Nondisturbing Extremum Seeking Control for Multiagent Industrial Systems

Mark Haring, Synne Fossøy, Thiago Lima Silva, and Alexey Pavlov, *Senior Member, IEEE*

***Abstract*—Industrial applications of extremum seeking control (ESC) can be a hit and miss affair. Although a gain in performance can be achieved, the dither applied to excite the system causes unwanted fluctuations in the performance of the system. The fluctuations in systems with a single extremum seeking loop are generally small. However, for systems with many extremum seeking loops, the fluctuations in each loop may add up to an intolerable amount of fluctuation in the total performance. In this article, we propose a method to cancel the dither-induced fluctuations in the overall system performance to a large extent by smartly constructing the dither signals in each extremum seeking loop using a centralized coordinator. The novelty of our method lies in the direct calculation of the dither signals that avoids the heavy computations required by other methods. Moreover, we provide a solvability analysis for the problem of cancelling dither-induced fluctuations in the total performance of the system. Furthermore, a complete stability analysis of the overall ESC scheme with dither coordination is given.**

***Index Terms*—Distributed systems, dither coordination, extremum seeking control (ESC), optimization, perturbation methods.**

I. Introduction

INDUSTRIAL applications of extremum seeking control (ESC) can be found in sectors that encounter a large level of uncertainty in daily operation. These uncertainties are often caused by inconsistencies in the composition of raw materials, such as in the process industry, or by changing environmental conditions, commonly encountered in the energy industry. Typical examples of industrial applications of ESC are power optimization of wind turbines [8], [12], [18] and photovoltaic arrays [7], [22], [23], oil well optimization [26], [29], and the operation of smelting furnaces [5]. For many of these applications, the system to be controlled can be divided into several subsystems, where each subsystem is locally controlled by ESC. Although controlled by individual local controllers, the subsystems may share input resources, output infrastructure, or system-wide constraints that must be honored while steering the control inputs toward their optimal values. Examples of distributed ESC for multiagent systems are presented in [9], [13], [27], [30], [39], and [40].

Although ESC comes in many forms (see, e.g., [3], [35], and [41] and references therein), if the uncertainties acting on the system are large and slowly time varying, dither is commonly used to track changes in the unknown optimal operating conditions. Too large or too fast dither signals in the input may be harmful to equipment, violate constraints, or cause unacceptably large fluctuations in the system's output. For individual systems controlled by ESC, dither is often chosen to be sufficiently small to not pose this problem, although a positive lower bound on the amount of dither commonly exists in practice due to the presence of measurement noise [14]. For large systems with many ESC loops, the combined effects and consequences of the introduced dither signals may be substantial, even if all dither amplitudes are small. For instance, for a multiagent system comprised of many parallel subsystems, the amplitude of the dither-induced fluctuation in the overall output of the system may be as large as the absolute sum of the fluctuation amplitudes of all individual subsystems [31]. Therefore, it can be challenging to choose suitable dither signals that are sufficiently large to ensure a proper tracking of changes in the optimum of each individual subsystem, and at the same time, are small enough to avoid losses and damages due to overly large fluctuations in the overall system output.

Previous works addressing dither signals concern mostly their shape, frequency, and amplitude. For instance, the use case in [36] presents results that show a relation between the size of dither amplitudes and the achieved convergence speed. Besides, the amplitudes should be large enough to satisfy the persistence of excitation requirement, and thus, ensure convergence to the optimum; see, e.g., [1] and [2]. When it comes to the dither frequencies, an important criterion is to select them such that time-scale separation is obtained between the dynamics of the system, the dither, and all other time scales of the controller;

Manuscript received 20 November 2021; accepted 10 February 2022. Date of publication 24 February 2022; date of current version 28 February 2023. This work was supported in part by the Research Council of Norway and in part by BRU21—NTNU's Research and Innovation Program in Digital and Automation Solutions for the Oil and Gas Industry. Recommended by Associate Editor Nathan van de Wouw. *(Corresponding author: Mark Haring).*

Mark Haring and Synne Fossøy are with the Department of Mathematics and Cybernetics, SINTEF Digital, 7037 Trondheim, Norway (e-mail: mark.haring@sintef.no; synne.fossoy@sintef.no).

Thiago Lima Silva is with the Department of Geoscience and Petroleum, Norwegian University of Science and Technology, 7491 Trondheim, Norway, and also with the Department of Sustainable Energy Technology, SINTEF Industry, 7031 Trondheim, Norway (e-mail: thiago.silva@sintef.no).

Alexey Pavlov is with the Department of Geoscience and Petroleum, Norwegian University of Science and Technology, 7031 Trondheim, Norway (e-mail: alexey.pavlov@ntnu.no).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TAC.2022.3153228.

Digital Object Identifier 10.1109/TAC.2022.3153228

see, e.g., [14], [20], and [33]. Performance improvement of multiagent systems can also be achieved by reusing agent frequencies that are not significantly coupled with regards to their steady-state performance; see, e.g., [21].

Preliminary ideas on cancelling dither-induced fluctuations in the total system input were reported in [27] for a multiagent resource allocation problem. In that article, the cancellation is achieved by coordinating phases of individual dither signals. Some form of minimizing dither signal amplitudes while ensuring an appropriate persistence of excitation condition was presented in [2]. However, that result does not guarantee cancellation of dither-induced fluctuations. The challenge of designing dither signals to cancel fluctuations both in the total input and total output of a multiagent system, while at the same time ensuring that each subsystem is sufficiently excited, was first treated in [31]. The solution presented in that article calculates dither signals by solving a nonconvex, yet computationally feasible optimization problem. The proposed algorithm implicitly coordinates both amplitudes and phases of individual dither signals to achieve minimization of dither-induced fluctuations both in the total input and output. This approach was then successfully applied to the case of a network of subsystems sharing a common constraint [32], with practical benefits demonstrated for the case of constrained ESC optimization for an oil production system.

This article focuses on minimizing dither-induced fluctuations in the total output of a multiagent system. It proposes an algorithm for computing the dither-signals for each subsystem's extremum seeking controller such that the fluctuations in the total output (defined as the sum of the outputs of individual subsystems) are as small as possible. Considerations to avoid rapid changes in dither signals, and to maintain a constant excitation level in each subsystem's input, make a complete cancellation in total output infeasible. In contrast to the method proposed in [31] and [32], we propose an algorithm for *direct calculation* of the dither signals, which is much lighter than the computationally heavy method proposed in these articles. Furthermore, the dither signal optimizer of [31] and [32] considers only sinusoidal dither signals, while the new proposed method holds for a larger class of dither signals. Finally, this article presents a solvability analysis for the problem of cancelling dither-induced fluctuations in the total output and a complete stability proof for each extremum seeking loop consisting of a subsystem, an extremum seeking controller, and the dither signal generated by the proposed algorithm.

This article is organized as follows. Preliminaries are given in Section II. The ESC problem is presented in Section III. Our method to solve the dither coordination problem is given in Section IV. In Section V, an example of a stabilizing extremum seeking controller is given, which is subsequently used in Section VI to show stability and convergence of the resulting extremum seeking loops. Section VII presents two use cases to illustrate our method. Finally, Section VIII concludes this article.

II. Preliminaries

The sets of real numbers, the set of nonnegative real numbers, and the set of positive real numbers are denoted by $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{R}_{>0}$, respectively. The sets of natural numbers (nonnegative integers) and positive integers are denoted by $\mathbb{N}$ and $\mathbb{N}_{>0}$. A scalar function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to belong to class $\mathcal{K}$ if it is continuous, strictly increasing and $\alpha(0) = 0$. A continuous function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to belong to the class $\mathcal{KL}$ if, for each fixed s, the mapping $\beta(r, s)$ belongs to the class $\mathcal{K}$ with respect to r, and for each fixed r, the mapping $\beta(r, s)$ is decreasing with respect to s, and $\beta(r, s) \to 0$ as $s \to \infty$ [19]. Let $\mathbf{x}$ be a vector in $\mathbb{R}^n$, where n is a positive integer. The transpose of the vector $\mathbf{x}$ is denoted by $\mathbf{x}^T$. The Euclidean norm of $\mathbf{x}$ is denoted by $\|\mathbf{x}\|$.

III. Extremum-Seeking Problem Formulation

Consider a system consisting of N parallel subsystems. We assume that there is no interaction between the subsystems, or at least that the interaction between the subsystems is negligibly small. The dynamics of each subsystem $i \in \{1, 2, \ldots, N\}$ are given by a function $\mathbf{f}_i : \mathbb{R}^{n_i} \times \mathcal{U}_i \to \mathbb{R}^{n_i}$ such that

$$\dot{\mathbf{x}}_i(t) = \mathbf{f}_i(\mathbf{x}_i(t), u_i(t)) \tag{1}$$

where $\mathbf{x}_i(t) \in \mathbb{R}^{n_i}$ is the state of the subsystem with dimension $n_i \in \mathbb{N}_{>0}$, $u_i(t) \in \mathcal{U}_i$ is a controllable input, and $t \in \mathbb{R}_{\geq 0}$ is the time. The set of possible input values $\mathcal{U}_i$ is assumed to be a closed interval defined by

$$\mathcal{U}_i = \left\{u_i \in \mathbb{R} : s_i^{\min} \leq u_i \leq s_i^{\max}\right\} \tag{2}$$

for some constants $s_i^{\min} < s_i^{\max}$. The measured output of each subsystem is given by

$$y_i(t) = h_i(\mathbf{x}_i(t)) + d_i(t) \tag{3}$$

for some function $h_i : \mathbb{R}^{n_i} \to \mathbb{R}$, where $d_i(t) \in \mathbb{R}$ is measurement noise. Due to limited knowledge about the system, we consider the state $\mathbf{x}_i(t)$ of each subsystem, the measurement noise $d_i(t)$, and the functions $\mathbf{f}_i$ in (1) and h_i in (3) to be unknown. Let us denote the true (noiseless) output of the subsystem by

$$z_i(t) = h_i(\mathbf{x}_i(t)). \tag{4}$$

The true output $z_i(t)$ is measure for the performance cost of the subsystem. The total performance cost of the system is defined as the sum of the true outputs of all subsystems

$$\check{z}(t) = \sum_{i=1}^{N} z_i(t). \tag{5}$$

To achieve optimal operating conditions under steady-state conditions, we are interested in minimizing the total steady-state performance cost. To clarify what we mean with the total steady-state performance cost, let us consider fixed input values $u_i(t) = u_i$. We assume the following.

Assumption 1: For each $i \in \{1, 2, \ldots, N\}$ and each fixed $u_i \in \mathcal{U}_i$, there exists a (constant) solution $\mathbf{X}_i(u_i) \in \mathbb{R}^{n_i}$ of the dynamics in (1) such that

$$\mathbf{0} = \mathbf{f}_i(\mathbf{X}_i(u_i), u_i). \tag{6}$$

In addition, we assume that the constant solution given by the mapping $\mathbf{X}_i : \mathcal{U}_i \to \mathbb{R}^{n_i}$ is locally attractive, uniformly on its domain as well as in time.

Assumption 2: The following holds for each $i \in \{1, 2, \ldots, N\}$. Let us define the closed ball

$$\mathcal{B}_i(u_i; r_i) = \{\mathbf{x}_i \in \mathbb{R}^{n_i} : \|\mathbf{x}_i - \mathbf{X}_i(u_i)\| \leq r_i\} \tag{7}$$

centered around $\mathbf{X}_i(u_i)$ with radius $r_i \in \mathbb{R}_{>0}$. There exists a radius $\rho_i^0 \in \mathbb{R}_{>0}$ and a function $\beta_{\mathbf{x}i} \in \mathcal{KL}$ such that

$$\|\mathbf{x}_i(t) - \mathbf{X}_i(u_i)\| \leq \beta_{\mathbf{x}i}(\|\mathbf{x}_i(0) - \mathbf{X}_i(u_i)\|, t) \tag{8}$$

for all $\mathbf{x}_i(0) \in \mathcal{B}_i(u_i; \rho_i^0)$, $u_i \in \mathcal{U}_i$, and $t \geq 0$.

Because all solutions of the subsystem that start sufficiently close to the constant solution $\mathbf{X}_i(u_i)$ eventually converge to it, we say that $\mathbf{X}_i(u_i)$ is the steady state of the subsystem. In correspondence to (4), the true steady-state output of the subsystem is given by

$$Z_i(u_i) = h_i(\mathbf{X}_i(u_i)). \tag{9}$$

Therefore, the total steady-state performance cost may be expressed as

$$\check{Z}(\mathbf{u}) = \sum_{i=1}^{N} Z_i(u_i) \tag{10}$$

with $\mathbf{u} = [u_1, u_2, \ldots, u_N]^T$. In order to minimize the total steady-state performance cost, we have to take into account all constraints on the system. As each controllable input u_i must remain within its feasible region $\mathcal{U}_i$ in (2), we may formulate the optimization problem as follows:

$$\min_{\mathbf{u} \in \mathcal{U}} \check{Z}(\mathbf{u}) \tag{11}$$

with

$$\mathcal{U} = \Big\{ [u_1, u_2, \ldots, u_N]^T \in \mathbb{R}^N : (\forall i \in \{1, 2, \ldots, N\})[s_i^{\min} \leq u_i \leq s_i^{\max}] \Big\}. \tag{12}$$

However, because the function h_i and the state $\mathbf{x}_i$ (as well as the steady-state $\mathbf{X}_i(u_i)$) of each subsystem are unknown, the cost function $\check{Z}$ in (10) is also unknown. To optimize the system's performance despite the limited knowledge about the system, we note that the optimizer of the optimization problem in (11) corresponds to the optimizers of the optimization problems

$$\min_{u_i \in \mathcal{U}_i} Z_i(u_i) \tag{13}$$

for each subsystem $i \in \{1, 2, \ldots, N\}$; see (10). Moreover, the output $y_i(t)$ in (3) is a reasonably accurate approximation of the steady-state performance cost $Z_i(u_i(t))$ in (9) for sufficiently low noise levels and near-steady-state conditions. Therefore, we may minimize the total steady-state performance cost by using N extremum seeking controllers to minimize the steady-state performance cost of each subsystem. A depiction of the system with N parallel control loops is given in Fig. 1.

