Statistical Power Grid Observability under NOMA-based Communication Constraints

Qiyan Zhan*, Nan Liu *, Zhiwen Pan *, Hongjian Sun †
*The National Mobile Communications Research Laboratory, Southeast University, Nanjing, China
†Department of Engineering, Durham University, Durham, UK
{qiyanzhan, nanliu, pzw}@seu.edu.cn, hongjian.sun@durham.ac.uk

***Abstract*—This paper studies the observability of the power grid by jointly considering the power system with the wireless communication system under the strict latency requirements of Phasor Measurement Units (PMUs), which is characterized via the theory of effective capacity. In order to meet the quality of service (QoS) requirements and save communication bandwidth at the same time, the technique of non-orthogonal multiple access (NOMA) is adopted. For practical purposes, we consider the case where each NOMA group consists of at most two PMUs. The problem is formulated as minimizing the required communication bandwidth while satisfying the observability constraint of the power grid, over all possible NOMA user pairing strategies, bandwidth allocation among NOMA groups, power allocation within each NOMA group and the normalized QoS exponents of each PMU. In order to solve this problem, we first derive the closed-form expressions of the effective capacity of the two-PMU NOMA pair, then the problem is solved by first fixing the NOMA user pairing strategy and the probability of successful transmission of each PMU, and finding the optimal bandwidth allocation among NOMA groups, power allocation within each NOMA group and the normalized QoS exponents of each PMU via the bisection search method. Then, the probability of successful transmission of each PMU is found via the simulated annealing algorithm for a fixed NOMA user pairing strategy. Finally, we discuss the benefit of using uniform channel gain difference (UCGD) pairing as the PMU-pairing strategy. Numerical results on the IEEE 14-bus power system demonstrate the significant bandwidth savings of the proposed algorithm compared with the orthogonal multiple access (OMA) method and the NOMA method with equal bandwidth and power allocation.**

***Index Terms*—observability, cyber-physical systems, NOMA, smart grid, effective capacity**

I. Introduction

In the power grid, Phasor Measurement Units (PMUs) can accomplish real-time measurements [1], such as bus voltage phasor measurements and branch current phasor measurements. By installing PMU devices on the buses of the smart grid, real-time measurement results can be uploaded to the control center, which can realize many delay-sensitive applications such as real-time stability enhancement and vulnerability assessment etc. [2]. An important feature of the PMU is that when installed on a bus, say Bus i, the PMU not only measures the information of Bus i, but also the information of the buses that connect to Bus i. As a result, in [3], the optimal installation algorithm of the PMUs in the IEEE smart grid system is proposed, such that a relatively small number of PMUs are used to monitor the operation state of the entire grid.

The information detected by the PMUs, such as voltage phasor and current phasor varies with time. Real-time and up-to-date information are the most useful in real-time control of the power grid. Therefore, the information collected by the PMU devices have an extremely strict maximum delay limit $D_{\max}$ for uploading to the power grid control center. Measurements received after $D_{\max}$ are considered outdated and may lower the monitoring performance of the entire grid. Hence, designing communications schemes to meet the latency requirements of the PMU measurements is a meaningful and important problem.

In this paper, we focus on using the cellular network to connect the PMUs to the base station, and then the base station would forward the information to the control center via a high-speed backbone network. Wireless communication has many advantages, such as flexibility and low cost. However, wireless communication is not only subject to random receiver noise but also random fluctuations of channel gains between the base station and the PMUs because of the channel fading effect [4]. The random fluctuations of the channel gains bring about random variations of transmission latency, which are time-varying, and a deterministic bound is hard to obtain. As a result, it is difficult to guarantee that the transmission latency of each PMU meets the maximum latency threshold $D_{\max}$, beyond which the information of the associated buses is considered outdated. Reference [5] formulates the observability problem of the power grid under the constraints of the fading channel, where effective capacity theory is adopted to perform a cross-layer statistical analysis in the communication system. In general, effective capacity is defined as the maximum achievable rate of a fading channel under the condition that the statistical delay constraint is satisfied [6]. Utilizing the notion of effective capacity, the observability performance metrics of observability redundancy (OR), observability sensitivity (OS) and observability probability (OP) are proposed, and the communication resource allocation is optimized in an frequency division multiple access (FDMA) network [5].

While FDMA has many advantages, such as no intra-cell

This work is partially supported by the National Key Research and Development Project under Grants 2019YFE0123600 and 2020YFB1806805, the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 872172 (TESTBED2 project: www.testbed2.org), and the Research Fund of National Mobile Communications Research Laboratory, Southeast University (No. 2022A03).

interference and low complexity, it also has the disadvantages of low spectral efficiency and system capacity compared to non-orthogonal multiple access (NOMA) [7]. NOMA is a multi access method [8], where users are first grouped so that different users within the same group use the same time-frequency resources, while different user groups use orthogonal time-frequency resources. NOMA has gained significant research interest, and through power-domain or code-domain techniques, a greater spectral efficiency and system capacity can be obtained [9], [10].

In this paper, the power-domain NOMA technique is adopted to improve the observability of the power grid while saving communication bandwidth. More specifically, we we consider the case where each NOMA group consists of at most two PMUs. The problem is formulated as minimizing the required communication bandwidth while satisfying the observability constraint of the power grid, over all possible NOMA user pairing strategies, bandwidth allocation among NOMA groups, power allocation within each NOMA group and the normalized QoS exponents of each PMU. In order to solve this problem, we first derive the closed-form expressions of the effective capacity of the two-PMU NOMA pair, then the problem is solved by first fixing the NOMA user pairing strategy and the probability of successful transmission of each PMU, and finding the optimal bandwidth allocation among NOMA groups, power allocation within each NOMA group and the normalized QoS exponents of each PMU via the bisection search method. Then, the probability of successful transmission of each PMU is found via the simulated annealing algorithm for a fixed NOMA user pairing strategy. Finally, we discuss the benefit of using uniform channel gain difference (UCGD) pairing [11] as the PMU-pairing strategy. The major contributions and novelties are the following:

1) Compared to the use of OMA in [5], We adopt the power-domain NOMA technique to allocate communication resources to PMU devices to improve smart grid observability while saving communication bandwidth. Compared with the traditional OMA method, significant bandwidth savings can be achieved under the same observability requirements.
2) We derive the approximate closed-form expressions of effective capacity of the NOMA two-PMU pairs, where the decoding order is based on the *average* SNR of the pair.
3) While the optimization problem involves many variables, by exploiting the nature of the problem, a low-complexity algorithm is found via the bisection method search for the optimal bandwidth allocation among NOMA groups, power allocation within each NOMA group and the normalized QoS exponents of each PMU, when a NOMA user pairing strategy and the probability of successful transmission of each PMU is fixed. Then, simulated annealing is used to find the probability of successful transmission of each PMU.
4) We find that devising new PMU-pairing strategy for the problem at hand offers no significant benefit, and UCGD pairing [11] serves well as the PMU-pairing strategy.

II. System Model

We describe the system model studied in this paper as follows.

A. Observability of a smart grid system

In a smart grid system of N buses with a fixed topology, the connection matrix $\mathbf{L}$ of size $N \times N$ is known, whose elements are given as:

$$l_{ij} = \begin{cases} 1 & \text{if } i = j \text{ or if bus } i \text{ and } j \text{ are connected,} \\ 0 & \text{otherwise.} \end{cases}$$

for $i, j = 1, 2, \cdots, N$. Suppose there are K PMUs installed to observe the performance of the smart grid system. We define the installation vector of the PMUs as $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \end{bmatrix}^T$, i.e.,

$$x_i = \begin{cases} 1 & \text{if a PMU is installed at bus } i, \\ 0 & \text{otherwise.} \end{cases} \qquad i = 1, 2, \cdots, N$$

The observability vector as defined in [12] is $\mathbf{b} = \begin{bmatrix} b_1 & b_2 & \cdots & b_N \end{bmatrix}^T$, where b_i represents the number of PMUs that can measure bus i, i.e.,

$$b_i = \sum_{j=1}^{N} l_{ij} X_j, \quad i = 1, 2, \cdots, N$$

or written in matrix form, we have

$$\mathbf{b} = \mathbf{L}\mathbf{x} \tag{1}$$

If bus i satisfies $b_i \geq 1$, then Bus i is observable, otherwise, $b_i = 0$ and Bus i is unobservable. When the communication link between each PMU and the control center is assumed perfect, i.e., noiseless and delay-free, the observability vector $\mathbf{b}$ would capture the quality of the grid monitoring.

