# Online power system parameter estimation and optimal operation

Xu Du, Alexander Engelmann, Timm Faulwasser and Boris Houska

*Abstract*— The integration of renewables into electrical grids calls for novel control schemes, which usually are model based. Classically, for power systems parameter estimation and optimization-based control are often decoupled, which may lead to increased cost of system operation during the estimation procedures. The present work proposes a method for simultaneously minimizing grid operation cost and estimating line parameters. To this end, we rely on methods from optimal design of experiments. This approach leads to a substantial reduction in cost for optimal estimation and in higher accuracy in the parameters compared with standard combination of optimal power flow and maximum-likelihood estimation. We illustrate the performance of the proposed method on simple benchmark system.

*Keywords:* Optimal Experiment Design, Power System Parameter Estimation, Admittance Estimation, Optimal Power Flow

## I. INTRODUCTION

Economical optimal and safe operation of power systems with a large share of renewables requires reliable grid models. While the grid topology is often known, the parameters are are frequently unknown or erroneous [1], [2]. A classical approach for the optimal operation of power systems is to run an estimation procedure obtaining grid parameters first, and secondly using these parameters in an Optimal Power Flow (OPF) problem for computing optimal generator set-points. Although this procedure is usually reliable, it may lead to high system operation cost until the estimation is converge to accurate parameters.

Established theory of maximum-likelihood estimation and Bayesian methods can be found in many textbooks [3], [4]. For static power system parameter estimation based on multiple measurement snapshots, recursive least-squares based techniques have been proposed in [5], [6], [7]. Approaches for combined parameter and topology estimation are considered in [8], [9]. These works do typically consider constant or given power injections.

Recently, parameter estimation based on techniques from Optimal Experiment Design (OED) have been suggested [10], [11].[1] Therein, the conceptually new idea is to compute set points for generators such that a maximum amount of

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

TF and AE are with the Institute of Energy Systems, Energy Efficiency and Energy Economics, TU Dortmund University, Germany. Main parts of this work have been conducted while AE was with the Institute for Automation and applied Informatics, Karlsruhe Institue of Technology, Germany. alexander.engelmann@tu-dortmund.de, timm.faulwasser@ieee.org

[1]Consider [12] for an excellent presentation of the foundations of OED.

information is extracted in each time instant leading to a fast and accurate estimation. To this end, an optimization problem is constructed, which minimizes the trace of the covariance matrix of the parameter estimates. Although this procedure leads to fast and accurate estimates, it also induces to high cost of system operation in the estimation procedure since it is agnostic to the associated economic costs of choosing the set points optimally with respect to the estimation variance.

A second branch considers an *economic* variant of OED by considering the cost of experiments in the design procedure [13]. This approach is combined with model predictive control [14] and with power system parameter estimation [10].

In the present work we propose a similar approach for power system parameter estimation aiming at lowering the operation cost in the estimation procedure. We compute a Pareto front trading off system operation cost via OPF versus the goal of obtaining highly accurate grid parameters. Based on this curve, we develop a scheme for adjusting the weighting parameter in the estimation step to reach a predefined accuracy in the parameter estimates after a desired number of sampling instants.

The remainder of this paper is organized as follows: Section II recalls AC grid modeling basics. In Section III we describe the main contribution of this paper: a method trading-off estimation with optimal system operation based on optimal design of experiments. Section IV shows promising numerical results in terms of a higher accuracy in the parameters compared with classical maximum-likelihood estimation methods and a substantially reduced operation cost compared to classical OED.

## II. THE AC GRID MODEL AND OPTIMAL POWER FLOW

This section recalls basics of AC power system modeling and the AC OPF problem, which serves as a basis for our developments.

