Statistical Power Grid Observability under Finite Blocklength

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*—We study the stochastic observability of the power grid system under communication constraints in the finite blocklength regime. Compared to the study under the assumption of infinite blocklength, we introduce two new elements: probability of decoding error and transmission delay. An optimization problem to maximize the observability of the smart grid over all possible bandwidth allocation is proposed, incorporating these two new elements. To solve the optimization problem, for a given bandwidth allocation, we first solve parallel subproblems, one for each synchronous phasor measurement unit (PMU), using alternating optimization, to find the optimal QoS exponent, transmission delay and probability of decoding error for each PMU. Then, simulated annealing method is used to find the optimal bandwidth allocation among PMUs. Numerical results verify that the assumption of infinite blocklength is indeed too optimistic and instead, finite blocklength should be studied. Large bandwidth saving gains of the proposed scheme are demonstrated compared to the equal bandwidth allocation scheme.**

***Index Terms*—observability, finite blocklength, decoding error, smart grid, effective capacity**

I. Introduction

In order to ensure the stability of the power grid, it is necessary to timely and accurately monitor the changes of the system states of the smart grid. Contributing to the evolution of this trend is the deployment of synchronous phasor measurement units (PMUs), which can accurately measure voltage and current phasors on the buses at high sampling rates. The goal of deploying a large number of PMU devices in a smart grid system is to be able to detect and cancel grid disturbances in real time. Therefore, a communication infrastructure is required to complete the transmission of information, which needs to meet the strict service requirements of PMU devices, namely, extremely high reliability and ultra-low (millisecond) latency.

The emergence of 5G communication network greatly promotes the development of distributed information collection and real-time information transmission and processing services required by smart grid systems. Conventionally, studies on the communication network and the smart grid system is done separately, for example [1], [2]. However, to better serve the needs of the smart grid system, joint study of the communication network and the power network is needed. In [3], a unified cyber-physical system model of the power network and the communication network is established, where the bandwidth allocation to the PMUs are studied so that the observability of the smart grid system is maximized. The definition of observability is extended to take into account the communication constraints, and effective capacity theory is used to carry out a cross-layer statistical analysis of the QoS delay requirements of the communication system. Two assumptions were made in [3]: 1) the dominant communication delay is the queueing delay, which is due to the fading characteristics of the wireless channel; and 2) the problem is under the infinite blocklength regime, and as a result, when applying effective capacity theory, the Shannon capacity formula is used and communication errors were not considered.

Due to the stringent delay requirements for the collected data of the PMU devices, for example, 10 milliseconds, and the amount of bandwidth allocated to each PMU device, around 20-30 kHz [3], the blocklength of the transmission is no more than 300. Thus, the assumption that the communication between the PMU and the base station (BS) is under the infinite blocklength regime with no communication error is not accurate. Hence, in this paper, we consider the observability of the smart grid under communication constraints in the finite blocklengh regime, in hope of providing a more accurate analysis and solutions that are more consistent with practical settings.

In this paper, the statistical power grid observability under communication constraints in the finite blocklength regime is studied. The major novelty and contributions are the following:

1) Finite blocklength theory [4] is used for the communication between the PMUs and the BS. As a result, communication error is incorporated into the definition of observability.
2) In terms of communication delay, we consider both the queueing delay as in [3], and the transmission delay, which dictates the blocklength of the transmission. With the requirement of total delay fixed, we optimize the amount of transmission delay and queueing delay of each PMU to maximize the observability of the smart grid.
3) The closed-form approximate expression of the effective capacity of the PMU under finite blocklength regime is

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).

derived. While previous papers that study the effective capacity under finite blocklength regime considers the use of the Automatic Repeat reQuest (ARQ) mechanism to get rid of decoding error [5]–[7], in the joint study of the communication network and power network of this paper, we exploit the redundancy of the PMU placement of the smart grid to overcome the effect of decoding error, thus avoiding employing ARQ and causing additional delay.

4) An optimization problem for maximizing the smart grid observability is proposed under the finite blocklength regime. For a fixed bandwidth allocation, we identify the optimal decoding error of each PMU as a function of its QoS exponent and transmission delay. Then, alternating optimization is proposed for finding the optimal QoS exponent and transmission delay, and this can be done independently for each PMU. Finally, the simulated annealing algorithm is employed to find the optimal bandwidth allocation among PMUs.
5) Numerical results show that there is a large gap between the performance of the system under the infinite blocklength assumption and the finite blocklength assumption. This validates the need to study the problem under the more accurate model of the finite blocklength. Also, the proposed bandwidth allocation scheme outperforms that of the equal bandwidth allocation scheme by a significant amount, which indicates that judicious bandwidth allocation is necessary for improving the observability of the smart grid and saving bandwidth under the finite blocklength regime.

II. System Model

A. *Observability of the smart grid under the finite blocklength regime*

For a smart grid system with N buses and a fixed topology, the connection matrix $\mathbf{L}$ is a $N \times N$-dimensional matrix, and its element values are as follows:

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

for $i, j = 1, 2, \cdots, N$.

Assume 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.,

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

When communication constraints are not considered, i.e., the channel between each PMU and the control center is considered an ideal channel, the smart grid observability vector can be expressed as:

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

If Bus i satisfies $b_i \geq 1$, then Bus i is observable, otherwise, $b_i = 0$ and Bus i is unobservable. The smart grid is observable if every bus is observable. Hence, 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.

