

RESEARCH ARTICLE

WILEY

On non-stationary polarization methods in FFT-based computational micromechanics

Matti Schneider 

Karlsruhe Institute of Technology (KIT),
Institute of Engineering Mechanics,
Karlsruhe, Germany

Correspondence

Matti Schneider, Karlsruhe Institute of
Technology (KIT), Institute of Engineering
Mechanics, Karlsruhe, Germany.
Email: matti.schneider@kit.edu

Abstract

Polarization-type methods are among the fastest solution methods for FFT-based computational micromechanics. However, their performance depends critically on the choice of the reference material. Only for finitely contrasted materials, optimum-selection strategies are known. This work focuses on adaptive strategies for choosing the reference material, details their efficient implementation, and investigates the computational performance. The case of porous materials is explicitly included. As a byproduct, we introduce a suitable convergence criterion that permits a fair comparison to strain-based FFT solvers and Eyre–Milton type implementations.

KEYWORDS

ADMM, computational micromechanics, FFT-based computational homogenization, polarization methods, porous materials

1 | INTRODUCTION

1.1 | State of the art

As a typical model for computational homogenization^{1,2} on a periodic cell Q and at small strains, we are concerned with the optimization problem

$$W(\varepsilon) \rightarrow \min_{\varepsilon \in \mathcal{K}} \quad (1)$$

seeking the total strain field ε on the cell Q minimizing the average elastic energy

$$W(\varepsilon) = \langle w(\cdot, \varepsilon) \rangle_Q$$

under the kinematic constraints \mathcal{K} which encode compatibility and fixed macroscopic strain $\bar{\varepsilon}$. A prototype for primal methods for solving the problem (1) is projected gradient descent, which iterates

$$\varepsilon^{k+1} = \mathcal{P}_{\mathcal{K}} \left(\varepsilon^k - s^k \nabla W(\varepsilon^k) \right)$$

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *International Journal for Numerical Methods in Engineering* published by John Wiley & Sons Ltd.

for step sizes s^k , where ∇W denotes the gradient of W and $\mathcal{P}_{\mathcal{K}}$ refers to the orthogonal projection onto the kinematic constraint \mathcal{K} . More explicitly, for the L^2 inner product, we are led to the basic scheme^{3,4} of Moulinec and Suquet

$$\varepsilon^{k+1} = \bar{\varepsilon} - \Gamma^k : (\sigma(\varepsilon^k) - \mathbb{C}^k : \varepsilon^k), \tag{2}$$

where $\bar{\varepsilon}$ is the average strain, \mathbb{C}^k is an elastic reference material and the Γ^k -operator is defined by $\Gamma^k = \nabla^s(\text{div } \mathbb{C}^k \nabla^s)^{-1} \text{div}$ in terms of the symmetrized gradient ∇^s and the associated divergence div . We denote by $\sigma = \partial w / \partial \varepsilon$ the stress operator, which maps a strain field to the associated stress field.

The basic scheme (2) is the prototype of a *primal* method, that is, a method where the kinematic constraints are satisfied at every iteration. Extensions of the basic scheme lead to fast and robust solution methods for FFT-based computational micromechanics. These include conjugate gradient methods,⁵⁻⁷ fast gradient methods^{8,9} as well as Newton,¹⁰⁻¹² and Quasi-Newton methods.¹³⁻¹⁵ Please note that purely *dual* methods may be developed as well.^{16,17} These are, however, conceptually similar to the primal methods.

An alternative to the primal methods are primal-dual methods, where primal and dual feasibility, that is, kinematic compatibility and stress equilibrium, are satisfied only at convergence, in general. Michel et al.^{18,19} suggested to rewrite the optimization problem (1) in split form

$$\iota_{\mathcal{K}}(\varepsilon) + W(e) \rightarrow \min_{\varepsilon=e}, \tag{3}$$

where $\iota_{\mathcal{K}}$ denotes the *indicator function* of the kinematic constraint, attaining the value 0 if $\varepsilon \in \mathcal{K}$ and $+\infty$ otherwise. For fixed reference material \mathbb{C}^0 , the associated augmented Lagrangian function reads

$$L(\varepsilon, e, \lambda) = \iota_{\mathcal{K}}(\varepsilon) + W(e) + \langle \lambda : (\varepsilon - e) \rangle_Q + \frac{1}{2} \langle (\varepsilon - e) : \mathbb{C}^0 : (\varepsilon - e) \rangle_Q,$$

where λ is the field of Lagrange multipliers. Michel et al.^{19(sec.4.1.2)} suggested to utilize the associated alternating direction method of multipliers (ADMM)²⁰⁻²²

$$\begin{aligned} \varepsilon^{k+1} &= \bar{\varepsilon} - \Gamma^0 : (\lambda^k - \mathbb{C}^0 : e^k), \\ \sigma(e^{k+1}) + \mathbb{C}^0 : e^{k+1} &= \lambda^k + \mathbb{C}^0 : \varepsilon^{k+1}, \\ \lambda^{k+1} &= \lambda^k + \mathbb{C}^0 : (\varepsilon^{k+1} - e^{k+1}). \end{aligned} \tag{4}$$

Upon convergence, the two strain fields coincide, that is, $\varepsilon = e$ holds, and the field λ represents the associated stress field $\lambda = \sigma(\varepsilon)$. Please note that the second line of ADMM (4) involves an *implicit* stress update, that is, the mapping $e \mapsto \sigma(e) + \mathbb{C}^0 : e$ needs to be inverted. This contrasts with the basic scheme (2) and its descendants which involve an *explicit* stress update $\varepsilon \mapsto \sigma(\varepsilon)$.

We distinguish the stress operator σ and the *stress field*, which we denote by λ . Upon convergence of the algorithm (4), the stress field λ arises from the strain field ε via the stress operator $\lambda = \sigma(\varepsilon)$. Taking care of this distinction in the notation helps to clarify the structure of polarization-type methods, as they crucially rely upon implicit stress updates.

For linear elastic material behavior, Monchiet and Bonnet²³ introduced a family of methods iterating on the stress polarization $P = \sigma(\varepsilon) + \mathbb{C}^0 : \varepsilon$. As pointed out by Moulinec and Silva,²⁴ these methods included the Eyre–Milton method^{19,25} and, for the linear elastic case, the ADMM (4) as special cases.^{23,24} For linear elastic material behavior and a finite contrast between the phases, Moulinec and Silva²⁴ showed convergence of polarization schemes for any reference material (and reasonable damping coefficient). Also, in the class of reference materials proportional to the identity, $\mathbb{C}^0 = \alpha_0 \text{Id}$, they identified the optimum convergence rate at $\alpha_0 = \sqrt{\alpha_- \alpha_+}$ and vanishing damping, where α_{\pm} refer to the minimum and maximum eigenvalues of the elastic stiffnesses of the phases.

Building upon earlier work for nonlinear polarization methods,^{26,27} Schneider et al.²⁸ considered nonlinear polarization methods. Based on optimum parameter selection for the Douglas–Rachford splitting²⁹ (equivalent to ADMM for our case), the parameter choice $\alpha_0 = \sqrt{\alpha_- \alpha_+}$ could also be established for nonlinear material behavior, where α_{\pm} refers to the modulus of strong convexity and the Lipschitz constant of the stress operator σ . Furthermore, Schneider et al.²⁸ generalized previous work^{19,27,30} and showed that the implicit stress update (4) can be evaluated explicitly for small-strain inelastic materials where the stress is related to the *elastic* strain by Hooke’s law.

Renewed interest in polarization methods arose due to possible advantages in parallelization³¹ and for treating non-differentiable energies.^{32,33} Still, treating microstructures with pores remained out of scope for polarization methods. For a start, it was not even clear whether polarization methods converged for porous microstructures, or, to be more precise, for which reference materials convergence could be ensured. Secondly, the simple rule $\alpha_0 = \sqrt{\alpha_- \alpha_+}$ for choosing the reference material does not make sense if $\alpha_- = 0$.

For primal methods, it was realized that the original spectral discretization of Moulinec and Suquet^{3,4,34} was numerically unstable for porous microstructures. By changing the discretization to finite difference^{35,36} or finite element schemes,^{37,38} convergent computational methods could be obtained. In particular, the convergence speed of the solution methods were reduced to judicious parameter selection.^{7,39}

Anderson acceleration^{40,41} is a Quasi-Newton method applicable to generic fixed point schemes that was successfully applied to primal FFT-based methods¹³⁻¹⁵. Combined with polarization schemes, the resulting Anderson-accelerated polarization schemes⁴² showed robust and fast convergence, also for porous microstructures. Unfortunately, Anderson acceleration comes with a severe memory demand. For this reason, renewed interest was sparked concerning an adaptive selection of the reference material in polarization methods.

1.2 | Contributions

In recent years, primal FFT-based computational homogenization methods have experienced renewed interest. Solution strategies are sought which unite fast convergence with a low memory footprint. For this reason, modifications and extensions of the basic scheme (2) of Moulinec and Suquet^{3,4} were studied.^{5-8,39} One of the insights of these works is that it is generally advantageous to select the solver parameters adaptively instead of requiring a clever user input. For instance, the Barzilai–Borwein method³⁹ is nothing but the basic scheme where the reference material is updated for every iteration based on the Barzilai–Borwein rule.⁴³

To the best of the author's knowledge, adaptively changing the reference material has not been considered for polarization methods, and the work at hand is aimed at closing this gap. We consider the general iterative scheme

$$\begin{aligned} \varepsilon^{k+\frac{1}{2}} &= \bar{\varepsilon} + \mathbb{D}^k : \bar{\sigma} - (\Gamma^k + \mathbb{D}^k : \mathbb{Q} : \langle \cdot \rangle_Q) : (\lambda^k - \mathbb{C}^k : e^k), \\ \varepsilon^{k+1} &= 2(1 - \gamma)\varepsilon^{k+\frac{1}{2}} - (1 - 2\gamma)e^k, \\ \sigma(e^{k+1}) + \mathbb{C}^k : e^{k+1} &= \lambda^k + \mathbb{C}^k : \varepsilon^{k+1}, \\ \lambda^{k+1} &= \lambda^k + \mathbb{C}^k : (\varepsilon^{k+1} - e^{k+1}). \end{aligned} \quad (5)$$

involving a damping (or relaxation) factor $\gamma \in [0, 1)$ and a reference material \mathbb{C}^k that may change from one iteration to the next, as for the basic scheme (2). We consider mixed “boundary conditions”⁴⁴ and $\mathbb{D}^k = (\mathbb{C}^k)^{-1}$. More precisely, the mixed boundary conditions are given in terms of a pair (\mathbb{P}, \mathbb{Q}) of complementary and orthogonal projectors and prescribed partial strains $\bar{\varepsilon}$ and stresses $\bar{\sigma}$ which satisfy the compatibility conditions

$$\mathbb{P} : \bar{\varepsilon} = \bar{\varepsilon} \quad \text{and} \quad \mathbb{Q} : \bar{\sigma} = \bar{\sigma}.$$

The algorithm (5) represents a (more or less trivial) extension of the one considered by Moulinec and Silva^{24(eq.13)} to nonlinear material behavior and mixed boundary conditions, with the parameter conversion $\alpha = 2(1 - \gamma)$ and $\beta = \alpha$. By a similar reasoning as in Moulinec and Silva,²⁴ it is not difficult to see that the algorithm (5) reduces to the classical ADMM scheme (4) for $\gamma = \frac{1}{2}$ and $\mathbb{C}^k \equiv \mathbb{C}^0$. The damping coefficient $\gamma \in [0, 1)$ allows to represent, both, the Eyre–Milton method²⁵ and the Monchiet–Bonnet scheme²³ as special cases, see Monchiet and Bonnet²³ and Moulinec and Silva²⁴. More generally, the abstract nonlinear polarization schemes considered in Schneider et al.²⁸ are included. The reason for considering the form (5) rather than the polarization form²⁸ is that the former allows for adaptively changing reference material in a more natural way. Indeed, for fixed strain field ε^* , but changing reference material \mathbb{C}^k , the polarization field $p^k = \sigma(\varepsilon^*) + \mathbb{C}^k : \varepsilon^*$ changes, as well.

