

# Random Linear Network Coding based Physical Layer Security for Relay-aided Device-to-Device Communication

Amjad Saeed Khan<sup>1\*</sup>, Ioannis Chatzigeorgiou<sup>2</sup>, Gan Zheng<sup>1</sup>, Bokamoso Basutli<sup>3</sup>, Joseph M Chuma<sup>3</sup>, and Sangarapillai Lambotharan<sup>1</sup>

<sup>1</sup> School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, LE11 3TU, UK

<sup>2</sup> School of Computing and Communications, Lancaster University, UK

<sup>3</sup> Electrical, Computer, and Telecommunications Engineering, Botswana International University of Science and Technology, Palapye Botswana

\* E-mail: a.khan@lboro.ac.uk

**Abstract:** We investigate physical layer security design, which employs random linear network coding with opportunistic relaying and jamming to exploit the secrecy benefit of both source and relay transmissions. The proposed scheme requires the source to transmit artificial noise along with a confidential message. Moreover, in order to further improve the dynamical behaviour of the network against an eavesdropping attack, aggregated power controlled transmissions with optimal power allocation strategy is considered. The network security is accurately characterized by the probability that the eavesdropper will manage to intercept a sufficient number of coded packets to partially or fully recover the confidential message.

## 1 Introduction

Device-to-device (D2D) communication has attracted enormous attention from both academia and industry and is considered as one of the promising technologies for achieving high spectral efficiency and ultra-densification in future wireless networks (5G and beyond) [1], [2]. However, one of the major challenges in D2D networks is the broadcast nature of wireless medium that makes the communication over this medium vulnerable to eavesdropping attacks, such that, any node in the coverage range of transmitting node will be able to listen and extract information [3]. Moreover, the efficiency of D2D communication reduces when the source and destination are not in proximity. For example, the outage probability of links connecting D2D pairs will increase if the nodes are farther apart. Furthermore, some destinations may not be within the range of transmitting devices because of transmit power constraints. This motivates the use of cooperative relays to enhance the range of communication and thus to improve the network efficiency [4], [5]. In this context, Random Linear Network Coding (RLNC) has been realized as one of the potential paradigms for D2D communications that not only supports cooperative communication for high throughput and efficiency but can also provide lightweight security in networks [6–9]. For example, RLNC contains an inherent feature of security to prevent information leakage to eavesdropper even if the eavesdropper overhears some of the source transmissions. Moreover, despite the use of numerous cryptographic methods, Physical Layer Security (PLS) is emerging as one of the promising solutions for ensuring secrecy in wireless networks [10], [11], [12]. The main idea is to exploit the physical characteristics of wireless channels including fading and noise to transmit an information message from a source to a legitimate destination while keeping it confidential from an eavesdropper. In particular, securing information with the injection of artificial noise to confuse the eavesdropper is one of the fundamental techniques in PLS which has been widely studied in the literature [13], [14].

There are several works studying the benefits of network coding in a PLS framework [15], [16], [17]. For example, Khan *et al.* in [15] studied the intrinsic nature of RLNC against eavesdropping attacks and identified the advantage of optimising the number of source transmissions based on feedback by the legitimate destination. Later

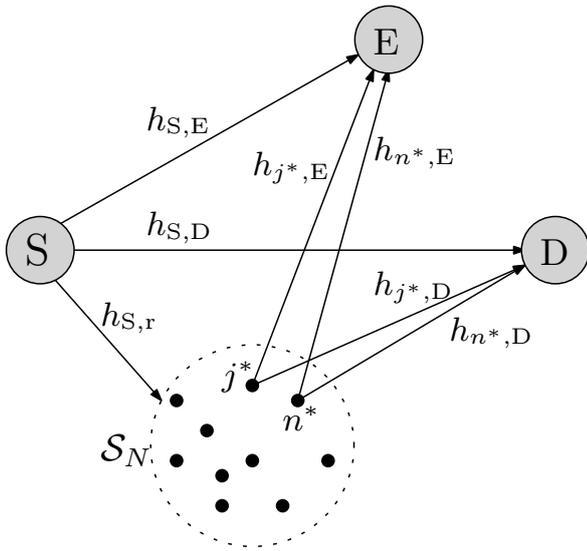
in [17], the authors proposed an RLNC based opportunistic relaying and jamming framework. Sun *et al.* proposed fountain-coding based cooperative jamming technique in [16]. Differently from the previous work, this work takes into account power controlled transmissions of both message and interfering signals, which improves the dynamic behaviour of the network against the eavesdropping attack. In addition, the security advantage of a direct source-to-destination link is exploited which is often ignored in PLS designs, while the source is allowed to superimpose the confidential message onto an artificial noise (AN) signal. Furthermore, optimal relay selection for signal jamming is carried out not only when the source signal is relayed to the destination but also when it is broadcast to the relays. Note that, unlike [16], [18], in this work the legitimate destination is not able to mitigate any AN and interference signals.

The main contributions are summarised as follows: (i) We investigate a physical layer security design including RLNC based opportunistic relaying and jamming, where source and relay nodes are jointly considered with adjustable transmission power for both security and reliability; (ii) we derive exact theoretical expressions of the outage probabilities at both the destination and the eavesdropper, and accurately characterise the network performance in terms of the  $\tau$ -intercept probability which quantifies the exact number of data packets. Extensive simulations are conducted to validate the accuracy of the derived expressions, and the network performance is investigated based on the optimal power allocations.

The rest of this paper is organised as follows: Section 2 describes the system model. A detailed description of the considered relay selection techniques is provided in Section 3 and outage probability expressions for both the legitimate destination and the eavesdropper are derived. Exact theoretical expressions for quantifying the secrecy of RLNC-enabled opportunistic relaying and jamming are derived in Section 3.2. Section 3.3 presents the resource allocation model to obtain the optimal values of power control coefficients. Results are discussed in Section 4 and conclusions are drawn in Section 5.

## 2 System Model

We consider a network with a source S, one destination D, one eavesdropper E, and a set of  $N$  trusted relays  $\mathcal{S}_N = \{1, \dots, N\}$ , as shown in Fig. 1. Each node is equipped with a single antenna and operates in half-duplex mode. All the links connecting the nodes are assumed to be independent but not identically distributed



**Fig. 1:** Block diagram of the system model. Note that,  $j_1^*$  represents  $j_1^*$  in the broadcast phase and relay  $j_2^*$  in the relay phase but  $j_1^*$  can be different to  $j_2^*$ .