In this paper, we consider the observability of the power grid under communication constraints under the finite blocklength regime. In the finite blocklength regime, communication constraints means both communication delay and communication error. In terms of communication delay, we consider both transmission delay and queueing delay, where queuing delay is caused by fading channels in wireless communication, and will be characterized via effective capacity theory, similar to [3]. Transmission delay is linked to the block length and thus also needs to be considered. More specifically, for a block length m transmitted over bandwidth B, the transmission delay is $\frac{m}{B}$ [8]. Furthermore, unlike the case of infinite blocklength in Shannon theory, where the probability of decoding error is negligible, the influence of decoding error needs to be considered in the finite blocklength regime.

Hence, we modify the observability vector in (1) to incorporate both communication delay and communication error as follows. To maintain real-time performance, information measurements for each PMU device will be valid within the delay threshold D_{max}. If the communication delay is longer than D_{max}, then these measurements are outdated and useless. Assuming that the transmission delay of PMU_k is D_k^t, then, the queueing delay can be no longer than D_k^q, where D_k^q is defined as

$$D_k^q \triangleq D_{\text{max}} - D_k^t \tag{2}$$

However, in most fading channel environments, it is not feasible to provide a definite queuing delay bound for the communication system [9]. Therefore, we consider a probability p_k to provide a statistical guarantee that the random queueing delay of PMU_k, denoted by D_k, is less than D_k^q, that is:

$$p_k = \text{Pr}\left\{D_k \leq D_k^q\right\} \tag{3}$$

Furthermore, even if the information of PMU_k is transmitted to the control center within the delay threshold D_{max}, when a decoding error occurs, the information is also useless. Therefore, the transmitted information is valid if and only if the delay constraint is satisfied *and* no decoding error occurs. This happens with probability

$$p_{\epsilon_k} \triangleq p_k(1 - \epsilon_k), \tag{4}$$

where ϵ_k denotes the probability of decoding error of the information sent by PMU_k, and p_k is given by (3).

Hence, in contrast to (1), the random observability vector under the communication constraints of a non-ideal channel is given as

$$\widetilde{\mathbf{b}} = \mathbf{L}\mathbf{\Lambda}_Q\mathbf{x}, \tag{5}$$

where $\mathbf{\Lambda}_Q$ is an $N \times N$-dimensional diagonal matrix, whose i-th diagonal element Q_i is a random variable if a PMU is installed on Bus i, otherwise, the i-th diagonal element is 0. $Q_i = 1$ means that the PMU installed on Bus i can transmit its information to the control center under the constraint of communication delay and without error, that is, if PMU_k is installed on Bus i, then

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

where P_i is defined as

$$P_i \triangleq p_{\epsilon_k}. \tag{6}$$

The expected observability vector of the smart grid can be expressed as:

$$\bar{\mathbf{b}} = \mathbb{E}\left[\tilde{\mathbf{b}}\right] = \mathbf{L}\mathbf{\Lambda}_P \mathbf{x} \tag{7}$$

where $\mathbf{\Lambda}_P$ is an $N \times N$-dimensional diagonal matrix, whose i-th diagonal element is P_i if a PMU is installed on Bus i.

Next, we characterize p_k, i.e., the probability that the information of PMU_k is delivered to the control center within queueing delay constraint D_k^q, via effective capacity theory [10]. More specifically, for PMU k, its effective capacity can be expressed as [10]:

$$\text{EC}_k = -\frac{1}{\theta_k T} \ln \mathbb{E}\left\{e^{-\theta_k T \tilde{r}_k}\right\} \tag{8}$$

where $\theta_k > 0$ is called the QoS exponent, T is the coherence time of the channel, and $\tilde{r}_k$ is the instantaneous transmission rate of PMU_k to the BS. Under the finite blocklength regime [4], we have

$$\begin{aligned}\tilde{r}_k =& B_k \log_2(1+\rho_k) \\ &- B_k \sqrt{\frac{1}{m}\left(1 - \frac{1}{(\rho_k+1)^2}\right)} Q^{-1}(\epsilon_k) \log_2 e\end{aligned} \tag{9}$$

where B_k is bandwidth occupied by PMU_k, ϵ_k is the probability of decoding error due to finite blocklength, and ρ_k is the instantaneous signal-to-noise ratio (SNR) of PMU_k. We assume that the channel fading is Rayleigh distributed, then the instantaneous SNR ρ_k obeys an exponential distribution with a mean λ_k, termed the average SNR of PMU_k. The probability distribution function (PDF) and the cumulative distribution function (CDF) of ρ_k are expressed as follows:

$$f(x) = \frac{1}{\lambda_k} e^{-\frac{x}{\lambda_k}}, \quad F(x) = 1 - e^{-\frac{x}{\lambda_k}}.$$

Note that $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} dt$ is the Gaussian Q function, whose inverse function is denoted as $Q^{-1}(\cdot)$. In (9), m is the finite block length which is equal to $\lfloor B_k D_k^t \rfloor \approx B_k D_k^t$ [8], hence, we have

$$\begin{aligned}\tilde{r}_k \approx& B_k \log_2(1+\rho_k) \\ &- \sqrt{\frac{B_k}{D_k^t}\left(1 - \frac{1}{(\rho_k+1)^2}\right)} Q^{-1}(\epsilon_k) \log_2 e\end{aligned} \tag{10}$$

Looking at (10), we see that the first term on the right-hand side (RHS) of this expression is the Shannon capacity, and the second term is the adjustment of the transmission rate due to the existence of the bit error rate ϵ_k. We require that the observability of the smart grid is high, which means that we are interested in the regime where $0 < \epsilon_k < \frac{1}{2}$, i.e., $Q^{-1}(\epsilon_k) > 0$. Therefore, the instantaneous transmission rate of the signal under the finite blocklength regime is smaller than the Shannon capacity.

