Preprint ANL/MCS 1303–1105

Mathematical Programming manuscript No.
(will be inserted by the editor)

Mihai Anitescu[1], Dan Negrut[1], Peter Zapol[2], and Anter El-Azab[3]

A Note on the Regularity of Reduced Models Obtained by Nonlocal Quasi-continuum-like Approaches

May 3, 2007

Abstract. The paper investigates model reduction techniques that are based on a nonlocal quasi-continuum-like approach. These techniques reduce a large optimization problem to either a system of nonlinear equations or another optimization problem that are expressed in a smaller number of degrees of freedom. The reduction is based on the observation that many of the components of the solution of the original optimization problem are well approximated by certain interpolation operators with respect to a restricted set of representative components.

Under certain assumptions, the "optimize and interpolate" and the "interpolate and optimize" approaches result in a regular nonlinear equation and an optimization problem whose solutions are close to the solution of the original problem, respectively. The validity of these assumptions is investigated by using examples from potential-based and electronic structure-based calculations in Materials Science models. A methodology is presented for using quasi-continuum-like model reduction for real-space DFT computations in the absence of periodic boundary conditions. The methodology is illustrated using a one-dimensional basic Thomas-Fermi-Dirac case study.

1. Introduction

This work investigates the optimization problems and nonlinear equations problem that appear in modern computational Materials Science as a result of applying quasi-continuum-like model reduction techniques. The original, full-resolution problems are optimization problems in their state variables (such as the atomic positions or distribution of electron density), in which an energy is minimized with respect to these variables and, sometimes, the constraints (such as boundary conditions or total electron density constraints).

The quasi-continuum approach [23,17] is a model reduction technique of increasing popularity in the computational materials science community. In the nonlocal form investigated here, the method is based on the observation that at the solution of the full-resolution problem many of the state variables can be well approximated by interpolation of a much smaller set of state variables called *representative variables*. In Materials Science, the state variables are the positions of nuclei and, sometimes, values of the electronic density. For example, for the simulation of the response of a crystal described by potentials to a nanoindenter, the full-resolution problem consists of minimizing the total energy of the system, which is the sum of pairwise atomic potentials

Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, e-mail: {anitescu,negrut}@mcs.anl.gov

Materials Science and Chemistry Divisions, Argonne National Laboratory, Argonne, IL 60439, e-mail: zapol@anl.gov

Materials Theory Group, College of Engineering, Florida State University, Tallahassee, FL 32310, e-mail: anter@eng.fsu.edu

[12], whereas the representative variables are the positions of atoms that are nodes in a mesh whose size is at the scale of the system to be simulated (the macro scale). The local quasi-continuum method was recently extended to include electronic density as a state variable [6], and nonzero temperature [21]. In the study of nanoindentation of Au, the quasi-continuum approach has resulted in a reduction from 2.5×10^{11} atomic positions to 25237 atomic positions, while achieving a reasonable accuracy [13].

This work investigates the regularity of the reduced problems generated by a quasi-continuum-like approach, regarded here as a reduction based on a fixed linear operator (interpolation operator). We note that other analytical results exist in the issue of regularity of quasi-continuum-like approaches applied to materials science problems. Such results include well-posedness and numerical analysis of quasi-continuum methods applied to one dimensional problems involving inter-atomic Lennard Jones potentials [16, 20], a study of cluster summation rules used in potential based quasi continuum methods [5], and the interaction between continuum and atomistic models as well as the accuracy of the continuum limit of both potential based approaches and density functional theory like approaches [4]. While our approach also analyzes the well-posedness problems opening by of quasi-continuum-like reduction, it applies to both potential-based and density-functional-theory- based approaches irrespective of dimension and includes the case with constraints (while at the same time providing results that are weaker than for the more restrictive case described above).

The paper is organized as follows. Section 2 describes the abstract framework for both the full-resolution problem and the two reduction techniques: the "optimize and interpolate" version that leads to a nonlinear equation, and the "interpolate and optimize" version that results in an optimization problem. The assumptions needed for regularity of the reduced problems are stated in Section 3, followed by an analysis of the two techniques. Section 4 presents two numerical experiments used for evaluating the validity of the assumptions made for the analytical analysis of the reduction techniques. The section concludes with a description of the nonlocal quasi-continuum approach extended for density functional theory (DFT) calculations.

Notation If $u_1, u_2, \ldots, u_q$ are column vectors, $(u_1; u_2; \ldots; u_q)$ denotes the column vector obtain by adjoining all the vectors. The full-resolution state vector is denoted by $x = (x_1; x_2) \in I\!R^n$, where $x_1 \in I\!R^m$ is the set of representative states. For a matrix N, we denote by $\sigma_m(N)$ its minimum singular value and by $||N||$ its norm. Hereafter, if $f = f(x_1, x_2)$ and T is the interpolation operator used in this work, applying the chain rule yields

$$\frac{df(x_1, Tx_1)}{dx_1} = \nabla_{x_1} f(x_1, Tx_1) + \nabla_{x_2} f(x_1, Tx_1)\, T.$$

In addition, we denote by $M(1)$ a matrix-valued function of several unspecified variables that satisfies $||M(1)|| \leq 1$ for any value of its variables.

2. Formulation of the Reduced Problems

Consider the optimization problem

$$\text{(O)} \qquad \begin{array}{rll} \min_{x_1,x_2} & f(x_1,x_2) & \\ & g_1(x_1) & = 0 \\ \text{s.t.} & g_2(x_2) & = 0 \\ & g_3(x_1,x_2) & = 0\ . \end{array}$$

The functions $g_1(x_1) : I\!R^m \to I\!R^{q_1}$, $g_2(x_2) : I\!R^{n-m} \to I\!R^{q_2}$ and $g(x_1,x_2) : I\!R^n \to I\!R^{q_3}$ are the constraint functions, which, together with the objective function $f(x_1,x_2) : I\!R^n \to I\!R$, are twice continuously differentiable.

In the original application of the quasi-continuum method [23], x_1 were positions of representative atoms that were nodes of a mesh on a scale much larger than the interatomic distance, whereas x_2 were the rest of the atomic positions. An example of one-dimensional application of the nonlocal quasi-continuum approach is provided in Section 4.1.

Using the notation $\lambda = (\lambda_1;\lambda_2;\lambda_3)$, $x = (x_1,x_2)$, one can define the regularity of the solution of the original problem in terms of the Lagrangian function

$$L(x,\lambda) = f(x_1,x_2) + \langle g_1(x_1),\lambda_1\rangle + \langle g_2(x_2),\lambda_2\rangle + \langle g_3(x_1,x_2),\lambda_3\rangle\,. \qquad (1)$$

Herein, the definition of regularity of the solution of the original problem is composed of the constraint qualification and the second-order sufficient conditions from classical nonlinear optimization theory [8, Lemma 9.2.2, Theorem 9.3.2].

Regularity Assumption: The following conditions hold at the solution (x^*,λ^*) of the problem (O):

1. Constraint Qualification Condition (CQC): The rows of the matrices $\nabla_x g_1(x_1)$, $\nabla_x g_2(x_2)$ and $\nabla_x g_3(x_1,x_2)$ are linearly independent. We denote by σ_g the minimum singular value of the Jacobian of the constraints.
2. Second-Order Sufficient Condition (SOSC): With the notation $\nabla_x g_1(x_1) = [\nabla_{x_1} g_1(x_1),\ 0_{q_1\times n-m}]$, $\nabla_x g_2(x_2) = [0_{q_2\times m},\ \nabla_{x_2} g_2(x_2)]$, the Hessian of the Lagrangian function satisfies

$$\left.\begin{array}{r} \nabla_x g_1(x_1^*)\Delta x = 0, \\ \nabla_x g_2(x_2^*)\Delta x = 0, \\ \nabla_x g_3(x_1^*,x_2^*)\Delta x = 0, \\ \Delta x \neq 0 \end{array}\right\} \Rightarrow \Delta x^T \nabla^2_{xx} L(x^*,\lambda^*)\Delta x \geq \sigma_L\,||\Delta x||^2 > 0.$$

Hereafter, the CQC or SOSC will be invoked for optimization problems other than (O) with the understanding that for the respective cases they convey the same meaning.

The key observation of the quasi-continuum approach [23, 12] is that at the solution of the problem (O) the position of the nonrepresentative degrees of freedom can be approximated by an interpolation operator, namely the linear interpolation operator with nodes at the representative atoms. This observation is formalized in the following.

Interpolation Assumption: At the optimal solution (x_1^*,x_2^*) of the problem (O), we have that

$$x_2^* \approx T(x_1^*)$$

where T is a linear operator identified with its matrix form $T(x_1) = Tx_1$.

This assumption is qualitative in nature and cannot be used for analytical estimates, although it describes quite well the nature of the approximation regime that we will be working in. When presenting our proofs, we will mention this assumption only as a mean to alert the reader that an approximation of the type $x_2^* = Tx_1^* + o(1)$ is involved with an error level to be specified in that context.

Introducing the Lagrange multipliers $\lambda_1 \in I\!R^{q_1}$, $\lambda_2 \in I\!R^{q_2}$ and $\lambda_3 \in I\!R^{q_3}$, applying the optimality conditions to the problem (O), and using the notation $\langle a, b\rangle = a^T b$, one obtains

$$\begin{aligned} \nabla_{x_1} f(x_1^*, x_2^*) + \nabla_{x_1} \langle g_3(x_1^*, x_2^*), \lambda_3\rangle + \nabla_{x_1} \langle g_1(x_1^*), \lambda_1\rangle &= 0 \\ \nabla_{x_2} f(x_1^*, x_2^*) + \nabla_{x_2} \langle g_3(x_1^*, x_2^*), \lambda_3\rangle + \nabla_{x_2} \langle g_2(x_2^*), \lambda_2\rangle &= 0 \\ g_1(x_1^*) &= 0 \\ g_2(x_2^*) &= 0 \\ g_3(x_1^*, x_2^*) &= 0\ . \end{aligned} \tag{2}$$

The interpolation assumption suggests two ways of creating a reduced problem. The "optimize and interpolate" (or "optimize and reduce") approach, in which one substitutes $x_2 = T(x_1)$ in the optimality conditions of (2), leads to the following reduced system of nonlinear equations:

$$\text{(RE)} \qquad \begin{aligned} \nabla_{x_1} f(x_1, Tx_1) + \nabla_{x_1} \langle g_3(x_1, Tx_1), \lambda_3\rangle + \nabla_{x_1} \langle g_1(x_1, Tx_1), \lambda_1\rangle &= 0 \\ g_1(x_1) &= 0 \\ g_3(x_1, Tx_1) &= 0. \end{aligned}$$

In the second approach, referred to as the "interpolate and optimize" (or "reduce and optimize") approach, one substitutes $x_2 = T(x_1)$ in the problem (O), resulting in the following optimization problem:

$$\text{(RO)} \qquad \begin{array}{ll} \min_{x_1} & f(x_1, Tx_1) \\ \text{s.t.} & g_1(x_1) = 0 \\ & g_3(x_1, Tx_1) = 0. \end{array}$$

Clearly, (RE) does not represent the optimality conditions of (RO) because it makes no direct reference to $\nabla_{x_2} f(x_1, x_2)$, which does appear if one writes the optimality conditions of (RO). In the application of the quasi-continuum methodology to the minimization of energies computed through pairwise potentials, the "optimize and interpolate" approach corresponds to the force-based quasi-continuum approach [12, 17], whereas the "interpolate and optimize" approach corresponds to the energy-based quasi-continuum approach [23, 17], except for the fact that in the respective references, further transformations are carried out to approximate the data of the problems (RE) and (RO), for reasons that will be discussed in Section 3.3.