Let the controllers that optimize the steady-state performance of each subsystem $i \in \{1, 2, \ldots, N\}$ be perturbation based. Although details may vary, the general idea behind a perturbation-based controller is to excite the subsystem by adding a dither signal to its input in order to retrieve sufficient information to approximate the first- and sometimes, second-order derivatives of its steady-state performance cost function Z_i. Subsequently, these derivative estimates are utilized to steer the nominal input of the subsystem toward its optimal value using a gradient-based [3], [35] or Newton-based [11], [25] optimizer. For simplicity, we limit ourselves to controllers that rely on gradient-based optimization in this work. A schematic representation of a gradient-based controller is given in Fig. 2, where $\hat{u}_i(t)$ denotes the nominal input of the subsystem and $p_i(t)$ denotes the dither signal. A detailed example of such a controller is given in Section V.

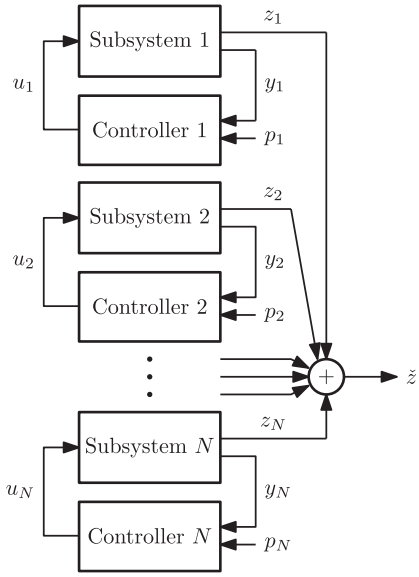

Fig. 1. System with N parallel subsystems with ESC loops.

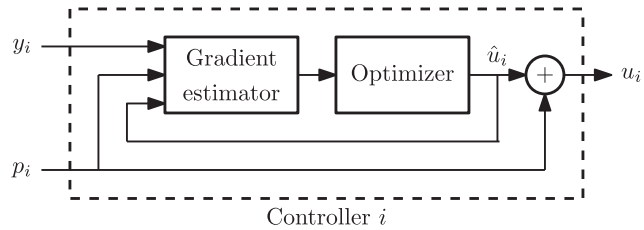

Fig. 2. Perturbation-based controller for the subsystem $i \in \{1, 2, \ldots, N\}$.

A. Dither-Induced Fluctuations

Although dither is often necessary to optimize the steady-state performance of the system, it introduces a fluctuation in the output of each subsystem. For subsystem $i \in \{1, 2, \ldots, N\}$, let the input $u_i(t)$ of the subsystem be the sum of its nominal value $\hat{u}_i(t)$ and its dither signal $p_i(t)$ as

$$u_i(t) = \hat{u}_i(t) + p_i(t). \tag{14}$$

Assuming that the subsystem is close to steady state and that the nominal input is much slower than the dither signal, we obtain from a first-order Taylor series approximation of Z_i in (9) that the subsystem's true output in (4) can be approximated by

$$z_i(t) \approx Z_i(\hat{u}_i(t)) + \frac{dZ_i}{du_i}(\hat{u}_i(t))p_i(t). \tag{15}$$

Here, $Z_i(\hat{u}_i(t))$ is the nominal value of the performance cost and $\frac{dZ_i}{du_i}(\hat{u}_i(t))p_i(t)$ is a dither-induced fluctuation. Because the total performance cost in (5) is the sum of all subsystem outputs,

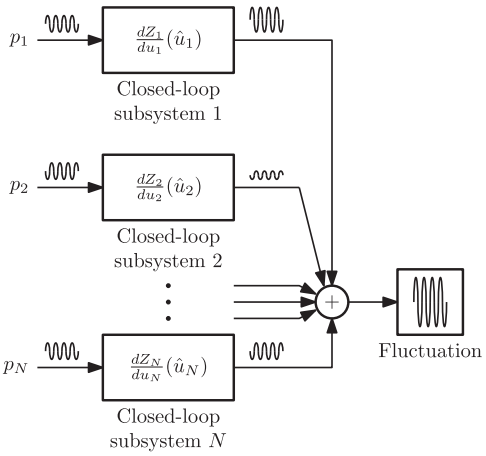

Fig. 3. Fluctuation in the total performance cost.

we obtain that

$$\check{z}(t) \approx \sum_{k=1}^{N} Z_k(\hat{u}_{\mathcal{I}(k)}(t)) + \sum_{i=1}^{N} \frac{dZ_i}{du_i}(\hat{u}_i(t))p_i(t) \tag{16}$$

where the first summation in the right-hand side of (16) is the nominal value of the total performance cost of the system, and the second summation is the corresponding fluctuation. Because the fluctuation of the total performance cost is the sum of the fluctuations of the outputs of all subsystems, we may have that the fluctuation of the total performance cost is large even if all fluctuations of the outputs of the subsystems are small. This situation is illustrated in Fig. 3. A large dither-induced fluctuation of the performance cost is generally undesirable as it may lead to damaged equipment, violated constraints, and large variations in production rates. A large performance fluctuation may rule out ESC as a feasible optimization method for industrial applications, especially for systems for which a steady production rate is critical, such as energy production systems.

Similarly to [31] and [32], we aim to minimize the fluctuation in the total performance cost by cleverly coordinating the dither signals of the subsystems such that the fluctuations of the outputs of the subsystems cancel each other out instead of amplifying each other. Our solution in the next section is simpler to compute than the computationally heavy solutions in [31] and [32]. Moreover, we consider a much broader class of dither signals than the single-frequency sinusoids used therein.

IV. Dither Signal Coordination

For any subsystem $i \in \{1, 2, \ldots, N\}$, we define the dither signal $p_i(t)$ in (14) as the linear sum of M time-varying basis functions, denoted by $b_j(\omega t)$, with $j \in \{1, 2, \ldots, M\}$ as

$$p_i(t) = \sum_{j=1}^{M} a_{ij}(t) b_j(\omega t) \tag{17}$$

where $a_{i,j}(t) \in \mathbb{R}$ represent the corresponding amplitudes for $i \in \{1, 2, \ldots, N\}$ and $j \in \{1, 2, \ldots, M\}$. By scaling the input of the basis functions by the positive parameter $\omega \in \mathbb{R}_{>0}$, we can select the frequency of the dither signal based on the value of ω. For simplicity, the basis functions are chosen to be periodic with common period $T \in \mathbb{R}_{>0}$ (for $\omega = 1$). In addition, they are designed such that

$$\frac{1}{T}\int_{t}^{t+T} b_j(\tau)d\tau = 0 \tag{18}$$

and

$$\frac{1}{T}\int_{t}^{t+T} b_j(\tau)b_k(\tau)d\tau = \begin{cases} 1, & \text{if } j = k \\ 0, & \text{if } j \neq k \end{cases} \tag{19}$$

for all $j, k \in \{1, 2, \ldots, M\}$ and all $t \geq 0$. Hence, the basis functions have zero mean and are orthogonal. Moreover, the basis functions and their first-order time derivatives are uniformly bounded. That means that there exist constants $\delta_b, \delta_{\dot{b}} \in \mathbb{R}_{>0}$, such that

$$|b_j(t)| \leq \delta_b \tag{20}$$

and

$$\left|\dot{b}_j(t)\right| \leq \delta_{\dot{b}} \tag{21}$$

for all $t \geq 0$ and all $j \in \{1, 2, \ldots, M\}$. As an example, the basis functions can be chosen as $b_1(\omega t) = \sqrt{2}\sin(\omega t), b_2(\omega t) = \sqrt{2}\cos(\omega t)$, $b_3(\omega t) = \sqrt{2}\sin(2\omega t)$, $b_4(\omega t) = \sqrt{2}\cos(2\omega t)$, $b_5(\omega t) = \sqrt{2}\sin(3\omega t)$, etc. Note that a wide class of dither signals can be captured by (17) given that M is sufficiently large.

Although the amplitudes $a_{i,j}(t)$ may change over time, we assume that this change is much slower than the variations in values of the basis functions; in the time scale of the basis functions, the amplitudes can be regarded as quasi-constant. As a measure for the magnitude of the fluctuation of the total performance cost [see (16)], we introduce

$$\gamma(t) = \sqrt{\frac{\omega}{T}\int_{t}^{t+\frac{T}{\omega}} \left(\sum_{i=1}^{N} \frac{dZ_i}{du_i}(\hat{u}_i(t)) \sum_{j=1}^{M} a_{ij}(t) b_j(\omega\tau)\right)^2 d\tau} \tag{22}$$

where we average over one period of the basis functions to obtain a measure that represents the average magnitude of the fluctuation instead of its instant value. Note that

$$\gamma(t) \approx \sqrt{\frac{\omega}{T}\int_{t}^{t+\frac{T}{\omega}} \left(\sum_{i=1}^{N} \frac{dZ_i}{du_i}(\hat{u}_i(\tau))p_i(\tau)\right)^2 d\tau} \tag{23}$$

if the amplitudes $a_{i,j}(t)$ and the gradients $\frac{dZ_i}{du_i}(\hat{u}_i(t))$ are slowly time varying with respect to the basis functions of the dither signals. Using the approximation in (23), $\gamma(t)$ is zero if the fluctuation of the total performance cost is zero, and positive otherwise. Utilizing the orthogonality property of the basis functions in (19), we obtain from (22) that

$$\gamma(t) = \sqrt{\sum_{j=1}^{M} \left(\sum_{i=1}^{N} \frac{dZ_i}{du_i}(\hat{u}_i(t))a_{ij}(t)\right)^2}. \tag{24}$$

For any gradients $\frac{dZ_i}{du_i}(\hat{u}_i(t))$, we aim to keep the fluctuation in the total performance cost at a minimum by selecting suitable amplitudes $a_{ij}(t)$ that result in a low value of the measure $\gamma(t)$. However, the selection of amplitudes should not have a profound effect on the optimization of the steady-state performance of

each subsystem $i \in \{1, 2, \ldots, N\}$. Motivated by the stability analysis in Section VI, the following three important requirements must be met when selecting the dither signals.

1) For a robust optimization process, there must exist a uniform positive lower bound on the level of excitation provided by the dither signals in order to estimate the gradient $\frac{dZ_i}{du_i}(\hat{u}_i(t))$ that is required for the optimization of the steady-state performance of each subsystem; see Section III.
2) All dither signals are required to be small; large dither signals result in inputs with large oscillations around their performance-optimal values, which may lead to a substantial performance loss on average.
3) The rate of change of all dither signals must be small for each subsystem to remain close to steady-state operation.

Despite these three requirements, there is a large freedom in selecting the amplitudes $a_{ij}(t)$ in (17). If the amplitudes of the dither signals are slowly time varying with respect to their basis functions, the level of excitation provided by the dither signals is given by

$$\frac{\omega}{T}\int_t^{t+\frac{T}{\omega}} p_i^2(\tau)d\tau \approx \sum_{j=1}^{M} a_{ij}^2(t) \tag{25}$$

for all $t \geq 0$ and all $i \in \{1, 2, \ldots, N\}$; see (17) and (19). To satisfy the first requirement, we fix the level of excitation of each dither signal as

$$\sum_{j=1}^{M} a_{ij}^2(t) = \nu^2 \tag{26}$$

for all $t \geq 0$ and $i \in \{1, 2, \ldots, N\}$, where $\nu \in \mathbb{R}_{>0}$ is a constant. Because (26) implies that $|a_{ij}(t)| \leq \nu$ for all $t \geq 0$, $i \in \{1, 2, \ldots, N\}$, and $j \in \{1, 2, \ldots, M\}$, and because all basis functions are uniformly bounded [see (20)], the second requirement is also satisfied by choosing a sufficiently small value of ν. From (24), it follows that $\gamma(t)$ is a function of the dither amplitudes $a_{ij}(t)$ and the gradients $\frac{dZ_i}{du_i}(\hat{u}_i(t))$. We aim to find mappings from the gradients $\frac{dZ_i}{du_i}(\hat{u}_i(t))$ to the amplitudes $a_{ij}(t)$ that lead to a small value of $\gamma(t)$ for each possible set of gradient values. For each $i \in \{1, 2, \ldots, N\}$ and each $j \in \{1, 2, \ldots, M\}$, let this mapping be denoted by $A_{ij} : \mathbb{R}^N \to \mathbb{R}$ as

$$a_{ij}(t) = A_{ij}\left(\mathbf{g}_Z(t)\right) \tag{27}$$

with

$$\mathbf{g}_Z(t) = \begin{bmatrix} \frac{dZ_1}{du_1}(\hat{u}_1(t)) & \frac{dZ_2}{du_2}(\hat{u}_2(t)) & \cdots & \frac{dZ_N}{du_N}(\hat{u}_N(t)) \end{bmatrix}. \tag{28}$$

We assume that each mapping is defined such that the fixed-excitation condition in (26) is satisfied for all possible gradient values. To satisfy the third requirement, we additionally assume that each mapping is globally Lipschitz so that

$$|\dot{a}_{ij}(t)| \leq L_A \left\|\dot{\mathbf{g}}_Z(t)\right\| \tag{29}$$

for some Lipschitz constant $L_A \in \mathbb{R}_{>0}$. In this case, the rate of change of all dither signals can be made arbitrarily small by tuning the extremum seeking controller of each subsystem such that the rate of change of all gradients is small, and by choosing a small value of ω such that the basis functions $b_j(\omega t)$ of the dither signals are slowly time varying; see (21). We note that the gradients $\frac{dZ_i}{du_i}(\hat{u}_i(t))$ are not known. However, as mentioned in Section III, for each $i \in \{1, 2, \ldots, N\}$, the extremum seeking controller of the subsystem i contains a gradient estimator that produces an estimate of $\frac{dZ_i}{du_i}(\hat{u}_i(t))$ to optimize the performance of the subsystem. This estimate is also used for computing the dither amplitudes.

A. Theoretical Lower Fluctuation Bound for a Fixed Excitation Level

As a benchmark for the reduction of the fluctuation of the total performance cost that can be achieved with the described approach, we determine the theoretical minimal value of $\gamma(t)$ for any (not necessarily Lipschitz) mappings in (27) that satisfy the fixed-excitation condition in (26). This theoretical lower bound on the value of $\gamma(t)$ is given in the following lemma.

