NOMA-Assisted Cognitive Short-Packet Communication with Node Mobility and Imperfect Channel Estimation

Chunli Xia, Zhongwu Xiang, Jin Meng, Hongbo Liu and Gaofeng Pan

Abstract—To address the high-performance requirements of internet of things (IoT) application scenarios, we design a non-orthogonal multiple access (NOMA) assisted cognitive short packet communication system in this paper, in which the primary user is a mobile node and the secondary network shares the spectrum of primary network for transmission subject to transmit power constraints. To characterize the performance of the proposed system, the average block error rate (BLER) and throughput performance of the secondary users are derived and analyzed, and the theoretical results are verified using Monte Carlo simulations. To facilitate the performance improvement of network communication, the analysis and insights on the node's mobility and imperfect channel estimation are provided. The results show that both node's mobility and imperfect channel estimation result in the degradation of system performance. In addition, our examination shows that the increase in distance between primary and secondary networks indeed reduces interference and improves performance. Interestingly, we further find that shorter blocklength enhances the throughput performance for a long-distance scenario between primary transmitter and secondary users. In contrast, the opposite result is obtained for the short-distance scenario, which means that short packet communication is prominent in our proposed system.

Index Terms—Non-orthogonal multiple access, short-packet communications, node's mobility, imperfect channel estimates.

I. INTRODUCTION

In recent years, with the continuous development of the fifth generation (5G) and beyond 5G communication, the network transmission capability has been increasing, which accelerates the development of internet of things (IoT) and provides ubiquitous connectivity for intelligent industries, smart transportation, smart home, etc., making the interconnection of everything possible [1], [2]. Many critical IoT applications require ultra-high reliability and low latency metrics [3]–[5]. While long packet communication, which is widely used, hasn't a relatively high transmission delay, as longer packet block lengths can make communication processes such as coding and decoding consume more time. By contrast, the amount of information transmitted can be controlled to hundreds of bits in short packet communication (SPC), which significantly reduces the transmission delay [6]–[9]. In addition, the transmitted data is usually commands or status feedback information in numerous IoT communication environments, and only a short data packet is required to satisfy the transmission. Therefore, SPC is one of the distinguishing features of IoT. Shannon's capacity formula is no longer applicable in SPC because the law of large numbers is invalid. Therefore, the achievable data rate at SPC needs to be reconsidered. In [10], Polyanskiy et al. pioneered the study of SPC in additive Gaussian white noise (AWGN) channels from an information-theoretic perspective and obtained an approximation of the maximum achievable rate as a complex function of signal-to-noise ratio (SNR), blocklength, and decoding error probability.

Although SPC can optimize the transmission time to meet the requirements of IoT, it cannot solve the problem of scarce spectrum resources due to the wireless connection of large-scale IoT devices. Non-orthogonal multiple access (NOMA) is considered a key multiple access technology for IoT to achieve massive connectivity and plays a vital role in attaining spectrum efficiency improvements and multi-user shared communications [11], [12]. Unlike the orthogonal multiple access (OMA) technology currently used, NOMA can serve more users simultaneously by efficiently using non-orthogonal resources to meet the demand for massive access [13], [14]. In power domain NOMA, simultaneous communication between a base station (BS) and multiple users is established by using superposition coding at the transmitter and successive interference cancellation (SIC) at the receiver [15]. In this case, the BS adjusts the power allocation factor according to the channel gain of different users, and less power is allocated to the stronger user to achieve user equity. In comparison, the weaker user is allocated more power. NOMA offers excellent advantages in terms of enabling large-scale connectivity and user fairness, which has been used in SPC networks to further improve the performance of the system [16]–[18]. In [19], the authors investigated the physical-layer transmission latency reduction enabled by NOMA in SPC and demonstrated the superior performance of NOMA over OMA. A downlink NOMA network with hybrid long-packet and short-packet communications was proposed in [20] and the proposed NOMA network achieved lower outage probabilities.
than the OMA network.

IoT is envisioned to provide massive connectivity to ubiquitous networked devices, and the significant growth in the number of devices will inevitably lead to spectrum scarcity. However, cognitive radio (CR) is a promising technology that can improve spectrum efficiency and enhance the number of connections and coverage of the IoT through collaborative transmission [21], [22]. Combining CR with NOMA, known as CR-NOMA, has great potential to further improve spectral efficiency and enhance wireless network system performance. CR networks can be divided into three main types according to dynamic spectrum access: interweave, overlay and underlay. In underlay mode, secondary users communicate simultaneously with primary users in the same frequency band under interference threshold constraints to ensure the primary user’s quality of service (QoS). While in overlay mode, primary and secondary users share the same spectrum for signal transmission, and in exchange, secondary users must assist primary users in signal relaying. In interweave mode, the secondary users can only transmit when the primary spectrum is not occupied. CR-NOMA has attracted the interest of many researchers. Resource allocation of the downlink CR network with NOMA was studied for maximizing the network throughput in [23]. In [24], the authors investigated the issue of joint power allocation and collaborative beamforming for physical layer security in underlay cognitive radio NOMA relay systems. While in [25], physical layer security in CR-inspired NOMA networks with multiple primary and secondary users was explored.

