

Distributed learning for optimal allocation in radial power systems

Citation for published version (APA):

Jouini, T., & Sun, Z. (2020). Distributed learning for optimal allocation in radial power systems. *arXiv, 2020*, [2009.13857]. <https://arxiv.org/abs/2009.13857>

Document status and date:

Published: 29/09/2020

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Distributed learning for optimal allocation in radial power systems

Taouba Jouini¹, Zhiyong Sun²

Abstract—We revisit the classical log-linear learning algorithm for optimal allocation of DC/AC converters and synchronous machines in radial power systems. The objective is to assign to each generator node a type; either a synchronous machine or a DC/AC converter in closed-loop with droop control, while minimizing the steady state angle deviation relative to an optimum associated with unknown optimal configuration of synchronous machines and DC/AC converters. Additionally, we study the robustness of the learning algorithm against uniform drop in the line susceptances and with respect to well-defined feasibility region describing admissible power deviations. We show guaranteed probabilistic convergence to maximizers of the perturbed potential function with feasible power flows and demonstrate our theoretical findings via simulative example of power network with six generation units.

I. INTRODUCTION

With increased share of renewable energy resources (wind, solar, fuel cells, etc.) in the electrical grid, it is of an uttermost importance to understand the ramification of heterogeneous power generation on the operation and maintenance of normal grid conditions. Since a large percentage of generation in power system will not be based on synchronous machines, the dominant dynamics justified by their presence, is not valid anymore [1]. This sheds new lights into the ongoing research to enhance the grid performance for efficient power generation using tools from optimization theory [2].

In particular, the field of game theory has gained more and more attention over the years as it intersects several disciplines. Game theory offers a rich set of model elements, solution concepts, and evolutionary notions. Within the realm of engineering systems, a key element in the use of game theory is to design incentives to obtain desirable behaviors [3], [4]. A game theoretical model is more than just a dynamic feedback system, as each player learns the environment, which in turn learns the player [4].

Different game theoretic concepts have previously been applied to formulate and solve optimization problems in power systems. For example, in [5], a feedback controller is proposed based on population games to achieve frequency regulation and economic efficiency. The authors of [6], [7] design distributed control laws for decision process of individual sources in small-scale DC power systems and propose a proportional allocation mechanism using non-cooperative

game theory, respectively. In [8], coalitional game approach allows consumers to cooperatively share storage with each other, while minimizing electricity consumption cost. Moreover, game theory has been leveraged for pricing and market mechanisms in power systems [9]–[11] as well as security assessment and mitigation of attacks [12], [13].

Learning in game theory provides a framework for designing, analyzing and controlling multi-agent systems [14] and is well-understood in the literature with guaranteed asymptotic results, for example, convergence to a Nash equilibrium. Learning algorithms have been extensively studied in potential games [15], [16] and are designed with the goal to implement a prescriptive control approach, where the guaranteed limiting behavior represents a desirable operating condition [17]. In particular, log-linear learning originally introduced in [18] is a learning algorithm that ensures that the action profiles that maximize the global objective of the multi-agent system coincide with the potential function maximizers. The inclusion of the noise function enables the players to occasionally make mistakes corresponding to sub-optimal actions. As the noise vanishes, the probability that a player selects a best response or an optimal action goes to one [17].

In this work, our contributions are put together as follows: First, we consider the log-linear learning algorithm in its classical formulation [18] through new lenses by applying it to an optimal allocation problem for radial power systems, where the goal is to assign a type (DC/AC converter in closed-loop with droop control or synchronous machine) to each generation unit, by minimizing steady state angle deviations with respect to an optimum. Our problem is motivated by the planning and construction phase of the power grid that focuses on a rational structure of the power network and presents a particular configuration of the network that can meet the special needs of the future smart grid [19]. Second, we investigate the robustness of the learning algorithm against uniform drop in the line susceptances, representing uncertainty in the knowledge of their exact value, while steady state power deviations are confined to well-defined feasibility region. This is performed by determining an upper bound on the allowed line susceptance drop. Additionally, we show that the learning algorithm for optimal allocation converges, in the probabilistic sense, to an optimal configuration that corresponds to maximizers of the perturbed potential function. Third, we validate our findings on a power system setup consisting of six generation units arranged according to a line graph. We simulate the power system with unperturbed as well as perturbed weight susceptances and discuss the convergence to an unknown

*This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No: 834142).

¹Taouba Jouini is with the Department of Automatic Control, LTH, Lund University, Ole Rmners vg 1, 22363 Lund, Sweden. ²Zhiyong Sun is with Department of Electrical Engineering, Eindhoven University of Technology, the Netherlands. E-mails: taouba.jouini@control.lth.se, z.sun@tue.nl.

optimal configuration for a susceptance perturbation within the derived theoretical bound.

The remainder of this paper unfurls as follows. Section II introduces the power systems model, as well as game-theoretic setup and formulates optimal allocation problem for radial power systems. In Section III, we study the robustness of the learning algorithm with respect to uniform drop in the line susceptances, provide interpretation to our results and link these to other well-studied optimization problems. Finally, Section IV illustrates our results by providing numerical simulations of the learning algorithm for optimal allocation on a network consisting of aligned six- generation units.

