Stability analysis of fast extremum seeking control for Wiener systems using online complex curve fitting

Juan Javier Palacios Roman, Matthijs van Berkel, *Member, IEEE*, Maurice Heemels, *Fellow, IEEE*, and Thijs van Keulen

***Abstract*— In this paper, we show uniform semi-global practical asymptotic stability of fast extremum seeking control (ESC) for single-input single-output Wiener systems. While classic ESC requires a time-scale separation between plant and dither, the fast ESC method circumvents this time-scale separation by exploiting limited knowledge of the frequency response of the linear part of the Wiener system, thereby achieving faster convergence. The assumptions under which the fast ESC method works are relaxed compared to existing work and explicit bounds on the design parameters of the fast ESC scheme are provided. A numerical case study illustrates the enhanced convergence and the robustness of the fast ESC method.**

***Index Terms*— Adaptive control, extremum seeking, Lyapunov methods, stability of nonlinear systems.**

I. INTRODUCTION

EXTREMUM seeking control (ESC) is an online, model-free optimization method used to improve steady-state behavior of dynamic systems. These systems have input signals that can be tuned to maximize a measurable performance output. ESC adaptively changes these input signals to maximize the performance output without requiring extensive system knowledge. Many applications and variants of ESC can be found in the literature, see, e.g., [1]–[3] for overviews.

In classic ESC [4] a sinusoidal wave, called the dither signal, is used to perturb the input of the system around a nominal value. In case the system is a dynamic system, the frequency of the dither signal must be chosen sufficiently low such that the dynamic system with measurable performance output can be approximated by a static input-output mapping on the time-scale of the perturbation. By filtering the response of the performance output using low- and high-pass filters, an estimate of the local derivative of the static input-output performance mapping can be derived. The cutoff frequency of these filters must be sufficiently low compared to the dither frequency. As a result, two time-scale separations are required: one between the system and the dither and one between the dither and the derivative estimator. Subsequently, the gradient ascent algorithm, which operates on the time-scale of the derivative estimator, is used to recursively update the nominal value of the input.

The two time-scale separations limit the rate of convergence to the optimum. Therefore, methods to alleviate these time-scale separations have been researched. In [5], the use of moving average filters instead of low- and high-pass filters in the derivative estimation was proposed to remove the time-scale separation between the dither and the derivative estimation. Similarly, the ESC scheme in [6] circumvents the same time-scale separation by using second-order averaging. In [7]–[10], so-called fast ESC schemes are proposed that remove the time-scale separation between the system and the dither. They achieve this by imposing a certain system structure, e.g., a Wiener system, and leveraging limited system knowledge that is assumed to be available.

Of particular interest in this paper is the fast ESC method from [10]. This method assumes that the system is a single-input single-output Wiener system and that a frequency-domain approximation of the linear part of the Wiener system is available. By fitting the frequency-domain approximation to online measurements of the input and output, a local derivative estimate of the underlying steady-state input-output mapping is obtained without requiring a time-scale separation between system and dither. The benefits presented by this method are as follows. Firstly, the method utilizes moving average filters, like in [5], thereby circumventing both time-scale separations required by classic ESC. Secondly, due to the prevalence of frequency-domain design tools in industry [11], it is often the case that limited knowledge of the frequency response of the system is available, see, e.g., [12]. Instead of neglecting this information, the method in [10] leverages this available knowledge to achieve faster convergence and, in industrial use cases, higher throughput than classic ESC. Thirdly, other fast ESC methods [7]–[9] have strict conditions, such as knowing the exact relative degree of the LTI part of

Manuscript received 12 September 2025; revised 16 November 2025; accepted 2 December 2025.

Juan Javier Palacios Roman, Maurice Heemels and Thijs van Keulen are with the Control Systems Technology group, Department of Mechanical Engineering, Eindhoven University of Technology, 5612 AE, Eindhoven, The Netherlands (email: j.j.palacios.roman@tue.nl; w.p.m.h.heemels@tue.nl; t.a.c.v.keulen@tue.nl)

Juan Javier Palacios Roman and Matthijs van Berkel are with the Dutch Institute for Fundamental Energy Research, 5612 AJ, Eindhoven, The Netherlands (email: J.J.PalaciosRoman@differ.nl; M.vanBerkel@differ.nl). DIFFER is part of the institutes organisation of NWO.

Matthijs van Berkel is with the Control Systems group, Department of Electrical Engineering, Eindhoven University of Technology, 5612 AE, Eindhoven, The Netherlands (email: m.v.berkel@tue.nl).

the system, which in practice can be difficult to determine, e.g., due to the presence of time delay in feedback-controlled systems. Similarly, gray-box ESC methods, see, e.g., [2] for an overview, assume that the entire model, except for a finite number of parameters, is known. In contrast, the method in [10] only requires approximate knowledge, making the method inherently more robust.

Current limitations of the fast ESC scheme in [10] are that the formal stability analysis still relies on a time-scale separation between system and dither and that a strict requirement is placed on the assumed available knowledge of the Wiener system. To address these limitations, the contributions of this paper are as follows: we provide a proof of uniform semi-global practical asymptotic stability (USGPAS) of the fast ESC method that does not require a time-scale separation assumption between the system and the dither signal. This removal of the time-scale separation, in addition to using functional differential equations over ordinary differential equations, causes conventional methods of establishing stability for ESC methods, see, e.g., [2], [4], [13], to not be applicable. Thus, a novel proof is needed. In addition, we also provide clear upper bounds on the design parameters of the fast ESC scheme that guarantee USGPAS. Furthermore, the conditons on the frequency-domain approximation are relaxed compared to [10], making the discussed method more widely applicable. Finally, the method is demonstrated using an illustrative example.

The remainder of this paper is structured as follows: Section II will provide some notation and definitions. In Section III a brief description of the system and the control problem will be given. In Section IV the fast ESC scheme will be discussed. USGPAS of the fast ESC scheme will be shown in Section V. An application of the fast ESC scheme and a comparison to classic ESC will be presented on a numerical example in Section VI. Finally, concluding remarks will be given in Section VII.

