2021 61st IEEE Conference on Decision and Control (CDC)
December 6-9, 2022. Cancún, Mexico

Approximations for Optimal Experimental Design in Power System Parameter Estimation

Xu Du, Alexander Engelmann, Timm Faulwasser and Boris Houska

Abstract— **This paper is about computationally tractable methods for power system parameter estimation and Optimal Experiment Design (OED). The main motivation of OED is to increase the accuracy of power system parameter estimates for a given number of batches. One issue in OED, however, is that solving the OED problem for larger power grids turns out to be computationally expensive and, in many cases, computationally intractable. Therefore, the present paper proposes three numerical approximation techniques, which increase the computational tractability of OED for power systems. These approximation techniques are benchmarked on a 5-bus and a 14-bus case study.**

Keywords: **Power Systems, Parameter Estimation, Optimal Experiment Design, Admittance Estimation**

I. Introduction

The power system industry is facing a variety of challenges such as supply diversification, reducing carbon emissions, secure network access for renewable energy and electricity market pricing. Techniques based on online optimization are among the most promising approaches for addressing these challenges. These techniques range from classical Optimal Power Flow (OPF) problems [1], over optimal reactive power dispatch [2], to reactive power planning [3].

In the above approaches, the admittance of the power grid is typically treated as known. However, parameters are often unknown in practice and may even vary over time, e.g., due to temperature changes. Hence, online parameter estimation is used. Approaches based on multiple measurements snapshots via Recursive Least Squares (RLS) has been considered in [4–7]. However, RLS-based estimation is sometimes limited in accuracy.

More recently, Optimal Experimental Design (OED) was applied as an alternative to pure RLS [8–10]. Here, the main idea is to choose active/reactive power inputs of the generators such that the amount of extracted information in the estimation step is maximized. However, doing so may lead to high costs for system operation, since OED neglects the cost of power generation. Having this in mind and inspired by [11], [12] proposes an adaptive method for trading-off OED and the OPF cost. This leads to an excitation of the system, mainly its reactive power, which reduces the cost of optimal estimation substantially compared with pure OED, but still leads to an improved performance compared with RLS. A drawback of all OED approaches outlined before is their computational intractability for larger grids. The objective function in OED requires inversion of the Fisher Information Matrix (FIM), which does not scale well with the number of buses. Moreover, the nonlinear power flow equations lead to additional complexity.

In the present paper, we propose approximation techniques to improve the numerical tractability of OED for power system parameter estimation. Specifically, we develop an approach based on a Newton-type iteration for inner OED approximation and two outer approximation techniques for approximating the inversion of the Fisher matrix. All approximations avoid symbolic matrix inversion in the implementation, which is potentially costly. In terms of controlled variables, we focus on the reactive power in order to keep the system operation cost low during the estimation.

The remaining chapters are structured as follows: Section II reviews basics of AC power system modeling. Section III presents the approximation techniques outlined above. Section VI presents numerical case studies for a 5-bus and a 14-bus power grid.

Notation: For $a \in \mathbb{R}^n$ and $\mathcal{C} \subseteq \{1, ..., n\}$, $(a_i)_{i\in\mathcal{C}} \in \mathbb{R}^{|\mathcal{C}|}$ collects all components of a whose index i is in $\mathcal{C}$. Similarly, for $A \in \mathbb{R}^{n\times l}$ and $\mathcal{S} \subseteq \{1,\dots,n\} \times \{1,\dots,l\}$, $(A_{i,j})_{(i,j)\in\mathcal{S}} \in \mathbb{R}^{|\mathcal{S}|}$ denotes the concatenation of $A_{i,j}$ for all $(i,j) \in \mathcal{S}$. $\mathrm{i} = \sqrt{-1}$ denotes the imaginary unit, such that $\mathrm{Re}(z) + \mathrm{i}\cdot\mathrm{Im}(z) = z \in \mathbb{C}$, and $\hat{a}$ denotes the estimated value of a.

II. AC Power System Model

Consider a power grid defined by the triple $(\mathcal{N}, \mathcal{L}, Y)$, where $\mathcal{N} = \{1, 2, \dots N\}$ represents the set of buses, $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$ specifies the transmission lines and $Y \in \mathbb{C}^{N\times N}$ denotes the complex admittance matrix

$$Y_{k,l} \doteq \begin{cases} \sum\limits_{i\neq k} (g_{k,i} + \mathrm{i}\, b_{k,i}) & \text{if } k = l, \\ -(g_{k,l} + \mathrm{i}\, b_{k,l}) & \text{if } k \neq l. \end{cases}$$

Here, $g_{k,l}$ and $b_{k,l}$ are the conductances and susceptances of the transmission line $(k,l) \in \mathcal{L}$, which we aim to estimate. Note that $Y_{k,l} = 0$ if $(k,l) \notin \mathcal{L}$. The set $\mathcal{G} \in \mathcal{N}$ collects all nodes equipped with generators. Figure 1 shows an exemplary 5-bus system with $\mathcal{N} = \{1,\dots,5\}$, and $\mathcal{G} = \{1, 3, 4, 5\}$.