### A. Power Grid Model

We consider a power grid $(\mathcal{N}, \mathcal{L})$, where $\mathcal{N} = \{1 \ldots N\}$ denotes the set of buses and $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$ represents the set of transmission lines. The physical properties of transmission lines are described by line conductances $g_{k,l}$ and line susceptances $b_{k,l}$ for all transmission lines $(k, l) \in \mathcal{L}$, which we would like to estimate. We set $g_{k,l} = b_{k,l} = 0$ for all $(k, l) \notin \mathcal{L}$ [15], [16]. We collect these line parameters in

$$y = \begin{pmatrix} g_{k,l} \\ b_{k,l} \end{pmatrix}_{(k,l) \in \mathcal{L}} \in \mathbb{R}^{2|\mathcal{L}|},$$

where $(v_k)_{k \in \mathcal{S}}$ denotes the vertical concatenation of vectors $v$ over the index set $\mathcal{S}$.

Denoting the voltage amplitude at node $k$ by $v_k$ and the voltage magnitude at node $k$ by $\theta_k$, we define the algebraic state of the system as

$$x = (v_2, \theta_2, v_3, \theta_3, \ldots, v_N, \theta_N)^\top .$$

The voltage magnitude and the voltage angle at the first node, the slack node, are assumed to be fixed and given

$$\theta_1 = 0 \quad \text{and} \quad v_1 = \text{const.}$$

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

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

Summing up the transmission line flows of all nodes neighbored to node $k$ yields the residual active and reactive power

$$S_k(x, y) = \sum_{l \in \mathcal{N}_k} \Pi_{k,l}(x, y).$$

Let $p_k^{\text{g}}$ and $q_k^{\text{g}}$ denote the active and reactive power generation at node $k$. Then we set $p_k^g = 0$ and $q_k^g = 0$ for all $k \notin \mathcal{G}$, where $\mathcal{G} \subseteq \mathcal{N}$ denotes the set of generators of the system. We denote the active power demand at node $k$ by $p_k^d$ and the reactive power demand at node $k$ by $q_k^d$. Both are assumed to be known and constant and we set $p_k^d = 0$ and $q_k^d = 0$ in case there is no consumer at node $k$. We assume that the active and reactive power generation of all generators except at the first node are the only values we can control. Hence, we introduce an input vector $u$ and a vector of power demands $d$ as

$$u = \begin{pmatrix} p_k^{\text{g}} \\ q_k^{\text{g}} \end{pmatrix}_{k \in \mathcal{N} \setminus \{1\}} \quad \text{and} \quad d = \begin{pmatrix} p_k^{\text{d}} \\ q_k^{\text{d}} \end{pmatrix}_{k \in \mathcal{N} \setminus \{1\}} .$$

With that, the so-called AC power flow equations [16] are given by

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

where

$$S(x, u) = \big(S_k(x, u)\big)_{k \in \mathcal{N} \setminus \{1\}}.$$

### B. Optimal Power Flow

Optimal Power Flow aims at minimizing the total cost of power generation in an electrical grid subject to the power flow equations and physical and technical limits such as voltage bounds, line limits, and generator limits [16], [17],

$$\min_{x, u} \ C(u)$$
$$\text{s.t.} \begin{cases} S(x, y) = u - d \\ H(x, u) \leq 0 \end{cases}. \tag{2}$$

Here, $S$ encodes the power flow equations (1). The cost function is typically quadratic in the active power generation

$$C(u) = \sum_{i \in \mathcal{G}} \alpha_i (p_i^g)^2 + \beta_i (p_i^g) , \tag{3}$$

where $\alpha_i$ and $\beta_i$ are given and positive cost coefficients. Other formulations are possible, for example aiming at minimizing grid losses [16]. Moreover, the limits on voltages and power generation are modelled as

$$H(x, u) = \begin{pmatrix} p_k^g - \bar{p}_k \\ q_k^g - \bar{q}_k \\ x_k - \bar{x}_k \\ \underline{p}_k - p_k^g \\ \underline{q}_k - q_k^g \\ \underline{x}_k - x_k \end{pmatrix}_{k \in \mathcal{N}} .$$

## III. Optimal Experiment Design for OPF

Next, we develop a method, which simultaneously estimates grid parameters and computes optimal generator set points. This leads to a higher accuracy in parameter estimates and lower operation cost compared to classical methods.