II. LEARNING FOR OPTIMAL ALLOCATION IN POWER SYSTEMS

A. Modeling of power systems

We consider a power network model, defined by a graph $G = (\mathcal{V}, \mathcal{E})$ in radial (or acyclic) undirected network, where \mathcal{V} is the set of n (possibly) heterogeneous generation units. We consider inductive loads with constant susceptances, absorbed in the lines. Define $\{1, \dots, n\}$ as the index set of all the generation units in the power system. The voltage magnitude V_i at the i -th bus is assumed to be constant and equal to 1 per unit. Let \mathcal{E} be the set of m edges (purely inductive transmission lines) with susceptance weight $b_e > 0$, $e \in \mathcal{E}$. We denote by $\mathcal{I} \in \mathbb{R}^{n \times m}$ the incidence matrix of the graph G , and by \mathcal{N}_i the neighbor set of the i -th generation unit (synchronous machine [20], or DC/AC converter in closed-loop with the droop control [21]). Under the assumption of quasi-stationary steady state, the swing equation of the i -th synchronous machine with inertial constant $m_i > 0$, damping coefficient $d_i > 0$ describes the i -th generation unit dynamics as follows,

$$m_i \ddot{\theta}_i(t) = -d_i \dot{\theta}_i(t) - \sum_{j \in \mathcal{N}_i} b_{ij} \sin(\theta_i(t) - \theta_j(t)) + P_{0,i}, \quad (1)$$

where $\theta_i(t) \in \mathbb{R}$ denotes the (virtual) voltage phase angle, $P_{0,i} \in \mathbb{R}$ is a constant power input that represents mechanical or DC side power input and $P_i = \sum_{j \in \mathcal{N}_i} b_{ij} \sin(\theta_i(t) - \theta_j(t))$ is the electrical power injected from the i -th generation into the neighbor set \mathcal{N}_i .

The dynamics in (1) describes both DC/AC converters that are equipped with droop control, or the rotor dynamics of synchronous machines, where the difference between the two models lies in the values of the damping d_i and inertia m_i .

In the sequel, we denote the phase angles of the generation units at steady state by $\theta^s = [\theta_1^s, \dots, \theta_n^s]^\top$, induced by configuration (or arrangement) of the generation units $s = [s_1, \dots, s_n]^\top$, where $s_i \in \mathcal{S} = \{M, C\}$ is the type (synchronous machine (M), or DC/AC converter (C)) of the i -th generation unit. The steady state angles are described by the following equation,

$$d_i \dot{\theta}_i^s = - \sum_{j \in \mathcal{N}_i} b_{ij} \sin(\theta_{ij}^s) + P_{0,i}, \quad \theta_{ij}^s = \theta_i^s - \theta_j^s. \quad (2)$$

After defining $P_0 = [P_{0,1}, \dots, P_{0,n}]^\top$, $D = \text{diag}\{d_i\}_{i=1 \dots n}$ and $\mathbf{1}_n$ as the vector of all ones, we make the following assumption.

Assumption 1 (Flow feasibility [21]). *Consider the power system model at steady state in (2). We assume that*

$$\|\text{diag}(\{b_{ij}\}_{(i,j) \in \mathcal{E}})^{-1} \cdot \xi^s\|_\infty < 1,$$

for all configurations s , where $\xi^s \in \mathbb{R}^m$ is the vector of edge flows satisfying,

$$P_0 - \omega_0 D \mathbf{1}_n = \mathcal{I} \xi^s, \quad (3)$$

and $\omega_0 = \sum_{i=1}^n P_{0,i} / \sum_{i=1}^n d_i$.

For a given configuration $s = [s_1, \dots, s_n]^\top$ through the damping matrix D , the steady state angles θ_{ij}^s are determined from,

$$\underline{\sin}(\mathcal{I}^\top \theta^s) = \text{diag}(\{b_{ij}\}_{(i,j) \in \mathcal{E}})^{-1} \xi^s, \quad (4)$$

with $\underline{\sin}(z) = [\sin(z_1), \dots, \sin(z_m)]^\top$ and the edge flow vector ξ^s is defined in (3).

It is important to mention that, the damping parameter of synchronous machines is considered to have higher value, than that of DC/AC converters in closed loop with droop control and enters the power balance equation (3) through the damping matrix D . This establishes the link between the chosen configuration $s = [s_1, \dots, s_n]$ and the resulting relative steady state angles θ_{ij}^s . Thus, the steady state angles depend on the type of the i -th generation unit s_i , through the damping values d_i and are obtained from the n - decoupled equations in (4).

Assumption 1 is equivalent to the existence of unique and synchronized steady state angles $\theta^s = [\theta_1^s, \dots, \theta_n^s]$, that are phase cohesive, i.e., $\theta_{ij}^s < \gamma$, $\gamma \in [0, \pi/2[$ satisfying (4). For more details about the derivation of (3) and (4), we refer the interested reader to Theorem 2 in [21].

Let $\theta_0 = [\theta_{01}, \dots, \theta_{0n}]^\top$ be the initial angle vector and $\theta^{s^*} = \omega_0 \mathbf{1}_n t + \theta_0$ denote the optimal phase angles $\theta^{s^*} = [\theta_1^{s^*}, \dots, \theta_n^{s^*}]^\top$ corresponding to an optimal, yet unknown configuration $[s_1^*, \dots, s_n^*]$, $s_i^* \in \mathcal{S} = \{M, C\}$. The optimal angles θ^{s^*} can be derived, from solving an optimal power flow [22] for the underlying power system model by setting $\underline{\sin}(\mathcal{I}^\top \theta^{s^*}) = \text{diag}(\{b_{ij}\}_{(i,j) \in \mathcal{E}})^{-1} \xi^{s^*}$, where ξ^{s^*} satisfies (3), falls the optimal configuration s^* of DC/AC converters and synchronous machines encoded in the damping matrix D^* , is known.