3. Analysis of the Reduced Problems

The goal of this section is to explore under what circumstances the reduced problems (RE) and (RO) are regular in a neighborhood of a regular solution of the original problem (O).

3.1. The Optimize and Interpolate Case: the Reduced Nonlinear Equation

The regularity of the reduced system of nonlinear equations (RE) requires two additional assumptions.

RE Constraint Form Assumption (RECF): The constraints of the problem (O) are separable; that is, $g_3 = \emptyset$. Likewise, the constraints $g_2(x_2) = 0$ are linear and satisfy

$$g_1(x_1) = 0 \Rightarrow g_2(Tx_1) = 0.$$

The second part of the assumption explains why it is possible to completely remove the constraints on the nonrepresentative variables from the reduced problems (RE) and (RO). Indeed, in most applications of the quasi-continuum approach, the boundary conditions result in constraints that satisfy the above assumption. For example, in the Au nanoindentation application [12], the bottom layer of atoms of a cubic-shaped crystal is fixed whereas its sides move only in the z direction. If representative atoms are located at the corners of the cube and the operator T are generated from linear interpolation with nodes at the representative atoms, then all of the constraints described in the preceding sentence satisfy the *RE constraint form assumption*. In practical terms, an interpolation operator that results in good approximation properties is bound to satisfy the second requirement in the assumption, since the degrees of freedom that could help enforce $g_2(x_2) = 0$ have disappeared in the reduced problem.

The second assumption plays a central role in proving the regularity results, and it relates the Hessian matrix and the interpolation operator T.

H-T Assumption: The Hessian of the Lagrange function satisfies

$$\nabla^2_{x_2x_2} L(x^*, \lambda^*)T + \nabla^2_{x_2x_1} L(x^*, \lambda^*) \approx 0$$

This assumption has a similar purpose as the interpolation assumption, it serves to alert the reader that

$$\nabla^2_{x_2x_2} L(x^*, \lambda^*)T + \nabla^2_{x_2x_1} L(x^*, \lambda^*) = o(1)$$

with a level of error in the approximation to be specified in the text surrounding the place where is the approximation will be used.

The main results of our work are presented in Theorems 1 and 2. The assumptions of the theorems include the satisfaction of one or both ofthe Interpolation Assumption and the H-T Assumption at sufficiently high levels of accuracy. The level of accuracy is characterized by the respective results and is a function of expressions depending on certain characteristic of the data of the problem. Note that the level of accuracy to which Interpolation Assumption and H-T Assumption are satisfied are a characteristic of the problem and not controllable by the user, and for a multiscale problem they may be an expression of how close such systems are to an appropriate (though perhaps hard to compute, which justifies the use of a quasi continuum-approach) continuum limit.

We also define the following quantity

$$\Gamma = \sup_{i=0,1,2,3,\left|\left|x_1-x_1^*\right|\right|\leq\epsilon_b,\left|\left|x_2-x_2^*\right|\right|\leq\epsilon_b} \left\{\max\left\{\left|\left|\nabla^i_{x^i} f(x_1,x_2)\right|\right|, \left|\left|\nabla^i_{x^i} g(x_1,x_2)\right|\right|\right\}\right\} \tag{3}$$

where ϵ_b is a fixed parameter.

Theorem 1. *If the regularity assumption and RE constraint form assumption hold at the solution $(x_1^*; x_2^*; \lambda_1^*; \lambda_2^*; \lambda_3^*)$, of (O), then if the interpolation assumption and H-T assumption are satisfied with sufficiently high accuracy, that is*

$$||x_2^* - Tx_1^*|| \leq \epsilon_0, \quad \left|\left|\nabla^2_{x_2x_2} L(x^*, \lambda^*)T + \nabla^2_{x_2x_1} L(x^*, \lambda^*)\right|\right| \leq \epsilon_0$$

where $\epsilon_0 \leq \frac{1}{\theta_0(\Gamma,T,\sigma_g,\sigma_L)}$, at (x_1^*, x_2^*), then the problem (RE) has a nonsingular Jacobian at $(x_1^*, \lambda_1^*, \lambda_3^*)$. If, in addition, we have that

$$\epsilon_0 \leq \frac{1}{2\theta_1(\Gamma, T, \sigma_g, \sigma_L)\theta_2(\Gamma, T, \sigma_g, \sigma_L)}$$

then (RE) has a unique solution in a neighborhood of the same point $(x_1^*, \lambda_1^*, \lambda_3^*)$. The size of the neighborhood is at most

$$\frac{1 - \sqrt{1 - 2\theta_1(\Gamma, T, \sigma_g, \sigma_L)\theta_2(\Gamma, T, \sigma_g, \sigma_L)\epsilon_0}}{\theta_1(\Gamma, T, \sigma_g, \sigma_L)}.$$

Here θ_0, θ_1, and θ_2 are nonnegative functions that depend only on Γ, T, σ_g, σ_L. Recall that σ_g, σ_L are quantities introduced at the regularity assumption .

A set of lemmas will be used in proving this result. The first lemma is essential in the study of augmented Lagrangians and is stated here (as well as its reciprocal) for completeness.

Lemma 1. *Let P and Q be symmetric $n \times n$ matrices, and assume that Q is positive semidefinite. Then there exists a scalar c such that $P + cQ$ is positive definite if and only if $x^T Px > 0$ whenever $x \neq 0$ and $x^T Qx = 0$. In addition, if σ_P is such that*

$$x \neq 0,\ x^T Qx = 0,\ \Rightarrow x^T Px \geq \sigma_P \left|\left|x\right|\right|^2,$$

then there exists a $c = c(||P||, ||Q||, \sigma_m(Q), \sigma_P)$ such that $\sigma_m(P + cQ) \geq \frac{\sigma_P}{2}$.

Proof. If $x^T Px > 0$ whenever $x \neq 0$ and $x^T Qx = 0$, then there exists a c such that $P + cQ$ is positive definite [3, Lemma 1.25]. The reciprocal is obvious. The second part of the proof follows the same way (for example applying the first part to the matrix $P - \frac{\sigma_P}{2} I$). □

Lemma 2. *Assume that the functions $g_1(x_1)$ and $g_2(x_2)$ are such that the following hold.*

1. *The Jacobian of the function $g_1(x_1)$ is full row rank.*
2. *The following relationship holds, $\forall x_1$:*

$$g_1(x_1) = 0 \Rightarrow g_2(Tx_1) = 0.$$

If Δx_1 is such that $\nabla_{x_1} g_1(x_1)\, \Delta x_1 = 0$, then for all $\lambda_2 \in \mathbb{R}^{q_2}$,

(i) $\nabla_{x1}(g_2(Tx_1))\, \Delta x_1 = \nabla_{x_2} g_2(Tx_1)\, T\Delta x_1 = 0$,

(ii) The following identity holds:

$$\nabla_{x_2} g_2(Tx_1)\, T = S(x_1) \nabla_{x_1} g(x_1),$$

where $S(x_1)$ is the differentiable matrix

$$S(x_1) = \nabla_{x_2} g_2(Tx_1)\, T\, \nabla_{x_1} g_1(x_1)^T \left(\nabla g_1(x_1)\, \nabla g_1(x_1)^T\right)^{-1}.$$

(iii) The following identity holds:

$$(T\Delta x_1)^T\, \nabla^2_{x_2 x_2} \left\langle g_2(Tx_1), \lambda_2 \right\rangle T\Delta x_1 = \Delta x_1 \nabla^2_{x_1 x_1} \left\langle g_1(x_1), S(x_1)^T \lambda_2 \right\rangle \Delta x_1,$$

where the entries of $S(x_1)$ are not differentiated in the last equation.

Proof. Consider an arc $x_1(t)$ that satisfies

$$g_1(x_1(t)) = 0, \forall\, t > 0 \text{ and } x_1(0) = x_1;\ \left.\frac{dx_1(t)}{dt}\right|_{t=0} = \Delta x_1. \tag{4}$$

Such an arc exists from the first assumption of the hypothesis. Then, from the second assumption,

$$\left.\frac{dg_2(Tx_1(t))}{dt}\right|_{t=0} = 0.$$

Using the definition of the arc $x_1(t)$ leads to

$$\nabla_{x_2} g_2(Tx_1) T\Delta x_1 = 0,$$

which proves (i).

From (i),

$$\nabla_{x_1} g_1(x_1)\Delta x_1 = 0 \Rightarrow \nabla_{x_2} g_2(Tx_1)\, T\Delta x_1 = 0,$$

and it follows, from Farkas' lemma [8, Lemma 9.2.4] and the subsequent Lagrange multiplier theory of constrained optimization applied to each row of $\nabla_{x_2} g_2(Tx_1)\, T$, that there exists a matrix $S(x_1)$ such that

$$\nabla_{x_2} g_2(Tx_1)\, T = S(x_1) \nabla_{x_1} g_1(x_1).$$

Since this displayed equation implies that the rows of $\nabla_{x_2} g_2(Tx_1)\, T$ are orthogonal to the kernel subspace of $\nabla_{x_1} g_1(x_1)$, it follows that $\nabla_{x_2} g_2(Tx_1)\, T$ coincides with its orthogonal projection on the space orthogonal to the same kernel subspace; that is,

$$\nabla_{x_2} g_2(Tx_1)\, T \left[I_{q_1} - \nabla_{x_1} g(x_1)^T \left(\nabla_{x_1} g(x_1)\, \nabla_{x_1} g(x_1)^T\right)^{-1} \nabla_{x_1} g(x_1)\right] = 0.$$

Herein, I_s is the identity matrix of dimension s. Expanding the left side of the displayed equation leads to conclusion (ii).

Consider again the arc (4) for which

$$\left.\frac{d^2 g_1(x_1(t))}{dt^2}\right|_{t=0} = 0 \text{ and } \left.\frac{d^2 \left\langle g_2(Tx_1(t)), \lambda_2 \right\rangle}{dt^2}\right|_{t=0} = 0.$$

Expanding these second time derivatives yields, for $i = 1, 2, \ldots, q_1$,

$$\nabla_{x_1} g_1^i(x_1)\ddot{x}_1(0) + \Delta x_1^T \nabla^2_{x_1 x_1} g_1^i(x_1)\, \Delta x_1 = 0 \qquad (5)$$

$$\nabla_{x_2} \langle g_2(Tx_1), \lambda_2\rangle T\ddot{x}_1(0) + (T\Delta x_1)^T \nabla_{x_2 x_2} \langle g_2(Tx_1), \lambda_2\rangle T\Delta x_1 = 0. \qquad (6)$$

Based on (ii), the first term in (6) can be expressed as

$$\nabla_{x_2} \langle g_2(Tx_1), \lambda_2\rangle T\ddot{x}_1(0) = \langle \nabla_{x_2} g_2(Tx_1)T, \lambda_2\rangle \ddot{x}_1(0) =$$
$$\langle S(x_1)\nabla_{x_1} g_1(x_1), \lambda_2\rangle \ddot{x}_1(0) = \left\langle \nabla_{x_1} g_1(x_1), S(x_1)^T\lambda_2\right\rangle \ddot{x}_1(0).$$

Multiplying each of the equations of (5) with the corresponding component of $S(x_1)^T\lambda_2$ and summing them, one obtains

$$\begin{aligned}\nabla_{x_2} \langle g_2(Tx_1), \lambda_2\rangle T\ddot{x}_1(0) &= \left\langle \nabla_{x_1} g_1(x_1), S(x_1)^T\lambda_2\right\rangle \ddot{x}_1(0) \\ &= -\Delta x_1^T \nabla^2_{x_1 x_1} \left\langle g_1(x_1), S(x_1)^T\lambda_2\right\rangle \Delta x_1,\end{aligned}$$

where the $\nabla_{x_1 x_1}$ operator does not act on $S(x_1)$. Conclusion (iii) is proved by replacing the left term from the last displayed equation with the right term in (6). □

Lemma 3. *Define the Lagrangian of the problem (O) that excludes the constraint* $g_2(x_2) = 0$*,*

$$\widehat{L}(x, \lambda) = f(x_1, x_2) + \langle g_1(x_1), \lambda_1\rangle + \langle g_3(x_1, x_2), \lambda_3\rangle . \qquad (7)$$

Define the matrix

$$J_O = \begin{bmatrix} \nabla_{x_1} g_1(x_1^*) \\ \nabla_{x_1} g_3(x_1^*, x_2^*) + \nabla_{x_2} g_3(x_1^*, x_2^*)T \end{bmatrix},$$

and assume that $\forall x_1$, $g_1(x_1) = 0 \Rightarrow g_2(Tx_1) = 0$*. Then, there exists a function* $\theta_3(\Gamma, T, \sigma_g, \sigma_L)$ *such that, if the regularity assumption holds, and interpolation assumption is sufficiently accurate, specifically,*

$$||x_2^* - Tx_1^*|| \leq \frac{1}{\theta_3(\Gamma, T, \sigma_g, \sigma_L)},$$

then one has the following.

