MIT Open Access Articles

Approximation of Parametric Derivatives by the Empirical Interpolation Method

Massachusetts Institute of Technology

Approximation of Parametric Derivatives by the Empirical Interpolation Method[*]

Jens L. Eftang[†]   Martin A. Grepl[‡]   Anthony T. Patera[§]

Einar M. Rønquist[¶]

Abstract

We introduce a general *a priori* convergence result for the approximation of parametric derivatives of parametrized functions. We consider the best approximations to parametric derivatives in a sequence of approximation spaces generated by a general approximation scheme, and we show that these approximations are convergent provided that the best approximation to the function itself is convergent. We also provide estimates for the convergence rates. We present numerical results with spaces generated by a particular approximation scheme — the Empirical Interpolation Method — to confirm the validity of the general theory.

Communicated by Ron DeVore

Keywords: function approximation; a priori convergence; Empirical Interpolation Method; parametric derivatives; sensitivity derivatives

AMS (MSC 2010) subject classification: G5G99, 65D05, 41A05, 41A25

1   Introduction

In contexts such as product design, shape optimization, and parameter estimation it is crucial to understand the behavior of a given physical system as a function of parameters that describe the system in terms of for example materials, shapes, or operation conditions. Typically, the goal is to minimize a

parameter dependent cost functional related to certain quantities derived from the state of the system. To this end an automatic optimization algorithm may be employed. Such algorithms typically necessitate calculation of derivatives of the cost functional with respect to the parameters (parametric or sensitivity derivatives). This calculation may be performed directly on the cost functional (by finite differences), or through parametric derivatives of the state [3].

The parameter dependent function that describes the state of a parametrized physical system is typically defined implicitly as the solution to a parametrized partial or ordinary differential equation. Here, we restrict attention to parameter dependent functions that are defined explicitly, i.e., functions that in addition to spatial variables have an explicit dependence on one or several scalar parameters.

In this paper, we develop a new *a priori* convergence theory for the approximation of parametric derivatives of parametrized functions. We consider the best approximations to parametric derivatives in a sequence of approximation spaces generated by a general approximation scheme, and we show that these approximations are convergent provided that the best approximation to the function itself is convergent. We also provide estimates for the convergence rates. The main limitations of the theory are related to regularity assumptions in space and parameter on the parametrized function, and on the particular norms that may be considered. However in principle our new theory also applies to cases in which the parametrized function is defined implicitly as the solution to a differential equation.

As a particular space-generating scheme we consider the Empirical Interpolation Method (EIM), introduced in [1, 8]. The EIM is an interpolation method developed specifically for the approximation of parametrized functions.[1] The new convergence theory in this paper is originally developed with the EIM in mind; in particular, we may relate the error in the EIM approximation to the error in the best approximation through the EIM Lebesgue constant. However, our theoretical results also apply to rather general approximation schemes *other* than the EIM; in principle, we may consider both projection-based and interpolation-based approximation.

The results in this paper have several useful implications. First — not exclusive to the EIM — if we consider an approximation scheme for evaluation of an objective function subject to optimization with respect to a set of parameters, our theory suggests that we may accurately compute the parametric Jacobian and Hessian matrices without expensive generation of additional approximation spaces. Second, the rigorous *a posteriori* bounds for the error in the EIM approximation recently introduced in [6] depend on the error in the EIM approximation of parametric derivatives at a finite number of points in the parameter domain; smaller errors for these EIM derivative approximations

---

[1]In particular, the EIM serves to construct parametrically affine approximations of parameter dependent non-affine or non-linear differential operators within the Reduced Basis (RB) framework for parametric reduced order modelling of partial differential equations [14]. An affine representation (or approximation) of the operator allows an efficient "offline-online" computational decoupling, which in turn is a crucial ingredient in the RB computational framework. We refer to [7, 8] for the application of the EIM for RB approximations.

imply sharper EIM error bounds.

The remainder of the paper is organized as follows. First, in Section 2 we introduce necessary notation and recall some results from polynomial approximation theory. Next, in Section 3, we present the new general *a priori* convergence result. Then, in Section 4 we review the EIM and apply the new convergence theory in this particular context. Subsequently, in Section 5, we validate the theory through two numerical examples where the functions under consideration have different parametric regularity. Finally, in Section 6, we provide some concluding remarks.

2 Preliminaries

2.1 Notation

We denote by $\Omega \subset \mathbb{R}^d$ the spatial domain ($d = 1, 2, 3$); a particular point $x \in \Omega$ shall be denoted by $x = (x_{(1)}, \ldots, x_{(d)})$. We denote by $\mathcal{D} = [-1, 1]^P \subset \mathbb{R}^P$ the parameter domain ($P \geq 1$); a particular parameter value $\mu \in \mathcal{D}$ shall be denoted by $\mu = (\mu_{(1)}, \ldots, \mu_{(P)})$. We choose $\mathcal{D} = [-1, 1]^P$ for the sake of simplicity in our theoretical arguments; our results remain valid for any parameter domain that maps to $[-1, 1]^P$ through an affine transformation.

We introduce a parametrized function $\mathcal{F} : \Omega \times \mathcal{D} \to \mathbb{R}$ for which we assume $\mathcal{F}(\cdot; \mu) \in L^\infty(\Omega)$ for all $\mu \in \mathcal{D}$; here, $L^\infty(\Omega) = \{v : \operatorname{ess\,sup}_{x \in \Omega} |v(x)| < \infty\}$. We then introduce a multi-index of dimension P,

$$\beta = (\beta_1, \ldots, \beta_P), \tag{1}$$

where the entries β_i, $1 \leq i \leq P$, are non-negative integers. We define for any multi-index β the parametric derivatives of $\mathcal{F}$,

$$\mathcal{F}^{(\beta)} = \frac{\partial^{|\beta|} \mathcal{F}}{\partial \mu_{(1)}^{\beta_1} \cdots \partial \mu_{(P)}^{\beta_P}}, \tag{2}$$

where

$$|\beta| = \sum_{i=1}^P \beta_i \tag{3}$$

is the length of β and hence the differential order. Given the parameter domain dimension P, we denote the set of all distinct multi-indices β of length p by $\mathcal{M}_p$.

For purposes of our theoretical arguments later we must require, for all $x \in \Omega$, that $\mathcal{F}(x; \cdot) \in C^1(\mathcal{D})$ and that $\sup_{\mu \in \mathcal{D}} |\mathcal{F}^{(\beta)}(x; \mu)| < \infty$ for all $\beta \in \mathcal{M}_2$. Here $C^s(\mathcal{D})$ denotes the space of functions with continuous order s parametric derivatives over $\mathcal{D}$.

Also for purposes of our theoretical arguments we shall write $\mathcal{D}$ as the tensor product $\mathcal{D} = \mathcal{D}_{(1)} \times \cdots \times \mathcal{D}_{(P)}$, where $\mathcal{D}_{(i)} = [-1, 1]$, $1 \leq i \leq P$. We shall further

consider any particular parameter dimension $\mathcal{S} \equiv \mathcal{D}_j$, $1 \leq j \leq P$. In this case we fix the $P-1$ parameter values $\mu_{(i)} \in \mathcal{D}_{(i)}$, $1 \leq i \leq P$, $i \neq j$, and we introduce for simplicity of notation the function $\mathcal{J}_{\beta,j} : \Omega \times \mathcal{S} \to \mathbb{R}$ defined for $x \in \Omega$ and $y \in \mathcal{S}$ by

$$\mathcal{J}_{\beta,j}(x;y) \equiv \mathcal{F}^{(\beta)}\big(x;(\mu_{(1)},\ldots,\mu_{(j-1)},y,\mu_{(j+1)},\ldots,\mu_{(P)})\big). \tag{4}$$

2.2 Polynomial Interpolation

In this section we first describe a general interpolation framework for which we state three hypotheses; these hypotheses are key ingredients in the proof of our new convergence theory in Section 3. We then recall some results from polynomial approximation theory that confirm the hypotheses under different regularity conditions.

Let $\Gamma = [-1, 1]$, and let $\mathcal{H}$ denote a family of functions such that each $h \in \mathcal{H}$ is a function $\Gamma \to \mathbb{R}$ with suitable regularity. We introduce $N + 1$ distinct interpolation nodes $y_{N,i} \in \Gamma$, $0 \leq i \leq N$, and $N + 1$ characteristic functions $\chi_{N,i}$, $0 \leq i \leq N$, that satisfy $\chi_{N,i}(y_{N,j}) = \delta_{i,j}$, $0 \leq i,j \leq N$; here, $\delta_{i,j}$ is the Kronecker delta symbol. We finally introduce an interpolation operator I_N defined for any function $h \in \mathcal{H}$ by $I_N h = \sum_{i=0}^{N} h(y_{N,i})\chi_{N,i}$. We may now formally state our three hypotheses.

Hypothesis 1. *For all $h \in \mathcal{H}$ the error in the derivative of the interpolant $I_N h$ satisfies*

$$|h'(y) - (I_N h)'(y)| \leq G(N), \quad \forall y \in \Gamma, \tag{5}$$

where the function $G : \mathbb{N} \to (0,\infty)$ is independent of h and satisfies $G(N) \to 0$ as $N \to \infty$.

Hypothesis 2. *The characteristic functions $\chi_{N,i}$, $0 \leq i \leq N$, satisfy*

$$\sum_{i=0}^{N} |\chi'_{N,i}(y)| \leq D(N), \quad \forall y \in \Gamma, \tag{6}$$

where the function $D : \mathbb{N} \to (0,\infty)$ satisfies $D(N) \to \infty$ as $N \to \infty$.