How to choose the reference material \mathbb{C}^k in the polarization scheme (5) is treated in Section 2. We discuss theoretical convergence results, available in the literature, specialized to our case. We also present general strategies for selecting the reference material in ADMM, including the classical residual-balancing strategy,⁴⁵ a procedure⁴⁶ developed for the non-stationary Douglas–Rachford scheme, and a Barzilai–Borwein stepsize.^{47,48} In Section 3, we discuss the efficient

implementation of the non-stationary polarization method (5), and also provide details on evaluating the residual during the iterations. In particular, comparability to strain-based solvers and the polarization schemes²⁸ is ensured. Section 4 contains computational examples, evaluating the benefits of non-stationary polarization methods, but also demonstrating their limitations.

2 | CONVERGENCE AND ADAPTIVE CHOICE OF THE REFERENCE MATERIAL

2.1 | Convergence statements

In this section, we provide a guide to convergence assertions for the ADMM (5) with adaptive relaxation parameter $\gamma_k \in [0, 1)$ and a reference material $\mathbb{C}^k = \alpha_k \text{Id}$ proportional to the identity, that is, a linear elastic reference material with vanishing Poisson's ratio and shear modulus $\mu_k = \alpha_k/2$. The results may be extended to truly isotropic reference materials provided the phases are isotropic as well. For the work at hand, we restrict to the choice $\mathbb{C}^k = \alpha_k \text{Id}$.

We assume that the stress operator σ arises from a (possibly condensed) free energy density w which we assume convex and continuously differentiable in the strain variable, that is,

$$\sigma(x, \varepsilon(x)) = \frac{\partial w}{\partial \varepsilon}(x, \varepsilon(x)), \quad x \in Q,$$

is assumed to hold for any square-integrable strain field. Furthermore, we suppose that w satisfies a growth condition,^{49(sec.2.1)} s.t. the optimization problem (1) admits a solution.

For the general case, suppose that the relaxation parameters γ_k and the coefficients α_k satisfy the conditions:

1. The relaxation parameter is restricted as follows: $0 < \gamma_k \leq 1/2$ for all k .
2. There is some $\alpha_\infty > 0$, s.t. $\alpha_k \geq \alpha_\infty$ for some k .
3. The summability condition

$$\sum_{k=1}^{\infty} \frac{1 - \gamma_k}{\gamma_k} \max \left\{ \frac{\alpha_k^2}{\alpha_{k-1}^2}, 0 \right\} < \infty$$

holds.

Then, for any initial point, the ADMM algorithm (5) converges in the sense that $\varepsilon^k - e^k$ goes to zero in L^2 and the effective free energy at e^k converges to a minimum. We refer to Xu et al.^{48(sec.5)} for a proof. Let us briefly comment on the conditions. The first condition excludes the case $\gamma_k = 0$, that is, damping is *necessary* for the general case. Condition 2 ensures that the shear moduli of the reference materials are bounded. Interpreted in terms of a step size,^{28(sec.3)} this ensures that the step size does not go to infinity. Condition 3 ensures that decreases in the reference constant α_k have a sufficiently fast decay. There is a version for increasing α_k ,^{48(Ass.2)} but this is not relevant for our problem. Indeed, the linear elastic stiffness typically provides a (pessimistic) upper bound for the reference material.

Please note that, for this general case, only a slow convergence rate can be expected. Indeed, the considered scenario covers also degenerate cases such as porous materials with fractal interfaces. For the general case, and fixed parameters, ADMM converges with a $1/k$ -rate.^{50,51} This rate is similar to what may be achieved by projected gradient descent.⁴⁹

After treating the general case, let us consider the best-case scenario.^{28,29} More precisely, suppose that there are positive constants α_- and α_+ , s.t. the stress operator σ is uniformly α_- -strongly convex

$$[\sigma(x, \varepsilon_1) - \sigma(x, \varepsilon_2)] : (\varepsilon_1 - \varepsilon_2) \geq \alpha_- \|\varepsilon_1 - \varepsilon_2\|^2, \quad x \in Q, \quad \varepsilon_1, \varepsilon_2 \in \text{Sym}(d), \quad (6)$$

and α_+ -Lipschitz-continuous

$$\|\sigma(x, \varepsilon_1) - \sigma(x, \varepsilon_2)\| \leq \alpha_+ \|\varepsilon_1 - \varepsilon_2\|, \quad x \in Q, \quad \varepsilon_1, \varepsilon_2 \in \text{Sym}(d), \quad (7)$$

for the Frobenius inner product $\|\varepsilon\| = \sqrt{\varepsilon : \varepsilon}$ on $\text{Sym}(d)$. Then, for any reference material $\mathbb{C}^k \equiv \mathbb{C}^0 = \alpha_0 \text{Id}$ with $\alpha_0 > 0$ and relaxation parameter $\gamma_k \equiv \gamma \in [0, 1)$, the iterates $(\varepsilon^k, e^k, \lambda^k)$ produced by the ADMM algorithm (5) converge to the

(unique) fixed point $(\varepsilon^*, \varepsilon^*, \sigma(\varepsilon^*))$. Moreover, the estimate

$$\|P^{k+1} - P^*\|_{L^2} \leq \left(\gamma + (1 - \gamma) \max \left\{ \frac{\alpha_+ - \alpha_0}{\alpha_+ + \alpha_0}, \frac{\alpha_0 - \alpha_-}{\alpha_0 + \alpha_-} \right\} \right) \|P^k - P^*\|_{L^2} \quad (8)$$

holds for the stress polarizations

$$P^k = \lambda^k + \alpha_0 e^k \quad \text{and} \quad P^* = \sigma(\varepsilon^*) + \alpha_0 \varepsilon^*.$$

The fastest convergence rate is achieved for $\gamma = 0$ and $\alpha_0 = \sqrt{\alpha_- \alpha_+}$. In particular, for fixed precision, the number of necessary iterations is proportional to the square root of the material contrast $\sqrt{\alpha_+ / \alpha_-}$. The linear elastic case was discussed by Moulinec and Silva.²⁴

Please keep in mind that the choice $\gamma = 0$ and $\alpha_0 = \sqrt{\alpha_- \alpha_+}$ leads to the fastest overall convergence rate. This does not necessarily correspond to the minimum iterations for prescribed tolerance. Rather, the rationale behind optimizing the convergence rate is that, for high accuracy, the number of required iterations depends on the convergence rate, in the first place.

The contraction estimate (8) was established for the Douglas–Rachford rewriting⁵² of the ADMM, which is more convenient to analyze. The assertion (8) reveals a few things. For a start, zero damping $\gamma = 0$, is permitted, and even favorable. This possibility was excluded for the general case treated earlier. In computational practice, however, small positive values of γ are typically preferred over $\gamma = 0$.

Secondly, by the same reasoning as for obtaining the estimate (8), it is possible to derive the bound

$$\|\lambda^{k+1} - \lambda^k + \alpha_k (e^{k+1} - e^k)\|_{L^2} \leq \left(\gamma + (1 - \gamma) \max \left\{ \frac{\alpha_+ - \alpha_k}{\alpha_+ + \alpha_k}, \frac{\alpha_k - \alpha_-}{\alpha_k + \alpha_-} \right\} \right) \|\lambda^k - \lambda^{k-1} + \alpha_k (e^k - e^{k-1})\|_{L^2}.$$

Thus, as long as the sequence α_k is convergent to $\alpha^* \in [0, 1)$, the adaptive ADMM converges linearly, as well.

The third insight from the estimate (8) is that a clever choice of the reference constant α_0 is crucial. For instance, inserting the upper bound $\alpha_0 = \alpha_+$ into the estimate (8) with $\gamma = 0$ produces

$$\|P^{k+1} - P^*\|_{L^2} \leq \frac{\alpha_+ - \alpha_-}{\alpha_+ + \alpha_-} \|P^k - P^*\|_{L^2}.$$

For fixed tolerance, the latter estimate implies that the necessary iteration count is proportional to the material contrast α_+ / α_- . Thus, no advantage is gained over the basic scheme.¹⁹ In particular, due to the character of the optimum reference constant

$$\alpha_0 = \sqrt{\alpha_- \alpha_+}, \quad (9)$$

expressed as a *geometric* mean of the constants α_{\pm} , a good guess for the lower bound α_- is critical. In contrast, a good upper bound α_+ is usually inferred from the *elastic* stiffness of the material.

To conclude this section, let us briefly comment on the case of porous microstructures, where the stress operator σ is identically zero on a non-trivial subdomain of the unit cell Q . Then, the stress operator is no longer strongly monotone, and the Giselsson–Boyd estimate (8) no longer applies. The general convergence statement discussed at the beginning of this section ensures convergence, at least in theory. However, this convergence behavior can be excruciatingly slow in practice, even leading to divergence in case of accumulated round-off errors. This is the reason for debates in the community concerning whether polarization schemes converge for porous materials or not.^{23,24,39}

Recently, the convergence behavior of primal algorithms for porous microstructures was analyzed.⁴⁹ It was found, both computationally and numerically, that the discretization scheme plays an important role in the stability of such numerical methods for porous microstructures. Furthermore, in the continuous (non-discretized) case, the convergence behavior of the basic scheme (and its descendants) is closely tied to a specific subspace of compatible strain fields, on which the stress operator is strongly monotone. The monotonicity constant, however, strongly depends on the geometry of the solid-pore interface, and is not readily accessible.

Without going into details, it is readily seen from the iterative scheme (5) that the strain field $\varepsilon^{k+\frac{1}{2}}$ is always compatible, and even lives in the specific subspace on which the stress operator is strongly monotone. Unfortunately, in the ADMM

scheme, the stress operator is only applied to the incompatible strain field e via $\lambda = \sigma(e)$, and the non-degeneracy of the stress operator σ cannot be transferred to the compatible strains, invoking a rigorous convergence statement. However, it seems reasonable that, if the residual $\left\| \varepsilon^{k+\frac{1}{2}} - e^k \right\|_{L^2}$ is sufficiently small (and a stable discretization is used), then ADMM (5) will enter a regime of linear convergence. This statement can be made rigorous, see Liang et al.⁵³ We will further study this issue numerically in Section 4.

The plan for the remainder of this section is to discuss a number of strategies for adaptively estimating the reference constant α_k , fixing the damping factor $\gamma_k \equiv \gamma$.

2.2 | Finite material contrast

Let us start with the simplest case of (non-degenerate) linear elastic materials.^{23,24,54} Denote by $\alpha_{\pm}(x)$ the smallest and largest eigenvalue of the local stiffness tensor $\mathbb{C}(x)$ at the point $x \in Q$. Then, we may compute

$$\alpha_- = \min_{x \in Q} \alpha_-(x) \quad \text{as well as} \quad \alpha_+ = \max_{x \in Q} \alpha_+(x),$$

and choose $\alpha_0 = \sqrt{\alpha_- \alpha_+}$. For the case of non-linear materials at small strains, the following strategy proves useful.²⁸ Denote by $\alpha_{k,\pm}(x)$ the smallest and the largest eigenvalue of the material tangent

$$\frac{\partial^2 w}{\partial^2 \varepsilon}(x, e^k(x)), \quad x \in Q,$$

evaluated for the current strain field e^k at iteration k . Then, we may set

$$\alpha_{k,-} = \min_{x \in Q} \{ \min \alpha_{k,-}(x), \alpha_{k-1,-} \} \quad \text{as well as} \quad \alpha_{k,+} = \max_{x \in Q} \{ \max \alpha_{k,+}(x), \alpha_{k-1,+} \}$$

and define $\alpha_{k+1} = \sqrt{\alpha_{k,-} \alpha_{k,+}}$. Typically, $\alpha_{k,+}$ is independent of k , and $\alpha_{k,-}$ is a monotonically decreasing sequence, compare condition 3 in Section 2.1.