(i) If g_2 *is a linear function, the matrix*

$$\widehat{L}_T = \begin{bmatrix} I_m \\ T \end{bmatrix}^T \nabla^2_{xx} \widehat{L}(x^*, \lambda^*) \begin{bmatrix} I_m \\ T \end{bmatrix}$$

is positive definite over the set

$$\mathcal{F} = \{\Delta x_1 | J_O \Delta x_1 = 0\},$$

and satisfies

$$\Delta x_1 \in \mathcal{F} \Rightarrow \Delta x_1 \widehat{L}_T \Delta x_1 \geq \frac{\sigma_L}{4} ||\Delta x_1||^2 .$$

(ii) If $\widetilde{\lambda} = \left(\lambda_1^* + S(x_1^*)^T\lambda_2^*, 0, \lambda_3^*\right)$*, the matrix*

$$\widetilde{L}_T = \begin{bmatrix} I_m \\ T \end{bmatrix}^T \nabla^2_{xx}\widehat{L}(x^*, \widetilde{\lambda}) \begin{bmatrix} I_m \\ T \end{bmatrix}$$

is positive definite over the set

$$\mathcal{F} = \{\Delta x_1 | J_O \Delta x_1 = 0\},$$

and satisfies

$$\Delta x_1 \in \mathcal{F} \Rightarrow \Delta x_1 \widetilde{L}_T \Delta x_1 \geq \frac{\sigma_L}{4} ||\Delta x_1||^2.$$

Proof. Consider the symmetric positive semidefinite matrix

$$\begin{aligned} Q = &[I_m; 0]^T \nabla_{x_1} g_1(x_1^*)^T \nabla_{x_1} g_1(x_1^*) [I_m; 0] \\ &+ [0; I_{n-m}]^T \nabla_{x_2} g_2(x_2^*)^T \nabla_{x_2} g_2(x_2^*) [0; I_{n-m}] + \nabla_x g_3(x^*)^T \nabla_x g_3(x^*). \end{aligned}$$

Since the regularity assumption holds, it follows from Lemma 1 with $P = \nabla^2_{xx} L(x^*, \lambda^*)$ that there exists a finite $c = c(\sigma_g, \sigma_L, T, \Gamma) > 0$ such that

$$\begin{aligned} L_c = &\nabla^2_{xx} L(x^*, \lambda^*) + c [I_m; 0]^T \nabla_{x_1} g_1(x_1^*)^T \nabla_{x_1} g_1(x_1^*) [I_m; 0] \\ &+ c [0; I_{n-m}]^T \nabla_{x_2} g_2(x_2^*)^T \nabla_{x_2} g_2(x_2^*) [0; I_{n-m}] + c \nabla_x g_3(x^*)^T \nabla_x g_3(x^*) \end{aligned}$$

satisfies $\sigma_m(L_c) \geq \frac{\sigma_L}{2}$. It is immediate that, the matrix

$$L_{c,T} = \begin{bmatrix} I_m \\ T \end{bmatrix}^T L_c \begin{bmatrix} I_m \\ T \end{bmatrix}$$

also satisfies $\sigma_m(L_c) \geq \frac{\sigma_L}{2}$. Considering the definition of the Lagrangian $\widehat{L}$ and of the matrix L_c, one has that

$$\begin{aligned} L_{c,T} = &\widehat{L}_T + c \nabla_{x_1} g_1(x_1^*)^T \nabla_{x_1} g_1(x_1^*) + c \begin{bmatrix} I_m \\ T \end{bmatrix}^T \nabla_x g_3^T(x^*) \nabla_x g_3(x^*) \begin{bmatrix} I_m \\ T \end{bmatrix} \\ &+ c T^T \nabla_{x_2} g_2(x_2^*)^T \nabla_{x_2} g_2(x_2^*) T + T^T \nabla^2_{x_2 x_2} \langle g(x_2^*), \lambda_2^* \rangle T. \end{aligned} \quad (8)$$

Define

$$U(x_1) = c T^T \nabla_{x_2} g_2(T x_1)^T \nabla_{x_2} g_2(T x_1) T.$$

Since g_2 is a linear function, and from an application of the higher-dimensional "mean value theorem" [19, Prop. 3.2.3] it follows from the interpolation assumption that the last two terms of (8) satisfy

$$\begin{aligned} &c T^T \nabla_{x_2} g_2(x_2^*)^T \nabla_{x_2} g_2(x_2^*) T + T^T \nabla^2_{x_2 x_2} \langle g(x_2^*), \lambda_2^* \rangle T \\ &= U(x_1) + \eta_1(\Gamma, T, \sigma_g, \sigma_L) M(1) ||x_2^* - T x_1^*||, \end{aligned}$$

for some expression $\eta_1(\cdot) \geq 0$.

Since $\sigma_m(L_{c,T}) \geq \frac{\sigma_L}{2}$, it follows that, whenever

$$||x_2^* - T x_1^*|| \leq \frac{\sigma_L}{4\eta_1(\Gamma, T, \sigma_g, \sigma_L)},$$

we have that

$$\widehat{L}_T + c\nabla_{x_1} g_1(x_1^*)^T \nabla_{x_1} g_1(x_1^*) + c \begin{bmatrix} I_m \\ T \end{bmatrix}^T \nabla_x g_3^T(x^*) \nabla_x g_3(x^*) \begin{bmatrix} I_m \\ T \end{bmatrix} + U(x_1^*),$$

has a minimum eigenvalue that exceeds $\frac{\sigma_L}{4}$. In turn, from Lemma 1, we obtain that the matrix $\widehat{L}_T + U(x_1^*)$ is positive definite over the set $\mathcal{F}$. If $\Delta x_1 \in \mathcal{F}$, then $\nabla_{x_1} g(x_1^*)\Delta x_1 = 0$; and based on Lemma (2)(i), $\Delta x_1^T U(x_1^*)\Delta x_1 = 0$, which completes the proof for (i), with the choice $\theta_3(\cdot) = \eta_1(\cdot)$.

For part (ii), the equivalent of (8) is

$$\begin{aligned} L_{c,T} = \widetilde{L}_T &+ c\, \nabla_{x_1} g_1(x_1^*)^T\, \nabla_{x_1} g_1(x_1^*) + c \begin{bmatrix} I_m \\ T \end{bmatrix}^T \nabla_x g_3(x^*)^T\, \nabla_x g_3(x^*) \begin{bmatrix} I_m \\ T \end{bmatrix} \\ &+ c\, T^T\, \nabla_{x_2} g_2(x_2^*)^T\, \nabla_{x_2} g_2(x_2^*)\, T + T^T\, \nabla^2_{x_2 x_2} \left\langle g_2(x_2^*), \lambda_2^* \right\rangle T \\ &- \nabla^2_{x_1 x_1} \left\langle g_1(x_1^*), S(x_1^*)^T \lambda_2^* \right\rangle, \end{aligned} \tag{9}$$

where, again, the entries of $S(\cdot)$ are not differentiated. Similar to the previous case, we have that $\sigma_m(L_{c,T}) \geq \frac{\sigma_L}{2}$. Based on the interpolation assumption , and the higher dimensional "mean value theorem" [19, Prop. 3.2.3], we obtain that

$$\begin{aligned} &cT^T \nabla_{x_2} g_2(x_2^*)^T \nabla_{x_2} g_2(x_2^*) T + T^T \nabla^2_{x_2 x_2} \left\langle g_2(x_2^*), \lambda_2^* \right\rangle T \\ &- \nabla^2_{x_1 x_1} \left\langle g_1(x_1^*), S(x_1^*)^T \lambda_2^* \right\rangle = U(x_1^*, \lambda_2^*) + \eta_2(\Gamma, T, \sigma_g, \sigma_L) M(1)\, ||x_2^* - T x_1^*||, \end{aligned} \tag{10}$$

for some expression $\eta_2(\cdot) \geq 0$, where

$$\begin{aligned} U(x_1, \lambda_2) = &cT^T \nabla_{x_2} g_2(Tx_1)^T \nabla_{x_2} g_2(x_1) T + T^T \nabla^2_{x_2 x_2} \left\langle g_2(Tx_1), \lambda_2 \right\rangle T \\ &- \nabla^2_{x_1 x_1} \left\langle g_1(x_1), S(x_1)^T \lambda_2 \right\rangle. \end{aligned}$$

From equations (10) and (9) it follows that, for

$$||x_2^* - T x_1^*|| \leq \frac{\sigma_L}{4\eta_2(\Gamma, T, \sigma_g, \sigma_L)},$$

the matrix

$$\widetilde{L}_T^c = \widetilde{L}_T + c\nabla_{x_1} g_1(x_1^*)^T \nabla_{x_1} g_1(x_1^*) + c \begin{bmatrix} I \\ T \end{bmatrix}^T \nabla_x g_3^T(x^*) \nabla_x g_3(x^*) \begin{bmatrix} I \\ T \end{bmatrix} + U(x_1^*, \lambda_2^*)$$

satisfies $\sigma_m(\widetilde{L}_T^c) \geq \frac{\sigma_L}{4}$. From Lemma 1, with the matrix Q given by

$$Q = \nabla_{x_1} g_1(x_1^*)^T \nabla_{x_1} g_1(x_1^*) + \begin{bmatrix} I \\ T \end{bmatrix}^T \nabla_x g_3^T(x^*) \nabla_x g_3(x^*) \begin{bmatrix} I \\ T \end{bmatrix}$$

it follows that the matrix $\widetilde{L}_T + U(x_1^*, \lambda_2^*)$ is positive definite over the set $\mathcal{F}$ and its associated quadratic form is bounded below by $\frac{\sigma_L}{4}$ over the same set. But for any $\Delta x_1 \in \mathcal{F}$, $\nabla_{x_1} g(x_1)\Delta x_1 = 0$ and, based on Lemma 2 (i) and (iii), $\Delta x_1^T U(x_1^*, \lambda_2^*)\Delta x_1 = 0$, which in turn implies that $\widetilde{L}_T$ is positive definite over the set $\mathcal{F}$. The conclusion follows after taking $\theta_3(\cdot) = \eta_2(\cdot)$. □

All the intermediary results needed to prove the main theorem are now available.

