INTERPOLATORY RATIONAL MODEL ORDER REDUCTION OF PARAMETRIC PROBLEMS LACKING UNIFORM INF-SUP STABILITY[*]

DAVIDE PRADOVERA[†]

Abstract. We present a technique for the approximation of a class of Hilbert space–valued maps which arise within the framework of model order reduction (MOR) for parametric partial differential equations, whose solution map has a meromorphic structure. Our MOR strategy consists in constructing an *explicit rational* approximation based on few snapshots of the solution in an interpolatory fashion. Under some restrictions on the structure of the original problem, we describe a priori convergence results for our technique, hereafter called *minimal rational interpolation*, which show its ability to identify the main features (e.g., resonance locations) of the target solution map. We also investigate some procedures to obtain a posteriori error indicators, which may be employed to adapt the degree and the sampling points of the minimal rational interpolant. Finally, some numerical experiments are carried out to confirm the theoretical results and the effectiveness of our technique.

Key words. model order reduction, rational approximation, Hilbert space–valued meromorphic maps, noncoercive parametric PDEs, Helmholtz equation, frequency response

AMS subject classifications. 30D30, 41A20, 41A25, 35J05, 35P15, 65D15

DOI. 10.1137/19M1269695

1. Introduction. Mathematical models based on PDEs are employed to analyze numerically a wide array of physical, financial, and engineering-related phenomena. In many situations, such models depend on one or more parameters, either because of uncertainties or for control/design purposes, and multiple solves of the model have to be performed for different parameter values.

Often the "naive" approach of solving the original problem (which we will refer to as the *full order model*, FOM) as many times as required is not feasible, for instance because the number of evaluations is too large or because a real-time solution is required. In such cases, model order reduction (MOR) is employed to reduce the computational effort needed for the solution of the FOM. This is achieved by constructing a surrogate model, whose solution is managed to be close to that of the FOM. The construction of the reduced model often requires a considerable computational cost, which, however, can be done *offline* in a preliminary phase, whereas its evaluation at any parameter value is quite cheap and can be carried out *online* in real time.

Many MOR techniques have been proposed for general FOMs, with the most notable and widely applied being the reduced basis (RB) method [15, 16, 17, 21, 26, 35, 36]. The RB method, in its simplest form, assumes that few samples of the solution of the FOM (snapshots) are enough to capture the main features of the solution manifold, i.e., the set of all solutions of the FOM for all parameter values. Exploiting this idea, a surrogate model is built by projecting the FOM onto a subspace generated by few precomputed snapshots. The effectiveness of the RB method has been verified in many cases. However, it is possible to find some parametric problems for which projection-based MOR does not perform optimally. Such examples are often

[†]CSQI, EPFL, Lausanne, Vaud, 1015, Switzerland (davide.pradovera@epfl.ch).

characterized by some irregularity of the FOM (e.g., nonsmooth dependence on the parameters, nonlinearity, lack of stability, etc., of the differential operator therein).

One simple case falling into this category, which is of primary importance in our discussion, is the time-harmonic wave (Helmholtz) equation with parametric wavenumber:

$$(1.1) \qquad \text{for } k \in K \subset \mathbb{C}, \text{ find } u(k) \in V \text{ s.t. } -\Delta u(k) - k^2 u(k) \overset{(W)}{=} f \in W,$$

where V and $W \supset V$ are some suitable Banach function spaces. Depending on the choice of K and of the boundary conditions which complement (1.1), the problem above may lack inf-sup stability over K [32], due to the presence of resonant wavenumbers. If one carelessly applies the RB method (in both its main versions, POD and greedy [35, 36]) to (1.1), one may observe the appearance of spurious ("nonphysical") resonances in the surrogate model.

Discussions concerning this effect can be found in works related to MOR methods for dynamical systems (e.g., Krylov subspace methods [20, 22]), where problems of similar form are quite common, due to the need for frequency-domain computations. Usually, if the number of parameters is small (often just 1, the frequency), explicit rather than implicit approximants are considered, with the surrogate model being built by enforcing interpolation or moment matching conditions at the sample points in the parameter space, possibly in a least squares (LS) way. The main representatives of this class of methods are the Loewner framework [1, 2, 6, 7, 29] and the vector fitting (VF) algorithm [18, 22, 23, 25, 30, 39], which are tailored to the specific structure of the FOM: rational functions are employed, with the objective of representing each resonance of the FOM by a pole of the approximant. This turns out to be particularly useful when the resonances of the FOM have a physical meaning, since their approximation is implicitly enclosed in the surrogate model, from which it can be obtained with little or no computational effort.

More recently, in the wake of these methods for dynamical systems, and somehow trying to profit from their main advantages, univariate LS Padé approximants have been introduced and studied, both in a standard [9, 11] and a fast [10] version, in the context of a single parameter. Such techniques are based on multiple solves of the FOM at a single parameter value, in the same spirit as Krylov subspace methods, but yield an explicit rational approximant like VF. Actually, LS Padé approximants are actually quite comparable to VF, the main differences between the methods being the following:

(i) In VF, one usually must achieve a reasonably high "sampling density" [18], namely use at least $2N$ samples to approximate N resonances; conversely, fast LS Padé approximants can get away with half as many samples.

(ii) The fast version of LS Padé approximants relies on the high dimensionality of the samples, namely on their linear independence, whereas VF can be applied even for scalar outputs (this is not actually a limitation of LS Padé approximants, since it suffices to build a rational approximation of the solution map and then to apply the desired linear functional to such a surrogate solution).

(iii) The error convergence of LS Padé approximants with respect to the number of samples and to the degree of the denominator is well understood for a reasonably broad class of problems [9, 10], while the convergence behavior of the VF error has, to our knowledge, yet to be analyzed thoroughly [18].

(iv) LS Padé approximants are computed from moments of the solution at a single parameter, the explicit computation of which may feature numerically instabilities

(see [12]); instead, VF uses distributed samples in an LS-Lagrange interpolation fashion, for which ill-conditioning issues are much milder.

In this paper, we wish to develop and analyze an extension of fast LS Padé approximants which overcomes (iv) by allowing a distributed parameter sampling while keeping (i) and (iii) intact. In particular, in section 2 we introduce our method, which we name *minimal rational interpolation*, and we state more precisely to which kinds of FOMs our theory applies. Then in section 3 we describe the main convergence results concerning minimal rational interpolation. Afterwards, in section 4 we discuss briefly the topic of a posteriori error/residual estimation for our technique. Possible extensions to multiparameter problems are described in section 5. In section 6, we show two numerical examples to verify the effectiveness of our method and of the presented error indicator, as well as the theoretical convergence rates. Some concluding remarks are in section 7.

2. Description of the method. Let $(V, \langle \cdot, \cdot \rangle_V)$ be a Hilbert space over $\mathbb{C}$, with induced norm $\|\cdot\|_V$, and $K \subset \mathbb{C}$ be compact. Our task is to find a surrogate for a given map $u : K \to V$. Notably, we seek an approach which builds a surrogate starting only from few evaluations (snapshots) of u. Assuming u to be the outcome of a complex computational model, such a technique qualifies as *nonintrusive*, in the sense that it can employ any available solver as a "black box."

We assume that $S > 0$ samples of u at the points $\Xi_S = \{\mu_j\}_{j=1}^S \subset K$ are available. In particular, under some additional regularity assumptions on u, we allow confluence of the sample points: if μ_j appears $K_j + 1$ times in Ξ_S, we assume to have evaluated u, as well as its derivatives with respect to μ up to order K_j, at the point μ_j.

In order to set up our approximation, it is necessary to specify how many "resonances" of u we wish to approximate. To this aim, we consider the integer $N \in \{0, \ldots, S-1\}$, as well as the set of normalized polynomials of degree at most N

$$(2.1) \qquad \mathbb{P}_N^\star(\mathbb{C}) = \{Q \in \mathbb{P}_N(\mathbb{C}), \ \|\|Q\|\|_N = 1\},$$

with $\mathbb{P}_N(\mathbb{C})$ the space of $\mathbb{C}$-valued polynomials of degree up to N and $\|\|\cdot\|\|_N$ an arbitrary (but fixed) Hilbertian norm on $\mathbb{P}_N(\mathbb{C})$. Our aim will be to approximate N resonances of u by the roots of some element of $\mathbb{P}_N^\star(\mathbb{C})$.

Let I^{Ξ_S} be the polynomial interpolation operator[1] for V-valued functions based on samples at the points Ξ_S. Namely, given $\phi : \Xi_S \to V$, the interpolant $I^{\Xi_S}(\phi)$ is the element of the space of V-valued polynomials

$$(2.2) \qquad \mathbb{P}_{S-1}(\mathbb{C}; V) = \left\{ \mu \mapsto \sum_{i=0}^{S-1} p_i \mu^i, \ p_i \in V \ \text{for} \ i = 0, \ldots, S-1 \right\}$$

such that $\phi(\mu') = I^{\Xi_S}(\phi)(\mu')$ for $\mu' \in \Xi_S$. In case of confluent sample points, such a interpolation condition has to be interpreted in a Hermite sense, with derivatives of the interpolant matching those of the original function.

DEFINITION 2.1 (*minimal rational interpolants*). *A minimal rational interpolant of u of type $[S-1/N]$ with samples at Ξ_S is a rational function*

$$u_N^{\Xi_S}(\mu) = I^{\Xi_S}\left(u Q_N^{\Xi_S}\right)(\mu) \Big/ Q_N^{\Xi_S}(\mu), \quad \mu \in \mathbb{C}, Q_N^{\Xi_S}(\mu) \neq 0,$$

---

[1] For stability reasons, it may be wise to replace I^{Ξ_S} with an LS-type interpolator. With some effort, our results can be generalized to such a case.

where the denominator $Q_N^{\Xi_S}$ minimizes the functional

$$(2.3) \qquad \mathbb{P}_N(\mathbb{C}) \ni Q \mapsto j^{\Xi_S}(Q)^2 = \left\| \frac{1}{(S-1)!} \frac{\mathrm{d}^{S-1}}{\mathrm{d}\,\mu^{S-1}} I^{\Xi_S}(uQ) \right\|_V^2$$

over $\mathbb{P}_N^\star(\mathbb{C})$.

The specific expression of j^{Ξ_S} is the natural generalization of the corresponding functional in fast LS Padé approximation [10]. We will see that this form is essential to prove all our convergence results, starting from the central results described in Lemma 3.5.

We remark that since j^{Ξ_S} is convex (see (2.4)), the existence of a minimizer $Q_N^{\Xi_S}$ is guaranteed by the compactness of $\mathbb{P}_N^\star(\mathbb{C})$. From a more practical perspective, the minimization of j^{Ξ_S} can be carried out as follows:

 (i) Build a V-orthonormal basis $\{\varphi_i\}_{i=1}^S$ of the span of the S samples $\{u(\mu_j)\}_{j=1}^S$, as well as the corresponding "component map"

$$\mathbf{w}(\mu_j) = (\langle u(\mu_j), \varphi_i \rangle_V)_{i=1}^S, \quad \text{so that} \quad u(\mu_j) = \sum_{i=1}^S w_i(\mu_j)\varphi_i \quad \text{for } j = 1, \dots, S;$$

   in practice, the basis and the component map can be computed via the Gram–Schmidt procedure or by a generalized QR decomposition of the snapshots.

 (ii) Fix a basis $\{\psi_l\}_{l=0}^N$ of $\mathbb{P}_N(\mathbb{C})$, orthonormal with respect to the scalar product associated to $\|\!|\cdot|\!\|_N$, and employ it to represent any arbitrary $Q \in \mathbb{P}_N^\star(\mathbb{C})$ by its component vector $\mathbf{q}$ (by construction, $\|\mathbf{q}\|_2 = 1$); for a more detailed discussion on practical choices of the polynomial basis, we refer the reader to subsection 3.1.

 (iii) Build the Gramian factor

$$\Psi = \left(\frac{1}{(S-1)!} \frac{\mathrm{d}^{S-1}}{\mathrm{d}\,\mu^{S-1}} I^{\Xi_S}(w_i \psi_l) \right)_{i=1,\dots,S,\, l=0,\dots,N} \in \mathbb{C}^{S \times (N+1)},$$

   which, by linearity, allows one to express the target functional as a quadratic form:

$$
\begin{aligned}
j^{\Xi_S}(Q)^2 &= \frac{1}{(S-1)!^2} \left\langle \frac{\mathrm{d}^{S-1}}{\mathrm{d}\,\mu^{S-1}} I^{\Xi_S}(uQ) \frac{\mathrm{d}^{S-1}}{\mathrm{d}\,\mu^{S-1}} I^{\Xi_S}(uQ) \right\rangle_V \\
&= \sum_{i,i'=1}^S \sum_{l,l'=0}^N \frac{1}{(S-1)!^2} q_l \frac{\mathrm{d}^{S-1}}{\mathrm{d}\,\mu^{S-1}} I^{\Xi_S}(w_i \psi_l) \overline{q_{l'} \frac{\mathrm{d}^{S-1}}{\mathrm{d}\,\mu^{S-1}} I^{\Xi_S}(w_{i'} \psi_{l'})} \langle \varphi_i, \varphi_{i'} \rangle_V \\
(2.4) \quad &= \sum_{i=1}^S \sum_{l,l'=0}^N q_l \overline{q_{l'}} \, \Psi_{il} \overline{\Psi_{il'}} = \mathbf{q}^H \Psi^H \Psi \mathbf{q}
\end{aligned}
$$

   ($\overline{\cdot}$ and $\cdot^H$ denote conjugation and conjugate transposition, respectively).

 (iv) By SVD, obtain the required minimizer as the minimal right singular vector of the Gramian factor Ψ.

We remark that introducing the orthogonal bases $\{\varphi_i\}_{i=1}^S$ and $\{\psi_l\}_{l=0}^N$ is not strictly necessary for the minimization of j^{Ξ_S}. However, they allow one to state the SVD problem in the usual Euclidean norm, and to reduce it to size S, independently of the dimension of V.

