

Interpolatory rational model order reduction of parametric problems lacking uniform inf-sup stability

Davide Pradovera^{§*}

[§] CSQI – MATH, École Polytechnique Fédérale de Lausanne
Station 8, CH-1015 Lausanne, Switzerland
davide.pradovera@epfl.ch

Keywords: model order reduction, rational approximation, eigenvalue estimation, Hilbert space-valued meromorphic maps, non-coercive parametric PDEs, Helmholtz equation, frequency response.

AMS Subject Classification: 30D30, 41A20, 41A25, 35J05, 35P15, 65D15.

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 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.

1 Introduction

Mathematical models based on Partial Differential Equations 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 *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 [9, 10, 11, 15, 21, 26, 27, 30, 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

*D. Pradovera has been funded by the Swiss National Science Foundation (SNSF) through project 182236.

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 pre-computed 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 characterized by some irregularity of the FOM (e.g. roughness, non-linearity, lack of stability, etc. of the differential operator therein).

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

$$\text{for } k \in K \subset \mathbb{C}, \text{ find } u(k) \in V \text{ s.t. } -\Delta u(k) - k^2 u(k) \stackrel{(W)}{=} f \in W, \quad (1)$$

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), the problem above may lack inf-sup stability over K [32], due to the presence of resonant wavenumbers. If one applies carelessly the RB method (in both its main versions, POD and greedy [35, 36]) to (1), one may observe the appearance of spurious (“non-physical”) resonances in the surrogate model. The roots of this issue can be traced back to linear algebra, in particular projection-based numerical methods for eigenvalue problems, such as the Arnoldi or Lanczos techniques [19], which show analogous behaviors.

Discussions concerning this effect can be found in works related to MOR methods for dynamical systems, where problems of similar form are quite common, due to the need for frequency-domain computations. For instance, in Krylov subspace methods [1, 13, 16, 18, 22] one employs RB-like projection-based ideas, quite often suffering from the issues described above.

Usually, if the number of parameters is small (often, just 1), 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 representative of this class of methods is the vector fitting (VF) algorithm [22, 23, 25, 31, 39]. VF is tailored to the specific structure of the FOM (or, more precisely, of its *solution map*, see Section 2). Indeed, rational functions are employed, with the objective of representing each resonance of the FOM by a pole of the approximant, i.e. a root of its denominator. 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 small 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 [5, 7] and a fast [6] 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:

- (i) LS Padé approximants, particularly in their fast version, do not intrinsically need any oversampling (i.e. taking more samples of the FOM than the actual amount needed to ensure existence and uniqueness of the surrogate model), and they have been shown to provide good results even without it; conversely, a heuristics commonly employed in VF is to choose a reasonably high “sampling density” [14], namely at least twice as many samples as the total number of resonances; in particular, VF requires, by construction, at least as many samples as the number of parameters in the approximant, while fast LS Padé approximants can get away with half as many;
- (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;

actually, if the (scalar) quantity of interest is a linear functional of the solution of the FOM, fast LS Padé approximants can still be employed: it suffices to build a rational approximation of the solution map, and then to apply the desired linear functional to such 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 [5, 6], while the convergence behavior of the VF error (even with respect to the number of VF iterations) has, to our knowledge, yet to be analyzed thoroughly [14];
- (iv) LS Padé approximants are computed from moments of the solution at a single parameter, the explicit computation of which may feature numerical instabilities; instead, VF uses distributed samples in a LS-Lagrange (or Hermite) interpolation fashion, which does not suffer from the same ill-conditioning issues.

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 to. 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. In Section 5 we show some numerical examples to verify our theoretical claims and show the effectiveness of our error indicator. Some concluding remarks are in Section 6.

Before we proceed, we wish to mention some details on the application of our technique to problems depending on multiple parameters. Some work on multi-variate rational approximation has already been carried out, considering either scalar Padé approximation [12, 24], or Krylov subspace methods for the bivariate case [39]. We do not discuss the multivariate case here, leaving as a possible future research direction the integration of our rational approximants in such frameworks.

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 $u : K \rightarrow V$.

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 technique qualifies as *non-intrusive*, 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 $N \in \{0, \dots, S - 1\}$, as well as the set of normalized polynomials of degree at most N

$$\mathbb{P}_N^*(\mathbb{C}) = \{Q \in \mathbb{P}_N(\mathbb{C}), \|Q\|_N = 1\}, \quad (2)$$

with $\|\cdot\|_N$ 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 $\mathbb{P}_N^*(\mathbb{C})$.

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

$$\mathbb{P}_{S-1}(\mathbb{C}; V) = \left\{ \mu \mapsto \sum_{i=0}^{S-1} p_i \mu^i, p_i \in V \text{ for } i = 0, \dots, S-1 \right\} \quad (3)$$

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

Definition 1 (Minimal rational interpolants) *The minimal rational interpolant of u of type $[S-1/N]$ with samples at Ξ_S is the rational function*

$$u_N^{\Xi_S}(\mu) = \frac{I^{\Xi_S}(uQ_N^{\Xi_S})(\mu)}{Q_N^{\Xi_S}(\mu)}, \quad \mu \in \mathbb{C}, Q_N^{\Xi_S}(\mu) \neq 0, \quad (4)$$

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

$$j^{\Xi_S}(Q)^2 = \left\| \frac{1}{(S-1)!} \frac{d^{S-1}}{d\mu^{S-1}} I^{\Xi_S}(uQ) \right\|_V^2, \quad (5)$$

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

We remark that, since j^{Ξ_S} is convex, the existence of a minimizer $Q_N^{\Xi_S}$ is guaranteed by the compactness of $\mathbb{P}_N^*(\mathbb{C})$. 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’’

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

- (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 $Q \in \mathbb{P}_N^*(\mathbb{C})$ by its component vector \mathbf{q} (by construction, $\|\mathbf{q}\|_2 = 1$);

- (iii) by linearity, rewrite the interpolant $I^{\Xi_S}(uQ)$ as $\sum_{i=1}^S \sum_{l=0}^N q_l I^{\Xi_S}(w_i \psi_l) \varphi_i$, so that

$$j^{\Xi_S}(Q)^2 = \sum_{i=1}^S \left| \sum_{l=0}^N \frac{1}{(S-1)!} \frac{d^{S-1}}{d\mu^{S-1}} I^{\Xi_S}(w_i \psi_l) q_l \right|^2 = \mathbf{q}^H \Psi^H \Psi \mathbf{q}$$

(H is employed to denote conjugate transposition), with

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

- (iv) after assembling the Gramian factor Ψ , obtain the required minimizer (more accurately, its components with respect to $\{\psi_l\}_{l=0}^N$) through a SVD.

For stability reasons it may be wise to replace I^{Ξ_S} with a Least Squares-type interpolator $I_D^{\Xi_S}$, which yields as interpolant a degree $D < S-1$ polynomial minimizing the approximation error, computed through some discrete norm at the points Ξ_S , e.g.

$$I_D^{\Xi_S}(\phi) = \arg \min_{P \in \mathbb{P}_D(\mathbb{C}; V)} \sum_{\mu \in \Xi_S} w_\mu^2 \|\phi(\mu) - P(\mu)\|_V^2.$$

In this case, the resulting minimal rational interpolant is of type $[D/N]$, and it may be necessary to perform small adjustments to the definitions and the properties discussed below.

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:

$$\text{for } \mu \in K \subset \mathbb{C}, \text{ find } u(\mu) \in \text{Dom}(\mathcal{F}_\mu) \text{ s.t. } \mathcal{F}_\mu(u(\mu)) = 0, \quad (6)$$

with $\mathcal{F}_\mu : V \supset \text{Dom}(\mathcal{F}_\mu) \rightarrow V$ a family of densely defined operators, depending continuously on μ . Provided the set $K^* \subset K$ of values of μ , for which a unique $u(\mu)$ solving (6) exists, is non-empty, we focus on the *solution map* associated to (6):

$$\begin{aligned} u : K^* &\rightarrow V \\ \mu &\mapsto u(\mu) \text{ solution of (6)}. \end{aligned} \quad (7)$$

In the following, we restrict our focus to a special class of family of operators $\{\mathcal{F}_\mu\}_{\mu \in K}$, namely diagonal perturbations $\mathcal{L} + \mu\mathcal{I}$ of a bijective densely defined operator $\mathcal{L} : V \supset \text{Dom}(\mathcal{L}) \rightarrow V$, whose inverse is compact and normal. Due to the spectral properties of \mathcal{L} , see [6], such operator families induce solution maps (7) that are meromorphic with simple poles over \mathbb{C} , and can be expressed as

$$u(\mu) = \sum_{\lambda \in \Lambda} \frac{v_\lambda}{\lambda - \mu}, \quad (8)$$

where $\{v_\lambda\}_{\lambda \in \Lambda}$ is a V -orthogonal set¹, and $\Lambda \subset \mathbb{C}$ is countable, with no finite limit point. If Λ is not finite, (8) has to be understood in the sense of V -convergence, for an arbitrary ordering of Λ . Moreover, the set of excluded parameters $K \setminus K^*$ can be characterized as the intersection between Λ and K . In the following we will refer to the elements of Λ interchangeably as poles or resonances of u or, by transitivity, of (6).

In our theoretical derivation, we will assume that (8) holds, with Λ having infinitely many elements, for the sake of notation. Analogous conclusions can be obtained for finite Λ , either by introducing slight modifications in the statements of the results, or, with an abuse of notation, by extending Λ to a countably infinite set through the addition of fictitious poles at ∞ .