(i.n.i.d) quasi-static Rayleigh fading channels, as in [19]. The channel gain between node  $i$  and  $j$  is represented as  $|h_{i,j}|$  with variance  $\sigma_{i,j}^2 = d_{i,j}^{-\eta_{i,j}}$ , where  $d_{i,j}$  and  $\eta_{i,j}$  are the Euclidean distance and the path loss exponent between the two nodes, respectively. Moreover, we assume that the transmission power of each node is limited to  $P$ .

Before the communication process, the source  $S$  first divides the message into  $K$  data packets. Afterwards, RLNC over some finite field  $\mathbb{F}_q$  [20] is employed to encode the packets into  $K$  linearly independent coded packets, where  $q$  represents the field size. These coded packets are then forwarded to the physical layer for further processing. At the physical layer, the modulation and coding scheme (MCS) converts the stream of  $K$  coded packets into a sequence of  $K$  signals, represented as  $\{x_{s,1}, x_{s,2}, \dots, x_{s,K}\}$  which will be sent over the wireless channel. The end-to-end communication is performed in two phases: the broadcast phase and the relay phase.

During the broadcast phase, the source generates a signal that is the superposition of  $x_{s,i}$  and the AN signal  $x_{A,i}$  and broadcasts it towards  $D$ . The transmitted signal can be represented as  $(\sqrt{P a_s} x_{s,i} + \sqrt{P \bar{a}_s} x_{A,i})$ , where  $a_s$  and  $\bar{a}_s$  are the power control coefficients, such that,  $a_s \geq \bar{a}_s$  and  $a_s + \bar{a}_s = 1$ . Meanwhile, in order to further combat the eavesdropper's attack and according to the security protocol, a selected relay  $j_1^*$  operates in *jamming mode*, that is, it radiates artificial interference in synchronization with the source signal to deteriorate the eavesdropper's channel. The signal transmitted by  $j_1^*$  can be expressed as  $\sqrt{P a_{j_1}} x_{j_1^*,i}$ , where  $a_{j_1}$  represents the power control coefficient. In order to protect the destination from severe artificial interference, we assume  $a_{j_1} \leq a_s$ . Note that a centralized control unit is considered for selecting an appropriate relay and controlling transmission powers for each transmission, based on the perfectly known channel state information of all the links. Additive white Gaussian noise is assumed at each node with zero mean and variance  $N_0$ . If we set  $\rho = P/N_0$ , the received SINR at node  $Z \in \mathcal{S}_N \setminus j_1^* \cup \{D, E\}$  can be expressed as:

$$\text{SINR}_Z^{(S)} = \frac{\rho a_s |h_{S,Z}|^2}{\rho \bar{a}_s |h_{S,Z}|^2 + \rho a_{j_1} |h_{j_1^*,Z}|^2 + 1}. \quad (1)$$

Both the destination  $D$  and eavesdropper  $E$  demodulate and store the correctly received coded packets for future RLNC decoding. On the other hand, each node in  $\mathcal{S}_N \setminus j_1^*$  demodulates and stores the correctly received coded packets for RLNC decoding at the end of the phase. For simplicity and to avail the possibility of selecting  $j_1^*$  as a potential relay in the next phase, we assume that all the relays in  $\mathcal{S}_N$  cooperate locally to exchange missing coded packets. Thus after collecting  $K$  linearly independent coded packets, at the end of this

**Table 1** Key parameters of the system model

Notation	Description
$K$	Number of data packets.
$N$	Number of relay nodes.
$\mathcal{S}_N$	Set of $N$ trusted relay nodes.
$P$	Total transmit power of a node.
$N_0$	Variance of the additive white Gaussian noise
$h_{i,j}$	Fading coefficient of the channel between nodes $i$ and $j$ .
$\gamma_{i,j}$	Instantaneous SNR of the link between nodes $i$ and $j$ .
$\lambda_{i,j}$	The inverse of the average SNR of the link between nodes $i$ and $j$ .
$\varepsilon_{ij}$	Outage probability of the link between nodes $i$ and $j$ .
$n^*$	Selected relay.
$j_1^*$	Selected jammer during the broadcast phase.
$j_2^*$	Selected jammer during the relay phase.
$n_T$	Number of transmitted coded packets.
$\eta_r$	Transmissions during the relay phase.
$N_T$	Maximum permitted number of transmissions.

phase, each relay in  $\mathcal{S}_N$  employs RLNC decoding and successfully retrieves the original data packets.

During the relay phase and according to the protocol, two selected relays transmit towards  $D$  and  $E$ . The first relay  $n^*$  generates a coded packet using RLNC over the same finite field  $\mathbb{F}_q$ , and forwards to the destination  $D$ . At the same time, the second relay  $j_2^*$  generates artificial interference for the same reason as that for  $j_1^*$ . The signal transmitted by both  $n^*$  and  $j_2^*$  can be expressed as  $\sqrt{P a_r} x_{n^*,i}$  and  $\sqrt{P a_{j_2}} x_{j_2^*,i}$ , respectively, where  $a_r$  and  $a_{j_2}$  are power control coefficients such that  $a_{j_2} \leq a_r$  and  $a_r + a_{j_2} = 1$ . The received SINR at node  $Z \in \{D, E\}$  is:

$$\text{SINR}_Z^{(R)} = \frac{\rho a_r |h_{n^*,Z}|^2}{\rho a_{j_2} |h_{j_2^*,Z}|^2 + 1}. \quad (2)$$

This process is repeated up to  $\eta_r$  times, and thus up to  $\eta_r$  coded packets are transmitted during this phase; each time, the appropriate relays  $n^*$  and  $j_2^*$  are selected from  $\mathcal{S}_N$ , depending on the instantaneous channel conditions. Both the destination  $D$  and eavesdropper  $E$  are needed to collect at least  $K$  linearly independent coded packets in order to reconstruct the source message. If the destination recovers the message before the set deadline of  $N_T = K + \eta_r$  total transmissions in both phases, it sends a notification to the control unit to terminate the relay selection and packet transmission process. For convenience, the key parameters of the system model are summarized in Table 1.

