IFAC Journal of Systems and Control 35 (2026) 100350

Contents lists available at ScienceDirect

IFAC Journal of Systems and Control

journal homepage: www.elsevier.com/locate/ifacsc

Full length article

Identification of reaction–diffusion systems from finitely many non-local noisy measurements via exponential fitting[☆,☆☆]

Rami Katz[a,*], Giulia Giordano[b], Dmitry Batenkov[c,d]

[a] School of Electrical and Computer Engineering, Tel Aviv University, Tel Aviv, Israel
[b] Department of Industrial Engineering, University of Trento, Trento, Italy
[c] Basis Research Institute, New York City, NY, USA
[d] Department of Applied Mathematics, School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel

ARTICLE INFO

Article history:
Received 15 August 2025
Received in revised form 5 November 2025
Accepted 30 November 2025
Available online 3 December 2025

Keywords:
System identification
Distributed parameter systems
Estimation
Exponential fitting
Partial differential equations
Reaction–diffusion systems
Reconstruction

ABSTRACT

Given a reaction–diffusion equation with unknown right-hand side, we consider the nonlinear inverse problem of estimating the associated leading eigenvalues and initial condition Fourier coefficients from a finite number of non-local noisy measurements. We define a reconstruction (i.e., estimation) criterion and, for small enough noise, we prove existence and uniqueness of the desired estimates. We derive closed-form expressions for the first-order condition numbers and bounds for their asymptotic behavior in a regime when the number of measured samples is fixed and the inter-sampling interval length is arbitrarily large. When computing the sought estimates numerically, our simulations show that the exponential fitting algorithm ESPRIT is first-order optimal, since its first-order condition numbers have the same asymptotic behavior as the analytic condition numbers in the considered regime.

1. Introduction

Reaction–diffusion equations (RDEs), which belong to the class of distributed parameter systems, are widely used to model phenomena in physics and engineering, including magnetized plasma, flame front propagation and chemical processes (Nicolaenko, 1986; Sivashinsky, 1977). Observation and control of RDEs has been investigated over the last decades (Balas, 1988; Harkort & Deutscher, 2011; Katz & Fridman, 2022), often by employing approaches based on modal decomposition (Christofides, 2001; Curtain, 1982; Katz & Fridman, 2021a). Existing control and observation techniques typically assume explicit knowledge of the spatial operator of the system or of the eigenvalue/eigenfunction pairs corresponding to its modes.

Estimation of unknown parameters in RDEs is a challenging problem, which has been mostly studied in an adaptive estimation framework (Banks & Kunisch, 2012; Demetriou & Rosen, 1994). Adaptive estimation relies on the assumption of *persistency of excitation* of a system input (Willems et al., 2005), 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 still an open problem. Other identification methods, accompanied by sound numerical algorithms, have been derived in the field of inverse problems (Kirsch, 2011; Lowe et al., 1992; Rundell & Sacks, 1992). These approaches aim at recovering the spatial operator of the system under the assumption of *complete knowledge* of its eigenvalues. However, this assumption is often non-realistic from a control theory perspective, where only discrete-time measurements of the state are available as datum. Hence, constructive and implementable identification techniques for RDEs are still a very active area of research.

We propose a novel approach inspired by the algebraic super-resolution literature (Batenkov et al., 2021), where *estimation* of unknown parameters goes under the name of *reconstruction*, while the obtained *estimates* are termed *approximations*, and we provide the following contributions:

- Differently from existing adaptive estimation algorithms, which assume the availability of a controlled input that

☆ This article is part of a Special issue entitled: 'IFAC WC 2026 - TC1.1' published in IFAC Journal of Systems and Control.

☆☆ R. Katz was supported by the Alon Fellowship from the Council of Higher Education in Israel. G. Giordano acknowledges support from 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 partially supported by Israel Science Foundation Grant 1793/20.

* Corresponding author.
E-mail addresses: ramkatsee@tauex.tau.ac.il (R. Katz), giulia.giordano@unitn.it (G. Giordano), dima@basis.ai (D. Batenkov).

https://doi.org/10.1016/j.ifacsc.2025.100350
2468-6018/

excites the system, and of measurements either for all $t \geq 0$ or at an infinite sequence of high-frequency discrete-time samples $\{t_k\}_{k=1}^{\infty}$, we assume the measurements are available at *finitely many* time steps. We also assume that the measurements are affected by structured noise, due to the unintentional measurement of 'undesirable' modes.
- We show that parameter estimation in RDEs can be cast as an *exponential fitting problem*, i.e., the recovery of the parameters $\left\{\alpha_j, \beta_j\right\}_{j=1}^{M} \subseteq \mathbb{C}$ from measurements of the form $\mu_k = \sum_{j=1}^{M} \alpha_j \beta_j^k$, $k = 1, \ldots K$, which is a classical topic in data analysis with numerous applications (Batenkov et al., 2021; Istratov & Vyvenko, 1999; Pereyra & Scherer, 2010). We then define a reconstruction criterion for the problem in the presence of structured noise, and prove the existence and uniqueness of the associated approximation for small enough noise intensity (see Theorem 1).
- We introduce first-order condition numbers that describe how the structured noise is amplified in the reconstruction procedure, thereby affecting the accuracy of the obtained parameter estimates. Then, we explicitly express the introduced condition numbers as linear combinations of interpolating Hermite polynomials.
- To investigate the dependence of the considered inverse problem on its parameters, we study the asymptotic behavior of the derived condition numbers in a specific parameter regime, and obtain rigorous asymptotic bounds on them. These bounds shed light on the sensitivity of the inverse problem with respect to its underlying parameters.
- Finally, we consider the problem of algorithmically computing the approximations (i.e., the estimates). Our numerical simulations show that the well-known ESPRIT algorithm (Roy & Kailath, 1989) achieves first-order optimality, meaning that the first-order condition numbers of the ESPRIT algorithm exhibit the same asymptotic behavior as the analytic condition numbers, in the considered regime. These numerical simulations validate our theoretical analysis and constitute a basis for a future algorithmic study of the considered reconstruction problem.

Our results, which for the first time analyze a parameter estimation problem for RDEs, in the presence of structured noise, via exponential fitting, provide a novel methodology that paves the way towards new directions in identification of RDEs.

Notation: Given $N \in \mathbb{N}$, $[N] = \{1, \ldots, N\}$. For $a \in \mathbb{R}$, $\lfloor a \rfloor = \max\{m \in \mathbb{Z}\,;\ m \leq a\}$. For matrices $\{A_i\}_{i=1}^k$ of appropriate size, diag $\{A_i\}_{i=1}^k$ is the block-diagonal matrix with the A_i's on the diagonal, whereas row $\{A_i\}_{i=1}^k$ (resp. col $\{A_i\}_{i=1}^k$) is the block matrix stacking the A_i's in consecutive columns (resp. rows). $L^2(0,1)$ is the Hilbert space of square-integrable scalar-valued functions on $(0,1)$ with inner product $\langle\cdot,\cdot\rangle$, while $\mathrm{H}^2(0,1)$ (resp. $\mathrm{H}_0^1(0,1)$) is the Sobolev space of functions f 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 $f(0) = f(1) = 0$).

2. Model and estimation problem

We consider the linear one-dimensional reaction–diffusion equation in Fickian form, with Dirichlet boundary conditions:

$$\begin{aligned} &z_t(x,t) = (p(x)z_x(x,t))_x + q(x)z(x,t), \\ &z(0,t) = 0, \quad z(1,t) = 0, \end{aligned} \tag{1}$$

along with a non-local measurement signal

$$y(t) = \int_0^1 c(x)z(x,t)\mathrm{d}x, \quad t \geq 0. \tag{2}$$

Here $x \in (0,1)$, $z(x,t) \in \mathbb{R}$, $p(x)$ and $q(x)$ are *unknown* smooth functions defined on $[0,1]$, and $\min_{x\in[0,1]} p(x) = \underline{p}$ for some *unknown* $\underline{p} > 0$. Finally, we assume that $c \in L^2(0,1)$.

Remark 1. The positive lower bound $\underline{p}$ on $p(x)$ is assumed for well-posedness only. However, knowledge of $\underline{p}$ is not needed for the parameter estimation problem below.

Before presenting in detail our parameter estimation problem, we briefly recall some properties of (1). The system can be associated with the Sturm–Liouville operator $\mathscr{A}$, defined as

$$\begin{aligned} &[\mathscr{A}h](x) = -\left(p(x)h'(x)\right)' - q(x)h(x), \quad x \in (0,1), \\ &\mathrm{Dom}(\mathscr{A}) = \left\{h \in \mathrm{H}^2(0,1)\,;\ h(0) = h(1) = 0\right\}. \end{aligned} \tag{3}$$

The operator $\mathscr{A}$ has an infinite and strictly monotonically increasing sequence of simple eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ that satisfy $\lim_{n\to\infty} \lambda_n = \infty$ (Orlov, 2017). The corresponding eigenvectors $\{\psi_n\}_{n=1}^{\infty}$ form a complete orthonormal basis in $L^2(0,1)$. The well-posedness of system (1) has been studied thoroughly (see, e.g., the arguments in Katz & Fridman, 2021b). In particular, given $z(\cdot,0) \in L^2(0,1)$, system (1) has a unique solution

$$z \in C([0,\infty), L^2(0,1)) \cap C^1((0,\infty), L^2(0,1))$$

such that $z(\cdot,t) \in \mathrm{Dom}(\mathscr{A})$ for all $t > 0$ (see Pazy, 1983 for more details). The solution of (1) can be presented as

$$z(x,t) = \sum_{n=1}^{\infty} z_n(t)\psi_n(x), \quad z_n(t) = \langle z(\cdot,t), \psi_n\rangle\,,\ n \in \mathbb{N}. \tag{4}$$

Differentiating under the integral sign and integrating by parts, we obtain $\dot{z}_n(t) = -\lambda_n z_n(t)$ for all $n \in \mathbb{N}$, and hence

$$z(x,t) = \sum_{n=1}^{\infty} z_n(0)e^{-\lambda_n t}\psi_n(x). \tag{5}$$

Substituting (5) into the measurement signal (2) yields

$$y(t) = \sum_{n=1}^{\infty} c_n z_n(0)e^{-\lambda_n t}, \quad c_n = \langle c, \psi_n\rangle\,,\ n \in \mathbb{N}. \tag{6}$$

Remark 2. The solution (5) to (1) is not assumed to be exponentially stable; in fact, for $\min_{x\in[0,1]} q(x)$ large enough, it may contain (finitely) many unstable eigenvalues among $\{\lambda_n\}_{n=1}^{\infty}$.

