

A Multivariable Extremum Seeking Control Approach using Fast Fourier Transform for Gradient Estimation

Dinesh Krishnamoorthy

Abstract—This paper considers a network of subsystems working towards optimizing the sum of their local objectives, where the overall objective is the only available measurement. This paper proposes a multivariable extremum seeking scheme using Fast Fourier Transform (FFT) for optimizing such multi-input single-output (MISO) systems. In the proposed method, the different inputs are perturbed with different dither frequencies, and the amplitude spectrum of the overall output signal obtained using FFT is used to estimate the steady-state cost gradient w.r.t. each input channel. The inputs for the subsystems are then updated using integral control in order to drive the respective gradients to zero. The paper provides analytical rules for designing the FFT-based gradient estimation algorithm and analyzes the stability properties of the proposed extremum seeking scheme. The effectiveness of the proposed FFT-based multivariable extremum seeking scheme is demonstrated using two examples: a wind farm power optimization problem, and a heat exchanger network for industrial waste-heat recovery.

Index Terms—Adaptive control, Extremum seeking control, Fourier transform, Real-time optimization

I. INTRODUCTION

EXTREMUM seeking control (ESC) has emerged as a popular real-time optimization (RTO) and adaptive control tool to optimize the steady-state output of a plant, despite lack of a mathematical model. In extremum seeking control, the steady-state gradient of the cost with respect to the input is directly estimated from the cost measurement, which is then driven to zero using an integral controller, thereby satisfying the necessary condition of optimality (NCO). The most popular approach to estimating the gradient is the classical sinusoidal perturbation-based extremum seeking control that dates back to the work of Draper and Li [1], and later gained popularity with the work of Krstic and Wang [2].

The classical extremum seeking control is based on perturbing the input with a slow sinusoidal dither signal, such that the plant appears as a static map. A high-pass filter is used to remove the DC-component of the cost measurement, and the resulting detrended sinusoidal response of the output signal (with zero mean) is correlated with the perturbation signal, known as demodulation. The product of the two sinusoids will have a DC component, which is extracted using a low-pass filter. This is approximately the steady-state cost gradient as explained in [2], [3]. The classical extremum seeking control

thus involves the use of a high-pass and low-pass filter such that the dither signal is in the pass band of the filters [3].

Although extremum seeking control was predominantly developed for single-input single-output (SISO) systems, recently there is a surge of interest in developing extremum seeking control for large-scale networks with several subsystems, such as wind farm power optimization [4]–[7], maximum power point tracking for photovoltaic grids [6], [8], oil production network [9], [10], heat exchanger networks [11], [12], formation flight control [13] etc. to name a few. Development of extremum seeking control for such multi-input systems can be broadly categorized into multivariable ESC and distributed ESC.

Multivariable extremum seeking control considers a multi-input single-output (MISO) case, where the objective is to estimate the steady-state gradient of the overall cost w.r.t. each input using a single extremum seeking loop [8], [14]. Distributed extremum seeking on the other hand assumes that each agent in the network has access to its local cost measurement. Each subsystem then locally employs single-input single-output (SISO) extremum seeking loop, and a consensus estimation algorithm is used for coordination among the different agents as explained in [4], [15].

This paper deals with the former class, where we assume that the different subsystems in a network work towards optimizing the sum of their local objectives, however the local cost of each subsystem is not available. This makes decentralized and distributed ESC approaches unsuitable for such systems. Maximizing the total power from a wind or solar farm, maximizing oil and gas production from a network of wells, maximizing heat recovery in a heat exchanger network, minimizing total energy consumption in a processing plant or in a commercial building are a few examples of such systems.

As noted in [2], [3], the choice of the filter cut-off frequencies in the classical extremum seeking approach depends on the input perturbation frequency. With multiple inputs and a single output, tuning and designing the filter cut-off frequencies to retrieve the gradient w.r.t each input channel from the single cost measurement may not be trivial. This is due to the fact that the output signal now contains the frequency components of the different input perturbations. Moreover, using the same low-pass filter on each channel may also lead to additional harmonics in the gradient estimation as noted in [8]. In short, good performance of the multivariable extremum seeking control using this approach requires good tuning of the filter cut-off frequencies, which can be challenging and time consuming.

A more natural approach to analyze the effect of differ-

Financial support from the Research Council of Norway through the IKTPLUSS program is gratefully acknowledged.

Dinesh Krishnamoorthy is with the Department of Chemical Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway dinesh.krishnamoorthy@ntnu.no

ent frequency components in a signal is to use the Fourier transform for spectral analysis, which maps the time series data into a series of frequencies. The Fourier transform and the inverse Fourier transform are complimentary mathematical pairs that allow the equivalent information to be represented in the time domain or the frequency domain. The proposed approach is based on the idea that the Fourier transform decomposes a signal and provides the magnitude and the phase of the different frequency components that make it up. One can then simply extract the frequency components of interest from the amplitude and phase spectrum to estimate the steady-state gradients.

Fast Fourier transform (FFT) is also an easy, robust, and effective numerical approach to estimate the steady-state gradient in the multi-input single-output case. In fact, the classical extremum seeking control can be seen as a special case of the Fourier transform, where instead of demodulating the cost measurement with sinusoids of different frequencies, in the classical ES scheme, the cost measurement is demodulated with a sinusoid of the same frequency as the input dither signal. Therefore, it is only natural to consider the more generic fast Fourier transform in the multivariable setting. More importantly, by using FFT, we eliminate the high pass and low pass filters, which otherwise complicates the tuning. On the other hand, since the proposed approach is based on the N -point DFT, the length of the window frame N is the only tuning parameter in the gradient estimation scheme. This paper also provides analytical rules on the choice of the window frame length N .

Related work: In the context of extremum seeking control, Fourier transform has been predominantly used to extract the phase information in order to compensate for the phase shift introduced by the plant dynamics, such that the static map assumption can be relaxed [16]–[18]. For example, the authors in [17] estimate the different harmonics using a Kalman filter, and drive the high frequency harmonics to zero when the perturbation and the plant dynamics are in the same timescale. On the other hand, in this paper we use FFT to directly extract the gradient information w.r.t multiple input channels from the amplitude spectrum using the two-timescale separation principle (just as in the case with classical extremum seeking control). The phase spectrum may be used only to estimate the sign of the gradient. The intuitive nature of FFT and the amplitude spectrum has previously led to the use of gradient estimation using FFT in a few engineering applications such as [19], [20]. However, these works provide little information into designing the FFT-based extremum seeking scheme and study its properties, which this paper aims to address.

In addition to extremum seeking control, estimating the steady-state cost gradients directly from the cost measurement is also used commonly in other real-time optimization frameworks commonly studied in the process control literature. For example NCO-tracking control [21] and modifier adaptation [22] rely on model-free estimation of the cost-gradient. These methods typically employ univariate model-free gradient estimation techniques [23], which may not be well suited in many large chemical processes, where the cost is typically measured downstream, and is often a sum of various flow

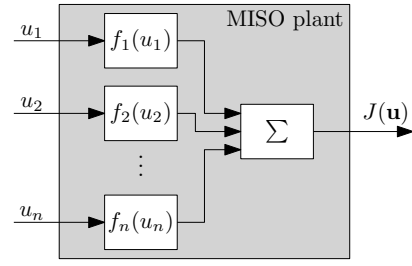


Fig. 1: Schematic representation of a multi-input single-output (MISO) network of subsystems working towards optimizing the sum of their local objectives. Everything inside the gray box is unknown and not accessible.

streams. The FFT-based gradient estimation proposed in this paper can also be used with other such RTO schemes that rely on model-free cost gradient estimation. Particularly in the context of a large-scale chemical processing networks, the overall cost is measured downstream, and the use of spectral analysis for gradient estimation makes it a very good tool to analyze the effect of various input perturbations in downstream measurements.

Contributions: The main contribution of this paper is a multivariable extremum seeking control framework using fast Fourier transform (FFT) for the gradient estimation. We show that the FFT is a natural and simple approach to extract the gradient of the cost with respect to each input channel. The paper also provides analytical tuning rules on the choice of the window frame length N , and analyzes the stability properties of the proposed approach, using a similar framework as used in [24]. The effectiveness of the FFT-based extremum seeking scheme is demonstrated using two case examples: 1) maximize total power generation from a network of wind turbines, 2) maximize heat recovery from a network of heat exchangers.

