FAS-assisted NOMA Short-Packet Communication Systems

Jianchao Zheng, Tuo Wu, Xiazhi Lai, Cunhua Pan, Maged Elkashlan, Kai-Kit Wong, *Fellow, IEEE*

Abstract—In this letter, we investigate a fluid antenna system (FAS)-assisted downlink non-orthogonal multiple access (NOMA) for short-packet communications. The base station (BS) adopts a single fixed antenna, while both the central user (CU) and the cell-edge user (CEU) are equipped with a FAS. Each FAS comprises N flexible positions (also known as ports), linked to N arbitrarily correlated Rayleigh fading channels. We derive expressions for the average block error rate (BLER) of the FAS-assisted NOMA system and provide asymptotic BLER expressions. We determine that the diversity order for CU and CEU is N, indicating that the system performance can be considerably improved by increasing N. Simulation results validate the great performance of FAS.

Index Terms—Block error rate (BLER), flexible-position antenna, fluid antenna system (FAS), non-orthogonal multiple access (NOMA), short-packet communication.

I. Introduction

The rapid advances in wireless technologies signal a new chapter in wireless communications. One technology leading to this change is non-orthogonal multiple access (NOMA), which offers improvements in energy and spectral efficiency [1], [2], [3], [4]. Two users are considered as the implementation complexity for user pairing is acceptable while delivering the benefits of NOMA as suggested in [4], [5]. The work on the downlink NOMA network with more than two users remains challenging deserves future work to determine its practicality. Recognizing the pressing need for reduced communication delay, Polyanskiy *et al.* introduced the concept of short-packet communications [6], where the block error rate (BLER) has become an important performance metric [5], [6], [7].

Besides, there have been some exciting developments in the field of antenna and radio frequency (RF) technologies. One innovative idea is to use flexible materials such as liquid metal to create unique antenna designs. More recently, pixel-based switchable antennas further provide delay-free reconfiguration of antennas. These novel designs are embraced in the concept of fluid antenna system (FAS) which includes all forms of movable and non-movable flexible-position antennas [8].

In [9], the ergodic capacity for a single-antenna FAS was first studied while [10] deepened the outage probability anal-

(Corresponding authors: Tuo Wu, Cunhua Pan and Kai-Kit Wong).

J. Zheng is with the School of Computer Science and Engineering, Huizhou University, Huizhou 516000, China (E-mail: zhengjch@hzu.edu.cn.). T. Wu and M. Elkashlan are with the School of Electronic Engineering and Computer Science at Queen Mary University of London, London E1 4NS, U.K. (Email:{tuo.wu, maged.elkashlan}@qmul.ac.uk). X. Lai is with the School of Computer Science, Guangdong University of Education, Guangzhou, Guangdong, China (E-mail: xzlai@outlook.com). C. Pan is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China. (e-mail: cpan@seu.edu.cn). K. K. Wong is with the Department of Electronic and Electrical Engineering, University College London, WC1E 6BT London, U.K., and also with the Yonsei Frontier Laboratory and the School of Integrated Technology, Yonsei University, Seoul 03722, South Korea (e-mail: kat-kit.wong@ucl.ac.uk).

ysis. Later work in [11] then offered a closed-form expression that adeptly characterizes the spatial correlation across all the positions (known as ports), directly linking it to the size of the fluid antenna. In [12], Khammassi *et al.* proposed better models to account for the impact of spatial correlation. Then [13] revisited the performance analysis of FAS and presented new results to quantify the diversity order of FAS. Broadening the scope of FAS applications, Wong *et al.* also delved into its potential for multiple access, as elucidated in [14], [15].

Utilizing FAS at mobile devices even with limited space offers a great spatial diversity gain, which matches perfectly with the principles of short-packet communication. Furthermore, when compared to the traditional orthogonal multiple access (OMA) system, combining FAS with NOMA will further elevate the performance. However, the integration of FAS into NOMA short-packet systems remains unexplored. To fill this research gap, this letter delves into exploring the potential of the FAS-assisted downlink NOMA short-packet system. The main contributions can be summarized as follows:

- We formulate a FAS-aided downlink NOMA short-packet system model consisting of a base station (BS), a central user (CU), and a cell-edge user (CEU). Both CU and CEU are equipped with fluid antennas, enabling dynamic repositioning amongst N preset locations (or ports). The BS transmits short packets of N_c bits to the CU and N_e bits to the CEU, both with a blocklength of L.

- We theoretically analyze the average BLER at the CU and the CEU based on the Jake's model in [13] representing rich scattering scenarios using both the linear approximation and Gauss-Chebyshev quadrature. Additionally, we present the asymptotic expressions for the BLER by using the first-order Riemann integral approximation.

- The results indicate that the diversity order for both CU and CEU is N. Consequently, a large N can significantly improve the system performance.

- We derive the minimum common blocklength at high SNRs. For comparison, the expressions for the minimum common blocklength without FAS and the minimum sum blocklength in OMA systems are derived in closed form.

- The simulation results corroborate the correctness of the derivations of our theoretical analysis.

II. System Model

Consider a FAS-assisted downlink NOMA communication system that includes a BS, a CU, and a CEU. All terminals have a single antenna. Both CU and CEU use a fluid antenna to receive signals from the BS while the BS transmits using a fixed antenna. In this system, the BS transmits packets of N_c bits with a blocklength of L to the CU and packets of N_e bits

with the same blocklength to the CEU. The fluid antennas at the CU and CEU are assumed to always switch their positions to the optimal position amongst N preset positions within a linear space of size $W\lambda$, where λ represents the radiation wavelength. The time delay for port switching is negligible, when pixel-based switchable antennas are considered [8].