Accurate estimates of the eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ and of the Fourier coefficients $\{z_n(0)\}_{n=1}^{\infty}$ of the initial condition $z(\cdot,0) \in L^2(0,1)$ are essential both for control applications and for estimating the operator $\mathscr{A}$ using methods from inverse problems (Kirsch, 2011; Lowe et al., 1992; Rundell & Sacks, 1992). Hence, our first goal (**Objective 1**) is to define a reconstruction criterion to estimate a subset of $\{\lambda_n\}_{n=1}^{\infty}$ and $\{z_n(0)\}_{n=1}^{\infty}$ from measurements $\{y(t_k)\}_{k=0}^{2N_1-1}$ of signal (6), available at a finite discrete set of times $\{t_k\}_{k=0}^{2N_1-1}$, in a scenario where:

- the Sturm–Liouville operator $\mathscr{A}$ in (3) and the initial condition $z(\cdot,0) \in L^2(0,1)$ are both unknown, whence also the associated eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ and Fourier coefficients $\{z_n(0)\}_{n=1}^{\infty}$ in (6) are unknown;
- the measurement kernel $c(x)$ in (2) is band-limited, i.e., there exist $N_1, N_2 \in \mathbb{N}$ such that $c_n = 0$ for all $n \geq N_1 + N_2$; also, the coefficients c_n with $n \in [N_1]$ are known and are much larger in magnitude than c_n with $n \in [N_1 + N_2] \setminus [N_1]$, as detailed precisely in Assumption 1 below.

Given the specified reconstruction criterion, our second goal (**Objective 2**) is to quantify the achieved estimation accuracy by

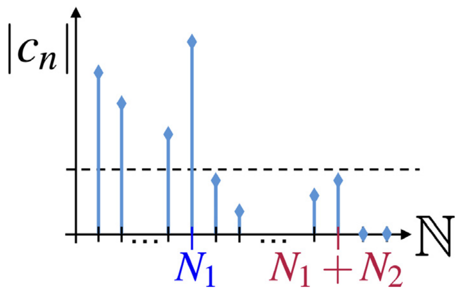


Fig. 1. The spectral structure of $c(x)$ in Assumption 1.

rigorously analyzing the associated first-order condition numbers, which capture the amplification of the measurement noise through the reconstruction procedure. In particular, we derive theoretical bounds for the asymptotic behavior of the condition numbers. Our third goal (**Objective 3**) is to show that the theoretical bounds can be numerically achieved; as shown by our numerical simulations in Section 5, they are indeed exhibited by the ESPRIT method (Roy & Kailath, 1989).

We now provide a formal description of the estimation problem and of its underlying assumptions, so as to formally state our objectives.

Assumption 1. There exist $N_1, N_2 \in \mathbb{N}$ such that the coefficients $\{c_n\}_{n=1}^{\infty}$ satisfy the following properties:

(I) for all $n \notin [N_1+N_2]$, $c_n = 0$;
(II) for all $n \in [N_1]$, c_n is *known* and satisfy $\ell^{-1} < |c_n| < \ell$, where $\ell > 0$ is a fixed constant;
(III) for all $n \in [N_1+N_2] \setminus [N_1]$, we can write $c_n = \epsilon \tilde{c}_n$, where $\epsilon > 0$ is unknown and $\tilde{c}_n$ are such that, for some fixed $M_c > 0$, $\frac{|\tilde{c}_n|}{|c_k|} \le M_c$ for all $k \in [N_1]$; therefore, $|c_n| < \epsilon M_c \ell$.

Part (I) of Assumption 1 poses no loss of generality. Since $c \in L^2(0,1)$, its Fourier coefficients decay to zero; hence, there exists some N_* such that $|c_n|$ is smaller than machine precision for all $n \ge N_*$ and will be rounded to zero in any computer-based estimation algorithm. Alternatively, such an assumption can be relaxed entirely, as discussed after Theorem 1 below. In part (II), the assumption that c_n with $n \in [N_1]$ are known could be removed: if both $\{c_n\}_{n=1}^{\infty}$ and $\{z_n(0)\}_{n=1}^{\infty}$ in (6) were unknown, our approach could still estimate the products $c_n z_n(0)$, and further information would be required to distinguish between the constituent Fourier coefficients. In part (III), ϵ represents the intensity of structured noise affecting the measurements: the majority of the energy of $c(x)$ is concentrated on the first N_1 modes, whereas the remaining $N_2 - N_1$ modes of $c(x)$ constitute a smaller "side-lobe" that introduces structured noise into the reconstruction problem; see (7) and (8) below. The spectral structure of $c(x)$ is visualized in Fig. 1.

Assumption 2. The initial condition $z(\cdot,0) \in L^2(0,1)$ is *unknown*. For all $n \in [N_1]$, $\ell_1^{-1} < |z_n(0)| < \ell_1$ for some fixed $\ell_1 > 0$. Furthermore, for all $n \in [N_1+N_2] \setminus [N_1]$, for some fixed $M_z > 0$, $\frac{|z_n(0)|}{|z_k(0)|} \le M_z$ for all $k \in [N_1]$, and hence $|z_n(0)| < M_z \ell_1$.

The boundedness of the ratios in Assumption 2 is guaranteed by the decay of the Fourier coefficients of $z(\cdot,0) \in L^2(0,1)$. Assumption 2 can be viewed as a *persistency of excitation* assumption for the initial condition, which guarantees that the modes one is interested in recovering are present in the sampled signal. This is different from assuming persistency of excitation for system inputs (see, e.g., Willems et al., 2005), which does not apply in our case, since system (1) is uncontrolled.

Assumption 3. Let $\Delta_* > 0$ be fixed and $\Delta \in [\Delta_*, \infty)$. The system measurements are available *only at a finite set of times* $\{t_k\}_{k=0}^{2N_1-1} := \{k\Delta\}_{k=0}^{2N_1-1}$, with step-size Δ.

Assumption 3 reflects the realistic set-up in which the measurement signal $y(t)$ is sampled in discrete-time with some minimal sampling step-size, and only finitely many ($2N_1$) snapshots (i.e., discrete-time measurements) of the signal are available for identification purposes.

Subject to Assumptions 1–3, the measurements of signal (6) at the available times $\{t_k\}_{k=0}^{2N_1-1}$ can be written as

$$y(t_k) = \underbrace{\sum_{n=1}^{N_1} y_n e^{-\lambda_n k\Delta}}_{y_{\text{main}}(t_k)} + \epsilon \underbrace{\sum_{n=N_1+1}^{N_1+N_2} y_n e^{-\lambda_n k\Delta}}_{y_{\text{tail}}(t_k)} \tag{7}$$

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

$$y_n := \begin{cases} c_n z_n(0), & n \in [N_1] \\ \tilde{c}_n z_n(0), & n \in [N_1+N_2] \setminus [N_1] \end{cases} \tag{8}$$

satisfy $\frac{|y_n|}{|y_k|} \le M_c M_z =: M_y$ for all $n \in [N_1+N_2] \setminus [N_1]$, with $k \in [N_1]$; in the measurements $y(t_k)$, components $y_{\text{main}}(t_k)$ contain $z_n(0) = \frac{y_n}{c_n}$ and λ_n, which we wish to estimate, while $y_{\text{tail}}(t_k)$ are structured measurement noise of intensity $\epsilon > 0$, stemming from the "side-lobe" of the measurement kernel $c(x)$.

We are now ready to state our objectives. We denote

$$\mathcal{N}_\eta = \left\{(n, N_1) \in \mathbb{N}^2 \; ; \; n \in [\lfloor \eta N_1 \rfloor]\right\}, \tag{9}$$

where $\eta \in (0,1]$ regulates the fraction of the parameters $\{\lambda_n\}_{n=1}^{N_1}$ and $\{y_n\}_{n=1}^{N_1}$ which we are interested in estimating.

Objective 1: Under Assumptions 1–3: (i) Formulate a reconstruction criterion to obtain estimates $\{\hat{y}_n, \hat{\lambda}_n\}$ of $\{y_n, \lambda_n\}$, with $(n, N_1) \in \mathcal{N}_1$, from the measurement data in (7). Given $\{\hat{y}_n, \hat{\lambda}_n\}$, since we know c_n by Assumption 1, we can estimate from (8) the projection coefficients of the initial condition $z(\cdot,0)$ via

$$\hat{z}_n(0) = \frac{\hat{y}_n}{c_n}, \quad n \in [N_1].$$

(ii) Show that the reconstruction criterion is well-posed (i.e., the estimates exist and are unique) for small noise intensity $\epsilon > 0$. (iii) Prove that the (first-order) condition numbers in ϵ, i.e. the first-order expansions of the reconstruction errors with respect to ϵ for λ_n and for y_n, can be defined as

$$\mathcal{K}_\lambda(n) := \lim_{\epsilon \to 0^+} \frac{\hat{\lambda}_n(\epsilon) - \lambda_n}{\epsilon}, \quad \mathcal{K}_y(n) := \lim_{\epsilon \to 0^+} \frac{\hat{y}_n(\epsilon) - y_n}{\epsilon}, \tag{10}$$

and derive their explicit expression.

Objective 2: Analyze the asymptotic behavior of $|\mathcal{K}_\lambda(n)|$ and $\left|\mathcal{K}_y(n)\right|$, for $(n, N_1) \in \mathcal{N}_\eta$, in the parameter regime with

$$N_1, N_2 \text{ fixed} \quad \text{and} \quad \Delta \gg 0. \tag{11}$$

Objective 3: Identify a numerical reconstruction algorithm achieving the reconstruction guarantees derived in **Objective 2**.

Reformulating the measurements as in (7) recasts our parameter estimation task into an exponential fitting problem. The task is highly challenging for two reasons. First, only *finitely many* measurements are available for the reconstruction procedure. Second, although (1) is a linear system, estimating $\{y_n, \lambda_n\}_{n=1}^{N_1}$ from the measurements (7) is a nonlinear inverse problem, as the measurements depend nonlinearly on the parameters.

Remark 3. For a *given* set of initial condition and kernel modes $(\{z_n(0)\}, \{c_n\})$, appropriate constants M_c and M_z as in Assumptions 1 and 2 can always be found. However, our problem formulation allows the initial condition and kernel modes $(\{z_n(0)\}, \{c_n\})$

to vary within a large set $\mathscr{I} \times \mathscr{M}$, consisting of *all pairs* satisfying Assumptions 1 and 2 for some fixed ℓ, ℓ_1, M_c and M_z. Our proposed estimation approach, as well as the derived upper bounds on the condition numbers, hold *uniformly for all* $(\{z_n(0)\}, \{c_n\})$ satisfying Assumptions 1 and 2.