Proof of Theorem 1

The Jacobian of (RE) at $(x_1^*, Tx_1^*, \lambda^*)$ is

$$J^{RE} = \begin{bmatrix} \nabla^2_{x_1x_1}\widehat{L}(x_1^*, Tx_1^*, \lambda^*) + \nabla^2_{x_1x_2}\widehat{L}(x_1^*, Tx_1^*, \lambda^*)T & \nabla_{x_1} g_1(x_1^*)^T \\ \nabla_{x_1} g_1(x_1^*) & 0 \end{bmatrix}.$$

For the upper left corner of the Jacobian, by virtue of the interpolation assumption and using the higher dimensional "mean value theorem" [19, Prop. 3.2.3], we obtain that

$$\begin{aligned} J^{RE}_{11} &= \nabla^2_{x_1x_1}\widehat{L}(x_1^*, Tx_1^*, \lambda^*) + \nabla^2_{x_1x_2}\widehat{L}(x_1^*, Tx_1^*, \lambda^*)T \\ &= \nabla^2_{x_1x_1}\widehat{L}(x_1^*, x_2^*, \lambda^*) + \nabla^2_{x_1x_2}\widehat{L}(x_1^*, x_2^*, \lambda^*)T + \eta_1(\Gamma, T, \sigma_g, \sigma_L)\,||x_2^* - Tx_1^*||\,M(1) \end{aligned}$$

Using the definition of $\widehat{L}$, invoking the H-T assumption and using the higher dimensional "mean value theorem" [19, Prop. 3.2.3], we obtain that

$$\begin{aligned} &\left|\left|\nabla^2_{x_2x_2}L(x^*, \lambda^*)T + \nabla^2_{x_2x_1}L(x^*, \lambda^*)\right|\right| ||T||\,M(1) \\ &= T^T\,\nabla^2_{x_2x_2}L(x^*, \lambda^*)\,T + T^T\,\nabla^2_{x_2x_1}L(x^*, \lambda^*) \\ &= T^T\,\nabla^2_{x_2x_2}\widehat{L}(x^*, \lambda^*)\,T + T^T\,\nabla^2_{x_2x_1}\widehat{L}(x^*, \lambda^*) + T^T\,\nabla^2_{x_2x_2}\left\langle g_2(x_2^*), \lambda_2^*\right\rangle\,T \\ &= T^T\nabla^2_{x_2x_2}\widehat{L}(x^*, \lambda^*)\,T + T^T\,\nabla^2_{x_2x_1}\widehat{L}(x^*, \lambda^*), \end{aligned}$$

for some expression $\eta_1(\cdot) \geq 0$, where the last step follows from the assumption that $g(x_2)$ is linear. Combining the last two displayed equations,

$$\begin{aligned} J^{RE}_{11} &= \begin{bmatrix} I \\ T \end{bmatrix}^T \begin{bmatrix} \nabla^2_{x_1x_1}\widehat{L}(x^*, \lambda^*) & \nabla^2_{x_1x_2}\widehat{L}(x^*, \lambda^*) \\ \nabla^2_{x_2x_1}\widehat{L}(x^*, \lambda^*) & \nabla^2_{x_2x_2}\widehat{L}(x^*, \lambda^*) \end{bmatrix} \begin{bmatrix} I \\ T \end{bmatrix} \\ &\quad + M(1)\eta_2(\Gamma, T, \sigma_g, \sigma_L)\max\{||x_2^* - Tx_1^*||\,, \left|\left|\nabla^2_{x_2x_2}L(x^*, \lambda^*)T + \nabla^2_{x_2x_1}L(x^*, \lambda^*)\right|\right|\} \\ &= \begin{bmatrix} I \\ T \end{bmatrix}^T \nabla^2_{xx}\widehat{L}(x^*, \lambda^*) \begin{bmatrix} I \\ T \end{bmatrix} \\ &\quad + M(1)\eta_2(\Gamma, T, \sigma_g, \sigma_L)\max\{||x_2^* - Tx_1^*||\,, \left|\left|\nabla^2_{x_2x_2}L(x^*, \lambda^*)T + \nabla^2_{x_2x_1}L(x^*, \lambda^*)\right|\right|\}, \end{aligned} \tag{11}$$

where

$$\eta_2(\Gamma, T, \sigma_g, \sigma_L) = \eta_1(\Gamma, T, \sigma_g, \sigma_L) + ||T||$$

From Lemma 3(i), it follows that, as soon as

$$||x_2^* - Tx_1^*|| \leq \frac{1}{\theta_3(\Gamma, T, \sigma_g, \sigma_L)},$$

the matrix

$$\widehat{L}_{MT} \begin{bmatrix} I \\ T \end{bmatrix}^T \nabla^2_{xx}\widehat{L}(x^*, \lambda^*) \begin{bmatrix} I \\ T \end{bmatrix}$$

is positive definite over the set

$$\mathcal{F}_1 = \{\Delta x_1 | \nabla_{x_1} g_1(x_1^*)\Delta x_1 = 0\}$$

and that it satisfies

$$\Delta x_1 \in \mathcal{F}_1 \Rightarrow \Delta x_1^T \widehat{L}_{MT}\Delta x_1 \geq \frac{\sigma_L}{4}\,||\Delta x_1||^2\,.$$

From equation (11) it then follows that, provided that

$$\max\left\{||x_2^* - Tx_1^*||\,, \left|\left|\nabla^2_{x_2x_2}L(x^*,\lambda^*)T + \nabla^2_{x_2x_1}L(x^*,\lambda^*)\right|\right|\right\} \leq \frac{\sigma_L}{8\eta_2(\Gamma,T,\sigma_g,\sigma_L)},$$

the matrix J^{11}_{RE} is positive definite (though not necessarily symmetric) over the set $\mathcal{F}_1$ with a reduced minimum singular value no less than $\frac{\sigma_L}{8}$. In turn, from the full rank property of $\nabla_x g_1(x_1^*)$ implied by RE constraint form assumption , this implies that the matrix

$$\begin{bmatrix} J^{11}_{RE} & \nabla_{x_1} g_1(x_1^*)^T \\ \nabla_{x_1} g_1(x_1^*) & 0 \end{bmatrix}$$

is not singular at x_1^* and has a minimum singular value bounded bellow by some positive $\eta_3(\Gamma,T,\sigma_g,\sigma_L)$, which concludes the first part of the proof, after taking

$$\theta_0(\Gamma,T,\sigma_g,\sigma_L) = \max\left\{\frac{8\eta_2(\Gamma,T,\sigma_g,\sigma_L)}{\sigma_L}, \theta_3(\Gamma,T,\sigma_g,\sigma_L)\right\}$$

For the second part of the proof, the focus shifts to the residual of the nonlinear equation (RE) at (x_1^*,λ_1^*), for $g_3 = \emptyset$. Based on the interpolation assumption and (2),

$$\nabla_{x_1} f(x_1^*, Tx_1^*) + \nabla_{x_1}\left\langle g_1(x_1^*)\lambda_1^*\right\rangle = \nabla_{x_1} f(x_1^*, x_2^*) + \nabla_{x_1}\left\langle g_1(x_1^*)\lambda_1^*\right\rangle = \eta_4(\Gamma,T,\sigma_g,\sigma_L)\,||x_2^* - Tx_1^*||\,M(1),$$
$$g(x_1^*) = 0,$$

for some expression $\eta_4(\cdot) \geq 0$. The conclusion of the second part of the proof follows from the fact that the Jacobian is not singular and from applying Kantorovich's theorem [19, Theorem 12.6.1] (RE). with the following identification (referring to the notations in that reference) $\gamma = \Gamma$, $\beta = \eta_4(\Gamma,T,\sigma_g,\sigma_L)$, $\eta = \epsilon_0\eta_4(\Gamma,T,\sigma_g,\sigma_L)$, and choosing $\theta_1 = \beta\gamma$, and $\theta_2 = \beta\gamma\eta_4(\Gamma,T,\sigma_g,\sigma_L)$. □

Theorem 1 therefore proves that the reduced nonlinear equation (RE) produced by the local quasi-continuum approach is regular, at least in the neighborhood of the solution of the original problem. As a result, local convergence of a Newton-type method to the solution of (RE) is guaranteed under the conditions of Theorem 1.

The H-T assumption assumption seems quite restrictive. Nonetheless, we present in Section 4.1 an example that satisfies it.

3.2. The Interpolate and Optimize Case: the Reduced Optimization Problem

Although (RO) and (RE) share a number of characteristics, (RE) does not represent the optimality conditions of (RO). The (RO) problem can be shown to be well posed under less restrictive assumptions.

RO Constraint Form Assumption: The constraints of the problem (O) are such that *(i)* the matrix

$$J_{RO} = \begin{bmatrix} \nabla_{x_1} g_1(x_1^*) \\ \frac{d}{dx_1} g_3(x_1^*, Tx_1^*) \end{bmatrix} = \begin{bmatrix} \nabla_{x_1} g_1(x_1^*) \\ \nabla_{x_1} g_3(x_1^*, Tx_1^*) + \nabla_{x_2} g_3(x_1^*, Tx_1^*)T \end{bmatrix}$$

has full row rank and *(ii)* the following condition holds:

$$g_1(x_1) = 0 \quad \Rightarrow \quad g_2(Tx_1) = 0, \forall x_1$$

.

Note that in the form of the matrix J_{RO} an assumption was made that both $g_1 \neq \emptyset$ and $g_3 \neq \emptyset$. If this is not the case, the constraints that are missing in the formulation are removed from the expression of J_{RO}.

Theorem 2. *If the regularity assumption and RE constraint form assumption hold at the solution $(x_1^*; x_2^*; \lambda_1^*; \lambda_2^*; \lambda_3^*)$, of (O), then if the interpolation assumption and H-T assumption are satisfied with sufficiently high accuracy, that is*

$$||x_2^* - Tx_1^*|| \leq \epsilon_0,$$

where $\epsilon_0 \leq \frac{1}{\theta_0(\Gamma,T,\sigma_g,\sigma_L)}$, at (x_1^*, x_2^*), then the problem (RO) satisfies both the SOSC and the CQC at x_1^* with multiplier $\left(\lambda_1^* + S(x_1^*)^T\lambda_2^*, \lambda_3^*\right)$. If, in addition, we have that

$$\epsilon_0 \leq \frac{1}{2\theta_1(\Gamma,T,\sigma_g,\sigma_L)\theta_2(\Gamma,T,\sigma_g,\sigma_L)}$$

then (RO) has a unique solution in a neighborhood of x_1^*.

The size of the neighborhood is at most

$$\frac{1-\sqrt{1-2\theta_1(\Gamma,T,\sigma_g,\sigma_L)\theta_2(\Gamma,T,\sigma_g,\sigma_L)\epsilon_0}}{\theta_1(\Gamma,T,\sigma_g,\sigma_L)}.$$

Here θ_0, θ_1, and θ_2 are nonnegative functions that depend only on Γ, T, σ_g, σ_L. Recall that σ_g, σ_L are quantities introduced at the regularity assumption .

Note The part involving the computation of the quantities needed for computing the size of the neighborhood with Kantorovich's Theorem is very similar to the one in the proof of Theorem 1. For brevity and clarity, we use estimates of the type $O(\epsilon_0)$ at certain parts of the proof, and we mean quantities that are bounded above by $\eta_1(\Gamma,T,\sigma_g,\sigma_L)\epsilon_0$, instead of going into the details on how expressions like $\eta_1(\cdot)$ may actually be formed.

