEMODULATION SCHEME

Zhai et al. [222] also address the problem of static thresholds on the receiver side. Zhai et al. introduce and evaluate the Increase Detection Algorithm (IDA), a detection technique for demodulating the received symbols to avoid ISI. The MC system follows the model from Fig. 3, i.e., consists of a point source and a spherical absorbing receiver. The system is based on time-slotted MC, i.e., a bit is transmitted with a fixed, constant time slot length T [222]. Zhai et al. assume that the transmitter and the receiver are perfectly synchronized [222]. To transmit a bit sequence $b[i]$, the transmitter releases a prescribed number N of molecules or no molecules, i.e., OOK modulation is used. The received bit sequence $\hat{b}[i]$ is demodulated using IDA [222].

In traditional DBMC systems using OOK modulation, a fixed threshold is used for demodulation [222], i.e., the receiver compares the determined concentration or number of molecules with a fixed threshold. If the determined concentration at the receiver is greater than the threshold, then a bit-1 results; and vice versa for a smaller value, a bit-0 results. In [222], the fixed threshold is set to half of the maximum molecule concentration, $C_{\max}/2$. In the first step of the demodulation algorithm, the duration τ [seconds] of the time interval from the start of a time slot to the maximum concentration is determined. The maximum molecule concentration $C_{\max}$ occurs when the derivative of the concentration function $C(t)$ with respect to time t equals 0 [222], i.e., at the time t_0 when

$$\frac{\partial C(t_0)}{\partial t} = 0. \tag{47}$$

Noting that based on Sec. II-A1, specifically Eq. (7),

$$t_0 = \frac{d^2}{6D}, \tag{48}$$

the maximum concentration $C_{\max}$ at time instant t_0 can be evaluated with Eq. (14) as

$$C_{\max} = C(t_0) = \frac{N}{(\frac{2}{3}\pi d^2)^{3/2}} e^{-\frac{3}{2}}, \tag{49}$$

whereby d denotes the distance between transmitter and receiver, D denotes the diffusion coefficient, and N denotes

the number of emitted molecules. If in the demodulation process, the maximum value of the received molecular concentration is smaller than $C_{\max}/2$, then the corresponding symbol is demodulated as bit-0. Otherwise, the difference of two adjacent symbols Dif_i of the molecular concentrations at the beginning of a time slot is calculated. If the difference Dif_i is greater than 0, then the concentration function is predicted to increase in that time slot. The count variable c is increased by one. Subsequently, the next difference value Dif_{i+1} is determined until the corresponding threshold for the count variable c is reached. If the value of the count variable c is greater than the threshold of the count variable after completion of the sampling process, then a bit-1 is demodulated, and vice versa a bit-0 [222].

The presented demodulation mechanism is validated using a macroscale testbed [222]. The experimental testbed is a tabletop testbed with MQ-3 metal-oxide sensors. Alcohol molecules are used as information carrier. The results demonstrate that the BER using IDA is smaller than with traditional demodulation techniques. IDA can mitigate ISI and increases the information rate by shortening the time slot length [222].

3) ISI-RESISTANT CODING SCHEMES

Keshavarz-Haddad et al. [223] take a different approach to ISI-resistant codes, namely the so-called distinct Hamming codes based on [221]. We first introduce the foundations of the distinct Hamming code approach and then survey the individual ISI-resistant codes proposed in [223]. In the distinct Hamming code system, the information is converted from binary to decimal form, where the number of ones in the code word corresponds to a decimal number. On the receiver side, the number of received ones is counted, and the value is converted back into a binary format. For example, the 5-bit information sequence '00110', which is associated with the decimal number '6', is encoded into a code word consisting of six ones. The receiver calculates the number of received ones and converts the decimal number into a binary number [223].

The system model is the one-dimensional motion in a fluid without consideration of noise, whereby one molecule is used for the transmission of one bit. If a code word contains both a bit-0 and a bit-1, then two different molecule types are transmitted; specifically, one molecule type corresponds to bit-0, and another molecule type corresponds to bit-1. To transmit a code word over the channel, time intervals of duration τ_b seconds are required for the transmission of each bit. A code word of the length ω bits requires $\tau_b\omega$ seconds for transmission [223]. Keshavarz-Haddad et al. distinguish between *(i)* crossovers in one code word, and *(ii)* crossovers between two code words [223]. To reduce the probability of crossovers occurring between two code words, a time-gap of duration τ_g is inserted between two consecutive code words. The distinct Hamming codes approach is used as a foundation for the proposed schemes in [223].

The first ISI-resistant coding scheme presented in [223] is Crossover Resistant Coding with Time-Gap (CRCTG), which uses two different molecule types. In $CRCTG(n, k, \rho)$ coding, a code word of length n bits is used to transmit k bits of information. Thereby, $\rho = \tau_g/\tau_b$ denotes the ratio between the time-gap τ_g and the time-bin τ_b for the transmission of a molecule [221]. A decimal value denoted by ω is associated with the k information bits, whereby ω has a range from 0 to $2^k - 1$. For CRCTG, $n = 2^k - 1$ bits are used for the code word. The code word starts with $[\frac{n-\omega}{2}]$ bit-0s (rounded down for an even decimal ω), followed by ω bit-1s, and ends with $[\frac{n-\omega}{2}]$ bit-0s (rounded up for an even decimal ω). For odd decimal numbers ω, the code words are symmetrical; for even decimal numbers ω, the code words are asymmetrical, as illustrated by an example with $n = 31$ (ω depends on the code word) in column two in Table 6. For example, the information bits '00011' represent the decimal number 3, so that $\omega = 3$ ones are placed in between the zeros. Errors occur in CRCTG in the decoding process when bit-0s are replaced by bit-1s and vice versa. In most cases, the crossover occurs between two consecutive code words [223].

The second encoding scheme in [223] is Modified Crossover Resistant Coding with Time-Gap (MCRCTG) [223]. The MCRCTG encoding technique uses only one molecule type. The number of bit-1s equal to the decimal value of the information (bit) sequence to be transmitted is increased by one, i.e., to send k bits, the corresponding decimal number ω of the information bits is increased by one [223]. This increased decimal value is denoted as ω^*. For example, if the information bit sequence '00110' is to be encoded, the corresponding decimal value, i.e., $\omega = 6$, is increased by one, so that $\omega^* = \omega + 1 = 6 + 1 = 7$. Table 6, column three, illustrates the MCRCTG encoding for the example bit sequence '00011' converted to a decimal number ($\omega = 3$), so that four ones ($\omega^* = \omega + 1 = 3+1 = 4$) are used for the code word. However, it is important to note that the duration of each code word is fixed. Errors occur when the number of molecules sent in a time interval of duration T does not correspond to the number of received molecules [223].

The third ISI-resistant decoding technique presented in [223] is Modified Crossover Resistant Coding with Adaptive Time-Gap (MCRCATG). The number of molecules for each code word corresponds to the Modified Crossover Resistant Coding with Time-Gap (MCRCTG) scheme, but the duration T of the code word is adapted [223]. The main reason for the adaptation of T is to increase the throughput by shortening the time-gap for short code words [223]. With α and β representing positive constants, the time-gap τ_g after each code word is defined by a linear function of the code word length ω_l:

$$\tau_g = (\alpha + \beta\omega_l)\tau_b. \tag{50}$$

By increasing α and β, the average length of the time-gap increases, and consequently the BER decreases [223]. Thereby, the decoder determines the end time of the time-gap for each code word. In the next step of the approach, the receiver starts a timer and searches for the lowest value of ω_l for which $1 \leq \omega_l \leq 2^k$ [223]. If the timer reaches, e.g.,

TABLE 6. Code word assignment in ISI-resistant *CRCTG*(n, k, ρ) and *MCRCTG*(n, k, ρ) coding [223], with code word lengths $n = 31$ [*CRCTG*$(31, 5, \rho)$] and $n = 32$ [*MCRCTG*$(32, 5, \rho)$] for $k = 5$ information bits that represent the example decimal numbers ω zero to three.

Information bits	ω	Code word $CRCTG(31, 5, \rho)$	Code word $MCRCTG(32, 5, \rho)$
00000	0	00000 00000 00000 0 00000 00000 00000	1
00001	1	00000 00000 00000 **1** 00000 00000 00000	11
00010	2	00000 00000 00001 **1** 00000 00000 00000	111
00011	3	00000 00000 00001 **1** 10000 00000 00000	1111

almost 90 percent of the value of $\omega_l \tau_b + (\alpha + \beta\omega_l)\tau_b$ seconds, the receiver has collected exactly ω_l molecules [223]. The demodulator can then set the length of the code word equal to ω_l and all further received molecules are added to the next code word [223]. Errors occur when the demodulator decodes the length ω_l of a code word incorrectly.

The last coding scheme in [223] is an adapted variant of the MCRCTG, the Zebra Modified Crossover Resistant Coding with Time-Gap (ZMCRCTG). The ZMCRCTG uses two different molecule types. The molecule type is changed in dependence of the respective time interval, e.g., molecules of type A are used in even time intervals, and molecules of type B are used in odd intervals. This technique significantly reduces the BER and increases throughput due to the very low probability of ISI occurring as different molecule types are used [223]. However, the higher complexity of the system due to the use of two molecule types is a disadvantage [223]. Errors occur in ZMCRCTG when the molecules of time interval T_i reach the receiver in the time interval T_{i+k} with $k \in \{2, 4, \ldots, 2i\}$.

Simulations were performed in [223] to evaluate the four coding schemes. For the time-bin duration (required time interval to transmit a molecule) $\tau_b = 0.3$ seconds, the optimal time-gaps τ_g were determined as 9×0.3, 10×0.3, 3×0.3, and 0×0.3 seconds for CRCTG, MCRCTG, MCRCATG, and ZMCRCTG, respectively [223]. All four featured channel codes achieve lower BER than the codes in [218]. The simulations also demonstrate that MCRCTG has a slightly higher BER than CRCTG. The BER of ZMCRCTG, on the other hand, is very low, so that a much higher throughput can be achieved than with the other techniques [223]. If the distance between the transmitter and the receiver is increased, then MCRCATG and ZMCRCTG achieve a lower BER than the other encoding schemes [223].

4) ISI-CHANNEL CODE DESIGN

In 2019, Kislal et al. [135] present three new algorithms for channel coding for DBMC channels, namely a greedy algorithm, a genetic algorithm, and a mixed integer programming algorithm, to mitigate the ISI. Kislal et al. [135] consider a three-dimensional DBMC system consisting of a point transmitter located at a fixed point and a spherical absorbing receiver. The modulation technique is BCSK (specifically OOK), i.e., the transmitter releases N molecules at the beginning of each time interval for a bit-1, and no molecules for a bit-0. For an (n, k) code family, there are 2^k code words of length n, which are selected from a total of 2^n possibilities to design a codebook [135]. The transition probability matrix $P_{i,j} = P(w_j|w_i)$, where w_i and w_j with $i, j = 1, 2, \ldots, 2^n$ denote the transmitted and the received binary sequence, i.e., the transition probability matrix

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \ldots & P_{1,2^n} \\ P_{2,1} & P_{2,2} & \ldots & P_{2,2^n} \\ \vdots & \vdots & & \vdots \\ P_{2^n,1} & P_{2^n,2} & \ldots & P_{2^n,2^n} \end{bmatrix} \tag{51}$$

is the basis for channel coding [135]. The calculation of the elements of the matrix turns out to be challenging [135].

The first algorithms presented for DBMC systems in [135] are the greedy algorithm and its improved version. The greedy algorithm is based on an iterative selection of the best option for each step. Generally, greedy algorithms are relatively easy to implement and computationally cheap, so they are widely applied. The selection of code words can be considered as a choice from a scattered set among all possible sets for minimizing the probabilities of incorrect decoding of the transmitted code words. However, the greedy algorithm presented in [135] does not even outperform Hamming codes, its improved version does. The reason lies in the neglect of the decoding region. Consequently, not only the code words but also the decoding regions should be as widely scattered as possible. Difficulties arise in determining the decoding region since all code words have to be determined first [135]. To solve this problem, according to Kislal et al. the possible decoding regions for each binary sequence should be predicted, and then the code words can be selected [135].

The second proposed algorithm is a genetic algorithm implemented in DBMC systems. Genetic algorithms are inspired by evolution and have been applied in a wide variety of contexts in networking and generally in engineering [244], [245], [246], [247], [248]. Genetic algorithms are based on the fundamental idea of optimization through natural section. An initial population should be created with a high diversity to allow parallel search strands in the search space. Good properties of the individuals (here proposed solutions) are to produce an optimal combination by interbreeding. The performance of the genetic algorithm does not primarily depend on the initial population [135]. An individual of the genetic algorithm for an (n, k) encoding is represented as a vector of length 2^n that maps the code word assignments for each binary sequence. Then, the corresponding code word is assigned to each binary sequence, considering the error probability [135].

The third and last algorithm presented in [135] is a Mixed Integer Programming (MIP) based algorithm. Generally,

constraints which are either linear or integer constraints based on a linear objective function with integer variables can be used for a mathematical optimization problem formulation as an Integer Programming (IP); the problem is a MIP if some variables are continuous. Problems in DBMC systems can be considered in using an IP model because of the (in principle) integer number of molecules; however, from a computational (processing time) perspective, the MIP variant is typically preferred, since results close to the optimum can be obtained with MIP solvers with reasonable time and memory constraints [135]. Indeed, nanoscale DBMC systems operate typically under time and memory constraints. In particular, for designing increasingly powerful codebooks, which require increasing n and k, the search space grows exponentially with, so that finding an optimal MIP solution may become prohibitive. For reasonable time and memory boundary conditions, depending on the nanomachines, MIP can still achieve results close to the optimum and is thus generally suitable for channel coding [135].

The simulations in [135] compare the BER of the presented algorithms with the BER of the Hamming code. As expected, the Hamming code has the highest BER, followed by the improved Greedy algorithm, the genetic algorithm, and, finally, the MIP, which is slightly better than the genetic algorithm and thus achieves the lowest BER [135]. MIP is therefore advantageous for the different cases (and thus the varying parameters), the Greedy algorithm is suitable for certain boundary conditions, and the genetic algorithm performs similarly to MIP [135].

In 2020, Kislal et al. build on the experience from [135] and propose a novel low-complexity channel coding scheme for DBMC systems [136]. The presented channel coding method can be applied to different channel types. For the final simulations, Kislal et al. use a three-dimensional environment with a point transmitter and a spherical absorbing receiver, analogous to Fig. 3. BCSK modulation, specifically OOK modulation, is used [136].

The proposed ISI-mitigating codes are represented as (n, k) *ISI-mtg* codes, where n denotes the length of the code word and k is the number of information bits. The presented *ISI-mtg* codes use an adaptive threshold, which is updated for each code word. In addition, the codes follow three properties [136]: *(i.)* The code word does not contain consecutive bit-1s to reduce the ISI due to the detrimental influence of consecutive bit-1s on following bit-0s during decoding [132], [249]. *(ii.)* The code word starts with a bit-0 to prevent consecutive bit-1s between two adjacent code words. *(iii.)* The code word must contain at least one bit-1, specifically for the detection algorithm presented in [136]. For $n = 3$, the set of code words is

$$CW_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \tag{52}$$

whereby each row represents a code word. The first column is filled with bit-0s to satisfy the second property. The matrix W_3 in turn starts with bit-1s, but also contains no consecutive bit-1s for $n = 3$ [136]:

$$W_3 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}. \tag{53}$$

By concatenating the two matrices CW_3 and W_3, CW_4 follows [136]:

$$CW_4 = \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] = \left[\begin{array}{c|ccc} 0 & \multicolumn{3}{c}{CW_3} \\ \hline 0 & \multicolumn{3}{c}{W_3} \end{array}\right]. \tag{54}$$

For $n > 3$, CW_{n+1} and W_{n+1} can be recursively constructed as [136]:

$$CW_{n+1} = \left[\begin{array}{c|c} 0 & CW_n \\ \hline 0 & W_n \end{array}\right], \tag{55}$$

$$W_{n+1} = \left[\begin{array}{c|c} 1 & 0 \\ \hline 1 & CW_n \end{array}\right]. \tag{56}$$

Turning to the decoding, based on the properties of the *ISI-mtg* codes [136], there are at least one bit-0 and one bit-1 in each code word. Kislal et al. use an adaptive threshold to differentiate between bit-0 and bit-1. Let $r^i = (r_2^i, r_3^i, \ldots, r_n^i)$ represent the number of received molecules of each bit of the i^{th} transmitted code word, excluding the first bit r_1^i (whereby r_1^i is a bit-0 because of the convention in [136]). Since a code word consists of both bit-1s and bit-0s, the optimal threshold must therefore lie between $d_0 = \min r^i$ and $d_1 = \max r^i$. With this assumption, the threshold can be adaptively determined to be [136]:

$$\tau^i = \alpha d_0 + (1 - \alpha) d_1, \tag{57}$$

where by α denotes a scaling factor, which is difficult to determine [136]. In the first step of the decoding algorithm, a decision is made about all received binary sequences z_l^i $(l = 2, 3, \ldots, n)$ in comparison with the adaptive threshold, i.e., whether the corresponding bit is decoded as a bit-0 or bit-1 according to $z_l^i \gtrless \tau^i$. The decoding algorithm is then used to determine the position of the bit with the highest number of received molecules.

Specifically, in the next step, the position of the bit with the highest number of received molecules is determined. This position is with the highest probability the position of a bit-1. The adjacent bits are then decoded as bit-0s in the following step, since consecutive bit-1s are not allowed as per property *(i.)* of *ISI-mtg* codes [136]. All steps are repeated for all code words. In the case of two bit-1s occurring in succession, at least one of the two bit-1s must be a bit-0, based on property *(i.)* of the proposed *ISI-mtg* codes [136].

The final simulation results in [136] demonstrate that the presented *ISI-mtg* coding scheme has a lower BER and performs significantly better than the uncoded case, especially for a low symbol duration of $T_S = 200$ ms. A key advantage of the *ISI-mtg* scheme in [136] is that, unlike other coding schemes, for low symbol durations and a high number of

transmitted molecules, the *ISI-mtg* scheme does not reach an error floor. Kislal et al. also present a macroscale testbed in [136] and implement the *ISI-mtg* codes in the testbed. The testbed results demonstrate that the *ISI-mtg* codes can be successfully implemented in an experimental setup and significantly reduce the BER [136].

C. TCH CODES

Figueiredo et al. [137] introduce a new approach for DBMC channel coding, namely the TCH codes. The advantage of TCH codes is their simplified detection [137]. The DBMC model is based on a three-dimensional communication channel. The source and sink in a DBMC system are a point transmitter and a spherical absorbing receiver [67], [250], [251], [252], which determines the number of absorbed molecules in a prescribed time window $[t_l, t_u]$. In [137], the two modulation techniques OOK and BCSK are considered.

In general, the concentration vector in the time slot k can be represented as [137]:

$$y[k] = \sum_{l=0}^{L-1} h[l,k] \cdot x[k-l] + n_e[k], \tag{58}$$

where L denotes the length of the ISI, $h[l,k]$ describes the number of the molecules observed by the receiver during the time slot k after N molecules have been released by the transmitter in a time slot $k-l$, $x[k-l]$ denotes the binary modulation symbol (based on concentration) that is emitted by the transmitter in time slot $k-l$, and $n_e[k]$ denotes the environmental noise in time slot k. Both $n_e[k]$ and $h[l,k]$ follow the Poisson random variable modeling, i.e., $n_e[k] \sim Poisson(\bar{n}_e)$ and $h[l,k] \sim Poisson(\bar{h}[l,k])$ [137]. The Poisson probability model outlined in Section II-A1 gives

$$\bar{h}[l,k] = N \cdot F((l+1)\cdot T_S,\ l \cdot T_S), \tag{59}$$

where F and T_S denote the hitting probability and the symbol duration. The concentration received at the receiver $y[k]$ in time slot k can be represented as [137]:

$$y[k] = \bar{h}[0,k] \cdot x[k] + v[k] + I[k] + n_e[k], \tag{60}$$

where v denotes the diffusion noise, and I the ISI. The desired signal is denoted by $\bar{h}[0,k]\cdot x[k]$. The diffusion noise $v[k]$ is also modeled by a Poisson random variable, $v[k] \sim Poisson(\bar{h}[0,k]\cdot x[k])$ [137].

Two different simple algorithms are used for detection and bit decision. The first algorithm is a threshold based detection, similar to the one in [136], which uses hard decision symbols [137]. In binary concentration based modulation, A_0 and A_1 represent the levels for a bit-0 and bit-1, so that each symbol comprising code word i can be represented as $x^i[k] \in \{A_0, A_1\}$.

For each individual bit, the second detection technique is based on ad hoc computation of log-likelihood ratios [137]. Thereby, estimates for probabilities of occurrence are computed, e.g., the probability for the occurrence of a certain symbol at a certain position. These soft estimates can then be used for soft-decision decoding, exploiting Euclidean distance minimization [137].

Let us now turn specifically to the TCH codes. In general, TCH codes are binary, non-systematic, non-linear, cyclic block codes of length $n = 2^m$ [137], whereby m denotes any positive integer. TCH codes are represented as $TCH(n,k,t)$, where n represents the code length, k represents the number of information bits in the code word, and t represents the error correction capacity. In TCH codes, the number of zeros is equal to the number of ones, so they are balanced codes, which is advantageous in DBMC systems. Furthermore, TCH codes can be created in almost any length [137]. An important subclass are the basic TCH polynomials with degree n [137]. These codes follow the equation

$$P(x) = \sum_{i=0}^{\frac{p-1}{p}-1} a_i x^{K_i}, \quad a_i \in GF(q),\ q = p^k,\ k \in \mathbb{N}. \tag{61}$$

The exponents K_i satisfy

$$a^{K_i} = 1 + a^{2i+1}, \quad i = 0, 1, \ldots, \frac{p-1}{p} - 1. \tag{62}$$

Thereby, p denotes a prime number following

$$p = n_{TCH} + 1 = 2^m + 1. \tag{63}$$

If Eq. (63) is satisfied, then the prime number can be written as [137]:

$$F_i = 2^{2^i} + 1, \quad i \in \mathbb{N}. \tag{64}$$

Consequently, the prime numbers are Fermat numbers. In total, only five code word lengths satisfy the rules, so that only code words of lengths 2, 4, 16, 256, or 65536 can be formed [137]. Basic TCH codes have very good cross and autocorrelation [137].

In [137], an end-to-end MC simulator is used to evaluate the TCH codes in a DBMC system. The performance of the TCH codes is compared with Hamming codes and with the ISI-mitigating codes proposed in [136]. To ensure comparability, similar sets of code rates, which include both lower and higher code rates, are considered. For the uncoded case and the Hamming code, threshold hard decisions are used in simulations. In contrast, the TCH codes and the ISI-mitigating codes use soft decisions. The simulations demonstrate that the channel codes give higher BER for OOK modulation than for BCSK modulation [137]. Importantly, TCH codes achieve lower BER than other codes with similar parameters. Furthermore, TCH codes with higher rates, e.g., $TCH(16,8)$, achieve lower BER than other codes with lower rates, e.g., $Hamming(3,1)$ [137]. The simulations also indicate that TCH codes with the same code rate, but longer code words achieve lower BER [137]. In summary, the TCH codes presented in [137] can achieve lower BER compared to other channel coding techniques, regardless of symbol duration, modulation, or pulse type, provided soft decisions are used.

Future research should conduct evaluations that consistently use hard decisions for all compared approaches, or soft

decisions for all approaches. In particular, the relative performance of TCH codes with soft decisions compared to uncoded and Hamming coding benchmarks with soft decisions is presently unknown. Also, the performance of TCH codes for hard decisions is presently unknown.

In addition to simulations, [137] introduces a macroscale experimental testbed setup, which uses pH levels to test molecules as information carriers. For more detailed information, the reader is referred to [137], as this survey specifically focuses on the nanoscale.

