Reduced-basis method for band structure calculations

George S. H. Pau*
Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 3-264, Cambridge, Massachusetts 02139, USA
(Received 18 June 2007; revised manuscript received 29 August 2007; published 17 October 2007)

We describe the application of the reduced-basis method in rapid and accurate determination of band energies in band structure calculations. The method is well suited for problems requiring repetitive evaluations of the band energies, especially in the many-query limit. We demonstrate the efficacy of the method in the determination of the spectral properties of crystalline silicon.

DOI: 10.1103/PhysRevE.76.046704 PACS number(s): 02.70.−c, 71.15.Dx, 71.15.Ap

I. INTRODUCTION

Studies of periodic structures frequently require evaluations of band energies $E_i(\mathbf{k})$, $1 \le i \le n_b$, at many different wave vectors $\mathbf{k}$, evaluations of which require solutions of an eigenvalue problem derived from a quantum model. Here, n_b is the number of lowest band energies we are interested in given a $\mathbf{k}$ point. In many cases, the number of evaluations required is large. For example, in [1], an accurate determination of the anomalous Hall conductivity requires solutions at millions of $\mathbf{k}$ points. In the determination of the dielectric function of nanostructures [2,3], the number of evaluations required is further augmented by the dependency of the dielectric function on space. Certainly, accurate yet rapid methods to evaluate band energies and associated eigenvectors are highly desirable.

In [4], the use of the maximally localized Wannier functions (MLWFs) [5] and the Slater-Koster interpolation scheme [6] was proposed for rapid evaluations of band energies. This approach is used in [1] to evaluate the anomalous Hall conductivity. However, the efficiency of the method is contingent on finding "good" MLWFs through a nonconvex optimization procedure. While difficulties related to nonconvex optimization are partially alleviated through the use of the simultaneous diagonalization procedure [7], the Slater-Koster interpolation scheme is a nonvariational approach—the interpolated solutions do not satisfy the model's governing equations. In addition, an *a posteriori* error estimation procedure is also absent.

In this paper, we introduce the reduced-basis method as a good alternative to the above approach. To motivate the application of the reduced-basis method, it is advantageous to first facilitate the band structure calculation with an input-output abstraction: the input parameter is the wave vector $\mathbf{k}$ and the outputs are the band energies $E_i(\mathbf{k})$ or functionals of the wave functions $u_i(\mathbf{x};\mathbf{k})$; to determine $E_i(\mathbf{k})$ and $u_i(\mathbf{x};\mathbf{k})$, we must solve a linear eigenvalue problem parametrized by $\mathbf{k}$. In this paper, we emphasize the rapid evaluations of functionals involving $u_i(\mathbf{x};\mathbf{k})$ and not on the evaluations of the wave functions $u_i(\mathbf{x};\mathbf{k})$, which are functions of the spatial variable $\mathbf{x}$, in addition to the parameter $\mathbf{k}$. To underscore this emphasis and for notational simplicity, we denote $u_i(\mathbf{x};\mathbf{k})$ by $u_i(\mathbf{k})$.

A. Background on reduced-basis method

The reduced-basis method exploits dimension reduction afforded by the *low-dimensional* and *smooth* parametrically induced solution manifold. More precisely, to approximate solutions of an underlying parametrized partial differential equation, we use a basis set consisting of solutions at a number of judiciously selected parameter points instead of using general basis sets consisting of, say, Fourier basis functions. An approximation is then obtained by a projection onto a finite and low-dimensional vector space spanned by the solutions at these selected points.

The reduced-basis method was first introduced in the late 1970s in the context of nonlinear structural analysis [8,9] and subsequently abstracted, analyzed, and extended to a much larger class of parametrized partial differential equations [10–14]. In the more recent past the reduced-basis approach, and in particular associated *a posteriori* error estimation procedures have been successfully developed for (i) linear elliptic and parabolic partial differential equations (PDEs) that are affine in the parameter [15–18]; (ii) PDEs that are at most quadratically nonlinear in the first argument [19–21]; and (iii) general nonaffine PDEs [22,23]. Application of the reduced-basis method and associated rigorous *a posteriori* error bounds to the determination of the first eigenvalue of a linear eigenvalue problem has also been examined in [16]. In these cases a very efficient offline-online computational strategy can be developed. The operation count for the online stage—in which, given a new parameter value, we calculate the reduced-basis output and associated error bound—is *independent* of $\mathcal{N}$, the dimension of the underlying "truth" approximation. This will be further elaborated in Sec. III.

B. Examples

To illustrate the utility of the reduced-basis method, we shall examine the determination of the integrated density of states,

$$I(E) = \int_{-\infty}^{E} \left(\frac{1}{|\mathcal{D}|} \sum_{i=1}^{n_b} \int_{\mathcal{D}} \delta(\mathcal{E} - E_i(\mathbf{k})) d\mathbf{k} \right) d\mathcal{E}; \qquad (1)$$

the joint density of states,

*Present address: Lawrence Berkeley National Laboratory, MS 50A-1488, One Cyclotron Road, Berkeley, CA 94720. gpau@lbl.gov

$$J(E) = \sum_{i=1}^{n_v} \sum_{j=n_v+1}^{n_b} \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} \delta(E_j(\mathbf{k}) - E_i(\mathbf{k}) - E) d\mathbf{k}; \quad (2)$$

and the complex component of the dielectric function,

$$\varepsilon_2(E) = \left(\frac{2\pi}{E}\right)^2 \sum_{i=1}^{n_v} \sum_{j=n_v+1}^{n_b} \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} \left(\sum_{\ell=1}^{3} \left| \int_{\Omega} \frac{\partial u_j(\mathbf{k})}{\partial x_\ell} u_i^*(\mathbf{k}) \right|^2 \right) \times \delta(E_j(\mathbf{k}) - E_i(\mathbf{k}) - E) d\mathbf{k}. \quad (3)$$

Here n_v is the number of valence bands, n_b is the total number of bands required to compute the quantities accurately, and $\mathcal{D}$ is the parameter domain in which $\mathbf{k}$ varies—in the above examples, $\mathcal{D}$ is given by the first irreducible Brillouin zone since, through symmetry arguments, any $\mathbf{k} \notin \mathcal{D}$ can be mapped to a point in $\mathcal{D}$. To evaluate (2) and (3), we employ the tetrahedron method [24,25]—we discretize $\mathcal{D}$ by constructing a tetrahedral mesh $\mathcal{T}$ consisting of n_k mesh points, evaluate $E_i(\mathbf{k})$ at these n_k $\mathbf{k}$ points, and assume a linear interpolation of $E_i(\mathbf{k})$ within each of the tetrahedra. Depending on the desired accuracy, n_k can be large—[26] found that 4000 $\mathbf{k}$ points are needed to sufficiently resolve the van Hove singularities in the density of states.

There are other applications in which the reduced-basis method for linear eigenvalue problems can be useful within the computational chemistry context. For example, in *ab initio* calculations based on density functional theory models, each fixed point iteration in the self-consistent field (SCF) scheme may require solutions to a linear eigenvalue problem at n_k $\mathbf{k}$ points—these solutions are then used to accurately determine the electron density and related functionals [27]. If n_k required is large, a reduced-basis approximation within each iteration can significantly speed up evaluations of the n_k eigensolutions, thus improving the overall efficiency of the SCF algorithm.

For simplicity, we shall perform band structure calculations based on the empirical pseudopotential model where the effective background potential is defined in [28]. We note that the method is not limited to this particular model. In Sec. V we examine how the current approach can be extended to more realistic density functional theory models, especially as a postprocessing tool. Note that we will work in atomic units but, in Sec. IV B, the atomic unit for the energy has been converted to eV to facilitate comparison with results in the existing literature.

II. PROBLEM FORMULATION

Consider a crystal structure defined by the Bravais lattice vectors $\{\mathbf{a}_i \in \mathbb{R}^3,\ 1 \le i \le 3\}$ and the basis vectors $\boldsymbol{\tau} \equiv (\tau_1, \ldots, \tau_{n_\tau})$. For any given $\mathbf{k} \equiv (k_1, k_2, k_3) \in \mathcal{D}$, we would like to find the band energies $E_i(\mathbf{k})$, $1 \le i \le n_b$, given by

$$E_i(\mathbf{k}) = \lambda_i(\mathbf{k}) + \frac{1}{2}|\mathbf{k}|^2, \quad (4)$$

where $\mathcal{D} \subset \mathbb{R}^3$ is a bounded domain given by the irreducible Brillouin zone of the Bravais lattice; and $[\hat{\mathbf{u}}(\mathbf{k}) \equiv (u_1(\mathbf{k}), \ldots, u_{n_b}(\mathbf{k})),\ \hat{\boldsymbol{\lambda}}(\mathbf{k}) \equiv (\lambda_1(\mathbf{k}), \ldots, \lambda_{n_b}(\mathbf{k}))] \in (Y^{n_b} \times \mathbb{R}^{n_b})$ satisfies

$$\left[-\frac{1}{2}\Delta - i\mathbf{k} \cdot \nabla + V_{\text{eff}}(\mathbf{x}; \boldsymbol{\tau}) \right] u_i(\mathbf{k}) = \lambda_i(\mathbf{k}) u_i(\mathbf{k}), \quad 1 \le i \le n_b,$$

$$\int_{\Omega} u_i^*(\mathbf{k}) u_j(\mathbf{k}) = \delta_{ij}, \quad 1 \le i \le j \le n_b. \quad (5)$$

Here, $Y \equiv H^1_{\text{per}}(\Omega)$ is the space of $\{\mathbf{a}_i\}_{i=1}^3$-periodic complex functions in $H^1(\mathbb{R}^3)$; Ω is the primitive unit cell; $\mathbf{x}$ is a point in Ω; $V_{\text{eff}}(\cdot; \boldsymbol{\tau}) \in C^m$ is a real periodic function dependent on $\boldsymbol{\tau}$; and $*$ denotes complex conjugation. Components in $\hat{\boldsymbol{\lambda}}(\mathbf{k})$ are real and arranged such that $\lambda_1(\mathbf{k}) \le \lambda_2(\mathbf{k}) \le \cdots \le \lambda_{n_b}(\mathbf{k})$. We note that (5) is a linear eigenvalue problem.