Let v_k denote the voltage amplitude at the k-th node and θ_k the corresponding voltage angle. Throughout this

XD and BH are with the School of Information Science and Technology, ShanghaiTech University, Shanghai, China. {duxu, borish}@shanghaitech.edu.cn

XD is also with Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, as well as the University of Chinese Academy of Sciences, China.

TF and AE are with the Institute for Energy Systems, Energy Efficiency and Energy Economics, TU Dortmund University, Germany. {alexander.engelmann, timm.faulwasser}@ieee.org,

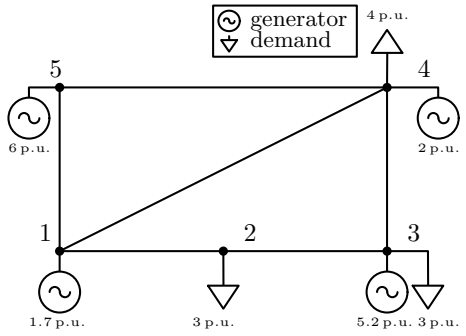

Fig. 1. Modified 5-bus system from [13] with 4 generators and 3 consumers.

paper, we assume that the voltage magnitude and the voltage angle at the first node (the slack node) are fixed, $\theta_1 = \text{const}$ and $v_1 = \text{const}$. This assumption can be made without loss of generality, since the power flow in the network depends on the voltage angle differences $\theta_k - \theta_l$. Since θ_1 and v_1 are given, we define the state of the system as $x \doteq (v_2, \theta_2, v_3, \theta_3, \dots, v_N, \theta_N)^\top$. Moreover, we have active and reactive power generation of generators p_k^g and q_k^g for all $k \in \mathcal{G}$. The tuple $d_k = (p_k^d,\ q_k^d)^\top$ denotes the active and reactive power demand at demand nodes $\mathcal{D} \subseteq \mathcal{N}$. As an input, we consider only the reactive powers at all but the first generator $u \doteq (p_1^g, (q_k^g)_{k\in\mathcal{G}})$ and we assume that the active power generation $\{p_k^g\}_{k\in\mathcal{G}\setminus\{1\}}$ is fixed. Moreover, $y \doteq (g_{k,l}, b_{k,l})_{(k,l)\in\mathcal{L}}^\top \in \mathbb{R}^{2|\mathcal{L}|}$ denotes the parameter vector. Note that transmission lines $(k,l) \notin \mathcal{L}$ are not considered in y, i.e., they are not estimated and thus the sparsity of Y is considered. The grid topology is assumed to be known in advance.

The active and reactive power flow over the transmission line $(k,l) \in \mathcal{L}$ is given by

$$\Pi_{k,l}(x,y) \doteq v_k^2 \begin{pmatrix} g_{k,l} \\ -b_{k,l} \end{pmatrix} - v_k v_l \begin{pmatrix} g_{k,l} & b_{k,l} \\ -b_{k,l} & g_{k,l} \end{pmatrix} \begin{pmatrix} \cos(\theta_k - \theta_l) \\ \sin(\theta_k - \theta_l) \end{pmatrix}.$$

The total power outflow from node $k \in \mathcal{N}$ is given by

$$\sum_{l\in\mathcal{N}_k} \Pi_{k,l}(x,y) = P_k(x,y) \doteq v_k^2 \sum_{l\in\mathcal{N}_k} \begin{pmatrix} g_{k,l} \\ -b_{k,l} \end{pmatrix} - v_k \sum_{l\in\mathcal{N}_k} v_l \begin{pmatrix} g_{k,l} & b_{k,l} \\ -b_{k,l} & g_{k,l} \end{pmatrix} \begin{pmatrix} \cos(\theta_k - \theta_l) \\ \sin(\theta_k - \theta_l) \end{pmatrix},$$

where $\mathcal{N}_k \doteq \{ l \in \mathcal{N} \mid (k,l) \in \mathcal{L} \}$ denotes the set of neighbors of node $k \in \mathcal{N}$. Thus, the power flow equations can be written in the form

$$P(x,y) = S(u)\ , \tag{1}$$

with $\dim(P) = 2|\mathcal{N}|$, where $P(x,y) \doteq \big(P_1(x,y)^\top, \dots, P_N(x,y)^\top\big)^\top$, and $S(u) \doteq \big(S_1(u)^\top, \dots, S_N(u)^\top\big)^\top$. Moreover,

$$S_k(u) \doteq \begin{pmatrix} p_k^g \\ q_k^g \end{pmatrix} - d_k,\ k \in \mathcal{D},\ S_k(u) \doteq \begin{pmatrix} p_k^g \\ q_k^g \end{pmatrix},\ k \notin \mathcal{D}.$$

