# Accelerating Extremum Seeking Convergence by Richardson Extrapolation Methods

Jan-Henrik Metsch[1], Jonathan Neuhauser[2], Jerome Jouffroy[3], Taous-Meriem Laleg-Kirati[4], Johann Reger[5]

***Abstract*— In this paper, we propose the concept of accelerated convergence that has originally been developed to speed up the convergence of numerical methods for extremum seeking (ES) loops. We demonstrate how the dynamics of ES loops may be analyzed to extract structural information about the generated output of the loop. This information is then used to distil the limit of the loop without having to wait for the system to converge to it.**

## I. Introduction

Extremum seeking is a model-free and robust scheme, originally proposed in 1922 by Leblanc (see [1]), to track an extremal operating point of an apparatus by adaptively shifting the operating point in the direction of greatest increase in some output function. The approach has been widely used in the control of systems with *a priori* unknown dynamics. A classical source for an in-depth reference is e.g. [2], where a proof of convergence is given. Tracking the extremal operating point is achieved by adding a sinusoidal perturbation to the input signal, comparing its phase to the one in the generated output and adjusting the current input based on the phase difference. This is a robust method of tracking an extremal state, but its convergence is rather slow. There are many approaches to analyzing and increasing the speed of convergence as well as eliminating oscillations around the limit available in the literature. Robustness of several ES methods in application to robotics are discussed in [3]. The influence of the loop parameters on the speed as well as the domain of convergence is studied in [4]. A method to eliminate oscillations around the limit and achieve asymptotic convergence by decreasing the dithering amplitude over time is presented in [5]. Faster convergence has also been established in [6] by the usage of fractional operators. Ref. [7] achieves enhanced convergence for small amplitude and low frequency perturbations by taking the entire plant parameter signals (instead of only the perturbation-related ones) as well as curvature information of the objective function into account. Quite recently Poveda and Kristić have introduced the concept of *'prescribed fixed time'*-ES (see [8], [9]). They accomplish convergence in a given finite time independent of the initial conditions by employing continuous gradient and Newton flows without a Lipschitz property.

In this article, we propose to extract the limit directly from the system dynamics. To achieve this, we conduct an in-depth study of the dynamics governing ES to deduce an asymptotic model for the generated output $y(t)$. We then solve the asymptotic model for its limit in terms of the output $y(t)$. This methodology is a form of Richardson extrapolation; a technique originally developed to speed up the convergence of sequences (see [10]). Similar ideas have found applications in a variety of fields such as perturbative quantum field theory (see e.g. [11], [12]) or machine learning (see e.g. [13]). The method is, to the best of our knowledge and exhaustive search through the literature, new and has not been applied in the context of control theory.

This paper is structured as follows: First, we discuss preliminaries by giving a short introduction to ES and then present the basic idea of accelerated convergence by discussing an ES loop in its most simple form. Next, we demonstrate how to analyze an ES loop theoretically to apply acceleration concepts. We then proceed with some numerical examples to illustrate the performance of the method and close with an outlook on possible future developments. After the bibliography we present detailed proofs.

## II. Preliminaries

### A. Problem formulation

We consider a function $f:\mathbb{R}\to\mathbb{R},\ f(x)=y$ with a local minimum at $x=L$ that we wish to find (for example to optimize a given objective). Such problems appear naturally in many situations such as tracking the optimal operating point of photovoltaic systems (see e.g. [14]) or controlling the optimal substrate flow in bioreactors (see

The authors gratefully acknowledge support by the German Academic Scholarship Foundation for organizing and funding the *Wissenschaftliches Kolleg* during which this project was started and the anonymous referees for their in-depth review. The forth and fifth author gratefully acknowledge funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 824046.

[1]J.-H. Metsch, Department of Mathematics, University of Freiburg, Germany (jan.metsch@math.uni-freiburg.de)

[2]J. Neuhauser, Institute of Fluid Mechanics, Karlsruhe Institute of Technology, Germany (jonathan.neuhauser@kit.edu)

[3]Jerome Jouffroy, Department of Mechanical and Electrical Engineering, University of Southern Denmark, Denmark (jerome@sdu.dk)

[4]Taous-Meriem Laleg-Kirati, Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Saudi Arabia (taousmeriem.laleg@kaust.edu.sa)

[5]Johann Reger is with the Control Engineering Group, Technische Universität Ilmenau, P.O. Box 10 05 65, 98684 Ilmenau, Germany (reger@ieee.org)

[⋆]Corresponding author: jan.metsch@math.uni-freiburg.de

e.g. [15]). Similar tasks arise in the backpropagation of neural networks (see e.g. [16], Chapter 4).

ES provides an algorithm that continuously improves an initial guess $x_0$ such that the resulting signal $x(t)$ converges exponentially to a neighbourhood of $L$. Intuitively this is achieved by the law

$$dx = -f'(x(t))dt. \tag{1}$$

To access the value $f'(x)$, a small oscillation $\epsilon\sin(\omega t)$ is added to $x$ leading to

$$f(x+\epsilon\sin(\omega t)) = f(x)+\epsilon f'(x)\sin(\omega t)+\mathcal{O}(\epsilon^2).$$

Running the output of $f$ through a high-pass filter and multiplying with $\sin(\omega t)$ produces the signal $\varphi(t) = \epsilon\sin^2(\omega t)f'(x)$. Replacing the actual gradient $f'(x(t))$ in (1) with $\varphi(t)$ gives the law $dx = -\varphi(t)dt$. A block diagram for this process is shown in Fig. 1. Closer analysis (see e.g. Chapter 1 in [2], Equation (1.9)) of this process suggests the approximate formula

$$x(t) \approx L+Ce^{-\epsilon bT}+\epsilon p(t), \tag{2}$$

where $p(t)$ is an oscillating function and $C$ and $b$ are constants. The two error terms '*compete*' with each other in the following sense: For large $\epsilon$ the exponential converges rapidly while the oscillating terms becomes large. For small $\epsilon$ the oscillation get suppressed while the exponential decay becomes slow.

This motivates studying the dynamics of the ES scheme described above in-depth to '*resolve*' the '*competing objectives*' in (2). The method we propose in this article is essentially designed to eliminate the exponential decay term in (2) which allows for fast convergence for sufficiently small values of $\epsilon$.

### B. *Accelerated convergence*

We present an easy example of accelerated convergence. A detailed review can be found in [17]. Consider the sequence $S_n := \sum_{j=1}^{n}\frac{1}{j^2}$. It is well known that $S_n \to \frac{\pi^2}{6}$. The convergence is very slow however as

$$\frac{\pi^2}{6} - S_n \sim \int_n^\infty \frac{dt}{t^2} = \frac{1}{n}. \tag{3}$$

To accelerate the convergence, we first construct an asymptotic model. Motivated by (3) it is reasonable to assume (and not too hard to prove) an expansion of the form

$$S_n = L+\sum_{j=1}^{\infty}\frac{a_j}{n^j}. \tag{4}$$

Here we abbreviated the limit of $S_n$ as $L := \frac{\pi^2}{6}$. A quick calculation shows that

$$\tilde{S}_n := \frac{1}{2}\left((n+2)^2S_{n+2}-2(n+1)^2S_{n+1}+n^2S_n\right) \tag{5}$$

satisfies $\tilde{S}_n = L+\mathcal{O}\left(\frac{1}{n^3}\right)$. Hence, the convergence has been accelerated. Indeed $L = 1.64493$, $\tilde{S}_{10} = 1.64481$ while $S_{10} = 1.54976$.

## III. Theory

We show how the concept of accelerated convergence may be applied to ES by studying two distinct loops starting with the easiest one and then demonstrating how a more complex situation may be analyzed. For the latter, we need to perform perturbation analysis to extract structural information about the dynamics. We remark that regular dependence of solutions on a perturbation parameter is a standard result and e.g. discussed in [18], Chapter 2, Section 9. The analysis essentially aims to derive a precise version of (2) similar to (4). Considering shifts in time $t \to t+T$ we then derive extraction schemes for the limit of the system, similar to (5). Finally, we point out that a similar analysis has been performed in [19] for the Mathieu equation (see Chapter 11, Section 4).

### A. *Basic model*

Let $a,b,L \in \mathbb{R}$ and $f(x) := a+b(x-L)^2$. Initially, we analyze the ES loop depicted in Fig. 1.

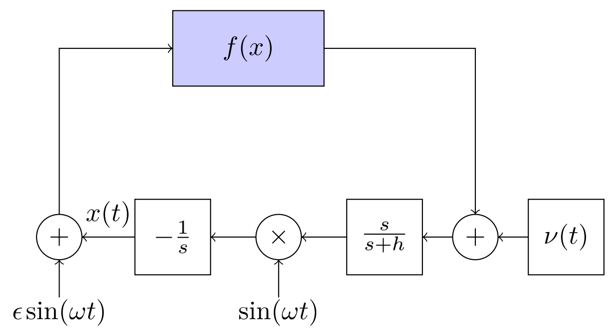


Fig. 1. Extremum seeking loop

$\nu(t)$ is a noise source, which will be included in the simulations in Section IV. Denoting the high-pass filter by $\mathcal{F}$, Figure 1 corresponds to the integral equation

$$x(t) = x(0) - \int_0^t \mathcal{F}\left[f\left(x(\tau)+\epsilon\sin(\omega\tau)\right) + \nu(t)\right]\sin(\omega\tau)d\tau. \tag{6}$$

**Proposition III.1.** *Let $T := \frac{2\pi}{\omega}$, $\theta := e^{-\epsilon bT}$ and $x : [0,\infty) \to \mathbb{R}$ be a solution to the loop in Fig. 1 with $\nu \equiv 0$. For any $t \geq 0$, put $x_n := x(t+nT)$. Then*

$$L = \frac{(x_0-x_1)x_2+\theta x_0(x_2-x_1)}{x_0-(1+\theta)x_1+\theta x_2}+\mathcal{O}(\epsilon^2).$$

*Additionally, putting*

$$g := \frac{(x_0-x_1)(x_2-x_3)}{(x_1-x_2)(x_0-x_3)}$$

*the following extraction law for $\theta$ holds:*

$$\theta = \frac{1-g}{2g} - \frac{1}{2g}\sqrt{-4g^2+(g-1)^2}+\mathcal{O}(\epsilon^2)$$

*Proof.* Let $y(t) := x(t) - L$. Then $\dot{y} = \dot{x}$. Differentiating (6) and using $n \equiv 0$ gives

$$\begin{aligned}&\dot{y} + \epsilon b(1-\cos(2\omega t))y + by^2\sin(\omega t)\\ =& -b\epsilon^2\sin(\omega t)^3.\end{aligned} \tag{7}$$

This is a Ricatti equation without a closed-form solution. We consider $\epsilon$ as a perturbative parameter and only study (7) to first order. This justifies dropping the $\epsilon^2$-term in (7) which gives a Bernoulli Equation. Putting

$$x_0(t) := \exp\left[-\epsilon bt + \frac{\epsilon b}{2\omega}\sin(2\omega t)\right]$$

we derive the following formula for its solution $x$ in Appendix B:

$$x(t) = L + \frac{x_0(t)}{C + b\int_0^t \sin(\omega s)x_0(s)ds}. \tag{8}$$

The constant $C$ is related to the initial value $x(0)$. Recalling $\theta = e^{-\epsilon bT}$, it is clear that $x_0(t+T) = \theta x_0(t)$. Let $\varphi(t) := C + b\int_0^t \sin(\omega s)x_0(s)ds$ so that $\dot{\varphi}(t+T) = \theta\dot{\varphi}(t)$. Lemma A.1 in Appendix A implies $\varphi(t) = \tilde{C} + X(t)$ for a constant $\tilde{C}$ and a function $X$ satisfying $X(t+T) = \theta X(t)$. This gives the following equations:

$$\begin{aligned}x(t) - L &= \frac{x_0(t)}{\tilde{C} + X(t)}\\ x(t+T) - L &= \theta\frac{x_0(t)}{\tilde{C} + \theta X(t)}\\ x(t+2T) - L &= \theta^2\frac{x_0(t)}{\tilde{C} + \theta^2 X(t)}\end{aligned} \tag{9}$$

If we regard $x(t+nT)$ as known parameters, (9) can be thought of as a nonlinear system of ordinary equations for $L, \tilde{C}, X(t)$ and $x_0(t)$. A solution for $L$ then gives a formula of the limit in terms of the values $x_n := x(t+nT)$. Direct computation shows

$$L = \frac{(x_0 - x_1)x_2 + \theta x_0(x_2 - x_1)}{x_0 - (1+\theta)x_1 + \theta x_2}. \tag{10}$$

Equation (10) uses the data points $x(t)$, $x(t+T)$ and $x(t+2T)$ and fits them onto the solution (8). It eliminates the unknown values $x_0(t)$, $\tilde{C}$ and $X(t)$ and hence requires three data points. Note however that $\theta = e^{-\epsilon bT}$ features in the extraction law. While $T$ and $\epsilon$ are part of the design of the loop and therefore known, the parameter $b$ is part of the function $f$ and in general not known. By incorporating a fourth data point into the analysis we can eliminate $\theta$ from (10). Indeed we note that (10) also holds for $t \to t+T$ and hence

$$\begin{aligned}L =& \frac{(x_0 - x_1)x_2 + \theta x_0(x_2 - x_1)}{x_0 - (1+\theta)x_1 + \theta x_2}\\ =& \frac{(x_1 - x_2)x_3 + \theta x_1(x_3 - x_2)}{x_1 - (1+\theta)x_2 + \theta x_3}.\end{aligned} \tag{11}$$

This is a quadratic equation for $\theta$ with two solutions. However, putting

$$g := \frac{(x_0 - x_1)(x_2 - x_3)}{(x_1 - x_2)(x_0 - x_3)} \tag{12}$$

we prove in Appendix C that

$$\theta = \frac{1-g}{2g} - \frac{1}{2g}\sqrt{-4g^2 + (g-1)^2} \tag{13}$$

by exploiting $\theta = e^{-\epsilon bT} \in (0,1)$. $\square$

We have derived an extraction scheme that uses four data points. It first applies (13) to find $\theta$ and then uses (10) to extract the limit $L$.