A. Parametrized weak form

The parametrized weak form of (5) is obtained as follows. For a given $\mathbf{k} \in \mathcal{D}$, find $(\hat{\mathbf{u}}(\mathbf{k}), \hat{\boldsymbol{\lambda}}(\mathbf{k})) \in (Y^{n_b} \times \mathbb{R}^{n_b})$ that satisfies

$$\langle v | \mathcal{A}(\mathbf{k}) | u_i(\mathbf{k}) \rangle = \lambda_i(\mathbf{k}) \langle v | u_i(\mathbf{k}) \rangle, \quad \forall\, v \in Y, \quad (6)$$

$$\langle u_j(\mathbf{k}) | u_i(\mathbf{k}) \rangle = \delta_{ij}, \quad i \le j \le n_b, \quad (7)$$

for $1 \le i \le n_b$ where

$$\langle v | \mathcal{A}(\mathbf{k}) | w \rangle \equiv \frac{1}{2} \int_{\Omega} \nabla w \nabla v^* + \int_{\Omega} V_{\text{eff}} w v^* - i \sum_{j=1}^{3} k_j \int_{\Omega} \frac{\partial w}{\partial x_j} v^*, \quad (8)$$

$$\langle v | w \rangle \equiv \int_{\Omega} w v^*, \quad (9)$$

for any $w \in Y$ and $v \in Y$.

B. Affine parameter dependence

We note that the functional form of $\mathcal{A}(\mathbf{k})$ is affine with respect to the parameter $\mathbf{k}$—we can express $\langle \cdot | \mathcal{A}(\mathbf{k}) | \cdot \rangle$ as

$$\langle v | \mathcal{A}(\mathbf{k}) | w \rangle = \langle v | \mathcal{A}_1 | w \rangle + \sum_{j=1}^{3} k_j \langle v | \mathcal{A}_2^j | w \rangle, \quad (10)$$

where the $\mathbf{k}$-independent forms $\langle v | \mathcal{A}_1 | w \rangle$, and $\langle v | \mathcal{A}_2^j | w \rangle$, $1 \le j \le 3$, are given by

$$\langle v | \mathcal{A}_1 | w \rangle \equiv \frac{1}{2} \int_{\Omega} \nabla w \nabla v^* + \int_{\Omega} t w v^*, \quad (11)$$

$$\langle v | \mathcal{A}_2^j | w \rangle \equiv -i \int_{\Omega} \frac{\partial w}{\partial x_j} v^*. \quad (12)$$

This affine parameter dependence property allows $\langle \cdot | \mathcal{A}(\mathbf{k}) | \cdot \rangle$ to be expressed as $\Sigma_{q=1}^{Q} \Theta_q(\mathbf{k}) \langle \cdot | \mathcal{A}_q | \cdot \rangle$ for some finite Q, where $\Theta_q : \mathcal{D} \to \mathbb{R}$, $1 \le q \le Q$, are smooth parameter-dependent functions, and $\langle \cdot | \mathcal{A}_q | \cdot \rangle : Y \times Y \to \mathbb{R}$, $1 \le q \le Q$, are parameter-independent continuous bilinear forms. Here $Q=4$, $\Theta_1 = 1$, $\Theta_2 = k_1$, $\Theta_3 = k_2$, and $\Theta_4 = k_3$. We note that

$\Theta_q(\mathbf{k})$, $1 \le q \le Q$, are usually simple algebraic expressions that can be readily evaluated in $O(1)$ operations. We will exploit this property in formulating an efficient computational strategy in Sec. III C.

C. Numerical example

We consider the band structure calculation for the diamond structure of silicon based on the empirical pseudopotential model in [28]. The Bravais lattice vectors are defined by

$$\mathbf{a}_1 = \frac{a}{2}(0,1,1), \quad \mathbf{a}_2 = \frac{a}{2}(1,0,1), \quad \mathbf{a}_3 = \frac{a}{2}(1,1,0), \quad (13)$$

where a is the lattice length, of magnitude 10.32. In addition, there are two atoms per unit cell and the set of basis vectors $\boldsymbol{\tau}$ is given by $(-\tau_0, \tau_0)$, where $\tau_0 = \Sigma_{i=1}^3 \mathbf{a}_i/8$. As such, $V_{\text{eff}}(\cdot;\boldsymbol{\tau})$ can be written as $V_{\text{eff}}(\cdot;\tau_0)$ for this particular example.

The parameter $\mathbf{k}$ lies in the domain $\mathcal{D}$ given by the irreducible Brillouin zone of the fcc structure defined by the polyhedron with vertices given by the high-symmetry points: $L \equiv (2\pi/a)(1/2,1/2,1/2)$, $\Gamma \equiv (2\pi/a)(0,0,0)$, $X \equiv (2\pi/a) \times (1,0,0)$, $W \equiv (2\pi/a)(3/4,3/4,0)$, $K \equiv (2\pi/a)(1,0,1/2)$, and $U \equiv (2\pi/a)(1,1/4,1/4)$. As described in Sec. I B, we will subsequently discretize $\mathcal{D}$ into a set of tetrahedra so that we can evaluate (2) and (3) by the tetrahedron method.

The effective potential V_{eff} (as defined in [28]) is given by

$$V_{\text{eff}}(\mathbf{x};\tau_0) = \sum_{\mathbf{G}} S(\mathbf{G};\tau_0) V(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{x}}. \quad (14)$$

Here $\mathbf{G} = \Sigma_{i=1}^3 m_i \mathbf{b}_i$ where $\{\mathbf{b}_i\}_{i=1}^3$ are the reciprocal lattice vectors satisfying the relations $\mathbf{a}_i \mathbf{b}_j = 2\pi\delta_{ij}$, $1 \le i,j \le 3$; $S(\mathbf{G};\tau_0) = \cos \mathbf{G}\tau_0$; and $V(\mathbf{G})$ is given by

$$V(\mathbf{G}) = \begin{cases} -0.21, & |\mathbf{G}|^2 = 3\left(\frac{2\pi}{a}\right)^2, \\ 0.04, & |\mathbf{G}|^2 = 8\left(\frac{2\pi}{a}\right)^2, \\ 0.08, & |\mathbf{G}|^2 = 11\left(\frac{2\pi}{a}\right)^2, \\ 0, & \text{otherwise}. \end{cases} \quad (15)$$

We note that $V_{\text{eff}}(\cdot;\tau_0)$ is smooth and represented by just 44 Fourier modes.

D. "Truth" approximation

We now consider the approximation of (7) by the plane-wave method. We define our Fourier approximation space $Y^{\mathcal{N}} \subset Y$ of dimension $\mathcal{N}$ as

$$Y^{\mathcal{N}} \equiv \text{span}\left\{ \varphi_{\mathbf{G}} \equiv e^{i\mathbf{G}\cdot\mathbf{x}} | \mathbf{x} \in \Omega, \frac{1}{2}|\mathbf{G}|^2 \le E_{\text{cut}} \right\} \quad (16)$$

where E_{cut} is a user-defined cutoff kinetic energy of the plane waves—$\mathcal{N}$ is then the number of $\mathbf{G}$'s that satisfy the inequality $\frac{1}{2}|\mathbf{G}|^2 \le E_{\text{cut}}$. Our plane-wave approximation to (7) is then given by: for a given $\mathbf{k} \in \mathcal{D}$, find $(\hat{\mathbf{u}}^{\mathcal{N}}(\mathbf{k}), \hat{\boldsymbol{\lambda}}^{\mathcal{N}}(\mathbf{k})) \in ((Y^{\mathcal{N}})^{n_b} \times \mathbb{R}^{n_b})$ that satisfies

$$\langle v | \mathcal{A}(\mathbf{k}) | u_i^{\mathcal{N}}(\mathbf{k}) \rangle = \lambda_i^{\mathcal{N}}(\mathbf{k}) \langle v | u_i^{\mathcal{N}}(\mathbf{k}) \rangle, \quad \forall \ v \in Y^{\mathcal{N}},$$

$$\langle u_j^{\mathcal{N}}(\mathbf{k}) | u_i^{\mathcal{N}}(\mathbf{k}) \rangle = \delta_{ij}, \quad i \le j \le n_b, \quad (17)$$

for $1 \le i \le n_b$. The above then gives an $\mathcal{N} \times \mathcal{N}$ algebraic system which can then be diagonalized to obtain the desired eigensolutions.

We now determine convergence of the solutions with respect to $\mathcal{N}$. In Fig. 1, we show the convergence of the absolute error

$$\max_{1 \le i \le n_b} |\lambda_i^{\mathcal{N}}(\mathbf{k}_0) - \lambda_i^{\text{box}}(\mathbf{k}_0)| \quad (18)$$

where $\mathbf{k}_0$ is the Baldereschi mean value point [29] given by $(2\pi/a)(0.6223, 0.2953, 0)$, and $\lambda_i^{\text{box}}(\cdot)$, $1 \le i \le n_b$, is a plane-wave approximation based on

$$Y^{\text{box}} \equiv \text{span}\{\varphi_{\mathbf{G}} \equiv e^{i\mathbf{G}\cdot\mathbf{x}} | \mathbf{x} \in \Omega, \ \mathbf{G} \equiv \sum_{i=1}^3 m_i \mathbf{b}_i, \ m_i \in \mathbb{Z}, \ -9 \le m_i \le 10, \ 1 \le i \le 3\}. \quad (19)$$

We see that the error (18) for $n_b = 20$ is of $O(10^{-12})$ at $\mathcal{N} = 1807$.

We now denote an approximation based on $Y^{\mathcal{N}_t}$ where $\mathcal{N}_t = 1807$ as the "truth" approximation; the subscript t denotes truth. The point of departure for the reduced-basis method is this truth approximation. We build our reduced-basis approximation on, and measure the error in the reduced-basis approximation relative to, this truth approximation. Note that, since the reduced-basis approximation is built upon this truth approximation, it cannot perform better than this truth approximation. Thus, $\mathcal{N}_t$ must usually be large in order to obtain an accurate reduced-basis approximation. Thankfully, however, we shall see that, once the reduced-basis approximation has been built, the computational costs will be independent of $\mathcal{N}_t$. To simplify the notation, we drop the superscript $\mathcal{N}_t$ from all subsequent formulations, with the understanding that the truth approximation in fact refers to

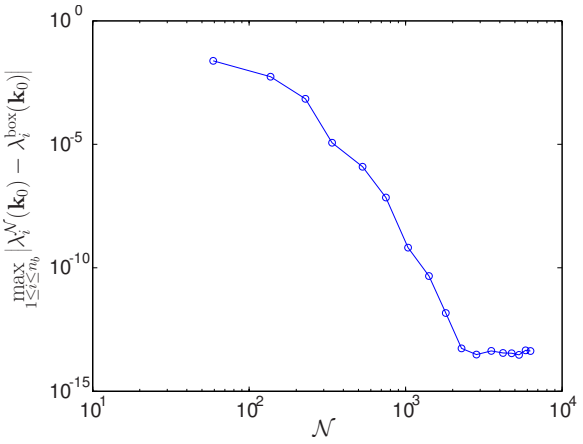

FIG. 1. (Color online) Convergence of the approximation error (18) with $\mathcal{N}$ for $n_b = 20$ and $\mathbf{k}_0 = (2\pi/a)(0.6223, 0.2953, 0)$.

