IEEE Open Journal of the Communications Society

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version 11 January, 2024.

Digital Object Identifier 10.1109/OJCOMS.2024.011100

Widely Linear Processing Improves the Throughput of Nonorthogonal User Access

A. A. Nasir[1], H. D. Tuan[2], H. V. Poor[3] *(Life Fellow, IEEE)*, AND L. Hanzo[4] *(Life Fellow, IEEE)*

[1]Department of Electrical Engineering and Center for Communication Systems and Sensing at King Fahd University of Petroleum and Minerals (KFUPM), Dhahran 31261, Saudi Arabia
[2]School of Electrical and Data Engineering, University of Technology Sydney, Broadway, NSW 2007, Australia
[3]The department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA.
[4]School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K.

CORRESPONDING AUTHOR: L. Hanzo (e-mail: lh@ecs.soton.ac.uk).

ABSTRACT The quality-of-service (QoS) provided by wireless communication networks can be upgraded by the technique of non-orthogonal multiple-access (NOMA) in tandem with widely linear beamforming (WLB), which uses a pair of beamformers for each information symbol. Conventionally, rate-fairness among the users is achieved by maximizing the users' minimal throughput (max-min throughput optimization). However, this is computationally challenging, as each iteration requires solving a high-dimensional convex optimization problem, even for small networks. We circumvent this by maximizing the geometric mean (GM) of the users' throughput (GM-throughput maximization) and design novel algorithms based on iterating closed-form expressions are developed, which are shown to be hundreds of times more computationally efficient than the existing algorithms that are based on convex-solvers. The proposed algorithms are developed for both conventional wireless networks and networks requiring ultra-reliable and low-latency communications (URLLC).

INDEX TERMS Widely linear beamforming, non-orthogonal multiple access, ultra-reliable and low-latency communication, geometric mean maximization.

I. Introduction

Non-orthogonal multiple access (NOMA) [1] has been recognized as a flexible and efficient technology to improve the quality-of-service (QoS) provided by wireless communication networks [2]. A key principle of NOMA is that the users (UEs) of interest also decode messages intended for other UEs for interference mitigation. More specifically, NOMA enables UEs of the same cluster to successively decode all the messages destined to all UEs to mitigate intra-cluster interference. This is in a sharp contrast to the conventional coordinated signaling (CoSig) [3], [4], where each UE decodes its intended message while treating interference from other UEs as noise.[1]

The most popular approach to implement NOMA is to pair UEs so that each of the paired UEs decodes the entire message destined for the paired UE. However, this NOMA scheme has been demonstrated to outperform CoSig in terms of the worst-case UE throughput only when the channel conditions of the paired UEs are sufficiently different [5]. Recently, a new-NOMA (n-NOMA) scheme was proposed in [6], which takes advantage of the Han-Kobayashi (HK) strategy [7]–[10], wherein only a portion of the message intended for one UE is designated as a common message to be decoded by both paired UEs. Traditional NOMA is thus a particular case of n-NOMA when the common message is the entire message for the two paired UEs. Likewise, CoSig can be viewed as another particular case of n-NOMA when there is no common message. It is shown in [6] that n-NOMA readily outperforms both NOMA and CoSig under diverse channel conditions.

Exploring transmit beamforming at the base station (BS) is essential to take full advantage of NOMA for improving the QoS in terms of the UEs' throughput [5], [11], [12]. However, transmit beamforming design problems are often high-dimensional nonconvex optimization tasks, even for small networks, and hence difficult to solve. The solution algorithms developed in [5], [11] invoke convex quadratic problems of the same size at each iteration to generate a

[1]Technically, CoSig is also a form of NOMA because all users share the same communication bandwidth.

TABLE 1: Comparison of this paper and relevant literature.

	This work	[5]	[13]	[14]	[6]
Single-antenna UEs	✓	✓	✓	✓	✓
Multi-antenna UEs	✓	✓	✓		
n-NOMA	✓				✓
Fair comparison to CoSig	✓	✓		✓	✓
WLB	✓			✓	✓
Polynomial-time complexity solution		✓		✓	✓
Low-complexity solution	✓				
URLLC	✓				

better feasible point. The computational complexity of these convex problems is polynomial in their size, and thus is still high even for small-size networks. The authors of [13] adopted the same beamformer for UEs in the same cluster and used convex relaxation at each iteration, which not only causes an increase in the dimension (size) of the problem but may also fail to output a feasible point. On the other hand, the authors of [14] and [6] have shown that the UEs' throughput thresholds in NOMA and n-NOMA networks can be increased by performing widely linear beamforming (WLB), which uses two beamformers for each information symbol [15]. WLB enables improper Gaussian signaling (IGS) by relaxing signal properness, allowing correlation with complex conjugates and yielding additional signaling degrees of freedom. IGS signals are generated from proper Gaussian sources using WLB [15]–[19]. Since the problem size in WLB design is twice that in linear beamforming (LB) design, the former is more computationally demanding than the latter [20], [21].

In other relevant research, NOMA has recently been exploited for minimizing either the transmission power [22] or probability of error [23] in ultra-reliable and low-latency communications (URLLC) [24]. The solutions in [22], [23] are, however, restricted to simple single-input single-output (SISO) systems, and hence do not use beamforming.

In general, enhancing the QoS for all UEs is hard. It has been recently shown that by employing geometric-mean (GM)-throughput optimization [25], [26], which optimizes the GM of the UEs' throughputs, one can naturally achieve fair throughput distribution, resulting in a low standard deviation among the UEs' throughputs, without enforcing specific QoS constraints. Thus, GM-throughput optimization has a potential to offer computationally tractable solutions for resource allocation problems and can be exploited for WLB designs in multiple-input-multiple-output (MIMO)-NOMA networks.

Against the above background, this paper aims to exploit both n-NOMA as well as WLB and obtain computationally-efficient beamforming solutions to enhance the QoS in large wireless networks. We also provide a computationally efficient WLB solution for enhancing the QoS of URLLC networks. Here, we would like to emphasize that this work is significantly different from [25], [26], as it is non-trivial to apply the concepts of GM-throughput optimization introduced in [25], [26] to an n-NOMA network under the proposed WLB framework. Moreover, additional novelty arises from extending the proposed WLB design to URLLC networks. Our contributions are contrasted to the state-of-the-art in Table 1 and are elaborated on further below.

- We propose a computationally efficient WLB design for both n-NOMA and NOMA networks, based on maximizing the geometric mean (GM) of the UEs' throughputs under a sum transmit power constraint. While the GM-throughput is a smooth function of the individual throughputs, the throughputs themselves are non-smooth with respect to the beamforming variables, making the optimization challenging. To address this, we develop novel algorithms that iterate over closed-form expressions, offering a computational efficiency that is hundreds of times greater than existing convex-solver-based methods. We also demonstrate that GM-throughput maximization not only leads to a fairer distribution of throughput among UEs, but also promotes more balanced power allocation across antennas—even without enforcing per-antenna power constraints. Compared to LB, the proposed WLB approach more than doubles the users' throughput and reduces the standard deviation of users' throughputs by over 50%. Additionally, we show that n-NOMA outperforms both CoSig and conventional NOMA schemes.
- The proposed WLB design is also extended to URLLC, which presents an even more challenging optimization problem due to the additional impact of channel dispersion on the throughput function. Nevertheless, we successfully develop a computationally efficient iterative procedure, based on closed-form expressions, for WLB computation in such wireless networks as well.

The remainder of the paper is structured as follows. Section II describes the system model and formulates the WLB design problem for GM-throughput maximization. The algorithm proposed for efficiently solving the problem is presented in Section III. Section IV extends the proposed algorithm for URLLC networks. Simulation results are provided in Section V. Finally, Section VI concludes the paper.

Notation. Design (decision) variables subject to optimization are represented in boldface. I_n denotes the identity matrix of dimension $n \times n$. The expressions X^T, X^H, X^*, $\langle X \rangle$, and $|X|$ correspond to the transpose, conjugate transpose, complex conjugate, trace, and determinant of matrix X, respectively, and $[X]^2 = XX^H$. The inner product $\langle X, Y \rangle$ for real matrices X and Y is given by trace$(X^H Y)$. The notation $|\cdot|$ refers to the Frobenius norm for matrices and the Euclidean norm for vectors. $\Re x$ extracts the real part of a complex number x. The expression $x \sim \mathcal{CN}(0, Z)$

TABLE 2: List of commonly used symbols.

Symbols	Description
N	number of UEs
$\mathcal{N}$	$\{1,\ldots,N\}$
$\mathcal{N}_1$ & $\mathcal{N}_2$	$\{1,\ldots,N/2\}$ & $\{N/2+1,\ldots,N\}$
$\mathcal{N}_E$	$\{1,\ldots,N+N/2\}$
s_n, $n\in\mathcal{N}$	private message for UE n
$\pi(n)$, $n\in\mathcal{N}_1$	$n+N/2$
s_{n+N}, $n\in\mathcal{N}_1$	common message for UE n & UE $\pi(n)$
$\mathbf{W}_{1,n}$ / $\mathbf{W}_{2,n}$, $n\in\mathcal{N}_E$	WLB for s_n / s_n^*
$\sum_{n\in\mathcal{N}_E} x_n$	transmitted signal
y_n, $n\in\mathcal{N}$	received signal by UE n
H_n, $n\in\mathcal{N}$	channel matrix from BS to UE n
n_n, $n\in\mathcal{N}$	AWGN at UE n
$\{\bar{s}_n,\bar{y}_n,\bar{H}_n,\mathbf{V}_n\}$	linear mapping of $\{x_n,y_n,H_n,\mathbf{W}_n\}$
$\frac{1}{2}\rho_{1,N+n}(\mathbf{V})$, $n\in\mathcal{N}_1$	throughput in decoding s_{N+n} by UE n
$\frac{1}{2}\rho_{2,N+n}(\mathbf{V})$, $n\in\mathcal{N}_1$	throughput in decoding s_{N+n} by UE $\pi(n)$
$\frac{1}{2}\rho_{\chi}(\mathbf{V})$, $\chi\in\{n,\pi(n)\}$	throughput in decoding s_χ by UE χ
$\frac{1}{2}\rho_{n}(\mathbf{V})$, $n\in\mathcal{N}_1$	throughput of UE n, $n\in\mathcal{N}_1$
$\min_{i=1,2}\frac{1}{2}\rho_{i,\pi(n)}(\mathbf{V})$	throughput of UE $\pi(n)$, $n\in\mathcal{N}_1$
$\rho_{i,\pi(n)}(\mathbf{V})$	$\triangleq\rho_{\pi(n)}(\mathbf{V})+\rho_{i,N+n}(\mathbf{V})$, $i=1,2$
P	transmit power budget
t_t	URLLC transmission duration
$\mathcal{B}$	communication bandwidth
ϵ_c	target decoding error probability
$\nu(\mathbf{V})$	channel dispersion
$\varphi(\mathbf{V})$	SINR

indicates that x is a circularly symmetric complex Gaussian random vector with zero mean and covariance matrix Z. A matrix $A\succ 0$ signifies that A is positive definite. The operator $\mathbb{E}\cdot$ denotes statistical expectation. $\binom{n}{k}=\frac{n!}{k!(n-k)!}$ is the combination formula to choose k items out of n items. The list of commonly used symbols is provided in Table 2.

Ingredient. According to [27, p. 366], a function $\tilde{f}$ is called a minorant (or majorant, respectively) of the function f over the domain $\mathsf{dom}(f)$ if $\tilde{f}(\mathbf{x})\le f(\mathbf{x})\ \forall\mathbf{x}\in\mathsf{dom}(f)$ ($\tilde{f}(\mathbf{x})\ge f(\mathbf{x})\ \forall\mathbf{x}\in\mathsf{dom}(f)$, resp.). In addition, if $\tilde{f}(\bar{x})=f(\bar{x})$ for some $\bar{x}\in\mathsf{dom}(f)$ then the former is called a tight minorant (tight majorant, resp.) of the latter at $\bar{x}$. Let $x_{\rm opt}=\arg\max_{\mathbf{x}\in\mathsf{dom}(f)}\tilde{f}(\mathbf{x})$, where $\tilde{f}$ is a tight minorant of f at $\bar{x}$. Then $f(x_{\rm opt})\ge\tilde{f}(x_{\rm opt})>\tilde{f}(\bar{x})=f(\bar{x})$ whenever $\tilde{f}(x_{\rm opt})\ne\tilde{f}(\bar{x})$, i.e. $x_{\rm opt}$ is a superior feasible point compared to $\bar{x}$ for maximizing f. This property does not hold when $\tilde{f}$ is only a minorant of f.

II. System Model and Problem Formulation

Consider downlink transmission in a wireless network, which consists of a base station equipped with N_t transmit antennas (TAs) to serve N UEs $n\in\mathcal{N}\triangleq\{1,2,\cdots,N\}$, each having N_r receive antennas. The UEs are grouped into two distinct subsets $\mathcal{N}_1\triangleq\{1,2,\cdots,N/2\}$ and $\mathcal{N}_2\triangleq\{N/2+1,\cdots,N\}$. Let UE $n\in\mathcal{N}_1$ be paired with UE $\pi(n)\in\mathcal{N}_2$ to create a virtual cluster. For ease of notation, the pairing is selected such that $\pi(n)=n+N/2$. We also define $\mathcal{N}_E\triangleq\{1,\ldots,N+N/2\}$. The joint optimization of user-pairing, decoding order, and beamforming is computationally intractable , especially for MIMO-NOMA networks [28], [29].[2] It is beyond the scope of this work, which focuses on the design of *low-complexity* and *widely linear* beamforming solution for MIMO-NOMA networks.

We adopt the following common strategy for user-pairing [30], [31]. First, the UEs are ordered in terms of their distances from the BS. Then we pair the first ordered UE with the $(N/2+1)$-th ordered UE. Similarly, the second ordered UE is paired with the $(N/2+2)$-th ordered UE, and so on. This strategy relies on the distances of users from the BSs, which ensures minimal overhead when updating users about their pairing strategy. Similarly, within the paired users, we assume that the decoding order is based on the ascending order of users' channel strength [13], [28], [29], [32], which is determined mainly by the users' distances from the BSs [5], [6], [14], [33]. It is very likely for our adopted user-pairing strategy, because the distance between the paired users is as large as possible. As a result, their channel strength is reflected mainly by their distances from the BS owing to the considerable difference in their path losses.

Under n-NOMA [6], in addition to a private message $s_n\sim\mathcal{CN}(0,I_{N_r})$, $n=1,\ldots,N$ that is decoded by the individual UE n's receiver, there is a common message $s_{N+n}\sim\mathcal{CN}(0,I_{N_r})$, $n=1,\ldots,N/2$, intended for UE $\pi(n)$, but it is decoded by both UE n's and UE $\pi(n)$'s receiver. This is demonstrated in the system model of Fig. 1 and further elaborated on and contrasted to both NOMA and CoSig in Table 3.[3] We know that UEs $\{N/2+1,\ldots,N\}$, which are far from the BS, experience higher path-loss and poorer throughput compared to the other UEs. In contrast to both NOMA and CoSig, n-NOMA serves these UEs $\{N/2+1,\ldots,N\}$ using both private and common messages, thus holds the promise of enhancing both the users' minimal throughput and the GM of the users' throughput.

With WLB, each $s_{n'}$, $n'=1,\ldots,N+N/2$ is processed as

$$x_{n'}=\mathbf{W}_{1,n'}s_{n'}+\mathbf{W}_{2,n'}s_{n'}^*,\tag{1}$$

where $\mathbf{W}_{i,n'}\in\mathbb{C}^{N_t\times N_r}$, $i=1,2$. Note that the above processing is different from the conventional LB, which applies a single beamformer for each information symbol. In other words, LB is a special case of WLB along with $\mathbf{W}_{2,n}\equiv 0$. The received signal at UE n when employing

[2]There are $\frac{1}{(N/2)!}\prod_{i=0}^{N/2-1}\binom{N-2i}{2}$ possible user-pairings and $2^{N/2}$ decoding order options [29]. For example, for $N=18$ UEs, we need to choose from 34459425 user-pairing and 512 decoding order options.

[3]Our description assumes that N is even. In case of odd N, one UE will be left unpaired and will be served with its private message alone.

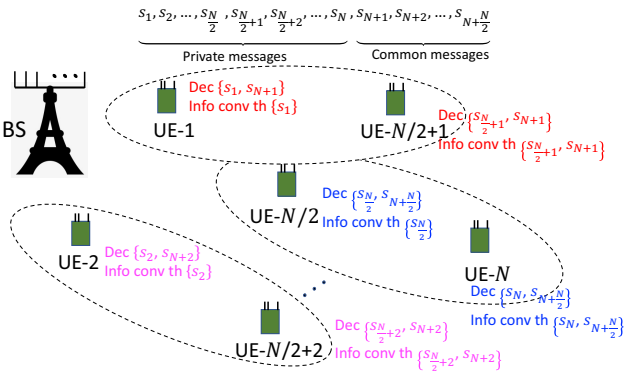

FIGURE 1: A system model illustration for n-NOMA scheme, where abbreviations "Dec" and "Info conv th" stand for "decodes" and "information is conveyed through", respectively.