3 Convergence theory

3.1 Preliminaries for convergence theory

In the next sections, 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 . We choose to follow a potential theory-inspired approach, and we start by summarizing some classical results that can be found e.g. in [38].

For completeness, we wish to remark that equivalent results could be obtained starting from properties of lemniscates instead of those of Green's functions, see e.g. [3, Section 6.6], [38, Section 3.3], and [33]. While the path we choose to follow leads to results which are somewhat nicer and more easily readable, the theory based on lemniscates applies also to more general sampling strategies (e.g. confluent points), which are implicitly excluded in our derivations, see Theorem 1.

¹Most of the results in the following rely quite heavily on this orthogonality, but can be generalized to the case of $\{v_\lambda\}_{\lambda \in \Lambda}$ being just linearly independent, e.g. when the perturbation of \mathcal{L} is not diagonal. The price to pay is a ‘‘collinearity’’ constant in front of most results. Moreover, all results in Section 3.2 hold with halved convergence rate, in direct analogy to the lack of symmetry/Hermiticity in the matricial case [19].

Assume that $K \subset \mathbb{C}$ is either the closure of the finite region bounded by a Jordan curve or a line segment with positive length. There exist a (unique) positive real number $\text{Cap}(K)$ and a continuous bijective conformal mapping

$$\phi_K : \overline{\mathbb{C} \setminus K} \rightarrow \mathbb{C} \setminus \mathcal{B}(0, \text{Cap}(K)) = \{z \in \mathbb{C}, |z| \geq \text{Cap}(K)\}, \quad (9)$$

holomorphic over $\mathbb{C} \setminus K$, such that $|\phi_K| = \text{Cap}(K)$ over ∂K , $\phi_K(\infty) = \infty$, and $\phi'_K(\infty) = 1$.

The value $\text{Cap}(K)$ is the *logarithmic capacity* of K , and captures the scaling of the set. It is a powerful concept in complex analysis, algebraic geometry, and approximation theory, and we refer to the monograph [28] for a thorough introduction to the subject of logarithmic capacity. 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 the complex plane: for any compact set $A \subset \mathbb{C}$ it holds

$$\sqrt{\frac{\ell(A)}{\pi}} \leq \text{Cap}(A) \leq \frac{1}{2} \text{diam}(A) = \frac{1}{2} \max_{z, w \in A} |z - w|, \quad (10)$$

with ℓ denoting the Lebesgue measure over \mathbb{C} . In particular, zero-measure sets are not guaranteed to have capacity zero (e.g. $\text{Cap}([-2, 2]) = 1$), and both bounds in (10) are attained if A is a closed disk.

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 :

$$\Phi_K(\mu) = \begin{cases} |\phi_K(\mu)| & \text{if } \mu \notin K, \\ \text{Cap}(K) & \text{if } \mu \in K \end{cases}. \quad (11)$$

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

$$\Phi_K(\lambda_1) \leq \Phi_K(\lambda_2) \leq \dots \quad (12)$$

Moreover, we require that $\Phi_K(\lambda_{N+1}) > \text{Cap}(K)$, i.e. that the number of poles within K is at most N , see (11) and (12). Indeed, we will show in the following that minimal rational interpolants fail to differentiate between poles of u within K ; hence, without the assumption $\Phi_K(\lambda_{N+1}) > \text{Cap}(K)$, the “unidentifiable” poles would distort the approximant and, in general, prevent convergence to the target solution map.

Now, for a given sample set Ξ_S , we define the corresponding nodal polynomial

$$\omega^{\Xi_S}(\mu) = \prod_{\mu' \in \Xi_S} (\mu - \mu'). \quad (13)$$

This allows us to state the result which motivates our interest in Green's potentials.

Theorem 1 ([17, 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 $(\Xi_S^K)_{S \in \mathbb{N}}$, called Fejér points for K , such that*

$$\limsup_{S \rightarrow \infty} \left| \omega^{\Xi_S^K}(\cdot) \right|^{1/S} \leq \Phi_K(\cdot) \quad \text{uniformly over any compact subset of } \mathbb{C}, \quad (14)$$

as well as

$$\lim_{S \rightarrow \infty} \left| \omega^{\Xi_S^K}(\cdot) \right|^{1/S} = \Phi_K(\cdot) \quad \text{uniformly over any compact subset of } \mathbb{C} \setminus \partial K. \quad (15)$$

The Fejér points of order S for the set K admit an explicit characterization in terms of the conformal mapping ϕ_K : if K is the closure of the finite region bounded by a Jordan curve, then

$$\Xi_S^K = \left\{ \phi_K^{-1} \left(\text{Cap}(K) \exp \left\{ \left(\frac{2\pi}{S} j + \theta \right) \iota \right\} \right), j = 1, \dots, S \right\}, \quad (16)$$

with ι the imaginary unit, and $\theta \in \mathbb{R}$ arbitrary; if K is a line segment, they simply coincide with the Chebyshev nodes of order S for K . \square

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. We stress once more that we introduce this assumption simply to streamline the derivation of the theoretical results. Our theory can be extended to more general choices of the sample set Ξ_S : the only requirement is the existence of some reasonable bound on the asymptotic behavior of the nodal polynomial, in the same spirit as (14). In fact, some results can be generalized even in absence of such control for large S .

A second class of considerations involves the denominator space $\mathbb{P}_N^*(\mathbb{C})$, whose definition revolves around 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 1 *Let $\mu_0 \in \mathbb{C}$ be arbitrary but fixed. There exist positive constants R_{μ_0} , c_{μ_0} , and C_{μ_0} , all independent of N , such that*

$$c_{\mu_0} \prod_{j=1}^D |\mu_0 - z_j| \leq \left\| \prod_{j=1}^D (\cdot - z_j) \right\|_N \leq C_{\mu_0} \prod_{j=1}^D (R_{\mu_0} + |\mu_0 - z_j|), \quad (17)$$

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

In [6] it was shown that Assumption 1 holds (with $R_{\mu_0} = c_{\mu_0} = C_{\mu_0} = 1$) for the “monomial” norm $\|\cdot\|_N = \|\cdot\|_{\mu_0, N}$ defined as

$$\|Q\|_{\mu_0, N} = \sqrt{\sum_{j=0}^N |q_j|^2}, \quad \text{where } Q = \sum_{j=0}^N q_j (\cdot - \mu_0)^j. \quad (18)$$

We describe in Appendix A a more general framework, not relying on the often ill-conditioned monomial basis, where Assumption 1 can still be shown to hold.

Many of our results do not rely on Assumption 1, but exploit the following weaker result.

Lemma 1 *Let $\mu_0 \in \mathbb{C}$ be arbitrary but fixed. There exist positive constants R_{μ_0} (independent of N), $c_{\mu_0, N}$, and $C_{\mu_0, N}$, such that*

$$c_{\mu_0, N} \prod_{j=1}^D |\mu_0 - z_j| \leq \left\| \prod_{j=1}^D (\cdot - z_j) \right\|_N \leq C_{\mu_0, N} \prod_{j=1}^D (R_{\mu_0} + |\mu_0 - z_j|), \quad (19)$$

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

Proof. As described above, $\|\cdot\|_N = \|\cdot\|_{\mu_0, N}$ satisfies (19). The result follows by the equivalence of all norms on a finite-dimensional space. \square

3.2 Convergence of minimal rational interpolant poles

In the present section we prove that the denominator $Q_N^{\Xi_S}$ described in Definition 1 performs well in identifying the approximate position of some of the poles Λ of the solution map, among which those within the parameter domain K . Results of similar flavor can be easily obtained for poles located outside K , provided they are close enough, in the sense of (12).

As a preliminary step, it is useful to exploit the definition of $Q_N^{\Xi_S}$ to bound the minimal value of j^{Ξ_S} over $\mathbb{P}_N^*(\mathbb{C})$.

Lemma 2 *Let $Q \in \mathbb{P}_N(\mathbb{C})$. Then*

$$j^{\Xi_S}(Q)^2 = \sum_{\lambda \in \Lambda} \frac{\|v_\lambda\|_V^2}{|\omega^{\Xi_S}(\lambda)|^2} |Q(\lambda)|^2. \quad (20)$$

Moreover, for any subset Λ_0 of Λ with cardinality at most N , it holds

$$j^{\Xi_S}(Q_N^{\Xi_S}) \leq \frac{\left(\sum_{\lambda \in \Lambda \setminus \Lambda_0} \|v_\lambda\|_V^2\right)^{1/2}}{\|\omega^{\Lambda_0}\|_N} \sup_{\Lambda \setminus \Lambda_0} \left| \frac{\omega^{\Lambda_0}}{\omega^{\Xi_S}} \right|. \quad (21)$$

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

$$j^{\Xi_S}(Q_N^{\Xi_S}) \leq (1 + \varepsilon)^S \frac{\left(\sum_{j > N} \|v_{\lambda_j}\|_V^2\right)^{1/2}}{\left\| \prod_{j=1}^N (\cdot - \lambda_j) \right\|_N} \frac{\prod_{j=1}^N |\lambda_{N+1} - \lambda_j|}{\Phi_K(\lambda_{N+1})^S} \quad \text{for } S > S_{\varepsilon, N}, \quad (22)$$

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 (12), for which (22) holds.

