Data-driven Delay Estimation in Reaction-Diffusion Systems via Exponential Fitting ⋆

Rami Katz * **Giulia Giordano** * **Dmitry Batenkov** **

* *Department of Industrial Engineering, University of Trento, Italy (e-mail: ramkatsee@gmail.com, giulia.giordano@unitn.it)*
** *Department of Applied Mathematics, Tel Aviv University, Israel (e-mail: dbatenkov@tauex.tau.ac.il)*

Abstract: For a reaction-diffusion equation with unknown right-hand side and non-local measurements subject to unknown constant measurement delay, we consider the nonlinear inverse problem of estimating the associated leading eigenvalues and measurement delay from a finite number of noisy measurements. We propose a reconstruction criterion and, for small enough noise intensity, prove existence and uniqueness of the desired approximation and derive closed-form expressions for the first-order condition numbers, as well as bounds for their asymptotic behavior in a regime when the number of measurements tends to infinity and the inter-sampling interval length is fixed. We perform numerical simulations indicating that the exponential fitting algorithm ESPRIT is first-order optimal, namely, its first-order condition numbers have the same asymptotic behavior as the analytic ones in this regime.

Keywords: Time-delay systems, Data-driven control, Estimation

1. INTRODUCTION

Reaction-diffusion equations (RDEs) are widely used to model phenomena in physics and engineering, including magnetized plasma, flame front propagation and chemical processes (Sivashinsky 1977; Nicolaenko 1986). RDEs belong to the class of distributed parameters systems, and their control and observation have been extensively investigated over the last decades, see e.g. Balas (1988); Harkort and Deutscher (2011); Katz and Fridman (2022). In particular, observation and control of RDEs through modal decomposition was employed e.g. by Christofides (2001); Curtain (1982); Katz and Fridman (2021a). Almost all existing control and observation techniques assume explicit knowledge of the spatial operator of the system or of the eigenvalue/eigenfunction pairs corresponding to its modes.

Identification of unknown parameters in RDEs is a challenging problem, mostly studied in an adaptive estimation framework (Demetriou and Rosen 1994; Banks and Kunisch 2012). Adaptive estimation relies on a persistency of excitation assumption, which may be difficult to verify in practice. It also requires continuous-time measurements of the state and has not been generalized so far to a sampled-data framework and/or to estimation from a *finite* number of measurements. Finally, translation of these theoretical methods into tractable and efficient algorithms is, to the best of our knowledge, still an open problem. Other identification methods, accompanied by sound numerical algorithms, have been derived in the field of inverse problems (Lowe et al. 1992; Rundell and Sacks 1992; Kirsch 2011). These approaches treat the problem of recovering the spatial operator of the system under the assumption of *complete knowledge* of its eigenvalues. However, this assumption is non-realistic from a control theory perspective, since often only discrete-time measurements of the state are available. Hence, constructive and implementable data-driven identification techniques for reaction-diffusion equations are still missing.

In this work, we consider a 1D reaction-diffusion equation with unknown right-hand side and non-local measurements subject to an unknown, but upper bounded, constant measurement delay. Our goal is to estimate the delay and a finite number of dominant modes, corresponding to the RDE right-hand side. Our main contributions are the following:

(1) Differently from existing adaptive estimation methods, which require measurements of the form $y(t)$, $t \geq 0$, or $\{y(s_k)\}_{k=1}^{\infty}$ with $\lim_{k\to\infty} s_k = \infty$ and inter-sampling periods of sufficiently small length, we assume that measurements are taken at *finitely many* uniformly distributed time steps with arbitrary inter-sampling period length. Moreover, the measurements contain structured noise, with intensity $\varepsilon > 0$, which emanates from measuring 'undesirable' system modes.

(2) We reformulate the identification task in the framework of *exponential fitting*, a central topic in data analysis (Istratov and Vyvenko 1999; Pereyra and Scherer 2010; Batenkov et al. 2021). We define a reconstruction criterion, and prove existence and uniqueness of the associated approximation, if the intensity ε of the structured noise is not too large (Theorem 1).

(3) For the exponential fitting problem, we introduce first-order condition numbers (Theorem 1), which describe how the ε-noise is amplified in the reconstruction errors, and provide explicit expressions for them and for their asymptotic behavior in a specific parameter regime (see (21) and Theorem 2).

(4) We numerically compute the approximations via the ESPRIT algorithm (Roy and Kailath 1989) and show that it achieves first-order optimality, meaning that its first-order con-

⋆ The work of R. Katz and G. Giordano was funded by the European Union through the ERC INSPIRE grant (project number 101076926); views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council; neither the European Union nor the granting authority can be held responsible for them. D. Batenkov was supported by Israel Science Foundation Grant 1793/20.

dition numbers exhibit the same asymptotic behaviour as the analytic ones, in the considered regime.

2. MODEL AND PRELIMINARIES

We consider the 1D reaction-diffusion equation

$$z_t(x,t) = (p(x)z_x(x,t))_x + q(x)z(x,t), \quad (1)$$

with $x \in (0,1)$, $z(x,t) \in \mathbb{R}$ and $z(0,t) = z(1,t) = 0$, subject to non-local measurements

$$y(t) = \int_0^1 c(x)z(x,t-D)dx \in \mathbb{R}, \quad t \geq 0, \quad (2)$$

where $c \in L^2(0,1)$ is *partially known* and belongs to a certain class of kernels (see Assumption 1 in Section 3). The *unknown* smooth functions $p,q : [0,1] \to \mathbb{R}$ satisfy the bounds

$$0 < \underline{p} \leq p(x) \leq \overline{p} < \infty, \quad \underline{q} \leq q(x) \leq \overline{q}, \quad x \in [0,1], \quad (3)$$

where the constants $\underline{p}, \underline{q}, \overline{p}, \overline{q}$ do not need to be known. The *unknown* constant delay $D > 0$ satisfies $D < D_{\max}$, with a *known* upper bound $D_{\max}$. The initial condition $z(\cdot,0)$ is known and belongs to a class of admissible initial conditions (see Assumption 2 in Section 3); we set $z(\cdot,t) = z(\cdot,0),\ t < 0$.

We denote by $\mathscr{H}^2(0,1)$ (resp. $\mathscr{H}_0^1(0,1)$) the Sobolev space of functions f defined on $[0,1]$ that are twice (resp. once) weakly differentiable with $f'' \in L^2(0,1)$ (resp. $f' \in L^2(0,1)$ and $f(0) = f(1) = 0$). Define the operator $\mathscr{A}$

$$[\mathscr{A}h](x) = -\left(p(x)h'(x)\right)' - q(x)h(x), \quad x \in (0,1),$$
$$\mathrm{Dom}(\mathscr{A}) = \left\{h \in \mathscr{H}^2(0,1)\ ;\ h(0) = h(1) = 0\right\}.$$

System (1) is well posed (Katz and Fridman 2021b). Given $z(\cdot,0) \in L^2(0,1)$, the unique solution $z \in C([0,\infty), \mathscr{H}_0^1(0,1)) \cap C^1((0,\infty), \mathscr{H}_0^1(0,1))$ is such that $z(\cdot,t) \in \mathrm{Dom}(\mathscr{A})$ for all $t > 0$.

The operator $\mathscr{A}$ has an infinite monotone sequence of simple eigenvalues $\{\lambda_n\}_{n=1}^\infty$ satisfying $\lim_{n\to\infty} \lambda_n = \infty$ (Orlov 2017). The eigenvectors $\{\psi_n\}_{n=1}^\infty$ form a complete orthonormal system in $L^2(0,1)$. Also, the following inequality holds (Orlov 2017)

$$\pi^2 n^2 \underline{p} + \underline{q} \leq \lambda_n \leq \pi^2 n^2 \overline{p} + \overline{q}, \quad n \in \mathbb{N}. \quad (4)$$

We denote $[n] = \{i \in \mathbb{N}\ ;\ 1 \leq i \leq n\}$. The next proposition estimates the differences between pairs of eigenvalues $\{\lambda_n\}_{n=1}^\infty$ and will be essential in deriving asymptotic rates for the condition numbers (see Lemma 1).