Due to the close proximity of the ports within the FAS, they exhibit a significant spatial correlation. Relying on the Jake's model [13], the spatial correlation between the m-th and n-th ports can be expressed as

$$J_{m,n} = \sigma^2 J_0\left(\frac{2\pi(m-n)W}{N-1}\right), \tag{1}$$

where σ^2 represents the large-scale fading effect and $J_0(\cdot)$ denotes the zero-order Bessel function of the first kind.

To better investigate the performance of the overall system, we define the correlation matrix as

$$\mathbf{J} = \begin{bmatrix} J_{1,1} & \ldots & J_{1,N} \\ \vdots & \ddots & \vdots \\ J_{N,1} & \ldots & J_{N,N} \end{bmatrix}. \tag{2}$$

For the elements within $\mathbf{J}$ in (2), we assume that $J_{m,n} = J_{n,m}$. Consequently, leveraging the eigenvalue decomposition, the matrix $\mathbf{J}$ can be decomposed as $\mathbf{J} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, where $\mathbf{U}$ is an $N \times N$ matrix with its n-th column, represented as $\mathbf{u}_n$. Concurrently, $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_N)$ is an $N \times N$ diagonal matrix, where the n-th diagonal entry corresponds to the eigenvalue of $\mathbf{u}_n$. For the sake of analytical convenience, it is assumed that the eigenvalues in $\boldsymbol{\Lambda}$ are sequenced in a descending order, i.e., $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$.

Concerning the spatial correlation mentioned above, the complex channel at the n-th port for the CU and CEU, denoted as $g_n^{(c)}$ and $g_n^{(e)}$, can be formulated as

$$g_n^{(c)} = \sum_{m=1}^{N} u_{n,m}\sqrt{\lambda_m}\omega_{cm}, \tag{3}$$

$$g_n^{(e)} = \sum_{m=1}^{N} u_{n,m}\sqrt{\lambda_m}\omega_{em}, \tag{4}$$

where $u_{n,m}$ represents the (n,m)-th component of $\mathbf{U}$. Besides, it is assumed that $\omega_m^{(c)} = a_m^{(c)} + jb_m^{(c)}$ and $\omega_m^{(e)} = a_m^{(e)} + jb_m^{(e)}$. The terms $a_m^{(c)}$, $b_m^{(c)}$, $a_m^{(e)}$, and $b_m^{(e)}$, $\forall m$ are assumed to be independent and identically distributed (i.i.d.) Gaussian random variables, each with a mean of zero and a variance of $\frac{1}{2}$, respectively.

Furthermore, we assume that the employed FAS within both the CU and the CEU is equipped with a single RF chain. Consequently, at any given time, only one port of the FAS can be active for communication with one user. Hence, the signals received at the n-th port for the CU and the CEU can be expressed as

$$y_n^{(c)} = g_n^{(c)} d_c^{-\frac{a}{2}}\left(\sqrt{\alpha_c P}s_c + \sqrt{\alpha_e P}s_e\right) + z_n^{(c)}, \tag{5}$$

$$y_n^{(e)} = g_n^{(e)} d_e^{-\frac{a}{2}}\left(\sqrt{\alpha_c P}s_c + \sqrt{\alpha_e P}s_e\right) + z_n^{(e)}, \tag{6}$$

where $n \in \{1, 2, \ldots, N\}$. The terms $z_n^{(c)}$ and $z_n^{(e)}$ denote the zero-mean complex Gaussian noise at the n-th port with a

variance of δ^2. The symbols s_c and s_e represent the information symbol transmitted from the BS to the CU and CEU, respectively. We assume that $\mathbb{E}[|s_c|^2] = 1$ and $\mathbb{E}[|s_e|^2] = 1$. The factors α_c and α_e represent power allocation and satisfy $\alpha_c + \alpha_e = 1$. Also, P denotes the BS transmit power, while d_c and d_e represent the distances from the BS to the CU and CEU, respectively, and a is the path loss exponent.

To optimize the system performance, it is assumed that users can dynamically adjust the fluid antenna to be connected to the most advantageous port. For the clarity of analysis, the overall communication performance is assessed based on the maximum values of $g_n^{(c)}$ and $g_n^{(e)}$, expressed as

$$\left|g_{\text{FAS}}^{(c)}\right| = \max\left\{|g_1^{(c)}|, |g_2^{(c)}|, \ldots, |g_N^{(c)}|\right\}, \tag{7}$$

$$\left|g_{\text{FAS}}^{(e)}\right| = \max\left\{|g_1^{(e)}|, |g_2^{(e)}|, \ldots, |g_N^{(e)}|\right\}. \tag{8}$$

Upon receiving the signal $y_n^{(c)}$, the CU prioritizes decoding the signal s_e by regarding s_c as interference. We can obtain the signal-to-interference-plus-noise ratio (SINR) for decoding s_e at the CU as

$$\gamma_{ce} = \frac{\alpha_e\rho d_c^{-a}|g_{\text{FAS}}^{(c)}|^2}{\alpha_c\rho d_c^{-a}|g_{\text{FAS}}^{(c)}|^2 + 1}, \tag{9}$$

with $\rho = \frac{P}{\delta^2}$ representing the ratio between the transmit power and noise variance. As given in [6], the BLER for decoding s_e at the CU can be approximated as

$$\epsilon_{ce} \approx \Psi(\gamma_{ce}, N_e, L) \triangleq Q\left(\frac{C(\gamma_{ce}) - N_e/L}{\sqrt{V(\gamma_{ce})/L}}\right), \tag{10}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} e^{-\frac{t^2}{2}} dt$, $C(\gamma) = \log_2(1+\gamma)$, and $V(\gamma) = (\log_2 e)^2 \cdot (1 - (1+\gamma)^{-2})$ denote the Gaussian Q-function, Shannon capacity formula, and the channel dispersion, respectively.

