C. R. Acad. Sci. Paris, Ser. I 355 (2017) 1161–1167

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Partial differential equations/Numerical analysis

An LP empirical quadrature procedure for parametrized functions

Une procédure de quadrature empirique par programmation linéaire pour les fonctions à paramètres

Anthony T. Patera [a], Masayuki Yano [b]

[a] *Room 3-266, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*
[b] *University of Toronto, 4925 Duffein Street, Toronto, ON, M3H 5T6, Canada*

ARTICLE INFO

Article history:
Received 27 August 2017
Accepted after revision 30 October 2017
Available online 10 November 2017

Presented by Olivier Pironneau

ABSTRACT

We extend the linear program empirical quadrature procedure proposed in [9] and subsequently [3] to the case in which the functions to be integrated are associated with a parametric manifold. We pose a discretized linear semi-infinite program: we minimize as objective the sum of the (positive) quadrature weights, an ℓ_1 norm that yields sparse solutions and furthermore ensures stability; we require as inequality constraints that the integrals of J functions sampled from the parametric manifold are evaluated to accuracy $\bar{\delta}$. We provide an *a priori* error estimate and numerical results that demonstrate that under suitable regularity conditions, the integral of any function from the parametric manifold is evaluated by the empirical quadrature rule to accuracy $\bar{\delta}$ as $J \to \infty$. We present two numerical examples: an inverse Laplace transform; reduced-basis treatment of a nonlinear partial differential equation.

RÉSUMÉ

Nous étendons la procédure de quadrature empirique par programmation linéaire proposée dans [9] et par la suite dans [3] au cas où les fonctions à intégrer sont associées à une variété paramétrique. Nous posons un problème de programmation linéaire discret et semi-infini : nous minimisons la fonction objectif, qui est la somme des poids (positifs) de quadrature, qui constitue une norme ℓ_1 menant à des solutions parcimonieuses et assurant la stabilité, les contraintes d'inégalité requises étant que les intégrales de J fonctions échantillonnées à partir de la variété soient évaluées à une précision $\bar{\delta}$. Nous fournissons un estimateur d'erreur a priori et des résultats numériques qui démontrent que, sous certaines conditions de régularité, toute fonction de la variété est évaluée par la méthode de quadrature empirique avec précision $\bar{\delta}$ quand $J \to \infty$. Nous présentons deux exemples nu-

E-mail addresses: patera@mit.edu (A.T. Patera), myano@utias.utoronto.ca (M. Yano).

https://doi.org/10.1016/j.crma.2017.10.020

mériques : une transformée inverse de Laplace et un traitement par base réduite d'une équation aux dérivées partielles non linéaire.

1. Introduction

In this note, we consider the integration of parameterized functions,

$$I(\mu) = \int_{\Omega} g(\mu;\xi)\,\mathrm{d}\xi\ , \tag{1}$$

for μ in the parameter domain $\mathcal{D} \subset \mathbb{R}^P$, $\Omega \subset \mathbb{R}^d$, and $g \in L^\infty(\mathcal{D}; L^\infty(\Omega))$. (We note that although Ω is parameter-independent, this spatial domain may be the result of a transformation from a parameter-dependent spatial domain through standard change-of-variable techniques.) Parameterized integrals arise in a variety of applications, from transform methods for ordinary differential equations, in which (say) ξ is frequency and μ includes time, to variational approximation of partial differential equations, in which ξ is a spatial coordinate and μ includes constitutive constants, sources, and geometric transformations.

We are interested in particular in the many-query context, in which $\mu \in \mathcal{D} \mapsto I(\mu)$ must be performed many times, often in real-time, for different values of μ in $\mathcal{D}$. We may thus gainfully consider offline–online approaches: an empirical quadrature rule — points and weights particularly optimized for (1) — is developed, once, in a relatively expensive offline stage; this efficient quadrature rule is then invoked, many times, in a very inexpensive online stage. The effort of the offline stage is justified, in fact amortized, over the many parameter queries $\mu \in \mathcal{D} \mapsto I(\mu)$ of the online stage.

One approach to (1) is interpolation-then-integration: we develop an interpolant for $g(\mu;\cdot)$ which then serves as a surrogate for $g(\mu;\xi)$ in (1); as an example of interpolation schemes for parametric functions, we cite the Empirical Interpolation Method [2]. Although interpolation-then-integration can be quite effective in practice, in fact the objectives and metrics associated with interpolation and integration are quite different, and thus a more direct approach — empirical quadrature rather than empirical interpolation — is also of interest.

An empirical quadrature procedure for parameterized functions is developed in [1] and further extended in [6]. These approaches consider an ℓ_2 framework and thus sparsity must be introduced explicitly, either through a heuristic sequential point selection process (as in [1]) or through an approximate ℓ_0 optimization (as in [6]); in both cases, a somewhat challenging non-negative least-squares problem must be addressed. In the current paper, we propose an ℓ_1 framework: a stronger norm which naturally yields sparse designs and which furthermore can be cast as a linear program (LP) efficiently treated by the dual simplex method. Our approach is an extension to the parametric context of the LP quadrature framework first proposed in [9] and further developed in [3].