Proposition 1. There exist constants $\upsilon, \Upsilon > 0$ such that

$$\upsilon\left(m^2 - n^2\right) \leq \lambda_m - \lambda_n \leq \Upsilon\left(m^2 - n^2\right) \quad (5)$$

holds for every choice of $1 \leq n < m$. $\diamond$

Proof. First, we show that there exists some constant $A_0 \in \mathbb{N}$ such that

$$\upsilon_0\left(m^2 - n^2\right) \leq \lambda_m - \lambda_n \leq \Upsilon_0\left(m^2 - n^2\right), \quad m > n \geq A_0, \quad (6)$$

for some $0 < \upsilon_0 < \Upsilon_0$. Indeed, by (Fulton and Pruess, 1994, Equation 4.21), the eigenvalues $\{\lambda_n\}_{n=1}^\infty$ have the asymptotic behavior $\lambda_n = \frac{\pi^2}{B^2}n^2 + a_0 + O\left(\frac{1}{n^2}\right)$, $n \geq 1$, where B and a_0 are *positive* constants. Taking the difference, we have

$$\lambda_m - \lambda_n = \frac{\pi^2}{B^2}\left(m^2 - n^2\right) + O\left(\frac{1}{n^2}\right) + O\left(\frac{1}{m^2}\right).$$

Let $m > n \geq A$, for a generic $A \in \mathbb{N}$. Since $\lim_{A\to\infty} O\left(\frac{1}{n^2}\right) + O\left(\frac{1}{m^2}\right) = 0$, there exists some $A_0 \in \mathbb{N}$ such that, for $m > n \geq A_0$,

$$-\frac{\pi^2}{2B^2} < O\left(\frac{1}{n^2}\right) + O\left(\frac{1}{m^2}\right) < 1.$$

Then, for this A_0 we have that any $m > n \geq A_0$ satisfy

$$\frac{\pi^2}{2B^2}\left(m^2 - n^2\right) \leq \lambda_m - \lambda_n \leq \left(\frac{\pi^2}{B^2} + 1\right)\left(m^2 - n^2\right).$$

Choosing $\upsilon_0 = \frac{\pi^2}{2B^2}$ and $\Upsilon_0 = \frac{\pi^2}{B^2} + 1$, we obtain (6).

Next, consider $m,n \in [A_0 - 1]$. Since $[A_0 - 1]$ is a finite set, there exist some $\upsilon_1 < \upsilon_0$ and $\Upsilon_0 < \Upsilon_1$ such that

$$\upsilon_1\left(m^2 - n^2\right) \leq \lambda_m - \lambda_n \leq \Upsilon_1\left(m^2 - n^2\right), \quad n < m \leq A_0 - 1. \quad (7)$$

In particular, (6) continues to hold with υ_0, Υ_0 replaced by υ_1, Υ_1, respectively. To finish the proof, we now show that there exist $\upsilon < \upsilon_1$ and $\Upsilon_1 < \Upsilon$ such that

$$\upsilon\left(m^2 - n^2\right) \leq \lambda_m - \lambda_n \leq \Upsilon\left(m^2 - n^2\right), n \leq A_0 - 1,\ m \geq A_0. \quad (8)$$

Assume υ cannot be found such that the lower bound in (8) holds. Then, setting $\upsilon_q = 2^{-q},\ q \in \mathbb{N}$, there exist $n_q \leq A_0 - 1$, $m_q \geq A_0$ such that

$$2^{-q} > \frac{\lambda_{m_q} - \lambda_{n_q}}{m_q^2 - n_q^2} \overset{(4)}{\geq} \frac{\pi^2\left(\underline{p}m_q^2 - \overline{p}A_0^2\right) + \underline{q} - \overline{q}}{m_q^2 - 1}.$$

Taking $q \to \infty$, which implies $m_q \to \infty$, we have $0 \geq \pi^2 \underline{p} > 0$, which is a contradiction. Similar arguments hold for Υ. ■

3. PROBLEM STATEMENT AND ASSUMPTIONS

Employing modal decomposition (Katz and Fridman 2021a), we present the solution to system (1) as

$$z(x,t) = \sum_{n=1}^\infty z_n(t)\psi_n(x), \quad z_n(t) = \langle z(\cdot,t), \psi_n \rangle, \quad n \in \mathbb{N}. \quad (9)$$

Differentiating under the integral sign and integrating by parts, we have $\dot{z}_n(t) = -\lambda_n z_n(t) \Longrightarrow z_n(t) = e^{-\lambda_n t} z_n(0)$ for all $n \in \mathbb{N}$, whence

$$z(x,t) = \sum_{n=1}^\infty z_n(0)e^{-\lambda_n t}\psi_n(x). \quad (10)$$

Substituting (10) into (2), for $t \geq D_{\max}$ we obtain

$$y(t) = \sum_{n=1}^\infty c_n z_n(0)e^{\lambda_n D}e^{-\lambda_n t}, \quad c_n = \langle c, \psi_n \rangle, \quad n \in \mathbb{N}. \quad (11)$$

Before formally stating our identification objective, we state our main assumptions on the system (1) and the measurements (2). There exist $N_1, N_2 \in \mathbb{N}$ such that the following properties hold.

Assumption 1. The measurement kernel $c \in L^2(0,1)$ belongs to the class of kernels whose coefficients $\{c_n\}_{n=1}^\infty$ satisfy

(a) $c_n = 0$ for all $n > N_1 + N_2$,
(b) c_n are *known and nonzero* for $n \in [N_1]$,
(c) $c_k = \varepsilon\tilde{c}_k$ for all $k \in [N_1 + N_2] \setminus [N_1]$, where for some $M_c > 0$ $|\tilde{c}_k| \leq M_c|c_n|$ for all $n \in [N_1]$ and all $k \in [N_1 + N_2] \setminus [N_1]$. $\diamond$

Assumptions 1(a) and 1(b) mean that $c \in L^2(0,1)$ is a bandlimited measurement kernel, supported on $\{\psi_n\}_{n=1}^{N_1+N_2}$, with known and nonzero projection coefficients on $\{\psi_n\}_{n=1}^{N_1}$. In addition, Assumption 1(c) means that the projection coefficients on $\{\psi_n\}_{n=N_1+1}^{N_2}$ are "small" in comparison to those on $\{\psi_n\}_{n=1}^{N_1}$. In signal processing terms, this means that the measurement kernel c (whose eigenstructure is shown in Figure 1) has a main lobe on the frequency domain $\{\lambda_n\}_{n=1}^{N_1}$ as well as an *undesirable* ε-small side lobe on the frequency domain $\{\lambda_n\}_{n=N_1+1}^{N_2}$, which is associated with *structured noise* in the measurements.

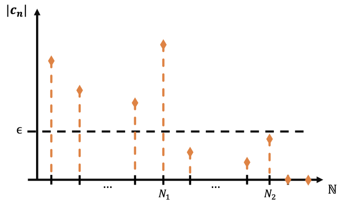


Fig. 1. Eigenstructure of the kernel $c \in L^2(0,1)$.