This approach comes with the advantage that, for inelastic materials featuring an elastic region, loadings at small strain lead to convergence rates in the polarization scheme like for the elastic case. Clearly, care has to be taken for materials whose material contrast goes to infinity for increased loading^{28(sec.7)}.

2.3 | Residual balancing

A classical approach, introduced by He et al.,⁴⁵ for adapting the step size in ADMM is residual balancing. For the case at hand (5) and unrelaxed ADMM $\gamma = 1/2$, we define the primal

$$p^{k+1} = \varepsilon^{k+1} - e^{k+1} \quad (10)$$

and the dual residual

$$d^{k+1} = \alpha_k (e^{k+1} - e^k). \quad (11)$$

The letters p and d were chosen to stand for primal and dual, respectively. In particular, no confusion with the stress polarization is intended. The quantity p^{k+1} (10) measures the feasibility condition $\varepsilon \stackrel{!}{=} e$ in the primal form of the optimization problem (1), whereas the increment d^{k+1} (11) may be interpreted as quantifying feasibility for the Lagrangian dual of problem (1), see Boyd et al.^{22(sec.3.3)}

In our case, please note the following. For a start, algebraic manipulations reveal the estimates

$$\left\| \bar{\varepsilon} - \mathbb{P} : \langle e^{k+1} \rangle_Q \right\|_{L^2} + \left\| (\text{Id} - \Gamma - \langle \cdot \rangle_Q) : e^{k+1} \right\|_{L^2} \leq \| p^{k+1} \|_{L^2} \quad (12)$$

and

$$\|\bar{\sigma} - \mathbb{Q} : \langle \lambda^{k+1} \rangle_Q\|_{L^2} + \|\Gamma : \lambda^{k+1}\|_{L^2} \leq \|d^{k+1}\|_{L^2}, \quad (13)$$

where $\Gamma \equiv \Gamma^k : \mathbb{C}^k$ is independent of α_k . We refer to Appendix A for a derivation. Thus, the primal and the dual residual measure the compatibility and the strain boundary condition as well as the equilibrium and the stress boundary condition *individually*. Also, due to the identity

$$p^{k+1} + d^{k+1}/\alpha_k = \varepsilon^{k+1} - e^k,$$

the primal and dual residuals may be interpreted as a decomposition of the total (consistent) residual (22), to be discussed below in Section 3.

He et al.⁴⁵ suggested that the primal and dual residuals should decay at the same rate in order to ensure fast convergence of ADMM, and to adapt the step size α_k if necessary. Some care has to be taken, as the primal residual (10) is a strain field and the dual residual (11) is a stress field. Due to the mismatch of dimensions, no direct comparison is possible. Rather, it is necessary to define a suitable interval of salient stress values $[\lambda_{\min}, \lambda_{\max}]$ that is considered appropriate and to ensure the condition

$$\frac{\|d^{k+1}\|_{L^2}}{\|p^{k+1}\|_{L^2}} \in [\lambda_{\min}, \lambda_{\max}]$$

directly via an amplification factor $A > 1$ in terms of the adaptive rule

$$\alpha_{k+1} = \begin{cases} \alpha_k/A, & \|d^{k+1}\|_{L^2} > \lambda_{\max} \|p^{k+1}\|_{L^2}, \\ A \alpha_k, & \|d^{k+1}\|_{L^2} < \lambda_{\min} \|p^{k+1}\|_{L^2}, \\ \alpha_k, & \text{otherwise.} \end{cases}$$

The rationale behind this strategy is that, if the primal residual is large, increasing the penalty factor α_k should lead to a decrease of the term penalizing the difference $\varepsilon - e$.

Typical values for parameters are $A = 2$, $\lambda_{\max} = \alpha_+$ and $\lambda_{\min} = \alpha_+/10^{22}$ (sec.3.4.1). The resulting strategy is called residual balancing. Computing the primal and dual residuals may be integrated into Algorithm 1 without memory overhead.

The original residual-balancing strategy of He et al.⁴⁵ was developed for $\gamma = 1/2$. We discuss a simple modification for $\gamma_k \equiv \gamma > 0$. Then, defining the primal and dual residuals by

$$p^{k+1} = \varepsilon^{k+\frac{1}{2}} - e^{k+1} \quad \text{and} \quad d^{k+1} = \alpha_k \left(e^{k+1} - e^k + \varepsilon^{k+\frac{1}{2}} - \varepsilon^{k+1} \right), \quad (14)$$

the identity

$$p^{k+1} + d^{k+1}/\alpha_k = 2\gamma \left(\varepsilon^{k+\frac{1}{2}} - e^k \right), \quad (15)$$

as well as the estimates (12) and (13) follow, see Appendix A. The residuals for $\gamma > 0$ may also be computed without additional memory overhead.

Please note the explicit expression for the dual residual

$$d^{k+1}/\alpha_k = e^{k+1} - (1 - 2\gamma) \varepsilon^{k+\frac{1}{2}} - 2\gamma e^k.$$

In particular, for $\gamma \equiv 0$, the identity $d^{k+1} = \alpha_k p^{k+1}$ holds, and the residual balancing strategy does not work.

2.4 | Lorenz–Tran-Dinh updating

Lorenz and Tran-Dinh⁴⁶ studied adaptive step-size strategies for the Douglas–Rachford splitting and the ADMM. They provided a novel convergence analysis for operator-splitting methods of Douglas–Rachford type with adaptive parameter

with absolutely summable increments. Based on a careful analysis of the linear case, they proposed a strategy which is valid for possibly multi-valued maximally monotone operators. For the iterative scheme (5) at hand, their suggestion takes the form

$$\alpha_k = \frac{\|\lambda^k\|_{L^2}}{\|e^k\|_{L^2}} \quad (16)$$

of a secant stiffness. For linear elasticity, the proposed reference material α_k is guaranteed to lie within the minimum and the maximum eigenvalues of the local stiffnesses.

2.5 | Barzilai–Borwein step size

Barzilai and Borwein⁴³ introduced an adaptive strategy for selecting the step size s^k in gradient descent methods

$$x^{k+1} = x^k - s^k \nabla f(x^k)$$

via the secant-type rule

$$s^{k+1} = \frac{\|x^{k+1} - x^k\|^2}{\langle \nabla f(x^{k+1}) - \nabla f(x^k), x^{k+1} - x^k \rangle}, \quad (17)$$

where angular brackets denote the inner product. The Barzilai–Borwein method turned out to be a powerful solution method for unconstrained smooth optimization,^{55–57} despite its apparent simplicity. The Barzilai–Borwein (BB) step-size rule (17) was shown to be successful in FFT-based computational micromechanics,³⁹ as well.

For the ADMM, Xu et al.^{47,48} studied step-size selection strategies of Barzilai–Borwein type. In this work, we consider a variation of the approach of Xu et al. More precisely, Xu et al. propose a BB strategy for the primal *and* the dual step size, and select the geometric mean of these two step sizes. For homogenization problems, it turns out to be more effective to rely upon a purely primal strategy. More precisely, the starting point of our discussion is the BB rule

$$\alpha_{k+1} = \frac{\langle (\lambda^{k+1} - \lambda^k) : (e^{k+1} - e^k) \rangle_Q}{\langle \|e^{k+1} - e^k\|^2 \rangle_Q}. \quad (18)$$

As the fields λ and e are related by the stress operator σ , we observe

$$\alpha_{k+1} = \frac{\langle (\sigma(e^{k+1}) - \sigma(e^k)) : (e^{k+1} - e^k) \rangle_Q}{\langle \|e^{k+1} - e^k\|^2 \rangle_Q}.$$

Thus, if the stress operator σ is α_- -monotone (6) and α_+ -Lipschitz (7), the inequalities

$$\alpha_- \leq \alpha_{k+1} \leq \alpha_+ \quad (19)$$

follow. Please note that Barzilai–Borwein⁴³ also introduced a second step-size rule, which, in our case, becomes

$$\tilde{\alpha}_{k+1} = \frac{\langle \|\lambda^{k+1} - \lambda^k\|^2 \rangle_Q}{\langle (\lambda^{k+1} - \lambda^k) : (e^{k+1} - e^k) \rangle_Q}.$$

The Cauchy–Schwarz inequality implies the ordering

$$\alpha_{k+1} \leq \tilde{\alpha}_{k+1}.$$

Thus, the second strategy $\tilde{\alpha}_{k+1}$ is more conservative than the first strategy (18). Quite often, the second strategy is slower in practice, and we will work with the first choice (18).

Please note that the bounds (19) are only effective for strongly convex stress operators. Indeed, if $\alpha_- = 0$, for instance in the case of porous microstructures, the reference material may decay to zero, leading to an unstable, and eventually diverging scheme. For this purpose, we augment the rule (18) by a safeguard proposed by Xu et al.^{47(sec.3.3)} For $\delta \in (0, 1)$, we define

$$\alpha_{k+1} = \begin{cases} \frac{\langle (\lambda^{k+1} - \lambda^k) : (e^{k+1} - e^k) \rangle_Q}{\langle \|e^{k+1} - e^k\|^2 \rangle_Q}, & \langle (\lambda^{k+1} - \lambda^k) : (e^{k+1} - e^k) \rangle_Q \geq \delta \|e^{k+1} - e^k\|_{L^2} \|\lambda^{k+1} - \lambda^k\|_{L^2}, \\ \alpha_k, & \text{otherwise.} \end{cases} \quad (20)$$

This safeguard ensures that the iteration-dependent lower bound

$$\alpha_{k+1} \geq \delta \frac{\|\lambda^{k+1} - \lambda^k\|_{L^2}}{\|e^{k+1} - e^k\|_{L^2}}$$

holds, which may be connected to the interface constant c_I .^{49(sec.2.2)} Please note that the strategy (20) may be integrated into the ADMM Algorithm 1 without memory overhead. Last but not least let us comment on the selection of δ . We follow the suggestion of Xu et al.⁴⁷ and fix $\delta = 0.2$.

3 | RESIDUALS AND IMPLEMENTATION

3.1 | The convergence criterion

In this section, we assume, for simplicity, that the reference materials \mathbb{C}^k commute with the projector \mathbb{P} . Otherwise, the formulae become lengthy.⁴⁴ To assess convergence, we utilize the \mathbb{C}^k - and \mathbb{D}^k -weighted L^2 -norms

$$\|\varepsilon\|_{\mathbb{C}^k} = \sqrt{\langle \varepsilon : \mathbb{C}^k : \varepsilon \rangle_Q} \quad \text{and} \quad \|\lambda\|_{\mathbb{D}^k} = \sqrt{\langle \lambda : \mathbb{D}^k : \lambda \rangle_Q}.$$

Please notice that, in view of the identification of the ADMM (5) with a damped Eyre–Milton scheme, see Moulinec and Silva²⁴, via the identification $P^k = \lambda^k + \mathbb{C}^k : e^k$, and as the identity $\lambda^k = \sigma(e^k)$ holds, it appears natural to consider the field e^k as the primary quantity on which the ADMM scheme (5) iterates. In particular, we should quantify the deviation from the following equations

$$\begin{aligned} \Gamma^k : \lambda^k & \stackrel{!}{=} 0 && \text{(equilibrium)} \\ (\text{Id} - \Gamma^k : \mathbb{C}^k - \langle \cdot \rangle_Q) : e^k & \stackrel{!}{=} 0 && \text{(kinematic compatibility)} \\ \mathbb{P} : \langle e^k \rangle_Q & \stackrel{!}{=} \bar{\varepsilon} && \text{(strain boundary condition)} \\ \mathbb{Q} : \langle \lambda^k \rangle_Q & \stackrel{!}{=} \bar{\sigma} && \text{(stress boundary condition)} \end{aligned} \quad (21)$$

Appendix B details the derivation of the identity

$$\begin{aligned} \|e^k - \varepsilon^{k+\frac{1}{2}}\|_{\mathbb{C}^k}^2 &= \|\Gamma^k : \lambda^k\|_{\mathbb{C}^k}^2 + \|(\text{Id} - \Gamma^k : \mathbb{C}^k - \langle \cdot \rangle_Q) : e^k\|_{\mathbb{C}^k}^2 \\ &+ \|\mathbb{P} : \langle e^k \rangle_Q - \bar{\varepsilon}\|_{\mathbb{C}^k}^2 + \|\bar{\sigma} - : \mathbb{Q} : \langle \lambda^k \rangle_Q\|_{\mathbb{D}^k}^2. \end{aligned} \quad (22)$$

Thus, we are led to the surprisingly simple result that the four different conditions (21) to be satisfied may be quantified by monitoring the quantity $e^k - \varepsilon^{k+\frac{1}{2}}$ only. Several remarks are in order.

