



HAL
open science

Rational kernel-based interpolation for complex-valued frequency response functions

Julien Bect, Niklas Georg, Ulrich Römer, Sebastian Schöps

► **To cite this version:**

Julien Bect, Niklas Georg, Ulrich Römer, Sebastian Schöps. Rational kernel-based interpolation for complex-valued frequency response functions. 2023. hal-04191766v1

HAL Id: hal-04191766

<https://centralesupelec.hal.science/hal-04191766v1>

Preprint submitted on 30 Aug 2023 (v1), last revised 4 Sep 2023 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

33 where s denotes the complex frequency variable, also known as the Laplace variable.
 34 Assuming a suitably normalized excitation $\hat{\mathbf{g}}(s)$, the frequency response function is
 35 defined as the value $\omega \mapsto \hat{f}(i\omega)$ of \hat{f} on the imaginary axis, where ω is called the
 36 angular frequency, and we are typically interested more specifically in its value on
 37 a certain interval $\Omega = [\omega_{\min}, \omega_{\max}] \subseteq [0, +\infty)$. In the following, we omit explicitly
 38 indicating frequency domain variables to simplify the notation.

39 The location of the poles of \hat{f} strongly depends on the properties of $\mathbf{K}, \mathbf{D}, \mathbf{M}$,
 40 see [38]. We assume, in particular, that no pole is placed on the frequency axis $i\mathbb{R}$
 41 and that the frequency response function is holomorphic on the shifted right half-
 42 plane $\Gamma_\alpha = \{s \in \mathbb{C} \mid \Re[s] > -\alpha\}$, $\alpha > 0$. The real parts of all poles are strictly
 43 negative for instance if $\mathbf{K}, \mathbf{D}, \mathbf{M}$ are symmetric positive definite, see Section 3 of [38].
 44 The same holds true if the homogeneous version of (1.1) is stable, in the sense that all
 45 solutions decay exponentially to zero as $t \rightarrow \infty$. The holomorphy of response functions
 46 has recently been studied also in the context of partial differential equations, see [7,
 47 Proposition 5.3] for instance. There, the frequency response map for an acoustic
 48 scattering problem was studied and appropriate damping terms ensured a locally
 49 holomorphic response function, with a negative real part for all poles¹.

50 Adopting a data-driven approach, (1.3) must be solved repeatedly on a set of
 51 interpolation/training points $\omega_i \in \Omega$, with $s_i = i\omega_i$. Numerical efficiency demands a
 52 small training set

$$53 \quad (1.4) \quad \{\omega_i, f(i\omega_i)\}_{i=1}^n, \quad \text{where } \omega_i \in \Omega, f(i\omega_i) \in \mathbb{C}, i = 1, \dots, n,$$

54 hence, there is a need for accurate interpolation in the frequency domain.

55 The data-driven approximation of frequency response functions has attracted con-
 56 siderable interest in the literature, see for instance [16, 23, 28] and the references
 57 therein. Among the numerous available approaches we mention vector fitting [16] and
 58 the adaptive Antoulas-Anderson method [28] in particular, which are widely used,
 59 state-of-the-art approximation methods.

60 Vector Fitting (VF) is a rational approximation technique, specifically tailored to
 61 functions in the frequency domain. It is based on a representation in terms of partial
 62 fractions as

$$63 \quad (1.5) \quad f(i\omega) \approx \sum_{m=1}^M \frac{r_m}{i\omega - p_m} + d + i\omega h,$$

64 where the M poles p_m are relocated in each iteration by solving a linear least-square
 65 problem, see [15, 16] for details. The implementation guarantees that all poles are
 66 stable, i.e. $\Re[p_m] < 0$, and are either real or come in complex-conjugate pairs.

67 The adaptive Antoulas-Anderson (AAA) method [28] employs the barycentric
 68 interpolation

$$69 \quad (1.6) \quad f(i\omega) \approx r(\omega) = \frac{n(\omega)}{d(\omega)} = \frac{\sum_{j \in J} \frac{w_j f(i\omega_j)}{\omega - \omega_j}}{\sum_{j \in J} \frac{w_j}{\omega - \omega_j}},$$

70 where $J \subseteq \{1, \dots, n\}$ has cardinality m . The rational function in (1.6) is of type
 71 $(m-1, m-1)$, which can be seen by multiplying both numerator and denominator
 72 by $\prod_{j \in J} (\omega - \omega_j)$. Moreover, $r(\omega_j) = f(i\omega_j)$ for all $j \in J$. The weights w_j and

¹Because of a different convention [7] establishes a negative *imaginary* part of the eigenvalues

73 nodes ω_j , $j \in J$, are determined adaptively in a two-step procedure, based on linear
 74 least squares problems and a greedy strategy [28].

75 Other data-driven approaches, related to rational interpolation and model order
 76 reduction are the Loewner framework [1] and the recent contribution [29], which
 77 employs the Heaviside representation. A Bayesian rational Polynomial Chaos-type
 78 model has been put forth in [37] to capture the effect of uncertain parameters, e.g.,
 79 on frequency response functions. A complex-valued version of support vector machine
 80 regression has been presented in [39], which is restricted to the so-called circular case
 81 with a single kernel only. Complex interpolation with a pair of kernels has been
 82 addressed in [6, 33] and also from a Gaussian process regression perspective in [5, 17].

83 Despite recent progress with complex kernel methods, a general framework with a
 84 complete mathematical background on the underlying reproducing kernel Hilbert spaces
 85 is missing. In this paper, we introduce a new kernel-based interpolation method
 86 which is well adapted to frequency responses. We will put special emphasis on the
 87 complex-valued setting and show that the data are used more efficiently if a dedi-
 88 cated kernel method is constructed and interpolation of the real- and imaginary part
 89 individually is avoided. To address problems with a few dominant poles we include
 90 a low-order rational basis into the kernel method and present a new model selection
 91 scheme. We compare our rational kernel-based interpolation method against both
 92 AAA and vector fitting and observe an improved or at least comparable performance
 93 for a variety of test cases. Finally, the paper develops the required notions of repro-
 94 ducing spaces and minimum norm interpolation for complex-valued kernel methods
 95 in general.

96 The material is structured in the following way. In Section 2 we introduce the
 97 concept of a complex/real kernel Hilbert space and consider the special case of fre-
 98 quency response functions as well as the connections to complex-valued Gaussian
 99 process regression. Section 3 introduces our new method, which employs a kernel,
 100 a pseudo-kernel and an additional rational basis for capturing dominant poles. Fi-
 101 nally, Section 4 reports several examples from PDE-based applications, comparing
 102 our method to AAA and vector fitting before conclusions are drawn.

103 *Nota bene: A method sharing some similarities with the one proposed in Section 3*
 104 *has been published recently in the automatic control literature [17]. We became aware*
 105 *of it at very late stage in the writing of the present article. After introducing our new*
 106 *method in Section 3, we discuss similarities and differences in Remark 3.3.*

107 **2. Complex/Real RKHS interpolation.** In order to address kernel-based
 108 interpolation of the frequency response function, we start by recalling basic facts
 109 on reproducing kernel Hilbert spaces (RKHSs); see, e.g., [30] for a comprehensive
 110 introduction to this topic.

111 **DEFINITION 2.1 (Complex RKHS).** *A complex RKHS H over a non-empty set \mathbb{S}*
 112 *is a complex Hilbert space of functions $\mathbb{S} \rightarrow \mathbb{C}$ such that, for all $s \in \mathbb{S}$, the evaluation*
 113 *functional $\delta_s : H \rightarrow \mathbb{C}$, $f \mapsto f(s)$, is continuous.*

114 The Riesz representation theorem implies that there exists a unique function
 115 $k : \mathbb{S} \times \mathbb{S} \rightarrow \mathbb{C}$, called the reproducing kernel of H , such that $k(\cdot, s) \in H$ and

116 (2.1)
$$f(s) = \delta_s(f) = \langle f, k(\cdot, s) \rangle_H$$

117 for all $s \in \mathbb{S}$ and $f \in H$, where $\langle \cdot, \cdot \rangle_H$ denotes the Hermitian inner product of H .
 118 Equation (2.1) is called the reproduction property, and it is easily seen that the

119 kernel k is *Hermitian* (i.e., $k(s, s_0) = k(s_0, s)^*$ for all $s, s_0 \in \mathbb{S}$) and *positive definite*:
 120 for all $n \in \mathbb{N}^*$ and all $(s_1, \alpha_1), \dots, (s_n, \alpha_n) \in \mathbb{S} \times \mathbb{C}$,

$$121 \quad (2.2) \quad \sum_{1 \leq i, j \leq n} \alpha_i^* \alpha_j k(s_i, s_j) \geq 0.$$

122 **THEOREM 2.2** (Moore-Aronszajn). *For any positive definite Hermitian kernel $k : \mathbb{S} \times \mathbb{S} \rightarrow \mathbb{C}$, there exists a unique complex Hilbert space H of functions on \mathbb{S} such that*
 123 *the reproduction property holds with reproducing kernel k .*
 124

125 Real RKHSs are defined similarly, replacing \mathbb{C} by \mathbb{R} in Definition 2.1: in this case
 126 H is a real Hilbert space, the reproducing kernel is symmetric positive definite, and
 127 a suitably modified statement of the Moore-Aronszajn theorem holds as well.

128 **THEOREM 2.3** (Interpolation). *Let H be a real or complex RKHS over \mathbb{S} with*
 129 *kernel $k : \mathbb{S} \times \mathbb{S} \rightarrow \mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or \mathbb{C} depending on the type of RKHS. Let $n \in \mathbb{N}^*$,*
 130 *$s_1, \dots, s_n \in \mathbb{S}$ and $y_1, \dots, y_n \in \mathbb{K}$. Then there exists a function $g \in H$ such that*
 131 *$g(s_i) = y_i$ for all $i \in \{1, \dots, n\}$ if, and only if, the system*

$$132 \quad (2.3) \quad \begin{bmatrix} k(s_1, s_1) & \dots & k(s_1, s_n) \\ \vdots & \ddots & \vdots \\ k(s_n, s_1) & \dots & k(s_n, s_n) \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

133 *admits a solution. Furthermore, for any solution of (2.3), $g = \sum_{i=1}^n \gamma_i k(\cdot, s_i)$ is the*
 134 *unique interpolant of the data $(s_1, y_1), \dots, (s_n, y_n)$ with minimal norm in H .*

135 A positive definite kernel is called *strictly positive definite* if the kernel matrix $K_n =$
 136 $(k(s_i, s_j))_{1 \leq i, j \leq n}$ is invertible (equivalently, if (2.2) is strict for all $(\alpha_1, \dots, \alpha_n) \neq 0$)
 137 whenever s_1, \dots, s_n are distinct points. This ensures that (2.3) has a unique solution.

138 We will proceed by introducing several complex RKHS and their kernels. For
 139 $s \in \mathbb{C}$, let $\Re[s]$ and $\Im[s]$ denote the real and imaginary part, respectively. An important
 140 example is the Hardy space $H^2(D)$ on the unit disc, where $D = \{s \in \mathbb{C} : |s| < 1\}$.
 141 This space plays a role in the analysis of the stability of discrete dynamical systems,
 142 see [2], for instance. Here, in the context of continuous-time dynamical systems, we
 143 are more interested in the corresponding Hardy space

$$144 \quad (2.4) \quad H^2(\Gamma_\alpha) = \left\{ f \in \text{Hol}(\Gamma_\alpha) : \|f\|_{H^2(\Gamma_\alpha)} = \sup_{x > -\alpha} \left(\int_{-\infty}^{\infty} |f(x + iy)|^2 dy \right)^{\frac{1}{2}} < \infty \right\},$$

145 where $\text{Hol}(\Gamma_\alpha)$ denotes the space of holomorphic functions on Γ_α . Note, that there
 146 is a Banach space isometry between the H^2 spaces on disc and half-plane, see [18,
 147 Chapter 8] for details.

148 **THEOREM 2.4.** *The space $H^2(\Gamma_\alpha)$ is a complex RKHS, with strictly positive def-*
 149 *inite reproducing kernel k given by*

$$150 \quad (2.5) \quad k_\alpha(s, s_0) = \frac{1}{2\pi(2\alpha + s + s_0^*)}, \quad s, s_0 \in \Gamma_\alpha.$$

151 A proof is given in Appendix A.1. Following standard terminology in complex
 152 analysis (see, e.g., [22]), we will refer to k_α as the *Szegő kernel* for the domain Γ_α .
 153 Evaluating (2.5) only on the imaginary axis $s = i\omega$, the expression simplifies to

$$154 \quad (2.6) \quad k_\alpha(i\omega, i\omega_0) = \frac{1}{2\pi(2\alpha + i(\omega - \omega_0))}, \quad \omega, \omega_0 \in \Omega.$$

155 We consider the stable spline kernel [34, 23] as another example. This kernel
 156 has been proposed in the time domain to model functions with a certain smoothness,
 157 which additionally incorporate impulse response stability [34]. The corresponding
 158 kernel for the frequency domain transfer function has been obtained in [23] and reads

159
 160 (2.7) $k_\alpha(i\omega, i\omega_0) = \frac{1}{2} \frac{1}{3\alpha + i(\omega - \omega_0)} \times$
 161 $\left(\frac{1}{2\alpha + i\omega} + \frac{1}{2\alpha - i\omega_0} - \frac{1}{3(3\alpha + i\omega)} - \frac{1}{3(3\alpha - i\omega_0)} \right).$
 162

163 Other related kernels can be found in the control literature, see [23, 17].

164 **2.1. Complex/real RKHS interpolation.** The frequency response function
 165 fulfills the symmetry property $f^*(s) = f(s^*)$ for all $s \in \Gamma_\alpha$, since it is the Laplace
 166 transform of a real-valued function. We are thus naturally led to cast our interpolation
 167 problem not in $H^2(\Gamma_\alpha)$ but in the subset

168 (2.8) $H_{\text{sym}}^2(\Gamma_\alpha) = \{f \in H^2(\Gamma_\alpha) : \forall s \in \Gamma_\alpha, f^*(s) = f(s^*)\}.$

169 This set of complex-valued functions, however, cannot be endowed with the structure
 170 of a complex RKHS. In fact, it is not even a vector space over \mathbb{C} : indeed, for any $f \in$
 171 $H_{\text{sym}}^2(\Gamma_\alpha)$ and $s \in \Gamma_\alpha$, we would have $(if)^*(s) = -if^*(s) = -if(s^*)$ and $(if)^*(s) =$
 172 $(if)(s^*) = if(s^*)$, which is a contradiction if $f(s^*) \neq 0$.

173 Observing that the subset of $H^2(\Gamma_\alpha)$ defined by (2.8) is a real vector space of
 174 complex-valued functions, we define in the following a new type of function space,
 175 which we call a complex/real RKHS.

176 **DEFINITION 2.5 (Complex/real RKHS).** *Let \mathbb{S} denote a non-empty set and let H*
 177 *denote a real Hilbert space of complex-valued functions on \mathbb{S} . We say that H is a*
 178 *complex/real RKHS if the evaluation functionals are continuous (i.e., for all $s \in \mathbb{S}$,*
 179 *the function $\delta_s : H \rightarrow \mathbb{C}, f \mapsto f(s)$, is continuous).*

180 In the remaining part of this section we will establish general results related to
 181 these spaces. Section 2.2 will then present consequences for the RKHS with the
 182 symmetry property $f^*(s) = f(s^*)$.

183 *Remark 2.6.* Any complex RKHS H (such as $H^2(\Gamma_\alpha)$) can be seen as a com-
 184 plex/real RKHS by forgetting the complex structure, i.e., by considering H as a real
 185 vector space, endowed with the real inner product $\langle f, g \rangle \mapsto \Re(\langle f, g \rangle_H)$. More gener-
 186 ally, any real subspace of H (such as $H_{\text{sym}}^2(\Gamma_\alpha)$), endowed with this inner product, is
 187 clearly a complex/real RKHS. The converse statement is false, however.