TABLE 3: Overview of how the information is conveyed.

Scheme	Source through which information is conveyed to	
	UEs $\{1, \ldots, N/2\}$	UEs $\{N/2+1, \ldots, N\}$
n-NOMA	private message	private & common messages
NOMA	private message	common message
CoSig	private message only	private message only

WLB can be expressed as

$$y_n = H_n \sum_{n' \in \mathcal{N}_E} x_{n'} + n_n \tag{2}$$

$$= H_n \sum_{n' \in \mathcal{N}_E} (\mathbf{W}_{1,n'} s_{n'} + \mathbf{W}_{2,n'} s_{n'}^*) + n_n, \tag{3}$$

where $H_n \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix from the BS to UE n, which is assumed to be known, and $n_n \sim \mathcal{CN}(0, \sigma I_{N_r})$ is additive white Gaussian noise (AWGN).

Define $\mathbf{W} \triangleq \{\mathbf{W}_{n'}, n' \in \mathcal{N}_E\}$, $\mathbf{W}_{n'} \triangleq \{\mathbf{W}_{i,n'}, i = 1, 2\}$, and the real-valued variables:

$$\bar{H}_n \triangleq \begin{bmatrix} \Re\{H_n\} & -\Im\{H_n\} \\ \Im\{H_n\} & \Re\{H_n\} \end{bmatrix} \in \mathbb{R}^{(2N_r)\times(2N_t)},$$
$$\bar{s}_{n'} \triangleq \begin{bmatrix} \Re\{s_{n'}\} \\ \Im\{s_{n'}\} \end{bmatrix} \in \mathbb{R}^{2N_r}, \ \bar{y}_n \triangleq \begin{bmatrix} \Re\{y_n\} \\ \Im\{y_n\} \end{bmatrix} \in \mathbb{R}^{2N_r}, \tag{4}$$
$$\bar{n}_n \triangleq \begin{bmatrix} \Re\{n_n\} \\ \Im\{n_n\} \end{bmatrix} \in \mathbb{R}^{2N_r}.$$

Then (3) can be rewritten using real-valued variables as:

$$\bar{y}_n = \bar{H}_n \sum_{n' \in \mathcal{N}_E} \mathbf{V}_{n'} \bar{s}_{n'} + \bar{n}_n, \tag{5}$$

where we have:

$$\mathbf{V}_{n'} \triangleq [\mathbf{V}_{n'}^{ij}]_{(i,j)\in\{1,2\}} = \begin{bmatrix} \Re\{\mathbf{W}_{1,n'} + \mathbf{W}_{2,n'}\} & -\Im\{\mathbf{W}_{1,n'} - \mathbf{W}_{2,n'}\} \\ \Im\{\mathbf{W}_{1,n'} + \mathbf{W}_{2,n'}\} & \Re\{\mathbf{W}_{1,n'} - \mathbf{W}_{2,n'}\} \end{bmatrix}, \tag{6}$$

with

$$\|\mathbf{V}_{n'}\|^2 = 2(\|\mathbf{W}_{1,n'}\|^2 + \|\mathbf{W}_{2,n'}\|^2). \tag{7}$$

The decoding order is enforced as follows. First, both UEs n and $\pi(n)$ decode the common message s_{N+n}, which is intended for UE $\pi(n) \in \mathcal{N}_2$, treating other messages $s_{n'}$, $n' \neq N+n$ as interferences. By writing (5) as

$$\bar{y}_n = \bar{H}_n \mathbf{V}_{N+n} \bar{s}_{N+n} + \sum_{n' \in \mathcal{N}_E \setminus \{N+n\}} \bar{H}_n \mathbf{V}_{n'} \bar{s}_{n'} + \bar{n}_n, \tag{8}$$

the rate of decoding s_{N+n} by UE n is $\frac{1}{2}\rho_{1,N+n}(\mathbf{V})$ with [34]

$$\rho_{1,N+n}(\mathbf{V}) \triangleq \ln \left| I_{2N_r} + [\bar{H}_n \mathbf{V}_{N+n}]^2 \Psi_{1,N+n}^{-1}(\mathbf{V}) \right|, \tag{9}$$

with $\Psi_{1,N+n}(\mathbf{V}) \triangleq \sum_{n' \in \mathcal{N}_E \setminus \{N+n\}} [\bar{H}_n \mathbf{V}_{n'}]^2 + \sigma I_{2N_r}$, which represents the covariance of interference-plus-noise signal. Analogously, by writing (5) as

$$\bar{y}_{\pi(n)} = \bar{H}_{\pi(n)} \mathbf{V}_{N+n} \bar{s}_{N+n} + \sum_{n' \in \mathcal{N}_E \setminus \{N+n\}} \bar{H}_{\pi(n)} \mathbf{V}_{n'} \bar{s}_{n'} + \bar{n}_{\pi(n)}, \tag{10}$$

the rate of decoding s_{N+n} by UE $\pi(n)$ is $\frac{1}{2}\rho_{2,N+n}(\mathbf{V})$ with [34]

$$\rho_{2,N+n}(\mathbf{V}) \triangleq \ln \left| I_{2N_r} + [\bar{H}_{\pi(n)} \mathbf{V}_{N+n}]^2 \Psi_{2,N+n}^{-1}(\mathbf{V}) \right|, \tag{11}$$

along with $\Psi_{2,N+n}(\mathbf{V}) \triangleq \sum_{n' \in \mathcal{N}_E \setminus \{N+n\}} [\bar{H}_{\pi(n)} \mathbf{V}_{n'}]^2 + \sigma I_{2N_r}$.

Denote the throughput in decoding the common message s_{N+n} by both UEs n and $\pi(n)$ as r_{N+n}, which we refer to as the "common-throughput". Based on information theory [34], we have:

$$r_{N+n} \leq \frac{1}{2} \rho_{i,N+n}(\mathbf{V}), i = 1, 2, \tag{12}$$

or equivalently,

$$r_{N+n} \leq \frac{1}{2} \rho_{N+n}(\mathbf{V}), \tag{13}$$

along with

$$\rho_{N+n}(\mathbf{V}) \triangleq \min_{i=1,2} \rho_{i,N+n}(\mathbf{V}). \tag{14}$$

Next, after decoding s_{N+n}, the UEs $n \in \mathcal{N}_1$ and $\pi(n) \in \mathcal{N}_2$ subtract it from their received signals $\bar{y}_n$ and $\bar{y}_{\pi(n)}$ before decoding their private messages s_n and $s_{\pi(n)}$. By (5), the equations of the signals received at UEs n and $\pi(n)$ after subtracting s_{N+n} become:

$$\bar{y}_n - \bar{H}_n \mathbf{V}_{N+n} \bar{s}_{N+n} = \bar{H}_n \mathbf{V}_n s_n + \sum_{n' \in \mathcal{N}_E \setminus \{N+n, n\}} \bar{H}_n \mathbf{V}_{n'} \bar{s}_{n'} + \bar{n}_n, \tag{15}$$

and

$$\bar{y}_{\pi(n)} - \bar{H}_{\pi(n)} \mathbf{V}_{N+n} \bar{s}_{N+n} = \bar{H}_{\pi(n)} \mathbf{V}_{\pi(n)} s_{\pi(n)} + \sum_{n' \in \mathcal{N}_E \setminus \{N+n, \pi(n)\}} \bar{H}_{\pi(n)} \mathbf{V}_{n'} \bar{s}_{n'} + \bar{n}_{\pi(n)}, \tag{16}$$

respectively. It follows from (15) that the throughput in decoding s_n by UE n, while treating the other messages $s_{n'}$, $n' \in \mathcal{N}_E \setminus \{N+n, n\}$ as interference is $\frac{1}{2}\rho_n(\mathbf{V})$, where

$$\rho_n(\mathbf{V}) \triangleq \ln \left| I_{2N_r} + [\bar{H}_n \mathbf{V}_n]^2 \Psi_n^{-1}(\mathbf{V}) \right|, \tag{17}$$

along with

$$\Psi_n(\mathbf{V}) \triangleq \sum_{n' \in \mathcal{N}_E \setminus \{n, N+n\}} [\bar{H}_n \mathbf{V}_{n'}]^2 + \sigma I_{2N_r}. \tag{18}$$

Analogously, it follows from (16) that the throughput in decoding $s_{\pi(n)}$ by UE $\pi(n)$, while treating other messages $s_{n'}$, $n' \in \mathcal{N}_E \setminus \{N+n, \pi(n)\}$ as interference is $\frac{1}{2}\rho_{\pi(n)}(\mathbf{V})$, where we have

$$\rho_{\pi(n)}(\mathbf{V}) \triangleq \ln \left| I_{2N_r} + [\bar{H}_{\pi(n)} \mathbf{V}_{\pi(n)}]^2 \Psi_{\pi(n)}^{-1}(\mathbf{V}) \right|, \tag{19}$$

with

$$\Psi_{\pi(n)}(\mathbf{V}) \triangleq \sum_{n' \in \mathcal{N}_E \setminus \{\pi(n), N+n\}} [\bar{H}_{\pi(n)} \mathbf{V}_{n'}]^2 + \sigma I_{2N_r}. \tag{20}$$

Since the information of the UE $\pi(n) \in \mathcal{N}_2$ is conveyed through both the private message $s_{\pi(n)}$ and common message s_{N+n}, its throughput, based on (13) and (19), is given by

$$\frac{1}{2}\rho_{\pi(n)}(\mathbf{V}) + r_{N+n}, \tag{21}$$

where r_{N+n} is subject to the constraint (13).

Our objective is to maximize the GM of the UEs' throughputs while satisfying the total power constraint. From (7), (19), and (21), such a problem is formulated as[4]

$$\max_{\mathbf{V}, r_{N+n}} \left[\prod_{n \in \mathcal{N}_1} \left(\frac{1}{2}\rho_n(\mathbf{V}) \left(\frac{1}{2}\rho_{\pi(n)}(\mathbf{V}) + r_{N+n} \right) \right) \right]^{1/N} \tag{22a}$$

$$\text{s.t.} \quad (13), \tag{22b}$$

$$\sum_{n' \in \mathcal{N}_E} \|\mathbf{V}_{n'}\|^2 \leq 2P, \tag{22c}$$

where P is the power budget. The advantage of considering GM-throughput maximization in (22) is that it ensures a fair distribution of throughput among UEs. This is due to the structure of GM, where the throughput functions of all UEs are multiplied together, preventing any individual UE from experiencing very low throughput. As a result, there is no need to impose additional computationally intractable QoS constraints, such as

$$\frac{1}{2}\rho_n(\mathbf{V}) \geq \bar{r}, \ n \in \mathcal{N}_1$$

and

$$\frac{1}{2}\rho_{\pi(n)}(\mathbf{V}) + r_{N+n} \geq \bar{r}, \ n \in \mathcal{N}_1,$$

where $\bar{r}$ is the throughput threshold, set for all users. We would like to add that there is no need to impose any constraint on the common throughput r_{N+n}, because a zero or nearly zero common throughput implies that the common message s_{N+n} is allocated zero power, i.e., it is effectively not included in the transmit signal. Later, we will demonstrate through simulations that non-zero common-throughput results in the superior performance of n-NOMA compared to CoSig.

We consider a single sum-power constraint in (22c) instead of considering N_t individual per-antenna power constraints, although the latter is helpful for achieving transmit-power balance across the transmit antennas, which may be required owing to physical antenna limitations. However, by solving the proposed GM-throughput optimization problem with a single sum-power constraint, we observe another advantage of achieving even power distribution across the transmit antennas without enforcing per-antenna power constraints. These advantages will be demonstrated in Section V through simulation results.

Since the optimal solution of (22) is attained at the equality sign of the constraint (13), the problem (22) is equivalent to

$$\max_{\mathbf{V}} \left[\prod_{n \in \mathcal{N}_1} \left(\frac{1}{2}\rho_n(\mathbf{V}) \left(\frac{1}{2}\rho_{\pi(n)}(\mathbf{V}) + \frac{1}{2}\rho_{N+n}(\mathbf{V}) \right) \right) \right]^{1/N} \tag{23a}$$

$$\text{s.t.} \quad (22c), \tag{23b}$$

with $\rho_{N+n}(\mathbf{V})$ defined from (14). It can be readily checked that

$$\frac{1}{2}\left(\rho_{\pi(n)}(\mathbf{V}) + \rho_{N+n}(\mathbf{V})\right) = \frac{\min\{\rho_{1,\pi(n)}(\mathbf{V}), \rho_{2,\pi(n)}(\mathbf{V})\}}{2} \tag{24}$$

with

$$\rho_{i,\pi(n)}(\mathbf{V}) \triangleq \rho_{\pi(n)}(\mathbf{V}) + \rho_{i,N+n}(\mathbf{V}), i = 1, 2.$$

Therefore, the problem (23a) can be written as :

$$\max_{\mathbf{V}} \ g_{\text{GM}}(\mathbf{V}) \triangleq \frac{1}{2}\left[\prod_{n \in \mathcal{N}_1} \left(\rho_n(\mathbf{V}) \min_{i=1,2} \rho_{i,\pi(n)}(\mathbf{V}) \right) \right]^{1/N} \tag{25a}$$

$$\text{s.t.} \quad (22c). \tag{25b}$$

As discussed before, NOMA is a particular case of n-NOMA with

$$\mathbf{V}_{\pi(n)} \equiv 0, n \in \mathcal{N}_1, \tag{26}$$

because both UE n and UE $\pi(n)$ fully decode the information message designated for UE $\pi(n)$. The throughput at UE $\pi(n)$ is defined accordingly as $(1/2)\min_{i=1,2}\rho_{i,N+n}(\mathbf{V})$. The GM-throughput maximization problem corresponding to NOMA is thus the following particular case of (25):

$$\max_{\mathbf{V}} \ g_{\text{GM}}(\mathbf{V}) \triangleq \frac{1}{2}\left(\prod_{n \in \mathcal{N}_1} \left(\rho_n(\mathbf{V}) \min_{i=1,2} \rho_{i,N+n}(\mathbf{V}) \right) \right)^{1/N}$$

$$\text{s.t.} \quad (22c), (26). \tag{27}$$

Conversely, CoSig, where every UE decodes only its designated information message and regards the others as

[4]The objective in (22a) is the GM of N users' throughput.

interference, represents a special scenario of the aforementioned n-NOMA framework with

$$\mathbf{V}_{N+n} \equiv 0, n \in \mathcal{N}. \tag{28}$$

The corresponding GM-throughput maximization problem is the following particular case of (25):

$$\max_{\mathbf{V}} \; g(\mathbf{V}) \triangleq \frac{1}{2}\left(\prod_{n\in\mathcal{N}} \rho_n(\mathbf{V})\right)^{1/N} \quad \text{s.t.} \quad (22c), (28), \tag{29}$$

with $\rho_n(\mathbf{V}) = \ln|I_{2N_r} + [\bar{H}_n\mathbf{V}_n]^2(\sum_{n'\in\mathcal{N}\setminus\{n\}}[\bar{H}_n\mathbf{V}_{n'}]^2 + \sigma I_{2N_r})^{-1}|$.

III. Proposed Solution

The non-smooth function $\min_{i=1,2}\rho_{i,\pi(n)}(\mathbf{V})$ in (25) presents a challenge for solving the problem in (25) with a computation procedure that is based on closed-form expressions. This challenge shall be tackled in this section.