Upon successfully decoding s_e at the CU, the CU subsequently decodes s_c. The instantaneous signal-to-noise ratio (SNR) for decoding s_c can be expressed as

$$\gamma_{cc} = \alpha_c\rho d_c^{-a}|g_{\text{FAS}}^{(c)}|^2. \tag{11}$$

Consequently, the instantaneous BLER when decoding s_c at the CU can be approximated by

$$\epsilon_{cc} \approx \Psi(\gamma_{cc}, N_c, L) \triangleq Q\left(\frac{C(\gamma_{cc}) - N_c/L}{\sqrt{V(\gamma_{cc})/L}}\right). \tag{12}$$

Therefore, the BLER of decoding s_c at the CU is given by

$$\epsilon_c = \epsilon_{ce} + (1 - \epsilon_{ce})\epsilon_{cc}. \tag{13}$$

For the CEU, after receiving the signal $y_n^{(e)}$, it decodes s_e by treating s_c as interference. The instantaneous SINR for decoding s_e at the CEU can be represented as

$$\gamma_e = \frac{\alpha_e\rho d_e^{-a}|g_{\text{FAS}}^{(e)}|^2}{\alpha_c\rho d_e^{-a}|g_{\text{FAS}}^{(e)}|^2 + 1}. \tag{14}$$

According to [6], the BLER for decoding y_e at the CEU is found as

$$\epsilon_e \approx \Psi(\gamma_e, N_e, L) \triangleq Q\left(\frac{C(\gamma_e) - N_e/L}{\sqrt{V(\gamma_e)/L}}\right). \tag{15}$$

III. THEORETICAL ANALYSIS OF AVERAGE BLER

In this section, we will derive the theoretical expression of the average BLER to decode s_e at the CEU and that to decode s_c at the CU. The theoretical analysis of the average BLER to decode s_c at the CU is obtained by

$$\mathbb{E}[\epsilon_c] = \mathbb{E}[\epsilon_{cc}] + \mathbb{E}[\epsilon_{ce}] - \mathbb{E}[\epsilon_{cc}\epsilon_{ce}], \tag{16}$$

where $\mathbb{E}[\epsilon_{cc}]$ is expressed as

$$\mathbb{E}[\epsilon_{cc}] \approx \int_0^\infty \Psi(\gamma_{cc}, N_c, L) f_{\gamma_{cc}}(t) dt. \tag{17}$$

To further derive $\mathbb{E}[\epsilon_{cc}]$, we first introduce the linear approximation of $\Psi(\gamma_{cc}, N_c, L)$ as [5], [7]

$$\Psi(\gamma_{cc}, N_c, L)$$
$$= \begin{cases} 1, & \gamma_{cc} \leq v_{N_c,L}, \\ \frac{1}{2} - \delta_{N_c,L}\sqrt{L}(\gamma_{cc} - \beta_{N_c,L}), & v_{N_c,L} < \gamma_{cc} < u_{N_c,L}, \\ 0, & \gamma_{cc} \geq u_{N_c,L}, \end{cases} \tag{18}$$

where $\beta_{N_c,L} = 2^{\frac{N_c}{L}} - 1$, $\delta_{N_c,L} = (2\pi(2^{\frac{2N_c}{L}} - 1))^{-\frac{1}{2}}$, $v_{N_c,L} = \beta_{N_c,L} - \frac{1}{2}\delta_{N_c,L}^{-1}L^{-\frac{1}{2}}$, and $u_{N_c,L} = \beta_{N_c,L} + \frac{1}{2}\delta_{N_c,L}^{-1}L^{-\frac{1}{2}}$. Substituting (18) into (17), we have

$$\mathbb{E}[\epsilon_{cc}] \approx \delta_{N_c,L}\sqrt{L} \int_{v_{N_c,L}}^{u_{N_c,L}} F_{\gamma_{cc}}(\tau) d\tau, \tag{19}$$

where $F_{\gamma_{cc}}(\tau)$ is the cumulative distribution function (CDF) of γ_{cc}. However, direct calculation of the integral for $F_{\gamma_{cc}}(\tau)$ proves to be challenging. To further derive the expression of $\mathbb{E}[\epsilon_{cc}]$, we introduce the following lemma.

Lemma 1: Let $D = \det(\mathbf{J})$, $P_t = \frac{(-K_{m,n})^{s_t^*}}{s_t^*! D}$, and $P_n = \frac{1}{2}\left(\frac{K_{n,n}}{D}\right)^{-\frac{\bar{s}_n}{2} - \frac{1}{2}}\left[\Gamma\left(\frac{1+\bar{s}_n}{2}\right) - \Gamma\left(\frac{1+\bar{s}_n}{2}, \frac{K_{n,n}d_c^a\tau}{\alpha_c\rho D}\right)\right]$, the CDF of γ_{cc} is given by

$$F_{\gamma_{cc}}(\tau) = \sum_{s_1=0}^{s_0} \cdots \sum_{s_T=0}^{s_{T-1}} \frac{g(\mathbf{s}^*)}{\pi^N D} \prod_{t=1}^T P_t \prod_{n=1}^N P_n, \tag{20}$$

where $T = \frac{N(N-1)}{2}$, and

$$g(\mathbf{s}^*) = \frac{1}{2^{\sum_{t=1}^T s_t^*}} \sum_{\mathbf{v} \in \mathcal{V}} \left[\prod_{t=1}^T \binom{s_t^*}{v_t}\right] (2\pi)^N \prod_{i=1}^N \mathbf{1}_{\{\Delta_i=0\}}, \tag{21}$$

and $\bar{s}_n = \sum_{i=1}^N S_{n,i}^* + \sum_{i=1}^{n-1} S_{i,n}^* + 1$, $S_{i,n}^*$ denotes the (i,n)-th entry of $\mathbf{S}^*$. Besides, $\mathbf{S}^*$ is given by

$$\mathbf{S}^* = \begin{bmatrix} 0 & s_1^* & \cdots & s_{N-1}^* \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & & s_T^* \end{bmatrix}, \tag{22}$$

where $s_t^* = s_t - s_{t+1}$ with $s_{T+1} = 0$. Additionally, s_0 is a finite constant which must be sufficiently large for the approximation to be accurate. The index t is defined as $t = n + (m-1)N - \frac{m(m+1)}{2}$, where $m < n$. The indices m and n can be derived from t, such that m is the minimal integer m' satisfying the condition $\sum_{i=1}^{m'}(N-i) > t$, and $n = t - (m-1)N + \frac{m(m+1)}{2}$.