Hypothesis 3. *Let $\epsilon \in \mathbb{R}^+$. As $\epsilon \to 0$ the solution $N_{\mathrm{bal}} = N_{\mathrm{bal}}(\epsilon) > 0$ to the equation*

$$G(N_{\mathrm{bal}}) = D(N_{\mathrm{bal}})\epsilon \tag{7}$$

satisfies

$$H(\epsilon) \equiv \epsilon D\big(N_{\mathrm{bal}}(\epsilon)\big) \to 0. \tag{8}$$

We next consider several interpolation schemes and in each case confirm the corresponding instantiations of our hypotheses under suitable regularity

conditions on the functions in the family $\mathcal{H}$. First, we assume for all $h \in \mathcal{H}$ that $h \in C^1(\Gamma)$ and furthermore that $\sup_{y \in \Gamma} |h''(y)| < \infty$; we then consider piecewise linear interpolation over equidistant interpolation nodes $y_{N,i} = (2i/N - 1) \in \Gamma$, $0 \le i \le N$. In this case the characteristic functions $\chi_{N,i}$ are continuous and piecewise linear "hat functions" with support only on the interval $[y_{N,0}, y_{N,1}]$ for $i = 0$, only on the interval $[y_{N,i-1}, y_{N,i+1}]$ for $1 \le i \le N - 1$, and only on the interval $[y_{N,N-1}, y_{N,N}]$ for $i = N$. For piecewise linear interpolation Hypothesis 1 and Hypothesis 2 obtain for

$$G(N) = c_{\mathcal{H}}^{\mathrm{lin}} N^{-1}, \tag{9}$$

$$D(N) = N, \tag{10}$$

respectively, where $c_{\mathcal{H}}^{\mathrm{lin}} \equiv 2 \sup_{h \in \mathcal{H}} \|h''\|_{L^\infty(\Gamma)}$. In this case (6) in Hypothesis 2 obtains with equality. We include the proofs in Appendix A.1. It is straightforward to demonstrate Hypothesis 3: we note that in this case (7) has the solution

$$N_{\mathrm{bal}}(\epsilon) = \left(\frac{c_{\mathcal{H}}^{\mathrm{lin}}}{\epsilon} \right)^{1/2} \tag{11}$$

and hence that

$$H(\epsilon) = (c_{\mathcal{H}}^{\mathrm{lin}} \epsilon)^{1/2} \to 0 \tag{12}$$

as $\epsilon \to 0$.

Next, we assume for all $h \in \mathcal{H}$ that $h \in C^2(\Gamma)$ and furthermore that $\sup_{y \in \Gamma} |h'''(y)| < \infty$; we then consider piecewise quadratic interpolation over equidistant interpolation nodes $y_{N,i} = (2i/N - 1) \in \Gamma$, $0 \le i \le N$. We assume that N is even such that we may divide Γ into $N/2$ intervals $[y_{N,i}, y_{N,i+2}]$, for $i = 0, 2, 4, \ldots, N - 2$. The characteristic functions are for $y \in [y_{N,i}, y_{N,i+2}]$ then given as

$$\chi_{N,i}(y) = \frac{(y - y_{N,i+1})(y - y_{N,i+2})}{2h^2}, \tag{13}$$

$$\chi_{N,i+1}(y) = \frac{(y - y_{N,i})(y - y_{N,i+2})}{-h^2}, \tag{14}$$

$$\chi_{N,i+2}(y) = \frac{(y - y_{N,i})(y - y_{N,i+1})}{2h^2}, \tag{15}$$

for $i = 0, 2, 4, \ldots, N - 2$, where $h = 2/N = y_{N,j+1} - y_{N,j}$, $0 \le j \le N - 1$. For piecewise quadratic interpolation Hypothesis 1 and Hypothesis 2 obtain for

$$G(N) = c_{\mathcal{H}}^{\mathrm{quad}} N^{-2}, \tag{16}$$

$$D(N) = \frac{5}{2} N, \tag{17}$$

respectively, where $c_{\mathcal{H}}^{\mathrm{quad}} \equiv 28 \sup_{h \in \mathcal{H}} \|h'''\|_{L^\infty(\Gamma)}$. We include the proofs in Appendix A.2. It is straightforward to demonstrate Hypothesis 3: we note that

in this case (7) has the solution

$$N_{\text{bal}}(\epsilon) = \left(\frac{2c_{\mathcal{H}}^{\text{quad}}}{5\epsilon} \right)^{1/3} \tag{18}$$

and hence that

$$H(\epsilon) = \frac{5}{2} \left(\frac{2}{5} c_{\mathcal{H}}^{\text{quad}} \right)^{1/3} \epsilon^{2/3} \to 0 \tag{19}$$

as $\epsilon \to 0$.

Finally, we assume that all $h \in \mathcal{H}$ are analytic over Γ and consider standard Chebyshev interpolation over the $N + 1$ Chebyshev nodes $y_{N,i} = -\cos(i\pi/N)$, $0 \le i \le N$. The characteristic functions are in this case the Lagrange polynomials $\chi_{N,i} \in \mathbb{P}_N(\Gamma)$ that satisfy $\chi_{N,i}(y_{N,j}) = \delta_{i,j}$, $0 \le i, j \le N$; here $\mathbb{P}_N(\Gamma)$ is the space of degree N polynomials over Γ. For Chebyshev interpolation Hypothesis 1 and Hypothesis 2 obtain for

$$G(N) = c_{\mathcal{H}}^{\text{Cheb}} N e^{-N \log(\rho_{\mathcal{H}})}, \quad (N > 0) \tag{20}$$

$$D(N) = N^2, \tag{21}$$

respectively, where $c_{\mathcal{H}}^{\text{Cheb}} > 0$ and $\rho_{\mathcal{H}} > 1$ depend only on $\mathcal{H}$.

A proof of (20) can be found in [12]; a similar but somewhat less optimal result is obtained in [15, Eq. (4.18)]. However the result in [15] holds only for the maximum error over the $N+1$ Chebyshev interpolation nodes. In [12] this result is improved and extended to a bound for the maximum pointwise error over the entire interval $[-1, 1]$. We note that results similar to (20) (arbitrarily high algebraic order convergence for smooth functions) are common in the literature for L^2 or Sobolev norms; see for example [2, 4, 15]. The pointwise exponential convergence estimate (20) required for our theoretical derivation in this paper proved more difficult to find.

The result (6) in Hypothesis 2 obtains in this case with equality. We refer to [13, pp. 119–121] for a proof.

We finally demonstrate Hypothesis 3: we let $c = c_{\mathcal{H}}^{\text{Cheb}}$, $\eta = \eta_{\mathcal{H}} = \log(\rho_{\mathcal{H}}) > 0$, and we note that in this case (7) yields the transcendental equation

$$cN e^{-N\eta} = N^2 \epsilon, \tag{22}$$

which admits the solution

$$N_{\text{bal}}(\epsilon) = \frac{1}{\eta} \mathcal{W} \left(\frac{c\eta}{\epsilon} \right); \tag{23}$$

here, $\mathcal{W}$ denotes the Lambert W function [5] defined by $\xi = \mathcal{W}(\xi) e^{\mathcal{W}(\xi)}$ for any $\xi \in \mathbb{C}$.

For real $\xi > e$, we have $\mathcal{W}(\xi) < \log(\xi)$; we demonstrate the proof in Appendix A.3. Thus, for sufficiently large ϵ such that $c\eta/\epsilon > e$, we obtain

$$N_{\text{bal}}(\epsilon) < \frac{1}{\eta} \log \left(\frac{c\eta}{\epsilon} \right) = \frac{1}{\eta} \left(\log(c\eta) + \log(1/\epsilon) \right) \le A \log(1/\epsilon) \tag{24}$$

for some sufficiently large constant A. Hence in this case

$$H(\epsilon) < \epsilon A^2 (\log(1/\epsilon))^2. \tag{25}$$

We now consider $H(\epsilon)$ as $\epsilon \to 0$. By application of l'Hôpital's rule twice (Eqs. (28) and (30) below) we obtain

$$\lim_{\epsilon \to 0} H(\epsilon) < A^2 \lim_{\epsilon \to 0} \epsilon (\log(1/\epsilon))^2 \tag{26}$$

$$= A^2 \lim_{\epsilon \to 0} \frac{(\log(\epsilon))^2}{1/\epsilon} \tag{27}$$

$$= A^2 \lim_{\epsilon \to 0} \frac{2\log(\epsilon)/\epsilon}{-1/\epsilon^2} \tag{28}$$

$$= 2A^2 \lim_{\epsilon \to 0} \frac{\log(\epsilon)}{-1/\epsilon} \tag{29}$$

$$= 2A^2 \lim_{\epsilon \to 0} \frac{1/\epsilon}{1/\epsilon^2} \tag{30}$$

$$= 2A^2 \lim_{\epsilon \to 0} \epsilon = 0. \tag{31}$$

Hypothesis 3 thus holds.

3   A General *A Priori* Convergence Result

We introduce an approximation space $W_M \equiv W_M(\Omega)$ of finite dimension M. For any $\mu \in \mathcal{D}$, our approximation to the function $\mathcal{F}(\cdot; \mu) : \Omega \to \mathbb{R}$ shall reside in W_M; the particular approximation scheme invoked is not relevant for our theoretical results in this section. We show here that if, for any $\mu \in \mathcal{D}$, the error in the best $L^\infty(\Omega)$ approximation to $\mathcal{F}(\cdot; \mu)$ in W_M goes to zero as $M \to \infty$, then, for any multi-index β, $|\beta| \geq 0$, the error in the best $L^\infty(\Omega)$ approximation to $\mathcal{F}^{(\beta)}(\cdot; \mu)$ in W_M also goes to zero as $M \to \infty$. Of course, only modest M are of interest in practice: the computational cost associated with the approximation is in general M-dependent. However, our theoretical results in this section provide some promise that we may in practice invoke the original approximation space W_M and approximation procedure also for the approximation of parametric derivatives.

We introduce, for any fixed $p \geq 0$ and any $M \geq 1$, the order p derivative error

$$e_M^p \equiv \max_{\beta \in \mathcal{M}_p} \max_{\mu \in \mathcal{D}} \inf_{w \in W_M} \|\mathcal{F}^{(\beta)}(\cdot; \mu) - w\|_{L^\infty(\Omega)}. \tag{32}$$

We then recall the definition of $\mathcal{J}_{\beta,j}$ from (4), and state

Proposition 1. *Let p be a fixed non-negative integer. Assume that Hypotheses 1, 2, and 3 hold for the family of functions $\mathcal{H}$ given by*

$$\mathcal{H} = \{\mathcal{J}_{\beta,j}(x; \cdot) : x \in \Omega, \beta \in \mathcal{M}_p, 1 \leq j \leq P\}. \tag{33}$$

7

In this case, if $e_M^p \to 0$ as $M \to \infty$, then

$$e_M^{p+1} \to 0 \tag{34}$$

as $M \to \infty$.