2. Formulation

We define a parameter domain $\mathcal{D} \subset \mathbb{R}^P$, a point in which will be denoted $\mu = (\mu_1, \ldots, \mu_P)$, and an integration domain $\Omega \subset \mathbb{R}^d$, a point in which will be denoted $\xi = (\xi_1, \ldots, \xi_d)$. We then introduce a set of parameterized functions $g_m : \mathcal{D} \times \Omega \to \mathbb{R}$, $\forall m \in \mathbb{M}$; here $\mathbb{M} \equiv \{1, \ldots, M\}$ for M a finite positive integer. We shall assume that our set of functions satisfies a Lipschitz condition,

$$\sup_{m\in\mathbb{M}} \sup_{\mu',\mu''\in\mathcal{D}^2} \|g_m(\mu';\cdot) - g_m(\mu'';\cdot)\|_{L^\infty(\Omega)} \le L_g \|\mu' - \mu''\|_2\ , \tag{2}$$

for L_g a finite constant and $\|z\|_2$ the usual Euclidean norm (here) for $z \in \mathbb{R}^P$.

We next define the set of integrals of interest:

$$I_m(\mu) = \int_{\Omega} g_m(\mu;\xi)\,\mathrm{d}\xi\ , \forall m \in \mathbb{M}, \forall \mu \in \mathcal{D}\,. \tag{3}$$

We shall also require a "truth" quadrature,

$$I_m^{\text{truth}}(\mu) = \sum_{i=1}^{\mathcal{N}} w_i^{\text{truth}} g_m(\mu;\xi_i^{\text{truth}}), \forall m \in \mathbb{M}, \forall \mu \in \mathcal{D}\ , \tag{4}$$

where $\{w_i^{\text{truth}}\}_{i=1,\ldots,\mathcal{N}}$ and $\{\xi_i^{\text{truth}}\}_{i=1,\ldots,\mathcal{N}}$ are the truth (non-negative) quadrature weights and truth quadrature points, respectively. We shall assume that (a) for ϵ a prescribed error tolerance,

$$|I_m(\mu) - I_m^{\text{truth}}(\mu)| \le \epsilon/2, \forall m \in \mathbb{M}, \forall \mu \in \mathcal{D}\,, \tag{5}$$

and (b) the truth quadrature rule integrates exactly the constant function so that

$$\sum_{i=1}^{\mathcal{N}} w_i^{\text{truth}} = |\Omega|\,, \tag{6}$$

where $|\Omega|$ denotes the measure of the domain of integration Ω.

We now search for an empirical quadrature rule, points $\{\xi_k^\nu\}_{k=1,\ldots,K^\nu}$, and associated non-negative weights $\{w_k^\nu\}_{k=1,\ldots,K^\nu}$, in terms of which we approximate our integrals as

$$I_m^\nu(\mu) = \sum_{k=1}^{K^\nu} w_k^\nu g_m(\mu; \xi_k^\nu), \forall m \in \mathbb{M}, \forall \mu \in \mathcal{D}\,; \tag{7}$$

the "hyperparameter" ν characterizes the procedure by which the empirical quadrature rule is derived. (Note that "hyperparameter" here refers to hyperreduction and, in particular, distinguishes the parameter ν that determines the quadrature rule from the parameter μ that defines the parameterized functions of interest.) We wish to find an empirical quadrature rule that is efficient, $K^\nu \ll \mathcal{N}$, and accurate,

$$|I_m^{\text{truth}}(\mu) - I_m^\nu(\mu)| \le \bar{\delta}, \forall m \in \mathbb{M}, \forall \mu \in \mathcal{D}\,; \tag{8}$$

we shall typically choose $\bar{\delta} = \epsilon/2$ such that, from (5) and (8), $|I_m(\mu) - I_m^\nu(\mu)| \le \epsilon$, $\forall m \in \mathbb{M}$, $\forall \mu \in \mathcal{D}$.

We consider an offline–online strategy: in the offline stage, the empirical quadrature rule, points and weights, is identified; in the online stage, the empirical quadrature rule is "queried,"

$$\mu \in \mathcal{D} \mapsto \{I_m^\nu(\mu)\}_{m\in\mathbb{M}}\,. \tag{9}$$

The offline stage will typically be expensive, but can be justified by a premium on real-time response or alternatively amortized over many parameter queries (9). In contrast, the online stage, (9) evaluated as (7), is inexpensive (under the assumption that $K^\nu \ll \mathcal{N}$): operation count $\mathcal{O}(K^\nu M)$. In the remainder of this paper, we focus on the offline stage.