188 **PROPOSITION 2.7.** *There exists a complex/real RKHS of dimension two over the*
 189 *reals that is not a real subspace of a complex RKHS.*

190 The elements of a complex/real RKHS are complex-valued functions over \mathbb{S} , but
 191 can be conveniently represented as real-valued functions over $\tilde{\mathbb{S}} = \mathbb{S} \times \{\mathbb{R}, \mathbb{I}\}$ through
 192 the mapping $\mathcal{A} : \mathbb{C}^{\mathbb{S}} \rightarrow \mathbb{R}^{\tilde{\mathbb{S}}}$ defined by

193 (2.9) $(\mathcal{A}f)(s, a) = G_a(f(s)),$

194 where $G_{\mathbb{R}}(s) = \Re(s)$ and $G_{\mathbb{I}}(s) = \Im(s)$. This mapping defines an isometric isomor-
 195 phism of real Hilbert spaces between H and the real vector space $\tilde{H} = \mathcal{A}H \subseteq \mathbb{R}^{\tilde{\mathbb{S}}}$,
 196 endowed with the image inner product. The image space \tilde{H} is easily seen to be a real

197 RKHS if and only if H is a complex/real RKHS: this observation will be useful both
 198 from a theoretical point of view, to establish properties of complex/real RKHSs, and
 199 from a practical point of view (see Section 4).

200 *Remark 2.8.* Complex/real RKHSs can also be seen a special case of vector-
 201 valued RKHSs [9, 26], through the usual identification of \mathbb{C} with \mathbb{R}^2 .

202 The term “functional” is used in a loose sense in Definition 2.5, since H is a real
 203 vector space while δ_s is a complex-valued function. Therefore, in contrast with the
 204 usual case of complex RKHSs, the continuous functionals δ_s , $s \in \mathbb{S}$, do not belong to
 205 the topological dual of H . The real and imaginary evaluation functions however—
 206 namely, $\Re \circ \delta_s$ and $\Im \circ \delta_s$ —do belong to the topological dual, and can thus be expressed
 207 through inner products.

208 PROPOSITION 2.9. *Let H be a complex/real RKHS on a set \mathbb{S} , and set*

$$209 \quad (2.10) \quad k_{aa_0}(s, s_0) = \tilde{k}((s, a), (s_0, a_0)), \quad s, s_0 \in \mathbb{S}, \quad a, a_0 \in \{\mathbb{R}, \mathbb{I}\},$$

210 *where \tilde{k} denotes the reproducing kernel of $\tilde{H} = \mathcal{A}H$. Then, for all $s \in \mathbb{S}$, we have*

$$211 \quad (2.11) \quad \delta_s = \underbrace{\langle \cdot, \varphi_{\mathbb{R}}(\cdot, s) \rangle_H}_{\Re \circ \delta_s} + i \underbrace{\langle \cdot, \varphi_{\mathbb{I}}(\cdot, s) \rangle_H}_{\Im \circ \delta_s},$$

212 *where $\varphi_{\mathbb{R}} = k_{\mathbb{R}\mathbb{R}} + i k_{\mathbb{I}\mathbb{R}}$ and $\varphi_{\mathbb{I}} = k_{\mathbb{R}\mathbb{I}} + i k_{\mathbb{I}\mathbb{I}}$.*

213 This result associates to each complex/real RKHS a pair $(\varphi_{\mathbb{R}}, \varphi_{\mathbb{I}})$ of kernels $\varphi_a : \mathbb{S} \times \mathbb{S} \rightarrow \mathbb{C}$, $a \in \{\mathbb{R}, \mathbb{I}\}$. Characterizing admissible choices for this pair of kernels, in the
 214 spirit of Theorem 2.2 for complex RKHSs, is possible but not convenient. Instead,
 215 motivated by the connection between complex/real RKHSs and complex Gaussian
 216 processes (to be discussed in Section 2.3), and in particular the work of Picinbono
 217 [32], we introduce another pair of kernels as follows.

218 DEFINITION 2.10. *Let H denote a complex/real RKHS and let $k_{\mathbb{R}\mathbb{R}}, k_{\mathbb{I}\mathbb{I}}, k_{\mathbb{R}\mathbb{I}}, k_{\mathbb{I}\mathbb{R}}$,
 219 $\varphi_{\mathbb{R}}$ and $\varphi_{\mathbb{I}}$ be defined as in Proposition 2.9. Then we define the complex kernel k of
 220 the complex/real RKHS as*

$$222 \quad (2.12) \quad k = (k_{\mathbb{R}\mathbb{R}} + k_{\mathbb{I}\mathbb{I}}) + i(k_{\mathbb{I}\mathbb{R}} - k_{\mathbb{R}\mathbb{I}}) = \varphi_{\mathbb{R}} - i\varphi_{\mathbb{I}},$$

223 *and its pseudo-kernel c as:*

$$224 \quad (2.13) \quad c = (k_{\mathbb{R}\mathbb{R}} - k_{\mathbb{I}\mathbb{I}}) + i(k_{\mathbb{I}\mathbb{R}} + k_{\mathbb{R}\mathbb{I}}) = \varphi_{\mathbb{R}} + i\varphi_{\mathbb{I}}.$$

225 PROPOSITION 2.11. *The functions of the form $\gamma k(\cdot, s_0) + \gamma^* c(\cdot, s_0)$, with $\gamma \in \mathbb{C}$
 226 and $s_0 \in \mathbb{S}$, span a dense subset of H .*

227 *Remark 2.12.* Proposition 2.11 suggests that the concept of a complex/real
 228 RKHS, introduced in this article, provides a rigorous formalization of the idea of a
 229 “wide-linear complex-valued RKHS” (WL-RKHS) proposed in [6] (see Definition 3.1).

230 It can be shown that the complex/real RKHS obtained by forgetting the complex
 231 structure of a complex RKHS with reproducing kernel k_0 , as described in Remark 2.6,
 232 is the complex/real RKHS with complex kernel $k = 2k_0$ and vanishing pseudo-kernel—
 233 which, borrowing terminology from the signal processing literature [32], can be called
 234 *circular*. The factor 2 in the relation between k and k_0 is the price to pay for the
 235 consistency of Definition 2.10 with the concepts of covariance and pseudo-covariance
 236 functions for complex Gaussian processes (see Section 2.3). More generally, we have
 237 the following characterization of the set of admissible (k, c) pairs.

238 THEOREM 2.13. For a given complex/real RKHS H , the kernels k and c intro-
 239 duced in Definition 2.10 satisfy the following:

- 240 i) k is complex-valued, Hermitian and positive definite.
- 241 ii) c is complex-valued and symmetric.

242 Moreover, for all $n \geq 1$ and all $s_1, \dots, s_n \in \mathbb{S}$:

- 243 iii) $\ker K_n \subseteq \ker C_n^*$ and,
 - 244 iv) if K_n is positive definite, $K_n^* - C_n^* K_n^{-1} C_n$ is positive semi-definite,
- 245 where $K_n = (k(s_i, s_j))_{1 \leq i, j \leq n}$ and $C_n = (c(s_i, s_j))_{1 \leq i, j \leq n}$.

246 Conversely, for any pair of functions $k, c : \mathbb{S} \times \mathbb{S} \rightarrow \mathbb{C}$ that satisfies these four
 247 properties, there exists a unique complex/real RKHS on \mathbb{S} with complex kernel k and
 248 pseudo-kernel c .

249 THEOREM 2.14 (Interpolation in a complex/real RKHS). Let H denote a com-
 250 plex/real RKHS over \mathbb{S} with complex kernel k and pseudo-kernel c . Let $n \in \mathbb{N}^*$,
 251 $s_1, \dots, s_n \in \mathbb{S}$ and $y_1, \dots, y_n \in \mathbb{C}$. Then there exists a function $g \in H$ such that
 252 $g(s_i) = y_i$ for all $i \in \{1, \dots, n\}$ if, and only if, the system

$$253 \quad (2.14) \quad K_n \gamma + C_n \gamma^* = y$$

254 admits a solution $\gamma \in \mathbb{C}^n$, where $K_n = (k(s_i, s_j))_{1 \leq i, j \leq n}$, $C_n = (c(s_i, s_j))_{1 \leq i, j \leq n}$, and
 255 $y = (y_1, \dots, y_n)^T$. Furthermore, for any solution of (2.14),

$$256 \quad (2.15) \quad g = \sum_{i=1}^n \gamma_i k(\cdot, s_i) + \sum_{i=1}^n \gamma_i^* c(\cdot, s_i)$$

257 is the unique interpolant of the data $(s_1, y_1), \dots, (s_n, y_n)$ with minimal norm in H .

258 For the usual setting of real or complex RKHSs, strictly positive definite kernels
 259 guarantee that the interpolation system (2.3) has a solution for any data y_1, \dots, y_n .
 260 This remains true for the system (2.14) in the case of a complex/real RKHS if the
 261 associated real kernel \tilde{k} is strictly positive definite on $\tilde{\mathbb{S}} = \mathbb{S} \times \{\mathbb{R}, \mathbb{I}\}$.

262 **2.2. Complex/real RKHS with symmetry condition.** We now character-
 263 ize, in full generality, the complex/real RKHSs where a symmetry condition of the
 264 form $f^*(s) = f(s^*)$ holds for all $f \in H$ and $s \in \mathbb{S}$. The following theorem provides a
 265 necessary and sufficient condition on k for such a space to exist and gives the expres-
 266 sion of the corresponding pseudo-kernel. The expression appeared previously in [23,
 267 Equations (48)–(49)] for a special type of kernel.

268 THEOREM 2.15. Let \mathbb{S} denote a non-empty set, equipped with an involution $s \mapsto s^*$
 269 and $k : \mathbb{S} \times \mathbb{S} \rightarrow \mathbb{C}$ denote a Hermitian positive definite kernel on \mathbb{S} . Then the following
 270 assertions are equivalent:

- 271 i) There exists a complex/real RKHS H on \mathbb{S} , with complex kernel k , such that

$$272 \quad (2.16) \quad \forall f \in H, \forall s \in \mathbb{S}, \quad f^*(s) = f(s^*).$$

- 273 ii) There exists a complex/real RKHS H on \mathbb{S} , with complex kernel k and pseudo-
 274 kernel c defined by

$$275 \quad (2.17) \quad \forall s, s_0 \in \mathbb{S}, \quad c(s, s_0) = k(s, s_0^*).$$

- 276 iii) $\forall s, s_0 \in \mathbb{S}, \quad k(s, s_0^*) = k(s_0, s^*)$.

277 *If any (and consequently all) of these assertions holds, then the complex/real*
 278 *RKHS H with complex covariance k and pseudo kernel (2.17) is the unique RKHS*
 279 *on \mathbb{S} with complex covariance k such that (2.16) holds. Moreover, denoting by $H_{\mathbb{C}}$ the*
 280 *complex RKHS with kernel k , we have $H_{\mathbb{C}} = H \oplus iH$, $H = \{f \in H_{\mathbb{C}} \mid (2.16) \text{ holds}\}$*
 281 *and $\langle f, g \rangle = \Re \langle f, g \rangle_{H_{\mathbb{C}}}$ for all $f, g \in H$.*

282 It follows from this theorem that $H_{\text{sym}}^2(\Gamma_{\alpha})$ can be characterized as the com-
 283 plex/real RKHS over Γ_{α} with complex kernel (2.5) and pseudo-kernel:

$$284 \quad (2.18) \quad c_{\alpha}(s, s_0) = \frac{1}{2\pi(2\alpha + s + s_0)}, \quad s, s_0 \in \Gamma_{\alpha}.$$

285 More generally, Theorem 2.15 shows that the problem of minimum-norm inter-
 286 polation in a complex RKHS, with a symmetry constraint of the form (2.16), can
 287 be solved by considering the equivalent problem of minimal-norm interpolation in
 288 the complex/real RKHS with the same complex kernel and the pseudo-kernel given
 289 by (2.17). In presence of the symmetry condition, even if the complex kernel k is
 290 strictly positive definite, \tilde{k} is not and an additional condition on the data is required
 291 to ensure that (2.3) has a solution.

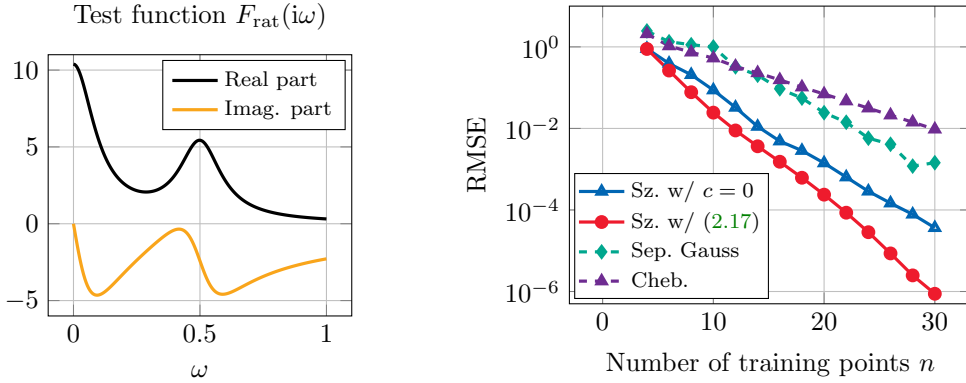
292 **THEOREM 2.16.** *In the setting of Theorem 2.15, assume that k is strictly positive*
 293 *definite, c is given by (2.17), and $s_1, \dots, s_n \in \mathbb{S}$ are distinct. Then (2.14) has a*
 294 *solution if, and only if, $y_j = y_i^*$ for all i, j such that $s_j = s_i^*$. When this holds, there*
 295 *is a unique solution such that $\gamma_i = \gamma_j^*$ for all i, j such that $s_j = s_i^*$.*

296 For illustration, we consider the third order rational function

$$297 \quad (2.19) \quad F_{\text{rat}}(i\omega) = \frac{1}{i\omega - (-0.1)} + \frac{0.5}{i\omega - (-0.1 - 0.5i)} + \frac{0.5}{i\omega - (-0.1 + 0.5i)}, \quad \omega \in [0, 1],$$

298 which is the Laplace transform of the real-valued function $t \mapsto e^{-0.1t}(1 + \cos(0.5t))$
 299 and thus belongs to $H_{\text{sym}}^2(\Gamma_{0.1+\epsilon}) \subseteq H^2(\Gamma_{0.1+\epsilon})$ for all $\epsilon > 0$. To illustrate the im-
 300 portance of the choice of pseudo-kernel, we conduct a convergence study in terms of
 301 the root-mean-square error (RMSE) of the approximations, using equidistant training
 302 ing points (details on the implementation and selection of hyper-parameters will be
 303 given in the following sections). In Figure 1 we demonstrate that choosing a suit-
 304 able pseudo-kernel might have a significant impact on the convergence properties of
 305 the (complex/real) RKHS interpolation. For the test function (2.19), the pseudo-
 306 kernel (2.17) improves the convergence significantly. Note that the test function is
 307 a low order rational function which is here only used to illustrate the impact of the
 308 pseudo-kernel. Accordingly, rational interpolation techniques as AAA or VF reach
 309 machine accuracy already with ≈ 8 training points and are hence excluded in the
 310 convergence plot for clarity. However, it can already be observed that complex/real
 311 RKHS interpolation with the Szegő kernel outperforms the alternative approach of
 312 separate kernel approximations for real and imaginary part with a Gaussian kernel,
 313 as well as polynomial interpolation on Chebyshev nodes.

314 **2.3. Relation to Gaussian process interpolation.** This section draws con-
 315 nections between minimum norm interpolation in a RKHS and the posterior mean
 316 prediction of a Gaussian process (GP), for both the complex and complex/real case.
 317 GPs are widely used, but to the authors knowledge this is the first time that the
 318 RKHS associated to any complex GP prediction is characterized. Another intention
 319 of this section is to make results from the GP literature available for interpolation



(a) Function $F_{\text{rat}}(i\omega)$ (with real-valued inverse Laplace transform).

(b) Convergence study of the RMSE for different approximations of $F_{\text{rat}}(i\omega)$.

Fig. 1: Left: Illustrations with the test function $F_{\text{rat}} \in H_{\text{sym}}^2(\Gamma_{0.1+\epsilon})$, $\epsilon > 0$ defined in (2.19). Right: Convergence of the RMSE as a function of the number of (equidistant) training points. Solid lines: complex/real interpolation with the Szegő kernel for $H^2(\Gamma_\alpha)$, combined with the zero pseudo-kernel (blue) and the pseudo-kernel (2.17) (red). Dashed lines: interpolation with a Gaussian kernel for the real and imaginary part separately (green) and polynomial interpolation on Chebyshev nodes (purple).

320 with a complex/real RKHS. In particular, we are interested in employing statistical
 321 methods for model selection (see, e.g., [31] and references therein)—this will be fur-
 322 ther developed in Section 3.2. We consider zero-mean processes in this section, for
 323 simplicity see Remark 2.19 below.