The parameter θ_k is related to the probability of successful transmission within the delay constraint as [10]

$$p_k = \Pr\{D_k \le D_k^q\} = 1 - e^{-\theta_k \text{EC}_k D_k^q} \tag{11}$$

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.,

$$\text{EC}_k \ge R_k^{\text{th}}. \tag{12}$$

B. Problem Formulation

We are interested in the problem of using orthogonal frequency division multiple access (OFDMA) and allocate communication bandwidth to different PMU devices, maximizing the observability of the smart grid over all possible bandwidth allocation strategies under the finite blocklength regime. Thus, the problem is formulated as

$$\max_{\substack{\{B_1,\cdots,B_K\},\{\epsilon_1,\cdots,\epsilon_K\},\\ \{\theta_1,\cdots,\theta_K\},\{D_1^t,\cdots,D_K^t\}}} \quad \text{Obs}(p_{\epsilon_1}, p_{\epsilon_2}, \cdots, p_{\epsilon_K}) \tag{13}$$

$$\text{s.t.} \quad \text{EC}_k = -\frac{1}{\theta_k T} \ln \mathbb{E}\left\{e^{-\theta_k T \tilde{r}_k}\right\}, \quad k = 1, \cdots, K, \tag{14}$$

$$p_{\epsilon_k} = (1-\epsilon_k)(1 - e^{-\theta_k \text{EC}_k (D_{\max} - D_k^t)}), k = 1, \cdots, K, \tag{15}$$

$$\text{EC}_k \ge R_k^{\text{th}}, \quad k = 1, \cdots, K, \tag{16}$$

$$\sum_{k=1}^K B_k \le B^{\text{total}} \tag{17}$$

$$0 < D_k^t < \min\{T, D_{\max}\}, \qquad \forall k = 1, \cdots, K \tag{18}$$

$$B_k \ge 0, \qquad \forall k = 1, \cdots, K \tag{19}$$

$$0 < \epsilon_k < \frac{1}{2}, \qquad \forall k = 1, \cdots, K \tag{20}$$

$$\theta_k > 0, \qquad \forall k = 1, \cdots, K \tag{21}$$

where (13) is the optimization objective function of the observability of the smart grid, which is a function of the probabilities p_{ϵ_k}, $k = 1, \cdots, K$, which is defined in (4), (14) follows from (8) with $\tilde{r}_k$ given in (10), (15) follows from (11), (2) and the definition of p_{ϵ_k} in (4), (16) follows from (12), the constraint (17) means that the total bandwidth consumed by all PMU devices should not exceed the given total bandwidth B^{total}, the constraint (18) indicates that the maximum delay $D_{\max}$ is split between the transmission delay D_k^t and the queueing delay $D_k^q = D_{\max} - D_k^t$, which is positive. Furthermore, since the transmission should be finished within the coherence time of the channel, we require $D_k^t < T$, and the constraint in (20) indicate that the decoding error ϵ_k is strictly larger than zero, since we are operating in the finite blocklength regime and zero error is not possible, and ϵ_k is no larger than $\frac{1}{2}$, otherwise the observability of the smart grid is too low to be of practical interest.

For power grid observability under communication constraints, we will consider the following three metrics proposed in [3]:

1) Observability redundancy (OR), which is defined as $\text{Obs}_1(p_{\epsilon_1}, p_{\epsilon_2}, \cdots, p_{\epsilon_K}) \triangleq \mathbf{1}_N^T \mathbf{L} \mathbf{\Lambda}_P \mathbf{x} - N$.This metric evaluates thte total smart grid system observability redundancy via the expected observability vector defined in (7).
2) Observability sensitivity (OS), which is defined as $\text{Obs}_2(p_{\epsilon_1}, p_{\epsilon_2}, \cdots, p_{\epsilon_K}) \triangleq \min_n \bar{b}_n$, where $\bar{b}_n$ is the n-th element of the expected power grid observability vector $\bar{\mathbf{b}}$ as defined in (7). The bus with the smallest expected observability value is generally the bus that is least likely to be observed, so it can be regarded as the performance bottleneck of the observability of the whole smart grid system, and maximizing the bottleneck will increase the observability performance of the whole grid.
3) Observability probability (OP), which is defined as $\text{Obs}_3(p_{\epsilon_1}, p_{\epsilon_2}, \cdots, p_{\epsilon_K}) \triangleq \Pr\left[\widetilde{\mathbf{b}} \geq \lambda\right]$. This is the probability that the observability random vector is larger than a threshold vector λ. Maximizing this probability will increase the observability of the power grid.