II. NOTATION & DEFINITIONS

Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$ and $\mathbb{C}$ denote the sets of real, nonnegative real, positive real and complex numbers, respectively. We denote by $|z|$, $\angle(z)$, $\mathrm{Re}\{z\}$ and $\mathrm{Im}\{z\}$ the absolute value, principal argument, real part and imaginary part of $z \in \mathbb{C}$, respectively. Furthermore, $\|x\|$ and $\|M\|$ denote the Euclidean norm of $x \in \mathbb{R}^n$ and the spectral norm of $M \in \mathbb{R}^{n\times n}$, respectively. We adopt the standard definitions for $\mathcal{K}$- and $\mathcal{KL}$-functions from [14, Definitions 4.2 and 4.3]. We say $f(x) = O(g(x))$, where $x \in \mathbb{R}$, if there exist positive constants E and ϵ such that $|f(x)| \leq E|g(x)|$ if $|x| < \epsilon$. We adopt the definition of uniform semi-global practical asymptotic stability (USGPAS) from [13, Definition 1].

III. SYSTEM DESCRIPTION

Consider the single-input single-output Wiener system

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t)\\ y(t) &= Cx(t) + Du(t)\\ z(t) &= h(y(t)),\end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}$ is the input, $y(t) \in \mathbb{R}$ is the output of the linear system and $z(t) \in \mathbb{R}$ is the output of the Wiener system, all at time $t \in \mathbb{R}_{\geq 0}$, and A, B, C and D are real matrices of appropriate dimensions. The transfer function of (1) is defined by

$$H(s) = C(sI - A)^{-1}B + D,$$

in which s is the Laplace variable and I the identity matrix of $\mathbb{R}^{n\times n}$. Finally, $h : \mathbb{R} \to \mathbb{R}$ is a static, nonlinear function and at least three times differentiable. We make the following assumptions on system (1).

Assumption 1: The matrix A is Hurwitz

The first assumption implies the existence of an asymptotically stable equilibrium point $\bar{x}(\bar{u}) := -A^{-1}B\bar{u}$ for any constant input $u(t) = \bar{u} \in \mathbb{R}$ for all $t \in \mathbb{R}_{\geq 0}$. Hence, there exists an asymptotically stable steady-state input-output mapping $Q : \mathbb{R} \to \mathbb{R}$ given by $Q(\bar{u}) := h(-CA^{-1}B\bar{u} + D\bar{u})$.

Assumption 2: The steady-state input-output mapping Q possesses a unique maximum at $u^\star$ and

$$\frac{\mathrm{d}Q}{\mathrm{d}\bar{u}}(\bar{u}) = 0 \text{ if and only if } \bar{u} = u^\star.$$

The second assumption implies that Q is unimodal, i.e.,

$$(\bar{u} - u^\star)\frac{\mathrm{d}Q}{\mathrm{d}\bar{u}}(\bar{u}) < 0 \quad \forall \bar{u} \neq u^\star, \tag{2}$$

which includes (but is not limited to) all strictly concave functions. These assumptions are analogous to the standard assumptions of extremum seeking control (ESC) in a non-local stability context, see, e.g., [13, Assumptions 1-3]. The control goal is to steer the input $u(t)$ to a neighborhood of $u^\star$ such that $x(t)$, $y(t)$ and $z(t)$ converge to a neighborhood of $x^\star := \bar{x}(u^\star)$, $-CA^{-1}Bu^\star + Du^\star$ and $Q(u^\star)$, respectively.

IV. FAST ESC DESCRIPTION

We consider the dither-based version of the fast ESC scheme in [10] with a single sine dither signal

$$\begin{aligned}u(t) &= \hat{u}(t) + d(t)\\ d(t) &= a\cos(\omega t)\\ \dot{\hat{u}}(t) &= \begin{cases} 0, & t < T\\ c\xi(t), & t \geq T\end{cases}\\ \xi(t) &= \hat{H}(0)\cdot \arg\min_{\theta\in\mathbb{R}} \left|Z(t) - \theta\hat{H}(i\omega)N(t)\right|^2\\ Z(t) &= \frac{1}{T}\int_{t-T}^{t} z(\tau)e^{-i\omega\tau}\mathrm{d}\tau\\ N(t) &= \frac{1}{T}\int_{t-T}^{t} d(\tau)e^{-i\omega\tau}\mathrm{d}\tau.\end{aligned} \tag{3}$$

Here, $\hat{u}(t) \in \mathbb{R}$ is the optimizer output, $d(t) \in \mathbb{R}$ is the dither signal (with amplitude $a \in \mathbb{R}_{>0}$, frequency $\omega \in \mathbb{R}_{>0}$ and period $T := 2\pi/\omega$), all at time $t \in \mathbb{R}_{\geq 0}$. The signals $Z(t) \in \mathbb{C}$ and $N(t) \in \mathbb{C}$ are the Fourier transforms of z and d, respectively, over the time-window $[t - T, t]$ evaluated at $i\omega$, and $\xi(t) \in \mathbb{R}$ is used as an estimate of $\mathrm{d}Q/\mathrm{d}\bar{u}(\hat{u}(t))$, all at time $t \geq T$. The parameter $c \in \mathbb{R}_{>0}$ is the optimizer gain and $\theta \in \mathbb{R}$ is a fitting parameter. The values $\hat{H}(0) \in \mathbb{R}$ and $\hat{H}(i\omega) \in \mathbb{C}$ are approximations of $H(s)$ evaluated at $s = 0$

and $s = i\omega$, respectively. We assume the following on the closeness of the estimate $\hat{H}$ to H.