Proof. For each $x \in \Omega$, and for given $\beta \in \mathcal{M}_p$ and j, $1 \leq j \leq P$, we first introduce the interpolant $\mathcal{J}_{N,\beta,j}(x; \cdot) \equiv I_N \mathcal{J}_{\beta,j}(x; \cdot) \in \mathbb{P}_N(\mathcal{S})$ given by

$$\mathcal{J}_{N,\beta,j}(x; \cdot) \equiv I_N \mathcal{J}_{\beta,j}(x; \cdot) = \sum_{i=0}^{N} \mathcal{J}_{\beta,j}(x; y_{N,i}) \chi_{N,i}(\cdot); \tag{35}$$

recall that here, $\mathcal{S} = \mathcal{D}_{(j)} = [-1, 1]$, and $\chi_{N,i} : \mathcal{S} \to \mathbb{R}$, $0 \leq i \leq N$, are characteristic functions that satisfy $\chi_{N,i}(y_{N,j}) = \delta_{i,j}$, $0 \leq i, j \leq N$.

We then introduce functions

$$\hat{w}_{\beta,j}^*(\cdot; \mu_{(1)}, \ldots, \mu_{(j-1)}, y, \mu_{(j+1)}, \ldots, \mu_{(P)}) \equiv w_{\beta,j}^*(\cdot; y)$$
$$\equiv \arg \inf_{w \in W_M} \|\mathcal{J}_{\beta,j}(\cdot; y) - w\|_{L^\infty(\Omega)} \tag{36}$$

for any $y \in \mathcal{S}$.[2] Next, for all $(x, y) \in \Omega \times \mathcal{S}$, we consider an approximation to $\mathcal{J}_{\beta,j}(x; y)$ given by $\sum_{i=0}^{N} w_{\beta,j}^*(x; y_{N,i}) \chi_{N,i}(y)$. Note that this approximation is just an interpolation between the optimal approximations in W_M at each of the interpolation nodes $y_{N,i} \in \mathcal{S}$. We next let $'$ denote differentiation with respect to the variable y and consider the error in the derivative of this approximation. By the triangle inequality we obtain

$$\left\| \mathcal{J}_{\beta,j}' - \sum_{i=0}^{N} w_{\beta,j}^*(\cdot; y_{N,i}) \chi_{N,i}' \right\|_{L^\infty(\Omega \times \mathcal{S})}$$
$$= \left\| \mathcal{J}_{N,\beta,j}' - \sum_{i=0}^{N} w_{\beta,j}^*(\cdot; y_{N,i}) \chi_{N,i}' + \mathcal{J}_{\beta,j}' - \mathcal{J}_{N,\beta,j}' \right\|_{L^\infty(\Omega \times \mathcal{S})}$$
$$\leq \left\| \mathcal{J}_{N,\beta,j}' - \sum_{i=0}^{N} w_{\beta,j}^*(\cdot; y_{N,i}) \chi_{N,i}' \right\|_{L^\infty(\Omega \times \mathcal{S})} + \left\| \mathcal{J}_{\beta,j}' - \mathcal{J}_{N,\beta,j}' \right\|_{L^\infty(\Omega \times \mathcal{S})}. \tag{37}$$

Here, $\mathcal{J}_{N,\beta,j}' \equiv (\mathcal{J}_{N,\beta,j})' = \sum_{i=0}^{N} \mathcal{J}_{\beta,j}(\cdot; y_{N,i}) \chi_{N,i}'(\cdot)$.

We first develop a bound for the first term on the right hand side of (37).

---

[2]Note that $w_{\beta,j}^*$ depends on all P parameter values $\mu_{(i)}$, $1 \leq i \leq P$. However we shall suppress the dependence on parameters $\mu_{(i)}$, $i \neq j$, for simplicity of notation.

By (35) and the triangle inequality we obtain

$$\left\|\mathcal{J}'_{N,\beta,j} - \sum_{i=0}^{N} w^*_{\beta,j}(\cdot\,; y_{N,i})\chi'_{N,i}\right\|_{L^\infty(\Omega\times\mathcal{S})} \tag{38}$$

$$= \left\|\sum_{i=0}^{N} (\mathcal{J}_{\beta,j}(\cdot\,; y_{N,i}) - w^*_{\beta,j}(\cdot\,; y_{N,i}))\chi'_{N,i}\right\|_{L^\infty(\Omega\times\mathcal{S})} \tag{39}$$

$$\leq \left\|\sum_{i=0}^{N} |\mathcal{J}_{\beta,j}(\cdot\,; y_{N,i}) - w^*_{\beta,j}(\cdot\,; y_{N,i})||\chi'_{N,i}|\right\|_{L^\infty(\Omega\times\mathcal{S})} \tag{40}$$

$$\leq \left\|\max_{0\leq i\leq N} |\mathcal{J}_{\beta,j}(\cdot\,; y_{N,i}) - w^*_{\beta,j}(\cdot\,; y_{N,i})|\sum_{j=0}^{N}|\chi'_{N,j}|\right\|_{L^\infty(\Omega\times\mathcal{S})}. \tag{41}$$

Further, by Hypothesis 2, by taking the maximum over the interval $\mathcal{S}$, by the definition of $w^*_{\beta,j}$ in (36), and finally the definition of e^p_M in (32), we obtain

$$\left\|\max_{0\leq i\leq N} |\mathcal{J}_{\beta,j}(\cdot\,; y_{N,i}) - w^*_{\beta,j}(\cdot\,; y_{N,i})|\sum_{j=0}^{N}|\chi'_{N,j}|\right\|_{L^\infty(\Omega\times\mathcal{S})} \tag{42}$$

$$\leq D(N)\max_{0\leq i\leq N}\|\mathcal{J}_{\beta,j}(\cdot\,; y_{N,i}) - w^*_{\beta,j}(\cdot\,; y_{N,i})\|_{L^\infty(\Omega)} \tag{43}$$

$$\leq D(N)\max_{y\in\mathcal{S}}\|\mathcal{J}_{\beta,j}(\cdot\,; y) - w^*_{\beta,j}(\cdot\,; y)\|_{L^\infty(\Omega)} \tag{44}$$

$$= D(N)\max_{y\in\mathcal{S}}\inf_{w\in W_M}\|\mathcal{J}_{\beta,j}(\cdot\,; y) - w\|_{L^\infty(\Omega)} \tag{45}$$

$$\leq D(N)e^p_M, \tag{46}$$

and hence

$$\left\|\mathcal{J}'_{N,\beta,j} - \sum_{i=0}^{N} w^*_{\beta,j}(\cdot\,; y_{N,i})\chi'_{N,i}\right\|_{L^\infty(\Omega\times\mathcal{S})} \leq D(N)e^p_M \tag{47}$$

We next develop a bound for the second term on the right hand side of (37). To this end we invoke the fact that for any $x \in \Omega$ the function $\mathcal{J}_{\beta,j}(x;\cdot)$ belongs to the family $\mathcal{H}$ defined by (33). We may thus invoke Hypothesis 1 to directly obtain

$$\|\mathcal{J}'_{\beta,j} - \mathcal{J}'_{N,\beta,j}\|_{L^\infty(\Omega\times\mathcal{S})} \leq G(N). \tag{48}$$

We may now combine (37) with (47) and (48) to obtain

$$\left\|\mathcal{J}'_{\beta,j} - \sum_{i=0}^{N} w^*_{\beta,j}(\cdot\,; y_{N,i})\chi'_{N,i}\right\|_{L^\infty(\Omega\times\mathcal{S})} \leq G(N) + D(N)e^p_M. \tag{49}$$

Next, we introduce $\beta^+_j = \beta + e_j$ where e_j is the canonical unit vector of dimension P with the j'th entry equal to unity; we recall that β has length

$|\beta| = p$ and hence β_j^+ has length $|\beta_j^+| = p + 1$. We note that the multi-index β, the parameter values $\mu_{(i)} \in \mathcal{D}_{(i)}$, $1 \leq i \leq P$, $i \neq j$, as well as the dimension j, were chosen arbitrarily above. We may thus conclude (recall above we wrote $y = \mu_{(j)}$ for each fixed j) that[3]

$$\max_{\beta \in \mathcal{M}_p} \max_{1 \leq j \leq P} \max_{\mu \in \mathcal{D}} \left\| \mathcal{F}^{(\beta_j^+)}(\cdot; \mu) - \sum_{i=0}^{N} w_{\beta,j}^*(\cdot; y_{N,i}) \chi_{N,i}'(\mu_{(j)}) \right\|_{L^\infty(\Omega)}$$
$$\leq G(N) + D(N) e_M^p. \quad (50)$$

We note that for any $\beta \in \mathcal{M}_p$, any $\mu_{(j)} \in \mathcal{D}_{(j)}$, and any $1 \leq j \leq P$, the function $\sum_{i=0}^{N} \chi_{N,i}'(\mu_{(j)}) w_{\beta,j}^*(\cdot; y_{N,i})$ is just one particular member of W_M. For the error e_M^{p+1} of the *best approximation* of any derivative of order $p + 1$ in W_M we thus obtain

$$e_M^{p+1} = \max_{\beta \in \mathcal{M}_{p+1}} \max_{\mu \in \mathcal{D}} \inf_{w \in W_M} \| \mathcal{F}^{(\beta)}(\cdot; \mu) - w \|_{L^\infty(\Omega)} \leq G(N) + D(N) e_M^p. \quad (51)$$

The final step is to bound the right-hand side of (51) in terms of e_M^p alone. To this end we note that we may choose N freely; for example we may choose N to minimize the right hand side of (51). However, we shall make a different choice for N: we choose $N = N_{\mathrm{bal}}(e_M^p)$ to balance the two terms on the right hand side of (51). With this choice we obtain

$$e_M^{p+1} \leq 2D(N_{\mathrm{bal}}(e_M^p)) e_M^p = 2H(e_M^p), \quad (52)$$

and thus $e_M^{p+1} \to 0$ as $e_M^p \to 0$ by Hypothesis 3. $\qquad \square$

We now provide three lemmas, each of which quantifies the convergence in Proposition 1 under different regularity conditions. The first lemma quantifies the convergence in Proposition 1 in the case that $\mathcal{F}(x; \cdot) \in C^1(\mathcal{D})$ for all $x \in \Omega$.