Assumption 2. The initial condition $z(\cdot,0) \in L^2(0,1)$, $z_n(0) \neq 0$, is *known* for $n \in [N_1]$. Furthermore, $\frac{|z_k(0)|}{|z_n(0)|} \leq M_z$ for some $M_z > 0, n \in [N_1]$ and $k \in [N_1+N_2] \setminus [N_1]$. $\diamond$

Assumption 3. The system measurements are taken *only* at times $\{t_k := k\Delta + D_{\max}\}_{k=0}^{2N_1-1}$, with step-size $\Delta > 0$. $\diamond$

Assumption 3 implies that we only have finitely many "snapshots" of the system output.

Subject to Assumptions 1-3, the measurements (11) at the available times $\{t_k\}_{k=0}^{2N_1-1}$ can be presented as

$$y(t_k) = \underbrace{\sum_{n=1}^{N_1} y_n e^{-\lambda_n k\Delta}}_{y_{\text{main}}(t_k)} + \varepsilon \underbrace{\sum_{n=N_1+1}^{N_1+N_2} y_n e^{-\lambda_n k\Delta}}_{y_{\text{tail}}(t_k)}, \tag{12}$$

for $k = 0, \ldots, 2N_1 - 1$, where

$$y_n := \begin{cases} c_n z_n(0) e^{\lambda_n (D-D_{\max})}, & n \in [N_1] \\ \tilde{c}_n z_n(0) e^{\lambda_n (D-D_{\max})}, & n \in [N_1+N_2] \setminus [N_1] \end{cases} \tag{13}$$

satisfy $\frac{|y_k|}{|y_n|} \leq M_c M_z =: M_y$ for all $n \in [N_1], k \in [N_1+N_2] \setminus [N_1]$, since $e^{(\lambda_k - \lambda_n)(D-D_{\max})} < 1$ for such indices, as $D < D_{\max}$.

Identification objective: Given the measurements (12) and $N_0 \in [N_1]$, estimate the eigenvalues $\{\lambda_n\}_{n=1}^{N_0}$ and the constant delay D.

The problem of recovering $\{y_n, \lambda_n\}_{n=1}^{N_1}$ from the measurements (12) is known as *exponential fitting* (Pereyra and Scherer 2010). The components $\{y_{\text{tail}}(t_k)\}_{k=0}^{2N_1-1}$ in the measurements constitute an "ε-structured" measurement noise, emanating from the fact that c is not a perfect filter (i.e., $c \notin \text{span}\{\psi_n\}_{n=1}^{N_0}$). For $\varepsilon = 0$, there exist multiple methods which recover $\{y_n, \lambda_n\}_{n=1}^{N_1}$ *exactly*; see the discussion of the ESPRIT algorithm (Roy and Kailath 1989) in Section 5. However, if $\varepsilon > 0$, the structured measurement noise introduces errors into the estimation. To the best of our knowledge, error estimates for exponential fitting in the presence of small structured noise do not exist currently in the literature (except in some very special cases such as Batenkov et al. 2021). The goal of this work is to study the analytic estimation errors due to noise, to first order in ε, and to show that the ESPRIT algorithm is first-order optimal in achieving the identification objective, thereby gaining insight into the system (1).

The considered problem is highly challenging for two reasons. First, we assume that only *finitely many* measurements are available for the reconstruction procedure, for any triplet (Δ, N_1, N_2). Second, although (1) is a linear system, the task of recovering $\{D, \lambda_n\}_{n=1}^{N_0}$ from the measurements (12) is a *nonlinear inverse problem*, as the measurements depend *nonlinearly* on these parameters.

4. IDENTIFICATION CRITERION AND ITS ANALYSIS

Given (12), we introduce the map

$$\mathscr{F}\left(\left\{\hat{y}_n, \hat{\lambda}_n\right\}_{n=1}^{N_1}; \varepsilon\right) = \text{col}\left\{\sum_{n=1}^{N_1} \hat{y}_n e^{-\hat{\lambda}_n k\Delta} - y(t_k)\right\}_{k=0}^{2N_1-1}. \tag{14}$$

Given an approximation candidate $\hat{P} := \left\{\hat{y}_n, \hat{\lambda}_n\right\}_{n=1}^{N_1}$, the function $\mathscr{F}(\hat{P};\varepsilon)$ returns the discrepancy between measurements $\{y(t_k)\}_{k=0}^{2N_1-1}$ and "virtual measurements" $\left\{\sum_{n=1}^{N_1} \hat{y}_n e^{-\hat{\lambda}_n k\Delta}\right\}_{k=0}^{2N_1-1}$. In particular, if $\varepsilon = 0$ (i.e., there is no structured noise in the measurements), we see that $\mathscr{F}\left(\{y_n, \lambda_n\}_{n=1}^{N_1}; 0\right) = 0$. As an identification criterion, we look for estimates $\hat{P}$ that maintain the equality $\mathscr{F}(\hat{P};\varepsilon) = 0$ even in the presence of noise $\varepsilon > 0$.

Definition 1. We say that $\hat{P}$ is an *ε-approximation* of $\{y_n, \lambda_n\}$ if $\mathscr{F}\left(\hat{P};\varepsilon\right) = 0$. To avoid ambiguity, we always assume that the elements of $\hat{P}$ are sorted in increasing order of eigenvalues, i.e. $\hat{\lambda}_k < \hat{\lambda}_{k+1}$ for all k. $\diamond$

Given an ε-approximation $\hat{P}$, one can generate

$$\hat{D}^{(n)} = \hat{\lambda}_n^{-1} \log\left(\frac{\hat{y}_n}{c_n z_n(0) e^{-\hat{\lambda}_n D_{\max}}}\right) \iff n \in \left\{m \in [N_1]\,;\ \frac{\hat{y}_m}{c_m z_m(0)} > 0\right\} =: \mathscr{Y}_D$$

provided $\mathscr{Y}_D \neq \emptyset$ (see Section 5). In that case, we propose to use $\hat{D}^{(n)}$ to approximate D, where $n = \min \mathscr{Y}_D$.

Remark 1. In (13) λ_n and y_n, $n \in [N_1]$, are not independent. However, in (14) we search for a candidate $\hat{P}$, where $\hat{\lambda}_n$ and $\hat{y}_n$, $n \in [N_1]$, are treated as independent. This approach may lead to a loss of structure, but it has the key advantage of yielding a well-posed inverse problem (see Theorem 1) for whose solution tractable numerical algorithms exist (see Section 5). $\diamond$

In the following analysis, only to keep the presentation simpler, we assume that $N_2 = 1$: the sum $y_{\text{tail}}(t_k)$, $k \in \{0\} \cup [2N_1 - 1]$, in (12) contains a single term. Our analysis and conclusions remain identical for an arbitrary fixed $N_2 \in \mathbb{N}$.
Hereafter, we use the notation

$$\phi_n := e^{-\lambda_n \Delta}, \quad \hat{\phi}_n = e^{-\hat{\lambda}_n \Delta}, \quad n \in [N_1+1]. \tag{15}$$

The measurements in (12) are then rewritten as

$$y(t_k) = \sum_{n=1}^{N_1} y_n \phi_n^k \; + \; \varepsilon \, y_{N_1+1} \, \phi_{N_1+1}^k, k \in \{0\} \cup [2N_1 - 1]. \tag{16}$$

To show that our criterion is well defined, we prove the existence and uniqueness of an ε-approximation, for small $\varepsilon > 0$.