Assumption 3: The approximations $\hat{H}(0)$ and $\hat{H}(i\omega)$ are nonzero and finite, $\hat{H}(0)$ and $H(0)$ have equal sign and

$$\left|\angle(H(i\omega)) - \angle(\hat{H}(i\omega))\right| < \frac{\pi}{2}. \quad (4)$$

Remark 1: Note that (4) only puts a restriction on one value of ω, not over a range of ω. Moreover, since $\angle(H(i\omega)) \in (-\pi, \pi]$, the error bound in (4) spans 50% of the range of $\angle(H(i\omega))$. Thus, Assumption 3 is a significant relaxation compared to [10, Assumption 3] and matches many practical scenarios, where a dither frequency can be selected based on frequency response data with uncertainty bounds.

We can simplify the expression for $\xi(t)$ in (3) by explicitly calculating the solution to the minimization problem using $N(t) = a/2$, which follows from Euler's formula. Then,

$$\begin{aligned}\xi(t) &= \hat{H}(0) \arg\min_{\theta\in\mathbb{R}} \left\| \begin{bmatrix}\mathrm{Re}\{Z(t)\}\\ \mathrm{Im}\{Z(t)\}\end{bmatrix} - \theta\frac{a}{2}\begin{bmatrix}\mathrm{Re}\{\hat{H}(i\omega)\}\\ \mathrm{Im}\{\hat{H}(i\omega)\}\end{bmatrix}\right\|^2 \\ &= K\begin{bmatrix}\mathrm{Re}\{Z(t)\}\\ \mathrm{Im}\{Z(t)\}\end{bmatrix}, \end{aligned} \quad (5)$$

where

$$K = \frac{2}{a}\cdot\frac{\hat{H}(0)}{\left|\hat{H}(i\omega)\right|^2}\begin{bmatrix}\mathrm{Re}\{\hat{H}(i\omega)\}\\ \mathrm{Im}\{\hat{H}(i\omega)\}\end{bmatrix}^\top. \quad (6)$$

Note that the Euclidean norm of K reads

$$\|K\| = \frac{2}{a}\left|\frac{\hat{H}(0)}{\hat{H}(i\omega)}\right|,$$

which we will use later in the stability proof.

V. STABILITY ANALYSIS

In this section we will show that the feedback loop in (1) and (3) is USGPAS for sufficiently small values of c and a. For most ESC variants USGPAS can be proven using singular perturbation and averaging, see, e.g., [4], [8], [13], [15]. However, the fast ESC method in (3) uses a moving average filter, which can equivalently be expressed as a functional differential equation, but not as an ordinary differential equation. For this type of system, the standard singular perturbation and averaging theorems in [14] are not directly applicable. Thus, like in [5], stability is shown using alternative arguments. We define $k = c/a^2$.

Theorem 1: There exist $\mathcal{KL}$-functions β_x and β_u such that for any $\Delta^x > \delta^x > 0$ and $\Delta^u > \delta^u > 0$, there exist parameters $a^* \in \mathbb{R}_{>0}$ and $k^* \in \mathbb{R}_{>0}$ such that for all $a \in (0, a^*)$ and $k \in (0, k^*)$ the solutions x and $\hat{u}$ of (1) and (3) satisfy for any $t \geq t_0 \geq 0$

$$\begin{aligned}\|x(t) - x^\star\| \leq \beta_x(\,&\|x(t_0) - x^\star\| + \\ &\|\hat{u}(t_0) - u^\star\|, a^2k(t-t_0)) + \delta^x \end{aligned} \quad (7)$$

$$\|\hat{u}(t) - u^\star\| \leq \beta_u(\,\|\hat{u}(t_0) - u^\star\|, a^2k(t-t_0)) + \delta^u \quad (8)$$

if $\|x(t_0) - x^\star\| \leq \Delta^x$ and $\|\hat{u}(t_0)) - u^\star\| \leq \Delta^u$.

Proof: The proof is structured as follows. Firstly, we approximate the response y with a quasi-static part $\hat{y}$ and a periodic part $\tilde{y}$ and we derive an explicit bound on the error of the approximation. Secondly, we approximate the response z around $\hat{y}$ over the time-window $[t-T, t]$ using Taylor's Theorem. Next, the approximation of z is used to calculate Z and how closely ξ approaches $\mathrm{d}Q/\mathrm{d}\bar{u}(\hat{u})$. Finally, USGPAS is shown using a Lyapunov-like argument.

As said, we analyze the response of y and decompose it into a quasi-static part $\hat{y}$, a periodic part $\tilde{y}$, and an error term. The response of y given initial condition $x(t_0)$ is

$$\begin{aligned} y(t) = &\int_{t_0}^{t} Ce^{A(t-\tau)}B\hat{u}(\tau)\mathrm{d}\tau + D\hat{u}(t) + \\ &\int_{t_0}^{t} Ce^{A(t-\tau)}Bd(\tau)\mathrm{d}\tau + Dd(t) + Ce^{A(t-t_0)}x(t_0). \end{aligned} \quad (9)$$

The first term is rewritten using integration by parts as

$$CA^{-1}\left(e^{A(t-t_0)}B\hat{u}(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}B\dot{\hat{u}}(\tau)\mathrm{d}\tau - B\hat{u}(t)\right).$$

Similarly, the third term on the right-hand side of (9) is rewritten using Euler's formula as

$$a\mathrm{Re}\{C(i\omega I - A)^{-1}(Be^{i\omega t} - e^{A(t-t_0)}Be^{i\omega t_0})\}.$$

Then, define

$$\begin{aligned}\hat{y}(t) &= -CA^{-1}B\hat{u}(t) + D\hat{u}(t) = H(0)\hat{u}(t),\\ \tilde{y}(t) &= a\mathrm{Re}\{C(i\omega I - A)^{-1}Be^{i\omega t}\} + Dd(t)\\ &= a\mathrm{Re}\{H(i\omega)e^{i\omega t}\},\\ \bar{y}(t) &= \int_{t_0}^{t} CA^{-1}e^{A(t-\tau)}B\dot{\hat{u}}(\tau)\mathrm{d}\tau,\\ \check{y}(t) &= Ce^{A(t-t_0)}x(t_0) + CA^{-1}e^{A(t-t_0)}B\hat{u}(t_0) - \\ &\quad aC\mathrm{Re}\{(i\omega I - A)^{-1}e^{i\omega t_0}\}e^{A(t-t_0)}B.\end{aligned}$$

and note that $y(t) = \hat{y}(t) + \tilde{y}(t) + \bar{y}(t) + \check{y}(t)$.