324 Complex GPs are covered for instance in [27]. A complex GP is a complex process,
 325 where the real and imaginary part considered jointly are a real GP. We consider a
 326 zero-mean complex-valued random process ξ on \mathbb{S} , with covariance function \mathfrak{k} and
 327 pseudo-covariance function \mathfrak{c} :

$$328 \quad (2.20) \quad \mathbf{E}(\xi(s)\xi(s_0)^*) = \mathfrak{k}(s, s_0),$$

$$329 \quad (2.21) \quad \mathbf{E}(\xi(s)\xi(s_0)) = \mathfrak{c}(s, s_0).$$

331 Relying on the mapping \mathcal{A} , we can work in a real-valued setting, i.e., with a real-
 332 valued GP ξ indexed on \mathbb{S} . In the real-valued case, it is well-known that the conditional
 333 mean of a GP is identical to the minimum-norm interpolant in the RKHS associated
 334 to its covariance function. Hence, using \mathcal{A} , the conditional mean of a complex GP ξ
 335 is also identical to a minimum-norm interpolant, but this time in a complex/real RKHS,
 336 the complex kernel k and pseudo-kernel c of which are equal to \mathfrak{k} and \mathfrak{c} respectively
 337 (this follows from Equations (2.12)–(2.13)). It is given by Equation (2.15) in general,
 338 which simplifies to

$$339 \quad (2.22) \quad \mathbf{E}(\xi(s)|y) = \sum_{i=1}^n \gamma_i k(s, s_i), \quad \text{with } K_n \gamma = y,$$

340 if the pseudo-covariance is zero (i.e., in the circular case).

341 *Remark 2.17.* A common approach to deal with complex data is to use GP inter-
 342 polation for the real and imaginary part separately (see, e.g., [14]). This corresponds,
 343 using notations from Proposition 2.9, to $k_{RI} = k_{IR} = 0$, and therefore to a complex
 344 GP with covariance $k = k_{RR} + k_{II}$ and pseudo-covariance $c = k_{RR} - k_{II}$.

345 *Remark 2.18.* GP regression with both covariance and pseudo-covariance function
 346 has also been considered under the name widely linear posterior mean. In [32] it is
 347 first shown that the posterior mean is widely linear [33], which leads to

$$348 \quad (2.23) \quad \mathbb{E}(\xi(s)|y) = (k_{s,n} - c_{s,n}K_n^{-*}C_n^H)P_n^{-*}y + (c_{s,n} - k_{s,n}K_n^{-1}C_n)P_n^{-1}y^*,$$

349 where $P_n = K_n^* - C_n^H K_n^{-1} C_n$ and P_n^{-*} denotes the complex conjugate of the inverse
 350 of P_n . The formulas for the circular and non-circular case can also be found in [5].

351 *Remark 2.19.* In practice, GP models often include a non-zero mean function m ,
 352 usually written as a linear combination $m(x) = \sum_{\ell=1}^L \beta_\ell h_\ell(x)$ of known basis func-
 353 tions h_ℓ , with unknown coefficients β_ℓ . If the coefficients are estimated by maximum
 354 likelihood (as in Section 3), the posterior mean of the GP is then equal to the inter-
 355 polant with minimal *semi-norm* in $G = V + H$, where $V = \text{span}\{h_1, \dots, h_L\}$ and the
 356 semi-norm is defined by $|g|_G = \inf_{v \in V} \|g - v\|_H$.

357 **3. Hybrid algorithm.** We focus from now, unless otherwise specified, on func-
 358 tions satisfying the property $f^*(s) = f(s^*)$, and we employ the Szegő kernel (2.5),
 359 together with the pseudo-kernel (2.18), for complex/real interpolation. In practice,
 360 the convergence of complex/real RKHS interpolation can be significantly slower than
 361 that of rational approximations techniques (such as AAA or VF) when the func-
 362 tion has a few dominant poles p_i , i.e., poles with small attenuation $\Re[p_i] \approx 0$. In
 363 this section, we discuss how complex/real RKHS interpolation with the Szegő kernel
 364 and associated pseudo-kernel can be combined with a small number of rational basis
 365 functions for the approximation of such frequency response functions.

366 **3.1. Gaussian process model.** We propose to use a complex GP model with
 367 rational mean function $m = \sum_{\ell=1}^L \beta_\ell h_\ell$ (cf. Remark 2.19), covariance function $\sigma^2 k_\alpha$
 368 and pseudo-covariance function $\sigma^2 c_\alpha$, where k_α denotes the Szegő kernel (2.5), c_α
 369 the associated pseudo-kernel (2.18), and σ^2 , α , β_1, \dots, β_L are real parameters with
 370 $\sigma^2 > 0$ and $\alpha > 0$. For the mean function m we assume a rational function satisfying
 371 the property $m^*(s) = m(s^*)$, of the form

$$372 \quad (3.1) \quad m(s) = \sum_{i=1}^K \left\{ \frac{1}{s - p_i} r_i + \frac{1}{s - p_i^*} r_i^* \right\},$$

373 with residues $r_1, \dots, r_K \in \mathbb{C}$ and (stable) complex conjugate poles $p_1, p_1^*, \dots, p_K, p_K^* \in$
 374 \mathbb{C} such that $\Re(p_i) < 0$ and $\Im(p_i) > 0$ for all i . This representation is similar to the
 375 one used in VF [15, 16]. Equation (3.1) can be rewritten as $m = \sum_{\ell=1}^L \beta_\ell h_\ell$ with
 376 $L = 2K$,

$$377 \quad \beta_\ell = \begin{cases} \Re(r_i) & \text{if } \ell = 2i - 1, \\ \Im(r_i) & \text{if } \ell = 2i, \end{cases} \quad \text{and} \quad h_\ell(s) = \begin{cases} \frac{1}{s - p_i} + \frac{1}{s - p_i^*} & \text{if } \ell = 2i - 1, \\ \frac{i}{s - p_i} - \frac{i}{s - p_i^*} & \text{if } \ell = 2i. \end{cases}$$

378 Note that m is an element of $H_{\text{sym}}^2(\Gamma_{\alpha'})$ with $\alpha' = \min_{1 \leq i \leq K} |\Re(p_i)| + \epsilon$, $\epsilon > 0$. For
 379 simplicity we only consider complex conjugate poles in (3.1), but real poles could be
 380 included as well, as in VF. In the context of the present work, we typically consider
 381 a small number K of pole pairs ($K \leq K_{\max} = \min(5, \lfloor n/4 \rfloor$) in the examples).

382 For a given number K of pole pairs, we select the hyper-parameters σ^2 , α ,
 383 $\mathbf{p} = (p_1, \dots, p_K)$ and $\mathbf{r} = (r_1, \dots, r_K)$ by maximization of a penalized log-likelihood
 384 function, where the penalty stems from a vague log-normal prior on α ; see Supple-
 385 mentary Material for details. An original procedure for the selection of an appropriate
 386 number K of pole pairs will be presented in the next section.

387 *Remark 3.1.* Note that we do not include a constant basis function, as is usually
 388 done in Gaussian process modeling, to ensure that the interpolant satisfies the desired
 389 property (namely, goes to zero) when $\omega \rightarrow \pm\infty$.

390 *Remark 3.2.* Alternatively, the residues r_1, \dots, r_K could be integrated out ana-
 391 lytically using a Gaussian prior, resulting in additional terms in the covariance and
 392 pseudo-covariance functions of the GP; see, e.g., [17]. This would allow the uncer-
 393 tainty on the residues, for a given set of poles, to be reflected in the uncertainty
 394 quantification (posterior variances) produced by the GP model. We do not pursue
 395 this idea further in this article, since our focus is on interpolation rather than uncer-
 396 tainty quantification.

397 **3.2. Adaptive pole selection.** Selecting a suitable number K of pole pairs to
 398 be included in the mean function (3.1) is a crucial step to ensure good accuracy of the
 399 proposed hybrid method. In this section we propose a model selection procedure to
 400 select this number automatically, in a data-driven manner. While this procedure relies
 401 on the well-established idea of (leave-one-out) cross-validation, it contains an original
 402 ingredient in the form an “instability penalty”, which will be described below.

403 First we build $K_{\max} + 1$ interpolants $f_n^{(K)}$, where the superscript K indicates the
 404 number of pole pairs, ranging from 0 (zero-mean Gaussian process model) to K_{\max} .
 405 Following standard VF practice [16], we begin with the maximum number of poles,
 406 $K = K_{\max}$, using an equidistant distribution of poles close to the frequency axis as
 407 a starting point for optimization. The other interpolants are then constructed itera-
 408 tively, going backwards: at each step optimization is initialized using K of the $K + 1$
 409 poles selected at the previous step, by removing the least relevant pole according to
 410 the (penalized) log-likelihood function.

411 Model selection is then based on leave-one-out (LOO) cross-validation, i.e., on
 412 the error indicators

$$413 \quad (3.2) \quad \epsilon_{\text{LOO}}^K = \frac{1}{n} \sum_{i=1}^n \left| f(i\omega_i) - \hat{f}_{n-1,i}^{(K)}(i\omega_i) \right|^2, \quad K = 0, 1, \dots, K_{\max},$$

414 where $\hat{f}_{n-1,i}^{(K)}$ denotes a model constructed without the i -th data point. Keeping the
 415 poles and kernel hyper-parameters fixed, when removing points, makes it possible to
 416 reduce the computational effort, but was found to introduce an undesired preference
 417 for models with a larger number of poles. Hence, we employ the LOO criterion with
 418 re-tuning, using the poles and hyper-parameters of $f_n^{(K)}$ as an initial guess when
 419 constructing $\hat{f}_{n-1,i}^{(K)}$, $1 \leq i \leq n$.

420 Furthermore, we introduce an additional penalty term, which also takes *global*
 421 model variations into account. This approach can be motivated by the example il-
 422 lustrated in Figure 2 (top). The corresponding vibro-acoustic benchmark model will
 423 be described in Section 4, however, here we simply consider the approximation of
 424 the dashed function, based on interpolation of the training points (black dots), as
 425 a general example. At the top, it can be observed that the LOO criterion (3.2)
 426 leads to the selection of a model (solid lines) $\hat{f}_n^{(5)}$ which wrongly identifies a pole

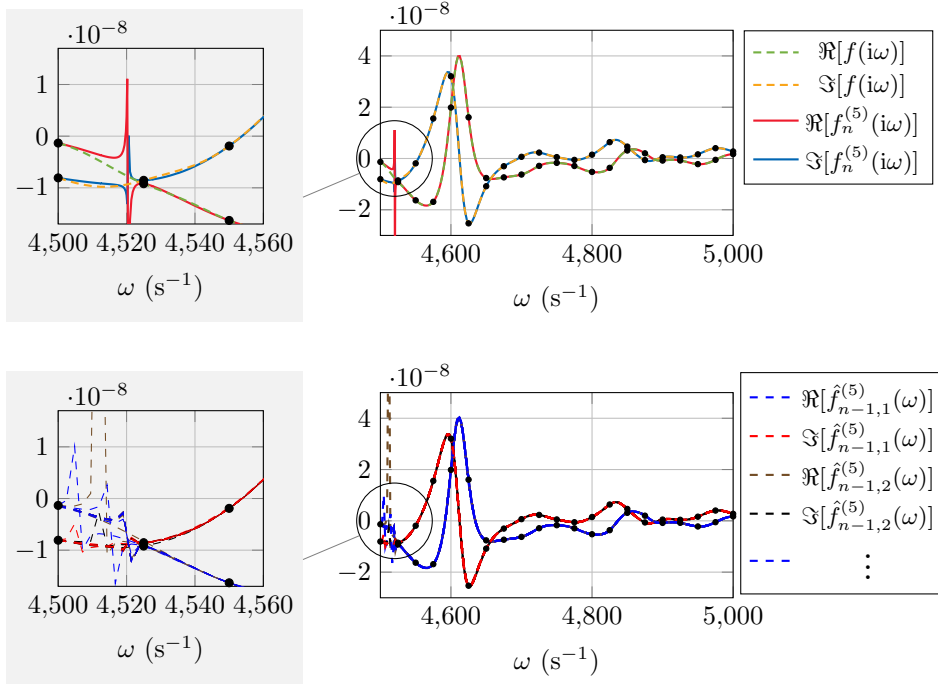


Fig. 2: Top: Dashed lines show the function to approximate. Black dots indicate the training data. Solid lines represent a *bad* approximation model which, however, is selected by the LOO criterion. Zoomed plot (gray background) highlights the influence of a wrongly identified pole. Bottom: Leave-on-out predictions, which show strong local variations between 4500 s^{-1} and 4520 s^{-1} . However, these variations do not significantly affect the values at the respective training points.

427 at $\approx 4520 \text{ s}^{-1}$. However, this effect is rather local, it mainly takes place between
 428 two training points (illustrated by black dots). At the bottom, we show the models
 429 $\hat{f}_{n-1,i}^{(K)}(\omega)$, $i = 1, \dots, n$, which show strong variations close to $\approx 4510 \text{ s}^{-1}$ but rather
 430 small errors at the training points ω_i . To take this into account, we introduce an
 431 instability penalty term, which leads to the criterion

$$432 \quad (3.3) \quad \epsilon_{\text{LOO,stab}}^K = \epsilon_{\text{LOO}}^K + \lambda \frac{1}{n} \frac{1}{M} \sum_{i=1}^n \sum_{j=1}^M \left| f_n^{(K)}(i\hat{\omega}_j) - \hat{f}_{n-1,i}^{(K)}(i\hat{\omega}_j) \right|^2,$$

433 where $\{\hat{\omega}_j\}_{j=1}^M$ denotes a fine grid on Ω (more precisely, an equidistant grid with
 434 $M = 10n + 1$ points). The weighting factor λ is chosen as

$$435 \quad (3.4) \quad \lambda = 0.2 \frac{\epsilon_{\text{LOO}}^0}{\frac{1}{n} \frac{1}{M} \sum_{i=1}^n \sum_{j=1}^M \left| f_n^{(0)}(i\hat{\omega}_j) - \hat{f}_{n-1,i}^{(0)}(i\hat{\omega}_j) \right|^2},$$

436 i.e., 0.2 after normalizing both terms w.r.t. the respective values of the purely kernel-
 437 based interpolation model. To our knowledge, this approach for model selection has
 438 not been considered before, although it is related to the continuously-defined LOO

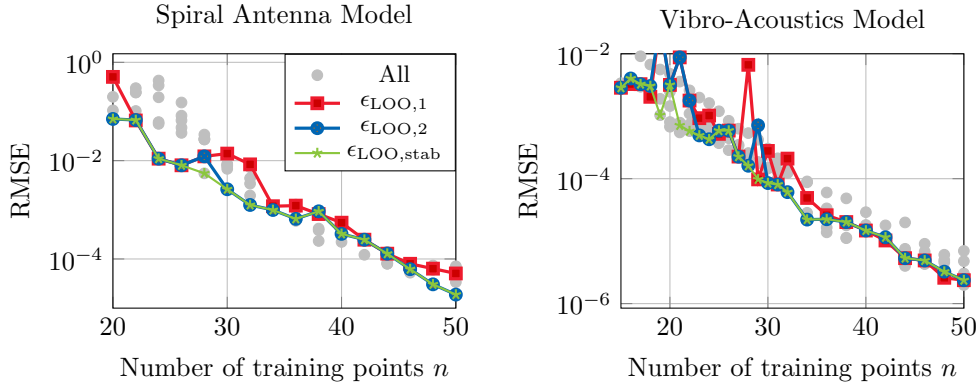


Fig. 3: Comparison of different model selection criteria for two benchmark problems. $\epsilon_{\text{LOO},1}$ and $\epsilon_{\text{LOO},2}$ denote the leave-on-out residual without and with retuning of hyper-parameters, respectively. The stabilized criterion $\epsilon_{\text{LOO},\text{stab}}$ (with retuning) defined in (3.3) gives the best results.

439 error [20, 21, 13]. The continuously-defined LOO error was employed for sequential
 440 sampling, while we propose to use it to construct an instability penalty for model
 441 selection. Stability selection [24, 25] is another related approach, which is also based
 442 on resampling of the data, but usually employed for variable selection.

443 Employing the stabilized criterion (3.3) for model selection gives satisfactory re-
 444 sults for the benchmark examples considered in this work. For illustration, we consider
 445 the convergence studies for two models, which will be described in Section 4. Figure 3
 446 shows the root-mean-square-errors (RMSEs) of the available models with gray dots
 447 and the accuracy of the selected models by the different criteria. It can be observed
 448 that the stabilized criterion $\epsilon_{\text{LOO},\text{stab}}^K$ gives the best results, while LOO residuals with
 449 retuning is superior to the approach without retuning.