III. Solving the Optimization Problem in (13)

The optimization problem defined in (13) is complex. Firstly, expectation exists in the effective capacity formula in (14), which is not in closed form. Secondly, the optimization problem requires the joint optimization of many optimization parameters, including the QoS exponent of each PMU, the probability of decoding error of each PMU's information, the transmission delay of each PMU, and the bandwidth allocation to each PMU. In this section, we will discuss in detail how to solve this problem step by step.

A. Approximate closed-form expression of effective capacity under the finite blocklength regime

In order to solve the optimization problem in (13), we first derive the approximate closed-form expression of the effective capacity of the PMU device under the finite blocklength regime, namely the RHS expression of (14).

More specifically, we have the following Lemma.

Lemma 1 *The approximate closed-form expression of the effective capacity of PMU$_k$, with bandwidth allocation B_k, average SNR λ_k, QoS exponent θ_k, transmission delay D_k^t, and probability of decoding error ϵ_k, is given as*

$$EC_k \approx EC_k^\infty - \sqrt{\frac{B_k}{D_k^t}} Q^{-1}(\epsilon_k) \log_2 e \tag{22}$$

where EC_k^∞ is the effective capacity for the case of infinite blocklenth, i.e., using Shannon capacity [3], i.e.,

$$EC_k^\infty = -\frac{1}{\theta_k T} \ln\left[\frac{1}{\lambda_k} U\left(1, 2+w_k, \frac{1}{\lambda_k}\right)\right]. \tag{23}$$

where $w_k = -\frac{\theta_k T B_k}{\ln 2}$, and $U(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function, that is:

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

and $\Gamma(\cdot)$ is the gamma function.

The proof for Lemma 1 is provided in Appendix A. Note that the first term on the RHS of (22) is a function of $(B_k, \lambda_k, \theta_k)$ only, while the second term on the RHS of (22) is a function of (B_k, D_k^t, ϵ_k) only.

B. Solving the subproblem for a given $\{B_1, \cdots, B_K\}$

When $\{B_1, \cdots, B_K\}$ is given, the problem in (13) becomes the following K parallel subproblems, which can be solved independently for each $k = 1, \cdots, K$:

$$\max_{\epsilon_k, \theta_k, D_k^t} p_{\epsilon_k} = (1-\epsilon_k) \cdot \left(1 - e^{-\theta_k \left(EC_k^\infty - \sqrt{\frac{B_k}{D_k^t}} Q^{-1}(\epsilon_k) \log_2 e\right)(D_{\max} - D_k^t)}\right) \tag{24}$$

$$\text{s.t.} \quad EC_k^\infty - \sqrt{\frac{B_k}{D_k^t}} Q^{-1}(\epsilon_k) \log_2 e \geq R_k^{\text{th}}, \tag{25}$$

$$0 < D_k^t < \min\{T, D_{\max}\}, \tag{26}$$

$$0 < \epsilon_k < \frac{1}{2}, \tag{27}$$

$$\theta_k > 0, \tag{28}$$

where we have used Lemma 1 to replace EC_k with its approximate expression. Since all three observability functions OR, OS and OP increases with each p_{ϵ_k}, $k = 1, \cdots, K$, maximizing the observability function is the same as maximize p_{ϵ_k} for each k, when $\{B_1, \cdots, B_K\}$ is given.

To solve the problem in (24) for each $k = 1, \cdots, K$, we first note that in order to maximize the cost function (24), the constraint in (25) must be satisfied with equality. This is because given $(B_k, \lambda_k, \epsilon_k, D_k^t)$, the approximate expression of EC_k, i.e., $EC_k^\infty - \sqrt{\frac{B_k}{D_k^t}} Q^{-1}(\epsilon_k) \log_2 e$, is a decreasing function of θ_k [3]. Hence, if (25) is a strict inequality, we may always increase θ_k, and this enlarges the cost function of (24), because for a given $(B_k, \lambda_k, \epsilon_k, D_k^t)$, $-\theta_k \left(EC_k^\infty - \sqrt{\frac{B_k}{D_k^t}} Q^{-1}(\epsilon_k) \log_2 e\right)$ is a decreasing function of θ_k [3]. Hence, to achieve the maximum cost function, (25) should be an equality.

Thus, the problem in (24) is equivalent to

$$\max_{\theta_k, D_k^t} \left(1 - e^{-\theta_k R_k^{\text{th}} (D_{\max} - D_k^t)}\right) \cdot \left[1 - Q\left(\sqrt{\frac{D_k^t}{B_k}} \ln 2 \left(EC_k^\infty - R_k^{\text{th}}\right)\right)\right] \tag{29}$$

$$\text{s.t.} \quad 0 < D_k^t < \min\{T, D_{\max}\}, \tag{30}$$

$$0 < \theta_k < \bar{\theta}_k. \tag{31}$$

where we have used the fact that (25) should be an equality which leads to ϵ_k satisfying

$$\epsilon_k = Q\left(\sqrt{\frac{D_k^t}{B_k}\ln 2\left(EC_k^{\infty} - R_k^{\text{th}}\right)}\right).$$

Note that we are operating in the regime of $0 < \epsilon_k < \frac{1}{2}$, which means we are interested in the case where $EC_k^{\infty} > R_k^{\text{th}}$. This is ensured by $\theta < \bar{\theta}_k$ in (31), where $\bar{\theta}_k$ denotes the root of the following equation for the fixed B_k,

$$EC_k^{\infty} = R_k^{\text{th}}. \tag{32}$$

Any $\theta < \bar{\theta}_k$ will satisfy $EC_k^{\infty} > R_k^{\text{th}}$, because given B_k, the approximate expression of EC_k, i.e., $EC_k^{\infty} - \sqrt{\frac{B_k}{D_k^t}}Q^{-1}(\epsilon_k)\log_2 e$, is an decreasing function of θ_k [3].