Theorem 1. There exist $\varepsilon_* > 0$ and unique continuously differentiable functions $\hat{P}(\varepsilon) := \left\{\hat{y}_n(\varepsilon), \hat{\lambda}_n(\varepsilon)\right\}$ such that $\hat{P}(0) = \{y_n, \lambda_n\}_{n=1}^{N_1}$ and for all $|\varepsilon| < \varepsilon_*$, $\mathscr{F}(\hat{P};\varepsilon) = 0 \iff \hat{P} = \hat{P}(\varepsilon)$. Furthermore, the components of $\hat{P}(\varepsilon)$ are continuously differentiable on $|\varepsilon| < \varepsilon_*$ and satisfy as $\varepsilon \to 0$

$$\begin{aligned} \hat{\lambda}_n(\varepsilon) - \lambda_n &= \mathscr{K}_\lambda(n; N_1, \Delta)\varepsilon + o_{n,N_1,\Delta}(\varepsilon), \\ \hat{y}_n(\varepsilon) - y_n &= \mathscr{K}_y(n; N_1, \Delta)\varepsilon + o_{n,N_1,\Delta}(\varepsilon), \end{aligned} \tag{17}$$

for

$$\begin{bmatrix} \mathscr{K}_y(n; N_1, \Delta) \\ \mathscr{K}_\lambda(n; N_1, \Delta) \end{bmatrix} = y_{N_1+1} \begin{bmatrix} H_{\Phi,n}(\phi_{N_1+1}) \\ -\dfrac{1}{\Delta y_n \phi_n} \tilde{H}_{\Phi,n}(\phi_{N_1+1}) \end{bmatrix}, \; n \in [N_1]. \tag{18}$$

Here $\left\{H_{\Phi,n}, \tilde{H}_{\Phi,n}\right\}_{n=1}^{N_1}$ are the Hermite interpolation basis polynomials, given in (A.2), associated with $\Phi = \{\phi_n\}_{n=1}^{N_1}$. $\diamond$

The terms $\mathscr{K}_y(n;N_1,\Delta)$ and $\mathscr{K}_\lambda(n;N_1,\Delta)$, $n \in [N_1]$ are the *first order (in ε) condition numbers* of the problem. Henceforth, we will suppress their dependence on N_1, Δ for brevity.

Proof. $\mathscr{F}(\hat{P},\varepsilon)$ is differentiable in all variables $(\hat{P},\varepsilon)$. We denote by $\partial_{\hat{P}}\mathscr{F}(P,0)$ its Jacobian with respect to $\hat{P}$ evaluated at $\hat{P} = \{y_n,\lambda_n\}_{n=1}^{N_1} =: P$ and $\varepsilon = 0$. A direct computation yields

$$\partial_{\hat{y}_j}\mathscr{F}(P,0) = \mathrm{col}\left\{e^{-\lambda_n k\Delta}\right\}_{k=0}^{2N_1-1} = \mathrm{col}\left\{\phi_j^k\right\}_{k=0}^{2N_1-1},$$
$$\partial_{\hat{\lambda}_j}\mathscr{F}(P,0) = \mathrm{col}\left\{-k\Delta y_j e^{-\lambda_j k\Delta}\right\}_{k=0}^{2N_1-1} = \mathrm{col}\left\{-k\Delta y_j \phi_j^k\right\}_{k=0}^{2N_1-1}.$$

Thus $\partial_{\hat{P}}\mathscr{F}(P,0) = J(P,0)D(P,0)$, with

$$J(P,0) = W_\Phi^\top, \quad D(P,0) = \mathrm{diag}\left\{\begin{bmatrix}1 & 0\\ 0 & -\Delta y_n\phi_n\end{bmatrix}\right\}_{n=1}^{N_1}, \tag{19}$$

where W_Φ, which is associated with Φ, is given in (A.3).

Since the eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ are simple, it follows from the uniqueness of Hermite interpolation that $W_\Phi^\top$ is invertible and $W_\Phi^{-\top} = \mathbb{H}_{N_1}(\Phi)$. In view of Assumptions 1-2 and of (13), we have that $y_n \neq 0$, $n \in [N_1]$, whereas $\phi_n \neq 0$, $n \in [N_1]$, by definition. Therefore, $\det(\partial_{\hat{P}}\mathscr{F}(P,0)) \neq 0$. The implicit function theorem (Spivak 2018) guarantees that there exist $\varepsilon_* > 0$ and unique continuously differentiable functions $\hat{P}(\varepsilon)$ such that $\hat{P}(0) = P$ and $\mathscr{F}(\hat{P};\varepsilon) = 0 \iff \hat{P} = \hat{P}(\varepsilon)$ for all $|\varepsilon| < \varepsilon_*$. Differentiating $\mathscr{F}(\hat{P}(\varepsilon),\varepsilon) = 0$ with respect to ε and substituting $\varepsilon = 0$, we obtain

$$\begin{aligned}\mathrm{col}\left\{\begin{bmatrix}\mathscr{K}_y(n)\\ \mathscr{K}_\lambda(n)\end{bmatrix}\right\}_{n=1}^{N_1} &= \partial_{\hat{P}}\mathscr{F}(P,0)^{-1} y_{N_1+1}\,\mathrm{col}\left\{\phi_{N_1+1}^k\right\}_{k=0}^{2N_1-1}\\ &= \mathrm{diag}\left\{\begin{bmatrix}y_{N_1+1} & 0\\ 0 & -\frac{y_{N_1+1}}{\Delta y_n\phi_n}\end{bmatrix}\right\}_{n=1}^{N_1}\mathbb{H}_{N_1}(\Phi)\,\mathrm{col}\left\{\phi_{N_1+1}^k\right\}_{k=0}^{2N_1-1}.\end{aligned} \tag{20}$$

Since (A.4) holds, we obtain the expression in (18). ∎

Equation (17) in Theorem 1 implies that a small ε-perturbation in the measurements is amplified in the reconstruction errors

$$\frac{e_\lambda(n)}{\varepsilon} := \frac{\hat{\lambda}_n(\varepsilon)-\lambda_n}{\varepsilon} = \mathscr{K}_\lambda(n) + o_{n,N_1,\Delta}(1),$$
$$\frac{e_y(n)}{\varepsilon} := \frac{\hat{y}_n(\varepsilon)-y_n}{\varepsilon} = \mathscr{K}_y(n) + o_{n,N_1,\Delta}(1)$$

by the condition numbers $\mathscr{K}_y(n)$ and $\mathscr{K}_\lambda(n)$, when seeking for the ε-approximation $\hat{P}$.

We now consider the **asymptotic analysis of the condition numbers $\mathscr{K}_y(n)$ and $\mathscr{K}_\lambda(n)$**, of which we wish to determine the dependence on N_1 and Δ, so as to better understand the reconstruction errors using the proposed criterion. There are many regimes that relate N_1 and Δ, which might be of interest. In this work we focus on

$$\underline{\text{Regime:}}\ \Delta \text{ fixed and } N_1 \to \infty \tag{21}$$

corresponding to growing support of the measurement kernel $c \in L^2(0,1)$ in the frequency domain, subject to Assumption 1.

Recall the Hermite interpolation polynomials in (18); see also the Appendix. Given $\Delta > 0$, $N_1 \in \mathbb{N}$ and $n \in [N_1]$, let

$$\begin{aligned}\xi_1 &= \prod_{j\neq n}(\phi_{N_1+1}-\phi_j)^2, \quad \xi_2 = \prod_{j=1}^{n-1}(\phi_n-\phi_j)^2,\\ \xi_3 &= \prod_{j=n+1}^{N_1}(\phi_n-\phi_j)^2, \quad \xi_4 = \sum_{k\neq n}|\phi_n-\phi_k|^{-1},\end{aligned} \tag{22}$$

where all summations/products range over indices in $[N_1]$, and we use the convention that $\prod_{j=l}^k b_j = 1$ and $\sum_{j=l}^k b_j = 0$ whenever $k < l$. We omit the dependence of functions on (n,N_1,Δ) for simplicity of notation.