Furthermore, $\mathbf{v} = [v_1, \ldots, v_T]^T$ and $\mathcal{V}$ represents the set of all possible permutations. The variable Δ_i is given by $\Delta_i =$

$\sum_{n=1}^N G_{i,n} - \sum_{n=1}^N G_{n,i} - G_{i,i}$. The entry $K_{m,n}$ is the (m,n)-th element of $\mathbf{K}$, where $\mathbf{K}$ is the co-factor of $\mathbf{J}$. Similarly, $G_{m,n}$ is the (m,n)-th element of $\mathbf{G}$, which is defined as

$$\mathbf{G} = \begin{bmatrix} 0 & \gamma_1 & \cdots & \gamma_{N-1} \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & & \gamma_T \end{bmatrix}, \tag{23}$$

where $\gamma_t = 2v_t - s_t^*$ and $\gamma_t \in \mathbb{Z}$. Here, $\Gamma(\cdot)$ stands for the Gamma function, while $\Gamma(a,b)$ represents the upper incomplete gamma function.

Proof: See Appendix A. ∎

Following *Lemma 1*, we can substitute (20) into (19) and utilize the Gauss-Chebyshev quadrature [16], which yields

$$\mathbb{E}[\epsilon_{cc}] \approx \frac{\pi}{2U_p} \sum_{p=1}^{U_p} \sqrt{1 - \eta_p^2} R, \tag{24}$$

where $R = \sum_{s_1=0}^{s_0} \cdots \sum_{s_T=0}^{s_{T-1}} \frac{g(\mathbf{s}^*)}{\pi^N D} \prod_{t=1}^T P_t \prod_{n=1}^N P_n$, and U_p represents the complexity-accuracy tradeoff parameters, $\eta_p = \cos\left(\frac{(2p-1)}{2U_p}\pi\right)$, and $y_{pc} = \frac{\eta_p}{2\delta_{N_c,L}\sqrt{L}} + \beta_{N_c,L}$.

Drawing a parallel, by invoking *Lemma 1*, when $\tau \leq \frac{\alpha_e}{\alpha_c}$, we have the CDF of γ_{ce} given by

$$F_{\gamma_{ce}}(\tau) = \sum_{s_1=0}^{s_0} \cdots \sum_{s_T=0}^{s_{T-1}} \frac{g(\mathbf{s}^*)}{\pi^N D} \prod_{t=1}^T P_t \prod_{n=1}^N P_n', \tag{25}$$

where

$$P_n' = \frac{1}{2}\left(\frac{Kn,n}{D}\right)^{-\frac{\bar{s}_n}{2} - \frac{1}{2}}$$
$$\times \left[\Gamma\left(\frac{1+\bar{s}_n}{2}\right) - \Gamma\left(\frac{1+\bar{s}n}{2}, \frac{K_{n,n}d_c^a\tau}{(\alpha_e\rho - \alpha_c\rho\tau)D}\right)\right]. \tag{26}$$

Similarly with $\mathbb{E}[\epsilon_{cc}]$, we have

$$\mathbb{E}[\epsilon_{ce}] \approx \delta\sqrt{m} \int_v^u F_{\gamma_{ce}}(\tau) d\tau, \tag{27}$$

where $\delta = (2\pi(2^{\frac{2N_e}{L}} - 1))^{-\frac{1}{2}}$, $v = \beta - \frac{1}{2}\delta^{-1}L^{-\frac{1}{2}}$, and $\beta = 2^{\frac{N_e}{L}} - 1$, respectively. Consequently, by substituting (25) into (27) and assuming $Q = \frac{g(\mathbf{s}^*)}{\pi^N D}$, we obtain

$$\mathbb{E}[\epsilon_{ce}] \approx \frac{\pi}{2U_p} \sum_{p=1}^{U_p} \sqrt{1 - \eta_p^2} \times \sum_{s_1=0}^{s_0} \cdots \sum_{s_T=0}^{s_{T-1}} Q \prod_{t=1}^T P_t \prod_{n=1}^N P_n', \tag{28}$$

where

$$y_{pe} = \frac{\eta_p}{2\delta_{N_e,L}\sqrt{L}} + \beta_{N_e,L}, \tag{29}$$

for $y_{pe} \leq \frac{\alpha_e}{\alpha_c}$; otherwise, $\mathbb{E}[\epsilon_{ce}] = 1$.