By the arithmetic-geometric mean inequality, it follows that

$$\frac{\rho_{1,\pi(n)}(\mathbf{V}) + \rho_{2,\pi(n)}(\mathbf{V})}{2} \geq \sqrt{\rho_{1,\pi(n)}(\mathbf{V})\rho_{2,\pi(n)}(\mathbf{V})}, \tag{30}$$

with equality holding at $\rho_{1,\pi(n)}(\mathbf{V}) = \rho_{2,\pi(n)}(\mathbf{V})$. This means that maximizing the left-hand side of (30) results in unequal values of $\rho_{1,\pi(n)}(\mathbf{V})$ and $\rho_{2,\pi(n)}(\mathbf{V})$, thereby reduces their minimum. By contrast, maximizing the right-hand side (RHS) of (30) ensure balanced values of $\rho_{1,\pi(n)}(\mathbf{V})$ and $\rho_{2,\pi(n)}(\mathbf{V})$ and thus improves their minimum. Therefore, we choose the RHS of (30) as a surrogate for $\min_{i=1,2}\rho_{i,\pi(n)}(\mathbf{V})$ in (25). Accordingly, we solve (25) by addressing the following surrogate problem:

$$\max_{\mathbf{V}} \; g(\rho(\mathbf{V})) \quad \text{s.t.} \quad (22c), \tag{31}$$

where $g(\rho(\mathbf{V}))$ is the composite of the function

$$g(\rho) \triangleq \left(\prod_{n\in\mathcal{N}_1}\left(\rho_n\sqrt{\rho_{1,\pi(n)}\rho_{2,\pi(n)}}\right)\right)^{1/N} \tag{32}$$

of $N + N/2$ variables and of the mapping

$$\begin{aligned}\rho(\mathbf{V}) \triangleq (&\rho_1(\mathbf{V}), \rho_{1,\pi(1)}(\mathbf{V}), \rho_{2,\pi(1)}(\mathbf{V}), \ldots,\\ &\rho_{N/2}(\mathbf{V}), \rho_{1,\pi(N/2)}(\mathbf{V}), \rho_{2,\pi(N/2)}(\mathbf{V})).\end{aligned} \tag{33}$$

We will now propose a low-complexity iterative method to compute (31). Starting with the feasible point $V^{(0)}$, for $p = 1, \ldots, p_{\max}$, let $V^{(p)}$ be its feasible point obtained from the $(p-1)$-st iteration, and $\rho^{(p)} \triangleq \rho(V^{(p)})$. At the p-th iteration, we seek for a steep ascent by considering the following problem of nonlinear function optimization

$$\max_{\mathbf{V}} \; \sum_{k=1}^{K}\left(\frac{\partial g(\rho^{(p)})}{\partial\rho_n}\rho_n(\mathbf{V}) + \sum_{i=1}^{2}\frac{\partial g(\rho^{(p)})}{\partial\rho_{i,\pi(n)}}\rho_{i,\pi(n)}(\mathbf{V})\right) \quad \text{s.t.} \quad (22c). \tag{34}$$

Since $\frac{\partial g(\rho^{(p)})}{\partial\rho_n} = \frac{g(\rho^{(p)})}{N\rho_n(V^{(p)})}$ and $\frac{\partial g(\rho^{(p)})}{\partial\rho_{i,\pi(n)}} = \frac{g(\rho^{(p)})}{2N\rho_{i,\pi(n)}(V^{(p)})}$, $i = 1, 2$, with $g(\rho^{(p)}) > 0$, the problem (34) is in fact the following problem

$$\max_{\mathbf{V}} \; g^{(p)}(\mathbf{V}) \triangleq \sum_{n\in\mathcal{N}_1}\left[\gamma_n^{(p)}\rho_n(\mathbf{V}) + \sum_{i=1}^{2}\gamma_{i,\pi(n)}^{(p)}\rho_{i,\pi(n)}(\mathbf{V})\right] \quad \text{s.t.} \quad (22c), \tag{35}$$

where we have:

$$\begin{aligned}\gamma_n^{(p)} &= \tfrac{\rho_{\max}^{(p)}}{\rho_n(V^{(p)})}, n \in \mathcal{N}_1,\\ \gamma_{i,\pi(n)}^{(p)} &= \tfrac{\rho_{\max}^{(p)}}{2\rho_{i,\pi(n)}(V^{(p)})}, i = 1, 2; n \in \mathcal{N}_1,\end{aligned} \tag{36}$$

for

$$\rho_{\max}^{(p)} \triangleq \max_{n\in\mathcal{N}_1}\max\{\rho_n(V^{(p)}), \rho_{1,\pi(n)}(V^{(p)}), \rho_{2,\pi(n)}(V^{(p)})\}. \tag{37}$$

The problem (35) is non-convex because of the non-concave nature of the objective function. By using the inequality (91) in the Appendix for $X = \bar{H}_n\mathbf{V}_n$, $Y = \Psi_n^{-1}(\mathbf{V})$, $\bar{X} = \bar{H}_nV^{(p)}$, and $\bar{Y} = \Psi_n^{-1}(V^{(p)})$ leads to the following tight concave quadratic minorant of $\rho_n(\mathbf{V})$ at $V^{(p)}$:

$$\begin{aligned}\rho_n^{(p)}(\mathbf{V}) \triangleq\; &\rho_n(V^{(p)}) - \langle[\bar{H}_nV_n^{(p)}]^2\Psi_n^{-1}(V^{(p)})\rangle + 2\langle B_n^{(p)}, \mathbf{V}_n\rangle\\ &- \left\langle C_n^{(p)}, \sum_{n'\in\mathcal{N}\setminus\{N+n\}}[\bar{H}_n\mathbf{V}_{n'}]^2 + \sigma I_{2N_r}\right\rangle \qquad (38)\\ =\; &a_n^{(p)} + 2\langle B_n^{(p)}, \mathbf{V}_n\rangle - \sum_{n'\in\mathcal{N}_E}\langle\mathcal{C}_{n,n'}^{(p)}, [\mathbf{V}_{n'}]^2\rangle, \qquad (39)\end{aligned}$$

where

$$\begin{aligned}a_n^{(p)} &\triangleq \rho_n(V^{(p)}) - \langle[\bar{H}_nV_n^{(p)}]^2\Psi_n^{-1}(V^{(p)})\rangle - \sigma\langle C_n^{(p)}\rangle,\\ B_n^{(p)} &\triangleq \bar{H}_n^H\Psi_n^{-1}(V^{(p)})\bar{H}_nV_n^{(p)},\\ C_n^{(p)} &\triangleq \Psi_n^{-1}(V^{(p)}) - \left(\Psi_n(V^{(p)}) + [\bar{H}_nV_n^{(p)}]^2\right)^{-1},\\ \mathcal{C}_{n,n'}^{(p)} &= \begin{cases} 0_{2N_r\times 2N_r} & \text{for } n' = N+n\\ \bar{H}_n^HC_n^{(p)}\bar{H}_n & \text{otherwise.}\end{cases}.\end{aligned} \tag{40}$$

We can observe from (39) that $\rho_n(\mathbf{V}) \geq \rho_n^{(p)}(\mathbf{V})$ for all $\mathbf{V}$ and $\rho_n(V^{(p)}) = \rho_n^{(p)}(V^{(p)})$, i.e., the inequality becomes equality at $\mathbf{V} = V^{(p)}$. Analogously, we derive the following tight concave quadratic minorant of $\rho_{i,N+n}(\mathbf{V})$ at $V^{(p)}$:

$$\begin{aligned}\rho_{i,N+n}^{(p)}(\mathbf{V}) \;\triangleq\;& a_{i,N+n}^{(p)} + 2\langle B_{i,N+n}^{(p)}, \mathbf{V}_{N+n}\rangle\\ &- \sum_{n'\in\mathcal{N}_E}\langle\mathcal{C}_{i,N+n,n'}^{(p)}, [\mathbf{V}_{n'}]^2\rangle,\end{aligned} \tag{41}$$

where

$$\begin{aligned}a_{i,N+n}^{(p)} &\triangleq \rho_{i,N+n}(V^{(p)}) - \sigma\langle C_{i,N+n}^{(p)}\rangle\\ &\quad -\langle[\bar{H}_{i,n}V_{N+n}^{(p)}]^2\Psi_{i,N+n}^{-1}(V^{(p)})\rangle,\\ B_{i,N+n}^{(p)} &\triangleq \bar{H}_{i,n}^H\Psi_{i,N+n}^{-1}(V^{(p)})\bar{H}_{i,n}V_{N+n}^{(p)},\\ C_{i,N+n}^{(p)} &\triangleq \Psi_{i,N+n}^{-1}(V^{(p)})\\ &\quad - \left(\Psi_{i,N+n}(V^{(p)}) + [\bar{H}_{i,n}V_{N+n}^{(p)}]^2\right)^{-1},\\ \mathcal{C}_{i,N+n,n'}^{(p)} &\equiv \bar{H}_{i,n}^HC_{i,N+n}^{(p)}\bar{H}_{i,n}, \quad n' \in \mathcal{N}_E,\end{aligned} \tag{42}$$

for $i = 1, 2$. From (39) and (41), a tight concave quadratic minorant of $\rho_{i,\pi(n)}$ at $V^{(p)}$ is given by

$$\begin{aligned}\rho^{(p)}_{i,\pi(n)}(\mathbf{V}) &\triangleq \rho^{(p)}_{\pi(n)}(\mathbf{V}) + \rho^{(p)}_{i,N+n}(\mathbf{V}) \\ &= a^{(p)}_{i,\pi(n)} + 2\langle B^{(p)}_{\pi(n)}, \mathbf{V}_{\pi(n)}\rangle + 2\langle B^{(p)}_{i,N+n}, \mathbf{V}_{N+n}\rangle \\ &\quad - \sum_{n' \in \mathcal{N}_E} \langle \mathcal{C}^{(p)}_{i,\pi(n),n'}, [\mathbf{V}_{n'}]^2\rangle, \end{aligned} \tag{43}$$

where $a^{(p)}_{i,\pi(n)} \triangleq a^{(p)}_{\pi(n)} + a^{(p)}_{i,N+n}$, and $\mathcal{C}^{(p)}_{i,\pi(n),n'} = \mathcal{C}^{(p)}_{\pi(n),n'} + \mathcal{C}^{(p)}_{i,N+n,n'}$, for $i = 1, 2$. In summary, a tight concave quadratic minorant of $g^{(p)}(\mathbf{V})$ at $V^{(p)}$ is

$$\begin{aligned}\tilde{g}^{(p)}(\mathbf{V}) &\triangleq \sum_{n \in \mathcal{N}_1} \left[\gamma^{(p)}_n \rho^{(p)}_n(\mathbf{V}) + \sum_{i=1}^{2} \gamma^{(p)}_{i,\pi(n)} \rho^{(p)}_{i,\pi(n)}(\mathbf{V})\right] \\ &= a^{(p)} + 2\sum_{n \in \mathcal{N}_E} \langle \tilde{B}^{(p)}_n, \mathbf{V}_n\rangle - \sum_{n \in \mathcal{N}_E} \langle \tilde{C}^{(p)}_n, [\mathbf{V}_n]^2\rangle, \end{aligned} \tag{44}$$

for

$$\begin{aligned} a^{(p)} &\triangleq \textstyle\sum_{n \in \mathcal{N}_1} \left(\gamma^{(p)}_n a^{(p)}_n + \sum_{i=1}^{2} \gamma^{(p)}_{i,\pi(n)} a^{(p)}_{i,\pi(n)}\right), \\ \tilde{B}^{(p)}_n &\triangleq \begin{cases} \gamma^{(p)}_n B^{(p)}_n & \text{for } n \in \mathcal{N}_1 \\ (\sum_{i=1}^{2} \gamma^{(p)}_{i,\pi(n-N/2)}) B^{(p)}_{\pi(n-N/2)} & \text{for } n \in \mathcal{N}_2 \\ \sum_{i=1}^{2} \gamma^{(p)}_{i,\pi(n-N)} B^{(p)}_{i,n} & \text{for } n > N, \end{cases} \\ \tilde{C}^{(p)}_n &\triangleq \textstyle\sum_{n' \in \mathcal{N}_1} \left(\gamma^{(p)}_{n'} \mathcal{C}^{(p)}_{n',n} + \sum_{i=1}^{2} \gamma^{(p)}_{i,\pi(n')} \mathcal{C}^{(p)}_{i,\pi(n'),n}\right), \\ &\qquad n \in \mathcal{N}_E. \end{aligned} \tag{45}$$

Hence, we address the following tight minorant maximization problem to compute the next iterative point $V^{(p+1)}$

$$\max_{\mathbf{V}} \quad \tilde{g}^{(p)}(\mathbf{V}) \quad \text{s.t.} \quad (22c), \tag{46}$$

which has a closed-form solution[5]

$$V^{(p+1)}_n = \begin{cases} (\tilde{C}^{(p)}_n)^{-1} \tilde{B}^{(p)}_n & \text{if } \Xi^{(p)} \le 2P, \\ (\tilde{C}^{(p)}_n + \lambda I_{2N_t})^{-1} \tilde{B}^{(p)}_n & \text{otherwise}, \end{cases} \tag{47}$$

where $\Xi^{(p)} \triangleq \sum_{n \in \mathcal{N}_E} ||(\tilde{C}^{(p)}_n)^{-1} \tilde{B}^{(p)}_n||^2$ and $\lambda > 0$ is found by bisection such that $\sum_{n \in \mathcal{N}_E} ||(\tilde{C}^{(p)}_n + \lambda I_{2N_t})^{-1} \tilde{B}^{(p)}_n||^2 = 2P$. The solution obtained in (47) can be explained as follows. If the power constraint (22c) is satisfied, we can obtain $V^{(p+1)}_n = (\tilde{C}^{(p)}_n)^{-1} \tilde{B}^{(p)}_n$ by computing the partial derivative of $\tilde{g}^{(p)}(\mathbf{V})$ with respect to $\mathbf{V}_n$ and setting it with zero. If the power constraint (22c) is violated, we can obtain $V^{(p+1)}_n$ by using the following Lagrangian function:

$$\begin{aligned}\mathcal{L}(\mathbf{V}, \lambda) = a^{(p)} &+ 2\sum_{n \in \mathcal{N}_E} \langle \tilde{B}^{(p)}_n, \mathbf{V}_n\rangle - \sum_{n \in \mathcal{N}_E} \langle \tilde{C}^{(p)}_n, [\mathbf{V}_n]^2\rangle \\ &- \lambda \left(\sum_{n \in \mathcal{N}_E} \|\mathbf{V}_n\|^2 - 2P\right) \end{aligned} \tag{48}$$

and taking its derivative with respect to $\mathbf{V}_n$, i.e.,

$$\frac{\partial \mathcal{L}(\mathbf{V}, \lambda)}{\partial \mathbf{V}_n} = 2\tilde{B}^{(p)}_n - 2\tilde{C}^{(p)}_n \mathbf{V}_n - 2\lambda \mathbf{V}_n. \tag{49}$$

[5] $(\tilde{C}^{(p)}_n)^{-1}$ is understood as the pseudo-inverse of $\tilde{C}^{(p)}_n$, when we have $\tilde{C}^{(p)}_n \succeq 0$.

Algorithm 1 GM throughput maximization algorithm for computing (31)

1: **Initialization:** Initialize a random feasible $V^{(0)}$ for (22c). Set $p = 0$.
2: **Repeat** the following until the objective function in (31) converges:
Update $\gamma^{(p)}_n$ and $\gamma^{(p)}_{i,\pi(n)}$ by (36). Update $\tilde{B}^{(p)}_n$ and $\tilde{C}^{(p)}_n$ by (45). Generate $V^{(p+1)}_n$ by (47). Reset $p := p+1$.
3: **Output** $V^{(p)}$.

Finally, we can obtain $V^{(p+1)}_n = (\tilde{C}^{(p)}_n + \lambda I_{2N_t})^{-1} \tilde{B}^{(p)}_n$ by solving $\frac{\partial \mathcal{L}(\mathbf{V}, \lambda)}{\partial \mathbf{V}_n} = 0$. The computational complexity of the closed-form based solution in (47) is $\mathcal{O}(3N_t \log_2(2N_t)N)$. The average number of iterations needed for convergence is reported later in the simulation results section, while its convergence is analyzed below:

Proposition 1:
Algorithm 1 produces a sequence $\{V^{(p+1)}\}$ of progressively improved points of (31), which converges to a Karush-Kuhn-Tucker (KKT) point.

Proof:
Since $\tilde{g}^{(p)}(\mathbf{V})$ is a tight concave quadratic minorant of $g^{(p)}(\mathbf{V})$ at $V^{(p)}$, we have

$$\begin{aligned} g^{(p)}(\mathbf{V}) &\ge \tilde{g}^{(p)}(\mathbf{V}) \\ g^{(p)}(V^{(p)}) &= \tilde{g}^{(p)}(V^{(p)}), \end{aligned}$$

Therefore,

$$\begin{aligned} g^{(p)}(V^{(p+1)}) &\ge \tilde{g}^{(p)}(V^{(p+1)}) \\ &> \tilde{g}^{(p)}(V^{(p)}) \\ &= g^{(p)}(V^{(p)}), \end{aligned} \tag{50}$$

where the second inequality arises from the fact that $V^{(p+1)}$ and $V^{(p)}$ are the optimal solution and a feasible point for (35), respectively. This demonstrates that $V^{(p+1)}$ is a better feasible point than $V^{(p+1)}$ for the problem (35). Moreover, the sequence $\{V^{(p)}\}$ is bounded by the constraints (22c). By Cauchy's theorem[6], there is a convergent subsequence $\{V^{(p_\nu)}\}$ with a limit point $\bar{V}$, i.e.,

$$\lim_{\nu \to +\infty} \left[g^{(p)}(V^{(p_\nu)}) - g^{(p)}(\bar{V})\right] = 0.$$

For every p, there is a ν such that $p_\nu \le p \le p_{\nu+1}$, so

$$\begin{aligned} 0 &= \lim_{\nu \to +\infty} \left[g^{(p)}(V^{(p_\nu)}) - g^{(p)}(\bar{V})\right] \\ &\le \lim_{p \to +\infty} \left[g^{(p)}(V^{(p)}) - g^{(p)}(\bar{V})\right] \\ &\le \lim_{\nu \to +\infty} \left[g^{(p)}(V^{(p_{\nu+1})}) - g^{(p)}(\bar{V})\right] \\ &= 0, \end{aligned}$$

[6] From a bounded (compact) sequence, there exists a convergent subsequence

which demonstrates that $\lim_{p\to+\infty} g^{(p)}(V^{(p)}) = g^{(p)}(\bar{V})$. Each accumulation point $\{\bar{V}\}$ of the sequence $\{V^{(p)}\}$ is a KKT-point, as stated in [35, Theorem 1].