III. Optimal Design of Experiments

In the following we assume that we can measure all states x and the power flow over the transmission lines $\pi_{k,l}$. Hence, the measurement function is defined by

$$M(x,y) \doteq \Big[x^\top, (\Pi_{k,l}(x,y))_{(k,l)\in\mathcal{L}}^\top\Big]^\top .$$

We assume additive Gaussian measurement noise with zero mean and given variance $\Sigma \in \mathbb{S}_{++}^m$, i.e. $\chi \sim \mathcal{N}(0,\Sigma)$. Hence, the measurements η are given by

$$\eta = M(x,y) + \chi.$$

An associated Maximum Likelihood Estimation (MLE) problem is given by [5]

$$\begin{aligned} \min_{x,y}\ & \frac{1}{2}\|M(x,y) - \eta\|_{\Sigma^{-1}}^2 + \frac{1}{2}\|y - \hat{y}\|_{\Sigma_0^{-1}}^2 \\ \text{s.t. }\ & P(x,y) = S(u), \quad \underline{x} \le x \le \overline{x}. \end{aligned} \tag{2}$$

Here, we assume that $\hat{y} \in \mathbb{R}^{2|\mathcal{L}|}$ is a given initial parameter estimate with given variance $\Sigma_0 \in \mathbb{S}_{++}^{2|\mathcal{L}|}$.

Remark 1 (Minimal Number of Measurements) *Note that the number of measurements we use here, $(4|\mathcal{L}|+2|\mathcal{N}|)$, is not minimal. A necessary condition to determine y uniquely is that there are at least $2|\mathcal{L}|$ measurements. This follows from the implicit function theorem [14, Thm 9.28]. However, the rank of $\frac{\partial M}{\partial y}$ also depends on the network topology (e.g. whether there exists islands) and the distribution of measurement devices in the network [15, Chap. 4], [16]. Approaches for reducing the number of measurements with appropriate measurement placement can be found in [17–19].*

A. The Fisher Information Matrix and OED

Next, we derive an approximation for the FIM, which we will use to compute inputs to maximize the information gained in an estimation step. The FIM characterizes the information content, which can be gained by an experiment. It can be expressed as [11]

$$\mathcal{F}(x,y,u) \doteq \Sigma_0^{-1} + \mathcal{T}(x,y,u)^\top \Sigma^{-1} \mathcal{T}(x,y,u)\,, \tag{3}$$

where

$$\mathcal{T}(x,y,u) \doteq \frac{\partial}{\partial y} M(x,y) + \frac{\partial}{\partial x} M(x,y) \frac{\partial}{\partial y} x^*(y,u). \tag{4}$$

Note that all derivatives are evaluated at the true parameter y. As the true parameter y is a priori unknown, we replace y by our current best guess $\hat{y}$ in the following, which is common practice in the context of OED [20].

The power flow equation (1) has in general multiple solutions. For example, this equation is invariant under voltage angle shifts. However, if the sensitivity matrix $\frac{\partial}{\partial x} P(x,y)$ has full rank at an optimal solution (x,y) of (2), we can use the implicit function theorem to show that a locally differentiable

parametric solution $x^*(y,u)$ of Equation (1) exists.[1] Thus, the last term in (4) reads

$$\frac{\partial}{\partial y}x^*(y,u) \doteq -\left[\frac{\partial}{\partial x}P(x,y)\right]^{-1}\frac{\partial}{\partial y}P(x,y). \tag{5}$$

Since $\mathcal{F}$ is a mapping $(\mathbb{R}^{n_x},\mathbb{R}^{2|\mathcal{L}|},\mathbb{R}^{n_u}) \to \mathbb{R}^{2|\mathcal{L}|\times 2|\mathcal{L}|}$, we have to choose a scalar criterion for characterizing the information content. Typical choices are the A-criterion, the D-criterion or the E-criterion. For details on their advantages/disadvantages and interpretations we refer to [20, 22]. Here, we choose the A-criterion, which minimizes the trace of the inverse of the (approximate) FIM. Thus, the OED problem is given by [8]

$$\begin{aligned}(x^*(\hat{y}),u^*(\hat{y})) = \arg\min_{x,u}\ & \operatorname{Tr}([\mathcal{F}(x,\hat{y},u)]^{-1}) \\ \text{s.t.}\ & \begin{cases} P(x,\hat{y}) = S(u) \\ \underline{u} \le u \le \overline{u} \\ \underline{x} \le x \le \overline{x} \end{cases}\end{aligned} \tag{6}$$

The overall OED algorithm is shown in Figure 2.[2]