### 3 Relay Selection and Outage Probabilities

This section presents the proposed relay selection schemes, and evaluates their performance in terms of outage probabilities at both the destination  $D$  and the eavesdropper  $E$ . For notational convenience, let us define  $\gamma_{i,j} = \rho |h_{i,j}|^2$  as the instantaneous SNR of the link between node  $i$  and  $j$ . The probability density function of  $\gamma_{i,j}$  follows the exponential distribution, that is:  $f_{\gamma_{i,j}}(x) = \lambda_{i,j} \exp(-x \lambda_{i,j})$ . If we define rate parameter  $\lambda_{i,j} = 1/\mathbb{E}\{\gamma_{i,j}\}$  with  $\mathbb{E}\{\cdot\}$  as the expectation operation, the cumulative distribution function of  $\gamma_{i,j}$  is equal to:

$$\Pr(\gamma_{i,j} \leq \hat{\gamma}) = 1 - e^{-\hat{\gamma} \lambda_{i,j}}. \quad (3)$$

**Opportunistic jammer:** This protocol incorporates the instantaneous channel quality of both relay to destination and eavesdropper links [19]. According to this scheme, an optimal jammer  $j_1^*$  is selected such that it generates maximum interference to the eavesdropper while causing least effect to the destination. Mathematically, it can be expressed as:

$$j_1^* = \arg \max_{j \in \mathcal{S}_N} \left( \frac{\gamma_{j,E}}{\gamma_{j,D}} \right). \quad (4)$$

Using the theory of order statistics (maximum among  $N$  independent and identically distributed (i.i.d) random variables) and taking integrals with respect to both  $\gamma_{j,E}$  and  $\gamma_{j,D}$  [21], we can obtain:

$$\Pr(j_1^* = j) = \int_0^\infty \int_0^\infty \prod_{\substack{i=1 \\ i \neq j}}^N \Pr\left(\frac{\gamma_{i,E}}{\gamma_{i,D}} \leq \frac{x}{y}\right) f_{\gamma_{j,E}}(x) f_{\gamma_{j,D}}(y) dx dy. \quad (5)$$

**Opportunistic relay and jammer:** For high reliability and for low complexity of selecting a relay-jammer pair, this protocol only takes into account the channel quality of relay-to-destination link, and selects a relay  $n^*$  and jammer  $j_2^*$ , such that:

$$n^* = \arg \max_{n \in \mathcal{S}_N} \gamma_{n,D}. \quad (6)$$

$$j_2^* = \arg \min_{m \in \mathcal{S}_N \setminus n^*} \gamma_{m,D}. \quad (7)$$

The probability of selecting the two nodes can be obtained by employing order statistics (maximum and minimum among  $N$  i.i.d random variables) and taking integrals with respect to both  $\gamma_{m,D}$  and  $\gamma_{n,D}$  [21], as follows:

$$\Pr[n^* = n, j_2^* = m] = \int_0^\infty \int_0^\infty \prod_{\substack{i=1 \\ i \neq n, \\ i \neq m}}^{N-2} \Pr(y \leq \gamma_{i,D} \leq x) f_{\gamma_{m,D}}(y) f_{\gamma_{n,D}}(x) dy dx. \quad (8)$$

Note that, several relay selection algorithms are available in the literature, aiming at promoting the assistance to the source as well as interference to the eavesdropper, but this discussion is beyond the scope of this paper.

### 3.1 Outage analysis

The outage event can be defined as the event when the instantaneous SINR drops below a predefined threshold. As shown in [22], the physical-layer MCS can be accurately characterized by an SNR threshold. Let us denote  $\hat{\gamma}_S$  and  $\hat{\gamma}_R$  as SNR thresholds for source and relay transmissions, respectively. The outage probability for the transmission between node  $i$  and  $j$  can be defined as:

$$\varepsilon_{ij} = \Pr(\text{SINR}_j^{(i)} \leq \hat{\gamma}_i).$$

**3.1.1 Broadcast Phase:** During this phase, the outage probability at D can be expressed as:

$$\begin{aligned} \varepsilon_{SD} &= \Pr(\text{SINR}_D^{(S)} \leq \hat{\gamma}_S) \\ &= \Pr\left(\frac{a_s \gamma_{S,D}}{\bar{a}_s \gamma_{S,D} + a_{j_1} \gamma_{j_1^*,D} + 1} \leq \hat{\gamma}_S\right) \\ &= \Pr(\gamma_{S,D} \leq A \gamma_{j_1^*,D} + B) \end{aligned} \quad (9)$$

where  $A = \frac{\hat{\gamma}_S a_{j_1}}{a_s - \bar{a}_s \hat{\gamma}_S}$  and  $B = \frac{\hat{\gamma}_S}{a_s - \bar{a}_s \hat{\gamma}_S}$ . Thus, the closed form expression of  $\varepsilon_{SD}$  can be obtained by exploiting the law of total

probability [23] over the joint probability of selecting  $j_1^*$  and the outage event at D, as follows:

$$\varepsilon_{SD} = \sum_{j=1}^{N-1} \Pr\left[(j_1^* = j) \cap (\gamma_{S,D} \leq A \gamma_{j_1^*,D} + B)\right]. \quad (10)$$

Using expression (5) and by properly setting the limits, we obtain

$$\varepsilon_{SD} = \sum_{j=1}^{N-1} \int_0^\infty \int_0^\infty \prod_{\substack{i=1 \\ i \neq j}}^N \Pr\left(\frac{\gamma_{i,E}}{\gamma_{i,D}} \leq \frac{x}{y}\right) f_{\gamma_{j,E}}(x) f_{\gamma_{S,D}}(z) f_{\gamma_{j,D}}(y) dx dz dy \quad (11)$$

By employing partial fractions, the product expression can be expanded as:

$$\prod_{\substack{i=1 \\ i \neq n}}^N \Pr\left(\frac{\gamma_{i,E}}{\gamma_{i,D}} \leq \frac{x}{y}\right) = 1 - \sum_{\substack{i=1 \\ i \neq n}}^N \frac{y}{x \Lambda_i + y} \prod_{k \notin \{n, i\}} \frac{-\Lambda_k}{\Lambda_i - \Lambda_k} \quad (12)$$

where  $\Lambda_i = \frac{\lambda_{i,E}}{\lambda_{i,D}}$ . Thus, (13) can be re-expressed as:

$$\varepsilon_{SD} = \sum_{j=1}^{N-1} \int_0^\infty \int_0^\infty \int_0^\infty \left[ 1 - \sum_{\substack{i=1 \\ i \neq n}}^N \frac{y}{x \Lambda_i + y} \prod_{k \notin \{n, i\}} \frac{-\Lambda_k}{\Lambda_i - \Lambda_k} \right] f_{\gamma_{j,E}}(x) f_{\gamma_{S,D}}(z) f_{\gamma_{j,D}}(y) dx dz dy. \quad (13)$$