450 *Remark 3.3.* The combination of kernel methods with a small number of rational
 451 basis functions has also been considered in [17] for data-driven modeling of frequency
 452 response functions. Therein, the authors employ first order stable spline kernels, which
 453 encode stability, causality and smoothness and add a rational basis for capturing the
 454 resonant poles of the transfer function. A prior is formed over the impulse responses
 455 linked to the resonant poles, which allows to derive additional kernels (one for each
 456 resonant pole) via the Fourier transform.

457 Our approach proceeds in a similar way, as our VF-inspired rational basis could
 458 also be transformed into additional kernels through a prior over β . Differences can be
 459 found in the model selection strategies, which are based on the local rational method in
 460 [17], whereas our approach is based on statistical model selection. Additionally, our fo-
 461 cus here is on providing a complete background on the RKHS concepts of complex/real
 462 interpolation, whereas [17] is additionally targeting uncertainty quantification for the
 463 data-driven modeling procedure.

464 **4. Numerical results.** We apply the presented approximation techniques to a
 465 number of benchmark functions from different fields. We always employ n training
 466 points $(\omega_i, f(i\omega_i))$, where the ω_i are equidistant frequency points in $[\omega_{\min}, \omega_{\max}]$, for
 467 simplicity. The accuracy of different approximations is then quantified in terms of

468 the root-mean-square error (RMSE), which is evaluated on a refined equidistant grid
 469 with 201 points for all numerical examples.

470 In the following we give a few details on the implementation. For AAA [28],
 471 we rely on the implementation of the `chebfun` toolbox [12]. For VF, we employ
 472 the `VectFit3` toolbox [16, 15, 11], where we use complex equidistant starting poles
 473 distributed according to the general recommendation, and always run 30 iterations.
 474 We apply the “relaxed non-triviality constraint” [15], include the constant but not
 475 the linear term, and enforce stable poles. The number of complex starting pole pairs
 476 is set to the maximum number of $2\lfloor \frac{n-1}{2} \rfloor$, which leads to the best results for the
 477 smooth test functions considered. For kernel interpolation we consider a separate
 478 interpolation of the real and imaginary part with the squared exponential kernel (SE)
 479 and complex/real interpolation with the Szegő kernel. The latter is also considered in
 480 combination with an adaptive rational basis (Sz.-Rat.) as described in Section 3. The
 481 implementation is done in `Matlab` as well, based on the `STK` toolbox [4]. To this end,
 482 we employ the mapping \mathcal{A} defined in (2.9) for the complex/real RKHS interpolation,
 483 which allows to realize the implementation based on real RKHS interpolation on
 484 an augmented input space $\Omega \times \{0, 1\}$. Note that this approach could be employed
 485 with any toolbox for real RKHS interpolation that provides the option to specify
 486 custom kernel functions. The tuning of the hyper-parameters and poles based on
 487 the likelihood function (see Section 3) is carried out using `fmincon` in `Matlab`, i.e.,
 488 gradient-based optimization (more precisely an interior point algorithm), which we
 489 combine with a multistart procedure; see Supplementary Material for more details.

490 *Remark 4.1.* By investigating the shape of the likelihood function for a number
 491 of benchmark problems, we have found that the logarithmic reparameterization, dis-
 492 cussed in [3] for instance, is not beneficial for the parameter α . Hence, it is only
 493 applied to the scaling parameter σ .

494 **4.1. Electric circuit (high order rational function).** We consider in the
 495 following a parallel connection of N underdamped series RLC circuits, as illustrated
 496 on the left side in Figure 4. The admittance is given as

$$497 \quad (4.1) \quad Y(s) = \sum_{i=1}^N \frac{s}{s^2 L_i + s R_i + C_i^{-1}} = \sum_{i=1}^N \frac{c_i}{s - a_i} + \frac{c_i^*}{s - a_i^*},$$

498 where $\Re[a_i] = -\frac{R_i}{2L_i}$ (an explicit representation of the poles a_i and residues c_i is given
 499 in the Supplementary Material) and we consider the frequency range [10 kHz, 25 kHz].
 500 First, we assume $N_1 = 1000$ random series RLC elements, where $C_i \sim \mathcal{U}(1, 20)$ μF
 501 and $L_i \sim \mathcal{U}(0.1, 2)$ mH , and we assume the resistance R_i to be roughly proportional
 502 to the inductance, with random variations of $\pm 20\%$: $R_i = L_i(1 + \Delta) \Omega(\text{mH})^{-1}$, where
 503 $\Delta \sim \mathcal{U}(-0.2, 0.2)$.

504 Note that for any combination of those parameters, the corresponding series RLC
 505 circuits are underdamped. For one particular realization, the distribution of the
 506 $2N = 2000$ poles is illustrated in Figure 4. The corresponding admittance $Y_1(i\omega)$
 507 is shown in Figure 5 with dashed black lines. We then conduct a convergence study
 508 for the particular realization of the electric circuit, which is shown in Figure 6 (top,
 509 left). We repeat the convergence study for 100 random realizations and depict the
 510 median RMSE at each point in Figure 6 (top, right). It can be observed that for
 511 the considered range of the number of training points (where $n \leq 60 \ll N$) the
 512 complex/real Szegő kernel-based interpolation outperforms AAA and VF. Employing
 513 the hybrid algorithm (Sz.-Rat.) does not yield an improvement, but leads to similarly

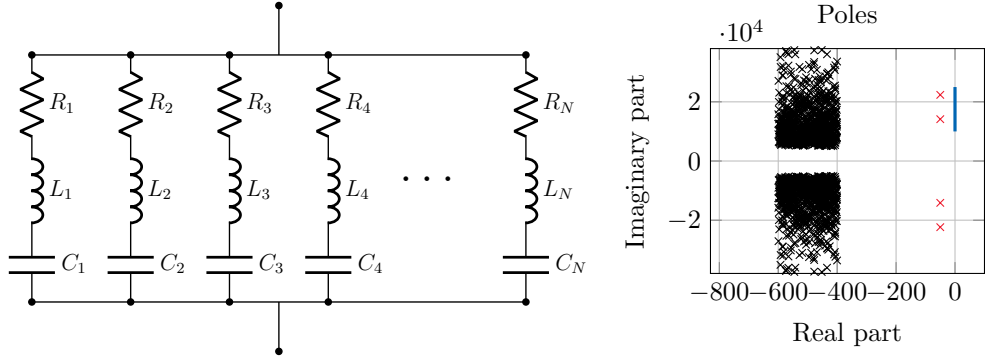


Fig. 4: Left: Parallel connection of (underdamped) series RLC circuits. Right: Black crosses indicate the distribution of $2N_1 = 2000$ poles of the circuit admittance Y_1 in the complex plane. Red crosses indicate the two additional poles considered for the circuit admittance Y_2 with $2N_2 = 2004$ poles. Blue line indicates the considered frequency range.

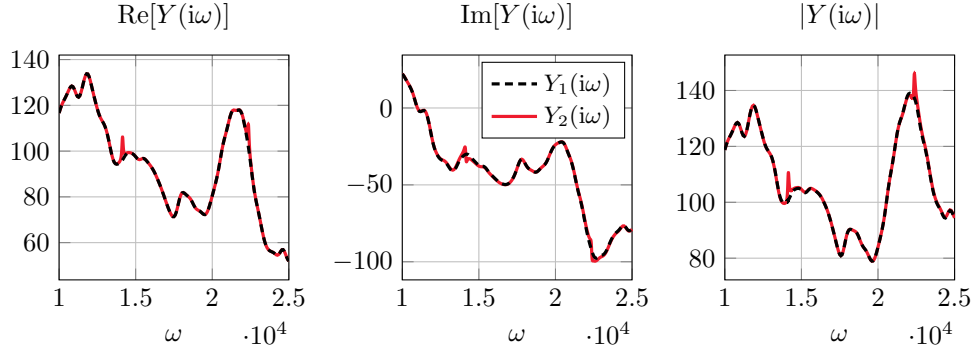


Fig. 5: Complex admittances Y_1 and Y_2 of the electric circuits versus frequency for a particular random parameter realization and $N_1 = 1000$ and $N_2 = 1002$, respectively.

514 good results.

515 In our second experiment, we introduce two additional circuit elements with a
 516 very small damping, i.e. we now consider $N_2 = 1002$ and

517 $C_{1001} = 5 \text{ pF}, \quad L_{1001} = 1 \text{ mH}, \quad R_{1001} = 0.1 \Omega,$
 518 $C_{1002} = 2 \text{ pF}, \quad L_{1002} = 1 \text{ mH}, \quad R_{1002} = 0.1 \Omega.$

520 This leads to two additional poles which are closer to the input domain, as illustrated
 521 by the red crosses in Figure 4. The corresponding admittance $Y_2(i\omega)$ differs very little
 522 from $Y_1(i\omega)$, except for two sharp peaks, as can be seen in Figure 5. However, the
 523 accuracy of the respective RKHS interpolation is significantly affected. In particular,
 524 at the bottom of Figure 6, it can be observed that the convergence order of Szegő
 525 kernel interpolation is significantly reduced. By adding the rational basis we are able
 526 to mitigate the impact of the two dominant poles: it exhibits fast convergence and an
 527 improvement w.r.t. AAA and VF can again be observed.

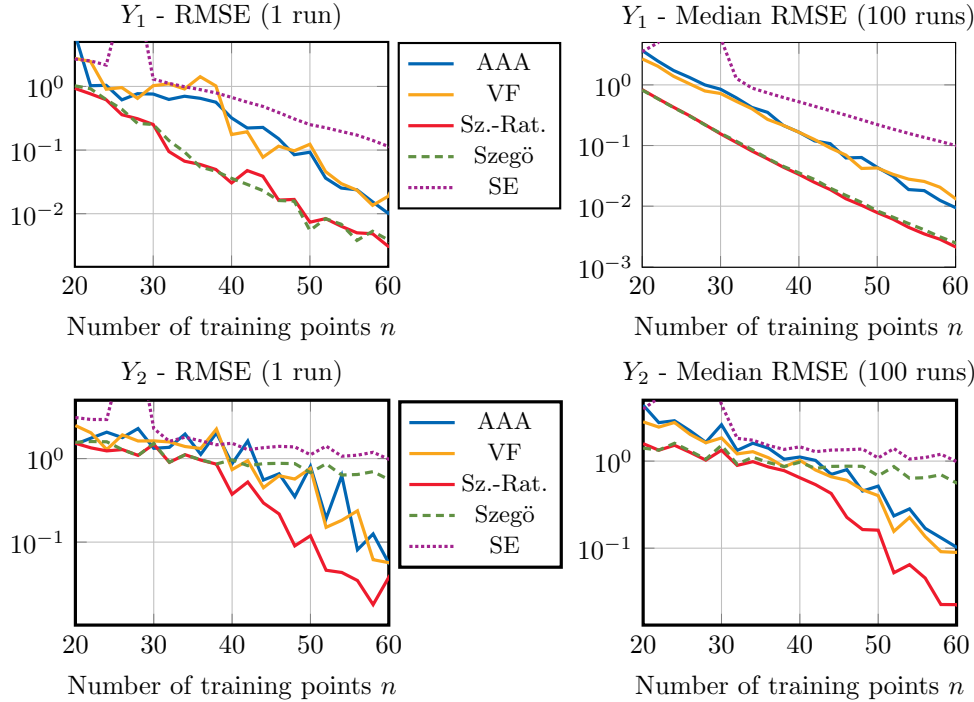


Fig. 6: Convergence study for admittances Y_1 (top) and Y_2 (bottom). Left: RMSE for one particular realization. Right: Median for 100 random realizations.

528 **4.2. PDE-based examples.** In the following, we investigate a number of PDE-
529 based examples. We start with the acoustic Helmholtz equation, in particular, the
530 PAC-MAN benchmark example, introduced in [40] which is also included in the plat-
531 form for benchmark cases in computational acoustics from the European Acoustics
532 Association [19]. The model, shown in Figure 7, has the PAC-MAN shape with an
533 opening angle of 30° and radius of 1 m. As in [40, Section 6.1], we consider as excita-
534 tion a vibration of the surface of the PAC-MAN with cylindrical modes and observe
535 the radiated field p_i at a point in 2 m distance at an angle of 10° . As in [19], the
536 computation was done based on the implementation of the analytical solution pro-
537 vided in [40] by replacing the python module `scipy` by `mpmath` for the computation
538 of higher order Bessel functions. In particular, we set the truncation order to 300.
539 The complex acoustic pressure field phasor p_i of the total sound-field versus the fre-
540 quency $f \in [2000 \text{ Hz}, 4000 \text{ Hz}]$ is shown in Figure 7 (top, right). We then conduct a
541 convergence study w.r.t. the number of training points, which is depicted in Figure 7
542 (bottom, left). It can be observed that the complex/real Szegő kernel-based inter-
543 polation outperforms the alternative approaches in the range up to about 40 training
544 points. Adding the rational mean function does not further improve the accuracy,
545 but does not harm the accuracy either.

546 Next, we consider an electromagnetic model problem, which is a demonstration
547 example of CST Microwave Studio [10], solving the full set of Maxwell equations in the
548 frequency domain. The model consists of a waveguide junction with 4 ports, which
549 contains a small metallic disk and is connected to an external cavity resonator (see

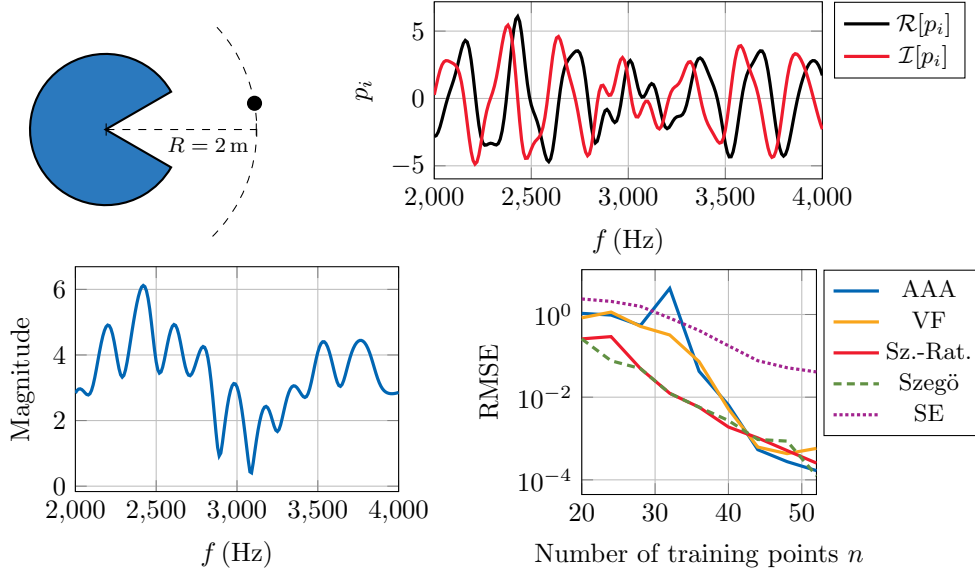


Fig. 7: Top left: We consider a surface vibration of the PAC-MAN model and evaluate the radiated acoustic field p_i at a point (black dot) in 2 m distance to the center. Top right: Complex frequency response function. Bottom left: Magnitude of frequency response function. Bottom right: Convergence study w.r.t. the number of training points.

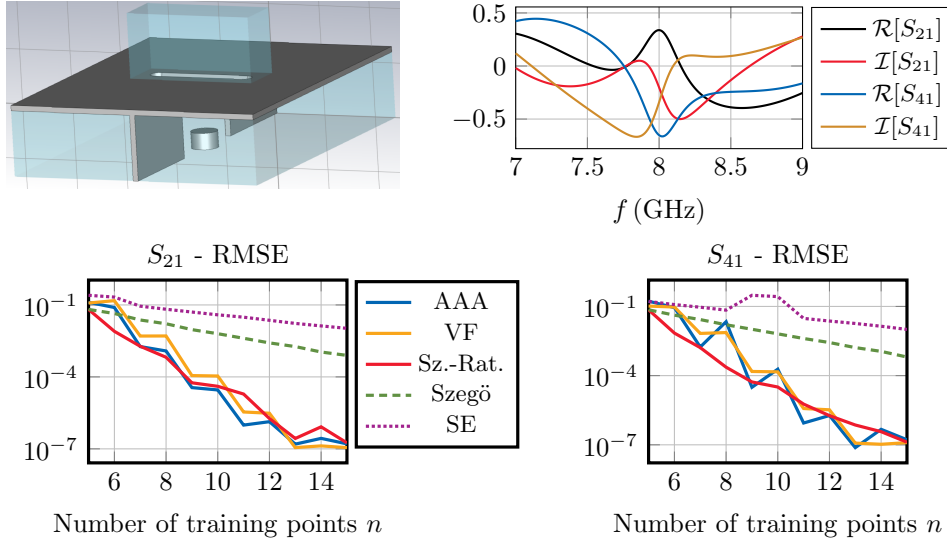


Fig. 8: Top left: Waveguide junction model, taken from CST Microwave Studio [10]. Top right: Complex frequency response functions. Bottom: Convergence studies w.r.t. the number of training points.