D. SUMMARY AND DISCUSSION

ISI-mitigation codes can be divided into ISI-free codes, ISI-aware codes, and TCH codes. ISI-free codes are constructed to be immune to the occurrence of ISI [219]. The ISI-aware approach determines the expected number of molecules delayed by ISI from the history of the previous symbols [219]. The receiver can thus subtract the number of delayed molecules from the number of received molecules, in order to increase the reliability of DBMC [219]. The ISI-aware approach can significantly reduce the BER and achieves the lowest BER compared to traditional convolutional codes [219]. Qiu et al. [220] demonstrated that ISI-free codes achieve significantly lower BER than Hamming codes and cyclic Reed-Muller codes, with a concurrent lower encoding rate [220]. The ISI-mitigation codes thus provide a promising starting point for applications in DBMC systems.

Tepekule et al. [132] develop MTSK, a physical-layer ISI-aware code that increases the data rate in DBMC systems without increasing the ISI. Simulations demonstrate that MTSK reduces the BER in MC-based communication systems, and achieves lower BER than the MoSK and BCSK modulation techniques [132].

Figueiredo et al. [137] presented a new approach to channel coding, namely the TCH codes. TCH codes have a simple detection compared to other channel codes [137]. Regardless of symbol duration, modulation technique, or pulse type, TCH codes achieve significantly lower BER than other channel codes [137].

The simulation results of Tepekule et al. [132] indicate that ISI-aware channel codes in combination with modulation techniques developed specifically for DBMC systems, e.g., MTSK, achieve lower BER than the well-known linear block codes. However, the evaluation results to date were mostly obtained from simulations and should be verified in the future on real testbeds, such as in [137], but at the nanoscale.

The high number of ISI-mitigation approaches in Table 5 underscores the importance of ISI-mitigation in DBMC systems. In particular, Multi-Input Multi-Output (MIMO) operation and the incomplete elimination (or reception) of all remaining molecules in the channel increases the ISI. The overview in Table 5 indicates that increasing complexity generally pays off in terms of increased error correction capability for ISI-mitigation approaches.

VIII. CHANNEL CODING: MISCELLANEOUS CODES

A. ERROR-FREE CONSTANT WEIGHT CODES

Einolghozati and Fekri [190], [224] research low-complexity error detection schemes. For this purpose, an algorithm is developed which dynamically determines the optimal code word [65]. The maximum error probability should approach zero. The modeled MC system operates over an asymmetric channel (see Section II-E), whereby the information is encoded into the concentration of the molecules, e.g., with OOK modulation [190], [224]. Einolghozati and Fekri [190], [224] find optimal error detection codes by two different theoretical measures, namely by maximizing the mutual information and by maximizing the code rate. The theoretical measures depend on the encoding and decoding complexity. For finding the optimal error detection codes by maximizing the mutual information, Einolghozati and Fekri consider an encoder generating the code words with any arbitrary relative frequency [190], [224]. For finding the optimal error detection codes by maximizing the code rate, Einolghozati and Fekri consider an encoder generating the code words equiprobably [190], [224].

The channel codes used by Einolghozati and Fekri, called constant-weight codes, are particularly suitable for MC due to their perfect error detection, but are only practical for very short code words [65], [190]. For decoding at the receiver side, the decoder assesses the weight of the received code word. If the code word mismatches with the weight of the default code words, then the received word is discarded [190], [224]. Einolghozati and Fekri also analyzed the optimal weight and length of the constant-weight codes. For example, for the weight $k = 2$ and for a small transition probability p_e from a bit-1 to a bit-0 in the asymmetric channel, the code rate $\omega = n/2$ should be chosen, whereby n denotes the block size [190].

A major challenge in MC is the random delay of the transmitted molecules due to the propagation via diffusion. If the molecules are encoded in a time-slotted system, then delayed molecule arrivals can lead to ISI, which increases the BER. In [225], the zero-error capacity of an MC channel is investigated and zero-error codes for the release of one type of molecules are presented. Abadi et al. [226] extend the work and introduce an upper bound for the zero-capacity and present a zero-error code for using multiple molecule types.

Specifically, in the communication channel of Kovačević and Popovski [225]: *(i.)* time is slotted, *(ii.)* N particles are sent in each time slot, *(iii.)* each particle in the communication channel is randomly delayed in the interval of $\{0, 1, \ldots, K\}, K \in \mathbb{N}$, and *(iv)* the particles are identical. Kovačević and Popovski show that the zero-error capacity of the channel is equal to $\log r$, where r is a unique positive real root of the polynomial $x^{K+1} - x^K - N$. For the in [225] presented case $N = 1$ and $K = 1$, the capacity is equal to $\log \phi$, where $\phi = (1 + \sqrt{5})/2$ is the golden ratio [225].

Error-free codes can be constructed according to [225] in two different ways: recursive construction or direct construction. In the recursive construction, general properties for

zero-padded zero-error codes are specified for DBMC, based on which the codes can be created in a straightforward way. The created codes can contain only code words with prefix N and $i \circ 0^K$, $i \in \{0, 1, \ldots, N-1\}$ without losing generality, whereby $\circ$ denotes concatenation [225]. For the zero-padded-zero error code $\mathcal{D}(n)$ [225]:

$$\mathcal{D} = \mathcal{D}^N \cup \bigcup_{i=0}^{N-1} \mathcal{D}^{i \circ 0^K}. \tag{65}$$

In the first step of the direct construction, in reverse lexicographic order, the sequences of length n over $0, 1, \ldots, N$ end with $\min\{n, K\}$ zeros. For $n > K$, the first sequence on the list is $N^{n-K} \circ 0^K$, the second is $N^{n-K-1} \circ (N-1) \circ 0^K$, and so on [225]. Through further repetitive processing of the sequences as detailed in [225], the error-free codes are then obtained.

Abadi et al. [226] extend the work in [225] to include several types of molecules that can be used for transmissions. The simulation results demonstrate that even for two or three molecule types, the rates and rate gains increase as N increases [226]. The proposed code outperforms a simple code for one, two, or three different molecule types [226].

B. MISCELLANEOUS CHANNEL ERROR MITIGATING CODES

This subsection is for completeness and covers miscellaneous codes that have been surveyed in [65]: single parity check [227] (which is a parity check code), the EGLDPC [133], parity check erasure [228] (which is a parity check code), and fair weak flip [229] (which is a binary non-linear code).

1) SINGLE PARITY CHECK

Marcone et al. [227] design a parity-check encoder with biological modulation, and an associated analog decoder based on MC in a biological cell. Marcone et al. used genetic circuits for the design using the reaction of molecular species, as well as activation and repression [227]. The proposed encoder is a parity check encoder inputting the concentrations $c(t)$ (as a function of the time t) of a cell and outputting the encoded bits x_i [227]. The encoded bits x_i are modulated into emission symbols $x_i(t)$ of the molecules based on the used modulation scheme (OOK). The parity-check encoder encodes two bits with a simple parity check code (with code word length $n = 3$). These two bits represent the cell state $c(t)$ at time t, e.g., $c(t) = (s_1(t),\ s_2(t))$ [227]. Furthermore, based on the propagated concentration, the decoder computes an *a posteriori* log-likelihood ratio of the information bits [227]. Thereby, the proposed parity check code only checks for errors—the parity check code does not correct errors [227]. The simulation results (biochemical simulation) demonstrate that the BER of a genetic circuit is close to the performance of an electrical network implementation.

2) EGLDPC

The design of EGLDPC codes is based on Euclidean geometry and its lines and points [133]. In [253] and [254], the EGLDPC codes are considered to be cyclic and two-dimensional. In general, the EGLDPC codes are represented as (n_L, k_L) codes, where $n_L = 2^{2s} - 1$ denotes the block length and $k_L = 2^{2s} - 3^s$ denotes the message length for $s \geq 2$ [253]. The minimum distance of the EGLDPC codes is $d_{L_{\min}} = 2^s + 1$ [253]. The encoding of the EGLDPC codes is based on the application of a simple feedback shift register, and decoding uses a one-step majority logic method [133].

Comparing Hamming codes, cyclic Reed Muller codes [237], and EGLDPC codes [110], Lu et al. [133] found that EGLDPC codes have the best energy efficiency, while Hamming codes are better suited for long distances in DBMC systems [65]. Furthermore, Lu et al. demonstrated that an LDPC code with $s = 4$ achieves a coding gain of 7.26 dB [133]. Especially for nanomachine-to-nanomachine communication in the BER interval $[10^{-6};\ 10^{-3}]$, Hamming codes with $m = 4$ are better suited for MC systems than EGLDPC codes. The lowest energy costs are achieved by LDPC codes for $s = 2$ [133]. Finally, for nanomachine-to-macromachine and macromachine-to-nanomachine communications, Lu et al. suggested using $s = 2$ and $s = 3$ LDPC codes [133].

3) PARITY CHECK ERASURE

Egashira et al. [228] introduce parity check erasure coding. Here, the transmitter performs XOR operations on a group of m molecular packets to generate k fixed-length parity bits and k additional packets [228]. These k additional packets are named parity packets and contain the parity codes. Finally, the code word consists of $m + k$ packets, which are transmitted to the receiver. The simulation results in [228] demonstrate that the protocol implemented in [228] is robust for DBMC systems and can reduce latency and jitter [65].

4) BINARY NON-LINEAR FAIR WEAK FLIP

Lin et al. [229] design error correcting codes for finite block lengths with applications that require short codes (ultra-small block codes). These codes are thus well suited for MC systems. A fundamental innovation in [229] is the column-wise reading of the codebook matrix, in contrast to the common row-wise reading (code words) of the matrices of nonlinear binary codes. The Hamming distance is consequently extended by a column-wise consideration. The scheme presented is named weak bit flip code [229]. Through theoretical analysis, Lin et al. show that the weak bit flip code is currently superior to most channel codes in terms of the BER performance [229]. However, since the code is still in an early stage of development, no details about its implementation complexity or quantitative performance are known at this time [65], [229].

C. SPATIAL CODES

Another channel coding approach is MIMO-based spatial coding, which have been explored by Damrath et al. [230], [231] and have not been previously surveyed. Yeh et al. proposed to employ MIMO techniques (coming from conventional wireless communications) in MC. Yeh et al.

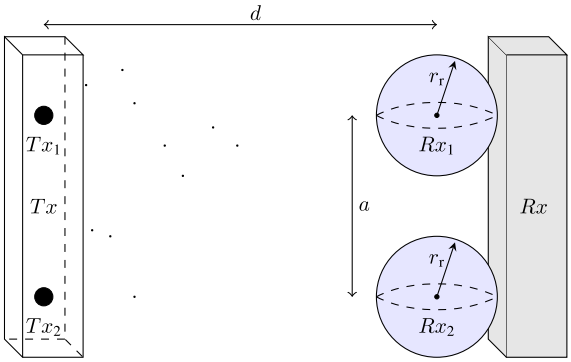


FIGURE 20. Diffusion-based molecular 2 × 2 MIMO system model, adapted from [256] as considered in the spatial coding studies [230], [231].

demonstrate that the MIMO mechanisms spatial multiplexing and selection diversity can be integrated into MC networks to improve the MC performance [219]. Additionally, Yeh et al. conducted initial explorations of further MIMO techniques for DBMC systems, e.g., maximum ratio combining and decision fusion, and indicated that these further MIMO techniques are promising approaches for future DBMC research [219].

In [230], two new coding techniques for spatial coding in DBMC systems were introduced: Alamouti-type coding and repetition MIMO coding. The system model is similar to the models introduced in [255] and [256], see Fig. 20. The model consists of the two point transmitters Tx_i, $i = 1, 2$, and the two receivers Rx_i in an infinite three-dimensional homogeneous fluid medium without drift. The spherical receivers with radius r_r are attached to the reflecting body of Rx. Tx_1 is aligned with Rx_1, and Tx_2 is aligned with Rx_2. d is the distance between Tx_1 and Rx_1, as well as between Tx_2 and Rx_2. The distance between the two receivers, Rx_1 and Rx_2, is given as a [230]. If the diffusing molecules hit Rx_1 or Rx_2, their number is determined, and they are absorbed, i.e., the molecules are removed from the channel. The OOK modulation technique is used.

Tx_i releases either no or N molecules at the beginning of a symbol period k of length T_S, depending on a bit-0 ($u_i[k] = 0$) or bit-1 ($u_i[k] = 1$). The emitted molecules of Tx_1 and Tx_2 are of the same type. The channel of the MIMO communication system can be divided into subchannels for each connection between Tx_i and Rx_i.

To achieve spatial diversity on the transmitter side, the information is transmitted via multiple antennas [230], whereby the transmitter into the MC channel is referred to as ''antenna''. Therefore, the binary data sequence $u[k]$ (with k denoting the symbol period index) is transmitted as a sequence of data symbols s_k [230].

1) ALAMOUTI-TYPE CODING

The Alamouti scheme is a space-time block code, which is mainly used in radio-based (EM) communication systems for spatial diversity [124]. The Alamouti scheme can be represented in its original form by the 2 × 2 matrix

$$G = \begin{bmatrix} s_k & s_{k+1} \\ -s_{k+1}^* & s_k^* \end{bmatrix}. \tag{66}$$

The columns of the matrix in Eq. (66) correspond to the transmitting antennas, the rows to the corresponding symbol periods k and $k+1$. Specifically, the transmissions in the first time slot are simultaneous: $x_1[k] = s_k$ is transmitted via the first antenna, and $x_2[k] = s_{k+1}$ via the second antenna. In the second time slot, the same procedure follows, $x_1[k+1] = -s_{k+1}^*$ is transmitted via Tx_1 and $x_2[k+1] = s_k^*$ via Tx_2. Spatial diversity can be ensured by transmitting the same information through both transmitters [230]. An advantage of the Alamouti scheme is the orthogonality of the code, whereby a maximum-likelihood detector can be implemented more easily, since Interlink Interference (ILI) can be completely neglected [230]. In MC, the data symbols, i.e., the number of released molecules, are non-negative real values (as opposed to complex values) [230].

For OOK, the Alamouti scheme was transferred by Damrath et al. to an Alamouti-type code, which avoids the negative values and conjugates complex values, conditioned by the boundary condition of the data symbols $s_k \in \{0, N\}$ [230]. The negative symbols can thus be replaced by $\tilde{s}_k := N - s_k$ and the complex conjugate operation will be discarded. Accordingly, the transmission matrix for the Alamouti-type code is [230]:

$$G = \begin{bmatrix} s_k & s_{k+1} \\ N - s_{k+1} & s_k \end{bmatrix}. \tag{67}$$

Further, $G^H G = 2I$, whereby G^H denotes the Hermitian of the matrix G and I denotes the identity matrix [230]. If the condition $G^H G = 2I$ is satisfied, then the matrices are orthogonal [230]. Orthogonality simplifies the implementation of a maximum-likelihood detector, because the ISI effect can be cancelled [230].

2) REPETITION MIMO CODING

An alternative to the Alamouti code is the repetition MIMO code [257]. With repetition MIMO coding, the information is distributed over all transmitters, i.e., all antennas transmit simultaneously the same data symbol. The transmission matrix for a 2 × 2 MIMO system is [230]:

$$G = [s_k \quad s_k]. \tag{68}$$

With repetition MIMO coding, a single detection algorithm can be applied even in the presence of ILI, which is one advantage of repetition MIMO coding [230].

3) RECEIVER COMBINING

If there are two or more receiver antennas in a communication system, then the received signal from each antenna must be combined or selected; this applies also for a DBMC system [231]. Only then the detection will follow. Basically, three selection techniques can be distinguished. With *i.)* selection diversity, the strongest signal from all

antennas is selected for detection [231]. However, in an MC system, with OOK modulation and the occurrence of ILI, it is difficult to determine the strongest signal. For $u[k] = 0$, the signal with the fewest received molecules would be the strongest, while for $u[k] = 1$, the signal with the most received molecules would be the strongest signal. In the second method, *ii.)* equal-gain combining, the received signals from all antennas are equally weighted and combined [231]. Weighing the individual subchannels according to a factor gives the third method, *iii.)* the maximum-ratio combining, which is similar to the maximum-likelihood receiver [231].

Simulations in [230] examined the impact of the spatial diversity on the BER. The results demonstrate that a repetition code with an adaptive threshold detector outperforms single-input single-output transmission for the release of few molecules [230]. The maximum improvement of repetition MIMO compared to the single-input single-output case is approximately one order of magnitude reduction of the BER [230]. However, Alamouti-type coding does not reduce the BER [230]. The reason for this is that in Alamouti-type coding, ILI is competitive and, as a result, more destructive in comparison to repetition MIMO codes, where the ILI contributes to the signal strength [230].

The influence of the transmission distance and the effect of antenna separation were additionally investigated through simulations in [231]. For the antenna separation, Damrath et al. [231] demonstrated that the BER decreases with an increasing distance a between the receivers. The reason for this is the spatial gain of the ILI, since an increasing a leads to a decrease of the ILI [231]. Also, a shorter transmission distance d results in more molecules being absorbed, and consequently increased signal strength [231].

D. MINIMUM ENERGY CODES (MEC)

Developing energy efficient and reliable communication techniques is essential for MC because of the limitations of molecular nanomachines, e.g., limited computational power [204]. Therefore, Bai et al. [204] investigated minimum energy channel codes for a diffusive channel, minimizing the energy while maintaining the reliability. Based on the proposal of MEC in [232] for nanocommunications, Bai et al. constructed MEC for MC. Due to the lower energy requirement for the transmission of a symbol '0' compared to the transmission of a symbol '1' (OOK modulation requires no energy for the transmission of a symbol '0' due to no release of molecules into the environment), lower weight code words reduce the energy demand [204]. Consequently, minimizing the code word weight minimizes the code word energy demand [204]. Thereby, following [204] the minimum code word length $n_{\min}$ is:

$$n_{\min} = d_m + (M - 2) \times \left(\frac{d_m}{2} \right), \tag{69}$$

whereby d_m denotes the minimum Hamming distance and M the number of code words. The simulations indicate that, on the one hand, MEC have a lower BER and lower energy demand than Hamming codes; on the other hand, MEC require larger code word lengths [204].

E. SUMMARY AND DISCUSSION

Error-free codes can be designed following [225] in two different ways: recursively and directly. The simulation results demonstrate that even for two or three molecule types, the rates and rate gains increase as the number N of arrivals per time slot increases [226]. Thereby, the proposed code outperforms a simple code [226]. Kovačević and Popovski dealt with a time-slotted MC system in [225]. In particular, the existing error-free codes are restricted to the respective modulation techniques, e.g., time-slotted for zero-error codes in [225] or OOK for spatial codes in [230] and [231]. In combination with different types of molecules, e.g., type A and B, and different pulse amplitudes, combined solutions should be developed in future research, which generate, e.g., four combinations (time-slotted MoSK).

Another channel coding approach is MIMO-based spatial coding, which has been explored by Damrath et al. [230], [231]. Thereby, in [230], two new coding techniques for spatial coding in DBMC systems were introduced, namely Alamouti-type coding and repetition MIMO coding. Damrath et al. demonstrate in [230] that the repetition code outperforms single-input single-output transmission for the release of a few molecules by a factor of 10 in terms of the BER; in contrast, Alamouti-type coding does not reduce the BER [230]. Furthermore, Damrath et al. [231] demonstrated that the BER can be decreased by increasing the distance between the receivers of MIMO DBMC systems. Also, a shorter transmission distance results in more molecules being absorbed, and consequently increased signal strength [231]. Analogously to ISI-free codes (Section VII-A), the existing spatial codes can be viewed as being focused on mitigating the ISI (and not directly focused on correcting channel errors).

Further miscellaneous channel coding approaches exist [65], [228], [229], but they are often at an early stage of development and no definitive conclusions about complexity or performance can yet be drawn. In particular, the error correcting codes for finite block lengths of [229] represent an avenue for future research. MC systems at the nanoscale are well suited for column-by-column reading of the codebook of the error correction codes by Lin et al. [229] due to the limited block length as well as the limited number of code words in MC systems [190]. The column-wise reading of codebooks instead of the conventional row-wise reading has general benefits for designing and analyzing codebooks with small numbers of code words [258]. Furthermore, column-wise reading allows the definition of further code families, e.g., codes based on the r-wise Hamming distance, which is an extension of the pairwise Hamming distance [229].

Overall, the number of studies on miscellaneous channel codes is relatively small, see Table 5. Future research should focus on the design of new miscellaneous codes for DBMC systems, focusing, e.g., on the further

TABLE 7. Overview of network coding studies in DBMC nanonetworks. For all presented approaches, the reception is additive with a threshold. In terms of complexity, the attribute "Low" describes the ease of implementation, application to molecular modulation techniques, and understanding the coding scheme. The "High" attribute, on the other hand, includes a more sophisticated implementation and more required computing power. "Medium" is in-between. For example, the duplex network coding scheme in Section IX-B is easy to understand, but the practical implementation, e.g., on a testbed, is difficult; therefore, we assigned the complexity attribute "Medium". All existing network coding studies in DBMC nanonetworks consider a diffusion channel model with 3D propagation.

Ref.	Year	Modul.	Backgr. Noise	Detection	# of Mol. Types	# of Rel.	Dyn.	Capac.	Complexity
1. Antecedents of Network Coding, Section IX-A									
[259]	2013	MoSK	Yes	Abs.	1	1	No	Low	Low
[260]	2014	OOK	Yes	Abs.	1	1	No	Low	Low
[261]	2015	OOK	Yes	Abs.	1	1	No	Low	Medium
2. Duplex Network Coding, Section IX-B									
[262]	2016	BCSK	No	Abs.	1	1	No	Medium	Medium
3. Physical-Layer Network Coding, Section IX-C									
[263]	2016	OOK	No	Abs./Reac.	3	1	No	Medium	Medium
[264]	2019	OOK	Yes	Abs./Reac.	3	1	No	Medium	Medium
4. Multi-hop Nanonetwork Coding, Section IX-D									
[265]	2018	OOK	No	Abs.	1	2–4	No	Medium	Medium
5. Performance Improvement of Network Coding, Section IX-E									
[266]	2020	OOK	Yes	Abs.	2	1	No	Medium	High
[267]	2021	OOK	No	Abs.	2	1	No	High	High
6. Network Coding in Mobile Nanonetworks, Section IX-F									
[268]	2022	D-MoSK	Yes	Abs.	2	1	Yes	Medium	Medium

Abbreviations: Abs.: Absorbing – BCSK: Binary Concentration-based Shift Keying – Backgr. Noise: Background Noise – Capac.: Capacity – Dyn.: Dynamic – Mol.: Molecules – Modul.: Modulation – MoSK: Molecule Shift Keying - OOK: On-Off Keying - Reac.: Reactive – Ref.: Reference – Rel.: Relay

distance functions for DBMC, e.g., the crossover-distance metric (see Section VI-C) or asymmetric-distance metric (see Section VI-D).

IX. NETWORK CODING IN DIFFUSION-BASED MOLECULAR NANONETWORKS

This section comprehensively surveys the field of network coding in DBMC nanonetworks. Table 7 gives a structured overview of the network coding studies to date.

A. ANTECEDENTS OF NETWORK CODING

As one of the first studies of network coding for DBMC nanonetworks, Unluturk et al. [259] investigated the influence of network coding on the bit rate of a DBMC system. For this purpose, a nanorelay was implemented in the system, as shown in Fig. 21. Two Transceivers (TRXs) transmit messages to each other. Due to the large distance d_{AB} and the low transmission rate (propagation delay problem), a nanorelay is used. The relay XORs the incoming messages from the two transceivers and sends the XORed message, a sequence of bits represented by MoSK modulation, to the TRXs.