1. For the basic scheme with mixed boundary conditions

$$\varepsilon^{k+1} = \bar{\varepsilon} + \mathbb{D}^k : \bar{\sigma} - (\Gamma^k + \mathbb{D}^k : \mathbb{Q} : \langle \cdot \rangle_Q) : (\sigma(\varepsilon^k) - \mathbb{C}^k : \varepsilon^k), \tag{23}$$

the identity

$$\|\varepsilon^{k+1} - \varepsilon^k\|_{\mathbb{C}^k}^2 = \|\Gamma^k : \sigma(\varepsilon^k)\|_{\mathbb{C}^k}^2 + \|\bar{\sigma} - : \mathbb{Q} : \langle \sigma(\varepsilon^k) \rangle_Q\|_{\mathbb{D}^k}^2 \tag{24}$$

holds.^{28(eq.5.7)} As each iterate ε^k ($k > 0$) of the basic scheme (23) is kinematically compatible and satisfies the \mathbb{P} -part of the strain boundary condition, the ADMM residual (22) appears as a natural extension of the residual (24) for the basic scheme. In fact, a convergence criterion implemented for the basic scheme may be reused for the ADMM, surprisingly.

2. For the pure ADMM case ($\gamma = 1/2$), Michel et al.^{19(sec.4.1.2)} suggested to monitor $\|e^k - \varepsilon^k\|$ and $\|\lambda^k - \sigma(\varepsilon^k)\|$. If the stress operator σ is Lipschitz continuous with constant L , we observe

$$\|\lambda^k - \sigma(\varepsilon^k)\| = \|\sigma(e^k) - \sigma(\varepsilon^k)\| \leq L \|e^k - \varepsilon^k\|.$$

Thus, monitoring the quantity $\lambda^k - \sigma(\varepsilon^k)$ is not necessary and may be avoided for performance reasons.

3. For non-changing reference material $\mathbb{C}^k \equiv \mathbb{C}^0$, the "natural" convergence criterion for the damped Eyre–Milton iteration²⁸

$$P^{k+1} = \gamma P^k + (1 - \gamma) [2\mathbb{C}^0 : \bar{\varepsilon} + 2\bar{\sigma} + Y^0 : Z^0(P^k)]$$

with mixed boundary conditions, involving the operators

$$Y^0 = \text{Id} - 2 (\mathbb{C}^0 : \Gamma^0 + \mathbb{Q} : \langle \cdot \rangle_Q) \quad \text{and} \quad Z^0 = \text{Id} - 2 \mathbb{C}^0 : (\sigma + \sigma^0)^{-1},$$

where $\sigma^0(\varepsilon) \equiv \mathbb{C}^0 : \varepsilon$, is motivated by the identity^{28(eq.5.8)}

$$\begin{aligned} \frac{1}{4(1 - \gamma)^2} \|P^{k+1} - P^k\|_{\mathbb{C}^0}^2 &= \|\Gamma^0 : \lambda^k\|_{\mathbb{C}^0}^2 + \|(\text{Id} - \Gamma^0 : \mathbb{C}^0 - \langle \cdot \rangle_Q) : e^k\|_{\mathbb{C}^0}^2 \\ &\quad + \|\mathbb{P} : \langle e^k \rangle_Q - \bar{\varepsilon}\|_{\mathbb{C}^0}^2 + \|\bar{\sigma} - : \mathbb{Q} : \langle \lambda^k \rangle_Q\|_{\mathbb{D}^0}^2 \end{aligned} \tag{25}$$

for the identification $P^k = \lambda^k + \mathbb{C}^0 : e^k$ with $\lambda^k = \sigma(e^k)$. Comparing to the expression for the ADMM (22), we notice the identity

$$\frac{1}{4(1 - \gamma)^2} \|P^{k+1} - P^k\|_{\mathbb{C}^0}^2 = \|e^k - \varepsilon^{k+\frac{1}{2}}\|_{\mathbb{C}^0}^2.$$

Thus, the residual for the ADMM is in fact *identical* to the residual for the damped Eyre–Milton method for non-changing reference material. Also, the established identity (22) may be considered as an extension of the identity (25), valid for polarization schemes, to non-stationary reference material.

3.2 | Efficient implementation

The algorithm (5) for mixed boundary conditions is straightforward to implement on three strain-like fields, see Algorithm 1. Thus, the damping parameter γ , whose clever choice may significantly decrease the time to solution, does not lead to additional memory requirements compared to the original version of Michel et al.¹⁹ The Algorithm 1 also indicates suitable instances when to update the reference material and when to compute the residual. Notice that no extra memory is required for evaluating the residual.

For suitable initialization and stationary reference material, the Algorithm 1 produces identical iterates (and identical residuals) as Alg. 1 in Schneider et al.²⁸ However, the implementation of Algorithm 1 requires an additional field to be stored. Still, the original implementation¹⁹ appears to be the most popular among practitioners of FFT-based methods, and is extended to Algorithm 1 with minimal effort.

Algorithm 1. ADMM with damping parameter $\gamma \in [0, 1)$

```

1: initialize  $\varepsilon, e, \lambda$  and  $\mathbb{C}^0$ 
2: repeat
3:    $\varepsilon \leftarrow \lambda - \mathbb{C}^0 : e$ 
4:    $\hat{\varepsilon} \leftarrow \text{DFT}(\varepsilon)$ 
5:    $\hat{\varepsilon}(\xi) \leftarrow -\hat{\Gamma}^0(\xi) : \hat{\varepsilon}(\xi), \quad \xi \neq 0$  ▷ Use your favorite  $\Gamma$ -operator
6:    $\hat{\varepsilon}(0) \leftarrow \bar{\varepsilon} + \mathbb{D}^0 : (\bar{\sigma} - \mathbb{Q} : \hat{\varepsilon}(0))$  ▷ Fix average strain
7:    $\varepsilon \leftarrow \text{DFT}^{-1}(\hat{\varepsilon})$ 
8:    $\bar{\lambda} \leftarrow \langle \lambda \rangle_Q$ 
9:    $\text{res} \leftarrow \|\varepsilon - e\|_{\mathbb{C}^0} / \|\bar{\lambda}\|_{\mathbb{D}^0}$ 
10:   $\varepsilon \leftarrow 2(1 - \gamma)\varepsilon - (1 - 2\gamma)e$ 
11:   $e \leftarrow (\sigma + \sigma^0)^{-1}(\lambda + \mathbb{C}^0 : \varepsilon)$  ▷ See Schneider et al.28(sec.6)
12:   $\lambda \leftarrow \lambda + \mathbb{C}^0 : (\varepsilon - e)$  ▷ Or as a byproduct in line 11
13:  Update reference material  $\mathbb{C}^0$ 
14: until  $\text{res} < \text{tol}$  ▷ Compare section~3.1
15: return  $(\varepsilon, \lambda)$ 

```

To complete the section on the implementation, let us specify that we use a discretization on a staggered grid,³⁶ see Schneider^{2(sec.2.5)} for the efficient evaluation of the Γ -operator. Please note, however, that the implementation is suitable for alternative discretizations, as well, see the overview article.^{2(sec.2)}

4 | COMPUTATIONAL EXAMPLES

4.1 | Setup

The methods described in this article were integrated into an in-house FFT-based computational micromechanics code (written in Python with Cython extensions) which is parallelized using OpenMP. All examples in this section are solved up to a tolerance of $\text{tol} = 10^{-5}$, measured in terms of the corresponding consistent convergence criteria. The run times refer to a desktop computer with six 3.70 GHz cores and 32 GB RAM. The different material parameters used in this article are summarized in Table 1, and we refer to Figure 1 for an impression on the considered microstructures.

4.2 | A short-fiber reinforced composite

Our first example concerns a short-fiber reinforced polymer. Such a test is a standard benchmark in FFT-based computational micromechanics because it permits studying finite material contrast, but complex microstructure features.

More precisely, we consider short glass fibers with a length of 200 μm and a diameter of 10 μm at 15% volume fraction inside a cubic volume element with edge length of 288 μm and an isotropic fiber orientation. The microstructure, see Figure 1A, was generated with the SAM algorithm⁶⁰ and comprises 229 fibers. We use a rasterized image with 192³ voxels, that is, a mesh size $h = 1.5 \mu\text{m}$.

TABLE 1 Material parameters used in this article, supplemented by their source

Material	Source	E in GPa	ν	Further parameters	
E-glass fibers	58	72	0.22		
Polymer matrix	58	3.0	0.35	$\sigma_Y = 20 \text{ MPa}$ $k_2 = 15 \text{ MPa}$	$k_1 = 1 \text{ MPa}$ $m = 150$
Ceramic particles	59	400	0.2		
Aluminum matrix	59	75	0.3	$\sigma_Y = 75 \text{ MPa}$ $m = 0.3895$	$k = 416 \text{ MPa}$

We furnish the phases with the material parameters of the polymer matrix and the E-glass fibers listed in Table 1. As a warm-up, we performed simulations in the linear elastic setting, see Table 2, for 5% uni-axial extension in x -direction. We study the performance of ADMM with different relaxation factors γ , corresponding to the accelerated scheme of Michel et al.¹⁹ for $\gamma = 1/2$, the Monchiet–Bonnet recommendation²⁷ $\gamma = 1/4$ and the Eyre–Milton scheme²⁵ for $\gamma = 0$. We observe that, for all adaptive choices of the reference material, the iteration count increases with γ , that is, in order to minimize the iteration count, setting $\gamma = 0$ is advised. In general, the optimum choice, see Section 2.2, which is available due to the finite material contrast, performs best. All other adaptive strategies are *initialized* with this theoretically optimum choice. Residual balancing, for instance, retains this optimum choice. Thus, its iteration counts are identical to those of the optimum choice (except for $\gamma = 1/2$ where the reference material is changed once, and directly changed back after the next iteration). Please also keep in mind that residual balancing only works for $\gamma \neq 0$.

In addition, we also considered the choice $\alpha_0 = \alpha_+/5$ which performs well in Monchiet–Bonnet.²⁷(fig.4). Indeed, the performance of the basic scheme and the ADMM will be quite similar if identical reference materials are chosen. Rather, the basic scheme is restricted by the condition $\alpha_0 > \alpha_+/2$ to retain stability,¹⁹ whereas no such limitation applies to the polarization schemes. Therefore, $\alpha_0 = \alpha_+/5$ represents a rather aggressive choice for the reference material in the polarization scheme. For the linear elastic example at hand, this choice leads to slightly slower iteration counts than the Lorenz–Tran–Dinh scaling.

Both, the Lorenz–Tran–Dinh and the Barzilai–Borwein strategies perform worse than the optimum choice, but not dramatically so. Compared to the primal solvers, where we include the "golden standard", the linear CG, and the primal Barzilai–Borwein method, ADMM performs admirably. For optimum parameter choice, the required iteration count lies below CG. These findings are illustrated in Figure 2A which shows the residual versus current iteration for the different strategies and the damping factor $\gamma = 0$. We observe that the convergence rate of ADMM with optimum parameters matches the rate of linear CG. Both the primal Barzilai–Borwein method and the Barzilai–Borwein strategy for the ADMM, lead to a non-monotone behavior of the residual. However, these oscillations are less pronounced for the primal-dual version. The Lorenz–Tran–Dinh strategy is the fastest for low accuracy, up to 10^{-3} , but experiences a decrease in convergence rate for higher accuracy.