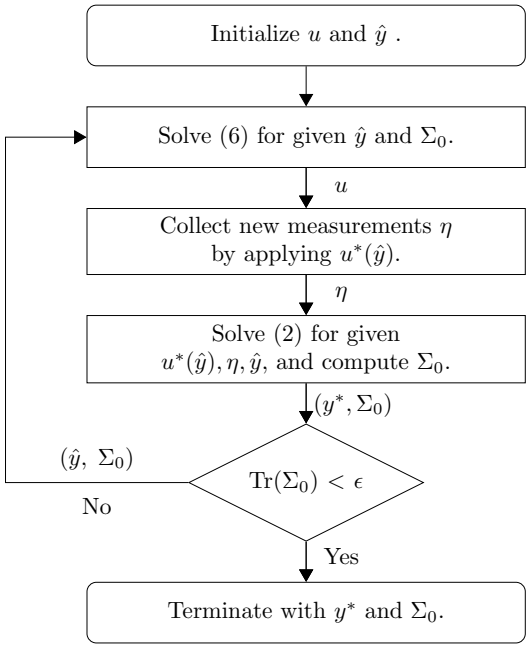

Fig. 2. Optimal experiment design for AC power grid admittance estimation.

IV. Approximation Techniques

There are two main computational difficulties, which make solving (6) challenging: a) we have to invert $\partial P/\partial x$ in (5) and, b) we have to compute the inverse of $\mathcal{F}$. In the following we will explore approaches for reducing the computational complexity.

A. Approximation via Inner Iterations

The inverse in (5) is difficult to compute in a symbolic or automatic differentiation context—especially if N is large. Therefore, we replace the decision variable x with an approximation of $\tilde{x}^*(\hat{x},\hat{y},\hat{u})$ that is defined by an iterative procedure.

Observe that (1) is a set of nonlinear equations, which can be solved locally using Newton-type method. Define $G(x,y,u) = P(x,y) - S(u)$. Newton-type method requires the evaluation and inversion of $\frac{\partial G}{\partial x}(x,y,u)$ in each step. This is expensive, and, hence we use a constant Jacobian approximation at the current iterate $\frac{\partial G}{\partial x}(x,y,u) \approx \frac{\partial G}{\partial x}(\hat{x},\hat{y},\hat{u})$ instead of an exact Jacobian. Algorithm 1 summarizes the resulting algorithm for K iterations. Note that from (4), we

Algorithm 1 Newton-type iteration for $\tilde{x}^*(\hat{x},\hat{y},u)$

Input: Current iterates $(\hat{x},\hat{y},\hat{x})$, set $x_0^* = \hat{x}$.
For $k = 0,\dots,K$**:**

$$x_{k+1}^* = x_k^* - \left[\frac{\partial G}{\partial x}(\hat{x},\hat{y},\hat{u})\right]^{-1} G(x_k^*,\hat{y},u)$$

$$k \leftarrow k+1$$

Output: $\tilde{x}^*(\hat{x},\hat{y},u) = x_{K+1}^*$

obtain an approximation of $\mathcal{T}$,

$$\tilde{\mathcal{T}}(\hat{x},\hat{y},u) \doteq \frac{\partial}{\partial y}M(\tilde{x}^*(\hat{x},\hat{y},u),y). \tag{7}$$

Observe that $\tilde{\mathcal{T}}$ is different from $\mathcal{T}$ since $\tilde{x}^*(\hat{x},y,u)$ is no longer an independent variable, which implicitly considers (1). Thus, we define an approximation of the FIM

$$\tilde{\mathcal{F}}(\hat{x},\hat{y},u) \doteq \Sigma_0^{-1} + \tilde{\mathcal{T}}(\hat{x},\hat{y},u)^\top \Sigma^{-1}\tilde{\mathcal{T}}(\hat{x},\hat{y},u). \tag{8}$$

B. Fisher linearized Approximation

Next, we derive an approximation of $\operatorname{Tr}(\tilde{\mathcal{F}}^{-1})$. We have

$$\begin{aligned}&\operatorname{Tr}(\tilde{\mathcal{F}}(\hat{x},\hat{y},u)^{-1}) \\ &= \operatorname{Tr}(\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}) + \operatorname{Tr}\left[\frac{\partial \tilde{\mathcal{F}}^{-1}}{\partial u}\cdot(u-\hat{u})\right] + O(\|u-\hat{u}\|^2) \\ &= \operatorname{Tr}(\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}) + O(\|u-\hat{u}\|^2) \\ &\quad - \operatorname{Tr}\left(\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}\left[\frac{\partial \tilde{\mathcal{F}}}{\partial u}\cdot(u-\hat{u})\right]\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}\right) \\ &= 2\operatorname{Tr}(\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}) \\ &\quad - \operatorname{Tr}(\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}\tilde{\mathcal{F}}(\hat{x},\hat{y},u)\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}) + O(\|u-\hat{u}\|^2),\end{aligned} \tag{9}$$

where we used a first-order Taylor expansion and $\operatorname{Tr}(A+B) = \operatorname{Tr}(A) + \operatorname{Tr}(B)$ in the first row, [23, Eq. (59)] in the

[1] Conditions under which the matrix $\frac{\partial}{\partial x}P(x,y)$ has full-rank can be found in [21], where linear independence constraint qualifications for AC power flow problems are discussed in a more general setting.