First, we consider the uncoded case. Suppose, that the two nano TRXs start the transmission of their message at the same time. Following Fig. 21, the left transceiver transmits molecules of type A (molecule A for brevity) and the right transceiver transmits molecules of type B (molecule B for brevity). However, both TRXs do not know at which position the other Transceiver (TRX) is located and merely release their molecules into the environment. Based on the diffusion of the molecules, the molecules A and B arrive at the relay at different times. After the message of a transceiver has arrived, i.e., molecule A or molecule B has been received, the relay starts forwarding the first received message (i.e., the first received molecule). Suppose that the molecules A from the left transceiver arrive before the molecules B from the right

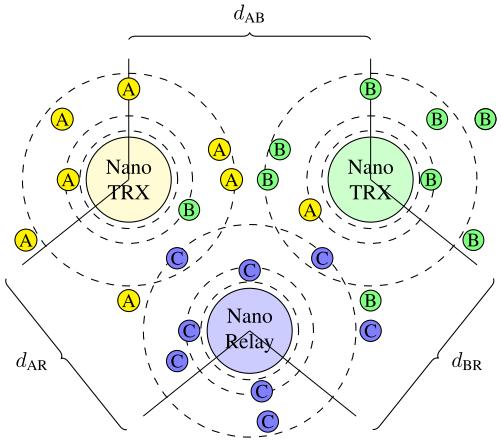


FIGURE 21. Network coding model for diffusion-based molecular nanonetworks, adapted from [259]. The two nanotransceivers (TRXs) each send and receive a message. At the nanorelay, the received messages from the two transceivers are combined (XOR) and retransmitted to the transceivers. The mathematical basis for this diffusion model in Section II-A2.

transceiver arrive at the relay. The relay will then first start forwarding the molecules from the left transceiver (transmitting molecule A) to the right transceiver. If, for example, a second message (from the right transceiver transmitting molecule B) arrives after the first message and thus during the forwarding process, then the second message is held in a queue. This queueing delays the forwarding of the second message [259].

Next, we consider the network-coded case. If the relay receives a message from a TRX, then the relay waits a prescribed time interval T_W for another message (from another TRX). If another message arrives within the time interval T_W, then the relay releases the XORed message C, following Fig. 21. This molecule of type C is either a third molecule type or a molecule formed from a chemical

reaction of the two initial molecules A and B. If no other message arrives within the time interval T_W, then the relay starts forwarding the (originally received) message (e.g., molecule A) separately [259].

This network coding significant increases the bit rate compared to the non-coded network. Following the network coding principles of Fig. 10(b), fewer time slots are required for network-encoded bidirectional transmission. For example, in 12 time slots, three bidirectional transmissions can take place without network coding, see Fig. 10(a); while four bidirectional transmissions are feasible in 12 time slots with network coding, see Fig. 10(b).

If the waiting time T_W is increased, then the waiting interval for another message from another TRX increases, and consequently, the probability of coding increases. The delay due to the nanorelay before coding increases the total delay. However, the nanorelay combines the received messages, thus decreasing the total number of messages. Concomitantly, the reception rate is increased.

The error performance of network coding in DBMC nanonetworks has been studied by Aijaz et al. [260]. Again, a relay, analogous to Fig. 21, is considered. The simulations examine the dependence of the error probability on the pulse amplitude, the probability of ISI, the diffusion coefficient of the medium, and the radius of the receiver. The simulation results indicate that the higher the pulse amplitude, i.e., the higher the number of emitted molecules per cubic micrometer, the lower the error probability in the transmission of a bit-1 in the on-off keying process. Furthermore, the greater the pulse duration, the lower the probability of ISI occurring [260].

Akdeniz et al. [261] use relays to counteract an obvious disadvantage of DBMC—-the limited range. A relay shortens the distance between the transmitter and the receiver, so that a larger communication range can be achieved, see Fig. 21. The two nano TRXs are each outside the communication range of the other TRX. By inserting a relay, the communication range can be extended, as communication between both TRXs becomes feasible via the relay. Also, higher data rates can be achieved using network coding. Of particular interest is the influence of the ISI on later received bits and time slots. In [261], a time slot of total duration T_S seconds for receiving a bit-1 was decomposed into time intervals of $\Delta t = 0.1$ seconds each. Tepekule et al. investigated how many molecules of the total number of received molecules arrive in which time interval (within a given time slot), examining the influence of the ISI on later received bits. In the first time interval, the received molecules were about 60% of the molecules received over the entire time slot; then, the received molecules dropped to 15%, then to 5%, and to below 2% in the subsequent time intervals of $\Delta t = 0.1s$, and became negligible thereafter.

B. DUPLEX NETWORK CODING

In 2016, Akdeniz et al. [262] presented two new approaches for network coding in DBMC nanonetworks. Both approaches simultaneously transmit over both the nanomachine-to-relay channel and the relay-to-nanomachine channel. The arrangement of the nanonetwork is again based on the structure in Fig. 10. Accordingly, network coding in MC communication systems can be divided into three basic types: conventional network coding, half-duplex network coding, and full-duplex network coding.

TABLE 8. Time scheme of conventional network coding, adapted from [260].

Time Slot	Time Interval	$A \to R$	$B \to R$	$R \to A, B$
i	$0 - T_s$	a_i		
	$T_s - 2T_s$		b_i	
	$2T_s - 3T_s$			$\tilde{a}_i \oplus b_i$
$i+1$	$3T_s - 4T_s$	a_{i+1}		
	$4T_s - 5T_s$		b_{i+1}	
	$5T_s - 6T_s$			$\tilde{a}_{i+1} \oplus b_{i+1}$

1) CONVENTIONAL NETWORK CODING

Following [260], in conventional network coding, the communication is split into $3T_S$ second time slots. In time slot i, nanomachine A wants to transmit a message $a_i = \{0, 1\}$ to B, and B wants to transmit a message $b_i = \{0, 1\}$ to A. As illustrated in Table 8, first, the nanomachine A transmits a_i to the nanorelay R in the time interval $[0 - T_s]$. This is followed by the transmission of b_i from the nanomachine B to the nanorelay R in $[T_s - 2T_s]$. The nanorelay encodes (source coding) a_i as $\tilde{a}_i$ and b_i as $\tilde{b}_i$ and transmits the network coded message $r_i = \tilde{a}_i \oplus \tilde{b}_i$ to both nanomachines in the slot $[2T_s - 3T_s]$. Based on the knowledge of the own sent message and the application of XORing, the message can be decoded and converted into the original message. Following the same principle, in time intervals $[3T_s - 4T_s]$ to $[5T_s - 6T_s]$, the next messages a_{i+1} and b_{i+1} can be exchanged between nanomachines A and B, as illustrated in Table 8.

2) HALF-DUPLEX NETWORK CODING

In half-duplex network coding, multiple nanomachines use the channel simultaneously. For a simple comparison with the conventional network coding, we set the time interval for a slot to $3T_s$, analogous to the conventional model illustrated in Table 8. In the first time interval $[0 - 1.5T_s]$ in half-duplex communications, nanomachines A and B simultaneously transmit bit messages a_i and b_i, see Table 9. The spherical absorbing receiver of the relay partially absorbs the molecules from the nanomachines A and B, see Fig. 21, so that no collisions occur. Between $[1.5T_s - 3T_s]$, the relay transmits the combined message r_i to the two nanomachines.

Half-duplex network coding also uses only one molecule type. If the number of received molecules at the relay is very low, both TRXs, see Fig. 21, have sent a bit-0 with high probability (i.e., $r_i = 0 \oplus 0 = 0$). If, on the other hand, the number of received molecules is very high, both TRX have released a bit-1, so that $r_i = 1 \oplus 1 = 0$. For $r_i = 0$, the relay does not release any molecules. If the number of received molecules is between two predefined thresholds, then one TRX has sent a bit-1 and one TRX has sent a bit-0, resulting

TABLE 9. Time scheme of half-duplex network coding, adapted from [262].

Time Slot	Time Interval	$A \to R$	$B \to R$	$R \to A, B$
i	$0 - 1.5T_s$	a_i	b_i	
	$1.5T_s - 3T_s$			$r_i = a_i \tilde{\oplus} b_i$
$i+1$	$3T_s - 4.5T_s$	a_{i+1}	b_{i+1}	
	$4.5T_s - 6T_s$			$r_{i+1} = a_{i+1} \tilde{\oplus} b_{i+1}$

TABLE 10. Time scheme of full-duplex network coding, adapted from [262].

Time Slot	Time Interval	$A \to R$	$B \to R$	$R \to A, B$
$i-1$	$0 - 3T_s$	a_i	b_i	$r_{i-1} = a_{i-1} \tilde{\oplus} b_{i-1}$
i	$3T_s - 6T_s$	a_{i+1}	b_{i+1}	$r_i = a_i \tilde{\oplus} b_i$

in the XORed message $r_i = 1 \oplus 0 = 1$, i.e., the release of the one molecule type by the relay [262].

3) FULL-DUPLEX NETWORK CODING

In the full-duplex communication system, both the nanorelay and the two nanomachines can use the channel simultaneously. Following the two previous communication systems, we set a time interval of $3T_s$ for illustrating full-duplex communication, see Table 10. In the first time step $[0 - T_s]$, nanomachines A and B simultaneously transmit messages a_i and b_i, using molecules of type 1, while in the same time step, the nanorelay transmits r_{i-1} using a different molecule type. Two different types of molecules do not lead to any interactions in the simultaneously used channel, so that the channel is effectively orthogonalized [262]. Since both nanomachines use the same molecule type, interference occurs at the relay [262]. Analogous to half-duplex network coding, receiving two bit-1s from the two nanomachines in time slot i leads to $r_{i+1} = 1 \oplus 1 = 0$; and for two bit-0s, $r_{i+1} = 0 \oplus 0 = 0$. For an interference value between two predefined thresholds, a different molecule type for $r_{i+1} = 1 \oplus 0 = 1$ is released in time slot $i + 1$.

As a result of the simultaneous use of the TRX1-relay channel and the TRX2-relay channel, longer symbol intervals can be used in each of the two communication channels. The longer symbol intervals reduce the ISI in all channels and enable communication over longer distances at higher data rates [262]. The simulation results in [262] demonstrate that the half- and full-duplex network coding approaches outperform the conventional network coding in terms of the BER, especially for long transmission distances [262]. If multiple transmissions are considered, the performance gain of the techniques presented by Akdeniz et al. [262] with respect to conventional network coding increases. The reason lies in the occurrence of ISI: The full-duplex system uses a time interval of $3T_s$ in total for one transmission (to achieve the same data rate as the half-duplex system) [262]. With multiple transmissions in full-duplex network coding, the relative influence of ISI is thus almost negligible [262]. In contrast, half-duplex network coding is more sensitive to ISI, since only time intervals of duration $1.5T_s$ are available. However, half-duplex network coding still outperforms conventional network coding in terms of the BER [262].

C. PHYSICAL-LAYER NETWORK CODING (PNC)

Classical wireless (EM) networks can use physical-layer network coding [269], [270], [271] based on the inverse additivity property of the signals in the medium. In MC, physical-layer network coding was not directly applied until 2016. Ghazani et al. [263] for the first time used different molecule types to implement the XOR operator on the MC physical layer, and this use of different molecule types implements the physical-layer network coding. PNC in MC has to date mainly been examined in the timing scheme of half-duplex network coding, analogous to Table 9.

1) STRAIGHTFORWARD NETWORK CODING (SNC)

In the typical Straightforward Network Coding (SNC), the relay decodes the messages of both transmitters, i.e., from the selected modulation technique, e.g., OOK, the signal, a bit-0, or a bit-1, is decoded. Subsequently, the relay emits the combined messages, recoded with the XOR operator, to the receivers. Each receiver can decode the message of the relay by XORing the received message with its own sent message [272]. Furthermore, if both transmitters send a bit-1, using on-off-keying modulation (i.e., via the emission of a prescribed number of molecules for a bit-1), the blocking of the different molecule types at the receivers is noted as an open challenge in [273].

2) XOR THROUGH CHEMICAL REACTIONS

To solve the open challenge from [273], Ghazani et al. [263] assign the molecule types X_1 and X_2 to the transceivers T_1 and T_2, respectively. Molecule types X_1 and X_2 can react with each other through an irreversible reaction:

$$X_1 + X_2 \xrightarrow{\gamma_{X_{1+2}}} X_{1+2}, \tag{70}$$

where $\gamma_{X_{1+2}}$ denotes the reaction rate of the molecules X_1 and X_2. Furthermore, the reaction molecule X_{1+2} cannot dock to the receptors of the nanorelay. The molecules of the corresponding type i, $i = 1, 2$, i.e., X_1 and X_2, react only with the molecule type R_i of the relay via the reversible reaction:

$$X_i + R_i \underset{\kappa_i}{\overset{\gamma_i}{\rightleftharpoons}} XR_i, \quad i \in \{1, 2\}. \tag{71}$$

whereby, $\kappa \geq 0$ and $\gamma \geq 0$ denote the constant forward and reverse reaction rates of the molecules. The system is designed such that $\gamma_{X_{1+2}} \gg \gamma_1, \gamma_2$. In the case where both transmitters transmit a one, molecules X_1 and X_2 reach the relay and react with each other following Eq. (70). As a result, the concentrations of the molecules X_1 and X_2 in the environment decrease and no molecules bind to the receptors of the nanorelay. In the complementary case, when only one transmitter, T_1 or T_2, emits a one, the corresponding molecules are bound to the relay, and the nanorelay releases molecules of type XR_i into the environment in the next time slot. Thus, Eq. (70) and Eq. (71) are logical XOR operators based on the reactions of the molecules [263].

The implementation of a physical-layer XOR through chemical reactions represents a simple implementation of the

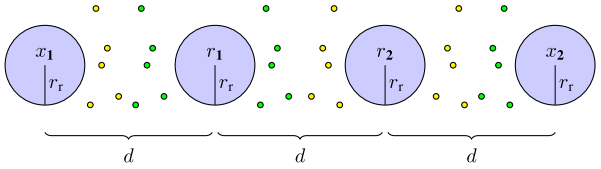


FIGURE 22. Scheme of a multi-hop communication system in nanonetworking, adapted from [265]. The transceiver x_1 wants to transmit a message to the transceiver x_2, and vice versa. Due to the large distance between the two transceivers, two relays r_1 and r_2 are used. The distance between the individual nodes of the nanocommunication system is d. Both the transceivers and the relays have a radius of r_r.

XOR operator. By means of reaction-based schemes on the physical layer, the error probability in two-way relay MC is reduced, and the complexity is reduced since no XOR gate (implemented in electronics) is required at the relay [263].

Physical-layer network coding has been further examined by Farahnak-Ghazani et al. [264], [274]. Analogous to [263], the network coding in [264] is implemented on the physical layer by executing an XOR operation inside the medium through chemical reactions. First, comparing the simulated Bit Error Probability (BEP) for a varied number of released molecules without considering the ISI, Farahnak-Ghazani et al. [264] demonstrate a significant BEP decrease for physical-layer network coding compared to SNC. Then, Farahnak-Ghazani et al. [264], [274] exploit the fast reversible reactions to reduce the ISI, i.e., reduce the number of (old) molecules remaining in the receiver environment through chemical reactions. Thereby, each transceiver of the DBMC system effectively cancels out the ISI of the other transceivers through a chemical reaction of the transmitted molecules [264]. For scenarios with ISI, Farahnak-Ghazani et al. [264] observe a reduced average BEP with PNC compared to SNC for the transmission of various numbers of molecules [264]. Whereby, Farahnak-Ghazani et al. [264] consider a benchmark SNC scheme with an adaptive transmission mechanism that compensates for the number of remaining molecules in the channel (ISI) by transmitting fewer molecules. In contrast, the transceivers in the PNC scheme in [264] transmit extra molecules for the purpose of chemically reacting with and thus effectively eliminating the remaining molecules (ISI) from previous time slots.

D. MULTI-HOP NETWORK CODING

In 2018, the expansion from relay-assisted networks to multi-hop nanonetworks took place in MC. Akdeniz et al. [265] presented a network coding method that can be applied to multi-hop nanocommunication systems. The nanoscale communication system involves two nanomachines that wish to exchange information via intervening nanorelays, see Figure 22. Each nanorelay encodes the symbols coming from its neighbors using XOR and forwards the XORed messages to its neighbors.

We visualize the proposed approach from [265] in Table 11. Whereby T_s denotes the time slot duration,

TABLE 11. Time scheme of multi-hop network coding, where $r_1^k = r_2^{k-1} \oplus x_1^{k-1}$ and $r_2^k = r_1^{k-1} \oplus x_2^{k-1}$, adapted from [265].

Time Slot	$x_1 \rightarrow r_1$	$r_1 \rightarrow r_2, x_1$	$r_1, x_2 \leftarrow r_2$	$r_2 \leftarrow x_2$
$0 - T_s$	x_1^1			x_2^1
$T_s - 2T_s$	x_1^2	r_1^1	r_2^1	x_2^2
$2T_s - 3T_s$	x_1^3	r_1^2	r_2^2	x_2^3
$3T_s - 4T_s$	x_1^4	r_1^3	r_2^3	x_2^4
$4T_s - 5T_s$	x_1^5	r_1^4	r_2^4	x_2^5
$5T_s - 6T_s$	x_1^6	r_1^5	r_2^5	x_2^6
$6T_s - 7T_s$	x_1^7	r_1^6	r_2^6	x_2^7

k denotes the time slot index, x_1 and x_2 denote the respective transceiver of a message, while r_1 and r_2 denote the relay nodes. The evaluation compares the Decode-and-Forward (DF) method [275] (which does not employ network coding at the relay; rather, the relay node decodes the molecules received from the source, re-encodes, and emits molecules to relay the message) with the network coding approach (XOR) for multi-hop nanocommunication systems. The evaluations found that the network coding approach achieves a significantly lower BER compared to DF; whereby, the BER difference increases with increasing number of molecules transmitted per bit-1 [265].

The reason for the BER reduction is the increased symbol duration with the network coding approach in Table 11, which employs the principles of full-duplex network coding, see Table 10. More specifically, the efficiency gains of network coding allow for doubling the symbol duration compared to DF for the same throughput rate of messages delivered to their respective destination transceiver per time slot: In steady state, DF delivers one symbol to a destination transceiver per time slot [265], compared to two symbols delivered to a destination transceiver with network coding, see Table 11 where r_2 delivers a symbol to x_2 and r_1 delivers a symbol to x_1 in each time slot in steady state. The longer symbol duration that is enabled by network coding induces more power into the signal and less power into the ISI. The evaluations in [265] also examined the BER increase with increasing number of nanorelays for up to four relays.

E. PERFORMANCE IMPROVEMENT OF NETWORK CODING

Goel et al. [266] present approaches for improving the performance of network coding in DBMC nanonetworks. Various factors that govern the performance of a molecular nanocommunication system were identified: *(i.)* the diffusion coefficient of the information carrier molecules, *(ii.)* the distance between the two communicating nanomachines, *(iii.)* the volume of the nanomachines, *(iv)* the time difference from emitting to absorbing the molecules, and *(v)* the concentration of each type of molecule in the surrounding fluid [266].

Cheng et al. [267] present new network coding approaches using a combination of the decode-and-forward relaying protocol, in which the same number of molecules is released by the nanomachines, and the network coding scheme, in which different numbers of molecules are released by the nanomachines. The DBMC system consists of two source nanomachines and a relay nanomachine in a three-

dimensional channel. The relay R is located between the two source transmitters S_1 and S_2. Furthermore, Cheng et al. assume that the transmitters are not within each other's transmission range [267]. The three nodes are passive absorbers, i.e., the molecules can diffuse through them without a reaction taking place. All nodes are also perfectly synchronized in terms of time. Analogous to Table 9, half-duplex network coding is considered. As a result, two time intervals (see Table 9) are required for a transmission, while four time intervals are required for a transmission without network coding ($S_1 \rightarrow R, R \rightarrow S_2, S_2 \rightarrow R, R \rightarrow S_1$).

The numerical simulations in [267] reveal a significantly better performance in terms of the average error probability for the network coding approach compared to no network coding [267]. As the number of released molecules increases, the average error probability decreases [267]. The higher the number of released molecules, the better the performance of the network coding approach [267]. The findings of the simulations can be explained on the one hand by the time required. Two time intervals for half-duplex network coding (see Table 9) lead to twice as many transmissions in a given time span than transmission without network coding (which requires four time intervals for one transmission, see Fig. 10(a)). Consequently, the error probability in absolute terms decreases as the transmission rate decreases. Also, an increase in the number of molecules leads to a decrease in the error probability, since the threshold value for the detection of a one is significantly exceeded or undershot.

In addition, an algorithm using Particle Swarm Optimization (PSO) [276] is applied in [267] to minimize the average error probability. More specifically, the PSO determines the limit value for OOK for all nanomachines in the system in order to minimize the average error probability. PSO searches for a solution to an optimization problem using biological swarm behavior. Analogous to the natural swarm phenomenon, a population of candidate solutions is moved through the search space to obtain a good solution. In each computational step, the position of each individual is recalculated [276]. The simulations demonstrate good convergence of the PSO algorithm [267]. Also, the introduced PSO algorithm has a significantly lower number of iterations compared to other algorithms, e.g., the Block Coordinate Descent Algorithm (BCDA) [277]. The average error probability of the DBMC system can be significantly reduced with the optimal threshold, both for the same number of molecules for the decode-and-forward relaying protocol and a different number of released molecules for the network coding scheme released by the nanomachines [267]. Cheng et al. also demonstrate through simulations that increasing the number of released molecules decreases the error probability, and increasing the distance between adjacent nodes increases the average error probability [267].

F. NETWORK CODING IN MOBILE NANONETWORKS

Cheng et al. [268] present for the first time a mobile DBMC system. The changing distance between the transmitters leads to time-varying channels [268]. Two source nanomachines, denoted as S_1 and S_2, exchange information using the nanorelay R, as illustrated in Fig. 23. In contrast to the previous models, the three nodes are mobile in a fluid environment [278]. In the initial feed sand, the relay R is located between the two nodes S_1 and S_2. In addition, it is assumed that S_1 is not in the transmission range of S_2 and the three nodes are time synchronized [268]. Quaternary MoSK modulation [279] is considered as a baseline modulation scheme. Quaternary MoSK uses four different molecule types for the associated bits, i.e., a node transmits the information bits 00, 10, 01, or 11, by releasing the molecule types A, B, C, or D, respectively.

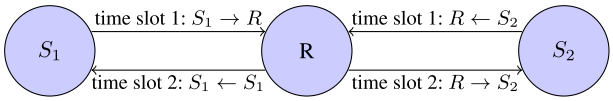


FIGURE 23. The mobile molecular communication system via diffusion using Amplify-and-Forward (AF) relaying and Analog Network Coding (ANC), adapted from [268]. Half-duplex network coding, analogous to Table 9, is utilized.

Cheng et al. consider also quaternary Depleted Molecule Shift Keying (D-MoSK) [280], [281], which we briefly outline for the example of a transmission from node S_2 to node S_1. If bits 00 are transmitted in time slot k, node S_2 does not release any molecules into the fluid environment. If, on the other hand, bits 01, 10, or 11 are transmitted by node S_2 in time slot k, then node S_2 releases, respectively, the molecule type A, molecule type B, or both molecule types A and B ($A+B$) simultaneously. After propagation through the channels and relaying by the relay R, the molecules arrive in the time slot $(k+1)$ at the receiver node S_1 and are decoded as follows. Let $Y_{S_1}[k+1]$ denote the detected symbol at node S_1 in time slot $(k+1)$. Let $N^A_{(S_1)}$ and $N^B_{(S_1)}$ denote the numbers of the received molecules of type A and B, respectively, also let $\theta^A_{S_1}$ and $\theta^B_{S_1}$ denote the optimal decision thresholds at node S_1 for molecule types A and B, respectively.