Evaluating the integrals and utilizing the relationships in [24, 25] leads to:

$$\begin{aligned} \varepsilon_{SD} &= \sum_{n=1}^N \left[ 1 - \frac{\lambda_{n,D} e^{-B \lambda_{S,D}}}{A \lambda_{S,D} + \lambda_{n,D}} - \sum_{j \neq n}^N \frac{\lambda_{n,E} \lambda_{n,D}}{\Lambda_j} \left[ \frac{1}{\alpha^2} \left\{ \ln\left(\frac{\Lambda_j}{\Lambda_n}\right) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{\Lambda_n}{\Lambda_j} - 1 \right\} - \frac{e^{-B \lambda_{S,D}}}{\alpha_1^2} \left\{ \ln\left(\frac{A \lambda_{S,D} \Lambda_j + \lambda_{n,D} \Lambda_j}{\lambda_{n,E}}\right) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{\lambda_{n,E}}{A \lambda_{S,D} \Lambda_j + \lambda_{n,D} \Lambda_j} - 1 \right\} \right] \right] \end{aligned} \quad (14)$$

where,  $\alpha = \lambda_{n,D} - \frac{\lambda_{n,E}}{\Lambda_j}$ ,  $\alpha_1 = A \lambda_{S,D} - \frac{\lambda_{n,E}}{\Lambda_j} + \lambda_{n,D}$ . Following the same line of thought, the outage probability at E is equal to:

$$\varepsilon_{SE} = \sum_{n=1}^{N-1} \int_0^\infty \int_0^\infty \int_0^\infty \prod_{\substack{i=1 \\ i \neq n}}^N \Pr\left(\frac{\gamma_{i,E}}{\gamma_{i,D}} \leq \frac{x}{y}\right) f_{\gamma_{S,E}}(z) f_{\gamma_{n,E}}(x) f_{\gamma_{n,D}}(y) dx dz dy$$

In order to derive an analytical expression of  $\varepsilon_{SE}$ , we follow a similar approach to the derivation of  $\varepsilon_{SD}$ , i.e., expansion of the product term and evaluation of the integrals. The closed form expression of  $\varepsilon_{SE}$  is:

$$\begin{aligned} \varepsilon_{SE} &= \sum_{n=1}^N \left[ 1 - \frac{\lambda_{n,E} e^{-B \lambda_{S,E}}}{A \lambda_{S,E} + \lambda_{n,E}} - \sum_{j \neq n}^N \frac{\lambda_{n,E} \lambda_{n,D}}{\Lambda_j} \left[ \frac{1}{\hat{\alpha}^2} \left\{ \ln\left(\frac{\Lambda_j}{\Lambda_n}\right) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{\Lambda_n}{\Lambda_j} - 1 \right\} - \frac{e^{-B \lambda_{S,E}}}{\hat{\alpha}_1^2} \left\{ \ln\left(\frac{\Lambda_j \lambda_{n,D}}{A \lambda_{S,E} + \lambda_{n,E}}\right) - 1 \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{A \lambda_{S,E} + \lambda_{n,E}}{\Lambda_j \lambda_{n,D}} \right\} \right] \right] \end{aligned} \quad (15)$$

where  $\hat{\alpha} = \lambda_{n,D} - \frac{\lambda_{n,E}}{\Lambda_j}$  and  $\hat{\alpha}_1 = \lambda_{n,D} - \frac{A \lambda_{S,E} + \lambda_{n,E}}{\Lambda_j}$ .

3.1.2 *Relay phase:* During this phase, the outage probability at D should take into account the joint probability of selecting a pair  $\{n^*, j^*\}$ , and the SINR at the destination not exceeding the SNR threshold  $\hat{\gamma}_R$ . Thus, by employing the law of total probability,  $\varepsilon_{RD}$  can be expressed as:

$$\varepsilon_{RD} = \sum_{n=1}^N \sum_{m \neq n}^N \Pr \left[ \left( n^* = n, j_2^* = m \right) \cap \left( \frac{a_r \gamma_{n,D}}{a_{j_2} \gamma_{m,D} + 1} \leq \hat{\gamma}_R \right) \right]. \quad (16)$$

Exploiting (8) and by properly setting the limits, we obtain:

$$\varepsilon_{RD} = \sum_{n=1}^N \sum_{m \neq n}^N \int_0^{\frac{(a_{j_2} y + 1) \hat{\gamma}_R}{a_r}} \int_y^{\infty} \prod_{\substack{i \neq n, \\ i \neq m}}^{N-2} \Pr(y \leq \gamma_{i,D} \leq x) f_{\gamma_{n,D}}(x) f_{\gamma_{m,D}}(y) dx dy$$

by invoking (3), we can obtain:

$$\varepsilon_{RD} = \sum_{n=1}^N \sum_{m \neq n}^N \int_0^{\frac{(a_{j_2} y + 1) \hat{\gamma}_R}{a_r}} \int_y^{\infty} \prod_{\substack{i \neq n, \\ i \neq m}}^{N-2} (e^{-y \lambda_{i,D}} - e^{-x \lambda_{i,D}}) f_{\gamma_{n,D}}(x) f_{\gamma_{m,D}}(y) dx dy. \quad (17)$$