Wireless communication is subject to the effect of fading, and it is not feasible to provide deterministic delay bounds for communication systems in most fading channel environments [13]. Thus, the quality of the grid monitoring can be captured by a statistical measure of the observable performance of the power system under the premise of satisfying communication constraints, in terms of both rate and delay requirements [5]. To represent the random observability vector, we first define an $N \times N$ diagonal probability matrix $\mathbf{\Lambda}_P$, where the i-th diagonal component is P_i, which is the probability that the information collected by the PMU installed on Bus i, is transmitted to the control center within the delay constraint $D_{\max}$. Thus, P_i is given as

$$P_i = \begin{cases} p_k & \text{if PMU}_k \text{ is installed at bus } i. \\ 0 & \text{otherwise,} \end{cases} \qquad i = 1, 2, \cdots, N$$

where p_k is the probability that PMU$_k$'s information is delivered within the delay constraint $D_{\max}$. More specifically, denote the random transmission delay incured for PMU$_k$ by D_k, then we have

$$p_k = \Pr\{D_k \leq D_{\max}\}$$

Further define another $N \times N$ diagonal communication constraint matrix $\mathbf{\Lambda}_Q$, where its i-th diagonal element Q_i is a random variable that indicates whether the PMU installed on Bus i can transfer its information to the control center within the delay constraint, then

$$\Pr\{Q_i = 1\} = P_i, \quad \Pr\{Q_i = 0\} = 1 - P_i$$

Thus, we obtain the smart grid observability random vector as

$$\widetilde{\mathbf{b}} = \mathbf{L}\mathbf{\Lambda}_Q\mathbf{x}, \tag{2}$$

and the expected power grid observability vector is given by

$$\bar{\mathbf{b}} = \mathbb{E}\left[\widetilde{\mathbf{b}}\right] = \mathbf{L}\mathbf{\Lambda}_P\mathbf{x}. \tag{3}$$

B. Effective Capacity

Similar to [5], we use the theory of effective capacity to characterize the probability that PMU_k's transmission delay is less than the threshold $D_{\max}$. More specifically, for PMU k, its effective capacity can be expressed as [14]:

$$\begin{aligned}\text{EC}_k &= -\frac{1}{\theta_k T}\ln\mathbb{E}\left\{e^{-\theta_k T B_k \log_2(1+\text{SINR}_k)}\right\} \\ &= \frac{B_k}{\beta_k}\log_2\mathbb{E}\left[(1+\text{SINR}_k)^{\beta_k}\right] \end{aligned} \tag{4}$$

where B_k is bandwidth occupied by PMU_k with a block transmission of duration T, and SINR_k is the signal-to-interference-and-noise ratio (SINR) of PMU_k. The parameter θ_k is called the QoS exponent and

$$\beta_k \triangleq -\frac{\theta_k T B_k}{\ln 2} \tag{5}$$

is the normalized QoS exponent [15]. The parameter θ_k is related to the probability of successful transmission within the delay constraint as

$$p_k = \Pr\{D_k \le D_{\max}\} = 1 - e^{-\theta_k \text{EC}_k D_{\max}} \tag{6}$$

$$= 1 - \left(\mathbb{E}\left[(1+\text{SINR}_k)^{\beta_k}\right]\right)^{\frac{D_{\max}}{T}} \tag{7}$$

It is assumed that PMU k take measurements and generate data at a rate of R_k^{th}. Thus, we require that the effective capacity to be no smaller than the data generating rate, i.e.,

$$\text{EC}_k \ge R_k^{\text{th}}. \tag{8}$$

From Lemma 1 of reference [5], we have following proposition.

Proposition 1 *[5] For a fixed B_k, EC_k is a monotonically decreasing function of θ_k, and for a fixed θ_k, EC_k is a monotonically increasing function of B_K.*

Using (4) and (6), the effective capacity theory provides a cross layer analysis framework for studying the effects of fading channels, latency bound and the related latency bound violation probability [16].

C. Uplink NOMA

In this paper, it is assumed that there are K PMUs who want to transmit information to the base station via uplink NOMA. The base station receives the superposition signal and uses the successive interference cancelation (SIC) principle to decode, and the decoding order is by descending average channel gains [17].

All PMUs transmit to the base station under a Rayleigh fading channel with independent channel states denote by h_i, $i = 1, 2, \cdots, K$ [18]. Thus, the corresponding power gain of the channel, $|h_i|^2$, follows the exponential distribution. As a result, we may express the probability distribution function (PDF) and the cumulative distribution function (CDF) of $|h_i|^2$ as

$$f_i(x) = \frac{1}{\lambda_i'}e^{-\frac{x}{\lambda_i'}}, \quad F_i(x) = 1 - e^{-\frac{x}{\lambda_i'}}, \quad i = 1, 2, \cdots, K \tag{9}$$

respectively, where λ_i' is the mean of $|h_i|^2$.

If too many PMUs are assigned to the same subcarrier, it may cause long delays and high complexity when applying the SIC principle [19]. For simplicity of analysis, we assume that there are two devices per group in NOMA, i.e., two devices will share the same time-frequency resources. We examine one of the groups consisting of PMU_a and PMU_b. Let the two unit-power symbols transmitted by PMU_a and PMU_b be S_a and S_b. For a fair comparison with OMA, we assume that the total transmission power of the two PMUs in the same pairing group is $2P$, while the transmission power of each PMU in OMA is P. Then the respective transmitted power of the two PMUs in the same group is $\alpha_i P$, $i = a, b$, where

$$\alpha_a + \alpha_b = 2. \tag{10}$$

The superposition signal of the two PMUs received by the base station is denoted as [20]

$$Y = \sqrt{\alpha_a P}h_a S_a + \sqrt{\alpha_b P}h_b S_b + N \tag{11}$$

where the receiver noise N is a Gaussian random variable with mean 0 and variance σ^2.

The receiver first sorts the two PMUs according to the average channel gains. Without loss of generality, we assume $\mathbb{E}\left[|h_a|^2\right] < \mathbb{E}\left[|h_b|^2\right]$, then according to SIC theory, the receiver will decode PMU_b's signal first while treating PMU_a's signal as interference [17]. Then, the receiver will subtract the decoded PMU_b's signal from its received signal and decode PMU_a's signal as in a single-user communication channel. Thus, the SINRs of the two PMUs are as follows

$$\text{SINR}_a = \frac{\alpha_a P|h_a|^2}{\sigma^2} = \alpha_a\rho_a \tag{12}$$

$$\text{SINR}_b = \frac{\alpha_b P|h_b|^2}{\alpha_a P|h_a|^2 + \sigma^2} = \frac{\alpha_b\rho_b}{1+\alpha_a\rho_a} \tag{13}$$

where we have defined $\rho_i = \frac{P|h_i|^2}{\sigma^2}$, $i = a, b$, which is exponentially distributed with mean

$$\lambda_i = \frac{P}{\sigma^2}\lambda_i', \tag{14}$$

which we call the average SNR of PMU i.

D. Problem Formulation

Let $\mathcal{G} = \left\{(i^+(g), i^-(g))|g = 1, 2, \cdots, \left\lceil \frac{K}{2} \right\rceil \right\}$ denote the device grouping strategy, which means identifying which two devices belong in the same group g, $g = 1, 2, \cdots, \left\lceil \frac{K}{2} \right\rceil$. For the two devices in Group g, we denote the device with a larger average SNR as $i^+(g)$ and the device with a smaller average SNR as $i^-(g)$. When the total number of devices K is odd, one of the devices will be in a group by itself. In this case, if Group g_0 contains a single device, for ease of presentation, set $i^-(g_0)$ to denote the device, and set $i^+(g_0) = 0$. Correspondingly, set $R^{\text{th}}_{i^+(g)} = 0$, $\lambda_{i^+(g)} = 0$. To satisfy the same power constraint, we set $\alpha_{i^-(g_0)} = \alpha_{i^+(g_0)} = 1$. Notice that each user can belong to one and only one group.