Since both (31) and (34)/(35) share the same first-order optimality condition, it can be easily shown that Algorithm 1 converges to a local solution that satisfies the first-order optimality condition of (31). Additionally, simulations consistently show that $g(V^{(p+1)}) > g(V^{(p)})$. Since Algorithm 1 is in fact a path-following algorithm generating progressively better feasible points, it often yields a global solution of (31) [36]. ■

Remark 1. For completeness, we develop a path-following algorithm below that uses a convex-solver to solve (25). This can serve as a benchmark for the aforementioned Algorithm 1, which is based on closed-form expressions for solving the surrogate problem (31). By introducing the slack variables $\lambda \triangleq (\lambda_1, \ldots, \lambda_N)$ satisfying the constraints

$$\lambda_i > 0, i = 1, \ldots, N, \tag{51}$$

we can express the problem (25) as

$$\max_{\mathbf{V},\lambda} \ \frac{1}{2}(\prod_{n=1}^{N} \lambda_n)^{1/N} \tag{52a}$$

$$\text{s.t. } (22c), (51), \tag{52b}$$

$$\rho_n(\mathbf{V}) \geq \lambda_n, n \in \mathcal{N}_1, \tag{52c}$$

$$\rho_{\pi(n)}(\mathbf{V}) + \rho_{i,N+n}(\mathbf{V}) \geq \lambda_{\pi(n)}, n \in \mathcal{N}_1, i = 1, 2. \tag{52d}$$

Initialized by a point $V^{(0)}$ feasible for (22c), let $V^{(p)}$ be a point feasible for (25) that is obtained from the $(p-1)$-th iteration. For the tight concave quadratic minorants $\rho_n^{(p)}(\mathbf{V})$ and $\rho_{i,N+n}^{(p)}(\mathbf{V})$ defined from (39) and (41), we solve the following convex problem at the p-th iteration, which provides a tight inner approximation of the nonconvex problem (52d), to generate the next feasible point $V^{(p+1)}$:

$$\max_{\mathbf{V},\lambda} \ \frac{1}{2}(\prod_{n=1}^{N} \lambda_n)^{1/N} \tag{53a}$$

$$\text{s.t. } (22c), (51), \tag{53b}$$

$$\rho_n^{(p)}(\mathbf{V}) \geq \lambda_n, n \in \mathcal{N}_1, \tag{53c}$$

$$\rho_{\pi(n)}^{(p)}(\mathbf{V}) + \rho_{i,N+n}^{(p)}(\mathbf{V}) \geq \lambda_{\pi(n)}, n \in \mathcal{N}_1, i = 1, 2. \tag{53d}$$

As long as $g_{GM}(V^{(p+1)}) \neq g_{GM}(V^{(p)})$, we have $g_{GM}(V^{(p+1)}) > g_{GM}(V^{(p)})$, indicating that the algorithm based on (53d) produces a path-following sequence of successively improved feasible points. This sequence is guaranteed to converge to at least a locally optimal solution [35], which in many cases has been observed to coincide with the global optimum [36]. Therefore, the path-following algorithm based on (53d) can serve as a benchmark for Algorithm 1 that is based on closed-form expressions. This is because the path-following algorithm directly addresses the original GM-throughput optimization problem (25) by solving the convex problem (53d), which is a tight inner approximation of (52d), or indirectly (25). On the other hand, Algorithm 1 addresses (25) by solving the surrogate problem (31). Therefore, we can expect some performance disadvantage by solving Algorithm 1 compared to the solution of the path-following algorithm. However, the former promises huge computational advantages as explained in the following paragraph.

The computational complexity of (53d) is $\mathcal{O}\left((6N_tN_rN + N)^3N\right)$ [37], which is huge compared to the computational complexity of the closed-form based solution in (47), which is as low as $\mathcal{O}\left(3N_t \log_2(2N_t)N\right)$.

A general-purpose projective gradient method for maximizing GM-throughput (see, for example, [38], which focuses on power allocation rather than beamforming) is excessively computationally complex. This is primarily because GM-throughput is a high-degree function of the decision variables, making it difficult to estimate the Lipschitz constant of its gradient—an essential parameter for ensuring convergence. Additionally, applying the Armijo rule for line search to adjust the step size presents further implementation challenges.

Remark 2. The conventional LB corresponds to $\mathbf{W}_{2,n'} \equiv 0$ in (3), so $\mathbf{V}_{n'}$ in (5) is structured as $\mathbf{V}_{n'} = \begin{bmatrix} \Re\{\mathbf{W}_{1,n'}\} & -\Im\{\mathbf{W}_{1,n'}\} \\ \Im\{\mathbf{W}_{1,n'}\} & \Re\{\mathbf{W}_{1,n'}\} \end{bmatrix}$. Therefore, LB cannot be implemented as a particular case of Algorithm 1, which is based on a structure-free $\mathbf{V}_{n'}$ in (5) to derive the closed-form expression (47) for the optimal solution of the problem (46). To overcome this, we introduce the real-composite form of $\mathbf{W}_{n'}$:

$$\mathbf{V}_{1,n'} \triangleq \begin{bmatrix} \mathbf{V}_{n'}^{11} \\ \mathbf{V}_{n'}^{21} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{W}_{1,n'}\} \\ \Im\{\mathbf{W}_{1,n'}\} \end{bmatrix}, n' \in \mathcal{N}_E. \tag{54}$$

By partitioning $\tilde{B}_n^{(p)}$ and $\tilde{C}_n^{(p)}$ defined from (45) as $\tilde{B}_n^{(p)} = [\tilde{B}_{n,ij}^{(p)}]_{(i,j)\in\{1,2\}}$, $\tilde{B}_{n,ij}^{(p)} \in \mathbb{R}^{N_t\times N_r}$, $\tilde{C}_n^{(p)} = [\tilde{C}_{n,ij}^{(p)}]_{(i,j)\in\{1,2\}}$, $\tilde{C}_{n,ij}^{(p)} \in \mathbb{R}^{N_t\times N_t}$, $i, j = 1, 2$, we can express $\tilde{g}^{(p)}(\mathbf{V})$ in (44) by

$$\tilde{g}^{(p)}(\mathbf{V}_1) = a^{(p)} + 2\sum_{n\in\mathcal{N}_E} \langle \tilde{B}_{1,n}^{(p)}, \mathbf{V}_{1,n}\rangle - \sum_{n\in\mathcal{N}_E} \langle \tilde{C}_{1,n}^{(p)}, [\mathbf{V}_{1,n}]^2\rangle, \tag{55}$$

for $\mathbf{V}_1 \triangleq \{\mathbf{V}_{1,n}, n \in \mathcal{N}_E\}$, and

$$\tilde{B}_{1,n}^{(p)} = \begin{bmatrix} \tilde{B}_{n,11}^{(p)} + \tilde{B}_{n,22}^{(p)} \\ \tilde{B}_{n,21}^{(p)} - \tilde{B}_{n,12}^{(p)} \end{bmatrix},$$
$$\tilde{C}_{1,n}^{(p)} = \begin{bmatrix} \tilde{C}_{n,11}^{(p)} + \tilde{C}_{n,22}^{(p)} & \tilde{C}_{n,12}^{(p)} - (\tilde{C}_{n,12}^{(p)})^T \\ (\tilde{C}_{n,12}^{(p)})^T - \tilde{C}_{n,12}^{(p)} & \tilde{C}_{n,11}^{(p)} + \tilde{C}_{n,22}^{(p)} \end{bmatrix}, n \in \mathcal{N}_E. \tag{56}$$

In place of (22c), we address the following problem, which maximizes a tight minorant, to obtain the next iterative point $V_1^{(p+1)}$

$$\max_{\mathbf{V}_1} \ \tilde{g}^{(p)}(\mathbf{V}_1) \quad \text{s.t.} \quad \sum_{n\in\mathcal{N}_E} ||\mathbf{V}_{1,n}||^2 \leq P. \tag{57}$$

Algorithm 2 GM throughput maximization algorithm for computing (31) under LB

1: **Initialization:** Initialize a feasible $V_{1,n}^{(0)} = \begin{bmatrix} V_n^{11,(0)} \\ V_n^{21,(0)} \end{bmatrix}$, $n \in \mathcal{N}_E$ for the sum power constraint $\sum_{n\in\mathcal{N}_E} ||\mathbf{V}_{1,n}||^2 \leq P$. Define $V_n^{(0)} \triangleq \begin{bmatrix} V_n^{11,(0)} & -V_n^{21,(0)} \\ V_n^{21,(0)} & V_n^{11,(0)} \end{bmatrix}$, $n \in \mathcal{N}_E$. Set $p = 0$.

2: **Repeat** the following until the objective function in (31) converges: Update $\gamma_n^{(p)}$ and $\gamma_{i,\pi(n)}^{(p)}$, for all $n \in \mathcal{N}_1$, $i = 1, 2$ by (36). Update $\tilde{B}_n^{(p)}$ and $\tilde{C}_n^{(p)}$ by (45) and then $\tilde{B}_{1,n}^{(p)}$ and $\tilde{C}_{1,n}^{(p)}$ by (56). Generate $V_{1,n}^{(p+1)} = \begin{bmatrix} V_n^{11,(p+1)} \\ V_n^{21,(p+1)} \end{bmatrix}$ by (58) and then define $V_n^{(p+1)} \triangleq \begin{bmatrix} V_n^{11,(p+1)} & -V_n^{21,(p+1)} \\ V_n^{21,(p+1)} & V_n^{11,(p+1)} \end{bmatrix}$. Reset $p := p + 1$.

3: **Output** $V_1^{(p)}$.

This admits the closed-form solution

$$V_{1,n}^{(p+1)} = \begin{cases} (\tilde{C}_{1,n}^{(p)})^{-1}\tilde{B}_{1,n}^{(p)} & \text{if } \tilde{\Xi}^{(p)} \leq P, \\ (\tilde{C}_{1,n}^{(p)} + \lambda I_{2N_t})^{-1}\tilde{B}_{1,n}^{(p)} & \text{otherwise,} \end{cases} \tag{58}$$

for $\tilde{\Xi}^{(p)} \triangleq \sum_{n\in\mathcal{N}_E} ||(\tilde{C}_{1,n}^{(p)})^{-1}\tilde{B}_{1,n}^{(p)}||^2$, where $\lambda > 0$ is found by bisection such that $\sum_{n\in\mathcal{N}_E} ||(\tilde{C}_{1,n}^{(p)} + \lambda I_{2N_t})^{-1}\tilde{B}_{1,n}^{(p)}||^2 = P$. Algorithm 2 provides a procedure for computing (31) under LB. The computational complexity of the closed-form based solution in (58) is $\mathcal{O}\left(3N_t \log_2(2N_t)N\right)$. The average number of iterations required before convergence is reported later in the simulation results section.

IV. Extension to URLLC Networks

In this section, by assuming that all UEs are equipped with single antennas, the proposed WLB design for maximizing the GM-throughput is extended to URLLC networks. With $\rho_{i,N+n}(\mathbf{V})$ defined from (9), the URLLC throughputs of decoding the message s_{N+n} by UEs n and $\pi(n)$ are $\frac{1}{2}\rho_{1,N+n}^{\mathrm{U}}(\mathbf{V})$ and $\frac{1}{2}\rho_{2,N+n}^{\mathrm{U}}(\mathbf{V})$ for [39]

$$\rho_{i,N+n}^{\mathrm{U}}(\mathbf{V}) \triangleq \rho_{i,N+n}(\mathbf{V}) - q\sqrt{\nu_{i,N+n}(\mathbf{V})}, \tag{59}$$

where $q \triangleq \frac{2}{\sqrt{\mathcal{B}t_t}} Q_G^{-1}(\epsilon_c)$, $\mathcal{B}$ is the communication bandwidth, t_t denotes the URLLC transmission duration, $Q_G^{-1}(\cdot)$ represents the inverse of the Gaussian Q-function, defined as $Q_G(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)\mathrm{d}t$, ϵ_c is the target decoding error probability, and $\nu_{i,N+n}(\mathbf{V})$ is the channel dispersion [39, eq. (27)] given by

$$\nu_{i,N+n}(\mathbf{V}) \triangleq 2\frac{\varphi_{i,n+N}(\mathbf{V})}{1+\varphi_{i,n+N}(\mathbf{V})} \tag{60}$$

under the signal-to-interference-plus-noise ratio (SINR), $\varphi_{i,n+N}(\mathbf{V})$, formulated as

$$\varphi_{i,n+N}(\mathbf{V}) = \frac{(1/2)\|H_{i,n}\mathbf{V}_{N+n}\|^2}{\sum_{n'\in\mathcal{N}_E\setminus\{N+n\}}(1/2)\|H_{i,n}\mathbf{V}_{n'}\|^2+\sigma}. \tag{61}$$

The achievable URLLC throughput of decoding the message s_{N+n} is $\frac{1}{2}\rho_{N+n}^{\mathrm{U}}(\mathbf{V})$ for

$$\rho_{N+n}^{\mathrm{U}}(\mathbf{V}) \triangleq \min_{i=1,2} \rho_{i,N+n}^{\mathrm{U}}(\mathbf{V}). \tag{62}$$

Similarly, the achievable URLLC throughput of decoding the message s_χ, for $\chi \in \{n, \pi(n)\}$, is given by $\frac{1}{2}\rho_\chi^{\mathrm{U}}(\mathbf{V})$, with

$$\rho_\chi^{\mathrm{U}}(\mathbf{V}) \triangleq \rho_\chi(\mathbf{V}) - q\sqrt{\nu_\chi(\mathbf{V})}, \tag{63}$$

where we have

$$\nu_\chi(\mathbf{V}) \triangleq 2\frac{\varphi_\chi(\mathbf{V})}{1+\varphi_\chi(\mathbf{V})} \tag{64}$$

under the SINR of

$$\varphi_\chi(\mathbf{V}) = \frac{(1/2)\|\bar{H}_\chi\mathbf{V}_\chi\|^2}{\sum_{n'\in\mathcal{N}_E\setminus\{\chi,N+n\}}(1/2)\|\bar{H}_\chi\mathbf{V}_{n'}\|^2+\sigma} \tag{65}$$

Thus, the URLLC throughput at UE $\pi(n)$ is $(1/2)\min\{\rho_{1,\pi(n)}(\mathbf{V}), \rho_{2,\pi(n)}(\mathbf{V})\}$ with

$$\rho_{i,\pi(n)}^{\mathrm{U}}(\mathbf{V}) \triangleq \rho_{\pi(n)}^{\mathrm{U}}(\mathbf{V}) + \rho_{i,N+n}^{\mathrm{U}}(\mathbf{V}), \quad i = 1, 2. \tag{66}$$

Similar to (25), the maximizing of the n-NOMA GM-throughput under URLLC constraints is cast as the following problem:

$$\max_{\mathbf{V}} \ \frac{1}{2}\left[\prod_{n\in\mathcal{N}_1}\left(\rho_n^{\mathrm{U}}(\mathbf{V}) \min_{i=1,2} \rho_{i,\pi(n)}^{\mathrm{U}}(\mathbf{V})\right)\right]^{1/N} \tag{67a}$$

$$\text{s.t.} \quad (22c). \tag{67b}$$

Note that the problem (67) is more challenging to solve compared to (25) because of the additional term $q\sqrt{\nu_n(\mathbf{V})}$ in the definition of the throughput function e.g., (63), which is a complex function of the beamformers. Recently, we proposed a convex-solver based computational procedure for LB design under URLLC [40]. However, a computationally efficient design based on the closed-form expressions of both LB and WLB is still open, which we now develop.