We first specify $\bar{\delta} \in \mathbb{R}_+$. We next define $\mathbb{J} \equiv \{1, \ldots, J\}$ for $J \in \mathbb{N}_+$ and provide a parameter training sample $\Xi_J^{\text{train}} \equiv \{\mu_j^{\text{train}} \in \mathcal{D}\}_{j\in\mathbb{J}}$ and an associated set of snapshots on the parametric manifold, $\{\phi_{m,j} \equiv g_m(\mu_j^{\text{train}}; \cdot)\}_{j\in\mathbb{J}, m\in\mathbb{M}}$. We may then define our hyperparameter $\nu = \{\bar{\delta}, J, \Xi_J^{\text{train}}\}$ and pose a linear program $\text{LP}_{\text{quad}}^\nu$: find a basic feasible vector $\rho_{\text{opt}}^\nu \in \mathbb{R}^{\mathcal{N}}$ that minimizes

$$\sum_{i=1}^{\mathcal{N}} \rho_i \tag{10}$$

subject to

$$\rho_i \ge 0, 1 \le i \le \mathcal{N}\,, \tag{11}$$

and the MJ "accuracy" (inequality) constraints

$$\left| \sum_{i=1}^{\mathcal{N}} w_i^{\text{truth}} \phi_{m,j}(\xi_i^{\text{truth}}) - \sum_{i=1}^{\mathcal{N}} \rho_i \phi_{m,j}(\xi_i^{\text{truth}}) \right| \le \bar{\delta}, \ \forall j \in \mathbb{J}, \ \forall m \in \mathbb{M}\,. \tag{12}$$

We know that $\text{LP}_{\text{quad}}^\nu$ is feasible: the truth quadrature rule satisfies (11) and (12). However, we search in particular for a solution of $\text{LP}_{\text{quad}}^\nu$ that is *basic feasible* with respect to the constraints (11)–(12); we shall ensure the latter condition by application of the (dual) simplex method. (Note that in practice (12) is unfolded into two one-sided constraints in order to cast the linear program in standard form.) We then identify the indices associated with non-zero values of ρ_{opt}^ν as i_k^ν, $1 \le k \le K^\nu$, and set $\xi_k^\nu = \xi_{i_k^\nu}^{\text{truth}}$, $w_k^\nu = (\rho_{\text{opt}}^\nu)_{i_k^\nu}$, $1 \le k \le K^\nu$. (We may also consider, under suitable hypotheses, a relative version of the constraint: in that case, $\bar{\delta}$ is multiplied by $|I_m^{\text{truth}}(\mu_j^{\text{train}})|$; we exercise this variant in our second example of Section 3.)

Our optimization problem $\text{LP}_{\text{quad}}^\nu$ is a discretization of a linear semi-infinite program with respect to the parametric manifold. We wish to integrate to accuracy $\bar{\delta}$ any function on the parametric manifold $\{g_m(\mu, \cdot) \,|\, \mu \in \mathcal{D}\}_{m\in\mathbb{M}}$; in that regard, we observe that our constraint (12) considers a finite sample of J functions, and it is thus plausible for a smooth manifold and J sufficiently large that we will indeed realize the desired accuracy. However, we also wish to obtain a solution of $\text{LP}_{\text{quad}}^\nu$ that corresponds to a sparse quadrature rule such that $K^\nu \ll \mathcal{N}$ (in fact, $K^\nu \ll \min\{MJ, \mathcal{N}\}$), even as we increase J and refine our discretization; in that regard, we observe that our objective function (10) is the ℓ_1 norm of ρ, and thus it is plausible that sparse (basic feasible) solutions will indeed exist.

We now compare $\text{LP}_{\text{quad}}^\nu$ to the approach proposed in [3]. In fact, both [3] and $\text{LP}_{\text{quad}}^\nu$ invoke the ℓ_1-norm objective function. However, [3] treats constraints in a fundamentally different fashion from (11)–(12):

(*i*) [3] considers a finite-dimensional linear space, W, rather than a parametric manifold;
(*ii*) [3] imposes equality constraints (such that any member of W, as represented in a basis, is integrated exactly) rather than inequality constraints.

We contend that the $\mathrm{LP}^{\nu}_{\mathrm{quad}}$ choices in (*i*) and (*ii*) are related: we prefer direct consideration of the parametric manifold to permit more explicit control of the error; inequality constraints are then required to preserve sparsity in the limit of refinement. To illustrate the latter, we consider $\bar{\delta} = 0$ (equivalent to equality constraints) in (12): a solution exists with $K^{\nu} = MJ$ [7,5], and furthermore, for independent constraints, (generically) no solution exists with $K^{\nu} < MJ$ [4] — clearly not sparse, and hence unacceptable, for J large; in contrast, with inequality constraints, $\bar{\delta} > 0$, we can obtain $K^{\nu} \ll \min\{MJ, \mathcal{N}\}$, and in particular K^{ν} will equal, not the number of constraints, but rather the number of *active* constraints. (Of course from the latter, we understand that we could in principle impose equality constraints at a few well-selected parameter values; however the identification of these parameter values — or alternatively, a good linear approximation space — perforce entails a search over a much larger set of parameter values.) Inequality constraints are also well motivated from the perspective of applications: the tolerance $\bar{\delta}$ is selected consistently with the desired accuracy, $\bar{\delta} = \epsilon/2$.