Proof. The interpolation operator I^{Ξ_S} can be expressed in barycentric coordinates as

$$I^{\Xi_S}(\phi) = \sum_{\mu \in \Xi_S} \frac{\phi(\mu) \omega^{\Xi_S}(\cdot)}{(\cdot - \mu) \omega^{\Xi_S'}(\mu)}. \quad (23)$$

Accordingly, since Q is interpolated exactly,

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

The first claim follows by orthonormality of $\{v_\lambda\}_{\lambda \in \Lambda}$.

The optimality of $Q_N^{\Xi_S}$ implies that

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

from which (21) follows.

Now, let $\varepsilon > 0$ and N be fixed, and define $\Lambda_0 = \{\lambda_j\}_{j=1}^N$. It remains to bound the supremum of $r^{\Lambda_0, \Xi_S} = |\omega^{\Lambda_0}/\omega^{\Xi_S}|$ over $\Lambda \setminus \Lambda_0$. 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 (15), there exists $S' = S'(\varepsilon, N, \Lambda, K)$ such that, for $S > S'$,

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

Assume that $\Phi_K(\lambda_{N+1}) < \Phi_K(\lambda_{N+2})$: since $|\omega^{\Lambda_0}|$ is independent of S , the last supremum is attained at λ_{N+1} , provided S is large enough. This yields (22).

If $\Phi_K(\lambda_{N+1}) = \Phi_K(\lambda_{N+\nu}) < \Phi_K(\lambda_{N+\nu+1})$ for some $\nu > 1$, the argument above still holds, provided λ_{N+1} maximizes $|\omega^{\Lambda_0}|$ over $\{\lambda_{N+k}\}_{k=1}^\nu$. \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 *Let N be fixed. For any $\varepsilon > 0$, there exists $S_{\varepsilon,N}$ such that*

$$\left| Q_N^{\Xi_S}(\lambda_j) \right| \leq C_j (1 + \varepsilon)^{4S} \left(\frac{\Phi_K(\lambda_j)}{\Phi_K(\lambda_{N+1})} \right)^{2S} \quad \text{for } S > S_{\varepsilon,N}, \quad (24)$$

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

Proof. Throughout this proof, we identify $\mathbb{P}_N(\mathbb{C})$ with \mathbb{C}^{N+1} (endowed with the Euclidean norm) through some isometry, so that, in particular, $\mathbb{P}_N^*(\mathbb{C})$ corresponds to the boundary of the unit sphere $\partial\mathcal{B}_{N+1}(\mathbf{0}, 1) \subset \mathbb{C}^{N+1}$. A simple duality argument guarantees the existence of a family $\{\omega_\lambda\}_{\lambda \in \Lambda} \subset \mathbb{C}^{N+1}$ such that

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

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

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

From here onward, the proof resembles quite closely that of Lemma 3 in [6], being essentially based on bounding 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 form: for $j = 1, \dots, N$, for large S , it holds

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

with C_j independent of S and ε . From here it suffices to bound $|\omega^{\Xi_S}(\lambda_j)|$ by $((1 + \varepsilon)\Phi_K(\lambda_j))^S$, thanks to (14). \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 (12).

Theorem 2 *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*

$$\min_{i=1, \dots, N} \left| \lambda_i^{\Xi_S} - \lambda_j \right| \leq C_j (1 + \varepsilon)^{4S} \left(\frac{\Phi_K(\lambda_j)}{\Phi_K(\lambda_{N+1})} \right)^{2S} \quad \text{for } S > S_{\varepsilon,N}, \quad (25)$$

for $j = 1, \dots, N$, 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

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

for any arbitrary $\mu_0 \in \mathbb{C}$, see Lemma 1.

From here it suffices to apply the same proof as for Theorem 2 in [6], of course with (24) replacing the corresponding bound for fast LS-Padé denominators. \square

Going back to our original problem formulation, let $0 \leq \nu \leq N$ be such that

$$\text{Cap}(K) = \Phi_K(\lambda_\nu) < \Phi_K(\lambda_{\nu+1}).$$

Theorem 2 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 $Q_{N^S}^{\Xi}$ converging to each of them at rate $\mathcal{O}((\text{Cap}(K)/\Phi_K(\lambda_{N+1}))^{2S})$.

3.2.1 Pole convergence with respect to denominator degree

The statements of Lemma 3 and Theorem 2 may seem somewhat limited due to the fact that the denominator degree N enters in the convergence bounds through the proportionality constants and the smallest value of S for which the bounds hold definitely. 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 overcome this shortcoming. The main result we prove is the following.

Theorem 3 *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 the 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 $Q_{N_k}^{\Xi^{(k)}}$.*

Let Assumption 1 be satisfied for some $\mu_0 \in \mathbb{C}$. Then, for all $j = 1, 2, \dots$, it holds

$$\lim_{k \rightarrow \infty} \min_{i=1, \dots, N_k} |\lambda_i^{(k)} - \lambda_j| = 0. \quad (26)$$

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

$$|\sigma_1 - \mu_0| \leq |\sigma_2 - \mu_0| \leq \dots \quad (27)$$

Let $j \in \mathbb{N}$ be arbitrary. Thanks to Assumption 1, it holds

$$\begin{aligned} |Q_{N_k}^{\Xi^{(k)}}(\sigma_j)| &= \frac{\prod_{i=1}^{N_k} |\sigma_j - \lambda_i^{(k)}|}{\left\| \prod_{i=1}^{N_k} (\cdot - \lambda_i^{(k)}) \right\|_{N_k}} \geq \frac{1}{C_{\mu_0}} \prod_{i=1}^{N_k} \frac{|\sigma_j - \lambda_i^{(k)}|}{R_{\mu_0} + |\sigma_j - \mu_0| + |\sigma_j - \lambda_i^{(k)}|} \\ &= \frac{1}{C_{\mu_0}} \prod_{i=1}^{N_k} \varphi_j(|\sigma_j - \lambda_i^{(k)}|), \end{aligned}$$

with $\varphi_j = \cdot / (R_{\mu_0} + |\sigma_j - \mu_0| + \cdot)$ non-negative and strictly increasing over \mathbb{R}^+ . On the other hand, thanks to Lemma 2 with $\Lambda_0 = \{\sigma_i\}_{i=1}^{N_k}$, as well as Assumption 1, it holds

$$\begin{aligned} j^{\Xi^{(k)}}(Q_{N_k}^{\Xi^{(k)}}) &\leq \frac{\left(\sum_{\lambda \in \Lambda \setminus \Lambda_0} \|v_\lambda\|_V^2 \right)^{1/2}}{\left\| \prod_{i=1}^{N_k} (\cdot - \sigma_i) \right\|_N} \sup_{l > N_k} \frac{\prod_{i=1}^{N_k} |\sigma_l - \sigma_i|}{|\omega^{\Xi^{(k)}}(\sigma_l)|} \\ &\leq \frac{\left(\sum_{\lambda \in \Lambda} \|v_\lambda\|_V^2 \right)^{1/2}}{c_{\mu_0}} \sup_{l > N_k} \left(\frac{1}{|\omega^{\Xi^{(k)}}(\sigma_l)|} \prod_{i=1}^{N_k} \left| \frac{\sigma_l - \sigma_i}{\mu_0 - \sigma_i} \right| \right). \end{aligned}$$

By observing that, thanks to (20),

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

the two bounds above can be combined to obtain

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

i.e., by monotonicity of φ_j ,

$$\varphi_j \left(\min_{i=1, \dots, N_k} \left| \sigma_j - \lambda_i^{(k)} \right| \right) \leq \left(C'_j \sup_{l > N_k} \left(\left| \frac{\omega^{\Xi^{(k)}}(\sigma_j)}{\omega^{\Xi^{(k)}}(\sigma_l)} \right| \prod_{i=1}^{N_k} \left| \frac{\sigma_l - \sigma_i}{\mu_0 - \sigma_i} \right| \right) \right)^{1/N_k}, \quad (28)$$

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 0, the claim follows if we can show that the right hand side of (28) 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 $|\mu_0 - \mu'|$ can be easily upper-bounded 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

$$\begin{aligned} \left| \frac{\omega^{\Xi^{(k)}}(\sigma_j)}{\omega^{\Xi^{(k)}}(\sigma_l)} \right| \prod_{i=1}^{N_k} \left| \frac{\sigma_l - \sigma_i}{\mu_0 - \sigma_i} \right| &\leq \left(\frac{|\mu_0 - \sigma_j| + R_{\mu_0, K}}{|\mu_0 - \sigma_l| - R_{\mu_0, K}} \right)^{S_k} \frac{2^{N_k} |\mu_0 - \sigma_l|^{N_k}}{\prod_{i=1}^{N_k} |\mu_0 - \sigma_i|} \\ &= 2^{N_k} \left(\frac{|\mu_0 - \sigma_j| + R_{\mu_0, K}}{|\mu_0 - \sigma_l| - R_{\mu_0, K}} \right)^{S_k - N_k} \left(\frac{|\mu_0 - \sigma_l|}{|\mu_0 - \sigma_l| - R_{\mu_0, K}} \right)^{N_k} \times \\ &\quad \times \prod_{i=1}^{N_k} \frac{|\mu_0 - \sigma_j| + R_{\mu_0, K}}{|\mu_0 - \sigma_i|}. \end{aligned}$$

Going back to the supremum, it is easy to see that all the terms depending on l in the right hand side above are maximized (over $l > N_k$) by the choice $l = N_k + 1$, thanks to (27). This, after a suitable rearrangement of the terms, leads to

$$\begin{aligned} \varphi_j \left(\min_{i=1, \dots, N_k} \left| \sigma_j - \lambda_i^{(k)} \right| \right) &\leq 2 C'_j{}^{1/N_k} \left(\frac{|\mu_0 - \sigma_j| + R_{\mu_0, K}}{|\mu_0 - \sigma_{N_k+1}| - R_{\mu_0, K}} \right)^{S_k/N_k - 1} \times \\ &\quad \times \frac{|\mu_0 - \sigma_{N_k+1}|}{|\mu_0 - \sigma_{N_k+1}| - R_{\mu_0, K}} \prod_{i=1}^{N_k} \left(\frac{|\mu_0 - \sigma_j| + R_{\mu_0, K}}{|\mu_0 - \sigma_i|} \right)^{1/N_k}. \quad (29) \end{aligned}$$

All of the factors appearing in (29), 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

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

To this aim, the Stolz-Cesàro [2] theorem can be applied to prove that

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

yielding the claim. \square

We can observe that this more general result does not provide any information about 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 (28) and (29) to be reasonably sharp, except in extremely pathological cases.