3. Reconstruction criterion and its analysis

To find a reconstruction criterion (**Objective 1**), we define the forward map $\mathscr{F}: \mathbb{R}^{2N_1+1} \to \mathbb{R}^{2N_1}$ as $\mathscr{F} = \mathrm{col}\,\{\mathscr{F}_k\}_{k=0}^{2N_1-1}$, with

$$\mathscr{F}_k\left(\left\{\hat{y}_n, \hat{\lambda}_n\right\}_{n\in[N_1]}, \epsilon\right) = \sum_{n=1}^{N_1} \hat{y}_n e^{-\hat{\lambda}_n k\Delta} - y(t_k), \tag{12}$$

where $k = 0, \ldots, 2N_1 - 1$. Each component $\mathscr{F}_k$ takes as input an *approximation candidate* $\hat{P} := \left\{\hat{y}_n, \hat{\lambda}_n\right\}_{n\in[N_1]}$ and a noise intensity ϵ, and computes the difference between the *approximate main* term $\sum_{n=1}^{N_1} \hat{y}_n e^{-\hat{\lambda}_n k\Delta}$, corresponding to measurement instant $t_k = k\Delta$, and the measurement $y(t_k) = y_{\mathrm{main}}(t_k) + \epsilon y_{\mathrm{tail}}(t_k)$ in (7); the latter introduces the ϵ-structured noise into $\mathscr{F}_k$.

In the absence of noise (i.e., when $\epsilon = 0$), if $\hat{P} = P := \{y_n, \lambda_n\}$, then $\mathscr{F}_k(P, 0) = 0$ for all $k = 0, \ldots, 2N_1 - 1$. As a reconstruction criterion, we therefore propose to seek an approximation $\hat{P}$ that satisfies $\mathscr{F}\left(\hat{P}, \epsilon\right) = 0$ also in the presence of structured noise (i.e., when $\epsilon > 0$), as per the following definition.

Definition 1. The approximation candidate $\hat{P} = \left\{\hat{y}_n, \hat{\lambda}_n\right\}_{n\in[N_1]}$ is an *ϵ-approximation* of $P = \{y_n, \lambda_n\}_{n\in[N_1]}$ if $\mathscr{F}\left(\hat{P}, \epsilon\right) = 0$.

Given the above reconstruction criterion, we now state our main results related to **Objectives 1** and **2**; all the proofs are reported in Section 4. We first show that an ϵ-approximation $\hat{P}$ always exists, and is unique, provided that the noise intensity $\epsilon > 0$ is small enough. We also justify the definition of the (first-order) condition numbers $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$, with $n \in [N_1]$, provided in (10). Throughout the paper, we use the notation

$$\phi_n := e^{-\lambda_n \Delta}, \quad n \in [N_1 + N_2]. \tag{13}$$

Theorem 1. *There exist $\epsilon_* > 0$ and continuously differentiable functions $\hat{y}_n, \hat{\lambda}_n : (-\epsilon_*, \epsilon_*) \to \mathbb{R}$ such that the approximation candidate $\hat{P}(\epsilon) = \left\{\hat{y}_n(\epsilon), \hat{\lambda}_n(\epsilon)\right\}_{n\in[N_1]}$ satisfies $\hat{P}(0) = P$ and*

$$\forall \epsilon \in (-\epsilon_*, \epsilon_*) \quad \mathscr{F}(\hat{P}, \epsilon) = 0 \iff \hat{P} = \hat{P}(\epsilon). \tag{14}$$

Furthermore, the condition numbers $\mathscr{K}_y(n)$ and $\mathscr{K}_\lambda(n)$ exist as defined in (10) and can be explicitly expressed as

$$\begin{bmatrix} \mathscr{K}_y(n) \\ \mathscr{K}_\lambda(n) \end{bmatrix} = \begin{bmatrix} \frac{\mathrm{d}\hat{y}_n}{\mathrm{d}\epsilon}(0) \\ \frac{\mathrm{d}\hat{\lambda}_n}{\mathrm{d}\epsilon}(0) \end{bmatrix} = \sum_{m=N_1+1}^{N_1+N_2} y_m \begin{bmatrix} H_n(\phi_m) \\ -\frac{1}{\Delta y_n \phi_n} \tilde{H}_n(\phi_m) \end{bmatrix}, \tag{15}$$

for $n \in [N_1]$, where $\left\{H_n, \tilde{H}_n\right\}_{n=1}^{N_1}$ are the Hermite interpolation basis polynomials associated with $\{\phi_n\}_{n=1}^{N_1+N_2}$.

Theorem 1 fulfills **Objective 1**. The necessary background on Hermite interpolation will be given in Section 4.1.

Remark 4. By Assumption 1, $c_n = 0$ for $n \notin [N_1 + N_2]$, meaning that c is band-limited; this assumption is justified from a practical and numerical perspective because of modal decay and numerical precision. From a theoretical perspective, our approach and its analysis can be extended to the case of non-band-limited $c \in L^2$. In fact, in such a scenario, the tail in (7) includes infinitely many terms, the Fourier coefficients $z_n(0)$ (whence also y_n) decay to zero, with rate higher or equal to $\frac{1}{n^2}$, and this allows one to show that all of the obtained series converge. For example, the right-hand side of (15) includes a series of vectors of Hermite polynomials H_n and $\tilde{H}_n$, evaluated at $\{\phi_m\}_{m=N_1+1}^{\infty}$. Then, combining the arguments of Theorem 2 below with the decay of $\{y_n\}_{n=1}^{\infty}$, one can show that our estimates on the condition numbers still hold for non-band-limited c. To avoid the extra technicalities introduced by series convergence, we proceed with our theoretical derivations under Assumption 1 and thus assume that c is band-limited. However, the numerical simulations in Fig. 6 demonstrate that our approach can handle the case of non-band-limited c.

The presentation (15) also allows us to establish the asymptotic behavior of the (first-order) condition numbers $\mathscr{K}_y(n)$ and $\mathscr{K}_\lambda(n)$, where $(n, N_1) \in \mathscr{N}_\eta$, $\eta \in (0, 1)$, in the regime (11) through the analysis of the corresponding Hermite interpolating polynomials, thereby fulfilling **Objective 2**.

Theorem 2. *Consider $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$ defined in (10), with $N_2 = 1$. Let $\eta \in (0, 1)$ and $(n, N_1) \in \mathscr{N}_\eta$. Then, in the regime (11), given any $\vartheta > 0$, there exist $N_1(\vartheta) \in \mathbb{N}$ and $\gamma_y(\eta, \vartheta), \gamma_\lambda(\eta, \vartheta) > 0$ such that, for all $N_1 \geq N_1(\vartheta)$,*

$$\begin{aligned} &\max_{n\in[\lfloor \eta N_1\rfloor]} \left|\mathscr{K}_y(n)\right| \leq \gamma_y(\eta, \vartheta) e^{-\vartheta\Delta}, \\ &\max_{n\in[\lfloor \eta N_1\rfloor]} \left|\mathscr{K}_\lambda(n)\right| \leq \frac{\gamma_\lambda(\eta, \vartheta)}{\Delta} e^{-\vartheta\Delta}, \end{aligned} \tag{16}$$

for $\Delta \gg \Delta_*$. In particular, the condition numbers $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$ decay exponentially with the step-size Δ.

Remark 5. We consider $\eta \in (0, 1)$ because, when $\eta = 1$, the proof of Theorem 2 can be shown to only yield $\mathscr{K}_\lambda(N_1) = O(1)$ and $\mathscr{K}_y(N_1) = O(1)$, uniformly in $N_1 \in \mathbb{N}$ and $\Delta \geq \Delta_* > 0$. This is further validated by our numerical experiments in Fig. 2(a).

4. Auxiliary results and proofs

Here, we provide the proofs of our main results, along with the needed preliminaries and auxiliary results.

4.1. Interpolation polynomials and proof of Theorem 1

We rely on interpolation theory (Quarteroni et al., 2006). Given a polynomial $q(z) = a_{2N_1-1}z^{2N_1-1} + \cdots + a_1 z + a_0$, we introduce the coordinate map

$$\mathscr{C}: q \mapsto \mathscr{C}_q := \mathrm{col}\left\{a_j\right\}_{j=0}^{2N_1-1}, \tag{17}$$

where $\mathscr{C}_q$ stacks the polynomial coefficients. Given distinct nodes $\{\chi_n\}_{n=1}^{N_1} \subseteq \mathbb{C}$ and the values $\left\{\vartheta_n, \vartheta_n'\right\}_{n=1}^{N_1} \subseteq \mathbb{C}$, there exists a unique polynomial $q(z)$ of degree at most $2N_1 - 1$ such that $q(\chi_n) = \vartheta_n$ and $q'(\chi_n) = \vartheta_n'$ for all $n \in [N_1]$. The coefficients of $q(z)$ satisfy the Hermite interpolation condition

$$\begin{aligned} &\mathscr{H}_{\{\chi_n\}}\mathscr{C}_q = \mathrm{col}\left\{\begin{bmatrix} \vartheta_n \\ \vartheta_n' \end{bmatrix}\right\}_{n=1}^{N_1}, \\ &\mathscr{H}_{\{\chi_n\}}^{\top} := \mathrm{row}\left\{\begin{bmatrix} 1 & 0 \\ \chi_j & 1 \\ \vdots & \vdots \\ \chi_j^{2N_1-1} & (2N_1-1)\chi_j^{2N_1-2} \end{bmatrix}\right\}_{j=1}^{N_1}, \end{aligned} \tag{18}$$

where $\mathscr{H}_{\{\chi_n\}}$ maps the coefficients of a polynomial to its values and the values of its derivative on the points $\{\chi_n\}_{n=1}^{N_1}$. The existence and uniqueness of the solution to the Hermite interpolation problem imply the invertibility of $\mathscr{H}_{\{\chi_n\}}$, whence $\mathscr{H}_{\{\chi_n\}}^{-1}$ maps the values $\left\{\vartheta_n, \vartheta_n'\right\}_{n=1}^{N_1}$ to the coefficients of the corresponding Hermite interpolation polynomial $q(z)$.