Like (31), we tackle (67) by considering the following surrogate formulation:

$$\max_{\mathbf{V}} \ g(\rho^{\mathrm{U}}(\mathbf{V})) \quad \text{s.t.} \quad (22c), \tag{68}$$

where $g(.)$ is defined from (32) and $\rho^{\mathrm{U}}(\mathbf{V}) \triangleq \big(\rho_1^{\mathrm{U}}(\mathbf{V}), \rho_{1,\pi(1)}^{\mathrm{U}}(\mathbf{V}), \rho_{2,\pi(1)}^{\mathrm{U}}(\mathbf{V}), \ldots, \rho_{N/2}^{\mathrm{U}}(\mathbf{V}), \rho_{1,\pi(N/2)}^{\mathrm{U}}(\mathbf{V}), \rho_{2,\pi(N/2)}^{\mathrm{U}}(\mathbf{V})\big)$.

We now proceed to design an iterative method for solving (68). Initialized by its feasible point $V^{(0)}$, for $p = 1, \ldots,$ let $V^{(p)}$ be its feasible point that is obtained from the $(p-1)$-st iteration, $\rho^{\mathrm{U},(p)} \triangleq \rho^{\mathrm{U}}(V^{(p)})$. Like (35), at the p-th iteration, we seek for a steep ascent by considering the following problem:

$$\max_{\mathbf{V}} \ g^{\mathrm{U},(p)}(\mathbf{V}) \triangleq \sum_{n\in\mathcal{N}_1}\Big[\gamma_n^{\mathrm{U},(p)}\rho_n^{\mathrm{U}}(\mathbf{V}) + \gamma_{1,\pi(n)}^{\mathrm{U},(p)}\rho_{1,\pi(n)}^{\mathrm{U}}(\mathbf{V}) + \gamma_{2,\pi(n)}^{\mathrm{U},(p)}\rho_{2,\pi(n)}^{\mathrm{U}}(\mathbf{V})\Big] \tag{69a}$$

$$\text{s.t.} \ (22c). \tag{69b}$$

where

$$\begin{aligned}\gamma_n^{\mathrm{U},(p)} &= \frac{\rho_{\max}^{\mathrm{U},(p)}}{\rho_n^{\mathrm{U}}(V^{(p)})}, n \in \mathcal{N}_1, \\ \gamma_{i,\pi(n)}^{\mathrm{U},(p)} &= \frac{\rho_{\max}^{\mathrm{U},(p)}}{2\rho_{i,\pi(n)}^{\mathrm{U}}(V^{(p)})}, i = 1,2; n \in \mathcal{N}_1,\end{aligned} \tag{70}$$

for $\rho_{\max}^{\mathrm{U},(p)} \triangleq \max_{n\in\mathcal{N}_1} \max\{\rho_n^{\mathrm{U}}(V^{(p)}), \rho_{1,\pi(n)}^{\mathrm{U}}(V^{(p)}), \rho_{2,\pi(n)}^{\mathrm{U}}(V^{(p)})\}$. Having $\rho_n^{(p)}(\mathbf{V})$ defined in (39), as a tight concave quadratic minorant of $\rho_n(\mathbf{V})$, we have to derive a tight convex quadratic majorant of $q\sqrt{\nu_n(\mathbf{V})}$, which is a complicated function of the beamformers $\mathbf{V}$. To this end, we use (64) and (65) to express $\nu_n(\mathbf{V})$ as $\nu_n(\mathbf{V}) = 2\left(1 - \frac{\alpha_n(\mathbf{V})}{\beta_n(\mathbf{V})}\right)$, for $\alpha_n(\mathbf{V}) \triangleq \sum_{n'\in\mathcal{N}_E\setminus\{n,N+n\}} \|\bar{H}_n\mathbf{V}_{n'}\|^2 + 2\sigma$ and $\beta_n(\mathbf{V}) \triangleq \sum_{n'\in\mathcal{N}_E\setminus\{N+n\}} \|\bar{H}_n\mathbf{V}_{n'}\|^2 + 2\sigma$. Then a tight convex quadratic majorant of $q\sqrt{\nu_n(\mathbf{V})}$ is obtained as

$$\begin{aligned} q\sqrt{\nu_n(\mathbf{V})} &\le q\frac{\sqrt{\nu_n(V^{(p)})}}{2}\left(1 + \frac{2}{\nu_n(V^{(p)})}\right) \\ &\quad - \frac{q}{\sqrt{\nu_n(V^{(p)})}}\frac{\alpha_n(\mathbf{V})}{\beta_n(\mathbf{V})} \qquad (71) \\ &\le q\frac{\sqrt{\nu_n(V^{(p)})}}{2}\left(1 + \frac{2}{\nu_n(V^{(p)})}\right) - \frac{q}{\sqrt{\nu_n(V^{(p)})}} \\ &\quad \times \left[\frac{\alpha_n^{(p)}}{\beta_n^{(p)}}\left(\frac{2\xi_n(\mathbf{V}) + 4\sigma}{\alpha_n^{(p)}} - \frac{\beta_n(\mathbf{V})}{\beta_n^{(p)}}\right)\right] \qquad (72) \\ &= \tilde{a}_n^{(p)} - 2\sum_{n'\in\mathcal{N}_E}\langle \tilde{B}_{n,n'}^{(p)}, \mathbf{V}_{n'}\rangle \\ &\quad - \sum_{n'\in\mathcal{N}_E}\langle \tilde{\mathcal{C}}_{n,n'}^{(p)}, [\mathbf{V}_{n'}]^2\rangle, \qquad (73) \\ &\triangleq \mathfrak{f}^{(p)}(\mathbf{V}), \qquad (74)\end{aligned}$$

where

$$\begin{aligned}
\xi_n(\mathbf{V}) &\triangleq \textstyle\sum_{n'\in\mathcal{N}_E\setminus\{n,N+n\}}\langle (V_n^{(p)})^H\bar{H}_n^H\bar{H}_n\mathbf{V}_{n'}\rangle \\
\tilde{a}_n^{(p)} &\triangleq \frac{q\sqrt{\nu_n(V^{(p)})}}{2}\left(1 + \frac{2}{\nu_n(V^{(p)})}\right) \\
&\quad + (2\sigma)\frac{q\alpha_n^{(p)}}{\beta_n^{(p)}\sqrt{\nu_n(V^{(p)})}}\left(-\frac{2}{\alpha_n^{(p)}} + \frac{1}{\beta_n^{(p)}}\right) \\
\tilde{B}_{n,n'}^{(p)} &\triangleq \begin{cases} \frac{q\bar{H}_n^H\bar{H}_nV_{n'}^{(p)}}{\beta_n^{(p)}\sqrt{\nu_n(V^{(p)})}} & \text{for } n' \in \mathcal{N}_E \neq \{n, N+n\}, \\ 0_{2N_t\times 2N_r} & \text{otherwise} \end{cases} \\
\tilde{\mathcal{C}}_{n,n'}^{(p)} &\triangleq \begin{cases} 0_{2N_r\times 2N_r} & \text{for } n' = N+n, \\ \frac{q}{\left(\beta_n^{(p)}\right)^2\sqrt{\nu_n(V^{(p)})}}\bar{H}_n^H\bar{H}_n & \text{otherwise} \end{cases} \\
\alpha_n^{(p)} &\triangleq \textstyle\sum_{n'\in\mathcal{N}_E\setminus\{n,N+n\}}\|\bar{H}_\chi V_{n'}^{(p)}\|^2 + 2\sigma \\
\beta_n^{(p)} &\triangleq \textstyle\sum_{n'\in\mathcal{N}_E\setminus\{N+n\}}\|\bar{H}_\chi V_{n'}^{(p)}\|^2 + 2\sigma,
\end{aligned} \tag{75}$$

where the inequality (71) is obtained by applying the inequality (92) in the Appendix for $x = \nu_n(\mathbf{V})$ and $\bar{x} = \nu_n(V^{(p)})$. Next, we obtain the inequality (72) from (71) by applying the inequality (93) in the Appendix. Note that the inequalities (71) and (72) become equalities at $\mathbf{V} = V^{(p)}$, which implies that $\mathfrak{f}^{(p)}(\mathbf{V})$ in (74) matches $q\sqrt{\nu_n(\mathbf{V})}$ at $\mathbf{V} = V^{(p)}$, so the former provides a tight majornat of the latter [27]. By combining (39) and (73), we achieve the following tight concave quadratic minorant of $\rho_n^{\mathrm{U}}(\mathbf{V})$:

$$\begin{aligned}\rho_n^{\mathrm{U},(p)}(\mathbf{V}) \triangleq a_n^{\mathrm{U},(p)} &+ 2\textstyle\sum_{n'\in\mathcal{N}_E}\langle B_{n,n'}^{\mathrm{U},(p)}, \mathbf{V}_{n'}\rangle \\ &- \textstyle\sum_{n'\in\mathcal{N}_E}\langle \mathcal{C}_{n,n'}^{\mathrm{U},(p)}, [\mathbf{V}_{n'}]^2\rangle,\end{aligned} \tag{76}$$

where

$$\begin{aligned}
a_n^{\mathrm{U},(p)} &\triangleq a_n^{(p)} - \tilde{a}_n^{(p)}, \\
B_{n,n'}^{\mathrm{U},(p)} &\triangleq \begin{cases} \frac{q\bar{H}_n^H\bar{H}_nV_{n'}^{(p)}}{\beta_n^{(p)}\sqrt{\nu_n(V^{(p)})}} & \text{for } n' \in \mathcal{N}_E \neq \{n, N+n\}, \\ B_n^{(p)} & \text{for } n' = n \\ 0_{2N_t\times 2N_r} & \text{otherwise} \end{cases} \\
\mathcal{C}_{n,n'}^{\mathrm{U},(p)} &\triangleq \begin{cases} 0_{2N_r\times 2N_r} & \text{for } n' = N+n \\ \bar{H}_n^H\left(C_n^{(p)} + \Upsilon_n^{(p)}\right)\bar{H}_n & \text{otherwise} \end{cases}
\end{aligned} \tag{77}$$

with $\Upsilon_n^{(p)} \triangleq \frac{q\alpha_n^{(p)}}{\left(\beta_n^{(p)}\right)^2\sqrt{\nu_n(V^{(p)})}}I_{2N_r}$. Similarly, the following tight concave minorant of $\rho_{i,N+n}^{\mathrm{U}}(\mathbf{V})$ is obtained:

$$\begin{aligned}\rho_{i,N+n}^{\mathrm{U},(p)}(\mathbf{V}) \triangleq a_{i,N+n}^{\mathrm{U},(p)} &+ 2\sum_{n'\in\mathcal{N}_E}\langle B_{i,N+n,n'}^{\mathrm{U},(p)}, \mathbf{V}_{n'}\rangle \\ &- \sum_{n'\in\mathcal{N}_E}\langle \mathcal{C}_{i,N+n,n'}^{\mathrm{U},(p)}, [\mathbf{V}_{n'}]^2\rangle,\end{aligned} \tag{78}$$

where

$$\begin{aligned}
a_{i,N+n}^{\mathrm{U},(p)} &\triangleq a_{i,N+n}^{(p)} - \tilde{a}_{i,N+n}^{(p)} \\
\tilde{a}_{i,N+n}^{(p)} &\triangleq \frac{q\sqrt{\nu_{i,N+n}(V^{(p)})}}{2}\left(1 + \frac{2}{\nu_{i,N+n}(V^{(p)})}\right) \\
&\quad + (2\sigma)\frac{q\alpha_{i,N+n}^{(p)}}{\beta_{i,N+n}^{(p)}\sqrt{\nu_{i,N+n}(V^{(p)})}} \\
&\quad \times \left(-\frac{2}{\alpha_{i,N+n}^{(p)}} + \frac{1}{\beta_{i,N+n}^{(p)}}\right) \\
B_{i,n+N,n'}^{\mathrm{U},(p)} &\triangleq \begin{cases} \frac{qH_{i,n}^HH_{i,n}V_{n'}^{(p)}}{\beta_{i,n+N}^{(p)}\sqrt{\nu_{i,n+N}(V^{(p)})}} & n' \in \mathcal{N}_E \neq \{n+N\}, \\ B_{i,n+N}^{(p)} & n' = n+N \end{cases} \\
\mathcal{C}_{i,n+N,n'}^{\mathrm{U},(p)} &\triangleq \begin{cases} 0_{2N_r\times 2N_r} & n' = N+n \\ H_{i,n}^H\Upsilon_{i,n+N}^{(p)}H_{i,n} & \text{otherwise} \end{cases} \\
\alpha_{i,N+n}^{(p)} &\triangleq \textstyle\sum_{n'\in\mathcal{N}_E\setminus\{N+n\}}\|H_{i,n}V_{n'}^{(p)}\|^2 + 2\sigma \\
\beta_{i,N+n}^{(p)} &\triangleq \textstyle\sum_{n'\in\mathcal{N}_E}\|H_{i,n}V_{n'}^{(p)}\|^2 + 2\sigma
\end{aligned} \tag{79}$$

with $\Upsilon_{i,n+N}^{(p)} \triangleq C_{i,N+n}^{(p)} + \frac{q\alpha_{i,N+n}^{(p)}}{\left(\beta_{i,N+n}^{(p)}\right)^2\sqrt{\nu_{i,N+n}(V^{(p)})}}I_{2N_r}$.

Therefore, a tight concave quadratic minorant of $\rho_{i,\pi(n)}^{\mathrm{U}}(\mathbf{V})$ defined from (66) is given by

$$\begin{aligned}\rho_{i,\pi(n)}^{\mathrm{U},(p)}(\mathbf{V}) \triangleq a_{i,\pi(n)}^{\mathrm{U},(p)} &+ 2\sum_{n'\in\mathcal{N}_E}\langle B_{i,\pi(n),n'}^{\mathrm{U},(p)}, \mathbf{V}_{n'}\rangle \\ &- \sum_{n'\in\mathcal{N}_E}\langle \mathcal{C}_{i,\pi(n),n'}^{\mathrm{U},(p)}, [\mathbf{V}_{n'}]^2\rangle\rangle,\end{aligned} \tag{80}$$

where $a_{i,\pi(n)}^{\mathrm{U},(p)} \triangleq a_{\pi(n)}^{\mathrm{U},(p)} + a_{i,N+n}^{\mathrm{U},(p)}$, $B_{i,\pi(n),n'}^{\mathrm{U},(p)} = B_{\pi(n),n'}^{\mathrm{U},(p)} + B_{i,N+n,n'}^{\mathrm{U},(p)}$, and $\mathcal{C}_{i,\pi(n),n'}^{\mathrm{U},(p)} = \mathcal{C}_{\pi(n),n'}^{\mathrm{U},(p)} + \mathcal{C}_{i,N+n,n'}^{\mathrm{U},(p)}$, for $i =$

Algorithm 3 URLLC throughput maximization algorithm for computing (68)

1: **Initialization:** Initialize a random feasible $V^{(0)}$ for (22c). Set $p=0$.
2: **Repeat** the following until the objective function in (68) converges:
Update $\gamma_n^{\mathrm{U},(p)}$ and $\gamma_{i,\pi(n)}^{\mathrm{U},(p)}$, for all $n\in\mathcal{N}_1$, $i=1,2$, using (70). Update $\tilde{B}_n^{\mathrm{U},(p)}$ and $\tilde{C}_n^{\mathrm{U},(p)}$, for all $n\in\mathcal{N}_E$, by (82). Generate $V^{(p+1)}$ by (84). Reset $p:=p+1$.
3: **Output** $V^{(p)}$.