### *B. Including a drift*

This Subsection demonstrates how to extend the analysis from Subsection III-A to other loops by considering an example. We modify the ES loop in Fig. 1 by taking $f(x,t) = (x - L - q(t))^2$ to be explicitly time dependent. We refer to the resulting loop as modified Fig. 1. Here $q(t) = q_0 e^{-\delta t}$ for a small positive drift parameter $\delta > 0$.

**Proposition III.2.** *Let $x$ be any solution of modified Fig. 1 with $\nu \equiv 0$ and put $z(t) := (x(t) - L - q(t))^{-1}$. Then*

$$z(t) = \sum_{j=0}^{\infty}\left[\delta^j e^{-j\delta t}\sum_{k=0}^{j+1} e^{k\epsilon t}p_{jk}(t)\right] + O(\epsilon^2).$$

*where all function $p_{jk}$ are $T$-periodic.*

*Proof.* We put $y := x - L - q$. Differentiating the analogue of (6) with time-dependent $f$ and exploiting and $\dot{q}(t) = -\delta q(t)$ gives

$$\dot{y}(t) + 2\epsilon\sin(\omega t)^2 + y^2\sin\omega t - \delta q = -\epsilon^2\sin(\omega t)^3 \tag{14}$$

After dropping $\epsilon^2$ as in the proof of Proposition III.1 and letting $z = \frac{1}{y}$, we get

$$\dot{z} - 2\epsilon\sin^2(\omega t)z + \delta q(t)z^2 = \sin(\omega t). \tag{15}$$

Equation (15) is another Riccati equation without closed-form solution. Still we may extract structural properties by perturbation analysis. Proposing $z(t) = \sum_{n\geq 0} z_n(t)\delta^n$ we get the following infinite system of linear ordinary differential equations: For $n = 0$:

$$\begin{cases}\dot{z}_0 - 2\epsilon\sin^2(\omega t)z_0 = \sin(\omega t)\\ z_0(0) = z(0)\end{cases} \tag{16}$$

For $n \geq 1$:

$$\begin{cases}\dot{z}_n - 2\epsilon\sin^2(\omega t)z_n = -q(t)\displaystyle\sum_{j=0}^{n-1} z_j z_{n-1-j}\\ z_n(0) = 0\end{cases} \tag{17}$$

Solving for $z_0$ is trivial. Working iteratively, the $n$-th equation is linear in $z_n$ with nonlinearities only in the already known functions $z_k$ with $k \leq n-1$. An inductive argument shows

$$z_n(t) = e^{\epsilon t}p_0^{(n)}(t) + \sum_{j=1}^{n}\sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t}p_{jk}^{(n)}(t) \tag{18}$$

with $T$-periodic functions $p_*^{(n)}$ for $n \geq 1$ and $z_0(t) = p_0^{(0)}(t) + e^{\epsilon t} p_1^{(0)}(t)$ with $T$-periodic functions $p_*^{(0)}$. Resumming gives the Lemma. □

To derive an exact extraction scheme from the expansion given in Proposition III.2, we would require infinitely many data points to eliminate all terms in the series. For small $\delta$ we may, however, truncate the perturbation series and construct a finite extraction scheme, which we demonstrate in the following Corollary.

**Corollary III.3.** *Let $A := e^{\epsilon T}$ and $x$ be any solution to the loop in modified Fig. 1 with $\nu \equiv 0$. Put $h(t) := x(t) - q(t)$ and $h_n := h(t+nT)$. Then*

$$L = \frac{h_1 h_0 - (1+A) h_2 h_0 + A h_2 h_1}{-h_2 + (1+A) h_1 - A h_0} \tag{19}$$

*up to an error of order $\mathcal{O}(\delta) + \mathcal{O}(\epsilon^2)$.*

*Proof.* As it is not entirely trivial, we also demonstrate how to derive a $\mathcal{O}(\delta^2)$-extraction law. Let $B := e^{\delta T}$ and $z_a := z(t + aT)$. It is readily checked that

$$\begin{aligned} 0 =& z_5 - (1 + A + B(1 + A + A^2) z_4 \\ &+ \big(A + B(1+A)A(1 + A + A^2) \\ &\qquad\qquad + B^2 A(1 + A + A^2)\big) z_3 \\ &- \big((AB(1 + A + A^2) + (A+1)B^2(A + A^2 + A^3) \\ &\qquad\qquad + A^3 B^3\big) z_2 \\ &+ (AB^2(A + A^2 + A^3) + (A+1)A^3 B^3) z_1 \\ &- A^4 B^3 z_0. \end{aligned}$$

Summarizing this as $\sum_{0 \leq i \leq 5} \mu_i z_i = 0$ and recalling the definition of $z$ we get the implicit extraction law

$$\sum_{i=0}^{5} \mu_i \prod_{\substack{j=0 \\ j \neq i}}^{5} (x_j - B^j q_0 - L) = 0.$$

For zero order extraction scheme one argues analogously. Solving the resulting implicit law gives (19). □

Note that extraction schemes for $q_0$ and $\delta$ are required, which we do not include here. To derive them, one employs the strategy that demonstrated following (11).

Considering the statement of Proposition III.2, we must have convergence of the series for its truncation to be a valid approximation. For the series to be convergent on $[0, \infty)$, demanding $\delta > \epsilon$ is plausible as the perturbation series grows exponentially otherwise. A sufficient but not necessary criterion to achieve convergence on $[0, \frac{1}{2\delta}]$ is

$$\Gamma := 24 e^{\frac{2\epsilon}{\omega}} |q_0| \left( |z(0)| + \frac{1}{\delta} \right) \overset{!}{<} 1. \tag{20}$$

To prove (20) one applies the variation of parameters formula to (17) and derives a recursive upper bound $u_n$ for $|z_n|$. Solving the recursion and demanding $\sum_{n \geq 0} u_n \delta^n$ to be convergent then gives (20).

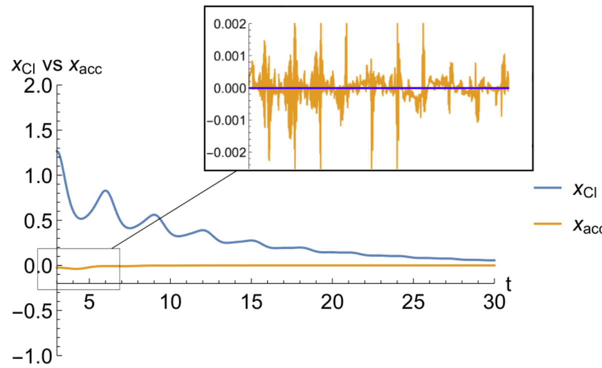


Fig. 2. Classical ES vs accelerated ES

## IV. Simulation

We implemented the equations studied above in Mathematica: All differential equations have been numerically solved using the `NDSolve` function. The following graphics are generated by evaluating the extraction schemes at $t \geq 0$ and plotting the result.

### A. Simple model

Fig. 2 shows the classical ES (as depicted in Fig. 1 without noise) versus the accelerated ES for parameters $T := 3$, $b := 2$, $\epsilon := .01$ and $L = 0$. The zoomed-in section of the figure shows that the accelerated curve oscillates around $L = 0$ with amplitude $\propto \epsilon^2$ as is to be expected from the theory. The initial conditions of the loop are absent in the accelerated scheme for $t \geq 0$. This is due to the extraction scheme using the data points $x(t + kT)$ with $k \leq 3$ (see (10) and (13)).

Fig. 3 demonstrates the extraction of $\theta$ and shows excellent agreement with the exact value $e^{-\epsilon bT} \approx .9418$.

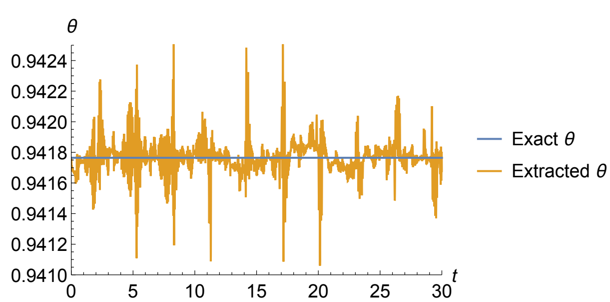


Fig. 3. Extraction of $\theta$

### B. Including noise

We now include the noise block in Fig. 1. The noise is realized as a piecewise constant function that takes randomized values in $[-N_0, N_0]$ on intervals of length $dt$. In all following simulations we use $b = 2$, $T = 3$, $\epsilon = .01$

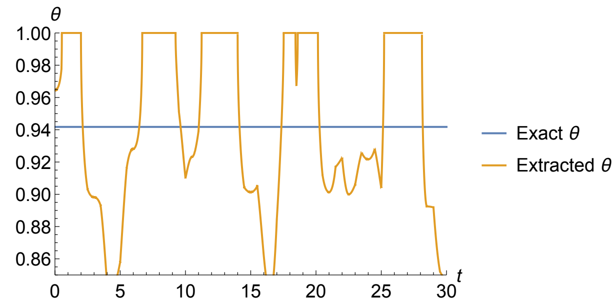


Fig. 4. Extraction of $\theta$ ($N_0 = \epsilon^2$).

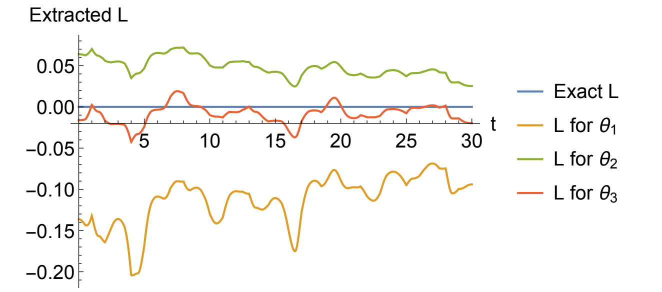


Fig. 6. Extraction of $L$ (averaged $\theta$, $N_0 = \epsilon^2$).

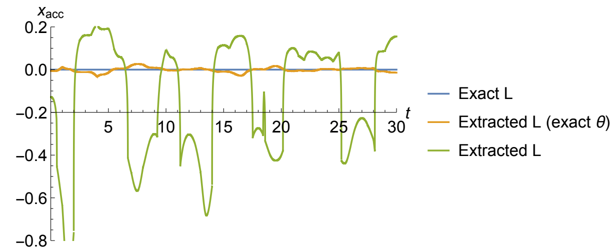


Fig. 5. Extraction of $L$ ($N_0 = \epsilon^2$).

and $dt = .5$. To explain the following simulation results, we remark that the inclusion of a noise source introduces a new term in (7):

$$\begin{aligned}&\dot{y} + \epsilon b(1 - \cos(2\omega t))y + by^2 \sin(\omega t)\\&= -b\epsilon^2 \sin(\omega t)^3 - \nu(t)\sin(\omega t)\end{aligned} \tag{21}$$

The analysis in Subsection III-A is based on dropping terms of order $\epsilon^2$ suggesting that noise of higher amplitude corrupts the method. Indeed, the scheme breaks down for $N_0 = \epsilon$. Taking $N_0 = \epsilon^2$ renders the *noise-term* in (21) to be of order $\epsilon^2$ suggesting the extraction schemes to work. Fig. 4 and Fig. 5 show the extraction of $\theta$ and $L$ with exact and extracted $\theta$ respectively. The cutoff visible in Fig. 4 is caused by cutting off $g$ at $g = \frac{1}{3}$ as larger values lead to complex $\theta$. Extraction of $L$ using the exact value of $\theta$ works fine. However, inclusion of noise causes noticeable oscillations in the extraction of $\theta$ which render the full extraction scheme for $L$ to work poorly. Averaging $\theta$ over time can, however, drastically improve this result. Fig. 6 shows the extracted value of $L$ that is obtained when using the average value $\theta_k$ of $\theta$ on $[0, kT]$ in (10).

Smaller $N_0$ such as $N_0 = \epsilon^{\frac{5}{2}}$ render the extraction of $\theta$ accurate enough to extract $L$ without having to resort to averaging procedures. Modifying $dt$ or adding an offset of order at most $\epsilon^2$ to the noise does not change the simulation results.

### C. Including a drift

For all following simulations, we choose $T = 3$, $L = 0$ and $z(0) = \frac{1}{2}$. Additionally, taking $\delta = .4$, $\epsilon = .1$ and $q_0 = .01$ gives $\Gamma = .79$ thereby ensuring the scheme to function properly as is verified in Fig. 7. Reusing the terminology from the previous Subsection, Fig. 7 also shows the effect of noise with $N_0 = \epsilon^2$ on the scheme. $\Gamma < 1$ is, however, not necessary: Taking $\epsilon = .2$, $q_0 = .01$ and $\delta \in \{1, .1, 10^{-9}\}$ produces accelerated convergence with high values of $\Gamma$ (see Fig. 8). However, taking $\delta = .1$, $\epsilon = .01$ and e.g. $q_0 \in \{.4, .05\}$ shows that that for $\Gamma > 1$ the acceleration scheme can in fact break down.

