INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2006; **00**:1–6 *Prepared using* nmeauth.cls *[Version: 2002/09/18 v2.02]*

Reduced basis approximation and *a posteriori* error estimation for stress intensity factors

D. B. P. Huynh[1] and A. T. Patera[2*†]

[1]*National University of Singapore, Singapore-MIT Alliance, Singapore*
[2]*Massachusetts Institute of Technology, U.S.A*

SUMMARY

We develop reduced basis approximations, rigorous *a posteriori* error bounds, and offline-online computational procedures for the accurate, fast, and reliable prediction of stress intensity factors or strain energy release rates for "Mode I" linear elastic crack problems relevant to fracture mechanics. We demonstrate the efficiency and rigor of our numerical method in several examples. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS: reduced basis approximation; *a posteriori* error estimation; offline-online procedures; fracture mechanics; stress intensity factor; energy release rate, linear elasticity; finite element method; quadratic outputs; compliance

1. INTRODUCTION

The stress intensity factor [22] is one of the most important quantities in Fracture Mechanics: it summarizes the stress, strain, and displacement fields in the near crack tip region. Stress intensity factors are crucial in the calculation of fatigue-induced crack growth [22] and in the prediction of (effectively) brittle failure [1]. In many calculations, the stress intensity factor (SIF) must be available in real–time (for example, for in-the-field Non–Destructive Evaluation and prognosis) or repeatedly (for example, in fatigue crack growth prediction); and most importantly, the SIF must be accurate and reliable in particular in situations related to failure. There are two main approaches to calculation of the stress intensity factor: we can either extract the SIF from reference handbook data (if available), or we can directly compute the SIF numerically.

*Correspondence to: Anthony T. Patera, Massachusetts Institute of Technology, Dept. of Mechanical Engineering, Room 3-266, 77 Mass. Ave., Cambridge, MA 02139, U.S.A
†Email: patera@mit.edu

Contract/grant sponsor: Singapore-MIT Alliance

Contract/grant sponsor: DARPA, AFOSR; contract/grant number: FA9550-05-1-0114

For some problems, stress intensity factors are provided in various formats in handbooks (e.g., [14]). The data is generated in a variety of ways based on analytical solutions, approximate models, and pre-computed values; the results are typically presented in interpolated form as plots and tables or empirical "fits." Not all problems of interest can be or have been treated, and in general the accuracy in particular of interpolated results can not be assessed or ensured.

The alternative method is to compute the stress intensity factor numerically "ad hoc." In this approach, the stress intensity factor is extracted [3] from the solution field of the partial differential equations (PDE) of linear elasticity obtained by (typically) the finite element method [22]. The numerical methods must often be enhanced (e.g., enriched finite element spaces [11, 12]) given the intrinsically singular origin of the SIF. This approach provides accurate results in general, but it is very expensive — thus precluding application in the real-time or many-query contexts.

Hence current stress intensity factor calculations are either not sufficiently accurate (handbook extraction) or not sufficiently fast (direct numerical evaluation) for many interesting applications; our goal is a computational procedure that provides both *fast* and *reliable* prediction of stress intensity factors. Our approach is based on the reduced basis method [5, 6, 7, 8, 9, 10, 17, 18, 23]. The main ingredients are (*i*) reduced basis approximations [16, 17, 23] that provide rapid and uniform convergence; (*ii*) *a posteriori* error estimators [23] that provide sharp and rigorous bounds for the error in the output (here the SIF); and (*iii*) offline/online computational strategies [16, 17, 23, 24, 27] that allow rapid calculation of both our output approximation and associated error bound.

In short, we propose an approach for stress intensity factor calculation which is guaranteed to have the high accuracy of a "full" numerical simulation — say, by the finite element method — but at the cost of a low-order model — (say) equivalent to the expense of graph/tabular (handbook) extraction. The key ingredient here is the *a posteriori* error estimator, which allows us to minimize the computational cost while ensuring the achievement of (effectively) finite element accuracy.

The paper is organized as follows. In Section 2 we introduce first the standard problem statement and then an equivalent new formulation that enables us to extract our output — the stress intensity factor — as a "compliant" energy release rate. In Section 3 we present several model problems/examples, numerical results for which shall be discussed and compared to available theoretical/numerical references in the subsequent sections. In Section 4, we discuss the extended finite element method — enriched spaces to accurately treat cracks [11, 12] — upon which we shall build our reduced basis approximation. In Section 5 we describe our reduced basis approximation with particular emphasis on those aspects tailored to the stress intensity factor. In Section 6, we present the reduced basis SIF *a posteriori* error estimation theory and associated computational procedures. And finally, in Section 7, we draw conclusions and discuss future work.

2. ABSTRACT STATEMENT

2.1. Classical Formulation

We consider a two-dimensional domain $\Omega \in \mathbb{R}^2$ with boundary $\partial\Omega$. We then introduce the Hilbert space

$$X = \{v \equiv (v_1, v_2) \in (H^1(\Omega))^2 | v_i = 0 \text{ on } \Gamma_D^i,\ i = 1, 2\}, \tag{1}$$

where Γ_D^i is the part of $\partial\Omega$ on which we shall impose homogeneous Dirichlet (zero displacement) boundary conditions. Here $H^1(\Omega) = \{v \in L^2(\Omega) \mid \nabla v \in (L^2(\Omega))^2\}$ where $L^2(\Omega)$ is the space of square-integrable functions over Ω. We equip our space with inner product and associated norm $(\cdot,\cdot)_X$ and $\|\cdot\|_X = \sqrt{(\cdot,\cdot)_X}$, respectively: a typical choice for $(\cdot,\cdot)_X$ is

$$(w, v)_X = \int_\Omega \frac{\partial w_i}{\partial x_j}\frac{\partial v_i}{\partial x_j} + w_i v_i, \tag{2}$$

where summation over repeated component indices is assumed. More general inner products and (equivalent) norms can be considered and are discussed later.

We next define our parameter set $\mathcal{D} \in \mathbb{R}^P$, a typical point in which shall be denoted $\mu \equiv (\mu_1, \ldots, \mu_P)$. The parameter describes the "input" for the problem, such as the physical properties or geometry of the model; in this paper P, the number of "input" parameters, will be $P = 2$ or $P = 3$. We further assume that the domain Ω contains a crack, and (for convenience) that μ_1 is the component of the parameter P–vector (or P–tuple) μ that represents the crack length.

We next introduce the "exact" two-dimensional plane-strain linear elasticity model (extension to plane-stress is of course straightforward [22]). We shall denote dimensional quantities with a superscript ˜; conversely, no superscript ˜ implies a non–dimensional quantity. We scale the dimensional spatial coordinates $\tilde{x}$ by a characteristic length $\widetilde{D}$ to obtain $x = (x_1, x_2)$ $(\in \Omega)$. We scale the dimensional displacement $\tilde{u}$ by $\widetilde{D}\tilde{\sigma}_0/\widetilde{E}_{\text{crack}}$ to obtain $u = (u_1, u_2)$; here $\tilde{\sigma}_0$ is a characteristic (imposed farfield) stress, and $\widetilde{E}_{\text{crack}}$ is the Young's modulus of the material in the vicinity of the crack. (We exploit linearity to eliminate the imposed stress and nominal Young's modulus as parameters.)

The displacement field $u(\mu) \in X$ satisfies the weak form

$$a(u(\mu), v; \mu) = f(v; \mu), \quad \forall\, v \in X; \tag{3}$$

here a is a parametrized bilinear form $a\colon X \times X \times \mathcal{D} \to \mathbb{R}$, and f is a parametrized linear form $f\colon X \times \mathcal{D} \to \mathbb{R}$. We assume that our bilinear form a is coercive, $a(w, w; \mu) \geq \alpha(\mu)\|w\|_X^2 \geq \alpha_0\|w\|_X^2$, $\forall\, w \in X$, $\forall\, \mu \in \mathcal{D}$, for some positive α_0; continuous, $a(w, v; \mu) \leq \gamma(\mu)\|w\|_X\|v\|_X \leq \gamma_0\|w\|_X\|v\|_X$, $\forall\, w, v \in X$, $\forall\, \mu \in \mathcal{D}$, for some finite γ_0; and symmetric, $a(w, v; \mu) = a(v, w; \mu)$, $\forall\, w, v \in X$. We also assume that our linear form f is bounded. Of course, a and f represent the standard linear elasticity weak form — particular instances of which we shall develop in the next section.

Moreover, we further require that $a(\cdot,\cdot;\mu)$ and $f(\cdot;\mu)$ are "affine" in the parameter: these

forms may be expressed as

$$a(w,v;\mu) = \sum_{q=1}^{Q^a} \Theta_q^a(\mu) a_q(w,v), \tag{4}$$

$$f(v;\mu) = \sum_{q=1}^{Q^f} \Theta_q^f(\mu) f_q(v), \tag{5}$$

where $\Theta_q^a(\mu)\colon \mathcal{D} \to \mathbb{R}$, $\Theta_q^f(\mu)\colon \mathcal{D} \to \mathbb{R}$ and $a_q(w,v)\colon X \times X \to \mathbb{R}$, $f_q(v)\colon X \to \mathbb{R}$ are parameter-dependent functions and parameter-independent continuous bilinear/linear forms, respectively. We shall further assume that the functions $\Theta_q^a(\mu)$ and $\Theta_f^q(\mu)$ are smooth, $\Theta_q^a(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q^a$, and $\Theta_q^f(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q^f$, and that the a_q, $1 \le q \le Q^a$, are symmetric.

We next define our output of interest $G(\mu)$: the energy release rate. The energy release rate is computed by an energy approach [3]. To begin, we define the total strain energy (per unit depth) $\Pi(\mu)$ for our model as

$$\Pi(\mu) = \frac{1}{2} a(u(\mu), u(\mu); \mu) - f(u(\mu); \mu), \tag{6}$$

where $u(\mu)$ is the solution of (3) for a particular parameter μ. The energy release rate (ERR) $G(\mu)$ is then derived from the total strain energy $\Pi(\mu)$ as

$$G(\mu) = -\frac{1}{2} \frac{\partial}{\partial \mu_1} a(u(\mu), u(\mu); \mu) + \frac{\partial}{\partial \mu_1} f(u(\mu); \mu), \tag{7}$$

where we recall that μ_1 is the parameter component in μ that represents the crack length. Note the partial derivatives in (7) act only on the "explicit" μ dependence through the last arguments of a and f.

We next invoke the affine nature of a and f, (4), (5) to obtain

$$G = -\frac{1}{2} \sum_{q=1}^{Q^a} \frac{\partial \Theta_q^a(\mu)}{\partial \mu_1} a_q(u(\mu), u(\mu)) + \sum_{q=1}^{Q^f} \frac{\partial \Theta_q^f(\mu)}{\partial \mu_1} f_q(u(\mu)). \tag{8}$$

We thus introduce a symmetric parametrized bilinear form $b\colon X \times X \times \mathcal{D}$ and a parametrized linear form $\ell\colon X \times \mathcal{D}$,

$$b(w,v;\mu) = -\frac{1}{2} \sum_{q=1}^{Q^a} \frac{\partial \Theta_q^a(\mu)}{\partial \mu_1} a_q(w,v), \tag{9}$$

$$\ell(v;\mu) = \sum_{q=1}^{Q^f} \frac{\partial \Theta_q^f(\mu)}{\partial \mu_1} f_q(v), \tag{10}$$

in terms of which our output can be expressed as $G(\mu) = b(u(\mu), u(\mu); \mu) + \ell(u(\mu); \mu)$. Note the output $G(\mu)$ is the sum of a *quadratic* output and a "*non–compliant*" linear output.

Our abstract statement is then: for any $\mu \in \mathcal{D}$, find $G(\mu) \in \mathbb{R}$ given by

$$G(\mu) = b(u(\mu), u(\mu); \mu) + \ell(u(\mu); \mu), \tag{11}$$

where the displacement field $u(\mu) \in X$ satisfies the equilibrium equations (3). This problem statement focuses on the energy release rate; however we can also readily extract the stress intensity factor (SIF) as a "derived" output.

In this paper we shall restrict our attention exclusively to "Mode I," or open-mode, fracture problems. It is known in fracture mechanics theory [22] that, for open-mode fracture problems, the non–dimensional stress intensity factor $K(\mu) = \widetilde{K}(\mu)/(\tilde{\sigma}_0\sqrt{\widetilde{D}})$ and energy release rate G are directly related as

$$K(\mu) = \sqrt{\frac{G(\mu)}{1-\nu_{\text{crack}}^2}}, \quad \text{plane strain}, \tag{12}$$

where ν_{crack} is the Poisson ratio in the vicinity of the crack. (Note the Young's modulus has "disappeared" due to our non-dimensional scaling of the displacement.) It is often stated in the literature [14, 22], and our numerical tests confirm, that the SIF is very insensitive to the Poisson ratio for non-homogeneous materials. Indeed, the SIFs are often presented (e.g., in the sources with which we compare in Section 4) without reference to any particular value of ν.

2.2. Expanded Formulation

We start by considering the following system of equations

$$\begin{aligned} a(u(\mu), v; \mu) &= f(v; \mu), && \forall\, v \in X, && (13)\\ a(z(\mu), v; \mu) &= b(u(\mu), v; \mu) + \frac{1}{2}\ell(v; \mu), && \forall\, v \in X. && (14) \end{aligned}$$

Note that z here is essentially the adjoint associated with our quadratic-linear output [13, 15].

Now we set

$$\begin{aligned} U_+(\mu) &= \frac{1}{2}(u(\mu) + z(\mu)), && (15)\\ U_-(\mu) &= \frac{1}{2}(u(\mu) - z(\mu)), && (16) \end{aligned}$$

in terms of which (13) and (14) can be written as

$$\begin{aligned} a(U_+, v; \mu) + a(U_-, v; \mu) &= f(v; \mu), && \forall\, v \in X,\\ a(U_+, v; \mu) - a(U_-, v; \mu) &= \frac{1}{2}\ell(v; \mu) + b(U_+, v; \mu) + b(U_-, v; \mu), && \forall\, v \in X. \quad (17) \end{aligned}$$

Subtracting and summing these two equations allow us to rewrite this system in a symmetric fashion: $\forall\, V_+, V_- \in X$,

$$\begin{aligned} -b(U_+, V_+; \mu) + 2a(U_+, V_+; \mu) - b(U_-, V_+; \mu) &= f(V_+; \mu) + \frac{1}{2}\ell(V+; \mu),\\ -b(U_+, V_-; \mu) - b(U_-, V_-; \mu) - 2a(U_-, V_-; \mu) &= -f(V_-; \mu) + \frac{1}{2}\ell(V_-; \mu). \quad (18) \end{aligned}$$

We now describe this new formulation more succinctly.

We define the parametrized symmetric bilinear form $\mathcal{A}\colon \mathcal{X} \times \mathcal{X} \times \mathcal{D} \to \mathbb{R}$ and parametrized linear form $\mathcal{F}\colon \mathcal{X} \times \mathcal{D} \to \mathbb{R}$ as

$$\begin{aligned} \mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) &= -b(W_+, V_+; \mu) + 2a(W_+, V_+; \mu) - b(W_-, V_+; \mu)\\ &\quad -b(W_+, V_-; \mu) - b(W_-, V_-; \mu) - 2a(W_-, V_-; \mu),\\ \mathcal{F}(\mathcal{V}; \mu) &= f(V_+; \mu) + \frac{1}{2}\ell(V_+; \mu) - f(V_-; \mu) + \frac{1}{2}\ell(V_-; \mu), \quad (19) \end{aligned}$$

where $\mathcal{X} \equiv X \times X$, and $\mathcal{W} \equiv (U_+, U_-)$, $\mathcal{V} \equiv (V_+, V_-)$. We equip $\mathcal{X}$ with inner product and associated norm $(\cdot,\cdot)_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{X}} = \sqrt{(\cdot,\cdot)_{\mathcal{X}}}$, respectively: our choice of $(\cdot,\cdot)_{\mathcal{X}}$ is $(\mathcal{W},\mathcal{V})_{\mathcal{X}} = a(W_+, V_+; \overline{\mu}) + a(W_-, V_-; \overline{\mu}), \forall\, \mathcal{W}, \mathcal{V} \in \mathcal{X}$; here $\overline{\mu}$ is a particular parameter value in $\mathcal{D}$. The choice of our inner product will not affect either the finite element or the reduced basis output predictions; it will *only* affect the quality and efficiency of the reduced basis *a posteriori* error estimator (primarily but not exclusively through the inf-sup lower bound described briefly in Section 6.1 and in greater detail in [25]). Unfortunately, we do not have any theory on which to base our choice of $\overline{\mu}$; in general, a choice of $\overline{\mu}$ near the center of the parameter domain typically works well.

We now observe that $\mathcal{F}(\mathcal{U}(\mu);\mu)$, $\mathcal{U} = (U_+, U_-)$, is equivalent to the energy release rate $G(\mu)$ in the "classical" formulation:

$$\begin{aligned}
\mathcal{F}(\mathcal{U}(\mu);\mu) &= f(U_+;\mu) + \frac{1}{2}\ell(U_+;\mu) - f(U_-;\mu) + \frac{1}{2}\ell(U_-;\mu) \\
&= f(z(\mu);\mu) + \frac{1}{2}\ell(u(\mu);\mu) \\
&= a(u(\mu), z(\mu);\mu) + \frac{1}{2}\ell(u(\mu);\mu) \\
&= a(z(\mu), u(\mu);\mu) + \frac{1}{2}\ell(u(\mu);\mu) \\
&= b(u(\mu), u(\mu);\mu) + \ell(u(\mu);\mu) \\
&= G(\mu). \qquad (20)
\end{aligned}$$

Here we have invoked (15)–(16), (13)–(14), symmetry, and (11).

Our abstract statement is thus: given $\mu \in \mathcal{D}$, find (the "compliant" output)

$$\mathcal{G}(\mu) = \mathcal{F}(\mathcal{U}(\mu);\mu), \qquad (21)$$

where $\mathcal{U}(\mu) \in \mathcal{X}$ satisfies

$$\mathcal{A}(\mathcal{U}(\mu), \mathcal{V};\mu) = \mathcal{F}(\mathcal{V};\mu), \quad \forall\, \mathcal{V} \in \mathcal{X}; \qquad (22)$$

here $\mathcal{A}$ and $\mathcal{F}$ are the "big" forms defined in (19).