Lemma 3: Let $M \geq N - 1$. Consider any gradients $\frac{dZ_i}{du_i}(\hat{u}_i(t)) \in \mathbb{R}$ for all $i \in \{1, 2, \ldots, N\}$. Let us renumber the subsystems. For all $i \in \{1, 2, \ldots, N\}$ and all $t \geq 0$, the number $\mathcal{I}(i) = \mathcal{I}(i, t)$ corresponds to the number of the subsystem with the ith largest gradient. Hence, we renumber the subsystems such that $|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}(t))| \geq |\frac{dZ_{\mathcal{I}(2)}}{du_{\mathcal{I}(2)}}(\hat{u}_{\mathcal{I}(2)}(t))| \geq \cdots \geq |\frac{dZ_{\mathcal{I}(N)}}{du_{\mathcal{I}(N)}}(\hat{u}_{\mathcal{I}(N)}(t))|$. Under the fixed-excitation condition in (26), the minimal value of $\gamma(t)$ that can be achieved for any (not necessarily Lipschitz) mappings of the dither amplitudes is given by

$$\gamma(t) = \max\Bigg\{0, \left(\left|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}(t))\right| - \sum_{i=2}^{N}\left|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)}(t))\right|\right)\nu\Bigg\}. \tag{30}$$

Proof: See the Appendix A.

It follows from Lemma 3 that a zero total fluctuation can only be achieved whenever the magnitude of the largest gradient is smaller than or equal to the sum of magnitudes of all other gradients. Only in that case, it is possible to choose the dither amplitudes such that the fluctuations of the performance cost of individual subsystems cancel each other out. If the magnitude of the largest gradient is larger than the sum of all others, complete cancellation is not possible under the fixed-excitation condition in (26). The smallest fluctuation of the total performance cost can then be achieved by cancelling out the fluctuation of the performance cost of the subsystem with the largest gradient as much as possible by choosing the same perturbation with an effective opposite sign for all other subsystems. This approach results in the minimal value of $\gamma(t)$ in (30).

The mappings of the dither amplitudes that correspond to the minimal value of $\gamma(t)$ in Lemma 3 are discontinuous, and therefore, not globally Lipschitz. These discontinuities occur at the zero crossings of the gradient values. Alterations to obtain globally Lipschitz mappings are described in the next section. Note that these alternations inevitably lead to a larger value of $\gamma(t)$.

B. Dither Coordination Method

In this section, we present our solution to the dither coordination problem. Our method requires that the number of the basis function M is equal to the number of subsystems N. Let $\mathcal{I}(i)$ denote the numbering of the subsystems from the largest to the smallest gradient value, as in Lemma 3. For each $i, j \in \{1, 2, \ldots, N\}$, let us define the mappings

$$A_{\mathcal{I}(i)\mathcal{I}(j)}\left(\mathbf{g}_Z(t)\right) = \begin{cases} \chi_i(t) + \sum_{k=i}^{N} \varphi_{ki}(t)\xi_k(t), & \text{if } i = j \\ -s_{ij}(t) \sum_{k=\max\{i,j\}}^{N} \frac{\varphi_{ki}(t)\xi_k(t)}{\max\{k-1,1\}}, & \text{if } i \neq j \end{cases} \tag{31}$$

where the nonnegative function $\varphi_{ki}(t)$ is given by

$$\varphi_{ki}(t) = \begin{cases} \frac{|\frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)}(t))|}{|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)}(t))|}, & \text{if } |\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)}(t))| > 0 \\ 1, & \text{if } |\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)}(t))| = 0 \end{cases} \tag{32}$$

for all $i \leq k \leq N$, where the nonnegative function $\chi_i(t)$ is given by

$$\chi_i(t) = \nu \max\left\{1 - \lambda_1 \left|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)}(t))\right|, 0\right\} \tag{33}$$

for all $i \in \{1, 2, \ldots, N\}$ and some shape parameter $\lambda_1 \in \mathbb{R}_{>0}$, and where the function $s_{ij}(t)$ is given by

$$s_{ij}(t) = \operatorname{sign}\left(\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)}(t))\right) \operatorname{sign}\left(\frac{dZ_{\mathcal{I}(j)}}{du_{\mathcal{I}(j)}}(\hat{u}_{\mathcal{I}(j)}(t))\right) \tag{34}$$

with

$$\operatorname{sign}(r) = \begin{cases} -1, & \text{if } r < 0 \\ 0, & \text{if } r = 0 \\ 1, & \text{if } r > 0. \end{cases} \tag{35}$$

Due to the difference in magnitude of the gradients, the dither amplitudes of the various subsystems may have different effects on the fluctuation of the total performance cost. The function $\varphi_{ki}(t)$ in (32) is introduced to counteract these differences. To ensure that the fluctuations cancel each other out instead of amplifying each other, the function $s_{ij}(t)$ in (34) is introduced to give each amplitude the appropriate sign. In addition, to guarantee that the resulting dither signals satisfy the excitation condition in (26), we define the scaling parameters $\xi_k(t) \in \mathbb{R}_{\geq 0}$ for all $k \in \{1, 2, \ldots, N\}$, such that

$$\sum_{j=1}^{N} \left(A_{\mathcal{I}(i)\mathcal{I}(j)}\left(\mathbf{g}_Z(t)\right)\right)^2 = \nu^2 \tag{36}$$

for all $i \in \{1, 2, \ldots, N\}$. We note that, for each $i \in \{1, 2, \ldots, N\}$, (36) is quadratic in the parameters $\xi_k(t)$ for all $k \geq i$. The values of $\xi_k(t)$ are the nonnegative (analytic) solutions of these equations, and can be computed in a recursive manner starting at $k = N$. If all gradients are large, then $\chi_i(t)$ in (33) is equal to zero for all $i \in \{1, 2, \ldots, N\}$. Without the function $\chi_i(j)$ in (33), the mappings in (31) are not globally Lipschitz (similar to the optimal mappings in Lemma 3). With $\chi_i(t)$, it can be verified that the mappings of the dither amplitudes in (31) are globally Lipschitz for any positive value of the shape parameter λ_1. By substituting $a_{ij}(t) = A_{ij}(\mathbf{g}_Z(t))$ into (24), we obtain that

$$\begin{aligned} \gamma(t) &= \left(\left(\tfrac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}(t))\right)\left(\xi_1(t) + \chi_1(t)\right)\right)^2 \\ &\quad + \sum_{j=2}^{N}\left(\tfrac{dZ_{\mathcal{I}(j)}}{du_{\mathcal{I}(j)}}(\hat{u}_{\mathcal{I}(j)}(t))\chi_j(t)\right)^2\Bigg)^{\frac{1}{2}} \\ &\leq \left(\sqrt{\left|\tfrac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}(t))\right|^2 + \frac{N-1}{16\lambda_1^2}}\right)\nu \end{aligned} \tag{37}$$

for all $t \geq 0$ and any set of gradient values. This indicates that, for large values of λ_1, the mappings A_{ij} in (31) result in a small total fluctuation that is reasonably close to the minimal total fluctuation computed in Lemma 3, while selecting a small value of λ_1 may lead to a significant increase in total fluctuation of the performance cost. However, the resulting global Lipschitz constant is large for small values of λ_1. As a result, this may require that the extremum seeking controller is tuned very conservatively to make gradients sufficiently slow to satisfy the third requirement in Section IV. Hence, there is a tradeoff to be found between maximizing the optimization speed of the total performance cost and minimizing its fluctuation when selecting the parameter λ_1.

Although the mappings in (31) are not optimal, it should be noted that the worst-case fluctuation of the total performance cost for the mappings in (31) is equal to and often much smaller than the worst-case fluctuation for fixed dither amplitudes. This claim is easy to prove by noting that, for any dither amplitudes that satisfy the excitation condition in (26), there exist gradients such that

$$\gamma(t) \geq \left(\sqrt{\sum_{j=1}^{N}\left(\tfrac{dZ_j}{du_j}(\hat{u}_j(t))\right)^2}\right)\nu \tag{38}$$

and that the right-hand side of (38) is equal to or larger than the equality (but not necessarily the inequality) in (37); see the definitions of $\chi_i(t)$ and $\xi_k(t)$.

As mentioned in Section IV, estimates of the gradients need to be used because their true values are unknown. Let the vector of gradient estimates be denoted by $\hat{\mathbf{g}}_Z(t) = [\hat{g}_{Z,1}(t), \hat{g}_{Z,2}(t), \ldots, \hat{g}_{Z,N}(t)]^T$. Although the true gradient values may be slowly time varying, their estimates may not be. To guarantee that the estimates are sufficiently slow, we propose the following rate limiters:

$$\dot{\bar{g}}_{Z,i}(t) = \lambda_2 \operatorname{sat}\left(\frac{\dot{\hat{g}}_{Z,i}(t) - \lambda_3\left(\bar{g}_{Z,i}(t) - \hat{g}_{Z,i}(t)\right)}{\lambda_2}\right) \tag{39}$$

for $i \in \{1, 2, \ldots, N\}$, with

$$\operatorname{sat}(r) = \begin{cases} -1, & \text{if } r < -1 \\ r_i, & \text{if } -1 \leq r \leq 1 \\ 1, & \text{if } r > 1 \end{cases} \tag{40}$$

where λ_2 and $\lambda_3 \in \mathbb{R}_{>0}$ are tuning parameters. Note that $|\dot{\bar{g}}_{Z,i}(t)| \leq \lambda_2$ for all $t \geq 0$. Hence, λ_2 determines the upper limit on magnitude of the time derivative of $\bar{g}_{Z,i}(t)$. The value of the parameter λ_3 should be large so that $\bar{g}_{Z,i}(t)$ converges fast to the gradient estimate $\hat{g}_{Z,i}(t)$ whenever $|\dot{\bar{g}}_{Z,i}(t)| \leq \lambda_2$.

The dither amplitudes that are used to attenuate the fluctuation of the total performance cost are thus given by

$$a_{ij}(t) = A_{ij}\left(\bar{\mathbf{g}}_Z(t)\right) \tag{41}$$

with $\bar{\mathbf{g}}_Z(t) = [\bar{g}_{Z,1}(t), \bar{g}_{Z,2}(t), \ldots, \bar{g}_{Z,N}(t)]^T$. Note that, for all $i, j \in \{1, 2, \ldots, N\}$, the following bound on the time derivatives of the dither amplitudes can be given:

$$|\dot{a}_{ij}(t)| \leq \lambda_2 L_A \tag{42}$$

where the parameter λ_2 is defined in (39) and L_A is the global Lipschitz constant of the mappings in (31).

V. Extremum Seeking Controller

The proposed dither coordination method can be applied to new and existing ESC schemes; see Section III. We introduce the following controllers as an illustration. For each subsystem $i \in \{1, 2, \ldots, N\}$, let us consider the following gradient estimator, similar to the one in [15]:

$$\begin{aligned}
\dot{m}_i(t) &= \eta\big(y_i(t) - m_i(t)\big) \\
\dot{q}_{1,i}(t) &= -\eta\big(p_i(t) + q_{1,i}(t)\big) + \dot{\hat{u}}_i(t) \\
\dot{q}_{2,i}(t) &= \eta q_{2,i}(t)\big(1 - \big(p_i(t) + q_{1,i}(t)\big)^2 q_{2,i}(t)\big) \\
\dot{\hat{g}}_{Z,i}(t) &= \eta\big(p_i(t) + q_{1,i}(t)\big) q_{2,i}(t) \\
&\quad \times \big(y_i(t) - m_i(t) - \big(p_i(t) + q_{1,i}(t)\big)\hat{g}_{Z,i}(t)\big).
\end{aligned} \tag{43}$$

Here, $\hat{g}_{Z,i}(t) \in \mathbb{R}$ is an estimate of the gradient $\frac{dZ_i}{du_i}(\hat{u}_i(t))$. The signals $m_i(t), q_{1,i}(t) \in \mathbb{R}$ and $q_{2,i}(t) \in \mathbb{R}_{>0}$ are internal variables of the gradient estimator. The tuning parameter $\eta \in \mathbb{R}_{>0}$ is chosen sufficiently small, such that the variable $q_{1,i}(t)$ remains close to zero and the time variations in the (positive) variable $q_{2,i}(t)$ are small. Let the optimizer for each subsystem be given by

$$\dot{\hat{u}}_i(t) = -\kappa_1 \operatorname{proj}_i(\operatorname{sat}(\kappa_2 \hat{g}_{Z,i}(t)), \hat{u}_i(t)) \tag{44}$$

where κ_1 and $\kappa_2 \in \mathbb{R}_{>0}$ are tuning parameters. The function proj_i is the projection operator that confines $\hat{u}_i(t)$ to a set $\hat{\mathcal{U}}_i \subset \mathcal{U}_i$. From (17), (20), and (26), it follows that there exists a constant $c_p, \in \mathbb{R}_{>0}$ such that

$$|p_i(t)| \leq \nu c_p \tag{45}$$

for all $i \in \{1, 2, \ldots, N\}$ and all $t \geq 0$. By defining the set $\hat{\mathcal{U}}_i$ as

$$\hat{\mathcal{U}}_i = \left\{\hat{u}_i \in \mathbb{R} : s_i^{\min} + \nu c_p \leq \hat{u}_i \leq s_i^{\max} - \nu c_p\right\} \tag{46}$$

we ensure that, if $\hat{u}_i(t) \in \hat{\mathcal{U}}_i$, then $u_i(t) \in \mathcal{U}_i$. Here, we assume that ν is sufficiently small, such that the set $\hat{\mathcal{U}}_i$ is nonempty. The projection operator proj_i can now be defined as

$$\operatorname{proj}_i(r_i, \hat{u}_i) = \begin{cases} 0, & \text{if } \hat{u}_i \geq s_i^{\max} - \nu c_p \text{ and } r_i > 0 \\ & \text{or if } \hat{u}_i \leq s_i^{\min} + \nu c_p \text{ and } r_i < 0 \\ r_i, & \text{otherwise} \end{cases} \tag{47}$$

for all $i \in \{1, 2, \ldots, N\}$. In the next section, we prove asymptotic stability under suitable conditions for each closed-loop subsystem consisting of the subsystem in (1) and (3), the controller in (43) and (44), and the dither signal in (17) with the amplitudes in (41).

VI. Stability Analysis

In order to proof convergence of each extremum seeking loop, we require that true steady-state output of each subsystem given by the mapping Z_i in (9) has a unique minimum that corresponds to the optimal performance.

Assumption 4: For each $i \in \{1, 2, \ldots, N\}$, there exist a constant $u_i^* \in \mathcal{U}_i$ and a function $\alpha_{Z,i} \in \mathcal{K}$ such that

$$(u_i - u_i^*)\frac{dZ_i}{du_i}(u_i) \geq \alpha_{Z,i}(|u_i - u_i^*|) \tag{48}$$

for all $u_i \in \mathcal{U}_i$. Hence, u_i^* is a unique minimizer of the function Z_i in (9) on the domain $\mathcal{U}_i$ in (2).