Lemma 1. *Assume for all $x \in \Omega$ that $\mathcal{F}^{(\beta)}(x; \cdot) \in C^1(\mathcal{D})$ and furthermore that all second order derivatives of $\mathcal{F}^{(\beta)}(x; \cdot)$ are bounded over $\mathcal{D}$. Then for any fixed $p = |\beta| \geq 0$ there is a constant $C_{p+1} > 0$ (independent of M) such that for any M*

$$e_M^{p+1} \leq C_{p+1} \sqrt{e_M^p}. \quad (53)$$

Proof. In this case we may invoke piecewise linear interpolation as our interpolation system in the proof of Proposition 1. By (12) and (52) we obtain $e_M^{p+1} \leq 2(c_{\mathcal{H}}^{\mathrm{lin}} e_M^p)^{1/2}$. The result follows for $C_{p+1} = 2(c_{\mathcal{H}}^{\mathrm{lin}})^{1/2}$. $\qquad \square$

The next lemma quantifies the convergence in Proposition 1 in the case that $\mathcal{F}^{(\beta)}(x; \cdot) \in C^2(\mathcal{D})$ for all $x \in \Omega$.

---

[3]Recall that $w_{\beta,j}^*$ depends implicitly on the parameter values $\mu_{(i)}$, $i \neq j$, through (36): $w_{\beta,j}^*(\cdot; y_{N,i}) \equiv \hat{w}_{\beta,j}^*(\cdot; \mu_{(1)}, \ldots, \mu_{(j-1)}, y_{N,i}, \mu_{(j+1)}, \ldots, \mu_{(P)})$.

Lemma 2. *Assume for all $x \in \Omega$ that $\mathcal{F}^{(\beta)}(x; \cdot) \in C^2(\mathcal{D})$ and furthermore that all third order derivatives of $\mathcal{F}^{(\beta)}(x; \cdot)$ are bounded over $\mathcal{D}$. Then for any fixed $p = |\beta| \geq 0$ there is a constant $C_{p+1} > 0$ (independent of M) such that for any M*

$$e_M^{p+1} \leq C_{p+1}(e_M^p)^{2/3}. \tag{54}$$

Proof. In this case we may invoke piecewise quadratic interpolation as our interpolation system in the proof of Proposition 1. By (19) and (52) we obtain $e_M^{p+1} \leq 5(2c_{\mathcal{H}}^{\text{quad}}/5)^{1/3}(e_M^p)^{2/3}$. The result follows for $C_{p+1} = 5(2c_{\mathcal{H}}^{\text{quad}}/5)^{1/3}$. $\qquad\square$

We make the following remark concerning Lemma 1 and Lemma 2 in the case of algebraic convergence.

Remark 1. *Let $|\beta| = p$, and assume for all $x \in \Omega$ that $\mathcal{F}^{(\beta)}(x, \cdot) \in C^{q_p}(\mathcal{D})$, $q_p \geq 1$, and furthermore that all $q_p + 1$ order derivatives of $\mathcal{F}^{(\beta)}(x, \cdot)$ are bounded over $\mathcal{D}$. Suppose in this case that $e_M^p \propto M^{-r_p}$ with $r_p > 0$.[4] For $q_p = 1$ we may invoke Lemma 1 to obtain*

$$e_M^{p+1} \leq C_{p+1}(e_M^p)^{\frac{1}{2}} \propto M^{-\frac{r_p}{2}} \propto M^{\frac{r_p}{2}} e_M^p. \tag{55}$$

Similarly, for $q_p = 2$ we may invoke Lemma 2 to obtain

$$e_M^{p+1} \leq C_{p+1}(e_M^p)^{\frac{2}{3}} \propto M^{-\frac{2r_p}{3}} \propto M^{\frac{r_p}{3}} e_M^p. \tag{56}$$

More generally, with higher-regularity versions of Lemma 1 and Lemma 2, we expect for any $q_p \geq 1$ that $H(\epsilon) \propto \epsilon^{1 - \frac{1}{1+q_p}}$ and hence that

$$e_M^{p+1} \leq C_{p+1}(e_M^p)^{1 - \frac{1}{q_p + 1}} \propto M^{-r_p\left(1 - \frac{1}{q_p + 1}\right)} \propto M^{\frac{r_p}{q_p + 1}} e_M^p. \tag{57}$$

We shall comment on these estimates further in our discussion of numerical results in Section 5.

The third lemma quantifies the convergence in Proposition 1 in the case that $\mathcal{F}(x, \cdot)$ is analytic over $\mathcal{D}$.

Lemma 3. *Assume for all $x \in \Omega$ that $\mathcal{F}(x, \cdot) : \mathcal{D} \to \mathbb{R}$ is analytic over $\mathcal{D}$. Then for any fixed $p \geq 0$ there exists a constant $C_{p+1} > 0$ (independent of M) such that for any $M > M_0$,*

$$e_M^{p+1} \leq C_{p+1} \log(e_M^p)^2 e_M^p. \tag{58}$$

Moreover, if for some p (independent of M)

$$e_M^p \leq \hat{c} M^\sigma e^{-\gamma M^\alpha} \tag{59}$$

where σ and α are constants and γ and $\hat{c}$ are positive constants, then there exists a constant $\hat{C}_{p+1}$ such that

$$e_M^{p+1} \leq \hat{C}_{p+1} M^{\sigma + 2\alpha} e^{-\gamma M^\alpha}. \tag{60}$$

---

[4]The convergence rate r_p will depend on the sequence of spaces W_M; we expect that the convergence rate also will depend on the parametric regularity q_p.

Proof. In this case we may invoke Chebyshev interpolation as our interpolation system in the proof of Proposition 1. For $M > M_0$ (i.e. sufficiently small e_M^p), we may use (25) and (52) to obtain $e_M^{p+1} < 2A^2(\log(1/e_M^p))^2 e_M^p$. The result (58) follows for $C_{p+1} = 2A^2$ since $(\log(1/e_M^p))^2 = (\log(e_M^p))^2$.

The additional result (60) follows from the assumption (59) since the right hand side of (58) decreases monotonically as $e_M^p \to 0$ for $e_M^p < e^{-2}$. We obtain in this case

$$
\begin{aligned}
e_M^{p+1} &\le C_{p+1} \log(\hat{c} M^\sigma e^{-\gamma M^\alpha})^2 \hat{c} M^\sigma e^{-\gamma M^\alpha} \\
&= C_{p+1} \hat{c} (\log \hat{c} + \sigma \log M - \gamma M^\alpha)^2 M^\sigma e^{-\gamma M^\alpha} \\
&\le \hat{C}_{p+1} M^{2\alpha} M^\sigma e^{-\gamma M^\alpha} \quad (61)
\end{aligned}
$$

for M sufficiently large. □

We make the following two remarks concerning Lemma 3.

Remark 2. *Note that Lemma 3, in contrast to 1 and 2, only holds for all $M > M_0$. The reason is that to obtain (58) we invoke the assumption $\mathcal{W}(\xi) < \log(\xi)$, which holds only for $\xi > e$ (see Appendix A.3 for a proof). Since here $\xi \propto 1/e_M^p$, this assumption will be satisfied when e_M is sufficiently small, i.e. when M is sufficiently large.*

Remark 3. *We note that we may invoke the result (60) recursively to obtain, for any fixed p and all $M > M_{0,p}$,*

$$
e_M^p \le \tilde{C}_p M^{\sigma+2\alpha p} e^{-\gamma M^\alpha} \tag{62}
$$

whenever $e_M^0 \le c M^\sigma e^{-\gamma M^\alpha}$. Here, $\tilde{C}_p$ and $M_{0,p}$ depend on p, but $\tilde{C}_p$ does not depend on M.

4   The Empirical Interpolation Method

In this section we first recall the Empirical Interpolation Method (EIM) [1, 8, 9] and then consider the convergence theory of the previous section applied to the EIM. In particular, we relate the error in the EIM approximation to the error in the best approximation in the EIM approximation space.

The EIM approximation space is spanned by precomputed snapshots of a parameter dependent "generating function" for judiciously chosen parameter values from a predefined parameter domain. Given any new parameter value in this parameter domain, we can construct an approximation to the generating function at this new parameter value — or in principle an approximation to any function defined over the same spatial domain — as a linear combination of the EIM basis functions. The particular linear combination is determined through interpolation at judiciously chosen points in the spatial domain. For parametrically smooth functions, the EIM approximation to the generating function yields rapid — typically exponential — convergence.

4.1 Procedure

We introduce the generating function $\mathcal{G} : \Omega \times \mathcal{D} \to \mathbb{R}$ such that for all $\mu \in \mathcal{D}$, $\mathcal{G}(\cdot; \mu) \in L^\infty(\Omega)$. We introduce a training set $\Xi_{\text{train}} \subset \mathcal{D}$ of finite cardinality $|\Xi_{\text{train}}|$ which shall serve as our computational surrogate for $\mathcal{D}$. We also introduce a triangulation $\mathcal{T}_\mathcal{N}(\Omega)$ of Ω with $\mathcal{N}$ vertices over which we shall in practice, for any $\mu \in \mathcal{D}$, realize $\mathcal{G}(\cdot; \mu)$ as a piecewise linear function.

Now, for $1 \leq M \leq M_{\max} < \infty$, we define the EIM approximation space $W_M^\mathcal{G}$ and the EIM interpolation nodes $T_M^\mathcal{G}$ associated with $\mathcal{G}$; here, $M_{\max}$ is a specified maximum EIM approximation space dimension. We first choose (randomly, say) an initial parameter value $\mu_1 \in \mathcal{D}$; we then determine the first EIM interpolation node as $t_1 = \arg\sup_{x \in \Omega} |\mathcal{G}(x; \mu_1)|$; we next define the first EIM basis function as $q_1 = \mathcal{G}(\cdot; \mu_1)/\mathcal{G}(t_1; \mu_1)$. We can then, for $M = 1$, define $W_M^\mathcal{G} = \text{span}\{q_1\}$ and $T_M^\mathcal{G} = \{t_1\}$. We also define a nodal value matrix B^1 with (a single) element $B_{1,1}^1 = q_1(t_1) = 1$.