We note that it directly follows from the definitions of $\mathcal{A}$ and $\mathcal{F}$ that both $\mathcal{A}$ and $\mathcal{F}$ are affine in the parameter. In particular,

$$\mathcal{A}(\mathcal{W}, \mathcal{V};\mu) = \sum_{q=1}^{Q^{\mathcal{A}}} \Theta_q^{\mathcal{A}}(\mu)\mathcal{A}_q(\mathcal{W}, \mathcal{V}), \qquad (23)$$

$$\mathcal{F}(\mathcal{V};\mu) = \sum_{q=1}^{Q^{\mathcal{F}}} \Theta_q^{\mathcal{F}}(\mu)\mathcal{F}_q(\mathcal{V}), \qquad (24)$$

where $\Theta_q^{\mathcal{A}}\colon \mathcal{D} \to \mathbb{R}$, $\Theta_q^{\mathcal{F}}\colon \mathcal{D} \to \mathbb{R}$ and $\mathcal{A}_q(\mathcal{W}, \mathcal{V})\colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, $\mathcal{F}_q(\mathcal{V})\colon \mathcal{X} \to \mathbb{R}$ are parameter-dependent functions and parameter-independent continuous bilinear/linear forms, respectively. Moreover, $\Theta_q^{\mathcal{A}}(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q^{\mathcal{A}}$, and $\Theta_q^{\mathcal{F}}(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q^{\mathcal{F}}$, and the $\mathcal{A}_q$, $1 \le q \le Q^{\mathcal{A}}$, are symmetric.

We next introduce the inf-sup parameter and continuity parameter as

$$\beta(\mu) \equiv \inf_{\mathcal{W}\in\mathcal{X}} \sup_{\mathcal{V}\in\mathcal{X}} \frac{\mathcal{A}(\mathcal{W}, \mathcal{V};\mu)}{\|\mathcal{W}\|_{\mathcal{X}}\|\mathcal{V}\|_{\mathcal{X}}}, \quad \forall\, \mu \in \mathcal{D}, \qquad (25)$$

and

$$\gamma(\mu) \equiv \sup_{\mathcal{W}\in\mathcal{X}} \sup_{\mathcal{V}\in\mathcal{X}} \frac{\mathcal{A}(\mathcal{W},\mathcal{V};\mu)}{\|\mathcal{W}\|_{\mathcal{X}}\|\mathcal{V}\|_{\mathcal{X}}} \quad \forall\, \mu \in \mathcal{D}. \tag{26}$$

We shall suppose that $0 < \beta_0 \le \beta(\mu) \le \gamma(\mu) < \gamma_0 < \infty$, $\forall\, \mu \in \mathcal{D}$. We further assume (in fact, this is readily proven given that f and ℓ are bounded) that $\mathcal{F}$ is a bounded linear functional. It then follows that our problem is well posed for all $\mu \in \mathcal{D}$.

We note that the new "expanded" formulation is no longer coercive, however, our output — the energy release rate — is now "compliant." The former is bad news; the latter is good news. However, the real merit of the expanded formulation is that we effectively eliminate the nonlinearity of the output which will in term permit us to develop much simpler, more efficient, and much sharper error bounds. It will be seen in the numerical results of Section 6 that the "expanded" formulation does indeed produce better results and in particular much *sharper* error estimators than the "classical" formulation [26].

3. MODEL PROBLEMS

In this section we shall introduce several "Mode I" linear fracture examples. These examples will serve to illustrate our methods.

3.1. The "two-layer" problem

In this example we consider a linear elasticity problem corresponding to a crack notch inside a two-layer plate: the left (surface) material may be viewed as a coating providing thermal protection [28]. We consider a geometry and loading, shown in Figure 1, which is symmetric about the crack, and which thus corresponds to a Mode I configuration. In consideration of the symmetry of the model about the centerline we treat only half of the original domain, as indicated by the dashed area in Figure 1.

We show in Figure 2 (left) the dimensional (superscript $\tilde{}$) domain $\widetilde{\Omega}(\tilde{d})$, consisting of two layers $\widetilde{\Omega}_1$ and $\widetilde{\Omega}_2$ corresponding to two different materials with Young modulus $\widetilde{E}_1$ and $\widetilde{E}_2$ $(= \widetilde{E}_{\text{crack}})$, respectively. We assume that the two materials share the same Poisson ratio $\nu_1 = \nu_2 = \nu = 0.3$. The two layers $\widetilde{\Omega}_1$ and $\widetilde{\Omega}_2$ are of width $\tilde{t}$ and $4\tilde{t}$, respectively. The crack is of length $\tilde{d}$, and the plate is of width $\tilde{w} \equiv 5\tilde{t}$ and of length $2\widetilde{L} \equiv 8\tilde{w}$. We impose a (normal) traction $\tilde{\sigma}_0$ at the top, $\widetilde{\Gamma}_T$, zero traction on the crack, $\widetilde{\Gamma}_{\text{crack}}$, zero traction on the exposed surface, $\widetilde{\Gamma}_s$, symmetry (zero normal displacement and tangential stress) on the plate centerline, $\widetilde{\Gamma}_C$, and homogeneous Dirichlet (zero displacement) boundary conditions on the right side of the plate, $\widetilde{\Gamma}_R$.

We choose for our characteristic length scale $\widetilde{D} = \tilde{t}$; hence $x^0 = \tilde{x}/\tilde{t}$, and $d = \tilde{d}/\tilde{t}$. Our non–dimensional displacement u^0 is related to the dimensional displacement $\tilde{u}$ by

$$u^0 = \frac{\tilde{u}\widetilde{E}_2}{\tilde{t}\tilde{\sigma}_0}. \tag{27}$$

We denote our new non–dimensional domain and boundaries corresponding to $\widetilde{\Omega}(\tilde{\mu})$, $\widetilde{\Omega}_1$, $\widetilde{\Omega}_2$, and $\widetilde{\Gamma}_{\text{crack}}$, $\widetilde{\Gamma}_C$, $\widetilde{\Gamma}_T$, $\widetilde{\Gamma}_s$, $\widetilde{\Gamma}_R$ as $\Omega^0(\mu)$, Ω^0_1, Ω^0_2, and Γ^0_{crack}, Γ^0_C, Γ^0_T, Γ^0_s, Γ^0_R respectively. The non–dimensional configuration $\Omega^0(d)$ is shown in Figure 2 (center).

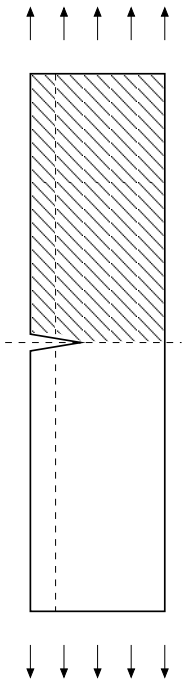


Figure 1. Example 3.1: Geometry model.

We shall consider $P = 2$ parameters: $\mu_1 \equiv d$ (the non–dimensional length of the crack) and $\mu_2 \equiv \kappa = \widetilde{E}_1/\widetilde{E}_2$ (the ratio of the two Young moduli). Our parameter domain is given by $\mathcal{D} = [2.0, 4.0] \times [0.1, 10]$, corresponding to relatively large variations in both crack length and material properties. We choose $\overline{\mu} = (3.0, 5.05)$ in the definition of our inner product; this point is the center point in our parameter domain. (Recall that $\overline{\mu}$ will only affect the sharpness of our reduced basis *a posteriori* error estimators.)

The governing equation is the partial differential equation of linear elasticity (for an isotropic material): the displacement field $u^0(x^0; \mu) \in X^0$ satisfies

$$\int_{\Omega_1^0} \kappa \frac{\partial u_i^0}{\partial x_j^0} C_{ijkl} \frac{\partial v_k}{\partial x_l^0} + \int_{\Omega_2^0} \frac{\partial u_i^0}{\partial x_j^0} C_{ijkl} \frac{\partial v_k}{\partial x_l^0} = \int_{\Gamma_T^0} v, \quad \forall\, v \in X^0, \tag{28}$$

where $C_{ijkl} = \overline{c}_1 \delta_{ij}\delta_{kl} + \overline{c}_2(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$ is the constitutive tensor. Here $\overline{c}_1$ and $\overline{c}_2$ are the Lamé constants for plain strain

$$\overline{c}_1 = \frac{\nu}{(1+\nu)(1-2\nu)} \quad , \quad \overline{c}_2 = \frac{1}{2(1+\nu)}; \tag{29}$$

recall that ν ($= 0.3$, here) is the Poisson ratio. Note that $X^0 = \{(v_1, v_2) \in (H^1(\Omega^0))^2 | v_2|_{\Gamma_C^0} = 0, v_1, v_2|_{\Gamma_R^0} = 0\}$.

In order to apply our methodology we map $\Omega^0(d)$ to the parameter-independent domain $\Omega \equiv \Omega^0(d = d_{\text{ref}} = 3.0)$ shown in Figure 2. The domain ($\Omega_2^0 \rightarrow$) Ω_2 is divided into three

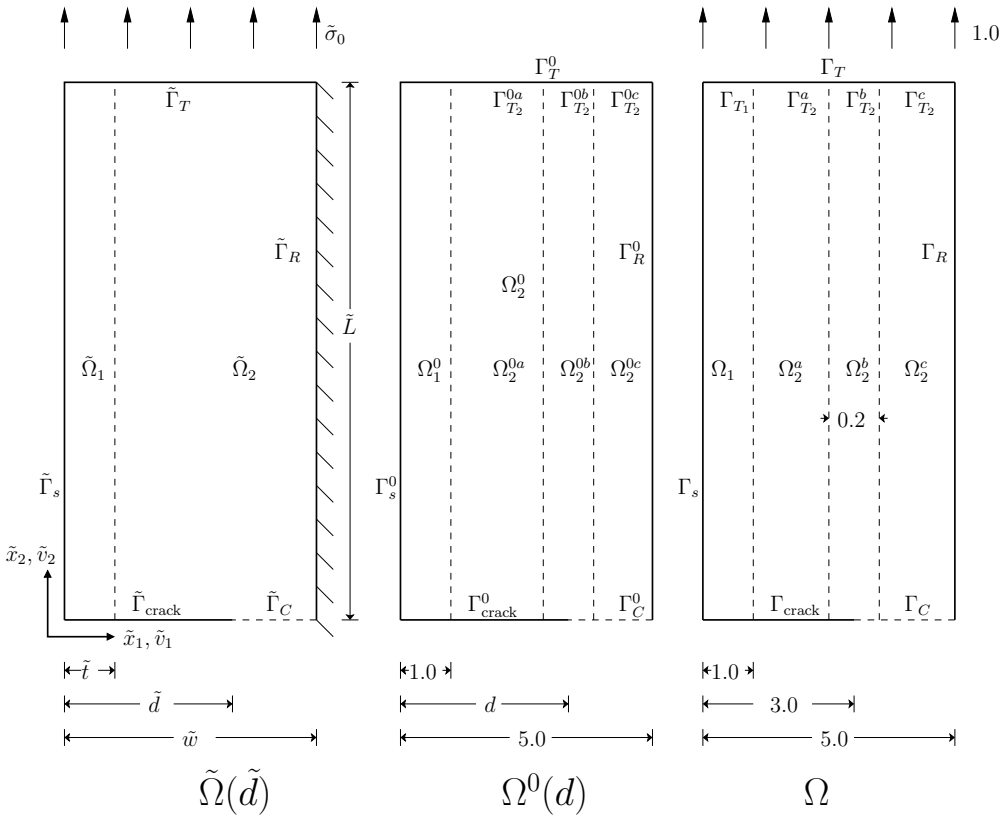


Figure 2. Example 3.1: Original dimensional domain (left); Non–dimensional domain (center); and Reference domain (right).

subdomains $(\Omega_2^{0a}, \Omega_2^{0b}, \Omega_2^{0c} \rightarrow)$ Ω_2^a, Ω_2^b, Ω_2^c, with associated top boundaries $\Gamma_{T_2}^a$, $\Gamma_{T_2}^b$ and $\Gamma_{T_2}^c$, respectively. The transformation is piecewise affine: an identity for Ω_1 and Ω_2^b; simple dilations (in x_1^0) for Ω_2^a and Ω_2^c. The reference domain shall serve all parameter values in $\mathcal{D}$: we choose the "center" $d_{\text{ref}} = 3.0$ to avoid excessive mesh distortion or skewed distribution; note the choice of d_{ref} affects only the accuracy of the finite element solution (on which we build our reduced basis approximation).

We now motivate the need for our domain decomposition: the important considerations are (*i*) the need for a fixed, parameter–independent, reference domain Ω on which to define a "common" reduced basis approximation, and (*ii*) the requirements imposed by the affine parameter–dependence assumption of (4). The domain Ω_1^0 $(= \Omega_1)$ "isolates" the coating layer and associated change in material property: Ω_1^0 is of fixed width and independent of the crack length variation. The three subdomains Ω_2^{0a}, Ω_2^{0b} and Ω_2^{0c}, each associated to a different transformation $(\rightarrow \Omega_2^a, \Omega_2^b, \Omega_2^c)$, "track" the crack: Ω_2^{0b} $(= \Omega_2^b)$ around the crack tip is of fixed width $2\Delta d = 0.2$ in order to preserve affine transformations even in the presence of enriched finite element approximations; Ω_2^{0a} and Ω_2^{0c} shrink/expand to accommodate crack length variation and preserve a fixed total domain width.

The abstract statement for our classical formulation of Section 2.1 is then recovered for $X = \{(v_1, v_2) \in (H^1(\Omega))^2 | v_2|_{\Gamma_C} = 0, v_1, v_2|_{\Gamma_R} = 0\}$, a given by (4) for Θ_q^a and a_q given in Table I, and f given by (5) for Θ_q^f and f_q given in Table II. As a result, we also recover $Q^{\mathcal{A}} = 8$ and $Q^{\mathcal{F}} = 4$ for our expanded formulation of Section 2.2.

q	$\Theta_q^a(\mu)$	$a_q(w, v)$
1	1	$\overline{c}_1 \int_{\Omega\setminus\Omega_2^a} \left(\frac{\partial v_1}{\partial x_1}\frac{\partial w_2}{\partial x_2} + \frac{\partial v_2}{\partial x_2}\frac{\partial w_1}{\partial x_1} \right) d\Omega$ $+ \overline{c}_2 \int_{\Omega\setminus\Omega_2^a} \left(\frac{\partial v_1}{\partial x_2}\frac{\partial w_2}{\partial x_1} + \frac{\partial v_2}{\partial x_1}\frac{\partial w_1}{\partial x_2} \right) d\Omega + \int_{\Omega_2^a} \frac{\partial v_i}{\partial x_j} C_{ijkl} \frac{\partial w_k}{\partial x_l} d\Omega$
2	κ	$\int_{\Omega_1} \frac{\partial v_i}{\partial x_j} C_{ijkl} \frac{\partial w_k}{\partial x_l} d\Omega$
3	t_x^1	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_2^b} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_2^b} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
4	$\frac{1}{t_x^1}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_2^b} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_2^b} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
5	t_x^2	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_2^c} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_2^c} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
6	$\frac{1}{t_x^2}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_2^c} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_2^c} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$

$$t_x^1 = \frac{d - 1.0 - \Delta d}{d_{\text{ref}} - 1.0 - \Delta d}, \quad t_x^2 = \frac{5.0 - d - \Delta d}{5.0 - d_{\text{ref}} - \Delta d}, \quad \Delta d = 0.1, \quad d_{\text{ref}} = 3.0$$

Table I. Example 3.1: Parametric functions $\Theta_q^a(\mu)$ and parameter-independent bilinear forms $a_q(w, v)$, $1 \le q \le Q^a = 6$.

q	$\Theta_q^f(\mu)$	$f_q(v)$
1	1	$\int_{\Gamma_{T_1}} v d\Gamma + \int_{\Gamma_{T_2}^b} v d\Gamma$
2	t_x^1	$\int_{\Gamma_{T_2}^a} v d\Gamma$
3	t_x^2	$\int_{\Gamma_{T_2}^c} v d\Gamma$

Table II. Example 3.1: Parametric functions $\Theta_q^f(\mu)$ and parameter-independent bilinear forms $f_q(v)$, $1 \le q \le Q^f = 3$.

3.2. The center-cracked tension specimen

In this example we consider a linear elasticity problem corresponding to a plate containing an internal center crack under tension. In consideration of the symmetry of the model about the centerlines we treat only a quarter of the original problem, as indicated by the dashed area in Figure 3.

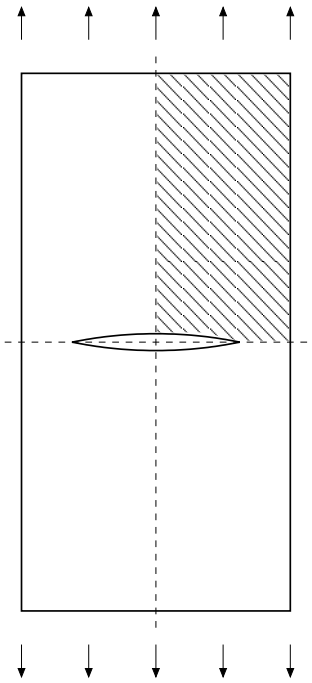


Figure 3. Example 3.2: Geometry model.

We show in Figure 4 the resulting domain $\widetilde{\Omega}(\tilde{d}, \widetilde{L})$. The crack is of length $2\tilde{d}$, and the plate is of width $2\tilde{w}$ and of length $2\widetilde{L}$. We impose (normal) traction $\tilde{\sigma}_0$ at the top, $\widetilde{\Gamma}_T$, zero traction on the crack, $\widetilde{\Gamma}_{\text{crack}}$, zero traction on the right side of the plate, $\widetilde{\Gamma}_R$, and symmetry boundary conditions on $\widetilde{\Gamma}_{C_1}$ and $\widetilde{\Gamma}_{C_2}$. We consider a homogeneous isotropic material characterized by Young's modulus $\widetilde{E} = \widetilde{E}_{\text{crack}}$ and Poisson ratio $\nu = 0.3$.

We choose for our characteristic length scale $\widetilde{D} = \tilde{w}$; hence $x^0 = \tilde{x}/\tilde{w}$, $L = \widetilde{L}/\tilde{w}$, and $d = \tilde{d}/\tilde{w}$. Our non–dimensional displacement u^0 is related to the dimensional displacement $\tilde{u}$ by

$$u^0 = \frac{\tilde{u}\tilde{E}}{\tilde{w}\tilde{\sigma}_0}. \tag{30}$$

We denote our new non–dimensional domain and our boundaries corresponding to $\widetilde{\Omega}(\tilde{\mu})$, and $\widetilde{\Gamma}_{\text{crack}}$, $\widetilde{\Gamma}_{C_1}$, $\widetilde{\Gamma}_{C_2}$, $\widetilde{\Gamma}_T$, $\widetilde{\Gamma}_R$ as $\Omega^0(\mu)$, and Γ^0_{crack}, $\Gamma^0_{C_1}$, $\Gamma^0_{C_2}$, Γ^0_T, Γ^0_R, respectively.