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

$$\left\| u(\mu) - u_N^{\Xi_S}(\mu) \right\|_V \leq C_E (1 + \varepsilon)^{2S} \left(\frac{\Phi_K(\mu)}{\Phi_K(\lambda_{N+1})} \right)^S \quad \text{for all } \mu \in E, \quad \text{for } S > S_{\varepsilon, N}, \quad (30)$$

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 (23) to obtain

$$\begin{aligned} Q_N^{\Xi_S}(\mu) \left(u(\mu) - u_N^{\Xi_S}(\mu) \right) &= u(\mu) I^{\Xi_S}(Q_N^{\Xi_S})(\mu) - I^{\Xi_S}(u Q_N^{\Xi_S})(\mu) \\ &= \sum_{\mu' \in \Xi_S} \sum_{\lambda \in \Lambda} \frac{Q_N^{\Xi_S}(\mu') \omega^{\Xi_S}(\mu)}{(\mu - \mu') \omega^{\Xi_S}(\mu')} \left(\frac{v_\lambda}{\lambda - \mu} - \frac{v_\lambda}{\lambda - \mu'} \right) \\ &= -\omega^{\Xi_S}(\mu) \sum_{\lambda \in \Lambda} \frac{v_\lambda}{\lambda - \mu} \sum_{\mu' \in \Xi_S} \frac{Q_N^{\Xi_S}(\mu')}{(\lambda - \mu') \omega^{\Xi_S}(\mu')} \\ &= -\omega^{\Xi_S}(\mu) \sum_{\lambda \in \Lambda} \frac{v_\lambda}{(\lambda - \mu) \omega^{\Xi_S}(\lambda)} Q_N^{\Xi_S}(\lambda). \end{aligned} \quad (31)$$

After dividing by $Q_N^{\Xi_S}(\mu)$, taking the norm, and applying Lemma 2, we obtain

$$\begin{aligned} \left\| u(\mu) - u_N^{\Xi_S}(\mu) \right\|_V &= \left| \frac{\omega^{\Xi_S}(\mu)}{Q_N^{\Xi_S}(\mu)} \right| \left(\sum_{\lambda \in \Lambda} \frac{\|v_\lambda\|_V^2}{|\lambda - \mu|^2 |\omega^{\Xi_S}(\lambda)|^2} \left| Q_N^{\Xi_S}(\lambda) \right|^2 \right)^{1/2} \\ &\leq \left| \frac{\omega^{\Xi_S}(\mu)}{Q_N^{\Xi_S}(\mu)} \right| \frac{1}{\min_\Lambda |\cdot - \mu|} j^{\Xi_S}(Q_N^{\Xi_S}) \\ &\leq \left| \frac{\omega^{\Xi_S}(\mu)}{Q_N^{\Xi_S}(\mu)} \right| \frac{1}{\min_\Lambda |\cdot - \mu|} (1 + \varepsilon)^S \frac{C'}{\Phi_K(\lambda_{N+1})^S}, \end{aligned} \quad (33)$$

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

Now, due to (14), the term $|\omega^{\Xi_S}|$ can be 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, a combination of Theorem 2 and of the triangular inequality ensures that, for S large enough, each root of $Q_N^{\Xi_S}$ has distance from E bounded away from 0. This yields the claim. \square

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

$$\max_{\mu \in E} \left\| u(\mu) - u_N^{\Xi_S}(\mu) \right\|_V \leq C_E \left(\frac{(1 + \varepsilon)^2 \text{Cap}(K)}{\Phi_K(\lambda_{N+1})} \right)^S. \quad (34)$$

Thus, provided N is chosen large enough, 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 Section 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, and is summarized in the following.

Theorem 5 *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 . As in Theorem 3, let $\Xi^{(k)}$ be the shorthand for Ξ_{S_k} . Let $E \subset \mathbb{C}$ be compact, and let Assumption 1 be satisfied for some fixed $\mu_0 \in \mathbb{C}$. Then, for all $\varepsilon > 0$, it holds*

$$\lim_{k \rightarrow \infty} \text{Cap} \left(\left\{ \mu \in E, \left\| u(\mu) - u_{N_k}^{\Xi^{(k)}}(\mu) \right\|_V \geq \varepsilon \right\} \right) = 0, \quad (35)$$

with Cap denoting the logarithmic capacity, see Section 3.1.

Proof. We rely on some concepts introduced in the proofs of Theorems 3 and 4, namely the pole ordering (27) and the intermediate bound (33). In particular, by exploiting (21) with $\Lambda_0 = \{\sigma_j\}_{j=1}^{N_k}$, (33) yields

$$\left\| u(\mu) - u_{N_k}^{\Xi^{(k)}}(\mu) \right\|_V \leq \frac{|\omega^{\Xi^{(k)}}(\mu)|}{\left| Q_{N_k}^{\Xi^{(k)}}(\mu) \right| \min_{\Lambda} |\cdot - \mu|} \frac{\left(\sum_{\lambda \in \Lambda} \|v_\lambda\|_V^2 \right)^{1/2}}{\left\| \prod_{i=1}^{N_k} (\cdot - \sigma_i) \right\|_{N_k}} \sup_{l > N_k} \frac{\prod_{i=1}^{N_k} |\sigma_l - \sigma_i|}{|\omega^{\Xi^{(k)}}(\sigma_l)|}.$$

By following the same steps as in the proof of Theorem 3, we can simplify the $\|\cdot\|_{N_k}$ -norm and the supremum to obtain

$$\left\| u(\mu) - u_{N_k}^{\Xi^{(k)}}(\mu) \right\|_V \leq \frac{2^{N_k} \left(\sum_{\lambda \in \Lambda} \|v_\lambda\|_V^2 \right)^{1/2}}{c_{\mu_0} \left| Q_{N_k}^{\Xi^{(k)}}(\mu) \right| \min_{\Lambda} |\cdot - \mu|} \left(\frac{R + R_{\mu_0, K}}{|\mu_0 - \sigma_{N_k+1}| - R_{\mu_0, K}} \right)^{S_k - N_k} \times \\ \times \left(\frac{|\mu_0 - \sigma_{N_k+1}|}{|\mu_0 - \sigma_{N_k+1}| - R_{\mu_0, K}} \right)^{N_k} \prod_{i=1}^{N_k} \frac{R + R_{\mu_0, K}}{|\mu_0 - \sigma_i|}, \quad (36)$$

with $R = \max_E |\cdot - \mu_0|$. In order to obtain the desired result, we need to manage carefully the two troublesome μ -dependent terms in the denominator.

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

$$\mathcal{E}_{k, \delta} = \left\{ \mu \in \mathbb{C}, \left(\prod_{\lambda' \in \Lambda_{in}^{(k)}} |\mu - \lambda'| \right) \left(\prod_{\lambda \in \Lambda \cap E} |\mu - \lambda| \right) \leq \delta^{\#\Lambda_{in}^{(k)} + \bar{N}} \right\}, \quad (37)$$

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

We remark that the exponent of δ in (37) is equal to the degree of the monic polynomial (in μ) whose magnitude is the left hand side of the inequality. Thus [3, Theorem 6.6.3], the logarithmic capacity of $\mathcal{E}_{k, \delta}$ is equal² to δ .

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;

²We are assuming without loss of generality that $\#\Lambda_{in}^{(k)} + \bar{N} > 0$. If this is not the case, $\mathcal{E}_{k, \delta}$ is empty for $\delta < 1$, and the claim holds quite trivially.