Proof

Consider the Lagrangian of problem (O) defined in (1). The fact that the constraint qualification holds is satisfied as an immediate conclusion to the RO constraint form assumption, since J_{RO} is the Jacobian of problem (RO). The Lagrangian of problem (RO) is

$$L^{RO}(x_1, \lambda) = f(x_1, Tx_1) + \langle g_1(x_1), \lambda_1\rangle + \langle g_3(x_1, Tx_1), \lambda_3\rangle .$$

At the solution of (O), using the interpolation assumption and the chain rule leads to

$$\frac{d^2}{dx_1^2}L^{RO}(x_1^*, \lambda_1^*+S(x_1^*)^T\lambda_2^*, \lambda_3^*) = \begin{bmatrix} I_m \\ T \end{bmatrix}^T \nabla^2_{xx}\widehat{L}(x^*, \lambda_1^*+S(x_1^*)^T\lambda_2^*, \lambda_3^*) \begin{bmatrix} I_m \\ T \end{bmatrix} + O(\epsilon_0).$$

Therefore, from Lemma 3(ii), for ϵ_0 sufficiently small, the matrix

$$\nabla^2_{x_1x_1} L^{RO}(x_1^*, \lambda_1^* + S(x_1^*)^T\lambda_2^*, \lambda_3^*) \text{ is } p.d. \text{ over } \{\Delta x_1 | J_O \Delta x_1 = 0\}. \quad (12)$$

Given the fact that J_{RO} has full row rank and that $||J_{RO} - J_O|| = O(\epsilon_0)$, it follows from the interpolation assumption and from (12) that for ϵ_0 sufficiently small J_O also has full row rank and that

$$\nabla^2_{x_1x_1} L^{RO}(x_1^*, \lambda_1^* + S(x_1^*)^T\lambda_2^*, \lambda_3^*) \text{ is } p.d. \text{ over } \{\Delta x_1 | J_{RO} \Delta x_1 = 0\}. \quad (13)$$

Since J_{RO} is assumed to have full row rank, it follows from (12) that for ϵ_0 sufficiently small the (RO) problem satisfies the SOSC and the CQC.

For the second part of the proof the focus shifts to the residual in the first-order conditions of (RO). From Lemma 2 and the *Interpolation Assumption*,

$$\begin{aligned}
&\nabla_{x_1} L^{RO}(x_1^*, \lambda_1^* + S(x_1^*)^T\lambda_2^*, \lambda_3^*) \\
&= \nabla_{x_1} f(x_1^*, Tx_1^*) + \nabla_{x_2} f(x_1^*, Tx_1^*)T + \nabla_{x_1} \langle g_1(x_1^*), \lambda_1^* \rangle \\
&+ \nabla_{x_2} \langle g_2(Tx_1^*), \lambda_2^* \rangle T + \nabla_{x_1} \langle g_3(x_1^*, Tx_1^*), \lambda_3^* \rangle + \nabla_{x_2} \langle g_3(x_1^*, Tx_1^*), \lambda_3^* \rangle T \\
&= O(\epsilon_0) + (\nabla_{x_1} f(x_1^*, x_2^*) + \nabla_{x_1} \langle g_1(x_1^*), \lambda_1^* \rangle + \nabla_{x_1} \langle g_3(x_1^*, x_2^*), \lambda_3^* \rangle) \\
&+ (\nabla_{x_2} f(x_1^*, x_2^*) + \nabla_{x_2} \langle g_2(Tx_1^*), \lambda_2^* \rangle + \nabla_{x_2} \langle g_3(x_1^*, x_2^*), \lambda_3^* \rangle) T \\
&= O(\epsilon_0)
\end{aligned}$$

where the result of Lemma 2(i)

$$\left\langle \nabla_{x_1} g_1(x_1^*), S(x^*)^T\lambda_2^* \right\rangle = \langle \nabla_{x_2} g_2(Tx_1^*), \lambda_2^* \rangle T,$$

was taken into account. In addition, it is immediate that $\nabla_{\lambda_1} L^{RO}(x_1^*, \lambda_1^* + S(x_1^*)^T\lambda_2^*, \lambda_3^*) = g_1(x_1^*) = 0$ and

$$\nabla_{\lambda_3} L^{RO}(x_1^*, \lambda_1^* + S(x_1^*)^T\lambda_2^*, \lambda_3^*) = g_3(x_1^*, Tx_1^*) = g_3(x_1^*, x_2^*) + O(\epsilon_0) = O(\epsilon_0).$$

Therefore,

$$\nabla_{(x_1,\lambda_1,\lambda_3)} L^{RO}\left(x_1^*, \lambda_1^* + S(x_1^*)^T\lambda_2^*, \lambda_3^*\right) = O(\epsilon_0).$$

Since the problem (RO) satisfies the QCQ and SOSC, it follows from the theory of constrained optimization that for ϵ_0 sufficiently small, the Jacobian of the nonlinear equation

$$\nabla_{(x_1,\lambda_1,\lambda_3)} L^{RO}(x_1, \lambda_1, \lambda_3) = 0$$

is nonsingular at $\left(x_1^*, \lambda_1^* + S(x_1^*)^T\lambda_2^*, \lambda_3^*\right)$. From Kantorovich's theorem [19, Theorem 12.6.1] it follows that this nonlinear equation has a solution in the neighborhood of $\left(x_1^*, \lambda_1^* + S(x_1^*)^T\lambda_2^*, \lambda_3^*\right)$ that, because of the positive definiteness of $\nabla^2_{x_1x_1} L^{RO}$ on the null space of the constraints, is a local solution of (RO). □

Note that the Lagrange multiplier of the constraint $g_1(x_1) = 0$ is sharply different in the solutions of the problems (O) and (RE), and that of the problem (RO), although the representative variables x_1^* are within $O(\epsilon_0)$. In the (RO) problem the respective constraints also carry the weight of the $g_2(x_2) = 0$ of (O), which does not occur in the (RE) problem.

Clearly, the conditions that render the (RO) problem well posed are less restrictive than the corresponding ones for the reduced problem (RE). In particular, it is unfortunate that a regularity result for the nonlinear equation (RE) for the case where $g_3 \neq \emptyset$ could not be provided. With the notation of the proof of Theorem 1, the difficulty originates in the fact that in this case the Jacobian of (RE) approaches

$$\begin{bmatrix} J_{RE}^{11} & \nabla_{x_1} g_1(x_1^*)^T & \nabla_{x_1} g_3(x_1^*, x_2^*)^T \\ \nabla_{x_1} g_1(x_1^*) & 0 & 0 \\ \nabla_{x_1} g_3(x_1^*, x_2^*) + \nabla_{x_2} g_3(x_1^*, x_2^*) T & 0 & 0 \end{bmatrix},$$

which is not a symmetric matrix. Therefore, one can no longer apply the proof technique, which relied on the fact that the positive definiteness of the upper left corner of the matrix with respect to the null space of the other rows of the matrix implies the nonsingularity of the corresponding symmetric indefinite matrix.

On the other hand, with techniques from the proofs of Theorems 1 and 2, it is immediate that if (RE) has a nonsingular Jacobian at $(x_1^*, \lambda_1^*, \lambda_3^*)$, then (RO) is also regular at x_1^* and both have primal solutions within O(ϵ_0) of x_1^*.

Note that our analysis refers to the regularity of local minima of the original optimization problem. Indeed, the regularity assumption that we make at the solution of the full optimization problem (O) are sufficient conditions for the existence of a local minimizer, and our theoretical results merely state that under appropriate conditions the reduced problems have a local solution in the neighborhood of that minimizer. In the case of the global minimum, it is immediate that the reduced optimization problem (RO) cannot introduce a spurious global minimum, since it is a minimization over a strictly smaller set (the one constrained by the interpolation relationship) and its minimum value must be larger. In addition, Theorem 2 can be used to state that (RO) has a local minimum in the neighborhood of the global minimum of the full optimization problem (O). At this time we cannot make similar statements about the reduced equation approach (RE).

Note also that in this work, much as in the references [4,16,5,12,13,17,23] we treat only "static" problems. An important line of research not discussed here is the one of the reduction of time-dependent problems, though some of the ideas seem readily extensible (if one considers for example implicit time stepping schemes to approximate time-dependent problems).

3.3. Further Computational Improvements

Problems (RO) and (RE) have a dimension of the variables space that is equal to the dimension of the variable x_1. Therefore any Newton-type methods applied to the reduced problems will work in a much smaller space than the ones applied to the original problem (O). This has two computational benefits.

– This makes available a larger variety of tools that perhaps do not scale to the size of the original problem. It is conceivable that even direct methods would be applicable in some configurations.

– Moreover, the condition number of the reduced approach may be much smaller and even matrix free iterative method may need less iterations to converge. While this is a difficult statement to prove rigorously for the general case we point to several pieces of evidence that indicate that this may be a wide occurrence.
  – The numerical results for the three dimensional density functional theory application that is presented in Subsection 4.3. In that section we use a matrix free method for solving the reduced optimization problem
  – The analysis of quasi-continuum approaches applied to material science problems with Lennard Jones potentials. The analysis indicates that the original optimization problem has a condition number that behaves like order $\frac{1}{q_1+q_2}$, whereas quasi-continuum approaches produce problems that approach the solution of continuum models [15], which means that one expects that their condition number is tied to the size off the macro scale mesh, $O(\frac{1}{q_1})$, and not the size of the interatomic distance.

Nonetheless, if an iterative method is used to solve the reduced problems one may need to explore at every iteration the entire (x_1, x_2) space, in order to compute the data of the reduced problems. We illustrate the situation with the following example. We represent the components of the vectors $x_1$and x_2 by

$$x_1 = (x_{11}, x_{12}, \ldots, x_{1q_1}),\ x_2 = (x_{21}, x_{22}, \ldots, x_{2q_2}).$$

Assume that the objective function has the following expression

$$f(x_1, x_2) = f^1(x_1, x_2) + f^2(x_1, x_2)$$

and that its first component can be written as

$$f^1(x_1, x_2) = \sum_{i=1}^{q_1} f^0(x_{1i}) + \sum_{i=1}^{q_2} f^0(x_{2i}).$$

Here, $f^0(\cdot)$ is a smooth function. Then, the substitution $x_2 = Tx_1$ prompted by the interpolation assumption results in $x_{2i} = T_i x_1,\ i = 1, 2, \ldots, q_2$ and

$$f^1(x_1, Tx_1) = \sum_{i=1}^{q_1} f^0(x_{1i}) + \sum_{i=1}^{q_2} f^0(T_i x_{2i}). \tag{14}$$

Here, $T_i,\ i = 1, 2, \ldots, q_2$ are q_1 dimensional vectors.

Then, evaluation of $f^1(x_1, Tx_1)$ and $\nabla_{x_1} f^1(x_1, Tx_1)$ needs to explore all the $q_1 + q_2$ elements of the above sum (1.1), even if the result is a function of only q_1 independent variables, and $q_1 \ll q_2$.

There exist exceptions to this situation. For example, in the (RE) approach for potential based configurations with cutoff, only a small number of atoms in the neighborhood of the positions of the representative atoms need to be explored in order to compute the vector function and its Jacobian. Therefore the number of operations needed by a matrix free approach at each iteration is proportional to the number of representatives degrees of freedom, that is, the dimension of the vector x_1. Nonetheless, such fortunate outcome cannot be expected in general as proven by the example above (14).