We shall consider $P = 2$ parameters: $\mu_1 \equiv d$ (the non–dimensional half–length of the crack) and $\mu_2 \equiv L$ (the non–dimensional half–length of the specimen); our parameter domain is

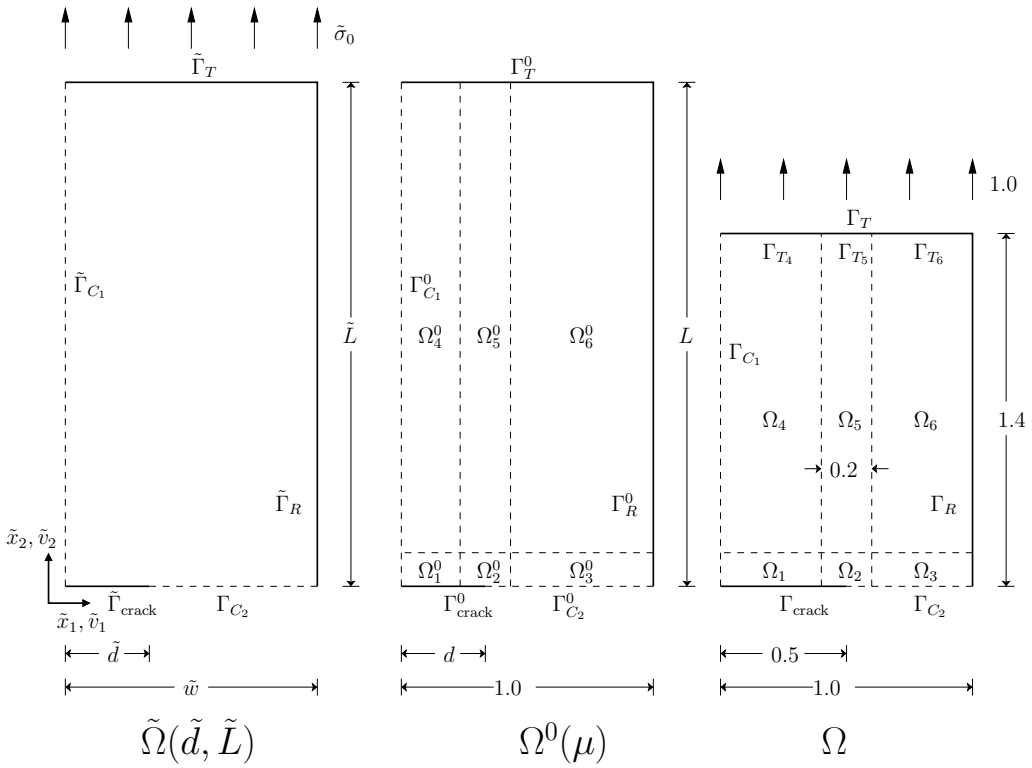


Figure 4. Example 3.2: Original dimensional domain (left); Non–dimensional domain (center); and Reference domain (right).

given by $\mathcal{D} = [0.3, 0.7] \times [0.4, 2.4]$. We emphasize that we do not assume a very long specimen; however, we shall show later that our parameter domain can in fact produce accurate results for the infinite specimen case. We choose $\overline{\mu} = (0.5, 1.4)$ — the center of our parameter domain — in the definition of our inner product.

The governing equation is the partial differential equations of linear elasticity: the displacement field $u^0(x; \mu) \in X^0$ satisfies

$$\int_{\Omega^0} \frac{\partial u_i^0}{\partial x_j^0} C_{ijkl} \frac{\partial v_k}{\partial x_l^0} = \int_{\Gamma_T^0} v, \quad \forall\, v \in X^0, \tag{31}$$

where $C_{ijkl} = \overline{c}_1 \delta_{ij}\delta_{kl} + \overline{c}_2(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$ is the constitutive tensor. Here $\overline{c}_1$ and $\overline{c}_2$ are the Lamé constants for plain strain

$$\overline{c}_1 = \frac{\nu}{(1+\nu)(1-2\nu)} \quad , \quad \overline{c}_2 = \frac{1}{2(1+\nu)}; \tag{32}$$

recall that $\nu = 0.3$ is the Poisson ratio. Note that $X^0 = \{(v_1, v_2) \in (H^1(\Omega^0))^2 | v_2|_{\Gamma^0_{C_2}} = 0, v_1|_{\Gamma^0_{C_1}} = 0\}$.

For reasons cited in Section 3.1, we now map our domain $\Omega^0(\mu)$ to a fixed parameter–independent reference domain, Ω. In particular, we apply the transformation $\Omega^0(\mu) \rightarrow \Omega \equiv$

$\Omega^0(\mu = \mu_{\rm ref} = (d_{\rm ref} = 0.5, L_{\rm ref} = 1.4))$: the mapping is piecewise-affine over 6 subdomains, $\Omega_i^0 \to \Omega_i$, $i = 1, \ldots, 6$ as shown in Figure 4. In addition, the boundary ($\Gamma_T^0 \to$) Γ_T is further divided into three segments Γ_{T_4}, Γ_{T_5} and Γ_{T_6}. The transformation is relatively simple: the identity for (the invariant domain of width $2\Delta d = 0.2$) Ω_2^0; dilation in the x_1^0-direction for Ω_1^0 and Ω_3^0 (to accommodate crack variation); dilation in the x_2^0-direction for Ω_5^0 (to accommodate specimen length variation); and dilations in both the x_1^0-direction and x_2^0-direction for Ω_4^0 and Ω_6^0 (to accommodate both crack and specimen length variation).

The abstract statement for our classical formulation is then recovered for $X = \{(v_1, v_2) \in (H^1(\Omega))^2 | v_2|_{\Gamma_{C_2}} = 0, v_1|_{\Gamma_{C_1}} = 0\}$, a given by (4) for Θ_q^a and a_q given in Table III, and f given by (5) for Θ_q^f and f_q given in Table IV. As a result, we also recover $Q^{\mathcal{A}} = 19$ and $Q^{\mathcal{F}} = 5$ for our expanded formulation of Section 2.2.

3.3. The crack-hole problem

In this example we consider a linear elasticity problem corresponding to a plate with a circular hole from which emanate two cracks. In consideration the symmetry of the model about the centerlines we treat only one quarter of the original problem, as indicated by the dashed area in Figure 5.

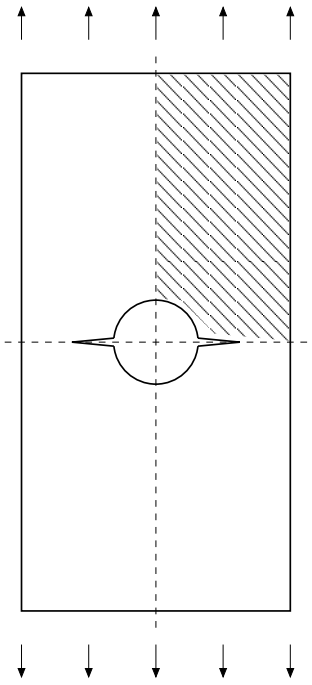


Figure 5. Example 3.3: Geometry model.

We show in Figure 6 the resulting domain $\widetilde{\Omega}^0(\tilde{d}, \widetilde{R}, \widetilde{L})$. The radius of the hole is $\widetilde{R}$, the length of each side crack (as measured from the center of the hole) is $\tilde{d}$, and the plate is of width $2\tilde{w}$ and of length $2\widetilde{L}$. We impose (normal) traction $\tilde{\sigma}_0$ at the top, $\widetilde{\Gamma}_T$, zero traction on the

q	$\Theta_q^a(\mu)$	$a_q^e(w,v)$
1	1	$\overline{c}_1 \int_{\Omega\setminus\Omega_2} \left(\frac{\partial v_1}{\partial x_1}\frac{\partial w_2}{\partial x_2} + \frac{\partial v_2}{\partial x_2}\frac{\partial w_1}{\partial x_1} \right) d\Omega$ $+ \overline{c}_2 \int_{\Omega\setminus\Omega_2} \left(\frac{\partial v_1}{\partial x_2}\frac{\partial w_2}{\partial x_1} + \frac{\partial v_2}{\partial x_1}\frac{\partial w_1}{\partial x_2} \right) d\Omega + \int_{\Omega_2} \frac{\partial v_i}{\partial x_j} C_{ijkl} \frac{\partial w_k}{\partial x_l} d\Omega$
2	t_x^1	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_1} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_1} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
3	$\frac{1}{t_x^1}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_1} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_1} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
4	t_x^2	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_3} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_3} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
5	$\frac{1}{t_x^2}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_3} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_3} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
6	$\frac{t_x^1}{t_y}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_4} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_4} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
7	$\frac{t_y}{t_x^1}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_4} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_4} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
8	$\frac{t_x^2}{t_y}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_6} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_6} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
9	$\frac{t_y}{t_x^2}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_6} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_6} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
10	$\frac{1}{t_y}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_8} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_8} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
11	t_y	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_8} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_8} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$

$$t_x^1 = \frac{d - \Delta d}{d_{\text{ref}} - \Delta d}, \quad t_x^2 = \frac{1.0 - d - \Delta d}{1.0 - d_{\text{ref}} - \Delta d}, \quad t_y = \frac{L - \Delta d}{L_{\text{ref}} - \Delta d},$$

$$\Delta d = 0.1, \qquad d_{\text{ref}} = 0.5, \qquad L_{\text{ref}} = 1.4$$

Table III. Example 3.2: Parametric functions $\Theta_q^a(\mu)$ and parameter-independent bilinear forms $a_q(w,v)$, $1 \le q \le Q^a = 11$.

q	$\Theta_q^f(\mu)$	$f_q(v)$
1	1	$\int_{\Gamma_{T_5}} v d\Gamma$
2	t_x^1	$\int_{\Gamma_{T_4}^a} v d\Gamma$
3	t_x^2	$\int_{\Gamma_{T_6}^c} v d\Gamma$

Table IV. Example 3.2: Parametric functions $\Theta_q^f(\mu)$ and parameter-independent bilinear forms $f^q(v)$, $1 \le q \le Q^f = 3$.

hole, $\widetilde{\Gamma}_{\text{hole}}$, zero traction on the crack, $\widetilde{\Gamma}_{\text{crack}}$, zero traction on the right side of the plate, $\widetilde{\Gamma}_R$, and symmetry boundary conditions on $\widetilde{\Gamma}_{C_1}$ and $\widetilde{\Gamma}_{C_2}$. We consider a homogeneous isotropic material characterized by Young's modulus $\widetilde{E} = \widetilde{E}_{\text{crack}}$ and Poisson ratio $\nu = 0.3$.

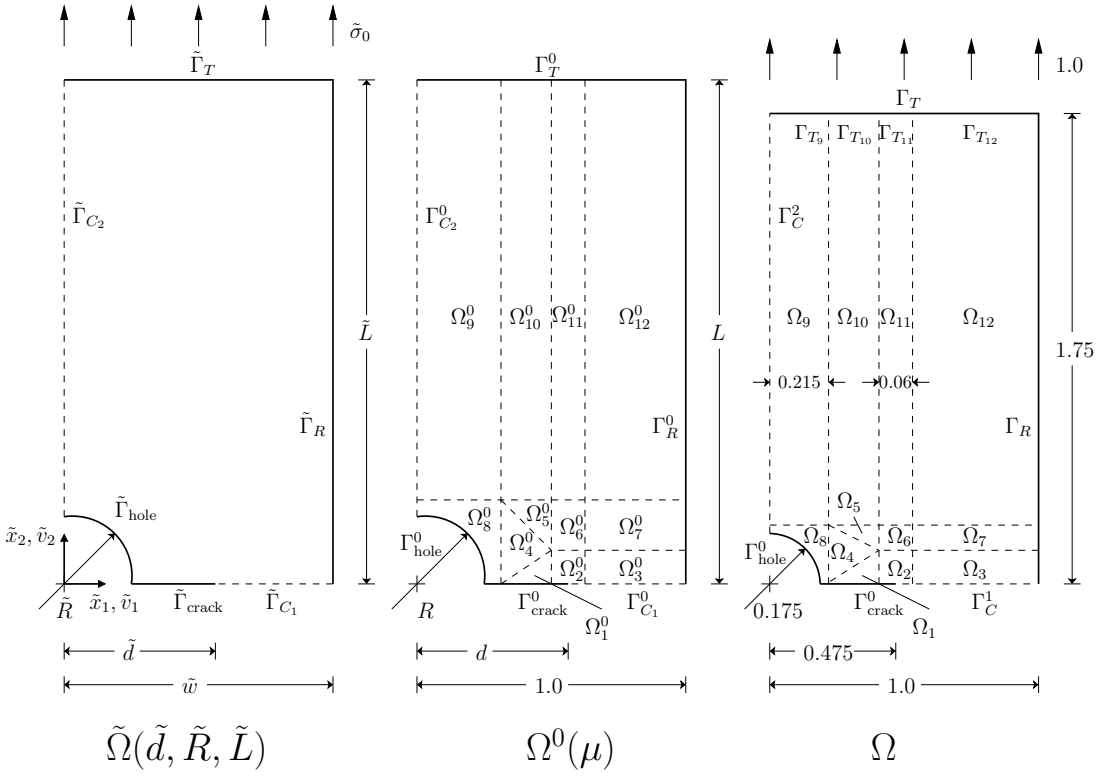


Figure 6. Example 3.3: Original dimensional domain (left); Non–dimensional domain (center) and Reference domain (right).

We choose for our characteristic length scale $\widetilde{D} = \tilde{w}$; hence $x^0 = \tilde{x}/\tilde{w}$, $L = \widetilde{L}/\tilde{w}$, $R = \widetilde{R}/\tilde{w}$,

and $d = \tilde{d}/\tilde{w}$. Our non–dimensional displacement u^0 is related to the dimensional displacement $\tilde{u}$ by

$$u^0 = \frac{\tilde{u}\widetilde{E}}{\tilde{w}\tilde{\sigma}_0}. \tag{33}$$

We denote our new non–dimensional domain and boundaries corresponding to $\widetilde{\Omega}(\tilde{\mu})$, and $\widetilde{\Gamma}_{\text{hole}}$, $\widetilde{\Gamma}_{\text{crack}}$, $\widetilde{\Gamma}_{C_1}$, $\widetilde{\Gamma}_{C_2}$, $\widetilde{\Gamma}_T$, $\widetilde{\Gamma}_R$ as $\Omega^0(\mu)$, and Γ^0_{hole}, Γ^0_{crack}, $\Gamma^0_{C_1}$, $\Gamma^0_{C_2}$, Γ^0_T, Γ^0_R, respectively.

In this example we shall consider $P = 3$ parameters: $\mu_1 \equiv d$ (the non–dimensional half–length of the crack), $\mu_2 \equiv R$ (the non–dimensional radius of the hole), and $\mu_3 \equiv L$ (the non–dimensional length of the specimen); our parameter domain is given by $\mathcal{D} = [0.325, 0.625] \times [0.1, 0.25] \times [1.5, 2.0]$. We choose $\overline{\mu} = (0.475, 0.175, 1.75)$ — again, the center of our parameter domain — in the definition of our inner product.

The governing equation is the partial differential equations of linear elasticity: the displacement field $u^0(x; \mu) \in X^0$ satisfies

$$\int_{\Omega^0} \frac{\partial u_i^0}{\partial x_j^0} C_{ijkl} \frac{\partial v_k}{\partial x_l^0} = \int_{\Gamma_T^0} v, \quad \forall\, v \in X^0, \tag{34}$$

where $C_{ijkl} = \overline{c}_1 \delta_{ij}\delta_{kl} + \overline{c}_2(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$ is the constitutive tensor. Here $\overline{c}_1$ and $\overline{c}_2$ are the Lamé constants for plain strain

$$\overline{c}_1 = \frac{\nu}{(1+\nu)(1-2\nu)} \quad , \quad \overline{c}_2 = \frac{1}{2(1+\nu)}; \tag{35}$$

recall that $\nu = 0.3$ is the Poisson ratio. Note that $X^0 = \{(v_1, v_2) \in (H^1(\Omega^0))^2 | v_2|_{\Gamma^0_{C_1}} = 0, v_1|_{\Gamma^0_{C_2}} = 0\}$.

In order to apply our methodology we map $\Omega^0(\mu) \rightarrow \Omega \equiv \Omega^0(\mu = \mu_{\text{ref}} = (d_{\text{ref}} = 0.475, R_{\text{ref}} = 0.175, L_{\text{ref}} = 1.75))$: the mapping is piecewise-affine over the 12 subdomains, $\Omega_i^0 \rightarrow \Omega_i$, $i = 1, \ldots, 12$, as shown in Figure 6. The boundary ($\Gamma_T^0 \rightarrow$) Γ_T is further divided into four segments Γ_{T_9}, $\Gamma_{T_{10}}$, $\Gamma_{T_{11}}$, and $\Gamma_{T_{12}}$. We note that $\Omega_1^0 = \Omega_1$ (identity mapping) is the invariant square domain around the crack tip (of size $2\Delta d = 0.06$); remark also that the near-hole subdomain $\Omega_8^0 \rightarrow \Omega_8$ is uniformly dilated in the x_1^0-direction and x_2^0-direction to preserve the circular image independent of μ. The remainder of the (admittedly complicated) details of the affine mappings, and the resulting parameter-dependent functions and parameter-independent forms, are provided in the Appendix.

4. FINITE ELEMENT METHOD

In this study our focus is on the reduced basis method, but because our reduced basis discretization is built upon, and measured (as regards accuracy) relative to, a "truth" approximation, we must first define this "truth." In this section we shall introduce our "truth" approximation, which is based on a (now relatively) standard extended/enriched finite element method.