2.1. Parametric problem framework. While the map u introduced above may be a very general Hilbert space–valued function, we are often interested in approximating maps of a quite specific type. Indeed, assume that we are given a parametric problem:

$$(2.5) \qquad \text{for } \mu \in K \subset \mathbb{C}, \text{ find } u(\mu) \in \text{Dom}(\mathcal{F}_\mu) \text{ s.t. } \mathcal{F}_\mu(u(\mu)) = 0,$$

with $\mathcal{F}_\mu : V \supset \text{Dom}(\mathcal{F}_\mu) \to V$ a family of densely defined operators, depending continuously on μ. Provided the set $K^\star \subset K$ of values of μ, for which a unique $u(\mu)$ solving (2.5) exists, is nonempty, we focus on the *solution map* associated to (2.5):

$$(2.6) \qquad \begin{aligned} u : K^\star &\to V, \\ \mu &\mapsto u(\mu) \text{ solution of } (2.5). \end{aligned}$$

In the following, we restrict our analysis to a special class of solution maps, namely meromorphic V-valued functions over $\mathbb{C}$ with simple poles:

$$(2.7) \qquad u(\mu) = \sum_{\lambda \in \Lambda} \frac{v_\lambda}{\lambda - \mu}, \quad \mu \in \mathbb{C}.$$

We denote by $\{v_\lambda\}_{\lambda \in \Lambda} \subset V$ the set of *residues*, which we assume to be V-square-summable: $\sum_{\lambda \in \Lambda} \|v_\lambda\|_V^2 < \infty$. Moreover, we assume that the set Λ of *poles*, or *resonances*, is countable, with no finite limit point. If Λ is not finite, (2.7) has to be understood in the sense of V-convergence for an arbitrary ordering of Λ. Without loss of generality, we will assume that Λ is unbounded: if this is not the case, one can always extend Λ by adding (countably) infinitely many fictitious poles at ∞ with arbitrary corresponding residues.

The bulk of our results will be obtained under the additional assumption that the residues form an orthogonal family, i.e., that $\langle v_\lambda, v_{\lambda'} \rangle_V = 0$ whenever $\lambda \neq \lambda'$. This allows considerable simplifications in our derivation: in particular, orthogonal residues have the great advantage of introducing no "correlation" (in the V-topology) between components of the solution corresponding to different resonances. Still, many of our conclusions can be generalized to the nonorthogonal case: we defer a discussion of such extensions to subsection 3.4.

To motivate our interest in this class of functions, we give here some examples of operator families $\{\mathcal{F}_\mu\}_{\mu \in K}$ which lead to meromorphic solution maps; for their approximation, it is reasonable to employ rational functions, as minimal rational interpolants do. However, we remark that the theory that we develop might not necessarily apply to all the cases which we list here:

- Frequency-domain dynamical systems: $\mathcal{F}_\mu(u) = (A - \mu E)u + B$, with $A, E \in \mathbb{C}^{n \times n}$, and $B \in V = \mathbb{C}^{n \times n_I}$ (one column per input). By employing a determinantal representation [5], one can easily show that the solution map is a rational function; in particular, if the pencil (A, E) admits n (not necessarily finite) eigenvalues counting multiplicity, then u is of the form (2.7) [4]. Depending on the spectral properties of A, one may be able to ensure residue orthogonality: for instance, it suffices for $E^\dagger A$ to be normal, where † denotes the Moore–Penrose pseudoinverse.
- Polynomial eigenvalue-like problems: $\mathcal{F}_\mu(u) = A(\mu)u + B$, where $A(\mu) \in \mathbb{C}^{n \times n}$ is a polynomial in μ, and $B \in V = \mathbb{C}^{n \times n_I}$. Using augmentation-based approaches [4, 24], one can cast the problem as a frequency-domain dynamical system of size nM, where M is the degree of A. This implies that the solution

map of the original problem (of size n) is a rational function; in order to check whether it is actually of the form (2.7), one has to investigate the spectral properties of the augmented $(nM) \times (nM)$ problem.

- Helmholtz-type problems: $\mathcal{F}_\mu(u) = (\mathcal{L} - \mu\mathcal{I})u - f$, where $\mathcal{L} : V \supset \mathrm{Dom}(\mathcal{L}) \to V$ is densely defined and bijective, with compact and normal inverse, and $f \in V$ (e.g., $V = L^2(\Omega)$, $\mathcal{L} = -\Delta$, and $\mathrm{Dom}(\mathcal{L}) = \Delta^{-1}(V)$). The solution map is of the form (2.7), and the normality of the resolvent $\mathcal{L}^{-1}$ ensures orthogonal residues [10].

- Analytic compact perturbations of bounded bijective operators: $\mathcal{F}_\mu(u) = (\mathcal{J} - \mathcal{L}(\mu))u - f$, where $\mathcal{J} : V \to V$ is bounded and bijective, $\mathcal{L}(\mu) : V \to V$ is compact for all $\mu \in \mathbb{C}$, and analytic with respect to μ, and $f \in V$; a proof of meromorphy is in [37]. Unfortunately, the form of the solution map is, in general, not (2.7): for instance, it cannot be guaranteed that the poles are simple, i.e., a λ-dependent exponent may appear in the denominator of (2.7).

3. Convergence theory. In this section, we investigate the convergence properties of minimal rational interpolants as the number of samples S increases. To this aim, it becomes necessary to introduce some assumptions on the asymptotic properties of the sample set Ξ_S. In order to streamline the derivation, we choose to follow a potential theory–inspired approach.

3.1. Preliminaries for convergence theory. Fix $K \subset \mathbb{C}$ either as the closure of the finite region bounded by a Jordan curve or a line segment with positive length. There exist [38] a (unique) positive real number $\mathrm{Cap}(K)$ and a continuous bijective conformal mapping

$$\phi_K : \overline{\mathbb{C} \setminus K} \to \mathbb{C} \setminus \mathcal{B}(0, \mathrm{Cap}(K)) = \{z \in \mathbb{C}, |z| \geq \mathrm{Cap}(K)\},$$

holomorphic outside K, such that $|\phi_K| = \mathrm{Cap}(K)$ over ∂K and $\lim_{\mu \to \infty} |\phi'_K(\mu)| = 1$.

The value $\mathrm{Cap}(K)$ is the *logarithmic capacity* of K. It is a powerful concept in complex analysis, algebraic geometry, and approximation theory, and we refer the reader to the monograph [27] for a thorough introduction to the subject. In its more general framework, logarithmic capacity is well defined for any compact set in the complex plane and coincides with the definition above under our assumptions.

Among its properties we remark some bounds for other metrics over complex sets: for any compact set $A \subset \mathbb{C}$, we have

$$(3.1) \qquad \sqrt{\frac{\ell(A)}{\pi}} \leq \mathrm{Cap}(A) \leq \frac{1}{2} \max_{z,w \in A} |z - w|,$$

with ℓ denoting the Lebesgue measure over $\mathbb{C}$. We remark that zero-measure sets are not guaranteed to have capacity zero: for instance, $\mathrm{Cap}([0, 4]) = 1$.

A crucial quantity for stating convergence results in approximation theory is the Green's potential, which is closely related to the complex magnitude of ϕ_K:

$$(3.2) \qquad \Phi_K(\mu) = \begin{cases} |\phi_K(\mu)| & \text{if } \mu \notin K, \\ \mathrm{Cap}(K) & \text{if } \mu \in K. \end{cases}$$

In particular, we use the Green's potential to induce an ordering in the set of poles $\Lambda = \{\lambda_j\}_{j=1,2,\dots}$:

$$(3.3) \qquad \Phi_K(\lambda_1) \leq \Phi_K(\lambda_2) \leq \cdots.$$

Moreover, we require that $\Phi_K(\lambda_{N+1}) > \operatorname{Cap}(K)$, i.e., that the number of poles within K is at most N; see (3.2) and (3.3). Indeed, we will show in the following that minimal rational interpolants fail to differentiate between poles of u within K; hence, without the assumption $\Phi_K(\lambda_{N+1}) > \operatorname{Cap}(K)$, the "unidentifiable" poles would distort the approximant and, in general, prevent convergence to the target solution map.

Now, for a given finite point set $\Xi \subset \mathbb{C}$, we define the corresponding nodal polynomial as $\omega^\Xi(\mu) = \prod_{\mu' \in \Xi} (\mu - \mu')$. This allows us to state the result which motivates our interest in Green's potentials.

THEOREM 3.1 (Fejér points [38]). *Let K be either the closure of the finite region bounded by a Jordan curve or a line segment with positive length. There exists a sequence of sample sets $\left(\Xi_S^K\right)_{S=1,2,\dots}$, whose elements are referred to as the Fejér points for K, such that*

$$(3.4) \qquad \limsup_{S \to \infty} \left| \omega^{\Xi_S^K} \right|^{1/S} \leq \Phi_K \quad \textit{uniformly over compact subsets of } \mathbb{C},$$

$$(3.5) \qquad \lim_{S \to \infty} \left| \omega^{\Xi_S^K} \right|^{1/S} = \Phi_K \quad \textit{uniformly over compact subsets of } \mathbb{C} \setminus \partial K.$$

This means that the Green's potential provides an asymptotic upper bound for the magnitude of the nodal polynomial, as long as the sample points are chosen in an optimal way. From here onward we assume that, given $S \in \mathbb{N}$, the sample set Ξ_S coincides with some set of Fejér points Ξ_S^K, allowing us to drop the superscript.

Remark 3.2. Fejér points are explicitly known only for very simple sets K; this is the case, for instance, in two situations which are very common in applications: the Fejér points of the unit disk are the roots of unity, and for a line segment they are the Chebyshev nodes [38]. However, our theory applies even for more general sampling strategies. The technical requirement is the following: there exists $\Phi : \mathbb{C} \to \mathbb{R}$ such that

- $\sup_K \Phi \leq \inf_{\mathbb{C} \setminus K} \Phi$,
- $\limsup_{S \to \infty} \left| \omega^{\Xi_S} \right|^{1/S} \leq \Phi$ uniformly over K,
- $\lim_{S \to \infty} \left| \omega^{\Xi_S} \right|^{1/S} = \Phi$ uniformly over compact subsets of $\mathbb{C} \setminus K$.

If this is the case, then one can replace the Green's potential by Φ in the entirety of our discussion; in particular, Φ should induce the pole ordering; cf. (3.3).

For instance, one could take any sequence which approximates asymptotically the Fejér points [13]. Another interesting case is that of repeated points, employing Lagrange-Hermite interpolation: given $\{\mu_i^0\}_{i=1}^L \subset \mathbb{C}$, one defines the sample points as $\mu_{nL+i} = \mu_i^0$ for all $n = 0, 1, \dots$ and $i = 1, \dots, L$. Then as long as K is defined as the discrete set $K = \{\mu_i^0\}_{i=1}^L$, the three conditions above hold with $\Phi(\mu) = \prod_{i=1}^L \left| \mu - \mu_i^0 \right|^{1/L}$. In particular, this allows one to analyze fast LS Padé approximants [10] as a special case of minimal rational interpolants, with $L = 1$ and $\Phi(\mu) = \left| \mu - \mu_1^0 \right|$.

A second class of considerations involves the denominator space $\mathbb{P}_N^\star(\mathbb{C})$, whose definition hinges upon the choice of the norm $\|\|\cdot\|\|_N$. Depending on the type of result we wish to prove, we may need to introduce the following assumption on the scaling of $\|\|\cdot\|\|_N$ with respect to N.

Assumption 3.3. There exist $\mu_0 \in \mathbb{C}$ and three positive constants R_{μ_0}, c_{μ_0}, and

C_{μ_0}, all independent of N, such that

$$c_{\mu_0} \prod_{j=1}^{D} |\mu_0 - z_j| \leq \left\| \left\| \prod_{j=1}^{D} (\cdot - z_j) \right\| \right\|_N \leq C_{\mu_0} \prod_{j=1}^{D} (R_{\mu_0} + |\mu_0 - z_j|)$$

for all $N \in \mathbb{N}$ and $\{z_j\}_{j=1}^{D} \in \mathbb{C}^D$, with $0 \leq D \leq N$.

In [10], it was shown that Assumption 3.3 holds (with $R_{\mu_0} = c_{\mu_0} = C_{\mu_0} = 1$) for the "monomial" norm $\|\|\cdot\|\|_N = \|\|\cdot\|\|_{\mu_0, N}$ defined as

$$(3.6) \qquad \|\|Q\|\|_{\mu_0, N} = \sqrt{\sum_{j=0}^{N} |q_j|^2}, \quad \text{where} \quad Q = \sum_{j=0}^{N} q_j \, (\cdot - \mu_0)^j.$$

Many of our results do not rely on Assumption 3.3 but exploit the following weaker result.

LEMMA 3.4. *For all $\mu_0 \in \mathbb{C}$, there exist positive constants R_{μ_0} (independent of N), $c_{\mu_0, N}$, and $C_{\mu_0, N}$, such that*

$$c_{\mu_0, N} \prod_{j=1}^{D} |\mu_0 - z_j| \leq \left\| \left\| \prod_{j=1}^{D} (\cdot - z_j) \right\| \right\|_N \leq C_{\mu_0, N} \prod_{j=1}^{D} (R_{\mu_0} + |\mu_0 - z_j|)$$

for all $N \in \mathbb{N}$ and $\{z_j\}_{j=1}^{D} \in \mathbb{C}^D$, with $0 \leq D \leq N$.

Proof. As described above, $\|\|\cdot\|\|_N = \|\|\cdot\|\|_{\mu_0, N}$ satisfies Lemma 3.4. The result follows by the equivalence of all norms on a finite-dimensional space. □

3.2. Convergence of minimal rational interpolant poles. In the present section, we prove that the denominator $Q_N^{\Xi_S}$ described in Definition 2.1 performs well in identifying the approximate position of some of the poles Λ of the solution map, among which are those within the parameter domain K. As a preliminary step, it is useful to exploit the definition of $Q_N^{\Xi_S}$ to derive a bound for the minimal value of j^{Ξ_S} over $\mathbb{P}_N^\star(\mathbb{C})$.

LEMMA 3.5. *Let $Q \in \mathbb{P}_N(\mathbb{C})$. Then*

$$j^{\Xi_S}(Q)^2 = \sum_{\lambda \in \Lambda} \frac{\|v_\lambda\|_V^2}{|\omega^{\Xi_S}(\lambda)|^2} |Q(\lambda)|^2.$$

Moreover, for any subset Λ_0 of Λ with cardinality at most N, we have

$$(3.7) \qquad j^{\Xi_S}(Q_N^{\Xi_S}) \leq \frac{\left(\sum_{\lambda \in \Lambda \setminus \Lambda_0} \|v_\lambda\|_V^2 \right)^{1/2}}{\|\|\omega^{\Lambda_0}\|\|_N} \sup_{\Lambda \setminus \Lambda_0} \left| \frac{\omega^{\Lambda_0}}{\omega^{\Xi_S}} \right|.$$

In particular, for any $\varepsilon > 0$, there exists $S_{\varepsilon, N}$ such that

$$(3.8) \quad j^{\Xi_S}(Q_N^{\Xi_S}) \leq (1 + \varepsilon)^S \frac{\left(\sum_{j > N} \|v_{\lambda_j}\|_V^2 \right)^{1/2}}{\left\| \left\| \prod_{j=1}^{N} (\cdot - \lambda_j) \right\| \right\|_N} \frac{\prod_{j=1}^{N} |\lambda_{N+1} - \lambda_j|}{\Phi_K(\lambda_{N+1})^S} \quad \text{for } S > S_{\varepsilon, N},$$

provided $\Phi_K(\lambda_{N+1}) < \Phi_K(\lambda_{N+2})$. (If this is not true, there exists a reordering of the poles, still satisfying (3.3), for which (3.8) holds.)