Remark 2. Recalling the Lagrange polynomials given in (A.1) we observe that $L_{\Phi,n}^2(\phi_{N_1+1}) = \frac{\xi_1}{\xi_2\xi_3}$ and $\left|L'_{\Phi,n}(\phi_n)\right| = \xi_4$. ◇

To prove our main result on the asymptotic behavior of the condition numbers, we need several lemmas. In the next, we employ the notations in Proposition 2 in the Appendix.

Lemma 1. The functions in (22) can be written as

$$\xi_1 = e^{-2\Delta\sum_{j\neq n}\lambda_j - 2\theta_1}, \quad \xi_2 = e^{-2\Delta\sum_{j=1}^{n-1}\lambda_j - 2\theta_2}, \quad n > 1,$$
$$\xi_3 = e^{-2\Delta(N_1-n)\lambda_n - 2\theta_3}, \ n < N_1,$$

where

$$\begin{aligned}\theta_1 &:= \sum_{j\neq n} -\log\left(1-e^{-\Delta(\lambda_{N_1+1}-\lambda_j)}\right) > 0,\\ \theta_2 &:= \sum_{j=1}^{n-1} -\log\left(1-e^{-\Delta(\lambda_n-\lambda_j)}\right) > 0, \quad n > 1,\\ \theta_3 &:= \sum_{j=n+1}^{N_1} -\log\left(1-e^{-\Delta(\lambda_j-\lambda_n)}\right) > 0, \ n < N_1\end{aligned} \tag{23}$$

satisfy the inequalities

$$\begin{aligned}&\theta_1 \leq \mathscr{J}_{0,\infty}(\Delta\upsilon N_1) + \log\left(1-e^{-\Delta\upsilon(N_1+1-n)(N_1+1)}\right),\\ &\theta_1 \geq \mathscr{J}_{1,2}(\Delta\Upsilon(2N_1+1)) + \log\left(1-e^{-\Delta\Upsilon(N_1+1-n)(2N_1+1)}\right),\\ &\mathscr{J}_{1,2}(\Delta\Upsilon(2n-1)) \leq \theta_2 \leq \mathscr{J}_{0,\infty}(\Delta\upsilon(n+1)),\\ &\mathscr{J}_{1,2}(\Delta\Upsilon(N_1+n+1)) \leq \theta_3 \leq \mathscr{J}_{0,\infty}(\Delta\upsilon(2n+1)),\end{aligned} \tag{24}$$

where the positive constants υ and Υ are those given in Proposition 1 and the function $\mathscr{J}_{w_1,w_2}$ is given in (B.1). ◇

Proof. We consider ξ_1 only. The results for ξ_2 and ξ_3 are proved similarly. We have

$$\log(\xi_1) = -2\Delta\sum_{j\neq n}\lambda_j - 2\theta_1,$$

with θ_1 given in (23). Employing (5) in (23), we obtain

$$\theta_1 \geq \sum_{j\neq n} -\log\left(1-e^{-\Delta\Upsilon((N_1+1)^2-j^2)}\right).$$

Let $\ell := \log\left(1-e^{-\Delta\Upsilon(N_1+1-n)(2N_1+1)}\right)$. Then, we have

$$\begin{aligned}&\sum_{j\neq n} -\log\left(1-e^{-\Delta\Upsilon((N_1+1)^2-j^2)}\right) - \ell \geq \sum_{j=1}^{N_1} -\log\left(1-e^{-\Delta\Upsilon j(2N_1+1)}\right)\\ &\geq \int_1^{N_1+1}\mathscr{Q}_{\Delta\Upsilon(2N_1+1)}(x)\mathrm{d}x = \mathscr{J}_{1,N_1+1}(\Delta\Upsilon(2N_1+1)) \geq \mathscr{J}_{1,2}(\Delta\Upsilon(2N_1+1)).\end{aligned}$$

The first inequality holds as $(N_1+1)^2-j^2 \leq (N_1+1-j)(2N_1+1)$, while the second one holds because the sum in the second row can be viewed as Riemannian sum of the positive and monotonically decreasing function $\mathscr{Q}_{\Delta\Upsilon(2N_1+1)}(x)$ over $x \in [1,N_1]$. Hence, the integral provides a lower bound for the sum. The upper bound is proved analogously, using $(N_1+1-j)(N_1+1) \leq (N_1+1)^2-j^2$, $\theta_1 \leq \sum_{j\neq n} -\log\left(1-e^{-\Delta\upsilon((N_1+1)^2-j^2)}\right)$ and

$$\begin{aligned}&\sum_{j\neq n} -\log\left(1-e^{-\Delta\upsilon((N_1+1)^2-j^2)}\right)\\ &\leq \int_0^{N_1}\mathscr{Q}_{\Delta\upsilon(N_1+1)}(x)\mathrm{d}x + \log\left(1-e^{-\Delta\upsilon(N_1+1-n)(N_1+1)}\right)\\ &\leq \mathscr{J}_{0,\infty}(\Delta\upsilon(N_1+1)) + \log\left(1-e^{-\Delta\upsilon(N_1+1-n)(N_1+1)}\right).\end{aligned}$$ ∎

Next, recall the Lagrange polynomials $\{L_{\Phi,n}\}_{n\in[N_1]}$ as in (A.1), where $\Phi = \{\phi_n\}_{n=1}^{N_1}$.

Lemma 2. For $n \in [N_1]$, we have

$$L_{\Phi,n}^2(\phi_{N_1+1}) = \begin{cases} e^{-2\Delta\sum_{j=n+1}^{N_1}(\lambda_j-\lambda_n)+\Theta}, & n < N_1,\\ e^{2\theta_2-2\theta_1}, & n = N_1,\end{cases} \tag{25}$$

where

$$\Theta = \begin{cases} -2(\theta_1-\theta_2-\theta_3), & n > 1,\\ -2(\theta_1-\theta_3), & n = 1.\end{cases} \tag{26}$$

Moreover, fixing $n \in [N_1-1]$, $\Delta > 0$ and denoting

$$\sigma_n^{N_1} := \frac{N_1(N_1+1)(2N_1+1)}{6} - \frac{n(n+1)(2n+1)}{6} - (N_1-n)n^2,$$

there exists a constant $M_\phi = M_\phi(\Delta) > 0$ such that

$$L^2_{\Phi,n}(\phi_{N_1+1}) \le M_\phi e^{-2\Delta\sigma_n^{N_1}} = M_\phi e^{-\frac{2\Delta\upsilon N_1^3}{3}\left(1+O_n\left(N_1^{-2}\right)\right)}. \quad \diamond \quad (27)$$

Proof. The equality (25) follows from Lemma 1 and the fact that $L^2_{\Phi,n}(\phi_{N_1+1}) = \frac{\xi_1}{\xi_2 \cdot \xi_3}$.
Fix $n \in [N_1-1]$ and consider (25). By Proposition 1,

$$-2\Delta \sum_{j=n+1}^{N_1} (\lambda_j - \lambda_n) \overset{(5)}{\le} -2\Delta\upsilon\sigma_n^{N_1} = -\frac{2\Delta\upsilon N_1^3}{3}\left(1+O_n\left(N_1^{-2}\right)\right), \quad (28)$$

as $N_1 \to \infty$ (note that $O_n\left(N_1^{-2}\right)$ is independent of Δ). On the other hand, consider the lower and upper bounds in (24). In view of these bounds and Proposition 2, θ_1, θ_2 and θ_3 are uniformly bounded in $N_1 \ge n$ (recall that Δ and n are considered fixed). Hence, Θ in (26) is uniformly bounded for all $N_1 \in \mathbb{N}$. Combining the latter with (25) and (28), we obtain (27). ■