We provide here a very brief overview of the final result. Given $\mu \in \mathcal{D}$, we define an extended/enriched finite element approximation $\mathcal{U}_h(\mu) \in \mathcal{X}_h$ to the exact solution $\mathcal{U}(\mu)$, where $\mathcal{X}_h$ is our finite element space of dimension $2\mathcal{N} = \dim(\mathcal{X}_h)$. We next choose our "truth"

finite element approximation $\mathcal{U}_{h_t}(\mu) \in \mathcal{X}_{h_t}$, where $\mathcal{X}_{h_t}$ is our finite element space of dimension $2\mathcal{N}_t = \dim(\mathcal{X}_{h_t})$; here we require that $\mathcal{X}_{h_t}$ is sufficiently rich that $\mathcal{U}_{h_t}(\mu)$ is sufficiently close to $\mathcal{U}(\mu)$ for all μ in the parameter domain $\mathcal{D}$. Then, given $\mu \in \mathcal{D}$, we find $\mathcal{U}_{h_t}(\mu) \in \mathcal{X}_{h_t}$ such that $\mathcal{A}_{h_t}(\mathcal{U}_{h_t}(\mu), \mathcal{V}; \mu) = \mathcal{F}_{h_t}(\mathcal{V}; \mu), \forall \mathcal{V} \in \mathcal{X}_{h_t}$; we can then evaluate the energy release rate as $\mathcal{G}_{h_t}(\mu) = \mathcal{F}_{h_t}(\mathcal{U}_{h_t}(\mu); \mu)$. The reader already familiar with the finite element method may choose in the first reading to proceed to Section 5 for the continued development of the reduced basis method. However the remaining material in Section 4 below is important in justifying and assessing the quality and relevance of the reduced basis results, and in understanding the computational effort and savings associated with reduced basis treatment.

4.1. Extended finite element method

The characteristic property of elliptic linear PDEs, such as the linear elasticity problems of interest here, is that the solution $u(\mu)$ is smooth if the domain boundary Γ and load/source $f(\cdot; \mu)$ are smooth; conversely, if the domain boundary Γ is *not* smooth — as in fracture-mechanics crack problems — singularities can occur, with corresponding detriment to convergence rates. This difficulty was of course recognized and demonstrated very early in the development of the finite element method [2].

One way to overcome this difficulty is to effectively or actually include the relevant singularities in the finite element space. There are two common treatments: the (*i*) "quarter-point" element [4]; and (*ii*) the enriched finite element method [2] and recent generalizations such as the extended finite element method [12]. In this paper we choose to pursue the latter, which exploits the partition of unity property [11] to enrich the region around the crack tip with appropriate asymptotic fields. It is shown that both enriched finite element method [2] and extended finite element method [19, 20] yield optimal convergence rates of h^m (in the X norm) for elements of polynomial order m.

For simplicity of exposition, we consider here a "rectilinear" domain Ω; a "shared" Dirichlet boundary $\Gamma_D^1 = \Gamma_D^2 = \Gamma_D$ such that $v \in X$ satisfies $v_1|_{\Gamma_D} = v_2|_{\Gamma_D} = 0$; a triangulation $\mathcal{T}_h$ that consists (only) of rectangular elements; and (only) bilinear elemental approximation spaces. In actual practice, we permit curved domains (e.g, for Example 3.3), "mixed" component (e.g, symmetry) boundary conditions, triangular and quadrilateral composite triangulations, (bi)linear and (bi)quadratic elements, and isoparametric elemental mappings [27]. We eschew these details here given that first, these embellishments are all standard and second, these methods are (a prerequisite for but) not the focus of this paper.

Under our assumptions of "expository convenience," we may write

$$\overline{\Omega} = \bigcup_{\overline{Q}_h \in \mathcal{T}_h} \overline{Q}_h, \tag{36}$$

where the Q_h are our rectangular elements; h here refers to the maximum sidelength over all $Q_h \in \mathcal{T}_h$. The usual (scalar) finite element bilinear approximation subspace (of dimension J) may then be expressed as

$$Y_h = \{v \in C^0(\Omega), v|_{\Gamma_D} = 0 \,|\, v_{Q_h} \in \mathbf{Q}_1, \, \forall \, Q_h \in \mathcal{T}_h\}, \tag{37}$$

where $\mathbf{Q}_1 = \text{span}\{x_1^i x_2^j, 0 \le i, j \le 1\}$. We choose the standard nodal basis, $\{\phi_i\}_{1 \le i \le J} : \phi_j \in Y_h$ and $\phi_j(x_i^v = (x_1^v, x_2^v)_i) = \delta_{ij}, 1 \le i, j \le J$; here the x_i^v are the vertices/nodes of $\mathcal{T}_h$ that do not reside on Γ_D.

We next define the displacement asymptotic crack tip fields, $\psi^k(r,\theta)$, $1 \le k \le 4$ [12]:

$$\begin{aligned}
\psi^1(r,\theta) &= \sqrt{r}\sin\frac{\theta}{2}\ , \\
\psi^2(r,\theta) &= \sqrt{r}\cos\frac{\theta}{2}\ , \\
\psi^3(r,\theta) &= \sqrt{r}\sin\frac{\theta}{2}\sin\theta\ , \\
\psi^4(r,\theta) &= \sqrt{r}\cos\frac{\theta}{2}\sin\theta\ ,
\end{aligned} \tag{38}$$

where (r,θ) are local polar coordinates relative to the crack tip (for a crack defined by $r > 0, \theta = -\pi$). We further introduce the index set $\mathcal{I}_M \in \{1,\ldots,J\}$ of nodes (more precisely, the numbers of nodes) near the crack tip that shall be enriched; we show in Figure 7 our typical choice for the enriched nodes x_m^v, $m \in \mathcal{I}_M$. Our enriched (scalar) finite element approximation subspace (of dimension $J^{\rm en} = J + 4M$) is then given by

$$\begin{aligned}
Y_h^{\rm en} &= \text{span}\{\phi_j, 1 \le j \le J; \phi_m\psi^k, \forall\, m \in \mathcal{I}_M, 1 \le k \le 4\} \\
&= \text{span}\{\varphi_j, 1 \le j \le J^{\rm en}\};
\end{aligned} \tag{39}$$

our basis functions φ are simply a relabelling of the $\phi_\cdot$ and $\phi_\cdot\psi^\cdot$.

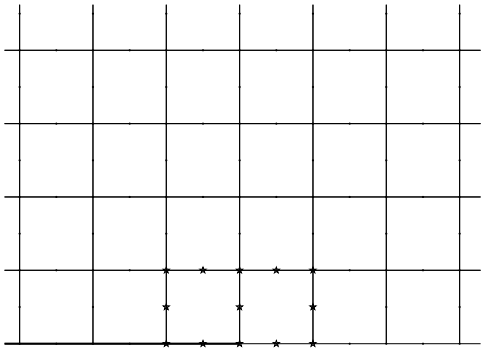


Figure 7. The "enriched" region around the crack tip region for $\mathbf{Q}_2$ elements: ($\star$) indicates nodes (whose indices are) in $\mathcal{I}_M$. Recall that we impose Neumann conditions on the crack surface.

We now define our enriched (vector) finite element approximation (of dimension $\mathcal{N} = 2J^{\rm en}$) as the product space

$$\begin{aligned}
X_h &= Y_h^{\rm en} \times Y_h^{\rm en} \\
&= \text{span}\{\varphi_j \mathbf{i}_l, 1 \le j \le J^{\rm en}, l = 1,2\},
\end{aligned} \tag{40}$$

where $\mathbf{i}_1 = (1,0)$ and $\mathbf{i}_2 = (0,1)$ are the unit vectors in $\mathbb{R}^2$.

Our Galerkin approximation to (11), (3) is then: given $\mu \in \mathcal{D}$, find

$$G_h(\mu) = b_h(u_h(\mu), u_h(\mu); \mu) + \ell_h(u_h(\mu); \mu), \tag{41}$$

where $u_h(\mu) \in X_h$ satisfies

$$a_h(u_h(\mu), v; \mu) = f_h(v; \mu), \quad \forall\, v \in X_h. \tag{42}$$

Here the subscript h in b_h, ℓ_h, a_h, and f_h refers to numerical quadrature of the integrals that define b, ℓ, a, and f, respectively; we exploit a $N_{\rm qu} \times N_{\rm qu}$ tensor-product Gauss scheme in each element Q_h, with $N_{\rm qu}$ as large as 15 for elements which contain enriched nodes (to accommodate the ψ).

To arrive at our discrete equations, we first write

$$u_h(x;\mu) = \sum_{l=1}^{2} \sum_{j=1}^{J^{\rm en}} \varphi_j \mathbf{i}_l (u_{hl})_j, \tag{43}$$

where the $(u_{hl})_j$, $1 \le j \le J^{\rm en}$, are the finite element nodal values (apart from the enrichment) of the $l^{\rm th}$ component of u_h. We then choose $v = \varphi_i \mathbf{i}_k$, $1 \le i \le J^{\rm en}$, $k = 1, 2$ to obtain

$$\sum_{l=1}^{2} \sum_{j=1}^{J^{\rm en}} a_h(\varphi_j \mathbf{i}_l, \varphi_i \mathbf{i}_k; \mu)(u_{hl})_j = f_h(\varphi_i \mathbf{i}_k; \mu), \quad 1 \le i \le J^{\rm en},\ k = 1, 2. \tag{44}$$

We may then evaluate G_h as

$$G_h(\mu) = \sum_{k,l=1}^{2} \sum_{i,j=1}^{J^{\rm en}} (u_{hk})_i b_h(\varphi_j \mathbf{i}_l, \varphi_i \mathbf{i}_k)(u_{hl})_j + \sum_{l=1}^{2} \sum_{j=1}^{J^{\rm en}} (u_{hl})_j \ell_h(\varphi_j \mathbf{i}_l). \tag{45}$$

We note that the introduction of the enriched finite element space also renders the discrete linear equations poorly conditioned, and thus a simple Gaussian elimination solver is not guaranteed to provide good accuracy; we solve the system by the Generalized Minimum Residual (GMRES) iterative algorithm.

We shall require the approximate solution $\mathcal{U}_h(\mu)$ to $\mathcal{U}(\mu)$ (of the "extended" formulation) in order to build our reduced basis. To obtain $\mathcal{U}_h(\mu)$, we apply the following procedure: (i) we first solve for $u_h(\mu)$ from (42); (ii) we next solve for the "adjoint" $z_h(\mu) \in X_h$ from $a_h(z_h(\mu), v; \mu) = b_h(u_h(\mu), v; \mu) + \frac{1}{2}\ell_h(v; \mu)$, $\forall v \in X_h$; (iii) we then obtain our approximations of U_{+h} and U_{-h} as

$$U_{+h}(\mu) = \frac{1}{2}(u_h(\mu) + z_h(\mu)), \tag{46}$$

$$U_{-h}(\mu) = \frac{1}{2}(u_h(\mu) - z_h(\mu)); \tag{47}$$

finally, (iv) we construct $\mathcal{U}_h(\mu) = (U_{+h}, U_{-h})$.

Our finite element approximation is thus: given $\mu \in \mathcal{D}$, find (the "compliant" output)

$$\mathcal{G}_h(\mu) = \mathcal{F}_h(\mathcal{U}_h(\mu); \mu), \tag{48}$$

where $\mathcal{U}_h(\mu) \in \mathcal{X}_h$ satisfies

$$\mathcal{A}_h(\mathcal{U}_h(\mu), \mathcal{V}; \mu) = \mathcal{F}_h(\mathcal{V}; \mu), \quad \forall\, \mathcal{V} \in \mathcal{X}_h; \tag{49}$$

here $\mathcal{X}_h \equiv X_h \times X_h$, and $\mathcal{A}_h$ and $\mathcal{F}_h$ are defined as

$$\begin{aligned} \mathcal{A}_h(\mathcal{W}, \mathcal{V}; \mu) &= -b_h(W_+, V_+; \mu) + 2a_h(W_+, V_+; \mu) - b_h(W_-, V_+; \mu) \\ &\quad -b_h(W_+, V_-; \mu) - b_h(W_-, V_-; \mu) - 2a_h(W_-, V_-; \mu), \\ \mathcal{F}_h(\mathcal{V}; \mu) &= f_h(V_+; \mu) + \frac{1}{2}\ell_h(V_+; \mu) - f_h(V_-; \mu) + \frac{1}{2}\ell_h(V_-; \mu). \end{aligned} \tag{50}$$

We denote by $\mathcal{A}_{hq}$, $1 \le q \le Q^{\mathcal{A}}$, and $\mathcal{F}_{hq}$, $1 \le q \le Q^{\mathcal{F}}$, our numerical quadratures of the bilinear/linear forms associated with our affine parameter decompositions $\mathcal{A}_q$, $1 \le q \le Q^{\mathcal{A}}$, and $\mathcal{F}_q$, $1 \le q \le Q^{\mathcal{F}}$, respectively.

4.2. Numerical results

For Example 3.1 we shall consider convergence of the energy release rate (ERR) as a function of refinement level. For Example 3.2 and Example 3.3 we shall directly compare our SIF predictions to available data. We recall that we can compute the SIF $K_h(\mu)$ from (12) as $K_h(\mu) = \sqrt{\mathcal{G}_h(\mu)/(1-\nu^2)}$ (in all cases for $\nu = 0.3$).

4.2.1. Example 3.1 We illustrate the finite element convergence for Example 3.1 for the case $\mu = (3.0, 1.0)$; this μ is simply a particular value at which we analyze the convergence of the finite element discretization. All meshes contain only quadrilateral elements, and the triangulation for refinement level $L_r + 1$ is constructed by sub-dividing all quadrilateral elements Q_h at level L_r into 4 smaller similar quadrilaterals; we show in Figure 8 the triangulation associated with the refinement level $L_r = 3$. We take as our reference solution, $\mathcal{U}_{h_{\rm ref}}$, a $\mathbf{Q}_2$ ($\equiv \text{span}\{x_1^i x_2^j,\ 0 \le i, j \le 2\}$) approximation on the (finest) $L_r = 6$ refinement corresponding to $\mathcal{N} = 198{,}330$ degrees of freedom.

We plot in Figure 9 $|\mathcal{F}_h(\mathcal{U}_h;\mu) - \mathcal{F}_{h_{\rm ref}}(\mathcal{U}_{h_{\rm ref}};\mu)|$ as calculated by the compliance formula (48) for both $\mathbf{Q}_1$ and $\mathbf{Q}_2$ elements: the convergence rate is improved relative to the classical FEM approach as we correctly capture the singularity fields around the crack. The convergence rate for the linear ($\mathbf{Q}_1$) approximation is close to order 2, and hence nearly optimal. However, the convergence rate of the quadratic approximation ($\mathbf{Q}_2$) is at best only of order slightly greater then 3 — and not the optimal 4. The sub-optimal rate may be due to the presence of only the first singularity in the expanded space; the incorporation of more enrichment functions can further improve the convergence rate, as shown in [20, 19]. (We believe that quadrature errors are negligible in our convergence tests.)

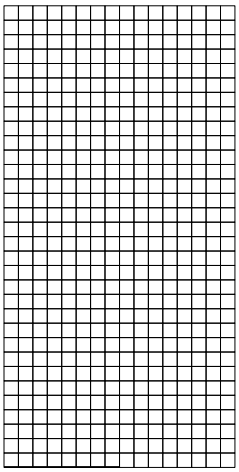


Figure 8. Example 3.1: Refinement level $L_r = 3$.

4.2.2. Example 3.2 We next present the numerical results for Example 3.2. For the infinite plate case ($L \to \infty$), the theoretical stress intensity factor values are available [14]; however, stress intensity factor values are not available for $L < 2.0$. In practice, it is known that the

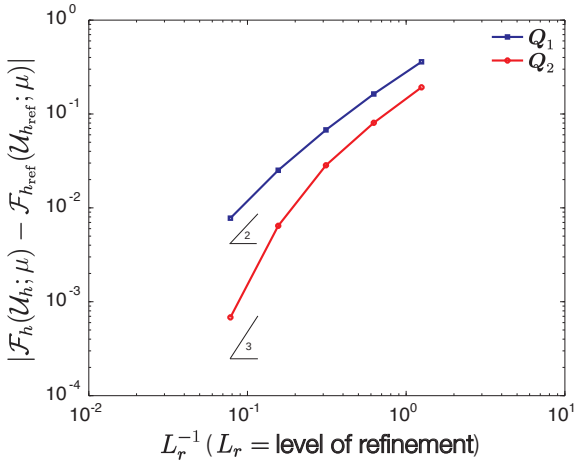


Figure 9. Example 3.1: Finite element convergence with refinement. Note the reference solution corresponds to $L_r = 6$.

stress intensity factor for a finite plate is quite close to the theoretical result for an infinite plate for $L \geq 2.0$.

For our finite element approximation we choose $\mathbf{Q}_2, \mathbf{P}_2$ elements on the "triangulation" shown in Figure 10 for a total of $\mathcal{N} = 4{,}490$ degrees of freedom; the enriched region is the first element ring around the crack tip. We present in Figure 11 the (non–dimensional) stress intensity factor $K_h(\mu)$. We note that the SIF varies very significantly for the parameter range considered — from 4.79 to 134.26. We also observe the expected effect of the length: L only affects the stress intensity factor for $L \lesssim 2$. Finally, our SIF results for $L = 2.4$ are quite close to the analytical solution for the infinite plate case [14]: the maximum difference between our FEM solution and the infinite-plate results is about 0.26% (which might be attributed to either numerical error and/or finite-plate effects).

4.2.3. Example 3.3 The SIF values for this problem are available for selected values of $\mu_2 = R$ in the form of a graph and table generated by numerical pre-solution of the elasticity PDE by a boundary element method [14].

For our finite element approximation we choose $\mathbf{Q}_2, \mathbf{P}_2$ elements for a total of $\mathcal{N} = 3{,}612$ degrees of freedom; the enriched region is the first element ring around the crack tip. Note that in this case, isoparametric mappings are necessary to accurately capture the curvature of the boundary. For this problem, we present our results in the form of a "boundary correction factor"

$$F(d, R, L) = \frac{K(d, R, L)}{\sqrt{d}}, \tag{51}$$

and associated finite element approximation $F_h(d, R, L)$ (for the particular case $\nu = 0.3$); for purposes of presentation, we fix $L = 2.0$. We observe in Figure 12 that our finite element results are in good agreement with the reference results (only) available for $R = 0.1$ and $R = 0.25$.

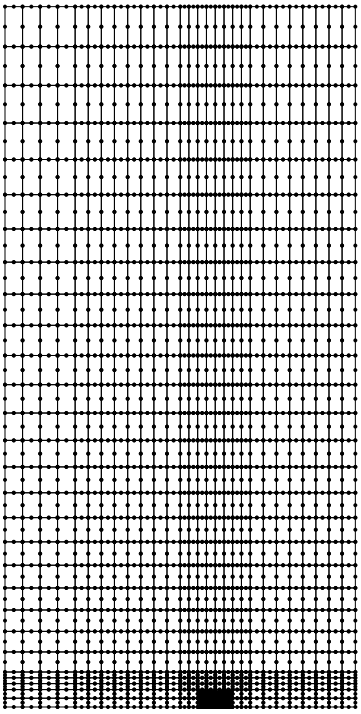


Figure 10. Example 3.2: Finite element mesh.