Now we wish to derive upperbounds on $\bar{y}(t)$ and $\check{y}(t)$. We define for some $\Gamma > \Delta^u$ the set

$$\mathcal{S}_u := \{u \in \mathbb{R} \mid |u| \leq \Gamma\}.$$

We assume and later on prove that $\hat{u}(t) - u^\star \in \mathcal{S}_u$ for all $t \geq t_0$. For now, given (9), $\hat{u}(t) - u^\star \in \mathcal{S}_u$, and $\|x(t_0)\| \leq \Delta^x + \|x^\star\|$, it follows that there exists a compact set $\mathcal{S}_y$ such that $y(t) \in \mathcal{S}_y$ for all $t \geq t_0$. Then, there exist constants $L_1, L_2, L_3 \in \mathbb{R}_{>0}$ such that

$$\left|h(y(t))\right| \leq L_1, \left|\frac{\mathrm{d}h}{\mathrm{d}y}(y(t))\right| \leq L_2, \left|\frac{\mathrm{d}^2h}{\mathrm{d}y^2}(y(t))\right| \leq L_3. \quad (10)$$

Then, $|Z(t)| \leq L_1$ and, by Cauchy-Schwarz inequality,

$$|\dot{\hat{u}}(t)| \leq |c\xi(t)| = \left|a^2kK\begin{bmatrix}\mathrm{Re}\{Z(t)\}\\ \mathrm{Im}\{Z(t)\}\end{bmatrix}\right| \leq a^2k\|K\|L_1$$

for all $t \geq t_0$. By the subaddivity of norms and commutativity of convolutions

$$\begin{aligned}|\bar{y}(t)| &\leq \int_{t_0}^{t}\left|CA^{-1}e^{A\tau}B\dot{\hat{u}}(t-\tau)\right|\mathrm{d}\tau\\ &\leq \int_{t_0}^{t}\left|CA^{-1}e^{A\tau}B\right|\mathrm{d}\tau\cdot a^2kL_1\|K\|\\ &\leq \int_{0}^{\infty}\left|CA^{-1}e^{A\tau}B\right|\mathrm{d}\tau\cdot a^2kL_1\|K\|\\ &= a^2kL_1L_4\|K\| = O(ak),\end{aligned}$$

where $L_4 := \int_0^\infty \left|CA^{-1}e^{A\tau}B\right| \mathrm{d}\tau$ is finite, because A is Hurwitz: then, there exist positive constants M and λ such that $\left\|e^{A\tau}\right\| \leq Me^{-\lambda\tau}$ [16, Theorem 1.9.2]. Note that in the definition of $\|K\|$ there is a term $2/a$ and thus we write $|\bar{y}(t)| = O(ak)$. Similarly, an upperbound on $\check{y}(t)$ follows from the Cauchy-Schwarz inequality and given bounded a is

$$|\check{y}(t)| \leq L_5 M e^{-\lambda(t-t_0)} = O(e^{-\lambda(t-t_0)}),$$

where $L_5 := \|C\| \|x(t_0)\| + \|C\| \left\|A^{-1}\right\| \|B\| \|\hat{u}(t_0)\| + \left\|aC\mathrm{Re}\{(i\omega I - A)^{-1}\}\right\| \|B\|$. Then,

$$y(t) = \hat{y}(t) + \tilde{y}(t) + O(ak + e^{-\lambda(t-t_0)}).$$

Secondly, we use Taylor's Theorem, see, e.g., [17, Theorem 4.12], to approximate the static nonlinear function h around $\hat{y}(t)$. Then, for all $\tau \in [t-T, t]$

$$\begin{aligned} z(\tau) &= h(\hat{y}(t)) + \frac{\mathrm{d}h}{\mathrm{d}y}(\hat{y}(t)) \cdot (y(\tau) - \hat{y}(t)) \\ &+ \frac{\mathrm{d}^2h}{\mathrm{d}y^2}(\hat{y}(t)) \cdot (y(\tau) - \hat{y}(t))^2 + O\left(|y(\tau) - \hat{y}(t)|^3\right). \end{aligned} \tag{11}$$

Note that the error term is $O(|\cdot|^3)$ because h is three-times differentiable as defined in Section III. Furthermore, $|\tilde{y}(\tau)| = O(a)$ and

$$\begin{aligned} |\hat{y}(\tau) - \hat{y}(t)| &= |H(0) \cdot (\hat{u}(\tau) - \hat{u}(t))| \\ &= |H(0)| \cdot |\hat{u}(\tau) - \hat{u}(t)| \\ &\leq |H(0)| \cdot T \cdot a^2 k L_1 \|K\| = O(ak). \end{aligned}$$

Thus, we can rewrite (11) as

$$\begin{aligned} z(\tau) &= h(\hat{y}(t)) + \frac{\mathrm{d}h}{\mathrm{d}y}(\hat{y}(t)) \cdot \tilde{y}(\tau) \\ &+ \frac{\mathrm{d}^2h}{\mathrm{d}y^2}(\hat{y}(t)) \cdot \tilde{y}(\tau)^2 + O(ak + e^{-\lambda(t-t_0)} + a^3) \end{aligned} \tag{12}$$

since $\mathrm{d}h/\mathrm{d}y(\hat{y}(t))$ and $\mathrm{d}^2h/\mathrm{d}y^2(\hat{y}(t))$ are bounded by (10).