However, the aforementioned studies mainly focused on the performance analysis of the system all the nodes are stationary. As a matter of fact, considering the case of mobile user is of great research interest in wireless communication networks. Since node movement leads to the time-selectivity of channels, which will bring challenges to the CR-inspired NOMA networks, what impact the mobile user has on the reliability becomes a question worth exploring. In [26], the authors studied an underlay device-to-device communication system in consideration with mobility of primary cellular user. Different from [26], multiple users are considered and NOMA is executed to improve spectrum efficiency in this paper. In addition, the channel estimation error is also a pivotal issue for CR systems, which may reduce the efficiency of spectrum sharing and thus lead to performance degradation of users in the IoT. Based on the above background, we consider a NOMA-assisted underlay network under imperfect channel estimates and node’s mobility in this paper. The contributions of this paper are as follows.

- We consider a NOMA-assisted underlay network with short packet communication where the user in the primary network is mobile, secondary users share the spectrum of the primary network to communicate subject to the interference threshold, and SIC is utilized to eliminate interference and enable multi-user communication simultaneously.
- Imperfect channel estimation is considered, and its impact on the system performance is analyzed. Specifically, it makes it challenging to fully guarantee the performance of the primary user and reduces the transmit power of the secondary transmitter, which in turn lowers the performance of the secondary user.
- The closed-form expressions of average BLER of secondary users are derived and analyzed over Nakagami-\(m\) channels in the proposed system. To further gain insights into the schemes, the average throughput performance is analyzed.
- The mobility of primary user deteriorates the performance of the proposed system to some extent, which is due to the time-selective fading caused by user movement. Under the NOMA protocol, the power allocation factor is positively correlated with the performance of users, and there is a trade-off between secondary users.
- Extensive evaluations of the impact of transmit power, blocklength, and channel parameters on the average BLER and throughput are presented. Results show that increasing distance between the primary and secondary networks improves the performance of the cognitive network. Besides, shorter blocklength will get better throughput performance when the distance between primary transmitter and secondary users is relatively long, while the opposite result is obtained in the closer region.

The rest of the paper is organized as follows. The network model and transmission scheme are introduced in Section II. The closed-form expressions for the average BLER of secondary users are derived at Section III, and the average throughput performance is also analyzed. Next, in Section IV, numerical results are presented to verify the derived closed-form expressions and some insights into system characteristics are presented. Finally, concluding remarks are provided in Section V.

II. NETWORK MODEL AND TRANSMISSION SCHEME

As shown in the Fig.1, we consider an underlay cognitive network in which multiple secondary users share the spectrum resources licensed to the primary user. The primary network consists of a primary transmitter node PT and a primary user P, and the primary user is mobile at a speed of \(\nu_{sp}\) relative to secondary transmitter. The secondary network consists of a primary transmitter node ST and its intended secondary users \(S_k \ (k \in \{1,...,i,...,j,...,K\})\). The presented network model can be applied to multiple IoT applications. Take the smart traffic system as an example. P can be a car, while \(S_k\)
can be some outdoor fixed installed sensor nodes or status monitoring nodes. Besides, the channel coefficients from PT to $S_k$, ST to $P$ ST to $S_i$ are denoted as $h_{ps}$, $h_{sp}$ and $h_{ss}$. And each node in the network is assumed to operate in half-duplex mode and equipped with a single antenna, the AWGN powers at each user and transmitter are $\sigma^2$. In the considered network, all channels are supposed to experience independent block Nakagami-m fading, and the channels between ST and $S_k$, i.e., $h_{ss}$ experience with $E[h_{ss}^2] = \Omega_{ss}$ and fading parameter $m_{ss}$. While due to the movement of P, channel $h_{sp}$ experiences time-selective Nakagami-m fading, and the correlation parameter of the link can be expressed as $\rho_{sp}$, whose value can be evaluated using the standard Jakes model as $\rho_{sp} = J_0(2\pi f\nu_{p} \Omega_{sp})$, where $J_0(\cdot)$ is the zero-order Bessel function of the first kind, $f_c$ is the carrier frequency, $\nu_p$ is the mobile speed, $R$ is the symbol rate and $\epsilon$ is the speed of light. We model the time-selective nature of link using the first order auto-regressive (AR1) process and expand $h_{sp}$ for $m$-th signaling position [27], [28] as

$$h_{sp}(m) = \rho_{sp}^{m-1} h_{sp}(1) + \frac{\rho_{sp}^{m-1} \hat{h}_{sp}(1) + \sqrt{1 - \rho_{sp}^2} \sum_{i=1}^{m-1} \rho_{sp}^{m-1-i} h_{sp}(i)}{h_{sp}(m)},$$

(1)

where $\hat{h}_{sp}(1)$ is the estimated channel coefficient over first signaling position with fading mean $\Omega_{sp}$ and fading parameter $m_{sp}$, $h_{sp}(1)$ is the error in estimation and $h_{sp}(1) \sim \mathcal{CN}(0, \Omega_{sp})$, $h_{sp}(i)$ is the time-varying component and assumed to be zero-mean complex Gaussian process with a density of $\mathcal{CN}(0, \Omega_{sp})$. The channel estimation error, the transmission of secondary according to the above power limit may still exceed the interference threshold. In order to enable the transmission of secondary network satisfies the power constraint as much as possible and reduce the performance loss caused by estimation error, an interference power control factor denoted as $\Psi$ is introduced, and $0 < \Psi < 1$. Therefore, the actual transmit power at ST is expressed as

$$P_s = \min\{P_s', P_{\max}\} = \min\left\{ \frac{\Psi I_{th}}{\rho_{sp}^{2(m-1)}|\hat{h}_{sp}(1)|^2 + \Omega_1}, P_{\max} \right\}$$

(7)

where $\Omega_1 = \left(1 - \frac{2(m-1)}{\rho_{sp}^2} \right) \Omega_{esp}$, $P_{\max}$ is the maximum power available at ST and for simplicity, $|\hat{h}_{sp}(1)|^2$ have been substituted by $|\hat{h}_{sp}|^2$ in the subsequent analysis. Under such a power constraint, the performance of the primary user can be greatly guaranteed, so the subsequent analysis mainly focuses on the performance of secondary users.