Utilizing (28) and (24) and following [17], we have the average BLER at the CU formulated as

$$\mathbb{E}[\epsilon_c] \geq \max\{\mathbb{E}[\epsilon_{cc}], \mathbb{E}[\epsilon_{ce}]\}. \tag{30}$$

For the derivation of $\mathbb{E}[\epsilon_e]$, when $\tau < \frac{\alpha_e}{\alpha_c}$, the CDF of γ_e can be expressed as

$$F_{\gamma_e}(\tau) = \sum_{s_1=0}^{s_0} \cdots \sum_{s_T=0}^{s_{T-1}} S \prod_{t=1}^T R_t \prod_{n=1}^N T_n U_n, \tag{31}$$

where $R_t = \frac{(-K_{m,n})^{s_t^*}}{s_t^*!D}$, $S = \frac{g(\mathbf{s}^*)}{\pi^N D}$, $T_n = \frac{1}{2}\left(\frac{K_{n,n}}{D}\right)^{-\frac{\bar{s}_n}{2}-\frac{1}{2}}$, and $U_n = \Gamma\left(\frac{1+\bar{s}_n}{2}\right) - \Gamma\left(\frac{1+\bar{s}_n}{2}, \frac{K_{n,n}d_e^a\tau}{(\alpha_e\rho-\alpha_c\rho\tau)D}\right)$. Otherwise, $F_{\gamma_e}(\tau) = 1$. Similar to $\mathbb{E}\left[\epsilon_{ce}\right]$, we have

$$\mathbb{E}\left[\epsilon_e\right] \approx \delta_{N_e,L}\sqrt{m}\int_{v_{N_e,L}}^{u_{N_e,L}} F_{\gamma_e}(\tau)d\tau. \tag{32}$$

Substituting (28) into (29), we obtain

$$\mathbb{E}\left[\epsilon_e\right] \approx \frac{\pi}{2U_p}\sum_{p=1}^{U_p}\sqrt{1-\eta_p^2}\sum_{s_1=0}^{s_0}\cdots\sum_{s_T=0}^{s_{T-1}} S\prod_{t=1}^{T} R_t\prod_{n=1}^{N} T_n U_n, \tag{33}$$

for $y_{pe} < \frac{\alpha_e}{\alpha_c}$; otherwise, $\mathbb{E}\left[\epsilon_e\right] = 1$.

IV. ASYMPTOTIC ANALYSIS

From [13], the CDF of $|g_{\text{FAS}}^{(c)}|$ and $|g_{\text{FAS}}^{(e)}|$ at high SNR is given by

$$F_{|g_{\text{FAS}}^{(c)}| \text{ or } |g_{\text{FAS}}^{(e)}|}(x) = \frac{1}{D}x^{2N} + o\left(\frac{1}{\rho^N}\right). \tag{34}$$

Consequently, by using the first-order Riemann integral approximation

$$\int_a^b f(\tau)d\tau = \frac{b-a}{2}f\left(\frac{b+a}{2}\right), \tag{35}$$

the asymptotic expression of $\mathbb{E}\left[\epsilon_{cc}\right]$, $\mathbb{E}\left[\epsilon_{ce}\right]$ and $\mathbb{E}\left[\epsilon_e\right]$ for high SNR can be approximated as

$$\mathbb{E}\left[\epsilon_{cc}\right] \approx \frac{1}{D}\left(\frac{d_c^a}{\alpha_c\rho}\beta_{N_c,L}\right)^N, \tag{36}$$

$$\mathbb{E}\left[\epsilon_{ce}\right] \approx \frac{1}{D}\left(\frac{d_c^a\beta_{N_e,L}}{\alpha_e\rho-\alpha_c\rho\beta_{N_e,L}}\right)^N, \tag{37}$$

$$\mathbb{E}\left[\epsilon_e\right] \approx \frac{1}{D}\left(\frac{d_e^a\beta_{N_e,L}}{\alpha_e\rho-\alpha_c\rho\beta_{N_e,L}}\right)^N, \tag{38}$$

for $\beta_{N_e,L} \leq \frac{\alpha_e}{\alpha_c}$; otherwise, $\mathbb{E}\left[\epsilon_{ce}\right] = 1$ and $\mathbb{E}\left[\epsilon_e\right] = 1$.

Remark 1: As we can see from the expressions above, we have the observation that the diversity order for both the CU and the CEU is directly dictated by N. This insight suggests that increasing N can play a significant role in enhancing the overall performance and robustness of the system.

V. BLOCKLENGTH OPTIMIZATION AT HIGH SNR

Given the asymptotic expression of $\mathbb{E}\left[\epsilon_{cc}\right]$, $\mathbb{E}\left[\epsilon_{ce}\right]$ and $\mathbb{E}\left[\epsilon_e\right]$ for high SNR, we aim to obtain the minimum common blocklength to improve the communication efficiency in this section. To this end, by assuming the average reliability constraints of the CU and the CEU as ϵ_c' and ϵ_e', respectively, we formulate the optimization problem as

$$\begin{aligned} \mathbf{P}: \min_{\alpha_c,\alpha_e} \quad & L^* = \Omega\left(\alpha_c, \alpha_e\right) \\ \text{s.t.} \quad & \mathbb{E}\left[\epsilon_c\right] \leq \epsilon_c', \\ & \mathbb{E}\left[\epsilon_e\right] \leq \epsilon_e', \\ & 0 < \alpha_c < 1, \\ & 0 < \alpha_e < 1, \\ & \alpha_c + \alpha_e = 1. \end{aligned} \tag{39}$$

As $\alpha_e = 1 - \alpha_c$, and $Q(x)$ is a decreasing function of x, (39) can be transformed as

$$\begin{aligned} \mathbf{P}: \min_{\alpha_c} \quad & L^* = \Omega\left(\alpha_c\right) \\ \text{s.t.} \quad & \mathbb{E}\left[\epsilon_c\right] = \epsilon_c', \\ & \mathbb{E}\left[\epsilon_e\right] = \epsilon_e', \\ & 0 < \alpha_c < 1. \end{aligned} \tag{40}$$

By substituting (38) into (40), i.e., $\mathbb{E}\left[\epsilon_e\right] = \epsilon_e'$ we can attain the minimum blocklength L_e for the CEU as

$$L_e \approx \frac{N_e}{\log_2\left(\frac{1+d_e^{-a}\rho \sqrt[N]{D\epsilon_e'}}{1+\alpha_c d_e^{-a}\rho \sqrt[N]{D\epsilon_e'}}\right)}. \tag{41}$$