[2] Note that in contrast to [8], we omit the regularization term to avoid input chattering. As an alternative one can simply stop the OED algorithm once the desired variance is reached.

second row, and again a Taylor expansion in the third row. Note that $\frac{\partial \tilde{\mathcal{F}}^{-1}}{\partial u} \in \mathbb{R}^{2|\mathcal{L}|\times 2|\mathcal{L}|\times \mathcal{G}}$ is a tensor with

$$\frac{\partial \tilde{\mathcal{F}}^{-1}}{\partial u} \cdot (u - \hat{u}) = \tilde{\mathcal{F}}^{-1} \left[\sum_i \frac{\partial \tilde{\mathcal{F}}}{\partial u_i} \cdot (u_i - \hat{u}_i) \right] \tilde{\mathcal{F}}^{-1}.$$

Remark 2 *Notice that* (9) *can be interpreted as a weighted* T-criterion *[24, Chapter 6.5].*

V. OED Reformulations

In this section we use the approximations from the previous section to reformulate the OED problems. Three options for doing so are discussed.

A. Fisher Linearized Approximation OED

Using both approximations from the previous section, we rewrite (6) as

$$\begin{aligned} \min_u \; & -\operatorname{Tr}(\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}\tilde{\mathcal{F}}(\hat{x},\hat{y},u)\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}) \\ \text{s.t.} \; & \begin{cases} \underline{u} \le u \le \overline{u}, \\ \underline{x} \le \tilde{x}^*(\hat{x},\hat{y},u) \le \overline{x}. \end{cases} \end{aligned} \tag{10}$$

Observe that we only have u as decision variables, which reduces the problem dimension. We neglect the term $2\operatorname{Tr}(\tilde{\mathcal{F}}(\hat{x},\hat{y},\hat{u})^{-1}$ in (9) since a constant offset in the objective does not change the minimizer. Note that the variables $(\hat{x},\hat{y},\hat{u})$ are fixed to their current iterates in the above problem, and $\tilde{x}^*(\hat{x},\hat{y},u)$ is given by Algorithm 1.

The essence of the above method is to hide the power flow equations via Algorithm 1. As the number of iterations K increases, the complexity of the objective function of (10) increases, although the external form is concise. In the next subsection, we use a different approach, which is less accurate but also less costly.

B. Quadratic Approximation OED

Observe that the objective in (10) is nonlinear. A quadratic approximation is given by

$$\begin{aligned} \min_u \; & \frac{1}{2} u^\top H(\hat{x},\hat{y},\hat{u})u + J(\hat{x},\hat{y},\hat{u})u \\ \text{s.t.} \; & \begin{cases} \underline{u} \le u \le \overline{u}, \\ \underline{x} \le C(\hat{x},\hat{y},\hat{u})u + \hat{x} \le \overline{x} \end{cases} \end{aligned} \tag{11}$$

where

$$\begin{aligned} J(\hat{x},\hat{y},\hat{u}) &\doteq \frac{\partial E}{\partial u}(\hat{x},\hat{y},\hat{u}), \quad H(\hat{x},\hat{y},\hat{u}) \doteq \frac{\partial^2 E}{\partial u^2}(\hat{x},\hat{y},\hat{u}) \\ C(\hat{x},\hat{y},\hat{u}) &\doteq \frac{\partial \tilde{x}^*(\hat{x},\hat{y},u)}{\partial u}(\hat{x},\hat{y},\hat{u}) \end{aligned}$$

and $E(\cdot)$ represents the objective function from (10). Note that the above problem is a Quadratic Program (QP), which can be solved efficiently by standard QP solvers.

C. Inner Linearized Approximation OED

Computing $x^*(\hat{x},\hat{y},u)$ via Algorithm 1 can lead to large memory requirements in the context of automatic differentiation, since the expression graph grows with the number of the Newton-type iterations K.

As an alternative, we use the nonlinear constraints from (6) combined with the objective approximation from (9). This leads to

$$\begin{aligned} \min_{x,u} \; & -\operatorname{Tr}(\mathcal{F}(\hat{x},\hat{y},\hat{u})^{-1}\mathcal{F}(x,\hat{y},u)\mathcal{F}(\hat{x},\hat{y},\hat{u})^{-1}) \\ \text{s.t.} \; & \begin{cases} P(x,\hat{y}) = S(u) \\ \underline{u} \le u \le \overline{u}, \qquad \underline{x} \le x \le \overline{x}. \end{cases} \end{aligned} \tag{12}$$

Here, we use (3) for evaluating the objective, where we substitute $\frac{\partial}{\partial y}\bar{x}^*(y,u) = -\left[\frac{\partial}{\partial x}P(\hat{x},\hat{y})\right]^{-1}\frac{\partial}{\partial y}P(x,y)$ in (4) for a fixed Jacobian evaluated at $(\hat{x},\hat{y})$.