According to the principle of SIC, $S_i$ detects its message by treating the signal of $S_j$ as interference. The received signal-to-interference-plus-noise ratio (SINR) for $x_{si}$ at $S_i$ is given by

$$\gamma_i = \frac{P_s \alpha_i |h_{ss_i}|^2}{P_s \alpha_i |h_{ss_i}|^2 + N_p + \sigma^2},$$

(4)

where $N_p$ denotes the average power of $I_p$, and can be expressed as $N_p = \delta \Omega_{ps}$, where $\delta$ is the interference coefficient and $\Omega_{ps}$ is the fading mean of the channel between primary transmitter and secondary users.

As for user $S_j$, before decoding $x_{sj}$, demodulation of the signal $x_{si}$ is firstly executed, the SINR of decoding $x_{si}$ at user $S_j$ is written as

$$\gamma_{j\rightarrow i} = \frac{P_s \alpha_j |h_{ss_j}|^2}{P_s \alpha_j |h_{ss_j}|^2 + \delta \Omega_{ps} + \sigma^2},$$

(5)

If user $S_j$ can decode and eliminate $x_{si}$ perfectly, the SINR of decoding $x_{sj}$ is given by

$$\gamma_{j\rightarrow j} = \frac{P_s \alpha_j |h_{ss_j}|^2}{\delta \Omega_{ps} + \sigma^2},$$

(6)

In addition, interference between the primary and secondary networks is unavoidable during spectrum sharing. The secondary transmitter should control transmit power to reduce the interference at user $P$. This indicates that the power of the secondary network should be limited to $E[h_{sp}(m)]^2 \leq I_{th}$, and maximum allowable power is $P_s = \frac{I_{th}}{\rho_{sp}^{2(m-1)}|\hat{h}_{sp}(1)|^2 + (1 - 2(m-1))\Omega_{sp}}$. However, due to the channel estimation error, the transmission of secondary according to the above power limit may still exceed the interference threshold. In order to enable the transmission of secondary network satisfies the power constraint as much as possible and reduce the performance loss caused by estimation error, an interference power control factor denoted as $\Psi$ is introduced, and $0 < \Psi < 1$. Therefore, the actual transmit power at ST is expressed as

$$P_s = \min\{P_s', P_{\max}\} = \min\left\{ \frac{\Psi I_{th}}{\rho_{sp}^{2(m-1)}|\hat{h}_{sp}(1)|^2 + \Omega_1}, P_{\max} \right\},$$

(7)

where $\Omega_1 = \left(1 - \frac{2(m-1)}{\rho_{sp}^2} \right) \Omega_{esp}$, $P_{\max}$ is the maximum power available at ST and for simplicity, $|\hat{h}_{sp}(1)|^2$ have been substituted by $|\hat{h}_{sp}|^2$ in the subsequent analysis. Under such a power constraint, the performance of the primary user can be greatly guaranteed, so the subsequent analysis mainly focuses on the performance of secondary users.
With the help of (7), we can rewrite the SINR expression (4) for $S_i$ as

$$\gamma_i = \min \left\{ \frac{\Psi I_{th} \alpha_i h_{ss_i}^2}{\Psi I_{th} \alpha_j h_{ss_j}^2 + Z \omega}, \frac{P_{\max} \alpha_i h_{ss_i}^2}{P_{\max} \alpha_j h_{ss_j}^2 + \alpha_j Z} \right\},$$

(8)

where $\omega = \rho_{sp}^2 (m-1) \gamma_{sp}^2 + \Omega_s + \delta \Omega_{ps} + \sigma^2$. And the SINR formulas (5) and (6) for $S_j$ can be rewritten as

$$\gamma_{j-i} = \min \left\{ \frac{\Psi I_{th} \alpha_i h_{ss_i}^2}{\Psi I_{th} \alpha_j h_{ss_j}^2 + Z \omega}, \frac{P_{\max} \alpha_i h_{ss_i}^2}{P_{\max} \alpha_j h_{ss_j}^2 + \alpha_j Z} \right\},$$

(9)

$$\gamma_{j-j} = \min \left\{ \frac{\Psi I_{th} \alpha_i h_{ss_i}^2}{\psi Z}, \frac{P_{\max} \alpha_i h_{ss_i}^2}{Z} \right\}. \quad (10)$$

III. PERFORMANCE ANALYSIS

In this section, we will derive closed-form expressions for the average block error rate (BLER) to look further into the reliability performance.