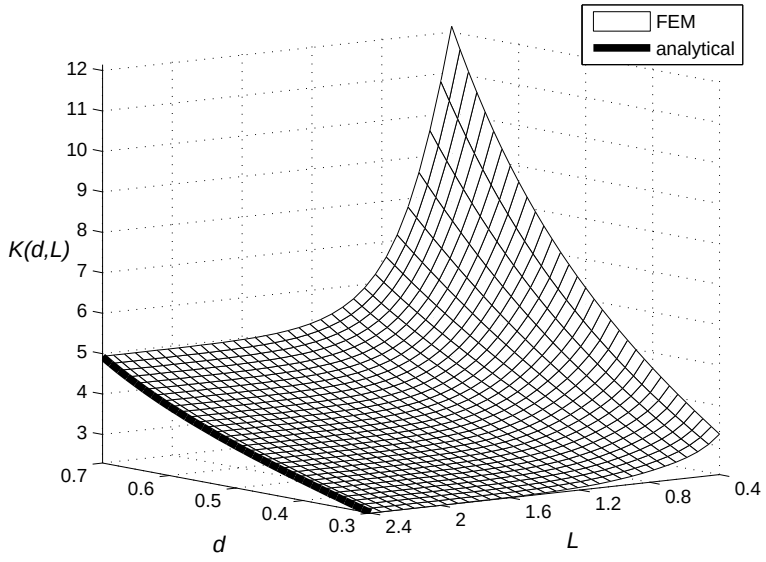


Figure 11. Example 3.2: Stress intensity factor results.

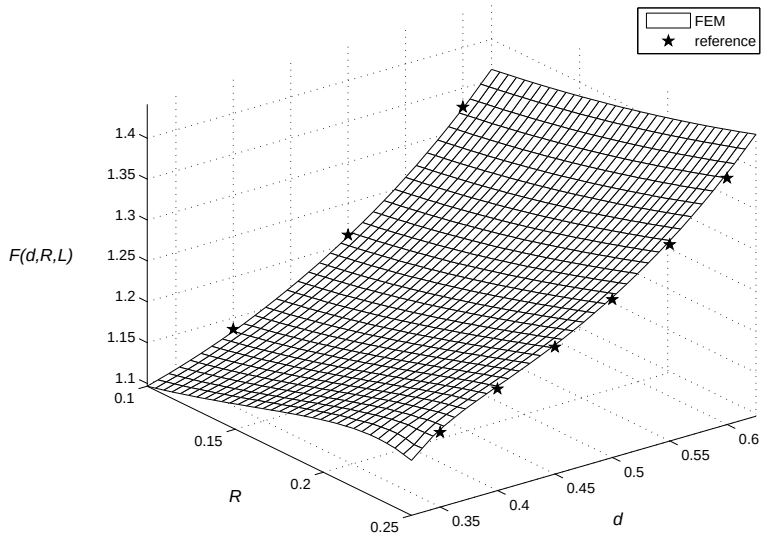


Figure 12. Example 3.3: Boundary correction factor results for $L = 2.0$.

4.3. Truth Approximation

4.3.1. Definition We shall build our reduced basis approximation upon, and measure the error in our reduced basis approximation relative to, a fixed "truth" finite element approximation $\mathcal{U}_{h_t} \in \mathcal{X}_{h_t}$ — a surrogate for the exact solution, $\mathcal{U}$. (In fact, we can also interpret the reduced basis development as an approximation to (any given) finite element model $\mathcal{U}_h$; however, in practice, the reduced basis approximation will only be relevant if $\mathcal{U}_h$ is suitably close to $\mathcal{U}$.) In general, we must anticipate that $2\mathcal{N}_t = \dim(\mathcal{X}_{h_t})$ will be very large, and we must hence require that our reduced basis approach is mathematically and computationally stable as $\mathcal{N} \to \infty$: to achieve the former we insist upon the correct (continuous-limit) norms and dual norms; to ensure the latter, we insist that our online/deployed operation count and storage is *independent* of $\mathcal{N}$. We may then choose our truth approximation *conservatively*.

We denote our truth approximation expanded space of dimension $2\mathcal{N}_t$ as $\mathcal{X}_{h_t} = X_{h_t} \times X_{h_t}$. We imbue $\mathcal{X}_{h_t}$ with the inner product

$$(\mathcal{W}, \mathcal{V})_{\mathcal{X}_{h_t}} \equiv a_{h_t}(W_+, V_+; \overline{\mu}) + a_{h_t}(W_-, V_-; \overline{\mu}), \quad \forall W = (W_+, W_-), V = (V_+, V_-) \in \mathcal{X}_{h_t}, \tag{52}$$

and induced norm $\|\mathcal{W}\|_{\mathcal{X}_{h_t}} = \sqrt{(\mathcal{W}, \mathcal{W})_{\mathcal{X}_{h_t}}}$. Then, given $\mu \in \mathcal{D}$, we find $\mathcal{U}_{h_t}(\mu) \in \mathcal{X}_{h_t}$ such that

$$\mathcal{A}_{h_t}(\mathcal{U}_{h_t}(\mu), \mathcal{V}; \mu) = \mathcal{F}_{h_t}(\mathcal{V}; \mu), \quad \forall\, \mathcal{V} \in \mathcal{X}_{h_t}; \tag{53}$$

we can then evaluate the energy release rate as

$$\mathcal{G}_{h_t}(\mu) = \mathcal{F}_{h_t}(\mathcal{U}_{h_t}(\mu); \mu). \tag{54}$$

In practice, we construct $\mathcal{U}_{h_t}(\mu)$ following the "triangular" procedure described at the end of Section 4.1.

Finally, we define the truth inf-sup and continuity "constants" as

$$\beta_{h_t}(\mu) \equiv \inf_{\mathcal{W}\in\mathcal{X}_{h_t}} \sup_{\mathcal{V}\in\mathcal{X}_{h_t}} \frac{\mathcal{A}_{h_t}(\mathcal{W},\mathcal{V};\mu)}{\|\mathcal{W}\|_{\mathcal{X}_{h_t}}\|\mathcal{V}\|_{\mathcal{X}_{h_t}}}, \tag{55}$$

and

$$\gamma_{h_t}(\mu) \equiv \sup_{\mathcal{W}\in\mathcal{X}_{h_t}} \sup_{\mathcal{V}\in\mathcal{X}_{h_t}} \frac{\mathcal{A}_{h_t}(\mathcal{W},\mathcal{V};\mu)}{\|\mathcal{W}\|_{\mathcal{X}_{h_t}}\|\mathcal{V}\|_{\mathcal{X}_{h_t}}}, \tag{56}$$

respectively. As we shall see in Section 6.2, we in fact construct a lower bound for β_{h_t} as part of our error estimation procedure, and hence ultimately we will numerically prove well-posedness of (at least) the truth approximation.

4.3.2. Model problems We now briefly characterize the truth approximations for each of our three examples.

Example 3.1 Our truth approximation space $\mathcal{X}_{h_t}$ is of dimension $2\mathcal{N}_t = 5{,}868$. The mesh is refined near the crack tip, and the enriched region is chosen as the first element ring around the crack tip ($\mathcal{I}_M$ contains $M = 27$ nodes). Although relatively coarse, this space gives very good accuracy: the maximum error (in output) relative to the reference mesh introduced in Section 4.2.1 is only 1.4%.

Example 3.2 Our truth approximation space $\mathcal{X}_{h_t}$ is of dimension $2\mathcal{N}_t = 8{,}990$. The mesh is refined near the crack tip, and the enriched region is chosen as the first element ring around the crack tip ($\mathcal{I}_M$ contains $M = 27$ nodes). We demonstrated the accuracy of this finite element space in Section 4.2.2; the numerical results are indistinguishable from the theoretical results (for large L).

Example 3.3 Our truth approximation space $\mathcal{X}_{h_t}$ is of dimension $2\mathcal{N}_t = 7{,}224$. The mesh is refined near the crack tip, and the enriched region is chosen as the first element ring around the crack tip ($\mathcal{I}_M$ contains $M = 27$ nodes). The accuracy of this finite element space is demonstrated in Section 4.2.3; our numerical predictions are in good agreement with the available reference results.

5. REDUCED BASIS APPROXIMATION

The reduced basis approximation was first introduced in the late 1970s [5, 6] for nonlinear structural analysis. The method has subsequently been analyzed and developed for a much larger class of parametrized PDEs [7, 8, 9, 10, 16, 18, 23, 24] — in particular, linear/nonlinear elliptic/parabolic PDEs — relevant to many engineering applications. The reduced basis effects dimension reduction by recognizing and exploiting the fact that the field variable does not only belong to the infinite-dimensional (or very large truth) space associated with the underlying PDE, but in fact resides on a much lower dimensional manifold induced by the parametric dependence.

In this section we shall develop the reduced basis approximation for our expanded formulation — a non–coercive symmetric elliptic PDE with "compliant" output. As already indicated, the reduced basis is constructed not as an approximation to the exact solution $\mathcal{U}(\mu)$, but rather as an approximation to the finite element truth approximation $\mathcal{U}_{h_t}(\mu)$.

5.1. Approximation

We shall denote by $2N$ the dimension of the reduced basis approximation space; we shall denote by $2N_{\max}$ the upper limit on the dimension of the reduced basis space — $N_{\max}$ determines the maximum reduced basis accuracy that can be achieved. We next introduce a set of nested samples[†] in parameter space,

$$S_N = \{\mu_1 \in \mathcal{D}, \ldots, \mu_N \in \mathcal{D}\}, \quad 1 \le N \le N_{\max}, \tag{57}$$

and an associated set of nested/hierarchical Lagrangian reduced basis approximation spaces

$$W_N = \text{span}\{(U_{h_t+}(\mu_n), 0), (0, U_{h_t-}(\mu_n)), 1 \le n \le N\}, \quad 1 \le N \le N_{\max}; \tag{58}$$

here $\mathcal{U}_{h_t}(\mu) = (U_{h_t-}(\mu), U_{h_t+}(\mu))(\mu)$ is the solution to (53) for $\mu = \mu_n$. Note that $W_N \subset \mathcal{X}_{h_t}$, and hence our reduced basis approximation is "conforming."

For our basis, we orthogonalize the snapshots with respect to the inner product $(\cdot,\cdot)_{\mathcal{X}_{h_t}}$ in order to preserve the good "conditioning" of the underlying PDE:

$$W_N = \text{span}\{\zeta_m, 1 \le m \le 2N\}, \quad 1 \le N \le N_{\max}, \tag{59}$$

where the basis functions ζ_n (respectively, ζ_{n+N}) are generated from the $(\mathcal{U}_{h_t+}(\mu_n), 0)$ (respectively, $(0, \mathcal{U}_{h_t-}(\mu_n))$) by a Gram-Schmidt process with respect to the inner product $(\cdot,\cdot)_{\mathcal{X}_{h_t}}$. Thus (for our inner product, (52))

$$(\zeta_i, \zeta_j)_{\mathcal{X}_{h_t}} = \delta_{ij}, \quad 1 \le i, j \le 2N_{\max}, \tag{60}$$

where δ_{ij} is the Kronecker delta.

Our reduced basis approximation $\mathcal{U}_N(\mu)$ is then obtained by a standard Galerkin projection (for other options, see [17]): for any $\mu \in \mathcal{D}$, we find $\mathcal{U}_N(\mu) \in W_N$ such that

$$\mathcal{A}_{h_t}(\mathcal{U}_N(\mu), \mathcal{V}; \mu) = \mathcal{F}_{h_t}(\mathcal{V}; \mu), \quad \forall\, \mathcal{V} \in W_N; \tag{61}$$

the reduced basis approximation $\mathcal{G}_N(\mu)$ to $\mathcal{G}_{h_t}(\mu)$ can then be evaluated as

$$\mathcal{G}_N(\mu) = \mathcal{F}_{h_t}(\mathcal{U}_N(\mu); \mu). \tag{62}$$

It is clear from the definition of W_N that the reduced basis space contains "snapshots" on the parametrically induced manifold $\mathcal{M}_{h_t} \equiv \{\mathcal{U}_{h_t}(\mu) \,|\, \mu \in \mathcal{D}\}$. We can anticipate that $\mathcal{M}_{h_t}$ is very *low-dimensional* and *smooth*, and hence $\mathcal{U}_N(\mu) \to \mathcal{U}_{h_t}(\mu)$ (and $\mathcal{G}_N(\mu) \to \mathcal{G}_{h_t}(\mu)$) very rapidly; we should thus realize $N \ll \mathcal{N}_t$.

5.2. A priori theory

We define our reduced basis inf-sup parameter $\beta_N(\mu) \in \mathbb{R}$ and continuity parameter $\gamma_N(\mu) \in \mathbb{R}$ as

$$\beta_N(\mu) \equiv \inf_{\mathcal{W}_N \in W_N} \sup_{\mathcal{V}_N \in W_N} \frac{\mathcal{A}_{h_t}(\mathcal{W}_N, \mathcal{V}_N; \mu)}{\|\mathcal{W}_N\|_{\mathcal{X}_{h_t}} \|\mathcal{V}_N\|_{\mathcal{X}_{h_t}}}, \tag{63}$$

[†]Details of the procedure by which we choose a *good* parameter set S_N will be discussed in Section 5.4.

and

$$\gamma_N(\mu) \equiv \sup_{\mathcal{W}_N \in W_N} \sup_{\mathcal{V}_N \in W_N} \frac{\mathcal{A}_{h_t}(\mathcal{W}_N, \mathcal{V}_N; \mu)}{\|\mathcal{W}_N\|_{\mathcal{X}_{h_t}} \|\mathcal{V}_N\|_{\mathcal{X}_{h_t}}}, \tag{64}$$

respectively; we shall make the assumption that, $0 < \beta_{0,N} \le \beta_N(\mu) \le \gamma_N(\mu) \le \gamma_{0,N} < \infty$, $\forall \mu \in \mathcal{D}$, which in turn implies that our discrete problem is well-posed. In practice, it is difficult to prove the condition $\beta_N(\mu) \ge \beta_{0,N} > 0$, $\forall \mu \in \mathcal{D}$; however we can provide numerical evidence. We consider Example 3.1 and present in Figure 13 $\beta_{h_t}(\mu)$ and $\beta_N(\mu)$ for the particular case $N = 20$. It is clear that both $\beta_{h_t}(\mu)$ and $\beta_N(\mu)$ are bounded away from zero; in particular, we observe $\min_{\mu \in \mathcal{D}} \beta_{h_t}(\mu) = 0.1417$ and $\min_{\mu \in \mathcal{D}} \beta_N(\mu) = 0.1431$. We also note that in this case $\beta_N(\mu) \ge \beta_{h_t}(\mu)$, $\forall\, \mu \in \mathcal{D}$, but this of course need not necessarily be the case due to the "inner sup."

We can show [27] that our reduced basis approximation $\mathcal{U}_N(\mu)$ is optimal in the $\mathcal{X}_{h_t}$-norm,

$$\|\mathcal{U}_{h_t}(\mu) - \mathcal{U}_N(\mu)\|_{\mathcal{X}_{h_t}} \le \left(1 + \frac{\gamma_{0,N}}{\beta_{0,N}}\right) \min_{\mathcal{W}_N \in W_N} \|\mathcal{U}_{h_t}(\mu) - \mathcal{W}_N\|_{\mathcal{X}_{h_t}}, \tag{65}$$

and that our compliance output approximation converges "quadratically,"

$$|\mathcal{G}_{h_t}(\mu) - \mathcal{G}_N(\mu)| \le \gamma_{0,N} \left(1 + \frac{\gamma_{0,N}}{\beta_{0,N}}\right) \min_{\mathcal{W}_N \in W_N} \|\mathcal{U}_{h_t}(\mu) - \mathcal{W}_N\|^2_{\mathcal{X}_{h_t}}. \tag{66}$$

To prove (65), we invoke (53), (61), (63), and (64) to obtain

$$\|\mathcal{U}_N(\mu) - \mathcal{W}_N\|_{\mathcal{X}_{h_t}} \le \frac{\gamma_N(\mu)}{\beta_N(\mu)} \|\mathcal{U}_{h_t}(\mu) - \mathcal{W}_N\|_{\mathcal{X}_{h_t}}, \quad \forall\, \mathcal{W}_N \in W_N; \tag{67}$$

the proof directly follows from the triangle inequality. To prove (66), we invoke (54), (61), the symmetry of $\mathcal{A}_{h_t}$, (64), and optimality of $\mathcal{U}_N(\mu)$, (65).

Unfortunately, at present we have no rigorous *a priori* result for the best approximation that appears in (65), (66). We can, however, demonstrate that $\mathcal{U}_{h_t}(\mu)$ is indeed smooth in μ, and hence rapid convergence is plausible (and, as we shall see, realized).

5.3. Offline/Online computational procedure

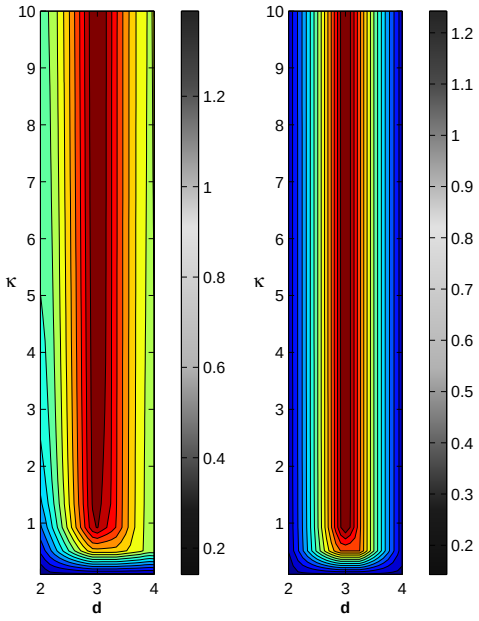
As we have argued, it is plausible that in order to obtain an accurate reduced basis approximation $\mathcal{U}_N(\mu)$, the dimension of W_N, $2N$ can be quite small. However, since the elements of W_N are still "large" in some sense (the representation of ζ_n is of length $2\mathcal{N}_t$), the computational savings are not self-evident. In this section we develop an offline-online procedure that will enable us to evaluate our approximations in real-time.