$1,2$. Using (76) and (80), a tight concave minorant of the objective function in (69a) is

$$\tilde{g}^{\mathrm{U},(p)}(\mathbf{V}) \triangleq a^{\mathrm{U},(p)} + 2\sum_{n\in\mathcal{N}_E}\langle \tilde{B}_n^{\mathrm{U},(p)}, \mathbf{V}_n\rangle - \sum_{n\in\mathcal{N}_E}\langle \tilde{C}_n^{\mathrm{U},(p)}, [\mathbf{V}_n]^2\rangle, \quad (81)$$

where

$$\begin{aligned} a^{\mathrm{U},(p)} &\triangleq \textstyle\sum_{n\in\mathcal{N}_1}\left(\gamma_n^{\mathrm{U},(p)}a_n^{\mathrm{U},(p)} + \sum_{i=1}^{2}\gamma_{i,\pi(n)}^{\mathrm{U},(p)}a_{i,\pi(n)}^{\mathrm{U},(p)}\right),\\ \tilde{B}_n^{\mathrm{U},(p)} &\triangleq \textstyle\sum_{n'\in\mathcal{N}_1}\left(\gamma_{n'}^{\mathrm{U},(p)}B_{n',n}^{\mathrm{U},(p)} + \sum_{i=1}^{2}\gamma_{i,\pi(n')}^{\mathrm{U},(p)}B_{i,\pi(n'),n}^{\mathrm{U},(p)}\right),\\ \tilde{C}_n^{\mathrm{U},(p)} &\triangleq \textstyle\sum_{n'\in\mathcal{N}_1}\left(\gamma_{n'}^{\mathrm{U},(p)}\mathcal{C}_{n',n}^{\mathrm{U},(p)} + \sum_{i=1}^{2}\gamma_{i,\pi(n')}^{\mathrm{U},(p)}\mathcal{C}_{i,\pi(n'),n}^{\mathrm{U},(p)}\right), \end{aligned} \quad (82)$$

for $n\in\mathcal{N}_E$. Therefore, we solve the following convex quadratic problem that maximizes the tight minorant, in order to determine the next feasible point $V^{(p+1)}$ for (67):

$$\max_{\mathbf{V}} \quad \tilde{g}^{\mathrm{U},(p)}(\mathbf{V}) \quad \text{s.t.} \quad (22c), \quad (83)$$

for which a closed-form solution can be derived as follows:

$$V_n^{(p+1)} = \begin{cases} (\tilde{C}_n^{\mathrm{U},(p)})^{-1}\tilde{B}_n^{\mathrm{U},(p)} & \text{if } \Xi^{\mathrm{U},(p)} \le 2P,\\ (\tilde{C}_n^{\mathrm{U},(p)} + \lambda I_{2N_t})^{-1}\tilde{B}_n^{\mathrm{U},(p)} & \text{otherwise,} \end{cases} \quad (84)$$

where $\Xi^{\mathrm{U},(p)} \triangleq \sum_{n\in\mathcal{N}_E}||(\tilde{C}_n^{\mathrm{U},(p)})^{-1}\tilde{B}_n^{\mathrm{U},(p)}||^2$ and $\lambda > 0$ is found by bisection such that $\sum_{n\in\mathcal{N}_E}||(\tilde{C}_n^{\mathrm{U},(p)} + \lambda I_{2N_t})^{-1}\tilde{B}_n^{\mathrm{U},(p)}||^2 = 2P$.

Algorithm 3 provides a procedure for addressing the problem (67). Similar to Algorithm 1, it guarantees convergence to at least a local optimum that fulfills the first-order optimality condition. The computational complexity of the closed-form based solution in (84) is $\mathcal{O}\left(3N_t\log_2(2N_t)N\right)$.

Remark 3. By introducing the slack variables λ, satisfying the constraints (51) we can express the problem (67) as

$$\max_{\mathbf{V},\lambda} \frac{1}{2}\left(\prod_{n=1}^{N}\lambda_n\right)^{1/N} \quad (85a)$$

$$\text{s.t. } (22c), (51), \quad (85b)$$

$$\rho_n^{\mathrm{U}}(\mathbf{V}) \ge \lambda_n, n\in\mathcal{N}_1, \quad (85c)$$

$$\rho_{\pi(n)}^{\mathrm{U}}(\mathbf{V}) + \rho_{i,N+n}^{\mathrm{U}}(\mathbf{V}) \ge \lambda_{\pi(n)}, n\in\mathcal{N}_1, i=1,2. \quad (85d)$$

Algorithm 4 GM throughput maximization algorithm for computing (68) under LB

1: **Initialization:** Initialize a feasible $V_{1,n}^{(0)} = \begin{bmatrix} V_n^{11,(0)} \\ V_n^{21,(0)}\end{bmatrix}$, $n\in\mathcal{N}_E$ for the sum power constraint $\sum_{n\in\mathcal{N}_E}||\mathbf{V}_{1,n}||^2 \le P$. Define $V_n^{(0)} \triangleq \begin{bmatrix} V_n^{11,(0)} & -V_n^{21,(0)} \\ V_n^{21,(0)} & V_n^{11,(0)}\end{bmatrix}$, $n\in\mathcal{N}_E$. Set $p=0$.
2: **Repeat** until the objective function in (68) converges:
Update $\gamma_n^{(p)}$ and $\gamma_{i,\pi(n)}^{(p)}$ by (70). Update $\tilde{B}_n^{\mathrm{U},(p)}$ and $\tilde{C}_n^{\mathrm{U},(p)}$ by (82) and then $\tilde{B}_{1,n}^{\mathrm{U},(p)}$ and $\tilde{C}_{1,n}^{\mathrm{U},(p)}$ by (88). Generate $V_{1,n}^{(p+1)} = \begin{bmatrix} V_n^{11,(p+1)} \\ V_n^{21,(p+1)}\end{bmatrix}$ by (90) and then define $V_n^{(p+1)} \triangleq \begin{bmatrix} V_n^{11,(p+1)} & -V_n^{21,(p+1)} \\ V_n^{21,(p+1)} & V_n^{11,(p+1)}\end{bmatrix}$. Reset $p:=p+1$.
3: **Output** $V_1^{(p)}$.

Initialized by a point $V^{(0)}$ feasible for (22c), let $V^{(p)}$ be a point feasible for (25) that is obtained from the $(p-1)$-st iteration. For $\rho_n^{\mathrm{U},(p)}(\mathbf{V})$ and $\rho_{i,N+n}^{\mathrm{U},(p)}(\mathbf{V})$ defined from (76) and (78), at the p-th iteration, we solve the following convex optimization problem to obtain the subsequent feasible point $V^{(p+1)}$:

$$\max_{\mathbf{V},\lambda} \frac{1}{2}\left(\prod_{n=1}^{N}\lambda_n\right)^{1/N} \quad (86a)$$

$$\text{s.t. } (22c), (51), \quad (86b)$$

$$\rho_n^{\mathrm{U},(p)}(\mathbf{V}) \ge \lambda_n, n\in\mathcal{N}_1, \quad (86c)$$

$$\rho_{\pi(n)}^{\mathrm{U},(p)}(\mathbf{V}) + \rho_{i,N+n}^{\mathrm{U},(p)}(\mathbf{V}) \ge \lambda_{\pi(n)}, n\in\mathcal{N}_1, i=1,2, \quad (86d)$$

which is a tight inner approximation for (85d). Its computational complexity is $\mathcal{O}\left((6N_tN_rN+N)^3N\right)$ [37], which is huge compared to the computational complexity of the closed-form based solution in (84), which is just $\mathcal{O}\left(3N_t\log_2(2N_t)N\right)$.

Remark 4. Recall that under $\mathbf{W}_{2,n'} \equiv 0$ for LB we introduce the real-composite form (54) of $\mathbf{W}_{1,n'}$. By partitioning $\tilde{B}_n^{\mathrm{U},(p)}$ and $\tilde{C}_n^{\mathrm{U},(p)}$ defined from (82) as $\tilde{B}_n^{\mathrm{U},(p)} = [\tilde{B}_{n,11}^{\mathrm{U},(p)}]_{(i,j)\in\{1,2\}}$, $\tilde{B}_{n,ij}^{\mathrm{U},(p)} \in \mathbb{R}^{N_t\times N_r}$, $\tilde{C}_n^{\mathrm{U},(p)} = [\tilde{C}_{n,ij}^{\mathrm{U},(p)}]_{(i,j)\in\{1,2\}}$, $\tilde{C}_{n,ij}^{\mathrm{U},(p)} \in \mathbb{R}^{N_t\times N_t}$, $i,j=1,2$, we can express $\tilde{g}^{\mathrm{U},(p)}$ in (81) as

$$\tilde{g}^{\mathrm{U},(p)}(\mathbf{V}_1) = a^{\mathrm{U},(p)} + 2\sum_{n\in\mathcal{N}_E}\langle \tilde{B}_{1,n}^{\mathrm{U},(p)}, \mathbf{V}_{1,n}\rangle - \sum_{n\in\mathcal{N}_E}\langle \tilde{C}_{1,n}^{\mathrm{U},(p)}, [\mathbf{V}_{1,n}]^2\rangle, \quad (87)$$

for $\mathbf{V}_1 \triangleq \{\mathbf{V}_{1,n}, n \in \mathcal{N}_E\}$, and

$$\tilde{B}_{1,n}^{\mathrm{U},(p)} = \begin{bmatrix} \tilde{B}_{n,11}^{\mathrm{U},(p)} + \tilde{B}_{n,22}^{\mathrm{U},(p)} \\ \tilde{B}_{n,21}^{\mathrm{U},(p)} - \tilde{B}_{n,12}^{\mathrm{U},(p)} \end{bmatrix},$$
$$\tilde{C}_{1,n}^{\mathrm{U},(p)} = \begin{bmatrix} \tilde{C}_{n,11}^{\mathrm{U},(p)} + \tilde{C}_{n,22}^{\mathrm{U},(p)} & \tilde{C}_{n,12}^{\mathrm{U},(p)} - (\tilde{C}_{n,12}^{\mathrm{U},(p)})^T \\ (\tilde{C}_{n,12}^{\mathrm{U},(p)})^T - \tilde{C}_{n,12}^{\mathrm{U},(p)} & \tilde{C}_{n,11}^{\mathrm{U},(p)} + \tilde{C}_{n,22}^{\mathrm{U},(p)} \end{bmatrix} \tag{88}$$

for $n \in \mathcal{N}_E$. Instead of (83), we solve the following optimization problem of tight minorant maximization to determine the next point $V_1^{(p+1)}$ in the iteration sequence:

$$\max_{\mathbf{V}_1} \quad \tilde{g}^{\mathrm{U},(p)}(\mathbf{V}_1) \quad \text{s.t.} \quad \sum_{n \in \mathcal{N}_E} ||\mathbf{V}_{1,n}||^2 \leq P, \tag{89}$$

which admits the closed-form solution

$$V_{1,n}^{(p+1)} = \begin{cases} (\tilde{C}_{1,n}^{\mathrm{U},(p)})^{-1} \tilde{B}_{1,n}^{\mathrm{U},(p)} & \text{if } \tilde{\Xi}^{\mathrm{U},(p)} \leq P, \\ (\tilde{C}_{1,n}^{\mathrm{U},(p)} + \lambda I_{2N_t})^{-1} \tilde{B}_{1,n}^{\mathrm{U},(p)} & \text{otherwise,} \end{cases} \tag{90}$$

where $\tilde{\Xi}^{\mathrm{U},(p)} \triangleq \sum_{n \in \mathcal{N}_E} ||(\tilde{C}_{1,n}^{\mathrm{U},(p)})^{-1} \tilde{B}_{1,n}^{\mathrm{U},(p)}||^2$ and $\lambda > 0$ is found by bisection such that $\sum_{n \in \mathcal{N}_E} ||(\tilde{C}_{1,n}^{\mathrm{U},(p)} + \lambda I_{2N_t})^{-1} \tilde{B}_{1,n}^{\mathrm{U},(p)}||^2 = P$. Algorithm 4 provides a procedure for computing (68) under LB. The computational complexity of the closed-form based solution in (90) is $\mathcal{O}\left(3N_t \log_2(2N_t)N\right)$.

Remark 5. We assume that the channel H_n spanning from the BS to UE n is quasi-static over a given coherent time interval (CTI). Accordingly, we design a widely linear beamforming (WLB) solution for each CTI, which must be updated as the channel changes in subsequent intervals. However, updating the WLB design for every CTI is not challenging with the aid of our proposed algorithm. This is because the solution from the previous CTI provides a good initialization, and more importantly, our method relies on solving computationally efficient closed-form expressions.

V. Simulation Results

For simulation, we consider a cell-radius of 300 meters. The channel H_n spanning from the BS to any UE n located at a distance d meters is modeled as $H_n = \sqrt{10^{-\sigma_{\mathrm{PL}}/10}} \tilde{H} n$, where $\sigma\mathrm{PL} = 38.46 + 10\beta \log_{10}(d)$ dB represents the path loss, and $\tilde{H}_n$ denotes the normalized channel gain. The normalized channel $\tilde{H}_n$ follows a Rayleigh fading model, with entries independently drawn from a complex Gaussian distribution. Here, 38.46 dB represents the free-space path loss at a reference distance of 1 meter and a carrier frequency of 2 GHz, while $\beta = 3.1$ denotes the path-loss exponent. The UEs are distributed randomly within the cell area. The noise power spectral density is configured as $\frac{\sigma}{B} = -174$ dBm/Hz with the bandwidth $\mathcal{B} = 1$ MHz. For the URLLC network, the decoding error probability ϵ_c is set to 10^{-5} and the transmission duration is set to $t_t = 0.1$ ms. Unless stated otherwise, the transmission power limit is assumed to be $P = 30$ dBm, the BS is equipped with $N_t = 13$ transmit antennas, while each UE is assigned $N_r = 2$ receive antennas, and the number of UEs in the system is $N = 18$. The convergence threshold for all algorithms is fixed at 10^{-3}, and the maximum number of iterations is limited to $p_{\max} = 1000$.

In presenting our simulation results, we use the phrase "n-NOMA-WLB" to refer to our proposed GM-throughput maximization Algorithms 1 and 3, where Algorithm 3 is the extension of Algorithm 1 for URLLC network. We use the phrases "NOMA-WLB" and "CoSig-WLB" to refer to the results of NOMA and CoSig implementation under WLB, which are obtained by tailoring the proposed algorithms to solve (27) and (29), respectively. We also simulate both the LB Algorithms 2 and 4 for comparison. The corresponding results will be referred by "n-NOMA-LB", "NOMA-LB", and "CoSig-LB".

In the following three sub-sections, we shall first compare the performance of the proposed closed-form expressions based algorithms to that of the convex-solver based implementation. Next, we will present detailed results for the proposed algorithms. Finally, we will compare the performance of the proposed GM-throughput maximization algorithms to the existing max-min throughput and sum-throughput optimization algorithms of [6].

A. Performance Comparison with Algorithms Based on Convex-Solver

In this subsection, we compare the performance of the proposed closed-form expressions-based algorithm with that of the convex-solver-based approach. Due to the high computational complexity of the convex-solver approach, we use a small-scale network with $N_t = 6$, $N = 8$, and $N_r = 1$ in this subsection.

Fig. 2 compares the achievable GM-throughput under n-NOMA between the proposed algorithm (Algorithm 1), which solves the tractable surrogate problem (31), and the convex-solver based algorithm (see Remark 1), which solves the GM rate optimization problem (25). First, we can observe from Fig. 2 that WLB clearly outperforms LB in terms of the achievable GM-throughput and the performance gap increases upon increasing the power budget P. The performance gain of the WLB design stems from the additional degrees of freedom provided by using a pair of beamformers for each information symbol, i.e., $\mathbf{W}_{1,n}$ and $\mathbf{W}_{2,n}$ for the message s_n. By contrast, the conventional LB approach applies a single beamformer, $\mathbf{W}_{1,n}$, for each information symbol s_n.

Secondly, we observe that the performance of the proposed closed-form-based solution is close to that of the convex-solver-based solution, particularly under the proposed WLB. Specifically, the performance gap is less than 1 dB, meaning that Algorithm 1 achieves comparable performance to the convex-solver-based approach with less than a 1 dB increase in power budget. Same conclusions can be drawn for the URLLC network, where the results are presented in Fig. 3 (obtained with Algorithm 3 and as outlined in Remark 2). Similarly, our performance comparisons between the proposed closed-form based solution with convex-solver based

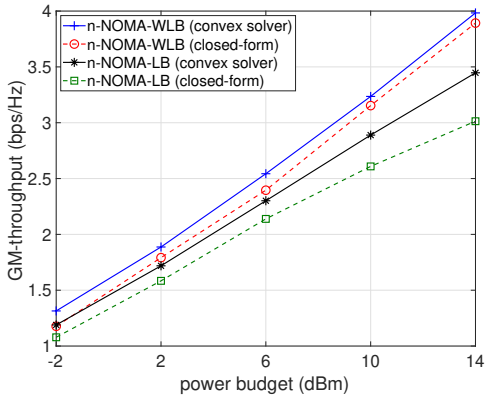

FIGURE 2: Comparison of the closed-form and convex-solver based solutions.

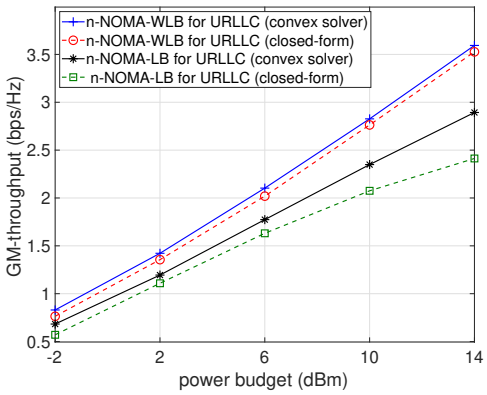

FIGURE 3: Comparison of the closed-form and convex-solver based solutions under URLLC.