To solve the problem in (29), we find via extensive numerical experiments that when given D_k^t, the cost function is a unimodal function θ_k, and when given θ_k, the cost function is a unimodal function of D_k^t. Thus, we propose to use alternating optimization to solve (29) and the details are given in Algorithm 1, where the golden section search method is used to find the maximum of a unimodal function.

Algorithm 1: Solving the optimization problem in (24) for each $k = 1, \cdots, K$.

Input: $B_k \leftarrow$ Bandwidth of PMU_k, $k = 1, \cdots, K$

initialization: $D_k^{t*} = \frac{\min\{T, D_{\max}\}}{2}$
repeat
  For $D_k^t = D_k^{t*}$, use the golden section search method to find the θ_k that maximizes the cost function in (29), with the corresponding maximum value denoted as f_1, and the optimal θ_k denoted as θ_k^*.
  For $\theta_k = \theta_k^*$, use the golden section search method to find the D_k^t that maximizes the cost function in (29), with the corresponding maximum value denoted as f_2, and the optimal D_k^t denoted as D_k^{t*}.
until f_1 and f_2 are close enough
Calculate $EC_k^{\infty*}$ according to (23) with $\theta_k = \theta_k^*$

Output: D_k^{*t}, θ_k^*, $\epsilon_k^* = Q\left(\sqrt{\frac{D_k^{t*}}{B_k}\ln 2\left(EC_k^{\infty*} - R_k^{\text{th}}\right)}\right)$

C. Solving the optimal problem in (13)

In this section, we mainly discuss how to allocate the communication bandwidth resources B^{total} among the PMU devices, thus, finding the optimal $\{B_1, \cdots, B_K\}$.

First, we deal with the constraints of the optimization problem by the interior penalty function method [11]. More specifically, the new objective function, denoted as $h(B_1, \cdots, B_K, \epsilon_1, \cdots, \epsilon_K, D_1^t, \cdots, D_K^t, \theta_1, \cdots, \theta_K)$, obtained by adding the constraint to the original objective function, is

$$\max_{B_1,\cdots,B_K} h(B_1, \cdots, B_K, \epsilon_1, \cdots, \epsilon_K, D_1^t, \cdots, D_K^t, \theta_1, \cdots, \theta_K)$$
$$\triangleq \text{Obs}(p_{\epsilon_1}, p_{\epsilon_2}, \cdots, p_{\epsilon_K}) + m_j\left[\sum_{k=1}^{K}(\max(-B_k, 0)^2) + (\sum_{k=1}^{K} B_k - B^{\text{total}})^2\right] \tag{33}$$

where the parameter m_j is called the penalty factor, which will continue to increase with iteration, whose index is given by j. The simulated annealing algorithm is proposed to solve the problem in (13), whose details are given by Algorithm 2.

Algorithm 2: Solving the optimal problem in (13)

Input: $N \leftarrow$ Number of buses
$K \leftarrow$ Number of PMUs
$B^{\text{total}} \leftarrow$ Constraint on total bandwidth
$\mathbf{L} \leftarrow$ Connection matrix
$\mathbf{x} \leftarrow$ PMU installation vector
$R_k^{\text{th}} \leftarrow$ Minimum data rate requirement of PMU k
$D_{\text{max}} \leftarrow$ Delay constraint
$T \leftarrow$ Coherence time
$\lambda_k \leftarrow$ Average SNR of PMU k
$T_{\text{initial}} \leftarrow$ Initial temperature
$T_{\text{final}} \leftarrow$ Stop temperature
$\gamma \leftarrow$ Attenuation coefficient

initialization:
$B_k = \frac{B^{\text{total}}}{K}, k = 1, \cdots, K$
Calculate the optimal $(\epsilon_k^*, D_k^{t*}, \theta_k^*)$ using Algorithm 1 for $k = 1, \cdots, K$
Set $H \triangleq h(B_1, \cdots, B_K, \epsilon_1^*, \cdots, \epsilon_K^*, D_1^{t*}, \cdots, D_K^{t*}, \theta_1^*, \cdots, \theta_K^*)$
$t = T_{\text{initial}}$
repeat
  $m = M \leftarrow$ Markov chain length
  repeat
    Generate a random neighbor $\{B_1', \cdots, B_K'\}$
    Calculate the optimal $(\epsilon_k^{*'}, D_k^{t*'}, \theta_k^{*'})$ using Algorithm 1 for $k = 1, \cdots, K$
    Set $H' \triangleq h(B_1', \cdots, B_K', \epsilon_1^{*'}, \cdots, \epsilon_K^{*'}, D_1^{t*'}, \cdots, D_K^{t*'}, \theta_1^{*'}, \cdots, \theta_K^{*'})$
    $\Delta E = H' - H$
    if $\Delta E > 0$ **then**
      $B_k = B_k', \quad \forall k = 1, 2, \cdots, K$, $H = H'$
    else
      with probability $e^{\frac{\Delta E}{t}}$, take $B_k = B_k'$, $\forall k = 1, 2, \cdots, K$
    end if
    $m = m - 1$
  until $m = 0$
  Update temperature $t = \gamma t$.
until $t \leq T_{\text{final}}$