To begin, we expand our reduced basis approximation as

$$\mathcal{U}_N(\mu) = \sum_{j=1}^{2N} \mathcal{U}_{N\,j}(\mu)\, \zeta_j. \tag{68}$$

By choosing $\mathcal{V} = \zeta_i, i = 1, \ldots, 2N$, in (61), it follows that the coefficients $\mathcal{U}_{N\,j}(\mu)$ satisfy the $2N \times 2N$ linear algebraic system

$$\sum_{j=1}^{2N} \left\{ \sum_{q=1}^{Q^{\mathcal{A}}} \Theta_q^{\mathcal{A}}(\mu) \mathcal{A}_{h_t q}(\zeta_j, \zeta_i) \right\} \mathcal{U}_{N\,j}(\mu) = \sum_{q=1}^{Q^{\mathcal{F}}} \Theta_q^{\mathcal{F}}(\mu) \mathcal{F}_{h_t q}(\zeta_i), 1 \le i \le 2N; \tag{69}$$



(a) (b)

Figure 13. Example 3.1: (a) Inf-sup parameter $\beta_{h_t}(\mu)$, and (b) Inf-sup parameter $\beta_N(\mu)$ for the case $N = 20$.

this representation is a direct result of the affine decomposition of a and f. The reduced basis output can then be calculated as

$$\mathcal{G}_N(\mu) = \sum_{j=1}^{2N} \left\{ \sum_{q=1}^{Q^{\mathcal{F}}} \Theta_{\mathcal{F}}^q(\mu) \mathcal{F}_{h_t q}(\zeta_j) \right\} \mathcal{U}_{N\,j}(\mu). \tag{70}$$

The offline/online computational strategy is now clear.

In the *offline* stage — performed *once* — we first solve for the ζ_j, $1 \le j \le 2N_{\max}$; we then form and store $\mathcal{A}_{h_t q}(\zeta_i, \zeta_j)$, $1 \le i, j \le 2N_{\max}$, $1 \le q \le Q^{\mathcal{A}}$; and finally we form and store $\mathcal{F}_{h_t q}(\zeta_j)$, $1 \le j \le 2N_{\max}$, $1 \le q \le Q^{\mathcal{F}}$. Note all quantities computed in this stage are *independent* of the parameter μ. This stage requires, in addition to the calculation of $N_{\max}$ truth solutions $\mathcal{U}_{h_t}$, $O(4Q^{\mathcal{A}} N_{\max}^2 \mathcal{N}_t)$ operations and $O(4Q^{\mathcal{A}} N_{\max}^2 + 4Q^{\mathcal{F}} N_{\max})$ "permanent" storage. Note the operation count is dominated by the formation of the $\mathcal{A}_{h_t q}(\zeta_i, \zeta_j)$ inner products, once the ζ_j have been obtained; in the former we exploit the sparsity in the finite element stiffness matrix.

In the *online* stage — performed *many times*, for each new value of μ — we first assemble and solve the $2N \times 2N$ linear algebraic system (69) to obtain $\mathcal{U}_{N\,j}$, $1 \le j \le 2N$; we next perform the summation in (70) to obtain $\mathcal{G}_N(\mu)$. The operation count in this stage is $O(4Q^{\mathcal{A}} N^2 + 2Q^{\mathcal{F}} N)$ and $O(8N^3)$ to assemble and solve the linear algebraic system, respectively, and finally $O(4Q^{\mathcal{F}} N^2)$ to evaluate the output. The essential point is that the complexity of the online stage is completely independent of $\mathcal{N}_t$; since $N \ll \mathcal{N}_t$, we expect — in the online/deployed

stage — significant computational savings relative to the classical direct approach. (Of course, the offline effort is considerable, and hence we must be in the many-query or real-time context to justify the reduced basis approach.)

5.4. Sample Construction

We now discuss the construction of the nested samples S_N and the associated reduced basis spaces W_N. For a given N, we would like to find the optimal sample S_N that maximizes our reduced basis accuracy. We cannot find the best space — a very difficult problem — however we can hope to find a good sub-optimal result. We shall assume here that we are able to calculate a rigorous, sharp and *inexpensive* (in the limit of many queries) upper bound $\Delta_N^{\mathcal{G}}(\mu)$ for $|\mathcal{G}_{h_t}(\mu) - \mathcal{G}_N(\mu)|$. (The procedure to construct our error estimators will be discussed in detail in Section 6.) We also shall need $\Xi_{\text{train}} \subset \mathcal{D}$, a sample of size $n_{\text{train}} \gg 1$ over the parameter domain $\mathcal{D}$.

We shall pursue a greedy algorithm [23] to find our set S_N. We denote the smallest error tolerance anticipated (in the online stage) by $\epsilon_{\text{tol,min}}$. We assume that we are given a sample S_N, and hence space W_N and associated reduced basis procedure/approximation $\mathcal{U}_N(\mu)$, $\forall\, \mu \in \mathcal{D}$; in practice, we choose $S_1 = \mu_1$ randomly. We then calculate

$$\begin{aligned} \epsilon_N^* &= \max_{\mu \in \Xi_{\text{train}}} \Delta_N^{\mathcal{G}}(\mu), \\ \mu_{N+1}^* &= \arg \max_{\mu \in \Xi_{\text{train}}} \Delta_N^{\mathcal{G}}(\mu); \end{aligned} \tag{71}$$

we next append μ_{N+1}^* to our sample S_N to form S_{N+1}, and hence W_{N+1}; we now continue this process until $N = N_{\max}$ such that $\epsilon_{N_{\max}}^* \le \epsilon_{\text{tol,min}}$. The strategy is rather heuristic, but in practice works very well.

The crucial point of this strategy is that the error estimator $\Delta_N^{\mathcal{G}}(\mu)$ can be computed "online-inexpensively," with marginal complexity independent of $\mathcal{N}_t$, and hence the cost in each greedy step does *not* scale as $n_{\text{train}}\mathcal{N}_t$. This permits us to perform a very exhaustive ($n_{\text{train}} \gg 1$) and inexpensive search for the best sample S_N: we compute the truth solution only for the winning candidate, namely, at the new sample point μ_{N+1}^*. We thus need only compute $N_{\max}$ truth solutions for our greedy process — plus an additional $O(N_{\max}Q^{\mathcal{A}})$ truth "pseudo–solutions" required by the error bounds. We shall discuss our error bound computation, and provide quantitative timing results, in Section 6.

5.5. Numerical Results

5.5.1. Example 3.1 We now apply the reduced basis approximation to Example 3.1. We first pursue the optimal sampling procedure described in the previous section for a random parameter train sample Ξ_{train} of size $n_{\text{train}} = 1,000$; we obtain $\epsilon_{N_{\max}}^* = 1.8\text{E-}04$ for $N_{\max} = 25$. We observe in Figure 14 that the distribution of parameter points is denser in the region $\mu_2 \le 2.0$; also there are more sample points near the boundary $\mu_1 = 2.0$. This is because of the effect of larger values of $\widetilde{E}_1$ (hence larger μ_2 values) on the stress intensity factor is not significant; furthermore, the effect becomes smaller as the distant between the crack tip and the first material layer increases (which is equivalent to large μ_1 values).

We next present, in Table V, the convergence of our output $\mathcal{G}_N(\mu)$. Here, the error $\mathtt{E}_N$ is defined as the maximum of the relative error, $|\mathcal{G}_{h_t}(\mu) - \mathcal{G}_N(\mu)|/|\mathcal{G}_{h_t}(\mu)|$, over a random parameter sample $\Xi_{\text{test}} \subset \mathcal{D}$ of size $n_{\text{test}} = 1{,}089$. We observe that the reduced

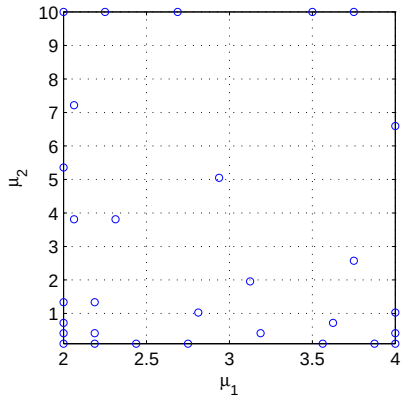


Figure 14. Example 3.1: Optimal sample set $S_{N_{\max}}$.

basis approximation converges very quickly. The test parameter sample Ξ_{test} here is introduced only to provide an independent sample with which to quantify the quality of the approximation/error bound. We emphasize that neither Ξ_{train} (for sample construction) nor Ξ_{test} (for presentation) are exhaustive, which in turn highlights the importance of online error bounds in which we (efficiently) evaluate $\Delta_N^{\mathcal{G}}(\mu)$ for any given value μ.

N	$\mathtt{E}_N$	Online Time
		$\mathcal{G}_N$
5	1.33E+00	3.96E-05
10	1.34E-02	8.27E-05
15	1.27E-03	1.14E-04
20	1.43E-04	1.82E-04
25	2.94E-05	2.54E-04

Table V. Example 3.1: Reduced basis convergence result and online time to evaluate $\mathcal{G}_N$ as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $\mathcal{G}_{h_t}$.

We also present in Table V the online reduced basis computational cost to evaluate $\mathcal{G}_N(\mu)$ compared to the finite element cost to evaluate $\mathcal{G}_{h_t}$ for any given μ. The computational savings is very high: our online evaluation is about 10,000 times faster than the conventional evaluation; moreover, it is seen in Table V that the reduced basis solution $\mathcal{G}_N(\mu)$ and the "truth" solution $\mathcal{G}_{h_t}(\mu)$ are indistinguishable for $N \geq 20$. Of course this timing comparison does not include the reduced basis offline effort and hence the results are only relevant to the many-query and real-time contexts. (We present offline costs in Section 6.4.)

5.5.2. Example 3.2 We show in Table VI the convergence rate of the reduced basis approximation $\mathcal{G}_N(\mu)$ for Example 3.2; the error $\mathtt{E}_N$ is defined as the maximum of the relative error, $|\mathcal{G}_{h_t}(\mu) - \mathcal{G}_N(\mu)|/|\mathcal{G}_{h_t}(\mu)|$, over a random parameter test sample $\Xi_{\text{test}} \subset \mathcal{D}$ of size

$n_{\text{test}} = 1{,}000$. We observe very rapid convergence with N: we need only $N = 20$ to obtain 10^{-5} accuracy. We also present in Table VI the online reduced basis computational cost to evaluate $\mathcal{G}_N(\mu)$ compared to the finite element cost to evaluate $\mathcal{G}_{h_t}$ for any given μ. The computational savings is very high: our online evaluation is about 10,000 times faster than the conventional evaluation; moreover, it is seen in Table V that the reduced basis solution $\mathcal{G}_N(\mu)$ and the "truth" solution $\mathcal{G}_{h_t}(\mu)$ are indistinguishable for $N \geq 15$.

N	$\mathtt{E}_N$	Online Time
		$\mathcal{G}_N$
5	1.69E-01	1.56E-05
10	4.22E-03	2.37E-05
15	5.23E-04	3.78E-05
20	8.80E-05	5.69E-05
25	1.09E-05	8.96E-05
30	1.11E-06	1.10E-04

Table VI. Example 3.2: Reduced basis convergence result and online time to evaluate $\mathcal{G}_N$ as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $\mathcal{G}_{h_t}$.

5.5.3. Example 3.3 The convergence rate for the reduced basis approximation $\mathcal{G}_N(\mu)$ for Example 3.3 is shown in Table VII. The error $\mathtt{E}_N$ is the maximum of the relative error, $|\mathcal{G}_{h_t}(\mu) - \mathcal{G}_N(\mu)|/|\mathcal{G}_{h_t}(\mu)|$ over a random parameter test sample $\Xi_{\text{test}} \subset \mathcal{D}$ of size $n_{\text{test}} = 3{,}000$. We observe very rapid convergence with N: even for this 3-parameter ($P = 3$) case, we need only $N = 30$ to obtain 10^{-4} accuracy. The computational savings is still very high (despite the large number of terms in our affine expansion): our online evaluation is about 10,000 times faster than the conventional evaluation; moreover, it is shown as in Table V that the reduced basis solution $\mathcal{G}_N(\mu)$ and the "truth" solution $\mathcal{G}_{h_t}(\mu)$ are indistinguishable for $N \geq 35$.

N	$\mathtt{E}_N$	Online Time
		$\mathcal{G}_N$
5	1.04E-01	1.80E-05
10	6.01E-02	2.32E-05
20	9.08E-03	5.95E-05
30	2.36E-03	1.92E-04
35	9.42E-05	2.12E-04
40	4.58E-05	2.78E-04

Table VII. Example 3.3: Reduced basis convergence result and online time to evaluate $\mathcal{G}_N$ as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $\mathcal{G}_{h_t}$.

6. *A POSTERIORI* ERROR ESTIMATION

From the last chapter we know that we can obtain $\mathcal{G}_N(\mu)$ very inexpensively: the online computational effort is independent of $\mathcal{N}_t$; and the dimension of the reduced basis space N, can be chosen quite small. But we do not know how small N should be chosen in order for the reduced basis method to produce the desired accuracy for all $\mu \in \mathcal{D}$: if N is too small our reduced basis approximation is unacceptably inaccurate — and may compromise safety in failure-related decisions; if N is too large our reduced basis approximation is unnecessarily expensive — and may compromise a timely response in critical situations. We thus need an *a posteriori* error estimator for $\mathcal{G}_N(\mu)$: not only will the *a posteriori* error estimator help us construct the reduced basis nested parameter set in the offline stage, it will also tell us how well we are doing for *any* $\mu \in \mathcal{D}$ in the online stage. In short, it will ensure that our reduced basis solution is *certified* and *reliable* (at least relative to the "truth" approximation).

Our *a posteriori* error estimator has many ingredients: the expanded formulation for our energy release rate output, the reduced basis approximation, the construction of a lower bound for the inf-sup parameter, the offline/online computational strategy, the application to (greedy) adaptivity, and the fact that we develop an error estimator for an SIF (or SIF–related) output of interest. Our approach shares the latter with the goal-oriented *a posteriori* error estimation adaptivity approaches of [29, 30, 31]. The difference, of course, is the discretization framework — finite element in [29, 30, 31] and reduced basis in our work — which in turn changes the adaptivity framework. (We might ask whether our expanded formulation might improve the finite element results? Perhaps not, as the error estimators in [30, 31] are already quite sharp: the expanded formulation is perhaps more crucial in the reduced basis context.)

In this section we shall discuss the development of *a posteriori* error estimators for reduced basis approximations. We require that the estimators are *inexpensive* — the online (marginal) computational complexity is independent of $\mathcal{N}_t$; *reliable* — an upper bound of the true error; and *sharp* — not overly conservative. We first briefly discuss the construction of a lower bound for the inf-sup parameter $\beta_{h_t}(\mu)$; we then develop our *a posteriori* error estimation procedures for our problem.

6.1. Lower Bound for the Inf-Sup Parameter

Our error bound requires an inexpensive (online) and reasonably accurate lower bound $\beta_{LB}(\mu)$ for the "truth" inf-sup stability parameter $\beta_{h_t}(\mu)$ as defined in (25). (We emphasize that $\beta_{LB}(\mu)$ is not a lower bound for the discrete inf-sup stability parameter $\beta_N(\mu)$ defined in (63).) The construction of the lower bound of $\beta_{LB}(\mu)$ is also decomposed into two stages: the offline stage — performed once, with computational cost depending on $\mathcal{N}_t$ and usually *expensive*; and the online stage — performed many times, with the computational cost independent of $\mathcal{N}_t$ and very *inexpensive*. There are several approaches to efficiently calculate an approximate lower bound of our parametric inf-sup parameter: the "natural norm" technique [25]; and the recently developed Successive Constraint Linear Optimization Method [32].

Here we construct our lower bound for the inf-sup parameter by the "natural norm" technique developed in [25]. We refer to the detailed procedure in [25, 27] and only summarize here the offline/online stages. In the offline stage we identify a sample of $\overline{K}$ parameters in $\mathcal{D}$, $\mathcal{V}^{\overline{K}}$, and we evaluate at each point $\overline{\mu}_k$ in $\mathcal{V}^{\overline{K}}$ a set of singular values and eigenvalues Λ_k, $1 \le k \le \overline{K}$. We can develop a good (positive) local lower bound to β_{h_t} in the neighborhood

of $\overline{\mu}_k \in \mathcal{V}^{\overline{K}}$ based on the Λ_k; $\mathcal{V}^{\overline{K}}$ is chosen to ensure that the union of these neighborhoods cover $\mathcal{D}$. (In practice, $\overline{K}$ is quite small for the technique of [25].) In the online stage, given a new μ, we simply search amongst a few $\overline{\mu}_k \in \mathcal{V}^{\overline{K}}$ near μ for the largest local lower bound; the complexity is independent of $\mathcal{N}_t$, and typically negligible compared to calculation of the dual norm of the residual (see Section 6.3).

6.2. Error bounds

We first define our error bound for the error in the output as

$$\Delta_N^{\mathcal{G}}(\mu) \equiv \frac{\varepsilon_N^2(\mu)}{\beta_{LB}(\mu)}, \tag{72}$$

where $\varepsilon_N(\mu)$ is the dual norm of the residual defined as

$$\varepsilon_N(\mu) = \sup_{\mathcal{V}\in\mathcal{X}_{h_t}} \frac{\mathcal{R}_{h_t}(\mathcal{V};\mu)}{\|\mathcal{V}\|_{\mathcal{X}_{h_t}}}; \tag{73}$$

here

$$\mathcal{R}_{h_t}(\mathcal{V};\mu) \quad = \quad \mathcal{F}_{h_t}(\mathcal{V}) - \mathcal{A}_{h_t}(\mathcal{U}_N(\mu),\mathcal{V};\mu), \quad \forall\, \mathcal{V} \in \mathcal{X}_{h_t}, \tag{74}$$

is the residual associated with $\mathcal{U}_N(\mu)$.

We next define our the effectivity associated with our error bound for the output as

$$\eta_N^{\mathcal{G}}(\mu) \equiv \frac{\Delta_N^{\mathcal{G}}(\mu)}{|\mathcal{G}_{h_t}(\mu) - \mathcal{G}_N(\mu)|}. \tag{75}$$

We can now demonstrate that the output effectivity satisfies

$$\eta_N^{\mathcal{G}}(\mu) \geq 1, \quad \forall\, \mu \in \mathcal{D}; \tag{76}$$

equivalently, $\Delta_N^{\mathcal{G}}(\mu)$ is a rigorous upper bound for the error in our reduced basis output. We cannot provide an upper bound for the effectivity: we return to this point in Section 6.4.