Next, for $2 \leq M \leq M_{\max}$, we first compute the empirical interpolation of $\mathcal{G}(\cdot; \mu)$ for all $\mu \in \Xi_{\text{train}}$: we solve the linear system

$$\sum_{j=1}^{M-1} \phi_j^{M-1}(\mu) B_{i,j}^{M-1} = \mathcal{G}(t_i; \mu), \quad 1 \leq i \leq M-1, \tag{63}$$

and compute the empirical interpolation $\mathcal{G}_{M-1}(\cdot; \mu) \in W_{M-1}^\mathcal{G}$ as

$$\mathcal{G}_{M-1}(\cdot; \mu) = \sum_{i=1}^{M-1} \phi_i^{M-1}(\mu) q_i, \tag{64}$$

for all $\mu \in \Xi_{\text{train}}$. We then choose the next parameter $\mu_M \in \mathcal{D}$ as the maximizer of the EIM interpolation error over the training set,

$$\mu_M = \arg\max_{\mu \in \Xi_{\text{train}}} \|\mathcal{G}_{M-1}(\cdot; \mu) - \mathcal{G}(\cdot; \mu)\|_{L^\infty(\Omega)}; \tag{65}$$

note that thanks to our piecewise linear realization of $\mathcal{G}(\cdot; \mu)$, the norm evaluation in (65) is a simple comparison of function values at the $\mathcal{N}$ vertices of $\mathcal{T}_\mathcal{N}(\Omega)$. We now choose the next EIM interpolation node as the point in Ω at which the EIM error associated with $\mathcal{G}_{M-1}(\mu_M)$ is largest,

$$t_M = \arg\sup_{x \in \Omega} |\mathcal{G}_{M-1}(x; \mu_M) - \mathcal{G}(x; \mu_M)|. \tag{66}$$

The next EIM basis function is then

$$q_M = \frac{\mathcal{G}_{M-1}(\cdot; \mu_M) - \mathcal{G}(\cdot; \mu_M)}{\mathcal{G}_{M-1}(t_M; \mu_M) - \mathcal{G}(t_M; \mu_M)}. \tag{67}$$

We finally enrich the EIM space: $W_M^\mathcal{G} = \text{span}\{q_1, \ldots, q_M\}$; expand the set of nodes: $T_M^\mathcal{G} = \{t_1, \ldots, t_M\}$; and expand the nodal value matrix: $B_{i,j}^M = q_j(t_i)$, $1 \leq i, j \leq M$.

Now, given any function $\mathcal{F} : \Omega \times \mathcal{D} \to \mathbb{R}$ (in particular, we shall consider $\mathcal{F} = \mathcal{G}^{(\beta)}$), we define for any $\mu \in \mathcal{D}$ and for $1 \leq M \leq M_{\max}$ the empirical interpolation of $\mathcal{F}(\cdot; \mu)$ in the space $W_M^{\mathcal{G}}$ (the space generated by $\mathcal{G}$) as

$$\mathcal{F}_M^{\mathcal{G}}(\cdot; \mu) = \sum_{i=1}^{M} \phi_i^M(\mu) q_i, \tag{68}$$

where the coefficients $\phi_i^M(\mu)$, $1 \leq i \leq M$, solve the linear system

$$\sum_{j=1}^{M} \phi_j^M(\mu) B_{i,j}^M = \mathcal{F}(t_i; \mu), \quad 1 \leq i \leq M. \tag{69}$$

We note that by construction the matrices $B^M \in \mathbb{R}^{M \times M}$, $1 \leq M \leq M_{\max}$, are lower triangular: by (63) and (64), $\mathcal{G}_{M-1}(t_j; \mu_M) = \mathcal{G}(t_j; \mu_M)$ for $j < M$. As a result, computation of the EIM coefficients ϕ_j^M, $1 \leq j \leq M$, in (69) and (63) are $\mathcal{O}(M^2)$ operations. We emphasize that the computational cost associated with the EIM approximation (68)–(69) (after snapshot precomputation) is independent of the number $\mathcal{N}$ of vertices in the triangulation $\mathcal{T}_{\mathcal{N}}(\Omega)$. The number $\mathcal{N}$ may thus be chosen conservatively.

We next note that, for any multi-index β,

$$(\mathcal{F}_M^{\mathcal{G}})^{(\beta)} = \left(\sum_{i=1}^{M} \phi_i^M(\mu) q_i \right)^{(\beta)} = \sum_{i=1}^{M} \varphi_i^M(\mu) q_i, \tag{70}$$

where $\varphi_i^M(\mu) = (\phi_i^M)^{(\beta)}(\mu)$, $1 \leq i \leq M$, solve the linear system (recall that the matrix B^M is μ-independent)

$$\sum_{j=1}^{M} \varphi_j^M(\mu) B_{i,j}^M = \mathcal{F}^{(\beta)}(t_i; \mu), \quad 1 \leq i \leq M. \tag{71}$$

Hence,

$$(\mathcal{F}_M^{\mathcal{G}})^{(\beta)} = (\mathcal{F}^{(\beta)})_M^{\mathcal{G}}, \tag{72}$$

that is, the parametric derivative of the approximation is equivalent to the approximation of the parametric derivative. We note that this equivalence holds since we invoke the same approximation space $W_M^{\mathcal{G}}$ for both EIM approximations $\mathcal{F}_M^{\mathcal{G}}$ and $(\mathcal{F}^{(\beta)})_M^{\mathcal{G}}$.

4.2 Convergence theory applied to the EIM

In this section we relate the error in the EIM approximation to the best approximation in the EIM approximation space. To this end we introduce the Lebesgue constants [11]

$$\Lambda_M = \sup_{x \in \Omega} \sum_{i=1}^{M} |V_i^M(x)|, \quad 1 \leq M \leq M_{\max}, \tag{73}$$

where $V_i^M \in W_M^{\mathcal{G}}$ are the characteristic functions associated with $W_M^{\mathcal{G}}$ and $T_M^{\mathcal{G}}$: $V_i^M(t_j) = \delta_{i,j}$, $1 \le i,j \le M$. Our theory of Section 3 considers the convergence in the best $L^\infty(\Omega)$ approximation. However, we can relate the EIM approximation to the best approximation of through

Lemma 4. *The error in the EIM derivative approximation satisfies*

$$\|\mathcal{F}^{(\beta)}(\cdot;\mu) - (\mathcal{F}^{(\beta)})_M^{\mathcal{G}}(\cdot;\mu)\|_{L^\infty(\Omega)}$$
$$\le (1 + \Lambda_M) \inf_{w \in W_M^{\mathcal{G}}} \|\mathcal{F}^{(\beta)}(\cdot;\mu) - w\|_{L^\infty(\Omega)}. \quad (74)$$

Proof. The proof is identical to [1, Lemma 3.1]. We first introduce $\mathcal{F}_M^{\mathcal{G},*}(\cdot;\mu) = \arg\inf_{w \in W_M^{\mathcal{G}}} \|\mathcal{F}(\cdot;\mu) - w\|_{L^\infty(\Omega)}$, and define coefficient functions $\omega_m^M(\mu)$, $1 \le m \le M$, such that $\mathcal{F}_M^{\mathcal{G}}(\cdot;\mu) - \mathcal{F}_M^{\mathcal{G},*}(\cdot;\mu) = \sum_{m=1}^M \omega_m^M(\mu)q_m$. By the interpolation property of the EIM we then obtain

$$\mathcal{F}(t_n;\mu) - \mathcal{F}_M^{\mathcal{G},*}(t_n;\mu) = \mathcal{F}_M^{\mathcal{G}}(t_n;\mu) - \mathcal{F}_M^{\mathcal{G},*}(t_n;\mu) = \sum_{m=1}^M \omega_m^M(\mu)q_m(t_n). \quad (75)$$

We then introduce $E_M^{\mathcal{G}}(\mu) = \|\mathcal{F}(\cdot;\mu) - \mathcal{F}_M^{\mathcal{G}}(\cdot;\mu)\|_{L^\infty(\Omega)}$ and $E_M^{\mathcal{G},*}(\mu) = \|\mathcal{F}(\cdot;\mu) - \mathcal{F}_M^{\mathcal{G},*}(\cdot;\mu)\|_{L^\infty(\Omega)}$ and note that

$$E_M^{\mathcal{G}}(\mu) - E_M^{\mathcal{G},*}(\mu) \le \Big\| \sum_{m=1}^M \omega_m^M(\mu)q_m \Big\|_{L^\infty(\Omega)} = \Big\| \sum_{n=1}^M \sum_{m=1}^M \omega_m^M(\mu)q_m(t_n)V_n^M \Big\|_{L^\infty(\Omega)}$$
$$= \Big\| \sum_{n=1}^M \big(\mathcal{F}(t_n;\mu) - \mathcal{F}_M^{\mathcal{G},*}(t_n;\mu)\big)V_n^M \Big\|_{L^\infty(\Omega)} \le \Lambda_M E_M^{\mathcal{G},*}(\mu) \quad (76)$$

since $|\mathcal{F}(t_n;\mu) - \mathcal{F}_M^{\mathcal{G},*}(t_n;\mu)| \le E_M^{\mathcal{G},*}(\mu)$, $1 \le n \le M$, and by the definition of Λ_M in (73). The result (74) follows for any β by replacing $\mathcal{F}$ by $\mathcal{F}^{(\beta)}$ in the arguments above. $\qquad\square$

It can be proven [1, 8] that $\Lambda_M < 2^M - 1$. However, in actual practice the growth of Λ_M is much slower than this exponential upper bound, as we shall observe below (see also results in [1, 8, 9]). Based on Lemma 4 and the anticipated slow growth of Λ_M, we expect the EIM approximation to any parametric derivative to be good as long as the best approximation is good.