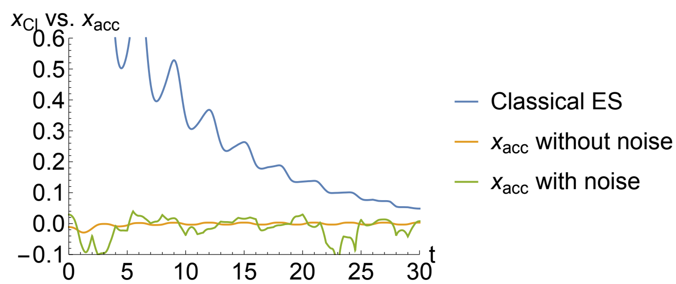


Fig. 7. Classical vs accelerated ES.

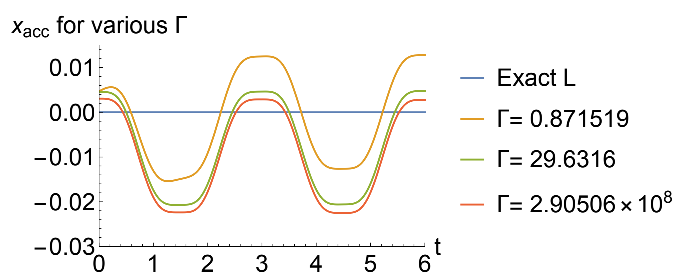


Fig. 8. Various values of $\Gamma/\ \delta$.

Fig. 8 is restricted to $0 \le t \le 6$ to make the differences between the curves visible. Again, modification of $dt$ and the inclusion of a small offset have no effect on the results.

## V. Summary and Outlook

We have demonstrated how ES loops can be analyzed by considering a perturbation expansion around simpler loops and how the resulting information can be used to derive extraction schemes that speed up the convergence drastically. This statement also holds in comparison to other acceleration schemes, such as fixed-time extremum seeking (see e.g. [9]). The obvious downside of the scheme is that it requires more information about the structure of the system that is to be optimized. The presented scheme is therefore suited to systems of which the physics (but not necessarily the system parameters!) are known and require fast convergence with little oscillations in the steady state, such as in robotics applications. There are still many open questions to be considered: General statements and formal proofs are needed to make the proof of concept presented here more rigorous. This also includes a detailed discussion concerning convergence. Experimental evidence is needed to show the suitability to real-world applications. Finally, additional generalizations such as multidimensional ES are still to be discussed.

## Appendix

### A. Calculus Lemmata

**Lemma A.1.** *Let $L>0$, $1\neq a\in\mathbb{R}^+$ and $y\in C^1(\mathbb{R})$ such that $y'(x+L)=ay'(x)$. Then*

$$y(x)=\alpha+a^{\frac{x}{L}}P(x)$$

*for some $\alpha\in\mathbb{R}$ and $L$-periodic $P\in C^1(\mathbb{R})$.*

*Proof.* We only prove the Lemma for $x\geq 0$. For $x<0$ one argues similarly. Since $(y(x)-ay(x-L))'=0$ there exists some $C\in\mathbb{R}$ such that $y(x)=C+ay(x-L)$. Let $x\geq 0$. There exist unique $n\in\mathbb{N}_0$ and $h\in[0,L)$ such that $x=nL+h$. Using $n=\frac{x-h}{L}$ we compute

$$\begin{aligned}
y(x)&=C+ay(x-L)=C(1+a)+a^2y(x-2L)\\
&=\ldots=C(1+a+\ldots+a^{n-1})+a^ny(h)\\
&=C\frac{a^n-1}{a-1}+a^ny(h)\\
&=-\frac{C}{a-1}+a^{\frac{x}{L}}a^{-\frac{h}{L}}\left(y(h)+\frac{C}{a-1}\right).
\end{aligned}$$

Setting $\alpha:=-\frac{C}{a-1}$ and $P(x):=a^{-\frac{h}{L}}(y(h)-\alpha)$ we get $y(x)=\alpha+a^{\frac{x}{L}}P(x)$. $P$ is $L$-periodic as $h(x+L)=h(x)$ and $P\in C^1$ follows from $P(x)=a^{-\frac{x}{L}}(y(x)-\alpha)$. $\square$

**Lemma A.2.** *Let $\eta,\omega,a\in\mathbb{R}$, $T:=\frac{2\pi}{\omega}$, $q\in C^0(\mathbb{R})$ be $T$-periodic and $y$ be a solution to*

$$\dot{y}(t)+2a\sin^2(\omega t)y(t)=e^{\eta t}q(t).$$

*Then $y(t)=e^{-at}p_1(t)+e^{\eta t}p_2(t)$ for some $T-$periodic functions $p_1$ and $p_2$.*

*Proof.* Using $2\sin^2(\omega t)=1-\cos(2\omega t)$ it is readily seen that

$$\frac{d}{dt}\left[y(t)e^{at-\frac{a\sin(2\omega t)}{2\omega}}\right]=q(t)e^{\eta t}e^{at-\frac{a\sin(2\omega t)}{2\omega}}.$$

Lemma A.1 implies the existence of a constant $\rho_0\in\mathbb{R}$ and a $T$-periodic function $\rho(t)$ such that

$$y(t)e^{at-\frac{a\sin(2\omega t)}{2\omega}}=\rho_0+e^{\eta t}e^{at}\rho(t).$$

This proves the Lemma. $\square$

### B. Proof of Equation (8)

As described in the paragraphs preceding (8) we study the ODE

$$\dot{y}+\epsilon b(1-\cos(2\omega t))y+by^2\sin(\omega t)=0.$$

We put $z:=\frac{1}{y}$ such that $\dot{z}=-y^{-2}\dot{y}$ and get

$$\dot{z}-\epsilon b(1-\cos(2\omega t))z=b\sin(\omega t).$$

Note that $x_0(t)=\exp(-\epsilon b(t-\frac{\sin(2\omega t)}{2\omega}))$ defines an integrating factor for the left hand side. Hence

$$\frac{d}{dt}\big(z(t)x_0(t)\big)=bx_0(t)\sin(\omega t).$$

Integrating from 0 to $t$ and abbreviating $z(0)x_0(0)=:C$ yields

$$z(t)x_0(t)=C+b\int_0^t x_0(s)\sin(\omega s)ds.$$

Equation (8) follows by definition of $z$.

### C. Proof of Equation (13)

Proving that (13) is true up to the sign in front of the square root is trivial. To prove that it is '$-$', we use $\theta=e^{-\epsilon bT}\in(0,1)$. We get $(g-1)^2\geq 4g^2$ and thus $-1\leq g\leq\frac{1}{3}$ as $\theta\in\mathbb{R}$. Additionally, $g=0$ is not possible by (12). Indeed, $\theta\neq 1$ and (9) imply $x_0\neq x_1$ and $x_2\neq x_3$. For $g\in[-1,0)$ we have $\frac{1-g}{2g}<0$. Hence

$$\pm\frac{1}{2g}\sqrt{-4g^2+(g-1)^2}\overset{!}{\geq}0$$

This implies that $-$ is the correct sign. For $g\in(0,\frac{1}{3})$ we note that $\frac{1-g}{2g}\geq 1$ and thus

$$\pm\frac{1}{2g}\sqrt{-4g^2+(g-1)^2}\overset{!}{\leq}0$$

implying again that $-$ is correct.

### D. Proof of Equations (17) and (18)

We substitute $z(t)=\sum_{n\geq 0}z_n(t)\delta^n$ into (15) to get

$$\begin{aligned}
\sum_{n=0}^{\infty}\left(\dot{z}_n-2z_n\epsilon\sin^2(\omega t)\right)\delta^n&=\sin(\omega t)\\
&-q(t)\sum_{n=0}^{\infty}\left[\delta^{n+1}\sum_{j=0}^{n}z_jz_{n-j}\right].
\end{aligned}$$

Equation (17) follows by comparing the coefficients of $\delta^n$. Equation (18) is proven inductively. For $n=0$ it follows by applying Lemma A.2 to

$$\dot{z}_0-2\epsilon\sin^2(\omega t)z_0=\sin(\omega t).$$

Supposing (18) for $z_0,\ldots,z_n$, it is checked by direct computation that there exist $T$-periodic functions $q_{\kappa,\beta}$ such that

$$q(t)\sum_{j=0}^{n}z_jz_{n-j}=\sum_{\kappa=1}^{n+1}\left[e^{-\kappa\delta}\sum_{\beta=0}^{\kappa+1}e^{\epsilon\beta t}q_{\kappa,\beta}(t)\right].$$

Using the linearity of (17) and Lemma A.2 readily implies (18) for $z_{n+1}$.

*E. Proof of Equation (20)*

**Lemma E.3.** *Let $\mu > 0$ and $f \in C^0(\mathbb{R})$ be positive. Then, for all $t \in [0, \frac{1}{2\mu}]$*

$$\int_0^t e^{-\mu s} f(s) ds \leq 2e^{-\mu t} \int_0^t f(s) ds.$$

*Proof.* Let $F(t) := \int_0^t e^{-\mu s} f(s) ds$. As $F$ is increasing and $F(0) = 0$, we may estimate

$$\begin{aligned}\int_0^t e^{-\mu s} f(s) ds &= e^{-\mu t} F(t) + \mu \int_0^t e^{-\mu s} F(s) ds \\ &\leq (e^{-\mu t} + \mu t) F(t).\end{aligned}$$

Using $x \leq e^{-x}$ for $x \leq \frac{1}{2}$, the Lemma follows. □

**Lemma E.4.** *Let $\epsilon > 0$, $\omega \in \mathbb{R}$, $R \in C^0(\mathbb{R})$ and $\xi$ solve $\dot{\xi}(t) - 2\epsilon \sin^2(\omega t)\xi(t) = R(t)$. Then, for $t \in [0, \frac{1}{2\epsilon}]$*

$$|\xi(t)| \leq |\xi(0)| e^{\frac{\epsilon}{2\omega}} e^{\epsilon t} + 2e^{\frac{\epsilon}{\omega}} \int_0^t |R(s)| ds.$$

*Proof.* It is clear that

$$\begin{aligned}\xi(t) =& e^{\epsilon t - \frac{\epsilon}{2\omega}\sin(2\omega t)} \Big[\xi(0) \\ &+ \int_0^t e^{-\epsilon s + \frac{\epsilon}{2\omega}\sin(2\omega s)} R(s) ds\Big].\end{aligned} \tag{22}$$

Estimating the second term using Lemma E.3 to the second term in (22) gives the Lemma. □

We now prove (20).

*Proof.* Put $t_0 := \frac{1}{2\delta}$ and $u_k := \sup_{0 \leq s \leq t_0} |z_k(s)|$. Applying Lemma E.4 to (16) gives

$$u_0 \leq |z(0)| e^{\frac{\epsilon}{2\omega} + \epsilon t_0} + 2e^{\frac{\epsilon}{\omega}} t_0 =: \alpha_0. \tag{23}$$

For $n \geq 0$, applying Lemma E.4 to (17) and subsequently using Lemma E.3 gives

$$\begin{aligned}u_{n+1} &\leq 2e^{\frac{\epsilon}{\omega}} |q_0| \sum_{j=0}^n \int_0^{t_0} e^{-\delta s} |z_j(s) z_{n-j}(s)| ds \\ &\leq 4e^{\frac{\epsilon}{\omega}} |q_0| e^{-\delta t_0} \sum_{j=0}^n \int_0^{t_0} |z_j(s) z_{n-j}(s)| ds \\ &\leq 4e^{\frac{\epsilon}{\omega}} |q_0| e^{-\delta t_0} t_0 \sum_{j=0}^n u_j u_{n-j}.\end{aligned}$$

Note $t_0 e^{-\delta t_0} \leq \delta^{-1}$, put $C := 4(e\delta)^{-1} e^{\frac{\epsilon}{\omega}} |q_0|$ and, for $n \geq 0$, define $\alpha_n$ by

$$\alpha_{n+1} = C \sum_{j=0}^n \alpha_j \alpha_{n-j}. \tag{24}$$

An inductive argument shows $u_n \leq \alpha_n$ and hence $\sum_{n \geq 0} z_n \delta^n$ converges absolutely when $\sum_{n \geq 0} \alpha_n \delta^n$ converges. Consider the generating function $A(x) := \sum_{j \geq 0} \alpha_j x^j$. Using (24) it is readily checked that $CxA^2(x) = A(x) - \alpha_0$ and hence

$$A(x) = \frac{1 - \sqrt{1 - 4C\alpha_0 x}}{2Cx}. \tag{25}$$

Expanding (25) and using Stirling's approximation gives

$$\alpha_n \sim \frac{(4C)^n}{\sqrt{\pi}(n+1)^{\frac{3}{2}}} \alpha_0^{n+1}.$$

Thus, $\sum_{n \geq 0} \alpha_n \delta^n$ converges when $4C\alpha_0\delta < 1$. Inserting $\alpha_0$ from (23) gives (20). □

## References


[1] M. Leblanc, "Sur i'electrification des chemins de fer au moyen de courants alternatifs de frequence elevee," *Revue générale de l'électricité*, 1922.