The problem we are interested in, is to minimize the total bandwidth over all possible device grouping strategies, power allocation within each group, and bandwidth allocation strategies among all the groups, such that the observability of the smart grid is above a given threshold. Hence, the optimization problem can be expressed as

$$\max_{\substack{\mathcal{G}, \{B_1, \cdots, B_{\left\lceil \frac{K}{2} \right\rceil}\} \\ \{\alpha_1, \cdots, \alpha_K\}, \{\beta_1, \cdots, \beta_K\}}} \sum_{g=1}^{\left\lceil \frac{K}{2} \right\rceil} B_g \tag{15}$$

$$\text{s.t.} \quad p_{i^-(g)} = 1 - \left(\mathbb{E}\left[\left(1 + \alpha_{i^-(g)} \rho_{i^-(g)}\right)^{\beta_{i^-(g)}} \right] \right)^{\frac{D_{\max}}{T}}, \tag{16}$$

$$p_{i^+(g)} = 1 - \left(\mathbb{E}\left[\left(1 + \frac{\alpha_{i^+(g)} \rho_{i^+(g)}}{1 + \alpha_{i^-(g)} \rho_{i^-(g)}} \right)^{\beta_{i^+(g)}} \right] \right)^{\frac{D_{\max}}{T}}, \tag{17}$$

$$\frac{B_g}{\beta_{i^-(g)}} \log_2 \mathbb{E}\left[\left(1 + \alpha_{i^-(g)} \rho_{i^-(g)}\right)^{\beta_{i^-(g)}} \right] \geq R^{\text{th}}_{i^-(g)}, \tag{18}$$

$$\frac{B_g}{\beta_{i^+(g)}} \log_2 \mathbb{E}\left[\left(1 + \frac{\alpha_{i^+(g)} \rho_{i^+(g)}}{1 + \alpha_{i^-(g)} \rho_{i^-(g)}} \right)^{\beta_{i^+(g)}} \right] \geq R^{\text{th}}_{i^+(g)}, \tag{19}$$

$$\alpha_{i^-(g)} + \alpha_{i^+(g)} = 2, \quad \forall g = 1, 2, \cdots, \left\lfloor \frac{K}{2} \right\rfloor \tag{20}$$

$$\alpha_{i^-(g)} = 1, \quad g = \left\lceil \frac{K}{2} \right\rceil \text{ and } K \text{ is odd} \tag{21}$$

$$\text{Obs}(p_1, p_2, \cdots, p_K) \geq O^{\text{th}} \tag{22}$$

$$B_g \geq 0, \quad \alpha_{i^-(g)} \in [0, 2], \qquad \forall g = 1, 2, \cdots, \left\lceil \frac{K}{2} \right\rceil \tag{23}$$

$$\beta_k \leq 0, \qquad \forall k = 1, 2, \cdots, K \tag{24}$$

where $\text{Obs}(p_1, p_2, \cdots, p_K)$ in (22) is the observability cost function, and it is a function of $(p_1, \cdots, p_K)$, i.e., the probability that each PMU's information is delivered within the delay constraint. In [5], the authors proposed three different performance metrics to characterize the power grid observability:

1) Observability redundancy (OR), which is defined as $\text{Obs}_1(p_1, p_2, \cdots, p_K) \triangleq \mathbf{1}_N^T \mathbf{L} \mathbf{\Lambda}_P \mathbf{x} - N$. When OR is large, it means that the expected sum of the number of PMUs successfully observing each bus is large.
2) Observability sensitivity (OS), which is defined as $\text{Obs}_2(p_1, p_2, \cdots, p_K) \triangleq \min_n \bar{b}_n$, where $\bar{b}_n$ is the n-th element of the expected power grid observability vector $\bar{\mathbf{b}}$ as defined in (3). The bus with the smallest expected power grid observability number is the one least likely to be observed on average, and therefore can be considered as a bottleneck of the observability of the system. Thus, maximizing the bottleneck will increase the observability of the entire grid.
3) Observability probability (OP), which is defined as $\text{Obs}_3(p_1, p_2, \cdots, p_K) \triangleq \Pr\left[\widetilde{\mathbf{b}} \geq \lambda\right]$. This is the probability that the observability random vector $\widetilde{\mathbf{b}}$ is larger than a threshold vector λ. Maximizing this probability will increase the observability of the power grid.

All three metrics characterize the observability of the power grid and we will consider all three in this paper. In terms of the constraints, (16) and (17) follows from (7), (18) and (19) follows from (4) and (8), (20) follows from the power split within the NOMA group, as discussed in (10), and (21) means that when K is odd, the PMU that is in a group by itself uses its own full power P. Note that (16)-(19) holds for all $g = 1, 2, \cdots \left\lceil \frac{K}{2} \right\rceil$. Finally, (22) means that the observability of the power grid is ensured, i.e., above a given threshold O^{th}.

III. Solving the Optimization Problem in (15)

The optimization problem in (15) is very challenging. First of all, there are expectations in the constraints which are hard to evaluate. Secondly, it involves the joint optimization of many parameters, such as the normalized QoS exponents of each PMU, NOMA user pairing, power allocation within each NOMA group, bandwidth allocation among NOMA groups. In this section, we will discuss these issues one-by-one.

A. Effective Capacity Of the Two-PMU Uplink NOMA System

In this subsection, we derive easy-to-evaluate expressions for the effective capacity of NOMA two-PMU pairs, i.e., the left-hand-side (LHS) of (18) and (19). Notice that the same pair of expectations exists in (16) and (17).

For ease of presentation, let us study a NOMA user group consisting of PMU_a and PMU_b, with $\lambda_a < \lambda_b$, i.e., the average SNR of PMU_b is larger and therefore is decoded first. We assume the group occupies a bandwidth of B_{ab}. Then, the effective capacity of the two PMUs can be written as

$$\text{EC}_a = \frac{B_{ab}}{\beta_a} \log_2 \mathbb{E}\left[\left(1 + \alpha_a \rho_a\right)^{\beta_a} \right], \tag{25}$$

$$\text{EC}_b = \frac{B_{ab}}{\beta_b} \log_2 \mathbb{E}\left[\left(1 + \frac{\alpha_b \rho_b}{\alpha_a \rho_a + 1} \right)^{\beta_b} \right] \tag{26}$$

We have the following lemma.

Lemma 1 *The effective capacity expressions of the two-user uplink NOMA system for PMU_a and PMU_b with $\lambda_a < \lambda_b$ is*

$$EC_a = \frac{B_{ab}}{\beta_a} \log_2 \left(\frac{1}{\lambda_a \alpha_a} U\left(1, 2 + \beta_a, \frac{1}{\lambda_a \alpha_a}\right) \right)$$

$$EC_b \approx \frac{B_{ab}}{\beta_b}\log_2\left(\frac{\alpha_a\lambda_a}{\alpha_b\lambda_b}\cdot\frac{1}{1-\beta_b}\cdot {}_2F_1\left(1,2;2-\beta_b;\frac{\alpha_b\lambda_b-\alpha_a\lambda_a}{\alpha_b\lambda_b}\right)\right) \quad (27)$$

where $U(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function, i.e.,

$$U(a,b,z)=\frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{\alpha-1}(1+t)^{b-a-1}dt, \quad (28)$$

$\Gamma(\cdot)$ is the gamma function, and ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ is the hypergeometric function, i.e.,

$$\int_0^\infty x^{a-1}e^{-sx}\Gamma(b,x)dx = \frac{\Gamma(a+b)}{a(1+s)^{a+b}}F\left(1,a+b;1+a;\frac{s}{1+s}\right) \quad (29)$$

and the approximation in (27) follows from omitting the noise in the SINR expression of PMU$_b$.