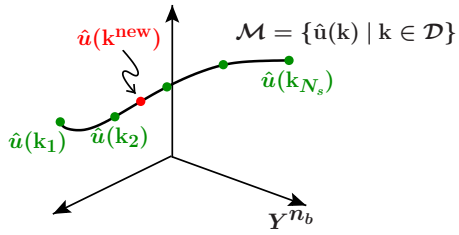

FIG. 2. (Color online) Conceptual drawing of the solution manifold $\mathcal{M}$. The solution $\hat{\mathbf{u}}(\mathbf{k}^{\mathrm{new}})$ is approximated by a reduced-basis space constructed from $\hat{\mathbf{u}}(\mathbf{k}_i)$, $1 \le i \le N_s$.

the plane-wave approximation with $\mathcal{N}=\mathcal{N}_t$. Thus, Y, $\hat{\mathbf{u}}$, and $\hat{\boldsymbol{\lambda}}$ will now be understood as $Y^{\mathcal{N}_t}$, $\hat{\mathbf{u}}^{\mathcal{N}_t}$, and $\hat{\boldsymbol{\lambda}}^{\mathcal{N}_t}$.

III. REDUCED-BASIS METHOD

In the plane-wave method described in Sec. II D, we have represented $u_i(\mathbf{k})$, $1 \le i \le n_b$, by a linear combination of $\varphi_{\mathbf{G}} \in Y$—$u_i(\mathbf{k})$ is an arbitrary member of Y. However, the solution $\hat{\mathbf{u}}(\mathbf{k})$ can in fact be localized to a much lower-dimensional manifold $\mathcal{M} \equiv \{\hat{\mathbf{u}}(\mathbf{k}),\ \mathbf{k} \in \mathcal{D}\}$ residing in Y^{n_b}. In the case of a single parameter, $\mathcal{M}$ can be visualized as a one-dimensional filament that winds through Y^{n_b} as sketched in Fig. 2. Presuming that $\mathcal{M}$ is sufficiently smooth, we can then look for an approximation of $\hat{\mathbf{u}}(\mathbf{k})$ in a finite-dimensional space spanned by elements in $\mathcal{M}$. The reduced-basis approach explicitly recognizes this computational opportunity.

To consolidate the above argument, we introduce the notion of the Kolmogorov N width d_N [30–32]:

$$d_N(A,Y) \equiv \inf_{Y_N \subset Y} \sup_{x \in A} \inf_{y \in Y_N} \|x - y\|_Y, \tag{20}$$

where A is a subset of Y and Y_N is an arbitrary N-dimensional subspace of Y. The Kolmogorov N width d_N measures the extent to which A may be approximated by a finite-dimensional space of dimension N in Y. We will have a rapidly convergent approximation if d_N approaches zero rapidly as N increases. For our case where $A \equiv \mathcal{M}$, we can attribute this to the smoothness of the solutions with respect to $\mathbf{k}$, as demonstrated for a single-parameter elliptic problem in [33]. In [34], it is further shown that d_N is almost realized if Y_N is spanned by elements in $\mathcal{M}$. The construction of $Y_N \subset \mathcal{M}$ that minimizes d_N is, however, combinatorially difficult. The reduced-basis method then provides an efficient procedure by which we can construct a good surrogate to Y_N.

A. Approximation

We first introduce nested sample sets

$$S_N = (\mathbf{k}_1, \ldots, \mathbf{k}_{N_s}), \quad 1 \le N_s \le N_{s,\max}, \tag{21}$$

and define the associated nested reduced-basis spaces as

$$\begin{aligned} W_N &= \mathrm{span}\{u_i(\mathbf{k}_j),\ 1 \le i \le n_b,\ 1 \le j \le N_s\} \\ &= \mathrm{span}\{\zeta_n,\ 1 \le n \le N \equiv N_s n_b\}, \end{aligned} \tag{22}$$

where $1 \le N_s \le N_{s,\max}$; $\hat{\mathbf{u}}(\mathbf{k}_j) \equiv (u_1(\mathbf{k}_j), \ldots, u_{n_b}(\mathbf{k}_j))$ is the solution of (7) at $\mathbf{k}=\mathbf{k}_j$; and ζ_n, $1 \le n \le N$, are basis functions obtained after $u_i(\mathbf{k}_j)$, $1 \le i \le n_b$, $1 \le j \le N_s$, are orthonormalized with respect to $\langle \cdot | \cdot \rangle$. An approximation of $u_i(\mathbf{k})$ in W_N is then given by $u_{N,i}(\mathbf{k}) = \Sigma_{n=1}^{N} \alpha_{in}(\mathbf{k}) \zeta_n$. By construction, W_N is hierarchical, i.e., $W_{n_b N_s} \subset W_{n_b(N_s+1)}$, and a different set of $W_{n_b N_s}$, $1 \le N_s \le N_{s,\max}$, are constructed for each n_b. Here, n_b can be specified according to the applications that we look at. For example, for studying ground state properties, $n_b = n_v = 4$ is sufficient. For studying optical properties, n_b may need to be as high as 10. In calculations involving metallic structure, it is necessary to predetermine the highest band number with the maximum band energy we are interested in. Of course, an approximation based on $W_{n_b N_s}$ can be used to approximate any ith eigensolutions for which $i \le n_b$.

The reduced-basis approximation to $(\hat{\mathbf{u}}(\mathbf{k}), \hat{\lambda}(\mathbf{k}))$ is given by: for a given $\mathbf{k} \in \mathcal{D}$, find $(\hat{\mathbf{u}}_N(\mathbf{k}), \hat{\boldsymbol{\lambda}}_N(\mathbf{k})) \in (W_N)^{n_b} \times \mathbb{R}^{n_b}$ such that

$$\begin{aligned} &\langle v | \mathcal{A}(\mathbf{k}) | u_{N,i}(\mathbf{k}) \rangle = \lambda_{N,i}(\mathbf{k}) \langle v | u_{N,i}(\mathbf{k}) \rangle, \quad \forall \ \ v \in W_N, \\ &\langle u_{N,j}(\mathbf{k}) | u_{N,i}(\mathbf{k}) \rangle = \delta_{ij}, \quad i \le j \le n_b, \end{aligned} \tag{23}$$

for $1 \le i \le n_b$.

B. Discrete equations

We expand our reduced-basis approximation as

$$u_{N,i}(\mathbf{k}) = \sum_{n=1}^{N} u_{N,in}(\mathbf{k}) \zeta_n, \quad 1 \le i \le n_b, \tag{24}$$

and insert this representation into (23) to obtain

$$\begin{aligned} &\sum_{n=1}^{N} \left(A_{m,n}^{N,1} + \sum_{\ell=1}^{3} k_\ell A_{m,n}^{N,2,\ell} \right) u_{N,in}(\mathbf{k}) \\ &\quad = \lambda_{N,i}(\mathbf{k}) \sum_{n=1}^{N} M_{m,n}^{N} u_{N,in}(\mathbf{k}), \quad 1 \le m \le N, \quad 1 \le i \le n_b; \\ &\sum_{n=1}^{N} \sum_{m=1}^{N} u_{N,in}^{*}(\mathbf{k}) M_{n,m}^{N} u_{N,i'm}(\mathbf{k}) = \delta_{i,i'}, \quad 1 \le i,i' \le n_b; \end{aligned} \tag{25}$$

where

$$A^{N,1} \in \mathbb{C}^{N \times N}, \tag{26}$$

$$A^{N,2,\ell} \in \mathbb{C}^{N \times N}, \quad 1 \le \ell \le 3, \tag{27}$$

$$M^{N} \in \mathbb{C}^{N \times N} \tag{28}$$

are given by

$$A_{m,n}^{N,1} = \langle \zeta_m | \mathcal{A}_1 | \zeta_n \rangle, \quad 1 \le m,n \le N, \tag{29}$$

$$A_{m,n}^{N,2,\ell} = \langle \zeta_m | \mathcal{A}_2^{\ell} | \zeta_n \rangle, \quad 1 \le m,n \le N, \tag{30}$$

$$M^N_{m,n} = \langle \zeta_m | \zeta_n \rangle, \quad 1 \le m,n \le N. \tag{31}$$

Since ζ_m, $1 \le m \le N$, are orthonormalized, M^N is an identity matrix. We can then solve (25) using any eigenvalue solver.

C. Offline-online computational framework

We observe that we can now develop an efficient offline-online computational strategy for the rapid evaluation of $\lambda_{N,i}(\mathbf{k})$ for any $\mathbf{k}$ in $\mathcal{D}$—a strategy where the operation count in the online stage is independent of $\mathcal{N}_t$ and dependent only on N, which we expect to be much smaller than $\mathcal{N}_t$.

In the offline stage—performed once—we generate nested reduced-basis spaces $W_N \equiv \{\zeta_1, \ldots, \zeta_N\}$, $1 \le N \le N_{\max}$, at the cost of $N_s \mathcal{N}_t^\bullet$—the $\bullet$ denotes the actual computational complexity of the truth approximation, which due to sparsity, should be less than 3. We then form and store $A^{N,1}$, $A^{N,2,\ell}$, and M^N at the cost of $(Q+1)N^2\mathcal{N}_t^\bullet$. The storage of each matrix requires a space of $N \times N$.