Firstly, considering Shannon Theory is not applicable in SPC and the coding error is inevitable, the maximal achievable rate can be expressed as [10], [32]

$$R = \log_2 (1 + \gamma) - \sqrt{\frac{V(\gamma) Q^{-1}(\varepsilon)}{N_d}} \ln 2,$$  

(11)

where $N_d$ is the packet length for the data transmission, it is necessary to clarify total packet length of one block is $N_S = N_i + N_d$ and $N_i$ denote the channel uses for training. $\gamma$ is receiving SINR, $\varepsilon$ is BLER. $C(\gamma) = \log_2 (1 + \gamma)$ is the Shannon capacity, $V(x) = 1 - (1 + x)^{-1}$ is the inverse of the Gaussian Q-function $Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$. From (11), we can compute the instantaneous BLER as follows:

$$\varepsilon = Q\left(\ln 2 \frac{\log_2 (1 + \gamma) - R}{\sqrt{V(\gamma)/N_d}}\right).$$

(12)

To facilitate subsequent analysis and derivation, (12) is tightly approximated as [33]

$$\varepsilon = \begin{cases} \frac{1}{2} - g \sqrt{N_d} (\gamma - h) & \gamma \leq A \\ A - \gamma < B , \end{cases} \quad (13)$$

where $g = \frac{\sqrt{2\pi}(2^{2\gamma} - 1)}{2\sqrt{\gamma} N_d}$, $h = \gamma^2 - 1$, $A = h - \frac{1}{2g\sqrt{N_d}}$ and $B = h + \frac{1}{2g\sqrt{N_d}}$. By this approximation, the average BLER can be calculated as

$$E[\varepsilon] = \int_0^B \varepsilon f_\gamma(x) dx = g \sqrt{N_d} \int_A^B F_\gamma(x) dx,$$  

(14)

where $f_\gamma(x)$ is the probability density function (PDF) of $\gamma$, $F_\gamma(x)$ is the cumulative distribution function (CDF) of $\gamma$. It should be noted that the main challenge in performance analysis lies in the derivation and integration of $F_\gamma(x)$.

A. Average BLER Analysis of $S_i$

Combining (8), (13) and (14), we can obtain the average BLER of $S_i$ in a closed-from expression and given in Theorem 1.

**Theorem 1:** Given the blocklength of data transmission $N_i$ and the maximal achievable rate $R_i$, the average BLER of $S_i$, i.e., $E[\varepsilon_i]$ can be expressed as (15), where $U(x) = \{1, x \geq 0\}$, $\gamma = 2^{R_i} - 1$, $A_i = h_i - \frac{1}{2g\sqrt{N_i}}$, $B_i = h_i + \frac{1}{2g\sqrt{N_i}}$, $\Lambda = \frac{\rho_{sp}^2 (m-1)}{\rho_{sp}^2 (m-1)} (\varphi_{th}^{I_{th}} - \Omega_i)$, $\omega_i = \frac{\pi}{n_g}$, $y_i = \frac{A_i + B_i}{2} + \frac{B_i - A_i}{4} x_i$, $x_i = \cos \left(\left(\frac{2\pi}{10}\right)\right)$,$\theta = c + \Lambda \theta$, $\theta = \frac{\sigma_{sp}^2}{\sigma_{sp}^2}$, $\beta_i = M_i \varsigma_i + \kappa_\alpha$.

**Proof:** See Appendix A.

**Remark 1:** It is noteworthy that the expression (15) in Theorem 1 holds only if $\alpha_i / \alpha_j > A_i$, otherwise $E[\varepsilon_i] = 1$. It indicates that larger power allocation factor $\alpha_i$ facilitates reliable transmission for user $S_i$. In particular, $E[\varepsilon_i]$ is a monotonically decreasing function of the power $P_s$ when $I_{th} \to \infty$, which reveals that high power assists the reliable transmission. Moreover, $E[\varepsilon_i]$ decreases with increasing $I_{th}$ when $P_{\max} \to \infty$, which demonstrates the higher power threshold allowed in the primary network contributes to the reliability improvement of $S_i$.

B. Average BLER Analysis of $S_j$

Different from the user $S_i$, the analysis of average BLER of $S_j$ consists of two parts, one is intended for demodulation of the signal of $x_{sj}$, which is pivotal for the successful execution of the SIC, the other is demodulation of its own desired signal $x_{sj}$. Analogously, the first part $E[\varepsilon_{sj-i}]$ can be obtained leveraging with (9), (13) and (14), which is described in Theorem 2, and the closed-form expression of $E[\varepsilon_{sj-j}]$ is given in Theorem 3.

**Theorem 2:** The closed-form expression for $E[\varepsilon_{sj-i}]$ is derived in (16), where $M_j = \frac{m_{ssj}}{\Omega_{ssj} \Omega_{th}}$, $M_i = \frac{m_{ssj}}{\Omega_{ssj} \Omega_{th} P_{\max}}$ and $\beta_j = M_j \varsigma_j + \kappa_\alpha$.

**Proof:** See Appendix B.

**Theorem 3:** Given the blocklength of data transmission $N_j$ and the maximal achievable rate $R_j$, the average BLER $E[\varepsilon_{sj-j}]$ can be expressed as (17), where $g_j = \frac{1}{2\pi(2^{R_j} - 1)}$, $h_j = 2^{R_j} - 1$, $A_j = h_j - \frac{1}{2g_j \sqrt{N_j}}$, $B_j = h_j + \frac{1}{2g_j \sqrt{N_j}}$, $y_j = \frac{A_j + B_j}{2} + \frac{B_j - A_j}{4} x_i$, $x_i = \cos \left(\left(\frac{2\pi}{10}\right)\right)$, $\beta_j = M_j \varsigma_j + \kappa_\alpha$.