Thirdly, with (12) we can derive expressions for $Z(t)$ and $\xi(t)$. From (3), (12) and the definition of $\tilde{y}(t)$ it follows that

$$Z(t) = \frac{a}{2} \cdot \frac{\mathrm{d}h}{\mathrm{d}y}(\hat{y}(t)) H(i\omega) + O(ak + e^{-\lambda(t-t_0)} + a^3). \tag{13}$$

Now, we can determine the accuracy with which $\xi(t)$ estimates $\mathrm{d}Q/\mathrm{d}\bar{u}(\hat{u}(t))$. By (5) and (13)

$$\begin{aligned} \xi(t) = K \cdot \frac{a}{2} \cdot \frac{\mathrm{d}h}{\mathrm{d}y}(\hat{y}(t)) \cdot \begin{bmatrix} \mathrm{Re}\{H(i\omega)\} \\ \mathrm{Im}\{H(i\omega)\} \end{bmatrix} \\ + K \cdot O(ak + e^{-\lambda(t-t_0)} + a^3). \end{aligned}$$

Substituting K with (6) results in

$$\begin{aligned} \xi(t) &= \frac{\mathrm{d}h}{\mathrm{d}y}(\hat{y}(t)) \cdot \frac{\hat{H}(0)}{\left|\hat{H}(i\omega)\right|^2} \begin{bmatrix} \mathrm{Re}\{\hat{H}(i\omega)\} \\ \mathrm{Im}\{\hat{H}(i\omega)\} \end{bmatrix}^\top \begin{bmatrix} \mathrm{Re}\{H(i\omega)\} \\ \mathrm{Im}\{H(i\omega)\} \end{bmatrix} \\ &+ O(k + e^{-\lambda t}/a + a^2) \\ &= \frac{\mathrm{d}h}{\mathrm{d}y}(\hat{y}(t)) \cdot H(0) \cdot \frac{\hat{H}(0)}{H(0)} \cdot \frac{\left|H(i\omega)\right|}{\left|\hat{H}(i\omega)\right|} \cdot \frac{1}{\left|H(i\omega)\right|\left|\hat{H}(i\omega)\right|} \\ &\begin{bmatrix} \mathrm{Re}\{\hat{H}(i\omega)\} \\ \mathrm{Im}\{\hat{H}(i\omega)\} \end{bmatrix}^\top \begin{bmatrix} \mathrm{Re}\{H(i\omega)\} \\ \mathrm{Im}\{H(i\omega)\} \end{bmatrix} + O(k + e^{-\lambda(t-t_0)}/a + a^2). \end{aligned}$$

Notice that $\mathrm{d}Q/\mathrm{d}\bar{u}(\hat{u}(t)) = \mathrm{d}h/\mathrm{d}y(\hat{y}(t)) \cdot H(0)$ and

$$\begin{aligned} &\cos\left(\angle(H(i\omega)) - \angle(\hat{H}(i\omega))\right) = \\ &\frac{1}{\left|H(i\omega)\right|\left|\hat{H}(i\omega)\right|} \begin{bmatrix} \mathrm{Re}\{\hat{H}(i\omega)\} \\ \mathrm{Im}\{\hat{H}(i\omega)\} \end{bmatrix}^\top \begin{bmatrix} \mathrm{Re}\{H(i\omega)\} \\ \mathrm{Im}\{H(i\omega)\} \end{bmatrix}. \end{aligned}$$

Thus,

$$\xi(t) = L_6 \frac{\mathrm{d}Q}{\mathrm{d}\bar{u}}(\hat{u}(t)) + O(k + e^{-\lambda(t-t_0)}/a + a^2),$$

where

$$L_6 := \frac{\hat{H}(0)}{H(0)} \frac{\left|H(i\omega)\right|}{\left|\hat{H}(i\omega)\right|} \cos\left(\angle(H(i\omega)) - \angle(\hat{H}(i\omega))\right),$$

which is positive under Assumption 3.

Finally, we show convergence of $x(t)$ and $\hat{u}(t)$ to a neighborhood of $x^\star$ and $u^\star$, respectively. The dynamics of (1) and (3) can now be described by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B\hat{u}(t) + Bd(t) \\ \dot{\hat{u}}(t) &= a^2kL_6 \frac{\mathrm{d}Q}{\mathrm{d}\bar{u}}(\hat{u}(t)) + a^2kO(k + e^{-\lambda(t-t_0)}/a + a^2) \end{aligned}$$

for all $t \geq T$. The second differential equation can be rewritten as

$$\dot{\tilde{u}}(t) = a^2kL_6 \frac{\mathrm{d}Q}{\mathrm{d}\bar{u}}(\tilde{u}(t) + u^\star) + a^2kw(t),$$

where $w(t)$ is a disturbance that satisfies

$$|w(t)| \leq E(k + e^{-\lambda(t-t_0)}/a + a^2) \tag{14}$$

if

$$k + e^{-\lambda(t-t_0)}/a + a^2 < \epsilon \tag{15}$$

for some positive constants E and ϵ, and $\tilde{u}(t) := \hat{u}(t) - u^\star$. We define a Lyapunov-like function

$$V(\tilde{u}(t)) = \frac{1}{2L_6}\tilde{u}(t)^2.$$

Then,

$$\dot{V}(\tilde{u}(t)) = a^2k\tilde{u}(t)\frac{\mathrm{d}Q}{\mathrm{d}\bar{u}}(\tilde{u}(t) + u^\star) + a^2k\frac{\tilde{u}(t)w(t)}{L_6}.$$

Define

$$t^*(a) = t_0 - \frac{1}{\lambda}\ln(\epsilon a \gamma_1) \tag{16}$$

where $\gamma_1 \in (0,1)$ and note that by (2), $\tilde{u}(t)\mathrm{d}Q/\mathrm{d}\bar{u}(\tilde{u}(t) + u^\star)$ is a negative definite function of $\tilde{u}(t)$. It follows from [14, Lemma 4.3] that there exists a $\mathcal{K}$-function α such that

$$\tilde{u}(t)\mathrm{d}Q/\mathrm{d}\bar{u}(\tilde{u}(t) + u^\star) \leq -\alpha(\|\tilde{u}(t)\|).$$