The proof for Lemma 1 is provided in Appendix A. Based on Lemma 1, we have the following easy-to-evaluate expression to substitute into constraints (16) and (18), i.e.,

$$\mathbb{E}\left[\left(1+\alpha_{i^-(g)}\rho_{i^-(g)}\right)^{\beta_{i^-(g)}}\right] = \frac{1}{\lambda_{i^-(g)}\alpha_{i^-(g)}}U\left(1,2+\beta_{i^-(g)},\frac{1}{\lambda_{i^-(g)}\alpha_{i^-(g)}}\right) \quad (30)$$

and the following expression to substitute into constraints (17) and (19), i.e.,

$$\mathbb{E}\left[\left(1+\frac{\alpha_{i^+(g)}\rho_{i^+(g)}}{1+\alpha_{i^-(g)}\rho_{i^-(g)}}\right)^{\beta_{i^+(g)}}\right] = \frac{\alpha_{i^-(g)}\lambda_{i^-(g)}}{\alpha_{i^+(g)}\lambda_{i^+(g)}}\cdot\frac{1}{1-\beta_{i^+(g)}}\cdot {}_2F_1\left(1,2;2-\beta_{i^+(g)};\frac{\alpha_{i^+(g)}\lambda_{i^+(g)}-\alpha_{i^-(g)}\lambda_{i^-(g)}}{\alpha_{i^+(g)}\lambda_{i^+(g)}}\right) \quad (31)$$

After we substitute (30) into constraints (16) and (18), and (31) into constraints (17) and (19), we can go on to solve the joint optimization of the normalized QoS exponents of each PMU, NOMA user pairing, power allocation within each NOMA group, and bandwidth allocation among NOMA groups.

B. Finding the optimal $\{\alpha_1,\cdots,\alpha_K\}$ and $\{\beta_1,\cdots,\beta_K\}$ for a given $\{p_1,\cdots,p_K\}$ and $\mathcal{G}$

In this subsection, we discuss how to find the optimal normalized QoS exponent of each PMU and the optimal power allocation within each NOMA group for a given set of probability of successful transmission, i.e., $\{p_1,\cdots,p_K\}$, that satisfies (22).

From (7), we have

$$\mathbb{E}\left[(1+\text{SINR}_k)^{\beta_k}\right]=(1-p_k)^{\frac{T}{D_{\max}}}, \quad k=1,\cdots,K \quad (32)$$

Using (32) together with (4), we have

$$\text{EC}_k=\frac{B_{g(k)}}{\beta_k}\log_2(1-p_k)^{\frac{T}{D_{\max}}} = \frac{B_{g(k)}T}{\beta_k D_{\max}}\log_2(1-p_k) \quad (33)(34)$$

where $g(k)$ is the index of the group to which PMU$_k$ belongs. From (8) and (34), we have

$$B_{g(k)}\geq\frac{D_{\max}}{T}\frac{\beta_k R_k^{\text{th}}}{\log_2(1-p_k)} \quad (35)$$

Since there are two users in each group g, for $g=1,\cdots,\lfloor\frac{K}{2}\rfloor$, the bandwidth of group g must satisfy (35) for each PMU in group g, i.e.,

$$B_{g(k)}\geq\frac{D_{\max}}{T}\max\left\{\frac{R^{\text{th}}_{i^-(g)}|\beta_{i^-(g)}|}{-\log_2(1-p_{i^-(g)})},\frac{R^{\text{th}}_{i^+(g)}|\beta_{i^+(g)}|}{-\log_2(1-p_{i^+(g)})}\right\} \quad (36)$$

Since the cost function is to minimize the sum of the bandwidth, for each group g, we would like to minimize $\max\left\{\frac{R^{\text{th}}_{i^-(g)}|\beta_{i^-(g)}|}{-\log_2(1-p_{i^-(g)})},\frac{R^{\text{th}}_{i^+(g)}|\beta_{i^+(g)}|}{-\log_2(1-p_{i^+(g)})}\right\}$ by varying over $\alpha_g\in[0,2]$.

When $\{p_1,\cdots,p_K\}$ is fixed, the relationship between $\alpha_{i^-(g)}$, $\beta_{i^-(g)}$ and $\beta_{i^+(g)}$ is given by (32), i.e., for each fixed $\alpha_{i^-(g)}$, $\beta_{i^-(g)}$ and $\beta_{i^+(g)}$ is the root of (32) and can be found by the bisection search method, more specifically, from (30), for the weak user $i^-(g)$, $\beta_{i^-(g)}$ is the root of the equation

$$(1-p_k)^{\frac{T}{D_{\max}}}=\frac{1}{\lambda_{i^-(g)}\alpha_{i^-(g)}}U\left(1,2+\beta_{i^-(g)},\frac{1}{\lambda_{i^-(g)}\alpha_{i^-(g)}}\right), \quad (37)$$

and for the strong user $i^+(g)$, $\beta_{i^+(g)}$ is the root of the equation

$$(1-p_k)^{\frac{T}{D_{\max}}}=\frac{\alpha_{i^-(g)}\lambda_{i^-(g)}}{\alpha_{i^+(g)}\lambda_{i^+(g)}}\cdot\frac{1}{1-\beta_{i^+(g)}}\cdot {}_2F_1\left(1,2;2-\beta_{i^+(g)};\frac{\alpha_{i^+(g)}\lambda_{i^+(g)}-\alpha_{i^-(g)}\lambda_{i^-(g)}}{\alpha_{i^+(g)}\lambda_{i^+(g)}}\right), \quad (38)$$

which follows from (31), and $\alpha_{i^+(g)}=2-\alpha_{i^-(g)}$.

For the weak user of Group g, when $\alpha_{i^-(g)}$ increases, its average SINR increases, so to satisfy (32), $|\beta_{i^-(g)}|$ decreases. For the strong user of Group g, when $\alpha_{i^-(g)}$ increases, its average SINR decreases, so to satisfy (32), $|\beta_{i^+(g)}|$ increases. Hence, the $\alpha_{i^-(g)}$ that minimizes $\max\left\{\frac{R^{\text{th}}_{i^-(g)}|\beta_{i^-(g)}|}{-\log_2(1-p_{i^-(g)})},\frac{R^{\text{th}}_{i^+(g)}|\beta_{i^+(g)}|}{-\log_2(1-p_{i^+(g)})}\right\}$ is the one that makes the two terms equal, i.e.,

$$\frac{R^{\text{th}}_{i^-(g)}\beta_{i^-(g)}}{\log_2(1-p_{i^-(g)})}=\frac{R^{\text{th}}_{i^+(g)}\beta_{i^+(g)}}{\log_2(1-p_{i^+(g)})} \quad (39)$$

Such an $\alpha_{i^-(g)}$ can be found by the bisection search method.

Algorithm 1 summarizes the above discussion and propose a way to find the optimal $\{\alpha_1,\cdots,\alpha_K\}$ and $\{\beta_1,\cdots,\beta_K\}$ for a given $\{p_1,\cdots,p_K\}$ and $\mathcal{G}$. Note that in Line 9 of Algorith 1, we have defined $f(\beta_{i^-(g)},\beta_{i^+(g)})$ as

$$f(\beta_{i^-(g)},\beta_{i^+(g)})\triangleq\frac{R^{\text{th}}_{i^-(g)}\beta_{i^-(g)}}{\log_2(1-p_{i^-(g)})}-\frac{R^{\text{th}}_{i^+(g)}\beta_{i^+(g)}}{\log_2(1-p_{i^+(g)})} \quad (40)$$

Algorithm 1: Finding the optimal $\{\alpha_1, \cdots, \alpha_K\}$ and $\{\beta_1, \cdots, \beta_K\}$ for a given $\{p_1, \cdots, p_K\}$ and $\mathcal{G}$

Input: $\mathcal{G} \leftarrow$ PMU pairing result
$p_k \leftarrow$ Probability of successful transmission of PMU$_k$ $k = 1, \cdots, K$
$R_k^{\text{th}} \leftarrow$ Minimum data requirmenet of PMU$_k$ $k = 1, \cdots, K$