Assumption 4 implies that, from any initial condition $u_i(0)$, we will end up at the optimal value u_i^* as long as we follow a gradient-descent path with the gradient-descent optimizer in (44). In addition to Assumption 4, we require that all functions are sufficiently smooth on their relevant domains. In order to define the relevant domain of the state vector $\mathbf{x}_i(t)$, let us introduce the set

$$\mathcal{X}_i = \bigcup_{u_i \in \mathcal{U}_i} \mathcal{B}_i(u_i; \rho_i) \tag{49}$$

where $\mathcal{B}_i(u_i; \rho_i)$ is the closed ball defined in (7) and $\rho_i = \beta_{\mathbf{x}i}(\rho_i^0, 0)$, with $\beta_{\mathbf{x}i}$ and ρ_i^0 defined in Assumption 2. We note that $\mathcal{X}_i$ is compact because the set $\mathcal{U}_i$ in (2) is a closed interval. We assume the following.

Assumption 5: The following holds for each $i \in \{1, 2, \ldots, N\}$. The function $\mathbf{f}_i(\mathbf{x}_i, u_i)$ in (1) is twice continuously differentiable in $\mathbf{x}_i$ on $\mathcal{X}_i$ and in u_i on $\mathcal{U}_i$. The function $h_i(\mathbf{x}_i)$ in (3) is twice continuously differentiable in $\mathbf{x}_i$ on $\mathcal{X}_i$. The mapping $\mathbf{X}_i(u_i)$ of Assumption 1 is twice continuously differentiable in u_i on $\mathcal{U}_i$.

With these assumptions, we are able to prove asymptotic convergence of each extremum seeking loop to a small region of the optimal steady-state conditions, and give a bound on the gradient estimation error.

Theorem 6: For any $i \in \{1, 2, \ldots, N\}$, suppose that the measurement noise $d_i(t)$ in (3) is uniformly bounded, i.e., there exists a constant $b_d \in \mathbb{R}_{>0}$, such that

$$|d_i(t)| \leq b_d \tag{50}$$

for all $t \geq 0$. Under Assumptions 1, 2, 4, and 5, there exist constants $\epsilon_1, \epsilon_2, \ldots, \epsilon_7 \in \mathbb{R}_{>0}$ and class-$\mathcal{K}$ functions α_u and $\alpha_{\dot{u}}$, such that for all initial conditions $\mathbf{x}_i(0) \in \mathcal{B}_i(u_i(0); \epsilon_1)$, $\hat{u}_i(0) \in \hat{\mathcal{U}}_i$, $m_i(0), \hat{g}_{Z,i}(0), \bar{g}_{Z,i}(0) \in \mathbb{R}$, $q_{1,i}(0) = 0$ and $q_{2,i}(0) = \frac{1}{\nu^2}$, and for all parameter values $\nu \leq \epsilon_2$, $\omega \leq \epsilon_3$, $\eta \leq \omega\epsilon_4$, $\kappa_1 \leq \eta\nu\epsilon_5$, $\lambda_1 \leq \epsilon_6$, $\lambda_2 \leq \eta\nu\epsilon_7$ and $\kappa_2, \lambda_3 \in \mathbb{R}_{>0}$, we have that the solution of the closed-loop system of the subsystem in (1) and (3), the controller in (43) and (44), and the dither signal in (17) with the amplitudes in (41) are uniformly bounded. Moreover, we have that

$$\begin{aligned}
&\limsup_{t\to\infty} \max\left\{|u_i(t) - u_i^*|, \left|\hat{g}_{Z,i}(t) - \frac{dZ_i}{du_i}(\hat{u}_i(t))\right|\right\} \\
&\leq \limsup_{t\to\infty} \alpha_u\left(\max\left\{\nu, \frac{\alpha_{\dot{u}}(\nu\omega)}{\nu}, \frac{|d_i(t)|}{\nu}\right\}\right).
\end{aligned} \tag{51}$$

Proof: See Section VI-A.

We note that we can make the bound in the right-hand side of (51) arbitrarily small in the absence of the disturbance $d_i(t)$ by choosing small values of ν and ω. In the presence of the disturbance $d_i(t)$, the bound cannot be made arbitrarily small. However, boundedness of the solutions can still be guaranteed. This stability result is similar to the ones in [14, Corollary 13] and [16, Th. 14].

A. Proof of the Stability Result

We begin the proof of Theorem 6 by deriving a bound on the magnitude of time derivative of the input $u_i(t)$. We obtain from (14) and (17) that

$$\dot{u}_i(t) = \dot{\hat{u}}_i(t) + \sum_{j=1}^{N} \left(\dot{a}_{ij}(t) b_j(\omega t) + \omega a_{ij}(t) \dot{b}_j(\omega t) \right) \quad (52)$$

where we have used that $M = N$ for the dither coordination method in Section IV-B. From (20), (21), $|a_{ij}(t)| \leq \nu$ in (26), (42), $|\dot{\hat{u}}_i(t)| \leq \kappa_1$ in (44), and the bounds on the parameters in Theorem 6, it subsequently follows that

$$|\dot{u}_i(t)| \leq \omega \nu c_{\dot{u}} \quad (53)$$

for all $t \geq 0$, all $\eta \leq \omega \epsilon_4$, all $\kappa_1 \leq \eta \nu \epsilon_5$ and all $\lambda_2 \leq \eta \nu \epsilon_7$ (where L_A in (42) is a global Lipschitz constant of the mappings in (31) for all $\lambda_1 \leq \epsilon_6$, for any given value of ϵ_6).

Now, let us focus our attention to the state of the subsystem in (1). Assumption 2 ensures convergence of the state $\mathbf{x}_i(t)$ to the steady state $\mathbf{X}_i(u_i)$ for fixed inputs u_i. In the following lemma, we prove convergence to a neighborhood of the steady state for time-varying inputs $u_i(t)$. Note that the size of this neighborhood depends on how fast the input $u_i(t)$ changes over time. Therefore, to ensure convergence to a small neighborhood, we must make sure that the time derivative of $u_i(t)$ is sufficiently small.

Lemma 7: Let us define

$$\tilde{\mathbf{x}}_i(t) = \mathbf{x}_i(t) - \mathbf{X}_i(u_i(t)). \quad (54)$$

Under the assumptions of Theorem 6, there exist class-$\mathcal{K}$ functions $\alpha_{\mathbf{x}1}$ and $\alpha_{\mathbf{x}2}$, such that the solutions of (1) satisfy

$$\sup_{t \geq 0} \|\tilde{\mathbf{x}}_i(t)\| \leq \max \left\{ \alpha_{\mathbf{x}1}(\|\tilde{\mathbf{x}}_i(0)\|), \sup_{t \geq 0} \alpha_{\mathbf{x}2}(|\dot{u}_i(\tau)|) \right\} \quad (55)$$

and

$$\limsup_{t \to \infty} \|\tilde{\mathbf{x}}_i(t)\| \leq \limsup_{t \to \infty} \alpha_{\mathbf{x}2}(|\dot{u}_i(t)|) \quad (56)$$

for all initial conditions $\mathbf{x}_i(0) \in \mathcal{B}_i(u_i; \epsilon_1)$, where the values of $\epsilon_1, \epsilon_2, \epsilon_3 > 0$ are sufficiently small to guarantee that $\mathbf{x}_i(t) \in \mathcal{X}_i$ for all $t \geq 0$, all $\nu \leq \epsilon_2$ and all $\omega \leq \epsilon_3$ by using the bound on $\dot{u}_i(t)$ in (53).

Proof: The proof of the lemma follows similar lines as the proofs of [37, Prop. 2] and [14, Lem. 8]. In short, first a converse Lyapunov theorem (e.g., [24] or [19, Th. 4.16]) is used to show that there exists a Lyapunov function for fixed inputs $u_i(t) = u_i$. Subsequently, this Lyapunov function is used to prove the input-to-state stability of the subsystem with respect to the time derivative $\dot{u}_i(t)$ of the input; see (55) and (56). For the construction of the converse Lyapunov function, we require that the functions $\mathbf{f}_i$ and $\mathbf{X}_i$ are Lipschitz on their domains. This is satisfied due to the differentiability of the functions (see Assumption 5) and the compactness of the sets $\mathcal{X}_i$ and $\mathcal{U}_i$. Further details of the proof are omitted for brevity.

The state error $\tilde{\mathbf{x}}_i(t)$ in (54) influences how close the measured output $y_i(t)$ in (3) of the subsystem is to the true steady-state output $Z_i(u_i(t))$ in (9). Let us define the measurement error

$$\tilde{y}_i(t) = y_i(t) - Z_i(u_i(t)). \quad (57)$$

Because $\mathbf{x}_i(t) \in \mathcal{X}_i$ and $u_i(t) \in \mathcal{U}_i$ for all $t \geq 0$ (see Lemma 7 and Section V), it follows (3), (9), and Assumption 5 that there exists a constant $c_{\tilde{y}} \in \mathbb{R}_{>0}$, such that

$$|\tilde{y}_i(t)| \leq c_{\tilde{y}} \|\tilde{\mathbf{x}}_i(t)\| + |d_i(t)| \quad (58)$$

for all $t \geq 0$. If $\tilde{\mathbf{x}}_i(t)$ and $d_i(t)$ are small, then the measurement error $\tilde{y}_i(t)$ is also small and the measurement output $y_i(t)$ is approximately equal to the true steady-state output $Z_i(u_i(t))$. In turn, the signal $\hat{g}_{Z_i}(t)$ produced by the estimator in (43) is an accurate approximation of the gradient $\frac{dZ_i}{du_i}(u_i(t))$, as shown next.

Lemma 8: Let us define

$$\tilde{g}_{Z,i}(t) = \hat{g}_{Z,i}(t) - \frac{dZ_i}{du_i}(\hat{u}_i(t)). \quad (59)$$

Under the assumptions of Theorem 6, there exists a constant $c_{\tilde{g}} \in \mathbb{R}_{>0}$, such that the solutions of the estimator in (43) are uniformly bounded and satisfy

$$\limsup_{t \to \infty} |\tilde{g}_{Z,i}(t)| \leq \limsup_{t \to \infty} \max \left\{ \nu, \frac{|\tilde{y}_i(t)|}{\nu} \right\} c_{\tilde{g}} \quad (60)$$

for all initial conditions $m_i(0) \in \mathbb{R}$, $\hat{g}_{Z,i}(0) \in \mathbb{R}$, $q_{1,i}(0) = 0$ and $q_{2,i}(0) = \frac{1}{\nu^2}$, and all parameter values $\eta \leq \omega \epsilon_4$, $\kappa_1 \leq \eta \nu \epsilon_5$ and $\lambda_2 \leq \eta \nu \epsilon_7$, with sufficiently small $\epsilon_4, \epsilon_5 > 0$.

Proof: See Appendix B.

Using an accurate gradient estimate, the gradient-descent optimizer in (44) steers the nominal input $\hat{u}_i(t)$ to a small neighborhood of the value

$$\hat{u}_i^* = \begin{cases} s_i^{\min} + \nu c_p, & \text{if } u_i^* < s_i^{\min} + \nu c_p \\ u_i^*, & \text{if } u_i^* \in \hat{\mathcal{U}}_i \\ s_i^{\max} - \nu c_p, & \text{if } u_i^* > s_i^{\max} - \nu c_p \end{cases} \quad (61)$$

see Assumption 4 and the definition of the set $\hat{\mathcal{U}}_i$ in (46). We obtain the following result.

Lemma 9: Under the conditions of Theorem 6, there exists a class-$\mathcal{K}$ function $\alpha_{\hat{u}}$, such that the solutions of the optimizer in (44) are uniformly bounded and satisfy

$$\limsup_{t \to \infty} |\hat{u}_i(t) - \hat{u}_i^*| \leq \limsup_{t \to \infty} \alpha_{\hat{u}}(|\tilde{g}_{Z,i}(t)|) \quad (62)$$

for all initial conditions $\hat{u}_i(0) \in \hat{\mathcal{U}}_i$, and all parameter values $\kappa_1 \leq \eta \nu \epsilon_5$ and $\kappa_2 \in \mathbb{R}_{>0}$.

Proof: Due to the use of the projection operator in (47), we have that the nominal input satisfies $\hat{u}_i(t) \in \hat{\mathcal{U}}_i$ for all $t \geq 0$, whenever $\hat{u}_i(0) \in \hat{\mathcal{U}}_i$. Thus, all solutions of the optimizer are uniformly bounded. Heavily relying on Assumption 4, the proof of the lemma follows similar lines as the proof of [37, Prop. 1]. Details of the proof are omitted for brevity.

It follows from (61) that $|u_i^* - \hat{u}_i^*| \leq \nu c_p$. Therefore, we obtain using (14) and (45) that

$$|u_i(t) - u_i^*| \leq |\hat{u}_i(t) - \hat{u}_i^*| + 2\nu c_p \quad (63)$$

for all $t \geq 0$. Note that all solutions of the extremum seeking scheme are bounded under the conditions of Theorem 6. The

TABLE I
PARAMETERS OF THE PHOTOVOLTAIC ARRAY MODEL

Parameters	Descriptions
$T_r = 298.15K$	Reference temperature
$I_s = 5.61A$	Reference short-circuit current at T_r
$I_0 = 1.13 \cdot 10^{-6}A$	Nominal reverse saturation current at T_r
$k_i = 1.96 \cdot 10^{-3}A/K$	Short-circuit current temperature coefficient
$N = 1.81$	Ideality factor
$E_g = 1.16eV$	Bandgap energy for Silisium
$k = 1.38 \cdot 10^{-23}J/K$	Boltzmann constant
$q = 1.60 \cdot 10^{-19}C$	Charge of an electron
$n_s = 72$	Number of photo-voltaic cells in series
$R_s = 2.83 \cdot 10^{-3}\Omega$	Series resistance
$R_p = 8.7\Omega$	Parallel resistance
$C_c = 1mF$	Converter's capacitance
$L_c = 5mH$	Converter's inductance
$R_c = 2\Omega$	Converter's resistance

existence of the bound in (51) follows from (53), (56), (58), and (60)–(63).