Remark 3 (Difference between $\tilde{\mathcal{F}}$ and $\mathcal{F}$) *Notice the difference between $\tilde{\mathcal{F}}$ and $\mathcal{F}$: Whereas $\tilde{\mathcal{F}}$ uses the approximation $\tilde{\mathcal{T}}$ from* (8) *including the Newton-type iteration from Algorithm 1, $\mathcal{F}$ in* (6) *and* (12) *uses* (3)*, and* (12) *with the approximation $\frac{\partial}{\partial y}\bar{x}^*(y,u) = -\left[\frac{\partial}{\partial x}P(\hat{x},\hat{y})\right]^{-1}\frac{\partial}{\partial y}P(x,y)$ in* (4)*.*

VI. Numerical Results

We illustrate the numerical performance of the approximations on a modified 5-bus power system from [13] (cf. Figure 1) and on a 14-bus power system [25].

A. Implementation and Data

The problem data is obtained from the `MATPOWER` dataset [26] ignoring shunt elements. The implementation of OED relies on `Casadi-v.3.4.5` with `IPOPT` [27] and `MATLAB 2020b`. We use Gaussian measurement noise with zero mean and a variance of 10^{-4}. The initial value of Σ_0^{-1} is set to $10^{-16}I$. We benchmark our method also against MLE with a standard Gaussian random input for each generator, scaled such that the total reactive power demand is met. For the 14-bus system, the input variance is $0.01\,(\text{p.u.})^2$.

B. Numerical Comparison

Next, we compare all OED variants numerically: a) classical OED from [8] using (6), b) *Fisher Linearized Approximation OED* (FLA-OED) from Section V-A, c) *Quadratic Approximation OED* (QA-OED) method from Section V-B, and, d) *Inner Linearized Approximation OED* (ILA-OED), from Section V-C.

TABLE I
Computation time for 200 Batches

OED Method	OED	FLA	QA	ILA
Time [sec]	19.7926	36.1696	3.9811	8.6328
Sum Variance [S^2]	0.0582	0.0281	0.0358	0.0583

Table I shows the computation times with 200 batches for the 5-bus system. We use $K = 20$ Newton-type iterations in

Algorithm 1 for FLA-OED and QA-OED. Compared with the classical OED, FLA-OED takes a longer computation time for the same number of batches, but it also achieves a higher accuracy. Table II shows the computation time for a

TABLE II

COMPUTATION TIME FOR A TARGET SUM VARIANCE OF 0.1.

Method	OED	FLA	QA	ILA
Time [sec]	11.1998	7.7718	1.7076	4.9239
Sum Variance [S^2]	0.0992	0.0996	0.0983	0.0990

given target sum variance of 0.1. Here, one can observe the trade-off between solution accuracy and computation time: the accurate classical OED method and ILA-OED are faster here, since they require a smaller number of batches to get to the desired variance. QA-OED still has the shortest operation time for the given target sum variance. QA-OED solves the above problem via quadratic approximation, which leads to fast computation. ILA-OED shows a similar accuracy as the standard OED with less computation time. Figure 3 shows

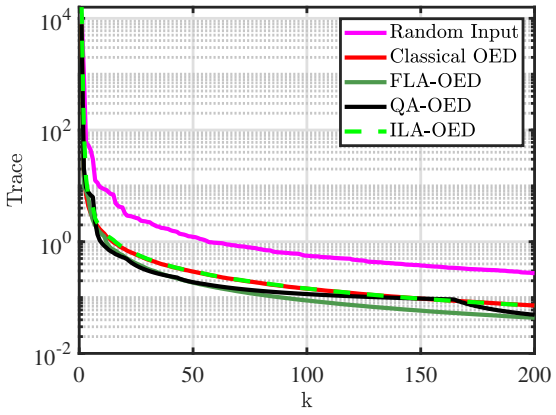

Fig. 3. Expected total variance comparison random Input, classical OED, FLA-OED, QA-OED as well as ILA-OED method for the 5-bus power system.

the expected total variance $\mathrm{Tr}(\mathcal{F}(\hat{x},\hat{u},\hat{y})^{-1})$, where $(\hat{x},\hat{u},\hat{y})$ are fixed to the current iterates. All four methods lead to similar levels of accuracy and they are more accurate than the pure standard RLS with random generator input.

Figure 4 compares the mean relative error (MRE)

$$\mathrm{MRE_g} = \frac{1}{|\mathcal{L}|}\sum_{(k,l)\in\mathcal{L}}\frac{|g_{k,l}-\bar{g}_{k,l}|}{|\bar{g}_{k,l}|},$$

$$\mathrm{MRE_b} = \frac{1}{|\mathcal{L}|}\sum_{(k,l)\in\mathcal{L}}\frac{|b_{k,l}-\bar{b}_{k,l}|}{|\bar{b}_{k,l}|},$$

for RLS with random input, for classical OED and QA-OED. The results of FLA-OED and ILA-OED are similar to the ones shown in Figure 4 and are thus omitted. Note that the value of $\mathcal{F}$ in (3) increases in each iteration because of the second term, which leads the decrease of its inverse. We refer to [28, Chapter 7] and [22] for further discussion.