[2] K. B. Ariyur and M. Krstić, *Real-Time Optimization by Extremum-Seeking Control*. John Wiley & Sons, 2003.

[3] B. Calli, W. Caarls, P. Jonker, and M. Wisse, "Comparison of extremum seeking control algorithms for robotic applications," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 3195–3202.

[4] D. Nešić, "Extremum seeking control: Convergence analysis," *European Journal of Control*, vol. 15, no. 3-4, pp. 331–347, 2009.

[5] S. J. Moura and Y. A. Chang, "Lyapunov-based switched extremum seeking for photovoltaic power maximization," *Control Engineering Practice*, vol. 21, no. 7, pp. 971–980, 2013.

[6] H. Malek and Y.-Q. Chen, "Fractional order extremum seeking control: Performance and stability analysis," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 3, pp. 1620–1628, 2016.

[7] M. A. Haring, "Extremum-seeking control: convergence improvements and asymptotic stability," Ph.D. dissertation, NTNU, 2016.

[8] J. I. Poveda and M. Krstić, "Fixed-time gradient-based extremum seeking," in *American Control Conference*, 2020, pp. 2838–2843.

[9] ——, "Fixed-time newton-like extremum seeking," in *IFAC World Congress*, 2020, pp. 5356–5361.

[10] L. F. Richardson, "The approximate arithmetical solution by finite differences of physical problems involving differential equations with an application to the stresses in a masonry dam," *Transactions of the Royal Society of London*, vol. Ser. A, no. 210, pp. 307–357, 1910.

[11] M. Borinsky, G. V. Dunne, and M. Meynig, "Semiclassical trans-series from the perturbative Hopf-algebraic Dyson-Schwinger equations: $\phi^3$ qft in 6 dimensions," *SIGMA*, vol. 17, p. 087 (26 pages), 2021.

[12] C. M. Bender, A. Pelster, and F. Weissbach, "Boundary-layer theory, strong-coupling series, and large-order behavior," *Journal of Mathematical Physics*, vol. 43, no. 8, pp. 4202–4220, 2002.

[13] F. Bach, "On the effectiveness of richardson extrapolation in machine learning," *SIAM Journal on Mathematics and Data Science*, vol. 3, no. 4, pp. 1251–1277, 2020.

[14] 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 Transactions on Power Electronics*, vol. 25, no. 10, pp. 2531–2540, 2010.

[15] H.-H. Wang, M. Krstić, and G. Bastin, "Optimizing bioreactors by extremum seeking," *International Journal of Adaptive Control and Signal Processing*, vol. 13, no. 8, pp. 651–669, 1999.

[16] O. Calin, *Deep Learning Architectures: A Mathematical Approach*. Springer, 2020.

[17] C. M. Bender and C. Heissenberg, "Convergent and divergent series in physics," 2017, Saalburg Summer School, Germany.

[18] H. Amann, *Ordinary Differential Equations: An Introduction to Nonlinear Analysis*. De Gruyter, 2011.

[19] C. M. Bender and S. A. Orszag, *Advanced mathematical methods for scientists and engineers: I: Asymptotic methods and perturbation theory.* Springer, 1999.

# Supplementary

## Structure of this Part

In the first section, we give more details on the technical Lemmas that are presented in Appendix A. Section 2 provides derivations that have been left out in Section III.A and the corresponding parts of the appendix. Similarly, Section 3 provides derivations that have been left out in Section III.B and the corresponding parts of the appendix. Finally, Section 4 describes how the simulations in Section IV have been generated.

## F. Technical Lemmas

**Lemma F.1.** *Let $L > 0$, $1 \neq a \in \mathbb{R}^+$ and $y \in C^1(\mathbb{R})$ such that $y'(x+L) = ay'(x)$. Then*

$$y(x) = \alpha + a^{\frac{x}{L}} P(x)$$

*for some $\alpha \in \mathbb{R}$ and $L$-periodic $P \in C^1(\mathbb{R})$.*

*Proof.* We only prove the Lemma for $x \geq 0$. For $x < 0$ one argues similarly. Since $(y(x) - ay(x-L))' = 0$ there exists some $C \in \mathbb{R}$ such that $y(x) = C + ay(x-L)$. Let $x \geq 0$. There exist unique $n \in \mathbb{N}_0$ and $h \in [0, L)$ such that $x = nL + h$. Using $n = \frac{x-h}{L}$ we compute

$$\begin{aligned} y(x) &= C + ay(x-L) \\ &= C(1+a) + a^2 y(x-2L) \\ &= \ldots \\ &= C(1+a+\ldots+a^{n-1}) + a^n y(h) \\ &= C\frac{a^n - 1}{a-1} + a^n y(h) \\ &= -\frac{C}{a-1} + a^{\frac{x}{L}} a^{-\frac{h}{L}} \left( y(h) + \frac{C}{a-1} \right). \end{aligned} \tag{26}$$

We now define

$$\alpha := -\frac{C}{a-1} \quad \text{and} \quad P(x) := a^{-\frac{h}{L}} \left( y(h) - \alpha \right).$$

Inserting these definitions into Equation (26), we get

$$y(x) = \alpha + a^{\frac{x}{L}} P(x) \tag{27}$$

$P$ is $L$-periodic. Indeed, if $x = nL + h$, then $x + L = (n+1)x + h$ and hence

$$P(x+L) = a^{-\frac{h}{L}} (y(h) - \alpha) = P(x).$$

To prove $P \in C^1$ we rewrite Equation (27) as

$$P(x) = a^{-\frac{x}{L}} (y(x) - \alpha) \in C^1.$$

In the last step we have used the regularity of $y$. □

**Lemma F.2.** *Let $\eta, \omega, a \in \mathbb{R}$, $T := \frac{2\pi}{\omega}$, $q \in C^0(\mathbb{R})$ be $T$-periodic and $y$ be a solution to*

$$\dot{y}(t) + 2a\sin^2(\omega t) = e^{\eta t} q(t).$$

*Then $y(t) = e^{-at} p_1(t) + e^{\eta t} p_2(t)$ for some $T-$periodic functions $p_1$ and $p_2$. Additionally, $p_2 = 0$ if $q = 0$.*

*Proof.* Using $2\sin^2(\omega t) = 1 - \cos(2\omega t)$, we compute

$$\begin{aligned} \frac{d}{dt}\left[ y(t) e^{at - \frac{a\sin(2\omega t)}{2\omega}} \right] &= e^{at - \frac{a\sin(2\omega t)}{2\omega}} \left[ \dot{y}(t) + y(t) \frac{d}{dt}\left( at - \frac{a\sin(2\omega t)}{2\omega} \right) \right] \\ &= e^{at - \frac{a\sin(2\omega t)}{2\omega}} \left[ \dot{y}(t) + y(t)\left( a - a\cos(2\omega t) \right) \right] \\ &= e^{at - \frac{a\sin(2\omega t)}{2\omega}} \left[ \dot{y}(t) + 2a\sin^2(\omega t) y \right] \\ &= e^{at - \frac{a\sin(2\omega t)}{2\omega}} e^{\eta t} q(t). \end{aligned} \tag{28}$$

In the last step we have used the ODE that $y$ solves. If $q = 0$, the right hand side in Equation (28) vanishes and we deduce that there is a constant $C$ such that

$$y(t)e^{at - \frac{a\sin(2\omega t)}{2\omega}} = C.$$

Hence

$$y(t) = e^{-at}Ce^{\frac{a\sin(2\omega t)}{2\omega}}$$

which is the claimed formula. For general $q$ we define

$$f(t) := e^{at - \frac{a\sin(2\omega t)}{2\omega}} e^{\eta t} q(t) \quad \text{and} \quad Y(t) := \frac{d}{dt}\left[ y(t)e^{at - \frac{a\sin(2\omega t)}{2\omega}} \right]. \tag{29}$$

Note that

$$f(t+T) = e^{\eta(t+T)}e^{a(t+T)}e^{-a\frac{\sin(2\omega(t+T))}{2\omega}} = e^{(a+\eta)T}f(t).$$

Using Equations (28) and (29) we get

$$Y'(t+T) - e^{(a+\eta)T}Y'(t) = f(t+T) - e^{(a+\eta)T}f(t) = 0.$$

Lemma F.1 implies the existence of a constant $\rho_0 \in \mathbb{R}$ and a $T$-periodic function $\rho(t)$ such that

$$y(t)e^{at - \frac{a\sin(2\omega t)}{2\omega}} = Y(t) = \rho_0 + e^{\eta t}e^{at}\rho(t).$$

This yields the claimed formula:

$$y(t) = e^{-at}\rho_0 e^{\frac{a\sin(2\omega t)}{2\omega}} + e^{\eta t}\rho(t)e^{\frac{a\sin(2\omega t)}{2\omega}}$$

□

**Proof of Equation (8)**

We define

$$x_0(t)\exp\left[-\epsilon b\left(t - \frac{\sin(2\omega t)}{2\omega}\right)\right] \tag{30}$$

**Lemma F.3.** *Let $\epsilon, b, \omega \in \mathbb{R}$ and $y$ be a solution of*

$$\dot{y} + \epsilon b(1 - \cos(2\omega t))y + by^2\sin(\omega t) = 0.$$

*Then there exists a constant $C$ such that*

$$y(t) = \frac{x_0(t)}{C + b\int_0^t \sin(\omega s)x_0(s)ds}.$$

*Proof.* We put $z := \frac{1}{y}$ such that $\dot{z} = -y^{-2}\dot{y}$ or equivalently $\dot{y} = -z^{-2}\dot{z}$. This gives

$$\begin{aligned}
\dot{z} - \epsilon b(1 - \cos(2\omega t))z &= -y^{-2}\dot{y} - \epsilon b(1 - \cos(2\omega t))y^{-1} \\
&= -y^{-2}(\dot{y} + \epsilon b(1 - \cos(2\omega t))y) \\
&= -y^{-2}\left(-by^2\sin(\omega t)\right) \\
&= b\sin(\omega t).
\end{aligned}$$

Using this ODE for $z$ we compute

$$\begin{aligned}
\frac{d}{dt}\big(z(t)x_0(t)\big) &= \dot{z}(t)x_0(t) + z(t)\dot{x}_0(t) \\
&= x_0(t)\left[\dot{z}(t) + z(t)(-\epsilon b + \epsilon b\cos(2\omega t))\right] \\
&= x_0(t)\left[\dot{z}(t) - \epsilon b z(t)(1 - \cos(2\omega t))\right] \\
&= x_0(t)b\sin(\omega t).
\end{aligned}$$

Integrating from 0 to $t$ and abbreviating $z(0)x_0(0) =: C$ yields

$$z(t)x_0(t) = C + b\int_0^t x_0(s)\sin(\omega s)ds.$$

The Lemma follows by inserting $z(t) = y(t)^{-1}$. □

## G. Supplementary Details to Section III.A

**Proof of Equation (10)**

We have

$$x_n := x(t+nT) = L + \frac{\theta^n x_0(t)}{\tilde{C} + \theta^n X(t)}.$$

Let $a := x_0(t)$, $b := X(t)$, $c := \tilde{C}$ and $y_n := x_n - L$. We compute

$$\begin{aligned}(y_0 - y_1)y_2 &= \left(\frac{a}{b+c} - \frac{a\theta}{c+b\theta}\right)\frac{a\theta^2}{c+b\theta^2} \\ &= \left(\frac{ac + ab\theta - a\theta b - a\theta c}{(b+c)(c+b\theta)}\right)\frac{a\theta^2}{c+b\theta^2} \\ &= \frac{ac - a\theta c}{(b+c)(c+b\theta)}\frac{a\theta^2}{c+b\theta^2}.\end{aligned}$$

Next, we compute

$$\begin{aligned}\theta(y_2 - y_1)y_0 &= \theta\left(\frac{a\theta^2}{b\theta^2 + c} - \frac{a\theta}{c+b\theta}\right)\frac{a}{c+b} \\ &= \left(\frac{a\theta^2 c + a\theta^3 b - ab\theta^3 - ac\theta}{(c+b\theta^2)(c+b\theta)}\right)\frac{\theta a}{c+b} \\ &= \frac{a\theta^2 c - ac\theta}{(c+b\theta^2)(c+b\theta)}\frac{\theta a}{c+b} \\ &= \frac{a\theta c - ac}{(c+b\theta^2)(c+b\theta)}\frac{\theta^2 a}{c+b}.\end{aligned}$$

Combining both equations we get

$$0 = (y_0 - y_1)y_2 + \theta(y_2 - y_1)y_0.$$

Since $y_n = x_n - L$ we have $y_k - y_l = x_k - x_l$. Consequently

$$\begin{aligned}0 &= (y_0 - y_1)y_2 + \theta(y_2 - y_1)y_0 \\ &= (x_0 - x_1)y_2 + \theta(x_2 - x_1)y_0 \\ &= (x_0 - x_1)x_2 + \theta(x_2 - x_1)x_0 - L\left((x_0 - x_1) + \theta(x_2 - x_1)\right).\end{aligned}$$

Rearranging gives

$$L = \frac{(x_0 - x_1)x_2 + \theta(x_2 - x_1)x_0}{x_0 - (1+\theta)x_1 + \theta x_2}.$$