1: **initialization:** $\alpha_g^{lo} = 0, \alpha_g^{hi} = 2$, $\forall g = 1, \cdots, \lfloor \frac{K}{2} \rfloor$
2: **for** $g = 1, \cdots, \lfloor \frac{K}{2} \rfloor$ **do**
3: **repeat**
4: $\alpha_g^{mid} = \alpha_g^{lo} + (\alpha_g^{hi} - \alpha_g^{lo})/2$
5: Find $\beta_{i^-(g)}^{mid}$ that satisfies (37) using the bisection search method for $\alpha_{i^-(g)} = \alpha_g^{mid}$
6: Find $\beta_{i^+(g)}^{mid}$ that satisfies (38) using the bisection search method for $\alpha_{i^-(g)} = \alpha_g^{mid}$ and $\alpha_{i^+(g)} = 2 - \alpha_g^{mid}$
7: Find $\beta_{i^-(g)}^{lo}$ that satisfies (37) using the bisection search method for $\alpha_{i^-(g)} = \alpha_g^{lo}$
8: Find $\beta_{i^+(g)}^{lo}$ that satisfies (38) using the bisection search method for $\alpha_{i^-(g)} = \alpha_g^{lo}$ and $\alpha_{i^+(g)} = 2 - \alpha_g^{lo}$
9: **if** $f(\beta_{i^-(g)}^{mid}, \beta_{i^+(g)}^{mid}) \cdot f(\beta_{i^-(g)}^{lo}, \beta_{i^+(g)}^{lo}) > 0$ **then**
10: $\alpha_g^{lo} = \alpha_g^{mid}$
11: **else**
12: $\alpha_g^{hi} = \alpha_g^{mid}$
13: **end if**
14: **until** $\left|\alpha_g^{hi} - \alpha_g^{lo}\right| \leqslant \epsilon$
15: **end for**
16: **for** $g = \lceil \frac{K}{2} \rceil$ when K is odd **do**
17: Find $\beta_{i^-(g)}^{mid}$ that satisfies (37) using the bisection search method for $\alpha_{i^-(g)} = 1$
18: **end for**
19: Calculate $B_g = \frac{D_{\max}}{T} \frac{R_{i^-(g)}^{\text{th}} \beta_{i^-(g)}^{mid}}{\log_2(1 - p_{i^-(g)})}$, $g = 1, \cdots, \lceil \frac{K}{2} \rceil$

Output: $\beta_{i^-(g)}, \beta_{i^+(g)}$ for all $g = 1, 2, \cdots, \lceil \frac{K}{2} \rceil$
$\alpha_{i^-(g)} = \alpha_g^{mid}$, $\alpha_{i^+(g)} = 2 - \alpha_g^{mid}$ for all $g = 1, 2, \cdots, \lfloor \frac{K}{2} \rfloor$;
$\alpha_{i^-(\lceil \frac{K}{2} \rceil)} = 1$ when K is odd;
B_g, $g = 1, \cdots, \lceil \frac{K}{2} \rceil$

Note that Line 2-Line 15 of Algorithm 1 is using the bisection search method to find $\alpha_{i^-(g)}$, for $g = 1, 2, \cdots, \lfloor \frac{K}{2} \rfloor$. When K is odd, we have a PMU who is in a NOMA group by itself, and its corresponding α is equal to 1. Line 16-18 describes how to find its normalized QoS exponent β. The output of Algorithm 1 is the optimal normalized QoS exponent of the PMUs, the optimal power allocation and the corresponding bandwidth for each NOMA group for a given $\{p_1, \cdots, p_K\}$ and $\mathcal{G}$.

C. Solving the optimal $\{p_1, \cdots, p_K\}$ for a fixed $\mathcal{G}$

The PMU groups interact with one another only through the constraint (22), and we propose to use the simulated annealing algorithm to find the optimal $\{p_1, \cdots, p_K\}$ for a fixed NOMA pairing result $\mathcal{G}$. First, we deal with the constraints of the optimization problem by the interior penalty function method [21]. More specifically, the new objective function, denoted as $h_{\mathcal{G}}\left(p_1, \cdots, p_K, \alpha_1, \cdots, \alpha_K, \beta_1, \cdots, \beta_K, B_1, \cdots, B_{\lceil \frac{K}{2} \rceil}\right)$, obtained by adding the constraint to the original objective function is

$$\begin{aligned} &\max_{p_1, \cdots, p_K} h_{\mathcal{G}}\Big(p_1, \cdots, p_K, \alpha_1, \cdots, \alpha_K, \beta_1, \cdots, \beta_K, \\ &\qquad B_1, \cdots, B_{\lceil \frac{K}{2} \rceil}\Big) \\ &\triangleq \sum_{g=1}^{\lceil \frac{K}{2} \rceil} B_g + m_j \Bigg[\left(\max(O^{\text{th}} - \text{Obs}(p_1, p_2, \cdots, p_K), 0)\right)^2 \\ &\qquad + \sum_{k=1}^{K} (\max(-p_k, 0))^2 + \sum_{k=1}^{K} (\max(p_k - 1, 0))^2 \Bigg] \end{aligned} \quad (41)$$

where the parameter m_j is called the penalty factor, which will continue to increase with iteration, whose index is given by j. The proposed algorithm to solve the problem in (15) for a fixed $\mathcal{G}$ is given by Algorithm 2.

To initialize, we assume that all p_ks are equal to a fixed number p_0, say 0.9. Given this set of p_ks, we find the value of the new objective function in (41) by first finding the corresponding optimal $\{\alpha_1^*, \cdots, \alpha_K^*\}$, $\{\beta_1^*, \cdots, \beta_K^*\}$ and $\{B_1, \cdots, B_{\lceil \frac{K}{2} \rceil}\}$ using Algorithm 1. The corresponding value of the new objective function is saved as H. For each temperature t, we run the algorithm M times, where during each time, a random neighbor $\{p_1', \cdots, p_K'\}$ is generated and its corresponding new objective function, denoted as H', is calculated. If H' is better than H, accept the neighbor as the optimal scheme and update H to be H', as indicated by Line 10. If H' is no better than H, accept the neighbor with a certain probability, as indicated on Line 12. Note that the worse H' is, and the lower the temperature, the smaller the probability that the worse neighbor will be accepted. Note that even when the worse neighborhood strategy is accepted, H still records the best value of the new objective function. After running the iteration M times, the temperature is lowered in Line 16, where $\gamma < 1$. The whole algorithm stops when the lowest temperature is reached, and the final bandwidth allocation strategy, and its corresponding power allocation and normalized QoS exponents obtained via Algorithm 1, are the output of the algorithm.

D. PMU Pairing Scheme

Finally, we discuss the PMU pairing scheme to obtain the PMU pairing result $\mathcal{G}$. PMU pairing is a discrete optimization problem and thus very challenging. Exhaustive search over all possible user pairings yields optimal results, but will result in very high computational complexity.

To determine whether a new NOMA pairing algorithm is necessary for the problem in (15), or existing NOMA pairing algorithms, such as near-far pairing algorithm, UCGD pairing

Algorithm 2: Algorithm to solve the problem in (15) for a fixed $\mathcal{G}$

Input: $\mathcal{G} \leftarrow$ PMU pairing result
$T_{\text{initial}} \leftarrow$ Initial temperature
$T_{\text{final}} \leftarrow$ Stop temperature
$\gamma \leftarrow$ Attenuation coefficient

initialization:
$p_k = p_0$, $k = 1, \cdots, K$
Run Algorithm 1 and find the corresponding optimal $\{\alpha_1^*, \cdots, \alpha_K^*\}$, $\{\beta_1^*, \cdots, \beta_K^*\}$ and $\{B_1, \cdots, B_{\lceil \frac{K}{2} \rceil}\}$
Set $H \triangleq h_{\mathcal{G}}(p_1, \cdots, p_K, \alpha_1^*, \cdots, \alpha_K^*, \beta_1^*, \cdots, \beta_K^*, B_1, \cdots, B_{\lceil \frac{K}{2} \rceil})$
$t = T_{\text{initial}}$
repeat
  $m = M \leftarrow$ Markov chain length
  repeat
    Generate a random neighbor $\{p_1', \cdots, p_K'\}$
    Calculate the optimal $\{\alpha_1^{'*}, \cdots, \alpha_K^{'*}\}$, $\{\beta_1^{'*}, \cdots, \beta_K^{'*}\}$ and $\{B_1', \cdots, B_{\lceil \frac{K}{2} \rceil}'\}$ according to Algorithm 1
    Set $H' \triangleq h_{\mathcal{G}}(p_1', \cdots, p_K', \alpha_1^{'*}, \cdots, \alpha_K^{'*}, \beta_1^{'*}, \cdots, \beta_K^{'*}, B_1', \cdots, B_{\lceil \frac{K}{2} \rceil}')$
    $\Delta E = H' - H$
    if $\Delta E < 0$ **then**
      $p_k = p_k'$, $\quad \forall k = 1, \cdots, K$, $H = H'$
    else
      with probability $e^{-\frac{\Delta E}{t}}$, take $p_k = p_k'$, $\forall k = 1, \cdots, K$
    end if
    $m = m - 1$
  until $m = 0$
  Update temperature $t = \gamma t$.
until $t \leq T_{\text{final}}$