To avoid the use of $q_1 + q_2$ operations at every step of an iterative method, the function $f^1(x_1, Tx_1)$ is further approximated in some quasi-continuum approaches [23]. For example, if $f^0(\cdot)$ is smooth, we can use the approximation

$$f^0(x_{2i}) \approx \hat{T}_i \begin{pmatrix} f^0(x_{11}) \\ f^0(x_{12}) \\ \vdots \\ f^0(x_{1q_1}) \end{pmatrix} = \sum_{j=1}^{q_1} \hat{t}_{ij} f^0(x_{1j})$$

Here

$$\hat{T} = \begin{bmatrix} \hat{T}_1 \\ \hat{T}_2 \\ \vdots \\ \hat{T}_{q_2} \end{bmatrix} = \begin{bmatrix} \hat{t}_{11} & \hat{t}_{12} & \hat{t}_{1q_1} \\ \hat{t}_{21} & \hat{t}_{22} & \hat{t}_{2q_1} \\ & & \\ \hat{t}_{q_2 1} & \hat{t}_{q_2} & \hat{t}_{q_2 q_1} \end{bmatrix}$$

is an interpolation operator (perhaps even the interpolation operator T).

In that case

$$f^1(x_1, Tx_1) \approx f^{1\hat{T}}(x_1) = \sum_{i=1}^{q_1} w_i f^0(x_{1i}),$$

where the weights w_j are defined as $w_j = 1 + \sum_{i=1}^{q_2} t_{ij}$ $j = 1, 2, \ldots q_1$. These weights are computed by exploring only once the degrees of freedom corresponding to the vector x_2, and then the function $f^{1\hat{T}}(x_1)$ is used as a surrogate for the objective function of the reduced problem, and needs only $O(q_1)$ operations to compute.

We note that if surrogate functions are used, the conclusions presented in Subsections 3.1 and 3.2 could still be reached provided that we can enforce the quality of the surrogate, such as

$$\begin{array}{l} \sup_{\|x_1 - x_1^*\| \le \delta_0} \left| f^{1\hat{T}}(x_1) - f^1(x_1, Tx_1) \right| \le \delta \\ \sup_{\|x_1 - x_1^*\| \le \delta_0} \left| \nabla_{x_1} f^{1\hat{T}}(x_1) - \frac{d}{dx_1} f^1(x_1, Tx_1) \right| \le \delta \\ \sup_{\|x_1 - x_1^*\| \le \delta_0} \left| \nabla^2_{x_1 x_1} f^{1\hat{T}}(x_1) - \frac{d^2}{dx_1^2} f^1(x_1, Tx_1) \right| \le \delta \end{array} ,$$

where the parameter δ_0 is fixed, provided that the parameter δ is sufficiently small with a size to be determined in the analysis. Such results would follow from the fact that the reduced problem is well-posed and stability results of nonlinear equations and nonlinear optimization problems [7]. Nonetheless, this would tremendously complicate the analysis. In addition, the variety of such surrogates is significant [23, 5] which makes their unified investigation non-trivial. We leave the development of an appropriate analysis framework to future research.

4. Numerical Experiments

All physical units used in this section are omitted and physical quantities involved are considered dimensionless.

4.1. Numerical Justification of the Assumptions: A Potential-based Calculation.

In this subsection the validity of the assumptions made in the previous sections is scrutinized. Particular attention is paid to the H-T assumption because in the context of the (RE) approach, it is the more unusual and restrictive of the assumptions made. The vehicle for this investigation will be a test case in which the objective function $f(x_1, x_2)$ in (O) is the total energy of a set of atoms represented in a one-dimensional setup, whose pairwise interaction is governed by the Lennard-Jones potential (see, for instance, [1]). The test is similar in spirit to, but simpler in complexity than, the more general three-dimensional ones presented in [12]. For this problem, $x = (r_1, \ldots, r_A)^T$, where r_i is the coordinate of atom i. The energy is defined in terms of a pairwise potential $V(\cdot)$. The total energy is $E(x) = \sum_{i}^{A} \sum_{j>i}^{A} V(r_i - r_j)$. The stable configuration of the atoms is obtained when the energy is minimized, which in turn implies that

$$0 = F(x) = \nabla E(x).$$

For a string of $A = 101$ atoms, the original problem (UO) (from unconstrained optimization), is solved using the (RE) approach. The representative atoms are the atoms 1, 2, 3, 4, 23, 42, 61, 80, 99, 100, 101. The atoms 4 through 99 are called "inner" atoms. In spite of being representative, the atoms 1, 2, 3, and 100, 101 are not used in the interpolation to prevent the boundary effects from crossing into the reconstruction process associated with the inner nonrepresentative atoms. The position of the 61st atom is fixed because the energy functional is translation invariant and it would thus have unbounded level sets, possibly compromising the global convergence properties of the algorithms. In the framework of problem (O), $x_1 = (r_1;\, r_2; r_3; r_4;\, r_{23}; r_{42}; r_{61}; r_{80};\, r_{99}; r_{100}; r_{101})$ and $x_2 = T\, x_1$. In addition, $f(x_1, x_2) = E(x)$, $g_1(x_1) = r_{61} - 61$, $g_2(x_2) = \emptyset$, $g_3(x_1, x_2) = \emptyset$. Both the RE constraint form assumption and RO constraint form assumption hold for this test, as well as the CQC part of the regularity assumption.

The solution is found with the package SNOPT [10] through the AMPL interface [9]; the solution was found in about 10 iterations. The expression of the Lennard-Jones potential considered was

$$V(r) = \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}, \quad \sigma = 1.122.$$

The problem was initialized with $r_i = i$, $i = 1, 2, \ldots, 101$. At the solution of (UO), the columns of

$$L_R = -\left[\nabla^2_{x_2 x_2} L(x^*, \lambda^*)\right]^{-1} \nabla^2_{x_2 x_1} L(x^*, \lambda^*)$$

that correspond to the atoms 4, 23, 42, 61, 80, 99 were calculated and displayed in Figure 1 (as a function of the index in the x_2 vector). The columns of L_R that correspond to the atoms 1,2,3, 100, 101 are negligible, in the sense that their norm is more than 100,000 times smaller than the one corresponding to the other columns. These results almost perfectly justify the H-T assumption, in that the columns of L_R are essentially identical to the ones of the linear interpolation operator with nodes at the inner representative atoms. Perhaps less surprisingly, the positions of the atoms themselves at the solution point also satisfy the same linear interpolation pattern and therefore justify the

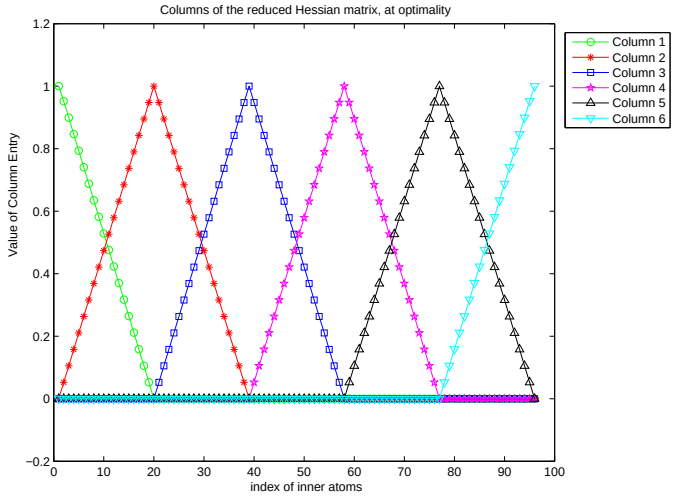


Fig. 1. Columns of L_R.

interpolation assumption. In addition, verifying the eigenvalues of the Hessian of the Lagrangian indicates that the SOSC part of the regularity assumption was satisfied.

For comparison, the same columns of L_R are evaluated away from the equilibrium (the configuration was first perturbed slightly as shown in Figure 3), and the results are displayed in Figure 2. The variation between two consecutive interatomic distances with respect to the original problem was no larger than 1.6%, and the end points were identical. Nevertheless, that pattern of the columns is now far away from the one corresponding to the interpolation operator T, which leads to the conclusion that the H-T assumption can be expected to be valid only near the solution of the original problem (O). The assumption is expected to be more accurate as the system size approaches the continuum limit.

In summary, the regularity assumption, interpolation assumption, H-T assumption, RE constraint form assumption, and RO constraint form assumption do apply, and therefore according to Theorems 1 and 2 the reduced problems (RE) and (RO) have a solution in the neighborhood of the solution of problem (UO).

4.2. Example Application of (RE) and (RO) to Density Functional Theory Computations

The model reduction techniques (RE) and (RO) are applied to solving an electronic structure computation problem. The purpose is to compute the electron density (which is a scalar function of the spatial variables) for a given position of the atoms and a given total number of electrons. A form of the local quasi-continuum method has been developed for electronic structure computation [6]. In that work, the local nature of the method required elements much larger than a crystal cell and the use of periodic boundary conditions. The approach proposed in this paper is not restricted by the use of periodic boundary conditions.

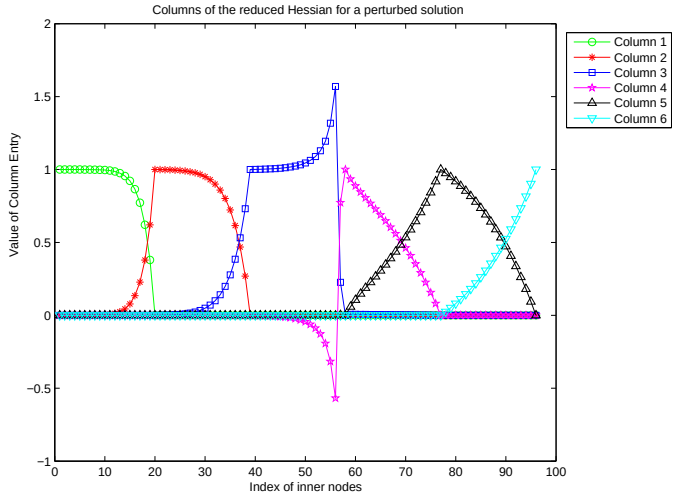


Fig. 2. Columns of L_R, perturbed configuration.

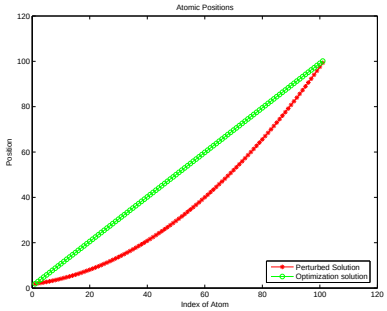


Fig. 3. Positions of original and perturbed solutions.

As a mathematical model, the problem is an optimization problem whose objective function is the total energy functional $E[\rho, \{R_A\}]$, where $\rho = \rho(r)$ is the variable electronic density function that is subject to the constraint that the total electronic density ($\int \rho(r)dr$) should add up to a prescribed number of electrons, and $\{R_A\}$ are the parameter atomic positions according to the Born-Oppenheimer assumption (see, for instance, [22]).