### A. Maximum Likelihood Parameter Estimation

Maximum likelihood estimation based on least-squares techniques is a standard method for estimation of line parameters in power systems [6], [7]. The estimates typically rely on (noisy) measurements of transmission line flows, voltage magnitude and angle measurements at buses. In this paper, we assume that all the active and reactive power flows through the transmission lines as well as the states of buses can be measured. Thus, the measurement function $M$ is

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

We assume additive Gaussian measurement noise with zero mean and a given variance $\Sigma \in \mathbb{S}_{++}^{|M|}$.

The corresponding Maximum Likelihood Estimation (MLE) problem for the line parameters $y$ is

$$\min_{x \in \mathbb{R}^{2|\mathcal{N}|-2}, y \in \mathbb{R}^{2|\mathcal{L}|}} \ \frac{1}{2} \|M(x, y) - \eta\|_{\Sigma^{-1}}^2 + \frac{1}{2} \|y - \hat{y}\|_{\Sigma_0^{-1}}^2$$
$$\text{s.t.} \begin{cases} S(x, y) = \hat{u} - d \\ H(x, u) \leq 0 \end{cases}. \tag{4}$$

Here, $\hat{y}$ is the given initial parameter estimate with a given variance $\Sigma_0 \in \mathbb{S}_{++}^{2|\mathcal{L}| \times 2|\mathcal{L}|}$ and $\eta$ are measured values and $\hat{u}$ donates the current system input.

### B. Optimal Experiment Design

Problem (4) depends the system inputs $u$. Optimal experiment design exploits this degree of freedom by choosing the inputs such that a "maximum amount of information" is extracted. The Fisher information matrix $\mathcal{F} \in \mathbb{S}_{++}^{2|\mathcal{L}| \times 2|\mathcal{L}|}$ of (4) encodes the information content in the parameter estimates [18] and is given by

$$\mathcal{F}(x, u, \hat{y}) = \Sigma_0^{-1} + M_p(x, u, \hat{y})^\top \Sigma^{-1} M_p(x, u, \hat{y}).$$

**3127**

Here, we use the shorthand

$$M_p(x_s(u, \hat{y}), u, \hat{y}) = -\frac{\partial M}{\partial x} \left(\frac{\partial S}{\partial x}\right)^{-1} \frac{\partial S}{\partial p},$$

where $x_s(u, \hat{y})$ is the implicit solution of the equation $S(x, \hat{y}) = u - d$, which we assume to be unique [19], [13]. Thus, the associated OED problem can be written as

$$
\begin{aligned}
\min_{x,u} \quad & \mathrm{Tr}(\mathcal{F}(x, u, \hat{y})^{-1}) + c \cdot \|p^g - p^{g-}\|_2^2 \\
\text{s.t.} \quad & \left\{ \begin{array}{l} S(x, \hat{y}) = u - d \\ H(x, u) \leq 0, \end{array} \right.
\end{aligned}
\tag{5}
$$

where $p^{g-}$ denotes the previous system active power input.

## C. Combining OPF and OED

Problem (5) computes optimal inputs $u$ such that the inverse of $\mathcal{F}$ (and thus the variance in the line parameters) is minimized. However, since (5) is agnostic to the induced extra cost of this approach, we introduce a combined OPF-OED problem next, which simultaneously minimizes system operation cost and variance in the line parameters.

For safety reasons, it is often required that the line parameters are known up to a certain accuracy (e.g. to avoid line congestion). Hence, one way to combine estimation with optimal operation is to pre-specify a certain target variance $\mathbb{V}_N^f$, which is necessary for safe operation and should be reached. Such a target variance can typically not be reached in one step. Hence, one approach is to specify that after $N$ time steps, the target variance $\mathbb{V}_N^f$ should be reached and that this target variance should be reached as cheaply as possible. In these time steps, some bounds might be violated. However, in many cases short-term overloading is possible due to thermal inertia of components.