**Proof:** See Appendix C.

Since the perfect SIC can not be guaranteed, the aggregate average BLER of decoding of $x_{sj}$ at $S_j$ can be expressed as

$$E[\varepsilon_j] = E[\varepsilon_{sj-i} + (1 - \varepsilon_{sj-j}) \varepsilon_{sj-j}] \approx E[\varepsilon_{sj-i}] + E[\varepsilon_{sj-j}].$$

(18)

Owing to BLER is small in IoT applications generally (e.g., $10^{-3}$ to $10^{-5}$) [34], so the approximation in (18) holds, where $E[\varepsilon_{sj-i}]$ and $E[\varepsilon_{sj-j}]$ are given in Theorem 2 and Theorem 3, respectively.
\[ E[\varepsilon_i] = 1 - \sum_{i=1}^{n_i} \sum_{r=0}^{m_{x,j} - 1} \frac{\omega_i \sqrt{(y_i - A_i)(B_i - y_i)} y_i \sqrt{N_i x^r}}{r!} \left[ \sum_{c=0}^{r} \left( \frac{r}{c} \right) \frac{\kappa_s^c r^c (\Omega_i)^r (\beta_i)^c}{\beta_i^c e^{M_i \Omega_i x^r}} + F[h_{sp}]^2 (\Lambda) M_m r^{-M_m x^r} \right] \]

(15)

\[ E[\varepsilon_{j \rightarrow i}] = 1 - \sum_{i=1}^{n_i} \sum_{r=0}^{m_{x,j} - 1} \frac{\omega_i \sqrt{(y_i - A_i)(B_i - y_i)} y_i \sqrt{N_i x^r}}{r!} \left[ \sum_{c=0}^{r} \left( \frac{r}{c} \right) \frac{\kappa_s^c r^c (\Omega_i)^r (\beta_i)^c}{\beta_i^c e^{M_i \Omega_i x^r}} + F[h_{sp}]^2 (\Lambda) M_m r^{-M_m x^r} \right] \]

(16)

\[ E[\varepsilon_{j \rightarrow j}] = 1 - \sum_{i=1}^{n_i} \sum_{r=0}^{m_{x,j} - 1} \frac{\omega_i \sqrt{(y_i - A_i)(B_i - y_i)} y_i \sqrt{N_i x^r}}{r!} \left[ \sum_{c=0}^{r} \left( \frac{r}{c} \right) \frac{\kappa_s^c r^c (\Omega_i)^r (\beta_i)^c}{\beta_i^c e^{M_i \Omega_i x^r}} + F[h_{sp}]^2 (\Lambda) M_m r^{-M_m x^r} \right] \]

(17)

**Remark 2:** The average BLER of $S_j$ consists of two components, described in (16) and (17), respectively. The expression (16) in **Theorem 2** holds only if $\frac{\alpha_i}{\alpha_j} > A_i$, otherwise $E[\varepsilon_{j \rightarrow i}]$ is 1. It is noteworthy that the larger power allocation factor $\alpha_i$ is necessary for user $S_j$ to perform SIC, but too large $\alpha_i$ may result in a decrease in the $E[\varepsilon_{j \rightarrow j}]$. Besides, the average BLER of $S_j$ ($E[\varepsilon_j]$) is a decreasing function with respect to blocklength, and in order to obtain satisfactory reliability, the blocklength in short packet communication cannot be set too short just to reduce the delay.

**C. Average Throughput**

Furthermore, considering both BLER and transmission rate, the throughput is adopted as the metric to evaluate the system performance with SPC. Mathematically, the average throughput for user $S_i$ is defined as

\[ \bar{T}_k = \left( 1 - \frac{N_i}{N_S} \right) R_k (1 - E[\varepsilon_k]). \]

(19)

**TABLE I**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fading means</td>
<td>$\Omega_{x,j} = -105$dBm, $\Omega_{x,j} = -92$dBm</td>
</tr>
<tr>
<td>Fading parameters</td>
<td>$m_{x,j} = 2$; $m_{x,j} = 2$</td>
</tr>
<tr>
<td>Number of information bits</td>
<td>$K_i = N_i R_i = 200$, $K_j = N_j = 200$</td>
</tr>
<tr>
<td>Average interference from PT</td>
<td>$N_p = -100$dBm</td>
</tr>
<tr>
<td>Average tolerable interference</td>
<td>$T_{th} = -100$dBm</td>
</tr>
<tr>
<td>Estimation error mean</td>
<td>$\Omega_{e,j} = -145$dBm</td>
</tr>
<tr>
<td>Noise power</td>
<td>$\sigma^2 = -121$dBm</td>
</tr>
<tr>
<td>Carrier frequency</td>
<td>$f_c = 2$GHz</td>
</tr>
<tr>
<td>Mobile speed of P</td>
<td>$v_m = 10$m/s</td>
</tr>
<tr>
<td>Maximum transmittable power</td>
<td>$P_{m,x} = 30$dBm</td>
</tr>
<tr>
<td>Interference power control factor</td>
<td>$\lambda = 0.8$</td>
</tr>
<tr>
<td>Power allocation coefficient</td>
<td>$\alpha_i = 0.85$</td>
</tr>
</tbody>
</table>

**IV. NUMERICAL RESULTS**

In this section, numerical and simulation results are presented to analyze the performance of the considered communication network in terms of reliable performance and throughput and to provide some valuable insights into the system design. The effect of key parameters on the system performance is investigated and the effect of node movement is analyzed. For clarity, unless otherwise stated, Table I summarizes the parameter settings used to evaluate the system performance.