Using the multinomial identity [26], we can expand the expression as:

$$\varepsilon_{RD} = \sum_{n=1}^N \sum_{m \neq n}^N \sum_{\ell=0}^{N-2} \sum_{|\mathcal{S}_\ell|=\ell} \int_0^{\frac{(a_{j_2} y + 1) \hat{\gamma}_R}{a_r}} \int_y^{\infty} (-1)^\ell e^{-x \sum_{i \in \mathcal{S}_\ell} \lambda_{i,D} - y \sum_{j \in \bar{\mathcal{S}}_\ell} \lambda_{j,D}} f_{\gamma_{n,D}}(x) f_{\gamma_{m,D}}(y) dx dy \quad (18)$$

where,  $\mathcal{X} = \mathcal{S}_N \setminus \{n, m\}$ , and  $\mathcal{S}_\ell \cup \bar{\mathcal{S}}_\ell = \mathcal{X}$ . By solving the integrals, the closed form expression can be obtained as:

$$\varepsilon_{RD} = \sum_{n=1}^N \sum_{m \neq n}^N \sum_{\ell=0}^{N-2} \sum_{|\mathcal{S}_\ell|=\ell} \frac{\lambda_{n,D} \lambda_{m,D}}{(-1)^\ell \varpi_n} \left\{ \frac{1}{\varpi_n + \varpi_m} - \frac{\exp\left(-\frac{1}{a_r} \hat{\gamma}_R \varpi_n\right)}{\hat{\gamma}_R \frac{a_{j_2}}{a_r} \varpi_n + \varpi_m} \right\} \quad (19)$$

where,  $\varpi_n = \sum_{i \in \mathcal{S}_\ell} \lambda_{n,E} + \lambda_{n,D}$ ,  $\varpi_m = \sum_{j \in \bar{\mathcal{S}}_\ell} \lambda_{m,E} + \lambda_{m,D}$ ,  $\hat{\gamma}_R \geq 0$ .

Given that the selection procedure of relay-jammer pair  $\{n^*, j_2^*\}$  is independent to the channel quality of relay-to-eavesdropper links, using the law of total probability likewise in the previous cases, the outage probability at E can be obtained as:

$$\varepsilon_{RE} = \sum_{n=1}^N \sum_{m \neq n}^N \Pr \left[ n^* = n, j_2^* = m \right] \Pr \left( \frac{a_r \gamma_{n,E}}{a_{j_2} \gamma_{m,E} + 1} \leq \hat{\gamma}_R \right). \quad (20)$$

By employing (8), the analytical expression of  $\varepsilon_{RE}$  can be expressed as:

$$\varepsilon_{RE} = \sum_{n=1}^N \sum_{m \neq n}^N \sum_{\ell=0}^{N-2} \sum_{|\mathcal{S}_\ell|=\ell} \left( 1 - \frac{\lambda_{m,E} e^{-\frac{\hat{\gamma}_R}{a_r} \lambda_{n,E}}}{\frac{a_{j_2} \hat{\gamma}_R}{a_r} \lambda_{n,E} + \lambda_{m,E}} \right) \frac{(-1)^\ell \lambda_{n,D} \lambda_{m,D}}{\sum_{j \in \bar{\mathcal{S}}_\ell} \lambda_{j,D} + \lambda_{m,D}} \left( \frac{1}{\sum_{i \in \mathcal{S}_\ell} \lambda_{i,D} + \lambda_{n,D}} - \frac{1}{\sum_{k=1}^N \lambda_{k,D}} \right). \quad (21)$$

### 3.2 Secrecy Analysis

This section quantifies the network performance in terms of probability that the eavesdropper will manage to recover at least  $\tau$  of the  $K$  data packets using Gaussian elimination, which is defined as  $\tau$ -intercept probability. Note that, a receiver requires to collect  $K$  linearly independent coded packets to recover the entire message composed of  $K$  data packets. To evaluate the network performance, let us assume that  $n_T \leq N_T$  transmissions are carried out during the communication process, where  $N_T$  represents the maximum permitted number of transmissions. The probability of successfully receiving  $n_R$  coded packets, and  $r \leq K$  of them are linearly independent is equal to:

$$\begin{aligned} \mathbb{P}_r^Z(n_R, a_s, a_{j_1}, a_r) &= \sum_{h_B=h_{\min}}^{\min(n_R, K)} \binom{K}{h_B} \binom{n_T - K}{n_R - h_B} \\ &\cdot \varepsilon_{SZ}(a_s, a_{j_1}, a_r)^{K-h_B} \varepsilon_{RZ}(a_s, a_{j_1}, a_r)^{n_T - K - n_R + h_B} \\ &\cdot (1 - \varepsilon_{SZ}(a_s, a_{j_1}, a_r))^{h_B} (1 - \varepsilon_{RZ}(a_s, a_{j_1}, a_r))^{n_R - h_B} \\ &\cdot P_{r-h_B}^Z(K - h_B, n_R - h_B) \end{aligned} \quad (22)$$

where  $Z \in \{D, E\}$ ,  $h_B$  represents the number of packets received during the broadcast phase with  $h_{\min} = \max(0, n_R - n_T + K)$ , and  $\varepsilon_{ij}$  can be evaluated using the outage probability expressions derived in Section 3.  $P_{r-h_B}$  is the probability of obtaining  $r - h_B$  linearly independent coded packets during the relay phase, can be obtained using [27]:

$$P_{r-h_B}^Z(K, n_R) = \frac{1}{q^{n_R K}} \left[ \begin{matrix} n_R \\ r - h_B \end{matrix} \right]_q \prod_{i=0}^{r-h_B-1} (q^K - q^i) \quad (23)$$

where  $q$  represents the Finite field size, and  $\left[ \begin{matrix} u \\ v \end{matrix} \right]_q$  is the  $q$ -binomial coefficient defined as [28, Eq. 1]. The probability that exactly  $\tau \leq r$  data packets can be recovered after collecting  $r$  linearly independent coded packets, is equal to [29]:

$$P(\tau, K|r) = \frac{\binom{K}{\tau}}{\binom{K}{r}} \sum_{j=0}^{K-\tau} (-1)^j \binom{K-\tau}{j} \binom{K-\tau-j}{r-\tau-j}_q. \quad (24)$$

In order to characterize the secrecy performance: let  $X_D$  represent the number of transmissions required by the destination D to recover the entire message, and denote  $X_E$  as the number of transmissions needed by the eavesdropper E to recover at least  $\tau$  data packets. We can express the cumulative distribution function of both  $X_D$  and  $X_E$  as follows:

$$\begin{aligned} F_Z(x, n_T) &= \Pr \{X_Z \leq n_T\} \\ &= \sum_{n_R=x}^{n_T} \sum_{r=x}^{\min(n_R, K)} \sum_{i=x}^r \mathbb{P}_r^Z(n_R, a_s, a_{j_1}, a_r) P(i, K|r) \end{aligned} \quad (25)$$

where  $Z \in \{D, E\}$ ,  $x$  is considered as  $\tau$  for E and  $K$  for D. The corresponding probability mass function is equal to:

$$\begin{aligned} f_Z(x, n_T) &= \Pr \{X_Z = n_T\} \\ &= F_Z(x, n_T) - F_Z(x, n_T - 1). \end{aligned} \quad (26)$$