If the number of the detected molecules of type A in time slot $(k+1)$ is higher than the optimal decision threshold, i.e., if $N^A_{(S_1)}[k+1] \geq \theta^A_{S_1}$, then node S_1 detects a one for the first digit, otherwise it detects a zero. Analogously, if the number of the detected molecules of type B in time slot $(k+1)$ is higher than the optimal decision threshold, i.e., if $N^B_{(S_1)}[k+1] \geq \theta^B_{S_1}$, then node S_1 detects a one for the second digit, otherwise it detects a zero, i.e.,

$$Y_{S_1}[k+1] = \begin{cases} 11, & \text{if } N^A_{(S_1)}[k+1] \geq \theta^A_{S_1}, \\ & N^B_{(S_1)}[k+1] \geq \theta^B_{S_1} \\ 00, & \text{if } N^A_{(S_1)}[k+1] < \theta^A_{S_1}, \\ & N^B_{(S_1)}[k+1] < \theta^B_{S_1} \\ 10, & \text{if } N^A_{(S_1)}[k+1] \geq \theta^A_{S_1}, \\ & N^B_{(S_1)}[k+1] < \theta^B_{S_1} \\ 01, & \text{if } N^A_{(S_1)}[k+1] < \theta^A_{S_1}, \\ & N^B_{(S_1)}[k+1] \geq \theta^B_{S_1}. \end{cases} \tag{72}$$

To simulate mobility, the positions of the molecules and the nodes are tracked in the simulation environment. At each

time step of the simulation, the molecules, node S_1, node S_2, and relay R are subject to a random walk. Let Δt denote the time interval duration, $D_{A,B}$ denote the diffusion coefficients of the molecule types A and B (which are assumed to be equal in [268]). Also, D_{S_1} and D_{S_2} denote the diffusion coefficients of nodes S_1 and S_2, respectively, and D_R denotes the diffusion coefficient of the relay R. The new positions are determined in each coordinate by sampling a random Gaussian variable with mean 0 and the respective standard deviation $\sqrt{2D_{A,B}\Delta t}$, $\sqrt{2D_{S_1}\Delta t}$, $\sqrt{2D_{S_2}\Delta t}$, and $\sqrt{2D_R\Delta t}$.

The evaluations in [268], which include numerical evaluation based on mathematical analysis and verifying simulations, compare the quaternary D-MoSK modulation with the quaternary MoSK modulation in terms of the average mutual information as the number of released molecules increases. Mutual information quantifies the information that one random variable conveys about another random variable [282]. Also, the evaluations in [268] compare an Amplify-and-Forward (AF) protocol with an Analog Network Coding (ANC) protocol. Generally, in an AF protocol, the relays linearly process the received signals and forward them to the receiver [283], i.e., the received signal is amplified and retransmitted without signal processing [284]; specifically, in [268], AF corresponds to communication without network coding, see Fig. 10(a). The ANC protocol in [268] corresponds to half-duplex network coding, see Table 9.

For quaternary D-MoSK (which uses only two different molecule types, A and B), an average achievable mutual information of 1 symbol per second can be achieved; while quaternary MoSK (which uses four different molecule types, A, B, C, D) achieves only 0.5 symbols per second [268]. Also, the smaller number of molecule types results in a lower decoding complexity of the quaternary D-MoSK. Furthermore, ANC achieves better performance in terms of average Symbol Error Probability (SEP) and average mutual information when compared with the AF protocol (i.e., no network coding) [268]. Compared to no network coding (see Fig. 10(a)), fewer time slots are needed for transmission with half-duplex network coding (see also Table 9). Consequently, more information can be transmitted in less time, and the average mutual information increases and converges more rapidly toward 100 percent for an increasing number of released molecules [268].

G. SUMMARY AND DISCUSSION

In [259], Unluturk et al. presented a simple MC system at the nanoscale, which takes into account two different cases, namely the network coded and the uncoded case. However, the simulation results demonstrate that a high data rate and negligible delay cannot be achieved simultaneously [259]. The Unluturk et al. study [259] served as a foundation for subsequent network coding approaches for DBMC systems. Aijaz et al. [260] present closed-form expressions for network coding in DBMC systems and its error performance. The simulation results in [260] demonstrate the dependence of the error probability on the ISI probability, the diffusion coefficient, the radius of the spherical receiver, and the pulse amplitude.

The study [260] was followed in 2016 by Akdeniz et al. [262] who present approaches for decoding the combined messages rather than decoding the individual messages, thus enabling network coding for duplex (simultaneous two-way) communication. Consequently, the data rate can be increased, or the reliability of the system can be improved at a constant data rate [262]. Simulation results also demonstrate that the schemes presented in [262] are better suited for communication with a longer transmission distance.

Further in 2016, Ghazani et al. [263] presented the first study of physical-layer network coding (PNC) for MC, which is also known as reaction-based physical-layer network coding. PNC uses different molecule types to form the XOR based on an irreversible chemical reaction. Consequently, no electronic XOR gate is needed at the relay, simplifying the implementation [263]. The simulation results in [263] demonstrate that this simple reaction-based physical-layer network coding achieves low error probabilities.

In summary, physical-layer network coding mainly focuses on implementing an XOR logical gate via an irreversible chemical reaction of two molecule types. In contrast, duplex network coding (Section IX-B) is not based on chemical reactions: With half-duplex network coding, one molecule type is sufficient to use the channel simultaneously. Full-duplex network coding uses two molecule types that do not react chemically and are only absorbed by the receiver to enable the transmission of three messages in the same time slot (see Table 10). Thus, duplex network coding does *not* involve chemical reactions. Rather in duplex network coding, the relay senses the sum (in XOR form, i.e., $\oplus$) of the bits (0 or 1) that the two transceivers transmitted (through the non-release or release of molecules); however, the released molecules do *not* chemically react with each other. In particular, for two transmitted bit-0s, the relay senses $0 \oplus 0 = 0$; for two transmitted bit-1s, the relay senses $r_{i+1} = 1 \oplus 1 = 0$; while for one transmitted bit-0 and one transmitted bit-1, the relay senses $0 \oplus 1 = 1$. Thus, duplex network coding effectively implements the XOR via the sensing mechanism in the relay.

We also note that PNC in MC has so far only been examined with the timing scheme of half-duplex network coding. Half-duplex network coding utilizes two time slots to exchange a bit transmission between two transceivers via a relay, see Table 9, i.e., achieves a throughput of one bit delivered to a destination transceiver per time slot (in steady state). Future research should explore PNC for the full-duplex network coding timing scheme. Full-duplex network coding timing would enable a (steady state) throughput of one bit delivered to the respective destination transceiver per time slot for exchanging a bit transmission between two transceivers via a relay. Thus, full-duplex network coding effectively delivers two bits to the respective two destination transceivers per time slot (in steady state).

Future research should also explore the combination of the chemical reaction based ISI mitigation approach in the PNC approach in [264] with alternative approaches, e.g., enzymatic reactions [285], [286] or photolysis reactions [287], which could further enhance the performance of PNC. Also, heterogeneity aspects of the MC system, e.g., asymmetry aspects of the relay which have been explored in [288], should be comprehensively investigated in future research.

In 2018, Akdeniz et al. present in [265] the first approaches for a multi-hop nanonetwork. Generally, if the distance between the transmitter and receiver in a molecular nanonetwork increases, then the probability of successful transmission of the molecules decreases [265]. Nanorelays can be used to increase the reliability of the system. In contrast to the Decode-and-Forward Method (DFM), in the network coding method presented in [265], the messages of the two neighboring nodes are received simultaneously and then XORed. Only then forwarding to the next nodes will be performed. This process continues from the transmitter to the receiver over all intermediate relay nodes. Finally, at the receiver, the message is decoded. The simulation results demonstrate a significant BER reduction compared to the DFM [265].

Cheng et al. [268] investigated network coding in a mobile MC system based on AF and ANC in 2020 (published in 2022). The two transceivers and the relay are free to move in a fluid environment. Simulation results demonstrate that network coding can substantially improve the performance in mobile DBMC nanonetworks; however, the simulations also indicate that the type of modulation and various other system parameters strongly influence the mobile nanonetwork performance. Comprehensive further studies of network coding in mobile nanonetworks need to follow up on the initial study by Cheng et al. [268] in this important type of DBMC nanonetwork. In [267], the average transmission error probability in an MC system is minimized with a PSO algorithm, which exhibits good convergence behavior [267].

Overall, Table 7 indicates that the complexity of the network coding approaches as well as their capacity has generally increased with the chronological order of the development. Currently, the PSO algorithm [267] represents the culmination of this chronological development of network coding approaches for DBMC systems, achieving the highest currently achievable capacity, while incurring the highest complexity among the existing network coding approaches.

X. OVERALL SUMMARY AND DISCUSSION

A. OVERVIEW + SOURCE CODING

This overall summary and discussion section concisely summarizes the main insights and lessons learned from the existing studies of each of the three main coding types (source, channel, and network coding) in the context of DBMC nanonetworks. Also, importantly, this section distills the overall inferences, limitations, and insights across the three coding types, i.e., source, channel, and network coding, as well as the overarching insights across the different categories of channel coding (from Sections V–VIII).

Source coding for DBMC systems is still a relatively novel research field. The DBMC source coding studies [98], [99] by Dhayabaran et al. date back to 2020 and 2021. In particular, Dhayabaran et al. demonstrate that the BER can be significantly reduced if the fraction of bit-1s in a code word to be transmitted is minimized [98], [99], which is a unique feature of source coding for MC communication that is distinct from source coding for conventional wireless (EM) communication.

B. CHANNEL CODING: FROM EM COMMUNICATIONS TO MC-SPECIFIC CODING

Most coding research in DBMC systems to date relates to channel and network coding. Generally, the MC channel and network coding schemes can be categorized into coding schemes that have been adapted from wireless communication and coding schemes that have been specifically designed for MC [62]. This discussion subsection views and interprets channel coding for MC from the perspective of the research and development progression from experimenting with conventional channel codes in DBMC settings to designing novel channel coding specifically for the unique characteristics of DBMC channels. Specifically, for MC channel coding, conventional channel codes from wireless EM communication, such as linear block codes (e.g., Hamming codes [128], [131], [133], [202], [203], [204], [205], [206], [207], Reed-Muller codes [209], [210], [211], and Reed-Solomon codes [133], [208]), convolutional codes, and some miscellaneous codes (e.g., constant-weight and parity check codes) have been integrated into DBMC systems and their BER performance has been evaluated. Within the context of channel coding schemes from wireless EM communication, it has been demonstrated that approaches that employ distance metrics specifically developed for DBMC, e.g., in the convolutional channel coding study [216], achieve lower BER than conventional general approaches (that are not DBMC system specific, such as approaches based on the Hamming distance).

There have also been contrasting developments of ISI-mitigation channel codes specifically for DBMC nanonetworks, e.g., in [217], [218], [221], and [136]. Specifically, Shih et al. [217], [218] proposed an (n, k, l) ISI-free coding scheme and extend their approaches up to a level-5 crossover ($l = 5$). Akhkandi et al. [221] introduced crossover resistant coding with time gap (CRCTG) for DBMC. Kislal et al. [136] proposed a low-complexity ISI-mitigation channel coding scheme for DBMC.

In general, channel codes for MC can be categorized into linear block codes (Section V), convolutional codes (Section VI), ISI-mitigation codes (Section VII), and miscellaneous codes (Section VIII). Linear block codes have a low to medium complexity and, concomitantly, only a low to medium error correction capability. In contrast, ISI-mitigation codes have a higher complexity, but also a higher error correction capability (except for ISI-free codes, which, strictly speaking, do not correct errors). Linear block codes and ISI-mitigation codes have received substantial attention

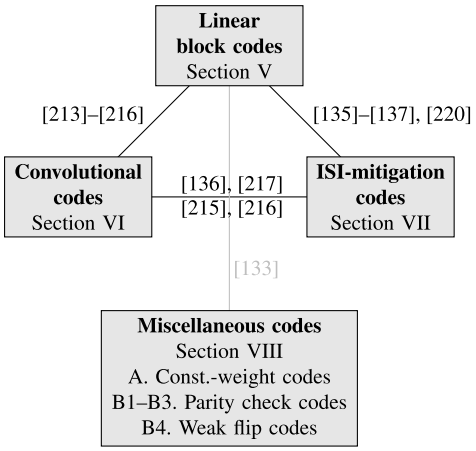


FIGURE 24. Summary of existing cross-category comparative evaluations for channel coding approaches.

in DBMC coding research to date, as indicated by the relatively large quantities of DBMC linear block and ISI-mitigation channel coding studies in Table 5. In contrast, convolutional DBMC channel coding has received only moderate research interest so far. Most recently, convolutional DBMC channel coding studies have focused on different distance metric functions (see Section VI-C or Section VI-D), for example, in [215] and [216]. Overall, convolutional codes have medium complexities and low to medium error correction capabilities.

Miscellaneous DBMC channel coding studies have examined a number of coding techniques for the main purpose of channel coding, namely reducing the errors introduced by signal propagation in the MC channel (medium), specifically constant weight codes (Section VIII-A), as well as parity check codes and weak flip codes (Section VIII-B). The error correction capabilities of these miscellaneous codes range from low to high, while their complexities range from low to medium.

In addition, some miscellaneous DBMC channel coding studies have examined other aspects of channel coding, namely spatial channel coding techniques (for MIMO) (Section VIII-C), and minimum energy channel codes (Section VIII-D) which are important for implementing channel coding in nanomachines due to limited resource availability. Thereby, spatial (MIMO) channel codes primarily address how to exploit a particular architecture set-up (configuration) consisting of multiple spatially diverse transmitters and receivers. Whereas, minimum energy codes address how to maintain a desired distance of the code words for ensuring reliability while minimizing the energy demand by reducing the number of bit-1s in the code words [204]. In future research, all channel coding approaches (linear block codes, convolutional codes, ISI-mitigation codes, constant weight codes, parity check codes, and weak flip codes) should be examined in the setting of spatially diverse transmitters and receivers (MIMO) and with respect to their energy efficiency. Also, the potential synergies of source coding for MC and energy efficient channel coding, which both strive to reduce the number of bit-1s over the channel, should be explored in future research.

C. CHANNEL CODING: PERFORMANCE COMPARISON PERSPECTIVE

This discussion subsection interprets channel coding from the perspective of performance comparisons between the different categories of channel coding approaches. In Fig. 24, we summarize the existing cross-category comparative evaluations of channel coding approaches. Fig. 24 indicates that linear block codes, e.g., Hamming codes, have been extensively used as benchmarks for the performance evaluation of convolutional codes [213], [214], [215], [216] and ISI-mitigation codes [135], [136], [137], [220]. More specifically, the comparative evaluations indicate that the BERs of the convolutional channel codes are not significantly lower than the uncoded systems for small constraint lengths, e.g., $K = 3$ [215], only the BERs of the convolutional channel codes with large constraint lengths, e.g., $K = 5$ and $K = 7$, are significantly lower than for the uncoded systems [215]; also, convolutional channel codes can achieve a higher energy efficiency [214]. Furthermore, the existing comparisons of ISI-mitigation channel codes (which have been specifically designed for DBMC), with the linear block codes (which originate from wireless (EM) communications), indicate that ISI-mitigation channel codes generally achieve a lower BER than linear block codes, e.g., Hamming codes [135], [220].

As Fig. 24 indicates, studies comparing all categories of channel coding approaches in DBMC are lacking. Only the studies [215], [216] have conducted cross-category comparisons of convolutional codes versus linear block codes and ISI-mitigation codes. Only the study [133] compared a miscellaneous code versus linear block codes. Specifically, Lu et al. [133] examined EGLDPC and C-RM codes for channel coding in an MC system. In the evaluations in [133], Lu et al. compared the examined EGLDPC and C-RM codes against the Hamming code and the uncoded system. Channel coding schemes in DBMC are often compared only to an uncoded transmission scenario, e.g., in [131], [202], [203], [210], [211], and [218] as well in some parts of [133], [216], and [217], or to different channel coding schemes within their category, e.g., for linear block codes between Hamming codes and Reed-Muller codes [133] or Reed-Solomon codes [210].

Thus, overall, we conclude that cross-category comparisons are relatively rare; whereby cross-category comparisons of miscellaneous codes against the other three code categories are especially rare. Therefore, the cross-category comparison of miscellaneous codes, i.e., error free constant-weight codes (Section VIII-A), parity check codes (Secs. VIII-B1–VIII-B3), and weak flip codes (Section VIII-B4), against linear block codes (Section V), convolutional codes (Section VI), and ISI-mitigation codes (Section VII) is wide open for future research and should be

incorporated into a comprehensive evaluation framework for coding in DBMC systems (Section XI-D2).

In summary, the limited comparative studies of DBMC channel coding approaches that are available to date indicate that channel codes that have been designed specifically for DBMC systems achieve lower BER and can be better adapted to the MC modulation techniques, e.g., OOK or MoSK, than conventional (EM) channel coding approaches. However, comprehensive comparisons of all existing channel coding categories are lacking. Such comprehensive comparisons are very important to conduct in future research as they could inform the scoping of future research directions in channel coding, e.g., focusing on a specific set of promising channel coding categories or approaches for the development of future DBMC channel codes.

D. NETWORK PERSPECTIVE

1) NETWORK CODING

Network coding is a promising coding technique for future DBMC networks. In particular, using the physical-layer network coding approach, biological XOR gates can be synthesized based on chemical reactions [263]. Approaches for multi-hop systems [265] to increase the transmission distance as well as half-duplex and full-duplex systems [262] to increase the data rate improve the performance of DBMC nanonetworks.

2) MOBILITY

Importantly, Cheng et al. [268] have investigated network coding for a mobile MC model, i.e., a communication system consisting of a mobile source nanomachines, a mobile relay nanomachine, and a mobile sink (destination) nanomachine. Aside from the Cheng et al. [268] study on network coding and the Qiu et al. [220] study on ISI-mitigating positional-distance codes in a mobile DBMC nanonetwork, there have only been few studies on other (non-coding) aspects of mobile DBMC nanonetworks, e.g., studies on power control and adaptive detection [289], [290]. Generally, investigations of all types of coding (source, channel, and network coding) for mobile nanomachines are of critical importance, as the nanomachines for a myriad of Internet of Bio-Nano Things applications, e.g., in-vessel nanomachines for intra-body communication, are mobile. However, such studies of coding for DBMC in mobile scenarios are lacking and present an important direction for future research.

E. PERSPECTIVE ON MOLECULES, MODULATION, AND INTEGRATED CODING

1) MOLECULES

In summary, MC is particularly suitable for nanoscale communication networks. In order to communicate on the nanoscale, particles or molecules, also on the nanoscale, are required. The presented coding techniques, i.e., source, channel, and network coding techniques, overall improve the performance of DBMC systems with respect to the ISI, and consequently with respect to the BER.

The chemical compositions and physical properties of molecules can be manipulated, as, e.g., noted in [259]; however, this molecule manipulation has not yet been exploited for coding in DBMC nanonetworks. Future research should explore deliberate manipulations of molecules to enhance their utility for DBMC nanonetworks.

Another limitation of the presented studies is inherent in the transmitted messages. All existing studies have considered the transmission of bit sequences, regardless of the respective specific application areas or use cases of the DBMC systems. However, specific application areas and use cases, such as body internal networking or environmental monitoring, involve use-case-specific messenger substances. More specifically, the human body, as sender and/or receiver of "data" does not communicate with abstract bit-1s and bit-0s; rather the human body communicates via the release of hormones or similar messenger substances (as sender). None of the existing coding studies has considered coding specifically for use-case-specific messenger substances in DBMC systems.

2) MODULATION

All the surveyed coding techniques are subject to the limitation that they can only be applied to the modulation technique considered in each respective study. That is, the surveyed coding techniques are not generally valid for arbitrary modulation techniques. Moreover, most of the evaluation results are based on simulations; only few studies, e.g., [136], [137], and [222], present testbeds for evaluating the proposed coding techniques.

3) INTEGRATED CODING

The existing coding studies are limited to addressing a single coding technique in isolation, i.e., each study addresses either source, channel, or network coding; none of the existing studies examine a combination of two or more coding techniques. That is, no MC nanosystem was simulated in its entirety, using multiple coding techniques, e.g., channel and network coding. Also, no uniform standards have been established yet, which could provide basic frameworks as well as structures and create a uniform common reference framework for the field of coding for DBMC nanonetworks, as elaborated in Section XI-D2.

XI. FUTURE RESEARCH DIRECTIONS

This section outlines future research directions for coding in DBMC nanonetworks, divided into the three types of coding techniques, namely source, channel, and network coding. Moreover, we outline overarching future research directions for coding in DBMC nanonetworks that span across the three coding types.

A. SOURCE CODING

Future source coding research on DBMC systems should include the development of new techniques to improve the channel capacity and the analysis of MISC techniques in a diffusive communication channel in a time-varying

manner [99]. With respect to the basic goal of source coding, namely the removal of redundancy, significant improvements can likely be achieved in MC nanonetworks in terms of transmission rate or latency. A major drawback of current DBMC networks is the slow transmission rate, which could be counteracted by compressing the information to be sent. Future research should therefore aim at enhancing the effectiveness of source coding in DBMC nanonetworks, focusing on approaches that increase the effective transmission bit rate and reduce the effective latency of the transmitted messages.

Additional future research directions arise in the development of new source coding protocols for DBMC-MIMO systems, also with the application of distributed source coding [94]. Another future research opportunity arises from timely lossless source coding symbols that arrive randomly [291]. Real-time communication systems based on DBMC, such as environmental monitoring or intra-body networks, require efficient compression at the source so that the receiver can reliably decode the received message despite limited network resources. In terms of a simple model, this problem corresponds to real-time compression over a resource-constrained data network. In contrast to traditional source coding, where the code word length is to be minimized, the goal of timely lossless source coding is to approximate the Shannon entropy of the source [291]. Approaches include asking what the optimal compression scheme would be if each update to the source is transmitted to the receiver over a binary channel using a fixed transmission rate so that the information at the receiver is still kept as up-to-date as possible. Whereby, such MC approaches focused on timeliness could incorporate the concept of age of information [292] in future research.

Furthermore, in recent years, the amount of data produced and stored by mankind in communication networks has increased immensely. This data is going to play a key role in multi-faceted aspects of human life, such as health care, economy, education, marketing, and recreation. Source encoding for DBMC systems should counteract the immense data traffic in the future, also from the point of view of the low data transmission rate of MC systems. One possible solution is based on DNA-based molecular computing, memory, and communication [293]. Another approach is the transition from semantic coding to semantic communication (joint source-channel coding) [294]. The basic idea is to *not* transmit the statistical probabilities, but rather the logical probabilities [295]. For the event "A and B" is true, the logical probability is 0.25 [295]. For the event "A or B", the logical probability is 0.75 [295]. The amount of semantic information H_S can be defined as a negative logarithmic value of the logical probability [296], so that $H_S(A \wedge B) = -\log_2(0.25) = 2$, and $H_S(A \vee B) = -\log_2(0.75) = 0.415$, while $H_S(A) = -\log_2(0.5) = 1$, and $H_S(B) = -\log_2(0.5) = 1$ [295].

Now, consider the statement "A and B". From an intuitive point of view, the statement "A and B" contains more information than the statement "A" [295]. If "A and B" is true, then "A" is true, but not vice versa. Bao et al. [295] define the logical probability of a sentence by the measurement of the likelihood that a sentence is true in all possible situations. Suppose that events "A" and "B" are independent of each other [295]. Both events are either true or false, analogous to a fair coin toss. Event "A" can be true or false, and event "B" also true or false, giving rise to four possible events, each with a statistical probability of 0.25 [295]. In conclusion, the largest value of the amount of semantic information H_S contains the most information, i.e., the statement $H_S(A \wedge B) = 2$ contains more information than just $H_S(A) = 1$ or $H_S(B) = 1$, since for a true $(A \wedge B)$ statement, both A and B are true. Furthermore, $H_S(A) = 1$ or $H_S(B) = 1$ contains more information than $H_S(A \vee B) = 0.415$, since only one statement (statement A or statement B) is correct.

B. CHANNEL CODING

This subsection outlines the future research directions of each category of channel coding, see Table 5.