Proof. The interpolation operator can be expressed in barycentric coordinates [8] as

$$(3.9) \qquad I^{\Xi_S}(\phi)(\mu) = \sum_{\mu' \in \Xi_S} \frac{\phi(\mu')\,\omega^{\Xi_S}(\mu)}{(\mu - \mu')\,\omega^{\Xi_S}{}'(\mu')}.$$

Accordingly, since Q is interpolated exactly,

$$\frac{1}{(S-1)!}\frac{\mathrm{d}^{S-1}}{\mathrm{d}\,\mu^{S-1}}I^{\Xi_S}(uQ) = \sum_{\mu' \in \Xi_S} \frac{u(\mu')Q(\mu')}{\omega^{\Xi_S}{}'(\mu')} = \sum_{\lambda \in \Lambda}\sum_{\mu' \in \Xi_S} \frac{v_\lambda Q(\mu')}{(\lambda - \mu')\,\omega^{\Xi_S}{}'(\mu')}$$
$$= \sum_{\lambda \in \Lambda} \frac{v_\lambda}{\omega^{\Xi_S}(\lambda)} I^{\Xi_S}(Q)(\lambda) = \sum_{\lambda \in \Lambda} \frac{v_\lambda}{\omega^{\Xi_S}(\lambda)} Q(\lambda).$$

The first claim follows by orthonormality of $\{v_\lambda\}_{\lambda \in \Lambda}$. From here, we remark that, by construction, $j^{\Xi_S}(Q_N^{\Xi_S}) \leq j^{\Xi_S}(Q)$ for all $Q \in \mathbb{P}_N^\star(\mathbb{C})$, so that, by choosing $Q = \omega^{\Lambda_0}/\||\omega^{\Lambda_0}\||_N$, we obtain

$$(3.10) \qquad j^{\Xi_S}(Q_N^{\Xi_S})^2 \leq j^{\Xi_S}\left(\frac{\omega^{\Lambda_0}}{\||\omega^{\Lambda_0}\||_N}\right)^2 = \sum_{\lambda \in \Lambda \setminus \Lambda_0} \frac{\|v_\lambda\|_V^2}{|\omega^{\Xi_S}(\lambda)|^2}\frac{|\omega^{\Lambda_0}(\lambda)|^2}{\||\omega^{\Lambda_0}\||_N^2},$$

from which (3.7) follows.

Now let $\varepsilon > 0$ and N be fixed, and define $\Lambda_0 = \{\lambda_j\}_{j=1}^N$. It remains to find a bound for the supremum of $r^{\Lambda_0,\Xi_S}(\lambda) = |\omega^{\Lambda_0}(\lambda)/\omega^{\Xi_S}(\lambda)|$ over $\lambda \in \Lambda \setminus \Lambda_0$. To this aim, let λ_S' maximize r^{Λ_0,Ξ_S} over $\Lambda \setminus \Lambda_0$ (for all S, the supremum must be attained, since $r^{\Lambda_0,\Xi_S}(\lambda')$ converges to 0 as $|\lambda'|$ diverges to infinity). The set of maximizers $\{\lambda_S'\}_{S>N}$ must be bounded, due to the polynomial degree of ω^{Λ_0} being fixed, and smaller than that of ω^{Ξ_S}. Thus, by (3.5) there exists $S' = S'(\varepsilon, N, \Lambda, K)$ such that, for $S > S'$,

$$\sup_{\Lambda \setminus \Lambda_0} r^{\Lambda_0,\Xi_S} = \sup_{\{\lambda_S'\}_{S>N}} r^{\Lambda_0,\Xi_S} \leq (1+\varepsilon)^S \sup_{\{\lambda_S'\}_{S>N}} \frac{|\omega^{\Lambda_0}|}{\Phi_K^S} \leq (1+\varepsilon)^S \sup_{\Lambda \setminus \Lambda_0} \frac{|\omega^{\Lambda_0}|}{\Phi_K^S}.$$

Let $\Phi_K(\lambda_{N+1}) = \Phi_K(\lambda_{N+\nu}) < \Phi_K(\lambda_{N+\nu+1})$ for some $\nu \geq 1$, and assume that λ_{N+1} maximizes $|\omega^{\Lambda_0}(\lambda)|$ over $\lambda \in \{\lambda_{N+k}\}_{k=1}^\nu$. Then since $|\omega^{\Lambda_0}|$ is independent of S, the last supremum is attained at λ_{N+1}, provided S is large enough. This yields (3.8). $\square$

Before proving convergence of the roots of the approximate denominator to poles of the solution map, we show as a preliminary result that $Q_N^{\Xi_S}$ takes small values close to some of the elements of Λ.

LEMMA 3.6. *Let N be fixed. For any $\varepsilon > 0$, there exists $S_{\varepsilon,N}$ such that*

$$(3.11) \qquad \left|Q_N^{\Xi_S}(\lambda_j)\right| \leq C_j(1+\varepsilon)^{4S}\left(\frac{\Phi_K(\lambda_j)}{\Phi_K(\lambda_{N+1})}\right)^{2S} \quad \textit{for } S > S_{\varepsilon,N},$$

for $j = 1, \ldots, N$, with C_j independent of S and ε.

Proof. Through some isometry, we identify $\mathbb{P}_N(\mathbb{C})$ with $\mathbb{C}^{N+1}$ (endowed with the Euclidean norm), so that, in particular, $\mathbb{P}_N^\star(\mathbb{C})$ corresponds to the boundary of the unit sphere $\partial \mathcal{B}_{N+1}(\mathbf{0}, 1) \subset \mathbb{C}^{N+1}$.

By duality, there exists a family $\{\boldsymbol{\omega}_\lambda\}_{\lambda \in \Lambda} \subset \mathbb{C}^{N+1}$ such that

$$\mathbb{P}_N(\mathbb{C}) \ni Q \leftrightsquigarrow \mathbf{q} \in \mathbb{C}^{N+1} \quad \text{implies} \quad \boldsymbol{\omega}_\lambda^H \mathbf{q} = Q(\lambda) \text{ for } \lambda \in \Lambda.$$

Let $\mathbf{q}_N^{\Xi_S} \in \partial\mathcal{B}_{N+1}(\mathbf{0}, 1)$ be the identification of $Q_N^{\Xi_S}$, and consider the Hermitian matrices $G^{\Xi_S}, G_N^{\Xi_S} \in \mathbb{C}^{(N+1)\times(N+1)}$ defined as

$$G^{\Xi_S} = \sum_{\lambda \in \Lambda} \frac{\|v_\lambda\|_V^2}{|\omega^{\Xi_S}(\lambda)|^2} \boldsymbol{\omega}_\lambda \boldsymbol{\omega}_\lambda^H \quad \text{and} \quad G_N^{\Xi_S} = \sum_{j=1}^N \frac{\|v_{\lambda_j}\|_V^2}{|\omega^{\Xi_S}(\lambda_j)|^2} \boldsymbol{\omega}_{\lambda_j} \boldsymbol{\omega}_{\lambda_j}^H.$$

From here onward, the proof resembles quite closely that of Lemma 5.4 in [10], being essentially based on deriving a bound for the smallest singular value of G^{Ξ_S} and on finding a rank-N decomposition of $G_N^{\Xi_S}$. The final result that can be obtained is of the following form: for $j = 1, \ldots, N$, for large S,

$$\left|Q_N^{\Xi_S}(\lambda_j)\right| = \left|\boldsymbol{\omega}_{\lambda_j}^H \mathbf{q}_N^{\Xi_S}\right| \leq C_j(1+\varepsilon)^{2S} \frac{\left|\omega^{\Xi_S}(\lambda_j)\right|^2}{\Phi_K(\lambda_{N+1})^{2S}},$$

with C_j independent of S and ε. From here it suffices to observe that $\left|\omega^{\Xi_S}(\lambda_j)\right|$ is bounded by $((1+\varepsilon)\Phi_K(\lambda_j))^S$, thanks to (3.4). $\square$

Finally, we are able to show that the poles of minimal rational approximants converge, as S increases, to some of the elements of Λ, namely the N poles which are the "closest" according to (3.3).

THEOREM 3.7. *Let N be fixed, and denote by $\{\lambda_j^{\Xi_S}\}_{j=1}^N$ the roots of $Q_N^{\Xi_S}$ (in case of deficient degree, we assume the missing roots to be ∞). For any $\varepsilon > 0$, there exists $S_{\varepsilon,N}$ such that*

$$(3.12) \qquad \min_{i=1,\ldots,N} \left|\lambda_i^{\Xi_S} - \lambda_j\right| \leq C_j(1+\varepsilon)^{4S}\left(\frac{\Phi_K(\lambda_j)}{\Phi_K(\lambda_{N+1})}\right)^{2S} \quad \text{for } S > S_{\varepsilon,N},$$

for $j = 1, \ldots, N$, with C_j independent of S and ε.

Proof. We start by observing that, by the normalization of $Q_N^{\Xi_S}$, it must hold that

$$(3.13) \quad \left|Q_N^{\Xi_S}(\lambda)\right| = \frac{\prod_{j=1}^S \left|\lambda - \lambda_j^{\Xi_S}\right|}{\left\|\left|\prod_{j=1}^S \left(\cdot - \lambda_j^{\Xi_S}\right)\right\|\right\|_N} \geq \frac{1}{C_{\mu_0,N}} \prod_{j=1}^N \frac{\left|\lambda - \lambda_j^{\Xi_S}\right|}{R_{\mu_0} + |\lambda - \mu_0| + \left|\lambda - \lambda_j^{\Xi_S}\right|}$$

for any arbitrary $\mu_0 \in \mathbb{C}$; see Lemma 3.4.

From here it suffices to apply the same proof as for Theorem 5.5 in [10], with (3.11) replacing the corresponding bound for fast LS-Padé denominators. The main steps are (i) using (3.11) and (3.13) to show that, for each $j = 1, \ldots, N$, there is a sequence of approximate poles $\{\lambda_{i(S)}^{\Xi_S}\}_{S>S_{\varepsilon,N}}$ which converges to λ_j; (ii) exploiting the fact that the poles of u are simple, concluding that $Q_N^{\Xi_S}$ cannot have, for large S, two roots too close to each other; and (iii) applying (3.11) locally to derive (3.12). $\square$

Going back to our original problem formulation, let $0 \leq \nu \leq N$ be such that $\mathrm{Cap}(K) = \Phi_K(\lambda_\nu) < \Phi_K(\lambda_{\nu+1})$. Theorem 3.7 shows that, for fixed N and large enough S, minimal rational interpolants approximate well the ν poles $\{\lambda_j\}_{j=1}^\nu$ within K, with (exactly) one root of $Q_N^{\Xi_S}$ converging to each of them at rate

$$(3.14) \qquad \mathcal{O}\left(\left(\mathrm{Cap}(K) \big/ \Phi_K(\lambda_{N+1})\right)^{2S}\right).$$

3.2.1. Pole convergence with respect to denominator degree. The statements of Lemma 3.6 and Theorem 3.7 may seem somewhat limited, due to the fact that the denominator degree N enters the convergence bounds in a complex way. This makes it impossible to determine how the pole approximation error behaves if N is increased together with S, namely if it diverges to infinity. In this section, we try to solve this shortcoming. The main result we prove is the following.

THEOREM 3.8. *Take two diverging sequences of integers $(N_k)_{k \in \mathbb{N}}$ and $(S_k)_{k \in \mathbb{N}}$, increasing and strictly increasing, respectively, such that $N_k < S_k$ for all k. Let $\Xi^{(k)}$ be shorthand for Ξ_{S_k}, so that $Q_{N_k}^{\Xi^{(k)}}$ is the denominator of the $[S_k - 1/N_k]$ minimal rational interpolant computed from samples of u at Ξ_{S_k}. We denote by $\{\lambda_j^{(k)}\}_{j=1}^{N_k}$ the roots of $Q_{N_k}^{\Xi^{(k)}}$. Also, let Assumption 3.3 be satisfied. For all $j = 1, 2, \ldots$, we have*

$$\lim_{k \to \infty} \min_{i=1,\ldots,N_k} \left| \lambda_i^{(k)} - \lambda_j \right| = 0.$$

Proof. Since we are interested in asymptotic properties of the poles, in the present proof we employ a slightly different ordering of the resonances $\Lambda = \{\sigma_j\}_{j=1,2,\ldots}$, such that

$$(3.15) \qquad |\sigma_1 - \mu_0| \le |\sigma_2 - \mu_0| \le \cdots.$$

Let $j \in \mathbb{N}$ be arbitrary. Thanks to Assumption 3.3, we can obtain the lower bound

$$\left| Q_{N_k}^{\Xi^{(k)}}(\sigma_j) \right| = \frac{\prod_{i=1}^{N_k} \left| \sigma_j - \lambda_i^{(k)} \right|}{\left\| \left| \prod_{i=1}^{N_k} \left(\cdot - \lambda_i^{(k)} \right) \right| \right\|_{N_k}} \ge \frac{1}{C_{\mu_0}} \prod_{i=1}^{N_k} \frac{\left| \sigma_j - \lambda_i^{(k)} \right|}{R_{\mu_0} + |\sigma_j - \mu_0| + \left| \sigma_j - \lambda_i^{(k)} \right|}$$

$$= \frac{1}{C_{\mu_0}} \prod_{i=1}^{N_k} \varphi_j \left(\left| \sigma_j - \lambda_i^{(k)} \right| \right),$$

with $\varphi_j = \cdot/(R_{\mu_0} + |\sigma_j - \mu_0| + \cdot)$ nonnegative and strictly increasing over $\mathbb{R}^+$. On the other hand, due to Assumption 3.3 and Lemma 3.5 with $\Lambda_0 = \{\sigma_i\}_{i=1}^{N_k}$, one has

$$\left| Q_{N_k}^{\Xi^{(k)}}(\sigma_j) \right| \le \frac{\left| \omega^{\Xi^{(k)}}(\sigma_j) \right|}{\| v_{\sigma_j} \|_V} j^{\Xi^{(k)}}(Q_{N_k}^{\Xi^{(k)}})$$

$$\le \frac{\left| \omega^{\Xi^{(k)}}(\sigma_j) \right|}{\| v_{\sigma_j} \|_V} \frac{\left(\sum_{\lambda \in \Lambda \setminus \Lambda_0} \| v_\lambda \|_V^2 \right)^{1/2}}{\left\| \left| \prod_{i=1}^{N_k} (\cdot - \sigma_i) \right| \right\|_N} \sup_{l > N_k} \frac{\prod_{i=1}^{N_k} |\sigma_l - \sigma_i|}{\left| \omega^{\Xi^{(k)}}(\sigma_l) \right|}$$

$$\le \frac{\left(\sum_{\lambda \in \Lambda} \| v_\lambda \|_V^2 \right)^{1/2}}{c_{\mu_0} \| v_{\sigma_j} \|_V} \sup_{l > N_k} \left(\left| \frac{\omega^{\Xi^{(k)}}(\sigma_j)}{\omega^{\Xi^{(k)}}(\sigma_l)} \right| \prod_{i=1}^{N_k} \left| \frac{\sigma_l - \sigma_i}{\mu_0 - \sigma_i} \right| \right).$$

The two bounds above can be combined with the monotonicity of φ_j to obtain

$$(3.16) \quad \varphi_j \left(\min_{i=1,\ldots,N_k} \left| \sigma_j - \lambda_i^{(k)} \right| \right) \le \left(C_j' \sup_{l > N_k} \left(\left| \frac{\omega^{\Xi^{(k)}}(\sigma_j)}{\omega^{\Xi^{(k)}}(\sigma_l)} \right| \prod_{i=1}^{N_k} \left| \frac{\sigma_l - \sigma_i}{\mu_0 - \sigma_i} \right| \right) \right)^{1/N_k},$$

with $C_j' = C_{\mu_0} (\sum_{\lambda \in \Lambda} \| v_\lambda \|_V^2)^{1/2}/(c_{\mu_0} \| v_{\sigma_j} \|_V)$.