Define $L_7 := \alpha(\gamma_2\delta^u)$, where $\gamma_2 \in (0,1)$. Suppose $\Gamma = \Delta^u + \gamma_2\delta^u$ and

$$a^2k \leq \min\left\{\lambda, \frac{\lambda\gamma_2\delta^u}{L_1 \|K\| \left|\ln(\epsilon a\gamma_1)\right|}\right\} \tag{17}$$

$$k + a^2 \leq \min\left\{\epsilon, \frac{L_6L_7}{\Gamma E}\right\} - \gamma_1\epsilon. \tag{18}$$

Using (16) and (17), for all $0 \leq t_0 \leq t \leq t^*(a)$

$$\begin{aligned} \|\tilde{u}(t)\| &\leq a^2 k L_1 \|K\| (t - t_0) + \|\tilde{u}(t_0)\| \\ &\leq a^2 k L_1 \|K\| (t^*(a) - t_0) + \|\tilde{u}(t_0)\| \\ &= -\frac{a^2 k L_1 \|K\|}{\lambda} \ln(\epsilon a \gamma_1) + \|\tilde{u}(t_0)\| \\ &\leq \gamma_2 \delta^u + \|\tilde{u}(t_0)\| \leq \Gamma - \Delta^u + \Delta^u = \Gamma, \end{aligned} \tag{19}$$

since $\|\tilde{u}(t_0)\| \leq \Delta^u$ and $|\dot{\tilde{u}}(t)| \leq a^2 k L_1 \|K\|$ if $\|\tilde{u}(t)\| \leq \Gamma$. This implies $\tilde{u}(t) \in \mathcal{S}_u$ for all $t_0 \leq t \leq t^*(a)$. Using (15), (16) and (18), for all $t \geq t^*(a)$

$$k + e^{-\lambda(t-t_0)}/a + a^2 \leq k + e^{-\lambda(t^*(a)-t_0)}/a + a^2 \leq \epsilon$$

and thus, using (14), for all $\|\tilde{u}(t)\| \leq \Gamma$,

$$\begin{aligned} \frac{\tilde{u}(t) w(t)}{L_6} &\leq \frac{\Gamma E(k + e^{-\lambda(t-t_0)}/a + a^2)}{L_6} \\ &\leq \frac{\Gamma E(k + e^{-\lambda(t^*(a)-t_0)}/a + a^2)}{L_6} \leq L_7. \end{aligned}$$

Then, for all $t \geq t^*(a)$ and $\gamma_2 \delta^u \leq \|\tilde{u}(t)\| \leq \Gamma$

$$\begin{aligned} \dot{V}(\tilde{u}(t)) &\leq -a^2 k \alpha(\|\tilde{u}(t)\|) + a^2 k L_7 \\ &= -a^2 k \alpha(\|\tilde{u}(t)\|) + a^2 k \alpha(\gamma_2 \delta^u) \\ &\leq -a^2 k \alpha(\|\tilde{u}(t)\|), \end{aligned}$$

since α is a strictly increasing function. Together with (19), this implies for all $t \geq t_0$ that $\tilde{u}(t) \in \mathcal{S}_u$ and that there exists $\mathcal{KL}$-function β_1 such that

$$\|\tilde{u}(t)\| \leq \begin{cases} \|\tilde{u}(t_0)\| + \gamma_2 \delta^u & \text{if } t \leq t^* \\ \beta_1(\|\tilde{u}(t_0)\|, a^2 k(t - t^*(a))) + \gamma_2 \delta^u & \text{if } t \geq t^*, \end{cases} \tag{20}$$

since $\|\tilde{u}(t^\star(a))\| \leq \|\tilde{u}(t_0)\| + \gamma_2 \delta^u$ and $\|\tilde{u}(t)\|$ is decreasing if $\dot{V}(\tilde{u}(t))$ is negative, i.e., if $\|\tilde{u}(t)\| \geq \gamma_2 \delta^u$, i.e., if $\|\tilde{u}(t_0)\| \geq 0$. Using the lowerbound $1 - 1/x \leq \ln(x)$, the second inequality in (17) can be simplified to

$$k < 4\lambda \epsilon^2 \gamma_1^2 \frac{\gamma_2 \delta_u}{L_1 \|K\|} \ \wedge \ a < \frac{1}{\epsilon \gamma_1}. \tag{21}$$

and $a^2 k(t - t^*(a))$ to

$$a^2 k(t - t^*(a)) \geq a^2 k(t - t_0) - \frac{k}{4\lambda \epsilon^2 \gamma_1^2}.$$

Given bounded k, there exists a function $\beta_u \in \mathcal{KL}$, independent from $t_0, a, k, \Delta^x, \Delta^u, \delta^x$ and δ^u, that upper bounds the right half of (20) such that (8) is satisfied [18, Lemma 4.1].

Similarly, we define $\tilde{x}(t) := x(t) - x^\star$. Then

$$\begin{aligned} \|\tilde{x}(t)\| &= \left\| e^{A(t-t_0)} \tilde{x}(t_0) + \int_{t_0}^{t} e^{A(t-\tau)} B[d(\tau) + \tilde{u}(\tau)] \mathrm{d}\tau \right\| \\ &\leq \underbrace{M e^{-\lambda(t-t_0)} \|\tilde{x}(t_0)\|}_{:=\beta_2(\|\tilde{x}(t_0)\|, \lambda(t-t_0))} + \underbrace{\frac{M}{\lambda} \|B\|}_{:=L_9} (a + \gamma_2 \delta^u) \\ &\quad + \underbrace{\int_{t_0}^{t} e^{A(t-\tau)} \|B\| \beta_u(\|\tilde{u}(t_0)\|, a^2 k(\tau - t_0)) \mathrm{d}\tau}_{\leq \beta_3(\|\tilde{u}(t_0)\|, a^2 k(t-t_0))} \end{aligned}$$