VII. EXAMPLES

A. Power Production of Solar Arrays

Consider three photovoltaic arrays with n_s cells in series affected by different sun light and temperature conditions. According to the model presented in [38], the light generated current i_s and the reverse saturation current i_o in a photovoltaic cell with temperature T and solar irradiance S can be described as

$$i_s = \left(I_s + k_i(T - T_r)\right)\frac{S}{1000}$$
$$i_0 = I_0\left(\frac{T}{T_r}\right)^3 e^{\frac{E_g}{NV_t}\left(\frac{T}{T_r}-1\right)}, \quad V_t = \frac{kT}{q} \tag{64}$$

where the parameters of the equations are given in Table I. The corresponding output current i of a photovoltaic array with n_s connected cells in series is

$$i = i_s - i_0\left(e^{\frac{v+iR_sn_s}{NV_tn_s}} - 1\right) - \frac{v + iR_sn_s}{R_pn_s} \tag{65}$$

where v is the output voltage. A dc–dc buck converter is used to connect each array to a dc load. The converter dynamics are represented by the model in [23] as

$$C_c\dot{v} = i(v;T,S) - i_Lu$$
$$L_c\dot{i}_L = -i_LR_c + vu \tag{66}$$

where $i(v;T,S)$ is the nonlinear mapping from the duty cycle u to the output current. This mapping can be computed from (64) and (65).

Three photovoltaic arrays are simulated for the temperatures T and irradiances S given in Table II. A subscript $i \in \{1,2,3\}$ is used to distinguish between the three arrays. The steady-state relation between the duty cycle and the power output of each array is depicted in Fig. 4. To maximize the total produced power, we apply the extremum seeking controllers in Section V to each of the photovoltaic arrays. Here, we use the negative of the power as a measure for the performance cost

$$y(t) = -P(t) = -i(t)v(t). \tag{67}$$

TABLE II
TEMPERATURE AND IRRADIANCE CONDITIONS FOR EACH PHOTOVOLTAIC ARRAY, AND CONTROLLER SETTINGS

Parameters	Descriptions
$T_1 = 296K$	Temperature of first array
$S_1 = 500Wm^{-2}$	Solar irradiance of first array
$T_2 = 295K$	Temperature of second array
$S_2 = 750Wm^{-2}$	Solar irradiance of second array
$T_3 = 298K$	Temperature of third array
$S_3 = 1000Wm^{-2}$	Solar irradiance of third array
$\nu = 0.01/\sqrt{3}$	Level of excitation in (26)
$\omega = 10$	Dither parameter in (17)
$\eta = 8$	Parameter of estimator in (43)
$k_1 = 1$	Parameter of optimizer in (44)
$k_2 = 7 \cdot 10^{-5}$	Parameter of optimizer in (44)
$\lambda_1 = 0.01$	Parameter of dither coordination in (33)
$\lambda_2 = 200$	Parameter of rate limiter in (39)
$\lambda_3 = 1000$	Parameter of rate limiter in (39)
$b_1 = \sqrt{3}\,\mathrm{tri}(4\omega t)$	First basis function dither
$b_2 = \sqrt{3}\,\mathrm{tri}(5\omega t)$	Second basis function dither
$b_3 = \sqrt{3}\,\mathrm{tri}(6\omega t)$	Third basis function dither

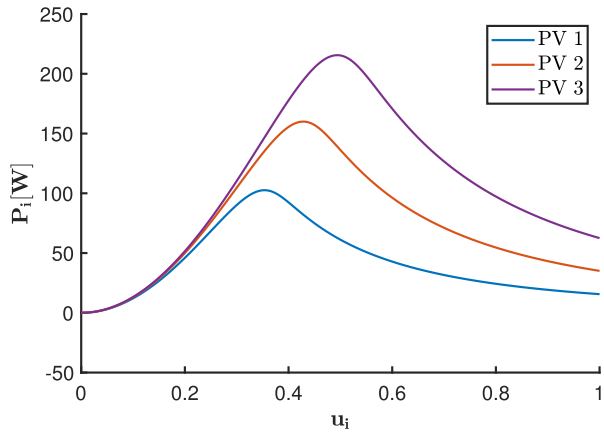

Fig. 4. Steady-state mappings of the effect of the duty cycle u_i on the produced power P_i for the three photovoltaic arrays $i \in \{1,2,3\}$ with different environmental conditions.

We apply the dither coordination method in Section IV-B to minimize the fluctuation of the total produced power. The base functions of the dither signals and the controller parameters are presented in Table II, where $\mathrm{tri}(\cdot)$ is the triangle wave function defined in [36]. We constrain the duty cycle of each array to be between 0.45 and 1. Therefore, two out of the three arrays are not able to produce at their peak values in Fig. 4. As a reference, we compare the results for the varying coordinated dither signals with those of fixed dither signals (where the amplitudes of the dither signals do not change over time). The fixed dither signals have the same basis functions and effective amplitude as the varying dither signals.

We observe from the simulation results in Fig. 5 that each extremum seeking controller steers the duty cycle u to its constrained optimal value. As a result, the power output P of each photovoltaic array goes to its constrained maximum. All optimal values are indicated by the dashed lines in Fig. 5. The figure shows the results for the varying dither signals. Due to their large similarity, the corresponding results for the fixed dither

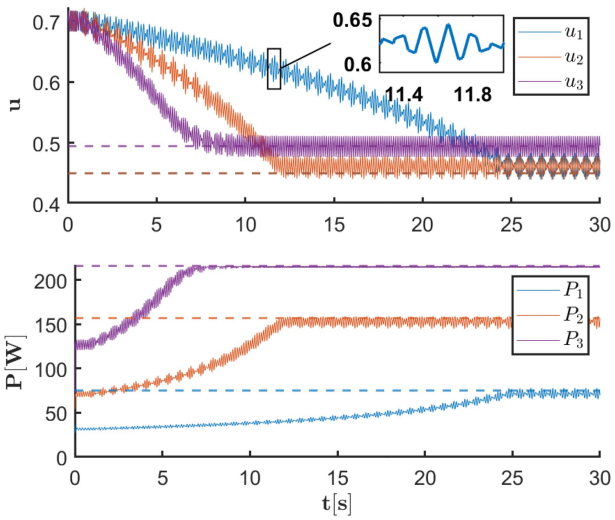

Fig. 5. Duty cycle u_i and power output P_i of each photovoltaic array $i \in \{1, 2, 3\}$ for the varying dither signals. The dashed lines correspond to the constrained optimal values.

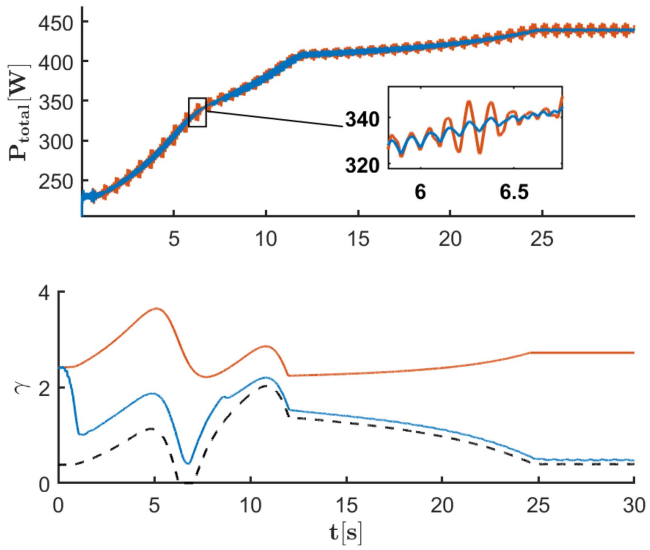

Fig. 6. Total power P_{total} and fluctuation measure γ for the three photovoltaic arrays with fixed dither signals and with varying dither signals. The blue lines correspond to the varying dither signals. The red lines correspond to the fixed dither signals. The dashed line corresponds to the theoretical lower bound on γ in Lemma 3.

signals are omitted. The main difference between the varying dither signals and the fixed dither signals is showcased in Fig. 6; the fluctuation in the total power P_{total} of the three arrays is visually much smaller for the varying dither signals with dither coordination than for the fixed dither signals without dither coordination. Due to the smaller fluctuation, we see in Fig. 6 that the value of the fluctuation measure γ in (24) is much smaller overall for the varying dither signals than for the fixed dither signals. Moreover, after the transient response has died out in the first two seconds, the value of γ for the varying dither signals remains close to the theoretical lower bound derived in Lemma 3. This lower bound is indicated by the dashed line in Fig. 6. Fig. 7

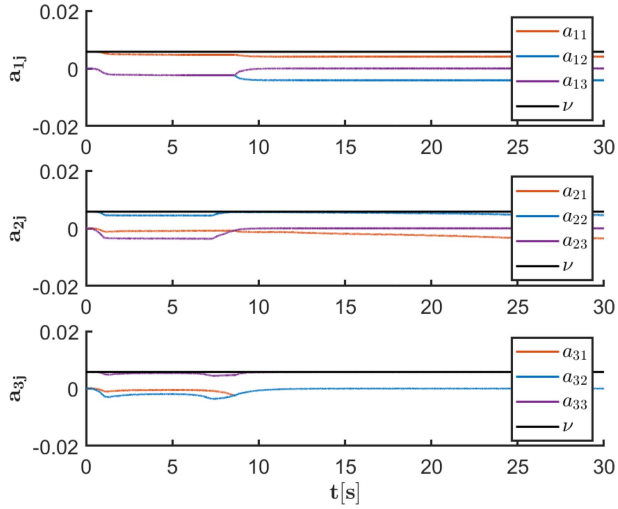

Fig. 7. Dither amplitudes a_{ij} for each photovoltaic array $i \in \{1, 2, 3\}$.

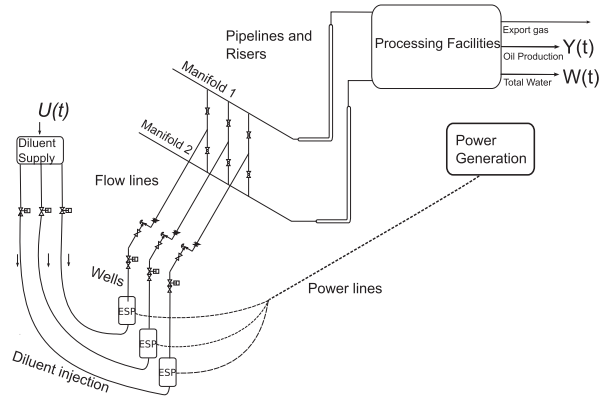

Fig. 8. ESP-lifted wells with diluent injection producing to an offshore platform.

presents the dither amplitude a_{ij} that are used to compute the varying dither signals. Note that there are no rapid changes in the values of a_{ij}, which is one of the three requirements in Section IV. Thus, the proposed dither coordination method in Section IV-B is able to drastically reduce the fluctuation in the total produced power, even in the presence of dynamics and local input constraints.

B. Power Optimization in Electric Submersible Pump (ESP)-Diluent Lifted Wells

In this example we consider oil wells equipped with ESPs [34] installed several hundred meters down in the well. High viscosity of the reservoir fluid (consisting of viscous oil and water) increased by emulsion formation at the ESP, significantly reduces ESP efficiency and increases its power consumption. To reduce the ESP power consumption, diluent (lighter oil) is injected upstream the ESP [10], [26]. A diagram of this production system for the case of several wells is shown in Fig. 8.

In the initial phase of oil production, when the percentage of water in the reservoir fluid is low, increasing the diluent injection rate initially reduces the ESP power consumption; after a certain optimal rate, it starts to increase the ESP power

consumption. The objective of diluent optimization is to find the optimal diluent injection rate that provides a minimal power consumption for the given ESP intake pressure setpoint.

Due to high uncertainties (emulsions can increase fluid viscosity up to six times [28]), extremum seeking control is a proper optimization tool to address this problem. However, dither-induced fluctuations in ESP power consumption from several ESP-lifted wells can add up to formidable power fluctuations negatively affecting power generators on the oil platform. The dither coordination method presented in this article can be applied to reduce these power fluctuations.

For simplicity, we present a static model of an ESP lifted well, which is derived from the dynamic model in [4]

$$\begin{aligned} p_{\text{in}} &= p_{\text{wh}} + \rho g h + \Delta p_f - \Delta p_p \\ q_r &= PI(p_r - p_{bh}) \\ q_p &= q_r + q_d \\ \Delta p_p &= \rho g C_H(\mu) H_0\left(\frac{q_p f_0}{C_q(\mu) f}\right)\left(\frac{f}{f_0}\right)^2 \\ P &= C_P(\mu) P_0\left(\frac{q_p f_0}{C_q(\mu) f}\right)\left(\frac{f}{f_0}\right)^3\left(\frac{\rho}{\rho_0}\right) \end{aligned} \tag{68}$$

where p_{in} is the ESP intake pressure; p_{wh} is the well-head pressure; $\rho g h$ is the hydrostatic pressure at the ESP intake for a given fluid density ρ and a vertical distance h; Δp_f is the frictional pressure loss (calculated based on the Darcy–Weisbach and Colebrook equations [17]); and Δp_p is the pressure increase yielded by the pump. The production rate of the reservoir fluid q_r is calculated as a product of the productivity index PI and the difference between the reservoir pressure p_r and the pressure p_{bh} at the bottom of the well. The total flow rate through the pump q_p consists of the reservoir inflow rate q_r and the diluent injection rate q_d.

The pressure increase provided by the ESP Δp_p and the ESP break horse power P depend on the pump speed f, flow rate through the pump q_p, the fluid density ρ, and the viscosity μ. Δp_p and P are calculated from the pump characteristics $H_0(q)$ and $P_0(q)$ corresponding to the pump performance at the pump speed f_0, the flow rate q, the fluid density ρ_0, and the viscosity $\mu = 1$ cP. The terms $C_H(\mu)$, $C_Q(\mu)$, and $C_P(\mu)$ are viscosity correction factors for using the model with fluids of higher viscosities; see [4] for more details on the model.

Diluent injection enters the model through the diluent injection rate q_d, the fluid viscosity μ, and the density ρ, which is calculated by mass balance equations from the densities and flow rates of the reservoir fluid and the diluent. The viscosity of the diluted fluid μ is calculated using formulas for oil–water emulsions from [6]. The effect of diluent depends on the water cut WC—The percentage of water in the reservoir fluid.

The ESP speed f is controlled by a proportional-integral controller that keeps ESP intake pressure p_{in} at a set-point p_{in}^{sp}, and thus, provides a constant inflow rate of the reservoir fluid consisting of oil and water.