Thanks to the bijectivity of φ_j, which maps 0 to itself, the claim follows if we can show that the right-hand side of (3.16) converges to 0 as k increases. To this aim, a first step is to remove the supremum, whose argument can be simplified by introducing the bounds

$$\prod_{i=1}^{N_k} \left| \frac{\sigma_l - \sigma_i}{\mu_0 - \sigma_i} \right| \leq \prod_{i=1}^{N_k} \frac{|\mu_0 - \sigma_l| + |\mu_0 - \sigma_i|}{|\mu_0 - \sigma_i|} \leq \prod_{i=1}^{N_k} \frac{2 |\mu_0 - \sigma_l|}{|\mu_0 - \sigma_i|} = \frac{2^{N_k} |\mu_0 - \sigma_l|^{N_k}}{\prod_{i=1}^{N_k} |\mu_0 - \sigma_i|}$$

and

$$\left| \frac{\omega^{\Xi^{(k)}}(\sigma_j)}{\omega^{\Xi^{(k)}}(\sigma_l)} \right| = \prod_{\mu' \in \Xi^{(k)}} \left| \frac{\mu' - \sigma_j}{\mu' - \sigma_l} \right| \leq \prod_{\mu' \in \Xi^{(k)}} \frac{|\mu_0 - \sigma_j| + |\mu_0 - \mu'|}{||\mu_0 - \sigma_l| - |\mu_0 - \mu'||},$$

where the term $|\mu_0 - \mu'|$ is bounded from above by $R_{\mu_0,K} = \max_K |\cdot - \mu_0|$.

Now let k be large enough, so that $|\mu_0 - \sigma_{N_k+1}| > R_{\mu_0,K}$. Then the absolute value in the last bound can be omitted, and it must hold that

$$(3.17) \qquad \left| \frac{\omega^{\Xi^{(k)}}(\sigma_j)}{\omega^{\Xi^{(k)}}(\sigma_l)} \right| \prod_{i=1}^{N_k} \left| \frac{\sigma_l - \sigma_i}{\mu_0 - \sigma_i} \right| \leq (C_{l,k,j,\mu_0})^{N_k} \prod_{i=1}^{N_k} \frac{|\mu_0 - \sigma_j| + R_{\mu_0,K}}{|\mu_0 - \sigma_i|},$$

with

$$C_{l,k,j,\mu_0} = \frac{2 |\mu_0 - \sigma_l|}{|\mu_0 - \sigma_l| - R_{\mu_0,K}} \left(\frac{|\mu_0 - \sigma_j| + R_{\mu_0,K}}{|\mu_0 - \sigma_l| - R_{\mu_0,K}} \right)^{S_k/N_k - 1}.$$

Going back to the supremum, it is easy to see that all the terms depending on l in the right-hand side of (3.17) are maximized (over $l > N_k$) by the choice $l = N_k + 1$, thanks to (3.15). This, after a suitable rearrangement of the terms, leads to
(3.18)

$$\varphi_j \left(\min_{i=1,\ldots,N_k} \left| \sigma_j - \lambda_i^{(k)} \right| \right) \leq C_j'^{\,1/N_k} \, C_{N_k+1,k,j,\mu_0} \prod_{i=1}^{N_k} \left(\frac{|\mu_0 - \sigma_j| + R_{\mu_0,K}}{|\mu_0 - \sigma_i|} \right)^{1/N_k}.$$

All of the factors appearing in (3.18), except for the last one, can be easily shown to be bounded as k (and consequently N_k) goes to ∞. Hence, it suffices to show that

$$\lim_{k \to \infty} \prod_{i=1}^{N_k} \left(\frac{|\mu_0 - \sigma_j| + R_{\mu_0,K}}{|\mu_0 - \sigma_i|} \right)^{1/N_k} = 0.$$

To this aim, the Stolz–Cesàro theorem [3] can be applied to prove that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^N \log \left(\frac{|\mu_0 - \sigma_j| + R_{\mu_0,K}}{|\mu_0 - \sigma_i|} \right) = \lim_{N \to \infty} \log \left(\frac{|\mu_0 - \sigma_j| + R_{\mu_0,K}}{|\mu_0 - \sigma_{N+1}|} \right) = -\infty,$$

yielding the claim. $\square$

We have thus shown that, by minimal rational interpolation, we can approximate an arbitrary number of poles of u, as long as we allow the denominator degree N to become sufficiently large. However, this more general result does not provide any information on the rate of convergence, at least not a meaningful one for typical applications. Indeed, most of the steps in the proof above introduce gross simplifications, which make it impossible for the intermediate bounds (3.16) and (3.18) to be reasonably sharp.

3.3. Convergence of minimal rational interpolants. As a complement to the previous results, in this section we investigate the approximation properties of the whole minimal rational interpolant $u_N^{\Xi_S}$, the reference being the map u. Similarly to the analysis of the denominator $Q_N^{\Xi_S}$, we start by fixing the denominator degree N and by proving convergence with respect to the number of samples S.

THEOREM 3.9. *Fix N and $\varepsilon > 0$, and let E be an arbitrary compact subset of $\{\mu \in \mathbb{C}, \Phi_K(\mu) < \Phi_K(\lambda_{N+1})\} \setminus \Lambda$. There exists $S_{\varepsilon,N,E}$ such that*

$$\left\| u(\mu) - u_N^{\Xi_S}(\mu) \right\|_V \leq C_E (1 + \varepsilon)^{2S} \left(\frac{\Phi_K(\mu)}{\Phi_K(\lambda_{N+1})} \right)^S \quad \text{for all } \mu \in E, \ S > S_{\varepsilon,N,E},$$

with C_E independent of S and ε.

Proof. We start by deriving a useful identity for the error $u - u_N^{\Xi_S}$: since $Q_N^{\Xi_S}$ is interpolated exactly by I^{Ξ_S}, we can exploit identity (3.9) to obtain

(3.19)

$$Q_N^{\Xi_S}(\mu) \left(u(\mu) - u_N^{\Xi_S}(\mu) \right) = u(\mu) I^{\Xi_S}(Q_N^{\Xi_S})(\mu) - I^{\Xi_S}(u Q_N^{\Xi_S})(\mu)$$

$$= \sum_{\mu' \in \Xi_S} \sum_{\lambda \in \Lambda} \frac{Q_N^{\Xi_S}(\mu') \omega^{\Xi_S}(\mu)}{(\mu - \mu') \, \omega^{\Xi_S}{}'(\mu')} \left(\frac{v_\lambda}{\lambda - \mu} - \frac{v_\lambda}{\lambda - \mu'} \right)$$

(3.20)

$$= \omega^{\Xi_S}(\mu) \sum_{\lambda \in \Lambda} \frac{v_\lambda}{(\lambda - \mu) \, \omega^{\Xi_S}(\lambda)} \underbrace{\sum_{\mu' \in \Xi_S} \frac{Q_N^{\Xi_S}(\mu') \omega^{\Xi_S}(\lambda)}{(\lambda - \mu') \, \omega^{\Xi_S}{}'(\mu')}}_{= I^{\Xi_S}(Q_N^{\Xi_S})(\lambda) = Q_N^{\Xi_S}(\lambda)}.$$

After dividing by $Q_N^{\Xi_S}(\mu)$, taking the norm, and applying Lemma 3.5, we obtain

$$\left\| u(\mu) - u_N^{\Xi_S}(\mu) \right\|_V = \left| \frac{\omega^{\Xi_S}(\mu)}{Q_N^{\Xi_S}(\mu)} \right| \left(\sum_{\lambda \in \Lambda} \frac{\|v_\lambda\|_V^2}{|\lambda - \mu|^2 \, |\omega^{\Xi_S}(\lambda)|^2} \left| Q_N^{\Xi_S}(\lambda) \right|^2 \right)^{1/2}$$

(3.21)

$$\leq \left| \frac{\omega^{\Xi_S}(\mu)}{Q_N^{\Xi_S}(\mu)} \right| \frac{1}{\min_\Lambda |\cdot - \mu|} j^{\Xi_S}(Q_N^{\Xi_S})$$

$$\leq \left| \frac{\omega^{\Xi_S}(\mu)}{Q_N^{\Xi_S}(\mu)} \right| \frac{1}{\min_\Lambda |\cdot - \mu|} (1 + \varepsilon)^S \frac{C'}{\Phi_K(\lambda_{N+1})^S}$$

for S large enough. In particular, the constant C' is independent of S and ε.

Now, due to (3.4), the term $\left| \omega^{\Xi_S} \right|$ is bounded by $(1 + \varepsilon)^S \Phi_K^S$ for S large enough. Thus, it just remains to show that $|Q_N^{\Xi_S}(\mu)| \min_\Lambda |\cdot - \mu|$ is bounded away from 0 for $\mu \in E$. Since E is a compact subset of $\mathbb{C} \setminus \Lambda$, the second factor is not troublesome. The first term is more problematic, since $Q_N^{\Xi_S}$ may have roots inside E. However, Theorem 3.7 and the triangular inequality ensure that, for S large enough, each root of $Q_N^{\Xi_S}$ has distance from E bounded away from 0. This yields the claim. □

If we restrict our interest to compact subsets E of $K \setminus \Lambda$, we can obtain the simplified result

(3.22)

$$\max_{\mu \in E} \left\| u(\mu) - u_N^{\Xi_S}(\mu) \right\|_V \leq C_E \left(\frac{(1 + \varepsilon)^2 \, \text{Cap}(K)}{\Phi_K(\lambda_{N+1})} \right)^S.$$

Thus, the rate of convergence of the approximation error is the same one that we could expect from polynomial approximants for a function holomorphic over $\{\mu \in \mathbb{C}, \Phi_K(\mu) < \Phi_K(\lambda_{N+1})\}$, having a pole or singularity at λ_{N+1}.

3.3.1. Error convergence with respect to denominator degree. Just like subsection 3.2.1 describes an extension for variable N of the convergence theory for approximate poles, here we consider a generalization of the error convergence theory in the case of diverging denominator degree. The result we are able to show is of a flavor similar to Theorem 3.8 and is summarized in the following.

THEOREM 3.10. *Let $(N_k)_{k\in\mathbb{N}}$, $(S_k)_{k\in\mathbb{N}}$, and $\Xi^{(k)}$ be as in Theorem* 3.8. *Fix $E \subset \mathbb{C}$ compact, and let Assumption* 3.3 *be satisfied for some fixed $\mu_0 \in \mathbb{C}$. Then, for all $\varepsilon > 0$, we have*

$$\lim_{k \to \infty} \mathrm{Cap}\left(\left\{\mu \in E, \left\|u(\mu) - u_{N_k}^{\Xi^{(k)}}(\mu)\right\|_V \geq \varepsilon\right\}\right) = 0,$$

with Cap *denoting the logarithmic capacity; see subsection* 3.1.

Proof. We rely on some concepts introduced in the proofs of Theorems 3.8 and 3.9, namely the pole ordering (3.15) and the intermediate bound (3.21). To start, we follow the same steps as in the proof of Theorem 3.8, exploiting (3.21) and (3.7) with $\Lambda_0 = \{\sigma_j\}_{j=1}^{N_k}$. The resulting bound is

$$(3.23) \qquad \left\|u(\mu) - u_{N_k}^{\Xi^{(k)}}(\mu)\right\|_V \leq \frac{C_{k,\mu_0,R}\left(\sum_{\lambda\in\Lambda}\|v_\lambda\|_V^2\right)^{1/2}}{c_{\mu_0}\left|Q_{N_k}^{\Xi^{(k)}}(\mu)\right|\min_\Lambda|\cdot-\mu|}\prod_{i=1}^{N_k}\frac{R+R_{\mu_0,K}}{|\mu_0-\sigma_i|},$$

with $R = \max_E|\cdot-\mu_0|$ and

$$C_{k,\mu_0,R} = \left(\frac{R+R_{\mu_0,K}}{|\mu_0-\sigma_{N_k+1}|-R_{\mu_0,K}}\right)^{S_k-N_k}\left(\frac{2|\mu_0-\sigma_{N_k+1}|}{|\mu_0-\sigma_{N_k+1}|-R_{\mu_0,K}}\right)^{N_k}.$$

In order to obtain the desired result, we need to manage carefully the two troublesome μ-dependent terms in the denominator of (3.23). To this aim, let $\Lambda^{(k)} = \{\lambda_j^{(k)}\}_{j=1}^{N_k}$ be the set of roots of $Q_{N_k}^{\Xi^{(k)}}$. We partition $\Lambda^{(k)}$ into two (potentially empty) sets $\Lambda_{in}^{(k)}$ and $\Lambda_{out}^{(k)}$ according to the characterization

$$\lambda_j^{(k)} \in \Lambda_{in}^{(k)} \quad \text{if and only if} \quad \left|\lambda_j^{(k)} - \mu_0\right| \leq 2R.$$

Given $\#\Lambda_{in}^{(k)}$ and $\overline{N}$ as the cardinalities of $\Lambda_{in}^{(k)}$ and $\Lambda \cap E$, respectively, we define the family of lemniscates

$$(3.24) \qquad \mathcal{E}_{k,\delta} = \left\{\mu \in \mathbb{C}, \left(\prod_{\lambda'\in\Lambda_{in}^{(k)}}|\mu-\lambda'|\right)\left(\prod_{\lambda\in\Lambda\cap E}|\mu-\lambda|\right) \leq \delta^{\#\Lambda_{in}^{(k)}+\overline{N}}\right\},$$

which depends on the index k and on the (small) value $\delta > 0$.