The average number of transmissions required by Z to recover at least  $x$  data packets can be expressed as:

$$E_x(N_T) = N_T - \sum_{v=0}^{D_T-1} F_Z(x, x+v) \quad (27)$$

where  $D_T$  represents the maximum permissible number of excess coded packet transmissions, that is,  $D_T = N_T - x$ . The  $\tau$ -intercept

probability that the eavesdropper will be successful in recovering at least  $\tau$  data packets from the intercepted coded packets can be expressed as [15]:

$$P_{\text{int}}(K, N_T, \tau, a_s, a_{j_1}, a_r) = F_E(\tau, N_T) [1 - F_D(K, N_T)] + \sum_{n_T=K}^{N_T} f_D(K, n_T) F_E(\tau, n_T) \quad (28)$$

The first term of  $P_{\text{int}}(K, N_T, \tau, a_s, a_{j_1}, a_r)$  calculates the probability that the eavesdropper will be successful in recovering at least  $\tau$  data packets from the intercepted coded packets but the destination will fail to reconstruct the message within the given transmissions bound. The second term evaluates the probability of the event that the destination will recover the complete message after the  $n_T$ -th coded packet packet has been transmitted but the eavesdropper has already recovered at least  $\tau$  data packets by that time.

### 3.3 Resource Allocation Model

This section aims to determine the optimum values of power controlled coefficients for minimizing the intercept probability while supporting the legitimate destination to successfully reconstruct the source message through a limited number of transmissions, as discussed in Section 2. Using the theoretical formulation of  $P_{\text{int}}(K, N_T, \tau, a_s, a_{j_1}, a_r)$ , the optimum values of power control coefficients can be obtained by

$$[a_s^*, a_{j_1}^*, a_r^*] = \arg \min_{a_s, a_{j_1}, a_r} P_{\text{int}}(K, N_T, \tau, a_s, a_{j_1}, a_r) \quad (29)$$

$$\text{subject to } F_D(K, N_T) \geq \hat{P}, N_T \leq \hat{N} \quad (30)$$

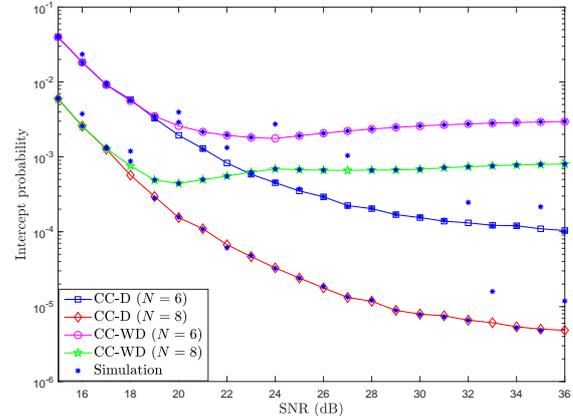
$$0.5 < a_s \leq 1, 0 < a_{j_1} \leq a_s, 0.5 < a_r \leq 1 \quad (31)$$

where (30) ensures that the message recovery probability of destination is at least  $\hat{P}$ , and prevents the uncontrollable increase of transmissions that is the number of intended coded packets is less than or equal to  $\hat{N}$ . Whereas (31) sets the suitable range of power control coefficients. Given the challenging nature of the optimisation problem and the lack of closed-form solutions, the direct search algorithm has been used and results are presented in Section 4. Note that the direct search algorithm is an iterative method of finding a globally optimal solution. In particular, instead of exploring all possible values of optimization parameters, the direct search algorithm converges quickly because during each iteration it searches both globally and locally. Once it finds the basin of convergence of the optimum, the algorithm automatically starts searching locally to find the optimum solution. A step-by-step description of the algorithm is provided in [30].

## 4 Results and Discussion

This section presents simulation results and validates the accuracy of the theoretical expressions. A Monte Carlo simulation platform representing the system model was developed in MATLAB. Instances where the eavesdropper successfully recovered at least  $\tau$  data packets were counted and averaged over  $10^5$  realizations to compute the intercept probability. The performance difference between the communication scenario when the direct links between source to destination and eavesdropper nodes are considered (which will be referred as CC-D) and when the direct links between the source to destination and eavesdropper nodes are not considered (which will be referred as CC-WD) while setting  $\epsilon_{SD} = \epsilon_{SE} = 1$  is also highlighted. Note that, for CC-WD, we assume that the direct links could be in deep shadowing or the destination and the eavesdropper could be outside the coverage area of the source. Let the pair  $\{d_{i,D}, d_{i,E}\}$  denote the distance of  $i^{\text{th}}$  relay node from D and E. The distance pairs in the simulation environment are configured as: (2, 1.3), (2, 1), (3, 3), (2, 3), (3, 4), (1.5, 1), (1.5, 1.1), (2.3, 1.5).

Without loss of generality, both source S and destination D are located at (0, 0) and (4, 0), respectively, and unless otherwise stated, the eavesdropper is considered at location (3, 0). In addition, in all the cases we consider  $\hat{P} = 0.90$ ,  $\hat{N} = 3K$ , and path loss exponent  $\eta_{i,j} = \eta = 3$ . Moreover for illustration purposes, we assume that the source employs convolutionally coded BPSK and the relay considers un-coded BPSK for physical layer transmissions, which are characterized by SNR thresholds  $\hat{\gamma}_S = -0.983$  dB and  $\hat{\gamma}_R = 5.782$  dB, respectively [22]. The term 'SNR' is used to refer to  $\rho = P/N_0$ , as defined in Section 2.