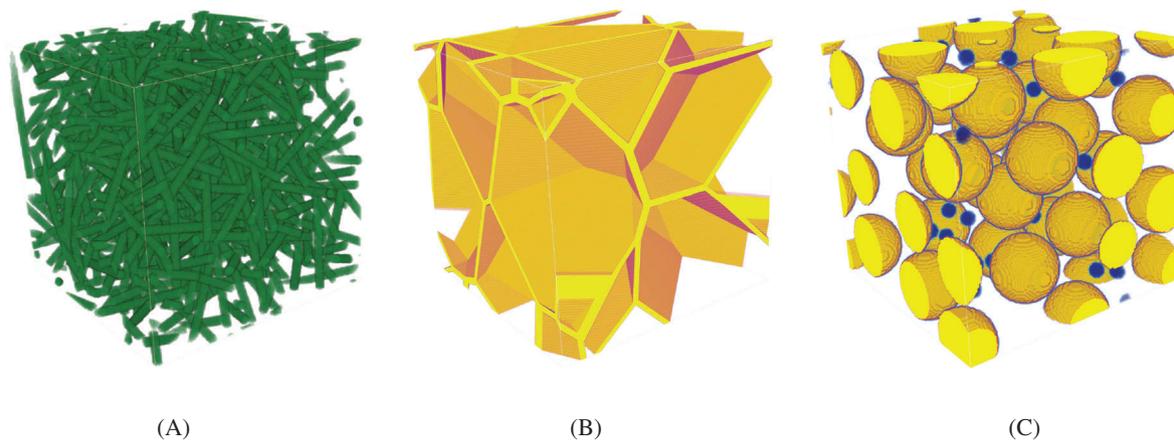


FIGURE 1 Microstructures considered in this article. (A) Short-fiber composite.⁶⁰ (B) Closed-cell foam.⁶¹ (C) Metal-matrix composite⁵⁹

TABLE 2 Iteration counts for the ADMM with different settings and the fiber composite with linear elastic constituents

ADMM	Optimum	$\alpha_+/5$	Residual balancing	Lorenz–Tran–Dinh	Barzilai–Borwein
$\gamma = 0$	32	47	–	40	41
$\gamma = 1/4$	43	61	43	55	55
$\gamma = 1/2$	66	94	69	85	73
Primal solvers					
BB	47				
CG	36				

Note: The primal solvers are shown for comparison. The bold values correspond to the lowest iteration count for the primal-dual and the primal solvers (hence on each).

Next, we study the inelastic behavior. Using the material parameters of Table 1 for the von-Mises elastoplastic matrix with exponential-linear hardening, we subject the composite to 5% uni-axial strain loading, distributed over 50 load steps. The mean iteration counts and the run times are listed in Table 3. Except for residual balancing, increasing the damping factor γ also leads to an increase of the iteration count for the ADMM. The Lorenz–Tran-Dinh strategy is significantly faster than the theoretically optimum choice. Also, the Barzilai–Borwein strategy almost matches the speed of the optimum choice. Residual balancing, on the other hand, does not appear competitive. The constant strategy $\alpha_0 = \alpha_+/5$ takes about twice as long as the optimum strategy.

Compared to the primal solvers, where we include the Barzilai–Borwein method,³⁹ nonlinear CG,⁷ and Newton-CG^{11,12} with the solver parameters identified in Wicht et al.,¹⁵ the ADMM compares favorably. Both Newton-CG (where we count, both, linear and nonlinear iterations) and nonlinear CG operate on a similar level as ADMM with the optimum choice. The (primal) Barzilai–Borwein method is significantly faster than the other primal methods, but cannot match the performance of ADMM with the Lorenz–Tran-Dinh scaling of the reference material. Indeed, the latter is twice as fast.

Let us take a look at the individual iteration counts per load step, see Figure 2B. Except for the very first steps, the Lorenz–Tran-Dinh scaling leads to the lowest iteration counts for every load step. This is consistent with the previous

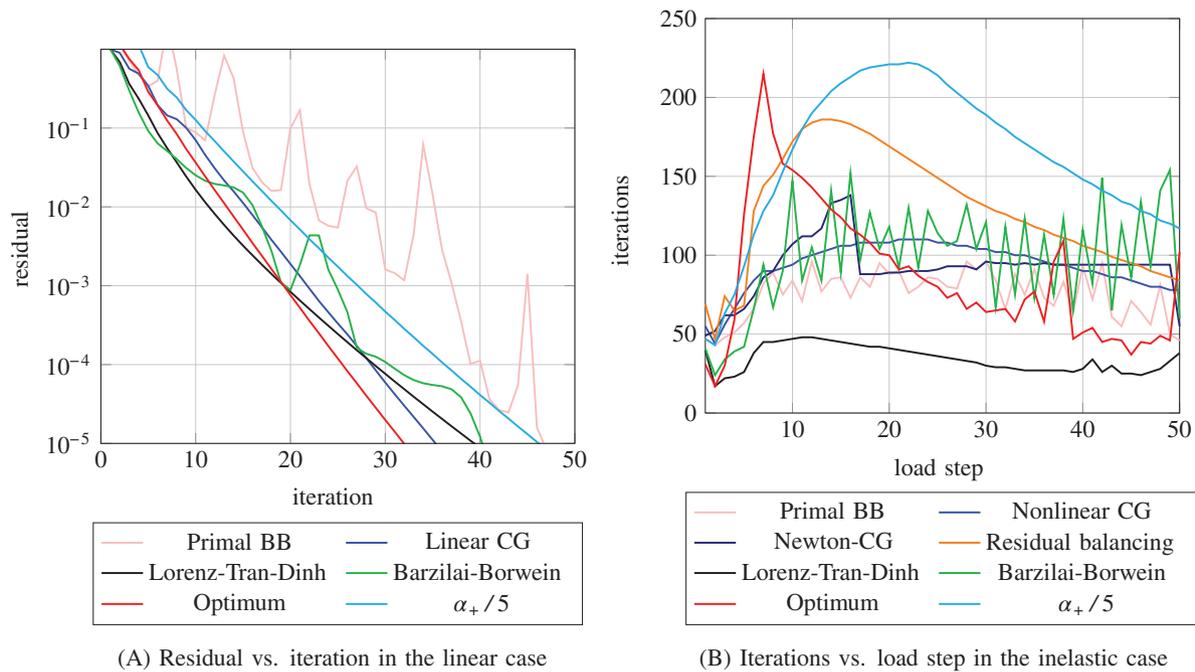


FIGURE 2 Studies for the fiber composite of Figure 1A. For the ADMM variants, we always use the damping γ leading to best performance

TABLE 3 Mean iterations and run time (in s) for the inelastic fiber composite

ADMM	Optimum	$\alpha_+/5$	Residual balancing	Lorenz–Tran-Dinh	Barzilai–Borwein
$\gamma = 0$	87.18 (4358.6)	162.70 (7990.5)	–	34.26 (1770.6)	98.08 (5078.3)
$\gamma = 1/4$	105.08 (5235.9)	216.9 (10618.9)	139.02 (7064.3)	41.18 (2102.4)	116.42 (6092.1)
$\gamma = 1/2$	157.36 (7794.4)	325.5 (15940.4)	129.28 (6556.0)	59.84 (3039.8)	173.30 (8968.7)
Primal solvers					
BB	75.48 (3320.7)				
CG	93.42 (4995.0)				
Newton	92.18 (4402.6)				

Note: The bold values correspond to the lowest iteration count for the primal-dual and the primal solvers (hence on each).

observations, where this method was able to reduce the residual quickly. Combined with the affine extrapolation, this initially fast decrease of the residual leads to low iteration counts for this method.

It is interesting to note that the “optimum choice” for the reference material in ADMM does not lead to the minimum iteration count, for instance exemplified by the peak at load step seven. The reasons are that, for a start, the optimum choice only realizes the optimum convergence rate. It is insensitive to the prefactor in front of the rate. In particular, combined to affine extrapolation, the optimum convergence rate is unable to come into play. Also, the optimum rate is only best if the bounds α_{\pm} are actually sharp. In any case, studying adaptive choices for the reference material in ADMM can be extremely productive, also in the case where the supposedly optimum parameter choice is known.

The constant strategy $\alpha_0 = \alpha_+/5$ performs well for the first ten load steps, but lags behind later on. Indeed, this strategy is fixed beforehand, and once and for all. In particular, the reference material cannot account for the increasing plastification for higher loading. All in all, this strategy (with $\gamma = 0$) is about 4.5 times slower than the Lorenz–Tran-Dinh scaling.

4.3 | A closed-cell foam

In this section, we consider a closed-cell foam, see Figure 1B. Following Abdullahi et al.,⁶¹ the microstructure was generated by shrinking the cells of a Laguerre tessellation, where we use the Newton algorithm discussed in Kuhn et al.⁶² to generate 27 Laguerre cells of equal volume. The resulting foam has a volume fraction of 11.6%, and was discretized by 128^3 voxels. Please note that we used a higher resolution for the visualization in Figure 1B.

This foam example is rather challenging for FFT-based solvers. In fact, FFT-based methods are certainly not the optimum for treating microstructures with such a high degree of porosity (88.4%). Rather, the foam structure should be regarded as a benchmark showing the limitations of FFT-based methods, in general, and the ADMM, in particular. We furnish the solid foam with the isotropic and linear elastic material parameters of aluminum listed in Table 1. The iteration counts for 5% uni-axial extension in x -direction are listed in Table 4. We prescribed a maximum of 5000 iterations, and ADMM with residual balancing and $\gamma = 1/4$ failed to converge within these limits. No damping ($\gamma = 0$), leads to divergence of the solver.

The conjugate gradient method converges in about 400 iterations. The primal BB method requires more than 1000 iterations, and is closely matched by the primal-dual BB. For this example, the Lorenz–Tran-Dinh scaling requires significantly more iterations, roughly by a factor of 3, than the BB scaling. The $\alpha_0 = \alpha_+/5$ strategy does not reach the prescribed tolerance within 5000 iterations.

Investigating the residuals more closely, see Figure 3C, we notice that the CG method does not converge monotonically, but rather shows a marked zig-zag pattern. This reflects the high condition number of the problem under consideration. Up to about 300 iterations, the CG residual is similar to BB's. After this point, CG significantly speeds up. All other considered ADMM schemes appear to have the same convergence rate, but with different prefactors. Please also note the severe oscillations in the residual of the primal BB scheme which cover three orders of magnitude.