The example is built around the Thomas-Fermi-Dirac form of the energy functional (see, for instance, [14]):

$$E[\rho, \{R_A\}] = E_{ne}[\rho, \{R_A\}] + J[\rho] + K[\rho] + T[\rho], \tag{15}$$

where

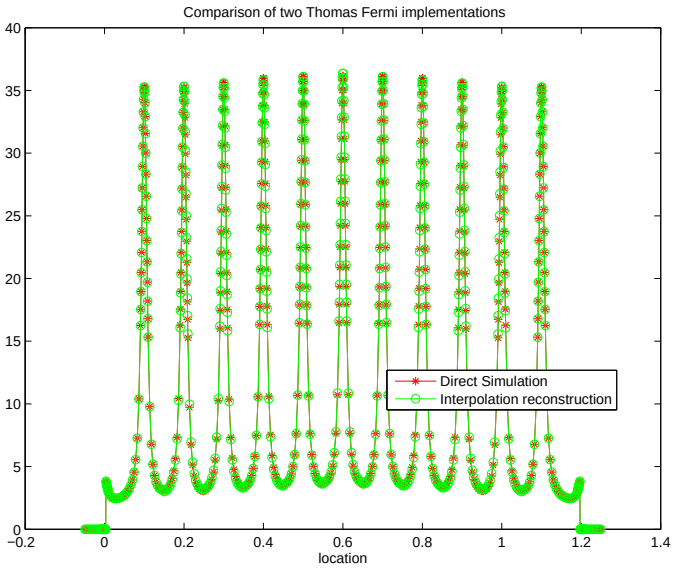


Fig. 4. Solution to (O) and (RO) problems.

$$E_{ne}\left[\rho,\{R_A\}\right] = -\sum_{A=1}^{M}\int \frac{Z_A\ \rho(r)}{\|R_A - r\|}\ \mathrm{d}r \tag{16a}$$

$$J\left[\rho\right] = \frac{1}{2}\int\int \frac{\rho(r)\ \rho(r')}{\|r - r'\|}\ \mathrm{d}r\ \mathrm{d}r' \tag{16b}$$

$$T\left[\rho\right] = C_F \int \rho^{\frac{5}{3}}(r)\ \mathrm{d}r \tag{16c}$$

$$K\left[\rho\right] = -C_x \int \rho^{\frac{4}{3}}(r)\ \mathrm{d}r. \tag{16d}$$

Here $C_F = \frac{3}{10}(3\pi^2)^{2/3}$, and $C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}$; E_{ne} is the energy corresponding to nucleus-electron interaction; J is the Coulomb energy; K represents the exchange energy; T is the kinetic energy; Z_A is the atomic number associated with nucleus A; r_i is the global position of electron i; R_A is the global position of nucleus of atom A; and $\int(\cdot)$ without integration limits is an integral over the entire domain.

It is well accepted that both for quantum chemistry and solid-state physics the Thomas-Fermi-Dirac functional is an inaccurate DF representation. This is less relevant in this context because the interest lies in evaluating the benefit of using a model reduction approach, rather than assessing the accuracy of the underlying DFT model. The purpose of the numerical experiment is to compare the solution of the full model with a prediction computed with the reduced model.

A detailed description of the reduction approach for an arbitrary domain and an arbitrary number of representative subdomains can be found in [18], and it is only

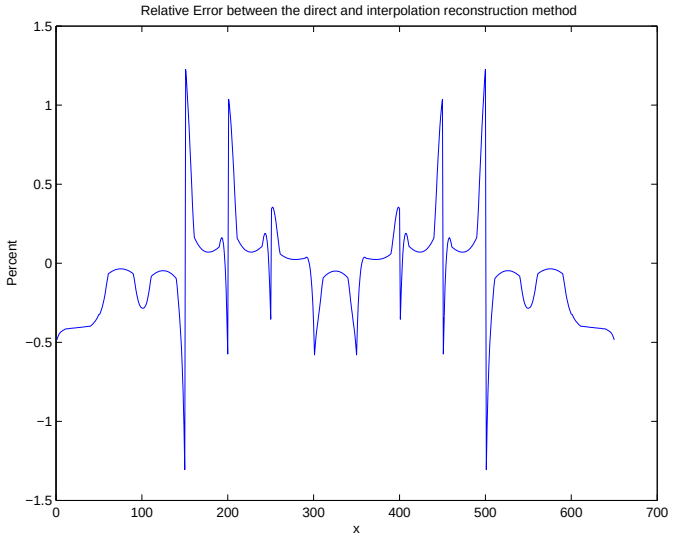


Fig. 5. Point-to-point relative error between (O) and (RE).

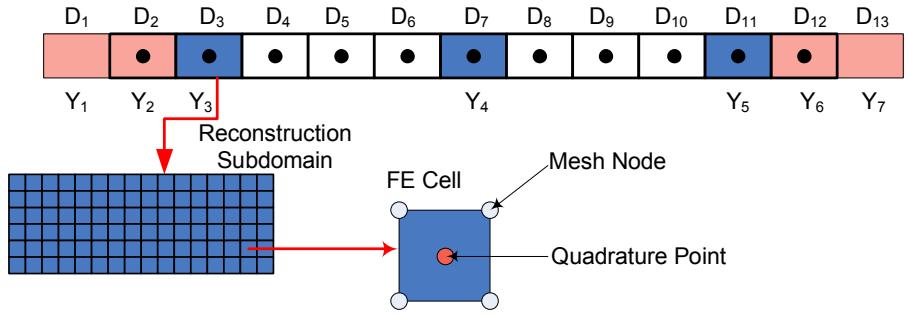


Fig. 6. Separation of the computational domain in representative and passive subdomains.

briefly discussed here. The computational domain is divided in subdomains D_i, $i = 1, 2, \ldots, u$, out of which p of them are chosen to be representative, and denoted by $Y_\alpha, \alpha \in \{1, \ldots, p\}$; the remaining $u - p$ subdomains, are called passive (the white subdomains in Figure 6). A choice of seven representative subdomains is presented in Figure 6. The density ρ_i on subdomain D_i is expressed by interpolation in terms of reference densities $\rho_\alpha \in Y_\alpha, \alpha \in \{1, \ldots, 7\}$. A set of weights ϑ determined based on the type of interpolation considered (linear, quadratic, etc.) is used to this end:

$$\rho_i(\Phi(\mathbf{r}^{0\prime}, t)) = \sum_{\alpha=1}^{p} \vartheta_\alpha(i)\rho_\alpha(\Phi(\mathbf{r}^{0\prime} + \mathbf{T}_{i\alpha}, t)) \tag{17}$$

where the vector $\mathbf{T}_{i\alpha}$ is the translation vector that takes the point $\mathbf{r}^{0\prime}$ in subdomain D_i to its image in the subdomain Y_α. The deformation mapping $\Phi(\cdot)$ is defined with respect to a "macroscale" mesh that contains many nuclei per element, much like in the

quasi-continuum method for potentials [23]. It describes the deformation of the subdomain (the relative displacement of the nuclei) with respect to a reference configuration. To simplify the definition of the translations, the nonrepresentative subdomains are assumed to correspond to a periodic reference configuration. In that case, in the reference configurations the subdomains D_i may be thought to be of identical shape, in which case, the interpolation approach is reminiscent of the gaptooth method [11] where the representative subdomains are the "teeth". In this work, however, the reconstruction by interpolation of the density is also carried out in the gaps, and not only at the boundary of the teeth due to the long-range electrostatic interactions.

For the interpolation ansatz to be reasonably accurate, regions that have dislocations, impurity atoms, or other irregularities must belong to representative subdomains. Therefore only some of the representative subdomains are used in the process of computing the value of the electron density in the passive subdomains, and these subdomains are called reconstruction subdomains. Among the representative subdomains, a non-zero value of the reconstruction weight in (17) is the defining attribute of a reconstruction subdomain.

For the test case considered, a one-dimensional subdomain contains 11 clamped nuclei with distance of 0.1 between consecutive nuclei and with unit charge $Z_A = 1$; the total number of electrons is $N = 11$. The atoms are at their reference positions and we have $\Phi(r) = r$. The location of the atoms is indicated by the small black circles in Figure 6. There are 11 subdomains D_2, D_3, ..., D_{12} of length 0.1 centered at the atomic positions, each with 50 nodes, of which 30 are equally spaced on an interval centered at the position of the atom and whose length is 1/5 of the distance between two atoms. In the 11 subdomains, the mesh is invariant by a translation of length 0.1. The trapezoidal rule was used for discretization of the integral operators (see, for instance, [2]). In order to allow the solution to relax near the boundary, two more boundary domains D_1 and D_{13}, of identical size and meshing but without any atoms, were added to D_2 and D_{12}, respectively. Restriction of electron density to a one-dimensional function has no physical meaning, but serves as illustration of the applicability of our interpolate-and-optimize approach.

In the framework of (O), (RE), and (RO), the representative variables x_1 are the electronic density values from subdomains Y_α, $\alpha = 1, \ldots, 7$. The values x_2 represent the electron density at nodes of the mesh from the rest of the subdomains. With the nodes of the mesh denoted by z_k, $k = 1, 2, \ldots, 650$, the interpolation operator is defined as follows:

$$(T\rho)(z_k) = \frac{4-i}{4}\,\rho\left(z_k - \frac{i}{10}\right) + \frac{i}{4}\,\rho\left(z_k + \frac{4-i}{10}\right), \quad z_k \in D_{3+i} \cup D_{7+i},\ i = 1, 2, 3. \tag{18}$$

The reconstruction subdomains are Y_3, Y_4, and Y_5; the other subdomains Y_α are representative subdomains , but not reconstruction subdomains, in order to prevent boundary effects from crossing into the reconstruction. In order to avoid the singularity brought about by the $\frac{1}{r}$ terms, a smoothing parameter $\delta = 10^{-4}$ was considered; terms like $1/||\cdot||$ were replaced with $1/||\cdot + \delta||$ (in two- and three-dimensional applications these singularities are integrable and can be treated by special approaches; this "smoothing" is actually not required).

The problem was modeled in the AMPL environment [9]; the resulting (O), (RE), and (RO) problems were solved with SNOPT (where the second was represented only as a nonlinear equation) [10]. All three formulations were successfully solved in a small number of major iterations (no more than 10). Note that the RO constraint form assumption holds because (a) the discretization of the constraint ($\int \rho(r)dr = N$) results in one linear constraint with positive coefficients and (b) T, seen as a matrix, has nonnegative entries. Then, $\nabla_{x_1} g_3(x_1^*, x_2^*) + \nabla_{x_2} g_3(x_1^*, x_2^*)T$ is a row vector with positive entries, which has rank one when seen as a matrix. Therefore, because the second-order sufficient condition of the regularity assumption has also been validated, the conclusions of Theorem 2 should hold. The assumptions of Theorem 1 could not be verified; nonetheless, the reduced nonlinear equation (RE) does give results of the same quality as (RO).

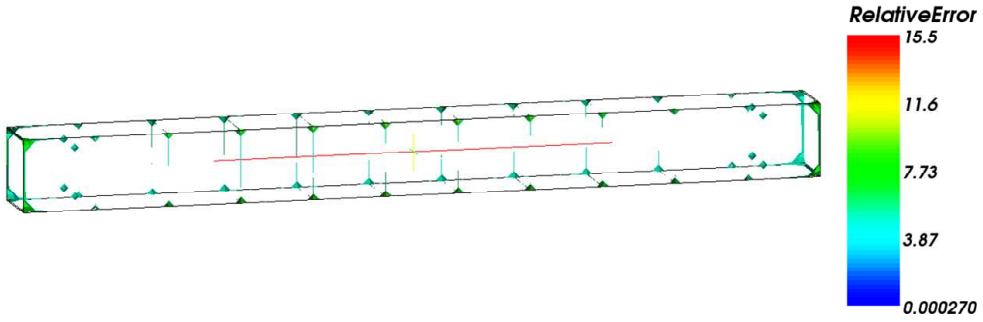
The solution of (O) and (RO) are provided in Figure 4, whereas the point-to-point solution error ($\frac{||\rho^{RE}(z_k) - \rho^{O}(z_k)||}{||\rho^{O}(z_k)||}$, at all grid points $z_k,\ k = 1, \ldots, 650$) between problems (O) and (RE) are displayed in Figure 5. The density plots of (O) and (RO) are essentially identical at visual accuracy, and the interpolation approach is successful in reconstructing the solution in the "gap" domains. Note, however, that the point-to-point solution error of (RO) is of the same order to the one of (RE) presented in Figure 5, that is, a maximum value of around 1% (though its uniformity is exceptional and is responsible for the remarkable apparent accuracy in Figure 4). The number of degrees of freedom of problems (RE) and (RO) is smaller by a factor of $7/13$. For larger, three dimensional configurations, the approach is expected to create an accurate reduced problem with an even smaller ratio of number of representative versus total number of degrees of freedom (a third power appears from the third-dimensional aspect alone, which is mitigated by the effect of the boundaries). The proposed approach does not have to apply only to a domain with a surface or a boundary. Indeed, one could treat much of the bulk with periodic boundary conditions and use the reconstruction technique only around defects.