Output: $p_1, \cdots, p_K$ and the corresponding optimal $\{\alpha_1^*, \cdots, \alpha_K^*\}$, $\{\beta_1^*, \cdots, \beta_K^*\}$ and $\{B_1, \cdots, B_{\lceil \frac{K}{2} \rceil}\}$ found via Algorithm 1

algorithm, and hybrid pairing algorithm [11], [22], [23], are good enough, we ran some numerical results on different parameter settings for the IEEE 14-bus power system with 9 PMUs. We find that even with exhaustive search over all user pairing, the performance gain over existing NOMA pairing algorithms is very slight, and thus, we conclude that proposing a new NOMA pairing algorithm for the problem in (15) is not needed. For example, for the case where the average SNR of all PMUs are set to 21, we ran Algorithm 2 for each NOMA user pairing possible, and we find that the best user pairing is $(2,5),(3,4),(8,6),(7,9),(1)$, which results in a required bandwidth of 214.2 kHz to achieve an OP of 99.99%, and the worst user pairing is $(3,4)(8,1)(7,5)(6,9),(2)$, which results in a required bandwidth of 224.1 kHz to achieve an OP of 99.99%. The saving of bandwidth is 4.42% when using the best pairing compared to the worst, and this saving is very slight. Since the averages SNRs are all the same, the slight saving comes from the heterogeneity of the PMUs where some PMUs contribute more to observability than others. But as can be seen, exploiting this heterogeneity of PMUs does not increase the performance significantly, even when the average SNRs of the PMUs are the same. When the average SNRs of the PMUs are different, the gain in performance will be even less. Hence, in this paper, we choose not to propose a new user pairing algorithm for the problem in (15).

Among existing PMU pairing schemes, we find that the UCGD performs slightly better than near-far and hybrid. Hence, for the problem in (15), we propose to use the UCGD user pairing algorithm for the problem in (15).

For completeness, we describe the UCGD pairing algorithm in the following. In a power grid system of N buses, where K PMUs are installed, for PMU$_k$, denote the bus it is installed on as i_k. Order the PMUs in an ascending order based on the average SNRs, i.e., suppose the j-th smallest average SNR belongs to PMU k_j, then we have $\lambda_{k_1} < \lambda_{k_2} < \cdots < \lambda_{k_K}$. Divide the PMUs into two sets according to the average SNR, i.e., the weak PMU set

$$\mathcal{S}_1 = \left\{k_1, k_2, \cdots, k_{\lfloor \frac{K}{2} \rfloor}\right\}, \tag{42}$$

and the strong PMU set

$$\mathcal{S}_2 = \left\{k_{\lceil \frac{K}{2} \rceil + 1}, k_{\lceil \frac{K}{2} \rceil + 2}, \cdots, k_K\right\}, \tag{43}$$

i.e., when K is even, the weak PMU set consists of the half of the weakest users, and the strong PMU set consists of the remaining users. When K is odd, the middle user is left out and does not belong to the weak PMU set or the strong PMU set. It will not be paired with another PMU for NOMA, and will transmit in the time-frequency block by itself. Once we have obtained the ordered sets of $\mathcal{S}_1$ and $\mathcal{S}_2$, the UCGD pairing principle is, users with a higher priority in one set are paired with those with a higher priority in another set, i.e., Pair$_i$ $= \left(k_i, k_{\lceil \frac{K}{2} \rceil + i}\right)$, $i = 1, 2, \cdots, \lfloor \frac{K}{2} \rfloor$. In other words, the NOMA user pairing is given as

$$\mathcal{G} = \left\{ \left(i^+(g), i^-(g)\right) \middle| i^-(g) = k_g, \right. \\ \left. i^+(g) = k_{\lceil \frac{K}{2} \rceil + g}, g = 1, 2, \cdots, \left\lfloor \frac{K}{2} \right\rfloor \right\}$$

when K is even, and when K is odd, we have

$$\mathcal{G} = \left\{\left(k_{\lceil \frac{K}{2} \rceil}, 0\right)\right\} \bigcup \left\{ \left(i^+(g), i^-(g)\right) \middle| i^-(g) = k_g, \right. \\ \left. i^+(g) = k_{\lceil \frac{K}{2} \rceil + g}, g = 1, 2, \cdots, \left\lfloor \frac{K}{2} \right\rfloor \right\}$$

IV. Numerical Results

In this section, we show through numerical results the performance of the NOMA-based communication for power grid observability proposed in this paper. The results are obtained using the IEEE 14-bus power system test case [24].

We assume that all PMUs generate data at the rate of 60 kbps and the maximum allowed latency bound for these measurement packages is set to be 10 ms [25]. The value of PMU installation vector is [12]

$$\mathbf{x} = [0 \;\; 1 \;\; 0 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 0 \;\; 1 \;\; 0 \;\; 1 \;\; 0]^T. \qquad (44)$$

For a fair comparison with orthogonal multiple access (OMA), we assume that the total transmission power of the two PMUs in the same pairing group is $2P$, while the transmission power of each PMU in OMA is P. The average SNR λ_k for PMU$_k$ defined in (14) is given in Table I.

TABLE I
INSTALLATION AND AVERAGE SNR FOR EACH PMU

PMU index k	1	2	3	4	5	6	7	8	9
Installed on Bus	2	4	5	6	7	8	9	11	13
Average SNR	13	32	31	12	10	25	29	30	11

Fig. 1 illustrates the bandwidth needed for different observability thresholds when OR observability is considered. We plot the performance of three algorithms: the red dotted-stared line corresponds to the performance of orthogonal multiple access (OMA) [5], the blue dash-dot-triangle line corresponds to the case of NOMA with equal bandwidth allocation among all groups of two users and equal power allocation within each group. Note that the group with one user uses half of the bandwidth of the groups with two users. The cyan solid-circle line illustrates the performance of the proposed NOMA-based resource allocation algorithm, i.e., Algorithm 2 with UCGD NOMA pairing. As can be seen, NOMA with equal bandwidth and power may perform even worse than OMA, which means that judicious allocation of power and bandwidth is important in applying the NOMA technique. The highest achievable OR observability is 25 and this can only be reached with infinite bandwidth. At the observability threshold of 24, the proposed NOMA strategy can achieve a bandwidth saving of 19.8% compared to OMA and 28.6% compared to NOMA with equal bandwidth and power. At the observability of threshold of $25 \cdot 99\% = 24.75$, the proposed NOMA strategy can achieve a bandwidth saving of 17.1% compared to OMA and 37.7% compared to NOMA with equal bandwidth and power. Similar performance gains may be observed when OS observability is used, as shown in Fig. 2. More specifically, with infinite bandwidth, the highest achievable OS is 2. At the observability threshold of 1.9, the proposed NOMA strategy can achieve a bandwidth saving of 23.1% compared to OMA and 30.7% compared to NOMA with equal bandwidth and power. At the observability of threshold of $2 \cdot 99\% = 1.98$, the proposed NOMA strategy can achieve a bandwidth saving of 21.5% compared to OMA and 42.2% compared to NOMA with equal bandwidth and power.

OP observability characterizes the probability that the observability of each individual bus exceeds the corresponding element of the threshold vector λ, i.e., Pr $\left[\widetilde{\mathbf{b}} \geq \lambda\right]$. Here, we take $\lambda = \mathbf{1}_N$. In addition to the performance shown in Fig. 3, we provide the following table which gives the bandwidth needed for achieving the OP observability of 99%, 99.9% and 99.99%, respectively. As can be seen, compared to OMA, at an OP observability threshold of 99%, 99.9% and 99.99%, the proposed NOMA achieves a bandwidth saving of 25%, 19.7% and 11.7%, respectively. We observe that for all OR, OS and OP observability, the gain of the proposed algorithm over OMA becomes smaller when the observability requirements becomes more stringent.