We remark that the exponent of δ in (3.24) is equal to the degree of the monic polynomial (in μ) whose magnitude is the left-hand side of the inequality. Thus [5, Theorem 6.6.3], the logarithmic capacity of $\mathcal{E}_{k,\delta}$ is equal to δ. (We are assuming without loss of generality that $\#\Lambda_{in}^{(k)} + \overline{N} > 0$. If this is not the case, $\mathcal{E}_{k,\delta}$ is empty for $\delta < 1$, and the claim holds quite trivially.)

The main structure of the remainder of the proof is the following:

(i) we define explicitly a sequence $(\delta_k)_{k=1,2,\dots}$ converging to 0;

(ii) we show that $\left\| u(\mu) - u_{N_k}^{\Xi^{(k)}}(\mu) \right\|_V < \varepsilon$ for all $\mu \in E \setminus \mathcal{E}_{k,\delta_k}$, for large k.

Then the claim follows, since, by inclusion,

$$0 \leq \lim_{k \to \infty} \mathrm{Cap}\left(\left\{ \mu \in E, \left\| u(\mu) - u_{N_k}^{\Xi^{(k)}}(\mu) \right\|_V \geq \varepsilon \right\} \right) \leq \lim_{k \to \infty} \mathrm{Cap}\left(\mathcal{E}_{k,\delta_k} \right) = \lim_{k \to \infty} \delta_k = 0.$$

However, before we can proceed with either step, it is necessary to introduce some additional bounds in (3.23). First, since $Q_{N_k}^{\Xi^{(k)}}$ is normalized, by Assumption 3.3 it must hold that

$$\left| Q_{N_k}^{\Xi^{(k)}}(\mu) \right| \geq \frac{1}{C_{\mu_0}} \left(\prod_{\lambda' \in \Lambda_{in}^{(k)}} \frac{|\mu - \lambda'|}{R_{\mu_0} + |\mu_0 - \lambda'|} \right) \left(\prod_{\lambda'' \in \Lambda_{out}^{(k)}} \frac{|\mu - \lambda''|}{R_{\mu_0} + |\mu_0 - \lambda''|} \right).$$

In the first group of factors, $|\mu_0 - \lambda'|$ is bounded from above by $2R$. In the second group, since $|\mu - \mu_0| \leq R$ for all $\mu \in E$ by definition, we have

$$\frac{|\mu - \lambda''|}{R_{\mu_0} + |\mu_0 - \lambda''|} \geq \frac{|\mu_0 - \lambda''|}{R_{\mu_0} + |\mu_0 - \lambda''|} - \frac{|\mu - \mu_0|}{R_{\mu_0} + |\mu_0 - \lambda''|} \geq \frac{R}{R_{\mu_0} + 2R}.$$

Hence,

$$\left| Q_{N_k}^{\Xi^{(k)}}(\mu) \right| \geq \frac{1}{C_{\mu_0}} \frac{R^{N_k - \#\Lambda_{in}^{(k)}}}{(R_{\mu_0} + 2R)^{N_k}} \prod_{\lambda' \in \Lambda_{in}^{(k)}} |\mu - \lambda'|,$$

which, provided $\delta \leq R$, for all $\mu \in E \setminus \mathcal{E}_{k,\delta}$, implies

$$\left| Q_{N_k}^{\Xi^{(k)}}(\mu) \right| \prod_{\lambda \in \Lambda \cap E} |\mu - \lambda| \geq \frac{R^{N_k - \#\Lambda_{in}^{(k)}}}{C_{\mu_0} (R_{\mu_0} + 2R)^{N_k}} \left(\prod_{\lambda' \in \Lambda_{in}^{(k)}} |\mu - \lambda'| \right) \left(\prod_{\lambda \in \Lambda \cap E} |\mu - \lambda| \right)$$

$$> \frac{R^{N_k - \#\Lambda_{in}^{(k)}} \delta^{\#\Lambda_{in}^{(k)} + \overline{N}}}{C_{\mu_0} (R_{\mu_0} + 2R)^{N_k}} \geq \frac{\delta^{N_k + \overline{N}}}{C_{\mu_0} (R_{\mu_0} + 2R)^{N_k}}.$$

Accordingly, we can simplify (3.23) for all $\mu \in E \setminus \mathcal{E}_{k,\delta}$:

$$\left\| u(\mu) - u_{N_k}^{\Xi^{(k)}}(\mu) \right\|_V < \frac{C_{\mu_0} \left(\sum_{\lambda \in \Lambda} \|v_\lambda\|_V^2 \right)^{1/2} \prod_{\lambda \in \Lambda \cap E} |\mu - \lambda|}{c_{\mu_0} \min_\Lambda |\cdot - \mu|}$$

$$\times C_{k,\mu_0,R} \frac{(R_{\mu_0} + 2R)^{N_k}}{\delta^{N_k + \overline{N}}} \prod_{i=1}^{N_k} \frac{R + R_{\mu_0,K}}{|\mu_0 - \sigma_i|}.$$

Since E is compact and Λ has no finite limit point, the distance between E and $\Lambda \setminus E$ is strictly positive. Hence, the quantity

$$C_{\mu_0} \left(\sum_{\lambda \in \Lambda} \|v_\lambda\|_V^2 \right)^{1/2} \prod_{\lambda \in \Lambda \cap E} |\mu - \lambda| \, \Big/ \, \left(c_{\mu_0} \min_\Lambda |\cdot - \mu| \right)$$

is bounded, for all $\mu \in E$, by a constant C' depending only on E, μ_0, and Λ.

Now it is trivial to achieve (ii) by setting

$$(3.25) \qquad \delta = \delta_k = \left(\frac{C'C_{k,\mu_0,R}\left(R_{\mu_0} + 2R\right)^{N_k}}{\varepsilon} \prod_{i=1}^{N_k} \frac{R + R_{\mu_0,K}}{|\mu_0 - \sigma_i|} \right)^{1/(N_k+\overline{N})}.$$

(We remark that we still need to guarantee that the working assumption $\delta \leq R$ is satisfied. However, since we are planning to prove (i), such a condition will trivially hold, provided k is large enough.)

At this point, it suffices to apply the same strategy employed in the final part of the proof of Theorem 3.8. In particular, we have

$$\lim_{k\to\infty} \left(\prod_{i=1}^{N_k} \frac{R + R_{\mu_0,K}}{|\mu_0 - \sigma_i|} \right)^{1/(N_k+\overline{N})} = \lim_{N\to\infty} \frac{R + R_{\mu_0,K}}{|\mu_0 - \sigma_{N+1}|} = 0,$$

while the other factors in (3.25) stay bounded as k increases.   □

Due to (3.1), Theorem 3.10 can be weakened to prove the convergence in probability of $u_{N_k}^{\Xi^{(k)}}$ to u when both are interpreted as functions from the probability space $(E, \mathcal{B}(E), \ell/\ell(E))$ to the Banach space $(V, \|\cdot\|_V)$.

Still, convergence in capacity is somewhat stronger than that in probability: for instance, it ensures that the set of parameter values for which the approximant yields an error above ε in the limit, namely

$$\bigcap_{l\in\mathbb{N}} \bigcup_{k\geq l} \left\{ \mu \in E \,:\, \left\| u(\mu) - u_{N_k}^{\Xi^{(k)}}(\mu) \right\|_V \geq \varepsilon \right\},$$

cannot include any curve in $\mathbb{C}$.

3.4. Generalizations to nonorthogonal residues. In the previous sections, we have been working under the assumption that the residues $\{v_\lambda\}_{\lambda\in\Lambda}$ form a V-orthogonal set. However, most of our results for fixed N can be extended to the general case, in which we impose no condition on the angles between residues. As such, in this section we remove the assumption of residue orthogonality.

We start by extending Lemma 3.5.

LEMMA 3.11. *Let Q, Λ_0, and ε be the same as in Lemma* 3.5. *Then*

$$j^{\Xi_S}(Q)^2 = \sum_{\lambda,\lambda'\in\Lambda} \frac{\langle v_\lambda, v_{\lambda'}\rangle_V}{\omega^{\Xi_S}(\lambda)\overline{\omega^{\Xi_S}(\lambda')}} Q(\lambda)\overline{Q(\lambda')}.$$

Moreover,

$$j^{\Xi_S}(Q_N^{\Xi_S}) \leq \frac{\sum_{\lambda\in\Lambda\setminus\Lambda_0} \|v_\lambda\|_V}{\|\!\|\omega^{\Lambda_0}\|\!\|_N} \sup_{\Lambda\setminus\Lambda_0} \left| \frac{\omega^{\Lambda_0}}{\omega^{\Xi_S}} \right|,$$

and there exists $S_{\varepsilon,N}$ such that

$$j^{\Xi_S}(Q_N^{\Xi_S}) \leq (1+\varepsilon)^S \frac{\sum_{j>N} \|v_{\lambda_j}\|_V}{\left\|\!\left\| \prod_{j=1}^N (\,\cdot - \lambda_j) \right\|\!\right\|_N} \frac{\prod_{j=1}^N |\lambda_{N+1} - \lambda_j|}{\Phi_K(\lambda_{N+1})^S} \quad \textit{for } S > S_{\varepsilon,N},$$

provided $\Phi_K(\lambda_{N+1}) < \Phi_K(\lambda_{N+2})$.

Proof (sketch). One can follow most of the proof of Lemma 3.5. However, (3.10) does not hold due to the lack of orthogonality. Still, we can replace it by

$$j^{\Xi_S}(Q_N^{\Xi_S})^2 \le j^{\Xi_S}\left(\frac{\omega^{\Lambda_0}}{\|\!|\omega^{\Lambda_0}\|\!|_N}\right)^2 \le \frac{1}{\|\!|\omega^{\Lambda_0}\|\!|_N^2}\left(\sum_{\lambda\in\Lambda\setminus\Lambda_0}\|v_\lambda\|_V\left|\frac{\omega^{\Lambda_0}(\lambda)}{\omega^{\Xi_S}(\lambda)}\right|\right)^2,$$

which holds by triangular inequality. □

We remark the the discrete 2-norm of the residue V-norms in Lemma 3.5 has been replaced by the stronger 1-norm. In order to derive convergence properties, we need to introduce the assumption that the V-norms of the residues are summable, so that the inequalities in Lemma 3.11 are nontrivial.

THEOREM 3.12. *Let N, $\{\lambda_j^{\Xi_S}\}_{j=1}^N$, and ε be the same as in Theorem 3.7. If $\sum_{\lambda\in\Lambda}\|v_\lambda\|_V < \infty$, there exists $S_{\varepsilon,N}$ such that*

$$\left|Q_N^{\Xi_S}(\lambda_j)\right| \le C_j(1+\varepsilon)^{4S}\left(\frac{\Phi_K(\lambda_j)}{\Phi_K(\lambda_{N+1})}\right)^S \quad \text{for } S > S_{\varepsilon,N}$$

and

$$\min_{i=1,\ldots,N}\left|\lambda_i^{\Xi_S} - \lambda_j\right| \le C_j(1+\varepsilon)^{4S}\left(\frac{\Phi_K(\lambda_j)}{\Phi_K(\lambda_{N+1})}\right)^S \quad \text{for } S > S_{\varepsilon,N},$$

for $j = 1,\ldots,N$, with C_j independent of S and ε.

Proof (sketch). The proof is similar to that of Lemma 3.6 and Theorem 3.7. However, we can observe that the exponents in the convergence rates are halved. This is caused by "interference" between the modes: the difference between the representative matrices G^{Ξ_S} and $G_N^{\Xi_S}$ is

$$G^{\Xi_S} - G_N^{\Xi_S} = \sum_{\substack{\lambda,\lambda'\in\Lambda \\ (\lambda,\lambda')\notin\{1,\ldots,N\}^2}} \frac{\langle v_\lambda, v_{\lambda'}\rangle_V}{\omega^{\Xi_S}(\lambda)\overline{\omega^{\Xi_S}(\lambda')}}\boldsymbol{\omega}_{\lambda'}\boldsymbol{\omega}_\lambda^H,$$

which, in particular, contains "off-diagonal" terms, for instance the one corresponding to $\lambda = 1$ and $\lambda' = N+1$. When computing an upper bound for a term of this form, we cannot hope to obtain more than a *single* factor of $\Phi_K(\lambda_{N+1})^{-S}$, whereas two such factors can be gained in the orthogonal case. Similarly, the correlation between the first N resonances causes one less factor of $\left|\omega^{\Xi_S}(\lambda_j)\right|$ to be obtainable from $G_N^{\Xi_S}$. □

THEOREM 3.13. *Let N, ε, and E be as in Theorem 3.9. If $\sum_{\lambda\in\Lambda}\|v_\lambda\|_V < \infty$, there exists $S_{\varepsilon,N,E}$ such that*

$$\left\|u(\mu) - u_N^{\Xi_S}(\mu)\right\|_V \le C_E(1+\varepsilon)^{4S}\left(\frac{\Phi_K(\mu)}{\Phi_K(\lambda_{N+1})}\right)^S \quad \text{for all } \mu \in E, \; S > S_{\varepsilon,N,E},$$

with C_E independent of S and ε.

Proof (sketch). Since Lemma 3.5 does not hold, it is impossible to apply the middle part of the proof of Theorem 3.9. Instead, we can split the sum in (3.20) into two parts: for the first N poles $\lambda_1,\ldots,\lambda_N$, Theorem 3.12 can be applied to bound $|Q_N^{\Xi_S}/\omega^{\Xi_S}|$ from above; for the remaining resonances, it suffices to use the argument from the last portion of Lemma 3.5, with some modifications to account for the presence of $Q_N^{\Xi_S}$. □

Unfortunately, convergence results for variable N do not appear to be possible in the nonorthogonal case, without some strong hypotheses on the "amount of correlation" between resonances: for instance, we believe that one could work under the assumption that the worst-case angle between residues

$$\alpha(\lambda) = \inf_{\lambda' \in \Lambda \setminus \{\lambda\}} \frac{\langle v_\lambda, v_{\lambda'} \rangle_V}{\|v_\lambda\|_V \|v_{\lambda'}\|_V}$$

is bounded away from 0 for all poles λ in the region of interest.

4. A posteriori error indicators. All convergence results presented above are a priori estimates. In particular, they contain quantities which are often not available, namely the number and locations of the poles inside (and of some of the ones outside) the parameter domain K. Moreover, the results in Theorems 3.7 and 3.10 hold only (with some extensions) if the solution map is of the form (2.7), whereas minimal rational interpolants can, in principle, be applied to more general parametric problems and sample points sets.