Fig. 2 plots the average BLERs of secondary users versus transmit power $P_s$ for different $\Psi$, where the solid lines and discrete markers represent the analysis and simulation results, respectively. Fig. 2 indicates that the simulation makers match well with the analysis curves, which verifies the correctness of the derivation. We can observe from this figure that the reliability of $S_i$ and $S_j$ enhances as transmit power $P_s$ increases in low power region, while reliability cannot be further improved because the power reaches the interference threshold of the primary network in the high-power region.
This reveals increasing transmit power improves reliability performance. On the other hand, as $\Psi$ decreases, the maximum power allowed to be transmitted at the ST decreases, which leads to an increase in the BLER of secondary users in the high power region. In contrast, the inset of Fig. 2 depicts the probability of secondary network transmission causing interference to the primary network (i.e., PU, which is defined as the probability that the secondary network transmission exceeds the interference threshold of the primary network and affects the transmission quality of the primary network) shows a decreasing trend. It is clear that the $\Psi$ value leads to a performance trade-off between primary and secondary users due to the fact better reliability performance of secondary network in the underlay CR mode requires sacrificing the primary network to bear the interference. Therefore, when the channel cannot be perfectly estimated, special attention should be paid to further control of the secondary transmit power. As also can be seen from the inset of Fig. 2, PU remains at zero in perfect channel estimation case ($\Omega_{s,p}=0$). And the channel estimation error results in an obvious increase in PU. Therefore, performance deterioration of the primary and secondary networks deepens as the channel estimation error increases, and it also predicts that it is crucial to choose the estimation method with high accuracy in channel estimation and focus on improving the channel estimation capability.

In Fig.3, we investigate the average BLER at secondary users versus $I_{th}$ for different power allocation factors. It can be seen from the Fig.3 that when the interference threshold $I_{th}$ is at a very low value, the average BLER of secondary users is 1. This is due to the fact that the maximum power value allowed for ST is too low to support normal and reliable communication at this time. In addition, as the allowable interference threshold increases, the average BLER decreases. This also shows that high power is beneficial to communication quality. However, it should be noted that although the increase in the interference threshold brings a significant improvement in the reliable communication of the secondary users, it also increases the detrimental effect on the transmission of the primary network. Therefore, it is necessary to consider the compromise between the primary and secondary users in the selection of the interference threshold. In addition, from the formula of the ST transmission power under the interference threshold limit, i.e., expression (7), it can be found that the lower the fading mean value of $h_{sp}$ link is more favorable to meet the power limit, so the longer distance between ST and the primary user will enhance the cognitive gains, and the communication reliability of the network can be further improved. Furthermore, as the power allocation factor $\alpha_i$ decreases, the average BLER of $S_i$ increases, while the average BLER of $S_j$ decreases. This observation can be explained by the fact that the reduction in $\alpha_i$ indicates an increase in the power allocated to user $S_i$ and a decrease in the power allocated to $S_j$. This also confirms the vital connection between high power and reliable communication.

Fig. 4 depicts average BLER at secondary users versus $\Omega_{ps}$ for different mobile speeds $\nu_p$. One can see from Fig. 4 that the average BLER increases when the fading mean of the link from the primary transmitter to secondary users increases. This is because the interference degree from the primary network increases when the $\Omega_{ps}$ enlarges, and the BLER naturally grows accordingly. Therefore, it reveals that the further the secondary users are from the primary transmitter, the communication reliability will be enhanced. Combining the previous findings, we can infer more conclusively that increasing the distance between the primary and secondary networks improves the performance of the cognitive network. Additionally, the mobility of primary user results in a decrease in the communication reliability of the secondary users, and the average BLER increases further with the increase of mobile speed. It is due to the time-selective fading caused by the primary user’s movement, and it will deepen as the speed increases, which corresponds to a decrease in the channel correlation coefficient in theoretical analysis. Although node mobility deteriorates the cognitive effect of the system, it also
reflects the important potential of sharing spectrum resources of mobile nodes in large-scale IoT.

In the subsequent simulations, the throughput performance of the secondary network is mainly analyzed. Fig. 5 illustrates the average throughput at secondary users versus $N_d$ for different training blocklength $N_t$. Herein, we set $N_i = N_j = N_d$, i.e., the data blocklength of two secondary users are equal, which is common and reasonable in IoT application scenarios. As shown in Fig. 5, as $N_d$ increases, the average throughput of secondary users rises first, reaching a maximum at $N_d = 90$ and then falls. This is due to the fact that BLER decreases as blocklength increases in the rising period, leading to a rise in throughput. And the later decrease mainly depends on the relation $R_m = K_m/N_m$ $(m \in \{i, j\})$, $R$ decreases as $N_d$ becomes further longer, the average throughput thereby decreases continuously. As such, the optimal throughput can be obtained by adjusting the data blocklength $N_d$. Furthermore, throughput performance deteriorates to some extent as $N_t$ enlarges, demonstrating that shorter training sequence is more beneficial to throughput performance. But, it is noteworthy that a shorter training sequence may affect the accuracy of the channel estimation.