5   Numerical Results

In this section we demonstrate the theory through two numerical examples. In each example, we consider a parametrized function $\mathcal{F}$ and we generate approximation spaces with the EIM for $\mathcal{G} = \mathcal{F}$ as the generating function. To confirm the theory we compute for a large number of parameter values in a test set

$\Xi_{\text{test}} \subset \mathcal{D}$, $\Xi_{\text{test}} \neq \Xi_{\text{train}}$, the best $L^\infty(\Omega)$ approximation of $\mathcal{F}$ and the parametric derivatives $\mathcal{F}^{(\beta)}$ in these spaces. We define the maximum best approximation error over the space of dimension M as

$$e^p_{M,\text{test}} \equiv \max_{\beta \in \mathcal{M}_p} \max_{\mu \in \Xi_{\text{test}}} \inf_{w \in W^{\mathcal{G}}_M} \|\mathcal{F}^{(\beta)}(\cdot; \mu) - w\|_{L^\infty(\Omega)} \tag{77}$$

(the test set Ξ_{test} will be different for each example problem). We note that thanks to the piecewise linear representation of $\mathcal{F}$ (and its parametric derivatives), determination of the best $L^\infty(\Omega)$ approximation (and associated error) is equivalent to the solution of a linear program for each $\mu \in \Xi_{\text{test}}$.

We shall also compute *error degradation factors*

$$\rho^p_{M,\text{test}} \equiv \frac{e^p_{M,\text{test}}}{e^0_{M,\text{test}}} \tag{78}$$

as a measure of how much accuracy we loose in the approximation of order p parametric derivatives.

We finally confirm for each example that the growth of the Lebesgue constant is only modest and hence that, by Lemma 4, the EIM approximation will be close to the best approximation.

5.1   Example 1: Parametrically smooth Gaussian surface

We introduce the spatial domain $\Omega = [0,1]^2$ and the parameter domain $\mathcal{D} = [0.4, 0.6]^2$. We consider the 2D Gaussian $\mathcal{F} : \Omega \times \mathcal{D} \to \mathbb{R}$ defined by

$$\mathcal{F}(x; \mu) = \exp\left(\frac{-(x_{(1)} - \mu_{(1)})^2 - (x_{(2)} - \mu_{(2)})^2}{2\sigma^2} \right) \tag{79}$$

for $x \in \Omega$, $\mu \in \mathcal{D}$, and $\sigma \equiv 0.1$. This function is thus parametrized by the location of the maximum of the Gaussian surface. We note that for all $x \in \Omega$ the function $\mathcal{F}(x; \cdot)$ is analytic over $\mathcal{D}$; we may thus invoke Lemma 3.

We introduce a triangulation $\mathcal{T}_{\mathcal{N}}(\Omega)$ with $\mathcal{N} = 2601$ vertices; we introduce an equi-distant training set "grid" $\Xi_{\text{train}} \subset \mathcal{D}$ of size $|\Xi_{\text{train}}| = 900 = 30 \times 30$. We then pursue the EIM with $\mathcal{G} = \mathcal{F}$ for $M_{\max} = 99$.

We now introduce a uniformly distributed random test set $\Xi_{\text{test}} \subset \mathcal{D}$ of size 1000 over which we compute best approximation errors e^p_M, $1 \leq M \leq M_{\max}$. In Figure 1 we show the maximum best approximation errors $e^p_{M,\text{test}}$ for $p = 0, 1, 2, 3$. We note that the convergence is exponential not only for the best approximation of $\mathcal{F}$ ($p = 0$), but also for the best approximation of its derivatives ($p > 0$). We also note that for large M, the (exponential) rates of convergence associated with the parametric derivatives are close to the rate associated with the generating function.

To provide for some theoretical explanation for these observations we make the assumption $e^0_M = \hat{c} M^\sigma e^{-\gamma M}$. An ordinary least squares linear regression on $\log(e^0_M)$ for $35 \leq M \leq M_{\max}$ provides estimates $\log \hat{c} \approx 4.4194$, $\sigma \approx -4.4611$,
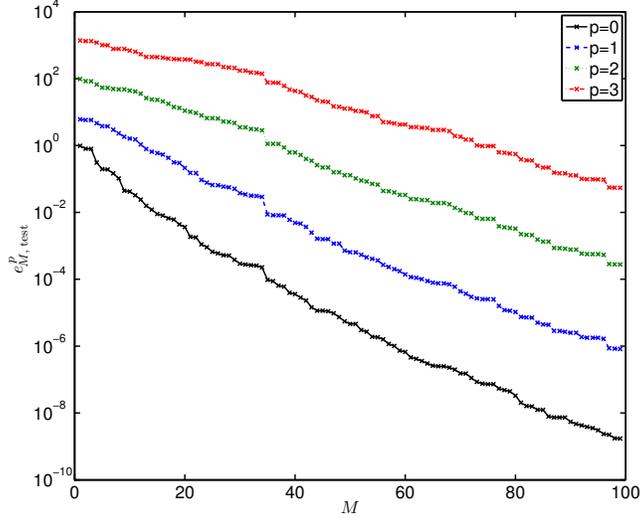
Figure 1: The maximum $L^\infty(\Omega)$ projection error over the test set, $e^p_{M,\text{test}}$, for $0 \leq p \leq 3$ for Example 1.

and $\gamma \approx 0.0436$. Based on these estimates and the relatively small associated standard errors[5] we may expect that this assumption holds. Hence we expect from Remark 3 that $e^p_M \leq \tilde{C}_p M^{\sigma+2p} e^{-\gamma M}$ also for $p > 0$. This result thus explains the exponential convergence associated with the parametric derivatives.

In Figure 2 we show the error degradation factors $\rho^p_{M,\text{test}}$ for $p = 1, 2, 3$ as functions of M. The plot suggests that indeed $\rho^p_{M,\text{test}} \leq \text{const} \cdot M^{2p}$ as predicted by the bounds $e^p_M \leq \tilde{C}_p M^{\sigma+2p} e^{-\gamma M}$, $p > 0$, obtained above. We note that had the result (60) of Lemma 3 (and (62)) been sharp we would have obtained $\rho^p_{M,\text{test}} \propto M^{2p}$. We conclude that, at least for the range for M considered for these computations and for this particular $\mathcal{F}$, the result (60) is not sharp.

We finally note that the factor M^2 in (60) originates from the *sharp* result (21); hence with our present strategy for the proof of Proposition 1 it is not clear how to sharpen (60). However, clearly our theory captures the correct qualitative behavior: we observe exponential convergence for the parametric derivatives and there is evidence of an algebraic degradation factor for the parametric derivative approximations.

Finally, in Figure 3, we report the Lebesgue constant Λ_M. We note that the growth of the Lebesgue constant is only modest. The EIM derivative approximation will thus be close to the best $L^\infty(\Omega)$ approximation in the space

---

[5]For the standard errors associated with $\log \hat{c}$, σ, and γ we obtain 0.6552, 0.4846, and 0.0033, respectively. We use these standard errors as a non-rigorous measure of the uncertainty in the estimated regression parameters; however we do not make particular assumptions on the regression error term and hence we can not assign any formal statistical interpretation to the standard errors.
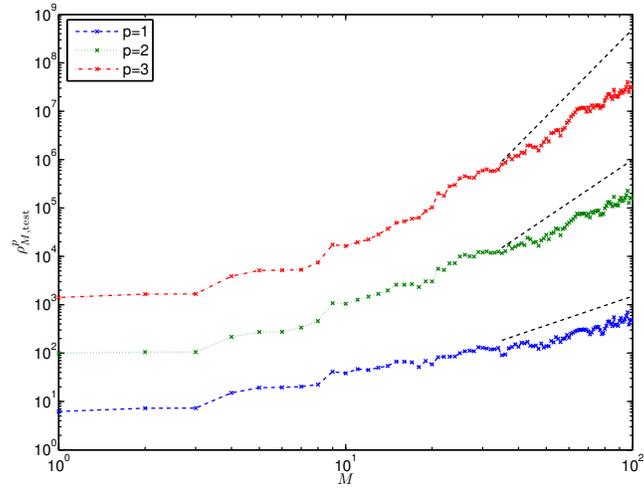
Figure 2: Error degradation factors $\rho^p_{M,\text{test}}$, $p = 1, 2, 3$, for Example 1. The shorter dashed lines are of slope M^{2p}.
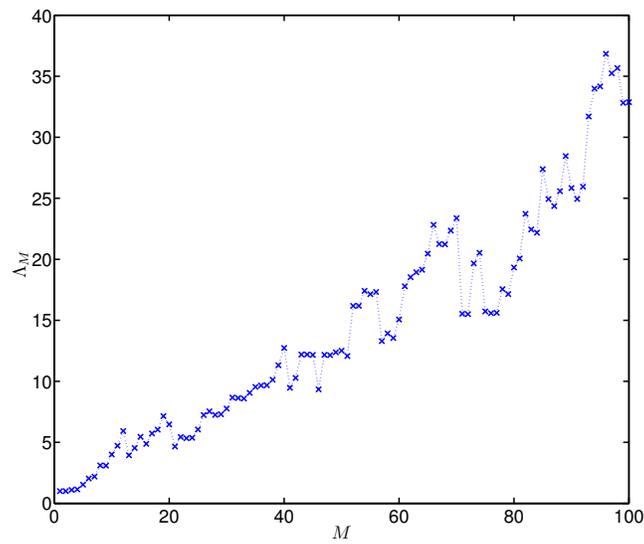

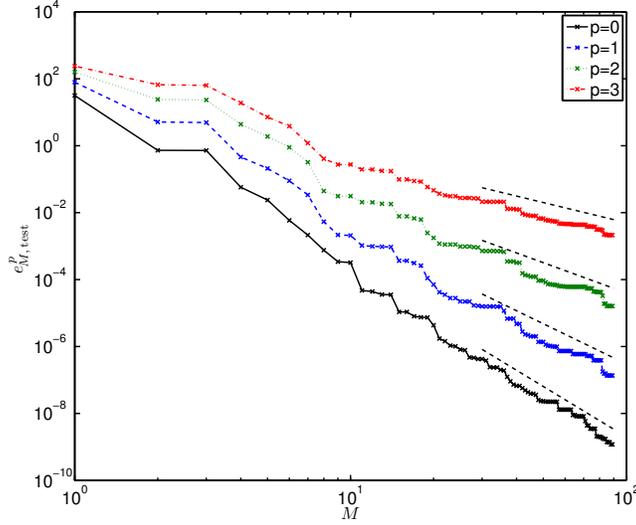Figure 3: The Lebesgue constant Λ_M for Example 1.