1) LINEAR BLOCK CODES

For Hamming codes, Lu et al. [131] showed that there is a barrier for the Hamming order, for which new codes have to be proposed, to achieve further gains in DBMC. Future research is needed to determine the barrier for other linear block codes. Also, for Reed Solomon codes, an adaptive threshold scheme for detection needs to be developed so that the optimal combination of the coding parameters n and k can be obtained for (n, k) RS codes [209]. With respect to the VLSI implementation of the Hamming and cyclic Reed-Muller codes, application-specific integrated circuits are an important future work direction [206], [208].

The existing linear block channel coding studies in DBMC systems indicate that the coding performance tradeoffs are strongly dependent on the coding parameters n and k. Since the different use cases of the different applications require different coding performance tradeoffs, it is important to thoroughly evaluate the full gamut of linear block codes for a wide range of combinations of the coding parameters n and k to determine the parameter combinations that can optimally support specific application use cases.

2) ISI-MITIGATION

Kislal et al. [135] noted that it is necessary to improve their ISI-mitigation algorithms, namely the Greedy algorithm, the genetic algorithm, and the MIP, in future work. Linear block codes need to be constructed, so that simple implementations for the use in future nanomachines become possible [135]. Keshavarz-Haddad et al. [223] propose to investigate the use of concentration modulation over a noisy communication channel in two- or three-dimensional environments in future research. Under the prevailing boundary conditions in such a noisy two- or three-dimensional environment, the number of molecules received at the receiver side varies with large amplitude; thus, new channel coding schemes,

especially for the occurrence of noise in the channel, must be developed [223].

Future research also needs to develop ISI-mitigation schemes for simultaneously deployed MC systems. Suppose that several local DBMC systems are communicating independently. The signals (molecules) of a specific given DBMC system should not be affected by the signals (molecules) of the other simultaneously operating DBMC systems.

3) MOLECULAR CODING DISTANCE FUNCTION AND CONVOLUTIONAL CODES

In 2012, the molecular coding distance function developed by Ko et al. [128] for OOK laid the foundation for research on specific molecular coding distance functions, thereby introducing a new channel coding concept for DBMC systems. Future research should extend the molecular coding distance function to other modulation schemes, e.g., level-based or type-based modulation schemes [238], in MC systems. Also, future research should develop efficient decoding algorithms based on the distance function and the state functional properties [128].

With the asymmetric-distance metric for DBMC systems and channel codes based on the asymmetric-distance metric presented by Li and Li [215] in 2018, the topic of distance functions for MC received renewed relevance. According to Li and Li, the metric could serve as a basis for future channel codes for MC systems. Future work should extend the asymmetric molecular coding distance function to other modulation techniques, beyond the MoSK scheme considered in [215].

4) MISCELLANEOUS CODES

a: SPATIAL CODES

Damrath et al. [230], [231] presented various spatial coding mechanisms for a DBMC 2×2 MIMO system in a three-dimensional environment. For future research, they suggested developing practical MIMO algorithms and efficient implementation strategies, including for asymmetric MC system arrangements [231]. The energy cost, the system complexity, and the implementation of strategies to overcome obstacles that interrupt or weaken the transmitted signal in the communication channel should also be investigated in future research [231]. An extension with respect to a larger number of transmitters and receivers (scalability to MIMO systems) is also mentioned by Damrath et al. as a goal for future research [231].

b: POLAR CODES

Polar coding represents a potential starting point for future research on channel coding in DBMC nanonetworks. To the best of our knowledge, polar codes, which are state-of-the-art block codes, see Fig. 6, have not yet been examined in detail in the context of DBMC nanonetworks. The polar code is a linear block error correction code [113]. The design of the polar code is based on multiple recursive concatenations of a short code, which in turn map the communication channel to multiple, virtual channels [113]. For a large number of recursions, the channels tend to either have a very high or very low reliability, i.e., the channels polarize. The bits to be transmitted are assigned to the most reliable channels [113]. Matsumine et al. [297] propose the first approaches of a polar coding method that uses molecular programming. At the nanoscale, chemical reaction networks can be developed without electrical energy sources to apply successive cancellation or maximum likelihood decoding methods for polar codes with short code word lengths [297]. These incipient molecular nanoscale approaches to polar coding open the new research field of polar coding for DBMC nanonetworks, which is still largely unexplored.

C. NETWORK CODING

The analysis of the rate-delay tradeoff in DBMC nanosystems with network coding in [259] shows that a high data rate and negligible latency cannot be achieved simultaneously [259]. For future research, the finding of Unluturk et al. should be exploited and serve as a foundation for optimizing nanoscale MC systems. For instance, this tradeoff should be explicitly considered in molecular applications in the nanoscale range that are sensitive to either a specific minimum transmission rate or a specific maximum delay [259].

Building on the optimization of the decision thresholds based on network coding for an MC system, Cheng et al. [267] note that future research should construct new relay schemes and further optimize the decision threshold algorithms [267]. Furthermore, in the future, chemical-reactive techniques should be developed to apply pilot-based channel estimation techniques in an experimental practical communication system using diffusion.

A final interesting starting point for future work in network coding of DBMC systems is sparse network coding [298], [299], [300], [301], [302], [303]. Sparse network coding is a technique that is intended to reduce the complexity of random linear network coding [304]. To implement the basic idea, a sparse coefficient matrix is selected for coding the packets. More specifically, the numbers of bit-1s can be deliberately reduced in the XOR network coding operations to make the coding sparse [300], [302]. In the context of DBMC nanonetworks, such a sparse coding could be explored as a mechanism to reduce the ISI, as fewer bit-1s that have to be communicated through diffusing molecules generally tend to reduce the ISI [98], [132], [249].

D. OVERARCHING FUTURE RESEARCH DIRECTIONS FOR CODING IN DBMC NANONETWORKS

This subsection outlines overarching future directions for coding in DBMC systems.

1) FLEXIBLE EFFICIENT MODULATION AND CODING

Reliability is a main open challenge for all types of coding, and is especially challenging for channel coding, which has to counteract the noisy channel effects. Future research needs to fundamentally advance the construction of source, channel, and network codes that operate reliably in DBMC nanonetworks. Importantly, future coding algorithms should

meet the constraints and characteristics of a wide variety of MC channels while ensuring highly reliable communication. Also, future coding techniques should incur only low complexity for the encoding and decoding processing in the end nodes and the recoding processing in the network nodes in case of network coding.

The existing coding schemes should be extended and complemented to accommodate a wide variety of modulation techniques, such as hybrid or timing-based modulation techniques, to ensure high levels of efficiency and reliability in a variety of MC systems. Generally, the modulation techniques of MC can be divided into concentration-based, type-based, timing-based, spatial, and hybrid modulation techniques [51]. The existing coding studies for DBMC nanonetworks focused on: *i.)* concentration-based modulation, specifically BCSK, OOK, MCSK, and PAM; as well as *ii.)* type-based modulation, specifically MTSK and MoSK. Also, a timing-based modulation technique is applied in Section VI-D, and the first approaches for spatial MIMO systems are applied in Section VIII-C. In future research, hybrid modulation techniques as well as timing-based modulation techniques should be given more attention.

Future research should develop overarching concepts and protocols for synergistic combinations of techniques for the three types (purposes) of coding to ensure optimal adaptation to MC systems. Currently, there are different (isolated) schemes for each type (purpose) of coding. Combinations of the three different coding schemes (integrated coding) could achieve synergies by letting source, channel, and network coding cooperate with mutual awareness and with exploitation of synergies versus the current approach of each type of coding working in isolation, without awareness of the bigger picture that incorporates the other types of coding. Analogous to the successes of cross-layer optimization in other types of networks [305], [306], [307], [308], [309], [310], [311] as well as specifically for different types of coding [312], [313], [314], [315], [316], a ''cross-coding type'' approach is expected to achieve substantial performance enhancements.

The existing coding techniques have been primarily designed and evaluated for point-to-point communication over an MC channel. Multiuser communication with multiple transmitters and/or multiple receivers [317], [318] has received very limited attention so far. To the best of our knowledge, only [210], [211], [230], and [231] have considered multiple transmitters in the context of Reed-Solomon channel coding [210], [211], see Section V-C, and in the context of spatial coding [230], [231], see Section VIII-C. Future research should develop and evaluate a full set of coding techniques for a wide range of multiuser DBMC nanonetworks, which have significant potential for addressing critical communication problems [319], [320], [321].

2) EVALUATION FRAMEWORKS, TESTBEDS, AND STANDARDS

The existing studies on coding in DBMC nanonetworks have typically employed simulation models or experimental setups that were developed by the respective research groups or one of the few simulation models described in the literature, e.g., [322]. Also, typically, only the proposed novel coding technique has been evaluated in comparison against a few benchmark coding approaches; there is no standardized comprehensive comparison framework. In order to conduct fair rigorous comprehensive evaluations of a proposed novel coding technique, each proposed novel coding technique should ideally be compared against a broad set of existing benchmark coding techniques in a well-defined evaluation framework. The evaluation framework should specify the evaluation setting and the evaluation metrics.

However, a generally accepted evaluation framework is currently missing for coding in DBMC nanonetworks. In order to enable fair comprehensive comparison studies, future research should define representative evaluation scenarios. The evaluation scenarios should reflect the characteristics of common practical use cases [259], such as Internet of Bio-Nano Things use cases [323], [324], [325], [326], [327], [328], [329], [330], e.g., for intra-body communication or communication with our animate environment. The evaluation framework should specify the fine-grained details of the application use cases and communication scenarios that provide the context for the evaluations.

Moreover, the evaluation framework should specify a comprehensive set of coding evaluation metrics. The existing studies have mainly focused on the BER, which relates to the reliability of the communication. However, there are a range of other metrics that are important for practical use cases. The transmission bit rate, and the related throughput of a DBMC nanonetwork are important, as is the latency, i.e., delay, of the transmitted information, and the related delay jitter (delay variation). Also, the computational complexity and energy usage are critical performance metrics for the implementation of coding techniques in nanocommunication networks. Furthermore, the scalability to a large number of transmitters and receivers or large transmission distances is important. For instance, in intra-body networks, large numbers of nanomachines and thus numerous transmitters and receivers may move within the human body. Coding schemes should be applicable to large numbers of nanomachines and to meshed networks in order to support such intra-body networks.

Overall, future research towards developing a comprehensive evaluation framework that encompasses both the evaluation settings and the evaluation metrics is urgently needed. The definition of the evaluation framework should be broadly supported by the nanonetwork research community. A community-supported comprehensive evaluation framework can provide a blueprint for developing open-source simulation frameworks as well as for building testbeds. The simulation frameworks and testbeds will facilitate insightful fair comprehensive evaluations of existing and newly proposed coding techniques for DBMC nanonetworks and help guide the future research to address any identified weaknesses

and limitations of the coding techniques to advance the field of coding for DBMC nanonetworks.

Generally, in communication systems, standards play an important role by ensuring interoperability of the different distributed components as well as different communications protocols and techniques. In order to ensure the correct communication (interoperation) between the different encoders and decoders, which could be manufactured by different companies, coding standards for DBMC systems should be developed. Suppose that manufacturer A and manufacturer B independently develop nanomachines for nanocommunication networks. Without standards, the encoders and decoders may not correctly follow a particular coding technique. With uniform standards, e.g., prescribed standards for particular channel coding approaches and modulation techniques, the source and the receiver can correctly communicate with each other. The standards could be based on the application areas of DBMC systems. For instance, for medical applications that require error-free and reliable communication, channel codes with high error-correction capabilities, e.g., TCH codes [137], can be standardized. Overall, the development of coding standards for DBMC can help design universal codes for MC systems and promote further research and development.

Testbeds in outer space in near-zero gravity can create direction-independent experimental conditions. Therefore, satellites and space stations could be equipped with MC systems to study communication strategies in space. In current terrestrial laboratory testbeds, gravity may hinder the molecules from spreading freely in all spatial directions during diffusion in the free, air-filled space. In contrast, most current simulation models of DBMC nanonetworks assume that the molecules spread out (diffuse) uniformly without the influence of gravity in three-dimensional space. Testbeds in space are very expensive, but could still be suitable for a few experiments. Spatial codes could be particularly suitable for the use in space due to their spatial structure. A few experiments could be conducted in space to verify the accuracy of the existing simulation models that ignore gravity to test the applicability of the existing gravity-free simulation models for in-space MC communications. Also, future research related to space may reveal whether MC systems are applicable for communicating in space or on other planets in the future. Concomitantly, it would be prudent to refine the existing simulation models to incorporate gravity and to verify the refined gravity-considering simulation models with experiments on earth.

3) EMERGING PARADIGMS: BEYOND SHANNON COMMUNICATION + CONFLUENCE OF COMMUNICATION AND COMPUTING

Encoding techniques in MC provide an alternative approach (next to EM communications) for modeling the principles of conventional Shannon information transmission via channels, as well as emerging semantic-information based transmission principles at the semantic level [331], [332]. Also, future research should explore how the synthetic fabrication of cells enables new possibilities for nanoscale communication and manipulation of information [332]. The broadening of the communications paradigms from the conventional Shannon information transmission to include the emerging semantic and goal-oriented communication paradigms opens up unprecedented opportunities for developing novel communication systems. Future research needs to investigate the synergies between MC and synthetic biology on one hand, and the emerging semantic communication paradigms on the other hand [332].

One example of semantic goal-oriented communication is identification via channels [333], [334], [335], [336], [337], [338], which has the potential for exponentially scaling performance improvements compared to conventional Shannon information transmission for communication tasks related to identifying messages or system states. DBMC presents interesting avenues for future research on identification via channels due to the fine-grained discriminatory power of biological systems [339]. Future research needs to develop coding techniques that efficiently and reliably support identification via MC channels in nanonetworks. More broadly, future research needs to develop and evaluate coding techniques for the full gamut of emerging Beyond Shannon communication paradigms for DBMC nanonetworks.

Advanced communication network systems increasingly incorporate communication, computation, and data storage in a unified manner [340], [341], [342], [343], [344]. Future research should explore molecular approaches for combining computing and communication. For instance, chemical reactions of molecules could implement computing functions, while the diffusion of the products of the chemical reaction carries out the communication task. Novel coding approaches could, for instance, be based on network coding techniques by expanding the recoding to incorporate elementary computing functions. More specifically, biological molecular or synthetic machines at the microscale—and in the future at the nanoscale—can be viewed as first applications for molecular approaches in communication, storage, as well as computing [345], [346], [347], [348]. For instance, emerging synthetic receptor designs can perform the computations for detecting received MC signals through chemical reactions [349]; such chemical reaction based computations in support of MC functions should be further explored in future research. Generally, molecular or synthetic machines can be employed to serve as interfaces between molecular nanoscale systems and conventional macroscale machines. Nanomotors, as well as nanomachines, can thereby be implemented through biological processes [348]. Broadly, future research should investigate communication between nanomachines, i.e., molecular or synthetic motors and machines at the molecular nanoscale, and conventional macroscale systems.

In a related development, the interfaces in conventional network nodes are increasingly taking over more computing functions, e.g., at the interfaces between the nodes and

the network in the form of smart network interface cards (NICs) [350], [351], [352], [353]. An interesting future research direction is to expand the functionalities of existing cyber-bio interfaces that interconnect conventional networks with DBMC nanonetworks, e.g., [191], [354], [355], [356], and [357], to incorporate computing functions. Throughout, for computing and decision-making in DBMC systems, the benefits and costs of utilizing emerging reinforcement learning approaches, such as [358] and [359], should be investigated.

Future research should explore the interactions of MC and quantum communications [360], [361], [362], [363], [364], [365]. Fundamental concepts of quantum mechanics could be invoked using molecular nanomagnets as test objects (through the physical properties that govern the magnetic behaviors of molecular nanoobjects) [366], [367], [368], [369]. Future research is needed to broadly explore the potential for technological applications of quantum mechanics based DBMC nanonetworks [366].

XII. CONCLUSION

This article presents a comprehensive in-depth up-to-date survey of the existing approaches for the three coding types, namely source, channel, and network coding, in diffusion-based molecular communication (DBMC) nanonetworks. The fundamentals of each coding approach are explained and related to communication systems based on molecular diffusion. For each coding type, we provide an overview of the historical development of the respective coding techniques in the field of DBMC nanonetworks from the very beginning to the latest studies by thoroughly surveying the existing literature. Furthermore, a comprehensive set of future research directions for the still nascent area of coding for DBMC nanonetworks is presented; specifically, research imperatives for each of the three main coding types, i.e., for source, channel, and network coding, as well as for overarching goals for future research are outlined.

REFERENCES

[1] R. P. Feynman, "There's plenty of room at the bottom," *Eng. Sci. Mag.*, vol. 13, no. 5, pp. 22–36, 1960.

[2] R. P. Feynman, "There's plenty of room at the bottom [data storage]," *J. Microelectromech. Syst.*, vol. 1, no. 1, pp. 60–66, Mar. 1992.

[3] *IEEE Recommended Practice for Nanoscale and Molecular Communication Framework*," IEEE Standard 1906.1-2015, IEEE 1906.1 Working Group, Jan. 2016, pp. 1–64.

[4] *Nanotechnologies-Vocabulary—Part 2: Nano-Objects*, Standard ISO/TS 80004-2:2015, 2015.

[5] M. Imran, R. Ghannam, and Q. Abbasi, *Engineering and Technology for Healthcare*. New York, NY, USA: IEEE Press, Feb. 2021.

[6] S. Jimenez-Falcao, D. Torres, P. Martínez-Ruiz, D. Vilela, R. Martínez-Máñez, and R. Villalonga, "Sucrose-responsive intercommunicated Janus nanoparticles network," *Nanomaterials*, vol. 11, no. 10, Sep. 2021, Art. no. 2492.

[7] M. Pierobon and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 602–611, May 2010.

[8] I. F. Akyildiz, F. Brunetti, and C. Blázquez, "Nanonetworks: A new communication paradigm," *Comput. Netw.*, vol. 52, no. 12, pp. 2260–2279, Aug. 2008.

[9] B. Atakan and S. Galmés, "Effects of framing errors on the performance of molecular communications with memory," *IEEE Access*, vol. 8, pp. 19970–19981, 2020.

[10] B. A. Bilgin, E. Dinc, and O. B. Akan, "DNA-based molecular communications," *IEEE Access*, vol. 6, pp. 73119–73129, 2018.

[11] F. Dressler and O. B. Akan, "A survey on bio-inspired networking," *Comput. Netw.*, vol. 54, no. 6, pp. 881–900, Apr. 2010.

[12] L. Lin, W. Li, R. Zheng, F. Liu, and H. Yan, "Diffusion-based reference broadcast synchronization for molecular communication in nanonetworks," *IEEE Access*, vol. 7, pp. 95527–95535, 2019.

[13] S. Liu, Z. Wei, X. Wang, and C. Zhao, "Channel capacity analysis of a comprehensive absorbing receiver for molecular communication via diffusion," *IEEE Access*, vol. 8, pp. 227152–227160, 2020.

[14] L. Mucchi, A. Martinelli, S. Jayousi, S. Caputo, and M. Pierobon, "Secrecy capacity and secure distance for diffusion-based molecular communication systems," *IEEE Access*, vol. 7, pp. 110687–110697, 2019.

[15] S. P. Singh, S. Yadav, R. K. Singh, V. Kansal, and G. Singh, "Secrecy capacity of diffusive molecular communication under different deployments," *IEEE Access*, vol. 10, pp. 21670–21683, 2022.

[16] E. Shitiri, H. B. Yilmaz, and H.-S. Cho, "A time-slotted molecular communication (TS-MOC): Framework and time-slot errors," *IEEE Access*, vol. 7, pp. 78146–78158, 2019.

[17] D. P. Trinh, Y. Jeong, and H. Shin, "Connectivity in molecular communication with random time constraints," *IEEE Access*, vol. 7, pp. 113121–113130, 2019.

[18] H.-H. Yen, X. Wang, D. Wang, and H.-T. Liaw, "A node activation-based routing scheme in Micro/Nanobots networks," *IEEE Access*, vol. 7, pp. 144075–144089, 2019.

[19] D. Bi, A. Almpanis, A. Noel, Y. Deng, and R. Schober, "A survey of molecular communication in cell biology: Establishing a new hierarchy for interdisciplinary applications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1494–1545, 3rd Quart., 2021.

[20] M. Egan, B. C. Akdeniz, and B. Q. Tang, "Stochastic reaction and diffusion systems in molecular communications: Recent results and open problems," *Digital Signal Processing*, vol. 124, pp. 103–117, May 2022.

[21] 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.

[22] L. Felicetti, M. Femminella, and G. Reali, "Directional receivers for diffusion-based molecular communications," *IEEE Access*, vol. 7, pp. 5769–5783, 2019.

[23] P. He, B. Pi, and Q. Liu, "Calcium signaling in mobile molecular communication networks: From a multimedia view," *IEEE Access*, vol. 7, pp. 164825–164834, 2019.

[24] N.-R. Kim, N. Farsad, C. Lee, A. W. Eckford, and C.-B. Chae, "An experimentally validated channel model for molecular communication systems," *IEEE Access*, vol. 7, pp. 81849–81858, 2019.

[25] A. M. Koshy, I. Jeerapan, and C. Manjunatha, "New insights on molecular communication in nano communication networks and their applications," *ECS Trans.*, vol. 107, no. 1, pp. 92–95, 2022.

[26] P. Lu, M. Veletic, J. Bergsland, and I. Balasingham, "Molecular communication aspects of potassium intracellular signaling in cardiomyocytes," *IEEE Access*, vol. 8, pp. 201770–201780, 2020.

[27] Y. Lu, R. Ni, and Q. Zhu, "Wireless communication in nanonetworks: Current status, prospect and challenges," *IEEE Trans. Mol., Biol. Multi-Scale Commun.*, vol. 6, no. 2, pp. 71–80, Nov. 2020.

[28] P. Manocha, G. Chandwani, and S. Das, "Characterization of dielectrophoresis based relay-assisted molecular communication using analogue transmission line," *IEEE Access*, vol. 8, pp. 33352–33359, 2020.

[29] A. Mukherjee, S. Das, and S. Chatterjee, "An architecture of calcium signaling for molecular communication based nano network," in *Modeling, Methodologies and Tools for Molecular and Nano-Scale Communications*. Cham, Switzerland: Springer, 2017, pp. 165–203.

[30] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[31] 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.

[32] T. Nakano, T. Suda, Y. Okaie, M. J. Moore, and A. V. Vasilakos, "Molecular communication among biological nanomachines: A layered architecture and research issues," *IEEE Trans. Nanobiosci.*, vol. 13, no. 3, pp. 169–197, Mar. 2014.

[33] T. Nakano, Y. Okaie, S. Kobayashi, T. Hara, Y. Hiraoka, and T. Haraguchi, "Methods and applications of mobile molecular communication," *Proc. IEEE*, vol. 107, no. 7, pp. 1442–1456, Jun. 2019.

[34] N. Rikhtegar and M. Keshtgary, "A brief survey on molecular and electromagnetic communications in nano-networks," *Int. J. Comput. Appl.*, vol. 79, no. 3, pp. 16–28, Oct. 2013.

[35] X. Qian, M. Di Renzo, and A. Eckford, "Molecular communications: Model-based and data-driven receiver design and optimization," *IEEE Access*, vol. 7, pp. 53555–53565, 2019.

[36] J. Wang, B. Yin, and M. Peng, "Diffusion based molecular communication: Principle, key technologies, and challenges," *China Commun.*, vol. 14, no. 2, pp. 1–18, Feb. 2017.

[37] L. P. Giné and I. F. Akyildiz, "Molecular communication options for long range nanonetworks," *Comput. Netw.*, vol. 53, no. 16, pp. 2753–2766, Nov. 2009.

[38] B. D. Unluturk and I. F. Akyildiz, "An end-to-end model of plant pheromone channel for long range molecular communication," *IEEE Trans. Nanobiosci.*, vol. 16, no. 1, pp. 11–20, Jan. 2016.