II. EXTREMUM SEEKING CONTROL USING FAST FOURIER TRANSFORM (FFT)

A. Problem formulation

Consider a network of n subsystems working towards a optimizing the sum of their local objectives as shown in Fig. 1. The inputs are denoted by $\mathbf{u} = [u_1, \dots, u_n]^T \in \mathbb{R}^n$, and the overall cost function $J(\mathbf{u}) \in \mathbb{R}$ is denoted by the static map

$$J(\mathbf{u}) = \sum_{i=1}^n f_i(u_i) \quad (1)$$

where $f_i(u_i)$ denotes the unknown functions representing the local cost for each agent i . The objective is to maximize the overall cost of the network

$$\max_{\mathbf{u} \in \mathbb{R}^n} J(\mathbf{u}) \quad (2)$$

Assumption 1: The overall cost J is the only measurement that is available. The local cost measurements $f_i(u_i)$ are not available.

The above assumption implies that we consider a multi-input single-output (MISO) plant. Without loss of generality, we state the following assumption for a maximization problem.

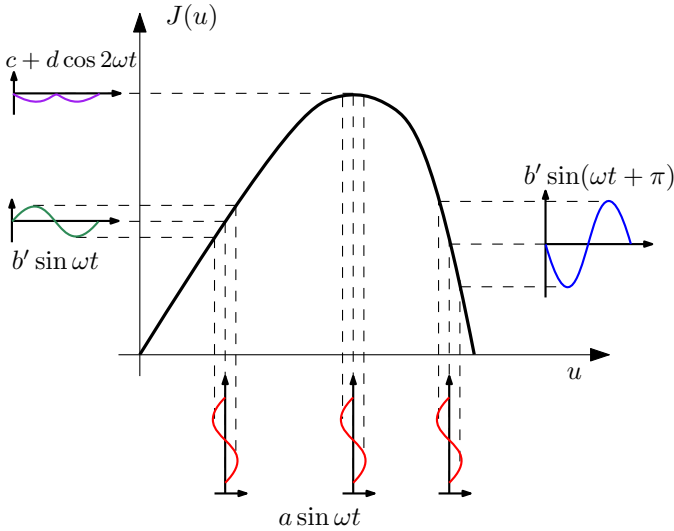


Fig. 2: Effect of sinusoidal perturbation of the input-output static map at different points.

Assumption 2: The nonlinear map (1) is sufficiently smooth and continuously differentiable such that $\mathbf{u} = \mathbf{u}^*$ is the unique maximizer of (2), such that

$$\frac{\partial J}{\partial u_i}(u_i^*) = 0 \quad \forall i = 1, \dots, n \quad (3)$$

$$\frac{\partial J}{\partial u_i}(u_i^* + \tilde{u}_i)\tilde{u}_i < 0 \quad \forall \tilde{u}_i \neq 0, \forall i = 1, \dots, n \quad (4)$$

where $\tilde{u}_i := u_i - u_i^*$

In this paper, we study the multivariable extremum seeking control of such a process, without the need for a mathematical model of the plant. The objective is to steer the system to the equilibrium \mathbf{u}^* using only the overall cost measurement J . To achieve optimal operation, we need a gradient estimator that continuously estimates the steady state cost gradient, which can then be driven to zero using an integral controller thereby satisfying (3).

B. Proposed Approach

The underlying idea behind the proposed FFT-based gradient estimation scheme is as follows: When we add a slow sinusoidal dither signal to the input, the output also varies periodically at the same frequency, but with a different amplitude based on the static map from the input to the output. The amplitude spectrum, which shows the strength of the variations of a signal at different frequencies, then shows the amplitude of the output signal variation at the perturbation frequency. From this we can estimate the magnitude of the cost gradient.

Consider a static map $J(u)$, as depicted in Fig. 2. For $u < u^*$, a sinusoidal perturbation in the input $a \sin \omega t$, generates a sinusoidal variation in the output at the same frequency $b \sin \omega t$ ¹. The power spectrum of the output signal would then have an amplitude of b at ω . Similarly, for $u > u^*$, a sinusoidal perturbation in the input $a \sin \omega t$, generates a sinusoidal variation in the output at the same frequency, shifted by 180°,

¹Note that for the static map setting, there will be no phase shift.

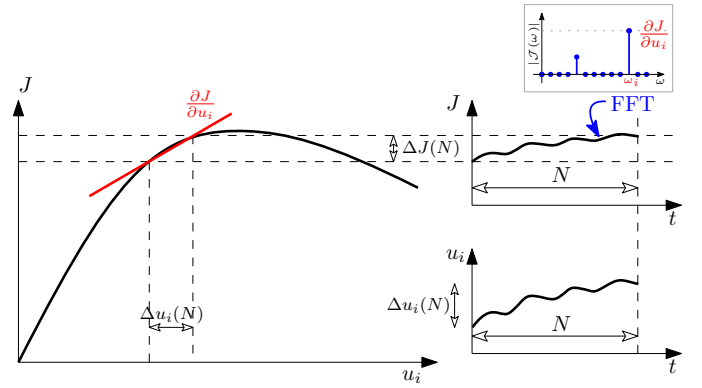


Fig. 3: Illustration of the local gradient estimation with respect to u_i using spectral analysis of the past N samples of cost measurement.

$b' \sin(\omega t + \pi)$. Also in this case, the power spectrum of the output signal would have an amplitude of b' at ω . For $u \approx u^*$, a sinusoidal perturbation in the input $a \sin \omega t$, generates a periodic signal with twice the frequency. Consequently, the power spectrum would indicate an amplitude of zero at ω , since power of the signal at this frequency is zero. Instead the power spectrum would have a non-zero amplitude at 2ω . For example, if we have a quadratic static map $y = (u - u^*)^2$, adding $a \sin \omega t$ at u^* would result in an output variation of $y = a^2 \sin^2(\omega t) = 0.5a^2(1 - \cos(2\omega t))$, which clearly only has a frequency component at 2ω .

In the multivariable case, each input channel is superimposed with a unique sinusoidal dither signal $a_i \sin \omega_i t$, with amplitude a_i and frequency ω_i . The overall output cost measurement J is then composed of the frequency components ω_i , for all $i = 1, \dots, n$. Using spectral analysis we can decompose the output signal J into different frequency components. To this end, we use a sliding window of fixed length with N past data points to get the N -point DFT using the finite set of cost measurement $[J(0), \dots, J(N-1)]^T$, and input signals $[u_i(0), \dots, u_i(N-1)]^T$ for all $i = 1, \dots, n$. The amplitude spectrum is then used to estimate the steady-state gradients $\frac{\partial J}{\partial u_i}$. This is also graphically illustrated in Fig. 3.

The proposed extremum seeking scheme is thus based on a two timescale separation. The perturbation frequencies are chosen to be sufficiently slow such that the dynamic plant behaves like a static map. The integral gain that drives the gradient to zero is chosen sufficiently small such that the static map can be assumed to be locally linear inside the window with the past N data points.

To perform fast Fourier transform (FFT), the signals need to be detrended such that they have zero mean. Hence, the DC-component from the time domain data is removed, for example using a moving average filter. The detrended signals with zero-mean are denoted by the superscript $(\cdot)^0$.

Remark 1: The detrending step is analogous to the high pass filter in the classical extremum seeking control, where the purpose of the high pass filter is also to remove the DC component from the cost measurement J [2].

Let the past N discrete samples of the detrended output cost measurement J^0 , be denoted by $[J^0(0), J^0(1), \dots, J^0(N-1)]^T$.

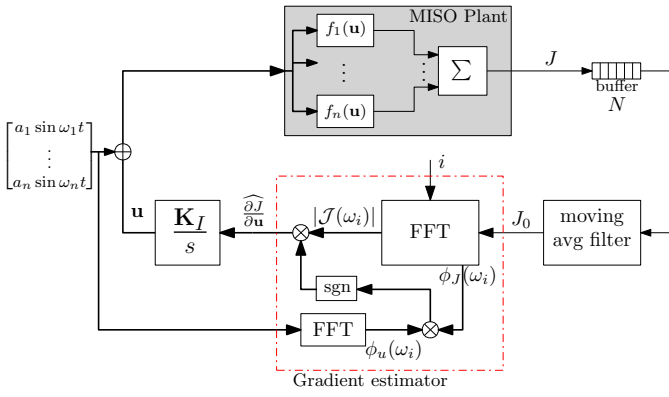


Fig. 4: Block diagram of the proposed FFT-based extremum seeking scheme.