The above problem can be formulated as a multi-stage OED-OPF problem

$$
\begin{aligned}
\min_{[x_k, u_k]} \quad & \sum_{k=1}^{N} C(u_k) \\
\text{s.t.} \quad & \left\{ \begin{array}{l} \mathrm{Tr}(\mathbb{V}_N(x_1, u_1, y_1, \ldots, x_N, u_N, y_N)) \leq \mathrm{Tr}(\mathbb{V}_N^f) \\ S(x_k, y) = u_k - d_k \\ H(x_k, u_k) \leq 0, \end{array} \right.
\end{aligned}
\tag{6}
$$

where

$$
\mathbb{V}_N(x_1, u_1, y_1, \ldots, x_N, u_N, y_N) =
$$
$$
\left( \Sigma_0^{-1} + \sum_{k=1}^{N} M_p(x_k, u_k, y_k)^\top \Sigma^{-1} M_p(x_k, u_k, y_k) \right)^{-1}
$$

is the predicted variance at time step $N$. Here, $[x_k, u_k] \in \mathbb{R}^{(4|\mathcal{N}|-4) \times |N|}$ indicate the decision variables of (6) for $N$ stages and $\Sigma_0^{-1}$ denotes the initial Fisher Information of the parameters.

In general it is hard to say whether the desired target variance is strictly reachable within $N$ steps since these predictions are supported by wrong parameters in each step. Moreover, such a multi-stage problem is also hard from a computational perspective—especially in case of large-scale

---

**Algorithm 1** Autotuned OED-OPF tradeoff for parameter estimation

**Input:** Initial guess of parameter $\hat{y}$ and variance $\mathbb{V}_0 \succ 0$, an initial $\rho > 0$, a termination tolerance $\epsilon > 0$, an initial generator set-point $\hat{u}$, and a terminal variance $\mathrm{Tr}(\mathbb{V}_N^f)$.
**Initialization:** $k = 1$.
**Repeat:**

1) *Collection of Measurements:* set the active and reactive power at the generators to u and take a measurement $\eta$.
2) *Maximum Likelihood Estimation:* Get new measurement $\eta$ and solve *Estimation problem* (4)

$$
\begin{aligned}
(x_s, \hat{y}^+) = \quad & \arg\min_{x,y} \quad \frac{1}{2}\|M(x,y) - \eta\|_{\Sigma^{-1}}^2 + \frac{1}{2}\|y - \hat{y}\|_{\Sigma_0^{-1}}^2 \\
& \text{s.t.} \left\{ \begin{array}{l} S(x, y) = u - d \\ H(x, u) \leq 0 \end{array} \right.
\end{aligned}
$$

3) *Set* $\mathbb{V}^+ = \mathrm{Tr}(\mathcal{F}(x_s, \hat{u}, \hat{y}^+)^{-1})$.
4) *Update the mean of the expectation gap for the remaining steps* $I^+ = \frac{1}{N-k}(\frac{1}{\mathrm{Tr}(\mathbb{V}_N^f)} - \frac{1}{\mathrm{Tr}(\mathbb{V}^+)})$.
5) *Update weight $\rho$:* $\rho \leftarrow \rho + K(I^+ - I_0)$.
6) *Experiment Design:* Solve OED+OPF problem (7) and perform a new experiment with $u^*(\hat{y})$.
7) *Termination Criterion:* If $\mathrm{Tr}(\mathbb{V}^+) < \epsilon$ for a small $\epsilon > 0$, stop.
8) *Update:* Otherwise, set $\Sigma_0 \leftarrow \mathbb{V}^+$, $u \leftarrow u^*(\hat{y})$ and $\hat{y} \leftarrow \hat{y}^+$ and return to step 1) with $k \leftarrow k + 1$.

---

grids. One way this issue is to relax the variance constraint to the objective function in a single-stage setting

$$
\begin{aligned}
(x^*, u^*) = \quad & \arg\min_{x,u} \quad C(u) + \frac{1}{\rho}\mathrm{Tr}(\mathbb{V}(x, u, \hat{y})) \\
& \text{s.t.} \left\{ \begin{array}{l} S(x, \hat{y}) = u - d \\ H(x, u) \leq 0 \end{array} \right.
\end{aligned}
\tag{7}
$$

where $\rho > 0$ is penalty parameter.[2]