In this work, we consider unknown optimal configuration s^* , and the optimal values of the phase angles θ^{s^*} can be obtained, for example, from historical records of the optimal power system operation.

B. Game-theoretic setup

Consider a game $(\mathcal{V}, \mathcal{A}, \{u_i\}_{i \in \mathcal{V}})$, where \mathcal{V} is a finite set of players and \mathcal{A} is a set of actions. Each player $i \in \{1, \dots, n\}$ has a utility function u_i (a.k.a reward or payoff function), which associates with every action $x \in \mathcal{X} = \mathcal{A}^\mathcal{V}$ the utility $u_i(x)$ that player i gets, when every other player $j \in \mathcal{N}_i$ is

playing action $x_j \in \mathcal{A}$. Let $x_{-i} = (x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ denote the profile of all players' actions, other than player i .

Definition II.1 (Potential game, [3]). A game $(\mathcal{V}, \mathcal{A}, \{u_i\}_{i \in \mathcal{V}})$ is called a potential game, if there exists a function $U : \mathcal{X} \rightarrow \mathbb{R}$ (referred to as the potential function of the game), such that for any two configurations $x, y \in \mathcal{X}$ and a player $i \in \mathcal{V}$, we have that,

$$u_i(y_i, x_{-i}) - u_i(x_i, x_{-i}) = U(y_i, x_{-i}) - U(x_i, x_{-i}).$$

A potential game as defined above requires perfect alignment between the global objective and the players local objective functions, so that the change in a player's utility that results from a unilateral change in strategy equals the change in the global utility [3].

Definition II.2 (Pure strategy Nash equilibrium). A (pure strategy) Nash equilibrium for the game $(\mathcal{V}, \mathcal{A}, \{u_i\}_{i \in \mathcal{V}})$ is an action configuration $x^* \in \mathcal{X}$, such that,

$$x_i^* \in \mathcal{B}_i(x_{-i}^*), i \in \mathcal{V},$$

where $\mathcal{B}_i(x_{-i}) = \operatorname{argmax}_{x_i \in \mathcal{A}} u_i(x_i, x_{-i})$ is the best response function.

A pure Nash equilibrium as given in Definition II.2, represents a configuration, in which no player has a unilateral incentive to deviate from his current action.

Definition II.3 (Noisy best response). Consider a game $(\mathcal{V}, \mathcal{A}, \{u_i\}_{i \in \mathcal{V}})$. The continuous-time asynchronous noisy best response dynamics is a Markov chain $X(t)$ with state space $\mathcal{X} = \mathcal{A}^{\mathcal{V}}$, where each player is equipped with an independent rate-1 Poisson clock. If the clock ticks at time t , the player i updates his actions to s , chosen from a conditional probability:

$$P(X_i(t+1) = s | X_i(t)) = \frac{e^{1/\tau(t) \cdot u_i(s, X_{-i})}}{\sum_{X_i \in \mathcal{A}_i} e^{1/\tau(t) \cdot u_i(X_i, X_{-i})}}, \quad (5)$$

where $\tau(t) > 0$ is the temperature or noise function that controls the smoothness of (5) and is a decreasing function of time [17].

The learning algorithm described in Definition II.3 is known as *log-linear* algorithm, and is well studied in game theory and classically described by the following rules [17]:

- Players' utility functions constitute a potential game.
- Players update their strategies one at a time, which is referred to as *asynchrony*.
- At any stage, a player can select any action in the action set.
- Each player assesses the utility for alternative actions, assuming that the actions of all other players remain fixed.

C. Optimal allocation problem

In this section, our goal is to assign a type $s_i \in \mathcal{S} = \{M, C\}$ to each generation unit $i = 1 \dots n$, represented by a Markov chain $X(t) = [X_1(t), \dots, X_n(t)]^\top$ with $X_i(t) = s_i \in \mathcal{S}$ being the generation unit type of player i at time t . This assignment

is performed so that, the phase angles at each unit are the closest to their given optimal relative angles θ_{ij}^* , derived from an unknown optimal configuration s^* of the converters and synchronous machines. In other words, we aim to find a valid generation unit assignment $s = [s_1, \dots, s_n]^\top$ by finding $X = [X_1, \dots, X_n]^\top$, so as to minimize the pairwise interaction cost,

$$c(X_i(t) = s_i, X_j(t)) = |\sin(\theta_{ij}^s) - \sin(\theta_{ij}^{s*})|, \quad (6)$$

where $\theta_{ij}^s = \theta_i^s - \theta_j^s$ and $\theta_{ij}^{s*} = \theta_i^{s*} - \theta_j^{s*}$ are steady state angle relative differences, resulting from choosing the type s_i for the i -th generation unit, given the generation types of the neighboring units $s_j, j \neq i$.

The cost function $c(X_i, X_j)$ in (6) depends implicitly on the configuration $X = s, s_i \in \mathcal{S}, i = 1 \dots n$, through (3) and (4).