Although we do not attempt to identify optimal parameter values for our constraint set, we do nevertheless anticipate that, for smooth manifolds, $\max_{m \in \mathbb{M}} \max_{\mu \in \mathcal{D}} |I^{\mathrm{truth}}_m(\mu) - I^{\nu}_m(\mu)|$ will tend to $\bar{\delta}$ rapidly as J increases: a small sample should suffice. Even for limited smoothness, we can demonstrate that the error should decrease as $1/J$. To begin, we provide the following general result.

Lemma 2.1. *For any $\mu \in \mathcal{D}$, and given hyperparameter ν,*

$$\max_{m \in \mathbb{M}} |I^{\mathrm{truth}}_m(\mu) - I^{\nu}_m(\mu)| \le \max_{m \in \mathbb{M}} \left(\inf_{\alpha \in \mathbb{R}^J} \left(\bar{\delta} \sum_{j=1}^{J} |\alpha_j| + 2|\Omega| \, \| g_m(\mu; \cdot) - \sum_{j=1}^{J} \alpha_j \phi_{m,j} \|_{L^\infty(\Omega)} \right) \right). \tag{13}$$

Proof. We first fix $m \in \mathbb{M}$. Then, for any $\alpha \in \mathbb{R}^J$,

$$\begin{aligned}
| I^{\mathrm{truth}}_m(\mu) - I^{\nu}_m(\mu)| &= \Big| \sum_{i=1}^{\mathcal{N}} w^{\mathrm{truth}}_i g_m(\mu; \xi^{\mathrm{truth}}_i) - \sum_{k=1}^{K^\nu} w^{\nu}_k g_m(\mu; \xi^{\nu}_k) \Big| \\
&\le \Big| \sum_{i=1}^{\mathcal{N}} w^{\mathrm{truth}}_i \sum_{j=1}^{J} \alpha_j \phi_{m,j}(\xi^{\mathrm{truth}}_i) - \sum_{k=1}^{K^\nu} w^{\nu}_k \sum_{j=1}^{J} \alpha_j \phi_{m,j}(\xi^{\nu}_k) \Big| \\
&\quad + \Big| \sum_{i=1}^{\mathcal{N}} w^{\mathrm{truth}}_i \Big(g_m(\mu; \xi^{\mathrm{truth}}_i) - \sum_{j=1}^{J} \alpha_j \phi_{m,j}(\xi^{\mathrm{truth}}_i) \Big) \Big| + \Big| \sum_{k=1}^{K^\nu} w^{\nu}_k \Big(g_m(\mu; \xi^{\nu}_k) - \sum_{j=1}^{J} \alpha_j \phi_{m,j}(\xi^{\nu}_k) \Big) \Big| \\
&\le \Big| \sum_{j=1}^{J} \alpha_j \Big(\sum_{i=1}^{\mathcal{N}} w^{\mathrm{truth}}_i \phi_{m,j}(\xi^{\mathrm{truth}}_i) - \sum_{k=1}^{K^\nu} w^{\nu}_k \phi_{m,j}(\xi^{\nu}_k) \Big) \Big| \\
&\quad + \Big| \sum_{i=1}^{\mathcal{N}} w^{\mathrm{truth}}_i \Big(g_m(\mu; \xi^{\mathrm{truth}}_i) - \sum_{j=1}^{J} \alpha_j \phi_{m,j}(\xi^{\mathrm{truth}}_i) \Big) \Big| + \Big| \sum_{k=1}^{K^\nu} w^{\nu}_k \Big(g_m(\mu; \xi^{\nu}_k) - \sum_{j=1}^{J} \alpha_j \phi_{m,j}(\xi^{\nu}_k) \Big) \Big| .
\end{aligned} \tag{14}$$

For the first term, we can now invoke (12) and apply Hölder's inequality ($p = 1, q = \infty$). For the second term, we appeal to (6) and, again, to Hölder's inequality. Finally, for the third term, we recall that the truth quadrature is in fact feasible; we may then invoke optimality — here equivalent to stability — to conclude that $\sum_{k=1}^{K} w^{\nu}_k \le \sum_{i=1}^{\mathcal{N}} w^{\mathrm{truth}}_i = |\Omega|$, and again apply Hölder's inequality. The result directly follows. □

We note that (13) of Lemma 2.1 quantifies how (first term) stably we can (second term) approximate, for any $m \in \mathbb{M}$ and any $\mu \in \mathcal{D}$, the function $g_m(\mu; \cdot)$ — and hence $I^{\mathrm{truth}}_m(\mu)$ — in terms of our snapshots $\{\phi_{m,j}\}_{j \in \mathbb{J}, m \in \mathbb{M}}$.

As presented, Lemma 2.1 is not actionable. We can, however, now choose an interpolation system and then quantify the error in terms of associated regularity estimates. Most simply, we can demonstrate the following theorem.