## D. Adaptive Strategy for $\rho$

We use an adaptive strategy for $\rho$ to reach our variance target. Note that the target information we would like to have after $N$ iterations is $\frac{1}{\mathrm{Tr}(\mathbb{V}_N)}$. Thus, the average information that we have to collect in each iteration is

$$
I_0 = \frac{1}{N}\left( \frac{1}{\mathrm{Tr}(\mathbb{V}_N^f)} - \frac{1}{\mathrm{Tr}(\mathbb{V}_0)} \right),
$$

where $\mathrm{Tr}(\mathbb{V}_0)$ is the trace of initial parameter variance. The information to be gathered after iteration $k$ to the final iteration $N$ thus is

$$
I^+ = \frac{1}{N-k}\left( \frac{1}{\mathrm{Tr}(\mathbb{V}_N^f)} - \frac{1}{\mathrm{Tr}(\mathbb{V}^+)} \right),
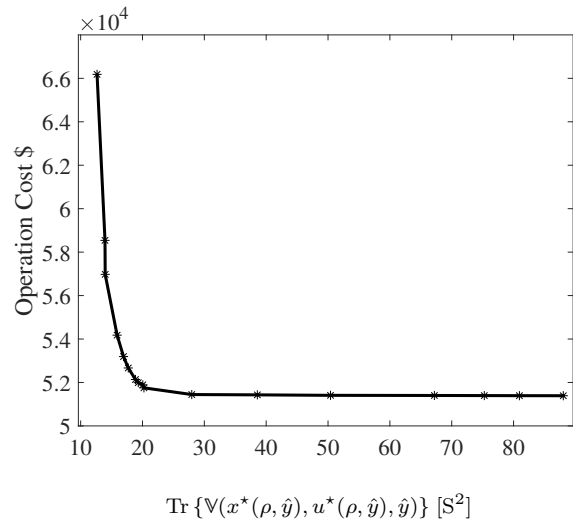$$

where $\mathbb{V}^+ = \mathcal{F}(x_s, \hat{u}, \hat{y})^{-1}$ is the realized variance after the estimation step (4), where $x_s$ denotes the solution of (4). With that, we introduce the update strategy

$$
\rho \leftarrow \rho + K(I^+)(I^+ - I_0),
\tag{8}
$$

---

[2]Note that we solve (7) to *local* optimality only. In the context of power systems, global optimality can usually not be guaranteed due to the non-convexity of the problem.
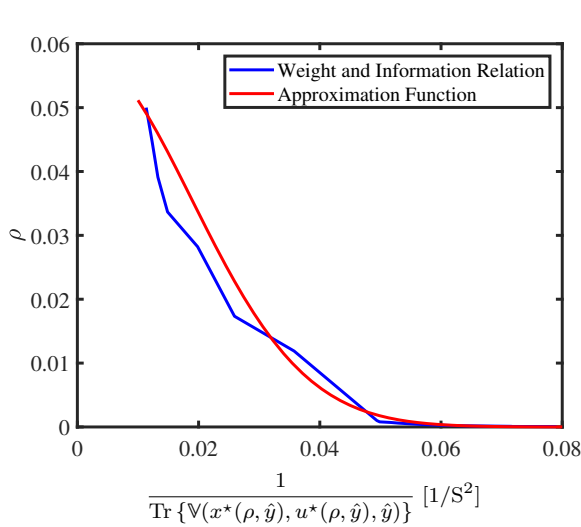
(a) Relationship between $\rho$ and $\frac{1}{\text{Tr}(\mathbb{V})}$.

(b) Pareto optimality curve corresponding with OED and OPF

Fig. 1: Infuence of the weighting parameter $\rho$ on $\text{Tr}(\mathbb{V})$ and on the operation cost.

tracking the average information we have to gather. Here, $K(I^+)$ is an information state dependent feedback gain that is set to

$$K(I^+) = \varphi'(I^+) \,,$$

where $\varphi : \mathbb{R} \to \mathbb{R}$ is an approximation of the inverse trade-off function such that

$$\varphi \left( \frac{1}{\text{Tr}\left\{\mathbb{V}(x^\star(\rho,\hat{y}), u^\star(\rho,\hat{y}),\hat{y})\right\}} \right) \approx \rho \,.$$

Notice that details on how to pre-compute approximations of this function as well the control gain $K(I^+)$ can be found in Section IV. In the above equation, $x^\star$ and $u^\star$ are the minimizers of (7) depending on $\rho$ and $\hat{y}$.