18

Figure 4: The maximum $L^\infty(\Omega)$ projection error over the test set, $e^p_{M,\text{test}}$, for $0 \le p \le 3$ for Example 2. The shorter dashed lines are of slope M^{-5+p}.

$W^{\mathcal{F}}_M$.

5.2 Example 2: A parametrically singular function

We introduce the spatial domain $\Omega = [-1, 1]$ and the parameter domain $\mathcal{D} = [-1, 1]$. We consider the function $\mathcal{F} : \Omega \times \mathcal{D} \to \mathbb{R}$ defined by

$$\mathcal{F}(x; \mu) = |x - \mu|^5 \tag{80}$$

for $x \in \Omega$ and $\mu \in \mathcal{D}$. The function thus has a singularity at $x = \mu$ for any $\mu \in \mathcal{D}$.

For any $x \in \Omega$ we have $\mathcal{F}^{(p)}(x; \cdot) \in C^{q_p}(\mathcal{D})$ for $q_p = 4 - p$ with $\mathcal{F}^{(5)}(x; \cdot)$ bounded over $\mathcal{D}$. Hence, to estimate e^1_M from e^0_M, we may as indicated in Remark 1 invoke a higher order version of Lemma 1 (and Lemma 2) using piecewise quartic interpolation. Similarly, to estimate e^2_M based on e^1_M, we may invoke a piecewise cubic version of Lemma 1 (and Lemma 2). To estimate e^3_M based on e^2_M, we may invoke Lemma 2 directly since $\mathcal{F}^{(2)}(x; \cdot) \in C^2(\mathcal{D})$ with its third order derivative bounded over $\mathcal{D}$.

We introduce a triangulation $\mathcal{T}_\mathcal{N}(\Omega)$ with $\mathcal{N} = 500$ vertices; we introduce an equi-distant training set "grid" $\Xi_{\text{train}} \subset \mathcal{D}$ of size $|\Xi_{\text{train}}| = 500$. We then pursue the EIM with $\mathcal{G} = \mathcal{F}$ for $M_{\max} = 89$.

We now introduce a uniformly distributed random test set $\Xi_{\text{test}} \subset \mathcal{D}$ of size 500. In Figure 4 we show the maximum best approximation errors $e^p_{M,\text{test}}$ for $p = 0, 1, 2, 3$. The convergence is algebraic: ordinary least squares best fits to the slopes for $30 \le M \le M_{\max}$ yield $e^0_{M,\text{test}} \approx \text{const} \cdot M^{-5.13}$, $e^1_{M,\text{test}} \approx$
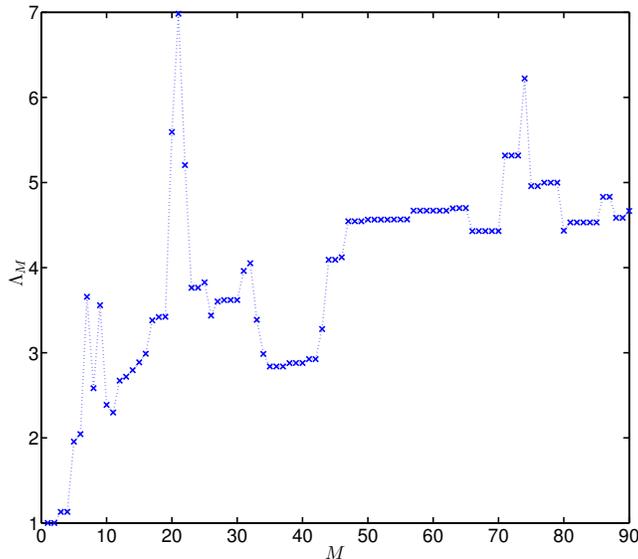
Figure 5: The Lebesgue constant Λ_M for Example 2.

const $\cdot M^{-4.27}$, $e^2_{M,\text{test}} \approx$ const $\cdot M^{-3.23}$, and $e^3_{M,\text{test}} \approx$ const $\cdot M^{-2.10}$ (the shorter dashed lines in the plot are of slope M^{-5+p}). These estimates suggest that $r_p = q_p + \omega$ where ω is somewhat larger than unity.

From Figure 4 we may also infer the approximate error degradation factors $\rho^p_{M,\text{test}}$ for $p = 1, 2, 3$ as functions of M: a rough estimate is $\rho^p_{M,\text{test}} \propto M^p$ since we loose approximately a factor M when p increases by one. We note that this is exactly what we expect from Remark 1 if $r_p = q_p + 1$ and the error estimates indicated in Remark 1 are sharp.

Finally, in Figure 5, we report the Lebesgue constant Λ_M: any growth of the Lebesgue constant is hardly present. The EIM derivative approximation will thus be close to the best $L^\infty(\Omega)$ approximation in the space $W^{\mathcal{F}}_M$.

6   Concluding remarks

We have introduced a new *a priori* convergence theory for the approximation of parametric derivatives. Given a sequence of approximation spaces, we have showed that the best approximation error associated with parametric derivatives of a function will go to zero provided that the best approximation error associated with the function itself goes to zero. We have also provided estimates for the convergence rates. In practice a method such as the EIM is used for the approximation of such functions, and hence the best approximation convergence result does not directly apply. However, thanks to the slowly growing Lebesgue constant associated with the EIM approximation scheme, we expect that the EIM approximation error will be small whenever the best approximation error

is small.

A natural approach to the EIM approximation of parametric derivatives would be to either enrich the original EIM space with snapshots of these parametric derivatives or to construct separate EIM spaces for each derivative, with this derivative as the generating function. The results in this paper, however, suggest that the EIM may be invoked in practice for the approximation of parametric derivatives without enrichment of the space or construction of additional spaces.

There are admittedly several opportunities for improvements of the theory. First, our numerical results of Section 5.1 *suggest* that the theoretical bounds for parametrically analytic functions are not sharp. The theory predicts an error degradation factor upper bound M^{2p}, but the numerical results show (for this particular example function $\mathcal{F}$) a smaller error degradation factor. It is not clear with the present strategy how to sharpen the theoretical bounds. Second, we would like to extend the validity of the theory to other (e.g. Sobolev) norms; in this case we may for example consider reduced basis approximations to parametric derivatives of solutions to parametrized partial differential equations [10, 14].

Acknowledgments

A Proofs for Hypotheses 1 and 2

A.1 Piecewise linear interpolation

We consider piecewise linear interpolation over the equidistant interpolation nodes $y_{N,i} = (2i/N - 1) \in \Gamma = [-1, 1]$, $0 \leq i \leq N$. In this case the characteristic functions $\chi_{N,i}$ are continuous and piecewise linear "hat functions" with support only on the interval $[y_{N,0}, y_{N,1}]$ for $i = 0$, on $[y_{N,i-1}, y_{N,i+1}]$ for $1 \leq i \leq N - 1$, and on $[y_{N,N-1}, y_{N,N}]$ for $i = N$.

We recall the results (9) and (10) from Section 2.2. Let $f : \Gamma \to \mathbb{R}$ with $f \in C^1(\Gamma)$ and assume that $\sup_{y \in \Gamma} |f''(y)| < \infty$. We then have, for any $y \in \Gamma$ and any $N \geq 0$,

$$|f'(y) - (I_N f)'(y)| \leq 2N^{-1} \|f''\|_{L^\infty(\Gamma)}. \tag{81}$$

Further, for all $y \in \Gamma$, the characteristic functions $\chi_{N,i}$, $0 \leq i \leq N$, satisfy

$$\sum_{i=0}^{N} |\chi'_{N,i}(y)| = N. \tag{82}$$

We first demonstrate (81) (and hence (9)). For $y \in [y_{N,i}, y_{N,i+1}]$, $0 \leq i \leq N - 1$, we have

$$(I_N f)'(y) = \frac{1}{h}\big(f(y_{N,i+1}) - f(y_{N,i})\big), \tag{83}$$

where $h = 2/N$. We next write $f(y_{N,i})$ and $f(y_{N,i+1})$ as Taylor series around y as

$$f(y_{N,i}) = \sum_{j=0}^{1} \frac{f^{(j)}(y)}{j!}(y_{N,i} - y)^j + \int_y^{y_{N,i}} f''(t)(y_{N,i} - t)\, dt, \tag{84}$$

$$f(y_{N,i+1}) = \sum_{j=0}^{1} \frac{f^{(j)}(y)}{j!}(y_{N,i+1} - y)^j + \int_y^{y_{N,i+1}} f''(t)(y_{N,i+1} - t)\, dt, \tag{85}$$

which we then insert in the expression (83) for $(I_N f)'$ to obtain

$$|(I_N f)'(y) - f'(y)| = \left| \frac{1}{h}\int_y^{y_{N,i+1}} f''(t)(y_{N,i+1}-t)\, dt - \frac{1}{h}\int_y^{y_{N,i}} f''(t)(y_{N,i}-t)\, dt \right|$$

$$\leq \frac{1}{h}\|f''\|_{L^\infty(\Gamma)} \max_{y\in[y_{N,i}, y_{N,i+1}]}\big(|y_{N,i+1} - y|^2 + |y_{N,i} - y|^2\big)$$

$$\leq h\|f''\|_{L^\infty(\Gamma)} = 2N^{-1}\|f''\|_{L^\infty(\Gamma)}. \tag{86}$$

We next demonstrate (82) (and hence (10)). It suffices to consider $y \in [y_{N,i}, y_{N,i+1}]$ for $0 \leq i \leq N - 1$. On $[y_{N,i}, y_{N,i+1}]$ only $|\chi'_{N,i}(y)|$ and $|\chi'_{N,i+1}(y)|$ contribute to the sum; furthermore we have $|\chi'_{N,i}(y)| = |\chi'_{N,i+1}(y)| = 1/h = N/2$, from where the result (82) follows.