In the online stage—performed many times for each new $\mathbf{k}$—we solve (25) for $\lambda_{N,i}(\mathbf{k})$, $1 \le i \le n_b$. The reconstruction of the reduced-basis system is QN^2 and the solution of the resulting discrete equations is of $O(N^3)$. The total operation count of the online stage is then $O(QN^2+N^3)$; we thus achieve a computational complexity that is independent of $\mathcal{N}_t$ and dependent only on N. The ability to calculate $\langle \zeta_m | \mathcal{A}_q | \zeta_n \rangle$ offline liberates the online computation from the $O(\mathcal{N}_t)$ complexity.

During the online stage, we also obtain the solution $u_{N,in}(\mathbf{k})$, $1 \le n \le N$, $1 \le i \le n_b$. The reduced-basis approximation of the wave functions is then simply given by (24). However, its evaluation will be of $O(N\mathcal{N}_t)$ due to the spatial dependence of the wave functions.

D. Convergence

For our convergence analysis, we introduce a test sample Ξ_T consisting of 488 $\mathbf{k}$ points distributed uniformly in $\mathcal{D}$. We will also define the reduced-basis approximation error in $\hat{\boldsymbol{\lambda}}_N$ as

$$\varepsilon^\lambda_{N,n_b} = \max_{\mathbf{k} \in \Xi_T} \epsilon^\lambda_{N,n_b}(\mathbf{k}), \tag{32}$$

where

$$\epsilon^\lambda_{N,n_b}(\mathbf{k}) = \max_{1 \le i \le n_b} |\lambda_{N,i}(\mathbf{k}) - \lambda_i(\mathbf{k})|. \tag{33}$$

From Fig. 3, we observe a rapidly convergent reduced-basis approximation as demonstrated by the convergence of $\varepsilon^\lambda_{N,n_b}$ at different n_b. In Table I, for a tolerance criterion of $\varepsilon^\lambda_{N,n_b} \le 1 \times 10^{-7}$, we see that N_s decreases with increasing n_b. However, for a coarser tolerance criterion of $\varepsilon^\lambda_{N,n_b} \le 1 \times 10^{-2}$, N_s remains the same for all n_b. This suggests two things. First, N_s must be above some critical value of N_s in order to achieve a reasonable approximation. Second, for N_s greater than this critical value, incremental improvement in the solutions can be obtained through inclusion of either solutions at more $\mathbf{k}$ points, higher eigenmodes, or both. This indicates that the modes have more general approximation properties—indeed, function approximation based on eigenmodes is a common technique in spectral methods, for example, expansion in Fourier modes or eigenfunctions of a suitable Sturm-Liouville problem [35].

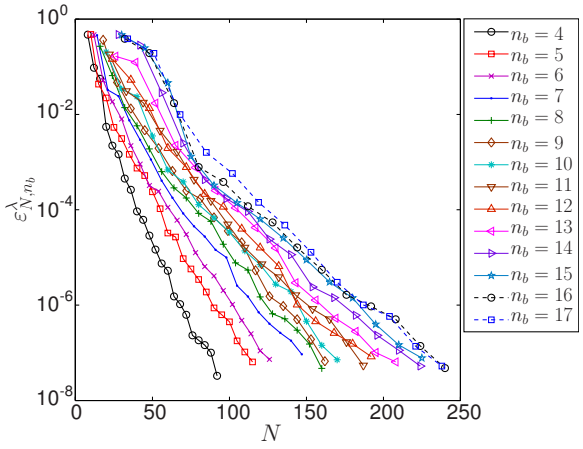

FIG. 3. (Color online) Convergence of the reduced-basis error in $\hat{\lambda}_N$, $\varepsilon^\lambda_{N,n_b}$ [given by (32)], with N for $4 \le n_b \le 17$.

We now examine the reduced-basis approximation error in $\hat{\mathbf{u}}_N$, defined as

$$\varepsilon^u_{N,n_b} = \max_{\mathbf{k} \in \Xi_T} \epsilon^u_{N,n_b}(\mathbf{k}), \tag{34}$$

where

$$\epsilon^u_{N,n_b}(\mathbf{k}) = \frac{\sum_{i=1}^{n_b} \| u_{N,i}(\mathbf{k}) - v_i^*(\mathbf{k}) \|_Y}{\sum_{j=1}^{n_b} \| u_j(\mathbf{k}) \|_Y} \tag{35}$$

and

$$v_i^*(\mathbf{k}) = \arg \min_{v \in P_i(\mathbf{k})} \| u_{N,i}(\mathbf{k}) - v \|_Y. \tag{36}$$

Here, $P_i(\mathbf{k})$ is the invariant eigensubspace associated with $\lambda_i(\mathbf{k})$. If $\lambda_i(\mathbf{k})$ has multiplicity n_p, $P_i(\mathbf{k})$ will be of dimension n_p and will consist of all eigenvectors $u_j(\mathbf{k})$ for which $\lambda_j(\mathbf{k}) = \lambda_i(\mathbf{k})$. Equations (35) and (36) have taken into account the degeneracy property of the wave functions. In particular, (36) removes the arbitrary phase factor present in the

TABLE I. N_s required to reduce the reduced-basis error in $\hat{\boldsymbol{\lambda}}_N$, $\varepsilon^\lambda_{N,n_b}$ [given by (32)], to below 1×10^{-2}, 1×10^{-4}, and 1×10^{-7} for $n_b = 4$, 8, 12, and 16.

	N_s		
n_b	$\varepsilon^\lambda_{N,n_b} < 1 \times 10^{-2}$	$\varepsilon^\lambda_{N,n_b} < 1 \times 10^{-4}$	$\varepsilon^\lambda_{N,n_b} < 1 \times 10^{-7}$
4	5	10	23
8	5	10	20
12	5	9	16
16	5	8	15

TABLE II. The reduced-basis error in $\hat{\mathbf{u}}_N$, ε^u_{N,n_b} [given by (34)] at different N_s for $n_b=4$. The reduced-basis error in $\hat{\boldsymbol{\lambda}}_N$, $\varepsilon^\lambda_{N,n_b}$, and $\ln(\varepsilon^u_{N,n_b})/\ln(\varepsilon^\lambda_{N,n_b})$ are listed as well for comparison.

N_s	ε^u_{N,n_b}	$\varepsilon^\lambda_{N,n_b}$	$\ln(\varepsilon^u_{N,n_b})/\ln(\varepsilon^\lambda_{N,n_b})$
4	5.71×10^{-1}	5.58×10^{-2}	0.19
8	1.33×10^{-2}	4.51×10^{-4}	0.56
12	2.99×10^{-3}	2.88×10^{-5}	0.56
16	8.04×10^{-4}	1.52×10^{-6}	0.53
20	2.33×10^{-4}	1.84×10^{-7}	0.54
23	9.20×10^{-5}	3.29×10^{-8}	0.54

numerical solutions, and determines a linear combination of eigenvectors in $P_i(\mathbf{k})$ such that the resulting v_i^* is optimally aligned with $u_{N,i}(\mathbf{k})$ before (35) is evaluated. Note that the determination of $v^*_{n_b}(\mathbf{k})$ may involve $P_{n_b}(\mathbf{k})$ consisting of $u_i(\mathbf{k})$ for which $i>n_b$. We determine $v_i^*(\mathbf{k})$ based on an alignment procedure detailed in [36].

Tables II and III show the error ε^u_{N,n_b} at different N_s for $n_b=4$ and 8. We observe that ε^u_{N,n_b} is of order $\sqrt{\varepsilon^\lambda_{N,n_b}}$ for sufficiently large N, consistent with the general approximation results for linear eigenvalue problem [37].

E. Sampling procedure

So far, we have not mentioned how the nested reduced-basis sample sets S_N are chosen. A sample set must be well chosen in order to obtain a rapidly convergent reduced-basis approximation, and a well-conditioned reduced-basis discrete system. In particular, we seek a sampling procedure that ensures "maximally independent" snapshots. We shall use the "greedy" adaptive sampling procedure outlined in [18,19,21].

We first assume that we are given a sample S_N and hence a reduced-basis space W_N, and the associated reduced-basis approximation (procedure to determine) $\hat{\mathbf{u}}_N(\mathbf{k})$ and $\hat{\boldsymbol{\lambda}}_N(\mathbf{k})$, $\forall\ \mathbf{k}\in\mathcal{D}$. We recall that $N=N_s n_b$. Then, for a suitably fine grid Ξ_T over the parameter space $\mathcal{D}$, we determine $\mathbf{k}^*_{N_s+1}=\arg\max_{\mathbf{k}\in\Xi_T}\epsilon^\bullet_N(\mathbf{k})$, where $\epsilon^\bullet_N(\mathbf{k})$ is an error measure of the reduced-basis approximation. Then we append $\mathbf{k}^*_{N_s+1}$ to S_N to form S_{N+n_b} and hence W_{N+n_b}. The procedure is repeated until $\varepsilon_{\max}=\epsilon^\bullet_N(\mathbf{k}^*_{N_s+1})$ is below $\varepsilon_{\mathrm{tol}}$, the tolerance we desire. This

TABLE III. The reduced-basis error in $\hat{\mathbf{u}}_N$, ε^u_{N,n_b} [given by (34)] at different N_s for $n_b=8$. The reduced-basis errors in $\hat{\boldsymbol{\lambda}}_N$, $\varepsilon^\lambda_{N,n_b}$, and $\ln(\varepsilon^u_{N,n_b})/\ln(\varepsilon^\lambda_{N,n_b})$ are listed as well for comparison.

N_s	ε^u_{N,n_b}	$\varepsilon^\lambda_{N,n_b}$	$\ln(\varepsilon^u_{N,n_b})/\ln(\varepsilon^\lambda_{N,n_b})$
4	3.73×10^{-1}	1.18×10^{-2}	0.22
8	1.33×10^{-2}	4.05×10^{-4}	0.55
12	2.97×10^{-3}	1.44×10^{-5}	0.52
16	2.36×10^{-4}	6.71×10^{-7}	0.59
19	9.92×10^{-5}	8.74×10^{-8}	0.57

Given S_{n_b}, W_{n_b};

Repeat $N_s=2,\ldots$

$\mathbf{k}^*_{N_s}=\arg\max_{\mathbf{k}\in\Xi_T}\epsilon^\bullet_{n_b(N_s-1)}(\mathbf{k})$;

$\varepsilon_{\max}=\epsilon^\bullet_{n_b(N_s-1)}(\mathbf{k}^*_{N_s})$;

$S_{n_bN_s}=S_{n_b(N_s-1)}\cup\mathbf{k}^*_{N_s}$;

$W_{n_bN_s}=W_{n_b(N_s-1)}+\mathrm{span}\ \{u_1(\mathbf{k}^*_{N_s}),\ldots,u_{n_b}(\mathbf{k}^*_{N_s})\}$;

until $\varepsilon_{\max}\le\varepsilon_{\mathrm{tol}}$.