1)]^T. The discrete-time Fourier transform (DTFT) reads as

$$\mathcal{J}(\omega) = \sum_{k=-\infty}^{\infty} J^0(k) e^{-j\omega k} \quad (5)$$

The N -point DFT samples $\mathcal{J}(\omega)$ at N frequency components given by,

$$\mathcal{J}(l) = \sum_{k=0}^{N-1} J^0(k) e^{-j\frac{2\pi}{N}lk} \quad \forall l = 0, \dots, N-1 \quad (6)$$

$\mathcal{J}(l)$ is a complex valued function, which has a corresponding magnitude spectrum denoted by $|\mathcal{J}(l)|$, and a phase spectrum denoted by $\phi_J(l)$. The magnitude of the cost function for the n input dither frequencies ω_i for all $i = 1, \dots, n$ can then be obtained from the single-sided amplitude spectrum $2|\mathcal{J}(l)|$ for all $l = 1, \dots, N/2$, where we extract the frequency components of interest from the N -point DFT. That is, the magnitude of the cost gradient w.r.t. the i^{th} input channel is given by

$$\left| \frac{\partial J}{\partial u_i} \right| = \frac{2}{a_i} |\mathcal{J}(\omega_i N/2\pi)| \quad \forall i = 1, \dots, n \quad (7)$$

Hereafter, we denote $|\mathcal{J}(\omega_i N/2\pi)|$ simply as $|\mathcal{J}(\omega_i)|$ for the sake of notational simplicity. Since the amplitude spectrum $|\mathcal{J}(\omega_i)| > 0$, the phase information $\phi_J(\omega_i)$ with respect to the phase of the input signal $\phi_{u_i}(\omega_i)$ is used to determine the sign of the gradient. The phase $\phi_J(\omega_i)$ is obtained from the FFT, and similarly, the phase of the input perturbations $\phi_{u_i}(\omega_i)$ can be obtained using FFT of the dither signals.

Together, the gradient of the cost w.r.t. to the i^{th} input is then estimated as follows:

$$\widehat{\frac{\partial J}{\partial u_i}} = \frac{2}{a_i} |\mathcal{J}(\omega_i)| \operatorname{sgn} \left[\frac{\phi_J(\omega_i)}{\phi_{u_i}(\omega_i)} \right], \quad \forall i = 1, \dots, n \quad (8)$$

This is schematically shown in Fig. 3 for the i^{th} input, where spectral analysis of the cost measurement from the past N samples is used to estimate the local gradient $\widehat{\frac{\partial J}{\partial u_i}}$. Alternatively, the estimated gradient can equivalently be computed directly as

$$\widehat{\frac{\partial J}{\partial u_i}} = \frac{\Re(\mathcal{J}(\omega_i))}{\Re(\mathcal{U}_i(\omega_i))}, \quad \forall i = 1, \dots, n$$

Algorithm 1 FFT-based Extremum seeking control

Input: $N, \omega_1, \dots, \omega_n, a_1, \dots, a_n$ and K_i

- 1: $J^0 \leftarrow$ Detrend the output cost measurement J
- 2: $\mathcal{J}(l) = \sum_{k=0}^{N-1} J^0(k) e^{-j\frac{2\pi}{N}lk}$ ▷ perform FFT
- 3: **for** $i = 1, \dots, n$ (in parallel) **do**
- 4: $\mathcal{U}_i(l) = \sum_{k=0}^{N-1} u_i^0(k) e^{-j\frac{2\pi}{N}lk}$ ▷ perform FFT
- 5: $|\mathcal{J}(\omega_i)|, \phi_J(\omega_i) \leftarrow$ amplitude and phase of the cost at the i^{th} input frequency
- 6: $|\mathcal{U}_i(\omega_i)|, \phi_{u_i}(\omega_i) \leftarrow$ amplitude and phase of the input at the i^{th} input frequency
- 7: $\widehat{\frac{\partial J}{\partial u_i}} \leftarrow \frac{2}{a_i} |\mathcal{J}(\omega_i)| \operatorname{sgn} \left[\frac{\phi_J(\omega_i)}{\phi_{u_i}(\omega_i)} \right]$
- 8: $\dot{u}_i = -K_i \widehat{\frac{\partial J}{\partial u_i}}$
- 9: **end for**

Output: \mathbf{u}

where $\Re(\cdot)$ denotes the real part of a complex number.

Remark 2: The Euler's formula establishes the fundamental relationship between the complex exponential function and the trigonometric functions in (6). Thus, the product of the detrended cost J_0 and exponential terms in (6) can be seen as analogous to the product of the detrended output signal and the sinusoidal dither in classical extremum seeking control, which provides the demodulation effect at the perturbation frequency.

Once the gradient is estimated, the inputs can be updated using a gradient descent step, which is simply an integral controller to drive the estimated gradients to zero,

$$\dot{u}_i = -K_i \widehat{\frac{\partial J}{\partial u_i}}, \quad \forall i = 1, \dots, n \quad (9)$$

where K_i is the integral gain (adaptation gain) for the i^{th} input. The perturbation signals are added to the nominal input \mathbf{u} computed by the integral controller, and the resulting input is applied on the plant, as shown in Fig. 4. The sketch of the proposed FFT-based extremum seeking control is shown in Algorithm 1.

III. STABILITY PROPERTIES

- Assumption 3 (Choice of dither signal):*
- (i) $\max(\{\omega_i\})$ is chosen such that the plant can be assumed to be a static map.
 - (ii) The output signal is persistently exciting of at least order $2n$.

Assumption 3(ii) can be satisfied if the perturbation frequencies of the different input channels are mutually independent, i.e. $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$, $\omega_i + \omega_j \neq \omega_k$ for any distinct i, j , and k . This means that the frequency of one input perturbation must not be the same or lie in the harmonics of perturbation of the other input channels. This is to ensure that unique persistence of excitation condition is satisfied [14].

When we use an N -point DFT to obtain the amplitude spectrum, it is based on spectral analysis on a finite set of data with length N . The choice of the window length is an important factor which affects the accuracy of the gradient estimation. If N is not carefully chosen, then it is possible

that the N -point DFT may not sample DTFT precisely at one or more of the perturbation frequencies. That is, the discrete frequency array $l = \{0, \dots, N-1\}$ may not contain one or more of the perturbation frequencies exactly (cf. (7)). This leads to spectral leakage resulting in bias in the spectral values, and consequently bias in the estimated gradient. The choice of the window length and its effect on the gradient estimation is formalized in the following theorem.

Theorem 1 (Minimum window length N): Given assumption 3, the length of the moving window frame N that ensures no spectral leakage is given by

$$N = \gamma \text{lcm} \left(\left\{ \frac{2\pi}{\omega_i} \right\}_{i=1}^n \right) \quad (10)$$

for some $\gamma \in \mathbb{Z}_+$, where $\text{lcm}(\cdot)$ denotes the least common multiple.