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

Then the claim follows, since, by inclusion,

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

However, before we can proceed with either step, it is necessary to introduce some additional bounds in (36). First, since $Q_{N_k}^{\Xi(k)}$ is normalized, by Assumption 1 it must hold

$$\begin{aligned} \left|Q_{N_k}^{\Xi(k)}(\mu)\right| &= \frac{\prod_{i=1}^{N_k} |\mu - \lambda_i^{(k)}|}{\left\|\prod_{i=1}^{N_k} (\cdot - \lambda_i^{(k)})\right\|_{N_k}} \geq \frac{1}{C_{\mu_0}} \prod_{i=1}^{N_k} \frac{|\mu - \lambda_i^{(k)}|}{R_{\mu_0} + |\mu_0 - \lambda_i^{(k)}|} \\ &= \frac{1}{C_{\mu_0}} \left(\prod_{\lambda' \in \Lambda_{in}^{(k)}} \frac{|\mu - \lambda'|}{R_{\mu_0} + |\mu_0 - \lambda'|} \right) \left(\prod_{\lambda'' \in \Lambda_{out}^{(k)}} \frac{|\mu - \lambda''|}{R_{\mu_0} + |\mu_0 - \lambda''|} \right). \end{aligned}$$

In the first group of factors, we can upper-bound $|\mu_0 - \lambda'|$ by $2R$. In the second group, since $|\mu - \mu_0| \leq R$ for all $\mu \in E$ by definition, it holds

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

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 leads to

$$\begin{aligned} \left|Q_{N_k}^{\Xi(k)}(\mu)\right| \prod_{\lambda \in \Lambda \cap E} |\mu - \lambda| &\geq \frac{R^{N_k - \#\Lambda_{in}^{(k)}}}{C_{\mu_0} (R_{\mu_0} + 2R)^{N_k}} \left(\prod_{\lambda' \in \Lambda_{in}^{(k)}} |\mu - \lambda'| \right) \left(\prod_{\lambda \in \Lambda \cap E} |\mu - \lambda| \right) \\ &> \frac{R^{N_k - \#\Lambda_{in}^{(k)}} \delta^{\#\Lambda_{in}^{(k)} + \bar{N}}}{C_{\mu_0} (R_{\mu_0} + 2R)^{N_k}} \geq \frac{\delta^{N_k + \bar{N}}}{C_{\mu_0} (R_{\mu_0} + 2R)^{N_k}}, \end{aligned}$$

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

Accordingly, we can simplify (36) for all $\mu \in E \setminus \mathcal{E}_{k,\delta}$:

$$\begin{aligned} \left\|u(\mu) - u_{N_k}^{\Xi(k)}(\mu)\right\|_V &< \frac{C_{\mu_0}}{c_{\mu_0}} \left(\sum_{\lambda \in \Lambda} \|v_\lambda\|_V^2 \right)^{1/2} \frac{\prod_{\lambda \in \Lambda \cap E} |\mu - \lambda| 2^{N_k} (R_{\mu_0} + 2R)^{N_k}}{\min_\Lambda |\cdot - \mu| \delta^{N_k + \bar{N}}} \times \\ &\times \left(\frac{R + R_{\mu_0,K}}{|\mu_0 - \sigma_{N_k+1}| - R_{\mu_0,K}} \right)^{S_k - N_k} \left(\frac{|\mu_0 - \sigma_{N_k+1}|}{|\mu_0 - \sigma_{N_k+1}| - R_{\mu_0,K}} \right)^{N_k} \prod_{i=1}^{N_k} \frac{R + R_{\mu_0,K}}{|\mu_0 - \sigma_i|}. \quad (38) \end{aligned}$$

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

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

can be upper-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

$$\begin{aligned} \delta = \delta_k &= \left(\frac{C' 2^{N_k} (R_{\mu_0} + 2R)^{N_k}}{\varepsilon} \left(\frac{R + R_{\mu_0,K}}{|\mu_0 - \sigma_{N_k+1}| - R_{\mu_0,K}} \right)^{S_k - N_k} \times \right. \\ &\quad \left. \times \left(\frac{|\mu_0 - \sigma_{N_k+1}|}{|\mu_0 - \sigma_{N_k+1}| - R_{\mu_0,K}} \right)^{N_k} \prod_{i=1}^{N_k} \frac{R + R_{\mu_0,K}}{|\mu_0 - \sigma_i|} \right)^{1/(N_k + \bar{N})}. \quad (39) \end{aligned}$$

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 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. In particular, thanks to the Stolz-Cesàro theorem, it can be proven that

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

while all other factors in (39) remain bounded as k increases. \square

Due to (10), Theorem 5 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} .

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 2–5 hold only (with some extensions) if the solution map is of the form (8), 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 (6) be given, along with some fixed tolerance $\varepsilon > 0$. The task is to obtain an approximate solution map \tilde{u} such that

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

In our framework, the presence of singularities in the true and, possibly, the approximate solution map within K may make the conditions above meaningless. In such cases, it may be preferable [11] to consider a residual-based accuracy requirement

$$\max_{\mu \in K} \|\mathcal{F}_\mu(\tilde{u}(\mu))\|_W \leq \varepsilon \quad \text{or} \quad \max_{\mu \in K} \frac{\|\mathcal{F}_\mu(\tilde{u}(\mu))\|_W}{\|\mathcal{F}_\mu(0)\|_W} \leq \varepsilon, \quad (40)$$

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

Now, let us take as approximate map \tilde{u} the minimal rational interpolant $u_{N_k}^{Z_S}$ 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 (40).

In particular, let us consider the problem of computing *a posteriori* the norm of the residual $\|\mathcal{F}_\mu(\tilde{u}(\mu))\|_W$ at some given point $\mu \in K$. In order to proceed, we will assume that the operator \mathcal{F}_μ 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

$$\mathcal{F}_\mu(v) = F(\mu)v - f(\mu) = \sum_{i=1}^{n_F} \theta_i^F(\mu) F_i v - \sum_{i=1}^{n_f} \theta_i^f(\mu) f_i. \quad (41)$$

This is the ideal situation to employ RB methods [10, 27, 30, 35, 36]. Many strategies have been devised to approximate general, possibly non-linear and non-separable, parametric problems into the form (41), see e.g. the (Discrete) Empirical Interpolation Method [9, 15, 21] and hyper-reduction techniques [8].

Now, $F(\mu)$ can be applied to both sides of (31) to obtain

$$\begin{aligned} Q_N^{Z_S}(\mu) \left(f(\mu) - F(\mu)u_N^{Z_S}(\mu) \right) &= f(\mu)I^{Z_S}(Q_N^{Z_S})(\mu) - F(\mu)I^{Z_S}(uQ_N^{Z_S})(\mu) \\ &= \sum_{\mu' \in Z_S} \frac{Q_N^{Z_S}(\mu') \omega^{Z_S}(\mu)}{(\mu - \mu') \omega^{Z_S'}(\mu')} (f(\mu) - F(\mu)u(\mu')) \\ &= \sum_{i=1}^{n_f} \Delta^{Z_S} \left(\theta_i^f Q_N^{Z_S} \right) (\mu) f_i \\ &\quad - \sum_{i=1}^{n_F} F_i \Delta^{Z_S} \left(\theta_i^F I^{Z_S} \left(uQ_N^{Z_S} \right) \right) (\mu). \end{aligned} \quad (42)$$

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, which consists of generalized moments of $Q_N^{Z_S}$ and $I^{Z_S}(uQ_N^{Z_S})$. Hence, its norm (squared) admits an affine decomposition with $(n_F + n_f)^2$ terms. In particular, if we assume an offline-online MOR framework [27, 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. If the samples are managed carefully, see e.g. [35, Section 3.7], the evaluation of the residual at a (new) parameter μ can be carried out online through low-dimensional operations only. The resulting complexity of such procedure is $\mathcal{O}((n_F S + n_f)^2)$, independently of the size of the original problem (41).

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* non-intrusive error indicator can be quite tricky. Here we propose a simplified approach, which is rigorous only for very specific problem structures, namely

$$\mathcal{F}_\mu(u) = F(\mu)u - f = (F_0 + \mu F_1)u - f. \quad (43)$$

Let the problem be of the form (43). Thanks to (23) and (42), it must hold

$$\begin{aligned} Q_N^{Z_S}(\mu) \left(f - F(\mu)u_N^{Z_S}(\mu) \right) &= -\omega^{Z_S}(\mu) F_1 \sum_{\mu' \in Z_S} \frac{Q_N^{Z_S}(\mu') u(\mu')}{\omega^{Z_S'}(\mu')} \\ &= -\frac{\omega^{Z_S}(\mu)}{(S-1)!} F_1 \frac{d^{S-1}}{d\mu^{S-1}} I^{Z_S}(uQ_N^{Z_S}). \end{aligned}$$

Now it suffices to take the $\|\cdot\|_W$ -norm and exploit (5) to obtain

$$\begin{aligned} \left\| \mathcal{F}_\mu \left(u_N^{Z_S}(\mu) \right) \right\|_W &\leq \left| \frac{\omega^{Z_S}(\mu)}{Q_N^{Z_S}(\mu)} \right| \|F_1\|_{\mathcal{L}(V;W)} \left\| \frac{1}{(S-1)!} \frac{d^{S-1}}{d\mu^{S-1}} I^{Z_S}(uQ_N^{Z_S}) \right\|_V \\ &= \left| \frac{\omega^{Z_S}(\mu)}{Q_N^{Z_S}(\mu)} \right| \|F_1\|_{\mathcal{L}(V;W)} j^{Z_S} \left(Q_N^{Z_S} \right). \end{aligned} \quad (44)$$

One obvious limitation of (44) is the need for some approximation of the $V \rightarrow W$ operator norm of F_1 , i.e.

$$\|F_1\|_{\mathcal{L}(V;W)} = \sup_{v \in V \setminus \{0\}} \frac{\|F_1 v\|_W}{\|v\|_V}.$$

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 [26, 29, 35, 36]. If one can overcome such limitation, see e.g. Section 5.2, bound (44) is certainly quite appealing from a computational standpoint, since its evaluation consists only of few scalar computations.