FIG. 4. The greedy sampling procedure to construct an optimal S_N and W_N where $N=n_bN_s$.

tolerance $\varepsilon_{\mathrm{tol}}$ determines the size of $N_{\max}$. Figure 4 summarizes the greedy sampling procedure.

We may define $\epsilon^\bullet_N(\mathbf{k})$ in several ways. For example, we may use the reduced-basis approximation error in $\hat{\boldsymbol{\lambda}}$, $\epsilon^\lambda_{N,n_b}(\mathbf{k})$, given by (33). However, evaluations of $\epsilon^\lambda_{N,n_b}(\mathbf{k})$ are in fact expensive since truth solutions must first be obtained for all $\mathbf{k}\in\Xi_T$. However, if an *a posteriori* error estimator is available for which the evaluation does not require knowledge of the truth solutions, a more efficient procedure is possible [18,19]. In this work, all quoted offline computational cost will be based on the use of the $O(\mathcal{N}_i)$-independent *a posteriori* error estimator in the greedy algorithm. The derivation of the error estimator and the offline-online computational procedure for efficient evaluation of the error estimator are detailed in the Appendix.

This sampling procedure is relatively insensitive to the starting sample set S_{n_b} and thus the starting reduced-basis space W_{n_b}. This is because subsequent $\mathbf{k}$ points are determined by the sampling procedure; the selection criteria used ensure that solutions at these $\mathbf{k}$ points will form a good approximation space for $\hat{\mathbf{u}}(\mathbf{k})$, $\mathbf{k}\in\mathcal{D}$. Even if $\mathbf{k}_1$ proves to be a poor choice, $\mathbf{k}_2$ will always lead to a better reduced-basis approximation space. In addition, the presence of $u_i(\mathbf{k}_1)$, $1\le i\le n_b$ in $W_{N>n_b}$ will not adversely affect subsequent approximation based on $W_{N>n_b}$ because the reduced-basis method is a projection method—it will simply select an optimal combination of the basis functions in W_N. The effect of a poor starting reduced-basis space is then limited to increasing the required N_s by 1. Indeed, for the case of $n_b=4$ with a requirement of $\varepsilon^\lambda_{N,n_b}<1\times10^{-4}$, the N_s required varies between 10 and 11 for 50 different starting $\mathbf{k}$ points.

IV. RESULTS

A. Comparison with plane-wave method

We first consider only the online computational cost needed to approximate $E_i(\mathbf{k})$, $1\le i\le n_b$, at a single $\mathbf{k}$ point. In Table IV, we compare the computational cost required by the reduced-basis method and the plane-wave method to achieve similar approximation errors for $\mathbf{k}=(2\pi/a)$

TABLE IV. Comparison of the computational cost of reduced-basis method and plane-wave method required to achieve a similar level of accuracy. Comparison is made based on a single **k** point given by $(2\pi/a)(0.2,0.2,0.2)$ and $n_b=4$.

	W_N			Plane wave	
N	Time (s)	$\varepsilon^{\lambda}_{N,n_b}$	$\mathcal{N}$	Time (s)	$\varepsilon^{\lambda}_{\mathcal{N},n_b}$
16	0.001	8.88×10^{-3}	65	0.029	9.37×10^{-3}
24	0.002	1.72×10^{-3}	113	0.053	1.95×10^{-3}
32	0.004	8.08×10^{-4}	137	0.063	6.99×10^{-4}
40	0.008	4.83×10^{-4}			
48	0.013	9.23×10^{-8}	531	0.457	9.34×10^{-8}
56	0.020	2.98×10^{-8}	609	0.584	1.76×10^{-9}

$\times(0.2,0.2,0.2)$ and $n_b=4$. For this purpose, we introduce the error measure

$$\varepsilon^{\lambda}_{\mathcal{N},n_b} = \max_{\mathbf{k}\in\Xi_T} \max_{1\le i\le n_b} |\lambda_i^{\mathcal{N}}(\mathbf{k}) - \lambda_i(\mathbf{k})|, \tag{37}$$

for the plane-wave approximation based on $Y^{\mathcal{N}}$. We observe that the computational saving achieved ranges from a factor of 15 to 35. For both, we use the eigenvalue solvers in MATLAB.

We now take the computational cost of the offline stage into consideration. Since we use an *a posteriori* error estimator, we only need to determine N_s truth solutions, where N_s is usually very small. The total offline computational cost is also determined by the maximum N_s, $N_{s,\max}$, usually chosen based on the highest accuracy we want for our approximation. In this section, we choose $N_{s,\max}$ to be 14 so that it corresponds to the maximum value of N in Table IV.

For $n_b=4$, the offline stage requires a total computational time of 67 s. Even with the *a posteriori* error estimation procedure, there must be a need to evaluate (7) at more than 1000 **k** points in order to justify the offline computational cost, assuming we only require $\varepsilon^{\lambda}_{N,n_b}$ to be of $O(10^{-4})$. This emphasizes the many-query limit where the reduced-basis method is most useful. We shall provide in the next section some examples where we indeed need to determine band energies at many **k** points.

B. Sample problems

To evaluate (1)–(3) efficiently, we approximate $E_i(\mathbf{k})$ by a reduced-basis approximant $E_{N,i}(\mathbf{k})$. In addition, for comparison we also introduce a plane-wave approximation based on $Y^{\mathcal{N}}$ for which $\mathcal{N}\le\mathcal{N}_t$. The truth approximations to the quantities (1)–(3) are obtained from a plane-wave approximation based on $Y^{\mathcal{N}_t}$ with $n_k=n_{k,t}=13\,200$. The offline computational cost is based on the greedy sampling algorithm outlined in Sec. III E with a convergence criterion given by $\Delta^{\lambda}_{N,n_b}\le 2\times10^{-2}$, where Δ^{λ}_{N,n_b} is the *a posteriori* error estimator.

1. Integrated density of states

Here, we look at a hypothetical problem of determining $I_{\text{dif}}(6\text{ eV})$ where $I_{\text{dif}}(E)=I(E)-I(E_4(0))$ and $E_4(0)$ is the highest valence band energy. Nevertheless, typical applications look at the inverse problem, i.e., determining an E_0 such that $I(E_0)$ is equivalent to a certain value—in the case of the Fermi level, we determine an E_0 that satisfies $I(E_0)=n_v$.

We denote a reduced-basis approximation to I_{dif} by $I_{\text{dif},n_k,N}$ and a plane-wave approximation by $I^{\mathcal{N}}_{\text{dif},n_k}$; n_k denotes the number of **k** points used. The truth approximation is denoted by $I^0_{\text{dif}}=I^{\mathcal{N}_t}_{\text{dif},n_{k,t}}$. To achieve a convergence criterion of $|I_{\text{dif},N}-I^0_{\text{dif}}|\le 0.01$, a reduced-basis approximation requires $N=36$ and $n_k=572$; the online computational cost is 3.8 s. For the plane-wave method, a combination of $\mathcal{N}=137$ and $n_k=572$ is required; the computational cost is 26 s. Thus, the reduced-basis method is seven times faster than the plane-wave method. Here, $n_b=9$.

However, we note that the offline computational cost of the reduced-basis approximation is 102 s. The reduced-basis method is thus competitive only if we need to evaluate the integrated density of states for larger n_k.

2. Joint density of states

We denote a reduced-basis approximation of J by J_{N,n_k}, a plane-wave approximation by $J^{\mathcal{N}}_{n_k}$, and a truth approximation by $J^0\equiv J^{\mathcal{N}_t}_{n_{k,t}}$. Figure 5 shows that a good approximation of J_{N,n_k} is obtained when $n_b=12$, $N=36$, and $n_k=8800$—the approximation is close to the truth solution for the entire range

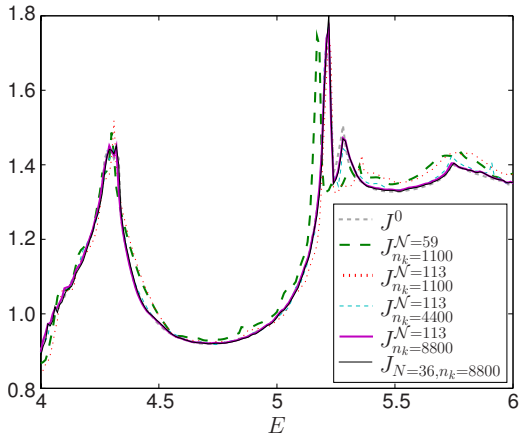

FIG. 5. (Color online) Different approximations to the joint density of states—J_{N,n_k}, $J^{\mathcal{N}}_{n_k}$, and J^0—versus energy E in eV.

TABLE V. Comparison of the cost required to compute $J_{N=36,n_k=8800}$ and $J^{\mathcal{N}=113}_{n_k=8800}$.

	J_{N,n_k}	$J^{\mathcal{N}}_{n_k}$
Dimension	$N=36$	$\mathcal{N}=113$
n_k	8800	8800
Computational time	Online: 123 s	
	Offline: 85 s	Total: 458 s
	Total: 208 s	

of E between 4 and 6 eV. Based on the same criterion, a good plane-wave approximation $J^{\mathcal{N}}_{n_k}$ is obtained when $n_b=12$, $\mathcal{N}=113$, and $n_k=8800$.

For the above two approximations, we show the computational cost in Table V. For the reduced-basis approximation, the convergence criterion used in the offline stage gives $N_{s,\max}=7$ and an offline computational cost of 85 s. The online computational cost is 123 s, thus giving a total computational cost of 208 s. For the plane-wave approximation, the computational cost is 458 s; we achieve a factor-of-2 saving in the computational cost.