A typical MOR problem can be cast in the following form: let a certain parameter set K and a parametric problem as in (2.5) be given, along with some fixed tolerance $\varepsilon > 0$. The task is to obtain an approximate solution map $\widetilde{u}$ such that

$$\max_{\mu \in K} \|u(\mu) - \widetilde{u}(\mu)\|_V \le \varepsilon \quad \text{or} \quad \max_{\mu \in K} \frac{\|u(\mu) - \widetilde{u}(\mu)\|_V}{\|u(\mu)\|_V} \le \varepsilon.$$

In our framework, the presence of singularities in the true and, possibly, the approximate solution maps within K may make the conditions above meaningless. In such cases, it may be preferable [17] to consider a residual-based accuracy requirement

$$(4.1) \qquad \max_{\mu \in K} \|\mathcal{F}_\mu(\widetilde{u}(\mu))\|_W \le \varepsilon \quad \text{or} \quad \max_{\mu \in K} \frac{\|\mathcal{F}_\mu(\widetilde{u}(\mu))\|_W}{\|\mathcal{F}_\mu(0)\|_W} \le \varepsilon,$$

with the $\|\cdot\|_W$-norm replacing $\|\cdot\|_V$ to account for regularity differences between residual and solution.

Now let us take as approximate map $\widetilde{u}$ the minimal rational interpolant $u_N^{Z_S}$ with a certain denominator degree N and samples at the S points Z_S, which need not be Fejér points for K. We face the task of understanding whether the required tolerance ε is achieved in the sense of (4.1).

In particular, let us consider the problem of computing a posteriori the norm of the residual $\|\mathcal{F}_\mu(\widetilde{u}(\mu))\|_W$ at some given point $\mu \in K$. In order to proceed, we will assume that the operator $\mathcal{F}_\mu$ is linear and has a separable form in μ, i.e., that there exist two families of complex-valued functions $\{\theta_i^F\}_{i=1}^{n_F}$ and $\{\theta_i^f\}_{i=1}^{n_f}$, a family of operators $\{F_i\}_{i=1}^{n_F}$ over (suitable subsets of) V, and a family $\{f_i\}_{i=1}^{n_f}$ of elements of V, such that

$$(4.2) \qquad \mathcal{F}_\mu(v) = F(\mu)v - f(\mu) = \sum_{i=1}^{n_F} \theta_i^F(\mu) F_i v - \sum_{i=1}^{n_f} \theta_i^f(\mu) f_i.$$

This is the ideal situation to employ RB methods [16, 26, 35, 36], and many strategies have been devised to approximate general parametric problems into the form (4.2), e.g., the (D)EIM [15, 21] and hyper-reduction techniques [14].

Now $F(\mu)$ can be applied to both sides of (3.19) to obtain

$$(4.3) \quad Q_N^{Z_S}(\mu)\mathcal{F}_\mu\left(u_N^{Z_S}(\mu)\right) = \sum_{\mu' \in Z_S} \frac{Q_N^{Z_S}(\mu')\,\omega^{Z_S}(\mu)}{(\mu - \mu')\,\omega^{Z_S}{}'(\mu')}\mathcal{F}_\mu\left(u(\mu')\right)$$

$$= \sum_{i=1}^{n_F} F_i\Delta^{Z_S}\left(\theta_i^F I^{Z_S}\left(uQ_N^{Z_S}\right)\right)(\mu) - \sum_{i=1}^{n_f} \Delta^{Z_S}\left(\theta_i^f Q_N^{Z_S}\right)(\mu)\,f_i,$$

with $\Delta^{Z_S}(\phi) = \phi - I^{Z_S}(\phi)$ the interpolation error for a function ϕ.

This, after dividing by $Q_N^{Z_S}(\mu)$, provides an affine decomposition of the residual with $n_F + n_f$ terms. Hence, its norm (squared) admits an affine decomposition with $(n_F + n_f)^2$ terms. In particular, assuming an offline-online framework [26, 35, 36], the separable terms of the residual can be precomputed offline, i.e., together with the S expensive evaluations of the target solution map u, without increasing the complexity of the overall MOR procedure.

Of course, the discussion above can intrinsically be applied only in an intrusive fashion, since we require access to the operators defining the parametric problem. In general, obtaining an a posteriori nonintrusive error indicator can be quite tricky. Here we propose a simplified approach, which is rigorous only for very specific problem structures, namely

$$(4.4) \qquad\qquad \mathcal{F}_\mu(u) = F(\mu)u - f = (F_0 + \mu F_1)u - f.$$

For such problems, thanks to (3.9) and (4.3), it must hold that

$$Q_N^{Z_S}(\mu)\mathcal{F}_\mu\left(u_N^{Z_S}(\mu)\right) = \omega^{Z_S}(\mu)F_1 \sum_{\mu' \in Z_S} \frac{Q_N^{Z_S}(\mu')u(\mu')}{\omega^{Z_S}{}'(\mu')}$$

$$= \frac{\omega^{Z_S}(\mu)}{(S-1)!}F_1\frac{\mathrm{d}^{S-1}}{\mathrm{d}\,\mu^{S-1}}I^{Z_S}(uQ_N^{Z_S}).$$

Now it suffices to take the $\|\cdot\|_W$-norm and exploit (2.3) to obtain

$$(4.5) \qquad \left\|\mathcal{F}_\mu\left(u_N^{Z_S}(\mu)\right)\right\|_W = \frac{1}{(S-1)!}\left\|F_1\frac{\mathrm{d}^{S-1}}{\mathrm{d}\,\mu^{S-1}}I^{Z_S}(uQ_N^{Z_S})\right\|_W\left|\frac{\omega^{Z_S}(\mu)}{Q_N^{Z_S}(\mu)}\right|$$

$$(4.6) \qquad\qquad \leq \left(\sup_{v \in V\setminus\{0\}}\frac{\|F_1v\|_W}{\|v\|_V}\right)j^{Z_S}\left(Q_N^{Z_S}\right)\left|\frac{\omega^{Z_S}(\mu)}{Q_N^{Z_S}(\mu)}\right|.$$

One obvious limitation of (4.5) and (4.6) is the need to evaluate or bound expensive norms (either the W one or the $V \to W$ operator one) involving F_1. We remark that this drawback can be interpreted as an issue in identifying the exact scaling of the error estimator, a common problem even for stable parametric problems, where the inf-sup (or coercivity) constant must be estimated to link residual and error [28, 35, 36]. If one can overcome such a limitation (see subsection 6.2), (4.5) and (4.6) are certainly quite appealing from a computational standpoint, since their evaluation consists only of few scalar computations.

If (4.4) does not hold, we can replace $F(\mu)$ by a suitable linear approximation (e.g., its first-order Taylor series around some parameter) in a spirit similar to that of the empirical interpolation method. Then (4.5) and (4.6) can be applied heuristically, with

an accuracy which may depend quite sharply on the smoothness of $\mathcal{F}_\mu$, in particular on its second-order variations over the parameter domain.

As a supplement to our discussion, we would like to note that a posteriori error indicators are often used in RB procedures not only to determine whether the required accuracy has been achieved but also to drive the selection of the sample points. For instance, in the weak-greedy RB approach [35, 36], one keeps adding a new snapshot at the parameter value which maximizes the a posteriori error indicator of choice, thus updating the surrogate model (and the error indicator as well) until convergence. We envision that a similar adaptive procedure could be devised also for minimal rational interpolation, with new parameter values being added to the sample set according to the same logic.

5. Extensions to multiple parameters. Throughout our discussion, we have only dealt with the approximation of univariate meromorphic solution maps. However, it is often of interest to approximate the solution map of multiparameter problems: for example, in the analysis of parametric dynamical systems, one may need to approximate the solution of

$$(5.1) \qquad (A(\mathbf{p}) - \mu E(\mathbf{p}))u(\mu, \mathbf{p}) = B(\mathbf{p}), \quad \text{with } \mathbf{p} \in \mathbb{C}^{d-1},$$

where $\mathbf{p}$ is a collection of physical or geometric parameters. For such problems, it is often quite complicated to prove properties of u, jointly in μ and $\mathbf{p}$; however, the analysis is much simpler if the extra parameters $\mathbf{p}$ are kept fixed ("marginalized out") and one looks at the univariate solution map $u_\mathbf{p}(\mu) : \mu \mapsto u(\mu, \mathbf{p})$.

This has lead to the development of *parametric MOR* (pMOR) [6, 19, 33, 40, 41], whose core idea is the following: take a set of parameter values $\{\mathbf{p}_1, \ldots, \mathbf{p}_T\} \subset \mathbb{C}^{d-1}$ and, for each $j = 1, \ldots, T$, build a surrogate $\widetilde{u}_{\mathbf{p}_j}$ of $u_{\mathbf{p}_j}$ (this involves computing the snapshots $u(\mu_i, \mathbf{p}_j)$ for $i = 1, \ldots, S_j$ and applying a single-parameter MOR strategy with respect to μ only); then the joint reduced model $\widetilde{u}$ of u is constructed as a suitable combination of $\widetilde{u}_{\mathbf{p}_1}, \ldots, \widetilde{u}_{\mathbf{p}_T}$.

In a pMOR framework, one could apply minimal rational approximation to obtain the surrogates in frequency only $\{\widetilde{u}_{\mathbf{p}_j}\}_{j=1}^T$. In particular, the results presented in the previous sections apply: for instance, the residual estimator in section 4 is still valid, allowing for a greedy approach with respect to μ. This line of approximation for multiparameter problems is currently under investigation.

There are several reasons to prefer pMOR over global MOR (in d parameters) for problems like (5.1). Many of them are related to the curse of dimensionality, which forces one to take a large number of snapshots, and to build a surrogate model of considerable size, undermining the cost effectiveness of MOR. We refer to the general pMOR literature cited above for a broader discussion.

Here we wish to remark that these disadvantages affect also the natural extension of minimal rational interpolation to d parameters, which we could define as follows: assume that the sets of polynomials $\mathbb{P}_N\left(\mathbb{C}^d; \mathbb{C}\right)$ and $\mathbb{P}_N^\star\left(\mathbb{C}^d; \mathbb{C}\right)$, as well as the interpolation operator I^{Ξ_S}, are suitably defined (for instance, $\mathbb{P}_N\left(\mathbb{C}^d; \mathbb{C}\right)$ could denote the space of polynomials of *total* degree at most N); the optimization problem which yields the surrogate denominator $Q_N^{\Xi_S}$ is now the following: find $Q_N^{\Xi_S} \in \mathbb{P}_N^\star\left(\mathbb{C}^d; \mathbb{C}\right)$, which minimizes over $\mathbb{P}_N^\star\left(\mathbb{C}^d; \mathbb{C}\right)$ the convex functional

$$Q \mapsto j^{\Xi_S}(Q)^2 = \sum_{\substack{\alpha_1 = 0,1\ldots}} \cdots \sum_{\substack{\alpha_d = 0,1\ldots \\ \alpha_1 + \cdots + \alpha_d = M}} \left\| \frac{1}{\alpha_1! \cdots \alpha_d!} \frac{\mathrm{d}^{\alpha_1}}{\mathrm{d}\,p_1^{\alpha_1}} \cdots \frac{\mathrm{d}^{\alpha_{d-1}}}{\mathrm{d}\,p_{d-1}^{\alpha_{d-1}}} \frac{\mathrm{d}^{\alpha_d}}{\mathrm{d}\,\mu^{\alpha_d}} I^{\Xi_S}(uQ) \right\|_V^2,$$

where M is the maximal total degree of the interpolant $I^{\Xi_S}(uQ)$, so that, in general, $S \sim M^d$. In practice, an algorithm for minimizing this functional can be found as a generalization of the strategy proposed in section 2; while we omit the details here, we remark that the Gramian matrix Ψ representing j^{Ξ_S} (see (2.4) for the univariate case) becomes of size approximately $S^{2-1/d} \times N^d$.

In addition to the other limitations of multivariate minimal rational interpolation, we also observe that it is impossible to extend directly our single-parameter theory to the multivariate case, since many of the core ideas (e.g., the barycentric expansion (3.9)) do not generalize nicely to more than one dimension. Overall, also considering the numerical issues intrinsic to high-dimensional interpolation, we believe minimal rational approximants to be an appropriate MOR technique only in low dimension: extensions to many parameters should rely on pMOR or similar approaches.

6. Numerical examples. In this section, we perform some numerical experiments to show the usefulness of minimal rational interpolation and to verify the theoretical convergence rates and error indicator. The first example satisfies the assumption under which our results have been derived; in particular, the residues are orthogonal. Instead, the solution map in the second example can be proven meromorphic, but not necessarily of the form (2.7). For the sake of reproducibility, the code used to run the simulations is available from [34].

6.1. Normal eigenvalue problem. Let $n = 100$, and take $A \in \mathbb{C}^{n \times n}$ as a normal matrix whose eigenvalues are randomly generated according to a uniform distribution over $\{x + \iota y, (x, y) \in [-5, 5]^2\}$ and whose eigenvectors are prescribed by orthonormalizing a matrix with random (complex) standard Gaussian entries. In particular, we order the eigenvalues $\{\lambda_j\}_{j=1}^n$ according to their distance from 0.

Our task is to estimate the eigenvalues of A which lie within the unit disk $K = \mathcal{B}(0, 1)$. To this aim, we fix $\mathbf{v} \in \mathbb{C}^n$ as a random Gaussian vector, and we approximate by minimal rational interpolants the solution map of the parametric problem

$$(6.1) \qquad \text{for } \mu \in K, \text{ find } \mathbf{u}(\mu) \in \mathbb{C}^n \text{ s.t. } (A - \mu I)\,\mathbf{u}(\mu) = \mathbf{v},$$

with $I \in \mathbb{C}^{n \times n}$ the identity matrix. We endow $\mathbb{C}^n$ with the Euclidean scalar product. As sample points Ξ_S, we take the Sth roots of unity, and we choose to employ the polynomial norm $\|\cdot\|_{0,N}$; see (3.6). This norm is induced by monomials, which are $L^2(\partial K)$-orthogonal. Moreover, the action of the interpolation operator I^{Ξ_S} can be evaluated quite efficiently as a fast Fourier transform.

Let us assume that our interest lies in approximating just the eigenvalues of A inside K. As their exact number is unknown, we can choose the denominator degree N according to different strategies:

- We fix N (in our case, we set $N = 10$) and increase progressively the number of samples S; to verify whether the starting guess on N was large enough, we can simply check whether at least one of the approximate poles has converged (for large S) to some point outside K; see Theorem 3.7.
- We can increase N along with the number of samples S, for instance by choosing $N = S - 1$ for all S; in this case, we are sure to approximate all poles; see Theorem 3.8.