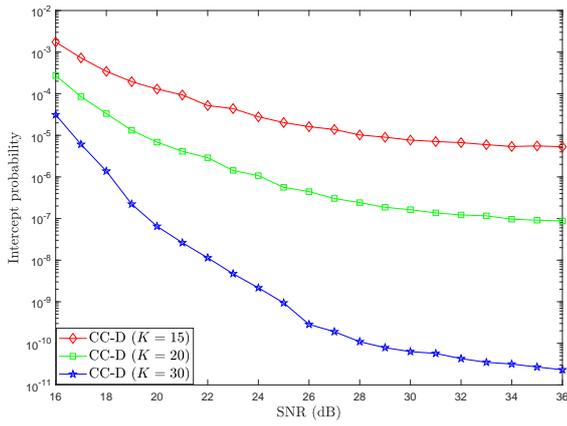


**Fig. 2:** Performance comparison between CC-D and CC-WD, when  $K = 15$ ,  $\tau = 11$ ,  $q = 4$ .

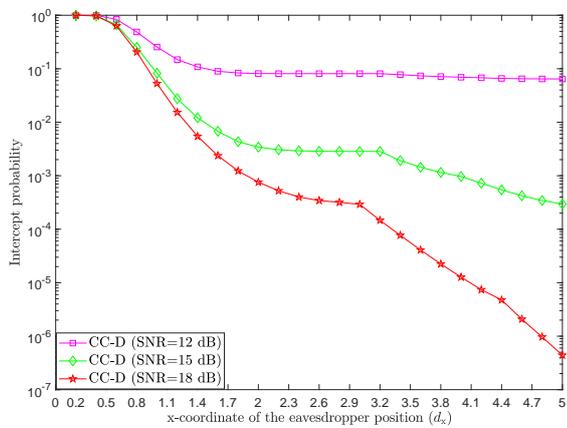
Fig. 2 shows the intercept probability as a function of transmitted SNR and the number of relays for both CC-D and CC-WD strategies. The analytical curves match well the simulation results, which confirms the correctness of our mathematical analysis. It can be seen that the intercept probability of CC-D reduces with increasing SNR, which opposes the conventional view that the intercept probability usually increases with the enhancement of SNR [31], [32]. This is due to the fact that the allocated power to interfering signals increases with SNR increasing, which effectively decreases the received SINR at the eavesdropper. On the other hand, at medium to high SNR regions, CC-WD follows the traditional pattern of intercept probability. Interestingly, at the low SNR region, the intercept probability of CC-WD reduces with the increase of SNR. This is because the eavesdropper links are less resistant to artificial interference at low SNR values, therefore optimal power allocation between relay and jammer leads to improved secrecy performance. Note that, because of the jamming effect, the intercept probability converges to a constant value, which specifies that a level of security can be offered even at high SNR values. It can also be noted that the secrecy performance of both CC-D and CC-WD is also affected by the number of relays  $N$ . Importantly, the performance gap between CC-D and CC-WD increases for high values of  $N$ . One of the intuitive reason is that CC-D exploits the diversity benefit of source-to-destination link, and utilizes the relays in both the broadcast and relay phase.

Fig. 3 demonstrates the impact of the number of data packets  $K$  on the secrecy performance. It is apparent that the intercept probability can be reduced by increasing the number of data packets  $K$ . This implies that the network security can be improved if the message to be transmitted is segmented into a larger number of shorter data packets. This intrinsic feature of network coding offers a flexibility in designing a physical layer security protocol, such that, it allows to adjust the secrecy level that meets the application requirement.

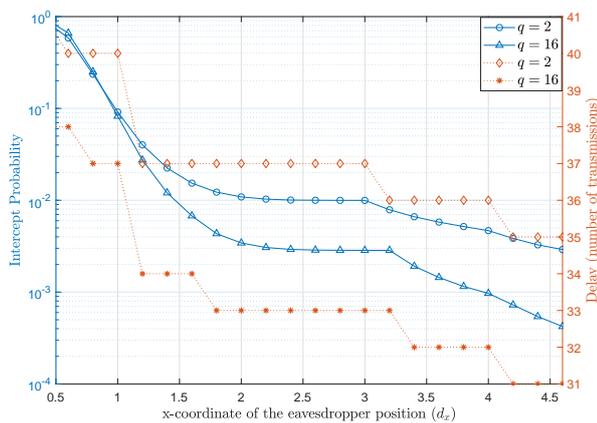
Fig. 4 exhibits the intercept probability as a function of the eavesdropper's position. The intercept probability increases when the distance between E and S decreases. The secrecy is achieved even when E is closer to S. But it is interesting to note that after a certain distance between S and E, the intercept probability decreases sharply. This clearly indicates the importance of superposition of AN



**Fig. 3:** Number of data packets  $K$  versus intercept probability, when  $\tau/K = 0.7$ ,  $q = 4$ .



**Fig. 4:** Eavesdropper position versus the intercept probability, when  $K = 15$ ,  $\tau = 11$ ,  $N = 10$ ,  $q = 16$ .



**Fig. 5:** Intercept probability and decoding delay as a function of Eavesdropper position and the finite field size  $q$ , when  $K = 20$ ,  $\tau = 11$ , and  $N = 10$ .

with the information bearing signal. For example, at SNR = 15 dB when E gets closer to S, the latter allocates more power to AN and less power to the information bearing signal. After a certain position of E (i.e.,  $d_x > 3$ ), S allocates less power to AN and offers more help to D in the accumulation of  $K$  linearly independent coded packets quickly. The results reaffirm the fact that remarkable secrecy gain can be achieved at high SNR values if power is optimally allocated among signals.

Fig. 5 shows the effect of the field size  $q$  on both the intercept probability and the delay performance in terms of the number of coded packets transmissions for D to successfully decode the source message. Clearly, the figure demonstrates that the intercept probability significantly increases when the field size reduces from  $q = 16$  to  $q = 2$ , and a comparatively low excess number of transmissions are required for D to recover the message. The performance gap increases with the increase of the distance between E and S. However, an increase in  $q$  also increases the overhead of RLNC and the decoding complexity of Gaussian elimination [33]. This yields a tradeoff between the complexity and the security performance.

## 5 Conclusion

This paper examined the network security in terms of  $\tau$ -intercept probability. The theoretical expressions match well with the simulation results. A key feature of this work is that RLNC enabled source transmissions leveraged the secrecy benefit of source-to-destination link, where the source transmits both AN and message signals. It has been demonstrated that the proposed framework can provide significantly high secrecy gain at large SNR values. Moreover, the network security can be further improved by increasing the number of relays and the number of data packets over which RLNC is performed.