where β_2, β_3 are $\mathcal{KL}$-functions and the existence of β_3 follows from the input-to-state stability of $\dot{\tilde{x}}(t) = A\tilde{x}(t) + B\tilde{u}(t)$ [14, Exercise 4.58]. Using $a^2 k \leq \lambda$ in (17) and the fact that $\mathcal{KL}$-functions are decreasing in their second argument, $\beta_2(\|\tilde{x}(t_0)\|, \lambda(t-t_0)) \leq \beta_2(\|\tilde{x}(t_0)\|, a^2 k(t-t_0))$. Then, there exists a $\mathcal{KL}$-function β_x such that (7) holds, if

$$a < \frac{\delta^x}{L_9} - \gamma_2 \delta^u. \tag{22}$$

From (17), (18), (21) and (22) it becomes clear that there are $\gamma_1, \gamma_2 \in (0,1)$ such that there exist a^* and k^*, independent from t_0, such that for all $a \in (0, a^*)$ and $k \in (0, k^*)$, upper bounds (7) and (8) are satisfied. ■

Note that (17), (18) and (22) provide clear upper bounds on the design parameters k and a. If k and a satisfy these conditions and Assumptions 1-3 are met, then $\xi(t)$ is approximately proportional to $\mathrm{d}Q/\mathrm{d}\bar{u}(\hat{u}(t))$ and $\hat{u}(t)$ will converge to an $O(k+a^2)$-neighborhood of $u^\star$, i.e., $\hat{u}(t) - u^\star = O(k+a^2)$ as $t \to \infty$. Similarly, $x(t)$ will converge to an $O(k+a)$-neighborhood of $x^\star$. These neighborhoods of convergence are equivalent to those of classic ESC. Thus, a simple yet effective tuning strategy is to reduce k and a until (17), (18) and (22) are satisfied and the feedback loop of (1) and (3) will become stable. Finally, note that the proof of Theorem 1 does not put any restrictions on the dither frequency ω. Thus, any dither frequency ω can be chosen, provided that the estimate $\hat{H}(i\omega)$ satisfies Assumption 3.

VI. NUMERICAL CASE STUDY

Consider the single-input single-output Wiener system

$$\begin{aligned} \dot{x}_1(t) &= x_2(t), \quad \dot{x}_2(t) = \omega_{p1}^2(u(t) - x_1(t)) - 2\beta_1 \omega_{p1} x_2(t), \\ \dot{x}_3(t) &= x_4(t), \quad \dot{x}_4(t) = \omega_{p2}^2(x_1(t) - x_3(t)) - 2\beta_2 \omega_{p2} x_4(t), \\ \dot{x}_5(t) &= x_6(t), \quad \dot{x}_6(t) = \omega_{p3}^2(x_3(t) - x_5(t)) - 2\beta_3 \omega_{p3} x_6(t), \\ y(t) &= x_5(t) \quad \text{and } z(t) = h(y(t)) := e^{-(y(t)-2)^2}, \end{aligned}$$

where $\omega_{p1} = 2\pi$, $\omega_{p2} = 10\pi$, $\omega_{p3} = 14\pi$, $\beta_1 = 0.2$ and $\beta_2 = \beta_3 = 0.06$. The system can be interpreted as three mass-spring-damper systems in series. The system satisfies Assumptions 1 and 2, where $Q(\bar{u}) = e^{-(\bar{u}-2)^2}$. We consider initial conditions $x(0) = 0$ and $\hat{u}(0) = 0$.

Firstly, we apply classic ESC [4] and the discussed fast ESC method (3) to the system. Both methods use a dither frequency $\omega \in \{1\pi, 2\pi, 4\pi\}$ rad/s, dither amplitude $a = 0.2$, optimizer gain $c = 0.7$. Furthermore, classic ESC uses filter frequencies $\omega_h = \omega_l = \omega/2$ rad/s. Fast ESC uses $\hat{H}(0) = H(0)$ and $\hat{H}(i\omega) = H(i\omega)$. See Figure 1. It can be seen that even with the same dither frequency the fast ESC method achieves faster convergence to the optimum than classic ESC. This is mainly due to the use of a moving average filter over low- and high-pass filters, which removes the time-scale separation between dither and derivative estimation. Furthermore, by increasing the dither frequency convergence is achieved within approximately 30 s with the fast ESC method. By contrast, classic ESC starts to diverge from the optimum when using higher dither frequencies.

Secondly, we consider the case where the discussed fast ESC method does not posses exact knowledge of the system, i.e., $L_6 \neq 1$. For simplicity, we keep $\hat{H}(0) = H(0)$ and $|\hat{H}(i\omega)| = |H(i\omega)|$. We only change the angle between $\hat{H}(i\omega)$

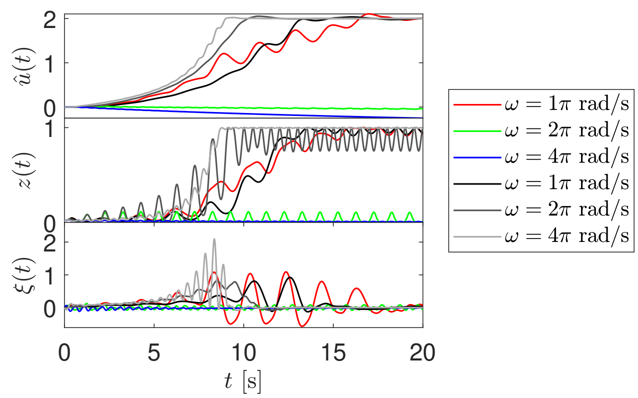

Fig. 1. Time trace of $\hat{u}(t)$, $z(t)$ and $\xi(t)$ for classic ESC (RGB) and fast ESC (gray) for multiple dither frequencies.