Proof: We can consider the finite segment of a signal $x_N(k)$ as a product of $x(k)$ and a rectangular data window

$$w(k) := \begin{cases} 1 & 0 \leq k \leq N-1 \\ 0 & \text{elsewhere} \end{cases}$$

The Fourier transform of $x_N(k)$ is the given by

$$\mathcal{X}_N(\omega) = \mathcal{X}(\omega) \otimes \mathcal{W}(\omega) \quad (11)$$

where for a rectangular window

$$\mathcal{W}(\omega) = \left[\frac{\sin(\omega N/2)}{\sin(\omega/2)} \right] e^{-j\omega(N-1)/2}$$

The periodic *sinc* function leads to mainlobes and sidelobes in the DTFT. Sampling this at $\omega = 2\pi l/N$ for $l = 0, \dots, N-1$ one can see that

$$\mathcal{W}(l) = \begin{cases} N & l = 0 \\ 0 & l = 1, \dots, N-1 \end{cases}$$

which implies that the sidelobes are sampled at their zero values. By choosing N as (10) we ensure $\omega_i \in \left\{ \frac{2\pi l}{N} \right\}_{l=0}^{N-1}$, and this ensures that there is no spectral leakage in (11). ■

An alternative way to understand the proof of Theorem 1 is by looking at the N -point DFT in (6), where it can be seen that this is periodic with period N . The inverse DFT of (6) expressed as

$$J'_0(k) = \frac{1}{N} \sum_{l=0}^{N-1} \mathcal{J}(l) e^{j \frac{2\pi}{N} l k} \quad (12)$$

tells us that $J'_0(k)$ is also periodic with period N . This is because sampling in the time domain results in periodicity in the frequency domain, and vice versa as given by the DFT sampling theorem. This implies that the two end points of the time series samples, $J_0(0)$ and $J_0(N)$, are interpreted as though they were connected together [25, Ch. 5]. Therefore, by choosing N as the least common multiple of the perturbation time periods (10), we can reconstruct $J'_0(k) = J_0(k)$ that contains integer number of full time periods of the all perturbation frequencies $\left\{ \frac{2\pi}{\omega_i} \right\}_{i=1}^n$. If N is not chosen according to (10), then the reconstructed $J'_0(k)$ contains discontinuities, which results in higher order harmonics in the power spectrum. In this

case, the DFT samples the sidelobes of the DTFT at non-zero values leading to *spectral leakage*. Therefore the minimum window length N_{min} that must used in the FFT-based gradient estimation is given by (10) with $\gamma = 1$.

Remark 3: The proposed approach can be seen as a variant of the Gabor transform. That is, the use of a moving window to estimate the time-varying local gradient using FFT is similar to using a Gabor transform [26]–[28], where instead of a moving Gaussian window, we use a moving rectangular window.

Due to the local linear approximation, as the nominal input u_i changes, using the past N data samples obtained at time t , denoted by $\Delta u_i(t) \geq 0$, may lead to some deviations in the estimated gradient $\widehat{\frac{\partial J}{\partial u_i}}$ from the true gradient $\frac{\partial J}{\partial u_i}(u_i(t))$ at time t . See Appendix I for an illustrative example that provides an intuitive understanding. This is now formalized in the following Lemma.

Lemma 1 (Boundedness of the gradient estimation error): Consider the FFT-based gradient estimation scheme given by (6)-(8), where the gradient estimated using the past N data samples is denoted as $\widehat{\frac{\partial J}{\partial u_i}}$. Let the true steady-state cost gradient w.r.t the inputs $u_i(t)$ be denoted as $\frac{\partial J}{\partial u_i}(u_i(t))$. Given Assumptions 2 and 3, and N chosen according to Theorem 1, the gradient estimation error

$$\epsilon_i(t) := \widehat{\frac{\partial J}{\partial u_i}} - \frac{\partial J}{\partial u_i}(u_i(t)) \quad (13)$$

has an upper bound given by,

$$|\epsilon_i(t)| \leq K_i N \sup_{\tau \in [0, N]} |\mathbf{H}_i(u_i(t - \tau))| \sup_{\tau \in [0, 2N]} \left| \frac{\partial J}{\partial u_i}(u_i(t - \tau)) \right| \quad (14)$$

where $\mathbf{H}_i := \frac{\partial^2 J}{\partial u_i^2}$ is the Hessian of the static map.

Proof: Since Assumption 2 holds, we can use the mean-value theorem to estimate the absolute error, where for some u_i in the interval $\Delta u_i(t) := \max_{\tau \in [0, N]} u_i(t - \tau) - \min_{\tau \in [0, N]} u_i(t - \tau)$, we have

$$|\epsilon_i(t)| \leq \sup_{\tau \in [0, N]} |\mathbf{H}_i(u_i(t - \tau))| \Delta u_i(t) \quad (15)$$

Furthermore, since u_i is given by the integral controller (9), it can be shown that

$$\Delta u_i(t) \leq K_i N \sup_{\tau \in [0, 2N]} \left| \frac{\partial J}{\partial u_i}(u_i(t - \tau)) \right| \quad (16)$$

Combining (15) and 16, we have

$$|\epsilon_i(t)| \leq K_i N \sup_{\tau \in [0, N]} |\mathbf{H}_i(u_i(t - \tau))| \sup_{\tau \in [0, 2N]} \left| \frac{\partial J}{\partial u_i}(u_i(t - \tau)) \right|$$

This implies that the upper bound on the gradient estimation error is higher for large values of N or K_i . It is also intuitive that the error bound is higher for large Δu_i as shown in (15). ■

Theorem 2 (Stability): Consider the FFT-based extremum seeking scheme given by (6)-(9). Let Assumptions 2 and 3

hold, and N chosen according to Theorem 1. Then there exists $K_i N$ small enough such that

$$\lim_{t \rightarrow \infty} u_i(t) = u_i^*$$

Proof: To analyze the stability of the optimum $u_i = u_i^*$, consider the candidate Lyapunov function,

$$V(u_i) = \frac{1}{2}(u_i - u_i^*)^\top (u_i - u_i^*) =: \frac{1}{2} \tilde{u}_i^\top \tilde{u}_i \quad (17)$$

with $\tilde{u}_i = u_i - u_i^*$, and the derivative is written as,

$$\begin{aligned} \dot{V}(u_i) &= \tilde{u}_i^\top \dot{\tilde{u}}_i \\ &= -\tilde{u}_i^\top K_i \frac{\partial J}{\partial u_i} \end{aligned} \quad (18)$$

$$= -\tilde{u}_i^\top K_i \left(\frac{\partial J}{\partial u_i}(u_i(t)) + \epsilon(t) \right) \quad (19)$$

Since this is a maximization problem, we have $K_i < 0$ for all i . From (4) it can be seen that the first term will always be non-negative, that is $\tilde{u}_i^\top K_i \frac{\partial J}{\partial u_i} \geq 0$. In the case with no gradient estimation error, i.e. $\frac{\partial J}{\partial u_i} = \frac{\partial J}{\partial u_i}$, we then have “nominal” stability, since $\dot{V} < 0$.

In the presence of gradient estimation error, for the system to be asymptotically stable, we require that $\dot{V} < 0$. From (4), we have $\dot{V} < 0$ when both $\frac{\partial J}{\partial u_i} < 0$ and $\epsilon < 0$ or $\frac{\partial J}{\partial u_i} > 0$ and $\epsilon > 0$.

However, when $\frac{\partial J}{\partial u_i} > 0$ and $\epsilon < 0$ or $\frac{\partial J}{\partial u_i} < 0$ and $\epsilon > 0$, then we need $|\epsilon| < \left| \frac{\partial J}{\partial u_i}(u_i(t)) \right|$ in order to ensure $\dot{V} < 0$. From Lemma 1, we can choose $K_i N$ small enough for a given N

$$K_i < \frac{1}{N \bar{H}_i} \quad (20)$$

with \bar{H}_i denoting the upper bound on the Hessian, such that this condition is satisfied, and we have $\dot{V} < 0$. Hence

$$\lim_{t \rightarrow \infty} u_i(t) = u_i^*$$

This can also be understood intuitively for a sufficiently small K_i . Assume that the controller drives the inputs u_i to some point u_i' . From Lemma 1, as $\Delta u_i(t) \rightarrow 0$, the gradient estimation error diminishes, and

$$\lim_{t \rightarrow \infty} \frac{\partial \hat{J}}{\partial u_i} = \frac{\partial J}{\partial u_i}(u_i(t))$$

and we have $\dot{V} < 0$.

If $\Delta u_i = 0$ while $\frac{\partial \hat{J}}{\partial u_i} \neq 0$, then this contradicts the controller dynamics (9). Therefore, from Assumption 2

$$\lim_{t \rightarrow \infty} \frac{\partial \hat{J}}{\partial u_i} = \frac{\partial J}{\partial u_i}(u_i^*) = 0$$

and

$$\lim_{t \rightarrow \infty} u_i(t) = u_i^*$$

To summarize, Theorem 1 provides a lower bound on the length of the window frame N as a function of the different input perturbation frequencies such that the gradient w.r.t all the inputs can be retrieved exactly from the amplitude spectrum of the output signal, and Theorem 2 provides an upper bound on the controller gains K_i such that the closed-loop system is stable.