Theorem 2.2. *Let*

$$\Delta^{\nu} \equiv \max_{\mu \in \mathcal{D}} \left(\min_{j \in \mathbb{J}} \| \mu - \mu^{\mathrm{train}}_j \|_2 \right). \tag{15}$$

Then, under the hypothesis of Lipschitz continuity, (2), for any $\mu \in \mathcal{D}$,

$$\max_{m \in \mathbb{M}} |I^{\mathrm{truth}}_m(\mu) - I^{\nu}_m(\mu)| \le \bar{\delta} + 2|\Omega| L_g \, \Delta^{\nu} . \tag{16}$$

Proof. We choose in (13) not the best $\alpha \in \mathbb{R}^J$, but rather the sub-optimal coefficients

$$\alpha_j^* = \begin{cases} 1 & \text{for } j = j^* \\ 0 & \text{for } j \in \mathbb{J} \setminus j^* \end{cases} \tag{17}$$

for $j^* = \arg\min_{j \in \mathbb{J}} \|\mu - \mu_j^{\text{train}}\|_2$. The result then directly follows from Lemma 2.1, (17), (2), and (15). □

We can further introduce a sampling hypothesis: $\Delta^\nu \to 0$ as $J \to \infty$; it then follows from Theorem 2.2 that for J sufficiently large $I_m^\nu(\mu)$ approximates $I_m^{\text{truth}}(\mu)$ to within the prescribed tolerance $\bar{\delta}$ for any $m \in \mathbb{M}$ and any $\mu \in \mathcal{D}$. Note here that ν depends on J; in actual practice, the quadrature scheme changes very little for sufficiently large J.

3. Examples

We consider two examples: one relates to the Fourier transform, the other to the reduced-basis method. All computations of $\text{LP}_{\text{quad}}^\nu$ are performed with the Matlab implementation of the dual simplex method on commodity laptops. Note we confirm in (selected) numerical tests that the number of quadrature points does indeed equal the number of active accuracy constraints.

We discuss in these examples both the offline stage, and in particular the generation of the quadrature rule, and then the online stage, which corresponds to the evaluation of the quadrature rule for particular instances. We emphasize that, in the online stage, we must evaluate $\mu \in \mathcal{D} \to g_m(\mu; \xi_k^\nu)$, $1 \le k \le K^\nu$, $\forall m \in \mathbb{M}$, which corresponds to (effectively) MK^ν function (g) evaluations. (A similar operation count would apply to integration-by-interpolation techniques, since function evaluations are required to determine the interpolation coefficients.) We do not discuss here the acceleration of the function evaluations; rather, empirical integration focuses on reduction of the *number* of function evaluations. In particular, we note that the savings (factor) associated with sparse quadrature relative to truth quadrature is independent of the cost of the individual function evaluation.

3.1. Fourier transform

We consider a function $f(\alpha, t)$ for α a real scalar parameter and t time, and the associated Laplace transform $\hat{f}(\alpha; s)$; we assume that all poles of $\hat{f}(\alpha; s)$ reside in the left-hand s-plane. The inverse Laplace transform relates f and $\hat{f}$ as

$$f(\alpha, t) = \int_0^\infty \frac{1}{\pi} \Re\Big(\exp(\mathrm{i}\omega t)\hat{f}(\alpha; \mathrm{i}\omega)\Big)\, \mathrm{d}\omega\,, \tag{18}$$

where $\mathrm{i} \equiv \sqrt{-1}$ and $\Re$ denotes the real part. We now map to our framework: $\xi \equiv \omega$, $\Omega \equiv [0, \infty)$, $\mu \equiv (\mu_1, \mu_2) \equiv (\alpha, t) \in \mathcal{P} \subset \mathbb{R}^{P=2}$ such that $f(\alpha, t) = I(\mu)$ of (1) for

$$g(\mu; \xi) \equiv \frac{1}{\pi} \Re\Big(\exp(\mathrm{i}\xi\mu_2)\hat{f}(\mu_1; \mathrm{i}\xi)\Big)\,. \tag{19}$$

We now introduce the truth quadrature by which we calculate $I^{\text{truth}}(\mu)$: the trapezoidal rule over $\mathcal{N}$ equi-spaced points on the interval $[0, \xi^{\max}]$. (In this example, $M = 1$, and hence we omit the subscript $m = 1$.) Note that in this case the truth quadrature introduces errors due to discretization, reflected in $\mathcal{N}$, as well as truncation, reflected in $\xi^{\max}$.