**Proof of Equation (13)**

We consider the Equation

$$\frac{(x_0 - x_1)x_2 + \theta x_0(x_2 - x_1)}{x_0 - (1+\theta)x_1 + \theta x_2} = \frac{(x_1 - x_2)x_3 + \theta x_1(x_3 - x_2)}{x_1 - (1+\theta)x_2 + \theta x_3} \tag{31}$$

where we recall that $\theta = e^{-\epsilon b T} \in (0,1)$, $x_n = x(t+nT)$ and

$$x_n = x(t+nT) = L + \frac{x_0(t+nT)}{\tilde{C} + X(t+nT)} = L + \frac{\theta^n x_0(t)}{\tilde{C} + \theta^n X(t)}. \tag{32}$$

$X(t)$ is a $T$-periodic function and $x_0(t)$ is as in Equation (30). Note that $x_n \to L$ as $n \to \infty$. This implies that $\tilde{C} \neq 0$ as otherwise

$$x_n = L + \frac{\theta^n x_0(t)}{\theta^n X(t)} = L + \frac{x_0(t)}{X(t)} \not\to L$$

as $x_0(t) > 0$. Using Equation (32), we get for $n \geq 0$ and $k \geq 1$

$$
\begin{aligned}
x_{n+k} - x_n &= \frac{\theta^{n+k} x_0(t)}{\tilde{C} + \theta^{n+k} X(t)} - \frac{\theta^n x_0(t)}{\tilde{C} + \theta^n X(t)} \\
&= \theta^k \frac{\theta^n x_0(t)}{\tilde{C} + \theta^n X(t)} \frac{\tilde{C} + \theta^n X(t)}{\tilde{C} + \theta^{n+k} X(t)} - \frac{\theta^n x_0(t)}{\tilde{C} + \theta^n X(t)} \\
&= \frac{\theta^n x_0(t)}{\tilde{C} + \theta^n X(t)} \left[ \theta^k \frac{\tilde{C} + \theta^n X(t)}{\tilde{C} + \theta^{n+k} X(t)} - 1 \right] \\
&= \frac{\theta^n x_0(t)}{\tilde{C} + \theta^n X(t)} \left[ \frac{\theta^k \tilde{C} + \theta^{n+k} X(t)}{\tilde{C} + \theta^{n+k} X(t)} - \frac{\tilde{C} + \theta^{n+k} X(t)}{\tilde{C} + \theta^{n+k} X(t)} \right] \\
&= \frac{\theta^n x_0(t)}{\tilde{C} + \theta^n X(t)} \frac{\tilde{C}(\theta^k - 1)}{\tilde{C} + \theta^{n+k} X(t)} \\
&\neq 0.
\end{aligned} \tag{33}
$$

We prove:

$$
\theta = \frac{1-g}{2g} - \frac{1}{2g} \sqrt{-4g^2 + (g-1)^2} \tag{34}
$$

*Proof.* We rewrite Equation Equation (31):

$$
\Big( (x_0 - x_1)x_2 + \theta x_0 (x_2 - x_1) \Big) \Big( x_1 - (1+\theta)x_2 + \theta x_3 \Big) = \Big( (x_1 - x_2)x_3 + \theta x_1 (x_3 - x_2) \Big) \Big( x_0 - (1+\theta)x_1 + \theta x_2 \Big)
$$

We further simplify by collecting the terms with and without $\theta$'s in the large parenthesis:

$$
\Big( (x_0 - x_1)x_2 + \theta x_0 (x_2 - x_1) \Big) \Big( x_1 - x_2 + \theta(x_3 - x_2) \Big) = \Big( (x_1 - x_2)x_3 + \theta x_1 (x_3 - x_2) \Big) \Big( x_0 - x_1 \theta(x_2 - x_1) \Big)
$$

Next we expand:

$$
\begin{aligned}
&(x_0 - x_1)x_2(x_1 - x_2) + \theta(x_0(x_2 - x_1)(x_1 - x_2) + (x_0 - x_1)x_2(x_3 - x_2)) + \theta^2 x_0 (x_2 - x_1)(x_3 - x_2) \\
=&(x_1 - x_2)x_3(x_0 - x_1) + \theta(x_1(x_3 - x_2)(x_0 - x_1) + (x_1 - x_2)x_3(x_2 - x_1)) + \theta^2 x_1 (x_3 - x_2)(x_2 - x_1)
\end{aligned}
$$

We subtract all terms in the second line and get:

$$
\begin{aligned}
0 =&(x_0 - x_1)(x_1 - x_2)(x_2 - x_3) \\
&+ \theta \left[ -(x_0 - x_3)(x_1 - x_2)^2 + (x_0 - x_1)(x_3 - x_2)(x_2 - x_1) \right] \\
&+ \theta^2 (x_0 - x_1)(x_1 - x_2)(x_2 - x_3)
\end{aligned}
$$

Considering Equation (33), we may divide by $(x_1 - x_2)^2(x_0 - x_3)$ and get

$$
\frac{(x_0 - x_1)(x_2 - x_3)}{(x_1 - x_2)(x_0 - x_3)}(1 + \theta + \theta^2) - \theta = 0.
$$

We define

$$
g := \frac{(x_0 - x_1)(x_2 - x_3)}{(x_1 - x_2)(x_0 - x_3)}.
$$

so that $g\theta^2 + (g-1)\theta + g = 0$. Hence

$$
\theta = \frac{1-g}{2} \pm \sqrt{\left( \frac{1-g}{2g} \right)^2 - 1} \overset{!}{=} \frac{1-g}{2} \pm \frac{1}{2g} \sqrt{(1-g)^2 - 4g^2}.
$$

In the last step (marked by !) we have pulled out $(2g)^{-2}$ out of the square root and written the factor $(2g)^{-1}$ in front of it. Really, we have to write $|2g|^{-1}$. However, we can absorb the potential sign difference in the still ambiguous $\pm$. Only now we determine the correct sign. To prove that is is '$-$', we use $\theta = e^{-\epsilon b T} \in (0,1)$. In particular, $\theta \in \mathbb{R}$ and so $(g-1)^2 \geq 4g^2$, which implies $-1 \leq g \leq \frac{1}{3}$ . Additionally, due to Equation (33), we deduce $g \neq 0$. Now we distinguish two cases.

1) For $g \in [-1,0)$ we have $\frac{1-g}{2g} < 0$. As $\theta \in (0,1)$ we deduce

$$\pm \frac{1}{2g}\sqrt{-4g^2+(g-1)^2} \overset{!}{\geq} 0.$$

This implies that $-$ is the correct sign.

2) For $g \in (0, \frac{1}{3})$ we note that $\frac{1-g}{2g} \geq 1$. As $\theta \in (0,1)$ we deduce

$$\pm \frac{1}{2g}\sqrt{-4g^2+(g-1)^2} \overset{!}{\leq} 0.$$

Again, this implies that $-$ is the correct sign.

$\square$

## H. Supplementary Details to Section III.B

For parameters $b, q_0, \omega \in \mathbb{R}$, $\delta, \epsilon > 0$ and we consider the Equation

$$\dot{z} - 2\epsilon \sin^2(\omega t) z + \delta q(t) z^2 = \sin(\omega t). \tag{35}$$

where $q(t) = q_0 e^{-\delta t}$. We treat $\delta$ as a perturbative parameter and propose the ansatz

$$z(t) = \sum_{n=0}^{\infty} z_n(t) \delta^n \tag{36}$$

with initial values $z_0(0) = z(0)$ and $z_n(0) = 0$ for all $n \geq 1$.

### Proof of Equations (17) and (18)

We claim that the following Equations follow:

$$\begin{cases} \dot{z}_0 - 2\epsilon \sin^2(\omega t) z_0 = \sin(\omega t), \\ z_0(0) = z(0). \end{cases} \tag{37}$$

For $n \geq 1$:

$$\begin{cases} \dot{z}_n - 2\epsilon \sin^2(\omega t) z_n = -q(t) \sum_{j=0}^{n-1} z_j z_{n-1-j}, \\ z_n(0) = 0. \end{cases} \tag{38}$$

*Proof.* Inserting the ansatz into Equation (35) gives

$$\sum_{n=0}^{\infty} \left(\dot{z}_n - 2 z_n \epsilon \sin^2(\omega t)\right) \delta^n + \delta q(t) \left(\sum_{l=0}^{\infty} z_l \delta^l\right)\left(\sum_{k=0}^{\infty} z_k \delta^k\right) = \sin(\omega t).$$

We write

$$\delta q(t) \left(\sum_{l=0}^{\infty} z_l \delta^l\right)\left(\sum_{k=0}^{\infty} z_k \delta^k\right) = q(t)\delta \sum_{n=0}^{\infty}\left(\delta^n \sum_{a=0}^{n} z_a z_{n-a}\right) = q(t) \sum_{n=0}^{\infty}\left(\delta^{n+1} \sum_{a=0}^{n} z_a z_{n-a}\right).$$

Inserting gives

$$\sum_{n=0}^{\infty} \left(\dot{z}_n - 2 z_n \epsilon \sin^2(\omega t)\right) \delta^n = \sin(\omega t) - q(t) \sum_{n=0}^{\infty}\left(\delta^{n+1} \sum_{a=0}^{n} z_a z_{n-a}\right).$$

Comparing coefficients gives Equations (37) and (38). $\square$

**Lemma H.1.** *There exist $T$-periodic functions $p_0^{(0)}$ and $p_1^{(0)}$ such that*

$$z_0(t) = p_0^{(0)}(t) + e^{\epsilon t} p_1^{(0)}(t).$$

*Further, for $n \geq 1$, $1 \leq j \leq n$ and $0 \leq k \leq j+1$ there exist $T$-periodic functions $p_0^{(n)}$ and $p_{jk}^{(n)}$ such that*

$$z_n(t) = e^{\epsilon t} p_0^{(n)}(t) + \sum_{j=1}^{n} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(n)}(t).$$

*Proof.* For $n=0$ we can apply Lemma F.2 with $y \to z_0$, $a \to -\epsilon$, $\eta \to 0$ and $q(t) \to \sin(\omega t)$ to obtain the claimed formula. For $n \geq 1$ we argue by induction. First we consider $n=1$. We have

$$\dot{z}_1 - 2\epsilon \sin^2(\omega t) z_1 = -q(t) z_0^2.$$

We insert $z_0(t) = p_0^{(0)}(t) + e^{\epsilon t} p_1^{(0)}(t)$ and use $q(t) = q_0 e^{-\delta t}$ to get

$$\dot{z}_1 - 2\epsilon \sin^2(\omega t) z_1 = -q_0 e^{-\delta t} \left[ \left(p_0^{(0)}(t)\right)^2 + e^{2\epsilon t} \left(p_1^{(0)}(t)\right)^2 + 2 e^{\epsilon t} p_0^{(0)}(t) p_1^{(0)}(t) \right].$$

The general solution to this Equation is given by

$$z_1(t) = z_{1,h}(t) + \sum_{k=0}^{2} z_{1,k}(t). \tag{39}$$

Here $z_{1,h}$ denotes a homogeneous solution and for $k=0,1,2$ the function $z_{1,k}$ is any solution of

$$\dot{z}_{1,k} - 2\epsilon \sin^2(\omega t) z_{1,k} = e^{-\delta t} e^{k\epsilon t} P_{1,k}(t)$$

where

$$P_{1,0} := -q_0 \left(p_0^{(0)}\right)^2, \qquad P_{1,1} = -2 q_0 p_0^{(0)} p_1^{(0)} \qquad \text{and} \qquad P_{1,2} := -q_0 \left(p_1^{(0)}\right)^2.$$

We can apply Lemma F.2 to obtain $T$-periodic functions $\pi_*$ ($*$ denotes arbitrary indices) such that

$$\begin{aligned} z_h(t) &= e^{\epsilon t} \pi_h(t), \\ z_k(t) &= e^{\epsilon t} \pi_{k,1}(t) + e^{k\epsilon t - \delta t} \pi_{k,2}(t). \end{aligned}$$

Using Equation (39), we get the claimed formula for $z_1$.