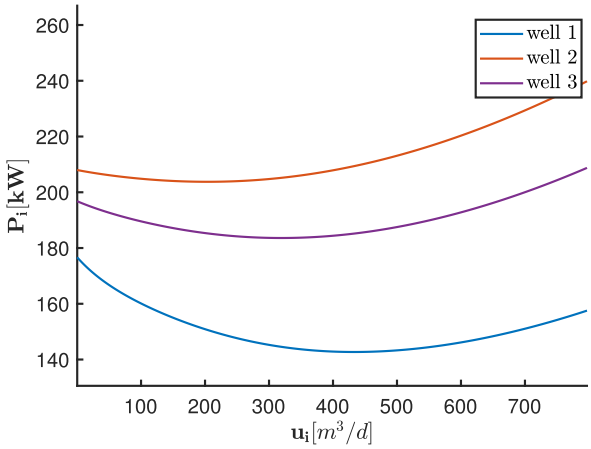

Fig. 9. Steady-state mappings describing the effect of the diluent injection u_i on the ESP power P_i for the three wells $i \in \{1, 2, 3\}$, with WC =32%, 40%, and 35%, respectively.

TABLE III
CONTROLLER PARAMETERS

Parameters	Descriptions
$\nu = 10/\sqrt{2}$	Level of excitation in (26)
$\omega = 0.1$	Dither parameter in (17)
$\eta = 5$	Parameter of estimator in (43)
$k_1 = 60$	Parameter of optimizer in (44)
$k_2 = 0.1$	Parameter of optimizer in (44)
$\lambda_1 = 100$	Parameter of dither coordination in (33)
$\lambda_2 = 5 \cdot 10^{-5}$	Parameter of rate limiter in (39)
$\lambda_3 = 5 \cdot 10^{-6}$	Parameter of rate limiter in (39)
$b_1 = \sqrt{2}\sin(3\omega t)$	First basis function dither
$b_2 = \sqrt{2}\cos(3\omega t)$	Second basis function dither
$b_3 = \sqrt{2}\sin(4\omega t)$	Third basis function dither

Based on the presented model, we calculate the curves describing the effect of diluent injection on the ESP power consumption for three identical wells with different water cuts: WC=32%, 40%, and 35%, respectively, as shown in Fig. 9. Numerical values for these curves are taken from [4] with the additional parameters being: the viscosities and densities $\mu_d = 1$ cP and $\rho_d = 800\,\text{kg/m}^3$ of the diluent, $\mu_o = 125$ cP and $\rho_o = 970\,\text{kg/m}^3$ of the oil, and $\mu_w = 1$ cP and $\rho_w = 1000\,\text{kg/m}^3$ of the water. In reality, these curves, apart from being convex, are highly uncertain due to inaccurate pump models, uncertainties in reservoir fluid properties and a varying diluent mixing efficiency at the ESP. Extremum seeking control is, thus, a suitable method to find the optimal diluent injection rate [26], although one needs to avoid or reduce the dither-induced fluctuation in the total ESP power consumption.

This can be achieved with the method presented in this article. The basis functions and parameters for the extremum seeking controller are presented in Table III. The basis functions have periods of about 15–20 min to ensure quasi steady-state operation of the well under these perturbations. The diluent injection is constrained to the interval of [100,400] m^3/d.

Fig. 10 shows the performance of the controller for the three ESP-lifted wells with varying dithers signals. As seen from the figure, the diluent injection rates $u_i = q_{d_i}$ for each well

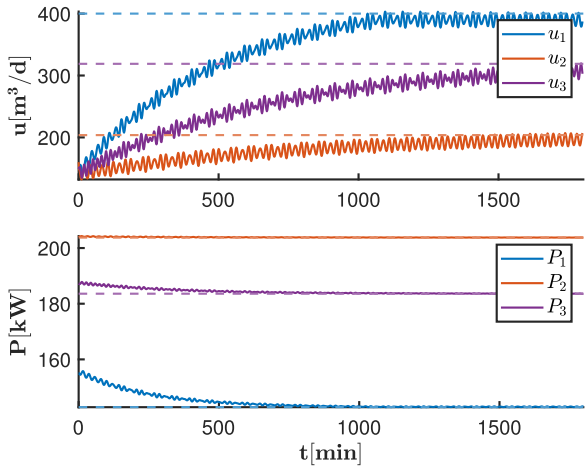

Fig. 10. Diluent injection u_i and consumed power P_i of each ESP $i \in \{1, 2, 3\}$ for the varying dither signals. The dashed lines correspond to the optimal values.

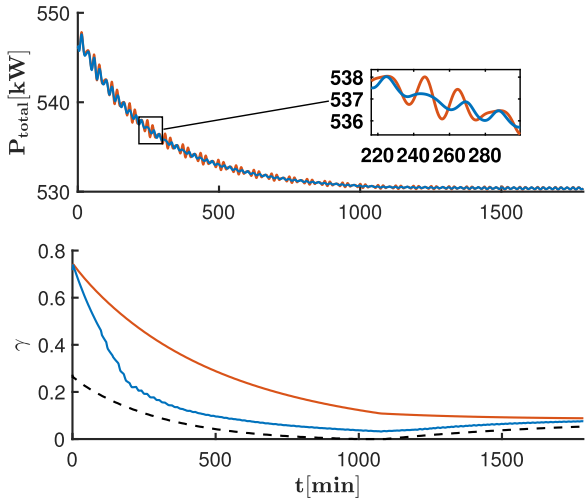

Fig. 11. Total power P_{total} and fluctuation measure γ for the three wells with fixed dither signals and with varying dither signals. The blue lines correspond to the varying dither signals. The red lines correspond to the fixed dither signals. The dashed line corresponds to the theoretical lower bound on γ in Lemma 3.

i are steered toward the optimal solutions by the extremum seeking controllers so that the power P_i consumed by each ESP reaches its minimum value. Fig. 11 presents the total power consumed by all the ESPs and the fluctuation measure γ in (24) for fixed and varying dither signals. As seen in the figure, the extremum seeking controllers with coordinated varying dither signals noticeably mitigate the fluctuation of the total power consumption compared to the standard method with fixed dither signals. The fluctuation measure γ is considerably lower for the coordinated dither signals, being much closer to the theoretical lower bound indicated by the dashed line in the figure. Fig. 12 depicts the dither amplitudes a_{ij} calculated for the coordinated dither signals shown in Fig. 10. The absence of spikes or abrupt changes in the dither amplitudes is essential for the smooth and safe operation of this algorithm in the presented oil production system.

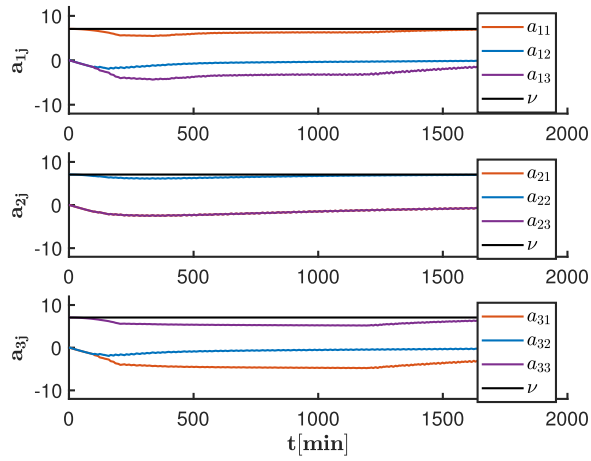

Fig. 12. Dither amplitudes a_{ij} for each well $i \in \{1, 2, 3\}$.

VIII. Conclusion

In this article, we have introduced a novel method to coordinate the dither signals utilized in the extremum seeking loops of individual subsystems in order to minimize the dither-induced fluctuation in the total output of a multiagent industrial system. Therefore, we do not only optimize the quantity of the total output by employing the optimizing capabilities of ESC but also its quality by minimizing the undesired dither-induced fluctuation. The presented method is computationally much lighter than current alternatives. This greatly benefits its practical deployment. The stability analysis in this article tells us that convergence to a small region of the optimum can be guaranteed under conditions that are similar to those of other ESC methods. Two case studies show that the presented dither-coordination method can drastically decrease the magnitude of the fluctuation in the total output, while having a negligible effect on the performance of the extremum seeking loop of each individual subsystem.

Appendix A
Proof of Lemma 3

For brevity of notation, we drop the time index t. The proof of the lemma consists of two parts. In the first part, we prove that, if $\sum_{i=2}^{N} |\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})| \geq |\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})|$, it is always possible to choose the dither amplitudes such that $\gamma = 0$ and $\sum_{j=1}^{M} a_{ij}^2 = \nu^2$ for all $i \in \{1, 2, \ldots, N\}$. In the second part, we prove that the minimal value of γ is $(|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})| - \sum_{i=2}^{N} |\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})|)\nu$, whenever $|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})| > \sum_{i=2}^{N} |\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})|$.

Part 1: For γ to be zero, we require that $\sum_{i=1}^{N} \frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)}) a_{\mathcal{I}(i)j} = 0$ for all $j \in \{1, 2, \ldots, M\}$; see (24). By squaring both sides of the equation, we can write this condition as $(\mathbf{a}_j^{\text{col}})^T \mathbf{M} \mathbf{a}_j^{\text{col}} = 0$, with $\mathbf{a}_j^{\text{col}} = [a_{\mathcal{I}(1)j}, a_{\mathcal{I}(2)j}, \ldots, a_{\mathcal{I}(N)j}]^T$ and $\mathbf{M} = (\frac{d\tilde{Z}}{d\mathbf{u}}(\hat{\mathbf{u}}))^T \frac{d\tilde{Z}}{d\mathbf{u}}(\hat{\mathbf{u}})$, with $\frac{d\tilde{Z}}{d\mathbf{u}}(\hat{\mathbf{u}}) = [\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}), \frac{dZ_{\mathcal{I}(2)}}{du_{\mathcal{I}(2)}}(\hat{u}_{\mathcal{I}(2)}), \ldots, \frac{dZ_{\mathcal{I}(N)}}{du_{\mathcal{I}(N)}}(\hat{u}_{\mathcal{I}(N)})]^T$. We obtain that $\gamma = 0$ if and only if $\mathbf{a}_j^{\text{col}} \in \ker(\mathbf{M})$ for all $j \in \{1, 2, \ldots, M\}$.

Due to the numbering specified in the lemma, we have that $\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}) = 0$ implies that $\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)}) = 0$ for all $i \in \{1, 2, \ldots, N\}$. If $\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}) = 0$, then $\mathbf{M}(t) = \mathbf{0}$. Therefore, any vector $\mathbf{a}_j^{\text{col}} \in \mathbb{R}^N$ is in the nullspace of $\mathbf{M}$. As a consequence, any dither amplitudes that satisfy $\sum_{j=1}^M a_{ij}^2 = \nu^2$ for all $i \in \{1, 2, \ldots, N\}$ will result in the theoretically minimal value $\gamma = 0$. We can always find dither amplitudes that satisfy these conditions. Thus, the minimal value of γ is zero if $\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}) = 0$.

Now, let $\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}) \neq 0$. The nullspace of the matrix $\mathbf{M}$ is given by

$$\ker(\mathbf{M}) = \text{span}\,\{\bar{\mathbf{v}}_2,\, \bar{\mathbf{v}}_3,\, \ldots,\, \bar{\mathbf{v}}_N\} \tag{69}$$

where for each $k \in \{2, 3, \ldots, N\}$, $\bar{\mathbf{v}}_k$ is a vector with two nonzero elements: its first element is $\frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)})$; its kth element is $-\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})$ (i.e., $\bar{\mathbf{v}}_2 = [\frac{dZ_{\mathcal{I}(2)}}{du_{\mathcal{I}(2)}}(\hat{u}_{\mathcal{I}(2)}), -\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}), 0, 0, \ldots, 0]^T$, $\bar{\mathbf{v}}_3 = [\frac{dZ_{\mathcal{I}(3)}}{du_{\mathcal{I}(3)}}(\hat{u}_{\mathcal{I}(3)}), 0, -\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}), 0, \ldots, 0]^T$, etc.). Therefore, any vector $\mathbf{a}_j^{\text{col}} \in \ker(\mathbf{M})$ may be written as $\mathbf{a}_j^{\text{col}} = \sum_{k=2}^N q_{jk}\bar{\mathbf{v}}_k$ for some coefficients $q_{jk} \in \mathbb{R}$. It follows from this and the definitions of $\mathbf{a}_j^{\text{col}}$ and $\bar{\mathbf{v}}_k$ that, if there exist dither amplitudes that satisfy $\gamma = 0$ and $\sum_{j=1}^M a_{ij}^2 = \nu^2$ for all $i \in \{1, 2, \ldots, N\}$, then there must exist coefficients q_{jk} such that

$$\sum_{j=1}^M a_{\mathcal{I}(1)j}^2 = \sum_{j=1}^M \left(\sum_{k=2}^N q_{jk} \frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)}) \right)^2 = \nu^2 \tag{70}$$

and

$$\sum_{j=1}^M a_{\mathcal{I}(k)j}^2 = \left(\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}) \right)^2 \sum_{j=1}^M q_{jk}^2 = \nu^2 \tag{71}$$

for all $k \in \{2, 3, \ldots, N\}$. Let us define the vectors $\mathbf{q}_k^{\text{col}} = [q_{1k}, q_{2k}, \ldots, q_{Mk}]^T$ for all $k \in \{2, 3, \ldots, N\}$. The equations in (70) and (71) can be rewritten as

$$\sum_{k=2}^N \sum_{l=2}^N \left(\mathbf{q}_k^{\text{col}}\right)^T \mathbf{q}_l^{\text{col}} \frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)}) \frac{dZ_{\mathcal{I}(l)}}{du_{\mathcal{I}(l)}}(\hat{u}_{\mathcal{I}(l)}) = \nu^2 \tag{72}$$

and

$$\left(\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}) \right)^2 \|\mathbf{q}_k^{\text{col}}\|^2 = \nu^2 \tag{73}$$

respectively. In turn, the condition in (73) is equivalent to $\|\mathbf{q}_k^{\text{col}}\| = \frac{\nu}{|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})|}$. Hence, $\mathbf{q}_k^{\text{col}}$ is a vector of length $\frac{\nu}{|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})|}$ for all $k \in \{2, 3, \ldots, N\}$. By combining the conditions in (72) and (73), we get that there exist dither amplitudes that satisfy the conditions $\gamma = 0$ and $\sum_{j=1}^M a_{ij}^2 = \nu^2$ for all $i \in \{1, 2, \ldots, N\}$ if and only if there exist vectors $\mathbf{r}_k^{\text{col}} \in \mathbb{R}^M$ of length one (i.e., $\|\mathbf{r}_k^{\text{col}}\| = 1$) for all $k \in \{2, 3, \ldots, N\}$ such that

$$\sum_{k=2}^N \sum_{l=2}^N \left(\mathbf{r}_k^{\text{col}}\right)^T \mathbf{r}_l^{\text{col}} \frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)}) \frac{dZ_{\mathcal{I}(l)}}{du_{\mathcal{I}(l)}}(\hat{u}_{\mathcal{I}(l)}) = \left(\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}) \right)^2. \tag{74}$$