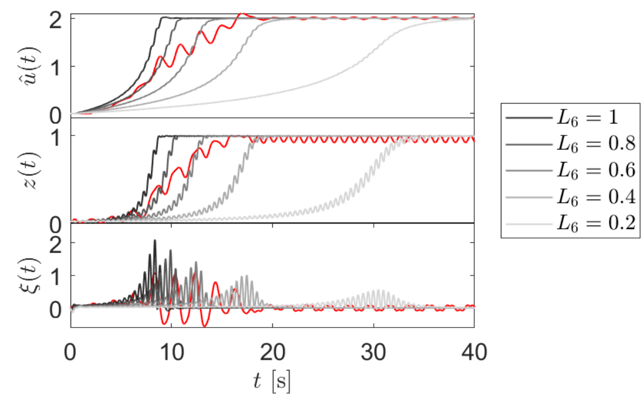

Fig. 2. Time trace of $\hat{u}(t)$, $z(t)$ and $\xi(t)$ for classic ESC (red) and fast ESC (gray) for multiple levels of L_6.

and $H(i\omega)$, i.e., $\hat{H}(i\omega) = H(i\omega)e^{i\phi}$ where $\phi \in [0, \pi/2)$, such that $L_6 = \cos(\phi)$. The used parameters are $\omega = 4\pi$ rad/s, $a = 0.2$ and $c = 0.2$. See Figure 2. The figure also shows the time trace of classic ESC with the parameters of the previous paragraph and $\omega = \pi$ rad/s. It can be seen that for low values of L_6 convergence to the optimum is still achieved but on a longer time-scale. Furthermore, for lower values of L_6 the control loop could become unstable if c and a are not lowered in tandem, as can be seen from (18). Note that the fast ESC method converges as fast as classic ESC when $L_6 \approx 0.6$. This suggests that for this particular example, the discussed fast ESC method is ensured to converge faster or just as fast as classic ESC if $\angle(H(i\omega))$ is known with an accuracy of $\pm 53°$, which is a generous margin of error.

VII. Conclusion

In this paper we have formally proven uniform semi-global practical asymptotic stability of a fast extremum seeking control (ESC) scheme. For single-input single-output Wiener systems, this scheme exploits a course approximation of the frequency response function at the dither frequency of the linear part of the Wiener system to remove the time-scale separation between system and dither and a moving average filter to remove the time-scale separation between dither and derivative estimator. Thus, both time-scale separations required by classic ESC are circumvented. The analysis of this fast ESC method required an inherently different approach from the stability analysis of classic ESC due to the absence of time-scale separation. Furthermore, the analysis led to a relaxation of the required assumptions and clear upperbounds on the design parameters of the fast ESC scheme. We demonstrated on a numerical example that the fast ESC scheme can achieve faster convergence than classic ESC due to the removal of the time-scale separation between dither and plant. Additionally, we showed the robustness of the method given an error in the approximation of the frequency response function.

For future work, we intend to extend the method to multi-input single-output Wiener systems and to higher-order derivative estimation. Moreover, it is of interest to make a comparison with other fast ESC methods on an experimental setup and to generalize the method to more general classes of nonlinear systems.

References

[1] Y. Tan, W. Moase, C. Manzie, D. Nešić, and I. Mareels, "Extremum seeking from 1922 to 2010," in *Proceedings of the 29th Chinese Control Conference*, 2010, pp. 14–26.

[2] D. Nešić, Y. Tan, C. Manzie, A. Mohammadi, and W. Moase, "A unifying framework for analysis and design of extremum seeking controllers," in *2012 24th Chinese Control and Decision Conference (CCDC)*, 2012, pp. 4274–4285.

[3] A. Scheinker, "100 years of extremum seeking: A survey," *Automatica*, vol. 161, p. 111481, 2024.

[4] 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.

[5] M. Haring, N. Van de Wouw, and D. Nesic, "Extremum-seeking control for nonlinear systems with periodic steady-state outputs," *Automatica*, vol. 49, no. 6, pp. 1883–1891, 2013.

[6] M. Abdelgalil and J. I. Poveda, "Initialization-free Lie-bracket extremum seeking," *Systems & Control Letters*, vol. 191, p. 105881, 2024.

[7] W. Moase and C. Manzie, "Fast extremum-seeking for Wiener–Hammerstein plants," *Automatica*, vol. 48, no. 10, pp. 2433–2443, 2012.

[8] K. Tourkey Atta and M. Guay, "Fast proportional integral phasor extremum seeking control for a class of nonlinear system," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 5724–5730, 2017, 20th IFAC World Congress.

[9] H. Liu, Y. Tan, T. Baček, D. Kulić, D. Oetomo, and C. Manzie, "Fast extremum seeking control for a class of generalized Hammerstein systems with the knowledge of relative degree," in *American Control Conference*, 2023, pp. 2405–2410.

[10] T. Van Keulen, R. Van der Weijst, and T. Oomen, "Fast extremum seeking using multisine dither and online complex curve fitting," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 5362–5367, 2020.

[11] G. F. Franklin, J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*, 8th ed. Harlow, UK: Pearson, 2020.

[12] R. Van der Weijst, T. van Keulen, and F. Willems, "Constrained multivariable extremum-seeking for online fuel-efficiency optimization of diesel engines," *Control Engineering Practice*, vol. 87, pp. 133–144, 2019.

[13] Y. Tan, D. Nešić, and I. Mareels, "On non-local stability properties of extremum seeking control," *Automatica*, vol. 42, no. 6, pp. 889–903, 2006.

[14] H. K. Khalil, *Nonlinear systems*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 2002.

[15] D. Nešić, A. Mohammadi, and C. Manzie, "A systematic approach to extremum seeking based on parameter estimation," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 3902–3907.

[16] L. Perko, *Differential equations and dynamical systems*, 3rd ed., ser. Texts in Applied Mathematics. New York, NY: Springer, 1991.

[17] R. A. Adams and C. Essex, *Calculus : A complete course*, 9th ed. Don Mills, Ontario: Pearson Canada, 2017.

[18] F. Albertini and E. Sontag, "Continuous control-Lyapunov functions for asymptotically controllable time-varying systems," *International Journal of Control*, vol. 72, 02 2000.