Now we consider the inductive step $n \to n+1$. Assume that the formulas for $z_k$ with $0 \leq k \leq n$ are already proven. We have to compute

$$\sum_{a=0}^{n} z_a z_{n-a} = 2 z_0 z_n + \sum_{a=1}^{n-1} z_a z_{n-a}.$$

For $1 \leq a \leq n-1$ we have the formulas

$$\begin{aligned} z_a(t) &= e^{\epsilon t} p_0^{(a)}(t) + \sum_{j=1}^{a} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(a)}(t) \\ z_{n-a}(t) &= e^{\epsilon t} p_0^{(n-a)}(t) + \sum_{j=1}^{n-a} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(n-a)}(t). \end{aligned}$$

Multiplying gives

$$\begin{aligned} z_a z_{n-a} = {} & e^{2\epsilon t} p_0^{(a)} p_0^{(n-a)} \\ & + e^{\epsilon t} p_0^{(a)}(t) \sum_{j=1}^{n-a} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(n-a)} \\ & + e^{\epsilon t} p_0^{(n-a)} \sum_{j=1}^{a} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(a)} \\ & + \left( \sum_{j=1}^{a} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(a)} \right) \left( \sum_{j=1}^{n-a} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(n-a)} \right) \end{aligned}$$

This sum is a linear combination of $e^{-\kappa\delta t}$ with $\kappa = 0, \ldots, n$. The coefficients of $e^{-\kappa\delta t}$ are linear combinations of functions of the form $e^{j\epsilon t} p(t)$ where $p$ stands for a general $T$-periodic function and $0 \leq j \leq \kappa + 2$. Consequently

$$\sum_{a=1}^{n-1} z_a z_{n-a} = \sum_{\kappa=0}^{n} \left[ e^{-\kappa\delta t} \sum_{j=0}^{\kappa+2} \pi_{\kappa j}(t) e^{j\epsilon t} \right]$$

where $\pi_{\kappa j}$ are some $T$-periodic functions. Now, we compute

$$\begin{aligned} z_0 z_n =&(p_0^{(0)} + e^{\epsilon t} p_1^{(0)}) \left( e^{\epsilon t} p_0^{(n)}(t) + \sum_{j=1}^{n} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(n)}(t) \right) \\ =& e^{\epsilon t} p_0^{(0)} p_0^{(n)}(t) + \sum_{j=1}^{n} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_0^{(0)} p_{jk}^{(n)}(t) \\ &+ e^{2\epsilon t} p_1^{(0)} p_0^{(n)}(t) + \sum_{j=1}^{n} \sum_{k=0}^{j+1} e^{((k+1)\epsilon - j\delta)t} p_1^{(0)} p_{jk}^{(n)}(t). \end{aligned}$$

Again, this is a linear combination of $e^{-\kappa\delta t}$ where $0 \le \kappa \le n$. The coefficients of $e^{-\kappa\delta t}$ are again linear combinations of functions $e^{j\epsilon t} p(t)$ where $0 \le j \le \kappa + 2$ and $p$ stands for a general $T$-periodic function. Therefore we have shown that

$$e^{-\delta t} \sum_{a=0}^{n} z_a z_{n-a} = \sum_{\kappa=0}^{n} e^{-(\kappa+1)\delta t} \sum_{j=0}^{\kappa+2} \tilde{\pi}_{\kappa j}(t) e^{j\epsilon t} = \sum_{\kappa=1}^{n+1} \left[ e^{-\kappa\delta t} \sum_{j=0}^{\kappa+1} \pi_{\kappa j}(t) e^{j\epsilon t} \right]$$

for some, potentially new. $T$-periodic functions $\pi_{\kappa j}$. Absorbing $-q_0$ into the definition of the functions $\pi_{\kappa j}$ we get

$$\dot{z}_{n+1} + 2\epsilon \sin^2(\omega t) = -q(t) \sum_{a=0}^{n} z_a z_{n-a} = \sum_{\kappa=1}^{n+1} \left[ e^{-\kappa\delta t} \sum_{j=0}^{\kappa+1} \pi_{\kappa j}(t) e^{j\epsilon t} \right]. \tag{40}$$

We now argue as we did for $n = 1$ and write

$$z_{n+1} = z_{n+1,h} + \sum_{\kappa=1}^{n+1} \sum_{j=0}^{\kappa+1} z_{\kappa,j}(t)$$

where $z_{n+1,h}$ is a homogeneous solution to Equation (40) and for $1 \le \kappa \le n+1$ and $0 \le j \le \kappa + 1$

$$\dot{z}_{\kappa,j} + 2\epsilon \sin^2(\omega t) z_{\kappa,j} = e^{-\kappa\delta t} e^{j\epsilon t} \pi_{\kappa j}.$$

Using Lemma F.2 and resumming we deduce that there exist periodic functions $p_0^{(n+1)}$ and $p_{jk}^{(n+1)}$ such that

$$z_{n+1}(t) = e^{\epsilon t} p_0^{(n+1)}(t) + \sum_{j=1}^{n+1} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(n+1)}(t).$$

This finishes the inductive argument. $\square$

The last step in the proof of Proposition III.2 is to resum the perturbative series. We compute

$$\begin{aligned} z(t) =& \sum_{n=0}^{\infty} z_n(t) \delta^n \\ =& p_0^{(0)} + e^{\epsilon t} p_1^{(0)} + \sum_{n=1}^{\infty} \delta^n \left[ e^{\epsilon t} p_0^{(n)}(t) + \sum_{j=1}^{n} \sum_{k=0}^{j+1} e^{(k\epsilon - j\delta)t} p_{jk}^{(n)}(t) \right] \\ =& p_0^{(0)} + e^{\epsilon t} \left[ p_1^{(0)} + \delta \sum_{n=1}^{\infty} p_0^{(n)} \right] + \sum_{n=1}^{\infty} \sum_{j=1}^{n} \sum_{k=0}^{j+1} \delta^n e^{(k\epsilon - j\delta)t} p_{jk}^{(n+1)}(t) \\ =& p_0^{(0)} + e^{\epsilon t} \left[ p_1^{(0)} + \delta \sum_{n=1}^{\infty} p_0^{(n)} \right] + \sum_{j=1}^{\infty} \sum_{n=j}^{\infty} \sum_{k=0}^{j+1} \delta^n e^{(k\epsilon - j\delta)t} p_{jk}^{(n+1)}(t) \\ =& p_0^{(0)} + e^{\epsilon t} \left[ p_1^{(0)} + \delta \sum_{n=1}^{\infty} p_0^{(n)} \right] + \sum_{j=1}^{\infty} \left[ \delta^j e^{-j\delta t} \sum_{k=0}^{j+1} \left( e^{k\epsilon t} \sum_{n=j}^{\infty} \delta^{n-j} p_{jk}^{(n+1)}(t) \right) \right]. \end{aligned}$$

For $j \ge 1$ and $0 \le k \le j+1$ we define

$$p_{jk}(t) := \sum_{n=j}^{\infty} \delta^{n-j} p_{jk}^{(n+1)}(t).$$

Note that $n-j \geq 0$, so $p_{jk} = \mathcal{O}(1)$ with respect to $\delta$. In particular, $p_{jk}$ do not blow up when $\delta \to 0$. Additionally, we put

$$p_{00} := p_0^{(0)} \qquad \text{and} \qquad p_{01} := p_1^{(0)} + \sum_{n=1}^{\infty} p_0^{(n)}.$$

Clearly, for all $0 \leq j$ and $0 \leq k \leq j+1$, the functions $p_{jk}$ are $T$-periodic. Also, we get

$$\begin{aligned} z(t) =& p_{00} + e^{\epsilon t} p_{01} + \sum_{j=1}^{\infty} \left[ \delta^j e^{-j\delta t} \sum_{k=0}^{j+1} e^{k\epsilon t} p_{jk}(t) \right] \\ =& \sum_{j=0}^{\infty} \left[ \delta^j e^{-j\delta t} \sum_{k=0}^{j+1} e^{k\epsilon t} p_{jk}(t) \right]. \end{aligned}$$

**Proof of Equation (20)**

**Lemma H.2.** *Let $\mu > 0$ and $f \in C^0(\mathbb{R})$ be positive. Then, for all $t \in [0, \frac{1}{2\mu}]$*

$$\int_0^t e^{-\mu s} f(s) ds \leq 2e^{-\mu t} \int_0^t f(s) ds.$$

*Proof.* Let $F(t) := \int_0^t e^{-\mu s} f(s) ds$. As $F$ is increasing and $F(0) = 0$, we may estimate

$$\begin{aligned} \int_0^t e^{-\mu s} f(s) ds &= \int_0^t e^{-\mu s} F'(s) ds \\ &= e^{-\mu s} F(s) \Big|_{s=0}^{s=t} - \int_0^t (-\mu) e^{-\mu s} F(s) ds \\ &= e^{-\mu t} F(t) - F(0) + \mu \int_0^t \underbrace{e^{-\mu s}}_{0 \leq \ldots \leq 1} \underbrace{F(s)}_{\geq 0} ds \\ &\leq e^{-\mu t} F(t) + \mu t \sup_{0 \leq s \leq t} F(s) \\ &\overset{!}{=} (e^{-\mu t} + \mu t) F(t) \\ &= (e^{-\mu t} + \mu t) \int_0^t f(s) ds. \end{aligned}$$

In the second to last step (marked by !) we have used that $F$ is increasing. For $x \in [0, \frac{1}{2}]$ we have $x \leq e^{-x}$. Indeed, $xe^x$ is increasing on $[0, \infty)$ and $\frac{1}{2} e^{\frac{1}{2}} = \frac{1}{2}\sqrt{e} \leq \frac{1}{2}\sqrt{4} = 1$. For $t \in [0, \frac{1}{2\mu}]$ we have $\mu t \in [0, \frac{1}{2}]$ and the Lemma follows by estimating $\mu t \leq e^{-\mu t}$. ☐

**Lemma H.3.** *Let $\epsilon > 0$, $\omega \in \mathbb{R}$, $R \in C^0(\mathbb{R})$ and $\xi$ solve $\dot{\xi}(t) - 2\epsilon \sin^2(\omega t)\xi(t) = R(t)$. Then, for $t \in [0, \frac{1}{2\epsilon}]$*

$$|\xi(t)| \leq |\xi(0)| e^{\frac{\epsilon}{2\omega}} e^{\epsilon t} + 2e^{\frac{\epsilon}{\omega}} \int_0^t |R(s)| ds.$$

*Proof.* Using $2\sin^2 x = 1 - \cos(2x)$, we compute

$$\begin{aligned} \frac{d}{dt}\left[\xi(t) e^{-\epsilon t + \frac{\epsilon \sin(2\omega t)}{2\omega}}\right] &= e^{-\epsilon t + \frac{\epsilon \sin(2\omega t)}{2\omega}} \left[\dot{\xi}(t) + \xi(t) \frac{d}{dt}\left(-\epsilon t + \frac{\epsilon \sin(2\omega t)}{2\omega}\right)\right] \\ &= e^{-\epsilon t + \frac{\epsilon \sin(2\omega t)}{2\omega}} \left[\dot{\xi}(t) + \xi(t)\left(-\epsilon + \epsilon \cos(2\omega t)\right)\right] \\ &= e^{-\epsilon t + \frac{\epsilon \sin(2\omega t)}{2\omega}} \left[\dot{\xi}(t) - 2\epsilon \sin^2(\omega t)\xi(t)\right] \\ &= e^{-\epsilon t + \frac{\epsilon \sin(2\omega t)}{2\omega}} R(t). \end{aligned} \tag{41}$$

Integrating gives

$$\xi(t) e^{-\epsilon t + \frac{\epsilon \sin(2\omega t)}{2\omega}} - \xi(0) = \int_0^t e^{-\epsilon s + \frac{\epsilon \sin(2\omega s)}{2\omega}} R(s) ds.$$

We now estimate

$$
\begin{aligned}
|\xi(t)| &\le e^{\epsilon t - \frac{\epsilon \sin(2\omega t)}{2\omega}}|\xi(0)| + e^{\epsilon t - \frac{\epsilon \sin(2\omega t)}{2\omega}} \int_0^t e^{-\epsilon s + \frac{\epsilon \sin(2\omega s)}{2\omega}} |R(s)| ds \\
&\le e^{\epsilon t + \frac{\epsilon}{2\omega}}|\xi(0)| + e^{\epsilon t + \frac{\epsilon}{2\omega}} \int_0^t e^{-\epsilon s + \frac{\epsilon}{2\omega}} |R(s)| ds \\
&\le e^{\epsilon t + \frac{\epsilon}{2\omega}}|\xi(0)| + e^{\epsilon t + \frac{\epsilon}{\omega}} \int_0^t e^{-\epsilon s} |R(s)| ds.
\end{aligned}
$$

To estimate further we use Lemma H.2 to estimate

$$
\int_0^t e^{-\epsilon s}|R(s)|ds \le 2e^{-\epsilon t}\int_0^t |R(s)|ds \qquad \text{for } t \in [0, \frac{1}{2\epsilon}].
$$

Inserting this Estimate gives

$$
|\xi(t)| \le |\xi(0)|e^{\frac{\epsilon}{2\omega}}e^{\epsilon t} + 2e^{\frac{\epsilon}{\omega}}e^{\epsilon t}e^{-\epsilon t}\int_0^t |R(s)|ds = |\xi(0)|e^{\frac{\epsilon}{2\omega}}e^{\epsilon t} + 2e^{\frac{\epsilon}{\omega}}\int_0^t |R(s)|ds.
$$

□

We now prove a criterion that ensures the convergence of the perturbation series in Equation (36).