Output: $B_1, \cdots, B_K$ and the corresponding optimal $\{\epsilon_1^*, \cdots, \epsilon_K^*\}$, $\{D_1^{t*}, \cdots, D_K^{t*}\}$ and $\{\theta_1^*, \cdots, \theta_K^*\}$

More specifically, to initialize, we assume that the K PMUs share the total bandwidth equally, and then find the corresponding optimal $(\epsilon_k^*, D_k^{t*}, \theta_k^*)$ using Algorithm 1 for

$k = 1, \cdots, K$. We calculate the value of the new objective function, and save it as H, thus completing the initialization process of the algorithm, i.e., Line 1. For each temperature t, run the simulated annealing algorithm M times, each of which does the following: firstly, a random new neighbor $\{B'_1, \cdots, B'_K\}$ is generated based on the solution $\{B_1, \cdots, B_K\}$, and then find the corresponding optimal $(\epsilon_k^{*'}, D_k^{t*'}, \theta_k^{*'})$ using Algorithm 1 for $k = 1, \cdots, K$. Based on this solution, the new objective function value is calculated, denoted as H'. If H' is better than H, accept the neighborhood bandwidth allocation scheme $\{B'_1, \cdots, B'_K\}$ as the optimal scheme and update H to be H', as indicated by Line 10. If H' is no better than H, accept the neighborhood strategy 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 neighborhood strategy 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 optimal QoS exponents, probabilities of decoding error and transmission delays obtained via Algorithm 1, are the output of Algorithm 2.

IV. Numerical Results

In this section, we simulate and verify the theories and algorithms proposed in this paper, and compare them with the corresponding results assuming infinite blocklength [3]. We further compare our result with the *equal bandwidth allocation* scheme: for each $k = 1, \cdots, K$, $B_k = \frac{B^{\text{total}}}{K}$, and for this equal bandwidth allocation, we calculate $\bar{\theta}_k$ as the root of (32) and set $\theta_k = \frac{\bar{\theta}_k}{2}$. Furthermore, set $D_k^t = \frac{\min\{T, D_{\max}\}}{2}$.

All simulations in this paper are based on the IEEE-14 bus power system standard test case, which has been widely used as a standard test case to verify power system performance [12], [13]. In this test case, there are 14 buses in the system, namely $N = 14$, and there are 9 PMUs installed in the system, namely $K = 9$. Assume that all PMU device have a source data rate of $R_k^{\text{th}} = 60$Kbps, $k = 1, \cdots, K$. The maximum delay allowed by this information is 10ms, namely $D_{\max} = 0.001$. The installation vector of the PMU on the smart grid system is taken as [14]

$$\mathbf{x} = [0\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 0\ \ 1\ \ 0\ \ 1\ \ 0]^T. \tag{34}$$

Assuming that the channel environment experienced by all PMUs is a Rayleigh fading channel, and the average SNR of the device PMU_k is shown in Table I. Based on the above parameters, the results obtained are as follows.

Fig. 1 illustrates the OR observability achieved by varying the total bandwidth B^{total}. We plot the performance of three cases: the red dotted-stared line corresponds to the performance under the assumption of infinite blocklength [3], the blue dash-dot-triangle line corresponds to the case of finite blocklength with equal bandwidth allocation among PMUs, and the cyan solid-circled line corresponds to the proposed scheme under the finite blocklength regime, i.e., Algorithm 2. As can be seen, the gap between the performance under the infinite blocklength assumption and that of the proposed scheme under finite blocklength regime is quite large. Since the infinite blocklength assumption does not hold in practice, the performance under the infinite blocklength assumption is too optimistic and indeed, finite blocklength regime should be studied, as is done in this paper.

TABLE I
Installation and Average SNR of each PMU

PMU index	1	2	3	4	
Installed on Bus	2	4	5	6	
Average SNR λ_k	18.25	32.00	29.25	15.50	
PMU index	5	6	7	8	9
Installed on Bus	7	8	9	11	13
Average SNR λ_k	10.00	21.00	23.75	26.50	12.75

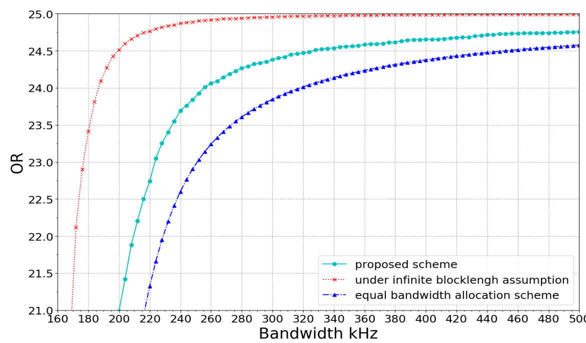

Fig. 1. OR observability vs. total bandwidth B^{total}

The highest achievable OR observability is 25 and this can only be reached with infinite bandwidth. To achieve the OR observability of 24, the equal bandwidth allocation scheme requires a total bandwidth of 320kHz, while the proposed algorithm only requires a total bandwidth of 254kHz, which amounts to a bandwidth saving of 20.6%. To achieve the OR observability of $25 \cdot 99\% = 24.75$, the equal bandwidth allocation scheme requires a total bandwidth of 706kHz, while the proposed algorithm only requires a total bandwidth of 440kHz, which amounts to a bandwidth saving of 37.7%.