A.2 Piecewise quadratic interpolation

We consider piecewise quadratic interpolation over equidistant interpolation nodes $y_{N,i} = (2i/N - 1) \in \Gamma$, $0 \leq i \leq N$. We consider N even such that we may divide Γ into $N/2$ intervals $[y_{N,i}, y_{N,i+2}]$, for $i = 0, 2, 4, \ldots, N - 2$. The characteristic functions $\chi_{N,i}$ are for $y \in [y_{N,i}, y_{N,i+2}]$ given as

$$\chi_{N,i}(y) = \frac{(y - y_{N,i+1})(y - y_{N,i+2})}{2h^2}, \tag{87}$$

$$\chi_{N,i+1}(y) = \frac{(y - y_{N,i})(y - y_{N,i+2})}{-h^2}, \tag{88}$$

$$\chi_{N,i+2}(y) = \frac{(y - y_{N,i})(y - y_{N,i+1})}{2h^2}, \tag{89}$$

for $i = 0, 2, 4, \ldots, N$, where $h = 2/N = y_{N,j+1} - y_{N,j}$, $0 \leq j \leq N - 1$.

We recall the results (16) and (17) from Section 2.2. Let $f : \Gamma \to \mathbb{R}$ with $f \in C^2(\Gamma)$ and assume that $\sup_{y\in\Gamma}|f'''(y)| < \infty$. We then have, for any $y \in \Gamma$ and any $N \geq 0$,

$$|f'(y) - (I_N f)'(y)| \leq 28\frac{\|f'''\|_{L^\infty(\Gamma)}}{N^2}. \tag{90}$$

Further, for all $y \in \Gamma$, the characteristic functions $\chi_{N,i}$, $0 \le i \le N$, satisfy

$$\sum_{i=0}^{N} |\chi'_{N,i}(y)| = \frac{5}{2}N. \tag{91}$$

We first demonstrate (90). It suffices to consider the interpolant $I_N f(y)$ for $y \in \Gamma_i \equiv [y_{N,i}, y_{N,i+2}]$, in which case

$$I_N f(y) = f(y_{N,i})\chi_{N,i}(y) + f(y_{N,i+1})\chi_{N,i+1}(y) + f(y_{N,i+2})\chi_{N,i+2}(y). \tag{92}$$

Insertion of (87)–(89) and differentiation yields

$$(I_N f)'(y) = \frac{1}{2h^2}\Big(f(y_{N,i})(2y - y_{N,i+1} - y_{N,i+2})$$
$$- 2f(y_{N,i+1})(2y - y_{N,i} - y_{N,i+2}) + f(y_{N,i+2})(2y - y_{N,i} - y_{N,i+1})\Big). \tag{93}$$

We next write $f(y_{N,i})$, $f(y_{N,i+1})$, and $f(y_{N,i+2})$ as Taylor series around y as

$$f(y_{N,i}) = \sum_{j=0}^{2} \frac{f^{(j)}(y)}{j!}(y_{N,i} - y)^j + \int_{y}^{y_{N,i}} f'''(t)\frac{(y_{N,i} - t)^2}{2}\, dt, \tag{94}$$

$$f(y_{N,i+1}) = \sum_{j=0}^{2} \frac{f^{(j)}(y)}{j!}(y_{N,i+1} - y)^j + \int_{y}^{y_{N,i+1}} f'''(t)\frac{(y_{N,i+1} - t)^2}{2}\, dt, \tag{95}$$

$$f(y_{N,i+2}) = \sum_{j=0}^{2} \frac{f^{(j)}(y)}{j!}(y_{N,i+2} - y)^j + \int_{y}^{y_{N,i+2}} f'''(t)\frac{(y_{N,i+2} - t)^2}{2}\, dt. \tag{96}$$

We may then insert the expressions (94)–(96) into (93) to obtain

$$(I_N f)'(y) - f'(y) = \frac{1}{2h^2}\Bigg((2y - y_{N,i+1} - y_{N,i+2})\int_{y}^{y_{N,i}} f'''(t)\frac{(y_{N,i} - t)^2}{2}\, dt$$

$$- 2(2y - y_{N,i} - y_{N,i+2})\int_{y}^{y_{N,i+1}} f'''(t)\frac{(y_{N,i+1} - t)^2}{2}\, dt$$

$$+ (2y - y_{N,i} - y_{N,i+1})\int_{y}^{y_{N,i+2}} f'''(t)\frac{(y_{N,i+2} - t)^2}{2}\, dt \Bigg). \tag{97}$$

(For $j = 0$ and $j = 2$ the terms on the right-hand-side of (93) cancel, and for $j = 1$ we obtain $f'(y)$.) We further bound (97) as

$$|(I_N f)'(y) - f'(y)| \le \frac{\|f'''\|_{L^\infty(\Gamma)}}{4h^2} \max_{y \in \Gamma_i}\Big(|2y - y_{N,i+1} - y_{N,i+2}||y_{N,i} - y|^3$$

$$+ 2|2y - y_{N,i} - y_{N,i+2}||y_{N,i+1} - y|^3 + |2y - y_{N,i} - y_{N,i+1}||y_{N,i+2} - y|^3 \Big)$$

$$\le \frac{\|f'''\|_{L^\infty(\Gamma)}}{4h} \max_{y \in \Gamma_i}\Big(3|y_{N,i} - y|^3 + 4|y_{N,i+1} - y|^3 + 3|y_{N,i+2} - y|^3 \Big)$$

$$\le \frac{\|f'''\|_{L^\infty(\Gamma)}}{4h}(3(2h)^3 + 4h^3) = 28\frac{\|f'''\|_{L^\infty(\Gamma)}}{N^2}, \tag{98}$$

23

which is the desired result.

We next demonstrate (91). It again suffices to consider $y \in \Gamma_i$. On Γ_i only $\chi'_{N,i}(y)$, $\chi'_{N,i+1}(y)$, and $\chi'_{N,i+2}(y)$ contribute to the sum. With $h = 2/N = y_{j+1} - y_j$, $0 \le j \le N - 1$, we have

$$\max_{y \in \Gamma_i} |\chi'_{N,i}(y)| = \frac{N^2}{8} \max_{y \in \Gamma_i} |2y - y_{N,i+1} - y_{N,i+2}| = \frac{3}{4}N, \tag{99}$$

$$\max_{y \in \Gamma_i} |\chi'_{N,i+1}(y)| = \frac{N^2}{4} \max_{y \in \Gamma_i} |2y - y_{N,i} - y_{N,i+2}| = N, \tag{100}$$

$$\max_{y \in \Gamma_i} |\chi'_{N,i+2}(y)| = \frac{N^2}{8} \max_{y \in \Gamma_i} |2y - y_{N,i} - y_{N,i+1}| = \frac{3}{4}N. \tag{101}$$

The result then follows.

A.3 Proof that $\mathcal{W}(\xi) < \log(\xi)$ for real $\xi > e$

We recall the definition of the LambertW function

$$\xi = \mathcal{W}(\xi)e^{\mathcal{W}(\xi)}, \quad \xi \in \mathbb{C}. \tag{102}$$

By implicit differentiation we obtain

$$\mathcal{W}'(\xi) = \frac{1}{e^{\mathcal{W}(\xi)} + \xi} \tag{103}$$

for $\xi \ne -1/e$. Further, $\mathcal{W}(\xi)$ is real-valued for real-valued $\xi > 0$. Hence

$$\mathcal{W}'(\xi) < \frac{1}{\xi} \tag{104}$$

for real $\xi > 0$. We then make the observation that

$$\mathcal{W}(e) = \log(e) = 1. \tag{105}$$

Hence for all $\xi > e$, we have $\mathcal{W}(\xi) < \log(\xi)$.

References

[1] M. Barrault, N. C. Nguyen, Y. Maday, and A. T. Patera. An "empirical interpolation" method: Application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique*, 339:667–672, 2004.

[2] C. Bernardi and Y. Maday. Spectral methods. In *Handbook of numerical analysis, Vol. V*, Handb. Numer. Anal., V, pages 209–485. North-Holland, Amsterdam, 1997.

[3] J. Borggaard and J. Burns. A PDE sensitivity equation method for optimal aerodynamic design. *J. Comput. Phys.*, 136(2):366–384, 1997.

[4] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral methods*. Scientific Computation. Springer-Verlag, Berlin, 2006. Fundamentals in single domains.

[5] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth. On the Lambert W function. *Advances in Computational Mathematics*, 5:329–359, 1996.

[6] J. L. Eftang, M. A. Grepl, and A. T. Patera. A posteriori error bounds for the empirical interpolation method. *Comptes Rendus Mathematique*, 348(9-10):575 – 579, 2010.

[7] M. Grepl. *A Posteriori* Error Bounds for Reduced-Basis Approximations of Nonaffine and Nonlinear Parabolic Partial Differential Equations. *Mathematical Models and Methods in Applied Sciences*, 22(3), 2012.

[8] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera. Efficient reduced basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM: M2AN*, 41(3):575–605, 2007.

[9] Y. Maday, N. C. Nguyen, A. T. Patera, and G. S. H. Pau. A general multipurpose interpolation procedure: The magic points. *Communications in pure and applied mathematics*, 8:383–404, 2009.

[10] I. Oliveira and A. Patera. Reduced-basis techniques for rapid reliable optimization of systems described by affinely parametrized coercive elliptic partial differential equations. *Optimization and Engineering*, 8:43–65, 2007. 10.1007/s11081-007-9002-6.

[11] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathematics*, volume 37 of *Texts Appl. Math.* Springer, New York, 1991.

[12] S. C. Reddy and J. A. C. Weideman. The accuracy of the Chebyshev differencing method for analytic functions. *SIAM J. Numer. Anal.*, 42(5):2176–2187 (electronic), 2005.

[13] T. J. Rivlin. *The Chebyshev polynomials*. Wiley-Interscience [John Wiley & Sons], New York, 1974. Pure and Applied Mathematics.

[14] G. Rozza, D. B. P. Huynh, and A. T. Patera. Reduced Basis Approximation and *a posteriori* Error Estimation for Affinely Parametrized Elliptic Coercive Partial Differential Equations. *Archives of Computational Methods in Engineering*, 15(3):229–275, 2008.

[15] E. Tadmor. The exponential accuracy of Fourier and Chebyshev differencing methods. *SIAM J. Numer. Anal.*, 23(1):1–10, 1986.