**Lemma H.4.** *If $\delta > \epsilon$ and*

$$
24e^{\frac{2\epsilon}{\omega}}|q_0|\left(|z(0)| + \frac{1}{\delta}\right) < 1,
$$

*the series in Equation (36) in convergent.*

*Proof.* Put $t_0 := \frac{1}{2\delta}$ and $u_k := \sup_{0\le s\le t_0}|z_k(s)|$. Applying Lemma H.3 to Equation (37) gives

$$
u_0 \le |z(0)|e^{\frac{\epsilon}{2\omega}+\epsilon t_0} + 2e^{\frac{\epsilon}{\omega}}t_0 =: \alpha_0. \tag{42}
$$

For $n \ge 0$, applying Lemma H.3 to Equation (38) and subsequently using Lemma H.2 gives

$$
\begin{aligned}
u_{n+1} &\le 2e^{\frac{\epsilon}{\omega}}|q_0|\sum_{j=0}^n \int_0^{t_0} e^{-\delta s}|z_j(s)z_{n-j}(s)|ds \\
&\le 4e^{\frac{\epsilon}{\omega}}|q_0|e^{-\delta t_0}\sum_{j=0}^n \int_0^{t_0} |z_j(s)z_{n-j}(s)|ds \\
&\le 4e^{\frac{\epsilon}{\omega}}|q_0|e^{-\delta t_0}t_0\sum_{j=0}^n u_j u_{n-j}.
\end{aligned} \tag{43}
$$

We use $xe^{-x} \le 1$ for all $x \ge 0$ to estimate $t_0 e^{-\delta t_0} \le \delta^{-1}$ and put $C := 4(e\delta)^{-1}e^{\frac{\epsilon}{\omega}}|q_0|$. For $n > 0$, we define $\alpha_n$ by

$$
\alpha_{n+1} = C\sum_{j=0}^n \alpha_j\alpha_{n-j}. \tag{44}
$$

We claim $u_n \le \alpha_n$ for all $n \ge 0$. For $n = 0$ this is true by definition and for $n \ge 1$ it follows inductively. Indeed, assuming $u_k \le \alpha_k$ for all $0 \le k \le n$ we estimate

$$
u_{n+1} \overset{(43)}{\le} C\sum_{j=0}^n u_j u_{n-j} \le C\sum_{j=0}^n \alpha_j\alpha_{n-j} \overset{(44)}{=} \alpha_{n+1}.
$$

Therefore $\sum_{n\ge 0} z_n\delta^n$ converges absolutely when $\sum_{n\ge 0}\alpha_n\delta^n$ converges. To derive a criterion for the convergence of

$\sum_{n\geq 0}\alpha_n\delta^n$ we consider the generating function $A(x) := \sum_{j\geq 0}\alpha_j x^j$. Using (44), we compute

$$\begin{aligned}
CxA^2(x) &= \sum_{n=0}^{\infty} Cx\left(\sum_{k=0}^{\infty}\alpha_k x^k\right)\left(\sum_{l=0}^{\infty}\alpha_l x^l\right) \\
&= Cx\sum_{n=0}^{\infty}\left[x^n\sum_{k=0}^{n}\alpha_k\alpha_{n-k}\right] \\
&= \sum_{n=0}^{\infty}\left[x^{n+1}\left(C\sum_{k=0}^{n}\alpha_k\alpha_{n-k}\right)\right] \\
&= \sum_{n=0}^{\infty}x^{n+1}\alpha_{n+1} \\
&= A(x)-\alpha_0.
\end{aligned}$$

This shows that $A(x)$ satisfies the quadratic equation $CxA(x)^2 - A(x) + \alpha_0 = 0$. Therefore

$$A(x) = \frac{1\pm\sqrt{1-4C\alpha_0 x}}{2Cx}.$$

The correct sign is $-$. Indeed, assume $+$ was correct. Then we get a contradiction as

$$\alpha_0 = A(0) = \lim_{x\to 0}A(x) \overset{!}{=} \lim_{x\to 0}\frac{1+\sqrt{1-4C\alpha_0 x}}{2Cx}\lim_{x\to 0}\frac{2}{2Cx}\qquad \text{, which is divergent.}$$

So

$$A(x) = \frac{1-\sqrt{1-4C\alpha_0 x}}{2Cx}. \tag{45}$$

We use the expansion

$$\sqrt{1-\epsilon} = (1-\epsilon)^{\frac{1}{2}} = \sum_{n=0}^{\infty}\binom{\frac{1}{2}}{n}(-\epsilon)^n$$

to expand

$$\begin{aligned}
A(x) &= \frac{1}{2Cx}\left[1-\sum_{n=0}^{\infty}\binom{\frac{1}{2}}{n}(-1)^n(4C\alpha_0 x)^n\right] \\
&= -\frac{1}{2Cx}\sum_{n=1}^{\infty}\binom{\frac{1}{2}}{n}(-1)^n(4C\alpha_0 x)^n \\
&= \sum_{n=0}^{\infty}-\frac{1}{2C}\binom{\frac{1}{2}}{n+1}(-4C\alpha_0)^{n+1}x^n.
\end{aligned}$$

By definition $A(x) = \sum_{n\geq 0}\alpha_n x^n$. Comparing coefficients, we get

$$\alpha_n = \frac{-1}{2C}\binom{\frac{1}{2}}{n+1}(-4C\alpha_0)^{n+1}.$$

We use Stirling's approximation to get an asymptotic expansion of $\alpha_n$:

$$\begin{aligned}
\binom{\frac{1}{2}}{n} &= \binom{2n}{n}\frac{(-1)^{n+1}}{4^n(2n-1)} \\
&= \frac{(2n)!}{(n!)^2}\frac{(-1)^{n+1}}{4^n(2n-1)} \\
&\sim \frac{(2n)^{2n)}}{e^{2n}}\sqrt{4\pi n}\frac{(e^n)^2}{(n^n)^2(\sqrt{2\pi n}^2}\frac{(-1)^{n+1}}{4^n(2n)} \\
&= \frac{4^n}{\sqrt{\pi n}}\frac{(-1)^{n+1}}{4^n(2n)} \\
&= \frac{(-1)^{n+1}}{2\sqrt{\pi}}\frac{1}{n^{\frac{3}{2}}}
\end{aligned}$$

Using this asymptotic formula we get

$$\alpha_n \sim \frac{-1}{2C}\frac{(-1)^n}{2\sqrt{\pi}}\frac{1}{\sqrt{(n+1)^{\frac{3}{2}}}}(-1)^{n+1}(4C\alpha_0)^{n+1} = \frac{1}{4C\sqrt{\pi}}\frac{(4C\alpha_0)^{n+1}}{(n+1)^{\frac{3}{2}}}.$$

To ensure convergence of $\sum_{n\geq 0}\alpha_n\delta^n$ we must require

$$1 > \lim_{n\to\infty}\frac{\alpha_{n+1}\delta^{n+1}}{\alpha_n\delta^n} = \delta\lim_{n\to\infty}\left[\frac{1}{4C\sqrt{\pi}}\frac{(4C\alpha_0)^{n+2}}{(n+2)^{\frac{3}{2}}}\left(\frac{1}{4C\sqrt{\pi}}\frac{(4C\alpha_0)^{n+1}}{(n+1)^{\frac{3}{2}}}\right)^{-1}\right] = (4C\alpha_0)\delta.$$

Inserting the definition of $C$ we get

$$1 \overset{!}{>} 4C\alpha_0\delta 4\left(4(\delta e)^{-1}e^{\frac{\epsilon}{\omega}}|q_0|\right)\delta = 16e^{-1}\alpha_0 e^{\frac{\epsilon}{\omega}}|q_0|.$$

By definition $\alpha_0 = u_0$. Hence

$$16e^{-1}\alpha_0 e^{\frac{\epsilon}{\omega}}|q_0| \leq 16e^{-1}e^{\frac{\epsilon}{\omega}}|q_0|\left(|z(0)|e^{\frac{\epsilon}{2\omega}+\epsilon t_0} + 2e^{\frac{\epsilon}{\omega}}t_0\right).$$

We have $t_0 = \frac{1}{2\delta} \leq \frac{1}{2\epsilon}$. The last step is justified by requiring $\epsilon < \delta$. Hence

$$\begin{aligned}
4C\alpha_0\delta &\leq 16e^{-1}\alpha_0 e^{\frac{\epsilon}{\omega}}|q_0| \\
&\leq 16e^{-1}e^{\frac{\epsilon}{\omega}}|q_0|\left(|z(0)|e^{\frac{\epsilon}{2\omega}+\epsilon t_0} + 2e^{\frac{\epsilon}{\omega}}t_0\right) \\
&\leq 16e^{-1}e^{\frac{\epsilon}{\omega}}|q_0|\left(|z(0)|e^{\frac{\epsilon}{2\omega}+\frac{1}{2}} + \frac{4}{\delta}e^{\frac{\epsilon}{\omega}}\right) \\
&\leq 16e^{-1}e^{\frac{\epsilon}{\omega}}|q_0|\left(|z(0)|e^{\frac{\epsilon}{2\omega}}\cdot 4 + \frac{4}{\delta}e^{\frac{\epsilon}{\omega}}\right) \\
&\leq 16e^{-1}e^{\frac{\epsilon}{\omega}}|q_0|\left(4e^{\frac{\epsilon}{\omega}}\right)\left(|z(0)| + \frac{1}{\delta}e^{\frac{\epsilon}{\omega}}\right) \\
&\leq 64e^{-1}e^{2\frac{\epsilon}{\omega}}|q_0|\left(|z(0)| + \frac{1}{\delta}\right) \\
&\leq 24e^{2\frac{\epsilon}{\omega}}|q_0|\left(|z(0)| + \frac{1}{\delta}\right).
\end{aligned}$$

In the last step we have used that $64e^{-1} = 23.544... \leq 24$. So, if

$$24e^{2\frac{\epsilon}{\omega}}|q_0|\left(|z(0)| + \frac{1}{\delta}\right) < 1,$$

the series $\sum_{n\geq 0}\alpha_n\delta^n$ converges and hence $\sum_{n\geq 0}z_n(t)\delta^n$ converges absolutely. □

**Detailed proof of Corollary III.3**
**Step 1: A Zeroth Order Extraction Law**
Including only the first term of the perturbation series gives

$$z(t) = p_0(t) + e^{-\epsilon t}p_1(t)$$

where $p_0$ and $p_1$ are $T$-periodic. We fix an arbitrary $t$ and put $p_0 := p_0(t)$, $p_1 := p_1(t)$ as well as $z_n := z(t+nT)$. Then

$$z_n = p_0 + A^n p_1.$$

We consider the following two equations:

$$\begin{aligned}
z_{n+1} - z_n &= A^n(A-1)p_1 \\
z_{n+2} - z_{n+1} &= A^{n+1}(A-1)p_1.
\end{aligned}$$

Subtracting $A$ times the first equation from the second gives

$$z_{n+2} - (1+A)z_{n+1} + Az_n = 0.$$

Recalling that $y = 1/z$, puttning $y_n := y(t+nT)$ and multiplying by $y_n y_{n+1} y_{n+2}$, we get

$$y_n y_{n+1} - (1+A)y_n y_{n+2} + Ay_{n+1}y_{n+2} = 0.$$

By definition $y = x - L - q = h - L$. Putting $h_n := h(t + nT)$, we get

$$\begin{aligned}
0 =& h_{n+2}h_{n+1} - (1+A)h_n h_{n+2} + Ah_{n+2}h_{n+2} \\
&+ L^2 - (1+A)L^2 + AL^2 \\
&- h_n L - Lh_{n+1} + (1+A)(Lh_{n+2} + h_n L) - A(h_{n+1}L + Lh_{n+2}) \\
=& h_{n+2}h_{n+1} - (1+A)h_n h_{n+2} + Ah_{n+2}h_{n+2} \\
&+ L(-h_n - h_{n+1} + h_{n+2} + h_n + Ah_{n+2} + Ah_n - Ah_{n+1} - Ah_{n+2}) \\
=& h_{n+2}h_{n+1} - (1+A)h_n h_{n+2} + Ah_{n+2}h_{n+2} \\
&+ L(-h_{n+1} + h_{n+2} + Ah_n - Ah_{n+1}).
\end{aligned}$$

Rearranging, we get

$$L = \frac{h_{n+2}h_{n+1} - (1+A)h_n h_{n+2} + Ah_{n+2}h_{n+2}}{-h_{n+2} + (1+A)h_{n+1} - Ah_n}.$$

**Step 2: A First Order Extraction Law**

Including the first two terms of the perturbation series gives

$$z(t) = p_0(t) + e^{\epsilon t}p_1(t) + e^{-\delta t}(p_2(t) + e^{\epsilon t}p_3(t) + e^{2\epsilon t}p_4(t))$$

where $p_k$ are $T$-periodic functions. Putting $z_n := z(t + nT)$ and $p_k := p_i(t)$ for $0 \leq k \leq 4$, we get

$$z_n = p_0 + A^n p_1 + B^n(p_2 + A^n p_3 + A^{2n}p_4).$$

We now methodically combine these equations for various $n$ to get an identity with right hand side 0. We begin by computing

$$z_{n+1} - z_n = A^n p_1(A-1) + B^n(B-1)p_2 + B^n A^n(BA-1)p_3 + B^n A^{2n}(BA^2-1)p_4.$$

Hence

$$\begin{aligned}
& z_{n+2} - z_{n+1} - A(z_{n+1} - z_n) \\
=& A^{n+1}p_1(A-1) + B^{n+1}(B-1)p_2 + B^{n+1}A^{n+1}(BA-1)p_3 + B^{n+1}A^{2(n+1)}(BA^2-1)p_4 \\
&- \left(A^{n+1}p_1(A-1) + AB^n(B-1)p_2 + AB^n A^n(BA-1)p_3 + AB^n A^{2n}(BA^2-1)p_4\right) \\
=& B^n(B-1)(B-A)p_2 + B^n A^{n+1}(B-1)(BA-1)p_3 + B^n A^{2n+1}(BA-1)(BA^2-1)p_4.
\end{aligned}$$

Next we compute

$$\begin{aligned}
& z_{n+3} - z_{n+2} - A(z_{n+2} - z_{n+1}) - B\left(z_{n+2} - z_{n+1} - A(z_{n+1} - z_n)\right) \\
=& B^{n+1}(B-1)(B-A)p_2 + B^{n+1}A^{n+2}(B-1)(BA-1)p_3 + B^{n+1}A^{2n+3}(BA-1)(BA^2-1)p_4 \\
&- \left(BB^n(B-1)(B-A)p_2 + BB^n A^{n+1}(B-1)(BA-1)p_3 + BB^n A^{2n+1}(BA-1)(BA^2-1)p_4\right) \\
=& B^{n+1}A^{n+1}(A-1)(B-1)(BA-1)p_3 + B^{n+1}A^{2n+1}(A^2-1)/BA-1)(BA^2-1)p_4.
\end{aligned}$$