If (43) 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 (44) can be applied heuristically, with an accuracy which may depend quite sharply on the smoothness of \mathcal{F}_μ , 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 Numerical examples

In this section we present two numerical experiments to verify some of the theoretical results presented above, as well as the effectiveness of the *a posteriori* error estimators described in the previous section.

5.1 Normal eigenvalue problem

Let $n = 100$ and take $A \in \mathbb{C}^{n \times n}$ a normal matrix whose eigenvalues are randomly generated according to a uniform distribution over $\{x + iy, (x, y) \in [-5, 5]^2\}$, and whose eigenvectors are set 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 $K = \overline{\mathcal{B}(0, 1)}$. To this aim, we fix $\mathbf{v} \in \mathbb{C}^n$ a random gaussian vector, and we approximate by minimal rational interpolants the solution map of the parametric problem

$$\text{for } \mu \in K, \text{ find } \mathbf{u}(\mu) \in \mathbb{C}^n \text{ s.t. } (A - \mu I) \mathbf{u}(\mu) = \mathbf{v}, \quad (45)$$

with $I \in \mathbb{C}^{n \times n}$ the identity matrix. As sample points Ξ_S , we choose the Fejér points of K , i.e. the roots of unity, and we choose to employ the polynomial norm $\|\cdot\|_{0,N}$, see (18). 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.

The poles of the solution map are represented in Figure 1, where we can observe that $\nu = 5$ poles are inside K . As expected, we can observe in Figure 2 that the (euclidean) norm of the solution map $\|\mathbf{u}(\mu)\|_2$ becomes unbounded if μ gets too close to a pole.

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:

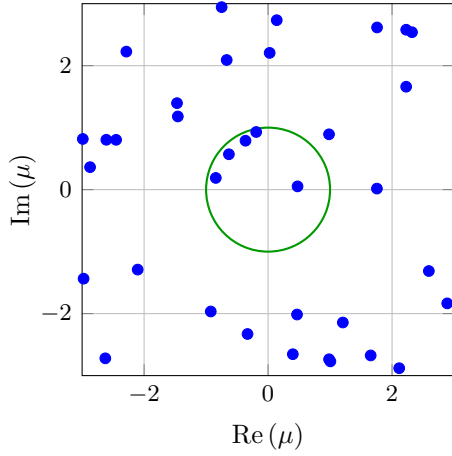


Figure 1: Eigenvalues of A near 0. In green, the boundary of the parameter domain.

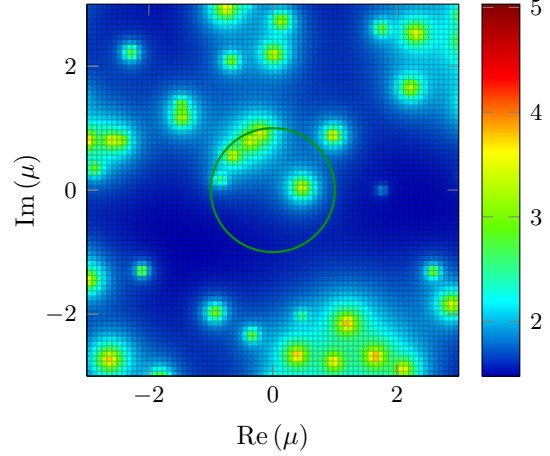


Figure 2: Logarithm of the euclidean norm of the solution map \mathbf{u} near 0.

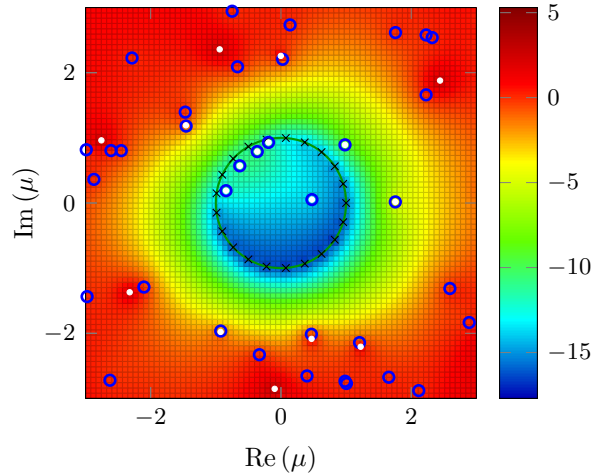
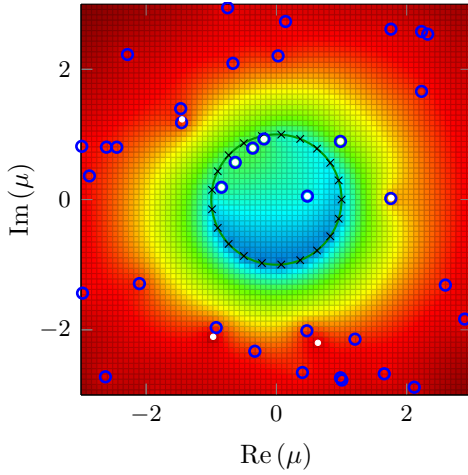


Figure 3: Logarithm of the relative error norm $\|\mathbf{u} - \mathbf{u}_N^{\Xi_S}\|_2 / \|\mathbf{u}\|_2$ achieved by minimal rational interpolants based on samples at $\Xi_S = \{e^{2j\pi i/21}\}_{j=1}^{21}$ (black crosses) with $N = 10$ (left) and $N = 20$ (right). The exact and approximated poles are represented in blue and white respectively.

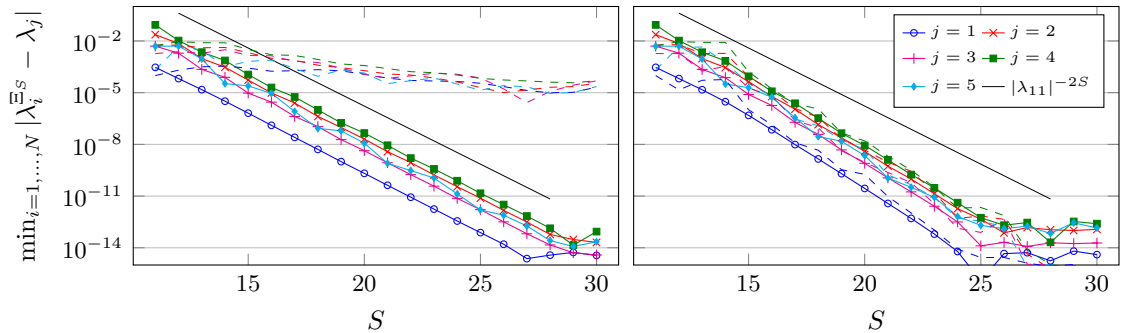


Figure 4: Pole approximation error for the poles inside K . On the left $N = 10$ is kept fixed, whereas $N = S - 1$ on the right. In black the theoretical rate (25). The dashed lines show the results obtained by projection on subspaces of dimension N computed through POD. The scale is the same for both graphs.

- we fix N (in our case, we set $N = 10$) and increase progressively the number of samples S ; to verify if the starting guess on N was large enough, we can simply check if at least one of the approximate poles has converged (for large S) to some point outside K , see Theorem 2.
- 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.

We compare visually the two approaches in Figure 3, where we show the error norm obtained by approximants of type $[20/10]$ and $[20/20]$, both based on $S = 21$ samples at Fejér points. We can see that the error increases much quicker near 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 4. A (possibly) more interesting observation is that the error appears to be globally smaller when we choose $N = 20$ (in fact, a closer look at the data shows that this is not the case only for 108 points in the uniform 75×75 grid used to obtain the plot). Overall, we can conclude that the choice $N = 20$ 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 (25), for the 5 poles within K . The results are shown in Figure 4. For fixed $N = 10$, the rate predicted by Theorem 2 can be observed (Green’s potential for the unit 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 4 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 problem 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.

5.2 Time-harmonic vibrations of a tuning fork

Let $\Omega \subset \mathbb{R}^3$ be the region occupied by a tuning fork, see Figure 5. We consider the following internal acoustic problem in the frequency domain. Assume that the tuning fork is clamped on a portion Γ_D of the handle. Let $\tilde{\mathbf{g}}(t, \mathbf{x}) = e^{i2\pi\nu t} \mathbf{g}(\mathbf{x})$ 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. 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 time-harmonic linear elasticity problem [20]

$$\begin{cases} -\nabla \cdot \sigma(\mathbf{u}(\nu)) - 2\pi\nu(2\pi\nu - \eta)\rho\mathbf{u}(\nu) = \mathbf{0} & \text{in } \Omega \\ \mathbf{u}(\nu) = \mathbf{0} & \text{on } \Gamma_D \\ \sigma(\mathbf{u}(\nu))\mathbf{n} = \mathbf{g} & \text{on } \Gamma_{\mathbf{g}} \\ \sigma(\mathbf{u}(\nu))\mathbf{n} = \mathbf{0} & \text{on } \Gamma_{free} = \partial\Omega \setminus \Gamma_D \setminus \Gamma_{\mathbf{g}} \end{cases}, \quad (46)$$

with η a damping parameter.