The combined OED-OPF procedure is illustrated in Algorithm 1. In the first step, the algorithm is initialized with trial values for the inputs $u^-$ and the line parameters $y^-$.

The guess for $u$ and $y$ are applied to the system and new measurements $\eta$ are arecollected. These measurements are put in to the maximum-likelihood estimation problem (4) yielding new estimates for the line parameters $y$ and an updated covariance matrix $\Sigma_0$. After running the adaption of $\rho$ from (8), new inputs are computed by means of the combined OED-OPF problem (5). After applying these new inputs, the algorithm starts from the beginning.

## IV. NUMERICAL CASE STUDY

Next, we illustrate the performance of Algorithm 1 on a 5-bus system shown in Figure 2.

### A. Implementation and Data

The problem data is obtained from the `MATPOWER` [20], where we neglect shunt capacities. The implementation of Algorithm 1 relies on `Casadi-v3.4.5` with `IPOPT` and `MATLAB 2019b`. The cost coefficients for the OPF cost $C$ from (3) are given in Table I.
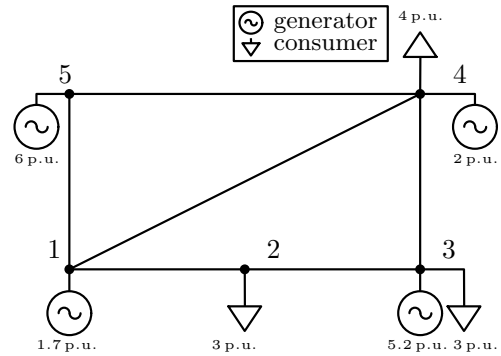


Fig. 2: Modied 5-bus system from Li and Bo (2010) with 4 generators and 3 consumers.

TABLE I: Generator cost coefficients for (3).

| Bus Number | $\alpha_i$ | $\beta_i$ |
|---|---|---|
| 1 | 0.1 | 15 |
| 3 | 0.11 | 30 |
| 4 | 0.12 | 40 |
| 5 | 0.13 | 10 |