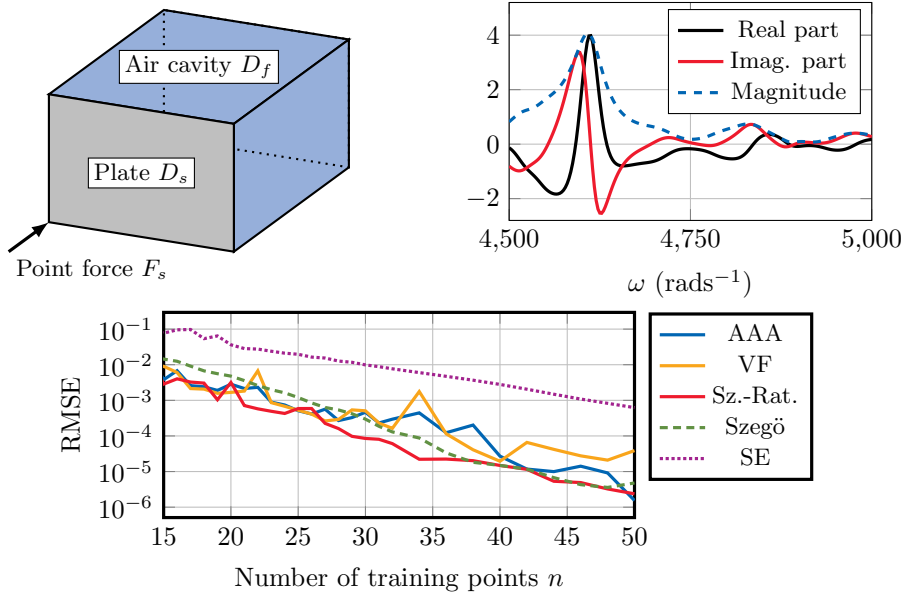


Fig. 9: Top left: Vibro-acoustic benchmark problem, based on [36]. Top right: Complex frequency response function. Bottom: Convergence study w.r.t. the number of training points.

550 Figure 8). The structure is excited at the first port and simulated using the finite
 551 element method in the frequency domain. In particular, we set the solver accuracy
 552 of the 3rd order solver to 10^{-6} and use a curved mesh with standard settings. We
 553 employ an initial adaptive mesh refinement at 9 GHz, where we set the scattering
 554 parameter criterion threshold with 2 subsequent checks to 10^{-4} . As quantity of inter-
 555 est we consider the scattering parameters on a frequency range of [7 GHz, 9 GHz]
 556 using equidistant sample points, where we restrict ourself to S_{21} and S_{41} for brevity,
 557 however, the results are qualitatively similar for all four scattering parameters. It
 558 can be seen that, the QoIs have a dominant pole at around 8 GHz. This causes the
 559 purely kernel-based interpolations to be inferior compared to the rational approxima-
 560 tions. However, the proposed combination of kernel-based interpolation and rational
 561 approximations leads to satisfactory results, with an accuracy comparable to that
 562 of AAA and VF.

563 The final test case is a vibroacoustic finite element model, taken from [36] and
 564 depicted in Figure 9. A 2D Mindlin plate (vibrating structure D_s) is excited by a
 565 point force and strongly coupled to a 3D acoustic domain (air cavity D_f). Then, the
 566 response at a point in the fluid is evaluated. See [36] for more details on the model.
 567 We consider the frequency response on a frequency interval $\omega \in [4500 \text{ s}^{-1}, 5000 \text{ s}^{-1}]$,
 568 shown in Figure 9 (top, right). The convergence study, given in Figure 9 (bottom),
 569 indicates that the proposed approach usually achieves an accuracy at least comparable
 570 to that of AAA and VF, with at certain points an improvement by about an order
 571 of magnitude can be observed. It can also be seen that the rational mean function
 572 improves the accuracy at the majority of points compared to the pure Szegő kernel-
 573 based interpolation.

574 **5. Conclusion.** We have presented a comprehensive framework for kernel-based
 575 interpolation of complex-valued functions and frequency response functions. In the
 576 complex-valued case, the pseudo-kernel is an additional ingredient, which can be used
 577 to improve the interpolation accuracy. We have introduced the concept of com-
 578 plex/real reproducing kernel Hilbert spaces to reveal the role of the pseudo-kernel
 579 and to establish results on minimum norm interpolation. Furthermore, we have pro-
 580 posed a hybrid method, which complements the kernel-interpolant with a low-order
 581 rational function and a new model selection criterion: this extension is crucial to
 582 account for dominant poles in applications.

583 The capabilities of the rational-kernel method have been illustrated with sev-
 584 eral examples, from circuits to frequency response functions originating from PDE
 585 problems. In all examples the performance was at least comparable, in some cases
 586 improved, compared to AAA and vector fitting on the same set of training data.

587 The kernel method was further linked to complex-valued Gaussian process regres-
 588 sion, which can be used in future work to include noise and adaptive sampling. A
 589 generalization to the multivariate case, where, e.g., uncertain parameters are consid-
 590 ered as well, and comparisons against multivariate AAA [35] or rational Polynomial
 591 Chaos [37], would also be of interest.

592 **Appendix A. Proofs.**

593 **A.1. Proof of Theorem 2.4.** We assume without loss of generality that $\alpha = 0$
 594 in this proof—i.e., we consider the case of the Hardy space $H^2(\Gamma_0)$ on the right half-
 595 plane $\Gamma_0 = \{s \in \mathbb{C} \mid \Re[s] > 0\}$. The general case follows by translation.

596 The fact that $H^2(\Gamma_0)$ is an RKHS is well known. Indeed, recall the one-sided
 597 Paley-Wiener theorem (see, e.g., Chapter 8 of [18]): for all $f \in H^2(\Gamma_0)$, there exists
 598 a unique $\hat{f} \in L^2(\mathbb{R}_+)$ such that

599 (A.1)
$$f(s) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \hat{f}(t) e^{-st} dt, \quad \forall s \in \Gamma_0,$$

600 and the mapping $f \mapsto \hat{f}$ is a surjective isometry: $\|f\|_{H^2(\Gamma_0)} = \|\hat{f}\|_{L^2(\mathbb{R}_+)}$. This proves
 601 that $H^2(\Gamma_0)$ is a Hilbert space, and a simple application of the Cauchy-Schwartz
 602 inequality for $s = x + iy \in \Gamma_0$ yields:

603
$$|f(s)| \leq \frac{1}{2\sqrt{\pi x}} \cdot \|\hat{f}\|_{L^2(\mathbb{R}_+)},$$

604 which proves that the evaluation functionals are continuous on $H^2(\Gamma_0)$.

605 Let us now determine the kernel k of this RKHS. Let $s_0 \in \Gamma_0$ and set $h = k(\cdot, s_0)$.
 606 Then, for any $f \in H^2(\Gamma_0)$, the reproduction property combined with (A.1) yields:

607
$$\langle f, h \rangle_{H^2(\Gamma_0)} = f(s_0) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \hat{f}(t) e^{-s_0 t} dt = \left\langle \hat{f}, \frac{1}{\sqrt{2\pi}} e^{-s_0^*(\cdot)} \right\rangle_{L^2(\mathbb{R}_+)},$$

608 which implies that $\hat{h} = \frac{1}{\sqrt{2\pi}} e^{-s_0^*(\cdot)}$ since $f \mapsto \hat{f}$ is an isometric isomorphism. The
 609 expression of the kernel follows:

610 (A.2)
$$k(s, s_0) = h(s) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \hat{h}(t) e^{-st} dt = \frac{1}{2\pi(s + s_0^*)}.$$

611 It remains to show that k is strictly positive definite. For any $m \geq 1$ and
 612 $s_1, \dots, s_m \in \Gamma_0$, the kernel matrix $K_m = (k(s_i, s_j))_{1 \leq i, j \leq m}$ can be seen as the conju-

613 gate Gram matrix of h_1, \dots, h_m in $L^2(\mathbb{R}_+)$, where $h_j(t) = \frac{1}{\sqrt{2\pi}} e^{-s_j^* t}$, $t \geq 0$. Assume
 614 that s_1, \dots, s_m are distinct. Then it is well known that the complex exponen-
 615 tials $e^{-s_1^*(\cdot)}, \dots, e^{-s_m^*(\cdot)}$ are linearly independent entire functions on \mathbb{C} . It follows,
 616 using the identity theorem, that h_1, \dots, h_m are linearly independent as well. The
 617 kernel matrix K_m is thus invertible and, consequently, positive definite. Therefore k
 618 is strictly positive definite.

619 *Remark A.1.* The expression of the reproducing kernel is also derived in [8, Theo-
 620 rem 2.12] (for the upper half-plane instead of Γ_0) using a different approach involving
 621 the kernel of the Hardy space of the unit disk. Note, however, that the factor 2π
 622 in the denominator of (A.2) is missing in [8, Equation (2.9)]; the discrepancy comes
 623 from a missing factor $\frac{1}{2\pi}$ in the definition of the norm on $H^p(\mathbb{D})$ on page 14.

624 **A.2. Proof of Proposition 2.7.** Take $H = \{\alpha f_0, \alpha \in \mathbb{C}\}$, where $f_0 : \mathbb{X} \rightarrow \mathbb{C}$
 625 is some fixed function, and define a real inner product over H by $\langle \alpha f_0, \beta f_0 \rangle :=$
 626 $\Re \alpha \cdot \Re \beta + 4 \Im \alpha \cdot \Im \beta$. Assuming that $f_0 \not\equiv 0$, the resulting space is complex/real
 627 RKHS of dimension two, spanned by $\{f_0, if_0\}$. (H is also a complex vector space of
 628 dimension 1.)

629 It not possible to embed H as a subspace of a complex Hilbert space $H_{\mathbb{C}}$ with
 630 inner product $\langle \cdot, \cdot \rangle_{\mathbb{C}}$ such that $\langle f, g \rangle = \Re \langle f, g \rangle_{\mathbb{C}}$ for all $f, g \in H$. To see it, note for
 631 instance that $\|f_0\| = 1$ while $\|if_0\| = 2$.

632 **A.3. Proof of Proposition 2.9.** Let $f \in H$, $s_0 \in \mathbb{S}$ and $a_0 \in \{\mathbb{R}, \mathbb{I}\}$. Then

$$633 \quad (\text{A.3}) \quad G_{a_0}(f(s_0)) = (\mathcal{A}f)(s_0, a_0) = \left\langle \mathcal{A}f, \tilde{k}(\cdot, (s_0, a_0)) \right\rangle_{\tilde{H}}$$

$$634 \quad (\text{A.4}) \quad = \left\langle f, \mathcal{A}^{-1} \left(\tilde{k}(\cdot, (s_0, a_0)) \right) \right\rangle_H.$$

636 Taking $a_0 = \mathbb{R}$, we have thus proved that $\Re \circ \delta_{s_0} = \langle \cdot, \varphi_{\mathbb{R}}(\cdot, s_0) \rangle_H$, where

$$637 \quad (\text{A.5}) \quad \varphi_{\mathbb{R}}(\cdot, s_0) = \mathcal{A}^{-1} \left(\tilde{k}(\cdot, (s_0, \mathbb{R})) \right) \in H$$

638 can be computed as follows:

$$639 \quad (\text{A.6}) \quad \Re[\varphi_{\mathbb{R}}(s, s_0)] = (\mathcal{A}[\varphi_{\mathbb{R}}(\cdot, s_0)])(s, \mathbb{R}) = \tilde{k}((s, \mathbb{R}), (s_0, \mathbb{R})) = k_{\mathbb{R}\mathbb{R}}(s, s_0),$$

$$640 \quad (\text{A.7}) \quad \Im[\varphi_{\mathbb{R}}(s, s_0)] = (\mathcal{A}[\varphi_{\mathbb{R}}(\cdot, s_0)])(s, \mathbb{I}) = \tilde{k}((s, \mathbb{I}), (s_0, \mathbb{R})) = k_{\mathbb{I}\mathbb{R}}(s, s_0).$$

642 The expression of $\varphi_{\mathbb{I}}(\cdot, s_0)$ is derived similarly by taking $a_0 = \mathbb{I}$ in (A.4).

643 **A.4. Proof of Proposition 2.11.** In a real or complex RKHS, it is well known
 644 that the partial kernel functions $k(\cdot, s_0)$, $s \in \mathbb{S}$, span a dense subset of the Hilbert
 645 space. Moreover, recall that the bijection \mathcal{A} defined in Section 2.1 is an isometric
 646 isomorphism between H and a real RKHS \tilde{H} on $\tilde{\mathbb{S}} = \mathbb{S} \times \{\mathbb{R}, \mathbb{I}\}$, whose kernel \tilde{k} can
 647 be recovered from k and c by inverting (2.9)–(2.10). The claim then follows from the
 648 observation that any function on $\tilde{\mathbb{S}}$ of the form

$$649 \quad \tilde{g} = \sum_{i=1}^n \alpha_i \tilde{k}(\cdot, (s_i, \mathbb{R})) + \sum_{i=1}^n \beta_i \tilde{k}(\cdot, (s_i, \mathbb{I})),$$

650 where $\alpha_1, \beta_1, \dots, \alpha_n, \beta_n \in \mathbb{R}$, corresponds to the image by \mathcal{A} of

$$\begin{aligned}
 651 \quad g &= \sum_{i=1}^n \alpha_i \mathcal{A}^{-1} \left(\tilde{k}(\cdot, (s_i, \mathbb{R})) \right) + \sum_{i=1}^n \beta_i \mathcal{A}^{-1} \left(\tilde{k}(\cdot, (s_i, \mathbb{I})) \right) \\
 652 \quad &= \sum_{i=1}^n \alpha_i \varphi_{\mathbb{R}}(\cdot, s_i) + \sum_{i=1}^n \beta_i \varphi_{\mathbb{I}}(\cdot, s_i) \\
 653 \quad &= \sum_{i=1}^n \gamma_i k(\cdot, s_i) + \sum_{i=1}^n \gamma_i^* c(\cdot, s_i), \quad \text{with } \gamma_i = \frac{1}{2}(\alpha_i + i\beta_i). \\
 654
 \end{aligned}$$

655 **A.5. Proof of Theorem 2.13.** Assume first that k and c are the complex
 656 kernel and pseudo-kernel associated to a given complex/real RKHS H . Let $\tilde{\xi}$ denote
 657 a zero-mean (e.g., Gaussian) real-valued random process indexed by \mathbb{S} with covariance
 658 function equal to the kernel \tilde{k} of the real RKHS $\tilde{H} = \mathcal{A}H$, and set $\xi = \tilde{\xi}(\cdot, \mathbb{R}) + i\tilde{\xi}(\cdot, \mathbb{I})$.
 659 Then ξ is a complex-valued random process on \mathbb{S} , with covariance function k and
 660 pseudo-covariance function c ; indeed, for all $s, s_0 \in \mathbb{S}$,

$$\begin{aligned}
 661 \quad \mathbb{E}(\xi(s)\xi(s_0)^*) &= \left(\tilde{k}((s, \mathbb{R}), (s_0, \mathbb{R})) + \tilde{k}((s, \mathbb{I}), (s_0, \mathbb{I})) \right) \\
 662 \quad &\quad + i \left(\tilde{k}((s, \mathbb{I}), (s_0, \mathbb{R})) - \tilde{k}((s, \mathbb{R}), (s_0, \mathbb{I})) \right) = k(s, s_0), \\
 663
 \end{aligned}$$

664 and similarly $\mathbb{E}(\xi(s)\xi(s_0)) = c(s, s_0)$. It follows readily that k is Hermitian and
 665 positive definite, and that c is symmetric, which proves i) and ii).