In actual practice, $I^{\text{truth}}(\mu)$ would serve to estimate (unknown) $f(\alpha, t)$ from known $\hat{f}(\alpha; s)$. In our example, we take a known pair,

$$f(\alpha, t) = \mathrm{e}^{-0.002t} \sin t + t^2 \mathrm{e}^{-\alpha t}, \;\; \hat{f}(\alpha, s) = \frac{1}{(s + 0.002)^2 + 1} + \frac{2}{(s+\alpha)^3}\,; \tag{20}$$

we further specify $\alpha \in [0.2, 2.0]$ and $t \in [0, 4]$ such that $\mathcal{P} \equiv [0.2, 2.0] \times [0, 4]$. Note that, for the parameter values of interest, $f(\alpha, t)$ is order unity, and hence absolute error and relative error are roughly equivalent. Finally, we choose for our truth quadrature $\xi^{\max} = 4$ and $\mathcal{N} = 1200$: we confirm that the latter choices yield an error in $I^{\text{truth}}(\mu)$ relative to $I(\mu)$ of several percent; we may thus meaningfully compare the computational cost for $\bar{\delta}$ on the order of 0.01 (but not for $\bar{\delta} \gg 0.01$, which implicitly places a larger burden than necessary on the truth approximation).

Our goal now is to develop an efficient empirical quadrature formula that introduces errors not greater than several percent. Towards that end, we let $J = (J')^2$ and create our train parameter sample Ξ_J^{train} as a $J' \times J'$ uniform grid over $\mathcal{D}$. We may then obtain, for any given J, our empirical quadrature rule $\{\xi_k^\nu\}_{k=1,\ldots,K^\nu}$, $\{w_k^\nu\}_{k=1,\ldots,K^\nu}$, as the post-processed solution to $\text{LP}_{\text{quad}}^\nu$; computation times range from ≈ 1 s ($J = 25^2$) to ≈ 8 s ($J = 45^2$). Finally, we further create a test parameter sample Ξ^{test} of size 100^2 constructed as the tensorization of a uniform random grid of size 100 in each of the two parameter directions; we then measure our error as

Table 1
Fourier transform example: number of quadrature points, K^ν, and error over test sample, E^ν, as a function of J' for $\bar{\delta}=0.1$ (left) and $\bar{\delta}=0.01$ (right). Note that our train sample for $\mathrm{LP}^\nu_{\mathrm{quad}}$, Ξ^{train}_J, is of size $J=(J')^2$, and our test sample for E^ν, Ξ^{test}, is of size 100^2.

J'	25	30	35	40	45
K^ν	10	11	11	12	11
E^ν	0.1578	0.1321	0.1009	0.1010	0.1011

J'	25	30	35	40	45
K^ν	14	14	14	16	15
E^ν	0.0234	0.0120	0.0131	0.0101	0.0102

$$E^\nu(\Xi^{\mathrm{test}}) = \max_{\mu\in\Xi^{\mathrm{test}}} |I^{\mathrm{truth}}(\mu) - I^\nu(\mu)| . \tag{21}$$

We present our results for a range of J' in Table 1 for $\bar{\delta}=0.1$ (left) and $\bar{\delta}=0.01$ (right).

We first observe that K^ν is (i) indeed very small, (ii) relatively insensitive to J, and (iii) only modestly dependent on $\bar{\delta}$. The actual quadrature point distribution is unremarkable, though we do note the clustering observed also in [9]. We next observe, consistent with Theorem 2.2, that E^ν rapidly approaches $\bar{\delta}$ as we increase J.

3.2. *Reduced-basis approximation of a nonlinear reaction–diffusion equation*

We consider a second example, related to the reduced-basis method, to further illustrate the approach. (We refer to [8] for a review of the reduced-basis method.) We introduce a spatial domain $\Omega \equiv (-1,1)\times(0,1) \subset \mathbb{R}^2$, which is split into two subdomains $\Omega_1 \equiv (-1,0)\times(0,1)$ and $\Omega_2 \equiv (0,1)\times(0,1)$. We also introduce a parameter domain $\mathcal{D} \equiv [1,1000]^2 \subset \mathbb{R}^{P=2}$. We then consider the following parametrized nonlinear reaction–diffusion equation: given $\mu\in\mathcal{D}$, find $u(\mu)\in\mathcal{V}\equiv H^1_0(\Omega)\equiv \{v\in H^1(\Omega)\,|\, v|_{\partial\Omega}=0\}$ such that

$$r(u(\mu), v; \mu) = 0 \quad \forall v\in\mathcal{V}, \tag{22}$$

where

$$r(z,v;\mu) \equiv \int_\Omega \nabla v(\xi)\cdot\nabla z(\xi)\,\mathrm{d}\xi + \int_\Omega v(\xi) z(\xi)^3\,\mathrm{d}\xi - \sum_{i=1}^{2}\mu_i \int_{\Omega_i} v(\xi)\,\mathrm{d}\xi = 0 \quad \forall z,v\in\mathcal{V}.$$

The space $\mathcal{V}$ is endowed with the standard $H^1(\Omega)$ inner product and norm.