Lemma 3. The term ξ_4 in (22) satisfies

$$\xi_4 \le M_\xi \frac{e^{\Delta\lambda_{N_1}}}{\Delta} \quad (29)$$

for some $M_\xi > 0$ independent of $\Delta > 0$. $\diamond$

Proof. We write $\xi_4 = \xi_{4,1} + \xi_{4,2}$, where

$$\xi_{4,1} = \sum_{k\in[n-1]} \frac{1}{|\phi_n - \phi_k|} \text{ and } \xi_{4,2} = \sum_{k=n+1}^{N_1} \frac{1}{|\phi_n - \phi_k|}.$$

For $\xi_{4,1}$ with $n > 1$, we have

$$\xi_{4,1} \le \frac{e^{\Delta\lambda_n}}{\Delta} \sum_{k=1}^{n-1} \frac{1}{\lambda_n - \lambda_k} \le \frac{e^{\Delta\lambda_n}}{\Delta\upsilon} \sum_{k=1}^{n-1} \frac{1}{n^2-k^2} \le \frac{e^{\Delta\lambda_n}}{\Delta\upsilon}\frac{\ln(2n)}{2n}, \quad (30)$$

where the first inequality follows from the application of Lagrange's theorem with the derivative computed at λ_n, the second follows from (5). The third inequality follows from comparison with the integral of the positive and monotonically increasing function $x \mapsto (n^2-x^2)^{-1}$ on $x \in [1, n-1]$. Analogously, for $n < N_1$ we obtain

$$\xi_{4,2} \le \Delta^{-1} \sum_{k=n+1}^{N_1} \frac{e^{\Delta\lambda_k}}{\lambda_k - \lambda_n} \overset{(5)}{\le} \frac{e^{\Delta\lambda_{N_1}}}{\Delta\upsilon} \sum_{k=n+1}^{N_1} \frac{1}{k^2-n^2} \le \frac{e^{\Delta\lambda_{N_1}}}{\Delta\upsilon}\frac{1+\ln(2n+1)}{2n}. \quad (31)$$

The result follows from (30), (31) since $\frac{\ln(2n)}{2n}$ and $\frac{1+\ln(2n+1)}{2n}$ are bounded on $n \in \mathbb{N}$. ■

We can now establish the asymptotic behaviour of the condition numbers $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$ in the regime (21).

Theorem 2. Recall the first-order condition numbers $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$ in (17). Let $n \in \mathbb{N}$. Given $\Delta > 0$, there exist some $\gamma_y(n,\Delta) > 0$ and $\gamma_\lambda(n,\Delta) > 0$ such that, as $N_1 \to \infty$,

$$\left|\mathscr{K}_y(n)\right| \le \gamma_y(n,\Delta) \cdot |z_{N_1+1}(0)\tilde{c}_{N_1+1}| e^{-\frac{2}{3}\Delta\upsilon N_1^3\left(1+O(N_1^{-1})\right)},$$
$$|\mathscr{K}_\lambda(n)| \le \gamma_\lambda(n,\Delta) \cdot e^{-\frac{2}{3}\Delta\upsilon N_1^3\left(1+O(N_1^{-2})\right)}. \quad \diamond$$

The condition numbers $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$ determine, to first order in ε, how much an ε-perturbation in the measurements is amplified in the ε-approximation $\hat{P}$; in view of Theorem 2, in the regime (21) they decay *super-exponentially* with $N_1 \to \infty$.

Proof. For $\mathscr{K}_y(n)$, in view of (18) and (A.2), we have

$$\left|\mathscr{K}_y(n)\right| \overset{\text{Remark2}}{\le} |y_{N_1+1}| \left(1 + 2e^{-\Delta\lambda_n}\xi_4\right) L^2_{\phi,n}(\phi_{N_1+1}).$$

Employing (27) and (29),

$$\left|\mathscr{K}_y(n)\right| \le M_\phi |y_{N_1+1}| \left(\Delta + 2e^{-\Delta\lambda_n} M_\xi e^{\Delta\lambda_{N_1}}\right)\Delta^{-1} e^{-2\Delta\upsilon\sigma_n^{N_1}}$$
$$= M_\phi \Delta^{-1}\left(\Delta e^{-\Delta\lambda_{N_1}} + 2e^{-\Delta\lambda_n} M_\xi\right)|y_{N_1+1}| e^{-2\Delta\left[\upsilon\sigma_n^{N_1} + \lambda_{N_1+1}\right]}.$$

Recalling (13), (28) and using the fact that, when $N_1 \to \infty$,

$$\upsilon\sigma_n^{N_1} + \lambda_{N_1+1}(D+1-D_{\max}) = \frac{2}{3}\upsilon N_1^3\left(1+O(N_1^{-1})\right),$$

we obtain the bound on $\mathscr{K}_y(n)$. Similarly, we have

$$|\mathscr{K}_\lambda(n)| = \frac{|y_{N_1+1}|}{\Delta|y_n|}\frac{|\phi_{N_1+1} - \phi_n|}{\phi_n} L^2_{\Phi,n}(\phi_{N_1+1}).$$

By Assumptions 1-3, $\frac{|y_{N_1+1}|}{|y_n|} \le M_y$, whereas $\frac{|\phi_{N_1+1}-\phi_n|}{\phi_n} \le 1$. Hence, from (27), we again have

$$|\mathscr{K}_\lambda(n)| \le M_y M_\phi \Delta^{-1} e^{-2\Delta\sigma_n^{N_1}}.$$

In light of (28), we obtain the bound on $\mathscr{K}_\lambda(n)$. ■

Remark 3. Theorem 2 continues to hold uniformly for $n \le \lfloor \beta N_1 \rfloor$, $\beta < 1$ as $N_1 \to \infty$: in this case, $\frac{\sigma_n^{N_1}}{N_1^3}$ is lower bounded by a positive constant for all $n \le \lfloor \beta N_1 \rfloor$, thereby the super-exponential decay rate is preserved. $\diamond$

5. NUMERICAL RESULTS

Our numerical simulations, implemented in Wolfram Mathematica, show that the first-order condition numbers of the ESPRIT algorithm (Roy and Kailath 1989) exhibit the same asymptotic behaviour as $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$ in the regime (21).

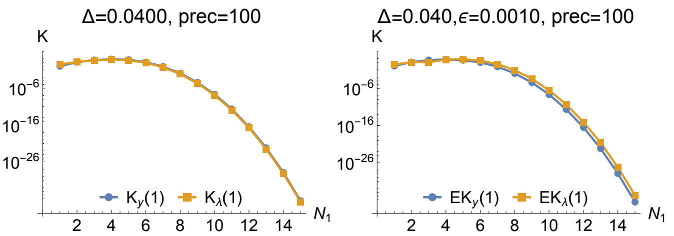
5.1 Multi-exponential model with structured perturbations

We start by examining the numerical conditioning of the multi-exponential model (12) with a structured (multi-exponential) perturbation term. These simulations are aimed at verifying the behavior of the analytic condition numbers of the multi-exponential model (12). A complete PDE delay estimation will be presented in the next subsection. We set $\lambda_n = n^2$, $y_n = 1$ for all n, and fix $\Delta = 0.04$, $N_2 = 1$. For various values of N_1, we compute the ideal condition numbers $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$, $n \in [N_1]$, given by (18). The results are shown in Fig. 2a. Super-exponential decay is clearly seen, as predicted by Theorem 2.