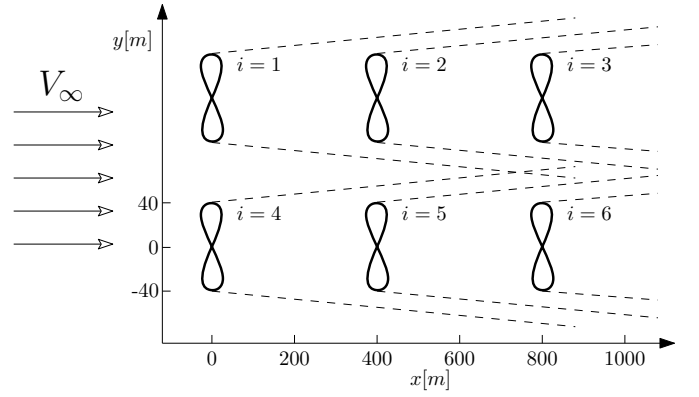


Fig. 5: Wind farm layout with $n = 6$ turbines with diameter 80 m, placed at (0,200), (400,200), (800,200) and (0,0), (400,0), (800,0).

IV. CASE EXAMPLES

In this section, we present two case examples that demonstrate the performance of the proposed FFT-based gradient estimation scheme in large-scale networks. The simulation results presented in this section were carried out using MATLAB v2020b, and the software can be found in the GitHub repository [29].

A. Example 1: Wind farm power optimization

In this example, we consider a wind farm with $n = 6$ identical wind turbines as shown in Fig. 5. We assume the wind direction is along the x co-ordinate and orthogonal to the y co-ordinate, as shown in Fig. 5. The wind turbines are uniform with diameter $D = 80m$ and blade roughness coefficient $k = 0.075$. The wind is in the direction of the positive x -axis with a steady wind velocity of $V_\infty = 8m/s$. The wind farm model used in the simulation is shown in the Appendix II.

The objective is to compute the optimal axial induction factor for the wind turbines $\mathbf{u} = [u_1, \dots, u_n]$, such that the total power generation from the wind farm is maximized. From (27), the maximum power coefficient is obtained when

$$\frac{dC_P}{du_i} = 1 - 4u_i + 3u_i^2 = 0$$

leading to $u_i^* = 0.333$. This would be the optimal axial induction factor for a wind farm without wake interactions. However, the wake interaction within a wind farm can significantly affect the overall power production. In order to maximize the overall power production, the axial induction factor of the upwind turbines must be adjusted.

The objective of the optimization problem is therefore, to maximize the total power generated from the wind farm with n turbines in the presence of wake interactions, by adjusting the axial induction factors.

$$\min_{\mathbf{u}} P(\mathbf{u}) \quad (21)$$

We assume that the total power produced by the wind farm P is the only measurement available.

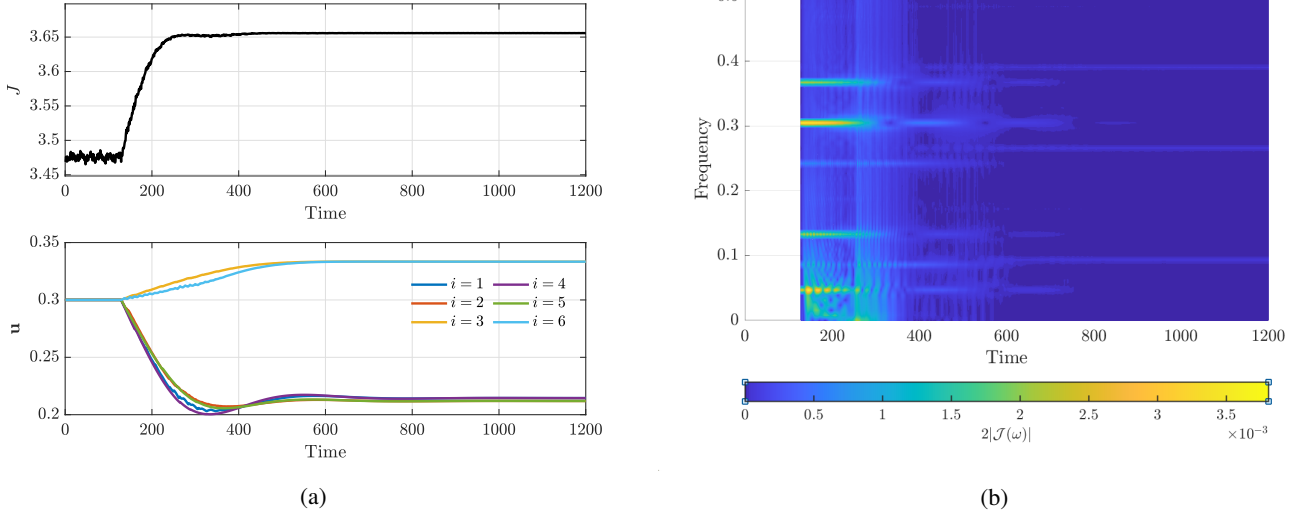


Fig. 6: Example 1 without noise: (a) Simulation results showing the overall power production from the wind farm (in MW), the optimal axial induction factors for the three turbines, and the corresponding steady-state gradients estimated using the proposed method. (b) Spectrogram showing the power spectrum as a function of time.

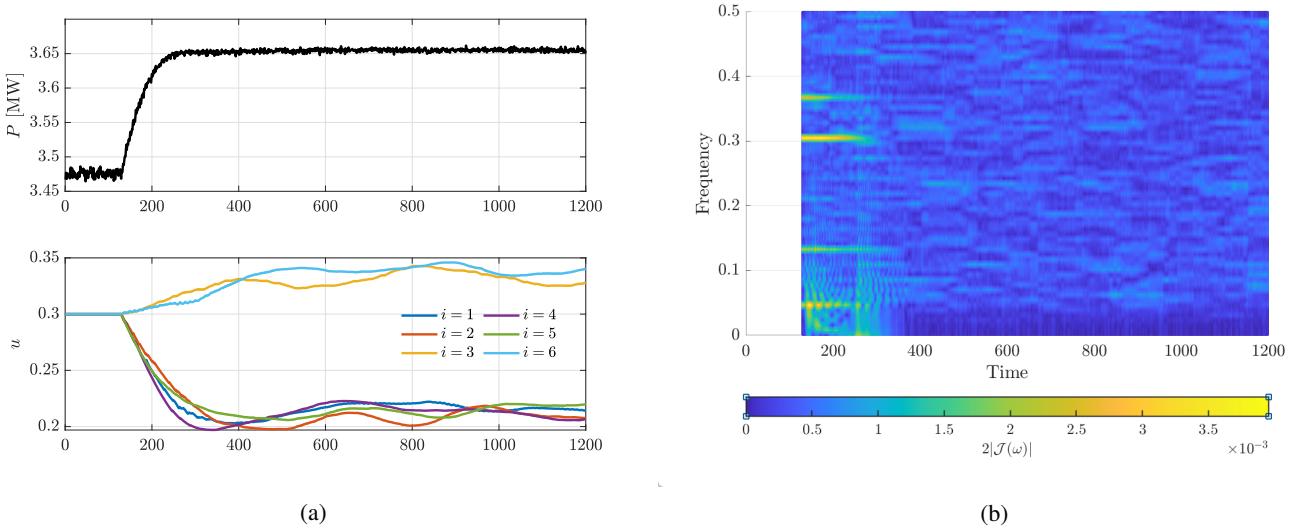


Fig. 7: Example 1 with noise: (a) Simulation results showing the overall power production from the wind farm (in MW), the optimal axial induction factors for the three turbines, and the corresponding steady-state gradients estimated using the proposed method. (b) Spectrogram showing the power spectrum as a function of time.

Since, accurate models of the aerodynamic interactions can be difficult to obtain, model-free optimization techniques such as extremum seeking control could be a suitable choice. Model-free extremum seeking control in a distributed manner using a consensus estimation algorithm for wind farm optimization was used by [4] and [5], where the authors assumed that each controller has access to its local power measurement. However, in this paper, we assume that the local power measurements are not available, and only the total overall power generated is measured.

We set the initial axial induction factors to $u_i = 0.3$ for all the wind turbines. The six inputs channels are superimposed

with sinusoidal signals with frequencies

$$f = [6 \ 17 \ 31 \ 39 \ 47 \ 11]^T / 128$$

$$= [0.0469 \ 0.1328 \ 0.2422 \ 0.3047 \ 0.3672 \ 0.0859]^T$$

and amplitude $a_i = 0.003$. The window length was chosen to be of $N = 128$ samples, according to Theorem 1.