We shall use the foam example to also investigate the influence of the discretization method and the appropriateness of the residual (22). Indeed, for finite material contrast, the various discretizations available for FFT-based computational micromechanics perform similarly well. Differences emerge for (highly) porous materials, like the open-cell foam. For this purpose, we re-ran the computational experiments with identical settings but different discretization schemes. We study the original discretization of Moulinec and Suquet,^{3,4} which might be interpreted as an underintegrated Fourier–, Galerkin

TABLE 4 Iteration counts for the ADMM with different settings and the closed-cell foam

ADMM	$\alpha_+/5$	Residual balancing	Lorenz–Tran-Dinh	Barzilai–Borwein
$\gamma = 1/4$	>5000	>5000	2905	1168
$\gamma = 1/2$	>5000	4083	4357	1753
Primal solvers				
BB	1042			
CG	397			

Note: The primal solvers are shown for comparison. The bold values correspond to the lowest iteration count for the primal-dual and the primal solvers (hence on each).

method,⁶³ and the discretization on a rotated staggered grid,^{64,65} introduced to FFT-based micromechanics by Willot.³⁵ The residuals versus iteration are shown in Figure 3. We observe that none of the considered solution schemes, be it primal or primal-dual, converges within 5 000 iterations. For the Moulinec–Suquet discretization, all solvers fail to advance significantly below a residual of 10^{-2} . More strikingly, the linear conjugate gradient method^{5,6} stagnates at a residual above 10%. For Willot's discretization, the solvers fare slightly better, reaching residuals which are about an order of magnitude lower than the Moulinec–Suquet discretization. Also, a stagnating convergence behavior is observed for linear CG. This convergence behavior of the solution methods is in stark contrast to their staggered grid counterparts, see Figure 3C, which perform significantly better. Recall that such a discrepancy between the performance of *primal* solvers and different discretization schemes was already observed in Schneider et al.³⁶

To study the interplay of the discretization scheme and the residual (22), we monitor the effective axial stress per iteration in Figure 4. Strikingly, for the Moulinec–Suquet discretization, see Figures 4A, the effective stress decreases monotonically, and is expected to converge to zero as the iteration count goes to infinity. After 1 000 iterations, all the considered solvers give rise to *different* predicted effective stresses. Such an anomalous behavior is characteristic for the Moulinec–Suquet discretization applied to porous microstructures.⁴⁹ For Willot's discretization, see Figure 4B, the effective stresses appear to converge to a unique value, as is the case for the staggered grid, see Figure 4C.

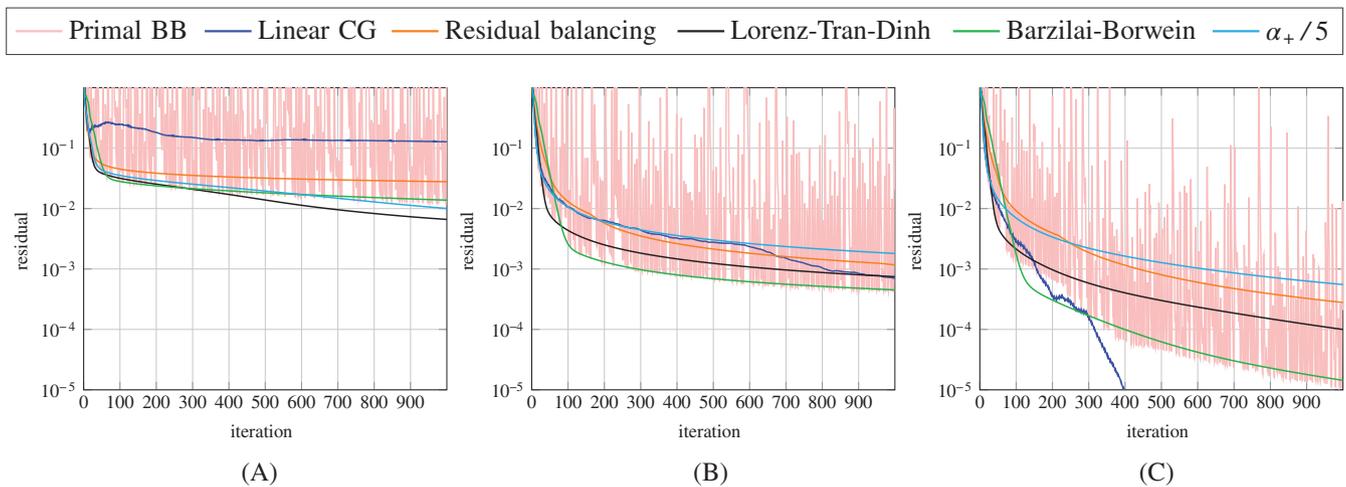


FIGURE 3 Residual versus iterations for the foam microstructure and different discretizations. For the ADMM variants, we always use the damping γ leading to best performance. (A) Moulinec–Suquet.^{3,4} (B) Willot.³⁵ (C) Staggered grid³⁶

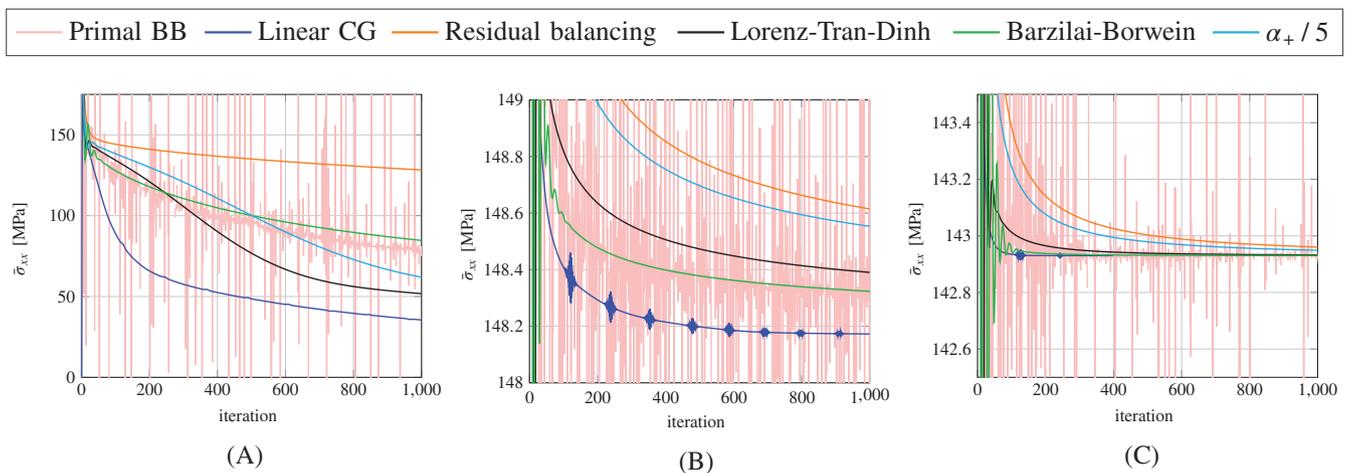


FIGURE 4 Evolution of the effective axial stress $\bar{\sigma}_{xx}$ for the foam microstructure, different discretizations and solvers. Notice the different scalings of the y-axes. (A) Moulinec–Suquet.^{3,4} (B) Willot.³⁵ (C) Staggered grid³⁶

TABLE 5 Mean iterations and run time (in s) for the porous Metal-Matrix Composite (MMC)

ADMM	$\alpha_+/5$	residual balancing	Lorenz-Tran-Dinh	Barzilai-Borwein
$\gamma = 1/4$	28.46 (735.3)	24.78 (645.4)	17.52 (469.7)	44.68 (1175.3)
$\gamma = 1/2$	42.52 (1068.2)	26.68 (698.8)	25.1 (660.0)	59.60 (1548.9)
Primal solvers				
BB	24.24 (483.2)			
CG	29.60 (684.9)			
Newton	33.86 (692.9)			

Note: The primal solvers are shown for comparison. The bold values correspond to the lowest iteration count for the primal-dual and the primal solvers (hence on each).

Actually, for the ADMM solvers and Willot's discretization, the change in the effective properties is about an order of magnitude higher than for the staggered grid, reflecting the different magnitudes of the residuals, which we observed in Figure 3.

To sum up, for this highly porous example, the influence of the discretization scheme becomes clear. Only for a well-conditioned discretization,⁴⁹ a robust convergence behavior of FFT-based solvers can be expected. This holds for primal solvers as well as for primal-dual solvers. In this regard, the staggered grid discretization shines, as it appears to be the most robust for highly porous materials among the studied discretization schemes.

The computational results confirm that the "natural" residual (22) serves as an appropriate indicator for convergence (or lack thereof), independent of the considered discretization scheme. Yet, when used in conjunction with the staggered grid discretization, ADMM appears to reach its limitations. It converges, but cannot be considered competitive to the fastest primal solver.

4.4 | A porous MMC

The last example concerns a Metal-Matrix Composite (MMC) with spherical pores (in blue), see Figure 1C. Apart from 29.5% spherical ceramic fillers, 0.5% spherical voids were included in the aluminum matrix. More precisely, 33 ceramic fillers of identical size were dispersed in the structure, together with 31 spherical pores with only 1/4th of the radius of the ceramic fillers. The spheres were placed inside the structure by the mechanical contraction method.⁶⁶ The structure was discretized by 128^3 voxels, and we consider uni-axial extension up to 1% strain, distributed over 50 equidistant load steps. We consider the ceramic fillers to be linear elastic, and furnish the aluminum matrix with a von Mises elastoplastic constitutive model and a power-law hardening, see Segurado et al.⁵⁹ The specific model parameters are listed in Table 1. Due to the exponent in the power-law hardening, the smallest eigenvalue of the material tangent in the aluminum matrix goes to zero as the strain loading goes to infinity. Thus, this setup challenges the optimum choice $\alpha_0 = \sqrt{\alpha_- \alpha_+}$ even in the non-porous case, see Schneider et al.^{28(sec.7.3)} Of course, including spherical pores creates additional difficulties for the FFT-based solvers. The performance of the considered solvers is summarized in Table 5. We first take a look at the primal solvers. Newton-CG requires the highest iteration count. This is not unexpected, because the principal advantage of Newton-CG rests upon the fact that the iterations of the linear CG can be much faster than the nonlinear iterations. For the problem at hand, the constitutive models are not much more expensive than the linear elastic model. Therefore, Newton-CG and nonlinear CG lead to pretty much the same run time. In contrast, the primal Barzilai-Borwein scheme is much faster in terms of run time. This roots in the relative simplicity of the BB update. Indeed, it is little more than the basic scheme, and does not feature the more involved linear algebra operations necessary for nonlinear CG and Newton-CG. As the constitutive updates are so cheap, the linear algebra operations need to be accounted for.

For the ADMM, the Barzilai-Borwein scaling does not perform well, at all. Residual balancing with $\gamma = 1/4$ leads to roughly the same iteration count as primal BB. However, the run time is much higher. Indeed, the linear algebra operations necessary for the ADMM significantly exceed those for primal BB. The Lorenz-Tran-Dinh scaling, on the other hand, slightly outperforms primal BB in terms of run time, and does so by requiring significantly less iterations. The iteration-independent choice $\alpha_0 = \alpha_+/5$ is about 56% slower than the fastest solver.

5 | CONCLUSIONS AND PERSPECTIVES

In this work, we studied strategies for choosing the reference material adaptively for the alternating direction method of multipliers introduced in the context of FFT-based computational micromechanics by Michel et al.^{18,19} The conclusions are the following.

1. In the literature, it was not fully clear whether the accelerated scheme converges for porous materials or not.^{23,24,39} The confusion may be attributed to different convergence criteria, a different choice of reference material, and usage of ill-conditioned discretization schemes.

In this work, we reported examples for a stable convergence process, and underlined these findings by theoretical arguments, see Section 2.1. More precisely, for a stable discretization scheme, for instance the staggered grid,³⁶ the damped ADMM Algorithm (5) converges for non-vanishing damping parameter γ . Under a geometric assumption on the solid-pore interface,⁴⁹ the convergence is linear.