666 Pick $s_1, \dots, s_n \in \mathbb{S}$, and set $K_n = (k(s_i, s_j))_{1 \leq i, j \leq n}$ and $C_n = (c(s_i, s_j))_{1 \leq i, j \leq n}$.
 667 Then K_n and C_n are respectively the covariance and pseudo-covariance matrix of the
 668 random vector $Z = (\xi(s_1), \dots, \xi(s_n))^{\mathbb{T}}$, and thus iv) is precisely the “only if” part
 669 the following result, due to [32].

670 **PROPOSITION A.2.** *Let $n \in \mathbb{N}^*$. Let K be a complex, Hermitian, positive defi-*
 671 *nite matrix of order n , and let C be a complex, symmetric matrix of the same size.*
 672 *Then there exists a complex random vector Z with covariance matrix K and pseudo-*
 673 *covariance matrix C if, and only if, $K^* - C^{\mathbb{H}}K^{-1}C$ is positive semi-definite.*

674 It remains to prove iii): let $u \in \ker K_n$. Then $u^{\mathbb{H}}K_n u = \mathbb{E}(|u^{\mathbb{H}}Z|^2) = 0$, therefore
 675 $u^{\mathbb{H}}Z = 0$ almost surely, and as a consequence:

$$676 \quad C_n^* u = \mathbb{E}(ZZ^{\mathbb{T}})^* u = \mathbb{E}(Z^* Z^{\mathbb{H}} u) = \mathbb{E}(Z^*(u^{\mathbb{H}}Z)^{\mathbb{H}}) = 0.$$

677 This completes the proof of i)–iv).

678 Conversely, assume now that k and c are two functions from $\mathbb{S} \times \mathbb{S}$ to \mathbb{C} , such that
 679 i)–iv) hold. Then it is easy to see that there is a unique function $\tilde{k} : \mathbb{S} \times \{\mathbb{R}, \mathbb{I}\} \rightarrow \mathbb{R}$
 680 such that (2.12)–(2.13) hold, given by

$$\begin{aligned}
 681 \quad k_{\mathbb{R}\mathbb{R}}(s, s_0) &= \frac{1}{2} \Re(k(s, s_0) + c(s, s_0)) \\
 682 \quad k_{\mathbb{I}\mathbb{I}}(s, s_0) &= \frac{1}{2} \Re(k(s, s_0) - c(s, s_0)) \\
 683 \quad k_{\mathbb{I}\mathbb{R}}(s, s_0) &= \frac{1}{2} \Im(k(s, s_0) + c(s, s_0)) = k_{\mathbb{R}\mathbb{I}}(s_0, s). \\
 684
 \end{aligned}$$

685 It remains to prove that \tilde{k} is positive definite. It is easy to see that this is true
 686 if, and only if, the matrices K_n and C_n defined above are the covariance and

687 pseudo-covariance matrices of a complex random vector Z , for any choice of the
 688 points $s_1, \dots, s_n \in \mathbb{S}$. Pick such a set of points, and let r denote the rank of K_n .
 689 Assume without loss of generality that

$$690 \quad (\text{A.8}) \quad K_n = \begin{pmatrix} K_{11} & K_{12} \\ K_{12}^H & K_{22} \end{pmatrix},$$

691 with K_{11} a positive definite $r \times r$ matrix. Then $K_{22} = K_{12}^H K_{11}^{-1} K_{12}$ and

$$692 \quad (\text{A.9}) \quad K_n = M \begin{pmatrix} K_{11} & 0 \\ 0 & 0 \end{pmatrix} M^H, \quad \text{where } M = \begin{pmatrix} I_r & 0 \\ K_{12}^H K_{11}^{-1} & I_{n-r} \end{pmatrix}$$

693 Denote by C_{11} the upper-left $r \times r$ block in C_n . Then it follows from iv) that
 694 $K_{11}^* - C_{11}^H K_{11}^{-1} C_{11}$ is positive semi-definite, and thus by Proposition A.2 there exists a
 695 complex random vector Z_1 of size r with covariance matrix K_{11} and pseudo-covariance
 696 matrix C_{11} . It is then clear from (A.9) that K_n is the covariance matrix of

$$697 \quad Z = M \begin{pmatrix} Z_1 \\ 0 \end{pmatrix}.$$

698 To complete the proof, it remains to observe that C_n is the pseudo-covariance matrix
 699 of Z :

$$700 \quad (\text{A.10}) \quad C_n = M \begin{pmatrix} C_{11} & 0 \\ 0 & 0 \end{pmatrix} M^T = \mathbf{E}(ZZ^T),$$

701 which follows from the facts that C_n is symmetric and that $\ker K_n \subseteq \ker C_n^*$, respec-
 702 tively by ii) and iii).

703 **A.6. Proof of Theorem 2.14.** Using the bijection \mathcal{A} defined in Section 2.1,
 704 the interpolation problem on \mathbb{S} with complex-valued data $(s_1, y_1), \dots, (s_n, y_n)$ can
 705 be reformulated as an interpolation problem on $\tilde{\mathbb{S}} = \mathbb{S} \times \{\mathbb{R}, \mathbb{I}\}$ with real-valued
 706 data $((s_1, \mathbb{R}), \Re(y_1)), ((s_1, \mathbb{I}), \Im(y_1)), \dots, ((s_n, \mathbb{R}), \Re(y_n)), ((s_n, \mathbb{I}), \Im(y_n))$. The claim
 707 then follows from Theorem 2.3 using, as in the proof of Proposition 2.11, the fact that
 708 \mathcal{A} is an isometric isomorphism between H and the real RKHS $\tilde{H} = \mathcal{A}(H)$.

709 **A.7. Proof of Theorem 2.15.** $i) \Rightarrow ii)$. Let H denote a complex/real RKHS
 710 on \mathbb{S} with complex kernel k , such that (2.16) holds. Let c denote the pseudo-covariance
 711 of H . Let $s_0 \in \mathbb{S}$. It follows from Proposition 2.11 that

$$712 \quad f_\gamma = \gamma k(\cdot, s_0) + \gamma^* c(\cdot, s_0)$$

713 is in H for all $\gamma \in \mathbb{C}$. Using (2.16), we see then that

$$714 \quad \begin{aligned} f_\gamma(s^*) &= \gamma k(s^*, s_0) + \gamma^* c(s^*, s_0) \\ 715 &= \gamma c(s, s_0)^* + \gamma^* k(s, s_0)^* = f_\gamma(s)^* \end{aligned}$$

717 holds for all $\gamma \in \mathbb{C}$. This yields in particular that $c(s, s_0) = k(s^*, s_0)^* = k(s_0, s^*)$, and
 718 the claim follows from the symmetry of c :

$$719 \quad c(s, s_0) = c(s_0, s) = k(s, s_0^*).$$

720 Note that we have actually proved a little more than $ii)$: if $i)$ holds, then $ii)$ holds
 721 for the *same* complex/real RKHS H . Since we will now prove that $ii) \Rightarrow iii) \Rightarrow i)$,

722 it follows that the complex/real RKHS with complex kernel k and pseudo-kernel c
 723 defined by (2.17), if it exists, is the only complex/real RKHS with complex kernel k
 724 such that (2.16) holds.

725 *ii*) \Rightarrow *iii*). Let H denote a complex/real RKHS on \mathbb{S} with complex kernel k .
 726 Assume that the pseudo-kernel c satisfies (2.17). Then, for all $s, s_0 \in \mathbb{S}$,

$$727 \quad k(s, s_0^*) = c(s, s_0) = c(s_0, s) = k(s_0, s^*).$$

728 *iii*) \Rightarrow *i*). Let k denote a Hermitian positive definite kernel on \mathbb{S} such that

$$729 \quad (\text{A.11}) \quad \forall s, s_0 \in \mathbb{S}, \quad k(s, s_0^*) = k(s_0, s^*).$$

730 Let $(H_{\mathbb{C}}, \langle \cdot, \cdot \rangle_{\mathbb{C}})$ denote the complex RKHS with kernel k and let $\langle \cdot, \cdot \rangle_{\mathbb{R}} = \Re \langle \cdot, \cdot \rangle_{\mathbb{C}}$.
 731 Then, as observed in Remark 2.6, $(H_{\mathbb{C}}, \langle \cdot, \cdot \rangle_{\mathbb{R}})$ is a complex/real RKHS. The associated
 732 real and imaginary evaluation kernels, which we denote by $\varphi_{\mathbb{R}}^{\diamond}$ and $\varphi_{\mathbb{I}}^{\diamond}$ respectively,
 733 are easily seen to be given by $\varphi_{\mathbb{R}}^{\diamond} = k$ and $\varphi_{\mathbb{I}}^{\diamond} = ik$, and the complex kernel and
 734 pseudo-kernel follow:

$$735 \quad k^{\diamond} = \varphi_{\mathbb{R}}^{\diamond} - i\varphi_{\mathbb{I}}^{\diamond} = 2k \quad \text{and} \quad c^{\diamond} = \varphi_{\mathbb{R}}^{\diamond} + i\varphi_{\mathbb{I}}^{\diamond} = 0.$$

736 Now let H denote the subset of all the functions $f \in H_{\mathbb{C}}$ that satisfy (2.16): H is
 737 clearly a real subspace of $H_{\mathbb{C}}$, and thus $(H, \langle \cdot, \cdot \rangle_{\mathbb{R}})$ is a complex/real RKHS as well.
 738 Moreover, for any $f \in H$,

$$\begin{aligned} 739 \quad \Re f(s) &= \Re \left\{ \frac{1}{2} (f(s) + f(s^*)^*) \right\} \\ 740 &= \frac{1}{2} \left\{ \langle f, \varphi_{\mathbb{R}}^{\diamond}(\cdot, s) \rangle_{\mathbb{R}} + \langle f, \varphi_{\mathbb{R}}^{\diamond}(\cdot, s^*) \rangle_{\mathbb{R}} \right\} \\ 741 &= \left\langle f, \frac{1}{2} (\varphi_{\mathbb{R}}^{\diamond}(\cdot, s) + \varphi_{\mathbb{R}}^{\diamond}(\cdot, s^*)) \right\rangle_{\mathbb{R}}. \\ 742 \end{aligned}$$

743 As a consequence of (A.11), the function $s \mapsto \frac{1}{2} (\varphi_{\mathbb{R}}^{\diamond}(\cdot, s) + \varphi_{\mathbb{R}}^{\diamond}(\cdot, s^*))$ in this inner
 744 product satisfies

$$\begin{aligned} 745 \quad &\frac{1}{2} (\varphi_{\mathbb{R}}^{\diamond}(s_0, s)^* + \varphi_{\mathbb{R}}^{\diamond}(s_0, s^*)^*) \\ 746 &= \frac{1}{2} (\varphi_{\mathbb{R}}^{\diamond}(s, s_0) + \varphi_{\mathbb{R}}^{\diamond}(s^*, s_0)) \\ 747 \end{aligned}$$

748 is an element of H , which proves that the real evaluation functional $\varphi_{\mathbb{R}}$ of $(H, \langle \cdot, \cdot \rangle_{\mathbb{R}})$
 749 is given by

$$750 \quad \varphi_{\mathbb{R}}(s, s_0) = \frac{1}{2} (\varphi_{\mathbb{R}}^{\diamond}(s, s_0) + \varphi_{\mathbb{R}}^{\diamond}(s, s_0^*)) = \frac{1}{2} (k(s, s_0) + k(s, s_0^*)).$$

751 Similarly for the imaginary evaluation functional $\varphi_{\mathbb{I}}$:

$$752 \quad \varphi_{\mathbb{I}}(s, s_0) = \frac{1}{2} (\varphi_{\mathbb{I}}^{\diamond}(s, s_0) - \varphi_{\mathbb{I}}^{\diamond}(s, s_0^*)) = \frac{i}{2} (k(s, s_0) - k(s, s_0^*)).$$

753 Therefore $\varphi_{\mathbb{R}} - i\varphi_{\mathbb{I}} = k$ is the complex kernel of $(H, \langle \cdot, \cdot \rangle_{\mathbb{R}})$, which proves *i*).

754 To prove the remaining assertions, assume that *i-iii*) hold. Let G denote the
 755 closed linear span of $\{k(\cdot, s_0); s_0 \in \mathbb{S}\}$ over \mathbb{R} . Then we have $G + iG = H_{\mathbb{C}}$, and it

756 follows from (A.11) that $G \subseteq H$. Observing that

$$757 \quad iH = \{f \in H_{\mathbb{C}} \mid \forall s \in \mathbb{S}, f(s^*) = -f(s)^*\},$$

758 we conclude that $H \cap iH = \{0\}$, therefore $G = H$ and $H \oplus iH = H_{\mathbb{C}}$.

759 **A.8. Proof of Theorem 2.16.** Observe first that, without loss of generality,
760 we can add m extra data points (s_i, y_i) , for some $m \leq n$, in such way that 1) the
761 points $s_i \in \mathbb{S}$ ($1 \leq i \leq n + m$) are still distinct, and 2) for each i we have $s_j = s_i^*$
762 and $y_j = y_i^*$ for some j .

763 *Existence.* Since k is strictly positive definite, we can find $\alpha_1, \dots, \alpha_{n+m} \in \mathbb{C}$ such
764 that $h = \sum_{i=1}^{n+m} \alpha_i k(\cdot, s_i)$ interpolates the extended data $(s_1, y_1), \dots, (s_{n+m}, y_{n+m})$.
765 This function h belongs to $H_{\mathbb{C}}$ but not in general to H . Set $g(s) = \frac{1}{2}(h(s) + h(s^*)^*)$.
766 Then g clearly satisfies the symmetry condition ($g(s^*) = g(s)^*$ for all $s \in \mathbb{S}$) and still
767 interpolates the extended data $(s_1, y_1), \dots, (s_{n+m}, y_{n+m})$. Moreover, using iii) from
768 Theorem 2.15, we obtain that

$$769 \quad g(s) = \frac{1}{2} \sum_{i=1}^{n+m} (\alpha_i k(s, s_i) + \alpha_i^* k(s, s_i^*)),$$

770 which shows that $g \in H_{\mathbb{C}}$, and thus $g \in H$. Besides, we easily see using (2.17) that:
771 if $s_i = s_i^*$ then

$$772 \quad (\text{A.12}) \quad \frac{1}{2} \left\{ \alpha_i k(s, s_i) + \alpha_i^* k(s, s_i^*) \right\} = \gamma_i k(s, s_i) + \gamma_i^* c(s, s_i)$$

773 with $\gamma_i = \frac{1}{2}\alpha_i$, and if $s_j = s_i^*$ with $i \neq j$ then

$$774 \quad (\text{A.13}) \quad \frac{1}{2} \left\{ (\alpha_i k(s, s_i) + \alpha_i^* k(s, s_i^*)) + (\alpha_j k(s, s_j) + \alpha_j^* k(s, s_j^*)) \right\} \\ = (\gamma_i k(s, s_i) + \gamma_i^* c(s, s_i)) + (\gamma_j k(s, s_j) + \gamma_j^* c(s, s_j))$$

775 with $\gamma_i = \frac{1}{2}(\alpha_i + \alpha_j^*)$ and $\gamma_j = 0$. It follows that g can be rewritten under the
776 form (2.15), using the fact that $\gamma_j = 0$ in (A.13) to get rid of the m extra terms. Thus
777 $\gamma = (\gamma_1, \dots, \gamma_n)^{\text{T}}$ solves (2.14), which proves the “existence” part of the theorem.

778 *Uniqueness.* Let $g \in H$ denote a function of the form (2.15), where the coeffi-
779 cients γ_i are such that (2.14) holds. Using the property that $c(s, s_i) = k(s, s_i^*)$, any
780 such function can be rewritten as $g = \sum_{i=1}^{n+m} \alpha_i k(\cdot, s_i)$. Moreover, since the s_i 's are
781 $n + m$ distinct points in \mathbb{S} and k is strictly positive definite, the coefficients $\alpha_i \in \mathbb{C}$ are
782 uniquely determined by the interpolation conditions: $g(s_i) = y_i$, $1 \leq i \leq n + m$. The
783 first n conditions come directly from (2.14), and the m additional conditions must
784 hold as well by symmetry, since $g \in H$.