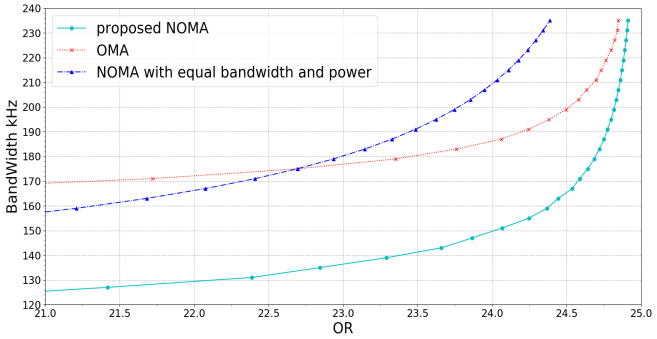

Fig. 1. OR observability

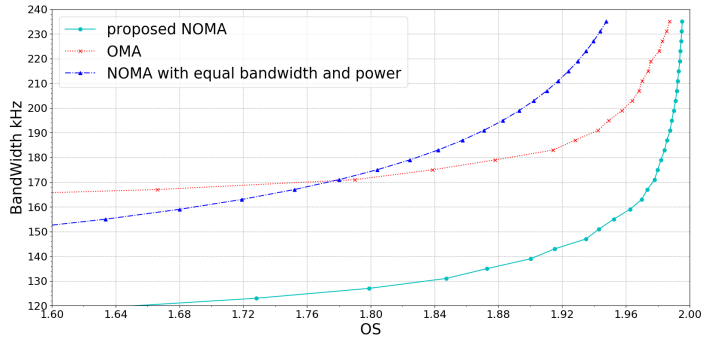

Fig. 2. OS observability

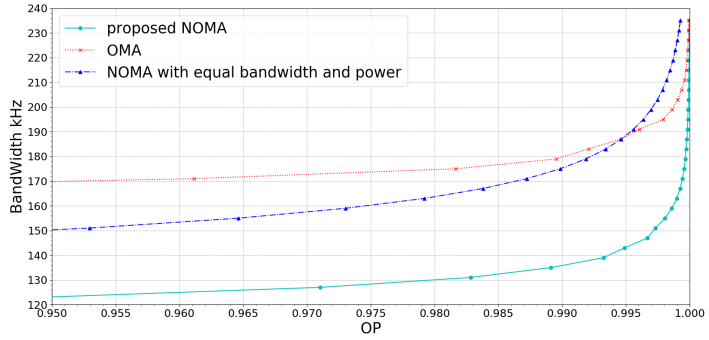

Fig. 3. OP observability

TABLE II
BANDWIDTH FOR DIFFERENT OP REQUIREMENTS

OP threshold	Proposed NOMA	NOMA with equal bandwidth and power	OMA
99%	135kHz	175kHz	180kHz
99.9%	163kHz	227kHz	203kHz
99.99%	203kHz	307kHz	230kHz

V. CONCLUSION

This paper studies the observability of the power grid by jointly considering the power system and the wireless communication system. In order to meet the observability constraint and save communication bandwidth, we adopt NOMA as the multiple access method. We derived the approximate closed-form effective capacity expressions of the NOMA user pairs. According to the three observability performance metrics of OR, OS and OP, the resource allocation problem of smart grid observability under communication constraints in the uplink NOMA system is proposed. Using the UCGD pairing as the PMU pairing strategy as it is both simple and offers good performance, we propose an algorithm to find the optimal bandwidth allocation among NOMA groups, power allocation within each NOMA group and the normalized QoS exponents of each PMU via the bisection search method and simulated annealing. Numerical results show that significant bandwidth savings can be achieved compared to the OMA method and the method of NOMA with equal bandwidth and power allocation.

APPENDIX A
PROOF OF LEMMA 1

A. Effective capacity of PMU$_a$

Let $c=\frac{1}{\lambda_a}$, the effective capacity of PMU$_a$ is:

$$\begin{aligned}\mathrm{EC}_a &= \frac{B_{ab}}{\beta_a}\log_2\mathbb{E}\left\{(1+\alpha_a\rho_a)^{\beta_a}\right\}\\ &= \frac{B_{ab}}{\beta_a}\log_2\int_0^\infty (1+\alpha_a x)^{\beta_a}\, ce^{-cx}dx\end{aligned}$$

Let $t=\alpha_a x$, then $x=\frac{1}{\alpha_a}t$, $dx=\frac{1}{\alpha_a}dt$, then we have:

$$\begin{aligned}\mathrm{EC}_a &= \frac{B_{ab}}{\beta_a}\log_2\int_0^\infty (1+t)^{\beta_a}\, e^{-\frac{c}{\alpha_a}t}\frac{c}{\alpha_a}dt\\ &= \frac{B_{ab}}{\beta_a}\log_2\left(\frac{c}{\alpha_a}U\left(1,2+\beta_a,\frac{c}{\alpha_a}\right)\right)\\ &= \frac{B_{ab}}{\beta_a}\log_2\left(\frac{1}{\lambda_a\alpha_a}U\left(1,2+\beta_a,\frac{1}{\lambda_a\alpha_a}\right)\right)\end{aligned}$$

where $\Gamma(\cdot)$ is the gamma function, and $U(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function, i.e.,

$$U(a,b,z)=\frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}dt$$

B. Effective capacity of PMU$_b$

The effective capacity of PMU$_b$ is:

$$\mathrm{EC}_b=\frac{B_{ab}}{\beta_b}\log_2\mathbb{E}\left[\left(1+\frac{\alpha_b\rho_b}{1+\alpha_a\rho_a}\right)^{\beta_b}\right]$$

In order to perform further derivations, we make an approximation by disregarding the noise, i.e., we omit the "1" in the denominator of SNR. This approximation is tight as we normally operate in an regime where the interference power is much larger than the noise power. In doing so, we have

$$\begin{aligned}\mathrm{EC}_b &\approx \frac{B_{ab}}{\beta_b}\log_2\mathbb{E}\left[\left(1+\frac{\alpha_b\rho_b}{\alpha_a\rho_a}\right)^{\beta_b}\right]\\ &= \frac{B_{ab}}{\beta_b}\log_2\left(\int_0^\infty\int_0^\infty\left(1+\frac{\alpha_b x_2}{\alpha_a x_1}\right)^{\beta_b}\frac{1}{\lambda_a\lambda_b}\right.\\ &\qquad\left.e^{-\frac{x_1}{\lambda_a}-\frac{x_2}{\lambda_b}}dx_1dx_2\right)\end{aligned}$$

Let's look at this integral expression first:

$$\begin{aligned}I_b &\triangleq \int_0^\infty\int_0^\infty\left(1+\frac{\alpha_b x_2}{\alpha_a x_1}\right)^{\beta_b}\frac{1}{\lambda_a\lambda_b}e^{-\frac{x_1}{\lambda_a}-\frac{x_2}{\lambda_b}}dx_1dx_2\\ &= \int_0^\infty\int_0^\infty\left(1+\frac{\alpha_b\lambda_b t_2}{\alpha_a\lambda_a t_1}\right)^{\beta_b}e^{-t_1}e^{-t_2}dt_1dt_2\end{aligned}$$

By letting $z=\alpha_a\lambda_a t_1+\alpha_b\lambda_b t_2$, i.e., $t_2=\frac{1}{\alpha_b\lambda_b}z-\frac{\alpha_a\lambda_a}{\alpha_b\lambda_b}t_1$, we have

$$I_b=\int_0^\infty\frac{e^{-t_1}e^{\frac{\alpha_a\lambda_a}{\alpha_b\lambda_b}t_1}}{(\alpha_a\lambda_a t_1)^{\beta_b}}\left(\int_{\alpha_a\lambda_a t_1}^\infty z^{\beta_b}e^{-\frac{1}{\alpha_b\lambda_b}z}\frac{1}{\alpha_b\lambda_b}dz\right)dt_1$$

Further let $\frac{z}{\alpha_b\lambda_b}=u$, we have:

$$I_b=\int_0^\infty\frac{(\alpha_b\lambda_b)^{\beta_b}}{(\alpha_a\lambda_a)^{\beta_b}}\frac{e^{-t_1}}{t_1^{\beta_b}}e^{\frac{\alpha_a\lambda_a}{\alpha_b\lambda_b}t_1}\left(\int_{\frac{\alpha_a\lambda_a}{\alpha_b\lambda_b}t_1}^\infty u^{\beta_b}e^{-u}du\right)dt_1$$