The simulation results with the proposed FFT-based multi-variable extremum seeking control is show in Fig. 6a, where it can be seen that the proposed approach is able to drive the different inputs to the optimum such that the total power generated by the wind farm is maximized. Fig. 6b shows the spectrogram, where it can be seen that the amplitude of the power spectrum at the perturbation frequencies become zero

as the system converges to the optimum. The same simulation results are repeated when the total power measurement is corrupted with a measurement noise of $w(t) \sim \mathcal{N}(0, 2 \times 10^{-3})$ MW. The time series as well as the spectrogram for this case is shown in Fig. 7.

B. Example 2: Maximizing heat recovery from a heat exchanger network

Consider an industrial cluster with n different processing plants, the details of which are not important. The different processes in the industrial cluster generates some amount of surplus heat, which is to be recovered using a heat exchanger network, and transferred to a local district heating system, as shown in Fig. 8. The cold stream inlet is split into eight different streams, one for each company in the industrial cluster. The objective here is to adjust the split ratio u_i for each stream in order to maximize the total heat recovery, which can be equivalently stated as maximizing the end temperature of the cold stream T_{end} [30]. The optimization problem can then be stated as,

$$\max_{\mathbf{u}} T_{end}(\mathbf{u}) \quad (22)$$

where $\mathbf{u} = [u_1, \dots, u_{n-1}]$ and $u_n = 1 - \sum_{i=1}^{n-1} u_i$.

As the temperature and flow rate of the hot stream from the industrial cluster varies, or if the overall heat transfer coefficient in the heat exchanger changes (e.g. due to heat exchanger fouling), this requires adjusting the split ratios in order to maximize the temperature. This process is inspired by the batch reactor heat recovery challenge problem from [31], where the authors considered a similar problem, and later proposed the concept of controlling the so-called *Jäschke Temperature* [30] in order to achieve optimal operation. The *Jäschke Temperature* is essentially the model gradient based on a simplified countercurrent heat exchanger model with arithmetic mean temperature difference.

In this industrial symbiosis network, we assume that no information regarding the heat exchanger network such as temperatures, flow rate, heat transfer coefficient etc. are available, and the model of the system is also not available. The end temperature T_{end} is the only measurement that is available for online process optimization. Note that one does not even need to know if the heat exchangers are of countercurrent or concurrent type. The details of the simulator model used in this section can be found in Appendix III. Note that this model is only used to simulate the process and no knowledge of this model is used by the extremum seeking controller.

We set the initial split ratios to $u_i = 0.1$ for $i = 1, \dots, 7$. The temperature of the hot streams are initially $T_{h,in} = [120 \ 130 \ 120 \ 140 \ 120 \ 125 \ 115 \ 110]^T$ °C. After 2000 time steps, the temperature of the first hot stream increases to $T_{h,in}^1 = 150$ °C. The seven inputs channels are superimposed with sinusoidal signals with frequencies $f = [6 \ 11 \ 17 \ 23 \ 31 \ 39 \ 47]^T / 128$ and amplitude $a_i = 0.003$ for $i = 1, 5, 6$ and $a_i = 0.002$ for $i = 2, 3, 4, 7$. The window length was chosen to be of $N = 128$ samples according to Theorem 1. The integral gain was chosen to be $K_i = 7.5 \times 10^{-6}$ for all $i = 1, \dots, 7$. The simulation

results with the proposed FFT-based multivariable extremum seeking control is shown in Fig. 9, where it can be seen that the proposed approach is able to drive the different inputs to the optimum such that the end temperature supplied to the district heating is maximized.

V. CONCLUSION

In this paper, we presented a novel multivariable extremum seeking scheme, where the gradient is estimated using spectral analysis obtained using fast Fourier transform (FFT) as shown in Algorithm 1. The proposed approach requires relatively less tuning parameters compared to the classical extremum seeking control. The minimum window length N used for gradient estimation is given by Theorem 1. Lemma 1 showed that the gradient estimation error is bounded and Theorem 2 provides an upper bound on K_i such that the algorithm asymptotically converges to the extremum. The wind farm optimization problem, as well as an industrial waste heat recovery problem showed that process can be driven to its optimum using the proposed extremum seeking scheme.

ACKNOWLEDGEMENTS

Useful discussions with Prof. Sigurd Skogestad at the Norwegian University of Science and Technology, as well as with Dr. Pål Kittilsen and Dr. Qin Li at Equinor Research Center are gratefully acknowledged.

REFERENCES

- [1] C. S. Draper and Y. T. Li, *Principles of optimizing control systems and an application to the internal combustion engine*. American Society of Mechanical Engineers, 1951.
- [2] M. Krstić and H.-H. Wang, "Stability of extremum seeking feedback for general nonlinear dynamic systems," *Automatica*, vol. 36, no. 4, pp. 595–601, 2000.
- [3] Y. Tan, W. Moase, C. Manzie, D. Nešić, and I. Mareels, "Extremum seeking from 1922 to 2010," in *Control Conference (CCC), 2010 29th Chinese*. IEEE, 2010, pp. 14–26.
- [4] A. Menon and J. S. Baras, "Collaborative extremum seeking for welfare optimization," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 346–351.
- [5] J. Ebeğbulen and M. Guay, "Power maximization of wind farms using discrete-time distributed extremum seeking control," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 339–344, 2018.
- [6] A. Ghaffari, M. Krstić, and S. Seshagiri, "Power optimization for photovoltaic microconverters using multivariable newton-based extremum seeking," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 6, pp. 2141–2149, 2014.
- [7] J. Creaby, Y. Li, and J. E. Seem, "Maximizing wind turbine energy capture using multivariable extremum seeking control," *Wind Engineering*, vol. 33, no. 4, pp. 361–387, 2009.
- [8] A. Ghaffari, S. Seshagiri, and M. Krstić, "Multivariable maximum power point tracking for photovoltaic micro-converters using extremum seeking," *Control Engineering Practice*, vol. 35, pp. 83–91, 2015.
- [9] D. Krishnamoorthy, J. Ryu, and S. Skogestad, "Dynamic extremum seeking scheme applied to gas lift optimization," *IFAC-PapersOnLine (DYCOPS)*, 2019.
- [10] A. Pavlov, M. Haring, and K. Fjalestad, "Practical extremum-seeking control for gas-lifted oil production," in *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*. IEEE, 2017, pp. 2102–2107.
- [11] B. Mu, Y. Li, J. M. House, and T. I. Salisbury, "Real-time optimization of a chilled water plant with parallel chillers based on extremum seeking control," *Applied Energy*, vol. 208, pp. 766–781, 2017.
- [12] B. Zitte, B. Hamroun, F. Couenne, and I. Pitault, "Extremum-seeking based distributed optimization of heat exchangers network," *IFAC-PapersOnLine*, vol. 51, no. 23, pp. 331–336, 2018.

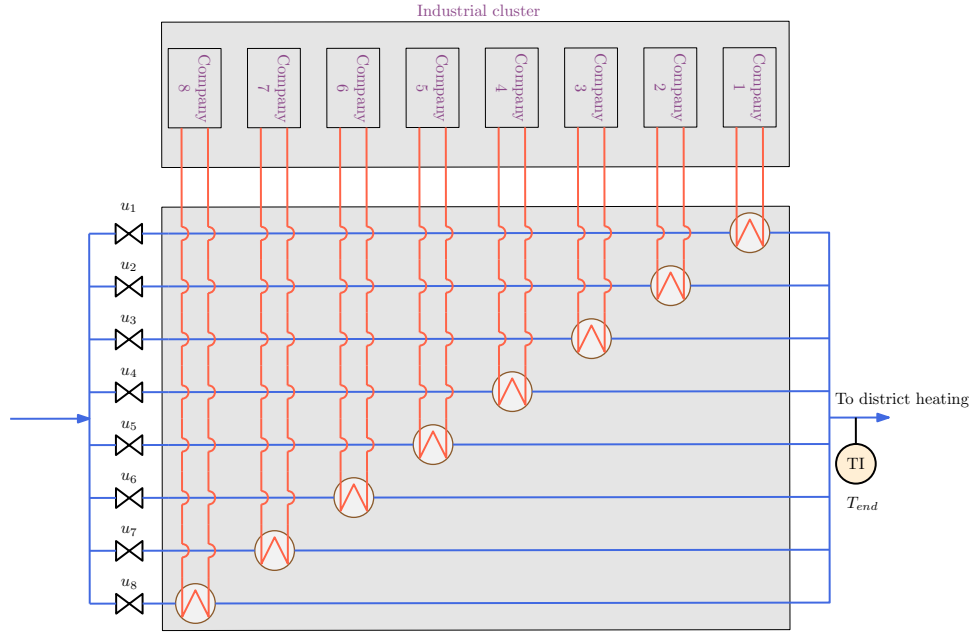


Fig. 8: Schematic representation of the heat recovery network from an industrial cluster.