We compare visually the two approaches in Figure 1, where we show the error norm obtained by approximants of types [20/10] and [20/20], both based on $S = 21$ samples at roots of unity. We can see that the error increases much more quickly near
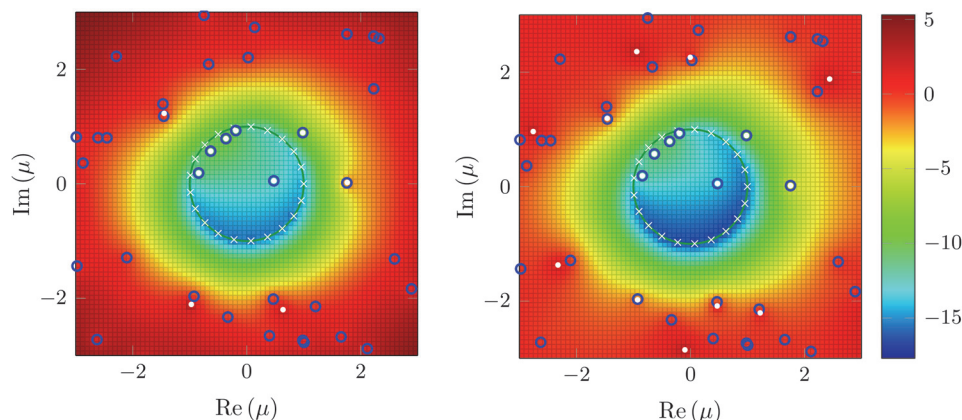
FIG. 1. *Natural logarithm of the relative error norm* $\|\mathbf{u} - \mathbf{u}_N^{\Xi_S}\|_2 / \|\mathbf{u}\|_2$ *achieved by minimal rational interpolants based on samples at* $\Xi_S = \{e^{2j\pi\iota/21}\}_{j=1}^{21}$ *(black crosses) with* $N = 10$ *(left) and* $N = 20$ *(right). The exact and approximated poles are represented in blue and white, respectively.*
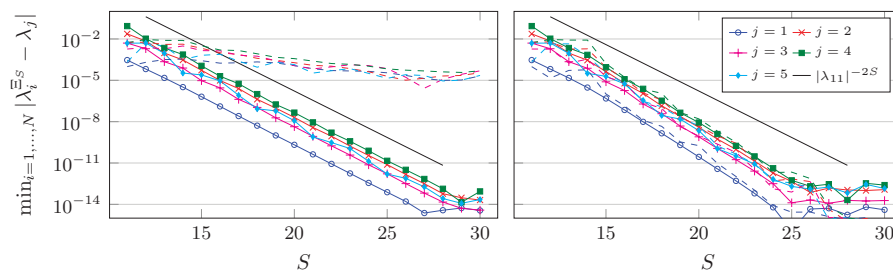

FIG. 2. *Pole approximation error for the poles inside* K. *On the left,* $N = 10$ *is kept fixed, whereas* $N = S - 1$ *on the right. In black is the theoretical rate* (3.14). *The dashed lines show the results obtained by projection on subspaces of dimension* N *computed through POD.*

the boundary of the sampled square in the case $N = 10$. This is to be expected, as the region of convergence increases together with N; see Theorem 3.9. A (possibly) more interesting observation is that the error appears to be globally smaller when we choose $N = 20$. Overall, we can conclude that the choice $N = S - 1$ leads to a better approximant.

Now it just remains to check how well each approach manages to capture the location of the eigenvalues of A. To this aim, we increase S from 11 to 30 and compute the pole approximation error, defined in the left-hand side of (3.12), for the five poles within K. The results are shown in Figure 2. For fixed $N = 10$, the rate predicted by (3.14) can be observed (the Green's potential for a centered disk coincides with the complex magnitude outside the disk). Remarkably, a similar, if not slightly better, rate of convergence is obtained for increasing $N = S - 1$, despite the absence of theoretical results in this regard.

In Figure 2, we also show the pole approximation error obtained through a RB-like projection technique: given the same sample set Ξ_S, we perform a proper orthogonal decomposition (POD) [35] of the samples, identifying the N dominant "sample directions"; then we project orthogonally the original problem on the subspace spanned by these N directions, and we use the solution of the resulting $N \times N$ eigenvalue prob-
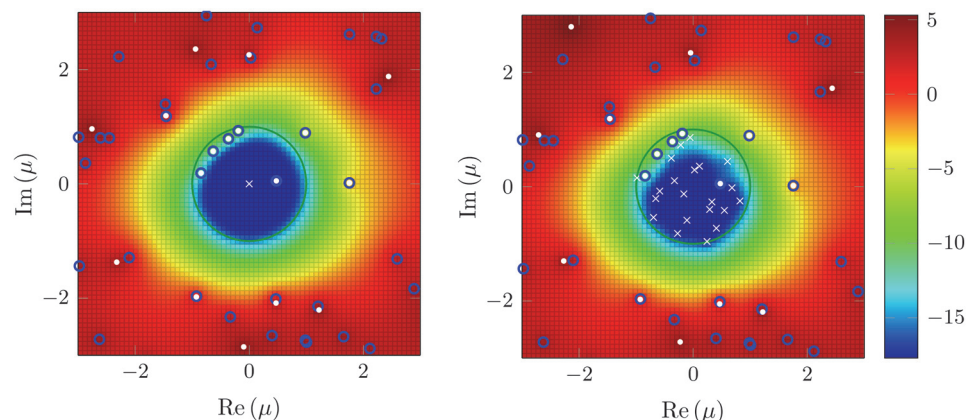
FIG. 3. *Logarithm of the relative error norm* $\|\mathbf{u} - \mathbf{u}_N^{\Xi_S}\|_2 / \|\mathbf{u}\|_2$ *achieved by minimal rational interpolants based on samples (white crosses) at the origin (left) and at quasi-random points (right) with* $N = 20$. *The exact and approximated poles are represented in blue and white, respectively.*
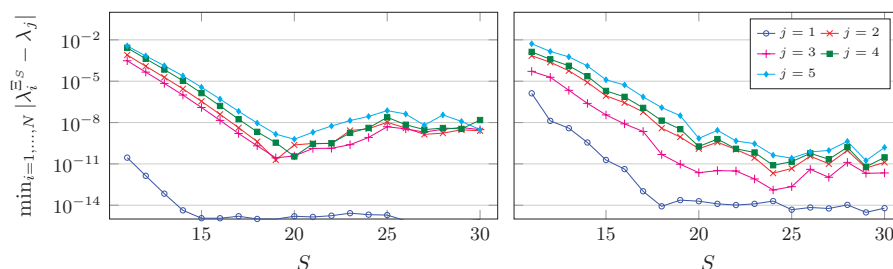

FIG. 4. *Pole approximation error for the poles inside* K. *The samples are all at the origin on the left, whereas on the right the samples are at quasi-random points. In both cases, we keep* $N = S - 1$.

lem to approximate the original one. We can observe that the projection technique performs much worse than minimal rational interpolants for fixed $N = 10$, whereas for $N = S - 1$ the two methods achieve very similar results.

Last, we wish to test two other approximation strategies: a fast LS Padé approximant [10] centered at the origin and a minimal rational interpolant based on quasi-random (Halton) points in the unit disk. We show the error norms and the approximated poles in Figure 3 for $N = S - 1 = 20$. Qualitatively, it appears that, for both choices, the accuracy in the approximation of the resonances is similar to that achieved with samples at the roots of unity. With respect to the [20/20] minimal rational interpolant in Figure 1, the errors for the two new approximants appear smaller within K but larger outside K; this is to be expected, since the new sample points are closer to the origin.

More quantitatively, we show in Figure 4 the convergence of the pole approximation error with respect to S (here as well, we keep $N = S - 1$). For fast LS Padé approximants, when S is small, the error converges quite quickly (the error is uniformly smaller than the one obtained with samples at the roots of unity), but then it stagnates and even increases starting from $S = 20$; this is a symptom of the instability which affects the construction of this centered approximant, namely the computation
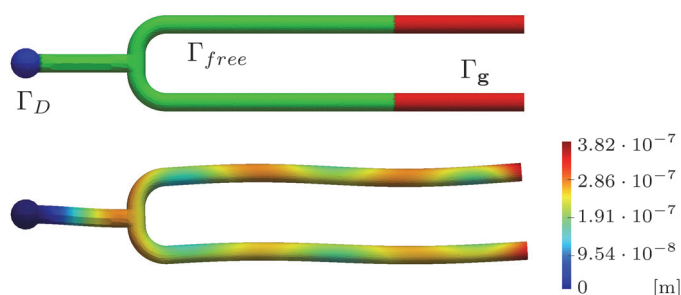
FIG. 5. *Reference domain (above) and the domain warped by the real part of the displacement, magnified by a factor of $3 \cdot 10^3$ (below). The displacement is the solution of (6.2) for frequency $\nu = 10$ kHz and damping coefficient $\eta = 0$ Hz. The total number of degrees of freedom of the discretized problem is* 36330.

of high-order derivatives of a meromorphic function.

The results for quasi-random sample points appear relatively similar, with the error flattening out starting from $S = 20$. In this case, the nonmonotone behavior of the error for large S is due to the interaction between two conflicting effects: the addition of sample points close to the resonances (which would, on its own, improve the accuracy of the approximation) and the ill-conditioning of the interpolation problem (for $S = 30$, the Lebesgue constant is $\sim 10^7$, as opposed to ~ 3 for the 30th roots of unity).

6.2. Time-harmonic vibrations of a tuning fork. Let $\Omega \subset \mathbb{R}^3$ be the region occupied by a tuning fork. Assume that the tuning fork is clamped on a portion Γ_D of the handle. Let $\widetilde{\mathbf{g}}(t, \mathbf{x}) = e^{\iota 2\pi \overline{\nu} t}\mathbf{g}(\mathbf{x})$ be a pressure pulse (sinusoidal in time and Gaussian in space) impinging on a portion $\Gamma_{\mathbf{g}}$ of the exterior of the tuning fork, which we assume to be sound-hard. See Figure 5 for a graphical representation of the setup.

We are interested in computing the time-harmonic deformation $\mathbf{u}(\nu)$ with frequency ν which the tuning fork undergoes. Assuming the tuning fork to have density ρ and linear stress constitutive relation $\sigma(\cdot)$, the displacement $\mathbf{u}(\nu)$ can be found by solving a (damped) time-harmonic linear elasticity problem

$$
(6.2) \quad
\begin{cases}
-\nabla \cdot \sigma(\mathbf{u}(\nu)) - 2\pi\nu(2\pi\nu - \eta\iota)\rho\mathbf{u}(\nu) = \mathbf{0} & \text{in } \Omega, \\
\mathbf{u}(\nu) = \mathbf{0} & \text{on } \Gamma_D, \\
\sigma(\mathbf{u}(\nu))\mathbf{n} = \mathbf{g} & \text{on } \Gamma_{\mathbf{g}}, \\
\sigma(\mathbf{u}(\nu))\mathbf{n} = \mathbf{0} & \text{on } \Gamma_{free} = \partial\Omega \setminus \Gamma_D \setminus \Gamma_{\mathbf{g}}.
\end{cases}
$$

As the V-inner product we choose the elastic energy inner product

$$
\langle \mathbf{u}, \mathbf{v} \rangle_V = \int_\Omega \sigma(\mathbf{u}) : \varepsilon(\mathbf{v}) \mathrm{d}\mathbf{x} = \int_\Omega \sigma(\mathbf{u}) : \frac{\nabla\mathbf{v} + \nabla\mathbf{v}^\top}{2} \mathrm{d}\mathbf{x}.
$$

In particular, we choose to approximate problem (6.2) using $\mathbb{P}_1$ FE on a sufficiently fine tetrahedral discretization of Ω. The FE discretization of (6.2) defines a parametric problem of the form (2.5), whose solution map $\mathbf{u}$ can be proven meromorphic using compactness arguments [11, 31]. Thus, we decide to approximate it for $\nu \in K = [10, 40]$ kHz with minimal rational interpolants relying on samples at the Chebyshev nodes of K. The chosen polynomial norm $\|\cdot\|_N$ is the one induced by the family of Chebyshev polynomials over K; see subsection 3.1.
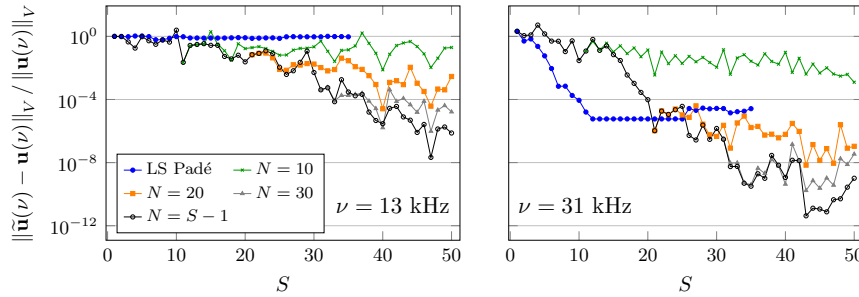
FIG. 6. *Convergence of the approximation error in the energy norm with respect to the number of samples at $\nu = 13$ kHz (left) and $\nu = 31$ kHz (right).*
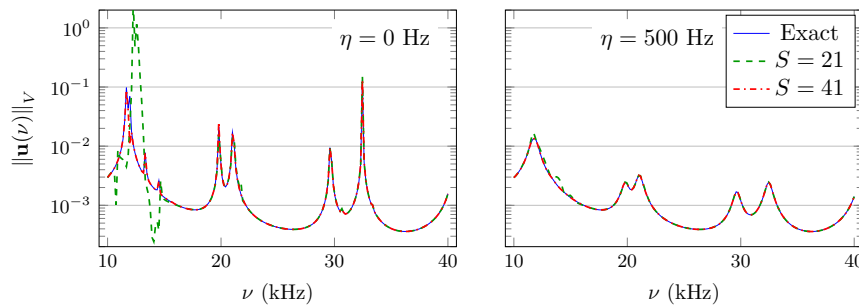

FIG. 7. *Energy norm of the solution map* **u** *of* (6.2) *in blue (solid) for* $\eta \in \{0, 500\}$ *Hz. In green (dashed) and red (dashed-dotted), the norm of the* [20/20] *and* [40/40] *minimal rational interpolants of* **u**.

First, we investigate the behavior of the relative approximation error $\frac{\|\widetilde{\mathbf{u}}(\nu) - \mathbf{u}(\nu)\|_V}{\|\mathbf{u}(\nu)\|_V}$ at $\nu = 13$ kHz and $\nu = 31$ kHz for $\eta = 0$ Hz. We apply different approximation strategies: for $S = 1, \ldots, 51$, we consider $[S - 1/S - 1]$ fast LS Padé approximants [10] with S samples at the center of K, $[S - 1/N]$ minimal rational approximants for fixed $N = 10, 20, 30$, and diagonal $[S - 1/S - 1]$ minimal rational approximants.

The results are shown in Figure 6. As expected from Theorem 3.9, minimal rational approximants of type $[S - 1/10]$ do not converge (K contains 16 resonances), whereas the ones of types $[S - 1/20]$ and $[S - 1/30]$ do; we can observe that the diagonal approximant $[S - 1/S - 1]$ yields the best results overall.