At every time instance t , one generation unit $g(t) \in \{1 \dots n\}$ chosen uniformly at random wakes up and updates its type $s_i \in \mathcal{S}$. The conditional probability of updating the i -th generation unit to type s_i is given by,

$$P(X_i(t+1) = s_i | X(t), g(t) = i) = \frac{e^{-1/\tau(t) \sum_{j \in \mathcal{N}_i} b_{ij} c(s_i, X_j(t))}}{\sum_{k \in \mathcal{S}} e^{-1/\tau(t) \sum_{j \in \mathcal{N}_i} b_{ij} c(k, X_j(t))}}, \quad (7)$$

where the function $\tau(t)$ is defined as in (5).

Note that $\tau(t)$ determines how likely is player i to select a sub-optimal action: As $\tau(t) \rightarrow \infty$, player i will select any unit type X_i with equal probability. As $\tau(t) \rightarrow 0$, player i will select a best response $X_i \in \mathcal{B}_i(X_{-i})$.

We assign to each player an objective function that captures the player's marginal contribution to the potential function. This translates to assigning to each player the following utility function

$$u_i(X_i, X_{-i}) = - \sum_{j \in \mathcal{N}_i} b_{ij} c(X_i, X_j). \quad (8)$$

Next, we define the potential function given by,

$$U(X) = -\frac{1}{2} \sum_{\substack{i, j \in \mathcal{V}, \\ (i, j) \in \mathcal{E}}} b_{ij} c(X_i(t), X_j(t)). \quad (9)$$

Following Assumption 1, the aforementioned potential function (9) achieves zero if and only if all generation units are aligned with an optimal configuration $s^* = [s_1^*, \dots, s_n^*]^\top$ of converters and synchronous machines. This can be seen from,

$$|\sin(\theta_{ij}^s) - \sin(\theta_{ij}^{s*})| = 0,$$

if and only if $\theta_{ij}^s = \theta_{ij}^{s*}$, for all $i = 1, \dots, n$ (due to phase angle cohesiveness). In fact, given optimal angles θ^{s*} , we can write (3) as

$$D \mathbb{1}_n = \frac{1}{\omega_0} \left[P_0 - \mathcal{I} \operatorname{diag}(\{b_{ij}\}) \sin(\mathcal{I}^\top \theta^{s*}) \right]$$

and deduce that $D = D^*$ and hence $s = s^*$.

Note that an optimal configuration $s^* = [s_1^*, \dots, s_n^*]^\top$ is a Nash equilibrium of the game, given by the utility function (8), because an optimal configuration maximizes the potential function (9). However, a Nash equilibrium $X^* = [X_1^*, \dots, X_n^*]^\top$ can be sub-optimal, i.e., $U(X^*) < 0$, and hence may fail to correspond to an optimal configuration.

In a repeated potential game, the stationary distribution $p_X(t) = P(X(t) = X)$ is given by [18],

$$p_X(t) = \frac{e^{1/\tau(t) \cdot U(X)}}{\sum_{X \in \mathcal{X}} e^{1/\tau(t) \cdot U(X)}}.$$

Here again, the greater the temperature as $\tau(t) \rightarrow \infty$, the closer the conditional probability (5) to a uniform distribution over the player action response. As $\tau(t) \rightarrow 0$, the only stochastically stable states (see Definition 4, in [23]) of the Markov process are the joint actions that maximize the potential function. This shows the probabilistic convergence of the log-linear learning algorithm, and hence that of the learning algorithm for optimal allocation, to a Nash equilibrium that maximizes the potential function [17], [24].

III. ROBUSTNESS OF OPTIMAL ALLOCATION PROBLEM

A. Drop in transmission line susceptances

In this section, we investigate the implication of an additive unknown perturbation in transmission line susceptances on the probabilistic convergence of the log-linear algorithm and its robustness with respect to a well-defined feasibility region of admissible power flows.

Under these settings, we define the utility function for the i -th generation unit as follows,

$$\hat{u}_i(X) = - \sum_{j \in \mathcal{N}_i} \hat{b}_{ij}(\delta) c(X_i(t), X_j(t)), \quad (10)$$

where $\hat{b}_{ij}(\delta) = b_{ij} - \delta$, $0 \leq \delta < b_{ij}$, as well as the perturbed potential function,

$$\hat{U}(X) = - \frac{1}{2} \sum_{\substack{i,j \in \mathcal{V}, \\ (i,j) \in \mathcal{E}}} \hat{b}_{ij}(\delta) c(X_i(t), X_j(t)). \quad (11)$$

Hence, δ represents a uniform drop in the line susceptances so that the perturbed weights are still positive [25]. A drop in the susceptance value can model uncertainty in the knowledge about the line susceptance, or also a uniform decrease in the susceptance value, due to the presence of identical capacitive load absorbed in the lines.

Note that for given optimal relative angles in the cost (6), satisfying (3),(4) with perturbed susceptances \hat{b}_{ij} , it holds that $\max_{X_i} u_i(X) = \max_{X_i} \hat{u}_i(X)$, for all $X_i \in \mathcal{B}_i(X_{-i})$.

In the sequel, we assume that Assumption 1 holds with perturbed weights, for all $0 \leq \delta < b_{ij}$, $(i, j) \in \mathcal{E}$ and configurations s , i.e.,

$$\|\text{diag}(\hat{b}_{ij})^{-1} \hat{\xi}^s\|_\infty < 1,$$

where $\hat{\xi}^s$ represent edge flows that solve the power balance equations (4) with the perturbed weights $\hat{b}_{ij} > 0$.