3. Dielectric function

Again, we shall denote a reduced-basis approximation of ε_2 by ε_{2,N,n_k}, a plane-wave approximation by $\varepsilon^{\mathcal{N}}_{2,n_k}$ and a truth approximation by $\varepsilon_2^0=\varepsilon^{\mathcal{N}_t}_{2,n_{k,t}}$. Compared to the determination of the joint density of states, we need to approximate $\int_\Omega[\partial u_j(\mathbf{k})/\partial x_\ell]u_i^*(\mathbf{k})$, $1\le i\le n_v$, $n_v+1\le j\le n_b$, in addition to $E_i(\mathbf{k})$. However, the approximation error in $\int_\Omega[\partial u_j^N(\mathbf{k})/\partial x_\ell]u_i^{N*}(\mathbf{k})$ (or $\int_\Omega[\partial u_j^{\mathcal{N}}(\mathbf{k})/\partial x_\ell]u_i^{\mathcal{N}*}(\mathbf{k})$) is of the same order as the error in $E_{N,i}$ (or $E_i^{\mathcal{N}}$); as such, the accuracy requirement is not higher than in the previous two problems.

We reuse the reduced-basis approximation we have constructed for the joint density of states. From Fig. 6, we see that a good reduced-basis approximation is obtained when $N=36$ and $n_k=8800$, while a good plane-wave approximation is obtained when $\mathcal{N}=113$ and $n_k=8800$. Table VI shows that we obtain a factor-of-1.5 saving in the computational

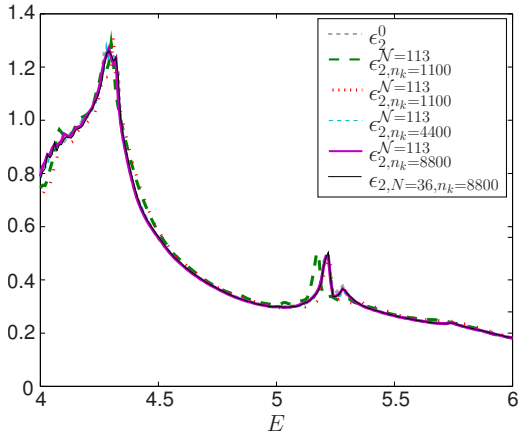

FIG. 6. (Color online) Different approximations to the dielectric function—ε_{2,N,n_k}, $\varepsilon^{\mathcal{N}}_{2,n_k}$, and ε_2^0—versus energy E in eV.

TABLE VI. Comparison of the cost required to compute $\varepsilon_{2,N=36,n_k=8800}$ and $\varepsilon^{\mathcal{N}=113}_{2,n_k=8800}$.

	ε_{2,N,n_k}	$\varepsilon^{\mathcal{N}}_{2,n_k}$
Dimension	$N=36$	$\mathcal{N}=113$
Computational time	Online: 266 s	
	Offline: 85 s	Total: 553 s
	Total: 351 s	

cost. The decrease in the gain obtained through the reduced-basis method is due to increase in the overhead computational cost unrelated to approximation of the u_i and λ_i. In the reduced-basis approximation of ε_2, this overhead cost amounts to nearly 80% of the total computational cost. On the other hand, for the reduced-basis approximation of J, this overhead cost is only 54% of the total computational cost.

V. EXTENSION

In a typical calculation based on pseudopotential density functional theory model [27], V_{eff} is either not explicitly constructed or not easily accessible to the user. The inaccessibility of V_{eff} does not allow the construction of the discrete reduced-basis matrix $A^{N,1}_{m,n}=\langle\zeta_m|\mathcal{A}_1|\zeta_n\rangle$, $1\le m$, $n\le N$, as outlined in Sec. III. Here we shall demonstrate a trick by which we obtain $A^{N,1}$ based solely on the solutions $(\hat{\mathbf{u}}(\mathbf{k}),\hat{\boldsymbol{\lambda}}(\mathbf{k}))$, $\mathbf{k}\in S_N$, which are typical outputs of any electronic structure calculation.

Suppose we are given a sample set $S_N=\{\mathbf{k}_1,\ldots,\mathbf{k}_{N_s}\}$ and associated solutions $(\hat{\mathbf{u}}(\mathbf{k}_n),\hat{\boldsymbol{\lambda}}(\mathbf{k}_n))\in(Y^{n_b}\times\mathbb{R}^{n_b})$, $1\le n\le N_s$. From (7), we can write

$$\langle u_{m'}(\mathbf{k}_{n'})|\mathcal{A}_1|u_m(\mathbf{k}_n)\rangle=\lambda_m(\mathbf{k}_n)\langle u_{m'}(\mathbf{k}_{n'})|u_m(\mathbf{k}_n)\rangle - \sum_{l=1}^{3} k_{l,n}\langle u_{m'}(\mathbf{k}_{n'})|\mathcal{A}_2^l|u_m(\mathbf{k}_n)\rangle, \tag{38}$$

for $1\le m$, $m'\le n_b$ and $1\le n$, $n'\le N_s$, since all $u_m(\mathbf{k}_n)$, $1\le m\le n_b$, $1\le n\le N_s$ reside in the same space Y. In addition,

$$\zeta_i=\sum_{m=1}^{n_b}\sum_{n=1}^{N}\alpha^i_{m,n}u_m(k_n),\quad 1\le i\le N, \tag{39}$$

where $N=N_sn_b$ and $\alpha^i_{m,n}$ are known from our orthogonalization procedure. The matrix $A^{N,1}$ is then simply given by

$$A^{N,1}_{i,j}=\sum_{m=1}^{n_b}\sum_{n=1}^{N}\sum_{m'=1}^{n_b}\sum_{n'=1}^{N}\alpha^i_{m,n}\alpha^j_{m',n}\langle u_{m'}(\mathbf{k}_{n'})|\mathcal{A}_1|u_m(\mathbf{k}_n)\rangle, \quad 1\le i,j\le N. \tag{40}$$

We have found the reduced-basis solutions obtained through this procedure to be identical to those from the procedure outlined in Sec. III. We note that the above construction is well defined only if Y remains the same for all $\mathbf{k}$ when determining the truth approximation.

ACKNOWLEDGMENTS

I would like to thank A. T. Patera, E. Cancès, C. Le Bris, and Y. Maday for their support and guidance, and N. Marzari for useful discussions. This work was supported by Singapore-MIT Alliance and by DARPA and AFOSR under Grant No. FA9550-05-1-0114.

APPENDIX: *A POSTERIORI* ERROR ESTIMATION

A posteriori error estimation procedures are well developed for algebraic eigenvalue problems [38–40] and approximation of eigenvalue problems based on, say, the finite-element method [41,42]. Simple error estimates for a computed eigenvalue can be determined from the residual vector. Within the reduced-basis context, asymptotic error bounds are first formulated for a symmetric positive definite eigenvalue problem in [16]. In addition, [16] provides a very efficient procedure by which these bounds can be computed through the offline-online computational framework. However, these error estimates usually do not provide rigorous bounds that can function as a certificate of fidelity for our reduced-basis approximation. Thus, previous work on reduced-basis approximation of partial differential equations [15,17–21] places significant emphasis on obtaining inexpensive and sharp error bounds for the output of interest.

Nonrigorous error bounds can nonetheless be very useful. In the greedy adaptive sampling procedure outlined in Sec. III E, an asymptotic error bound may be sufficient to serve as a guide in the construction of the reduced-basis sample set. Here, we shall construct an asymptotic *a posteriori* error bound for $\hat{\boldsymbol{\lambda}}_N(\mathbf{k})$ to be used in our sampling procedure. The development of the bound parallels that of algebraic eigenvalue problems.

1. Derivation

For $i=1,\ldots,n_b$, we define the residual as

$$R_i(v;\mathbf{k}) = \langle v|\mathcal{A}(\mathbf{k})|u_{N,i}(\mathbf{k})\rangle - \lambda_{N,i}(\mathbf{k})\langle v|u_{N,i}(\mathbf{k})\rangle, \quad \text{(A1)}$$

for $\forall\ v\in Y$. We also define a reconstructed error $\hat{e}_i$ in Y, such that

$$\langle v|\hat{\mathcal{A}}|\hat{e}_i\rangle = R_i(v;\mathbf{k}), \quad \forall\ v\in Y, \quad \text{(A2)}$$

where

$$\langle v|\hat{\mathcal{A}}|w\rangle = \langle v|\mathcal{A}_1|w\rangle + \gamma\langle w|v\rangle, \quad \gamma = 1+|\lambda_1(0)|, \quad \text{(A3)}$$

$$\|R_i(\cdot;\mathbf{k})\| \equiv \sup_{v\in Y}\frac{R_i(v;\mathbf{k})}{\langle v|\hat{\mathcal{A}}|v\rangle^{1/2}} = \langle \hat{e}_i|\hat{\mathcal{A}}|\hat{e}_i\rangle^{1/2}, \quad \text{(A4)}$$

and $\|\cdot\| = \langle\cdot|\hat{\mathcal{A}}|\cdot\rangle^{1/2}$.

We now define $\langle v|\mathcal{A}^+(\mathbf{k})|w\rangle = \langle v|\mathcal{A}(\mathbf{k})|w\rangle + \gamma\langle v|w\rangle$ and introduce the following eigenvalue problem: for $\mathbf{k}\in\mathcal{D}$, find $(\hat{\mathbf{u}}^+(\mathbf{k}),\hat{\boldsymbol{\lambda}}^+(\mathbf{k}))\in(Y^{n_b}\times\mathbb{R}^{n_b})$ such that

$$\langle v|\mathcal{A}^+(\mathbf{k})|u_i^+(\mathbf{k})\rangle = \lambda_i^+(\mathbf{k})\langle v|u_i^+(\mathbf{k})\rangle, \quad \forall\ v\in Y, \quad \text{(A5)}$$

$$\langle u_j^+(\mathbf{k})|u_i^+(\mathbf{k})\rangle = \delta_{ij}, \quad i\le j\le n_b, \quad \text{(A6)}$$

for $1\le i\le n_b$. It is clear that $\hat{\mathbf{u}}^+(\mathbf{k})=\hat{\mathbf{u}}(\mathbf{k})$ and $\lambda_i^+(\mathbf{k})=\lambda_i(\mathbf{k})+\gamma$.