Similar performance gains may be observed when OS observability is considered, as shown in Fig. 2. More specifically, with infinite bandwidth, the highest achievable OS is 2. To achieve an OS observability of 1.9, the equal bandwidth allocation scheme requires a total bandwidth of 300kHz, while the proposed algorithm only requires a total bandwidth of 234kHz, which amounts to a bandwidth saving of 22%. To achieve an OS observability of $2 \cdot 99\% = 1.98$, the equal bandwidth allocation scheme requires a total bandwidth of 690kHz, while the proposed algorithm only requires a total bandwidth of 424kHz, which amounts to a bandwidth saving of 38.6%.

OP observability characterizes the probability that the ob-

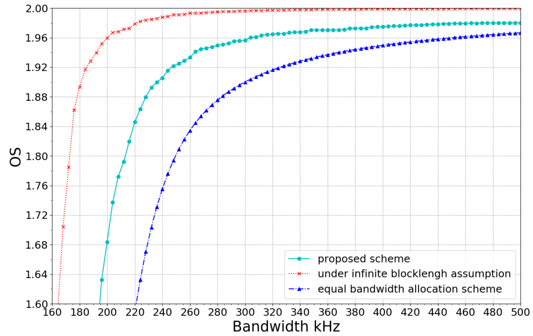

Fig. 2. OS observability

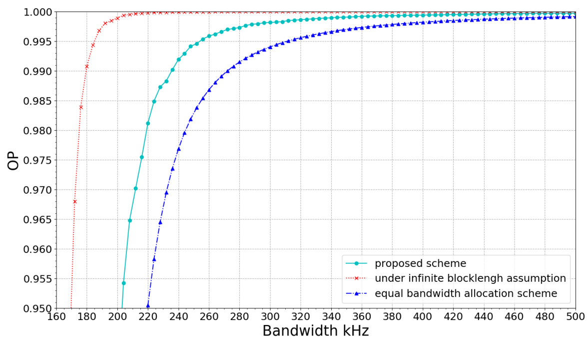

Fig. 3. OP observability

servability 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 Table II which provides the total bandwidth needed for achieving the OP observability of 99%, 99.9% and 99.99%, respectively. As can be seen, compared to the equal

TABLE II
BANDWIDTH FOR DIFFERENT OP REQUIREMENTS

OP threshold	proposed scheme	equal bandwidth allocation scheme	infinite blocklength assumption
99%	234kHz	272kHz	180kHz
99.9%	338kHz	468kHz	201kHz
99.99%	725kHz	1060kHz	232kHz

bandwidth allocation scheme, to reach the OP observability of 99%, 99.9% and 99.99%, the proposed scheme achieves a bandwidth saving of 14%, 27.8% and 31.6%, respectively. In Table II, again we see that under the infinite blocklength assumption, the total bandwidth required is much smaller than what is actually needed under the finite blocklength regime, and this gap increases as the a higher OP observability is needed.

We observe that for all OR, OS and OP observability, the bandwidth savings of the proposed scheme over the equal bandwidth allocation scheme is large, and furthermore, the savings are more when the observability requirement is more stringent. This justifies that judicious allocation of bandwidth is important in the observability of the smart grid under communication constraints in the finite blocklength regime. We also notice that the amount of bandwidth required increases quite drastically when a higher requirement for observability is needed.

V. CONCLUSION

In this paper, we consider the finite blocklength regime and analyze the transmission process of the PMUs, so as to ensure the observability performance of the smart grid system. Compared to the problem under the infinite blocklength assumption, we introduce two new elements: probability of decoding error and transmission delay. According to the three observability metrics of OR, OS and OP, an optimization problem incorporating these two elements is proposed. Alternating optimization and simulated annealing methods are used to solve the optimization problem. Numerical results verify that the assumption of infinite blocklength is indeed too optimistic and instead, finite blocklength should be studied. Large bandwidth saving gains of the proposed scheme are demonstrated compared to the equal bandwidth allocation scheme.

APPENDIX A
PROOF OF LEMMA 1

According to (9), we have that $e^{-\theta_k T\tilde{r}_k}$ can be expressed as follows:

$$
\begin{aligned}
& e^{-\theta_k T\tilde{r}_k} \\
&= e^{-\theta_k TB_k \log_2(1+\rho_k)+\theta_k TB_k\sqrt{\frac{1}{m}\left(1-\frac{1}{(\rho_k+1)^2}\right)}Q^{-1}(\epsilon_k)\log_2 e} \\
&= e^{-\theta_k TB_k \log_2(1+\rho_k)} \times e^{\theta_k TB_k\sqrt{\frac{1}{m}\left(1-\frac{1}{(\rho_k+1)^2}\right)}Q^{-1}(\epsilon_k)\log_2 e} \\
&= \underbrace{e^{-\theta_k TB_k \log_2(1+\rho_k)}}_{(a)} \times \underbrace{e^{\theta_k TB_k\sqrt{\frac{1}{m}}Q^{-1}(\epsilon_k)\log_2 e\sqrt{1-\frac{1}{(\rho_k+1)^2}}}}_{(b)}
\end{aligned} \tag{35}
$$