Still, optimized primal solvers outperformed the considered primal-dual solvers. It remains to investigate whether an improved selection strategy for the reference material permits outperforming primal solvers for porous microstructures, as can be done for finite material contrast.^{28,42}

2. We devoted attention to adaptive choices of the reference material in the damped ADMM. Such adaptive choices were shown to represent memory-efficient alternatives to the Anderson-accelerated polarization methods,⁴² sometimes even improving upon their performance.
3. For the ADMM, the residual-balancing strategy⁴⁵ is a classic, and widely used. We discussed the strategy for the problem at hand, see Section 2.3. In the process, we extended the strategy to variable damping, and identified the individual primal and dual residuals. It is interesting to note that, for the cell problem of micromechanics, the primal and dual residual vectors may be linearly combined to the consistent residual vector. This appears not to be the case in general.^{22(sec.3.3)}
4. The strategy proposed by Lorenz and Tran-Dinh,⁴⁶ see Section 2.4, turned out to be both simple to implement and effective. For the computational examples considered, the Lorenz–Tran-Dinh strategy performed unreasonably well. Based on these results, we recommend this strategy.
5. We investigated a Barzilai–Borwein step size for the ADMM,⁴⁷ see Section 2.5. In contrast to the Barzilai–Borwein primal solver,³⁹ the BB strategy leads to significantly less oscillations in the associated residuals. Still, the strategy inherits the convergence speed from the primal setting, and is also recommended.
6. For primal FFT-based solvers, it was shown⁷ that choosing the proper step size may be of secondary importance, as long as the momentum term is chosen in an adaptive manner. For the work at hand, we fixed the damping parameter γ . It might be of interest to investigate whether choosing the damping parameter adaptively can be beneficial, see Xu et al.⁴⁸
7. In this work, we considered convex and differentiable objective functions. The ADMM may be applied to non-differentiable objective functions, as well.^{32,33} Furthermore, for computational homogenization at finite strains¹⁰, ADMM was applied successfully,^{13,31} as well. Adaptive strategies for the reference material should be of interest for these scenarios. Similarly, treating non-convex problems like phase-field fracture on microstructures^{9,67} by ADMM could be pursued.

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (DFG) within the International Research Training Group “Integrated engineering of continuous-discontinuous long fiber reinforced polymer structures” (GRK 2078). The anonymous reviewers deserve gratitude for providing constructive feedback.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Matti Schneider  <https://orcid.org/0000-0001-7017-3618>

REFERENCES

1. Matouš K, Geers MGD, Kouznetsova VG, Gillman A. A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials. *J Comput Phys*. 2017;330:192-220.
2. Schneider M. A review of non-linear FFT-based computational homogenization methods. *Acta Mech*. 2021;1-51.
3. Moulinec H, Suquet P. A fast numerical method for computing the linear and nonlinear mechanical properties of composites. *Comptes Rendus de l'Académie des Sciences Série II*. 1994;318(11):1417-1423.
4. Moulinec H, Suquet P. A numerical method for computing the overall response of nonlinear composites with complex microstructure. *Comput Methods Appl Mech Eng*. 1998;157:69-94.
5. Zeman J, Vondřejc J, Novák J, Marek I. Accelerating a FFT-based solver for numerical homogenization of periodic media by conjugate gradients. *J Comput Phys*. 2010;229(21):8065-8071.
6. Brisard S, Dormieux L. FFT-based methods for the mechanics of composites: a general variational framework. *Comput Mater Sci*. 2010;49(3):663-671.
7. Schneider M. A dynamical view of nonlinear conjugate gradient methods with applications to FFT-based computational micromechanics. *Comput Mech*. 2020;66:239-257.
8. Schneider M. An FFT-based fast gradient method for elastic and inelastic unit cell homogenization problems. *Comput Methods Appl Mech Eng*. 2017;315:846-866.
9. Ernesti F, Schneider M, Böhlke T. Fast implicit solvers for phase field fracture problems on heterogeneous microstructures. *Comput Methods Appl Mech Eng*. 2020;363:112793.
10. Lahellec N, Michel JC, Moulinec H, Suquet P. Analysis of inhomogeneous materials at large strains using fast Fourier transforms. In: Miehe C, ed. *Proceedings of the IUTAM Symposium on Computational mechanics of Solid Materials at Large Strains. Solid Mechanics and Its Applications*. Vol 108. Springer; 2003:247-258.
11. Gélébart L, Mondon-Cancel R. Non-linear extension of FFT-based methods accelerated by conjugate gradients to evaluate the mechanical behavior of composite materials. *Comput Mater Sci*. 2013;77:430-439.
12. Kabel M, Böhlke T, Schneider M. Efficient fixed point and Newton-Krylov solvers for FFT-based homogenization of elasticity at large deformations. *Comput Mech*. 2014;54(6):1497-1514.
13. Shantraj P, Eisenlohr P, Diehl M, Roters F. Numerically robust spectral methods for crystal plasticity simulations of heterogeneous materials. *Int J Plast*. 2015;66:31-45.
14. Chen Y, Gélébart L, Chateau C, Bornert M, Sauder C, King A. Analysis of the damage initiation in a SiC/SiC composite tube from a direct comparison between large-scale numerical simulation and synchrotron X-ray micro-computed tomography. *Int J Solids Struct*. 2019;161:111-126.
15. Wicht D, Schneider M, Böhlke T. On Quasi-Newton methods in FFT-based micromechanics. *Int J Numer Methods Eng*. 2020;121(8):1665-1694.
16. Bhattacharya K, Suquet P. A model problem concerning recoverable strains of shape-memory polycrystals. *Proc Royal Soc A*. 2005;461:2797-2816.
17. Wicht D, Schneider M, Böhlke T. An efficient solution scheme for small-strain crystal-elasto-viscoplasticity in a dual framework. *Comput Methods Appl Mech Eng*. 2020;358:112611.
18. Michel JC, Moulinec H, Suquet P. A computational method based on augmented Lagrangians and fast Fourier transforms for composites with high contrast. *Comput Model Eng Sci*. 2000;1(2):79-88.
19. Michel JC, Moulinec H, Suquet P. A computational scheme for linear and non-linear composites with arbitrary phase contrast. *Int J Numer Methods Eng*. 2001;52:139-160.
20. Glowinski R, Marrocco A. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *ESAIM Math Model Numer Anal - Modélisation Mathématique et Analyse Numérique*. 1975;9:41-76.
21. Gabay D, Mercier B. A dual algorithm for the solution of nonlinear variational problems via finite element approximations. *Comput Math Appl*. 1976;2(1):17-40.
22. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found Trends Mach Learn*. 2010;3(1):1-122.
23. Monchiet V, Bonnet G. A polarization-based FFT iterative scheme for computing the effective properties of elastic composites with arbitrary contrast. *Int J Numer Methods Eng*. 2012;89:1419-1436.
24. Moulinec H, Silva F. Comparison of three accelerated FFT-based schemes for computing the mechanical response of composite materials. *Int J Numer Methods Eng*. 2014;97:960-985.
25. Eyre DJ, Milton GW. A fast numerical scheme for computing the response of composites using grid refinement. *Eur Phys J Appl Phys*. 1999;6(1):41-47.
26. Vinogradov V, Milton GW. An accelerated FFT algorithm for thermoelastic and non-linear composites. *Int J Numer Methods Eng*. 2008;76:1678-1695.
27. Monchiet V, Bonnet G. Numerical homogenization of nonlinear composites with a polarization-based FFT iterative scheme. *Comput Mater Sci*. 2013;79:276-283.
28. Schneider M, Wicht D, Böhlke T. On polarization-based schemes for the FFT-based computational homogenization of inelastic materials. *Comput Mech*. 2019;64(4):1073-1095.

29. Giselsson P, Boyd S. Linear convergence and metric selection for Douglas-Rachford splitting and ADMM. *IEEE Trans Autom Control*. 2017;62:532-544.
30. Lebensohn RA, Kanjarla AK, Eisenlohr P. An elasto-viscoplastic formulation based on fast Fourier transforms for the prediction of micromechanical fields in polycrystalline materials. *Int J Plast*. 2012;32-33:59-69.
31. Zhu H, Bhattacharya K. Accelerated computational micromechanics. arXiv e-prints, vol. 2010.06697; 2020:1-36.
32. Schneider M. An FFT-based method for computing weighted minimal surfaces in microstructures with applications to the computational homogenization of brittle fracture. *Int J Numer Methods Eng*. 2020;121(7):1367-1387.
33. Willot F. The effective conductivity of strongly nonlinear media: the dilute limit. *Int J Solids Struct*. 2020;184:287-295.
34. Schneider M. Convergence of FFT-based homogenization for strongly heterogeneous media. *Math Methods Appl Sci*. 2015;38(13):2761-2778.
35. Willot F. Fourier-based schemes for computing the mechanical response of composites with accurate local fields. *Comptes Rendus Mécanique*. 2015;343:232-245.
36. Schneider M, Ospald F, Kabel M. Computational homogenization of elasticity on a staggered grid. *Int J Numer Methods Eng*. 2016;105(9):693-720.
37. Schneider M, Merkert D, Kabel M. FFT-based homogenization for microstructures discretized by linear hexahedral elements. *Int J Numer Methods Eng*. 2017;109:1461-1489.
38. Leuschner M, Fritzen F. Fourier-accelerated nodal solvers (FANS) for homogenization problems. *Comput Mech*. 2018;62:359-392.
39. Schneider M. On the Barzilai-Borwein basic scheme in FFT-based computational homogenization. *Int J Numer Methods Eng*. 2019;118(8):482-494.
40. Anderson DG. Iterative procedures for nonlinear integral equations. *J ACM*. 1965;12(4):547-560.
41. Fang H-R, Saad Y. Two classes of multisecond methods for nonlinear acceleration. *Numer Linear Algebra Appl*. 2009;16:197-221.
42. Wicht D, Schneider M, Böhlke T. Anderson-accelerated polarization schemes for FFT-based computational homogenization. *Int J Numer Methods Eng*. 2021;1-30.
43. Barzilai J, Borwein JM. Two-point step size gradient methods. *IMA J Numer Anal*. 1988;8:141-148.
44. Kabel M, Fliegner S, Schneider M. Mixed boundary conditions for FFT-based homogenization at finite strains. *Comput Mech*. 2016;57(2):193-210.
45. He BS, Yang H, Wang SL. Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities. *J Optim Theory Appl*. 2000;106(2):337-356.
46. Lorenz DA, Tran-Dinh Q. Non-stationary Douglas-Rachford and alternating direction method of multipliers: adaptive step-sizes and convergence. *Comput Optim Appl*. 2019;74:67-92.
47. Xu Z, Figueiredo MAT, Goldstein T. Adaptive ADMM with spectral penalty parameter selection. In: Singh A, Zhu J, eds. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research*. Vol 54. PMLR; 2017:718-727.
48. Xu Z, Figueiredo MAT, Yuan X, Studer C, Goldstein T. Adaptive relaxed ADMM: convergence theory and practical implementation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE; 2017:7234-7243.
49. Schneider M. Lippmann-Schwinger solvers for the computational homogenization of materials with pores. *Int J Numer Methods Eng*. 2020;121(22):5017-5041.
50. He BS, Yuan XM. On the $O(1/t)$ convergence rate of the alternating direction method. *SIAM J Numer Anal*. 2012;50:700-709.
51. He BS, Yuan XM. On non-ergodic convergence rate of Douglas-Rachford alternating directions method of multipliers. *Numer Math*. 2015;130(3):567-577.
52. Eckstein J, Bertsekas D. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math Program*. 1992;55:293-318.
53. Liang J, Fadili J, Peyré G. Local convergence properties of douglas-rachford and alternating direction method of multipliers. *J Optim Theory Appl*. 2017;172:874-913.
54. Milton GW. *The Theory of Composites*. Cambridge University Press; 2002.
55. Raydan M. The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. *SIAM J Optim*. 1997;7:26-33.
56. Dai YH, Liao LZ. R-linear convergence of the Barzilai and Borwein gradient method. *IMA J Numer Anal*. 2002;22:1-10.
57. Fletcher R. On the Barzilai-Borwein method. In: Qi L, Teo K, Yang X, eds. *Optimization and Control with Applications*. Springer; 2005:235-256.
58. Wu L, Adam L, Doghri I, Noels L. An incremental-secant mean-field homogenization method with second statistical moments for elasto-visco-plastic composite materials. *Mech Mater*. 2017;114:180-200.
59. Segurado J, Llorca J, González C. On the accuracy of mean-field approaches to simulate the plastic deformation of composites. *Scr Mater*. 2002;46(7):525-529.
60. Schneider M. The sequential addition and migration method to generate representative volume elements for the homogenization of short fiber reinforced plastics. *Comput Mech*. 2017;59:247-263.
61. Abdullahi HS, Liang Y, Gao S. Predicting the elastic properties of closed-cell aluminum foams: a mesoscopic geometric modeling approach. *SN Applied Sciences*. 2019;1:380.