The error achieved by fast LS Padé approximants converges extremely quickly at $\nu = 31$ kHz for small S, but no convergence can be observed at $\nu = 13$ kHz; this is reasonable, since this type of approximant, relying on information at a single point, can deliver an extremely accurate surrogate close to the center, but the quality of the approximation degrades as we move further away. Also, we observe that the convergence plots for fast LS Padé approximants stop decreasing already at $S = 11$ and terminate abruptly at $S = 35$: for $S \geq 36$, the construction of the surrogate model fails, due to conditioning issues in the evaluation of the derivatives of the solution map and in the application of the (Taylor) interpolation operator [10, 12].

The norm of the solution map is shown in Figure 7 for $\eta \in \{0, 500\}$ Hz, along with the approximated norm profiles, obtained by minimal rational interpolants of types [20/20] and [40/40]. If $S = 21$, for both values of η we can observe that the approximation of $\|\mathbf{u}\|_V$ is quite accurate, except for the low-frequency region, where
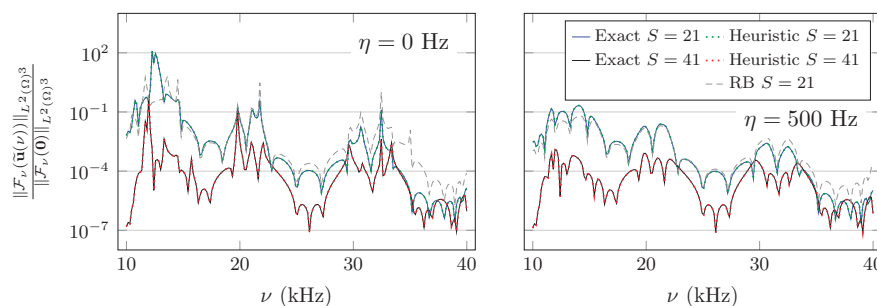
FIG. 8. *Relative $L^2(\Omega)^3$-norm of the residual for problem* (6.2) *and* $\widetilde{\mathbf{u}} = \mathbf{u}_{20}^{\Xi_{21}}$ *(blue) and* $\widetilde{\mathbf{u}} = \mathbf{u}_{40}^{\Xi_{41}}$ *(black). In green and red are the heuristic approximations* (4.5) *of the residual. The dashed grey line is the residual of the RB approximation of* $\mathbf{u}$ *with samples at* Ξ_{21}.

the approximant for $\eta = 0$ Hz actually appears quite unstable. This behavior can be improved by increasing the number of samples to $S = 41$.

In a (more) realistic application, in order to determine whether there is an actual need for more than 21 samples, one could rely on residual-based estimators such as (4.1). To this aim, we can apply (4.3) to obtain an expression for the residual in the $L^2(\Omega)^3$-norm. The evaluation of (4.3) on a fine grid is shown in Figure 8.

In the same figure, we also show the values obtained through the simplified residual estimator (4.6). In particular, in order to evaluate the estimator, we need to compute or approximate the operator norm $\|F_1\|_{\mathcal{L}(V;L^2(\Omega)^3)}$, with F_1 the derivative of the operator in (6.2) at some point μ'; see section 4. We manage to avoid this, thus keeping the estimator as nonintrusive as possible, by the following argument: if we assume that (4.5) holds approximately (the problem is quadratic and not linear in the parameter), we can identify a constant C, independent of μ, for which, employing the notation of section 4,

$$\left\|\mathcal{F}_\mu\left(u_N^{Z_S}(\mu)\right)\right\|_W \approx C\left|\frac{\omega^{Z_S}(\mu)}{Q_N^{Z_S}(\mu)}\right|.$$

In practice, the value of C can be easily found by evaluating (offline and at a relatively low computational cost) the exact value of the residual at some point $\mu' \in K \setminus Z_S$, so that the value of C can be set to

$$C = \left\|\mathcal{F}_\mu\left(u_N^{Z_S}(\mu')\right)\right\|_W\left|\frac{Q_N^{Z_S}(\mu')}{\omega^{Z_S}(\mu')}\right|.$$

In our example, we choose $\mu' = 25.5$ kHz, and we can observe the corresponding residual estimator to be extremely accurate. Indeed, the values obtained with the two estimators never differ by more than 1%, and the corresponding curves cannot be distinguished in the plot.

Finally, in Figure 8 we also perform a comparison between minimal rational interpolants and RBs by showing the norm of the residual obtained with the RB approximant of (6.2) computed from samples at Ξ_{21}. We can observe that the two residuals are quite similar in terms of both order of magnitude and behavior with respect to ν. In particular, simplifying grossly, minimal rational interpolants seem to have a very slight edge on RBs at higher frequencies, while the opposite appears to be true at lower frequencies.

7. Conclusions. In this paper, we have presented minimal rational interpolants as the natural extension of fast LS-Padé approximants [10], and we have shown that most of the theoretical properties of this latter technique generalize nicely to the former. The new method can be considered superior to the original one for two main reasons:

- The stability is greatly improved, since computing derivatives of the solution map accumulates numerical errors; the Arnoldi-like reorthogonalization technique described in [10] can somehow limit but not totally eliminate this issue.
- The distributed sampling helps to achieve a globally (e.g., using the $L^\infty(K)$-metric) small approximation error, which is the target in most applications; of course, an interpolatory approach cannot be as accurate close to the samples as a Taylor-type approximation, assuming the number of samples to be the same.

Our numerical example has shown minimal rational interpolants to perform well in the MOR of a normal eigenproblem and of a time-harmonic problem. In particular, we have shown the effectiveness of a heuristic and cheap to compute a posteriori residual estimator. Accordingly, we deem it of interest to investigate a possible greedy-type algorithm, in the same spirit as the one for the RB method, but with minimal rational interpolants replacing Galerkin projection. In particular, it is still unclear how well the approximation properties of the method are preserved if the samples are not computed at optimal (Fejér) points.

REFERENCES

[1] A. ANTOULAS, *Approximation of large-scale dynamical systems: An overview*, IFAC Proc. Vol., 37 (2004), pp. 19–28, https://doi.org/10.1016/s1474-6670(17)31584-7.

[2] A. ANTOULAS, C. BEATTIE, AND S. GUGERCIN, *Interpolatory Model Reduction of Large-Scale Dynamical Systems*, Springer, Boston, MA, 2010, pp. 3–58, https://doi.org/10.1007/978-1-4419-5757-3_1.

[3] J. ASH, A. BERELE, AND S. CATOIU, *Plausible and genuine extensions of L'Hospital's rule*, Math. Mag., 85 (2012), pp. 52–60.

[4] Z. BAI, J. DEMMEL, J. DONGARRA, A. RUHE, AND H. VAN DER VORST, *Templates for the Solution of Algebraic Eigenvalue Problems*, SIAM, Philadelphia, 2000, https://doi.org/10.1137/1.9780898719581.

[5] G. A. BAKER AND P. R. GRAVES-MORRIS, *Padé Approximants*, Cambridge University Press, Cambridge, UK, 1996.

[6] U. BAUR, C. BEATTIE, P. BENNER, AND S. GUGERCIN, *Interpolatory projection methods for parameterized model reduction*, SIAM J. Sci. Comput., 33 (2011), pp. 2489–2518, https://doi.org/10.1137/090776925.

[7] C. BEATTIE AND S. GUGERCIN, *Model reduction by rational interpolation*, Comput. Sci. Eng., 15 (2014), pp. 297–334.

[8] J.-P. BERRUT AND L. N. TREFETHEN, *Barycentric Lagrange interpolation*, SIAM Rev., 46 (2004), pp. 501–517, https://doi.org/10.1137/s0036144502417715.

[9] F. BONIZZONI, F. NOBILE, AND I. PERUGIA, *Convergence analysis of Padé approximations for Helmholtz frequency response problems*, ESAIM Math. Model. Numer. Anal., 52 (2018), pp. 1261–1284, https://doi.org/10.1051/m2an/2017050.

[10] F. BONIZZONI, F. NOBILE, I. PERUGIA, AND D. PRADOVERA, *Fast least-squares Padé approximation of problems with normal operators and meromorphic structure*, Math. Comp., 89 (2020), pp. 1229–1257, https://doi.org/10.1090/mcom/3511.

[11] F. BONIZZONI, F. NOBILE, I. PERUGIA, AND D. PRADOVERA, *Least-squares Padé approximation of parametric and stochastic Helmholtz maps*, Adv. Comput. Math., 46 (2020), 46, https://doi.org/10.1007/s10444-020-09749-3.

[12] F. Bonizzoni and D. Pradovera, *Distributed sampling for rational approximation of the acoustic scattering of an airfoil*, PAMM Proc. Appl. Math. Mech., 19 (2019).

[13] L. Bos and N. Levenberg, *On the calculation of approximate Fekete points: The univariate case*, Electron. Trans. Numer. Anal., 30 (2008), pp. 377–397.

[14] K. Carlberg, C. Farhat, J. Cortial, and D. Amsallem, *The GNAT method for nonlinear model reduction: Effective implementation and application to computational fluid dynamics and turbulent flows*, J. Comput. Phys., 242 (2013), pp. 623–647, https://doi.org/10.1016/j.jcp.2013.02.028.

[15] S. Chaturantabut and D. C. Sorensen, *Nonlinear model reduction via discrete empirical interpolation*, SIAM J. Sci. Comput., 32 (2010), pp. 2737–2764, https://doi.org/10.1137/090766498.

[16] Y. Chen, J. S. Hesthaven, Y. Maday, and J. Rodríguez, *Certified reduced basis methods and output bounds for the harmonic Maxwell's equations*, SIAM J. Sci. Comput., 32 (2010), pp. 970–996, https://doi.org/10.1137/09075250X.

[17] M. Chevreuil and A. Nouy, *Model order reduction based on proper generalized decomposition for the propagation of uncertainties in structural dynamics*, Internat. J. Numer. Methods Engrg., 89 (2012), pp. 241–268, https://doi.org/10.1002/nme.3249.

[18] Z. Drmač, S. Gugercin, and C. Beattie, *Quadrature-based vector fitting for discretized $\mathcal{H}_2$ approximation*, SIAM J. Sci. Comput., 37 (2015), pp. A625–A652, https://doi.org/10.1137/140961511.

[19] F. Ferranti, L. Knockaert, and T. Dhaene, *Passivity-preserving parametric macromodeling by means of scaled and shifted state-space systems*, IEEE Trans. Microw. Theory Tech., 59 (2011), pp. 2394–2403, https://doi.org/10.1109/TMTT.2011.2164551.

[20] R. W. Freund, *Model reduction methods based on Krylov subspaces*, Acta Numer., 12 (2003), pp. 267–319, https://doi.org/10.1017/S0962492902000120.

[21] M. Grepl, Y. Maday, N. Nguyen, and A. Patera, *Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations*, ESAIM Math. Model. Numer. Anal., 41 (2007), pp. 575–605, https://doi.org/10.1051/m2an:2007031.

[22] E. Grimme, *Krylov Projection Methods for Model Reduction*, Ph.D. thesis, University of Illinois at Urbana-Champaign, Champaign, IL, 1997.

[23] S. Grivet-Talocia and B. Gustavsen, *Passive Macromodeling: Theory and Applications*, Wiley Ser. Microw. Opt. Eng., John Wiley & Sons, Hoboken, NJ, 2015, https://doi.org/10.1002/9781119140931.

[24] P. Guillaume, *Nonlinear eigenproblems*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 575–595, https://doi.org/10.1137/S0895479897324172.

[25] B. Gustavsen and A. Semlyen, *Rational approximation of frequency domain responses by vector fitting*, IEEE Trans. Power Deliv., 14 (1999), pp. 1052–1061, https://doi.org/10.1109/61.772353.

[26] J. Hesthaven, G. Rozza, and B. Stamm, *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, BCAM SpringerBriefs, Bilbao, Spain, 2016, https://doi.org/10.1007/978-3-319-22470-1.

[27] E. Hille, *Analytic Function Theory*, 2nd ed., Chelsea, New York, 1965.

[28] D. Huynh, G. Rozza, S. Sen, and A. Patera, *A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants*, C. R. Math. Acad. Sci. Paris, 345 (2007), pp. 473–478, https://doi.org/10.1016/j.crma.2007.09.019.

[29] A. C. Ionita and A. C. Antoulas, *Data-driven parametrized model reduction in the Loewner framework*, SIAM J. Sci. Comput., 36 (2014), pp. A984–A1007, https://doi.org/10.1137/130914619.

[30] C. Lei, Y. Wang, Q. Chen, and N. Wong, *On Vector Fitting Methods in Signal / Power Integrity Applications*, Lect. Notes Eng. Comput. Sci. 2181, Springer, Berlin, 2010.

[31] M. Lenoir, M. Vullierme-Ledard, and C. Hazard, *Variational formulations for the determination of resonant states in scattering problems*, SIAM J. Math. Anal., 23 (1992), pp. 579–608, https://doi.org/10.1137/0523030.

[32] P. Monk, *Finite Element Methods for Maxwell's Equations*, Clarendon, Oxford, UK, 2003.

[33] H. Panzer, J. Mohring, R. Eid, and B. Lohmann, *Parametric model order reduction by matrix interpolation*, At-Automatisierungstechnik, 58 (2010), pp. 475–484, https://doi.org/10.1524/auto.2010.0863.

[34] D. Pradovera, *Interpolatory rational model order reduction of parametric problems lacking uniform inf-sup stability - numerical tests*, (2020), https://doi.org/10.5281/ZENODO.3653702.

[35] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced Basis Methods for Partial Differential*

Equations: An Introduction, Unitext 92, Springer, Cham, 2016.

[36] G. ROZZA, D. B. HUYNH, AND A. T. PATERA, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics, Arch. Comput. Methods Eng., 15 (2008), pp. 229–275, https://doi.org/10.1007/s11831-008-9019-9.

[37] S. STEINBERG, Meromorphic families of compact operators, Arch. Rational Mech. Anal., 31 (1968), pp. 372–379, https://doi.org/10.1007/BF00251419.

[38] J. L. WALSH, Interpolation and Approximation by Rational Functions in the Complex Domain, AMS, Providence, RI, 1969.

[39] D. WEILE, E. MICHIELSSEN, E. GRIMME, AND K. GALLIVAN, A method for generating rational interpolant reduced order models of two-parameter linear systems, Appl. Math. Lett., 12 (1999), pp. 93–102, https://doi.org/10.1016/S0893-9659(99)00063-4.

[40] Y. Q. XIAO, S. GRIVET-TALOCIA, P. MANFREDI, AND R. KHAZAKA, A novel framework for parametric Loewner matrix interpolation, IEEE Trans. Compon. Packag. Manuf. Technol., 9 (2019), pp. 2404–2417, https://doi.org/10.1109/tcpmt.2019.2948802.

[41] Y. YUE, L. FENG, AND P. BENNER, Reduced-order modelling of parametric systems via interpolation of heterogeneous surrogates, Adv. Model. Simul. Eng. Sci., 6 (2019), 10, https://doi.org/10.1186/s40323-019-0134-y.