Let $\hat{\theta}_{ij}^s$ denote the relative steady state angles resulting from solving for the power balance equations (3), (4) with

the perturbed line susceptances \hat{b}_{ij} , for all $i, j \in \mathcal{V}$. Next, we introduce a set of steady state angles, characterizing a feasibility region described by the maximal steady state power injected by each of the converters, relative to their optimal power output, i.e., corresponding to an optimal configuration s^* of the unperturbed susceptances b_{ij} , as follows,

$$\mathcal{P}^\alpha(\hat{\theta}^s) = \left\{ \hat{\theta}^s \in \mathbb{R}^n, \max_{i=1, \dots, n} \left| \sum_{j \in \mathcal{N}_i} \hat{b}_{ij}(\delta) \sin(\hat{\theta}_{ij}^s) - \sum_{j \in \mathcal{N}_i} b_{ij} \sin(\theta_{ij}^{s^*}) \right| < \alpha \right\}, \quad (12)$$

where $\sum_{j \in \mathcal{N}_i} \sin(\hat{\theta}_{ij}^s) > 0$ for all s (by phase angle cohesiveness of $\hat{\theta}_{ij}^s$). We choose $\alpha > 0$, so that,

$\max_{i=1, \dots, n} \left| \sum_{j \in \mathcal{N}_i} b_{ij} \left(\sin(\hat{\theta}_{ij}^s) - \sin(\theta_{ij}^{s^*}) \right) \right| < \alpha$, for all s . Thus, the steady state angles resulting from unperturbed susceptances ($\delta = 0$) pertain to the feasibility region (12). In fact, for all feasible flows (satisfying Assumption 1) the region \mathcal{P}^α in (12) can also be expressed in terms of the perturbed edge flows as $\{\hat{\xi}^s \in \mathbb{R}^m, \|\mathcal{I}(\hat{\xi}^s - \xi^{s^*})\|_\infty < \alpha\}$.

Definition III.1. *The optimal allocation robustness margin is defined by,*

$$\bar{R} = \inf\{|\delta|, \hat{\theta}^s \notin \mathcal{P}^\alpha(\hat{\theta}^s)\}. \quad (13)$$

This definition implies that the robustness margin is defined by the smallest susceptance drop δ , that steers the steady state angles $\hat{\theta}^s$, outside of their feasibility region \mathcal{P}^α .

At this stage, we ask two fundamental questions:

- For $\tau(t) \rightarrow 0$, does the distributed learning algorithm for optimal allocation by means of the perturbed utility function (10), converge in probability towards maximizers of the perturbed potential function (11)?
- Can we identify the robustness margin \bar{R}^s and determine its dependence on network parameters/configuration?

We provide answers to these questions in the following theorem.

Theorem III.2. *Consider the power system at steady state (2) with perturbed weights $\hat{b}_{ij} = b_{ij} - \delta$ and the utility function (10). Let,*

$$\bar{R}^s = \min_{i=1, \dots, n} \frac{\alpha + \sum_{j \in \mathcal{N}_i} \left(b_{ij} (\sin(\hat{\theta}_{ij}^s) - \sin(\theta_{ij}^{s^*})) \right)}{\sum_{j \in \mathcal{N}_i} \sin(\hat{\theta}_{ij}^s)}. \quad (14)$$

Additionally, consider the cost function (6) with optimal relative angles, corresponding to unknown configuration of the power system with perturbed weights \hat{b}_{ij} . Then, the following holds:

- 1) *The distributed learning algorithm converges as $\tau(t) \rightarrow 0$, to the uniform probability over the set of best responses, which are maximizers of the perturbed potential function (11).*
- 2) *For every configuration s , the power system model is robust against perturbations in the susceptances with*

a robustness margin given by \bar{R}^s in (14), if and only if $\delta < \bar{R}^s$.

Proof. Consider the cost function (6) with optimal steady state angles, that solve (3) and (4) with the perturbed weights $\hat{b}_{ij} = b_{ij} - \delta$. To prove the first statement, we follow the same lines of the proof of Proposition 3.1 from [17]. For this, we adopt the proof of Lemma 3.1 to our setup as follows: The perturbed learning algorithm for optimal allocation induces a finite, irreducible, aperiodic process over the state space. By introducing $\varepsilon = e^{-1/\tau}$, we denote by P^ε the corresponding transition matrix.

By using the perturbed utility function in (10), we have

$$\lim_{\varepsilon \rightarrow 0} \frac{P_{X \rightarrow X'}^\varepsilon}{\varepsilon^{\max_{X_i} u_i(X) - \hat{u}_i(X')}} = \frac{1}{n} \frac{1}{\sum_{X_i \in \mathcal{A}_i} \varepsilon^{\max_{X_i} u_i(X) - \hat{u}_i(X')}} < \infty, \quad (15)$$

with $X = (X_i, X_{-i})$, $X' = (X'_i, X_{-i})$, if and only if, $\max_{X_i} u_i(X) - \hat{u}_i(X) \geq 0$. This means that, for all $X_i \notin \mathcal{B}_i(X_{-i})$, we have

$$\frac{u_i(X) - \max_{X_i} u_i(X)}{\sum_{j \in \mathcal{N}_i} c(X_i, X_j)} < \delta, \quad (16)$$

and taking the supremum, yields the lower bound,

$$\underline{R} = \max_{i=1, \dots, n} \frac{u_i(X) - \max_{X_i} u_i(X)}{\sum_{j \in \mathcal{N}_i} c_i(X_i, X_j)} < 0. \quad (17)$$

Since $0 \leq \delta$, this shows that $\underline{R} < \delta$. Hence, Lemma 3.1 holds with the resistance $r(X \rightarrow X') = \max_{X_i} u_i(X) - \hat{u}_i(X')$. Together with Lemma 3.2, Proposition 3.1 in [17] implies that stochastically stable states are the set of perturbed potential function maximizers given in (11).