We use additive white Gaussian measurement noise with zero mean and variance $10^{-4}$, which is frequently considered used in context of power system parameter estimation [21]. We choose $\rho^0 = 10^{-4}$ for initialization and a sampling time of 15min. We initialize the parameters $y$ with the average true values of the admittance and the initial covariance matrix is set to $\Sigma_0 = 10^{20} \cdot I \in \mathbb{R}^{2|\mathcal{L}| \times 2|\mathcal{L}|}$. Moreover, we set $\text{Tr}(\mathbb{V}_N^f) = 10^0 \, \text{S}^2$ as the target variance after 25 iterations.
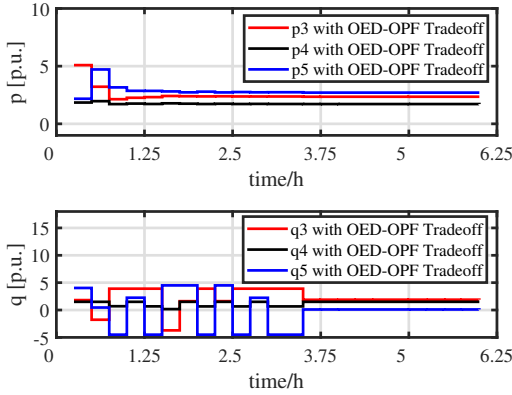
Figure 1a shows the dependency of $\frac{1}{\text{Tr}(\mathbb{V})}$ on $\rho$ for our numerical example after the first iteration in blue. Moreover, Figure 1b shows the corresponding Pareto-optimal curve of problem (7) for different values of $\rho$, where we used the
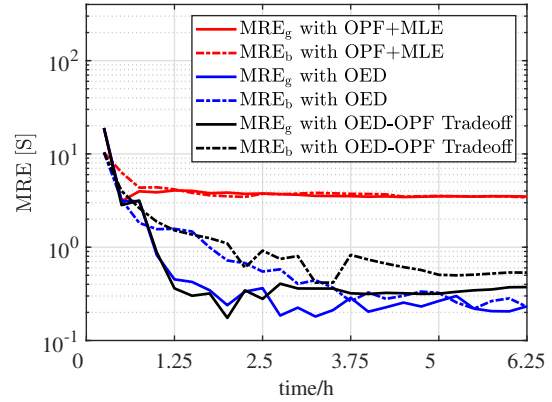
**3129**
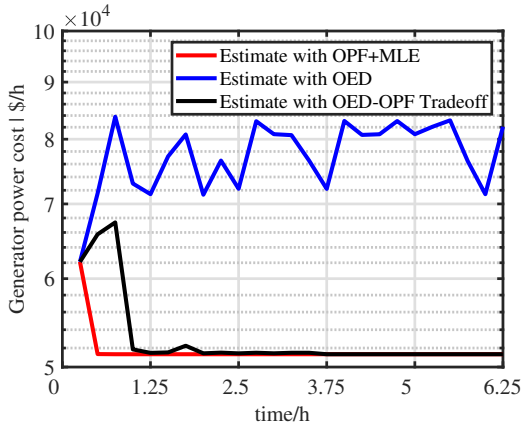
(a) Operation cost of these three methods.
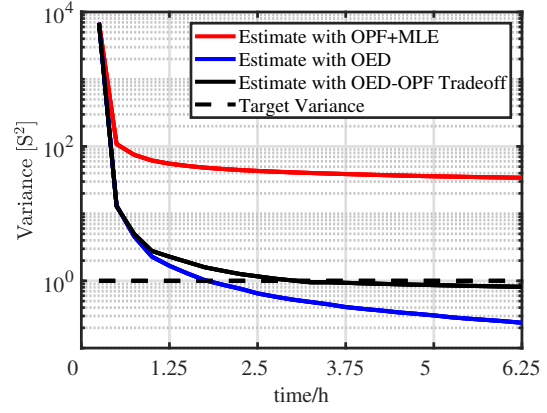
(b) Variance of the line parameters $y$.

(c) Optimal active and reactive power injection from Algorithm 1.

(d) Mean relative errors of the line parameters $g$ and $b$.

Fig. 3: Estimation performance and operation cost for all considered estimation algorithms.

Pareto filter[3] from [22] to remove local optima. We fit an exponential function

$$\varphi\left(\frac{1}{\mathbb{V}}\right) \approx 0.05891 \exp^{-1411(\mathbb{V})^2}$$

for the usage in Algorithm 1 shown in Figure 1a (red).

*B. Numerical Comparison*

Next, we compare the performance of Algorithm 1 to pure OPF combined with maximum likelihood estimation and pure OED in terms of operation cost and variance of the line parameters.

Figure 3 shows the estimation performance and associated costs for all algorithms. With a desired target variance of $\mathrm{Tr}(\mathbb{V}_N^f) = 10^0\,\mathrm{S}^2$, Algorithm 1 leads to a substantially reduced cost compared with classical OED, see Figure 3a. Classical OPF with MLE on the other hand is agnostic to the estimation variance and optimizes only with respect to the associated cost. Hence, classical OPF with MLE leads to a slightly improved cost compared to Algorithm 1 but it comes

with the disadvantage of a significantly worse estimation performance, see Figure 3b. The desired target variance is reached after 6.26 hours which is only 25 iterations with Algorithm 1.

Figure 3c depicts the optimal active and reactive power input of the generators for Algorithm 1. One can see that the active power is changed only in the first few steps and stays almost constant after that. The reactive power varies for more iterations. Pure OED, however, leads to very frequent set-point adjustments in reactive *and* active power, since it is agnostic to the associated economic cost [10]. This behavior is beneficial from a practical perspective: in contrast to changing the active power set-points, changing the reactive power set points it is technically much simpler and cheaper.