[39] Y. Chahibi, M. Pierobon, S. O. Song, and I. F. Akyildiz, "A molecular communication system model for particulate drug delivery systems," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3468–3483, Dec. 2013.

[40] Y. Chahibi, "Molecular communication for drug delivery systems: A survey," *Nano Commun. Netw.*, vol. 11, pp. 90–102, Mar. 2017.

[41] H. Chen, W. Guo, J. Shen, L. Wang, and J. Song, "Structural principles analysis of host-pathogen protein-protein interactions: A structural bioinformatics survey," *IEEE Access*, vol. 6, pp. 11760–11771, 2018.

[42] U. A. K. Chude-Okonkwo, R. Malekian, B. T. Maharaj, and A. V. Vasilakos, "Molecular communication and nanonetwork for targeted drug delivery: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 3046–3096, 4th Quart., 2017.

[43] L. Felicetti, M. Femminella, G. Reali, and P. Liò, "Applications of molecular communications to medicine: A survey," *Nano Commun. Netw.*, vol. 7, pp. 27–45, Mar. 2016.

[44] V. Loscrí, A. M. Vegni, and G. Fortino, "On the interaction between a nanoparticulate system and the human body in body area nanonetworks," *Micromachines*, vol. 6, no. 9, pp. 1213–1235, Aug. 2015.

[45] J. L. Marzo, J. M. Jornet, and M. Pierobon, "Nanonetworks in biomedical applications," *Current Drug Targets*, vol. 20, no. 8, pp. 800–807, May 2019.

[46] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim, "The future of healthcare Internet of Things: A survey of emerging technologies," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1121–1167, 2nd Quart., 2020.

[47] S. Qiu, W. Guo, S. Wang, N. Farsad, and A. Eckford, "A molecular communication link for monitoring in confined environments," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, Jun. 2014, pp. 718–723.

[48] M. Pierobon and I. F. Akyildiz, "Capacity of a diffusion-based molecular communication system with channel memory and molecular noise," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 942–954, Feb. 2013.

[49] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and B. Özerman, "Energy model for communication via diffusion in nanonetworks," *Nano Commun. Netw.*, vol. 1, no. 2, pp. 86–95, Jun. 2010.

[50] C. Rose, I. S. Mian, and M. Ozmen, "Capacity bounds on point-to-point communication using molecules," *Proc. IEEE*, vol. 107, no. 7, pp. 1342–1355, Jul. 2019.

[51] 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.

[52] M. H. Kabir, S. M. R. Islam, A. P. Shrestha, F. Ali, M. A. Badsha, M. J. Piran, and D.-T. Do, "Electromagnetic nanocommunication networks: Principles, applications, and challenges," *IEEE Access*, vol. 9, pp. 166147–166165, 2021.

[53] F. Lemic, S. Abadal, W. Tavernier, P. Stroobant, D. Colle, E. Alarcón, J. Marquez-Barja, and J. Famaey, "Survey on terahertz nanocommunication and networking: A top-down perspective," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1506–1543, Apr. 2021.

[54] S. Ghafoor, N. Boujnah, M. H. Rehmani, and A. Davy, "MAC protocols for terahertz communication: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2236–2282, 4th Quart., 2020.

[55] X.-W. Yao, Y.-C.-G. Wu, and W. Huang, "Routing techniques in wireless nanonetworks: A survey," *Nano Commun. Netw.*, vol. 21, Sep. 2019, Art. no. 100250.

[56] A. Al-Helali, B. Liang, and N. Nasser, "Novel molecular signaling method and system for molecular communication in human body," *IEEE Access*, vol. 8, pp. 119361–119375, 2020.

[57] P. Kulakowski, K. Turbic, and L. M. Correia, "From nano-communications to body area networks: A perspective on truly personal communications," *IEEE Access*, vol. 8, pp. 159839–159853, 2020.

[58] W. Pan, X. Chen, X. Yang, N. Zhao, L. Meng, and F. H. Shah, "A molecular communication platform based on body area nanonetwork," *Nanomaterials*, vol. 12, no. 4, 2022, Art. no. 722.

[59] K. Yang, D. Bi, Y. Deng, R. Zhang, M. M. U. Rahman, N. A. Ali, M. A. Imran, J. M. Jornet, Q. H. Abbasi, and A. Alomainy, "A comprehensive survey on hybrid communication in context of molecular communication and terahertz communication for body-centric nanonetworks," *IEEE Trans. Mol., Biol. Multi-Scale Commun.* vol. 6, no. 2, pp. 107–133, Aug. 2020.

[60] L. F. Borges, M. T. Barros, and M. Nogueira, "Toward reliable intra-body molecular communication: An error control perspective," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 114–120, May 2021.

[61] M. C. Gursoy, M. Nasiri-Kenari, and U. Mitra, "Towards high data-rate diffusive molecular communications: A review on performance enhancement strategies," *Digit. Signal Process.*, vol. 124, May 2022, Art. no. 103161.

[62] M. Damrath, "Channel coding in molecular communication," Ph.D. dissertation, Fac. Eng., Dept. Elect. Inf. Eng., Chair Inf. Coding Theory, Christian-Albrechts-Universität zu Kiel, Kiel, Germany, 2020. [Online]. Available: https://macau.uni-kiel.de/receive/macau_mods_00000779

[63] 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, Jul. 2019.

[64] M. Kuscu and O. B. Akan, "On the physical design of molecular communication receiver based on nanoscale biosensors," *IEEE Sensors J.*, vol. 16, no. 8, pp. 2228–2243, Sep. 2016.

[65] A. M. Darya, H. Vakani, and Q. Nasir, "Error control codes for molecular communication channels: A survey," in *Proc. Int. Conf. Commun., Signal Process., Appl. (ICCSPA)*, Mar. 2019, pp. 1–4.

[66] C. A. Söldner, E. Socher, V. Jamali, W. Wicke, A. Ahmadzadeh, H.-G. Breitinger, A. Burkovski, K. Castiglione, R. Schober, and H. Sticht, "A survey of biological building blocks for synthetic molecular communication systems," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2765–2800, 4th Quart., 2020.

[67] 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, Jun. 2019.

[68] 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.

[69] X. Huang, Y. Fang, and N. Yang, "A survey on estimation schemes in molecular communications," *Digit. Signal Process.*, vol. 124, pp. 103–163, May 2022.

[70] C. McBride, R. Shah, and D. D. Vecchio, "The effect of loads in molecular communications," *Proc. IEEE*, vol. 107, no. 7, pp. 1369–1386, May 2019.

[71] F. Ratti, M. Magarini, and D. Del Vecchio, "What is the trait d'Union between retroactivity and molecular communication performance limits?" *Molecules*, vol. 27, no. 10, May 2022, Art. no. 3130.

[72] J. M. Jornet, Y. Bae, C. R. Handelmann, B. Decker, A. Balcerak, A. Sangwan, P. Miao, A. Desai, L. Feng, E. K. Stachowiak, and M. K. Stachowiak, "Optogenomic interfaces: Bridging biological networks with the electronic digital world," *Proc. IEEE*, vol. 107, no. 7, pp. 1387–1401, Jul. 2019.

[73] M. Veletić and I. Balasingham, "Synaptic communication engineering for future cognitive Brain–Machine interfaces," *Proc. IEEE*, vol. 107, no. 7, pp. 1425–1441, Jul. 2019.

[74] K. F. Hassan and S. B. Sadkhan, "Security of nanoscale communication networks–status, challenges and future trends," in *Proc. Muthanna Int. Conf. Eng. Sci. Technol. (MICEST)*, Mar. 2022, pp. 224–229.

[75] V. Loscri and A. M. Vegni, "Enabling molecular communication through chirality of enantiomers," *ITU J. Future Evolving Technol.*, vol. 2, no. 3, pp. 1–8, Jun. 2021.

[76] *Diffusion*. Accessed: Feb. 21, 2022. [Online]. Available: https://www.chemie.de/lexikon/Diffusion.html

[77] K. Darchini and A. S. Alfa, "Molecular communication via microtubules and physical contact in nanonetworks: A survey," *Nano Commun. Netw.*, vol. 4, no. 2, pp. 73–85, Jun. 2013.

[78] S. Bhattacharjee, M. Damrath, L. Stratmann, P. A. Hoeher, and F. Dressler, "Digital communication techniques in macroscopic air-based molecular communication," *IEEE Trans. Mol., Biol. Multi-Scale Commun.*, vol. 8, no. 4, pp. 276–291, Dec. 2022.

[79] Y. Chen, Y. Li, L. Lin, and H. Yan, "Parameter estimation of diffusive molecular communication with drift," *IEEE Access*, vol. 8, pp. 142704–142713, 2020.

[80] B.-H. Koo, C. Lee, A. E. Pusane, T. Tugcu, and C.-B. Chae, "MIMO operations in molecular communications: Theory, prototypes, and open challenges," *IEEE Commun. Mag.*, vol. 59, no. 9, pp. 98–104, Sep. 2021.

[81] T. Tugcu. (Nov. 2013). *CmpE59G (Fall2013): NanoNetworking (Lecture 4)*. Accessed: Feb. 21, 2022. [Online]. Available: https://www.youtube.com/watch?v=goahWa0FVoE

[82] H. B. Yilmaz, C.-B. Chae, B. Tepekule, and A. E. Pusane, "Arrival modeling and error analysis for molecular communication via diffusion with drift," in *Proc. 2nd Annu. Int. Conf. Nanosc. Comput. Commun.*, Sep. 2015, pp. 1–6.

[83] H. B. Yilmaz and C.-B. Chae, "Arrival modelling for molecular communication via diffusion," *Electron. Lett.*, vol. 50, no. 23, pp. 1667–1669, Nov. 2014.

[84] I. F. Akyildiz, M. Pierobon, and S. Balasubramaniam, "An information theoretic framework to analyze molecular communication systems based on statistical mechanics," *Proc. IEEE*, vol. 107, no. 7, pp. 1230–1255, Jul. 2019.

[85] Georg-August-Universität Göttingen. *LP—Lösung der Diffusionsgleichung*. Accessed: Feb. 22, 2022. [Online]. Available: https://lp.uni-goettingen.de/get/text/7026

[86] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Oct. 1948.

[87] J. Kieffer, "History of source coding," *IEEE Inf. Theory Soc. Newslett.*, vol. 43, no. 2, pp. 13–15, Jun. 1993.

[88] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proc. Inst. Radio Eng.*, vol. 40, no. 9, pp. 1098–1101, Sep. 1952.

[89] R. Gallager, "Variations on a theme by Huffman," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 6, pp. 668–674, Nov. 1978.

[90] B. Mcmillan, "The basic theorems of information theory," *Ann. Math. Statist.*, vol. 24, no. 2, pp. 196–219, Jun. 1953.

[91] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 129–137, Mar. 1982.

[92] J. C. Kieffer, "A survey of the theory of source coding with a fidelity criterion," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1473–1490, Sep. 1993.

[93] C. Shannon, "Coding theorems for a discrete source with a fidelity criteria," *IRE Nat. Conv. Rec.*, vol. 7, pp. 325–350, Jan. 1959.

[94] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.

[95] L. D. Davisson, "Universal source coding," in *Advances in Source Coding*. Vienna, Austria: Springer, 1975, pp. 41–70.

[96] J. Rissanen, "Generalized Kraft inequality and arithmetic coding," *IBM J. Res. Develop.*, vol. 20, no. 3, pp. 198–203, May 1976.

[97] R. C. Pasco, "Source coding algorithms for fast data compression," Ph.D. dissertation, Electrical Eng., Stanford Univ., Stanford, CA, USA, 1976.

[98] B. Dhayabaran, G. T. Raja, and M. Magarini, "Modified inverse source coding for diffusion based molecular communication system," in *Proc. IEEE 21st Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, May 2020, pp. 1–5.

[99] B. Dhayabaran, G. T. Raja, and M. Magarini, "Low complex receiver design for modified inverse source coded diffusion-based molecular communication systems," *IEEE Trans. Mol., Biol. Multi-Scale Commun.*, vol. 7, no. 4, pp. 239–252, Dec. 2021.

[100] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 14th ed. New York, NY, USA: Wiley, 1991.

[101] D. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[102] B. Furht, "Huffman coding," in *Encyclopedia of Multimedia*. Boston, MA, USA: Springer, 2006, pp. 278–280.

[103] M. Esponda. (2013). *Funktionale Programmierung: Huffman-Kodierung (Lecture)*. Accessed: Jul. 6, 2022. [Online]. Available: http://www.inf.fu-berlin.de/lehre/WS12/ALP1/lectures/V16_ALPI_Huffman-Kodierung_2013.pdf

[104] A. Moffat, "Huffman coding," *ACM Comput. Surv.*, vol. 52, no. 4, p. 85, Jul. 2020.

[105] W. Ullah, "Quasi-cyclic low-density parity-check codes: Applications and iterative decoding hardware implementations," Ph.D. dissertation, Nanjing Univ. Aeronaut. Astronaut., College Electron. Inf. Eng., Nanjing, China, Mar. 2012.

[106] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Tech. J.*, vol. 29, no. 2, pp. 147–160, Apr. 1950.

[107] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 2, pp. 260–269, Apr. 1967.

[108] D. Forney, "Concatenated codes," *Scholarpedia*, vol. 4, no. 2, 2009, Art. no. 8374.

[109] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 2, Geneva, Switzerland, May 1993, pp. 1064–1070.

[110] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.

[111] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electron. Lett.*, vol. 33, no. 6, pp. 457–458, 1997.

[112] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.

[113] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2008.

[114] P. Elias, "Coding for noisy channels," *IRE Conv. Rec*, vol. 3, no. 4, pp. 37–46, 1955.

[115] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.

[116] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 1, pp. 55–67, Jan. 1982.

[117] W. Koch and A. Baier, "Optimum and sub-optimum detection of coded data disturbed by time-varying intersymbol interference (applicable to digital mobile radio receivers)," in *Proc. IEEE Global Telecommun. Conf. Exhib.*, vol. 3, Dec. 1990, pp. 1679–1684.

[118] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[119] E. R. Berlekamp, *Algebraic Coding Theory*. Singapore: World Scientific, 2015.

[120] J. Massey, "Shift-register synthesis and BCH decoding," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 1, pp. 122–127, Jan. 1969.

[121] D. Chase, "Class of algorithms for decoding block codes with channel measurement information," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 170–182, Jan. 1972.

[122] J. Wolf, "Efficient maximum likelihood decoding of linear block codes using a trellis," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 76–80, Jan. 1978.

[123] R. M. Pyndiah, "Near-optimum decoding of product codes: Block turbo codes," *IEEE Trans. Commun.*, vol. 46, no. 8, pp. 1003–1010, Aug. 1998.

[124] S. Alamouti, "A simple transmit diversity technique for wireless communication," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[125] P. Robertson and T. Wörz, "Turbo trellis-coded modulation (TTCM) employing parity bit puncturing and parallel concatenation," in *Wiley Encyclopedia of Telecommunications*. Hoboken, NJ, USA: Wiley, 2003, pp. 2737–2753.

[126] T. D. Schneider, "Theory of molecular machines. I. Channel capacity of molecular machines," *J. Theor. Biol.*, vol. 148, no. 1, pp. 83–123, Jan. 1991.

[127] M. Sartipi and F. Fekri, "Source and channel coding in wireless sensor networks using LDPC codes," in *Proc. 1st Annu. IEEE Commun. Soc. Conf. Sensor Ad Hoc Commun. Netw., IEEE SECON.*, Oct. 2004, pp. 309–316.
[128] P.-Y. Ko, Y.-C. Lee, P.-C. Yeh, C.-H. Lee, and K.-C. Chen, "A new paradigm for channel coding in diffusion-based molecular communications: Molecular coding distance function," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 3748–3753.
[129] T. Nakano and T. Suda, "NSF workshop on molecular communication/biological communications technology," Nat. Sci. Found., Arlington, VA, USA, Tech. Rep., Feb. 2008.
[130] T. D. Schneider, "A brief review of molecular information theory," *Nano Commun. Netw.*, vol. 1, no. 3, pp. 173–180, Sep. 2010.
[131] Y. Lu, M. D. Higgins, and M. S. Leeson, "Diffusion based molecular communications system enhancement using high order Hamming codes," in *Proc. 9th Int. Symp. Commun. Syst., Netw. Digit. Sign (CSNDSP)*, Jul. 2014, pp. 438–442.
[132] B. Tepekule, A. E. Pusane, H. B. Yilmaz, C.-B. Chae, and T. Tugcu, "ISI mitigation techniques in molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 1, no. 2, pp. 202–216, Jun. 2015.
[133] Y. Lu, M. D. Higgins, and M. S. Leeson, "Comparison of channel coding schemes for molecular communications systems," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3991–4001, Nov. 2015.
[134] H. Arjmandi, M. Movahednasab, A. Gohari, M. Mirmohseni, M. Nasiri-Kenari, and F. Fekri, "ISI-avoiding modulation for diffusion-based molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 3, no. 1, pp. 48–59, Mar. 2017.
[135] A. O. Kislal, H. B. Yilmaz, A. E. Pusane, and T. Tugcu, "ISI-aware channel code design for molecular communication via diffusion," *IEEE Trans. Nanobiosci.*, vol. 18, no. 2, pp. 205–213, Apr. 2019.
[136] A. O. Kislal, B. C. Akdeniz, C. Lee, A. E. Pusane, T. Tugcu, and C.-B. Chae, "ISI-mitigating channel codes for molecular communication via diffusion," *IEEE Access*, vol. 8, pp. 24588–24599, 2020.
[137] S. Figueiredo, N. Souto, and F. Cercas, "Low-complexity channel codes for reliable molecular communication via diffusion," *Sensors*, vol. 22, no. 1, p. 41, Dec. 2022.
[138] B. Sklar and F. J. Harris, "The ABCs of linear block codes," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 14–35, Jul. 2004.
[139] E. Galois and P. M. Neumann, *The Mathematical Writings of Évariste Galois*, vol. 6. Helsinki, Finland: European Mathematical Society, 2011.
[140] S. P. Singh, S. Singh, W. Guo, S. Mishra, and S. Kumar, "Radiation absorption noise for molecular information transfer," *IEEE Access*, vol. 8, pp. 6379–6387, 2020.
[141] B. Sklar, *Digital Communications: Fundamentals and Applications*. Upper Saddle River, NJ, USA: Prentice & Hall, 2001.
[142] A. K. Singh, "Error detection and correction by Hamming code," in *Proc. Int. Conf. Global Trends Signal Process., Inf. Comput. Commun. (ICGTSPICC)*, Dec. 2016, pp. 35–37.
[143] R. C. Bose and D. K. Ray-Chaudhuri, "On a class of error correcting binary group codes," *Inf. Control*, vol. 3, no. 1, pp. 68–79, Mar. 1960.
[144] R. C. Bose and D. K. Ray-Chaudhuri, "Further results on error correcting binary group codes," *Inf. Control*, vol. 3, no. 3, pp. 279–290, Sep. 1960.
[145] S. A. Vanstone and P. C. Van Oorschot, *An Introduction to Error Correcting Codes With Applications*, vol. 71. New York, NY, USA: Springer, 2013.
[146] A. Bensky, "Introduction to information theory and coding," in *Short-Range Wireless Communication*, 3rd ed., A. Bensky, Ed. Oxford, U.K.: Newnes, 2019, pp. 211–236.
[147] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
[148] R. Bassoli, H. Marques, J. Rodriguez, K. W. Shum, and R. Tafazolli, "Network coding theory: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 5, pp. 1950–1978, 4th Quart., 2013.
[149] D. E. Lucani, M. V. Pedersen, D. Ruano, C. W. Sørensen, F. H. Fitzek, J. Heide, O. Geil, V. Nguyen, and M. Reisslein, "Fulcrum: Flexible network coding for heterogeneous devices," *IEEE Access*, vol. 6, pp. 77890–77910, 2018.
[150] M. Medard and A. Sprintson, *Network Coding: Fundamentals and Applications*. Oxford, U.K.: Elsevier Science, 2012.
[151] A. A. AbuDaqa, A. Mahmoud, M. Abu-Amara, and T. Sheltami, "Survey of network coding based P2P file sharing in large scale networks," *Appl. Sci.*, vol. 10, no. 7, Mar. 2020, Art. no. 2206.
[152] F. Gabriel, S. Wunderlich, S. Pandi, F. H. Fitzek, and M. Reisslein, "Caterpillar RLNC with feedback (CRLNC-FB): Reducing delay in selective repeat ARQ through coding," *IEEE Access*, vol. 6, pp. 44787–44802, 2018.
[153] M. Hayashi, "Secure physical layer network coding versus secure network coding," *Entropy*, vol. 24, no. 1, Dec. 2021, Art. no. 47.
[154] S. Laurindo, R. Moraes, C. Montez, and F. Vasques, "Combining network coding and retransmission techniques to improve the communication reliability of wireless sensor network," *Information*, vol. 12, no. 5, Apr. 2021, Art. no. 184.
[155] G. Peralta, R. G. Cid-Fuentes, J. Bilbao, and P. M. Crespo, "Homomorphic encryption and network coding in IoT architectures: Advantages and future challenges," *Electronics*, vol. 8, no. 8, Jul. 2019, Art. no. 827.
[156] H. Shin and J.-S. Park, "Optimizing random network coding for multimedia content distribution over smartphones," *Multimedia Tools Appl.*, vol. 76, pp. 19379–19395, Oct. 2017.
[157] D. Vukobratovic, A. Tassi, S. Delic, and C. Khirallah, "Random linear network coding for 5G mobile video delivery," *Information*, vol. 9, no. 4, p. 72, Mar. 2018.
[158] S. Wunderlich, F. Gabriel, S. Pandi, F. H. Fitzek, and M. Reisslein, "Caterpillar RLNC (CRLNC): A practical finite sliding window RLNC approach," *IEEE Access*, vol. 5, pp. 20183–20197, 2017.
[159] S. Wunderlich, J. A. Cabrera, F. H. P. Fitzek, and M. Reisslein, "Network coding in heterogeneous multicore IoT nodes with DAG scheduling of parallel matrix block operations," *IEEE Internet Things J.*, vol. 4, no. 4, pp. 917–933, Aug. 2017.
[160] L. Lin, C. Yang, and M. Ma, "Maximum-likelihood estimator of clock offset between nanomachines in bionanosensor networks," *Sensors*, vol. 15, no. 12, pp. 30827–30838, Dec. 2015.
[161] P. Gong, T. M. Chen, and P. Xu, "Resource-conserving protection against energy draining (RCPED) routing protocol for wireless sensor networks," *Network*, vol. 2, no. 1, pp. 83–105, Feb. 2022.
[162] Y.-M. Huang, M.-Y. Hsieh, and F. E. Sandnes, "Wireless sensor networks and applications," in *Sensors: Advancements in Modeling, Design Issues, Fabrication and Practical Applications*, S. Mukhopadhyay and R. Huang, Eds. Berlin, Germany: Springer, 2008, pp. 199–219.
[163] A. Toledo and X. Wang, "Efficient multipath in sensor networks using diffusion and network coding," in *Proc. 40th Annu. Conf. Inf. Sci. Syst.*, Mar. 2006, pp. 87–92.
[164] I.-H. Hou, Y.-E. Tsai, T. F. Abdelzaher, and I. Gupta, "AdapCode: Adaptive network coding for code updates in wireless sensor networks," in *Proc. IEEE INFOCOM 27th Conf. Comput. Commun.*, Apr. 2008, pp. 1517–1525.
[165] Z. Guo, B. Wang, P. Xie, W. Zeng, and J.-H. Cui, "Efficient error recovery with network coding in underwater sensor networks," *Ad Hoc Netw.*, vol. 7, no. 4, pp. 791–802, Jun. 2009.
[166] Z. Guo, P. Xie, J.-H. Cui, and B. Wang, "On applying network coding to underwater sensor networks," in *Proc. 1st ACM Int. Workshop Underwater Netw. WUWNet*, 2006, pp. 109–112.
[167] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C.-B. Chae, "Three-dimensional channel characteristics for molecular communications with an absorbing receiver," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 929–932, Jun. 2014.
[168] M. S. Kuran, H. B. Yilmaz, and T. Tugcu, "A tunnel-based approach for signal shaping in molecular communication," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, Jun. 2013, pp. 776–781.
[169] C. Fragouli and E. Soljanin, "Network coding applications," *Found. Trends Netw.*, vol. 2, no. 2, pp. 135–269, Jun. 2008.
[170] H.-M. Wang, X.-G. Xia, and Q. Yin, "A linear analog network coding for asynchronous two-way relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3630–3637, Dec. 2010.
[171] W. Huang and L. Zhang, "An improved protocol based on directed diffusion routing protocol with network coding," in *Proc. 8th Int. Conf. Netw., Commun. Comput.*, Dec. 2019, pp. 159–162.
[172] P. A. Chou and Y. Wu, "Network coding for the internet and wireless networks," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 77–85, Sep. 2007.
[173] S. Deb, M. Effros, T. Ho, D. Karger, R. Kötter, D. Lun, M. Médard, and N. Ratnakar, (Jan. 2005). *Network Coding for Wireless Applications: A Brief Tutorial*. Accessed: Oct. 4, 2022. [Online]. Available: https://www.crab.rutgers.edu/dslun/docs/iwwan2005.pdf

[174] M. Di Renzo, M. Iezzi, and F. Graziosi, "Beyond routing via network coding: An overview of fundamental information-theoretic results," in *Proc. 21st Annu. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2010, pp. 2745–2750.