Similarly, by combining (30), (36), (37), and $\mathbb{E}\left[\epsilon_c\right] = \epsilon_c'$, $\mathbb{E}\left[\epsilon_c\right]$ can be approximated as

$$\mathbb{E}\left[\epsilon_c\right] = \epsilon_c' \approx \max\left\{\frac{1}{D}\left(\frac{d_c^a}{\alpha_c\rho}\beta_{N_c,L}\right)^N, \left(\frac{d_c}{d_e}\right)^{aN}\epsilon_e'\right\}. \tag{42}$$

Hence, the minimum blocklength L_c for the CU is given by

$$L_c \approx \frac{N_c}{\log_2\left(1+\alpha_c\rho d_c^{-a}\sqrt[N]{D\epsilon_c'}\right)}. \tag{43}$$

To solve (40), we introduce the following lemma.

Lemma 2: The optimal power allocation coefficient α_c^* can be solved under the condition that $L_e = L_c$.

Proof: The optimal blocklength L^*, meeting user reliability constraints, is given by $L^* = \min(\max(L_c, L_e))$. Since $\frac{\partial L_c}{\partial \alpha_c} < 0$ (implying L_c decreases with α_c) and $\frac{\partial L_e}{\partial \alpha_c} > 0$ (indicating L_e increases with α_c), the optimal α_c^* minimizing L^* is found where $L_c = L_e$. This has completed the proof.

Accordingly, the optimal power allocation coefficient α_c^* at high SNR can be approximated as

$$\alpha_c^* \approx \left(\frac{d_c^a}{\rho \sqrt[N]{\epsilon_c' D}}\right)^{\frac{N_e}{N_e+N_c}}. \tag{44}$$

By substituting (44) into (41), we can approximate the minimum common blocklength as

$$L^* \approx \frac{N_c}{\log_2\left(1+\alpha_c^*\rho d_c^{-a}\sqrt[N]{D\epsilon_c'}\right)}. \tag{45}$$

To evaluate the transmission gain offered by the proposed systems, we will give the expressions of the minimum common blocklength without FAS and the minimum sum blocklength in OMA systems as follows.

In the case without FAS, the optimal power allocation coefficient $\bar{\alpha}_c^*$ at high SNR without FAS is approximated as

$$\bar{\alpha}_c^* \approx \left(\frac{d_c^a}{\rho\epsilon_c'}\right)^{\frac{N_e}{N_e+N_c}}. \tag{46}$$

Accordingly, the minimum common blocklength without FAS is approximated as

$$\bar{L}^* \approx \frac{N_c}{\log_2\left(1+\bar{\alpha}_c^*\rho d_c^{-a}\epsilon_c'\right)}. \tag{47}$$
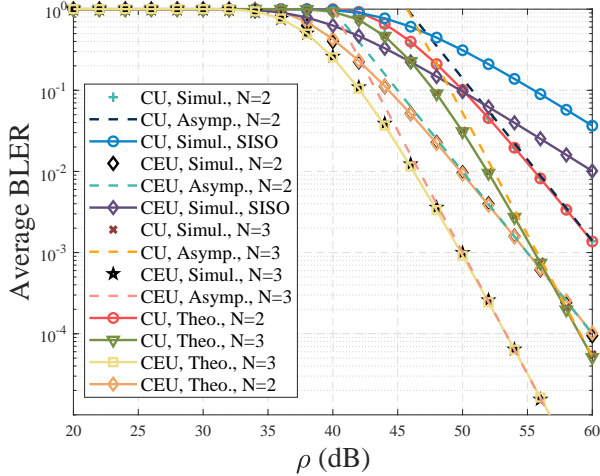
Fig. 1: Average BLER versus ρ, Parameters: $L = 100$, $N_c = 300$ bits, $N_e = 100$ bits, $\{N = 2, W = 5\}$ or $\{N = 3, W = 10\}$, $\alpha_c = 0.1$, $\alpha_e = 0.9$.

Fig. 2: Average BLER versus α_c, Parameters: $L = 100$, $N_c = 300$ bits, $N_e = 100$ bits, $\rho = 50$ dB, $\{N = 2, W = 5\}$ or $\{N = 3, W = 10\}$.

In the case of OMA systems, the minimum sum blocklength $\hat{L}^*$ is the summation of $\hat{L}_c$ and $\hat{L}_e$, where $\hat{L}_c$ and $\hat{L}_e$ are the minimum blocklength for CU and CEU in OMA, respectively. Similarly, at high SNR, $\hat{L}_c$ and $\hat{L}_e$ is approximated as

$$\hat{L}_e \approx \frac{N_e}{\log_2\left(1 + d_e^{-a}\rho\sqrt[N]{D\epsilon'_e}\right)}, \tag{48}$$

$$\hat{L}_c \approx \frac{N_c}{\log_2\left(1 + d_c^{-a}\rho\sqrt[N]{D\epsilon'_c}\right)}. \tag{49}$$

Thus, the minimum sum blocklength $\hat{L}^*$ is given by

$$\hat{L}^* \approx \frac{N_e}{\log_2\left(1 + d_e^{-a}\rho\sqrt[N]{D\epsilon'_e}\right)} + \frac{N_c}{\log_2\left(1 + d_c^{-a}\rho\sqrt[N]{D\epsilon'_c}\right)}. \tag{50}$$

VI. NUMERICAL RESULTS

Here, we provide numerical results to validate the accuracy of the analytically derived BLERs. For our simulations, we adopt the following parameters: $d_c = 5$ m, $d_e = 10$ m, $a = 3.9$, and $U_p = 10$. It is imperative to highlight that as we increase the value of U_p, our results show even better accuracy. We set the blocklength to $L = 100$, and the respective number of data bits allocated for CU and CEU are $N_c = 300$ bits and $N_e = 100$ bits, following the setup suggested by [17]. For each simulated result, we average over 10^5 randomly generated channel realizations to compute the average BLERs.