We next introduce a quadratic finite-element space $\mathcal{V}_h \equiv \{v\in\mathcal{V}\,|\,v|_\kappa\in\mathbb{P}^2(\kappa),\ \forall\kappa\in\mathcal{T}_h\}\subset\mathcal{V}$ over a triangulation $\mathcal{T}_h$ of Ω that comprises $16\times 8\times 2$ triangular elements. We may then state the associated finite-dimensional problem: given $\mu\in\mathcal{D}$, find $u_h(\mu)\in\mathcal{V}_h$ such that

$$r(u_h(\mu), v; \mu) = 0 \quad \forall v\in\mathcal{V}_h; \tag{23}$$

the cubic reaction term $\int_\Omega v(\xi)z(\xi)^3\,\mathrm{d}\xi$, whose integrand is a piecewise eighth-degree polynomial, is integrated exactly by a (truth) quadrature rule that consists of $\mathcal{N}=4864$ points.

We now consider a reduced-basis approximation of (23). Towards this end, we introduce a reduced-basis space $\mathcal{V}_N \equiv \mathrm{span}\{u_h(\mu)\}_{\mu\in\Xi^{\mathrm{rb}}_N}\subset\mathcal{V}_h$ associated with a snapshot parameter set $\Xi^{\mathrm{rb}}_N\subset\mathcal{D}$ of size N. We can then state our reduced-basis problem (exact quadrature): given $\mu\in\mathcal{D}$, find $u_N(\mu)\in\mathcal{V}_N$ such that

$$r(u_N(\mu), v; \mu) = 0 \quad \forall v\in\mathcal{V}_N. \tag{24}$$

We identify the hierarchical snapshot parameter sets $\Xi^{\mathrm{rb}}_{N=1}\subset\cdots\subset\Xi^{\mathrm{rb}}_{N=N_{\max}}$ and the associated reduced-basis approximation spaces $\mathcal{V}_{N=1}\subset\cdots\subset\mathcal{V}_{N=N_{\max}}$ by application of the strong greedy procedure [8] over a training set $\Xi^{\mathrm{rb,train}}\subset\mathcal{D}$, which consists of $|\Xi^{\mathrm{rb,train}}|\equiv 20^2$ points uniformly distributed over $\log(\mathcal{D})$; we require the relative $\mathcal{V}$-norm of the error to be less than 0.05 for all $\mu\in\Xi^{\mathrm{rb,train}}$, which yields $N_{\max}=7$.

We then consider an empirical quadrature approximation of (24) and introduce the following residual form:

$$r^\nu(z,v;\mu) \equiv \int_\Omega \nabla v(\xi)\cdot\nabla z(\xi)\,\mathrm{d}\xi + \sum_{k=1}^{K^\nu} w^\nu_k v(\xi^\nu_k) z(\xi^\nu_k)^3 - \sum_{i=1}^{2}\mu_i\int_{\Omega_i} v(\xi)\,\mathrm{d}\xi = 0 \quad \forall z,v\in\mathcal{V}.$$

The reduced-basis approximation associated with the reduced quadrature rule is defined as follows: given $\mu\in\mathcal{D}$, find $u^\nu_N(\mu)\in\mathcal{V}_N$ such that

$$r^\nu(u^\nu_N(\mu), v; \mu) = 0 \quad \forall v\in\mathcal{V}_N. \tag{25}$$

We next train the quadrature rule: given $\mathcal{V}_N\equiv\mathrm{span}\{u_h(\mu)\}_{\mu\in\Xi^{\mathrm{rb}}_N}$, we require the residual and Jacobian of the nonlinear reaction terms,

Table 2
Reduced-basis method example: the number of reduced-basis functions N, the number of reduced quadrature points K^ν, the maximum quadrature error for the cubic reaction term over the test set Ξ^{test}, $E^\nu(\Xi^{\text{test}})$, the maximum difference in the reduced-basis approximation using the exact quadrature and reduced quadrature, $\max_{\mu\in\Xi^{\text{test}}} \frac{\|u_N(\mu)-u_N^\nu\|_{\mathcal{V}}}{\|u_h(\mu)\|_{\mathcal{V}}}$, and the error in the reduced-basis approximation using the reduced quadrature, $\max_{\mu\in\Xi^{\text{test}}} \frac{\|u_h(\mu)-u_N^\nu(\mu)\|_{\mathcal{V}}}{\|u_h(\mu)\|_{\mathcal{V}}}$.

N	K^ν	$E^\nu(\Xi^{\text{test}})$	$\max_{\mu\in\Xi^{\text{test}}} \frac{\|u_N(\mu)-u_N^\nu\|_{\mathcal{V}}}{\|u_h(\mu)\|_{\mathcal{V}}}$	$\max_{\mu\in\Xi^{\text{test}}} \frac{\|u_h(\mu)-u_N^\nu(\mu)\|_{\mathcal{V}}}{\|u_h(\mu)\|_{\mathcal{V}}}$
1	1	1.00×10^{-4}	2.41×10^{-5}	7.79×10^{-1}
2	5	1.00×10^{-4}	2.13×10^{-4}	7.15×10^{-1}
3	15	1.00×10^{-4}	3.69×10^{-4}	3.01×10^{-1}
4	27	1.00×10^{-4}	1.38×10^{-3}	1.56×10^{-1}
5	33	1.01×10^{-4}	1.57×10^{-3}	7.87×10^{-2}
6	42	1.02×10^{-4}	2.39×10^{-3}	5.74×10^{-2}
7	42	1.01×10^{-4}	4.00×10^{-3}	3.63×10^{-2}