[175] C. Fragouli, J. Y. Le Boudec, and J. Widmer, "Network coding: An instant primer," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 1, pp. 63–68, 2006.

[176] T. Ho and D. S. Lun, *Network Coding: An Introduction*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[177] M. Sanna and E. Izquierdo, "A survey of linear network coding and network error correction code constructions and algorithms," *Int. J. Digit. Multimedia Broadcast.*, vol. 2011, May 2011, Art. no. 857847.

[178] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, "Network coding theory," in *Foundations and Trends in Communications and Information Theory*. vol. 2, nos. 4–5. Hanover, MA, USA: Now, 2005, pp. 241–381.

[179] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A survey on network codes for distributed storage," *Proc. IEEE*, vol. 99, no. 3, pp. 476–489, Mar. 2011.

[180] T. Matsuda, T. Noguchi, and T. Takine, "Survey of network coding and its applications," *IEICE Trans. Commun.*, vol. 94, no. 3, pp. 698–717, Mar. 2011.

[181] W. Guo, C. Mias, N. Farsad, and J.-L. Wu, "Molecular versus electromagnetic wave propagation loss in macro-scale environments," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 1, no. 1, pp. 18–25, Mar. 2015.

[182] D. C. Cox and R. Leck, "Distributions of multipath delay spread and average excess delay for 910-MHz urban mobile radio paths," *IEEE Trans. Antennas Propag.*, vol. AP-23, no. 2, pp. 206–213, Mar. 1975.

[183] S. M. Mustam, S. K. Syed-Yusof, and S. Zubair, "Capacity and delay spread in multilayer diffusion-based molecular communication (DBMC) channel," *IEEE Trans. Nanobiosci.*, vol. 15, no. 7, pp. 599–612, Oct. 2016.

[184] H. ShahMohammadian, G. G. Messier, and S. Magierowski, "Nano-machine molecular communication over a moving propagation medium," *Nano Commun. Netw.*, vol. 4, no. 3, pp. 142–153, Sep. 2013.

[185] H. Bai and M. Atiquzzaman, "Error modeling schemes for fading channels in wireless communications: A survey," *IEEE Commun. Surveys Tuts.*, vol. 5, no. 2, pp. 2–9, 4th Quart., 2003.

[186] A. Al-Smadi and M. M. Smadi, "Study of the reliability of a binary symmetric channel under non-Gaussian disturbances," *Int. J. Commun. Syst.*, vol. 16, no. 10, pp. 893–908, 2003.

[187] W. K. Harrison, T. Fernandes, M. A. C. Gomes, and J. P. Vilela, "Generating a binary symmetric channel for wiretap codes," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 8, pp. 2128–2138, Aug. 2019.

[188] H. S. Wang and N. Moayeri, "Finite-state Markov channel—A useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163–171, Feb. 1995.

[189] D. Arifler, "Capacity analysis of a diffusion-based short-range molecular nano-communication channel," *Comput. Netw.*, vol. 55, no. 6, pp. 1426–1434, 2011.

[190] A. Einolghozati and F. Fekri, "Analysis of error-detection schemes in diffusion-based molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 615–624, Mar. 2016.

[191] D. T. Mcguiness, S. Giannoukos, A. Marshall, and S. Taylor, "Modulation analysis in macro-molecular communications," *IEEE Access*, vol. 7, pp. 11049–11065, 2019.

[192] A. Einolghozati, M. Sardari, and F. Fekri, "Design and analysis of wireless communication systems using diffusion-based molecular communication among bacteria," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6096–6105, Dec. 2013.

[193] P. Chen, Z. Xie, Y. Fang, Z. Chen, S. Mumtaz, and J. Rodrigues, "Physical-layer network coding: An efficient technique for wireless communications," *IEEE Netw.*, vol. 34, no. 2, pp. 270–276, Mar./Apr. 2019.

[194] S. C. Liew, S. Zhang, and L. Lu, "Physical-layer network coding: Tutorial, survey, and beyond," *Phys. Commun.*, vol. 6, pp. 4–42, Mar. 2013.

[195] J. K. Kwon, "Inverse source coding for dimming in visible light communications using NRZ-OOK on reliable links," *IEEE Photon. Technol. Lett.*, vol. 22, no. 19, pp. 1455–1457, Oct. 1, 2010.

[196] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. Boston, MA, USA: McGraw-Hill, 2008.

[197] 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.

[198] B. Sklar, "How I learned to love the trellis," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 87–102, May 2003.

[199] N. Farsad, W. Guo, and A. W. Eckford, "Tabletop molecular communication: Text messages through chemical signals," *PLoS ONE*, vol. 8, no. 12, Dec. 2013, Art. no. e82935.

[200] P. Hofmann, J. T. Gómez, F. Dressler, and F. H. Fitzek, "Testbed-based receiver optimization for SISO molecular communication channels," in *Proc. Int. Balkan Conf. Commun. Netw. (BalkanCom)*, 2022, pp. 120–125.

[201] A. K. Shrivastava, D. Das, and R. Mahapatra, "Adaptive threshold detection and ISI mitigation in mobile molecular communication," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.

[202] M. S. Leeson and M. D. Higgins, "Error correction coding for molecular communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 6172–6176.

[203] M. S. Leeson and M. D. Higgins, "Forward error correction for molecular communications," *Nano Commun. Netw.*, vol. 3, no. 3, pp. 161–167, 2012.

[204] C. Bai, M. S. Leeson, and M. D. Higgins, "Minimum energy channel codes for molecular communications," *Electron. Lett.*, vol. 50, no. 23, pp. 1669–1671, 2014.

[205] A. Marcone, M. Pierobon, and M. Magarini, "A biological circuit design for modulated parity-check encoding in molecular communication," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–7.

[206] R. Rai, S. Pratap Singh, M. Lakshmanan, and V. K. Pandey, "VLSI implementation of Hamming code for molecular communication," in *Advances in Smart Communication and Imaging Systems*, vol. 721, R. Agrawal, C. Kishore Singh, and A. Goyal, Eds. Singapore: Springer, 2021, pp. 685–691.

[207] H.-Y. Cheong, "A study on the decoding of Hamming codes using soft values on the molecular communication channel," *J. Korea Inst. Inf., Electron., Commun. Technol.*, vol. 13, no. 5, pp. 338–343, 2020.

[208] R. Rai, S. Pratap Singh, M. Lakshmanan, and V. K. Pandey, "Implementation of cyclic Reed–Müller code for molecular communication," in *Advances in Smart Communication and Imaging Systems*, vol. 721, R. Agrawal, C. Kishore Singh, and A. Goyal, Eds. Singapore: Springer, 2021, pp. 637–644.

[209] M. B. Dissanayake, Y. Deng, A. Nallanathan, E. M. N. Ekanayake, and M. Elkashlan, "Reed Solomon codes for molecular communication with a full absorption receiver," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1245–1248, Jun. 2017.

[210] M. B. Dissanayake, Y. Deng, A. Nallanathan, M. Elkashlan, and U. Mitra, "Enhancing the reliability of large-scale multiuser molecular communication systems," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.

[211] M. B. Dissanayake, Y. Deng, A. Nallanathan, M. Elkashlan, and U. Mitra, "Interference mitigation in large-scale multiuser molecular communication," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4088–4103, Jun. 2019.

[212] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "Performance analysis of convolutional coding techniques in diffusion-based concentration-encoded PAM molecular communication systems," *BioNanoScience*, vol. 3, no. 3, pp. 270–284, 2013.

[213] Y. Lu, M. D. Higgins, and M. S. Leeson, "Self-orthogonal convolutional codes (SOCCs) for diffusion-based molecular communication systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 1049–1053.

[214] Y. Lu, X. Wang, M. D. Higgins, A. Noel, N. Neophytou, and M. S. Leeson, "Energy requirements of error correction codes in diffusion-based molecular communication systems," *Nano Commun. Netw.*, vol. 11, pp. 24–35, Mar. 2017.

[215] H. Li and Q. Li, "The crossover-distance for ISI-correcting decoding of convolutional codes in diffusion-based molecular communications," 2018, *arXiv:1812.11273*.

[216] Q. Li, "The asymmetric-distance metrics for decoding of convolutional codes in diffusion-based molecular communications," *IEEE Trans. Nanobiosci.*, vol. 18, no. 3, pp. 469–481, Jul. 2019.

[217] P.-J. Shih, C.-H. Lee, and P.-C. Yeh, "Channel codes for mitigating intersymbol interference in diffusion-based molecular communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 4228–4232.

[218] P.-J. Shih, C.-H. Lee, P.-C. Yeh, and K.-C. Chen, "Channel codes for reliability enhancement in molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 857–867, Dec. 2013.

[219] P.-C. Yeh, K.-C. Chen, Y.-C. Lee, L.-S. Meng, P.-J. Shih, P.-Y. Ko, W.-A. Lin, and C.-H. Lee, "A new frontier of wireless communication theory: Diffusion-based molecular communications," *IEEE Wireless Commun.*, vol. 19, no. 5, pp. 28–35, Oct. 2012.

[220] S. Qiu, T. Asyhari, and W. Guo, "Mobile molecular communications: Positional-distance codes," in *Proc. IEEE 17th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2016, pp. 1–5.

[221] P. Akhkandi, A. Keshavarz-Haddad, and A. Jamshidi, "A new channel code for decreasing inter-symbol-interference in diffusion based molecular communications," in *Proc. 8th Int. Symp. Telecommun. (IST)*, Sep. 2016, pp. 277–281.

[222] 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.

[223] A. Keshavarz-Haddad, A. Jamshidi, and P. Akhkandi, "Inter-symbol interference reduction channel codes based on time gap in diffusion-based molecular communications," *Nano Commun. Netw.*, vol. 19, pp. 148–156, Mar. 2019.

[224] A. Einolghozati and F. Fekri, "Error detection in diffusion-based molecular communication," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 1128–1133.

[225] M. Kovačević and P. Popovski, "Zero-error capacity of a class of timing channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6796–6800, Nov. 2014.

[226] N. Abadi, A. A. Gohari, M. Mirmohseni, and M. Nasiri-Kenari, "Zero-error codes for multi-type molecular communication in random delay channel," in *Proc. Iran Workshop Commun. Inf. Theory (IWCIT)*, Apr. 2018, pp. 1–6.

[227] A. Marcone, M. Pierobon, and M. Magarini, "Parity-check coding based on genetic circuits for engineered molecular communication between biological cells," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6221–6236, Dec. 2018.

[228] H. Egashira, J. Suzuki, J. S. Mitzman, T. Nakano, and H. Fukuda, "Robust directional-diffusive hybrid molecular communication with parity-check erasure coding," in *Proc. Joint 17th World Congr. Int. Fuzzy Syst. Assoc. 9th Int. Conf. Soft Comput. Intell. Syst. (IFSA-SCIS)*, Jun. 2017, pp. 1–8.

[229] H.-Y. Lin, S. M. Moser, and P.-N. Chen, "Weak flip codes and their optimality on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, pp. 5191–5218, Jul. 2018.

[230] M. Damrath, H. B. Yilmaz, C.-B. Chae, and P. A. Hoeher, "Spatial coding techniques for molecular MIMO," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Nov. 2017, pp. 324–328.

[231] M. Damrath, H. B. Yilmaz, C.-B. Chae, and P. A. Hoeher, "Array gain analysis in molecular MIMO communications," *IEEE Access*, vol. 6, pp. 61091–61102, 2018.

[232] M. Kocaoglu and O. B. Akan, "Minimum energy channel codes for nanoscale wireless communications," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1492–1500, Apr. 2013.

[233] J. L. Massey and D. J. Costello, "Nonsystematic convolutional codes for sequential decoding in space applications," *IEEE Trans. Commun. Technol.*, vol. CT-19, no. 5, pp. 806–813, Oct. 1971.

[234] S. Lin and D. J. Costello, *Error Control Coding*, 2nd ed. New York, NY, USA: Pearson, 2004.

[235] R. Chien, "Cyclic decoding procedures for Bose-Chaudhuri-Hocquenghem codes," *IEEE Trans. Inf. Theory*, vol. IT-10, no. 4, pp. 357–363, Oct. 1964.

[236] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *J. Soc. Ind. Appl. Math.*, vol. 8, no. 2, pp. 300–304, Jun. 1960.

[237] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *IRE Prof. Group Inf. Theory*, vol. 4, no. 4, pp. 38–49, Sep. 1954.

[238] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2011, pp. 1–5.

[239] M. Egan, T. Q. Duong, and M. D. Renzo, "Biological circuits for detection in MoSK-based molecular communication," *IEEE Access*, vol. 7, pp. 21094–21102, 2019.

[240] K. V. Srinivas, A. W. Eckford, and R. S. Adve, "Molecular communication in fluid media: The additive inverse Gaussian noise channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4678–4692, Jul. 2012.

[241] W. Guo, T. Asyhari, N. Farsad, H. B. Yilmaz, B. Li, A. Eckford, and C.-B. Chae, "Molecular communications: Channel model and physical layer techniques," *IEEE Wireless Commun. Mag.*, vol. 23, no. 4, pp. 120–127, Aug. 2016.

[242] H. Arjmandi, A. Gohari, M. N. Kenari, and F. Bateni, "Diffusion-based nanonetworking: A new modulation technique and performance analysis," *IEEE Commun. Lett.*, vol. 17, no. 4, pp. 645–648, Apr. 2013.

[243] A. Ahmadzadeh, A. Noel, and R. Schober, "Analysis and design of multi-hop diffusion-based molecular communication networks," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 1, no. 2, pp. 144–157, Jun. 2015.

[244] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: Past, present, and future," *Multimedia Tools Appl.*, vol. 80, no. 5, pp. 8091–8126, 2021.

[245] K. Poon, A. Conway, G. Wardrop, and J. Mellis, "Successful application of genetic algorithms to network design and planning," *BT Technol. J.*, vol. 18, no. 4, pp. 32–41, 2000.

[246] B. Shahi, S. Dahal, A. Mishra, S. B. V. Kumar, and C. P. Kumar, "A review over genetic algorithm and application of wireless network systems," *Proc. Comput. Sci.*, vol. 78, pp. 431–438, Jan. 2016.

[247] J. Shi, M. Habib, and H. Yan, "A review paper on different application of genetic algorithm for mobile ad-hoc network (MANET)," *Int. J. Online Biomed. Eng.*, vol. 16, no. 5, pp. 119–139, 2020.

[248] H.-S. Yang, M. Maier, M. Reisslein, and W. M. Carlyle, "A genetic algorithm-based methodology for optimizing multiservice convergence in a metro WDM network," *J. Lightw. Technol.*, vol. 21, no. 5, pp. 1114–1133, May 1, 2003.

[249] B. Tepekule, A. E. Pusane, H. B. Yilmaz, and T. Tugcu, "Energy efficient ISI mitigation for communication via diffusion," in *Proc. IEEE Int. Black Sea Conf. Commun. Netw. (BlackSeaCom)*, May 2014, pp. 33–37.

[250] A. Akkaya, H. B. Yilmaz, C.-B. Chae, and T. Tugcu, "Effect of receptor density and size on signal reception in molecular communication via diffusion with an absorbing receiver," *IEEE Commun. Lett.*, vol. 19, no. 2, pp. 155–158, Feb. 2014.

[251] Y. Deng, A. Noel, W. Guo, A. Nallanathan, and M. Elkashlan, "Analyzing large-scale multiuser molecular communication via 3-D stochastic geometry," *IEEE Trans. Mol., Biol. Multi-Scale Commun.*, vol. 3, no. 2, pp. 118–133, Jun. 2017.

[252] H. B. Yilmaz, G.-Y. Suk, and C.-B. Chae, "Chemical propagation pattern for molecular communications," *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 226–229, Apr. 2017.

[253] Y. Kou, S. Lin, and M. P. C. Fossorier, "Low-density parity-check codes based on finite geometries: A rediscovery and new results," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2711–2736, Nov. 2001.

[254] W. Ryan and S. Lin, *Channel Codes: Classical and Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[255] C. Lee, B. Koo, N.-R. Kim, B. Yilmaz, N. Farsad, A. Eckford, and C.-B. Chae, "Molecular MIMO communication link," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2015, pp. 13–14.

[256] B.-H. Koo, C. Lee, H. B. Yilmaz, N. Farsad, A. Eckford, and C.-B. Chae, "Molecular MIMO: From theory to prototype," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 600–614, Mar. 2016.

[257] S. G. Wilson, M. Brandt-Pearce, Q. Cao, and M. Baedke, "Optical repetition MIMO transmission with multipulse PPM," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1901–1910, Sep. 2005.

[258] P.-N. Chen, H.-Y. Lin, and S. M. Moser, "Optimal ultrasmall block-codes for binary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7346–7378, Nov. 2013.

[259] B. D. Unluturk, D. Malak, and O. B. Akan, "Rate-delay tradeoff with network coding in molecular nanonetworks," *IEEE Trans. Nanotechnol.*, vol. 12, no. 2, pp. 120–128, Mar. 2013.

[260] A. Aijaz, A. H. Aghvami, and M. R. Nakhai, "On error performance of network coding in diffusion-based molecular nanonetworks," *IEEE Trans. Nanotechnol.*, vol. 13, no. 5, pp. 871–874, Sep. 2014.

[261] B. C. Akdeniz, B. Tepekule, A. E. Pusane, and T. Tugcu, "Network coding applications in molecular communication," in *Proc. 23rd Signal Process. Commun. Appl. Conf. (SIU)*, May 2015, pp. 1212–1215.

[262] B. C. Akdeniz, B. Tepekule, A. E. Pusane, and T. Tuğcu, "Novel network coding approaches for diffusion-based molecular nanonetworks," *Trans. Emerg. Telecommun. Technol.*, vol. 28, no. 7, Jul. 2017, Art. no. e3105.

[263] M. F. Ghazani, G. Aminian, M. Mirmohseni, A. Gohari, and M. N. Kenari, "Physical layer network coding in molecular two-way relay networks," in *Proc. Iran Workshop Commun. Inf. Theory (IWCIT)*, May 2016, pp. 1–6.

[264] M. Farahnak-Ghazani, G. Aminian, M. Mirmohseni, A. Gohari, and M. Nasiri-Kenari, "On medium chemical reaction in diffusion-based molecular communication: A two-way relaying example," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1117–1132, Feb. 2019.

[265] B. C. Akdeniz, A. E. Pusane, and T. Tugcu, "A network coding approach for multi-hop nanonetworks in molecular communication," in *Proc. 15th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2018, pp. 1–6.

[266] P. K. Upadhyay, "Improving performance of relay-assisted molecular communication systems using network coding," in *Modelling, Simulation and Intelligent Computing*, N. Goel, S. Hasan, and V. Kalaichelvi, Eds. Singapore: Springer, 2020, pp. 184–193.

[267] Z. Cheng, Y. Tu, K. Chi, and M. Xia, "Optimization of decision thresholds in two-way molecular communication via diffusion with network coding," *IEEE Trans. Mol., Biol. Multi-Scale Commun.*, vol. 8, no. 4, pp. 249–262, Dec. 2022.

[268] Z. Cheng, J. Yan, Y. Tu, K. Chi, and M. Xia, "Mobile two-way molecular communication via diffusion using amplify-and-forward and analog network coding," *IEEE Trans. Nanobiosci.*, vol. 21, no. 2, pp. 273–285, Apr. 2022.

[269] H. S. Ghazi and K. Wesołowski, "Application of an interference cancellation detector in a two-way relaying system with physical network coding," *Electronics*, vol. 10, no. 11, May 2021, Art. no. 1294.

[270] Q. Liu, W. Zhang, S. Ding, H. Li, and Y. Wang, "Novel secure group data exchange protocol in smart home with physical layer network coding," *Sensors*, vol. 20, no. 4, Feb. 2020, Art. no. 1138.

[271] H. Qin, Y. Cao, X. Peng, and Z. Zhang, "Research on performance of cooperative FSO communication system based on hierarchical modulation and physical layer network code," *Sensors*, vol. 22, no. 18, Sep. 2022, Art. no. 6912.

[272] A. Credi, V. Balzani, S. J. Langford, and J. F. Stoddart, "Logic operations at the molecular Level. An XOR gate based on a molecular machine," *J. Amer. Chem. Soc.*, vol. 119, no. 11, pp. 2679–2681, Mar. 1997.

[273] G. Aminian, M. F. Ghazani, M. Mirmohseni, M. Nasiri-Kenari, and F. Fekri, "On the capacity of point-to-point and multiple-access molecular communications with ligand-receptors," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 1, no. 4, pp. 331–346, Dec. 2016.

[274] H. Abin, A. Gohari, and M. Nasiri-Kenari, "An analytical model for molecular communication over a non-linear reaction-diffusion medium," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8042–8054, Dec. 2021.

[275] P. Hou, A. W. Eckford, and L. Zhao, "Analysis and design of two-hop diffusion-based molecular communication with ligand receptors," *IEEE Access*, vol. 8, pp. 189458–189470, 2020.

[276] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, vol. 4, May 1995, pp. 1942–1948.

[277] N. Tavakkoli, P. Azmi, and N. Mokari, "Optimal positioning of relay node in cooperative molecular communication networks," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5293–5304, Dec. 2017.

[278] Y. Okaie and T. Nakano, "Mobile molecular communication through multiple measurements of the concentration of molecules," *IEEE Access*, vol. 8, pp. 179606–179615, 2020.

[279] H. ShahMohammadian, G. G. Messier, and S. Magierowski, "Optimum receiver for molecule shift keying modulation in diffusion-based molecular communication channels," *Nano Commun. Netw.*, vol. 3, no. 3, pp. 183–195, 2012.

[280] M. H. Kabir, S. M. R. Islam, and K. S. Kwak, "D-MoSK modulation in molecular communications," *IEEE Trans. Nanobiosci.*, vol. 14, no. 6, pp. 680–683, Sep. 2015.

[281] R. Mosayebi, A. Gohari, M. Mirmohseni, and M. Nasiri-Kenari, "Type-based sign modulation and its application for ISI mitigation in molecular communication," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 180–193, Jan. 2018.

[282] P. E. Latham and Y. Roudi, "Mutual information," *Scholarpedia*, vol. 4, no. 1, p. 1658, 2009.

[283] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "A tutorial on the optimization of amplify-and-forward MIMO relay systems," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1331–1346, Sep. 2012.