To prove the second claim, we note that $\hat{\theta}^s \notin \mathcal{P}^\alpha$ in (12), if and only if,

$$\left| \sum_{j \in \mathcal{N}_i} \hat{b}_{ij}(\delta) \sin(\hat{\theta}_{ij}^s) - \sum_{j \in \mathcal{N}_i} b_{ij} \sin(\theta_{ij}^{s*}) \right| \geq \alpha,$$

for some $i = 1 \dots n$. From $\hat{b}_{ij} = b_{ij} - \delta$, we solve for δ , the resulting inequality,

$$\left| \sum_{j \in \mathcal{N}_i} b_{ij} (\sin(\hat{\theta}_{ij}^s) - \sin(\theta_{ij}^{s*})) - \delta \sum_{j \in \mathcal{N}_i} \sin(\hat{\theta}_{ij}^s) \right| \geq \alpha.$$

This shows that,

$$\delta \geq \frac{\alpha + \sum_{j \in \mathcal{N}_i} [b_{ij} (\sin(\hat{\theta}_{ij}^s) - \sin(\theta_{ij}^{s*}))]}{\sum_{j \in \mathcal{N}_i} \sin(\hat{\theta}_{ij}^s)},$$

or

$$\delta \leq \frac{-\alpha + \sum_{j \in \mathcal{N}_i} [b_{ij} (\sin(\hat{\theta}_{ij}^s) - \sin(\theta_{ij}^{s*}))]}{\sum_{j \in \mathcal{N}_i} \sin(\hat{\theta}_{ij}^s)}.$$

By choice of α and definition of the robustness margin in (13), only the first inequality holds and we find (14) with $\bar{R}^s > 0$. \square

B. Discussion

We make the following remarks:

- 1) Theorem 1 establishes a robustness margin \bar{R}^s for the radial power system with respect to a uniform drop in the susceptance. The robustness margin \bar{R}^s depends on the configuration s and on admissible maximal deviation of steady state electrical power from its optimal value described in (12) and increases with larger α . Note that, optimal steady state angles of the power system with perturbed susceptances are in general unknown, because they require exact knowledge of δ . In this case, the learning algorithm with perturbed susceptances does not converge to a maximizer of $\hat{U}(t)$. See later Section IV.
- 2) A suitable choice for the temperature or noise function $\tau(t)$, as a decreasing function of time, is crucial for the improvement of the algorithm convergence properties. The function $\tau(t)$ shapes the learning parameter or exploration rate of the learning algorithm [23].
- 3) It is worth noting that the convergence in unperturbed (see Section II-B) as well as perturbed case (in Theorem III.2) is understood only in a probabilistic sense, that is, a Markov chain $X(t)$ itself can converge to a Nash equilibrium that is *not a maximizer* of the potential function $U(X)$. Hence, the potential function $U(X)$ is not always necessarily zero at the end of the learning algorithm.
- 4) The optimal allocation problem can be regarded in a broader perspective as a resource allocation problem [3], where the cost function (6) represents the welfare function. This can also be formulated as a task assignment [26], [27] for dynamic multi-agent networks, where each node selects its task, i.e., type from an admissible set and the cost function is interpreted as utility of an assignment profile for each node with a specific role. An illustrative example of the distributed learning in potential games is graph coloring, whose goal is to find a color assignment, such that none of two neighboring nodes have the same color. One could interpret the type set $\mathcal{S} = \{M, C\}$ as the colors with difference being in the cost function typically used for coloring algorithm that explicitly depend on the current configuration, see Section 6.1 in [3].

IV. SIMULATIONS

To illustrate the distributed learning for optimal allocation algorithm, we study an example of power systems represented by a line graph of six generation units, associated with power input vector $P_0 = [0.77778, 0.7, 0.798889, 0.7, 0.798889, 0.7]^\top$. The goal is to assign a type $s_i \in \mathcal{S}$ to each generation unit $i = 1, \dots, 6$. The optimal but unknown configuration s^* shown in Figure 1 induces given optimal steady state angles, which are minimizers of the potential function $U(t)$ in (9). The optimal relative angle differences $\{\theta_{ij}^{s*}\}_{i,j \in \mathcal{V}}$ are given by $[-0.0157, -0.0354, 0.0081, 0.0750]^\top$.

The distributed learning algorithm for optimal allocation aims to correctly assign a type (synchronous machine or DC/AC converter in closed-loop with droop control) and hence to find the optimal configuration s^* , based on given optimal angle differences $\{\theta_{ij}^*\}_{i,j \in \mathcal{V}}$.

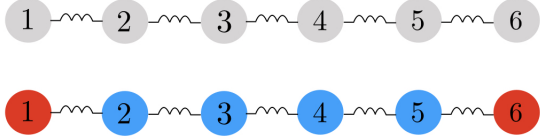


Fig. 1. Top: Power system represented by a line graph consisting of 6 generation units connected via inductive lines with susceptances $b_{ij} > 0$. Bottom: Unknown optimal configuration $s^* = [M, C, C, C, C, M]$ corresponding to given optimal relative angles $\{\theta_{ij}^*\}_{i,j \in \mathcal{V}}$. The color *red* refers to a synchronous machine and *blue* to DC/AC converters (in closed-loop with droop control).