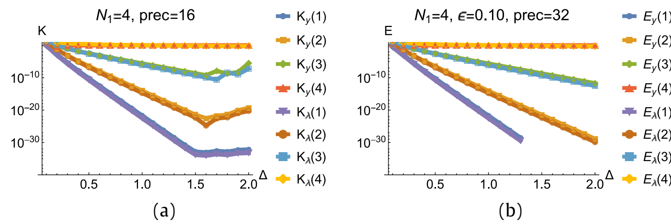


Fig. 2. 2(a) $\mathscr{K}_\lambda$, $\mathscr{K}_y$ for $N_1 = 4$ and $\lambda_n = n^2$. The asymptotics break down for large Δ, due to inversion of badly conditioned matrices in finite precision computations (16 decimal digits). 2(b) ESPRIT algorithm rescaled errors in (43), applied to the sequence $\{y(\Delta k)\}_{k=0}^{2N_1-1}$ with $N_1 = 4$ and $\lambda_n = n^2$. Here we used 32 decimal digits of precision, for better accuracy.

We can write $q(z) = \sum_{n=1}^{N}\left(\vartheta_n H_n(z) + \vartheta_n' \tilde{H}_n(z)\right)$ as the linear combination of the Hermite interpolation basis polynomials

$$H_n(z) := \left[1 - 2(z-\chi_n)L_n'(\chi_n)\right] L_n^2(z),\qquad \widetilde{H}_n(z) := (z-\chi_n)L_n^2(z), \tag{19}$$

where $L_n(z)$ is a polynomial of the corresponding Lagrange interpolation basis:

$$L_n(z) := \prod_{j\neq n} \frac{z-\chi_j}{\chi_n - \chi_j}, \quad n \in [N_1]. \tag{20}$$

The polynomials (19) allow one to explicitly express $\mathscr{H}_{\{\chi_n\}}^{-1}$: by choosing $\vartheta_j = 1$, $\vartheta_j' = 0$ and $\vartheta_n = \vartheta_n' = 0$, $n \neq j$, we have that

$$\mathrm{col}\left\{\begin{bmatrix} \vartheta_n \\ \vartheta_n' \end{bmatrix}\right\}_{n=1}^{N_1} = \mathrm{e}_{2j-1}$$

is a standard basis vector in $\mathbb{R}^{2N_1}$. Thus, $\mathscr{H}_{\{\chi_n\}}^{-1}\mathrm{e}_{2j-1}$ is the $(2j-1)$-th column of $\mathscr{H}_{\{\chi_n\}}^{-1}$. On the other hand, $\mathscr{H}_{\{\chi_n\}}^{-1}\mathrm{e}_{2j-1}$ is the vector of coefficients of the polynomial $q(z)$ for which $q(\chi_j) = 1$, $q'(\chi_j) = 0$ and $q(\chi_n) = q'(\chi_n) = 0$, $n \neq j$. Thus, we have $q(z) = H_n(z)$ and

$$\mathrm{col}_{2j-1}\left(\mathscr{H}_{\{\chi_n\}}^{-1}\right) = \mathscr{C}_{H_n}.$$

Similar arguments with $\vartheta_j = 0$, $\vartheta_j' = 1$ and $\vartheta_n = \vartheta_n' = 0$ yield

$$\mathrm{col}_{2j}\left(\mathscr{H}_{\{\chi_n\}}^{-1}\right) = \mathscr{C}_{\tilde{H}_n}.$$

We are now ready to provide the proof of Theorem 1, which relies on two key steps. First, we employ the implicit function theorem to show the existence and continuous differentiability of an approximation candidate $\hat{P}(\epsilon)$. Then, by taking a total derivative of $\mathscr{F}\left(\hat{P}(\epsilon), \epsilon\right) = 0$ in ϵ and substituting $\epsilon = 0$, we relate the condition numbers $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$ with an appropriate interpolation problem, allowing us to deduce (15).

Proof of Theorem 1. $\mathscr{F}$ in (12) is differentiable in all the variables $(\hat{P}, \epsilon)$. We can decompose the Jacobian matrix $J_{\hat{P}}(P, 0) := \frac{\partial \mathscr{F}}{\partial \hat{P}}(P, 0)$ as

$$J_{\hat{P}}(P,0) = \mathscr{H}_{\{\phi_n\}_{n\in[N_1]}}^{\top} \cdot D_{\hat{P}}(P,0),$$

where $\mathscr{H}_{\{\phi_n\}_{n\in[N_1]}}^{\top}$ is defined in (18) and

$$D_{\hat{P}}(P,0) := \mathrm{diag}\left\{\begin{bmatrix} 1 & 0 \\ 0 & -\Delta y_n \phi_n \end{bmatrix}\right\}_{n=1}^{N_1}. \tag{21}$$

Since the points $\{\phi_n\}_{n=1}^{N_1}$ are distinct, due to the strict monotonicity of the eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ (which are real, since the Sturm–Liouville operator $\mathscr{A}$ is self-adjoint), existence and uniqueness of the solution to the Hermite interpolation problem (18) guarantees that $\det\left(\mathscr{H}_{\{\phi_n\}_{n\in[N_1]}}^{\top}\right) \neq 0$, whereas $\det\left(D_{\hat{P}}(P,0)\right) \neq 0$ since $y_n \neq 0$ due to Assumptions 1 and 2. The implicit function theorem (Krantz & Parks, 2002) guarantees the existence of $\epsilon_* > 0$ and of the continuously differentiable functions $\hat{P}(\epsilon)$ such that $\hat{P}(0) = P$ and (14) is satisfied.

To compute the condition numbers, we differentiate the equality $\mathscr{F}(\hat{P}(\epsilon), \epsilon) = 0$ with respect to ϵ and we substitute $\epsilon = 0$:

$$J_{\hat{P}}(P,0) \cdot \mathrm{col}\left\{\begin{bmatrix} \frac{\mathrm{d}\hat{y}_n}{\mathrm{d}\epsilon}(0) \\ \frac{\mathrm{d}\hat{\lambda}_n}{\mathrm{d}\epsilon}(0) \end{bmatrix}\right\}_{n=1}^{N_1} = \mathrm{col}\left\{\sum_{m=N_1+1}^{N_1+N_2} y_m \phi_m^k\right\}_{k=0}^{2N_1-1}, \tag{22}$$

whence,

$$\mathrm{col}\left\{\begin{bmatrix} \frac{\mathrm{d}\hat{y}_n}{\mathrm{d}\epsilon}(0) \\ \frac{\mathrm{d}\hat{\lambda}_n}{\mathrm{d}\epsilon}(0) \end{bmatrix}\right\}_{n=1}^{N_1} = D_{\hat{P}}^{-1}(P,0)\left(\mathscr{H}_{\{\phi_n\}_{n\in[N_1]}}^{\top}\right)^{-1} \times \sum_{m=N_1+1}^{N_1+N_2} y_m\, \mathrm{col}\left\{\phi_m^k\right\}_{k=0}^{2N_1-1}. \tag{23}$$

By the properties of the Hermite interpolation (18), and specifically in view of the discussion after (20), which shows how to construct the column vectors $\mathscr{C}_{H_n}$ and $\mathscr{C}_{\tilde{H}_n}$,

$$\left(\mathscr{H}_{\{\phi_n\}_{n\in[N_1]}}^{\top}\right)^{-1} = \mathrm{col}\left\{\begin{bmatrix} \mathscr{C}_{H_n}^{\top} \\ \hline \mathscr{C}_{\tilde{H}_n}^{\top} \end{bmatrix}\right\}_{n=1}^{N_1},$$

where we used the notation (17) and $\mathscr{C}_{H_n}^{\top}$, $\mathscr{C}_{\tilde{H}_n}^{\top}$ are row vectors. Since

$$\begin{bmatrix} \mathscr{C}_{H_n}^{\top} \\ \hline \mathscr{C}_{\tilde{H}_n}^{\top} \end{bmatrix} \mathrm{col}\left\{\phi_m^k\right\}_{k=0}^{2N_1-1} \overset{(17)}{=} \begin{bmatrix} H_n(\phi_m) \\ \tilde{H}_n(\phi_m) \end{bmatrix}$$

and $D_{\hat{P}}^{-1}(P,0) := \mathrm{diag}\left\{\begin{bmatrix} 1 & 0 \\ 0 & -(\Delta y_n \phi_n)^{-1} \end{bmatrix}\right\}_{n=1}^{N_1}$, (23) yields (15). ■

4.2. Preliminaries and proof of Theorem 2

The proof of Theorem 2 relies on the explicit representation of the condition numbers $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$ in (15). For simplicity of presentation only, we assume that $N_2 = 1$ and proceed with the following steps: (i) show that $\lambda_m - \lambda_n \propto m^2 - n^2$ for $m, n \in \mathbb{N}$ with $m \geq n$ (Proposition 1); (ii) consider $L_n^2(z)$ in (19) and employ step (i) to upper bound $L_n^2(\phi_{N_1+1})$ by a decaying exponential of Δ (Lemma 2); (iii) employ (15) and step (ii) to bound $H_n(\phi_{N_1+1})$ and $\tilde{H}_n(\phi_{N_1+1})$, thereby deducing (16).

First, we provide bounds on the difference between pairs of eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ of the Sturm–Liouville operator $\mathscr{A}$ in (3).

Proposition 1. *There exist constants $\upsilon, \Upsilon > 0$, which depend on the Sturm–Liouville operator (3), such that*

$$\forall 1 \leq n < m: \quad \upsilon\left(m^2 - n^2\right) \leq \lambda_m - \lambda_n \leq \Upsilon\left(m^2 - n^2\right). \tag{24}$$

Proof. Employing Equation 4.21 in Fulton and Pruess (1994), the eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ have the asymptotic behavior $\lambda_n = \frac{\pi^2}{B^2}n^2 + a_0 + \alpha(n)$, $n \geq 1$, where B and a_0 are positive constants and $\alpha : \mathbb{N} \to \mathbb{R}$ satisfies $|\alpha(n)| \leq Cn^{-2}$ for some $C > 0$. Take $L \in \mathbb{N}$ such that for all $n \geq L$, $|\alpha(n)| \leq \min\left(\frac{1}{2}, \frac{\pi^2}{4B^2}\right)$. Then, for $L \leq n \leq m$, $\upsilon := \frac{\pi^2}{2B^2}$ and $\Upsilon := \frac{\pi^2}{B^2} + 1$, the asymptotic behavior of the eigenvalues yields

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

since $\lambda_m - \lambda_n = \frac{\pi^2}{B^2}\left(m^2 - n^2\right) + \alpha(m) - \alpha(n)$. Fixing this $L \in \mathbb{N}$, since $\{\lambda_n\}_{n=1}^{\infty}$ are distinct, by increasing Υ and decreasing υ, we can also guarantee (24) in the finite set $1 \leq n \leq m \leq L$. We finally show that Υ and υ can be further tuned such that (24) holds for $1 \leq n \leq L$ and $m > L$. Assume by contradiction that it is not the case for the lower bound. Then, for any $j \in \mathbb{N}$, we can set $\upsilon := \frac{1}{j}$ and there exist $m_j > L$ and $n_j \in [L]$ such that

$$\frac{1}{j} > \frac{\lambda_{m_j} - \lambda_{n_j}}{m_j^2 - n_j^2} = \frac{(\pi/B)^2 m_j^2 + a_0 + \alpha(m_j) - \lambda_{n_j}}{m_j^2 - n_j^2}.$$

By the pigeonhole principle and selection of a subsequence, without loss of generality, $n_j \equiv n \in [L]$ and $\lim_{j\to\infty} m_j = \infty$. Taking the limit $j \to \infty$ yields the contradiction $(\pi/B)^2 \leq 0$. The same arguments apply to the upper bound. ■

We now define a function $\mathscr{G}_{w_1,w_2}(\zeta)$ and show its properties, which are essential in deriving bounds on the condition numbers in (15).