[284] I. Krikidis and J. S. Thompson, "Overview of amplify-and-forward relaying," in *Cooperative Communications for Improved Wireless Network Transmission*. Hershey, PA, USA: IGI Global, 2010, pp. 29–61.

[285] A. Noel, K. C. Cheung, and R. Schober, "Improving receiver performance of diffusive molecular communication with enzymes," *IEEE Trans. Nanobiosci.*, vol. 13, no. 1, pp. 31–43, Mar. 2014.

[286] H. B. Yilmaz, Y.-J. Cho, W. Guo, and C.-B. Chae, "Interference reduction via enzyme deployment for molecular communication," *Electron. Lett.*, vol. 52, no. 13, pp. 1094–1096, Jun. 2016.

[287] O. A. Dambri and S. Cherkaoui, "Performance enhancement of diffusion-based molecular communication," *IEEE Trans. Nanobiosci.*, vol. 19, no. 1, pp. 48–58, Jan. 2020.

[288] J. Angjo, A. E. Pusane, H. B. Yilmaz, E. Basar, and T. Tugcu, "Asymmetrical relaying in molecular communications," *IEEE Trans. Nanobiosci.*, vol. 21, no. 4, pp. 570–574, Oct. 2022.

[289] D. Jing, Y. Li, and A. W. Eckford, "Power control for ISI mitigation in mobile molecular communication," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 460–464, Feb. 2020.

[290] G. Chang, L. Lin, and H. Yan, "Adaptive detection and ISI mitigation for mobile molecular communication," *IEEE Trans. Nanobiosci.*, vol. 17, no. 1, pp. 21–35, Jan. 2017.

[291] J. Zhong, R. D. Yates, and E. Soljanin, "Timely lossless source coding for randomly arriving symbols," in *Proc. IEEE Inf. Theory Workshop*, 2018, pp. 1–5.

[292] J. T. Gómez, K. Pitke, L. Stratmann, and F. Dressler, "Age of information in molecular communication channels," *Digit. Signal Process.*, vol. 124, May 2022, Art. no. 103108.

[293] Q. Liu, K. Yang, J. Xie, and Y. Sun, "DNA-based molecular computing, storage, and communications," *IEEE Internet Things J.*, vol. 9, no. 2, pp. 897–915, Jan. 2022.

[294] K. Lu, Q. Zhou, R. Li, Z. Zhao, X. Chen, J. Wu, and H. Zhang, "Rethinking modern communication from semantic coding to semantic communication," *IEEE Wireless Commun.*, early access, May 9, 2022, doi: 10.1109/MWC.013.2100642.

[295] J. Bao, P. Basu, M. Dean, C. Partridge, A. Swami, W. Leland, and J. A. Hendler, "Towards a theory of semantic communication," in *Proc. IEEE Netw. Sci. Workshop*, Jun. 2011, pp. 110–117.

[296] R. Carnap and Y. Bar-Hillel, "An outline of a theory of semantic information," *J. Symbolic Log.*, vol. 19, no. 3, pp. 230–232, 1954.

[297] T. Matsumine, T. Koike-Akino, and Y. Wang, "Polar coding with chemical reaction networks for molecular communications," in *Proc. IEEE Globecom*, Dec. 2020, pp. 1–6.

[298] P. Garrido, D. E. Lucani, and R. Agüero, "Markov chain model for the decoding probability of sparse network coding," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1675–1685, Apr. 2017.

[299] Y. Li, J. Wang, S. Zhang, Z. Bao, and J. Wang, "Efficient coastal communications with sparse network coding," *IEEE Netw.*, vol. 32, no. 4, pp. 122–128, Jul./Aug. 2018.

[300] V. Nguyen, E. Tasdemir, G. T. Nguyen, D. E. Lucani, F. H. P. Fitzek, and M. Reisslein, "DSEP Fulcrum: Dynamic sparsity and expansion packets for Fulcrum network coding," *IEEE Access*, vol. 8, pp. 78293–78314, 2020.

[301] E. Tasdemir, M. Tömösközi, J. A. Cabrera, F. Gabriel, D. You, F. H. P. Fitzek, and M. Reisslein, "SpaRec: Sparse systematic RLNC recoding in multi-hop networks," *IEEE Access*, vol. 9, pp. 168567–168586, 2021.

[302] E. Tasdemir, V. Nguyen, G. T. Nguyen, F. H. P. Fitzek, and M. Reisslein, "FSW: Fulcrum sliding window coding for low-latency communication," *IEEE Access*, vol. 10, pp. 54276–54290, 2022.

[303] A. Zarei, P. Pahlevani, and D. E. Lucani, "An analytical model for sparse network codes: Field size considerations," *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 729–733, Apr. 2020.

[304] A. Zarei, P. Pahlevani, and M. Davoodi, "On the partial decoding delay of sparse network coding," *IEEE Commun. Lett.*, vol. 22, no. 8, pp. 1668–1671, Aug. 2018.

[305] X. Chen, Y. Xu, and A. Liu, "Cross layer design for optimizing transmission reliability, energy efficiency, and lifetime in body sensor networks," *Sensors*, vol. 17, no. 4, Apr. 2017, Art. no. 900.

[306] K. Kumar, S. Kumar, O. Kaiwartya, Y. Cao, J. Lloret, and N. Aslam, "Cross-layer energy optimization for IoT environments: Technical advances and opportunities," *Energies*, vol. 10, no. 12, Dec. 2017, Art. no. 2073.

[307] H. S. Mansour, M. H. Mutar, I. A. Aziz, S. A. Mostafa, H. Mahdin, A. H. Abbas, M. H. Hassan, N. F. Abdulsattar, and M. A. Jubair, "Cross-layer and energy-aware AODV routing protocol for flying ad-hoc networks," *Sustainability*, vol. 14, no. 15, Jul. 2022, Art. no. 8980.

[308] A. Musaddiq, Z. Nain, Y. Ahmad Qadri, R. Ali, and S. W. Kim, "Reinforcement learning-enabled cross-layer optimization for low-power and lossy networks under heterogeneous traffic patterns," *Sensors*, vol. 20, no. 15, Jul. 2020, Art. no. 4158.

[309] A. Stamou, G. Kakkavas, K. Tsitseklis, V. Karyotis, and S. Papavassiliou, "Autonomic network management and cross-layer optimization in software defined radio environments," *Future Internet*, vol. 11, no. 2, Feb. 2019, Art. no. 37.

[310] Y. S. Yang and Y. Kim, "Low-power cross-layer error management using MIMO-LDPC iterative decoding for video processing," *IEEE Access*, vol. 9, pp. 133062–133075, 2021.

[311] D. Zhao, G. Lun, R. Xue, and Y. Sun, "Cross-layer-aided opportunistic routing for sparse underwater wireless sensor networks," *Sensors*, vol. 21, no. 9, May 2021, Art. no. 3205.

[312] A. J. Aljohani and S. X. Ng, "Distributed joint source-channel coding-based adaptive dynamic network coding," *IEEE Access*, vol. 8, pp. 86715–86731, 2020.

[313] J. Balsa, Ó. Fresnedo, J. A. García-Naya, T. Domínguez-Bolaño, and L. Castedo, "JSCC-cast: A joint source channel coding video encoding and transmission system with limited digital metadata," *Sensors*, vol. 21, no. 18, Sep. 2021, Art. no. 6208.

[314] H. Bao, C. Zhang, and S. Gao, "Design and analysis of joint source-channel code system with fixed-length code," *Information*, vol. 13, no. 6, May 2022, Art. no. 281.

[315] Z. Hong, Q. Yan, Z. Li, T. Zhan, and Y. Wang, "Photon-counting underwater optical wireless communication for reliable video transmission using joint source-channel coding based on distributed compressive sensing," *Sensors*, vol. 19, no. 5, Mar. 2019, Art. no. 1042.

[316] W. Zhao and X. Chen, "Zero-delay joint source channel coding for a bivariate Gaussian source over the broadcast channel with one-bit ADC front ends," *Entropy*, vol. 23, no. 12, Dec. 2021, Art. no. 1679.

[317] F. Zabini, "Spatially distributed molecular communications: An asynchronous stochastic model," *IEEE Commun. Lett.*, vol. 22, no. 7, pp. 1326–1329, Jul. 2018.

[318] M. Crippa, C. Perego, A. L. de Marco, and G. M. Pavan, "Molecular communications in complex systems of dynamic supramolecular polymers," *Nature Commun.*, vol. 13, no. 1, pp. 1–12, Apr. 2022.

[319] X. Chen, M. Wen, C.-B. Chae, L.-L. Yang, F. Ji, and K. K. Igorevich, "Resource allocation for multiuser molecular communication systems oriented to the Internet of Medical Things," *IEEE Internet Things J.*, vol. 8, no. 21, pp. 15939–15952, Nov. 2021.

[320] L. Chouhan and M.-S. Alouini, "Rescaled Brownian motion of molecules and devices in three-dimensional multiuser mobile molecular communication systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10472–10488, Dec. 2022.

[321] E. Dinc and O. B. Akan, "Theoretical limits on multiuser molecular communication in Internet of Nano-Bio Things," *IEEE Trans. Nanobiosci.*, vol. 16, no. 4, pp. 266–270, Jun. 2017.

[322] F. Dinc, M. Medvidovic, and L. Thiele, "Effective geometry Monte Carlo: A fast and reliable simulation framework for molecular communication," *IEEE Access*, vol. 7, pp. 28635–28650, 2019.

[323] I. F. Akyildiz, J. Chen, M. Ghovanloo, U. Guler, T. Ozkaya-Ahmadov, M. Pierobon, A. F. Sarioglu, and B. D. Unluturk, "Microbiome-Gut-Brain axis as a biomolecular communication network for the Internet of Bio-NanoThings," *IEEE Access*, vol. 7, pp. 136161–136175, 2019.

[324] I. F. Akyildiz, M. Ghovanloo, U. Guler, T. Ozkaya-Ahmadov, A. F. Sarioglu, and B. D. Unluturk, "PANACEA: An Internet of Bio-NanoThings application for early detection and mitigation of infectious diseases," *IEEE Access*, vol. 8, pp. 140512–140523, 2020.

[325] M. M. Al-Zu'bi and A. S. Mohan, "Modelling of implantable drug delivery system in tumor microenvironment using molecular communication paradigm," *IEEE Access*, vol. 7, pp. 141929–141940, 2019.

[326] T. Islam, E. Shitiri, and H.-S. Cho, "A simultaneous drug release scheme for targeted drug delivery using molecular communications," *IEEE Access*, vol. 8, pp. 91770–91778, 2020.

[327] T. Islam, E. Shitiri, and H.-S. Cho, "A molecular communication-based simultaneous targeted-drug delivery scheme," *IEEE Access*, vol. 9, pp. 96658–96670, 2021.

[328] T. Islam, E. Shitiri, and H.-S. Cho, "In-body sequential multidrug delivery scheme using molecular communication," *IEEE Access*, vol. 10, pp. 39975–39985, 2022.

[329] S. Javaid, Z. Wu, H. Fahim, M. M. S. Fareed, and F. Javed, "Exploiting temporal correlation mechanism for designing temperature-aware energy-efficient routing protocol for intrabody nanonetworks," *IEEE Access*, vol. 8, pp. 75906–75924, 2020.

[330] D. T. Mcguiness, V. Selis, and A. Marshall, "Molecular-based nano-communication network: A ring topology nano-bots for in-vivo drug delivery systems," *IEEE Access*, vol. 7, pp. 12901–12913, 2019.

[331] A. Kolchinsky and D. H. Wolpert, "Semantic information, autonomous agency and non-equilibrium statistical physics," *Interface Focus*, vol. 8, no. 6, Dec. 2018, Art. no. 20180041.

[332] M. Magarini and P. Stano, "Synthetic cells engaged in molecular communication: An opportunity for modelling Shannon- and semantic-information in the chemical domain," *Frontiers Commun. Netw.*, vol. 2, Sep. 2021, Art. no. 724597.

[333] J. A. Cabrera, H. Boche, C. Deppe, R. F. Schaefer, C. Scheunert, and F. H. Fitzek, "6G and the Post-Shannon theory," in *Shaping Future 6G Networks: Needs, Impacts, and Technologies*. New York, NY, USA: IEEE Press, 2021, pp. 271–294.

[334] R. Ferrara, L. Torres-Figueroa, H. Boche, C. Deppe, W. Labidi, U. Mönich, and A. Vlad-Costin, "Practical implementation of identification codes," 2021, *arXiv:2107.06801*.

[335] W. Haselmayr, A. Springer, G. Fischer, C. Alexiou, H. Boche, P. A. Hoeher, F. Dressler, and R. Schober, "Integration of molecular communications into future generation wireless networks," in *Proc. 1st IEEE 6G Wireless Summit*, Mar. 2019, pp. 1–2.

[336] C. Von Lengerke, A. Hefele, J. Cabrera, O. Kosut, M. Reisslein, and F. Fitzek, "Identification codes: A topical review with design guidelines for practical systems," *IEEE Access*, vol. 11, 2023.

[337] C. Von Lengerke, A. Hefele, J. Cabrera, M. Reisslein, and F. Fitzek, "Beyond the bound: A new performance perspective for identification via channels," *IEEE JSAC*, 2023.

[338] M. J. Salariseddigh, U. Pereg, H. Boche, and C. Deppe, "Deterministic identification over channels with power constraints," *IEEE Trans. Inf. Theory*, vol. 68, no. 1, pp. 1–24, Jan. 2021.

[339] M. J. Salariseddigh, U. Pereg, H. Boche, C. Deppe, V. Jamali, and R. Schober, "Deterministic identification for molecular communications over the Poisson channel," 2022, *arXiv:2203.02784*.

[340] T. V. Doan, G. T. Nguyen, M. Reisslein, and F. H. P. Fitzek, "FAST: Flexible and low-latency state transfer in mobile edge computing," *IEEE Access*, vol. 9, pp. 115315–115334, 2021.

[341] T. V. Doan, G. T. Nguyen, M. Reisslein, and F. H. P. Fitzek, "SAP: Subchain-aware NFV service placement in mobile edge cloud," *IEEE Trans. Netw. Service Manage.*, early access, Aug. 24, 2023, doi: 10.1109/TNSM.2022.3201388.

[342] K. Kamran, E. Yeh, and Q. Ma, "DECO: Joint computation scheduling, caching, and communication in data-intensive computing networks," *IEEE/ACM Trans. Netw.*, vol. 30, no. 3, pp. 1058–1072, Jun. 2022.

[343] M.-C. Lee and A. F. Molisch, "Optimal delay-outage analysis for noise-limited wireless networks with caching, computing, and communications," *IEEE Trans. Wireless Commun.*, early access, Sep. 14, 2023, doi: 10.1109/TWC.2022.3204738.

[344] Y. Zhou, L. Liu, L. Wang, N. Hui, X. Cui, J. Wu, Y. Peng, Y. Qi, and C. Xing, "Service-aware 6G: An intelligent and open network based on the convergence of communication, computing and caching," *Digit. Commun. Netw.*, vol. 6, no. 3, pp. 253–260, Aug. 2020.

[345] E. R. Kay, D. A. Leigh, and F. Zerbetto, "Synthetic molecular motors and mechanical machines," *Angew. Chem. Int. Ed.*, vol. 46, nos. 1–2, pp. 72–191, Jan. 2007.

[346] K. Konstas, S. J. Langford, and M. J. Latter, "Advances towards synthetic machines at the molecular and nanoscale level," *Int. J. Mol. Sci.*, vol. 11, no. 6, pp. 2453–2472, Jun. 2010.

[347] D. Blackiston, E. Lederer, S. Kriegman, S. Garnier, J. Bongard, and M. Levin, "A cellular platform for the development of synthetic living machines," *Sci. Robot.*, vol. 6, no. 52, Mar. 2021, Art. no. eabf1571.

[348] S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan, and A. L. Nussbaumer, "Artificial molecular machines," *Chem. Rev.*, vol. 115, no. 18, pp. 10081–10206, 2015.

[349] M. Kuscu and O. B. Akan, "Detection in molecular communications with ligand receptors under molecular interference," *Digit. Signal Process.*, vol. 124, May 2022, Art. no. 103186.

[350] X. Fei, F. Liu, Q. Zhang, H. Jin, and H. Hu, "Paving the way for NFV acceleration: A taxonomy, survey and future directions," *ACM Comput. Surveys (CSUR)*, vol. 53, no. 4, p. 73, 2020.

[351] L. Linguaglossa, S. Lange, S. Pontarelli, G. Rétvári, D. Rossi, T. Zinner, R. Bifulco, M. Jarschel, and G. Bianchi, "Survey of performance acceleration techniques for network function virtualization," *Proc. IEEE*, vol. 107, no. 4, pp. 746–764, Apr. 2019.
[352] P. Shantharama, A. S. Thyagaturu, and M. Reisslein, "Hardware-accelerated platforms and infrastructures for network functions: A survey of enabling technologies and research studies," *IEEE Access*, vol. 8, pp. 132021–132085, 2020.
[353] A. S. Thyagaturu, P. Shantharama, A. Nasrallah, and M. Reisslein, "Operating systems and hypervisors for network functions: A survey of enabling technologies and research studies," *IEEE Access*, vol. 10, pp. 79825–79873, 2022.
[354] S. M. Abd El-atty, R. Bidar, and E.-S.-M. El-Rabaie, "MolCom system with downlink/uplink biocyber interface for Internet of Bio-NanoThings," *Int. J. Commun. Syst.*, vol. 33, no. 1, p. e4171, Jan. 2020.
[355] Y. Koucheryavy, A. Yastrebova, D. P. Martins, and S. Balasubramaniam, "A review on bio-cyber interfaces for intrabody molecular communications systems," 2021, *arXiv:2104.14944*.
[356] S. Mohamed, J. Dong, S. M. A. El-Atty, and M. A. Eissa, "Bio-cyber interface parameter estimation with neural network for the Internet of Bio-Nano Things," *Wireless Pers. Commun.*, vol. 123, no. 2, pp. 1245–1263, Mar. 2022.
[357] S. Zafar, M. Nazir, T. Bakhshi, H. A. Khattak, S. Khan, M. Bilal, K.-K.-R. Choo, K.-S. Kwak, and A. Sabah, "A systematic review of bio-cyber interface technologies and security issues for Internet of Bio-Nano Things," *IEEE Access*, vol. 9, pp. 93529–93566, 2021.
[358] I. Isik, M. B. Er, and E. Isik, "Analysis and classification of the mobile molecular communication systems with deep learning," *J. Ambient Intell. Humanized Comput.*, vol. 13, no. 5, pp. 2903–2919, May 2022.
[359] 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. 70279–70286, 2019.
[360] A. Allevi, S. Olivares, and M. Bondani, "Special issue on basics and applications in quantum optics," *Appl. Sci.*, vol. 11, no. 21, Oct. 2021, Art. no. 10028.
[361] F. Cavaliere, E. Prati, L. Poti, I. Muhammad, and T. Catuogno, "Secure quantum communication technologies and systems: From labs to markets," *Quantum Rep.*, vol. 2, no. 1, pp. 80–106, Jan. 2020.
[362] I. B. Djordjevic, "On global quantum communication networking," *Entropy*, vol. 22, no. 8, Jul. 2020, Art. no. 831.
[363] I. B. Djordjevic, "Hybrid CV-DV quantum communications and quantum networks," *IEEE Access*, vol. 10, pp. 23284–23292, 2022.
[364] M. Khawasik, W. Elsayed, M. Rashad, and A. Younes, "A secured quantum two-bit commitment protocol for communication systems," *IEEE Access*, vol. 10, pp. 50218–50226, 2022.
[365] A. Manzalini, "Quantum communications in future networks and services," *Quantum Rep.*, vol. 2, no. 1, pp. 221–232, Mar. 2020.
[366] E. Garlatti, A. Chiesa, T. Guidi, G. Amoretti, P. Santini, and S. Carretta, "Unravelling the spin dynamics of molecular nanomagnets with four-dimensional inelastic neutron scattering," *Eur. J. Inorganic Chem.*, vol. 2019, no. 8, pp. 1106–1118, Feb. 2019.
[367] D. Aravena and E. Ruiz, "Spin dynamics in single-molecule magnets and molecular qubits," *Dalton Trans.*, vol. 49, no. 29, pp. 9916–9928, 2020.
[368] M. L. Baker, T. Guidi, S. Carretta, J. Ollivier, H. Mutka, H. U. Güdel, G. A. Timco, E. J. L. McInnes, G. Amoretti, R. E. P. Winpenny, and P. Santini, "Spin dynamics of molecular nanomagnets unravelled at atomic scale by four-dimensional inelastic neutron scattering," *Nature Phys.*, vol. 8, no. 12, pp. 906–911, Dec. 2012.
[369] M. N. Leuenberger, F. Meier, and D. Loss, "Quantum spin dynamics in molecular magnets," in *Molecular Magnets Recent Highlights*. 2002, pp. 101–117.

PIT HOFMANN (Graduate Student Member, IEEE) received the Dipl.-Ing. degree from Technische Universität Dresden, Dresden, Germany, in 2021. He is currently pursuing the joint Ph.D. degree with the Deutsche Telekom Chair of Communication Networks, Faculty of Electrical and Computer Engineering, Technische Universität Dresden, and the Centre for Tactile Internet with Human-in-the-Loop (CeTI), Cluster of Excellence, Dresden. He continues his studies in the field of communications networks. His research interests include molecular communications and nanonetworking.

JUAN A. CABRERA received the bachelor's degree in electronics engineering from Simon Bolivar University, Caracas, Venezuela, in 2013, the master's degree in wireless communications systems from Aalborg University, Aalborg, Denmark, in 2015, and the Dr.-Ing. degree from Technische Universität Dresden, in 2022. He is currently the Group Leader of Post-Shannon Communication with the Deutsche Telekom Chair of Communication Networks, Faculty of Electrical and Computer Engineering, Technische Universität Dresden. He is also with the Centre for Tactile Internet with Human-in-the-Loop (CeTI), Cluster of Excellence, Dresden, Germany.

RICCARDO BASSOLI (Member, IEEE) received the Ph.D. degree from the 5G Innovation Centre (5GIC), University of Surrey, U.K., in 2016. He is currently an Assistant Professor with the Deutsche Telekom Chair of Communication Networks, Faculty of Electrical and Computer Engineering, Technische Universität Dresden, Germany, and the Centre for Tactile Internet with Human-in-the-Loop (CeTI), Cluster of Excellence, Dresden, Germany. He is also with the Quantum Communication Networks (QCNets) Research Group, Technische Universität Dresden.

MARTIN REISSLEIN (Fellow, IEEE) received the Ph.D. degree in systems engineering from the University of Pennsylvania, Philadelphia, PA, USA, in 1998. He is currently a Professor in electrical engineering and the Program Chair of Computer Engineering (CEN) with the School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ, USA. He received the IEEE Communications Society Best Tutorial Paper Award, in 2008, the Friedrich Wilhelm Bessel Research Award from the Alexander von Humboldt Foundation, in 2015, and a DRESDEN Senior Fellowship, in 2016 and 2019.

FRANK H. P. FITZEK (Senior Member, IEEE) received the Diploma (Dipl.-Ing.) degree in electrical engineering from the University of Technology Rheinisch-Westfälische Technische Hochschule, Aachen, Germany, in 1997, the Ph.D. (Dr.-Ing.) degree in electrical engineering from Technical University Berlin, Germany, in 2002, and the Honorary degree (Doctor Honoris Causa) from the Budapest University of Technology and Economics, in 2015. He was an Adjunct Professor with the University of Ferrara, Italy, in 2002. In 2003, he joined Aalborg University as an Associate Professor, where he became a Professor. He co-founded several start-up companies, starting with acticom GmbH, Berlin, in 1999. He is currently a Professor and also the Head of the Deutsche Telekom Chair of Communication Networks, Technische Universität Dresden, Germany, where he coordinates the 5G Laboratory, Germany.