Figure 3d shows the mean relative errors

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

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

in the estimation to $g_{k,l}$ and $b_{k,l}$ for all three methods.

Table II shows the ground truth $\bar{y}$ and the OED estimation result after 25 iterations. One can see that in all cases the

[3]In multi-objective optimization problems, the optimal solution is usually not a single one but a set of local optimal solutions of non-convex problems. A Pareto filter is used to filter out partial local optimal solutions to obtain monotonically decreasing Pareto front.

TABLE II: Estimation results for Algorithm 1.

| Line Index | Conductance true val. [S] | Conductance estimate [S] | Susceptance true val. [S] | Susceptance estimate [S] |
|---|---|---|---|---|
| $(1,2)$ | 3.523 | 3.529 | -35.235 | -35.322 |
| $(1,4)$ | 3.257 | 3.167 | -32.569 | -32.703 |
| $(1,5)$ | 15.470 | 14.445 | -154.703 | -153.152 |
| $(2,3)$ | 9.168 | 9.429 | -91.676 | -91.746 |
| $(3,4)$ | 3.334 | 3.896 | -33.337 | -32.474 |
| $(4,5)$ | 3.334 | 3.223 | -33.337 | -33.305 |

relative error is below $10\%$.

## V. Summary and Outlook

This paper presented a parameter estimation method for simultaneously minimizing operation cost based on optimal power flow and estimating line parameters. An example shows that Algorithm 1 achieves a higher estimation accuracy compared with classical estimation methods and at the same time it is cheaper than strategies purely based on the optimal design of experiments.

Future work will consider advanced weighting strategies aiming at replacing the offline scheme adopted in this article. A numerical comparison between multi-stage OED-OPF problem and the penalty-function modified single-stage model will also be discussed.

## References

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

[2] G. L. Kusic and D. L. Garrison. Measurement of transmission line parameters from SCADA data. In *IEEE PES Power Systems Conference and Exposition*, pages 440–445, Oct 2004.

[3] S. M. Kay. *Fundamentals of statistical signal processing*. Prentice Hall PTR, 1993.

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

[5] 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*, 26(3):1209–1220, Aug 2011.

[6] I. W. Slutsker, S. Mokhtari, and K. A. Clements. Real time recursive parameter estimation in energy management systems. *IEEE Transactions on Power Systems*, 11(3):1393–1399, Aug 1996.

[7] T. Van Cutsem and V. H. Quintana. Network parameter estimation using online data with application to transformer tap position estimation. *IEEE Proceedings C - Generation, Transmission and Distribution*, 135(1):31–40, Jan 1988.

[8] D. Deka, S. Backhaus, and M. Chertkov. Learning topology of distribution grids using only terminal node measurements. In *IEEE International Conference on Smart Grid Communications*, pages 205–211, Nov 2016.

[9] S. Park, D. Deka, and M. Chertkov. Exact topology and parameter estimation in distribution grids with minimal observability. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–6, June 2018.

[10] 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.

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

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

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

[14] X. Feng and B. Houska. Real-time algorithm for self-reflective model predictive control. *Journal of Process Control*, 65:68–77, 2018.

[15] J.J. Grainger and W.D. Stevenson. *Power system analysis*. McGraw-Hill series in electrical and computer engineering: Power and energy. McGraw-Hill, 1994.

[16] S. Frank and S. Rebennack. An introduction to optimal power flow: Theory, formulation, and examples. *IIE Transactions*, 48(12):1172–1197, 2016.

[17] J. Zhu. *Optimization of Power System Operation*, volume 47. John Wiley & Sons, 2015.

[18] F. Pukelsheim. *Optimal Design of Experiments*. John Wiley & Sons, Inc., New York, 1993.

[19] 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*, 23:613–629, 2013.

[20] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas. Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1):12–19, February 2011.

[21] A. J. Wood and B. F. Wollenberg. *Power generation, operation, and control*. John Wiley & Sons, 2013.

[22] F. Logist, M. Vallerio, B. Houska, M. Diehl, and J. Van Impe. Multi-objective optimal control of chemical processes using acado toolkit. *Computers and Chemical Engineering*, 37:191–199, 2012.