In Fig. 1, the relationship between the average BLER and SNR ρ is depicted for both CU and CEU under various parameters. We have set the power allocation coefficients to $\alpha_c = 0.1$ and $\alpha_e = 0.9$ [17], with either $N = 2$ or $N = 3$. "CU, Simul.", "CU, Theo.", and "CU, Asymp." represent the simulated, theoretical, and asymptotic average BLER outcomes for CU. Similarly, "CEU, Simul.", "CEU, Theo.", and "CEU, Asymp." denote the related results for CEU. For comparison, we also depict the simulated average BLER for CU and CEU in a single antenna setup without FAS, labeled as "CU, SISO, Simul." and "CEU, SISO, Simul.".
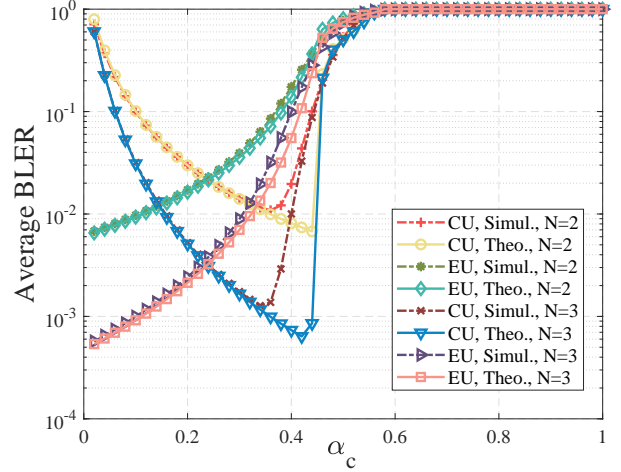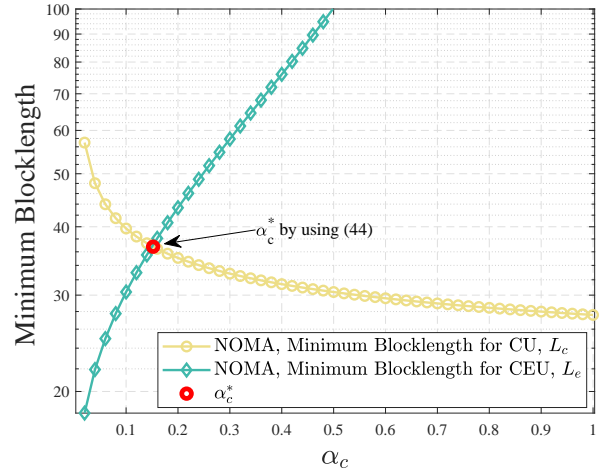

Fig. 3: Minimum Blocklength versus α_c, Parameters: $N_c = 300$ bits, $N_e = 100$ bits, $N = 2, W = 5$, $\rho = 80$ dB, $\epsilon'_c = 10^{-4}$ and $\epsilon'_e = 10^{-3}$.

From Fig. 1, it is seen that as SNR ρ increases, there is a remarkable decline in the average BLER. Also, the theoretical lines closely match the simulated results. A comparison with the non-FAS highlights the improved performance of the FAS-assisted NOMA short-packet system.

In Fig. 2, the average BLER is shown as a function of α_c for a fixed $\rho = 50$ dB and $N = 3$. The results in Fig. 2 demonstrate the consistency between the theoretical analysis and the simulated outcomes for the average BLERs of both CU and CEU, confirming the robustness of our proposed model.

In Fig. 3, we present the minimum common blocklength for both the CU and CEU across varying power allocation coefficients, denoted as α_c. It is observed that L_c decreases with an increase in α_c, whereas L_e exhibits an opposite trend, increasing as α_c rises. Notably, the optimal power allocation coefficient α_c^*, as derived from equation (45), corresponds to
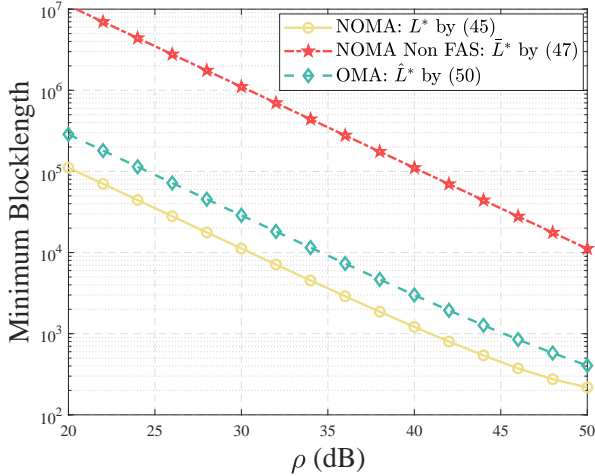
Fig. 4: Minimum Blocklength versus ρ, Parameters: $N_c = 300$ bits, $N_e = 100$ bits, $N = 2, W = 5, \epsilon_c' = 10^{-4}$ and $\epsilon_e' = 10^{-3}$.

the point where $L_c = L_e$.

In Fig. 4, we compare the minimum sum blocklength of NOMA with those of systems without FAS and of OMA. It is observed that the minimum blocklength L^* in NOMA is significantly lower than $\bar{L}^*$ in the system without FAS and $\hat{L}^*$ in OMA. Specifically, the transmission gain attributable to FAS is substantially greater than that achievable through the NOMA scheme.