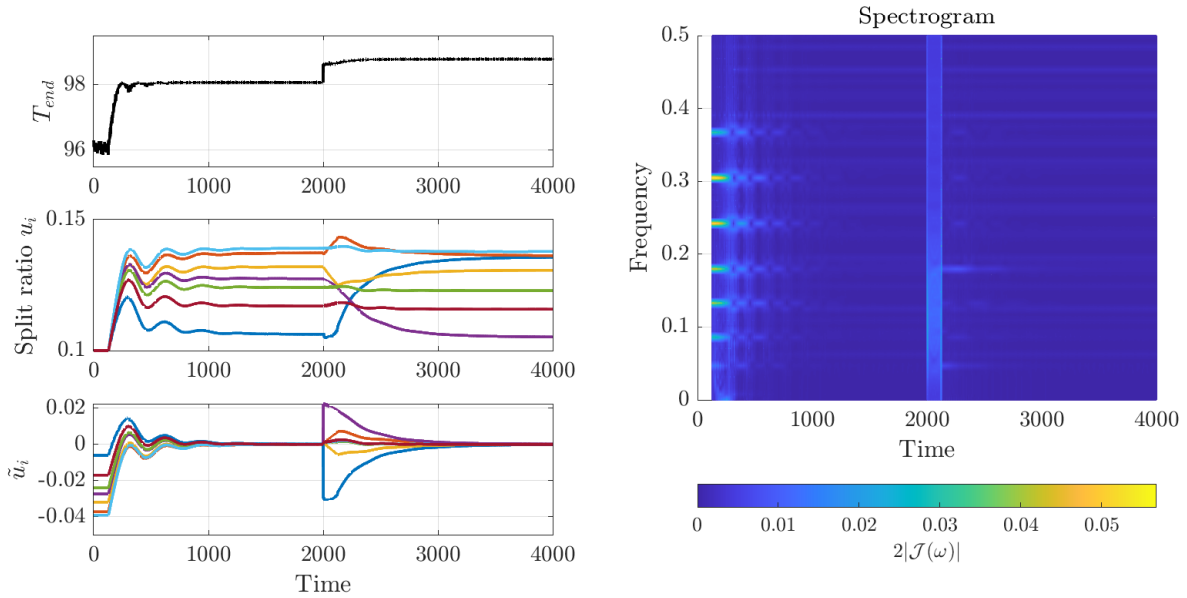


Fig. 9: Example 2: Simulation results showing the end temperature recovered from the heat exchanger network, the optimal split ratios for the different flow streams, and the spectrogram showing the power spectrum of the end temperature measurement as a function of time.

- [13] P. Binetti, K. B. Ariyur, M. Krstic, and F. Bernelli, "Formation flight optimization using extremum seeking feedback," *Journal of Guidance, Control, and Dynamics*, vol. 26, no. 1, pp. 132–142, 2003.
- [14] A. Ghaffari, M. Krstić, and D. Nešić, "Multivariable newton-based extremum seeking," *Automatica*, vol. 48, no. 8, pp. 1759–1767, 2012.
- [15] S. Dougherty and M. Guay, "Extremum-seeking control of distributed systems using consensus estimation," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 3450–3455.
- [16] I. Munteanu, A. I. Bratcu, and E. Ceangă, "Wind turbulence used as searching signal for mppt in variable-speed wind energy conversion systems," *Renewable Energy*, vol. 34, no. 1, pp. 322–327, 2009.
- [17] K. T. Atta, A. Johansson, and T. Gustafsson, "Extremum seeking control based on phasor estimation," *Systems & Control Letters*, vol. 85, pp. 37–45, 2015.
- [18] N. Bizon, "Global extremum seeking control of the power generated by a photovoltaic array under partially shaded conditions," *Energy Conversion and Management*, vol. 109, pp. 71–85, 2016.
- [19] E. Corti, C. Forte, G. Mancini, and D. Moro, "Automatic combustion phase calibration with extremum seeking approach," *Journal of Engineering for Gas Turbines and Power*, vol. 136, no. 9, 2014.
- [20] J.-F. Beaudoin, O. Cadot, J.-L. Aider, and J. E. Wesfreid, "Bluff-body drag reduction by extremum-seeking control," *Journal of fluids and structures*, vol. 22, no. 6-7, pp. 973–978, 2006.
- [21] G. François, B. Srinivasan, and D. Bonvin, "Use of measurements for enforcing the necessary conditions of optimality in the presence of constraints and uncertainty," *Journal of Process Control*, vol. 15, no. 6, pp. 701–712, 2005.
- [22] A. Marchetti, B. Chachuat, and D. Bonvin, "Modifier-adaptation

methodology for real-time optimization,” *Industrial & engineering chemistry research*, vol. 48, no. 13, pp. 6022–6033, 2009.

- [23] G. François, B. Srinivasan, and D. Bonvin, “Comparison of six implicit real-time optimization schemes,” *Journal Européen des Systèmes Automatisés*, vol. 46, no. EPFL-ARTICLE-170545, pp. 291–305, 2012.
- [24] B. Hunnekens, M. Haring, N. van de Wouw, and H. Nijmeijer, “A dither-free extremum-seeking control approach using 1st-order least-squares fits for gradient estimation,” in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 2679–2684.
- [25] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing*, 2000.
- [26] D. Gabor, “Theory of communication. part 1: The analysis of information,” *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, vol. 93, no. 26, pp. 429–441, 1946.
- [27] E. Sejdić, I. Djurović, and J. Jiang, “Time–frequency feature representation using energy concentration: An overview of recent advances,” *Digital signal processing*, vol. 19, no. 1, pp. 153–183, 2009.
- [28] H. Rotstein and S. Raz, “Gabor transform of time-varying systems: Exact representation and approximation,” *IEEE transactions on automatic control*, vol. 44, no. 4, pp. 729–741, 1999.
- [29] D. Krishnamoorthy, “dinesh-krishnamoorthy/Extremum-Seeking-FFT: v1.0.0,” Feb. 2021. [Online]. Available: <https://doi.org/10.5281/zenodo.4540283>
- [30] J. Jäschke and S. Skogestad, “Optimal operation of heat exchanger networks with stream split: Only temperature measurements are required,” *Computers & chemical engineering*, vol. 70, pp. 35–49, 2014.
- [31] —, “A batch reactor heat recovery challenge problem,” in *5th Nordic Optimization Symposium*, 2012.
- [32] I. Katic, J. Højstrup, and N. O. Jensen, “A simple model for cluster efficiency,” in *European wind energy association conference and exhibition*. A. Raguzzi, 1987.
- [33] T. Sørensen, M. L. Thøgersen, P. Nielsen, and N. Jernesvej, “Adapting and calibration of existing wake models to meet the conditions inside offshore wind farms,” *EMD International A/S. Aalborg*, 2008.

APPENDIX I ILLUSTRATIVE TOY EXAMPLE

Consider an univariate case, where the static map is given by

$$J = -100(u - 0.5)^2 \quad (23)$$

We use the proposed FFT-based extremum scheme from Algorithm 1, where the input is perturbed with a frequency of $f = 0.125$ Hz and amplitude of $a = 0.01$. We use a window length of $N = 128$. Fig. 10 shows the raw and detrended signal, and the single-sided amplitude spectrum of the input u (left hand side subplots) and the cost J (right hand side subplots) captured at a given time t . Here we see that the input u is moving from 0.2 to 0.4, and the single-sided amplitude spectrum has a magnitude of $2|U(\omega)| = 0.01$ at $\omega = 0.125$ that corresponds to the input perturbation as expected. As the input u changes from 0.2 to 0.4, the amplitude in the output changes from 0.6 to 0.2 respectively. The tapering of the amplitude is clearly seen in the detrended cost signal in Fig. 10 (right middle subplot). With a perturbation amplitude of 0.01, this implies that the gradient changes from 60 to 20 in this interval. Since spectral analysis shows the power of a signal as the mean squared amplitude at each frequency, the output power spectrum at $\omega = 0.125$ indicates that the cost varies with an amplitude close to 0.4 (i.e. gradient = 40). Thus, we see that the estimated amplitude in the output power spectrum corresponds to the true amplitude for some input over the last N samples. This example provides an intuitive understanding of the error bound on the gradient estimate formalized in Lemma 1.