Proposition 2. *For $\zeta \in (0, \infty)$, define*

$$\mathscr{G}_{w_1,w_2}(\zeta) := \int_{w_1}^{w_2} \mathfrak{g}(\zeta x)\mathrm{d}x > 0, \quad 0 \leq w_1 < w_2 \leq \infty,$$
$$\mathfrak{g}(x) := -\ln\left(1 - e^{-x}\right) > 0, \quad x \in (0, \infty). \tag{25}$$

The integral $\mathscr{G}_{w_1,w_2}(\zeta)$ is finite (it is an improper integral for $w_1 = 0$ and/or $w_2 = \infty$) and decreasing in ζ, and $\lim_{\zeta\to\infty} \mathscr{G}_{w_1,w_2}(\zeta) = 0$. Also, for any $x > 0$, the function $\mathfrak{g}(x)$ is strictly monotonically decreasing.

Proof. We consider the case $w_1 = 1$ and $w_2 = \infty$ (the complementary case $w_1 = 0$ and $w_2 = 1$ is similar). Integrating by parts, we have

$$\begin{aligned}\mathscr{G}_{1,\infty}(\zeta) &= \ln\left(1 - e^{-\zeta}\right) - \lim_{x\to\infty} x\ln\left(1 - e^{-\zeta x}\right) + \int_1^{\infty} \frac{\zeta x}{e^{\zeta x} - 1}\mathrm{d}x \\ &= \ln\left(1 - e^{-\zeta}\right) + \int_1^{\infty} \frac{\zeta x}{e^{\zeta x} - 1}\mathrm{d}x < \infty,\end{aligned}$$

where we used the fact that

$$\lim_{x\to\infty} x\ln\left(1 - e^{-\zeta x}\right) = -\zeta \lim_{x\to\infty} \frac{x^2}{1 - e^{-\zeta x}} e^{-\zeta x} = 0.$$

Since $\mathfrak{g}'(x) = -\frac{1}{e^x - 1} < 0$ for all $x > 0$, the function $\mathfrak{g}(x)$ is strictly monotonically decreasing, whence so is $\zeta \mapsto \mathscr{G}_{1,\infty}(\zeta)$, since $\mathfrak{g}(x)$ is continuous. Finally, $\lim_{\zeta\to\infty} \ln\left(1 - e^{-\zeta}\right) = 0$ while $\lim_{\zeta\to\infty} \int_1^{\infty} \frac{\zeta x}{e^{\zeta x} - 1}\mathrm{d}x = \lim_{\zeta\to\infty} \frac{1}{\zeta}\int_{\zeta}^{\infty} \frac{y}{e^y - 1}\mathrm{d}y = 0$, and hence we obtain $\lim_{\zeta\to\infty} \mathscr{G}_{1,\infty}(\zeta) = 0$. ■

Next, the asymptotic analysis requires two preliminary lemmas. For simplicity of presentation only, we henceforth assume $N_2 = 1$, meaning that the structured error term in (7) contains a single element. The case of arbitrary $N_2 \in \mathbb{N}$ follows from identical, but more tedious, arguments, which involve estimating each summation term in (15) separately.

Our first lemma leverages Propositions 1 and 2.

Lemma 1. *The following identities hold for $(n, N_1) \in \mathscr{N}_1$:*

$$\begin{aligned}\mathscr{Z}^1_{n,\Delta,N_1} &:= \prod_{j\in[N_1]\setminus\{n\}} (\phi_{N_1+1} - \phi_j)^2 = e^{-2\Delta\sum_{j\in[N_1]\setminus\{n\}}\lambda_j - 2\theta^1_{n,\Delta,N_1}}, \\ \mathscr{Z}^2_{n,\Delta} &:= \prod_{j=1}^{n-1}(\phi_n - \phi_j)^2 = e^{-2\Delta\sum_{j=1}^{n-1}\lambda_j - 2\theta^2_{n,\Delta}}, \\ \mathscr{Z}^3_{n,\Delta,N_1} &:= \prod_{j=n+1}^{N_1} (\phi_n - \phi_j)^2 = e^{-2\Delta(N_1-n)\lambda_n - 2\theta^3_{n,\Delta,N_1}},\end{aligned} \tag{26}$$

where, whenever $k < l$, $\prod_{j=l}^k b_j = 1$ and $\sum_{j=l}^k b_j = 0$, while

$$\theta^1_{n,\Delta,N_1} := \sum_{j\in[N_1]\setminus\{n\}} \mathfrak{g}(\Delta(\lambda_{N_1+1} - \lambda_j)),$$
$$\theta^2_{n,\Delta} := \sum_{j=1}^{n-1} \mathfrak{g}(\Delta(\lambda_n - \lambda_j)), \quad \theta^3_{n,\Delta,N_1} := \sum_{j=n+1}^{N_1} \mathfrak{g}(\Delta(\lambda_j - \lambda_n)),$$

with $\mathfrak{g}(x)$ defined in Proposition 2. For $\Delta \in [\Delta_*, \infty)$, the terms θ^1_{n,Δ,N_1}, $\theta^2_{n,\Delta}$ and θ^3_{n,Δ,N_1} are $O(1)$ (hence, bounded), uniformly in $(n, N_1) \in \mathscr{N}_1$. If $\Delta \to \infty$, then θ^1_{n,Δ,N_1}, $\theta^2_{n,\Delta}$ and θ^3_{n,Δ,N_1} are $o(1)$, uniformly in $(n, N_1) \in \mathscr{N}_1$.

Proof. We only report the calculations for $\mathscr{Z}^1_{n,\Delta,N_1}$, as the proof for the other terms in (26) is similar. We write

$$\begin{aligned}\ln\left(\mathscr{Z}^1_{n,\Delta,N_1}\right) &= 2\sum_{j\neq n} \ln\left(e^{-\Delta\lambda_j}\left(1 - e^{-\Delta(\lambda_{N_1+1} - \lambda_j)}\right)\right) \\ &\overset{(25)}{=} -2\Delta\sum_{j\neq n}\lambda_j - 2\theta^1_{n,\Delta,N_1}.\end{aligned}$$

By Proposition 2, the function $\mathfrak{g}$ is strictly monotonically decreasing. Combining this with (24) in Proposition 1, we have

$$\begin{aligned}&\sum_{j\neq n} \mathfrak{g}\left(\Delta\Upsilon((N_1+1)^2 - j^2)\right) \\ &\quad \leq \theta^1_{n,\Delta,N_1} \leq \sum_{j\neq n} \mathfrak{g}\left(\Delta\upsilon((N_1+1)^2 - j^2)\right).\end{aligned} \tag{27}$$

To convert the lower and upper bounds in (27) into Riemann sums, we note that $(N_1 + 1 - j)(N_1 + 1) \leq (N_1 + 1)^2 - j^2 \leq (N_1 + 1 - j)(2N_1 + 1)$, whence

$$\begin{aligned}\sum_{j\neq n} \mathfrak{g}\left(\Delta\Upsilon((N_1+1)^2 - j^2)\right) &\geq \sum_{j\neq n} \mathfrak{g}\left(\Delta\Upsilon(N_1 + 1 - j)(2N_1 + 1)\right) \\ &\overset{N_1+1-j=m}{=} \sum_{m\in[N_1]} \mathfrak{g}\left(\Delta\Upsilon(2N_1+1)m\right) - \mathfrak{g}_{n,\Upsilon} \\ &\geq \int_1^{N_1+1} \mathfrak{g}\left(\Delta\Upsilon(2N_1+1)x\right)\mathrm{d}x - \mathfrak{g}_{n,\Upsilon} \\ &= \mathscr{G}_{1,N_1+1}(\Delta\Upsilon(2N_1+1)) - \mathfrak{g}_{n,\Upsilon},\end{aligned} \tag{28}$$

where $\mathfrak{g}_{n,\Upsilon} := \mathfrak{g}\left(\Delta\Upsilon(N_1 + 1 - n)(2N_1 + 1)\right)$. The second inequality in (28) follows from comparison of the sum to an integral, using the properties of $\mathscr{G}_{w_1,w_2}$ and $\mathfrak{g}$ in Proposition 2.

Similarly, denoting $\mathfrak{g}_{n,\upsilon} := \mathfrak{g}\left(\Delta\upsilon(N_1 + 1 - n)(N_1 + 1)\right)$,

$$\sum_{j\neq n} \mathfrak{g}\left(\Delta\upsilon((N_1+1)^2 - j^2)\right) \leq \mathscr{G}_{0,N_1}(\Delta\upsilon(N_1+1)) - \mathfrak{g}_{n,\upsilon}. \tag{29}$$

By Proposition 2, both the lower bound (28) and the upper bound (29) on θ^1_{n,Δ,N_1} are bounded for $\Delta \geq \Delta_* > 0$ and tend to zero as $\Delta \to \infty$, uniformly in $(n, N_1) \in \mathscr{N}_1$. Thus, the same properties also hold for θ^1_{n,Δ,N_1}, which concludes the proof. ■

Our second lemma requires the following proposition.

Proposition 3. *Let*

$$\Theta(n) := \sum_{j=n+1}^{N_1} (j^2 - n^2), \quad n \in [N_1]. \tag{30}$$

$\Theta(n) \geq 0$ for all $(n, N_1) \in \mathcal{N}_\eta$ with $\eta \in (0, 1]$. Moreover, $\Theta(n)$ is monotonically decreasing with $\Theta(N_1) = 0$. If $\eta \in (0, 1)$, there exists $a(\eta) > 0$ such that $\Theta(\lfloor \eta N_1 \rfloor) \geq a(\eta) N_1^3$ for $N_1 \gg 1$.