Proposition 1. Given $\langle v|\hat{\mathcal{A}}|w\rangle = \langle v|\mathcal{A}_1|w\rangle + \gamma\langle v|w\rangle$ and $\gamma = 1+|\lambda_1(0)|$, we have

$$\langle v|\mathcal{A}^+(\mathbf{k})|v\rangle \equiv \langle v|\hat{\mathcal{A}}|v\rangle \ge \langle v|v\rangle \ge 0. \quad \text{(A7)}$$

Proof. First, we note that $\langle v|\mathcal{A}_2^j|v\rangle = 0$, for $j=1,\ldots,3$: let $v=v_1+\mathrm{i}v_2$, and $v_1,v_2\in\mathbb{R}$; then

$$\begin{aligned}\langle v|\mathcal{A}_2^j|v\rangle &= -\mathrm{i}\int_\Omega\left(\frac{\partial v_1}{\partial x_j}+\mathrm{i}\frac{\partial v_2}{\partial x_j}\right)(v_1-\mathrm{i}v_2)\\ &= -\mathrm{i}\int_\Omega\left(\frac{\partial v_1}{\partial x_j}v_1+\frac{\partial v_2}{\partial x_j}v_2\right)\\ &\quad -\int_\Omega\frac{\partial v_1}{\partial x_j}v_2+\int_\Omega\frac{\partial v_2}{\partial x_j}v_1 = 0, \quad \text{(A8)}\end{aligned}$$

since

$$\int_\Omega\frac{\partial v_1}{\partial x_j}v_2 = -\int_\Omega\frac{\partial v_2}{\partial x_j}v_1,$$

$$\int_\Omega\frac{\partial v_1}{\partial x_j}v_1 = 0, \quad \int_\Omega\frac{\partial v_2}{\partial x_j}v_2 = 0.$$

We can now prove the left equality:

$$\begin{aligned}\langle v|\hat{\mathcal{A}}|v\rangle &= \langle v|\mathcal{A}_1|v\rangle + \gamma\langle v|v\rangle = \langle v|\mathcal{A}(\mathbf{k})|v\rangle + \gamma\langle v|v\rangle\\ &= \langle v|\mathcal{A}^+(\mathbf{k})|v\rangle, \quad \text{(A9)}\end{aligned}$$

since $\langle v|\hat{\mathcal{A}}|v\rangle = \langle v|\mathcal{A}_1|v\rangle + \Sigma_{j=1}^3\langle v|\mathcal{A}_2^j|v\rangle$ based on (A8)

To prove the right inequality, we note that

$$\langle v|\mathcal{A}_1|v\rangle \ge \lambda_1(0)\langle v|v\rangle. \quad \text{(A10)}$$

Then,

$$\begin{aligned}\langle v|\hat{\mathcal{A}}|v\rangle &= \langle v|\mathcal{A}_1|v\rangle + [1+|\lambda_1(0)|]\langle v|v\rangle\\ &\ge [1+\lambda_1(0)+|\lambda_1(0)|]\langle v|v\rangle \ge \langle v|v\rangle, \quad \text{(A11)}\end{aligned}$$

since $\lambda_1(0)\le 0$. This concludes the proof of Proposition 1.■

Proposition 2. Assume that our reduced-basis approximation is convergent in the sense that

$$\lambda_{N,i}(\mathbf{k}) \to \lambda_i(\mathbf{k}), \quad 1\le i\le n_b \quad \text{as} \quad N\to\infty. \quad \text{(A12)}$$

Then, for large N and $i=1,\ldots,n_b$,

$$\left|\frac{\lambda_i(\mathbf{k})-\lambda_{N,i}(\mathbf{k})}{\lambda_i(\mathbf{k})+\gamma}\right| \le \frac{\|R_i(\cdot;\mathbf{k})\|}{(\lambda_{N,i}(\mathbf{k})+\gamma)^{1/2}}. \quad \text{(A13)}$$

Proof. For $i=1,\ldots,n_b$, we define $\tilde{e}\in Y$ as

$$\langle v|\mathcal{A}^+(\mathbf{k})|\tilde{e}_i\rangle = R_i(v;\mathbf{k}), \quad \forall\ v\in Y, \quad \text{(A14)}$$

$$\|\|R_i(\cdot;\mathbf{k})\|\| \equiv \sup_{v\in Y} \frac{R_i(v;\mathbf{k})}{\langle v|\mathcal{A}^+(\mathbf{k})|v\rangle^{1/2}} = \langle \tilde{e}_i|\mathcal{A}^+(\mathbf{k})|\tilde{e}_i\rangle^{1/2}, \tag{A15}$$

and $\|\|\cdot\|\|=\langle\cdot|\mathcal{A}^+(\mathbf{k})|\cdot\rangle^{1/2}$. From (A7), we then have

$$\|\|R_i(\cdot;\mathbf{k})\|\| \le \|R_i(\cdot;\mathbf{k})\|. \tag{A16}$$

Let $u_{N,i}(\mathbf{k})=\Sigma_{j=1}^{\mathcal{N}_t}\alpha_j u_j(\mathbf{k})$ and $\tilde{e}_i=\Sigma_{j=1}^{\mathcal{N}_t}\beta_j u_j(\mathbf{k})$. By substituting them into (A14), we obtain

$$\beta_j = \alpha_j\left(\frac{\lambda_j(\mathbf{k})-\lambda_{N,i}(\mathbf{k})}{\lambda_j^+(\mathbf{k})}\right). \tag{A17}$$

Then,

$$\|\|R_i(\cdot;\mathbf{k})\|\|^2 = \langle \tilde{e}_i|\mathcal{A}^+(\mathbf{k})|\tilde{e}_i\rangle = \sum_{j=1}^{\mathcal{N}_t}\beta_j^2\lambda_j^+(\mathbf{k})\langle u_j(\mathbf{k})|u_j(\mathbf{k})\rangle = \sum_{j=1}^{\mathcal{N}_t}\alpha_j^2\left(\frac{\lambda_j(\mathbf{k})-\lambda_{N,i}(\mathbf{k})}{\lambda_j^+(\mathbf{k})}\right)^2\lambda_j^+(\mathbf{k}). \tag{A18}$$

Dividing by $\lambda_{N,i}(\mathbf{k})+\gamma$, we obtain

$$\frac{\|\|R_i(\cdot;\mathbf{k})\|\|^2}{\lambda_{N,i}(\mathbf{k})+\gamma} = \frac{1}{\langle u_{N,i}(\mathbf{k})|\mathcal{A}^+(\mathbf{k})|u_{N,i}(\mathbf{k})\rangle}\sum_{j=1}^{\mathcal{N}_t}\alpha_j^2\left(\frac{\lambda_j(\mathbf{k})-\lambda_{N,i}(\mathbf{k})}{\lambda_j^+(\mathbf{k})}\right)^2\lambda_j^+(\mathbf{k}) \ge \min_{1\le j\le\mathcal{N}_t}\left(\frac{\lambda_j(\mathbf{k})-\lambda_{N,i}(\mathbf{k})}{\lambda_j^+(\mathbf{k})}\right)^2\frac{\sum_{j=1}^{\mathcal{N}_t}\alpha_j^2\lambda_j^+(\mathbf{k})}{\sum_{j'=1}^{\mathcal{N}_t}\alpha_{j'}^2\lambda_{j'}^+(\mathbf{k})} = \min_{1\le j\le\mathcal{N}_t}\left(\frac{\lambda_j(\mathbf{k})-\lambda_{N,i}(\mathbf{k})}{\lambda_j^+(\mathbf{k})}\right)^2. \tag{A19}$$

Therefore, in the asymptotic limit as defined by (A12), $i=\arg\min_{1\le j\le n_b}|[\lambda_j(\mathbf{k})-\lambda_{N,i}(\mathbf{k})]/\lambda_j^+(\mathbf{k})|$ and

$$\left|\frac{\lambda_i(\mathbf{k})-\lambda_{N,i}(\mathbf{k})}{\lambda_i(\mathbf{k})+\gamma}\right| \le \frac{\|\|R_i(\cdot;\mathbf{k})\|\|}{(\lambda_{N,i}(\mathbf{k})+\gamma)^{1/2}} \le \frac{\|R_i(\cdot;\mathbf{k})\|}{(\lambda_{N,i}(\mathbf{k})+\gamma)^{1/2}}, \tag{A20}$$

from (A16). This proves (A13). ■

Remark 1. In the asymptotic limit defined by (A12), we can also write (A13) as

$$|\lambda_i(\mathbf{k})-\lambda_{N,i}(\mathbf{k})| \le \|R_i(\cdot;\mathbf{k})\|(\lambda_{N,i}(\mathbf{k})+\gamma)^{1/2}. \tag{A21}$$

2. Offline-online computational framework

We can also construct very efficient offline-online computational strategies for the evaluation of our error estimators. From (A2) and our reduced-basis approximation, we have

$$\langle v|\hat{\mathcal{A}}|\hat{e}_i\rangle = \sum_{q=1}^{Q}\Theta_q(\mathbf{k})\langle v|\mathcal{A}_q|u_{N,i}(\mathbf{k})\rangle - \lambda_{N,i}(\mathbf{k})\langle v|u_{N,i}(\mathbf{k})\rangle, \quad v\in Y, \tag{A22}$$

for $1\le i\le n_b$. It follows from linear superposition that

$$\hat{e}_i(\mathbf{k}) = \sum_{q=1}^{Q}\sum_{n=1}^{N}\Theta_q(\mathbf{k})u_{N,in}(\mathbf{k})\xi_n^q - \lambda_{N,i}(\mathbf{k})\sum_{n=1}^{N}u_{N,in}(\mathbf{k})\xi_n^0, \tag{A23}$$

where

$$\langle v|\hat{\mathcal{A}}|\xi_n^q\rangle = \langle v|\mathcal{A}_q|\zeta_n\rangle, \quad 1\le q\le Q, \tag{A24}$$

$$\langle v|\hat{\mathcal{A}}|\xi_n^0\rangle = \langle v|\zeta_n\rangle, \tag{A25}$$

for $1\le n\le N$ and $\forall\ v\in Y$. Then, $\|R_i(\cdot;\mathbf{k})\|\equiv\langle\hat{e}_i|\hat{\mathcal{A}}|\hat{e}_i\rangle$ is given by

$$\|R_i(\cdot;\mathbf{k})\|^2 = \sum_{n=1}^{N}\sum_{n'=1}^{N}\sum_{q=1}^{Q}\sum_{q'=1}^{Q}u_{N,in}(\mathbf{k})u_{N,in'}(\mathbf{k})\Theta_q(\mathbf{k})\Theta_{q'}(\mathbf{k})\hat{A}_{n,n'}^{q,q'} + \sum_{n=1}^{N}\sum_{n'=1}^{N}\lambda_{N,i}^2(\mathbf{k})u_{N,in}(\mathbf{k})u_{N,in'}(\mathbf{k})\hat{A}_{n,n'}^{0,0} + \sum_{n=1}^{N}\sum_{n'=1}^{N}\sum_{q=1}^{Q}u_{N,in}(\mathbf{k})\lambda_{N,i}(\mathbf{k})\Theta_q(\mathbf{k})\hat{A}_{n,n'}^{q,0}; \tag{A26}$$

TABLE VII. Variation of the actual error $\varepsilon^{\lambda}_{N,n_b}$, $\max_{\mathbf{k}\in\Xi_T}\Delta^{\lambda}_{N,n_b}(\mathbf{k})$, and $\bar{\eta}^{\lambda}_{N,n_b}$ with N for $n_b=4$.