For simplicity of presentation, we choose between two colors: *red* for synchronous machines (M) and *blue* for DC/AC converters (C) with respective damping values (in p.u.) $d_C = 15$ and $d_M = 25$. The susceptances $\{b_{ij}\}_{(i,j) \in \mathcal{E}}$ are in p.u. and given by the matrix $\Gamma = \text{diag}\{[15.2631, 4.2350, 4.8156, 15.2631, 4.2350]\}$. The values are taken from the IEEE 14-bus system [28].

We begin with the unperturbed ($\delta = 0$) cost function (6). At initialization, all generation units are assumed to be synchronous machines, so that $X_i(t) = M, i = 1, \dots, n$. At each time $t > 0$, a generation unit $i \in U(t)$ is chosen uniformly (at random) and updates its type s_i (and hence color), according to the conditional probability (7). In Figure 2, we plot the time evolution of the potential function in (9), for two different realizations of the inverse of the temperature function $\eta(t) = \tau^{-1}(t)$. We notice that an increase in the slope of $\eta(t)$ is accompanied by an increase in the convergence rate to an optimal configuration.

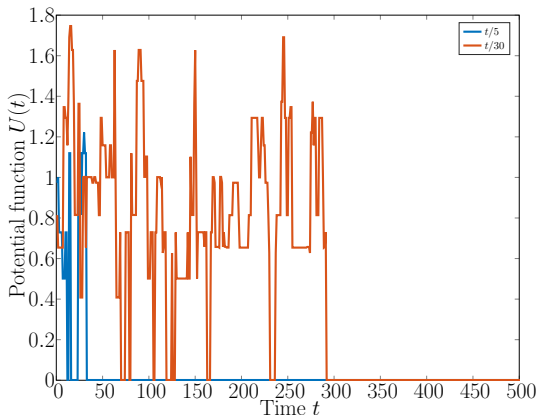


Fig. 2. Evolution of the potential function $U(t)$ in (9) corresponding to two different realizations of the inverse of the temperature function $\eta_1(t) = t/5$ and $\eta_2(t) = t/30$.

In the sequel, we consider the distributed learning algo-

rithm for optimal allocation with $\eta(t) = t/5$. We calculate

$$\max_{s_i \in \mathcal{S}} \max_{i=1 \dots n} \left| \sum_{j \in \mathcal{N}_i} b_{ij} (\sin(\theta_{ij}^s) - \sin(\theta_{ij}^*)) \right| = 0.8142,$$

and hence pick $\alpha = 1.5 > 0.8142$, so that $\theta^s \in \mathcal{P}^\alpha$ in (12), for all configurations s . This choice accounts for a large margin of admissible power deviations induced by bounded disturbances that might affect power system operation [20].

Next, we perturb the susceptance values given by $\hat{b}_{ij} = b_{ij} - \delta$, where $0 \leq \delta < b_{ij}$ is a uniformly randomly generated perturbation. The upper bound on the susceptance drop δ from Theorem III.2 is given by $K = \min_{\substack{s, s_i \in \mathcal{S} \\ i=1, \dots, n}} \bar{R}^s = 0.4723$. In

this case, the resulting steady state angles remain inside the feasibility region \mathcal{P}^α , where for example for $\delta = 0.0027$, we have that,

$$\max_{\substack{s, s_i \in \mathcal{S} \\ i=1, \dots, n}} \max_{i=1 \dots n} \left| \sum_{j \in \mathcal{N}_i} \hat{b}_{ij}(\delta) \sin(\hat{\theta}_{ij}^s) - \sum_{j \in \mathcal{N}_i} b_{ij} \sin(\theta_{ij}^*) \right| = 0.8862.$$

Since the optimal configuration $s^* = [s_1^*, \dots, s_n^*]$ in Figure 1, corresponds to steady state angles (3), (4) with unperturbed line susceptances b_{ij} , Figure 3 shows the perturbed potential function (11) (for $\delta = 0.3065$) that does not converge to zero, but rather to a Nash equilibrium that is not a minimizer of the perturbed potential $\hat{U}(t)$ in (11).

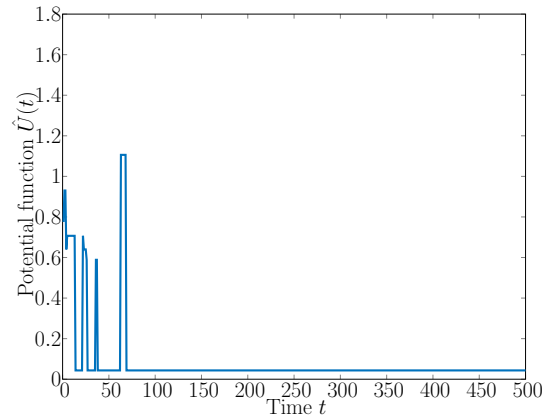


Fig. 3. Convergence of the perturbed potential function $\hat{U}(t)$ in (11) for $\delta = 0.3065$, to a non-zero value corresponding to a configuration of a Nash equilibrium, that is not a maximizer of the perturbed potential function.

V. CONCLUSION

We revisited a distributed learning algorithm for optimal allocation in radial power systems based on log-linear learning with guaranteed probabilistic convergence to Nash equilibrium. Moreover, we investigated its robustness against drops in the line susceptances with respect to feasible region of power deviations at steady state. We validated our results in simulations of line graph with six generation units. Future investigations involve generalization of the network topology and inclusion of more detailed models of synchronous machines and DC/AC converters in closed-loop with suitable control.