Proof. The first part follows from the definition (30). For the second part, let $N_1 \gg 1$. Then, a computation shows that

$$\Theta(\lfloor \eta N_1 \rfloor) \geq \frac{N_1^3}{3}\left(1 + 2\eta^3 - 3\eta^2\right) + O(N_1^2).$$

The function $a(x) = \frac{1}{6} + \frac{1}{3}x^3 - \frac{1}{2}x^2$ is strictly decreasing on $x \in [0, 1]$ with $a(1) = 0$, whence $a(\eta) > 0$. Therefore, $\Theta(\lfloor \eta N_1 \rfloor) \geq 2a(\eta) N_1^3 + O(N_1^2) \geq a(\eta) N_1^3$ for $N_1 \gg 1$. ■

We now recall the interpolation polynomials in Section 4.1 to state and prove our second lemma.

Lemma 2. *Consider the Lagrange polynomials L_n defined in (20), with the nodes $\{\chi_n\}_{n=1}^{N_1}$ replaced by $\{\phi_n\}_{n=1}^{N_1}$, respectively. Let $\Delta \geq \Delta_* > 0$ and $(n, N_1) \in \mathcal{N}_1$. Recalling (30), for any fixed $\eta \in (0, 1]$ and $N_1 \in \mathbb{N}$,*

$$\max_{n \in [\lfloor \eta N_1 \rfloor]} L_n^2\left(\phi_{N_1+1}\right) \leq e^{-2\upsilon \Delta \Theta(\lfloor \eta N_1 \rfloor) + o(1)}, \tag{31}$$

for $\Delta \gg \Delta_*$. In particular, for $\eta = 1$, $\max_{n \in [N_1]} L_n^2\left(\phi_{N_1+1}\right) = O(1)$. Here,

$$L_n^2\left(\phi_{N_1+1}\right) = e^{-2\Delta \sum_{j=n+1}^{N_1}(\lambda_j - \lambda_n)} \times \begin{cases} e^{-2\theta^1_{1,\Delta,N_1} + 2\theta^3_{1,\Delta,N_1}}, & n = 1, \\ e^{-2\theta^1_{N_1,\Delta,N_1} + 2\theta^2_{N_1,\Delta}}, & n = N_1, \\ e^{-2\theta^1_{n,\Delta,N_1} + 2\theta^2_{n,\Delta} + 2\theta^3_{n,\Delta,N_1}}, & \textit{otherwise}. \end{cases} \tag{32}$$

Proof. To derive (32), employing (20) and (26), we have

$$L_n^2\left(\phi_{N_1+1}\right) = \frac{\mathscr{L}^1_{n,\Delta,N_1}}{\mathscr{L}^2_{n,\Delta}\mathscr{L}^3_{n,\Delta,N_1}}.$$

Therefore, for $n \in [N_1 - 1] \setminus \{1\}$, we can write

$$L_n^2\left(\phi_{N_1+1}\right) = e^{-2\Delta \sum_{j=n+1}^{N_1}(\lambda_j - \lambda_n) - 2\theta^1_{n,\Delta,N_1} + 2\theta^2_{n,\Delta} + 2\theta^3_{n,\Delta,N_1}}. \tag{33}$$

Recalling the summation convention below (26), for $n = 1$ we obtain the expression in (33) with $2\theta^2_{n,\Delta}$ omitted, while for $n = N_1$ we have $L_n^2\left(\phi_{N_1+1}\right) = e^{-2\theta^1_{N_1,\Delta,N_1} + 2\theta^2_{N_1,\Delta}}$. This yields (32). Next, note that for $n \in [\lfloor \eta N_1 \rfloor]$, $\eta \in (0, 1]$,

$$\begin{aligned} -2\Delta \sum_{j=n+1}^{N_1}\left(\lambda_j - \lambda_n\right) &\overset{(24)}{\leq} -2\upsilon\Delta \sum_{j=n+1}^{N_1}\left(j^2 - n^2\right) \\ &\overset{(30)}{=} -2\upsilon\Delta\Theta(n) \leq -2\upsilon\Delta\Theta(\lfloor \eta N_1 \rfloor), \end{aligned} \tag{34}$$

since Θ is monotonically decreasing as per Proposition 3. Combining (34) with the last statement of Lemma 1 yields (31). In particular, for $\eta = 1$, since $\Theta(N_1) = 0$ by Proposition 3 and $\theta^1_{n,\Delta,N_1}, \theta^2_{n,\Delta}, \theta^3_{n,\Delta,N_1}$ are $O(1)$, uniformly in $(n, N_1) \in \mathcal{N}_1$, by Lemma 1, we obtain that $\max_{n \in [N_1]} L_n^2\left(\phi_{N_1+1}\right) = O(1)$. ■

We are now ready to prove Theorem 2.

Proof of Theorem 2. By Theorem 1,

$$\begin{aligned} \left|\mathscr{K}_y(n)\right| &= \left|y_{N_1+1}\right| \left|H_n(\phi_{N_1+1})\right| \\ &\overset{(19)}{=} \left|y_{N_1+1}\right| \left|1 - 2(\phi_{N_1+1} - \phi_n) L_n'(\phi_n)\right| L_n^2(\phi_{N_1+1}). \end{aligned} \tag{35}$$

Taking the derivative of the Lagrange polynomials $L_n(z)$ in (20) with $\{\chi_n\}_{n=1}^{N_1}$ replaced by $\{\phi_n\}_{n=1}^{N_1}$ and substituting $z = \phi_n$ yields

$$L_n'(\phi_n) = \sum_{k \in [N_1] \setminus \{n\}} (\phi_n - \phi_k)^{-1}. \tag{36}$$

In (35), $\left|1 - 2(\phi_{N_1+1} - \phi_n) L_n'(\phi_n)\right| \leq 1 + 2\left|\left(\phi_{N_1+1} - \phi_n\right) L_n'(\phi_n)\right|$. Substituting (36) yields

$$\begin{aligned} \left|\left(\phi_{N_1+1} - \phi_n\right) L_n'(\phi_n)\right| &\leq \sum_{k \in [N_1] \setminus \{n\}} \left|\frac{\phi_{N_1+1} - \phi_n}{\phi_n - \phi_k}\right| \\ &= \sum_{k=1}^{n-1} \frac{\phi_n - \phi_{N_1+1}}{\phi_k - \phi_n} + \sum_{k=n+1}^{N_1} \frac{\phi_n - \phi_{N_1+1}}{\phi_n - \phi_k}, \end{aligned}$$

where the summation convention below (26) holds. Then,

$$\begin{aligned} \sum_{k=1}^{n-1} \frac{\phi_n - \phi_{N_1+1}}{\phi_k - \phi_n} &= \sum_{k=1}^{n-1} \frac{e^{-\Delta\lambda_n} - e^{-\Delta\lambda_{N_1+1}}}{e^{-\Delta\lambda_k} - e^{-\Delta\lambda_n}} \\ &= \sum_{k=1}^{n-1} e^{-\Delta(\lambda_n - \lambda_k)} \frac{1 - e^{-\Delta\left(\lambda_{N_1+1} - \lambda_n\right)}}{1 - e^{-\Delta(\lambda_n - \lambda_k)}}. \end{aligned}$$

For $k \in [n-1]$, we have $e^{-\Delta(\lambda_n - \lambda_k)} \overset{(24)}{\leq} e^{-\Delta\upsilon(n^2 - k^2)} \leq e^{-3\Delta\upsilon}$, while $1 - e^{-\Delta\left(\lambda_{N_1+1} - \lambda_n\right)} \leq 1$ and $1 - e^{-\Delta(\lambda_n - \lambda_k)} \geq 1 - e^{-3\Delta\upsilon}$. Therefore, we can write

$$\sum_{k=1}^{n-1} \frac{\phi_n - \phi_{N_1+1}}{\phi_k - \phi_n} \leq \frac{(n-1)e^{-3\Delta\upsilon}}{1 - e^{-3\Delta\upsilon}} = \frac{n-1}{e^{3\Delta\upsilon} - 1} \leq \frac{(n-1)e^{3\Delta\upsilon}}{e^{3\Delta\upsilon} - 1}, \tag{37}$$

and, by similar arguments,

$$\begin{aligned} \sum_{k=n+1}^{N_1} \frac{\phi_n - \phi_{N_1+1}}{\phi_n - \phi_k} &= \sum_{k=n+1}^{N_1} \frac{e^{-\Delta\lambda_n} - e^{-\Delta\lambda_{N_1+1}}}{e^{-\Delta\lambda_n} - e^{-\Delta\lambda_k}} \\ &= \sum_{k=n+1}^{N_1} \frac{1 - e^{-\Delta\left(\lambda_{N_1+1} - \lambda_n\right)}}{1 - e^{-\Delta\left(\lambda_k - \lambda_n\right)}} \leq \frac{(N_1 - n)e^{3\Delta\upsilon}}{e^{3\Delta\upsilon} - 1}. \end{aligned}$$

Then, by denoting $\Xi(\Delta) := \frac{e^{3\Delta\upsilon}}{e^{3\Delta\upsilon} - 1}$, we have

$$\left|\left(\phi_{N_1+1} - \phi_n\right) L_n'(\phi_n)\right| \leq (N_1 - 1)\Xi(\Delta) \leq N_1 \Xi(\Delta_*), \tag{38}$$

because $\Xi(\Delta)$ is monotone decreasing and $\Delta \in [\Delta_*, \infty)$. Hence, by (35) and Lemma 2, for $\Delta \gg \Delta_*$,

$$\begin{aligned} \max_{n \in [\lfloor \eta N_1 \rfloor]} \left|\mathscr{K}_y(n)\right| &\leq \left|y_{N_1+1}\right| \left(1 + 2N_1\Xi(\Delta_*)\right) e^{-2\upsilon\Delta\Theta(\lfloor \eta N_1 \rfloor) + o(1)} \\ &\leq \left|y_{N_1+1}\right| \left(1 + 2\Xi(\Delta_*)\right) N_1 e^{-2\upsilon\Delta\Theta(\lfloor \eta N_1 \rfloor) + o(1)}. \end{aligned} \tag{39}$$

Given $\vartheta > 0$, choose $N_1(\vartheta) \in \mathbb{N}$ so that, for all $N_1 \geq N_1(\vartheta)$, we have $\upsilon\Theta(\lfloor \eta N_1 \rfloor) \geq \vartheta$, whence

$$e^{-\upsilon\Delta\Theta(\lfloor \eta N_1 \rfloor)} \leq e^{-\vartheta\Delta}. \tag{40}$$