785 For each i such that $s_i = s_i^*$, it is easily seen that $\alpha_i = \gamma_i + \gamma_i^*$ is real, and thus the
786 value of γ_i is uniquely determined by α_i and the additional condition that $\gamma_i = \gamma_i^*$.
787 Similarly, if $s_i = s_j^*$ for some $i, j \leq n$, $i \neq j$, then $\alpha_i = \gamma_i + \gamma_j^*$, $\alpha_j = \gamma_i^* + \gamma_j$, and
788 therefore γ_i, γ_j are uniquely determined by α_i, α_j and the condition $\gamma_i = \gamma_j^*$. Finally,
789 if $s_i = s_j^*$ for some $i \leq n$ and $j > n$, then $\alpha_i = \gamma_i$. We have thus proved that there is
790 a unique $\gamma = (\gamma_1, \dots, \gamma_n)^{\text{T}}$, with the property that $\gamma_i = \gamma_j^*$ when $s_i = s_j^*$, such that
791 (2.14) holds.

792 **Acknowledgments.** We thank Christopher Blech, Harikrishnan Sreekumar and
793 Sabine Langer for suggesting acoustic benchmarks and for providing the implementa-
794 tion of the finite element solver used to compute the vibroacoustics data set.

795

REFERENCES

796 [1] A. C. ANTOULAS, S. LEFTERIU, A. C. IONITA, P. BENNER, AND A. COHEN, *A tutorial intro-*
 797 *duction to the loewner framework for model reduction*, in *Model Reduction and Approx-*
 798 *imation: Theory and Algorithms*, vol. 15, SIAM, 2017, pp. 335–376.
 799 [2] L. BARATCHART, M. CARDELLI, AND M. OLIVI, *Identification and rational L^2 approxi-*
 800 *a gradient algorithm*, *Automatica*, 27 (1991), pp. 413–417.
 801 [3] S. BASAK, S. PETIT, J. BECT, AND E. VAZQUEZ, *Numerical issues in maximum likelihood*
 802 *parameter estimation for Gaussian process regression*, arXiv:2101.09747, (2021).
 803 [4] J. BECT, E. VAZQUEZ, ET AL., *STK: a Small (Matlab/Octave) Toolbox for Kriging. Release*
 804 *2.8, 2023*, <https://github.com/stk-kriging/stk/>.
 805 [5] R. BOLOIX-TORTOSA, J. J. MURILLO-FUENTES, F. J. PAYÁN-SOMET, AND F. PÉREZ-CRUZ,
 806 *Complex Gaussian processes for regression*, *IEEE Trans. Neural Netw. Learn. Syst.*, 29
 807 (2018), pp. 5499–5511.
 808 [6] R. BOLOIX-TORTOSA, J. J. MURILLO-FUENTES, I. SANTOS, AND F. PÉREZ-CRUZ, *Widely linear*
 809 *complex-valued kernel methods for regression*, *IEEE Trans. Signal. Process.*, 65 (2017),
 810 pp. 5240–5248.
 811 [7] F. BONIZZONI, F. NOBILE, I. PERUGIA, AND D. PRADOVERA, *Least-squares Padé approximation*
 812 *of parametric and stochastic Helmholtz maps*, *Adv. Comput. Math.*, 46 (2020), pp. 1–28.
 813 [8] J. BONYO, *Reproducing kernels for Hardy and Bergman spaces of the upper half plane*, *Com-*
 814 *munic. Adv. Math. Sci.*, 3 (2020), pp. 13–23.
 815 [9] J. BURBEA AND P. R. MASANI, *Banach and Hilbert spaces of vector-valued functions: their*
 816 *general theory and applications to holomorphy*, Pitman Publishing, 1984.
 817 [10] DASSAULT SYSTÈMES, *CST STUDIO SUITE 2019*, <https://www.cst.com>.
 818 [11] D. DESCHRIJVER, M. MROZOWSKI, T. DHAENE, AND D. DE ZUTTER, *Macromodeling of mul-*
 819 *tiport systems using a fast implementation of the vector fitting method*, *IEEE Microw.*
 820 *Wireless Compon. Lett.*, 18 (2008), pp. 383–385.
 821 [12] T. A. DRISCOLL, N. HALE, AND L. N. TREFETHEN, *Chebfun guide*, 2014.
 822 [13] J. N. FUHG, A. FAU, AND U. NACKENHORST, *State-of-the-art and comparative review of adaptive*
 823 *sampling methods for Kriging*, *Arch. Comput. Methods Eng.*, (2020), pp. 1–59.
 824 [14] M. FUHLÄNDER AND S. SCHÖPS, *A blackbox yield estimation workflow with Gaussian process*
 825 *regression for industrial problems*, *J. Math. Ind.*, 10 (2020), p. 25.
 826 [15] B. GUSTAVSEN, *Improving the pole relocating properties of vector fitting*, *IEEE Trans. Power*
 827 *Deliv.*, 21 (2006), pp. 1587–1592.
 828 [16] B. GUSTAVSEN AND A. SEMLYEN, *Rational approximation of frequency domain responses by*
 829 *vector fitting*, *IEEE Trans. Power Deliv.*, 14 (1999), pp. 1052–1061.
 830 [17] N. HALLEMANS, R. PINTELO, B. JOUKOVSKY, D. PEUMANS, AND J. LATAIRE, *FRF estimation*
 831 *using multiple kernel-based regularisation*, *Automatica*, 136 (2022), p. 110056.
 832 [18] K. HOFFMAN, *Banach spaces of analytic functions*, Prentice Hall, 1962.
 833 [19] M. HORNIKX, M. KALTENBACHER, AND S. MARBURG, *A platform for benchmark cases in com-*
 834 *putational acoustics*, *Acta Acust. united Ac.*, 101 (2015), pp. 811–820.
 835 [20] R. JIN, W. CHEN, AND A. SUDJANTO, *On sequential sampling for global metamodeling in engi-*
 836 *neering design*, in *International Design Engineering Technical Conferences and Computers*
 837 *and Information in Engineering Conference*, vol. 36223, 2002, pp. 539–548.
 838 [21] B. KIM, Y. LEE, AND D.-H. CHOI, *Construction of the radial basis function based on a sequen-*
 839 *tial sampling approach using cross-validation*, *J. Mech. Sci. Technol.*, 23 (2009), pp. 3357–
 840 3365.
 841 [22] S. G. KRANTZ, *Function theory of several complex variables. Second Edition*, AMS Chelsea
 842 Publishing, 2001.
 843 [23] J. LATAIRE AND T. CHEN, *Transfer function and transient estimation by Gaussian process*
 844 *regression in the frequency domain*, *Automatica*, 72 (2016), pp. 217–229.
 845 [24] Z. LIU, D. LESSELIER, B. SUDRET, AND J. WIART, *Surrogate modeling based on resampled*
 846 *polynomial chaos expansions*, *Reliab. Eng. Syst. Saf.*, 202 (2020), p. 107008.
 847 [25] N. MEINSHAUSEN AND P. BÜHLMANN, *Stability selection*, *J. R. Stat. Soc., B: Stat. Methodol.*,
 848 72 (2010), pp. 417–473.
 849 [26] C. A. MICHELLI AND M. PONTIL, *On learning vector-valued functions*, *Neural Comput.*, 17
 850 (2005), pp. 177–204.
 851 [27] K. S. MILLER, *Complex Gaussian processes*, *SIAM Rev.*, 11 (1969), pp. 544–567.
 852 [28] Y. NAKATSUKASA, O. SÈTE, AND L. N. TREFETHEN, *The AAA algorithm for rational approxi-*
 853 *mation*, *SIAM J. Sci. Comput.*, 40 (2018), pp. A1494–A1522.
 854 [29] F. NOBILE AND D. PRADOVERA, *Non-intrusive double-greedy parametric model reduction by*
 855 *interpolation of frequency-domain rational surrogates*, arXiv:2008.10864, (2020).

- 856 [30] V. I. PAULSEN AND M. RAGHUPATHI, *An introduction to the theory of reproducing kernel Hilbert*
857 *spaces*, vol. 152, Cambridge University Press, 2016.
- 858 [31] S. PETIT, J. BECT, P. FELIOT, AND E. VAZQUEZ, *Parameter selection in Gaussian process*
859 *interpolation: an empirical study of selection criteria*. arXiv:2107.06006v4, 2022.
- 860 [32] B. PICINBONO, *Second-order complex random vectors and normal distributions*, IEEE Trans.
861 *Signal Process.*, 44 (1996), pp. 2637–2640.
- 862 [33] B. PICINBONO AND P. CHEVALIER, *Widely linear estimation with complex data*, IEEE Trans.
863 *Signal Process.*, 43 (1995), pp. 2030–2033.
- 864 [34] G. PILLONETTO AND G. DE NICOLAO, *A new kernel-based approach for linear system identifi-*
865 *cation*, *Automatica*, 46 (2010), pp. 81–93.
- 866 [35] A. C. RODRIGUEZ AND S. GUGERCIN, *The p-AAA algorithm for data driven modeling of para-*
867 *metric dynamical systems*, arXiv:2003.06536, (2020).
- 868 [36] U. RÖMER, M. BOLLHÖFER, H. SREEKUMAR, C. BLECH, AND S. LANGER, *An adaptive sparse*
869 *grid rational arnoldi method for uncertainty quantification of dynamical systems in the*
870 *frequency domain*, *Int. J. Numer. Meth.*, (2021), pp. 5487–5511.
- 871 [37] F. SCHNEIDER, I. PAPAIOANNOU, AND G. MÜLLER, *Sparse bayesian learning for complex-valued*
872 *rational approximations*, *International Journal for Numerical Methods in Engineering*, 124
873 (2023), pp. 1721–1747.
- 874 [38] F. TISSEUR AND K. MEERBERGEN, *The quadratic eigenvalue problem*, *SIAM Rev.*, 43 (2001),
875 pp. 235–286.
- 876 [39] F. TREVISO, R. TRINCHERO, AND F. G. CANAVERO, *Multiple delay identification in long inter-*
877 *connects via LS-SVM regression*, *IEEE Access*, 9 (2021), pp. 39028–39042.
- 878 [40] H. ZIEGELWANGER AND P. REITER, *The PAC-MAN model: Benchmark case for linear acoustics*
879 *in computational physics*, *J. Comput. Phys.*, 346 (2017), pp. 152–171.

880
881
882
883

SUPPLEMENTARY MATERIAL

884 **SM1. Introduction.** This supplementary material is structured in the following
885 way. We first present additional material on the new method presented in the paper:
886 Section SM2.1 provides details on a non-intrusive implementation, and Section SM2.2
887 contains specifics on parameter optimization. We then collect material relevant to
888 the examples. In particular, a partial fraction representation of the circuit model in
889 Section SM3.1 and an additional numerical example in Section SM3.2. The algorithm
890 applied to this example, a spiral antenna, yields results which are comparable to the
891 pacman model shown in the main paper. Finally, some theoretical results regarding
892 circular complex/real RKHS are collected in Section SM4.1.

893 **SM2. Implementation details.**
894 **SM2.1. Non-intrusive implementation.**

895 **SM2.1.1. Zero-mean case.** The main idea is to construct an isomorphic real-
896 valued GP $\tilde{g}(\tilde{\mathbf{x}}) \sim \text{GP}(0, \tilde{k})$ on an *augmented input space* $\tilde{\mathbf{x}} \in (\mathbb{R}^n \times \{0, 1\})$, s.t.,

897 (SM2.1)
$$\begin{aligned} \tilde{g}(\begin{bmatrix} \mathbf{x} & 0 \end{bmatrix}) &= \text{Re}[g(\mathbf{x})], \\ \tilde{g}(\begin{bmatrix} \mathbf{x} & 1 \end{bmatrix}) &= \text{Im}[g(\mathbf{x})]. \end{aligned}$$

898 The augmented training data $(\tilde{\mathbf{x}}, \tilde{y}) \in (\mathbb{R}^n \times \{0, 1\}) \times \mathbb{R}$ is for each observation
899 $(\mathbf{x}^{(i)}, y^{(i)}) \in \mathbb{R}^n \times \mathbb{C}$ obtained as:

900
$$\begin{aligned} \tilde{\mathbf{x}}^{(i,1)} &= \begin{bmatrix} \mathbf{x}^{(i)} & 0 \end{bmatrix}, \quad \tilde{y}^{(i,1)} = \text{Re}[y^{(i)}], \\ \tilde{\mathbf{x}}^{(i,2)} &= \begin{bmatrix} \mathbf{x}^{(i)} & 1 \end{bmatrix}, \quad \tilde{y}^{(i,2)} = \text{Im}[y^{(i)}]. \end{aligned}$$

903 The *new* covariance function \tilde{k} can be derived by enforcing (SM2.1),

904
$$\tilde{k}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}') = \tilde{k}(\begin{bmatrix} \mathbf{x} & j \end{bmatrix}, \begin{bmatrix} \mathbf{x}' & j' \end{bmatrix}) = \begin{cases} \frac{1}{2} \text{Re}[k(\mathbf{x}, \mathbf{x}') + c(\mathbf{x}, \mathbf{x}')] & j = j' = 0 \\ \frac{1}{2} \text{Re}[k(\mathbf{x}, \mathbf{x}') - c(\mathbf{x}, \mathbf{x}')] & j = j' = 1 \\ \frac{1}{2} \text{Im}[-k(\mathbf{x}, \mathbf{x}') + c(\mathbf{x}, \mathbf{x}')] & j = 0, j' = 1 \\ \frac{1}{2} \text{Im}[k(\mathbf{x}, \mathbf{x}') + c(\mathbf{x}, \mathbf{x}')] & j = 1, j' = 0 \end{cases},$$

905

906 Note that this approach requires to define the modified covariance function \tilde{k} , however,
907 no (other) internal functions of existing GP implementations need to be able to cope
908 with complex numbers, which is why we refer to the implementation as *non-intrusive*.

909 **SM2.1.2. Linear model in the mean function.** Consider now the superpo-
910 sition

911
$$\underline{g}(\mathbf{x}) = g(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T \mathbf{b}$$

912 of a mean-free (real/)complex Gaussian Process $g(\mathbf{x}) \sim \text{CGP}(0, k, c)$ and a complex
913 linear model, where $h(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{C}^m$ denote explicit basis functions and $\mathbf{b} \in \mathbb{C}^m$ the
914 corresponding coefficients. Define the augmented process

915
$$\underline{\tilde{g}}(\tilde{\mathbf{x}}) = \tilde{g}(\tilde{\mathbf{x}}) + \tilde{\mathbf{h}}(\tilde{\mathbf{x}})^T \tilde{\mathbf{b}},$$

916 where $\tilde{g}(\tilde{\mathbf{x}}) \sim \text{GP}(0, \tilde{k})$, $\tilde{\mathbf{b}} \in \mathbb{R}^{2m}$, $\tilde{\mathbf{h}}(\tilde{\mathbf{x}}) : \mathbb{R}^n \times \{0, 1\} \rightarrow \mathbb{R}^{2m}$ and we require, similarly
 917 as in the last subsection, that

$$\begin{aligned} 918 \quad \tilde{g}([\mathbf{x} \ 0]) &= \text{Re}[g(\mathbf{x})], \\ 919 \quad \tilde{g}([\mathbf{x} \ 1]) &= \text{Im}[g(\mathbf{x})]. \end{aligned}$$

921 Incorporating (SM2.1), we can conclude that

$$\begin{aligned} 922 \quad \tilde{\mathbf{h}}([\mathbf{x} \ 0])^T \tilde{\mathbf{b}} &= \text{Re}[\mathbf{h}(\mathbf{x})^T \mathbf{b}] = \text{Re}[\mathbf{h}(\mathbf{x})]^T \text{Re}[\mathbf{b}] - \text{Im}[\mathbf{h}(\mathbf{x})]^T \text{Im}[\mathbf{b}] \\ 923 \quad \tilde{\mathbf{h}}([\mathbf{x} \ 1])^T \tilde{\mathbf{b}} &= \text{Im}[\mathbf{h}(\mathbf{x})^T \mathbf{b}] = \text{Im}[\mathbf{h}(\mathbf{x})]^T \text{Re}[\mathbf{b}] + \text{Re}[\mathbf{h}(\mathbf{x})]^T \text{Im}[\mathbf{b}]. \end{aligned}$$

925 needs to be fulfilled. This can be achieved by setting

$$926 \quad \tilde{\mathbf{h}}(\tilde{\mathbf{x}}) = \tilde{\mathbf{h}}([\mathbf{x} \ j]) = \begin{bmatrix} + \text{Re}[\mathbf{h}(\mathbf{x})] & j = 0 \\ + \text{Im}[\mathbf{h}(\mathbf{x})] & j = 1 \end{bmatrix} \begin{bmatrix} - \text{Im}[\mathbf{b}] & j = 0 \\ + \text{Re}[\mathbf{b}] & j = 1 \end{bmatrix}^T,$$

927 which leads to the coefficients vector

$$928 \quad \tilde{\mathbf{b}} = \begin{bmatrix} \text{Re}[\mathbf{b}] \\ \text{Im}[\mathbf{b}] \end{bmatrix}.$$

929 SM2.2. Details on parameter optimization.

930 **SM2.2.1. Penalized log-likelihood criterion.** We select the parameters α
 931 and σ^2 of the the Szegö kernel $\sigma^2 k_\alpha$ (cf. Section 3 in the article), together with the
 932 dominant poles $\mathbf{p} = (p_1, \dots, p_K)$ in the case of the hybrid method, by maximizing a
 933 penalized log-likelihood criterion:

$$934 \quad (\text{SM2.2}) \quad J(\alpha, \sigma^2, \mathbf{p}) = \max_{\mathbf{r} \in \mathbb{C}^K} \ln(p(\mathbf{y}|\alpha, \sigma^2, \mathbf{p}, \mathbf{r})) + \ln(\rho(\alpha)),$$

935 where the first term is the log-likelihood of the model, maximized (profiled) analyti-
 936 cally with respect to the residues \mathbf{r} of the rational mean function model m , and the
 937 second term is a penalty term, designed to pull α away from 0.