Let $w_k = -\frac{\theta_k TB_k}{\ln 2}$. For term (a) in (35), we have:

$$
e^{-\theta_k TB_k \log_2(1+\rho_k)} = (1+\rho_k)^{w_k} \tag{36}
$$

In the power grid system discussed in this paper, in general, the SNR of PMU is relatively large, i.e. $\rho_k \gg 1$, then [15]:

$$
\sqrt{1-\frac{1}{(\rho_k+1)^2}} \approx 1 \tag{37}
$$

Let $v_k = \theta_k B_k T\sqrt{\frac{1}{m}}Q^{-1}(\epsilon_k)\log_2 e$, for term (b) in (35), using (37), we have:

$$
\begin{aligned}
e^{\theta_k TB_k\sqrt{\frac{1}{m}}Q^{-1}(\epsilon_k)\log_2 e\sqrt{1-\frac{1}{(\rho_k+1)^2}}} &= e^{v_k\sqrt{1-\frac{1}{(\rho_k+1)^2}}} \\
&\approx e^{v_k}
\end{aligned} \tag{38}
$$

Combining (35), (36) and (38), we have

$$e^{-\theta_k T \tilde{r}_k} = (1+\rho)^{w_k} \times e^{v_k} \tag{39}$$

Hence, the effective capacity of PMU_k can be expressed as:

$$\begin{aligned} \text{EC}_k &= -\frac{1}{\theta_k T} \ln \mathbb{E}\left\{e^{-\theta_k T \tilde{r}_k}\right\} \\ &\approx -\frac{1}{\theta_k T} \ln \mathbb{E}\left\{(1+\rho_k)^{w_k} e^{v_k}\right\} \end{aligned} \tag{40}$$

Assuming that the channel is Rayleigh fading, the instantaneous SNR ρ_k obeys an exponential distribution with a mean value of λ_k. Thus, we have:

$$\begin{aligned} \text{EC}_k &\approx -\frac{1}{\theta_k T} \ln \mathbb{E}\left\{(1+\rho_k)^{w_k} e^{v_k}\right\} \\ &= -\frac{1}{\theta_k T} \ln \int_0^\infty e^{v_k} (1+x)^{w_k} \frac{1}{\lambda_k} e^{-\frac{1}{\lambda_k} x} dx \\ &= -\frac{1}{\theta_k T} \ln \left(\frac{1}{\lambda_k} e^{v_k} \int_0^\infty (1+x)^{w_k} e^{-\frac{1}{\lambda_k} x} dx \right) \\ &= -\frac{1}{\theta_k T} \ln \left(\frac{1}{\lambda_k} e^{v_k} U\left(1, 2+w_k, \frac{1}{\lambda_k}\right)\right) \end{aligned} \tag{41}$$

where $U(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function. This completes the proof of Lemma 1.

References

[1] Yu-Jia Chen, Li-Yu Cheng, and Li-Chun Wang. Prioritized resource reservation for reducing random access delay in 5G URLLC. In *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–5, 2017.

[2] Karim Amiri, Rasool Kazemzadeh, and Omid Eghbali. A novel optimal PMU location method in smart grids. In *2018 Smart Grid Conference (SGC)*, pages 1–5, 2018.

[3] 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.

[4] Yury Polyanskiy, H. Vincent Poor, and Sergio Verdu. Channel coding rate in the finite blocklength regime. *IEEE Transactions on Information Theory*, 56(5):2307–2359, 2010.

[5] Mustafa Cenk Gursoy. Throughput analysis of buffer-constrained wireless systems in the finite blocklength regime. In *2011 IEEE International Conference on Communications (ICC)*, pages 1–5, 2011.

[6] Mohammad Shehab, Endrit Dosti, Hirley Alves, and Matti Latva-aho. On the effective capacity of MTC networks in the finite blocklength regime. In *2017 European Conference on Networks and Communications (EuCNC)*, pages 1–5, 2017.

[7] Maria Cecilia Fernández Montefiore, Gustavo González, F. J. Lopez-Martinez, and Fernando Gregorio. Effective capacity of finite-blocklength NOMA and OMA with heterogeneous latency requirements. In *2021 Argentine Conference on Electronics (CAE)*, pages 78–81, 2021.

[8] Hong Ren, Cunhua Pan, Yansha Deng, Maged Elkashlan, and Arumugam Nallanathan. Joint power and blocklength optimization for URLLC in a factory automation scenario. *IEEE Transactions on Wireless Communications*, 19(3):1786–1801, 2020.

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

[10] 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.

[11] 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.

[12] Nabil H. Abbasy and Hanafy Mahmoud Ismail. A unified approach for the optimal PMU location for power system state estimation. *IEEE Transactions on Power Systems*, 24(2):806–813, 2009.

[13] 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.

[14] Xu Bei, Yeo Jun Yoon, and Ali Abur. Optimal placement and utilization of phasor measurements for state estimation. *PSERC Publication*, 2005.

[15] Muhammad Amjad, Leila Musavian, and Sonia Aissa. NOMA versus OMA in finite blocklength regime: link-layer rate performance. *IEEE Transactions on Vehicular Technology*, 69(12):16253–16257, 2020.