We simplify

$$\begin{aligned}
& z_{n+3} - z_{n+2} - A(z_{n+2} - z_{n+1}) - B\left(z_{n+2} - z_{n+1} - A(z_{n+1} - z_n)\right) \\
=& z_{n+3} - (1+A+B)z_{n+2} + z_{n+1}(A+B+AB) - ABz_n.
\end{aligned}$$

So we get

$$\begin{aligned}
& z_{n+3} - (1+A+B)z_{n+2} + z_{n+1}(A+B+AB) - ABz_n \\
=& B^{n+1}A^{n+1}(A-1)(B-1)(BA-1)p_3 + B^{n+1}A^{2n+1}(A^2-1)(BA-1)(BA^2-1)p_4.
\end{aligned}$$

Now we compute

$$\begin{aligned}
& z_{n+4} - (1+A+B)z_{n+3} + z_{n+2}(A+B+AB) - ABz_{n+1} \\
&- AB\left(z_{n+3} - (1+A+B)z_{n+2} + z_{n+1}(A+B+AB) - ABz_n\right) \\
=& B^{n+2}A^{2n+1}(A^2-1)(A^2-1)(BA-1)(BA^2-1)p_4.
\end{aligned}$$

So, we arrive at the identity

$$
\begin{aligned}
& z_{n+5}-(1+A+B)z_{n+4}+z_{n+3}(A+B+AB)-ABz_{n+2} \\
& -AB\left(z_{n+4}-(1+A+B)z_{n+3}+z_{n+2}(A+B+AB)-ABz_{n+1}\right) \\
& -A^2B\left[z_{n+4}-(1+A+B)z_{n+3}+z_{n+2}(A+B+AB)-ABz_{n+1}\right. \\
& \left.-AB\left(z_{n+3}-(1+A+B)z_{n+2}+z_{n+1}(A+B+AB)-ABz_n\right)\right] \\
& =0.
\end{aligned}
$$

We collect terms

$$
\begin{aligned}
0 = & + z_{n+5} \\
& - z_{n+4}(1+A+B+AB+A^2B) \\
& = +z_{n+3}(A+B+AB+AB(1+A+B)+A^2B(1+A+B)+A^3B^2) \\
& = -z_{n+2}(AB+AB(A+B+AB)+A^2B(A+B+AB)+A^3B^2(1+A+B)) \\
& = +z_{n+1}(A^2B^2+A^3B^2+A^3B^2(A+B+AB)) \\
& = -A^4B^3z_n
\end{aligned}
$$

. We can simplify further by taking $n=0$ and combining terms:

$$
\begin{aligned}
0 = & + z_5 \\
& - z_4(1+A+B(1+A+A^2)) \\
& = +z_3(A+B(1+A)A(1+A+A^2)+B^2A(1+A+A^2)) \\
& = -z_2(AB(1+A+A^2)+(A+1)B^2(A+A^2+A^3)+A^3B^3) \\
& = +z_1(AB^2(A+A^2+A^3)+A^3B^3(1+A)) \\
& = -A^4B^3z_0
\end{aligned}
$$

## H. Simulations **Simple model with/without noise**

We first define all parameters

```
ϵ= .01;
b = 2;
T = 3;
ω = 2 π/T;
θ = Exp[-ϵ b T];
```

Mathematica Code 1. Definitions

Next, we define the noise function. To do so, we first choose a parameter $dt>0$ and define the function

$$
\mathrm{Bumb}(t)=\begin{cases}1 & \text{if } 0\le t\le dt \\ 0 & \text{else}\end{cases}
$$

We now generate a random sequence of numbers $R_i$ and define the noise function

$$
n(t):=\epsilon^2\sum_i R_i\,\mathrm{Bump}(t-idt).
$$

```
dt = .5;
Bump[t_] = UnitStep[t] - UnitStep[t - dt];
TableR = RandomReal[{-1, 1}, 90/dt];
n[t_] = ϵ^(2) Sum[TableR[[i]] Bump[t - i dt], {i, 1, 90/dt}];
```

Mathematica Code 2. Definitions

Next, we implement the extremum seeking ODE

$$
\dot{y}+\epsilon b(1-\cos(2\omega t))y+by^2\sin(\omega t)=-b\epsilon^2\sin(\omega t)^3+n(t)\sin(\omega t).
$$

To generate the graphics without noise the last term in this equation must simply be dropped.

```
s = NDSolve[{y'[t] + ϵ b (1 - Cos[2 ω t]) y[t] +
      b y[t]^2 Sin[ω t] == -b ϵ² Sin[ω t]^3 +
      n[t] Sin[ω t], y[0] == 1.3}, y, {t, 0, 100},
   AccuracyGoal -> 15];
sol[t_] = y[t] /. s;

(*Plot*)
Plot[sol[t], {t, 0, 30}]
```

Mathematica Code 3. Definitions

Next, we implement the extraction scheme

$$L = \frac{(x_0 - x_1)x_2 + \theta(x_2 - x_1)x_0}{x_0 - (1+\theta)x_1 + \theta x_2}.$$

This formula requires the knowledge of $\theta$. To extract $\theta$, we first define

$$g := \frac{(x_0 - x_1)(x_2 - x_3)}{(x_1 - x_2)(x_0 - x_3)}$$

and obtain $\theta$ as

$$\theta = \frac{1-g}{2} - \frac{1}{2g}\sqrt{(1-g)^2 - 4g^2}.$$

We define extraction formulas for $L$ with the extracted and the exact value of $\theta$. The first is called ExtraL($t$) and the second one ExtraLCheat($t$).

```
g[t_] = Min[
   1/3, ((sol[t] - sol[t + T]) (sol[t + 2 T] -
        sol[t + 3 T]))/((sol[t + T] - sol[t + 2 T]) (sol[t] -
        sol[t + 3 T]))];

Extraθ[t_] = -(g[t] - 1)/(2 g[t]) -
   1/(2 g[t]) Sqrt[(g[t] - 1)^2 - 4 g[t]^2];

ExtraL[
   t_] = ((sol[t][[1]] - sol[t + T][[1]]) sol[t + 2 T][[1]] +
     Extraθ[t] sol[t][[
       1]] (sol[t + 2 T][[1]] - sol[t + T][[1]]))/(sol[t][[
      1]] - (1 + Extraθ[t]) sol[t + T][[1]] +
     Extraθ[t] sol[t + 2 T][[1]]);

ExtraLCheat[t_] = ((sol[t][[1]] - sol[t + T][[1]]) sol[t + 2 T][[
       1]] + θ sol[t][[
       1]] (sol[t + 2 T][[1]] - sol[t + T][[1]]))/(sol[t][[
      1]] - (1 + θ) sol[t + T][[1]] + θ sol[t + 2 T][[
       1]]);

(*Plot for extracted θ*)
Ptheta = Plot[{θ, Extraθ[t], θ}, {t, 0, 30} ,
  PlotRange -> {{0, 30}, {.86, 1}},
  AxesLabel -> {t,  "θ"}, LabelStyle -> {FontSize -> 15},
  PlotLegends -> {
    "Exact␣θ", "Extracted␣θ"},
  PlotStyle -> {RGBColor[0.368417, 0.506779, 0.709798],
    RGBColor[0.880722, 0.611041, 0.142051],
    RGBColor[0.368417, 0.506779, 0.709798]}]

(*Plot for g*)
Plot[g[t], {t, 0, 30}]

(*Plot for extracted L*)
PComp = Plot[{sol[t], ExtraL[t]}, {t, 0, 30}, PlotRange -> All,
  AxesLabel -> {t,
    "\!\(\*SubscriptBox[\(x\),␣\(Cl\)]\)␣vs␣\!\(\*SubscriptBox[\(x\),␣\
\(acc\)]\)"}, LabelStyle -> {FontSize -> 15}, AxesOrigin -> {0, -.2},
  PlotLegends -> {"\!\(\*SubscriptBox[\(x\),␣\(Cl\)]\)",
    "\!\(\*SubscriptBox[\(x\),␣\(acc\)]\)"}]
```

Mathematica Code 4. Definitions

The extraction scheme is also employed with an averaged value of $\theta$. First, we define an averaged value of $\theta$, that is obtained by averaging the extracted $\theta$ over intervals $[0, kT]$ for $k = 1,2,3$. Afterwards, the extraction formula for $L$ is implemented using these averaged values.

```
avθ1 =
 1/(T) NIntegrate[Extraθ[s], {s, 0, T}, AccuracyGoal -> 5,
   WorkingPrecision -> 10]
avθ2 =
 1/(2 T) NIntegrate[Extraθ[s], {s, 0, 2 T}, AccuracyGoal -> 5,
   WorkingPrecision -> 10]
avθ3 =
 1/(3 T) NIntegrate[Extraθ[s], {s, 0, 3 T}, AccuracyGoal -> 5,
   WorkingPrecision -> 10]

ExtraLnewmean1[
   t_] = ((sol[t][[1]] - sol[t + T][[1]]) sol[t + 2 T][[1]] +
     avθ1 sol[t][[
       1]] (sol[t + 2 T][[1]] - sol[t + T][[1]]))/(sol[t][[
      1]] - (1 + avθ1 ) sol[t + T][[1]] +
     avθ1 sol[t + 2 T][[1]]);
ExtraLnewmean2[
   t_] = ((sol[t][[1]] - sol[t + T][[1]]) sol[t + 2 T][[1]] +
     avθ2 sol[t][[
       1]] (sol[t + 2 T][[1]] - sol[t + T][[1]]))/(sol[t][[
      1]] - (1 + avθ2) sol[t + T][[1]] +
     avθ2 sol[t + 2 T][[1]]);
ExtraLnewmean3[
   t_] = ((sol[t][[1]] - sol[t + T][[1]]) sol[t + 2 T][[1]] +
     avθ3 sol[t][[
       1]] (sol[t + 2 T][[1]] - sol[t + T][[1]]))/(sol[t][[
      1]] - (1 + avθ3) sol[t + T][[1]] +
     avθ3 sol[t + 2 T][[1]]);

(*Plot*)
Plot[{ExtraLnewmean1[t], ExtraLnewmean2[t],
   ExtraLnewmean3[t]}, {t, 0, 30},
  PlotLegends -> {"L␣for␣\!\(\*SubscriptBox[\(θ\), \(1\)]\)",
    "L␣for␣\!\(\*SubscriptBox[\(θ\), \(2\)]\)",
    "L␣for␣\!\(\*SubscriptBox[\(θ\), \(3\)]\)"},
  AxesLabel -> {"t", "Extracted␣L"}, LabelStyle -> {FontSize -> 15}]
```

Mathematica Code 5. Definitions

### Including a Drift

We first define all parameters

```
T = 3;
ω = 2 π/T;
ϵ = .1;
δ = .4;
y0 = 2;
A = Exp[ϵ T];
q0 = .01;
A = Exp[ϵ T];
q[t_] = q0 Exp[-δ t];
```

Mathematica Code 6. Definitions

Next, we define the noise in the same way as we did before

```
dt = .5;
Bump[t_] = UnitStep[t] - UnitStep[t - dt];
TableR = RandomReal[{-1, 1}, 90/dt];
n[t_] = ϵ^(2) Sum[TableR[[i]] Bump[t - i dt], {i, 1, 90/dt}];
```

Mathematica Code 7. Definitions

We now implement the ODE

$$\dot{y} + 2\epsilon \sin^2(\omega t) y + y^2 \sin(\omega t) = -\epsilon^2 \sin(\omega t)^3 + \delta q(t) - \sin(\omega t) n(t).$$

To generate graphics without noise it again suffices to just drop the last term in this equation. We also define $h(t) := y(t) - q(t)$ and implement the extraction scheme

$$L = \frac{h_{n+2}h_{n+1} - (1+A)h_n h_{n+2} + A h_{n+2} h_{n+2}}{-h_{n+2} + (1+A)h_{n+1} - A h_n}.$$

To generate graphics without noise it again suffices to just drop the last term in this equation.

```
s = NDSolve[{y'[t] + 2 ϵ Sin[ω t]^2 y[t] +
      y[t]^2 Sin[ω t] == - ϵ^2 Sin[ω t]^3 + \
δ q[t] - Sin[ω t] n[t], y[0] == y0}, y, {t, 0, 70}];
sol[t_] = y[t] /. s;
h[t_] = sol[t] - q[t];
L[t_] = (h[t + T] h[t] - (1 + A) h[t + 2 T] h[
       t] + A h[t + 2 T] h[t + T])/(-h[t + 2 T] +
     h[t + T] (1 + A) - A h[t]);

(*Plot *)
Plot[{sol[t], L1[t]}, {t, 0, 30}, AxesOrigin -> {0, -.1},
 PlotRange -> {{0, 30}, {-.1, .6}}, LabelStyle -> {FontSize -> 20},
 AxesLabel -> {"t",
   "\!\(\*SubscriptBox[\(x\),␣\(\(Cl\)\(\\\␣\\\␣\)\)]\)vs.␣\
\!\(\*SubscriptBox[\(x\),␣\(acc\)]\)␣"},
 PlotLegends -> {
   Row[{"Classical␣ES"}],
   Row[{"\!\(\*SubscriptBox[\(x\),␣\(acc\)]\)␣with␣noise"}]
   }]
```

Mathematica Code 8. Definitions

Obtaining the graphics for various $\Gamma$ is achieved by generating plots for the various choices of parameters described in Subsection IV.C and combining the plots.