Remark 2: Given the specific modulation scheme and the channel coding, the simulation results of the average BLER give the BLER lower bound. Also, the actual coding and modulation schemes are not included here.

VII. CONCLUSION

This letter derived the average BLER of the FAS-assisted downlink NOMA short-packet communication system. Numerical results showed that the average BLERs of both the CU and CEU are nearly identical. Moreover, the FAS-assisted NOMA short-packet system outperforms its non-FAS counterpart, demonstrating great performance enhancement.

VIII. APPENDIX A

The joint CDF of $|g_{c1}|, |g_{c2}|,\ldots,|g_{cN}|$ is given by [13]

$$
\begin{aligned}
&F_{|g_{c1}|,|g_{c2}|,\ldots,|g_{cN}|}(r_1, r_2, \ldots, r_N) \\
&= \sum_{s_1=0}^{s_0} \sum_{s_2=0}^{s_1} \cdots \sum_{s_T=0}^{s_{T-1}} \frac{g(\mathbf{s}^*)}{\pi^N \det(\mathbf{J})} \prod_{t=1}^{T} \frac{(-K_{m,n})^{s_t^*}}{s_t^*! \det(\mathbf{J})} \\
&\times \prod_{n=1}^{N} \frac{1}{2} \left(\frac{K_{n,n}}{\det(\mathbf{J})} \right)^{-\frac{\bar{s}_n}{2}-\frac{1}{2}} \\
&\times \left[\Gamma\left(\frac{1+\bar{s}_n}{2} \right) - \Gamma\left(\frac{1+\bar{s}_n}{2}, \frac{K_{n,n}r_n^2}{\det(\mathbf{J})} \right) \right].
\end{aligned}
$$
(51)

With the joint cdf of $|g_{c1}|, |g_{c2}|,\ldots,|g_{cN}|$, we use the following expression to compute the CDF of γ_{cc} as

$$F_{\gamma_{cc}}(\tau) = \Pr\left(\gamma_{cc} < \tau \right).$$
(52)

Adding (11) into (52), we have

$$
\begin{aligned}
F_{\gamma_{cc}}(\tau) &= \Pr\left(|g_{\text{FAS}}^{(c)}| < \sqrt{\frac{d_c^a \tau}{\alpha_c \rho}} \right) \\
&= \Pr\left(|g_{c1}| < \sqrt{\frac{d_c^a \tau}{\alpha_c \rho}}, \ldots, |g_{cN}| < \sqrt{\frac{d_c^a \tau}{\alpha_c \rho}} \right).
\end{aligned}
$$
(53)

The CDF expression can be found by substituting $r_1 = r_2 = \cdots = r_N = \sqrt{\frac{d_c^a \tau}{\alpha_c \rho}}$ into the joint CDF (51), which completes the proof.

REFERENCES

[1] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of nonorthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[2] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Commun. Surv. Tut.*, vol. 19, no. 12, pp. 721–742, Second Quarter 2016.

[3] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[4] X. Sun, S. Yan, N. Yang, Z. Ding, C. Shen and Z. Zhong, "Short-packet downlink transmission with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4550–4564, Jul. 2018.

[5] Y. Yu, H. Chen, Y. Li, Z. Ding, and B. Vucetic, "On the performance of non-orthogonal multiple access in short-packet communications," *IEEE Commun. Lett.*, vol. 22, no. 3, pp. 590–593, Mar. 2018.

[6] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[7] B. Makki, T. Svensson, and M. Zorzi, "Finite block-length analysis of the incremental redundancy HARQ," *IEEE Wireless Commun. Lett.*, vol. 3, no. 5, pp. 529–532, Oct. 2014.

[8] K. K. Wong, K. F. Tong, Y. Shen, Y. Chen, and Y. Zhang, "Bruce Lee-inspired fluid antenna system: Six research topics and the potentials for 6G," *Frontiers Commun. & Netw., section Wireless Commun.*, vol. 3, no. 853416, Mar. 2022.

[9] K.-K. Wong, A. Shojaeifard, K.-F. Tong and Y. Zhang, "Performance limits of fluid antenna systems," *IEEE Wireless Commun. Lett.*, vol. 24, no. 11, pp. 2469–2472, Nov. 2020.

[10] K.-K. Wong, A. Shojaeifard, K.-F. Tong, and Y. Zhang, "Fluid antenna systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1950–1962, Mar. 2021.

[11] K.-K. Wong, K. F. Tong, Y. Chen, and Y. Zhang, "Closed-form expressions for spatial correlation parameters for performance analysis of fluid antenna systems," *IET Electron. Lett.*, vol. 58, no. 11, pp. 454-457, May, 2022.

[12] M. Khammassi, A. Kammoun, and M.-S. Alouini, "A new analytical approximation of the fluid antenna system channel," *IEEE Trans. Wireless Commun.*, early access, DOI:10.1109/TWC.2023.3266411.

[13] W. K. New, K.-K. Wong, H. Xu, K.-F. Tong, and C.-B. Chae, "Fluid antenna system: New insights on outage probability and diversity gain," *IEEE Trans. Wireless Commun.*, early access, May 19, 2023, doi: 10.1109/TWC.2023.3276245.

[14] K.-K.Wong and K.-F. Tong, "Fluid antenna multiple access," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4801–4815, Jul. 2022.

[15] K.-K. Wong, D. Morales-Jimenez, K.-F. Tong, and C.-B. Chae, "Slow fluid antenna multiple access," *IEEE Trans. Commun.*, vol. 71, no. 5, pp. 2831–2846, May 2022.

[16] E. Hildebrand, *Introduction to Numerical Analysis*. New York, NY, USA: Dover, 1987.

[17] X. Lai, Q. Zhang, and J. Qin, "Cooperative NOMA short-packet communications in flat Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6182–6186, Jun. 2019.