In particular, we choose to approximate problem (46) using \mathbb{P}_1 Finite Elements (FE) on a sufficiently fine tetrahedral discretization of Ω . The FE discretization of (46) defines a parametric problem of the form (6), whose solution map \mathbf{u} can be proven meromorphic using compactness arguments [7, 37]. Thus, we decide to approximate it for $\nu \in K = [1, 4] \cdot 10^4$ Hz with minimal

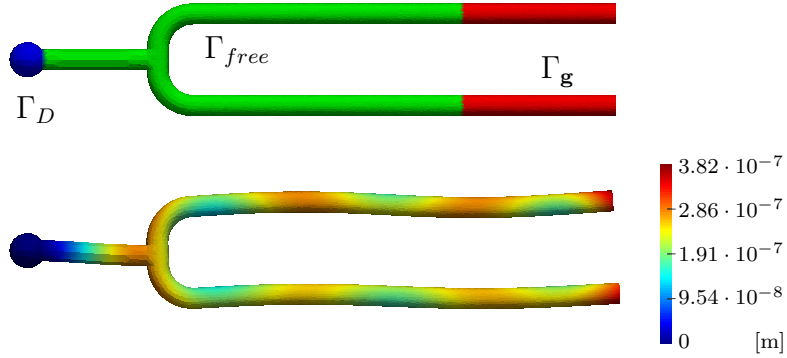


Figure 5: Reference domain (above) and domain warped by the real part of the displacement, magnified by a factor of $3 \cdot 10^3$ (below). The displacement is the solution of (46) for frequency $\nu = 10^4$ Hz and damping coefficient $\eta = 0$ Hz. The total number of degrees of freedom of the discretized problem is 36330.

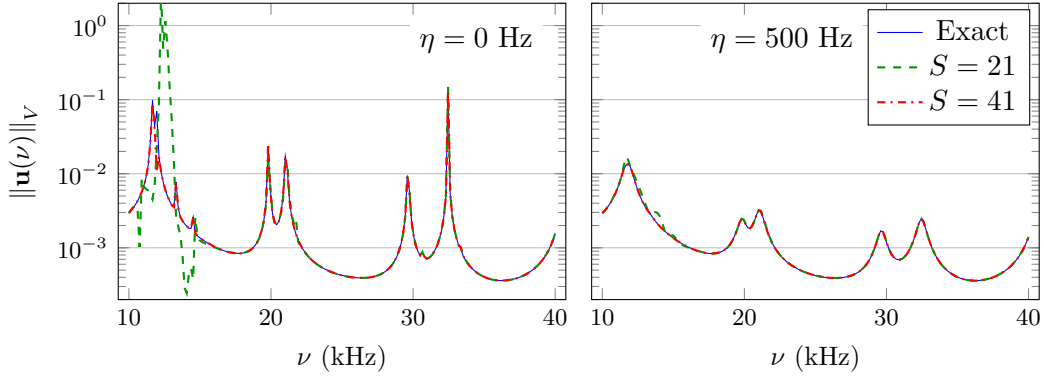


Figure 6: Energy norm of the solution map \mathbf{u} of (46) in blue, for $\eta \in \{0, 500\}$ Hz. In green and red the norm of the [20/20] and [40/40] minimal rational interpolant of \mathbf{u} obtained from 21 and 41 samples at Chebyshev nodes respectively.

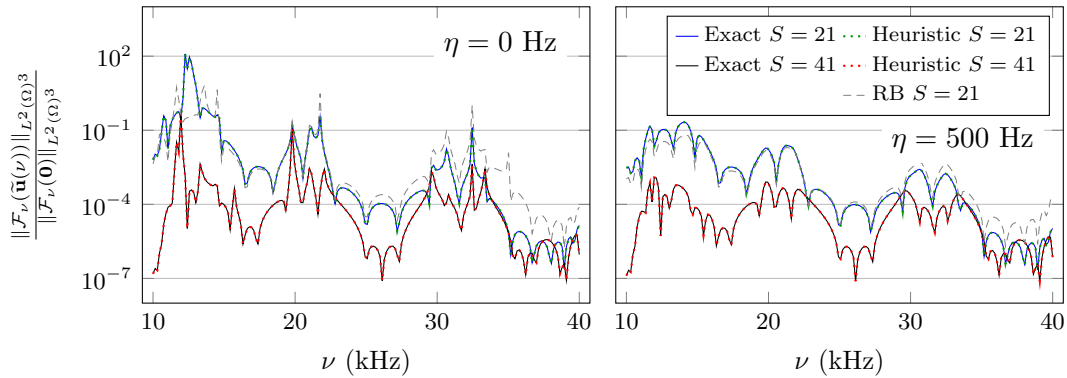


Figure 7: Relative $L^2(\Omega)^3$ -norm of the residual for problem (46) and $\tilde{\mathbf{u}} = \mathbf{u}_{20}^{\Xi_{21}}$ (blue) and $\tilde{\mathbf{u}} = \mathbf{u}_{40}^{\Xi_{41}}$ (black), for $\eta \in \{0, 500\}$ Hz. In green and red the approximations of the residual obtained through the heuristic approach (44). The dashed grey line is the residual obtained when $\tilde{\mathbf{u}}$ is the RB approximant of \mathbf{u} with samples at Ξ_{21} .

rational interpolants relying on samples at the Chebyshev nodes of K , which are the Fejér points for K . The chosen polynomial norm $\|\cdot\|_N$ is the one induced (see Appendix A) by the family of Chebyshev polynomials over K .

The (elastic energy) norm of the solution map is shown in Figure 6 for $\eta \in \{0, 500\}$ Hz, along with the approximated norm profiles, obtained by minimal rational interpolants of type [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 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 need for more than 21 samples, one could rely on residual-based estimators such as (40). To this aim, we can apply (42) to obtain an exact (since the problem depends affinely on the parameter) expression for the residual in the $L^2(\Omega)^3$ -norm. The evaluation of (42) on a fine grid is shown in Figure 7.

In the same Figure, we also show the values obtained through the simplified residual estimator (44). 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 (46) at some point μ' , see Section 4. We manage to avoid this (thus keeping the estimator as non-intrusive as possible) by the following heuristic argument. We assume that the right hand side of (44) is able to capture well the behavior of the residual with respect to the parameter, so that we just need to identify a constant C for which (employing the notation of Section 4)

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

This can be easily achieved by evaluating (offline and at a relatively low computational cost) the exact value of the residual at some point³ $\mu' \in K \setminus Z_S$, 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 set $\mu' = 25.5 \cdot 10^3$ Hz, 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.

By looking at the curve for $S = 21$ in the left plot of Figure 7, we can deduce that the approximation is particularly poor in the region of lowest frequency for $\eta = 0$ Hz. This could lead us to conclude that more sample points are needed, as is indeed the case. The same reasoning can be applied, more mildly, in the damped case.

Finally, in Figure 7 we also perform a comparison between minimal rational interpolants and RBs, by showing the norm of the residual obtained with the RB approximant of (46) 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.

6 Conclusions

In this paper we have presented minimal rational interpolants as the natural extension of fast LS-Padé approximants [6], 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:

³For stability reasons it may be advisable to take the average of more than one residual sample, provided the additional computation cost is acceptable.

- the stability is greatly improved, since computing (many) derivatives of the solution map accumulates numerical errors; the Arnoldi-like reorthogonalization technique described in [6] 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 examples have shown minimal rational interpolants to perform well when applied to normal eigenvalue problems and MOR of time-harmonic problems. In particular, in our second example we have shown the effectiveness of a heuristic *a posteriori* residual estimator, which is very inexpensive to compute. Accordingly, we deem 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 projections.

While the proven theoretical properties seem certainly appealing, we find quite interesting that the numerical results for denominator degree as large as possible ($N = S - 1$) manage to beat the exponential rate proven and achieved for fixed N . Still in connection with the greedy idea described just above, we believe that it could be interesting to determine whether this “superconvergence” is preserved if the samples are not computed at the Fejér points.

Concerning extensions to the multivariate case, it is our opinion that explicit rational approximants may not be well suited for high-dimensional problems, mostly because of ill-conditioning issues, almost entirely due to the curse of dimensionality. We believe that hybrid techniques, combining explicit (interpolatory) and implicit (projection-based) methods, as is done in parametric MOR approaches [4, 34], may yield better results in such cases.

Acknowledgments

The author gratefully wishes to thank Fabio Nobile for the many fruitful discussions and for his helpful advice at several stages of this work.

A Polynomial norm bounds

Let $\Gamma \subset \mathbb{C}$ be a Jordan curve, and consider $\{\varphi_j\}_{j=0,1,\dots}$ a hierarchical ($\text{degree}(\varphi_j) = j$ for all j) family of polynomials, orthonormal over $L^2(\Gamma)$:

$$\frac{\int_{\Gamma} \varphi_j \overline{\varphi_i} |dl|}{\int_{\Gamma} |dl|} = \delta_{ji} \quad \forall (i, j) \in \{0, 1, \dots\}^2. \quad (47)$$

This, for all $N \in \mathbb{N}$, induces a norm over $\mathbb{P}_N(\mathbb{C})$ in the following sense:

$$\| \| Q \| \|_N = \sqrt{\sum_{j=0}^N |q_j|^2}, \quad \text{where} \quad Q = \sum_{j=0}^N q_j \varphi_j. \quad (48)$$

The following holds.