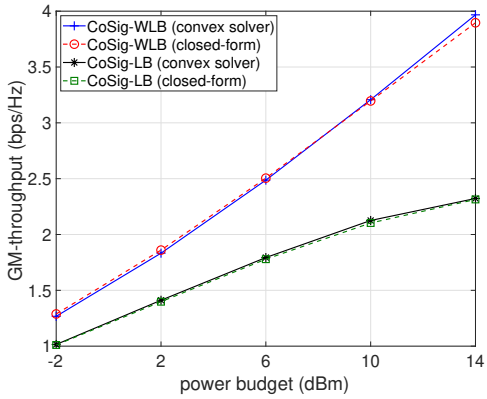

FIGURE 4: Closed-form and convex-solver based solution for CoSig scheme.

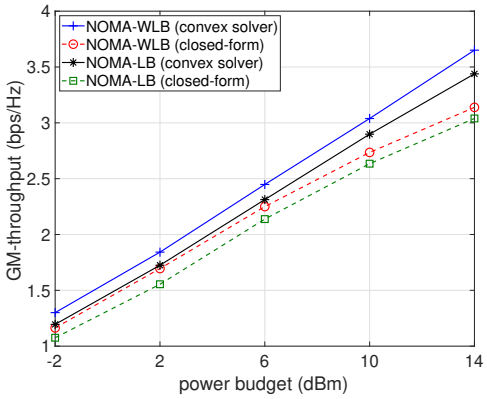

FIGURE 5: Closed-form and convex-solver based solution for NOMA scheme.

TABLE 4: Average computational time (CPU processing time in seconds) required by the proposed closed-form and convex-solver based solutions for $N_t = 6$, $N = 8$, and $N_r = 1$.

	CPU processing time in seconds		
Algorithms	CoSig-WLB	NOMA-WLB	n-NOMA-WLB
Proposed closed-form sol. (Algorithm 1)	0.068	0.108	0.1310
Convex-solver based sol. (Remark 1)	36.49	38.62	56.69
Proposed closed-form sol. (Algorithm 3)	0.0904	0.133	0.1742
Convex-solver based sol. (Remark 3)	47.88	50.54	75.39

solution under CoSig and NOMA are shown in Figures 4 and 5, respectively. We can observe that the two solution approaches perform similarly under CoSig, but there is a noticeable performance gap under NOMA. Particularly, we observe from Figures 2 and 5 that for n-NOMA-LB, NOMA-WLB, and NOMA-LB, the performance gap between the closed-form and convex-solver based solutions increases with the increase in P. This is because the convex-solver based solution is capable of taking advantage of the increase in the power, since it directly addresses the GM-throughput optimization problem (25), while the closed-form based solution is obtained by solving the surrogate problem (31). However, it is noteworthy from Fig. 2 that n-NOMA-WLB is capable of maintaining a constant performance gap between the closed-form and convex-solver based solution. This is because n-NOMA-WLB simultaneously offers throughput enhancements for the UEs by using both private and common messages for the distant UEs and degree of freedom by using pair of beamformers.

The better performance achieved by the algorithm based on convex-solver over the proposed algorithm under both n-NOMA and NOMA (as seen in Figures 2 and 5) is expected because the proposed algorithm relies on the solution of the surrogate problem (31). However, the main advantage of the proposed algorithm based on closed-form expressions is its significantly lower computational complexity. Specifically,

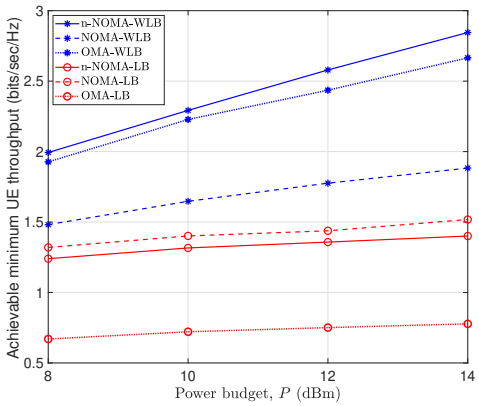

FIGURE 6: Achievable minimum UE throughput versus the transmit power budget P.

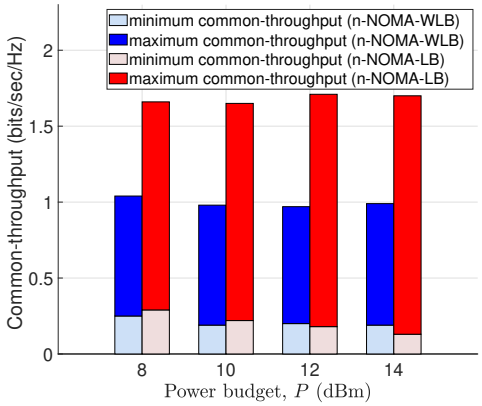

FIGURE 7: Achievable common-throughput versus the transmit power budget P.

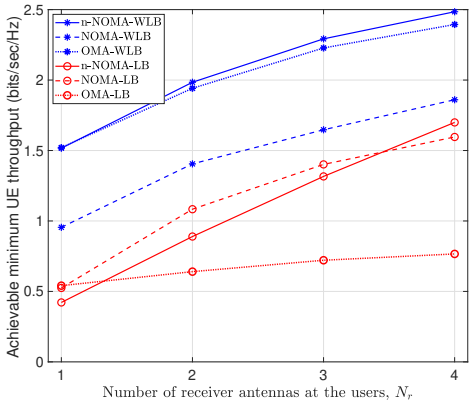

FIGURE 8: Achievable minimum UE throughput versus the number of antennas at the UEs' receiver, N_r.

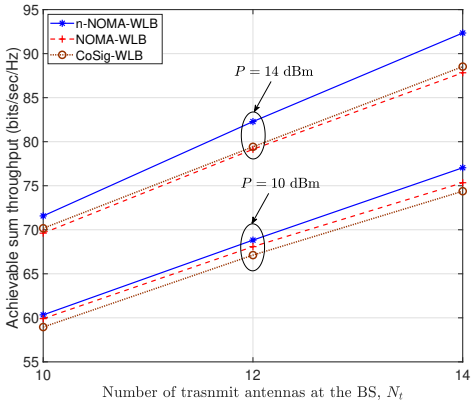

FIGURE 9: Achievable sum throughput versus the number of transmit antennas at the BS, N_t.

it is nearly 450 times more computationally efficient than the convex-solver-based algorithm. Detailed comparisons are provided in Table 4 for both Algorithm 1 and Algorithm 3. Note that the general computational complexity comparison is discussed at the end of Remarks 1 and 2.

B. Detailed results for the proposed Algorithms

In this subsection, we provide detailed results for our proposed closed-form based solution over varying values of transmit power budget P, transmit antennas N_t, UEs N, and receive antennas N_r. Unless specified otherwise, we set $P = 10$ dBm, $N_t = 6$, $N = 8$, and $N_r = 3$.

Fig. 6 plots the achievable minimum UE throughput versus the transmit power budget P, which clearly shows the advantage of employing WLB over LB. Fig. 6 also shows that the performance gain of n-NOMA-WLB and CoSig-WLB over the counterpart schemes increases upon increasing P.

Fig. 7 plots the achievable common-throughput, defined by r_{N+n} in (12), versus the transmit power budget P. While employing n-NOMA-WLB and n-NOMA-LB, Fig. 7 plots the minimum and maximum common-throughput. For example, the minimum common-throughput by n-NOMA-WLB is around 0.2 bps/Hz, while the maximum common-throughput is around 1 bps/Hz. Fig. 7 shows that the common-throughput by n-NOMA-LB is larger than that by n-NOMA-WLB. However, the total achievable throughput of the former is outperformed by the latter. This can be verified by referring to Fig. 6, which illustrates that the minimum UE throughput of n-NOMA-WLB is nearly double that of n-NOMA-LB. The presence of non-zero common-throughput in Fig. 7 reinforces the notion that the achievable throughput of n-NOMA exceeds that of CoSig, and it can be verified from Fig. 6.

Fig. 8 plots the achievable minimum UE throughput versus the number of antennas N_r at the UEs' receiver. As expected, the minimum UE throughput improves upon increasing N_r due to the increase in the resources. Fig. 8 also shows that "n-NOMA-WLB" and "CoSig-WLB" clearly outperform the counterpart schemes in terms of the achievable minimum UE throughput.

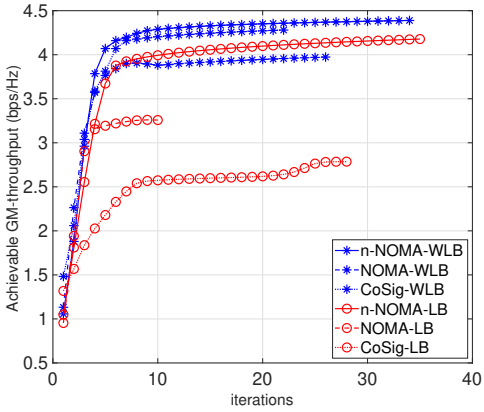

FIGURE 10: Convergence of proposed algorithms.

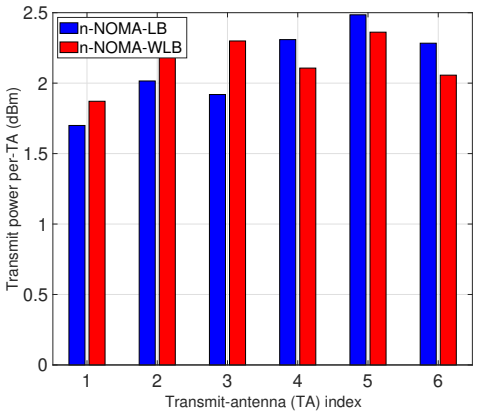

FIGURE 11: Transmit-power at different TAs.

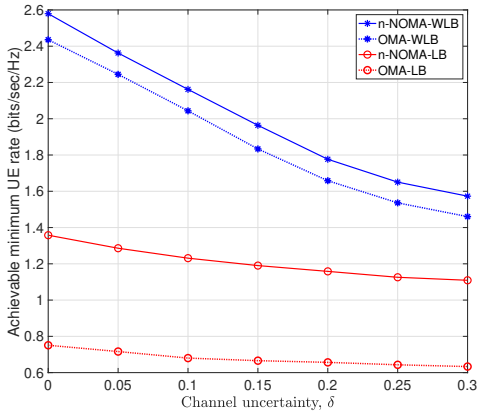

FIGURE 12: Achievable minimum UE throughput versus relative CSI uncertainty, δ.

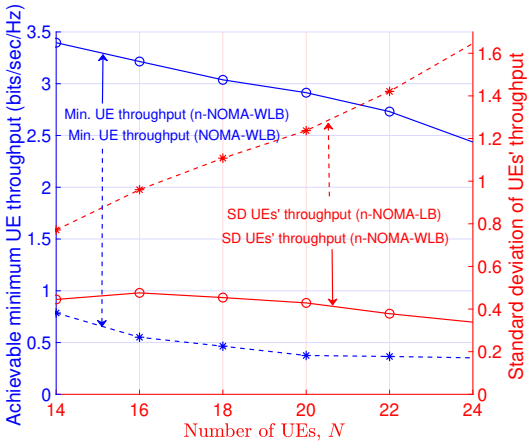

FIGURE 13: Achievable minimum UE throughput and SD among the UEs' throughput versus number of UEs N.

TABLE 5: Average number of iterations required by the proposed algorithms for convergence, under the simulation parameters $N_t = 6$, $N = 8$, $N_r = 3$, and $P = 10$ dBm.

Algorithms	CoSig-LB	NOMA-LB	n-NOMA-LB	CoSig-WLB	NOMA-WLB	n-NOMA-WLB
Avg. # of iter	22.7	23.2	30.7	22.6	27.3	36.4

Fig. 9 illustrates how the achievable sum throughput varies with the number of transmit antennas N_t. As N_t increases, the sum throughput also increases, owing to the greater availability of spatial resources. Fig. 9 shows the supremacy of n-NOMA-WLB over both the NOMA-WLB and CoSig-WLB implementations at different values of N_t and P and this performance gain by n-NOMA-WLB improves upon increasing N_t or P.

Fig.10 presents the convergence behavior of the proposed algorithms, clearly demonstrating the advantage of using WLB over LB in terms of the achieved GM-throughput. Notably, the WLB-based algorithms converge efficiently, typically within 25 to 30 iterations. The average iteration counts required for convergence are summarized in Table5.

Fig. 11 plots the transmit-power at different transmit antennas, while employing n-NOMA-WLB and n-NOMA-LB. Fig. 11 demonstrates that, even though we impose a single sum-power constraint in (22c) instead of individual power constraints for each transmit antenna, the transmit power is still distributed effectively across all the antennas. Observe that the variance of among the transmit-power at different TAs is 0.04 and 0.08 under n-NOMA-WLB and n-NOMA-LB, respectively.

Fig. 12 illustrates the impact of imperfect CSI on the minimum achievable throughput of UEs. To model this effect,

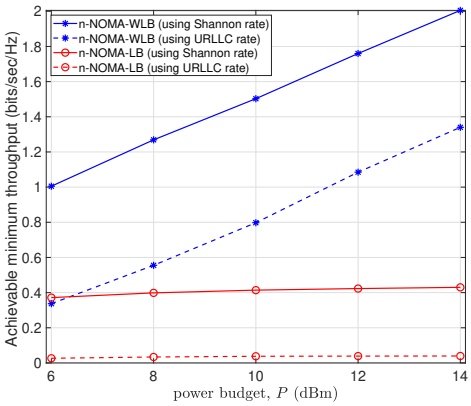

FIGURE 14: Achievable minimum UE throughput based on Shannon and URLLC rates versus P.

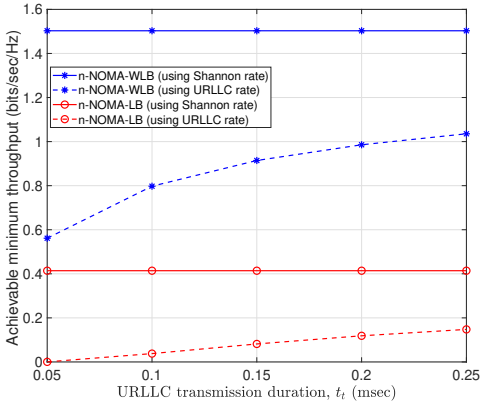

FIGURE 15: Achievable minimum UE throughput based on Shanon and URLLC rates versus t_t.

TABLE 6: Computational complexity comparison of the proposed and existing algorithms for n-NOMA-WLB implementation, under the simulation parameters $N_t = 3$, $N = 4$, and $N_r = 1$.

Algorithms	Complexity	CPU processing time (sec)
Proposed GM-throughput max.	$\mathcal{O}\left(3N_t \log_2(2N_t)N\right)$	0.1126
max-min throughput. opt. [6]	$\mathcal{O}\left((3N_tN)^3\right)$	51.9
sum-throughput max. using [6]	$\mathcal{O}\left((3N_tN)^3\right)$	40.20

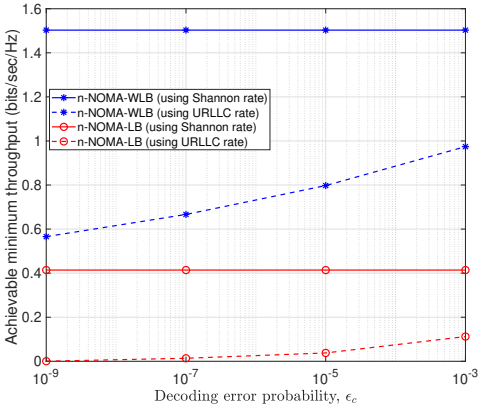

FIGURE 16: Achievable minimum UE throughput based on Shanon and URLLC rates versus ϵ_c.

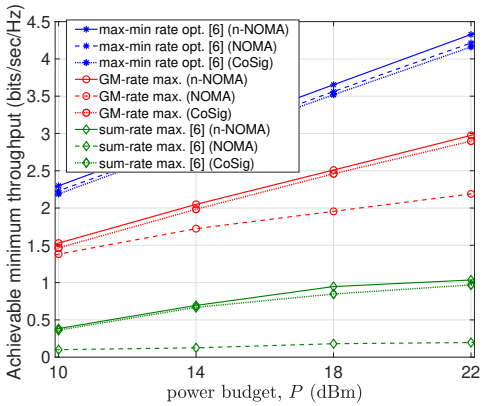

FIGURE 17: Achievable minimum UE throughput versus the transmit power budget P under WLB.

we introduce random channel estimation errors, with their magnitudes bounded by a factor of δ relative to the corresponding estimated channel coefficients. Explicitly, δ denotes the level of relative CSI uncertainty [41]. As anticipated, increasing the CSI uncertainty leads to a decrease in the minimum UE throughput. However, this decline is relatively modest. Specifically, even under a high uncertainty level of $\delta = 0.2$, the worst-case throughput of the n-NOMA-WLB scheme only drops by about 30%, highlighting the robustness of our proposed algorithms against CSI imperfections

Figure 13 examines a dense user deployment scenario, where, unlike the previous results that considered up to 8 UEs, we now randomly deploy up to $N = 24$ UEs in the network. The figure shows the achievable minimum UE throughput on the left y-axis and the standard deviation (SD) of UE throughputs on the right y-axis. The results clearly demonstrate the superior performance of the WLB scheme over LB, both in terms of the achievable minimum UE throughput and the fairness (measured by SD) among UEs. Notably, under dense user deployment (i.e., $N \geq 16$), WLB achieves approximately a seven-fold improvement in

the minimum UE throughput and over a five-fold reduction in throughput SD compared to LB. Interestingly, while the SD of UE throughputs increases with the number of UEs in the LB scheme, it remains nearly constant—or even slightly decreases—for the WLB scheme as N increases.