Such $N_1(\vartheta)$ exists in view of Proposition 3. On the other hand, by Assumptions 1 and 2, $|y_{N_1+1}|$ is bounded. Moreover,

$$\begin{aligned} N_1 e^{-\upsilon\Delta\Theta(\lfloor \eta N_1 \rfloor) + o(1)} &\leq N_1 e^{-\upsilon\Delta_*\Theta(\lfloor \eta N_1 \rfloor) + O(1)} \\ &\leq \max_{N_1 \geq N_1(\vartheta)} \left(N_1 e^{-\upsilon\Delta_*\Theta(\lfloor \eta N_1 \rfloor) + O(1)}\right) < \infty. \end{aligned} \tag{41}$$

The first inequality in (41) follows from the fact that $\Delta \geq \Delta_* > 0$ and from replacing $o(1)$ with $O(1)$, which is a uniform bound on θ^1_{n,Δ,N_1}, $\theta^2_{n,\Delta}$ and θ^3_{n,Δ,N_1}, as per Lemma 1. Since $\eta \in (0, 1)$, the maximum in (41) is finite by Proposition 3. Then, we obtain the first asymptotic bound in (16) by combining (39), (40) and (41), and defining

$$\gamma_y(\eta,\vartheta) := \left|y_{N_1+1}\right| (1+2\Xi(\Delta_*)) \times \max_{N_1\geq N_1(\vartheta)} \left(N_1 e^{-\upsilon\Delta_*\Theta(\lfloor\eta N_1\rfloor)+O(1)}\right).$$

Similarly, for $N_1 \geq N_1(\vartheta)$ and $\Delta \gg \Delta_*$,

$$\begin{aligned}\max_{n\in[\lfloor\eta N_1\rfloor]} |\mathscr{K}_\lambda(n)| &= \max_{n\in[\lfloor\eta N_1\rfloor]} \frac{\left|y_{N_1+1}\right|}{\Delta|y_n|\phi_n} \left|\tilde{H}_n(\phi_{N_1+1})\right| \\ &\overset{(19)}{\leq} \max_{n\in[\lfloor\eta N_1\rfloor]} \frac{M_y}{\Delta\phi_n} \left|\phi_{N_1+1}-\phi_n\right| L_n^2(\phi_{N_1+1}) \\ &\leq \frac{M_y}{\Delta} e^{-2\upsilon\Delta\Theta(\lfloor\eta N_1\rfloor)+o(1)} \leq \frac{\gamma_\lambda(\eta,\vartheta)}{\Delta} e^{-\vartheta\Delta},\end{aligned} \tag{42}$$

where $\gamma_\lambda(\eta,\vartheta) := M_y \max_{N_1\geq N_1(\vartheta)} e^{-\upsilon\Delta_*\Theta(\lfloor\eta N_1\rfloor)+O(1)}$ and we used the facts that $\frac{|y_{N_1+1}|}{y_n} \leq M_y$, as mentioned after (8), and $\phi_n^{-1}\left|\phi_{N_1+1}-\phi_n\right| = 1-e^{-\Delta(\lambda_{N_1+1}-\lambda_n)} \leq 1$. ∎

Remark 6. The analysis in the proof of Theorem 2 is conservative. The numerical simulations in the next section show that the predicted exponential decay rate is obtained for values of N_1 that are smaller than $N_1(\vartheta)$, defined in the proof of Theorem 2.

Remark 7. Our results do not depend on stability of the system. Although λ_n can be negative, our analysis only involves the differences $\lambda_m - \lambda_n$, with $m > n$, which are always positive because the eigenvalues form a monotonically increasing sequence. Intuitively speaking, as in other "resolution" problems, a successful approximation of the eigenvalues depends on the separation between the ϕ_n's (see, e.g., Batenkov et al., 2021). If some of the ϕ_n's tend to infinity with $\Delta \to \infty$, then they actually become well separated from their neighbors, making it easier to approximate the corresponding λ_n's. The reconstruction becomes difficult when multiple positive λ_n's exist: in such a scenario, the corresponding ϕ_n's cluster with growing Δ. Our results show that, for such λ_n's, one needs to analyze the relative rate of convergence to zero of the corresponding ϕ_n's, which is captured by the ratios $\frac{\phi_m}{\phi_n} = e^{-\Delta(\lambda_m-\lambda_n)}$.

Remark 8. We are interested in considering large sampling step-sizes $\Delta \to \infty$ for a two-fold reason. First, in applications that involve sampled-data, exogenous sensing constraints can prevent rapid and frequent sampling of the output, for example, if the cost of making a measurement is high and/or the set-up that is required to perform a measurement requires a long preparation time. This is the case, for instance, in meteorological sensing, where reaction–diffusion systems play a key role in modeling weather patterns; also in estimation over networks, requests to perform a measurement are sent to the sensors through a communication network that may involve transmission delays, thereby leading to large time spacing between the instants when the requests are received by the sensors. Second, from a theoretical perspective, we show that increasing the sampling step Δ is beneficial in reducing the approximation condition numbers. This does not contradict Shannon's theorem; differently from Shannon's sampling theory, we consider a number of measurements that is fixed, and hence does not increase when the sampling step-size is smaller. Our results show that for all λ_n's (even for multiple positive λ_n's, whose corresponding ϕ_n's cluster with growing Δ, as per Remark 7), the accuracy of the estimation, which is reflected in the decay of first-order condition numbers, improves when Δ increases.

5. Numerical simulations

In this section, we provide numerical examples to validate our theory and address **Objective 3** above, by showing that the numerical reconstruction algorithm ESPRIT (Roy & Kailath, 1989) achieves the theoretical reconstruction guarantees.

5.1. Theoretical condition numbers and their decay rate

Consider the model (7) with $N_1 = 4$, $N_2 = 1$, $\lambda_n = n^2$, $y_n = 1$ for all n and $t_k = k\Delta$. We use (15) to compute the condition numbers $\mathscr{K}_\lambda(n)$ and $\mathscr{K}_y(n)$, $n \in [N_1]$. The results in Fig. 2(a) show exponential decay of the condition numbers for increasing Δ when $1 \leq n < N_1$, with the decay becoming faster as n decreases, up to when the computational precision limit is reached. Differently from Theorem 2, here we fix N_1 *a priori*, and still we obtain an exponential decay even for a small value of N_1 (see Remark 6). For $n = N_1$ no decay is observed, consistently with our theory (see Remark 5).

5.2. ESPRIT algorithm

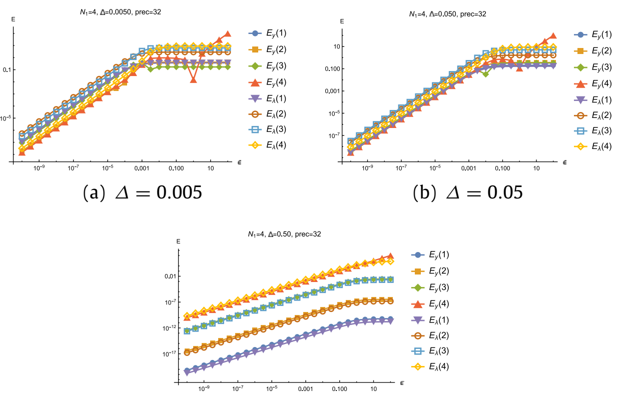
The ESPRIT algorithm (Roy & Kailath, 1989) is one of the best-performing methods for exponential fitting of general damped complex exponential signals, and is applicable in particular to real exponential signals such as (7). ESPRIT requires at least $2N_1$ equispaced samples of a signal $y(t)$ of the form (7), and produces estimates of the parameters $\{y_n, \lambda_n\}_{n=1}^{N_1}$. It provides exact solutions in the noiseless case (i.e. $\epsilon = 0$). Moreover, its performance is close to the best theoretically guaranteed in the presence of noise, in the context of the so-called super-resolution problem in applied harmonic analysis (Batenkov et al., 2021), which is essentially different from the case we consider here.

We apply ESPRIT to the sequence $\{y(\Delta k)\}_{k=0}^{2N_1-1}$, with the same setup as in Section 5.1. In Fig. 2(b), we see that the performance of the ESPRIT algorithm is consistent with Theorem 2 and the computed condition numbers in Section 5.1. We plot the rescaled ESPRIT errors $\mathscr{E}_\lambda^{ESP}(n)$ and $\mathscr{E}_y^{ESP}(n)$, which can be compared to the first-order condition numbers in (10):

$$\mathscr{E}_\lambda^{ESP}(n) := \frac{\left|\hat{\lambda}_n^{ESP} - \lambda_n\right|}{\epsilon}, \qquad \mathscr{E}_y^{ESP}(n) = \frac{\left|\hat{y}_n^{ESP} - y_n\right|}{\epsilon}, \tag{43}$$

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 index-matched to the true λ_n's. Here $\epsilon = 10^{-1}$, and the results were computed with 32 decimal digits of precision. The ESPRIT reconstruction errors exhibit the same asymptotic behavior as the analytic condition numbers in (10), thus satisfying **Objective 3**.

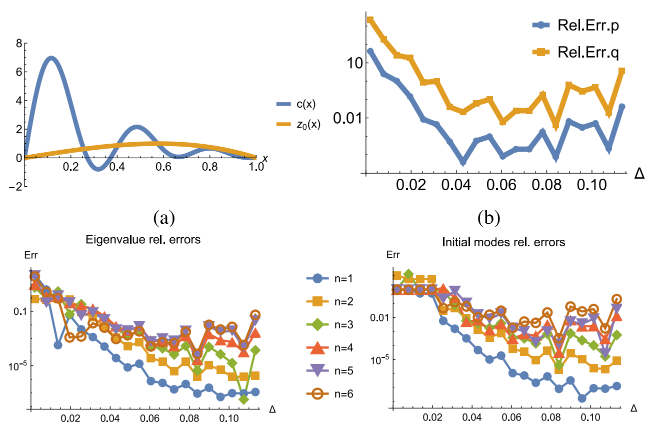
Since the forward map $\mathscr{F}$ is nonlinear, it is nontrivial to provide an effective bound on the neighborhood where the inverse function theorem holds, i.e., to estimate the largest possible value of ϵ_* in Theorem 1. According to a somewhat informal general principle in numerical analysis, this quantity is inversely proportional to the condition number (Demmel, 1987) and thus, in our setting, it is expected to increase with Δ. While a rigorous proof of such a result is outside the scope of the present paper and is usually associated with significant technical hurdles (see, e.g., Batenkov et al., 2021), we demonstrate this effect numerically in Fig. 3, which shows the ESPRIT errors $\epsilon\mathscr{E}_\lambda^{ESP}(n)$ and $\epsilon\mathscr{E}_y^{ESP}(n)$ as a function of ε for different values of the sampling step-size Δ.