This work does not address the energy minimization for both electronic density and atomic positions, which is the case in [6]. On the other hand, the method can be readily adapted to that case by using an interpolation based on a macroscale deformation of the crystal. Details are presented in [18].

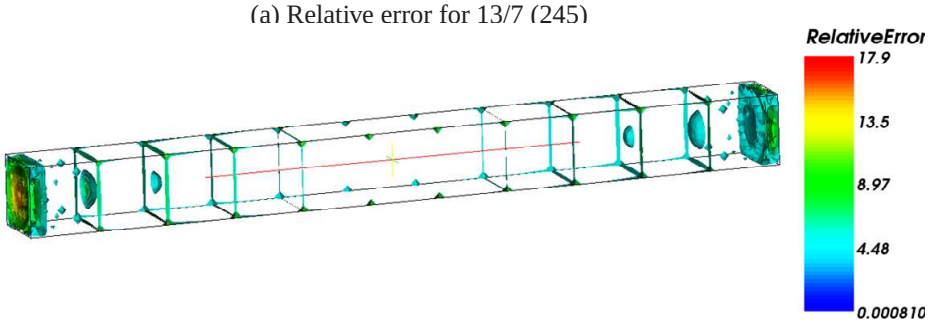
4.3. Three Dimensional String of Atoms Example

Our example is a three-dimensional variation of the one dimensional case analyzed in the previous section. The size of each of the 3D subdomains surrounding a hydrogen atom is $3 \times 3 \times 3$ (all units henceforth are atomic units). A full simulation with no reconstruction is provided as the reference solution.

In this case, the modeling technique described in Subsection 3.3 was used for the kinetic term $T[\rho]$ and exchange energy term $K[\rho]$ in (16), whereas the Coulomb term $J[\rho]$ and the nuclei-electron term E_{ne} are computed by exploring the nonrepresentative degrees of freedom (coresponding to the entries of x_2) only once. For the latter case, appropriate kernel matrices of dimensions $q_1 \times q_1$ are computed and used for each function and derivative evaluation of the problem (RO).



(a) Relative error for 13/7 (245)



(b) Relative error for 13/5 (221)

Fig. 7. Relative error surface for the 13-subdomain scenarios using (a) 7 and (b) 5 active subdomains. In parentheses we show the number of optimization iterations. The number of active subdomains considered in the algorithm reflects in the quality of the numerical solution: more active subdomains result in a larger number of degrees of freedom, which positively impacts ability to relax to lower energy levels and reduces boundary artifacts.

Two scenarios with seven and five active subdomains were subsequently considered for reduction to the problem; all meshes in this numerical experiment are uniform. In the first scenario, the subdomains D_1, D_2, D_3, D_7, D_{11}, D_{12}, and D_{13} were active; only D_3, D_7, and D_{11} were used for reconstruction. In the second scenario, the subdomains D_1, D_2, D_7, D_{12}, and D_{13} were active; only D_2, D_7, and D_{12} were used for reconstruction. For this test, the number of nodes/cells in the active subdomains is as follows: 28561/22464 for the nonreconstruction case (13/13), 15379/12096 for the 7/13, and 10985/8640 for the 5/13 case. We have used an interpolation operator defined by (17), similar to to (18), modified for the three-dimensional case. Specifically, the rule is

$$(T\rho)(\mathbf{z}) = \tfrac{l-i}{l}\,\rho\left(\mathbf{z}-(3i,0,0)\right) + \tfrac{i}{l}\,\rho\left(\mathbf{z}-(3(l-i),0,0)\right), \quad \mathbf{z}\in D_{j_\alpha+i}\cup D_{j_\alpha+i+l},\ i=1,2,\ldots,l-1.$$

In the case of 7 reconstruction domains, we have that j_α is one of 3 and 7 and $l=4$, whereas in the case of 5 reconstruction domains, j_α is one of 2 and 7 and $l=5$. Therefore the case $l=5$ uses less domains where the electronic density is represented and more domains where is reconstructed and is thus expected to have larger error.

We have used a hexahedral (cubic) mesh. Figure 7 displays the relative errors; shown are only the regions where the relative error is larger than 5%. The results show a slight improvement in the seven-subdomain case; as the number of active subdomains

Active Subdomains	13	7	5
Number of Iterations	605	245	221
Total Energy	-14.257	-14.256	-14.256

Table 1. Summary of the results. TAO-BLMVM stoping criteria are are 10^{-6} absolute and 10^{-5} relative gradient error.

increases, the quality of the results improve. Because of the dimension reduction, the size of the optimization problem decreases, thereby leading to a reduction in the number of iterations. Moreover, each iteration is computationally less expensive. The large relative errors are explained by the small values assumed by the electron density away from the nuclei where in practice it is expected to be zero. This and the boundary artifacts explain the accumulation of the 5% relative error isosurfaces far away from the nuclei and close to the boundary of the solution domain. While an exact quantitative characterization of the boundary artifacts remains to be produced, they are traced back to at least two sources. First, the small pockets of nonzero electron density are explained by a slow convergence rate of the optimization algorithm that currently does not use Hessian information and stops before clearing these pockets in remote corners of the nanostructure. Second, and more important, the assumption of underlying approximate periodicity of the solution when used in conjunction with a small number of reconstruction subdomains (few degrees of freedom) limits the capacity of the electron density to relax due to these periodicity constraints that must be numerically satisfied. As expected and illustrated in the results corresponding to the 5 active subdomains case, the situation is exacerbated as fewer degrees of freedom are available in the energy minimization step of the method. In spite of these boundary artifacts, it should be noted that the differences in total energy are small for both the 7 and 5 active subdomain cases (about 0.007%; see Table 1). The results reported were obtained by running in parallel with 13 processes on a Linux cluster.

5. Conclusion and Future Work

Model reduction (or reconstruction) techniques in computational materials science based on nonlocal quasi-continuum-like approach produce reduced optimization or nonlinear equations problems with a substantially smaller number of degrees of freedom. We show that, under certain assumptions, the reduced problem is well posed. Several potential and density-functional examples validate our findings.

A three-dimensional parallel computational environment that supports the (RO) approach is currently developed in a fashion that includes both explicit DFT approaches (such as the OFDFT [24]) and more elaborate Kohn-Sham approaches in which the kinetic energy functional and its derivatives are not explicitly available.

Acknowledgments

The authors would like to thank Todd Munson for valuable input. Mihai Anitescu and Dan Negrut were supported by Contract No. W-31-109-ENG-38 of the U.S. Department

of Energy. We thank the anonymous referees for useful suggestions and pointing us to the work of the authors of [4,5,16,20].

References

1. M. P. ALLEN AND D. J. TILDESLEY, *Computer Simulation of Liquids*, Clarendon Press, Oxford, 1987.
2. K. E. ATKINSON, *An Introduction to Numerical Analysis*, John Wiley & Sons Inc., New York, second ed., 1989.
3. D. P. BERTSEKAS, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.
4. X. BLANC, C. LEBRIS, AND P.-L. LIONS, *Atomistic to continuum limits for computational materials science*, Mathematical Modeling and Numerical Analysis, (2007). to appear.
5. W. E, J. LU, AND J. Z. YANG, *Uniform accuracy of the quasicontinuum method*, Physical Review B, 74 (2006), p. 214115.
6. M. FAGO, R. HAYES, E. CARTER, AND M. ORTIZ, *Density-functional-theory-based local quasicontinuum method: Prediction of dislocation nucleation*, PHYSICAL REVIEW B, 70 (2004).
7. A. V. FIACCO, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, 1983.
8. R. FLETCHER, *Practical Methods of Optimization*, John Wiley & Sons, Chichester, 1987.
9. R. FOURER, D. M. GAY, AND B. W. KERNIGHAN, *AMPL: A modeling language for mathematical programming*, Thomson, Toronto, Canada, 2nd ed., 2003. First chapter, software, and other material available at http://www.ampl.com.
10. P. E. GILL, W. MURRAY, AND M. A. SAUNDERS, *User's guide for SNOPT 5.3: A fortran package for large-scale nonlinear programming*, Report NA 97-5, Department of Mathematics, University of California, San Diego, 1997.
11. Y. KEVREKIDIS, C. W. GEAR, AND J. LI, *The gaptooth method in particle simulations*, Physics Letters A, 190 (2003).
12. J. KNAP AND M. ORTIZ, *An analysis of the quasicontinuum method*, J MECH PHYS SOLIDS, 49 (2001), pp. 1899–1923.
13. J. KNAP AND M. ORTIZ, *Effect of indenter-radius size on Au(001) nanoindentation*, Physical Review Letters, 90(22) (2003), pp. 226102–1– 226102–4.
14. W. KOCH AND M. C. HOLTHAUSEN, *A Chemist's Guide to Density Functional Theory*, John Wiley & Sons Inc., New York, second ed., 2001.
15. I. KUNIN, *Elastic Media with microstructure, I One-Dimensional Models*, Springer-Verlag, 1982.
16. P. LIN, *Theoretical and numerical analysis for the quasi-continuum approximation of a material particle model*, Mathematics of Computation, 72 (2002), pp. 657–675.
17. R. E. MILLER AND E. B. TADMOR, *The quasicontinuum method: Overview, applications and current directions*, Journal of Computer-Aided Materials Design, 9 (2002), pp. 203–239.
18. D. NEGRUT, M. ANITESCU, T. MUNSON, AND P. ZAPOL, *Simulating nanoscale processes in solids using DFT and the quasicontinuum method (IMECE2005-81755*, in Proceedings of IMECE 2005, ASME International Mechanical Engineering Congress and Exposition, 2005.
19. J. ORTEGA AND W. RHEINBOLDT, *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic Press, New York, 1972.
20. C. ORTNER AND E. SULI, *A-priori analysis of the quasicontinuum method in one dimension*, Tech. Rep. NA-06/12, Oxford University, Computing Laboratory, Oxford, UK, 2006.
21. D. RODNEY, *Mixed atomistic/continuum methods: static and dynamic quasi continuum methods*, in Proceedings of the NATO Conference in Thermodynamics, Microstructures and Plasticity, A. Finel, D. Maziere, and M. Veron, eds., Dordrecht, 2003, Kluwer.
22. A. SZABO AND N. OSTLUND, *Modern Quantum Chemistry*, Dover, 1989.
23. E. TADMOR, M. ORTIZ, AND R. A. PHILLIPS, *Quasicontinuum analysis of defects in solids*, PHILOS MAG A, 73 (1996), pp. 1529–1563.
24. Y. WANG, N. GOVIND, AND E. CARTER, *Orbital-free kinetic-energy density functionals with a density-dependent kernel*, Phys. Rev. B, 60 (1999), pp. 16350–16358.