Figure 5 shows the optimal reactive power inputs for all methods. Here one can see that the reactive power update

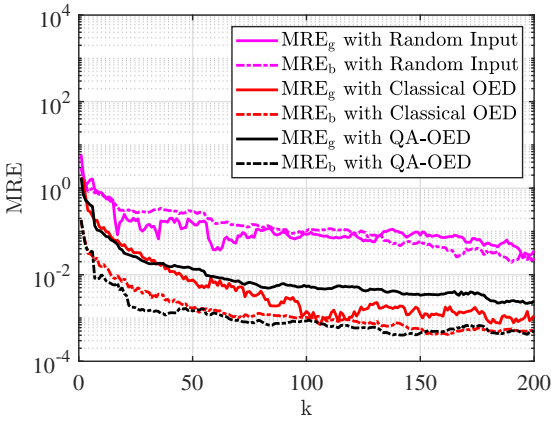

Fig. 4. Mean relative errors of $\mathrm{MRE_g}$ (solid line) and $\mathrm{MRE_b}$ (dashed line) for the 5-bus power system.

is less frequent in QA-OED compared with the other three methods, which is benefitial in grid operation.

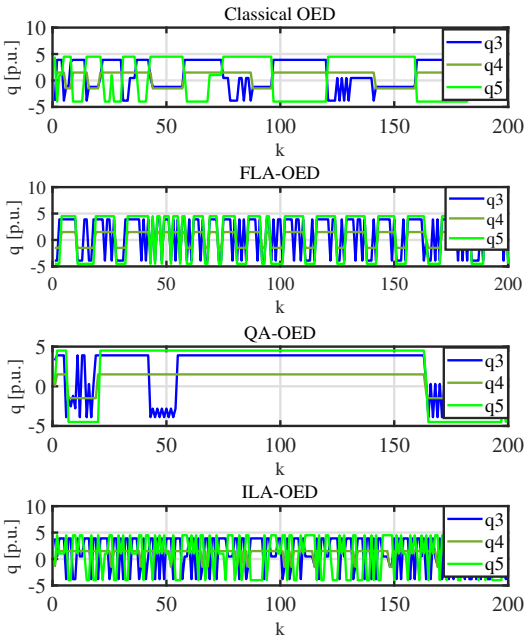

Fig. 5. Optimal reactive power inputs for all generators as obtained by proposed methods the 5-bus power system.

Table III shows the relative loss of optimality $\Delta(u(\cdot))$ of the first iteration with different methods based on the optimal objective $T^\star$ of classical OED,

$$\Delta(u(\cdot)) = \frac{\mathrm{Tr}(\mathcal{F}^{-1}(x(u(\cdot),\hat{y}),\hat{y},u(\cdot)) - T^\star}{T^\star}.$$

Here, $u(\cdot)$ denotes inputs for different OED variants. The

TABLE III

RELATIVE LOSS OF OPTIMALITY FOR THE FIRST ITERATION

Method	Random	FLA	QA	ILA
$\Delta(u(\cdot))$	8.45	0.36	0.36	4.94

relative loss of optimality for the random input is computed

as an average of 100 samples. One can observe the benefit of OED against a random input especially in early iterations.

Figure 6 shows the total variance $\mathrm{Tr}(\mathcal{F}(\hat{x},\hat{u},\hat{y})^{-1})$ with QA-OED applied to a modified 14-bus network. In order to have a larger generator-to-node ratio, we add extra generators on bus 9-13. With our best approximation method. QA-OED requires 15.2507 seconds for computing 200 estimation steps, while FLA-OED and ILA-OED are still not computational tractable for the 14-bus power system.

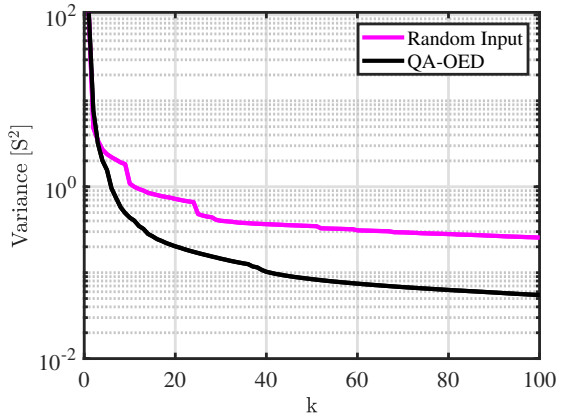

Fig. 6. Expected total variance comparison between random input and QA-OED for the 14-bus power system.

VII. Summary and Outlook

This paper has presented approximation methods for optimal experiment design operation of power grids in order to accelerate computation. We have shown that different approximations are possible, and first numerical results indicate that a combination of inner Newton-type iterations with quadratic approximation is promising. However, further improving the scalability of OED seems crucial for application in practice. Approaches based on semidefinte programming will be considered in future work.