The ESPRIT algorithm (Roy and Kailath 1989) is one of the best-performing methods for exponential fitting. It requires at least $2N_1$ equispaced samples of the signal $y(t)$ of the form (12), and produces estimates of the parameters $\{y_n, \lambda_n\}_{n=1}^{N_1}$. It provides exact solutions when $\varepsilon = 0$, and performs close to optimal in the presence of noise, in the context of the so-called super-resolution problem in applied harmonic analysis (Batenkov et al. 2021). We apply ESPRIT to the sequence $\{y(t_k)\}_{k=0}^{2N_1-1}$, with the same setup as described above. In Fig. 2b, we see that the conditioning of the ESPRIT algorithm is consistent with Theorem 2 and the computed condition numbers in Fig. 2a. We plot the rescaled errors (recall (17)),

$$\mathscr{E}\mathscr{K}_\lambda(n) = \varepsilon^{-1}\left|\hat{\lambda}_n^{ESP} - \lambda_n\right|, \ \mathscr{E}\mathscr{K}_y(n) = \varepsilon^{-1}\left|\hat{y}_n^{ESP} - y_n\right|, \quad (32)$$

where $\hat{\lambda}_n^{ESP}$ and $\hat{y}_n^{ESP}$ are the parameter values recovered by ESPRIT, and, furthermore, the $\hat{\lambda}_n^{ESP}$'s have been sorted in increasing order. Here $\varepsilon = 10^{-3}$, and the results were computed with 100 decimal digits of precision.



(a) Condition numbers of λ_1, y_1.

(b) ESPRIT conditioning.

Fig. 2. Condition numbers and ESPRIT performance.

5.2 PDE delay estimation

We test the complete procedure on a PDE identification problem. We consider the RDE (1) with constant $p \equiv q \equiv 0.1$. The eigenvalues and eigenfunctions are explicitly given by $\lambda_n = n^2\pi^2 - q$ and $\psi_n(x) = \sqrt{2}\sin(n\pi x)$. The initial condition is set to satisfy $z_n(0) = (-1)^{n+1}\left(\sqrt{2}n^3\right)^{-1}$. To solve the RDE, we use the method of lines for space discretization with $N_x = 40$ collocation points and 32nd order finite difference approximation to attain high accuracy. The resulting ODE system is integrated for $t \in [0, 1.5]$, with the resulting solution and the initial condition plotted in Fig. 3. Our implementation utilized the `NDSolve` library function.

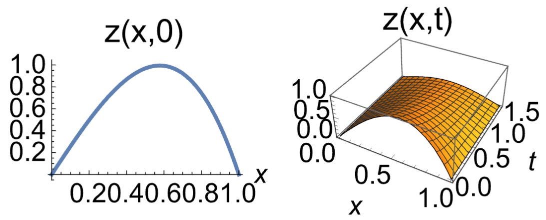


Fig. 3. The RDE initial condition and the solution.

Next, we consider the measurement model (2) with (a-priori unknown) delay $D = 1/12$. We fix the sampling step size to be $\Delta = 1/25$. We further fix $D_{\max} = 1/10$, for each $N_1 = 1, \ldots, N_{\max} = 10$. First, the measurement function is chosen with random coefficents $c_n \in [1, 2]$ and $\varepsilon = 0.01$:

$$c(x) = \sum_{n=1}^{N_1} c_n \psi_n(x) + \varepsilon \sum_{n=N_1+1}^{N_1+N_2} c_n \psi_n(x).$$

Next, the measurement function (2) is computed by global adaptive quadrature as implemented in `NIntegrate` library function, and sampled at the points $t_k = D_{\max} + k\Delta$, for $k = 0, \ldots, 2N_1 - 1$, giving the measurement vector $\operatorname{col}\{y(t_k)\}_{k=0}^{2N_1-1} =$

$$\sum_{n=1}^{N_1} c_n z_n(0) e^{-\lambda_n (D_{\max} - D)} e^{-\lambda_n k\Delta} + \varepsilon \sum_{n=N_1+1}^{N_1+N_2} c_n z_n(0) e^{-\lambda_n (D_{\max} - D)} e^{-\lambda_n k\Delta}.$$

Finally, we apply the ESPRIT algorithm to the measurement vector $\operatorname{col}\{y(t_k)\}_{k=0}^{2N_1-1}$ and recover $\{\hat{\lambda}_n^{ESP}\}_{n=1}^{N_1}$ directly. Then, since $c_n, z_n(0), D_{\max}$ are known, we can recover the approximation to D from the coefficients $\hat{y}_n^{ESP}$ as

$$D \approx \hat{D}_{ESP}^{(1)} := \frac{1}{\hat{\lambda}_1^{ESP}} \log \frac{\hat{y}_1^{ESP}}{c_1 z_1(0) e^{-\hat{\lambda}_1^{ESP} D_{\max}}}.$$

Whenever the argument of the logarithm is negative, we consider the reconstruction to be unsuccessful.

The errors $|\hat{\lambda}_1^{ESP} - \lambda_1|$ and $|\hat{D}_{ESP}^{(1)} - D|$ are shown in Fig. 4. The overall shape of the error curves is consistent with the theoretical predictions in Theorem 2 and the numerical conditioning in the previous section.

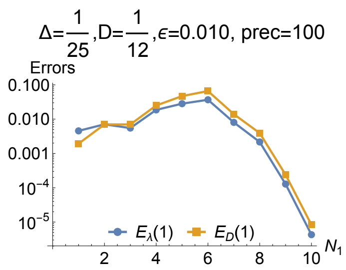


Fig. 4. Errors in λ_1 and the estimated delay.

REFERENCES

Balas, M.J. (1988). Finite-dimensional controllers for linear distributed parameter systems: exponential stability using residual mode filters. *Journal of Mathematical Analysis and Applications*, 133(2), 283–296.

Banks, H.T. and Kunisch, K. (2012). *Estimation techniques for distributed parameter systems*. Springer Science & Business Media.

Batenkov, D., Goldman, G., and Yomdin, Y. (2021). Super-resolution of near-colliding point sources. *Information and Inference: A Journal of the IMA*, 10(2), 515–572. doi: 10.1093/imaiai/iaaa005.

Christofides, P. (2001). *Nonlinear and Robust Control of PDE Systems: Methods and Applications to transport reaction processes*. Springer.

Curtain, R. (1982). Finite-dimensional compensator design for parabolic distributed systems with point sensors and boundary input. *IEEE Transactions on Automatic Control*, 27(1), 98–104.

Demetriou, M. and Rosen, I. (1994). Dynamic identification of implicit parabolic systems. *Lecture Notes in Pure and Applied Mathematics*, 153–153.

Fulton, C.T. and Pruess, S.A. (1994). Eigenvalue and eigenfunction asymptotics for regular Sturm-Liouville problems. *Journal of Mathematical Analysis and Applications*, 188(1), 297–340.

Harkort, C. and Deutscher, J. (2011). Finite-dimensional observer-based control of linear distributed parameter systems using cascaded output observers. *International Journal of Control*, 84(1), 107–122.

Istratov, A.A. and Vyvenko, O.F. (1999). Exponential analysis in physical phenomena. *Review of Scientific Instruments*, 70(2), 1233–1257. doi:10.1063/1.1149581.

Katz, R. and Fridman, E. (2021a). Finite-dimensional boundary control of the linear Kuramoto-Sivashinsky equation under point measurement with guaranteed L^2-gain. *IEEE Transactions on Automatic Control*, 67(10), 5570–5577.

Katz, R. and Fridman, E. (2021b). Finite-dimensional control of the heat equation: Dirichlet actuation and point measurement. *European Journal of Control*, 62, 158–164.

Katz, R. and Fridman, E. (2022). Delayed finite-dimensional observer-based control of 1D parabolic PDEs via reduced-order LMIs. *Automatica*, 142, 110341.

Kirsch, A. (2011). *An introduction to the mathematical theory of inverse problems*, volume 120. Springer.

Lowe, B.D., Pilant, M., and Rundell, W. (1992). The recovery of potentials from finite spectral data. *SIAM Journal on Mathematical Analysis*, 23(2), 482–504.