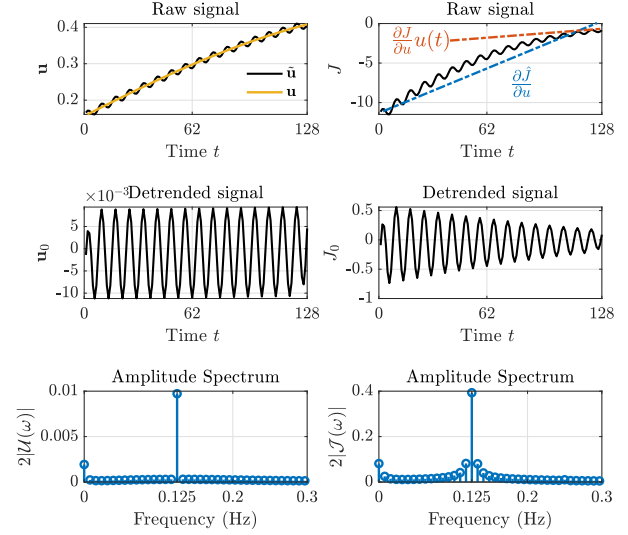


Fig. 10: Illustrative example showing the amplitude error bound when $\Delta u > 0$.

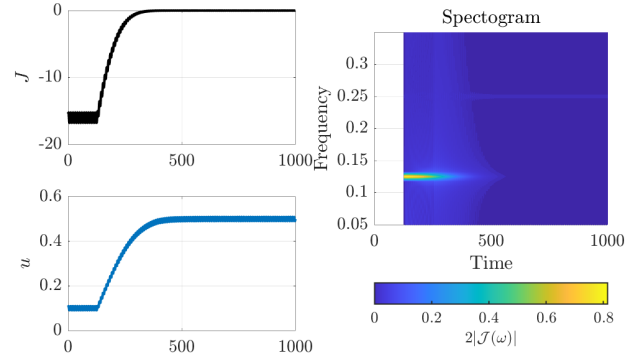


Fig. 11: Illustrative Example: Simulation results using the proposed FFT-based extremum seeking scheme. Right subplot shows the spectrogram.

In this simple example, the Hessian $\mathbf{H} = -200$ and with $\Delta u = 0.4 - 0.2 = 0.2$, the upper bound on the gradient estimation error is $|\epsilon| \leq 40$, which is reasonable given that the true gradient may be between 60 and 20 as noted.

The simulation results with an integral gain of $K_I = 1.5E-5$ is shown in Fig. 11. The spectrogram which shows the power spectrum as a function of time is shown in the right subplot in Fig. 11, where it can be seen that the amplitude at $f = 0.125$ goes to zero, since at the optimum, the input perturbation leads to a periodic variation of the output at twice the perturbation frequency. This is also seen in the spectrogram at $f = 0.25$.

APPENDIX II WIND FARM MODEL

In this section we present the wind farm model used in Section IV-A. Note that the model is used only to simulate the plant, and no information about the model is used by the extremum seeking algorithm.

Consider a wind farm with n wind turbines denoted by the set $\mathcal{W} = \{1, \dots, n\}$, placed at the co-ordinates (x, y) relative to a common vertex. The turbines are assumed to be uniform with a diameter D . The axial induction factor for the wind turbines are denoted by $\mathbf{u} = [u_1, \dots, u_n]$, which are the manipulated variables (degrees of freedom) for optimization. The set of admissible axial induction factors is given by $0 \leq u_i \leq 0.5$. Physically, the axial induction factor can be changed by adjusting the blade pitch angle.

A. Wake interaction model

When wind passes through a turbine, it creates turbulence, which is carried downwards. This is known as wake effect, which affects the wind turbines that are downwind. To model the wake interaction, we use the momentum balance and the blade element and momentum (BEM) theory [32]. The aggregate wind velocity for any turbine $i \in \mathcal{W}$ is expressed as,

$$V_i(\mathbf{u}) = V_\infty (1 - \delta V_i(\mathbf{u})) \quad (24)$$

where V_∞ denotes the free stream wind speed (in m/s), and $\delta V_i(\mathbf{u})$ denotes the aggregated velocity deficit seen by the turbine i due to the wake generated by the other upwind turbines. This is given by the expression,

$$\delta V_i(\mathbf{u}) = 2 \sqrt{\sum_{j \in \mathcal{W}: x_j < x_i} \left(a_j \left(\frac{D_j}{D_j + 2k(x_i - x_j)} \right)^2 \frac{A_{j,i}^{overlap}}{A_i} \right)^2} \quad (25)$$

where A_i is the sweep area of turbine i and $A_{j,i}^{overlap}$ is the overlap of the sweep area between turbine i and turbine j at a distance x downstream. k is the roughness co-efficient that defines the angle at which the wake expands out of the turbine. The roughness co-efficient typically takes the value $k = 0.075$ for farmlands and $k = 0.04$ for offshore wind turbines [33].

B. Power model

The power (in W) generated by any turbine $i \in \mathcal{W}$ takes the form,

$$P_i(\mathbf{u}) = \frac{1}{2} \rho_{air} A_i C_P(u_i) V_i(\mathbf{u})^3 \quad (26)$$

where, $\rho_{air} = 1.225 \text{ kg/m}^3$ is the air density and $V_i(\mathbf{u})$ is the wind speed at turbine i given by (24) and (25). C_P is the power efficiency co-efficient characterized by,

$$C_P(u_i) = 4u_i(1 - u_i)^2 \quad (27)$$

The total power generated by the wind farm is then simply given by,

$$P(\mathbf{u}) = \sum_{i \in \mathcal{W}} P_i(\mathbf{u}) \quad (28)$$

APPENDIX III

HEAT EXCHANGER NETWORK MODEL

Consider a heat exchanger network with n flow streams with no accumulation as shown in Fig. 9. Cold water at temperature $T_{c,in}$ with a flow rate w_c is split into n streams, with a split

ratio given by u_i for all $i = 1, \dots, n$. The mass balance dictates that

$$u_n = \sum_{i=1}^{n-1} u_i$$

Each flow stream has an heat exchanger with a hot stream at temperature $T_{h,in}^{(i)}$ and flow rate $w_h^{(i)}$. The outlet temperatures of the cold and hot stream from the i^{th} heat exchanger is given by,

$$T_{c,out}^{(i)} = T_{c,in} + \frac{UA^{(i)} \Delta T_{lm}^{(i)}}{u_i w_c c_p}$$

$$T_{h,out}^{(i)} = T_{h,in}^{(i)} - \frac{UA^{(i)} \Delta T_{lm}^{(i)}}{w_h^{(i)} c_p}$$

where $UA^{(i)}$ is the overall heat transfer coefficient of heat exchanger i (in W/°C), c_p is the specific heat capacity in kJ/kg/°C,

$$\Delta T_{lm}^{(i)} := \frac{\Delta T_1^{(i)} - \Delta T_2^{(i)}}{\ln \left(\Delta T_1^{(i)} / \Delta T_2^{(i)} \right)}$$

$$\approx \left[\Delta T_1^{(i)} \Delta T_2^{(i)} \left(\frac{\Delta T_1^{(i)} + \Delta T_2^{(i)}}{2} \right) \right]^{1/3}$$

is the log-mean temperature difference, and for counter-current flow $\Delta T_2^{(i)} := T_{h,out}^{(i)} - T_{c,in}^{(i)}$ and $\Delta T_1^{(i)} := T_{h,in}^{(i)} - T_{c,out}^{(i)}$. Assuming linear mixing, the end temperature T_{end} is given by

$$T_{end} = \sum_{i=1}^n u_i T_{c,out}^{(i)} \quad (29)$$