To start, we note that the error $e(\mu) = \mathcal{U}_{h_t}(\mu) - \mathcal{U}_N(\mu)$ satisfies

$$\mathcal{A}_{h_t}(e(\mu),\mathcal{V};\mu) = \mathcal{R}_{h_t}(\mathcal{V};\mu), \quad \forall\, \mathcal{V} \in \mathcal{X}_{h_t}. \tag{77}$$

We also observe from standard duality arguments that,

$$\varepsilon_N(\mu) = \|\hat{e}_{h_t}(\mu)\|_{\mathcal{X}_{h_t}}, \tag{78}$$

where $\hat{e}_{h_t}(\mu) \in \mathcal{X}_{h_t}$ is given by

$$(\hat{e}_{h_t}(\mu),\mathcal{V})_{\mathcal{X}_{h_t}} = \mathcal{R}_{h_t}(\mathcal{V};\mu), \quad \forall\, \mathcal{V} \in \mathcal{X}_{h_t}. \tag{79}$$

We then conclude from (77) and the inf-sup parameter definition that

$$\|e(\mu)\|_{\mathcal{X}_{h_t}} \leq \frac{\|\hat{e}_{h_t}(\mu)\|_{\mathcal{X}_{h_t}}}{\beta_{LB}(\mu)}. \tag{80}$$

Finally, it directly follows from the symmetry of $\mathcal{A}_{h_t}$, the definition of the "compliance" output $\mathcal{G}_{h_t}(\mu)$, Galerkin orthogonality, the dual norm of the residual, and (77) that

$$\begin{aligned}
|\mathcal{G}_{h_t}(\mu) - \mathcal{G}_N(\mu)| &= |\mathcal{F}_{h_t}(e(\mu);\mu)| \\
&= |\mathcal{A}_{h_t}(\mathcal{U}_{h_t}(\mu), e(\mu);\mu)| \\
&= |\mathcal{A}_{h_t}(e(\mu), \mathcal{U}_{h_t}(\mu);\mu)| \\
&= |\mathcal{A}_{h_t}(e(\mu), \mathcal{U}_{h_t}(\mu) - \mathcal{U}_N(\mu);\mu)| \\
&= |\mathcal{A}_{h_t}(e(\mu), e(\mu);\mu)| \\
&= |\mathcal{R}_{h_t}(e(\mu);\mu)| \\
&\leq \varepsilon_N(\mu)\|e(\mu)\|_{\mathcal{X}_{h_t}} \\
&\leq \frac{\varepsilon_N^2(\mu)}{\beta_{LB}(\mu)}, \qquad (81)
\end{aligned}$$

which concludes our proof.

We can further define our approximation of the (non–dimensional) stress intensity factor $K_N(\mu)$ and associated error bounds $\Delta_N^K(\mu)$ based on (12). We assume here that we first verify $\Delta_N^{\mathcal{G}}(\mu) \leq \mathcal{G}_N(\mu)$: for any given μ this condition is readily confirmed (or not) online; in general, the inequality will be satisfied for all but the smallest N. We then define

$$K_N(\mu) = \frac{1}{2\sqrt{1-\nu^2}}\left\{\sqrt{\mathcal{G}_N(\mu) + \Delta_N^{\mathcal{G}}(\mu)} + \sqrt{\mathcal{G}_N(\mu) - \Delta_N^{\mathcal{G}}(\mu)}\right\}, \qquad (82)$$

$$\Delta_N^K(\mu) = \frac{1}{2\sqrt{1-\nu^2}}\left\{\sqrt{\mathcal{G}_N(\mu) + \Delta_N^{\mathcal{G}}(\mu)} - \sqrt{\mathcal{G}_N(\mu) - \Delta_N^{\mathcal{G}}(\mu)}\right\}. \qquad (83)$$

It readily follows [27] that

$$K_N(\mu) - \Delta_N^K(\mu) \leq K_{h_t}(\mu) \leq K_N(\mu) + \Delta_N^K(\mu), \quad \forall\, \mu \in \mathcal{D}, \qquad (84)$$

where $K_{h_t}(\mu) = \sqrt{\mathcal{G}_{h_t}(\mu)/(1-\nu^2)}$. These lower and in particular upper bounds for the SIF are extremely useful in applications that require *highly accurate* and typically *conservative* stress intensity factor evaluations such as Non–Destructive Evaluation (NDE), crack growth prediction, or brittle failure applications [27].

6.3. Offline/Online Computational Procedure

It remains to develop associated offline-online computational procedures for the evaluation of $\Delta_N^{\mathcal{G}}(\mu)$ and in particular $\varepsilon_N(\mu)$, the dual norm of the residual. We begin from our reduced basis approximation $\mathcal{U}_N(\mu) = \sum_{n=1}^{2N} \mathcal{U}_{Nn}(\mu)\zeta_n$ and affine decomposition to express the residual as

$$\mathcal{R}_{h_t}(\mathcal{V};\mu) = \sum_{q=1}^{Q^{\mathcal{F}}} \Theta_q^{\mathcal{F}}(\mu)\mathcal{F}_{h_t q}(\mathcal{V}) - \sum_{q=1}^{Q^{\mathcal{A}}}\sum_{n=1}^{2N} \Theta_q^{\mathcal{A}}(\mu)\mathcal{U}_{Nn}(\mu)\mathcal{A}_{h_t q}(\zeta_n, \mathcal{V}), \quad \forall\, \mathcal{V} \in \mathcal{X}_{h_t}. \qquad (85)$$

It is clear from linear superposition that we can express $\hat{e}_{h_t}(\mu) \in \mathcal{X}_{h_t}$ as

$$\hat{e}_{h_t}(\mu) = \sum_{q=1}^{Q^{\mathcal{F}}} \Theta_q^{\mathcal{F}}(\mu)\mathcal{C}_q + \sum_{q=1}^{Q^{\mathcal{A}}}\sum_{n=1}^{2N} \Theta_q^{\mathcal{A}}(\mu)\mathcal{U}_{Nn}(\mu)\mathcal{L}_{qn}, \qquad (86)$$

where $(\mathcal{C}_q, \mathcal{V})_{\mathcal{X}_{h_t}} = \mathcal{F}_{h_t q}(\mathcal{V}),\ \forall\, \mathcal{V} \in \mathcal{X}_{h_t},\ 1 \le q \le Q^{\mathcal{F}}$, and $(\mathcal{L}_{qn}, \mathcal{V})_{\mathcal{X}_{h_t}} = -\mathcal{A}_{h_t q}(\zeta_n, \mathcal{V}),\ \forall\, \mathcal{V} \in \mathcal{X}_{h_t},\ 1 \le n \le 2N,\ 1 \le q \le Q^{\mathcal{A}}$. Note that the latter problems are simple parameter-independent Poisson problems (albeit over a somewhat complicated enriched space); the $\mathcal{C}_q$, $\mathcal{L}_q$ are the truth "pseudo–solutions" to which we alluded in Section 5.4.

It thus directly follows that

$$\begin{aligned}\|\hat{e}_{h_t}(\mu)\|^2_{\mathcal{X}_{h_t}} \;=\; & \sum_{q=1}^{Q^{\mathcal{F}}}\sum_{q'=1}^{Q^{\mathcal{F}}} \Theta^{\mathcal{F}}_q(\mu)\Theta^{\mathcal{F}}_{q'}(\mu)(\mathcal{C}_q, \mathcal{C}_{q'})_{\mathcal{X}_{h_t}} \\ & + 2\sum_{q=1}^{Q^{\mathcal{F}}}\sum_{q'=1}^{Q^{\mathcal{A}}}\sum_{n=1}^{2N} \Theta^{\mathcal{F}}_q(\mu)\Theta^{\mathcal{A}}_{q'}(\mu)\mathcal{U}_{Nn}(\mu)(\mathcal{C}_q, \mathcal{L}_{q'n})_{\mathcal{X}_{h_t}} \\ & + \sum_{q=1}^{Q^{\mathcal{A}}}\sum_{q'=1}^{Q^{\mathcal{A}}}\sum_{n=1}^{2N}\sum_{n'=1}^{2N} \Theta^{\mathcal{A}}_q(\mu)\Theta^{\mathcal{A}}_{q'}(\mu)\mathcal{U}_{Nn}(\mu)\mathcal{U}_{Nn'}(\mu)(\mathcal{L}_{qn}, \mathcal{L}_{q'n'})_{\mathcal{X}_{h_t}}, \quad (87)\end{aligned}$$

in terms of which we can then evaluate $\varepsilon_N(\mu) = \sqrt{\|\hat{e}_{h_t}(\mu)\|^2_{\mathcal{X}_{h_t}}}$. The expression (87) is simply a summation of products of parameter-dependent functions and parameter-independent inner products. The offline-online decomposition is now clear.

In the offline stage, we first obtain the "pseudo–solutions" $\mathcal{C}_q,\ 1 \le q \le Q^{\mathcal{F}}$, and $\mathcal{L}_{qn},\ 1 \le n \le 2N_{\max},\ 1 \le q \le Q^{\mathcal{A}}$; we then perform and store the parameter-independent inner products, $(\mathcal{C}_q, \mathcal{C}_{q'})_{\mathcal{X}_{h_t}},\ 1 \le q, q' \le Q^{\mathcal{F}}$, $(\mathcal{C}_q, \mathcal{L}_{q'n})_{\mathcal{X}_{h_t}},\ 1 \le n \le 2N_{\max},\ 1 \le q \le Q^{\mathcal{F}},\ 1 \le q' \le Q^{\mathcal{A}}$, and $(\mathcal{L}_{qn}, \mathcal{L}_{q'n'})_{\mathcal{X}_{h_t}},\ 1 \le n, n' \le 2N_{\max},\ 1 \le q, q' \le Q^{\mathcal{A}}$. This requires $O(2N_{\max}Q^{\mathcal{A}} + Q^{\mathcal{F}})$ "truth" pseudo–solutions and $O(4N^2_{\max}(Q^{\mathcal{A}})^2 + 2N_{\max}Q^{\mathcal{A}}Q^{\mathcal{F}} + (Q^{\mathcal{F}})^2)$ "$\mathcal{N}_t$ inner products."

In the online stage, given a new parameter value μ, we simply evaluate the sum (87) in terms of $\Theta^{\mathcal{A}}_q(\mu)$, $\Theta^{\mathcal{F}}_q(\mu)$, and $\mathcal{U}_{Nn}(\mu)$ and the *pre-computed* parameter-independent inner products. The operation count for this stage is only $O(N^2(4Q^{\mathcal{A}})^2 + 2NQ^{\mathcal{A}}Q^{\mathcal{F}} + (Q^{\mathcal{F}})^2)$ — independent of $\mathcal{N}_t$. We do note that for our more complicated (affine) geometric mappings, $Q^{\mathcal{A}}$ can be quite large: we thus expect — due to the $(Q^{\mathcal{A}})^2$ scaling — that $\Delta^{\mathcal{G}}_N$ will be more expensive to evaluate (online) than $\mathcal{G}_N$; we confirm this in the next section.

We also note that the linear superposition (87) involves the extensive cancellation of large terms, and hence round-off errors may "pollute" our error estimator if $\|\hat{e}_{h_t}(\mu)\|_{\mathcal{X}_{h_t}}$ is too small. The latter often obtains when the parameter μ is too near an element of the set S_N. In practice this precision issue is not a major problem in particular for compliant outputs, as it usually arises only when $\Delta^{\mathcal{G}}_N$ is already much smaller than typically required in the engineering context.

6.4. Numerical results

Example 3.1 We present in Table VIII the error bounds and effectivities for Example 3.1 as a function of N. The error bound reported, $\mathcal{E}_N$, is the maximum of the relative error bound, $\Delta^{\mathcal{G}}_N(\mu)/|\mathcal{G}_{h_t}(\mu)|$, over the same test sample Ξ_{test} of Section 5.5. We denote by $\overline{\eta}^{\mathcal{G}}_N$ the average of the effectivity, $\eta^{\mathcal{G}}_N(\mu)$, over Ξ_{test}. (Note that we reject from our sample Ξ_{test} those parameter points for which the error bound $\Delta^{\mathcal{G}}_N(\mu)$ is smaller than machine precision, as for these parameter points the calculation is contaminated by round-off.) We observe relatively good effectivity: our $\overline{\eta}^{\mathcal{G}}_N$ is usually of order $O(10\text{–}100)$; given the very rapid convergence of the

N	$\mathtt{E}_N$	$\mathcal{E}_N$	$\overline{\eta}_N^{\mathcal{G}}$	Online Time	
				$\mathcal{G}_N$	$\Delta_N^{\mathcal{G}}$
5	1.33E+00	1.70E+01	49.93	3.96E-05	1.40E-03
10	1.34E-02	4.96E+00	69.19	8.27E-05	1.80E-03
15	1.27E-03	9.99E-02	47.29	1.14E-04	2.18E-03
20	1.43E-04	1.97E-02	60.74	1.82E-04	2.41E-03
25	2.94E-05	3.24E-03	53.42	2.54E-04	2.79E-03

Table VIII. Example 3.1: Reduced basis error bound and effectivity, and online time to evaluate $\mathcal{G}_N$, $\Delta_N^{\mathcal{G}}$, as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $\mathcal{G}_{h_t}$.

reduced basis approximation, $O(50)$ effectivities are not too worrisome — readily absorbed by only a relatively slight increase in N. The "expanded" formulation improves our error estimator by converting the output into linear and compliant form: our effectivity is significantly better than the $O(100\text{–}10{,}000)$ effectivities of the "classical" formulation [26, 27], in which the output is *nonlinear*.

The computational savings are very high: one online evaluation for an output and the associated error bound for $N = 25$ is about 40–45ms compared to approximately 6–8s for a "truth" solution; our online evaluation (even with error estimation) is still about 140–200 times faster than conventional techniques. We also note that the timing scales as expected: the computational time required to compute $\mathcal{G}_N$ is smaller than, but grows faster than, the time required to compute $\Delta_N^{\mathcal{G}}$ — this reflects the respective computational costs of order $O(8N^3)$ and $O(4(Q^{\mathcal{A}})^2N^2)$. We do not necessarily observe exactly the expected scalings with N due to memory access and other overhead not accounted for in our complexity estimates; the correct scalings are expected for larger N, for which the overhead is relatively less.

The total offline computation time for our reduced basis and error bounds for $N = 15$ — including all 259 truth solutions and "pseudo-solutions" and all $\mathcal{N}_t$ inner product and greedy searches (71) — is 287; recall that all computations times are normalized with respect to the time to calculate one truth solution. Clearly, in this case, the offline time is dominated by the truth (pseudo–)solutions: indeed, the time for the search (71) (per greedy cycle) is only 0.12; the latter emphasizes the importance of the *a posteriori* error estimators even in the offline stage — and in this particular case suggests that we can easily conduct a more exhaustive search (i.e, choose n_{train} larger). For our particular example, we conclude that "many queries" — the point at which the reduced basis approach make sense — should be interpreted as roughly 300 queries. (Of course, in the real time context, the criterion is different.)

Example 3.2 We present in Table IX the error bounds and effectivities for Example 3.2 as a function of N. The error bound reported, $\mathcal{E}_N$, is the maximum of the relative error bound, $\Delta_N^{\mathcal{G}}(\mu)/|\mathcal{G}_{h_t}(\mu)|$ over the same test sample Ξ_{test} of Section 5.5. We denote by $\overline{\eta}_N^{\mathcal{G}}$ the average of the effectivity, $\eta_N^{\mathcal{G}}(\mu)$, over Ξ_{test}. (As before, we reject from our sample Ξ_{test} those parameter points for which the error bound $\Delta_N^{\mathcal{G}}(\mu)$ is smaller than machine precision, as for these parameter points the calculation is contaminated by round-off.)

We observe relatively good effectivity: our $\overline{\eta}_N^{\mathcal{G}}$ is usually of order $O(10\text{–}100)$. We also note that the reduced basis error bound is rather "crude" for small N (in particular, $N \leq 10$);

N	$\mathtt{E}_N$	$\mathcal{E}_N$	$\overline{\eta}_N^{\mathcal{G}}$	Online Time	
				$\mathcal{G}_N$	$\Delta_N^{\mathcal{G}}$
5	1.69E-01	2.20E+01	60.89	1.56E-05	2.06E-03
10	4.22E-03	3.40E-01	60.75	2.37E-05	2.69E-03
15	5.23E-04	9.70E-03	33.05	3.78E-05	3.65E-03
20	8.80E-05	2.72E-03	48.79	5.69E-05	4.03E-03
25	1.09E-05	5.86E-04	44.77	8.96E-05	5.25E-03
30	1.11E-06	1.69E-04	59.87	1.10E-04	6.15E-03

Table IX. Example 3.2: Reduced basis error bound and effectivity, and online time to evaluate $\mathcal{G}_N$, $\Delta_N^{\mathcal{G}}$, as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $\mathcal{G}_{h_t}$.

N	$\mathtt{E}_N$	$\mathcal{E}_N$	$\overline{\eta}_N^{\mathcal{G}}$	Online Time	
				$\mathcal{G}_N$	$\Delta_N^{\mathcal{G}}$
5	1.04E-01	1.66E+02	37.19	1.80E-05	2.12E-03
10	6.01E-02	8.72E+01	30.67	2.32E-05	2.37E-03
20	9.08E-03	4.39E-01	41.86	5.95E-05	3.96E-03
30	2.36E-03	1.31E-01	53.17	1.92E-04	5.03E-03
35	9.42E-05	3.17E-02	51.40	2.12E-04	7.10E-03
40	4.58E-05	8.86E-03	42.42	2.78E-04	7.69E-03

Table X. Example 3.3: Reduced basis error bound and effectivity, and online time to evaluate $\mathcal{G}_N$, $\Delta_N^{\mathcal{G}}$, as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $\mathcal{G}_{h_t}$.

however, for larger N, $\Delta_N^{\mathcal{G}}$ is more representative of the true error. The computational savings are very high: one online evaluation for an output and the associated error bound for $N = 25$ is about 35–40ms compared to approximately 10–12s for a "truth" solution; our online evaluation (even with error estimation) is about 250–300 times faster than the classical approach.

Example 3.3 We present in Table X the error bounds and effectivities for Example 3.3 as a function of N. The error bound reported, $\mathcal{E}_N$, is the maximum of the relative error bound, $\Delta_N^{\mathcal{G}}(\mu)/|\mathcal{G}_{h_t}(\mu)|$, over the same test sample Ξ_{test} of Section 5.5. We denote by $\overline{\eta}_N^{\mathcal{G}}$ the average of the effectivity, $\eta_N^{\mathcal{G}}(\mu)$, over Ξ_{test}. (As before, we reject from our sample Ξ_{test} those parameter points for which the error bound $\Delta_N^{\mathcal{G}}(\mu)$ is smaller than machine precision, as for these parameter points the calculation is contaminated by round-off.) We observe relatively good effectivity: our $\overline{\eta}_N^{\mathcal{G}}$ is usually of order $O(10\text{–}100)$.