$$g_{m_{\text{res}}(\lambda')}(\mu;\xi) = u(\lambda';\xi)(u_N(\mu;\xi))^3 \quad \forall\lambda'\in\Xi_N^{\text{rb}},$$
$$g_{m_{\text{Jac}}(\lambda',\lambda'')}(\mu;\xi) = u(\lambda';\xi)(u_N(\mu;\xi))^2 u(\lambda'';\xi) \quad \forall\lambda',\lambda''\in\Xi_N^{\text{rb}},$$

to be integrated to within accuracy $\bar\delta \equiv 10^{-4}$ in the relative sense for all $\mu\in\Xi_J^{\text{train}} \equiv \Xi^{\text{rb,train}}$; the formulation corresponds to $M = N + N^2$ and $J = |\Xi_J^{\text{train}}| = |\Xi^{\text{rb,train}}|$, the latter of which we determine to be sufficiently large to ensure a relative error close to $\bar\delta$ for any $\mu\in\mathcal{D}$. We note that the quadrature points are not hierarchical: we obtain different points for each N. In the online stage, the evaluation of the residual and Jacobian of the cubic reaction term requires $\mathcal{O}(K^\nu N)$ storage and $\mathcal{O}(K^\nu N^2)$ operations.

We summarize in Table 2 the result of solving the nonlinear reaction–diffusion equation using the reduced-basis method with the empirical quadrature rule for $\bar\delta = 10^{-4}$. We first observe that, in general, K^ν increases with N because the number of functions we wish to integrate exactly increases as $M = N + N^2$; nevertheless, we observe $K^\nu \ll \mathcal{N} = 4864$ for each N. We now introduce a set Ξ^{test} that consists of 1000 uniformly distributed random points over $\log(\mathcal{D})$. We first observe that the maximum quadrature error for both the residual and Jacobian of the cubic reaction term, $E^\nu(\Xi^{\text{test}})$, is close to $\bar\delta = 10^{-4}$. We next observe (the fourth column of Table 2) that the maximum difference between the reduced-basis approximations $u_N(\mu)$ and $u_N^\nu(\mu)$, which use the exact and empirical quadrature rules, respectively, is also small: an order of magnitude smaller than the error in the reduced-basis approximation itself (the fifth column of Table 2).

We finally remark on the cost. The offline computational times for $\text{LP}^\nu_{\text{quad}}$ with $\mathcal{N} = 4864$ ranges from ≈ 1 s ($N = 1$, $MJ = 800$) to ≈ 70 s ($N = 7$, $MJ = 22400$). In the online stage, rapid and accurate solution of the parameterized nonlinear reaction–diffusion equation is effected in $\mathcal{O}(NK^\nu)$ storage and $\mathcal{O}(N^2K^\nu)$ operations for formation of the linear system, and $\mathcal{O}(N^3)$ operations for solution of the linear system. We emphasize that the online storage and operation count are independent of both the dimension of the finite-element space $\dim(\mathcal{V}_h)$ and the number of truth quadrature points $\mathcal{N}$.

Acknowledgements

We thank Professor Albert Cohen of the 'Université Pierre-et-Marie-Curie' for valuable comments and suggestions. This work was supported by ONR Grant N00014-17-1-2077 (AT Patera) and NSERC Discovery Grant RGPIN-2017-06740 (M Yano).

References

[1] S.S. An, T. Kim, D.L. James, Optimizing cubature for efficient integration of subspace deformations, ACM Trans. Graph. 27 (5) (2008) 165.
[2] M. Barrault, Y. Maday, C.N. Nguyen, A.T. Patera, An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations, C. R. Acad. Sci. Paris, Ser. I 339 (2004) 667–672.
[3] R. DeVore, S. Foucart, G. Petrova, P. Wojtaszczyk, Computing a quantity of interest from observational data, preprint, 2017.
[4] C. Dossal, A necessary and sufficient condition for exact sparse recovery by ℓ_1 minimization, C. R. Acad. Sci. Paris, Ser. I 350 (1–2) (2012) 117–120.
[5] M. Elad, Sparse and Redundant Representations, Springer, 2010.
[6] C. Farhat, T. Chapman, P. Avery, Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models, Int. J. Numer. Methods Eng. 102 (2015) 1077–1110.
[7] S. Foucart, H. Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhäuser, 2013.
[8] G. Rozza, D.B.P. Huynh, A.T. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations; application to transport and continuum mechanics, Arch. Comput. Methods Eng. 15 (2008) 229–275.
[9] E.K. Ryu, S.P. Boyd, Extensions of Gauss quadrature via linear programming, Found. Comput. Math. 15 (2015) 953–971.