Fig. 6 plots the average throughput at secondary users versus $P_s$ for different power allocation factor $\alpha_i$. We can observe from the figure that the throughput performance of $S_i$ and $S_j$ can be improved with the increase in $P_s$, and reaching a plateau value eventually limited by the power interference threshold. To this end, the improvement of average throughput can be enhanced by boosting the power of ST and relaxing the power constraints on the secondary network. Furthermore, as $\alpha_i$ increases, the throughput of $S_i$ grows up, whereas the throughput of $S_j$ decreases, which means that more power allocated to the user is helpful in achieving high throughput, and the allocation factor allows for a trade-off in throughput performance between $S_i$ and $S_j$.

In Fig. 7, we consider the average throughput at secondary users versus $I_{th}$ for different mobile speeds of $P$. As the maximum allowable interference threshold of the primary network $I_{th}$ increases, the throughput of secondary users enlarges and eventually reaches a plateau value of constant, which is due to the fact that the maximum transmission power of secondary transmitter has been achieved, depending on the performance of transmission device itself. This indicates that the power limitation condition set by the primary network has a great impact on the performance of the secondary network. The threshold value is higher, better performance for the secondary users will be achieved, but the higher threshold may be detrimental to the communication of the primary user. In addition, compared with immobility of node $P$ ($v_p = 0 m/s$), the movement of $P$ deteriorates the throughput performance, because the mobility of node brings the Doppler effect, which causes the time-selective fading of the transmission channel. Moreover, the increase in $v_p$ leads to more obvious time-varying characteristics of the channel, so the performance is...
We investigated a NOMA-assisted underlay network with SPC, in which the primary user is in mobility, and the secondary network shares the spectrum of the primary network for transmission on the stringent interference threshold imposed by the primary user. We provided the analytical approximation closed-form expressions of secondary users for average BLER and analyzed the throughput performance. Finally, the effects of node’s movement, transmit power, blocklength, imperfect channel estimation, and channel parameters on the reliability and throughput performance of the cognitive network are revealed and relevant key insights are presented. Our analysis shows that both node’s mobility and imperfect channel estimation deteriorate the system performance. Moreover, when the secondary network shares the spectrum of the primary network, the longer distance will reduce the interference level and improve the network performance. It is noteworthy that shorter blocklength will get better throughput performance when the link between primary transmitter and secondary users is relatively long, and the opposite result is obtained in the closer range region.

V. CONCLUSIONS

Further reduced.

Fig.8 depicts average overall throughput of secondary network versus $\Omega_{ps}$ for different blocklength $N_S$. Similar to the premise of the previous discussion, we assume that the blocklength of secondary users is equal and both the training sequence $N_t$ and data length $N_d$ are equal, respectively. More precisely, when $N_S$ is reduced from 220 to 110, the data lengths are 200,150,100 respectively. We can see from this figure that when the fading mean of $h_{ps}$ link gradually increases, the overall throughput is showing a decreasing trend, and this phenomenon can be explained by the fact that if the distance between the primary transmitter with secondary users is closer, secondary users will experience more severe interference. Besides, in the low $\Omega_{ps}$ region, longer blocklength corresponds to the smaller throughput, which is a reasonable phenomenon through the analysis of Fig.5, but as the $\Omega_{ps}$ becomes larger, average overall throughput when $N_S$ is equal to 110 decreases sharply. From the insets, it can be found that the decrease in total throughput is mainly caused by the faster grow down in $S_t$, with increasing $\Omega_{ps}$, while $S_j$ shows a significant decrease only in the high $\Omega_{ps}$ region. Moreover, in the high $\Omega_{ps}$ region, it is found that the shorter blocklength will get the better throughput performance, which means that in the low $\Omega_{ps}$ region, the longer blocklength will get better throughput performance, while the opposite overcome is obtained in the high $\Omega_{ps}$ region. Remarkably, lower $\Omega_{ps}$ corresponds to the longer distance between primary transmitter and secondary users. Intuitively, shorter blocklength has advantage in throughput performance for a long distance scenario between primary transmitter and secondary users, in conjunction with the previous finding that the increasing distance between primary and secondary networks facilitates the performance improvement of cognitive network, we can observe that short packet communication has a prominent performance in proposed network.

APPENDIX A

PROOF OF THEOREM 1

Utilizing (13), the instantaneous BLER $\varepsilon_i$ can be approximated as

$$
\varepsilon_i = \begin{cases} 
\frac{1}{2} - g_i \sqrt{\frac{1}{N_t} (\gamma_i - h_i)} & \gamma_i \leq A_i \\
0 & \gamma_i > B_i 
\end{cases} \quad \text{for } i = 1, 2, \ldots, N 
$$

(20)

Utilizing (14) and (20), the average BLER of $S_i$ can be expressed as

$$
E[\varepsilon_i] = \int_0^\infty \varepsilon_i f_{\gamma_i}(x) dx = g_i \sqrt{\frac{1}{N_t}} \int_{A_i}^{B_i} F_{\gamma_i}(x) dx 
$$

(21)

For $S_i$, it should be noted that the (21) equation holds only if $\frac{\alpha_i}{m_i} > A_i$, otherwise $E[\varepsilon_i]$ is 0. $F_{\gamma_i}(x)$ can be further expressed as

$$
F_{\gamma_i}(x) = P\{\gamma_i \leq x\} = P\left\{\psi I_{th}\alpha_i |h_{ss_i}|^2 \leq \frac{P_{\max} \alpha_i |h_{ss_i}|^2}{xZ} \right\} 
$$

$$
= P\left\{|h_{ss_i}|^2 \leq \frac{xZ}{\psi I_{th}(\alpha_i - x \alpha_j)} \right\} 
$$

$$
= P\left\{|h_{ss_i}|^2 \leq \frac{\Delta_1}{\psi I_{th}(\alpha_i - x \alpha_j)} \right\} 
$$

$$
+ P\left\{|h_{ss_i}|^2 \leq \frac{\Delta_2}{\psi I_{th}(\alpha_i - x \alpha_j)} \right\} 
$$