Figures 14-16 plot the achievable minimum thorughput for a URLLC network, which is obtained with the proposed Algorithm 3 and $N_r = 1$ receive antenna. The results are plotted with respect to the power budget P in Fig. 14, the URLLC transmission duration t_t in Fig. 15, and the decoding error probability ϵ_c in Fig. 16. For comparison, we also plot the throughput obtained by the proposed Algorithm 1. As expected, the achievable GM-throughput of Algorithm 1 (based on Shannon rate expression) is higher than that of Algorithm 3 (based on URLLC rate expression). This is due to the channel dispersion factor, which is the second factor in the URLLC rate expression (59) and it is subtracted from the Shannon rate. It is noteworthy that although the original optimization problem (67) constructed for a URLLC network is non-convex and the second factor in (59) makes the problem (67) more challenging, our proposed Algorithm 3, which employs n-NOMA-WLB, achieves an increase in the minimum throughput upon increasing P. Figures 14-16 shows a huge performance gap between WLB and LB, which also manifests the advantage of employing the proposed WLB over conventional LB in URLLC networks.

C. Comparison with existing results

In this section, we evaluate the performance of the proposed GM-throughput maximization algorithms in comparison to the existing max-min throughput optimization and sum-throughput maximization algorithms in [6]. We consider a small-scale network as that is considered in [6] with $N_t = 3$ antennas, $N = 4$ UEs, and $N_r = 1$ antenna. This is because the optimization algorithms of [6] rely on convex-solver based iterations, which are computationally complex.

Table 6 lists the computational complexities and average computational times (CPU processing times) required by the proposed and existing algorithms for convergence. The computation time is measured on a machine equipped with a 2.3 GHz Intel Core i9 processor and 16 GB of RAM. Table 6 also shows that the proposed GM-throughput maximization algorithms are almost 400 times more computationally efficient than the max-min throughput optimization algorithms in [6]. This huge computational saving is the benefit of the closed-form expressions we developed for implementing the proposed algorithms, which is in a sharp contrast to the use of the complicated convex-solver for the algorithms in [6]. Note that while Table 6 specifically mentions the computational complexity of the convex-solver-based approach in [6], other relevant existing works, such as [5] and [14], also propose convex-solver-based methods with the same computational complexity of $\mathcal{O}\left((3N_tN)^3\right)$ when applied to a single-cell network.

Figs. 17 and 18 plot the achievable minimum-throughput and achievable sum-throughput, respectively, for all the algorithms (with WLB) under comparison versus P. As expected, the max-min throughput optimization algorithms in [6] achieve the highest minimum UE throughput (see Fig. 17), whereas the sum-throughput maximization algorithms in [6] achieve the highest sum-throughput (see Fig. 18). On the other hand, the proposed GM-throughput maximization algorithms (especially implemented with n-NOMA), although not explicitly targeting the max-min throughput optimization nor the sum-throughput maximization, are able to deliver a throughput that is not very far from the best achievable max-min throughput and the best achievable sum-throughput. Finally, in Fig. 19, we plot the individual user-rate distribution of the proposed Algorithm 1 and that of the max-min throughput optimization algorithm in [6] for the n-NOMA-WLB implementation at $P = 18$ dBm. We can observe that the individual user rate of the proposed Algorithm 1 is only marginally compromised compared to that by the max-min throughput optimization in [6]. On the other hand, the former is almost 400 times more computationally efficient compared to the latter. Considering the modest performance erosion in terms of both the achievable minimum UE-throughput and the achievable sum-throughput and the huge computational saving, the proposed GM-throughput maximization algorithms are very attractive alternatives to the existing algorithms of [6].

VI. Conclusions

Aiming for enhancing the QoS provided by large multi-user networks, this paper has proposed novel algorithms for designing WLB in conjunction with NOMA to maximize the geometric mean of the UEs' throughput subject to a sum transmit power constraint. The core in our proposed algorithms is a set of closed-form expressions, which can be conveniently and efficiently used in the iterations of the proposed algorithms. Compared to existing algorithms that are based on convex-solver in their iterations, the proposed algorithms are hundreds of times more computationally efficient, which makes them very attractive for WLB design in large multi-user NOMA networks. Extensive simulation results shown that n-NOMA clearly outperforms the popular NOMA and the traditional CoSig. Furthermore, WLB is clearly superior to LB. Moreover, the WLB design based on GM-throughput maximization not only results in fair throughput distribution among UEs, but also in even power distribution across transmit antennas without enforcing any per-antenna power constraints. The proposed WLB designs were developed for both conventional and URLLC networks. This work has assumed single-antenna UEs to develop WLB design for URLLC networks and its extension to multiple-antenna UEs could be the subject of our future research.

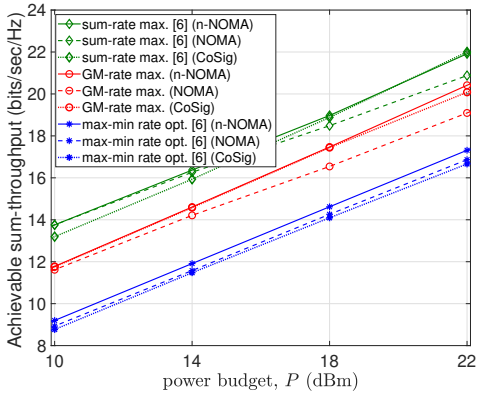

FIGURE 18: Achievable sum-throughput versus the transmit power budget P under WLB.

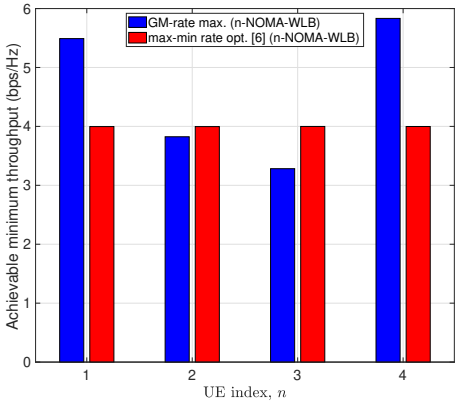

FIGURE 19: Individual user-throughput distribution by GM-rate max. and max-min rate opt. in [6].

Appendix: Fundamental inequalities

The following inequality holds for all real matrices X and $\bar{X}$ of dimensions $n \times m$, and real positive definite matrices Y and $\bar{Y}$ of dimensions $m \times m$ [42], [8, Theorem 1]

$$\begin{aligned}\ln\left|I_n + [X]^2 Y^{-1}\right| \geq \ln\left|I_n + [\bar{X}]^2 \bar{Y}^{-1}\right| - \langle [\bar{X}]^2, \bar{Y}^{-1}\rangle \\ + 2\langle \bar{Y}^{-1}\bar{X}, X\rangle \\ - \langle \bar{Y}^{-1} - \left(\bar{Y} + [\bar{X}]^2\right)^{-1}, [X]^2 + Y\rangle.\end{aligned} \tag{91}$$

Using the concavity of $\sqrt{x}$, the following inequality holds for all $x > 0$ and $\bar{x} > 0$

$$\sqrt{x} \leq \frac{\sqrt{\bar{x}}}{2}\left(1 + \frac{x}{\bar{x}}\right). \tag{92}$$

Using the convexity of the function x^2/y the following inequality is obtained for the matrices X and $\bar{X}$ of the same size, and positive scalars y and $\bar{y}$:

$$\frac{\|X\|^2}{y} \geq \frac{2\langle \bar{X}^H X\rangle}{\bar{y}} - \frac{\|\bar{X}\|^2}{\bar{y}^2} y. \tag{93}$$

ACKNOWLEDGMENT

This work was supported in part by the Deanship of Research Oversight and Coordination (DROC) at KFUPM for funding under the Interdisciplinary Research Center for Communication Systems and Sensing through project No. INCS2511, in part by the U.S National Science Foundation under Grants CNS-2128448 and ECCS-2335876, and in part by the Engineering and Physical Sciences Research Council projects EP/W016605/1, EP/X01228X/1, EP/Y026721/1 and EP/W032635/1 as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028)

REFERENCES

[1] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. 2013 IEEE 77th Veh. Technol. Conf. (VTC Spring)*, June 2013, pp. 1–5.

[2] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2294–2323, 3rd quarter 2018.

[3] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wirel. Commun.*, vol. 9, no. 5, pp. 1748 –1759, May 2010.

[4] G. C. Alexandropoulos, P. Ferrand, J.-M. Gorce, and C. B. Papadias, "Advanced coordinated beamforming for the downlink of future LTE cellular networks," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 54–60, 2016.

[5] V. D. Nguyen, H. D. Tuan, T. Q. Duong, H. V. Poor, and O. S. Shin, "Precoder design for signal superposition in MIMO-NOMA multicell networks," *IEEE J. Select. Areas Commun.*, vol. 35, no. 12, pp. 2681–2695, Dec 2017.

[6] A. A. Nasir, H. D. Tuan, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Signal superposition in NOMA with proper and improper Gaussian signaling," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6537–6551, Oct. 2020.

[7] E. Che, H. D. Tuan, H. H. M. Tam, and H. H. Nguyen, "Successive interference mitigation in multiuser MIMO interference channels," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2185–2199, June 2015.

[8] H. D. Tuan, H. H. M. Tam, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Superposition signaling in broadcast interference networks," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4646–4656, Nov. 2017.

[9] Z. Sheng, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "Secure UAV-enabled communication using Han–Kobayashi signaling," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 5, pp. 2905–2919, May 2020.

[10] H. D. Tuan, A. A. Nasir, M.-N. Nguyen, and M. Masood, "Han–Kobayashi signaling in MIMO broadcasting," *IEEE Commun. Lett.*, vol. 23, no. 5, pp. 855–858, May 2019.

[11] T. K. Nguyen, H. H. Nguyen, and H. D. Tuan, "Max-min QoS power control in generalized cell-free massive MIMO-NOMA with optimal backhaul combining," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 10 949–10 964, Oct. 2020.

[12] Y. Xu, C. Shen, Z. Ding, X. Sun, S. Yan, G. Zhu, and Z. Zhong, "Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4874–4886, 2017.

[13] J. Cui, Z. Ding, and P. Fan, "Outage probability constrained MIMO-NOMA designs under imperfect CSI," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 12, pp. 8239–8255, 2018.

[14] H. D. Tuan, A. A. Nasir, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Non-orthogonal multiple access with improper Gaussian signaling," *IEEE J. Sel. Areas Signal Process.*, vol. 13, no. 3, pp. 496–507, Jun. 2019.

[15] Y. Zeng, R. Zhang, E. Gunawan, and Y. L. Guan, "Optimized transmission with improper Gaussian signaling in the K-user MISO

interference channel," *IEEE Trans. Wirel. Commun.*, vol. 12, no. 12, pp. 6303–6313, December 2013.
[16] S. Lagen, A. Agustin, and J. Vidal, "On the superiority of improper Gaussian signaling in wireless interference MIMO scenarios," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3350–3368, Aug. 2016.
[17] M. Soleymani, I. Santamaria, and P. J. Schreier, "Improper gaussian signaling for the k-user MIMO interference channels with hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11 632–11 645, 2020.
[18] H. Yu, H. D. Tuan, A. A. Nasir, T. Q. Duong, and L. Hanzo, "Improper Gaussian signaling for computationally tractable energy and information beamforming," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13 990–13 995, Nov. 2020.
[19] H. Yu, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "Joint design of reconfigurable intelligent surfaces and transmit beamforming under proper and improper Gaussian signaling," *IEEE J. Select. Areas Commun.*, vol. 38, no. 11, pp. 2589–2603, Nov. 2020.
[20] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge University Press, 2010.
[21] S. Javed, O. Amin, B. Shihada, and M.-S. Alouini, "A journey from improper gaussian signaling to asymmetric signaling," *IEEE Comun. Surveys Tuts.*, vol. 22, no. 3, pp. 1539–1591, 2020.
[22] R. Kotaba, C. N. Manchón, T. Balercia, and P. Popovski, "How URLLC can benefit from NOMA-based retransmissions," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 3, pp. 1684–1699, 2021.
[23] X. Xie, X. Ou, H. Lu, and Q. Huang, "Joint uplink and downlink resource allocation in NOMA for end-to-end URLLC services," *IEEE Commun. Lett.*, vol. 25, no. 12, pp. 3942–3946, 2021.
[24] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proc. IEEE*, vol. 106, no. 10, pp. 1834–1853, 2018.
[25] H. Yu, H. D. T. E. Dutkiewicz, H. V. Poor, and L. Hanzo, "Maximizing the geometric mean of user-rates to improve rate-fairness: Proper vs. improper Gaussian signaling," *IEEE Trans. Wirel. Commun.*, vol. 21, pp. 295–309, Jan. 2022.
[26] W. Zhu, H. D. Tuan, E. Dutkiewicz, and L. Hanzo, "Collaborative beamforming aided fog radio access networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7805–7820, July 2022.
[27] H. Tuy, *Convex Analysis and Global Optimization (second edition)*. Springer International, New York, USA, 2016.
[28] Y. Liu, H. Xing, C. Pan, A. Nallanathan, M. Elkashlan, and L. Hanzo, "Multiple-antenna-assisted non-orthogonal multiple access," *IEEE Wirel.*, vol. 25, no. 2, pp. 17–23, 2018.
[29] B. Clerckx, Y. Mao, R. Schober, E. A. Jorswieck, D. J. Love, J. Yuan, L. Hanzo, G. Y. Li, E. G. Larsson, and G. Caire, "Is NOMA efficient in multi-antenna networks? A critical look at next generation multiple access techniques," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1310–1343, 2021.
[30] Z. Ding et al., "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct 2017.
[31] W. Liang, Z. Ding, Y. Li, and L. Song, "User pairing for downlink non-orthogonal multiple access networks using matching algorithm," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5319–5332, Dec 2017.
[32] J. Zhu, J. Wang, Y. Huang, K. Navaie, Z. Ding, and L. Yang, "On optimal beamforming design for downlink MISO NOMA systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3008–3020, 2020.
[33] Q. Sun, S. Han, C. L. I, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 405–408, Aug. 2015.
[34] T. M. Cover and J. A. Thomas, *Elements of Information Theory (second edition)*. John Wileys & Sons, 2006.
[35] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programms," *Operations Research*, vol. 26, no. 4, pp. 681–683, Jul 1978.
[36] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Secrecy rate beamforming for multicell networks with information and energy harvesting," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 677–689, 2017.
[37] D. Peaucelle, D. Henrion, and Y. Labit, "Users guide for SeDuMi interface 1.03," 2002. [Online]. Available: http://homepages.laas.fr/peaucell/software/sdmguide.pdf
[38] M. Farooq, H. Q. Ngo, E.-K. Hong, and L.-N. Tran, "Utility maximization for large-scale cell-free massive MIMO downlink," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 7050–7062, 2021.
[39] J. Scarlett, V. Y. F. Tan, and G. Durisi, "The dispersion of nearest-neighbor decoding for additive non-Gaussian channels," *IEEE Trans. Infor. Theory*, vol. 63, no. 1, pp. 81–92, Jan. 2017.
[40] A. A. Nasir, H. D. Tuan, H. H. Nguyen, M. Debbah, and H. V. Poor, "Resource allocation and beamforming design in the short blocklength regime for URLLC," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 2, pp. 1321–1335, 2021.
[41] A. A. Nasir, H. D. Tuan, E. Dutkiewicz, H. V. Poor, and L. Hanzo, "Low-resolution RIS-aided multiuser MIMO signaling," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6517–6531, 2022.
[42] H. H. M. Tam, H. D. Tuan, and D. T. Ngo, "Successive convex quadratic programming for quality-of-service management in full-duplex MU-MIMO multicell networks," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2340–2353, June 2016.