Here, $\mathbf{r}_k^{\text{col}} = \frac{\mathbf{q}_k^{\text{col}}}{\|\mathbf{q}_k^{\text{col}}\|}$ for all $k \in \{2, 3, \ldots, N\}$. Noting that $-1 \leq (\mathbf{r}_k^{\text{col}})^T \mathbf{r}_l^{\text{col}} \leq 1$ for all $k, l \in \{2, 3, \ldots, N\}$, it is not difficult to show that there exist vectors $\mathbf{r}_k^{\text{col}} \in \mathbb{R}^M$ for $k \in \{2, 3, \ldots, N\}$ for which the left-hand side of (74) attains its maximal value

$$\sum_{k=2}^N \sum_{l=2}^N \left(\mathbf{r}_k^{\text{col}}\right)^T \mathbf{r}_l^{\text{col}} \frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)}) \frac{dZ_{\mathcal{I}(l)}}{du_{\mathcal{I}(l)}}(\hat{u}_{\mathcal{I}(l)}) = \left(\sum_{k=2}^N \left| \frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)}) \right| \right)^2. \tag{75}$$

Moreover, we may always select vectors $\mathbf{r}_k^{\text{col}}$ for $k \in \{2, 3, \ldots, N\}$ such that

$$\begin{aligned} &\sum_{k=2}^N \sum_{l=2}^N \left(\mathbf{r}_k^{\text{col}}\right)^T \mathbf{r}_l^{\text{col}} \frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)}) \frac{dZ_{\mathcal{I}(l)}}{du_{\mathcal{I}(l)}}(\hat{u}_{\mathcal{I}(l)}) \\ &\quad = \left(\left| \tfrac{dZ_{\mathcal{I}(2)}}{du_{\mathcal{I}(2)}}(\hat{u}_{\mathcal{I}(2)}) \right| - \left| \tfrac{dZ_{\mathcal{I}(3)}}{du_{\mathcal{I}(3)}}(\hat{u}_{\mathcal{I}(3)}) \right| \right)^2 \\ &\qquad + \left(\left| \tfrac{dZ_{\mathcal{I}(4)}}{du_{\mathcal{I}(4)}}(\hat{u}_{\mathcal{I}(4)}) \right| - \left| \tfrac{dZ_{\mathcal{I}(5)}}{du_{\mathcal{I}(5)}}(\hat{u}_{\mathcal{I}(5)}) \right| \right)^2 + \cdots \\ &\qquad + \begin{cases} \left(\left| \tfrac{dZ_{\mathcal{I}(N-1)}}{du_{\mathcal{I}(N-1)}}(\hat{u}_{N-1}) \right| - \left| \tfrac{dZ_{\mathcal{I}(N)}}{du_{\mathcal{I}(N)}}(\hat{u}_{\mathcal{I}(N)}) \right| \right)^2, & \text{if } N \text{ is odd} \\ \left| \tfrac{dZ_{\mathcal{I}(N)}}{du_{\mathcal{I}(N)}}(\hat{u}_{\mathcal{I}(N)}) \right|^2, & \text{if } N \text{ is even.} \end{cases} \end{aligned} \tag{76}$$

Note that the right-hand side of (76) is smaller than or equal to $(\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}))^2$ using the specified numbering of the gradients. Moreover, noting that $\mathbf{r}_k^{\text{col}}$ for $k \in \{2, 3, \ldots, N\}$ are $N-1$ vectors on an M-dimensional unit sphere, $M \geq N-1$ implies that there exist a continuous path on the M-dimensional unit sphere between the vector values that correspond to (75) and the vector values that correspond to (76). Therefore, following this path, the value of the left-hand side of (74) will change from the value $(\sum_{k=2}^N |\frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)})|)^2$ to a value that is smaller than or equal to $(\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)}))^2$ in a continuous manner. Thus, somewhere on this path, we will always encounter values of $\mathbf{r}_k^{\text{col}}$ for $k \in \{2, 3, \ldots, N\}$ for which (74) holds if and only if $\sum_{k=2}^N |\frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)})| \geq |\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})|$. Hence, we obtain that there exist dither amplitudes for which $\gamma = 0$ and $\sum_{j=1}^M a_{ij}^2 = \nu^2$ for all $i \in \{1, 2, \ldots, N\}$ if and only if $\sum_{k=2}^N |\frac{dZ_{\mathcal{I}(k)}}{du_{\mathcal{I}(k)}}(\hat{u}_{\mathcal{I}(k)})| \geq |\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})|$.

Part 2: Let us define $\mathbf{a}_i^{\text{row}} = [a_{\mathcal{I}(i)1}, a_{\mathcal{I}(i)2}, \ldots, a_{\mathcal{I}(i)M}]^T$. The excitation condition $\sum_{j=1}^M a_{ij}^2 = \nu^2$ for all $i \in \{1, 2, \ldots, N\}$ implies that $\|\mathbf{a}_i^{\text{row}}\| = \nu$ for all $i \in \{1, 2, \ldots, N\}$. From (24), it follows that

$$\gamma^2 = \left\| \sum_{i=1}^N \frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)}) \mathbf{a}_i^{\text{row}} \right\|^2. \tag{77}$$

Using the sign function, we may rewrite this equation as follows:

$$\gamma^2 = \|\bar{\mathbf{w}}_1 + \bar{\mathbf{w}}_2\|^2 \tag{78}$$

with

$$\bar{\mathbf{w}}_1 = \left(\left|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})\right| - \sum_{i=2}^{N}\left|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})\right|\right)\mathbf{a}_1^{\text{row}}$$
$$\bar{\mathbf{w}}_2 = \sum_{i=2}^{N}\left|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})\right| \times \left(\text{sign}\left(\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})\right)\mathbf{a}_i^{\text{row}} + \mathbf{a}_1^{\text{row}}\right). \tag{79}$$

Because $\|\mathbf{a}_1^{\text{row}}\| = \nu$, we have

$$\|\bar{\mathbf{w}}_1\|^2 = \left(\left|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})\right| - \sum_{i=2}^{N}\left|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})\right|\right)^2 \nu^2 \tag{80}$$

regardless of the choice of $\mathbf{a}_1^{\text{row}}$. Note that, for any vectors $\mathbf{a}_1^{\text{row}}$ and $\mathbf{a}_i^{\text{row}}$, we have

$$(\mathbf{a}_1^{\text{row}})^T\left(\text{sign}\left(\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})\right)\mathbf{a}_i^{\text{row}} + \mathbf{a}_1^{\text{row}}\right) \geq 0 \tag{81}$$

because $\|\mathbf{a}_1^{\text{row}}\| = \|\mathbf{a}_i^{\text{row}}\|$. For $|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})| > \sum_{i=2}^{N}|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})|$, the inequality in (81) implies that $\bar{\mathbf{w}}_1^T\bar{\mathbf{w}}_2 \geq 0$; see (79). In addition, because $\|\bar{\mathbf{w}}_2\|^2 \geq 0$, we obtain from (78) that $\gamma \geq \|\bar{\mathbf{w}}_1\|$ for all dither amplitudes that satisfy $\sum_{j=1}^{M} a_{ij}^2 = \nu^2$ for all $i \in \{1, 2, \ldots, N\}$, under the condition that $|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})| > \sum_{i=2}^{N}|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})|$. The equality $\gamma = \|\bar{\mathbf{w}}_1\|$ is obtained if and only if $\bar{\mathbf{w}}_2 = \mathbf{0}$. Given any vector $\mathbf{a}_1^{\text{row}}$, by choosing $\mathbf{a}_i^{\text{row}}$ to be equal to $\mathbf{a}_1^{\text{row}}$ multiplied by an appropriate sign for all $i \in \{2, 3, \ldots, N\}$, we can always get $\bar{\mathbf{w}}_2 = \mathbf{0}$; see (79). Therefore, it follows from $\gamma = \|\bar{\mathbf{w}}_1\|$ and (80) that the minimal value of γ is $(|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})| - \sum_{i=2}^{N}|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})|)\nu$ if $|\frac{dZ_{\mathcal{I}(1)}}{du_{\mathcal{I}(1)}}(\hat{u}_{\mathcal{I}(1)})| > \sum_{i=2}^{N}|\frac{dZ_{\mathcal{I}(i)}}{du_{\mathcal{I}(i)}}(\hat{u}_{\mathcal{I}(i)})|$.

Appendix B Proof of Lemma 8

The proof consists of three parts. First, we show that

$$|q_{1,i}(t)| \leq \max\left\{\frac{\eta}{\omega}, \frac{\kappa_1}{\eta\nu}\right\}\nu c_{q1} \tag{82}$$

for all $t \geq 0$ and some constant $c_{q1} \in \mathbb{R}_{>0}$. Second, we demonstrate that

$$\frac{1}{2\nu^2} \leq q_{2,i}(t) \leq \frac{3}{2\nu^2} \tag{83}$$

for all $t \geq 0$. Third, we prove the statement in (60) using the derived bounds in the first two parts.

Part 1: Let us define the variable

$$\tilde{q}_{1,i}(t) = q_{1,i}(t) + \eta\sum_{j=1}^{M} a_{i,j}(t)\int_0^t b_j(\omega\tau)d\tau. \tag{84}$$

By differentiating $\tilde{q}_{1,i}$ with respect to time and using (17) and (43), we obtain

$$\dot{\tilde{q}}_{1,i}(t) = -\eta\tilde{q}_{1,i}(t) + \dot{\hat{u}}_i(t) + \eta\sum_{j=1}^{M}(\eta a_{i,j}(t) + \dot{a}_{i,j}(t))\int_0^t b_j(\omega\tau)d\tau. \tag{85}$$

We note that $|\dot{\hat{u}}_i(t)| \leq \kappa_1$ [see (44)], $|a_{i,j}(t)| \leq \nu$ [see (26)], $|\dot{a}_{ij}(t)| \leq \lambda_2 L_A$ [see (42)], and that it follows from (18) and (20) that

$$\left|\int_0^t b_j(\omega\tau)d\tau\right| \leq \frac{c_b}{\omega} \tag{86}$$

for all $t \geq 0$ and some constant $c_b \in \mathbb{R}_{>0}$. Therefore, the last two terms in the right-hand side of (85) can be regarded as uniformly bounded disturbances. Using that $\tilde{q}_{1,i}(0) = q_{1,i}(0) = 0$, we may use a Lyapunov stability analysis with the candidate function $V_{q1}(\tilde{q}_{1,i}) = \tilde{q}_{1,i}^2$ to show that

$$|\tilde{q}_{1,i}(t)| \leq \max\left\{\frac{\eta}{\omega}, \frac{\kappa_1}{\eta\nu}\right\}\nu c_{q1,1} \tag{87}$$

for all $t \geq 0$, all $\lambda_2 \leq \eta\nu\epsilon_7$ and some constant $c_{\tilde{q}1} \in \mathbb{R}_{>0}$. By applying the same bounds as before, it follows from (84) that

$$|q_{1,i}(t)| \leq |\tilde{q}_{1,i}(t)| + \frac{\eta}{\omega}\nu c_{q1,2} \tag{88}$$

for all $t \geq 0$. Combining (87) and (88) gives the bound in (82).

Part 2: We define the variable

$$\tilde{q}_{2,i}(t) = q_{2,i}^{-1}(t) - \eta\sum_{j=1}^{M}\sum_{k=1}^{M} a_{ij}(t)a_{ik}(t)B_{jk}(t) - \nu^2 \tag{89}$$

with

$$B_{jk}(t) = \begin{cases} \int_0^t (b_j(\omega\tau)b_k(\omega\tau) & \text{if } j = k \\ \int_0^t b_j(\omega\tau)b_k(\omega\tau)d\tau, & \text{if } j \neq k. \end{cases} \tag{90}$$

Using (43) and the excitation condition in (26), the time derivative of $\tilde{q}_{2,i}$ in (89) is given by

$$\dot{\tilde{q}}_{2,i}(t) = -\eta\tilde{q}_{2,i}(t) + 2\eta p_i(t)q_{1,i}(t) + \eta q_{1,i}^2(t) - \eta\sum_{j=1}^{M}\sum_{k=1}^{M}\left(\eta a_{ij}(t)a_{ik}(t) + \frac{d}{dt}\left(a_{ij}(t)a_{ik}(t)\right)\right)B_{jk}(t). \tag{91}$$

From the definition of B_{jk} in (90), the orthogonality condition of the basis functions in (19) and the bound in (20), we have that

$$|B_{jk}(t)| \leq \frac{c_B}{\omega} \tag{92}$$

for some constant $c_B \in \mathbb{R}_{>0}$. In addition, using (17), (82), and the same bounds as in Part 1, we obtain from a Lyapunov stability analysis with candidate function $V_{q2,i}(\tilde{q}_{2,i}) = \tilde{q}_{2,i}^2$ that

$$|\tilde{q}_{2,i}(t)| \leq \frac{\nu^2}{4} \tag{93}$$

for all $t \geq 0$, $\eta \leq \omega\epsilon_4$, and $\kappa_1 \leq \eta\nu\epsilon_5$, where $\epsilon_4, \epsilon_5 > 0$ are sufficiently small. Subsequently, we obtain from (89) that

$$\left|q_{2,i}^{-1}(t) - \nu^2\right| \leq \frac{1}{2}\nu^2 \tag{94}$$

for all $t \geq 0$ and all $\eta \leq \omega\epsilon_4$, with a sufficiently small value of $\epsilon_4 > 0$, where we used the bounds in (26) and (92). The bounds in (83) follow from the inequality in (94).