(22)

where $\psi = \rho_{sp}^{2(m-1)} |\hat{h}_{sp}|^2 + \Omega_1$. The expression in (22) consists of two parts, i.e., $\Delta_1$ and $\Delta_2$. Invoking the probability density function (PDF) of $|h_{ss_i}|^2$, i.e., $f_{|h_{ss_i}|^2}(x) =$
\[
\sum_{c=0}^{r} \binom{r}{c} \kappa_g c^c(\Omega^c) e^{-c \left( \frac{M_i c h_s}{\alpha_i - \zeta_{uy}} + \kappa_a \right)} \Gamma \left( \theta, \Lambda \left( \frac{M_i c h_s}{\alpha_i - \zeta_{uy}} + \kappa_a \right) \right).
\]

With the help of Gauss-Chebyshev integration, the integral \[B_i \int \Delta 1 dx\] can be easily derived as (31).

Finally, invoking (29) and (31) into (21) and further with some manipulations, the closed-form expression of \( E[\varepsilon_i] \) is obtained in (15).

**Appendix B**

**Proof of Theorem 2**

Considering BLER as a function with respect to \( N_d, R \) and \( \gamma \) as shown in (12), we can find \( E[\varepsilon_\gamma] \) is easily derived impelled by the preceding discussion with \( E[\varepsilon_i] \). Specifically, the expression of \( \varepsilon_\gamma \) shown in (9) is similar to \( \varepsilon_i \) shown in (8), since the signal decoded by both user 1 and 2, achievable rate \( R \) and packet length \( N_d \) are naturally equal. Thus, \( E[\varepsilon_\gamma] \) will have the similar expression as \( E[\varepsilon_i] \). We can obtain the closed-form expression of \( E[\varepsilon_\gamma] \) as shown in (16) by replacing \( |h_{ss}|^2 \) with \( |h_{ss^j}|^2 \), \( m_{ss} \) with \( m_{ss^j} \), \( \Omega_{ss} \) with \( \Omega_{ss^j} \).

**Appendix C**

**Proof of Theorem 3**

Leveraging (13), the instantaneous BLER \( \varepsilon_\gamma \) can be approximated as

\[
\varepsilon_\gamma = \begin{cases} \frac{1}{2} - g_j \sqrt{\frac{1}{N_j}} (\gamma_\gamma - h_j) & \gamma_\gamma - h_j \leq A_j, \\ \frac{1}{2} & \gamma_\gamma - h_j > B_j \end{cases}
\]

Utilizing (14) and (32), the average BLER can be expressed as

\[
E[\varepsilon_\gamma] = \int_0^\infty \varepsilon_\gamma f_{\gamma_\gamma} \left( \varepsilon \right) dx = g_j \sqrt{\frac{1}{N_j}} \int_{A_j}^{B_j} F_{\gamma_\gamma} (x) dx,
\]

where \( F_{\gamma_\gamma} (x) \) denotes the CDF of \( \gamma_\gamma \) and it can be written as

\[
F_{\gamma_\gamma} (x) = P \left\{ \varepsilon_\gamma \leq x \right\} = P \left\{ \min \left\{ \psi_{h_s, \Omega_{ss^j}}, m_{ss^j} - c m_{ss} \right\} \right\} \leq x \}
\]

\[
= P \left\{ \frac{x}{\Psi_{h_s, \Omega_s}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\} \leq x \}
\]

\[
= P \left\{ \frac{1}{\sqrt{2}}, \frac{x}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\} \leq \Lambda \}
\]

\[
= P \left\{ \frac{1}{\sqrt{2}}, \frac{x}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\} < \Lambda \}
\]
\[ \int_{A_j} \nabla 1 dx = \left[ -1 + F_{\hat{h}_{sp}}^2 (\Lambda) \right] \left( B_j - A_j \right) - U \left( B_j - A_j \right) \sum_{r=0}^{m_{sp} - 1} \omega_1 \sqrt{(y_j - A_j) (B_j - y_j)} \sum_{r=0}^{m_{sp} - 1} (M_k y_j) \frac{1}{r!} e^{-M_k \Omega_j y_j} \right] \times \frac{r}{c} \kappa_g \frac{c}{(\Omega_1)^{r-c}} (M_k y_j \varsigma + \kappa_a) \Gamma \left( \frac{1}{r} \Lambda (M_k y_j \varsigma + \kappa_a) \right). \] (37)

Finally, invoking (37) and (39) into (34) and further with some mathematical operations, the closed-form expression of \( E [c_{j \rightarrow j}] \) is obtained in (17).

### References


\[ \int \nabla 2 d x = F_{s, p}^2 (A_j) (B_j - A_j) - F_{s, p}^2 (A_j) U (B_j - A_j) \sum_{i=1}^{n} \sum_{r=0}^{m_{s, p} - 1} \omega_i \sqrt{(y_j - A_j) (B_j - y_j) (M_j y_j)} \]

(39)