By using the incomplete gamma function, i.e., $\Gamma(s,x)=\int_x^\infty m^{s-1}e^{-m}dm$, we have:

$$I_b=\int_0^\infty\frac{(\alpha_b\lambda_b)^{\beta_b}}{(\alpha_a\lambda_a)^{\beta_b}}\frac{e^{-t_1}}{t_1^{\beta_b}}e^{\frac{\alpha_a\lambda_a}{\alpha_b\lambda_b}t_1}\cdot\Gamma\left(\beta_b+1,\frac{\alpha_a\lambda_a}{\alpha_b\lambda_b}t_1\right)dt_1$$

By letting $\frac{\alpha_a\lambda_a}{\alpha_b\lambda_b}t_1=m$, we have:

$$I_b=\frac{\alpha_b\lambda_b}{\alpha_a\lambda_a}\int_0^\infty m^{-\beta_b}e^{-\left(\frac{\alpha_b\lambda_b}{\alpha_a\lambda_a}-1\right)m}\cdot\Gamma(\beta_b+1,m)dm$$

Note that $\int_0^\infty x^{a-1}e^{-sx}\Gamma(b,x)dx=\frac{\Gamma(a+b)}{a(1+s)^{a+b}}{}_2F_1(1,a+b;1+a;\frac{s}{1+s})$, where ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ is hypergeometric function. Thus,

$$I_b=\frac{\alpha_a\lambda_a}{\alpha_b\lambda_b}\cdot\frac{1}{1-\beta_b}\cdot{}_2F_1\left(1,2;2-\beta_b;\frac{\alpha_b\lambda_b-\alpha_a\lambda_a}{\alpha_b\lambda_b}\right)\tag{45}$$

Then, based on (45), the effective capacity of PMU$_2$ is approximated as

$$\mathrm{EC}_b\approx\frac{B_{ab}}{\beta_b}\log_2(I_b)$$

which completes the proof of Lemma 1.

REFERENCES

[1] Sanaz Sadeghi, Mohammad Hossein Yaghmaee Moghddam, Maliheh Bahekmat, and Ashraf Sadat Heydari Yazdi. Modeling of Smart Grid traffics using non-preemptive priority queues. In *Iranian Conference on Smart Grids*, pages 1–4, 2012.

[2] Arun G Phadke and John Samuel Thorp. *Synchronized Phasor Measurements and Their Applications*. Springer US, 2008.

[3] Xu Bei, Yeo Jun Yoon, and Ali Abur. Optimal placement and utilization of phasor measurements for state estimation. *The 15th Power Systems Computation Conference, PSCC 2005*, (August):22–26, 2005.

[4] Vehbi C. Gungor, Dilan Sahin, Taskin Kocak, Salih Ergut, Concettina Buccella, Carlo Cecati, and Gerhard P. Hancke. Smart Grid Technologies: Communication Technologies and Standards. *IEEE Transactions on Industrial Informatics*, 7(4):529–539, 2011.

[5] Minglei You, Jing Jiang, Andrea M. Tonello, Tilemachos Doukoglou, and Hongjian Sun. On statistical power grid observability under communication constraints (invited paper). *IET Smart Grid*, 1(2):40–47, 2018.

[6] Minglei You, Hongjian Sun, Jing Jiang, and Jiayi Zhang. Unified Framework for the Effective Rate Analysis of Wireless Communication Systems over MISO Fading Channels. *IEEE Transactions on Communications*, 65(4):1775–1785, 2017.

[7] Mohammed Al-Imari, Pei Xiao, Muhammad Ali Imran, and Rahim Tafazolli. Uplink non-orthogonal multiple access for 5G wireless networks. *The 11th International Symposium on Wireless Communications Systems, ISWCS 2014 - Proceedings*, (November):781–785, 2014.

[8] Hina Tabassum, Md Shipon Ali, Ekram Hossain, Md Jahangir Hossain, and Dong In Kim. Uplink Vs. Downlink NOMA in Cellular Networks: Challenges and Research Directions. *IEEE Vehicular Technology Conference*, 2017-June, 2017.

[9] Zhiguo Ding, Xianfu Lei, George K. Karagiannidis, Robert Schober, Jinhong Yuan, and Vijay K. Bhargava. A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends. *IEEE Journal on Selected Areas in Communications*, 35(10):2181–2195, 2017.

[10] Linglong Dai, Bichai Wang, Yifei Yuan, Shuangfeng Han, Chih-lin I, and Zhaocheng Wang. Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. *IEEE Communications Magazine*, 53(9):74–81, 2015.

[11] Md Shipon Ali, Hina Tabassum, and Ekram Hossain. Dynamic User Clustering and Power Allocation for Uplink and Downlink Non-Orthogonal Multiple Access (NOMA) Systems. *IEEE Access*, 4:6325–6343, 2016.

[12] Mostafa Beg Mohammadi, Rahmat-Allah Hooshmand, and Fariborz Haghighatdar Fesharaki. A New Approach for Optimal Placement of PMUs and Their Required Communication Infrastructure in Order to Minimize the Cost of the WAMS. *IEEE Transactions on Smart Grid*, 7(1):84–93, 2016.

[13] Jia Tang and Xi Zhang. Cross-layer modeling for quality of service guarantees over wireless links. *Wireless Communications IEEE Transactions on*, 6(12):4504–4512, 2007.

[14] Dapeng Wu and Rohit Negi. Effective capacity: A wireless link model for support of quality of service. *IEEE Transactions on Wireless Communications*, 2(4):630–643, 2003.

[15] Mouktar Bello, Arsenia Chorti, Inbar Fijalkow, Wenjuan Yu, and Leila Musavian. Asymptotic Performance Analysis of NOMA Uplink Networks Under Statistical QoS Delay Constraints. *IEEE Open Journal of the Communications Society*, 1:1691–1706, 2020.

[16] Leila Musavian and Qiang Ni. Effective capacity maximization with statistical delay and effective energy efficiency requirements. *IEEE Transactions on Wireless Communications*, 14(7):3824–3835, 2015.

[17] Zheng Yang, Zhiguo Ding, Pingzhi Fan, and Naofal Al-Dhahir. A General Power Allocation Scheme to Guarantee Quality of Service in Downlink and Uplink NOMA Systems. *IEEE Transactions on Wireless Communications*, 15(11):7244–7257, 2016.

[18] Hong-Chuan Yang and Mohamed-Slim Alouini. *Order statistics in wireless communications: diversity, adaptation, and scheduling in MIMO and OFDM systems*. Cambridge University Press, 2011.

[19] Mahrukh Liaqat, Kamarul Ariffin Noordin, Tarik Abdul Latef, and Kaharudin Dimyati. Power-domain non orthogonal multiple access (PD-NOMA) in cooperative networks: an overview. *Wireless Networks*, 26(1):181–203, 2020.

[20] Wenjuan Yu, Leila Musavian, and Qiang Ni. Link-Layer Capacity of NOMA under Statistical Delay QoS Guarantees. *IEEE Transactions on Communications*, 66(10):4907–4922, 2018.

[21] Margaret Wright. The interior-point revolution in optimization: history, recent developments, and lasting consequences. *Bulletin of the American mathematical society*, 42(01):39–56, 2005.

[22] Muhammad Basit Shahab, Mohammad Irfan, Md Fazlul Kader, and Soo Young Shin. User pairing schemes for capacity maximization in non-orthogonal multiple access systems. *Wireless Communications and Mobile Computing*, 16(17):2884–2894, 2016.

[23] Fahri Wisnu Murti and Soo Young Shin. User pairing schemes based on channel quality indicator for uplink non-orthogonal multiple access. *International Conference on Ubiquitous and Future Networks, ICUFN*, pages 225–230, 2017.

[24] IEEE Guide for the Interoperability of Energy Storage Systems Integrated with the Electric Power Infrastructure. *IEEE Std 2030.2-2015*, pages 1–138, 2015.

[25] IEEE Standard for Synchrophasor Data Transfer for Power Systems. *IEEE Std C37.118.2-2011 (Revision of IEEE Std C37.118-2005)*, pages 1–53, 2011.