(a) $\Delta = 0.005$ (b) $\Delta = 0.05$

(c) $\Delta = 0.5$

Fig. 3. ESPRIT errors $\epsilon\mathscr{E}_{\lambda}^{ESP}(n)$ and $\epsilon\mathscr{E}_{y}^{ESP}(n)$ as a function of ε for different values of the sampling step-size Δ. The breakdown point, i.e., the largest value of ε_* at which the error ceases to scale linearly with ϵ, increases with Δ, which is consistent with the decrease in the condition numbers.



(a) (b)

(c)

Fig. 4. ESPRIT reconstruction for RDE parameter estimation. 4(a) Measurement kernel $c(x)$ and initial condition $z_0(x)$, for $x \in [0, 1]$. 4(b) Reconstruction errors of p and q, estimated from $\{\hat{\lambda}_n^{ESP}\}_{n=1}^{N_1}$ by a linear least squares fit, for increasing Δ. 4(c) Reconstruction errors of RDE eigenvalues and Fourier coefficients (modes) of the initial condition, for increasing Δ.

5.3. RDE parameter estimation

We test the complete procedure on an RDE 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 $\psi_n(x) = \sqrt{2}\sin(n\pi x)$, $n \in \mathbb{N}$. The initial condition is set to satisfy $z_n(0) = (-1)^{n+1}\left(\sqrt{2}n^3\right)^{-1}$ for all n. To solve the RDE, we use the method of lines for space discretization with $N_x = 60$ collocation points and 4th order finite difference approximation, and the resulting ODE system is integrated for $t \in [0, 2]$. Let $c(x) = \sum_{n=1}^{N_1} c_n\psi_n(x) + \epsilon\sum_{n=N_1+1}^{N_1+2} c_n\psi_n(x)$ be the measurement kernel, where $\{c_n\}_{n=1}^{N_1+2}$, $c_n \in [1, 2]$ are randomly chosen, and $\epsilon = 10^{-4}$. The measurement signal $y(t)$ in (2) is computed using global adaptive quadrature as implemented in `NIntegrate` library function. Finally, $y(t)$ is sampled at 1025 equispaced points in $[0, 2]$, thus giving a minimal sampling step-size of $\Delta_* := \frac{1}{512}$. The initial condition $z_0(x)$ and the measurement kernel $c(x)$ are plotted in Fig. 4(a).

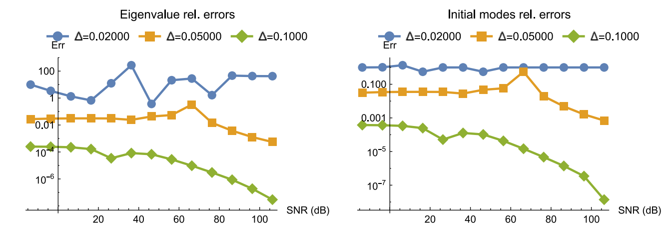


Fig. 5. ESPRIT reconstruction errors, as a function of the SNR computed as in (44), for varying values of Δ. Only the errors in the first eigenvalue and in the first mode of the initial condition are plotted. As expected, increasing Δ improves the accuracy of the reconstruction.

We apply the ESPRIT algorithm on $\{y(k\Delta)\}_{k=0}^{2N_1-1}$ with varying $\Delta \geq \Delta_*$. The relative errors in the estimated eigenvalues and in the Fourier coefficients (modes) of the initial condition are plotted in Fig. 4(c). The deterioration of the error when Δ crosses a certain threshold is due to the finite precision in the computations. Here, all computations are done with 100 decimal digits of precision. The ESPRIT reconstruction errors have exactly the asymptotic behavior predicted by our theory.

Following our observation regarding the robustness to noise of the recovery criterion in Definition 1 (see Fig. 3), we have run the RDE parameter estimation pipeline for varying values Δ and of the SNR

$$\mathrm{SNR} = 20\log_{10}\left(\frac{\left\|\{|y_n|\}_{n=1}^{N_1}\right\|}{\epsilon\left\|\{|y_n|\}_{n=N_1+1}^{N_2}\right\|}\right), \tag{44}$$

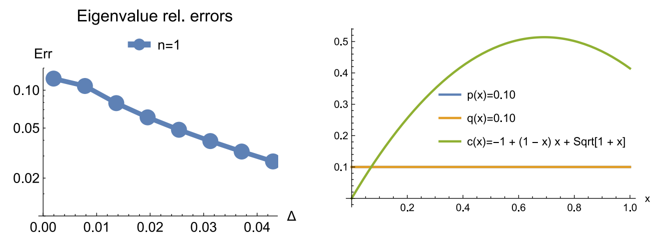
which depends on ε. Our results, shown in Fig. 5, confirm the robustness of the algorithm to noise.

Going beyond the estimation of the eigenvalues $\{\lambda_n\}_{n=1}^{\lfloor\eta N_1\rfloor}$ and modes of the initial condition $\{z_n(0)\}_{n=1}^{\lfloor\eta N_1\rfloor}$, in this example, where p and q are constant, we can also attempt to estimate p, q from the reconstructed eigenvalues, using the relationship $\lambda_n = \pi^2n^2p - q$, by applying linear least squares regression to $\{\hat{\lambda}_n^{ESP}\}_{n=1}^{N_1}$. The errors in the estimated parameters are plotted in Fig. 4(b). Moreover, choosing an optimal value of $\Delta \approx 0.085$ yields an estimate of $p = q = 0.1$ with a relative error of less than 1%. This result is expected to improve with higher computational precision.

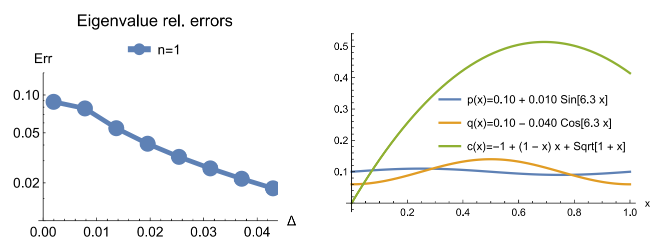
Our method is also applicable to the cases where $c(x)$ is not band-limited (see Remark 4), i.e., when it is given directly instead of using its Fourier coefficients $\{c_n\}$, as well as to the case of varying $p(x)$ and $q(x)$. To show the efficacy of our approach, in Fig. 6 we present examples of a reconstruction of the first eigenvalue in such cases. The computation of the measurement function $y(t)$ was carried out as before, using global adaptive quadrature. In Fig. 6(b), the eigenvalue λ_1 was computed numerically by a finite element method.

6. Conclusion

We have proposed a new approach for the estimation of leading eigenvalues and initial condition Fourier coefficients of an unknown reaction–diffusion equation from a finite number of non-local noisy measurements. Inspired by the algebraic super-resolution literature, we have cast it as an exponential fitting problem. We have proposed an estimation criterion and shown its well posedness for small enough noise intensity. We have also quantified the accuracy of the obtained estimates through the condition numbers, which we have computed explicitly and



(a)



(b)

Fig. 6. Reconstruction of the first eigenvalue in the case of non-band-limited $c(x)$, with 6(a) constant $p(x)$ and $q(x)$; 6(b) space-varying $p(x)$ and $q(x)$.

for which we have provided rigorous asymptotic bounds when the sampling step-size is large; our simulations have shown that the ESPRIT algorithm achieves the optimal asymptotic behavior. Our results build on two key characteristics of the considered problem: (i) The right-hand side of the equation is a Riesz spectral operator with associated eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ tending to infinity; (ii) An estimate on the asymptotic behavior of the eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ allows to derive bounds similar to those in Proposition 1. We believe that any equation in one spatial dimension that satisfies these properties is amenable to a similar analysis.

CRediT authorship contribution statement

Rami Katz: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization, Funding acquisition. **Giulia Giordano:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Funding acquisition, Formal analysis. **Dmitry Batenkov:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Rami Katz reports financial support was provided by the Council of Higher Education in Israel. Giulia Giordano reports financial support was provided by the European Research Council. Dmitry Batenkov reports financial support was provided by the Israel Science Foundation. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

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., & Kunisch, K. (2012). *Estimation techniques for distributed parameter systems*. Springer Science & Business Media.
Batenkov, D., Goldman, G., & Yomdin, Y. (2021). Super-resolution of near-colliding point sources. *Information and Inference: A Journal of the IMA*, *10*(2), 515–572.
Christofides, P. D. (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. A., & Rosen, I. G. (1994). *Lecture notes in pure and applied mathematics*, *Dynamic identification of implicit parabolic systems*. MARCEL DEKKER AG, 153–153.
Demmel, J. (1987). The geometry of ill-conditioning. *Journal of Complexity*, *3*(2), 201–229.
Fulton, C. T., & 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., & 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., & Vyvenko, O. F. (1999). Exponential analysis in physical phenomena. *Review of Scientific Instruments*, *70*(2), 1233–1257. http://dx.doi.org/10.1063/1.1149581.
Katz, R., & 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., & Fridman, E. (2021b). Finite-dimensional control of the heat equation: Dirichlet actuation and point measurement. *European Journal of Control*, *62*, 158–164.
Katz, R., & Fridman, E. (2022). Delayed finite-dimensional observer-based control of 1D parabolic PDEs via reduced-order LMIs. *Automatica*, *142*, Article 110341.
Kirsch, A. (2011). *vol. 120*, *An introduction to the mathematical theory of inverse problems*. Springer.
Krantz, S. G., & Parks, H. R. (2002). *The implicit function theorem: History, theory, and applications*. Springer Science & Business Media.
Lowe, B. D., Pilant, M., & 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.
Pazy, A. (1983). *vol. 44*, *Semigroups of linear operators and applications to partial differential equations*. Springer New York.
Pereyra, V., & Scherer, G. (2010). *Exponential data fitting and its applications*. Bentham Science Publishers.
Quarteroni, A., Sacco, R., & Saleri, F. (2006). *vol. 37*, *Numerical mathematics*. Springer Science & Business Media.
Roy, R., & 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., & Sacks, P. E. (1992). Reconstruction techniques for classical inverse Sturm-Liouville problems. *Mathematics of Computation*, *58*(197), 161–183.
Sivashinsky, G. I. (1977). Nonlinear analysis of hydrodynamic instability in laminar flames–I. Derivation of basic equations. *Acta Astronautica*, *4*, 1177–1206.
Willems, J. C., Rapisarda, P., Markovsky, I., & De Moor, B. L. (2005). A note on persistency of excitation. *Systems & Control Letters*, *54*(4), 325–329.