A Future direction on this topic could involve the use of machine learning techniques for resource allocations in a multi-source multi-relay network while considering the presence of multiple eavesdroppers. Moreover, analysing the network performance corresponding to secrecy diversity order would also be an interesting future direction.

## Acknowledgment

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) under grants EP/R006385/1 and EP/N007840/1.

## 6 References

- 1 T. Liu, J. C. Lui, X. Ma, and H. Jiang, "Enabling relay-assisted D2D communication for cellular networks: algorithm and protocols," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 3136–3150, 2018.
- 2 M. Höyhty, O. Apilo, and M. Lasanen, "Review of latest advances in 3GPP standardization: D2D communication in 5G systems and its energy consumption models," *Future Internet*, vol. 10, no. 1, p. 3, 2018.
- 3 P. Gandotra, R. K. Jha, and S. Jain, "A survey on device-to-device (D2D) communication: Architecture and security issues," *Journal of Network and Computer Applications*, vol. 78, pp. 9–29, 2017.
- 4 X. Liu and N. Ansari, "Green relay assisted D2D communications with dual batteries in heterogeneous cellular networks for IoT," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1707–1715, 2017.
- 5 F. Jiang, Y. Liu, B. Wang, and X. Wang, "A relay-aided device-to-device-based load balancing scheme for multi-tier heterogeneous networks," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1537–1551, 2017.
- 6 T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.
- 7 A. S. Khan and I. Chatzigeorgiou, "Performance analysis of random linear network coding in two-source single-relay networks," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, 2015, pp. 991–996.
- 8 L. Lima, M. Médard, and J. Barros, "Random linear network coding: A free cipher?" in *Proc. IEEE Int. Symp. on Inform. Theory*, Nice, France, Jun. 2007, pp. 546–550.
- 9 K. Bhattad and K. R. Narayanan, "Weakly secure network coding," in *Proc. 1st Workshop on Network Coding, Theory and Applications (NetCod)*, Riva del Garda, Italy, Apr. 2005.
- 10 T. Zhang, H. Wen, J. Tang, H. Song, R. Liao, Y. Chen, and Y. Jiang, "Analysis of the physical layer security enhancing of wireless communication system under the random mobile," *IET Communications*, vol. 13, no. 9, pp. 1164–1170, 2019.
- 11 P. Maji, S. Dhar Roy, and S. Kundu, "Physical layer security in cognitive radio network with energy harvesting relay and jamming in the presence of direct link," *IET Communications*, vol. 12, no. 11, pp. 1389–1395, 2018.
- 12 A. Pittolo and A. M. Tonello, "Physical layer security in power line communication networks: an emerging scenario, other than wireless," *IET Communications*, vol. 8, no. 8, pp. 1239–1247, 2014.
- 13 S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE transactions on wireless communications*, vol. 7, no. 6, pp. 2180–2189, 2008.

- 14 Y.-S. Shiu, S. Y. Chang, H.-C. Wu, S. C.-H. Huang, and H.-H. Chen, "Physical layer security in wireless networks: A tutorial," *IEEE wireless Communications*, vol. 18, no. 2, pp. 66–74, 2011.
- 15 A. S. Khan, A. Tassi, and I. Chatzigeorgiou, "Rethinking the intercept probability of random linear network coding," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1762–1765, Oct. 2015.
- 16 L. Sun, P. Ren, Q. Du, and Y. Wang, "Fountain-coding aided strategy for secure cooperative transmission in industrial wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 12, no. 1, pp. 291–300, Feb. 2016.
- 17 A. S. Khan and I. Chatzigeorgiou, "Opportunistic relaying and random linear network coding for secure and reliable communication," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 223–234, 2018.
- 18 L. Dong, H. Yousefi'zadeh, and H. Jafarkhani, "Cooperative jamming and power allocation for wireless relay networks in presence of eavesdropper," in *ICC*, 2011.
- 19 I. Krikidis, J. S. Thompson, and S. Mclaughlin, "Relay selection for secure cooperative networks with jamming," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5003–5011, Oct. 2009.
- 20 T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.
- 21 A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.
- 22 I. Chatzigeorgiou, I. J. Wassell, and R. Carrasco, "On the frame error rate of transmission schemes on quasi-static fading channels," in *Proc. 42nd Conf. on Inform. Sciences and Systems (CISS)*, Princeton, USA, Mar. 2008, pp. 577–581.
- 23 C. M. Grinstead and J. L. Snell, *Introduction to probability*. American Mathematical Soc., 2012.
- 24 A. Jeffrey and D. Zwillinger, *Table of integrals, series, and products*. Academic Press, 2007.
- 25 M. Geller and E. W. Ng, "A table of integrals of the exponential integral," *Journal of Research of the National Bureau of Standards*, vol. 71, pp. 1–20, 1969.
- 26 A. Bletsas, A. G. Dimitriou, and J. N. Sahalos, "Interference-limited opportunistic relaying with reactive sensing," *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 14–20, 2010.
- 27 È. M. Gabidulin, "Theory of codes with maximum rank distance," *Problems of Information Transmission*, vol. 21, no. 1, pp. 1–12, Jan. 1985.
- 28 B. Handa and S. Mohanty, "On q-binomial coefficients and some statistical applications," *SIAM Journal on Mathematical Analysis*, vol. 11, no. 6, pp. 1027–1035, 1980.
- 29 J. Claridge and I. Chatzigeorgiou, "Probability of partially solving random linear systems in network coding," *IEEE Communications Letters*, vol. 21, no. 9, pp. 1945–1948, Sep. 2017.
- 30 D. R. Jones, "Direct global optimization algorithm," *Encyclopedia of optimization*, pp. 431–440, 2001.
- 31 J. M. Moualeu, W. Hamouda, and F. Takawira, "Intercept probability analysis of wireless networks in the presence of eavesdropping attack with co-channel interference," *IEEE Access*, vol. 6, pp. 41 490–41 503, 2018.
- 32 Q. F. Zhou, Y. Li, F. C. Lau, and B. Vucetic, "Decode-and-forward two-way relaying with network coding and opportunistic relay selection," *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3070–3076, Nov. 2010.
- 33 J. Heide, M. V. Pedersen, F. H. Fitzek, and M. Médard, "On code parameters and coding vector representation for practical RLNC," in *2011 IEEE international conference on communications (ICC)*, 2011, pp. 1–5.