ACKNOWLEDGMENT

The authors would like to thank Prof. Emma Tegling for the insightful comments and discussions.

REFERENCES

- [1] M. Paolone, T. Gaunt, X. Guillaud, M. Liserre, S. Meliopoulos, A. Monti, T. Van Cutsem, V. Vittal, and C. Vournas, "Fundamentals of power systems modelling in the presence of converter-interfaced generation," *Electric Power Systems Research*, vol. 189, p. 106811, 2020.
- [2] T. Jouini and Z. Sun, "Performance analysis and optimization of power systems with spatially correlated noise," *IEEE Control Systems Letters*, vol. 5, no. 1, pp. 361–366, 2021.
- [3] J. R. Marden and A. Wierman, "Distributed welfare games," *Operations Research*, vol. 61, no. 1, pp. 155–168, 2013.
- [4] D. Bauso, *Game theory with engineering applications*. SIAM, 2016.
- [5] J. Barreiro-Gomez, F. Dörfler, and H. Tembine, "Distributed robust population games with applications to optimal frequency control in power systems," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 5762–5767.
- [6] P. Chakraborty, E. Baeyens, and P. P. Khargonekar, "Distributed control of flexible demand using proportional allocation mechanism in a smart grid: Game theoretic interaction and price of anarchy," *Sustainable Energy, Grids and Networks*, vol. 12, pp. 30–39, 2017.
- [7] W. W. Weaver and P. T. Krein, "Game-theoretic control of small-scale power systems," *IEEE Transactions on Power Delivery*, vol. 24, no. 3, pp. 1560–1567, 2009.
- [8] P. Chakraborty, E. Baeyens, K. Poolla, P. P. Khargonekar, and P. Varaiya, "Sharing storage in a smart grid: A coalitional game approach," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4379–4390, 2018.
- [9] Q. Xu, N. Zhang, C. Kang, Q. Xia, D. He, C. Liu, Y. Huang, L. Cheng, and J. Bai, "A game theoretical pricing mechanism for multi-area spinning reserve trading considering wind power uncertainty," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1084–1095, 2016.
- [10] B. K. Poolla, S. Bolognani, L. Na, and F. Dörfler, "A market mechanism for virtual inertia," *IEEE Transactions on Smart Grid*, 2020.
- [11] N. Forouzandehmehr, Z. Han, and R. Zheng, "Stochastic dynamic game between hydropower plant and thermal power plant in smart grid networks," *IEEE Systems Journal*, vol. 10, no. 1, pp. 88–96, 2014.
- [12] A. Farraj, E. Hammad, A. A. Daoud, and D. Kundur, "A game-theoretic analysis of cyber switching attacks and mitigation in smart grid systems," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1846–1855, 2016.
- [13] T. Alpcan and T. Başar, *Network security: A decision and game-theoretic approach*. Cambridge University Press, 2010.
- [14] D. Fudenberg, F. Drew, D. K. Levine, and D. K. Levine, *The theory of learning in games*. MIT press, 1998, vol. 2.
- [15] D. Monderer and L. S. Shapley, "Potential games," *Games and economic behavior*, vol. 14, no. 1, pp. 124–143, 1996.
- [16] J. R. Marden, G. Arslan, and J. S. Shamma, "Joint strategy fictitious play with inertia for potential games," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 208–220, 2009.
- [17] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation," *Games and Economic Behavior*, vol. 75, no. 2, pp. 788–808, 2012.
- [18] L. E. Blume *et al.*, "The statistical mechanics of strategic interaction," *Games and economic behavior*, vol. 5, no. 3, pp. 387–424, 1993.
- [19] R. Zhang, Y. Du, and L. Yuhong, "New challenges to power system planning and operation of smart grid development in china," in *2010 International Conference on Power System Technology*. IEEE, 2010, pp. 1–8.
- [20] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*. McGraw-hill New York, 1994, vol. 7.
- [21] J. W. Simpson-Porco, F. Dörfler, and F. Bullo, "Synchronization and power sharing for droop-controlled inverters in islanded microgrids," *Automatica*, vol. 49, no. 9, pp. 2603–2611, 2013.
- [22] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, 2011.
- [23] M. Hasanbeig and L. Pavel, "From game-theoretic multi-agent log linear learning to reinforcement learning," *arXiv preprint arXiv:1802.02277*, 2018.
- [24] J. R. Marden, G. Arslan, and J. S. Shamma, "Cooperative control and potential games," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 6, pp. 1393–1407, 2009.
- [25] T. Ding, R. Bo, Y. Yang, and F. Blaabjerg, "Impact of negative reactance on definiteness of b-matrix and feasibility of dc power flow," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1725–1734, 2017.
- [26] W. Gong, B. Zhang, and C. Li, "Task assignment in mobile crowd-sensing: Present and future directions," *IEEE Network*, vol. 32, no. 4, pp. 100–107, 2018.
- [27] G. Qu, D. Brown, and N. Li, "Distributed greedy algorithm for multi-agent task assignment problem with submodular utility functions," *Automatica*, vol. 105, pp. 206–215, 2019.
- [28] W. Group, "Common format for exchange of solved load flow data," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-92, no. 6, pp. 1916–1925, 1973.