It is seen that the computational savings are still very high: one online evaluation for an output and the associated error bound for $N = 30$ is about 80–100ms compared to approximately 9–12s for a "truth" solution; our online evaluation (even with error estimation) is still about 90–150 times faster than the classical approach. We also note from the timings

that the cost of $\Delta_N^{\mathcal{G}}$ is significantly greater than the cost of $\mathcal{G}_N$: the reason is that $Q^{\mathcal{A}} = 40$ is relatively large for this problem due to the more complicated (affine) geometric mappings near the hole. As a result, the computational time for $\Delta_N^{\mathcal{G}}$ (which has dominant complexity order $O(4(Q^{\mathcal{A}})^2 N^2)$) is much greater than the computational time for $\mathcal{G}_N$ (which is $O(8N^3)$) for small N.

7. CONCLUSIONS AND FUTURE WORKS

We conclude that, for all our Example model problems, our approach provides stress intensity factors at certified accuracy (relative to the "truth" finite element approximation) of 0.1% or 0.01% at less than $(1/100)^{\text{th}}$ the online cost of conventional numerical techniques. The savings would be even larger for problems with more complex geometry and solution structures, and in particular in three space dimensions with correspondingly larger $\mathcal{N}_t$. We emphasize that the stress intensity factor/energy release rate is obtained very inexpensively but also reliably — thanks to the rigorous and relatively sharp *a posteriori* error bounds.

In some sense, our method is very similar to a numerically based handbook approach. The important difference is that our "interpolation formula" — the reduced basis prediction — is extremely accurate for all $\mu \in \mathcal{D}$, and furthermore is endowed with a rigorous certificate of fidelity — our *a posteriori* error bound.

Many applications in fracture mechanics which require either real-time computation (for example, Non–Destructive Evaluation or failure prediction) or many-query computation (fatigue crack growth prediction, say) of the stress intensity factor/energy release rate can benefit from our approach [27]. We emphasize that the method is *not* useful if there is no real–time imperative and if only one (or a few) SIF evaluations are needed, due to the high computational cost of the offline stage. Typically, the break–even point (see Section 6.4) is $O(100 - 1000)$ input-output evaluations.

There are still many aspects of this work which require further investigation. First, the method is not very efficient for the treatment of many parameters. As P (the number of independent parameters) increases we anticipate that the reduced basis approximation will still converge quickly, however the offline computational cost for the inf-sup lower bound may grow exponentially (with P). The Successive Constraint Linear Optimization Method [32] can reduce the offline (and online) effort, however for large P the problem remains very challenging.

Second, problem set–up remains cumbersome. In this work, we exploit piecewise affine transformations; this is a good but very complicated tool that is not always practicable with very complex geometry or non–homogeneous material distribution. These mapping issues perhaps can be partially overcome by the "empirical interpolation" approach [21], or by introduction of domain decomposition ideas such as those proposed in [33].

The third and final aspect is the possible application of the method to more general crack problems, in particular "mixed-mode" situations. It should not be difficult to extend our methodology for "pure" Mode I cracks to the case of "pure" Mode II cracks. However, the general mixed–mode problem is much more difficult. It may be possible to first apply either the displacement decomposition approach or the J–integral approach [34] to reduce the requisite Mode I and Mode II Stress Intensity Factors to corresponding (continuous) quadratic outputs; the latter could then be treated by the expanded formulation — and associated reduced basis approximations and *a posteriori* error estimators — introduced in our current

paper. Nevertheless, in the mixed–mode case, calculation of both the SIFs and (ultimately) the subsequent crack/geometry evolution pose significant mathematical and computational challenges.

ACKNOWLEDGEMENTS

We would like to thank Professor David M. Parks and Professor J. Peraire of MIT, Professor G. R. Liu of NUS, and Professor David Clarke and Professor Anthony Evans of USCB for many helpful recommendations. This work was supported by the Singapore-MIT Alliance, and by DARPA and AFOSR under Grant FA9550-05-1-0114.

REFERENCES

1. A. A. Griffith. The phenomena of rupture and flaw in solids. *Transactions, Royal Society of London* 1920; Vol. A-221.
2. G. Strang, G. J. Fix. An Analysis of the Finite Element Method. Prentice-Hall: 1973.
3. D. M. Parks. Stiffness derivative finite-element technique for determination of crack tip stress intensity factors. *International Journal of Fracture* 1974; **10**(4):487–502.
4. Roshdy S. Barsoum. Use of isoparametric finite-elements in linear-fracture mechanics. *International Journal for Numerical Methods in Engineering* 1975; **10**(1):25–37.
5. B. O. Almroth, P. Stern and F. A. Brogan. Automatic choice of global shape functions in structural analysis. *AIAA Journal* 1978; **16**:525–528.
6. A. K. Noor and J. M. Peters. Reduced basis technique for nonlinear analysis of structures. *AIAA Journal* 1980; **18**(4):455–462.
7. J. P. Fink and W. C. Rheinboldt. On the error behavior of the reduced basis technique for non–linear finite-element approximations. *Zeitschrift fur Angewandte Mathematik und Mechanik* 1983; **63**(1):21–28.
8. T. A. Porsching. Estimation of the error in the reduced basis method solution of nonlinear equations. *Mathematics of Computation*, 1985; **45**(172):487–496.
9. M. D Gunzburger. *Finite Element Methods for Viscous Incompressible Flows: A Guide to Theory, Practice, and Algorithms*, Academic Press (Boston), 1989.
10. J. S. Peterson. The reduced basis method for incompressible viscous-flow calculations. *SIAM Journal of Scientific and Statistical Computing*, 1989; **10**(4):777–786.
11. I. Babuska, J. M. Melenk. The partition of unity method. *International Journal for Numerical Methods in Engineering* 1997; **40**(4):727–758.
12. N. Moës, J. Dolbow, T. Belytschko. A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering* 1999; **46**(1):131–150.
13. N. A. Pierce, M. B. Giles. Adjoint recovery of superconvergent functionals from PDE approximations. *SIAM Review* 2000; **42**(2):247–264.
14. Y. Murakami, ed. in chief. *Stress Intensity Factors Handbook*. Elsevier: 2001.
15. Heintz, F. Larsson, K. Hansbo and K. Runesson. Adaptive strategies and error control for computing material forces in fracture mechanics. *International Journal for Numerical Methods in Engineering* 2004; **60**(7):1287–1299.
16. C. Prud'homme, D.V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, G. Turinici. Reliable real-time solution of parametrized partial diffirential equations: Reduced-basis output bound methods. *Journal of Fluids Engineering* 2002; **124**(1):70–80.
17. Y. Maday, A. T. Patera, and D. V. Rovas. A blackbox reduced-basis output bound method for noncoercive linear problems. in *Nonlinear Partial Differential Equations* College de France Seminar XIV, D. Cioranescu and J.-L. Lions (editors), *Studies in Mathematics and Its Applications*, Vol. 31, Elsevier Science B.V, 2002: 533–569.
18. S. S. Ravindran. Adaptive reduced-order controllers for a thermal flow system using proper orthogonal decomposition. *SIAM Journal of Scientific Computing*, 2002; **23**(6):1924–1942.
19. E. Béchet, H. Minnebo, N. Moës, B. Burgardt. Improved implementation and robustness study of the X-FEM for stress analysis around cracks. *International Journal for Numerical Methods in Engineering* 2005; **64**(8):1033–1056.
20. P. Laborde, J. Pommier, Y. Renard, M. Salaün. High-order extended finite element method for cracked domains. *International Journal for Numerical Methods in Engineering* 2005; **64**:354–381.

21. M. Barrault, N. C. Nguyen, Y. Maday, A. T. Patera. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique* 2004; **339**(9):667–672.
22. T. L. Anderson. *Fracture Mechanics: Fundamentals and Applications* (3rd edn), CRC, 2004.
23. N. C. Nguyen, K. Veroy, A. T. Patera. Certified Real-Time Solution of Parametrized Partial Differential Equations. *Handbook of Materials Modeling*, pp. 1523–1558, Springer: 2005.
24. M. A. Grepl, A. T. Patera. *A posteriori* error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *M2AN. Mathematical Modelling and Numerical Analysis*, 2005; **39**(1):157–181.
25. S. Sen, K. Veroy, D. B. P. Huynh, S. Deparis, N. C. Nguyen, A. T. Patera. Natural Norm *A Posteriori* Error Estimators for Reduced Basis Approximations. *Journal of Computational Physics* 2006; **217**(1):37–62.
26. D. B. P. Huynh, J. Peraire, A. T. Patera and G. R. Liu. Real-Time Reliable Prediction of Linear-Elastic Mode-I Stress Intensity Factors for Failure Analysis. SMA Symposium, 2006.
27. D. B. P. Huynh. Reduced-Basis Approximation and Application to Fracture and Inverse Problems. Ph.D Thesis, National University of Singapore, 2007.
28. D. Clarke and A. Evans. Private communication, 2005.
29. L. Gallimard and J. Panetier. Error estimation of stress intensity factors for mixed-mode cracks. *International Journal for Numerical Methods in Engineering* 2006; **68**(3):299–316.
30. M. Ruter and E.Stein. Goal-oriented a posteriori error in linear elastic fracture mechanics. *Computer methods in applied mechanics and engineering* 2006; **195**:251–278.
31. Z. C. Xuan, N. Parés and J. Peraire. Computing upper and lower bounds for the J-integral in two-dimensional linear elasticity. *Computer methods in applied mechanics and engineering* 2006; **195**:430–443.
32. D. B. P. Huynh, G. Rozza, S. Sen, A. T. Patera. A Successive Constraint Linear Optimization Method for Lower Bounds of Parametric Coercivity and Inf-Sup Stability Constraints. *Comptes Rendus Mathematique* 2006, submitted.
33. Y. Maday and E. Rønquist. The Reduced Basis Element Method: Application to a Thermal Fin Problem. *SIAM Journal of Scientific Computing* 2004; **26**(1):240–258.
34. A. R. Ingraffea and P. A. Wawrzynek. Finite Element Methods for Linear Elastic Fracture Mechanics. *Comprehensive Structural Integrity.* Elsevier Science Ltd., Oxford, England: 2003

APPENDIX

The mapping of Example 3.3

Here we describe our affine transformations for the Example 3.3. Here $\Omega^0(\mu) \rightarrow \Omega \equiv \Omega^0(\mu = \mu_{\text{ref}} = (d_{\text{ref}} = 0.475, R_{\text{ref}} = 0.175, L_{\text{ref}} = 1.75))$. The transformation is piecewise affine: the identity for (the squared invariant domain of size $2\Delta d$) Ω_2; dilation in the x_1^0-direction for Ω_1 and Ω_3; dilation in the x_2^0-direction for Ω_6, Ω_{11}; dilations in both x_1^0-direction and x_2^0-direction for Ω_5, Ω_7, Ω_8, Ω_9, Ω_{10}, Ω_{12}; and dilations in both the x_1^0-direction and x_2^0-direction and shear in the x_2^0-direction for Ω_4. The domain Ω_8 is subjected to dilations in both the x_1^0-direction and the x_2^0-direction, but with the same scaling factor; the isoparametric transformations of elements around the hole thus still remain affine. The abstract statement for our classical formulation is then recovered for $Q^a = 24$, $Q^f = 4$, as listed in Table XI and Table XII. As a result, we also recover $Q^{\mathcal{A}} = 40$ and $Q^{\mathcal{F}} = 6$ for our expanded formulation of Section 2.2.

q	$\Theta_q^a(\mu)$	$a_q^e(w,v)$
1	1	$\overline{c}_1 \int_{\Omega\setminus\{\Omega_2 \bigcap \Omega_8\}} \left(\frac{\partial v_1}{\partial x_1}\frac{\partial w_2}{\partial x_2} + \frac{\partial v_2}{\partial x_2}\frac{\partial w_1}{\partial x_1}\right) d\Omega + \int_{\Omega_2} \frac{\partial v_i}{\partial x_j} C_{ijkl} \frac{\partial w_k}{\partial x_l} d\Omega$ $+ \overline{c}_2 \int_{\Omega\setminus\{\Omega_2 \bigcap \Omega_8\}} \left(\frac{\partial v_1}{\partial x_2}\frac{\partial w_2}{\partial x_1} + \frac{\partial v_2}{\partial x_1}\frac{\partial w_1}{\partial x_2}\right) d\Omega + \int_{\Omega_8} \frac{\partial v_i}{\partial x_j} C_{ijkl} \frac{\partial w_k}{\partial x_l} d\Omega$
2	$\frac{1}{t_y^1}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_6} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_6} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
3	t_y^1	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_6} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_6} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
4	$\frac{1}{t_y^2}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_{11}} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_{11}} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
5	t_y^2	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_{11}} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_{11}} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
6	$\frac{t_{xy}^0}{t_y^2}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_9} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_9} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
7	$\frac{t_y^2}{t_{xy}^0}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_9} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_9} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
8	t_x^1	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_1} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_1} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
9	$\frac{1}{t_x^1}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_1} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_1} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
10	$\frac{t_x^1}{t_y^1}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_5} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_5} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
11	$\frac{t_y^1}{t_1^1}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_5} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_5} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
12	$\frac{t_x^1}{t_y^2}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_{10}} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_{10}} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$
13	$\frac{t_y^2}{t_x^1}$	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_{10}} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1} d\Omega + \overline{c}_2 \int_{\Omega_{10}} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1} d\Omega$
14	t_x^2	$(\overline{c}_1 + 2\overline{c}_2) \int_{\Omega_3} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2} d\Omega + \overline{c}_2 \int_{\Omega_3} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2} d\Omega$

15	$\frac{1}{t_x^2}$	$(\bar{c}_1 + 2\bar{c}_2)\int_{\Omega_3} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1}d\Omega + \bar{c}_2 \int_{\Omega_3} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1}d\Omega$
16	$\frac{t_x^2}{t_y^1}$	$(\bar{c}_1 + 2\bar{c}_2)\int_{\Omega_7} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2}d\Omega + \bar{c}_2 \int_{\Omega_7} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2}d\Omega$
17	$\frac{t_y^1}{t_x^2}$	$(\bar{c}_1 + 2\bar{c}_2)\int_{\Omega_7} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1}d\Omega + \bar{c}_2 \int_{\Omega_7} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1}d\Omega$
18	$\frac{t_x^2}{t_y^2}$	$(\bar{c}_1 + 2\bar{c}_2)\int_{\Omega_{12}} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2}d\Omega + \bar{c}_2 \int_{\Omega_{12}} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2}d\Omega$
19	$\frac{t_y^2}{t_x^2}$	$(\bar{c}_1 + 2\bar{c}_2)\int_{\Omega_{12}} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1}d\Omega + \bar{c}_2 \int_{\Omega_{12}} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1}d\Omega$
20	$\frac{t_x^2}{t_{xy}^0}$	$(\bar{c}_1 + 2\bar{c}_2)\int_{\Omega_4} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2}d\Omega + \bar{c}_2 \int_{\Omega_4} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2}d\Omega$
21	$\frac{t_{xy}^0}{t_x^2}$	$(\bar{c}_1 + 2\bar{c}_2)\int_{\Omega_4} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1}d\Omega + \bar{c}_2 \int_{\Omega_4} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1}d\Omega$
22	$\frac{t_z}{t_{xy}^0}$	$-(\bar{c}_1 + 2\bar{c}_2)\Big(\int_{\Omega_4} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_2}d\Omega + \int_{\Omega_4} \frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_1}d\Omega\Big)$ $-\bar{c}_2\Big(\int_{\Omega_4} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_2}d\Omega + \int_{\Omega_4} \frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_1}d\Omega\Big)$
23	$\frac{t_z}{t_x^2}$	$-(\bar{c}_1 + 2\bar{c}_2)\Big(\int_{\Omega_4} \frac{\partial v_1}{\partial x_1}\frac{\partial w_2}{\partial x_1}d\Omega + \int_{\Omega_4} \frac{\partial v_2}{\partial x_1}\frac{\partial w_1}{\partial x_1}d\Omega\Big)$
24	$\frac{(t_z)^2}{t_x^2 t_{xy}^0}$	$(\bar{c}_1 + 2\bar{c}_2)\int_{\Omega_4} \frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1}d\Omega + \bar{c}_2 \int_{\Omega_4} \frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1}d\Omega$

$$t_{xy}^0 = \frac{R + \frac{R}{R_{\text{ref}}}\Delta R}{R_{\text{ref}} + \Delta R}, \qquad t_x^1 = \frac{1.0 - d - \Delta d}{1.0 - d_{\text{ref}} - \Delta d}, \qquad t_x^2 = \frac{d - R - \frac{R}{R_{\text{ref}}}\Delta R - \Delta d}{d_{\text{ref}} - R_{\text{ref}} - \Delta R - \Delta d}, \quad \Delta d = 0.03$$

$$t_y^1 = \frac{R + \frac{R}{R_{\text{ref}}}\Delta R - \Delta d}{R_{\text{ref}} + \Delta R - \Delta d}, \quad t_y^2 = \frac{L - R - \frac{R}{R_{\text{ref}}}\Delta R}{L_{\text{ref}} - R_{\text{ref}} - \Delta R}, \quad t_z = \frac{(1 - t_{xy}^0)\Delta d}{d_{\text{ref}} - R_{\text{ref}} - \Delta d - \Delta R}, \quad \Delta R = 0.04$$

$$d_{\text{ref}} = 0.475, \qquad R_{\text{ref}} = 0.175, \qquad L_{\text{ref}} = 1.75$$

Table XI. Example 3.3: Parametric functions $\Theta_q^a(\mu)$ and parameter-independent bilinear forms $a^q(w, v)$, $1 \le q \le Q^a = 24$.

q	$\Theta_q^f(\mu)$	$f_q(v)$
1	1	$\int_{\Gamma_{T_{11}}} v d\Gamma$
2	t^0_{xy}	$\int_{\Gamma^a_{T_9}} v d\Gamma$
3	t^1_x	$\int_{\Gamma^c_{T_{10}}} v d\Gamma$
4	t^2_x	$\int_{\Gamma^c_{T_{12}}} v d\Gamma$

Table XII. Example 3.3: Parametric functions $\Theta_q^f(\mu)$ and parameter-independent bilinear forms $f^q(w)$, $1 \le q \le Q^f = 4$.