N	$\varepsilon^{\lambda}_{N,n_b}$	$\max_{\mathbf{k}\in\Xi_T}\Delta^{\lambda}_{N,n_b}$	$\bar{\eta}^{\lambda}_{N,n_b}$
16	5.58×10^{-2}	1.78×10^{-1}	18.0
32	4.51×10^{-5}	2.26×10^{-2}	98.7
48	2.88×10^{-5}	5.83×10^{-3}	555
64	1.52×10^{-6}	1.45×10^{-3}	2535
80	1.84×10^{-7}	5.04×10^{-4}	6876

TABLE VIII. Variation of the actual error $\varepsilon^{\lambda}_{N,n_b}$, $\max_{\mathbf{k}\in\Xi_T}\Delta^{\lambda}_{N,n_b}(\mathbf{k})$, and $\bar{\eta}^{\lambda}_{N,n_b}$ with N for $n_b=8$.

N	$\varepsilon^{\lambda}_{N,n_b}$	$\max_{\mathbf{k}\in\Xi_T}\Delta^{\lambda}_{N,n_b}$	$\bar{\eta}^{\lambda}_{N,n_b}$
32	1.39×10^{-3}	1.26×10^{-2}	21.9
64	2.89×10^{-4}	1.95×10^{-2}	203.2
96	1.89×10^{-5}	5.56×10^{-3}	1063
128	6.59×10^{-7}	1.10×10^{-3}	4286

where $\hat{A}^{q,q'}\in\mathbb{R}^N\times\mathbb{R}^N$, $0\le q,\ q'\le Q$ are given by $\hat{A}^{q,q'}_{n,n'}=\langle\xi^{q'}_{n'}|\hat{\mathcal{A}}|\xi^{q}_{n}\rangle$, $0\le q,q'\le Q$, $1\le n,n'\le N$. We now see that the dual norm of the residual is the sum of products of parameter-dependent functions and parameter-independent functionals. The offline-online decomposition is now clear.

In the offline stage, we compute ξ^q_n, $0\le q\le Q$, $1\le n\le N$, based on (A22) at the cost of $O((Q+1)N\mathcal{N}^{\bullet}_t)$, where the $\bullet$ denotes the computational complexity of the linear solver used to obtain ξ^q_n. We then evaluate $\hat{A}^q$ and $\hat{M}$ at the cost of $O((Q+1)N^2\mathcal{N}^2_t)$. We store the matrices $\hat{A}^q$ and $\hat{M}$ at a total cost of $(Q+1)N^2$.

In the online stage, we simply evaluate the sum (A23) for a given $u_{N,i}$ and $\lambda_{N,i}$, $1\le i\le n_b$. The operation count is only $O(n_bQ^2N^2)$. The online complexity is thus independent of $\mathcal{N}_t$. Unless Q is large, the online cost to compute the error estimator is then a fraction of the cost required to obtain $u_{N,i}$ and $\lambda_{N,i}$.

3. Numerical results

We define our error estimator $\Delta^{\lambda}_{N,n_b}(\mathbf{k})$ as

$$\Delta^{\lambda}_{N,n_b}(\mathbf{k})=\max_{1\le i\le n_b}\|R_i(\cdot;\mathbf{k})\|(\lambda_{N,i}(\mathbf{k})+\gamma)^{1/2}, \tag{A27}$$

and the effectivity measure as

$$\eta^{\lambda}_{N,n_b}(\mathbf{k})=\frac{\Delta^{\lambda}_{N,n_b}(\mathbf{k})}{\varepsilon^{\lambda}_{N,n_b}(\mathbf{k})}. \tag{A28}$$

For brevity, we present in Tables VII and VIII the results for $n_b=4$ and 8 and define $\bar{\eta}^{\lambda}_{N,n_b}$ as the mean of $\eta^{\lambda}_{N,n_b}(\mathbf{k})$ for $\mathbf{k}\in\Xi_T$. We obtain an error estimator with effectivity closer to 1 at smaller N. However, this effectivity diverges as N increases since $|\lambda_{N,i}(\mathbf{k})-\lambda_i(\mathbf{k})|$ is of $O(\|u_{N,i}(\mathbf{k})-u_i(\mathbf{k})\|^2_Y)\approx O(\|R_i(\cdot;\mathbf{k})\|^2)$ and $\Delta^{\lambda}_{N,n_b}(\mathbf{k})$ is of $O(\|R_i(\cdot;\mathbf{k})\|)$. As a result, the use of (A27) as an error measure in the greedy sampling procedure may lead to unnecessarily large N. Thus, we would like to emphasize that, for the current problem, the error estimation procedure is only used to determine a good set of sample points given $N_{\max}$; it is not used to determine the size of $N_{\max}$.

[1] X. Wang, J. R. Yates, I. Souza, and D. Vanderbilt, Phys. Rev. B **74**, 195118 (2006).
[2] S. Hamel, A. Williamson, H. Wilson, F. Gygi, G. Galli, E. Ratner, and D. Wack http://meetings.aps.org/link/BAPS.2007.MAR.L43.7.
[3] H. Wilson, G. Galli, F. Gygi, S. Hamel, A. Williamson, E. Ratner, and D. Wack http://meetings.aps.org/link/BAPS.2007.MAR.J41.15.
[4] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2001).
[5] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).
[6] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
[7] F. Gygi, J.-L. Fattebert, and E. Schwegler, Comput. Phys. Commun. **155**, 1 (2003).
[8] B. O. Almroth, P. Stern, and F. A. Brogan, AIAA J. **16**, 525 (1978).
[9] A. K. Noor and J. M. Peters, AIAA J. **18**, 455 (1980).
[10] J. P. Fink and W. C. Rheinboldt, Z. Angew. Math. Mech. **63**, 21 (1983).
[11] M. D. Gunzburger, *Finite Element Methods for Viscous Incompressible Flows* (Academic Press, New York, 1989).
[12] K. Ito and S. S. Ravindran, J. Comput. Phys. **143**, 403 (1998).
[13] J. S. Peterson, SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **10**, 777 (1989).
[14] T. A. Porsching, Math. Comput. **45**, 487 (1985).
[15] M. A. Grepl, N. C. Nguyen, K. Veroy, A. T. Patera, and G. R. Liu, in *Towards Real-Time PDE-Constrained Optimization* (SIAM, Philadelphia, 2007), pp. 197–216.
[16] L. Machiels, Y. Maday, I. B. Oliveira, A. Patera, and D. Rovas, C. R. Acad. Sci., Ser. I: Math. **331**, 153 (2000).
[17] Y. Maday, A. T. Patera, and G. Turinici, C. R. Acad. Sci., Ser. I: Math.. **335**, 289 (2002).
[18] C. Prud'homme, D. Rovas, K. Veroy, Y. Maday, A. Patera, and G. Turinici, J. Fluids Eng. **124**, 70 (2002).
[19] N. C. Nguyen, K. Veroy, and A. T. Patera, in *Handbook of Materials Modeling*, edited by S. Yip (Springer, Berlin, 2005), pp. 1523–1558.
[20] K. Veroy and A. T. Patera, Int. J. Numer. Methods Fluids **47**, 773 (2005).
[21] K. Veroy, C. Prud'homme, D. V. Rovas, and A. T. Patera, in *Proceedings of the 16th AIAA Computational Fluid Dynamics Conference* (AIAA, Orlando, 2003), paper 2003–3847.
[22] M. Barrault, N. C. Nguyen, Y. Maday, and A. T. Patera, C. R. Acad. Sci., Ser. I: Math. **339**, 667 (2004).
[23] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera, Math. Modell. Numer. Anal. **41**, 575 (2007).

[24] P. E. Blöchl, O. Jepsen, and O. K. Andersen, Phys. Rev. B **49**, 16223 (1994).
[25] O. Jepson and O. K. Anderson, Solid State Commun. **9**, 1763 (1971).
[26] G. Gilat and L. J. Raubenheimer, Phys. Rev. **144**, 390 (1966).
[27] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, U.K., 2004).
[28] M. L. Cohen and T. K. Bergstresser, Phys. Rev. **141**, 789 (1966).
[29] A. Baldereschi, Phys. Rev. B **7**, 5212 (1973).
[30] A. D. Izaak, Math. Notes **55**, 30 (1994).
[31] S. N. Kudryavtsev, Math. Notes **77**, 494 (2005).
[32] A. Pinkus, *Widths in Approximation Theory* (Springer, Berlin, 1985).
[33] Y. Maday, A. Patera, and G. Turinici, J. Sci. Comput. **17**, 437 (2002).
[34] Y. Maday, in *International Congress of Mathematicians* (Eur. Math. Soc., Zürich, 2006), Vol. III, pp. 1255–1270.
[35] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods in Fluid Dynamics* (Springer, New York, 1987).
[36] G. S. H. Pau, Ph.D. thesis, Massachusetts Institute of Technology, 2007.
[37] G. Strang and G. J. Fix, *An Analysis of the Finite Element Method* (Prentice-Hall, Englewood Cliffs, NJ, 1973).
[38] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide* (SIAM, Philadelphia, 2000).
[39] E. Isaacson and H. B. Keller, *Computation of Eigenvalues and Eigenvectors, Analysis of Numerical Methods* (Dover, New York, 1994).
[40] B. N. Parlett, *The Symmetric Eigenvalue Problem* (SIAM, Philadelphia, 1998).
[41] M. Ainsworth and J. T. Oden, *A Posteriori Error Estimation in Finite Analysis* (Wiley, New York, 2000).
[42] A. K. Noor and I. Babuška, Finite Elem. Anal. Design **3**, 1 (1987).