938 More precisely, we take for $\rho(\alpha)$ the probability density function (pdf) of a log-
 939 normal random variable with parameters μ_α (to be specified) and $\sigma_\alpha = 3$; in other
 940 words, we use a “vague” prior distribution on α , such that $\log(\alpha)$ is Gaussian with
 941 mean μ_α and variance σ_α^2 . The parameter μ_α is chosen in such a way that the log-
 942 normal density for α has its mode at $|\Omega| = \omega_{\max} - \omega_{\min}$, or equivalently that the prior
 943 density for $\alpha/|\Omega|$ has its mode at 1. Using that the mode of the lognormal density is
 944 at $e^{\mu_\alpha - \sigma_\alpha^2}$, we deduce that $\mu_\alpha = \sigma_\alpha^2 + \ln|\Omega|$. The resulting pdf

$$945 \quad (\text{SM2.3}) \quad \rho(\alpha) = \frac{1}{\alpha \sigma_\alpha \sqrt{2\pi}} \exp\left(-\frac{(\ln(\alpha) - \mu_\alpha)^2}{2\sigma_\alpha^2}\right)$$

946 is shown in Figure SM1 for $|\Omega| = 1$. It can be seen that the chosen parameters allow
 947 for the choice of α within a range of several orders of magnitude.

948 The penalized log-likelihood criterion (SM2.2) is maximized numerically using
 949 bound-constrained gradient-based optimization—more precisely, interior point algo-
 950 rithm available from Matlab’s `fmincon` function—with a multistart procedure. Details
 951 about the bounds for the search domain and the initial points for the local search are
 952 provided in Sections SM2.2.2–SM2.2.3.

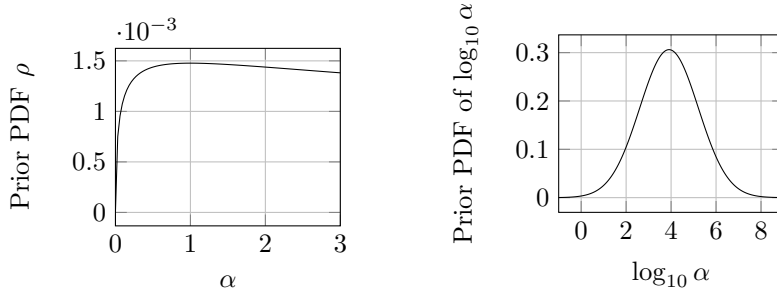


Fig. SM1: Log-normal prior on hyper-parameter α for $|\Omega| = 1$. Left: log-normal prior density of α . Note that the mode of the density is indeed at $\alpha = |\Omega| = 1$. Right: prior density of $\log_{10}(\alpha)$. This is a Gaussian density with mean $\mu_\alpha/\ln(10) \approx 3.91$ and standard deviation $\sigma_\alpha/\ln(10) \approx 1.30$.

953 *Remark SM2.1.* This parameter selection procedure can be considered as *maxi-*
 954 *mum a posteriori* estimate in the Bayesian sense. Indeed, the penalized log-likelihood
 955 criterion (SM2.2) can be seen as the log-posterior density, up to a constant, assuming
 956 a lognormal prior for α and an improper uniform prior for all the other parameters.

957 *Remark SM2.2.* Even when the complex kernel k is strictly positive definite, the
 958 distribution of data under the GP model does not always admit a probability density
 959 function with respect to Lebesgue’s measure on \mathbb{R}^{2n} (cf. related discussion regarding
 960 the strict positive definiteness of \hat{k} in Section 2.2 of the article). When this happens, a
 961 suitable reference measure has to be used in order to define the likelihood function. For
 962 instance, when the pseudo-kernel $c(s, s_0) = k(s, s_0^*)$ is used to enforce the symmetry
 963 condition, the value at $\omega = 0$ must be real, which yields a degenerate distribution if
 964 the response is evaluated at $\omega = 0$: the solution is simply to remove the imaginary
 965 part of the response at this point from the vector of observed variables. See Section 2
 966 of [23] for related considerations.

967 **SM2.2.2. Bounds for the search domain.** We optimize with respect to the
 968 transformed kernel parameters

$$969 \theta_1 = \ln\left(\frac{\sigma^2}{2\pi}\right),$$

$$970 \theta_2 = \alpha,$$

972 within the optimization bounds

$$973 -15 \leq \theta_1 \leq 15,$$

$$974 10^{-6}|\Omega| \leq \theta_2 \leq |\Omega|,$$

976 where $|\Omega| = \omega_{\max} - \omega_{\min} = \max\{\omega_i, 1 \leq i \leq n\} - \min\{\omega_i, 1 \leq i \leq n\}$.

977 For the hybrid model, the poles are optimized simultaneously with the kernel
 978 hyper-parameters, within the bounds

$$979 -|\Omega| \leq \Re(p_i) \leq -10^{-6}|\Omega|,$$

$$980 \max\left\{10^{-6}|\Omega|, \omega_{\min} - \frac{|\Omega|}{3}\right\} \leq \Im(p_i) \leq \omega_{\max} + \frac{|\Omega|}{3},$$

981

982 where we set the maximum number of poles pairs to $K_{\max} = \min\{5, \lfloor n/4 \rfloor\}$. The
 983 bounds are enlarged, if needed, in such a way that all the poles in the starting point
 984 of the optimization are contained within them.

985 **SM2.2.3. Starting point(s).** We use a multistart procedure to optimize the
 986 kernel parameters α and σ^2 . More precisely, we start $N_{\text{ms}} = 20$ local optimizations,
 987 with the initial value of α uniformly distributed between $10^{-6}|\Omega|$ and $|\Omega|$. For a given
 988 value of α , and a given choice of poles in the case of the hybrid algorithm, the GLS
 989 (generalized least squares) estimate is used as a starting point for σ^2 .

990 For the poles in the hybrid algorithm, we start from equidistant poles p_1, \dots, p_K
 991 close to the frequency axis:

$$992 \quad p_k = -\delta_{\Re} |\Omega| + i \left(\omega_{\min} + \left(k - \frac{1}{2} \right) \delta_{\Im} \right), \quad 1 \leq k \leq K,$$

993 where $\delta_{\Re} = 10^{-3}$ (weak attenuation) and $\delta_{\Im} = |\Omega|/K_{\max}$. The kernel parameters
 994 are initialized as described previously (with the GLS estimate for σ^2 and a multi-start
 995 procedure for α).

996 **SM3. Complements for the results section.**

997 **SM3.1. Partial fraction representation of RLC circuit.** The residues c_i, c_i^*
 998 and poles a_i, a_i^* of the partial fraction representation of the electric circuit admittance

$$999 \quad (\text{SM3.1}) \quad Y(s) = \sum_{i=1}^N \frac{c_i}{s - a_i} + \frac{c_i^*}{s - a_i^*},$$

1000 are given as

$$1001 \quad (\text{SM3.2}) \quad a_i = \frac{-R_i}{2L_i} + i \sqrt{\frac{1}{L_i C_i} - \left(\frac{R_i}{2L_i} \right)^2},$$

$$1002 \quad (\text{SM3.3}) \quad c_i = \frac{a_i}{L(a_i - a_i^*)} = \frac{\sqrt{\frac{1}{L_i C_i} - \left(\frac{R_i}{2L_i} \right)^2} + \frac{R_i}{2L_i} i}{2L_i \sqrt{\frac{1}{L_i C_i} - \left(\frac{R_i}{2L_i} \right)^2}},$$

1003

1004 where we assumed an underdamped system, i.e.

$$1005 \quad (\text{SM3.4}) \quad \frac{R_i}{2} \sqrt{\frac{C_i}{L_i}} < 1,$$

1006 which implies that the argument of the square roots is positive.

1007 **SM3.2. Additional numerical example.** The model is a spiral antenna, depicted in Figure SM2, where we consider the reflection coefficient S_{11} on a frequency range of [4 GHz, 6 GHz] as quantity of interest. The data sets are obtained using the boundary element method in CST Microwave Studio [10]. The results are qualitatively the same as for the PAC-MAN model, see Fig. SM2 (bottom).

1012 **SM4. Complements for the theoretical section.**

1013 **SM4.1. Circular complex/real RKHSs.**

1014 **DEFINITION SM4.1.** *We say that a complex/real RKHS is circular if it has a*
 1015 *vanishing pseudo-kernel.*

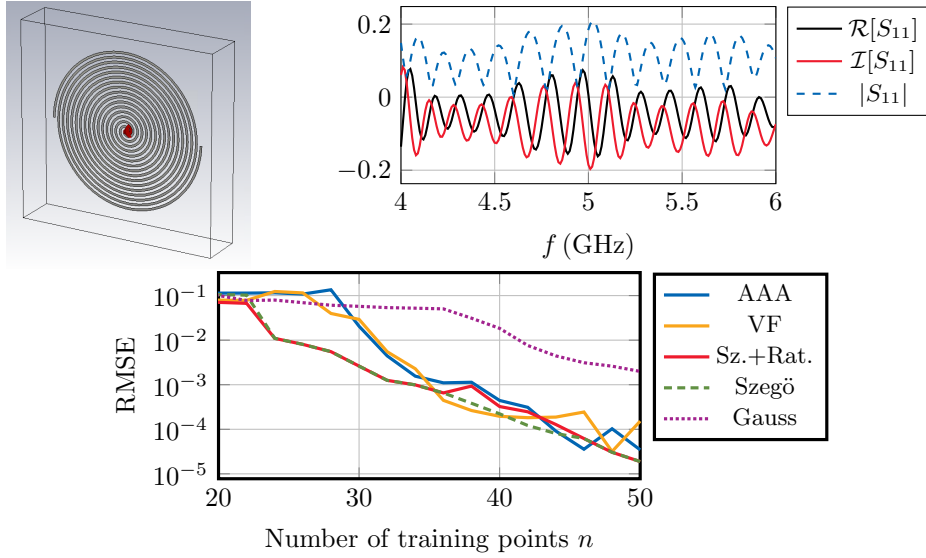


Fig. SM2: Top left: Spiral antenna model, taken from CST Microwave Studio [10]. Top right: Complex frequency response function S_{11} . Bottom: Convergence study w.r.t. the number of training points.

1016 The terms “proper” or “strictly complex” are also sometimes used instead of
 1017 “circular”, in the statistics and signal processing literature, for the case where the
 1018 pseudo-covariance of a complex Gaussian random vector or function vanishes (see,
 1019 e.g., [32, 5]).

1020 **THEOREM SM4.2.** *Let $H_{\mathbb{C}}$ denote a complex RKHS with kernel k_0 , and let H*
 1021 *denote the complex/real RKHS obtained by considering $H_{\mathbb{C}}$ as a real vector space,*
 1022 *endowed with the inner product: $\langle f, g \rangle \mapsto \Re(\langle f, g \rangle_{\mathbb{C}})$. Then H is the circular com-*
 1023 *plex/real RKHS with complex kernel $k = 2k_0$.*

1024 Since a complex/real RKHS is uniquely characterized by its (k, c) pair, the con-
 1025 verse holds as well: given a circular complex/real RKHS H with complex kernel k ,
 1026 there is a unique complex RKHS $H_{\mathbb{C}}$ (namely, the complex RKHS with kernel $k_0 = \frac{1}{2}k$)
 1027 such that H is obtained from $H_{\mathbb{C}}$ as in Theorem SM4.2.

1028 *Proof.* The main idea is already included in the proof of Theorem 2.14 of the
 1029 article, but we give here a slightly more detailed version. Let $\varphi_{\mathbb{R}}$ and $\varphi_{\mathbb{I}}$ denote the
 1030 real and imaginary evaluation kernels of H . Then, for all $f \in H$ and $s \in \mathbb{S}$,

$$1031 \quad \langle f, k_0(\cdot, s) \rangle = \Re(\langle f, k_0(\cdot, s) \rangle) = \Re(f(s))$$

1032 and

$$1033 \quad \begin{aligned} \langle f, ik_0(\cdot, s) \rangle &= \Re(\langle f, ik_0(\cdot, s) \rangle) = \Re(-i \langle f, k_0(\cdot, s) \rangle) \\ 1034 &= \Re(-i f(s)) = \Im(f(s)), \end{aligned}$$

1036 which proves that $\varphi_{\mathbb{R}} = k_0$ and $\varphi_{\mathbb{I}} = ik_0$. The complex kernel and pseudo-kernel of H

1037 are thus given by

$$\begin{aligned}
 1038 \quad k &= \varphi_{\mathbb{R}} - i\varphi_{\mathbb{I}} = 2k_0, \\
 1039 \quad c &= \varphi_{\mathbb{R}} + i\varphi_{\mathbb{I}} = 0. \quad \square
 \end{aligned}$$

1041 **SM4.2. A relation between the Szegő and rational quadratic kernels.**

1042 Consider the Szegő kernel on $H^2(\Gamma_\alpha)$:

$$\begin{aligned}
 1043 \quad k_\alpha(s, s_0) &= \frac{1}{2\pi(2\alpha + s + s_0^*)} \\
 1044 \quad &= \frac{1}{2\pi} \frac{(2\alpha + x + x_0) - i(y - y_0)}{(2\alpha + x + x_0)^2 + (y - y_0)^2}, \\
 1045
 \end{aligned}$$

1046 where $s = x + iy$, $s_0 = x_0 + iy_0 \in \Gamma_\alpha$. In the circular case, the corresponding kernels
 1047 for the real and imaginary parts are given by:

$$1048 \quad k_{\mathbb{R}}(s, s_0) = k_{\mathbb{I}}(s, s_0) = \frac{1}{4\pi} \frac{2\alpha + x + x_0}{(2\alpha + x + x_0)^2 + (y - y_0)^2},$$

1049 For a fixed value of $x = x_0 > -\alpha$, this is of the form

$$1050 \quad (y, y_0) \mapsto \frac{1}{4\pi} \frac{A}{A^2 + (y - y_0)^2}, \quad \text{with } A = 2\alpha + x + x_0 > 0,$$

1051 which is a special case of the so-called *rational quadratic* kernel (see, e.g., [43, 44]),
 1052 also called *generalized inverse multiquadric* kernel (see, e.g., [42, 41]).

1053

ADDITIONAL REFERENCES

1054 [41] M. BOZZINI, M. ROSSINI, AND R. SCHABACK, *Generalized Whittle–Matérn and polyharmonic*
 1055 *kernels*, *Advances in Computational Mathematics*, 39 (2013), pp. 129–141.
 1056 [42] X.-G. HU, T.-S. HO, AND H. RABITZ, *The collocation method based on a generalized inverse*
 1057 *multiquadric basis for bound-state problems*, *Computer physics communications*, 113 (1998),
 1058 pp. 168–179.
 1059 [43] C. E. RASMUSSEN AND C. K. I. WILLIAMS, *Gaussian Processes for Machine Learning*, MIT
 1060 Press, 2006.
 1061 [44] P. SOLLICH AND C. WILLIAMS, *Using the equivalent kernel to understand Gaussian process*
 1062 *regression*, *Advances in Neural Information Processing Systems*, 17 (2004).