Part 3: Let us define

$$\tilde{m}_i(t) = m_i(t) - Z_i(\hat{u}_i(t)) + q_{1,i}(t)\frac{dZ_i}{du_i}(\hat{u}_i(t)). \tag{95}$$

In addition, we define the Lyapunov-function candidate as

$$V_g(\tilde{m}_i, \tilde{g}_{Z,i}, q_{2,i}) = \tilde{m}_i^2 + \frac{\tilde{g}_{Z,i}^2}{q_{2,i}} \tag{96}$$

with $\tilde{g}_{Z,i}(t)$ in (59). For brevity, we omit the time argument in the following expressions. Using a first-order Taylor expansion of $Z_i(u_i)$, it follows that

$$Z_i(u_i) = Z_i(\hat{u}_i) + p_i \frac{dZ_i}{du_i}(\hat{u}_i) + w_i \tag{97}$$

with

$$w_i = p_i^2 \int_0^1 (1-s) \frac{d^2 Z_i}{du_i^2}(\hat{u}_i + sp_i) ds \tag{98}$$

see (17). By combining (43), (59), (95), and (97), we get that the time derivative of the Lyapunov-function candidate in (96) can be written as

$$\begin{aligned} \dot{V}_g(\tilde{m}_i, \tilde{g}_{Z,i}, q_{2,i}) &= -\eta(\tilde{m}_i + (p_i + q_{1,i})\tilde{g}_{Z,i})^2 \\ &\quad - \eta\left(\tilde{m}_i^2 + \frac{\tilde{g}_{Z,i}^2}{q_{2,i}}\right) + 2\eta(\tilde{m}_i + (p_i + q_{1,i})\tilde{g}_{Z,i})(\tilde{y}_i + w_i) \\ &\quad + 2\left(\tilde{m}_i q_{1,i} - \frac{\tilde{g}_{Z,i}}{q_{2,i}}\right) \frac{d^2 Z_i}{du_i^2}(\hat{u}_i)\dot{\hat{u}}_i. \end{aligned} \tag{99}$$

Subsequently, applying Young's inequality yields

$$\begin{aligned} \dot{V}_g(\tilde{m}_i, \tilde{g}_{Z,i}, q_{2,i}) &\leq -\frac{\eta}{2} V_g(\tilde{m}_i, \tilde{g}_{Z,i}, q_{2,i}) \\ &\quad + \eta(\tilde{y} + w_i)^2 + \frac{2}{\eta}\left(q_{1,i}^2 + \frac{1}{q_{2,i}}\right)\left(\frac{d^2 Z_i}{du_i^2}(\hat{u}_i)\dot{\hat{u}}_i\right)^2. \end{aligned} \tag{100}$$

Because the function Z_i in (9) is twice continuously differentiable (see Assumption 5), and because its domain $\mathcal{U}_i$ is compact, we have that $\frac{d^2 Z_i}{du_i^2}$ is uniformly bounded. Using this and (45), we obtain that w_i in (98) can be bounded by

$$|w_i| \leq \nu^2 c_w \tag{101}$$

for some constant $c_w \in \mathbb{R}_{>0}$. By applying the comparison lemma [19, Lem. 3.4], and by using $|\dot{u}_i(t)| \leq \kappa_1$ [see (44)] and the bounds in (82), (83), and (101), we obtain from (100) that

$$\begin{aligned} V_g(\tilde{m}_i(t), \tilde{g}_{Z,i}(t), q_{2,i}(t)) \leq \max\Big\{ &\nu^4, \sup_{\tau\in[0,t)} |\tilde{y}(\tau)|^2 \\ &e^{-\frac{\eta}{2}t} V_g(\tilde{m}_i(0), \tilde{g}_{Z,i}(0), q_{2,i}(0)) \Big\} c_V \end{aligned} \tag{102}$$

for all $t \geq 0$, $\eta \leq \omega\epsilon_4$, $\kappa_1 \leq \eta\nu\epsilon_5$, and some constant $c_V \in \mathbb{R}_{>0}$. From (53), (55), and (58), it follows that $\tilde{y}_i(t)$ is uniformly bounded under the conditions of Theorem 6. Therefore, we obtain from (83), (96), and (102) that $\tilde{m}_i(t)$ in (95) and $\tilde{g}_{Z,i}(t)$ in (59) are uniformly bounded. Moreover, the bound on the gradient estimation error in (60) follows from the same arguments.

References

[1] V. Adetola, D. DeHaan, and M. Guay, "Adaptive extremum-seeking receding horizon control of nonlinear systems," in *Proc. Amer. Control Conf.*, vol. 4, 2004, pp. 2937–2942.

[2] V. Adetola and M. Guay, "Parameter convergence in adaptive extremum-seeking control," *Automatica*, vol. 43, no. 1, pp. 105–110, 2007.

[3] K. B. Ariyur and M. Krstić, *Real-Time Optimization by Extremum-Seeking Control*. Hoboken, NJ, USA: Wiley, 2003.

[4] B. J. T. Binder, A. Pavlov, and T. A. Johansen, "Estimation of flow rate and viscosity in a well with an electric submersible pump using moving horizon estimation," *IFAC-PapersOnLine*, vol. 48, no. 6, pp. 140–146, 2015.

[5] P. Borg, T. Moen, and J. Aalbu, "Adaptive control of alumina reduction cell with point feeders," *Model., Identification, Control*, vol. 7, no. 1, pp. 45–56, 1986.

[6] H. C. Brinkman, "The viscosity of concentrated suspensions and solutions," *J. Chem. Phys.*, vol. 20, no. 4, pp. 571–571, 1952.

[7] S. L. Brunton, C. W. Rowley, S. R. Kulkarni, and C. Clarkson, "Maximum power point tracking for photovoltaic optimization using ripple-based extremum seeking control," *IEEE Trans. Power Electron.*, vol. 25, no. 10, pp. 2531–2540, Oct. 2010.

[8] J. Creaby, Y. Li, and J. E. Seem, "Maximizing wind turbine energy capture using multivariable extremum seeking control," *Wind Eng.*, vol. 33, no. 4, pp. 361–388, 2009.

[9] J. Ebegbulem and M. Guay, "Resource allocation for a class of multi-agent systems with unknown dynamics using extremum seeking control," in *Proc. IEEE Conf. Decis. Control*, Dec. 2018, pp. 2496–2501.

[10] R. D. Ensalzado, "Optimal diluent allocation in production systems with diluent-ESP-lifted wells," M.S. thesis, Dept. of Geosci. Petroleum, Norwegian Univ. of Sci. Technol., Trondheim, Norway, 2016.

[11] A. Ghaffari, M. Krstić, and D. Nešić, "Multivariable Newton-based extremum seeking," *Automatica*, vol. 48, no. 8, pp. 1759–1767, 2012.

[12] A. Ghaffari, M. Krstić, and S. Seshagiri, "Power optimization and control in wind energy conversion systems using extremum seeking," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 5, pp. 1684–1695, Sep. 2014.

[13] M. Guay, I. Vandermeulen, S. Dougherty, and P. J. McLellan, "Distributed extremum-seeking control over networks of dynamically coupled unstable dynamic agents," *Automatica*, vol. 93, pp. 498–509, 2018.

[14] M. Haring and T. A. Johansen, "Asymptotic stability of perturbation-based extremum-seeking control for nonlinear plants," *IEEE Trans. Autom. Control*, vol. 62, no. 5, pp. 2302–2317, May 2017.

[15] M. Haring and T. A. Johansen, "On the accuracy of gradient estimation in extremum-seeking control using small perturbations," *Automatica*, vol. 95, pp. 23–32, 2018.

[16] L. Hazeleger, M. Haring, and N. van de Wouw, "Extremum-seeking control for optimization of time-varying steady-state responses of nonlinear systems," *Automatica*, vol. 119, 2020, Art. no. 109068.

[17] J. D. Jansen, *Nodal Analysis of Oil and Gas Production Systems*. Richardson, TX, USA: Society of Petroleum Engineers, 2017.

[18] K. E. Johnson and G. Fritsch, "Assessment of extremum seeking control for wind farm energy production," *Wind Eng.*, vol. 36, no. 6, pp. 701–716, 2012.

[19] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2002.

[20] M. Krstić, "Performance improvement and limitations in extremum seeking control," *Syst. Control Lett.*, vol. 39, no. 5, pp. 313–326, 2000.

[21] R. J. Kutadinata, W. H. Moase, and C. Manzie, "Dither re-use in Nash equilibrium seeking," *IEEE Trans. Autom. Control*, vol. 60, no. 5, pp. 1433–1438, May 2015.

[22] R. Leyva, C. Alonso, I. Queinnec, A. Cid-Pastor, D. Lagrange, and L. Martínez-Salamero, "MPPT of photovoltaic systems using extremum-seeking control," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 1, pp. 249–258, Jan. 2006.

[23] X. Li, Y. Li, J. E. Seem, and P. Lei, "Detection of internal resistance change for photovoltaic arrays using extremum-seeking control MPPT signals," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 1, pp. 325–333, Jan. 2016.

[24] Y. Lin, E. D. Sontag, and Y. Wang, "A smooth converse Lyapunov theorem for robust stability," *SIAM J. Control Optim.*, vol. 34, no. 1, pp. 124–160, 1996.

[25] W. H. Moase, C. Manzie, and M. J. Brear, "Newton-like extremum-seeking for the control of thermoacoustic instability," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2094–2105, Sep. 2010.

[26] A. Pavlov and K. Fjalestad, "Method and system for the optimisation of the addition of diluent to an oil well comprising a downhole pump," U.S. Patent App. 15/769 997, Feb. 2019.

[27] A. Pavlov, M. Haring, and K. Fjalestad, "Practical extremum-seeking control for gas-lifted oil production," in *Proc. IEEE 56th Annu. Conf. Decis. Control*, 2017, pp. 2102–2107.

[28] A. Pavlov and B. H. Sannæs, "Experimental studies of ESP performance with two-phase fluids with live viscous oils," in *Proc. World Heavy Oil Congr.*, 2012.

[29] A. J. Peixoto, D. Pereira-Dias, A. F. S. Xaud, and A. R. Secchi, "Modelling and extremum seeking control of gas lifted oil wells," *IFAC-PapersOnLine*, vol. 48, no. 6, pp. 21–26, 2015.

[30] J. Poveda and N. Quijano, "Distributed extremum seeking for real-time resource allocation," in *Proc. Amer. Control Conf.*, 2013, pp. 2772–2777.

[31] T. L. Silva and A. Pavlov, "Dither signal optimization for multi-agent extremum seeking control," in *Proc. Eur. Control Conf.*, 2020, pp. 1230–1237.
[32] T. L. Silva and A. Pavlov, "Dither signal optimization in constrained multi-agent extremum seeking control," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 1633–1639, 2020.
[33] R. Suttner, "Extremum seeking control with an adaptive dither signal," *Automatica*, vol. 101, pp. 214–222, 2019.
[34] G. Takacs, *Electrical Submersible Pumps Manual: Design, Operations, and Maintenance*. Houston, TX, USA: Gulf, 2017.
[35] Y. Tan, W. H. Moase, C. Manzie, D. Nešić, and I. M. Y. Mareels, "Extremum seeking from 1922 to 2010," in *Proc. 29th Chin. Control Conf.*, Jul. 2010, pp. 14–26.
[36] Y. Tan, D. Nesic, and I. Mareels, "On the choice of dither in extremum seeking systems: A case study," *Automatica*, vol. 44, no. 5, pp. 1446–1450, 2008.
[37] Y. Tan, D. Nešić, and I. M. Y. Mareels, "On non-local stability properties of extremum seeking control," *Automatica*, vol. 42, no. 6, pp. 889–903, 2006.
[38] G. Vachtsevanos and K. Kalaitzakis, "A hybrid photovoltaic simulator for utility interactive studies," *IEEE Trans. Energy Convers.*, vol. EC-2, no. 2, pp. 227–231, Jun. 1987.
[39] D. Wang, M. Chen, and W. Wang, "Distributed extremum seeking for optimal resource allocation and its application to economic dispatch in smart grids," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 10, pp. 3161–3171, Oct. 2019.
[40] M. Ye and G. Hu, "A distributed extremum seeking scheme for networked optimization," in *Proc. 54th IEEE Conf. Decis. Control*, 2015, pp. 4928–4933.
[41] C. Zhang and R. Ordóñez, *Extremum-Seeking Control and Applications*. Berlin, Germany: Springer-Verlag, 2012.

Mark Haring received the B.Sc. and M.Sc. degrees in mechanical engineering from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2008 and 2011, respectively, and the Ph.D. degree in engineering cybernetics from the Norwegian University of Science and Technology, Trondheim, Norway, in 2016.

He is currently working as a Research Scientist with the Department of Mathematics and Cybernetics, SINTEF Digital, Trondheim. His research interests include automatic control, adaptive control, machine learning, and estimation.

Synne Fossøy received the M.Sc. degree in cybernetics and robotics from the Norwegian University of Science and Technology, Trondheim, Norway, in 2020.

She is currently working as a researcher with the Department of Mathematics and Cybernetics, SINTEF Digital, Trondheim.

Thiago Lima Silva received the B.Sc. degree in computer science from the Federal University of Bahia, Salvador, Brazil, in 2010, and the M.Sc. and Ph.D. degrees in systems and automation engineering from the Federal University of Santa Catarina (UFSC), Florianópolis, Brazil, in 2012 and 2017, respectively.

He is currently a Research Scientist with the Department of Sustainable Energy Technology, SINTEF Industry, Trondheim, Norway. He was a Visiting Researcher with the Department of Engineering Cybernetics, Norwegian University of Science and Technology (NTNU), from 2015 to 2016, and a Postdoctoral Fellow with the Department of Systems and Automation Engineering, UFSC, from 2017 to 2018, and the Department of Geoscience and Petroleum, NTNU, from 2018 to 2021. His current research interests include systems optimization, adaptive control, and machine learning.

Alexey Pavlov (Senior Member, IEEE) received the M.Sc. degree in applied mathematics from St. Petersburg State University, Saint Petersburg, Russia, in 1998, and the Ph.D. degree in mechanical engineering from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2004.

He is currently a Full Professor with the Department of Geoscience and Petroleum, Norwegian University of Science and Technology (NTNU), Trondheim, Norway, and leads the NTNU Research and Innovation Program in Digital and Automation Solutions for the Oil and Gas Industry. He was a Researcher with the Russian Academy of Sciences, St. Petersburg, in 1999, a Visiting Researcher with the Ford Research Laboratory, Dearborn, MI, USA, in 2000, and a Postdoctoral with the Eindhoven University of Technology in 2005 and the NTNU from 2005 to 2009. From 2009 to 2016, he was a Principal Researcher with Equinor Research Center, Porsgrunn, Norway. He has authored and coauthored a large number of papers including the book *Uniform Output Regulation of Nonlinear Systems: A Convergent Dynamics Approach*, with N. van de Wouw and H. Nijmeijer (Birkhäuser, 2005), and holds four patents. His current research interests include data-driven optimization, industrial control applications, and optimization of energy production systems.

Prof. Pavlov was the recipient of the IFAC World Congress Application Paper Award in 2011 and IEEE Control Systems Technology Award in 2015.