Nicolaenko, B. (1986). Some mathematical aspects of flame chaos and flame multiplicity. *Physica D: Nonlinear Phenomena*, 20(1), 109–121.

Orlov, Y. (2017). On general properties of eigenvalues and eigenfunctions of a Sturm-Liouville operator: comments on "ISS with respect to boundary disturbances for 1-D parabolic PDEs". *IEEE Transactions on Automatic Control*, 62(11), 5970–5973.

Pereyra, V. and Scherer, G. (2010). *Exponential Data Fitting and Its Applications*. Bentham Science Publishers.

Roy, R. and Kailath, T. (1989). ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7), 984–995.

Rundell, W. and Sacks, P.E. (1992). Reconstruction techniques for classical inverse Sturm-Liouville problems. *Mathematics of Computation*, 58(197), 161–183.

Sivashinsky, G. (1977). Nonlinear analysis of hydrodynamic instability in laminar flames–I. Derivation of basic equations. *Acta Astronautica*, 4, 1177–1206.

Spivak, M. (2018). *Calculus on manifolds: a modern approach to classical theorems of advanced calculus*. CRC press.

Appendix A. LAGRANGE AND HERMITE INTERPOLATION

Consider a set of distinct nodes $\chi = \{\chi_n\}_{n\in[S]} \subseteq \mathbb{R}$ and values $\{f_n\}_{n\in[S]} \subseteq \mathbb{R}$. The classical Lagrange interpolation problem seeks a polynomial $q(z)$, $\deg(q) \le S-1$ satisfying $q(\chi_n) = f_n$ for all $n \in [S]$. It is well known that the solution to the problem is given by $q(z) = \sum_{n\in[S]} f_n L_{\chi_n}(z)$, where the Lagrange interpolation basis is

$$L_{\chi,n}(z) = \prod_{j\neq n} \frac{z-\chi_j}{\chi_n-\chi_j}, \quad n \in [S]. \tag{A.1}$$

Given the Lagrange interpolation basis (A.1), one can construct the corresponding Hermite interpolation basis is

$$\begin{aligned} H_{\chi,n}(z) &= \left[1-2(z-\chi_n)L'_{\chi,n}(\chi_n)\right] L^2_{\chi,n}(z), \\ \widetilde{H}_{\chi,n}(z) &= (z-\chi_n)L^2_{\chi,n}(z), \quad n \in [S]. \end{aligned} \tag{A.2}$$

The Hermite basis is related to this interpolation problem: introducing further $\{f'_n\}_{n\in[S]}$, one seeks a polynomial $r(z)$, $\deg(r) \le 2S-1$ satisfying $r(\chi_n) = f_n$, $r'(\chi_n) = f'_n$ for all $n \in [S]$. The solution is given by $r(z) = \sum_{n\in[S]} \left(f_n H_{\chi,n}(z) + f'_n \tilde{H}_{\chi,n}(z)\right)$.

An alternative solution to the Hermite interpolation problem can be formulated as follows. Given polynomial $p(z) = \sum_{j=0}^{2S-1} a_j z^j$, we introduce the coordinate map

$$\mathfrak{C}(p) = [a_0, \dots, a_{2S-1}]^\top.$$

Then, the Hermite interpolating polynomial $r(z)$ satisfies

$$\begin{aligned} W_\chi \mathfrak{C}(r) &= \operatorname{col}\left\{ \begin{bmatrix} f_n \\ f'_n \end{bmatrix} \right\}_{n=1}^S, \\ W_\chi &:= \operatorname{col}\left\{ \begin{bmatrix} 1 & \chi_n & \chi_n^2 & \dots & \chi_n^{2S-1} \\ 0 & 1 & 2\chi_n & \dots & (2S-1)\chi_n^{2S-2} \end{bmatrix} \right\}_{n=1}^S \end{aligned} \tag{A.3}$$

The matrix W_χ maps $\mathfrak{C}(r)$ to the values of $p(z)$ and $p'(z)$ at the interpolation nodes $\{\chi_n\}_{n=1}^S$. Since the Hermite interpolation problem is always solvable, W_χ is invertible. Moreover, it can be verified that

$$\mathbb{H}_S(\chi) := W_\chi^{-\top} = \left(\operatorname{row}\left\{\left[\mathfrak{C}(H_{\chi,n}) \ \ \mathfrak{C}(\tilde{H}_{\chi,n})\right]\right\}_{n=1}^S\right)^\top \in \mathbb{R}^{2S\times 2S}$$

is the unique matrix satisfying

$$\mathbb{H}_S(\chi) \operatorname{col}\left\{\zeta^j\right\}_{j=0}^{2S-1} = \operatorname{col}\left\{ \begin{bmatrix} H_{\chi,n}(\zeta) \\ \tilde{H}_{\chi,n}(\zeta) \end{bmatrix} \right\}_{n=1}^S \tag{A.4}$$

for all $\zeta \in \mathbb{R}$.

Appendix B. INTEGRAL CONVERGENCE

Proposition 2. Let $0 \le w_1 < w_2 \le \infty$ and define

$$\mathscr{Q}_\alpha(x) = -\log\left(1-e^{-\alpha x}\right) > 0, \ x \in (0,\infty),$$

where $\alpha > 0$. The integrals

$$\mathscr{I}_{w_1,w_2}(\alpha) := \int_{w_1}^{w_2} \mathscr{Q}_\alpha(x)\mathrm{d}x \tag{B.1}$$

are finite, decreasing, and $\lim_{\alpha\to\infty} \mathscr{I}_{w_1,w_2}(\alpha) = 0$.

Proof. We prove the result for $w_1 = 1$ and $w_2 = \infty$ (other cases are similar). Integrating by parts, we have

$$\mathscr{I}_{1,\infty}(\alpha) = \left[-x\log\left(1-e^{-\alpha x}\right)\right]_1^\infty + \alpha\int_1^\infty \frac{x}{1-e^{-\alpha x}} e^{-\alpha x}\mathrm{d}x.$$

The first term on the right-hand side has

$$\lim_{x\to\infty} x\log\left(1-e^{-\alpha x}\right) = -\alpha \lim_{x\to\infty} \frac{x^2}{1-e^{-\alpha x}} e^{-\alpha x} = 0,$$

whereas

$$0 < \int_1^\infty \frac{x}{1-e^{-\alpha x}} e^{-\alpha x}\mathrm{d}x \le \frac{e^\alpha}{e^\alpha-1}\int_1^\infty x e^{-\alpha x}\mathrm{d}x < \infty.$$

Hence, $\mathscr{I}_{1,\infty}(\alpha) < \infty$.

Next, for a fixed $x \in (0,\infty)$,

$$(0,\infty) \ni \alpha \mapsto -\log\left(1-e^{-\alpha x}\right) \in (0,\infty)$$

is decreasing, whence $\mathscr{I}_{1,\infty}(\alpha)$ is also decreasing. Let $\varepsilon > 0$, $\alpha > 1$ and $M > 1$. Then,

$$\begin{aligned} \mathscr{I}_{1,\infty}(\alpha) &\le \mathscr{I}_{1,M}(\alpha) + \mathscr{I}_{M,\infty}(1) \\ &= \int_1^M -\log\left(1-e^{-\alpha x}\right)\mathrm{d}x + \int_M^\infty -\log\left(1-e^{-x}\right)\mathrm{d}x. \end{aligned}$$

Choosing M so that the rightmost integral is smaller than $\frac{\varepsilon}{2}$ and then α_* such that $\mathscr{I}_{1,M}(\alpha) < \frac{\varepsilon}{2}$ for $\alpha > \alpha_*$, we have that $\alpha > \min(1,\alpha_*)$ implies $0 < \mathscr{I}_{1,\infty}(\alpha) < \varepsilon$.