References

[1] J. Zhu, *Optimization of power system operation*. John Wiley & Sons, 2015.

[2] S. Frank, I. Steponavice, and S. Rebennack, "Optimal power flow: A bibliographic survey i," *Energy systems*, vol. 3, no. 3, pp. 221–258, 2012.

[3] S. Frank, I. Steponavice, and S. Rebennack, "Optimal power flow: A bibliographic survey ii," *Energy systems*, vol. 3, no. 3, pp. 259–289, 2012.

[4] X. Bian, X. R. Li, H. Chen, D. Gan, and J. Qiu, "Joint estimation of state and parameter with synchrophasors—part ii: Parameter tracking," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1209–1220, 2011.

[5] I. W. Slutsker, S. Mokhtari, and K. A. Clements, "Real time recursive parameter estimation in energy management systems," *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1393–1399, 1996.

[6] O. Lateef, R. G. Harley, and T. G. Habetler, "Bus admittance matrix estimation using phasor measurements," in *2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, IEEE, 2019, pp. 1–5.

[7] M. Saadeh, M. Alsarray, and R. McCann, "Estimation of the bus admittance matrix for transmission systems from synchrophasor data," in *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, IEEE, 2016, pp. 1–5.

[8] X. Du, A. Engelmann, Y. Jiang, T. Faulwasser, and B. Houska, "Optimal experiment design for ac power systems admittance estimation," in *In Proceedings of the 21st IFAC World Congress, Berlin, Germany*, 2020.

[9] E. Fabbiani, P. Nahata, G. De Nicolao, and G. Ferrari-Trecate, "Identification of ac networks via online learning," *arXiv preprint arXiv:2003.06210*, 2020.

[10] E. Fabbiani, P. Nahata, G. De Nicolao, and G. Ferrari-Trecate, "Identification of ac distribution networks with recursive least squares and optimal design of experiment," *IEEE Transactions on Control Systems Technology*, 2021.

[11] B. Houska, D. Telen, F. Logist, M. Diehl, and J. F. V. Impe, "An economic objective for the optimal experiment design of nonlinear dynamic processes," *Automatica*, vol. 51, pp. 98 –103, 2015.

[12] X. Du, A. Engelmann, T. Faulwasser, and B. Houska, "Online power system parameter estimation and optimal operation," in *In Proceedings of the American Control Conference, New Orleans, USA*, 2021, pp. 3126–3131.

[13] F. Li and R. Bo, "Small test systems for power system economic studies," in *IEEE PES General Meeting*, 2010, pp. 1–4.

[14] W. Rudin, *Principles of Mathematical Analysis*. Chennai: Example Product Manufacturer, 2013.

[15] A. Abur and A. G. Expósito, *Power System State Estimation: Theory and Implementation* (Power Engineering). CRC Press, 2004.

[16] T. Baldwin, L. Mili, M. Boisen, and R. Adapa, "Power system observability with minimal phasor measurement placement," *IEEE Transactions on Power Systems*, vol. 8, no. 2, pp. 707–715, 1993.

[17] A. Pal, A. K. S. Vullikanti, and S. S. Ravi, "A pmu placement scheme considering realistic costs and modern trends in relaying," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 552–561, 2016.

[18] N. M. Manousakis, G. N. Korres, and P. S. Georgilakis, "Taxonomy of pmu placement methodologies," *IEEE Transactions on power Systems*, vol. 27, no. 2, pp. 1070–1077, 2012.

[19] P. L. Donti, Y. Liu, A. J. Schmitt, A. Bernstein, R. Yang, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2520–2530, 2019.

[20] D. Telen, F. Logist, E. Van Derlinden, I. Tack, and J. Van Impe, "Optimal experiment design for dynamic bioprocesses: A multi-objective approach," *Chemical Engineering Science*, vol. 78, pp. 82–97, 2012.

[21] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, "Generic existence of unique lagrange multipliers in ac optimal power flow," *IEEE Control Systems Letters*, vol. 2, no. 4, pp. 791–796, 2018.

[22] D. Telen, B. Houska, F. Logist, E. Van Derlinden, M. Diehl, and J. Van Impe, "Optimal experiment design under process noise using riccati differential equations," *Journal of Process Control*, vol. 23, pp. 613–629, 2013.

[23] K. B. Petersen and M. S. Pedersen, *The matrix cookbook*, Version 20121115, 2012.

[24] F. Pukelsheim, *Optimal design of experiments*. SIAM, 2006.

[25] R. Christie, "Power systems test case archive," *Electrical Engineering dept., University of Washington*, vol. 108, 2000.

[26] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2010.

[27] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi: A software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.

[28] L. Ljung, *System Identification - Theory for the User*, 2nd ed. New Jersey: Prentice Hall, 1999.