Lemma 4 *The norm $\| \cdot \|_N$ in (48) satisfies Assumption 1 for any μ_0 chosen within the interior of Γ .*

Proof. For a given $N \in \mathbb{N}$, let $0 \leq D \leq N$, $\{z_j\}_{j=1}^D \subset \mathbb{C}^D$, and $Q = \prod_{j=1}^D (\cdot - z_j)$. Due to orthonormality, it holds

$$\| \| Q \| \|_N^2 \int_{\Gamma} |dl| = \int_{\Gamma} |Q|^2 |dl| = \int_{\Gamma} \prod_{j=1}^D |\cdot - z_j|^2 |dl|. \quad (49)$$

On one side, let $R_{\mu_0} = \max_{\Gamma} |\cdot - \mu_0|$. We can use the triangular inequality to obtain

$$\begin{aligned} \|Q\|_N^2 \int_{\Gamma} |dl| &\leq \int_{\Gamma} \prod_{j=1}^D (|\cdot - \mu_0| + |\mu_0 - z_j|)^2 |dl| \\ &\leq \prod_{j=1}^D (R_{\mu_0} + |\mu_0 - z_j|)^2 \int_{\Gamma} |dl|. \end{aligned} \quad (50)$$

On the other side, let $c_{\mu_0} = (2\pi \min_{\Gamma} |\cdot - \mu_0| / \int_{\Gamma} |dl|)^2$. We can exploit the Cauchy-Schwarz inequality to show that

$$\left(\int_{\Gamma} \prod_{j=1}^D |\cdot - z_j|^2 |dl| \right) \left(\int_{\Gamma} \frac{|dl|}{|\cdot - \mu_0|^2} \right) \geq \left| \int_{\Gamma} \frac{\prod_{j=1}^D (\cdot - z_j)}{\cdot - \mu_0} dl \right|^2,$$

so that, due to the Cauchy integral formula,

$$\begin{aligned} \|Q\|_N^2 \int_{\Gamma} |dl| &\geq \left| \int_{\Gamma} \frac{\prod_{j=1}^D (\cdot - z_j)}{\cdot - \mu_0} dl \right|^2 \left(\int_{\Gamma} \frac{|dl|}{|\cdot - \mu_0|^2} \right)^{-1} \\ &\geq 4\pi^2 \prod_{j=1}^D |\mu_0 - z_j|^2 \left(\min_{\Gamma} |\cdot - \mu_0| \right)^2 \left(\int_{\Gamma} |dl| \right)^{-1} \\ &\geq c_{\mu_0} \prod_{j=1}^D |\mu_0 - z_j|^2 \int_{\Gamma} |dl|. \end{aligned} \quad (51)$$

Dividing (50) and (51) by $\int_{\Gamma} |dl|$ yields the proof. \square

References

- [1] A.C. Antoulas. Approximation of Large-Scale Dynamical Systems: An Overview. *IFAC Proceedings Volumes*, 37(11):19–28, 2017.
- [2] J.M. Ash, A. Berele, and S. Catoiu. Plausible and genuine extensions of L’Hospital’s rule. *Mathematics Magazine*, 85(1):52–60, 2012.
- [3] G. A. Baker and P. R. Graves-Morris. *Padé approximants*. Cambridge University Press, 1996.
- [4] U. Baur, C. Beattie, P. Benner, and S. Gugercin. Interpolatory Projection Methods for Parameterized Model Reduction. *SIAM Journal on Scientific Computing*, 33(5):2489–2518, 2011.
- [5] F. Bonizzoni, F. Nobile, and I. Perugia. Convergence analysis of Padé approximations for Helmholtz frequency response problems. *ESAIM: M2AN.*, 2017. DOI: 10.1051/m2an/2017050.
- [6] F. Bonizzoni, F. Nobile, I. Perugia, and D. Pradovera. Fast Least-Squares Padé approximation of self-adjoint problems with meromorphic structure. *ArXiv e-prints*, 2018. DOI: arXiv/1808.03112.
- [7] F. Bonizzoni, F. Nobile, I. Perugia, and D. Pradovera. Least-Squares Padé approximation of parametric and stochastic Helmholtz maps. *ArXiv e-prints*, 2018. DOI: arXiv/1805.05031.
- [8] 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. *Journal of Computational Physics*, 242:623–647, 2013.
- [9] S. Chaturantabut and D.C. Sorensen. Nonlinear Model Reduction via Discrete Empirical Interpolation. *SIAM Journal on Scientific Computing*, 32(5):2737–2764, 2010.
- [10] 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(2):970–996, 2010.
- [11] M. Chevreuil and A. Nouy. Model order reduction based on proper generalized decomposition for the propagation of uncertainties in structural dynamics. *International Journal for Numerical Methods in Engineering*, 89(2):241–268, 2012.

- [12] A.A.M. Cuyt and X. Yang. A practical error formula for multivariate rational interpolation and approximation. *Numerical Algorithms*, 55(2):233–243, 2010.
- [13] L. Daniel, O.C. Siong, L.S. Chay, K.H. Lee, and J. White. A multiparameter moment-matching model-reduction approach for generating geometrically parameterized interconnect performance models. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 23(5):678–693, 2004.
- [14] Z. Drmač, S. Gugercin, and C. Beattie. Quadrature-Based Vector Fitting for Discretized \mathcal{H}_2 Approximation. *SIAM Journal on Scientific Computing*, 37(2):625–652, 2015.
- [15] M. Drohmann, B. Haasdonk, and M. Ohlberger. Reduced Basis Approximation for Nonlinear Parametrized Evolution Equations based on Empirical Operator Interpolation. *SIAM Journal on Scientific Computing*, 34(2):937–969, 2012.
- [16] O. Farle, V. Hill, P. Ingelstrom, and R. Dyczij-Edlinger. Multi-parameter polynomial order reduction of linear finite element models. *Mathematical and Computer Modelling of Dynamical Systems*, 14(5):421–434, 2008.
- [17] B. Fischer and L. Reichel. Newton Interpolation in Fejer and Chebyshev Points. *Math. Comp.*, 53(187):265–279, 1989.
- [18] Roland W. Freund. Model reduction methods based on Krylov subspaces. *Acta Numerica*, 12(2003):267–319, 2003.
- [19] G.H. Golub and C.F.V. Loan. *Matrix Computations*, 1996.
- [20] D. Gottlieb and J.S. Hesthaven. Spectral Methods for Hyperbolic Problems. *Handbook of Numerical Analysis*, 17:441–466, 2016.
- [21] M.A. Grepl, Y. Maday, N.C. Nguyen, and A.T. Patera. Efficient reduced-basis treatment of non-affine and nonlinear partial differential equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41(3):575–605, 2007.
- [22] E. Grimme. *Krylov projection methods for model reduction*. Electrical engineering, University of Illinois at Urbana-Champaign, 1997.
- [23] S. Grivet-Talocia and B. Gustavsen. *Passive Macromodeling: Theory and Applications*. Wiley series in microwave and optical engineering. John Wiley & Sons, Inc, Hoboken New Jersey, 2015.
- [24] P. Guillaume, A. Huard, and V. Robin. Generalized multivariate Padé approximants. *Journal of Approximation Theory*, 95(2):203 – 214, 1998.
- [25] B. Gustavsen and A. Semlyen. Rational approximation of frequency domain responses by vector fitting. *IEEE Transactions on Power Delivery*, 14(3):1052–1061, jul 1999.
- [26] S. Hain, M. Ohlberger, M. Radic, and K. Urban. A Hierarchical A-Posteriori Error Estimator for the Reduced Basis Method. *ArXiv e-prints*, 2018.
- [27] J.S. Hesthaven, G. Rozza, and B. Stamm. *Certified reduced basis methods for parametrized partial differential equations*. SpringerBriefs in Mathematics. Springer Cham; BCAM Basque Center for Applied Mathematics, Bilbao, 2016.
- [28] E Hille. *Analytic function theory*. Chelsea Publishing Company, New York, N.Y, second ed. edition, 1965.
- [29] D.B.P. Huynh, G. Rozza, S. Sen, and A.T. Patera. A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. *Comptes Rendus Mathematique*, 345(8):473–478, oct 2007.
- [30] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza. Generalized reduced basis methods and n -width estimates for the approximation of the solution manifold of parametric PDEs. In *Analysis and numerics of partial differential equations*, volume 4 of *Springer INdAM Ser.*, pages 307–329. Springer, Milan, 2013.
- [31] C. Lei, Y. Wang, Q. Chen, and N. Wong. On Vector Fitting Methods in Signal / Power Integrity Applications. II, 2010.
- [32] P. Monk. *Finite Element Methods for Maxwell’s Equations*. Clarendon, Oxford, 2003.

- [33] J. Nuttall. The convergence of Padé approximants of meromorphic functions. *Journal of Mathematical Analysis and Applications*, 31(1):147–153, jul 1970.
- [34] H. Panzer, J. Mohring, R. Eid, and B. Lohmann. Parametric model order reduction by matrix interpolation. *At-Automatisierungstechnik*, 58(8):475–484, 2010.
- [35] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced Basis Methods for Partial Differential Equations: An Introduction*. UNITEXT. Springer International Publishing, 2015.
- [36] G. Rozza, D. B.P. 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. *Archives of Computational Methods in Engineering*, 15(3):229–275, 2008.
- [37] S. Steinberg. Meromorphic families of compact operators. *Archive for Rational Mechanics and Analysis*, 31:372–379, dec 1968.
- [38] J L Walsh. *Interpolation and Approximation by Rational Functions in the Complex Domain*. American Mathematical Society, 1969.
- [39] D.S. Weile, E. Michielssen, E. Grimme, and K. Gallivan. A method for generating rational interpolant reduced order models of two-parameter linear systems. *Applied Mathematics Letters*, 12(5):93–102, 1999.