62. Kuhn J, Schneider M, Sonnweber-Ribic P, Böhlke T. Fast methods for computing centroidal Laguerre tessellations for prescribed volume fractions with applications to microstructure generation of polycrystalline materials. *Comput Methods Appl Mech Eng.* 2020;369:113175.
63. Vondřejc J, Zeman J, Marek I. An FFT-based Galerkin method for homogenization of periodic media. *Comput Math Appl.* 2014;68(3):156-173.
64. Saenger EH, Gold N, Shapiro SA. Modeling the propagation of elastic waves using a modified finite-difference grid. *Wave Motion.* 2000;31:77-92.
65. Saenger EH, Bohlen T. Finite-difference modeling of viscoelastic and anisotropic wave propagation using the rotated staggered grid. *Geophysics.* 2004;69:583-591.
66. Williams S, Philipse A. Random packings of spheres and spherocylinders simulated by mechanical contraction. *Phys Rev E.* 2003;67:1-9.
67. Chen Y, Vasiukov D, Gélébart L, Park CH. A FFT solver for variational phase-field modeling of brittle fracture. *Comput Methods Appl Mech Eng.* 2019;349:167-190.

How to cite this article: Schneider M. On non-stationary polarization methods in FFT-based computational micromechanics. *Int J Numer Methods Eng.* 2021;122:6800-6821. <https://doi.org/10.1002/nme.6812>

APPENDIX A. IDENTITIES FOR PRIMAL AND DUAL RESIDUALS

In this section, we wish to show the estimates (12)

$$\|\bar{\varepsilon} - \mathbb{P} : \langle e^{k+1} \rangle_Q\|_{L^2} + \|(\text{Id} - \Gamma - \langle \cdot \rangle_Q) : e^{k+1}\|_{L^2} \leq \|p^{k+1}\|_{L^2} \quad (\text{A1})$$

and (13)

$$\|\bar{\sigma} - \mathbb{Q} : \langle \lambda^{k+1} \rangle_Q\|_{L^2} + \|\Gamma : \lambda^{k+1}\|_{L^2} \leq \|d^{k+1}\|_{L^2}, \quad (\text{A2})$$

for the primal and dual residuals (14),

$$p^{k+1} = \varepsilon^{k+\frac{1}{2}} - e^{k+1} \quad \text{and} \quad d^{k+1} = \alpha_k \left(e^{k+1} - e^k + \varepsilon^{k+\frac{1}{2}} - \varepsilon^{k+1} \right), \quad (\text{A3})$$

respectively, and the general algorithm (5)

$$\begin{aligned} \varepsilon^{k+\frac{1}{2}} &= \bar{\varepsilon} + \mathbb{D}^k : \bar{\sigma} - (\Gamma^k + \mathbb{D}^k : \mathbb{Q} : \langle \cdot \rangle_Q) : (\lambda^k - \mathbb{C}^k : e^k), \\ \varepsilon^{k+1} &= 2(1 - \gamma_k)\varepsilon^{k+\frac{1}{2}} - (1 - 2\gamma_k)e^k, \\ \sigma(e^{k+1}) + \mathbb{C}^k : e^{k+1} &= \lambda^k + \mathbb{C}^k : \varepsilon^{k+1}, \\ \lambda^{k+1} &= \lambda^k + \mathbb{C}^k : (\varepsilon^{k+1} - e^{k+1}). \end{aligned} \quad (\text{A4})$$

To derive the estimate for the primal residual (A1), we utilize the compatibility of $\varepsilon^{k+\frac{1}{2}}$

$$\varepsilon^{k+\frac{1}{2}} = \mathbb{P} : \bar{\varepsilon} + \mathbb{Q} : \langle \varepsilon^{k+\frac{1}{2}} \rangle_Q + \Gamma^k : \mathbb{C}^k : \varepsilon^{k+\frac{1}{2}},$$

compare Equation (B4), and write

$$\begin{aligned} &(\text{Id} - \Gamma^k : \mathbb{C}^k - \langle \cdot \rangle_Q) : e^{k+1} + \mathbb{P} : \langle e^{k+1} \rangle_Q - \bar{\varepsilon} \\ &= e^{k+1} - \bar{\varepsilon} - \mathbb{Q} : \langle e^{k+1} \rangle_Q - \Gamma^k : \mathbb{C}^k : e^{k+1} \\ &= (\text{Id} - \Gamma^k : \mathbb{C}^k - \mathbb{Q} : \langle \cdot \rangle_Q) : \left(e^{k+1} - \varepsilon^{k+\frac{1}{2}} \right). \end{aligned}$$

Taking norms on both sides and using the Helmholtz decomposition (B2) shows the estimate (A1). For the dual residual, the last line of the iteration (A4) implies

$$\lambda^k - \mathbb{C}^k : e^k = \lambda^{k+1} - \mathbb{C}^k : (\varepsilon^{k+1} - e^{k+1}) - \mathbb{C}^k : e^k.$$

In view of Equation (B4) in the form

$$\varepsilon^{k+\frac{1}{2}} = \bar{\varepsilon} + (\Gamma^k + \mathbb{D}^k : \mathbb{Q} : \langle \cdot \rangle_Q) : \mathbb{C}^k : \varepsilon^{k+\frac{1}{2}},$$

we may insert the latter identity into the first line (A4) to obtain

$$\begin{aligned} \varepsilon^{k+\frac{1}{2}} &= \bar{\varepsilon} + \mathbb{D}^k : \bar{\sigma} - (\Gamma^k + \mathbb{D}^k : \mathbb{Q} : \langle \cdot \rangle_Q) : (\lambda^k - \mathbb{C}^k : e^k) \\ 0 &= \mathbb{D}^k : \bar{\sigma} - (\Gamma^k + \mathbb{D}^k : \mathbb{Q} : \langle \cdot \rangle_Q) : (\lambda^k - \mathbb{C}^k : e^k + \mathbb{C}^k : \varepsilon^{k+\frac{1}{2}}), \end{aligned}$$

that is, the identity

$$\Gamma^k : \lambda^{k+1} + \mathbb{D}^k : (\mathbb{Q} : \langle \lambda^k \rangle_Q - \bar{\sigma}) = \Gamma^k : \mathbb{C}^k : (\varepsilon^{k+1} - e^{k+1} + e^k - \varepsilon^{k+\frac{1}{2}}) = -\Gamma^k : \mathbb{C}^k : d^{k+1} / \alpha_k$$

holds. Taking norms on both sides and using the non-expansivity of $\Gamma^k : \mathbb{C}^k$ yields the dual estimate (A2).

APPENDIX B. DETAILS ON THE CONVERGENCE CRITERION

We wish to show the identity

$$\begin{aligned} \|e^k - \varepsilon^{k+\frac{1}{2}}\|_{\mathbb{C}^k}^2 &= \|\Gamma^k : \lambda^k\|_{\mathbb{C}^k}^2 + \|(\text{Id} - \Gamma^k : \mathbb{C}^k - \langle \cdot \rangle_Q) : e^k\|_{\mathbb{C}^k}^2 \\ &\quad + \|\mathbb{P} : \langle e^k \rangle_Q - \bar{\varepsilon}\|_{\mathbb{C}^k}^2 + \|\bar{\sigma} - \mathbb{Q} : \langle \lambda^k \rangle_Q\|_{\mathbb{D}^k}^2. \end{aligned} \quad (\text{B1})$$

Before coming to the details, let us record the following consequence of the Helmholtz decomposition in (linearized) elasticity for mixed boundary conditions^{28(Apx.)}

$$\|\varepsilon\|_{\mathbb{C}^k}^2 = \|\Gamma^k : \mathbb{C}^k : \varepsilon\|_{\mathbb{C}^k}^2 + \|(\text{Id} - \Gamma^k : \mathbb{C}^k - \langle \cdot \rangle_Q) : \varepsilon\|_{\mathbb{C}^k}^2 + \|\mathbb{P} : \langle \varepsilon \rangle_Q\|_{\mathbb{C}^k}^2 + \|\mathbb{Q} : \langle \varepsilon \rangle_Q\|_{\mathbb{C}^k}^2, \quad (\text{B2})$$

valid for any square-integrable strain field, measuring (in order) equilibrium, compatibility and the magnitude of the mean, decomposed into the \mathbb{P} - and \mathbb{Q} -parts.

Let us investigate the first line in the Algorithm (5)

$$\varepsilon^{k+\frac{1}{2}} = \bar{\varepsilon} + \mathbb{D}^k : \bar{\sigma} - (\Gamma^k + \mathbb{D}^k : \mathbb{Q} : \langle \cdot \rangle_Q) : (\lambda^k - \mathbb{C}^k : e^k)$$

in more detail. Separating the homogeneous and non-homogeneous parts, we observe

$$\varepsilon^{k+\frac{1}{2}} = \bar{\varepsilon} + \mathbb{Q} : \langle e^k \rangle_Q + \mathbb{D}^k : (\bar{\sigma} - \mathbb{Q} : \langle \lambda^k \rangle_Q) - \Gamma^k : (\lambda^k - \mathbb{C}^k : e^k). \quad (\text{B3})$$

As $\Gamma^k : \mathbb{C}^k$ is the \mathbb{C}^k -orthogonal projector onto kinematically compatible fields, it follows

$$\varepsilon^{k+\frac{1}{2}} = \bar{\varepsilon} + \mathbb{Q} : \langle e^k \rangle_Q + \mathbb{D}^k : (\bar{\sigma} - \mathbb{Q} : \langle \lambda^k \rangle_Q) + \Gamma^k : \mathbb{C}^k : \varepsilon^{k+\frac{1}{2}}. \quad (\text{B4})$$

In order to assess the validity of the last three Equations 21, we need to monitor the quantity

$$(\text{Id} - \Gamma^k : \mathbb{C}^k - \langle \cdot \rangle_Q) : e^k + \mathbb{P} : \langle e^k \rangle_Q - \bar{\varepsilon} - \mathbb{D}^k : (\bar{\sigma} - \mathbb{Q} : \langle \lambda^k \rangle_Q).$$

In view of the identity (B4), the latter quantity may be rewritten in the form

$$\begin{aligned}
 & (\text{Id} - \Gamma^k : \mathbb{C}^k - \langle \cdot \rangle_Q) : e^k + \mathbb{P} : \langle e^k \rangle_Q - \bar{\varepsilon} - \mathbb{D}^k : (\bar{\sigma} - \mathbb{Q} : \langle \lambda^k \rangle_Q) \\
 &= e^k - \Gamma^k : \mathbb{C}^k : e^k - \mathbb{P} : \langle e^k \rangle_Q - \mathbb{Q} : \langle e^k \rangle_Q + \mathbb{P} : \langle e^k \rangle_Q - \bar{\varepsilon} - \mathbb{D}^k : (\bar{\sigma} - \mathbb{Q} : \langle \lambda^k \rangle_Q) \\
 &= e^k - \Gamma^k : \mathbb{C}^k : e^k - \mathbb{Q} : \langle e^k \rangle_Q - \bar{\varepsilon} - \mathbb{D}^k : (\bar{\sigma} - \mathbb{Q} : \langle \lambda^k \rangle_Q) \\
 &= e^k - \Gamma^k : \mathbb{C}^k : e^k - \varepsilon^{k+\frac{1}{2}} + \Gamma^k : \mathbb{C}^k : \varepsilon^{k+\frac{1}{2}} \\
 &= (\text{Id} - \Gamma^k : \mathbb{C}^k) : (e^k - \varepsilon^{k+\frac{1}{2}}).
 \end{aligned} \tag{B5}$$

To assess the equilibrium of the stress field λ^k , comparing the Equations (B3) and (B4) yields the identity

$$-\Gamma^k : (\lambda^k - \mathbb{C}^k : e^k) = \Gamma^k : \mathbb{C}^k : \varepsilon^{k+\frac{1}{2}},$$

which might be rearranged into the form

$$\Gamma^k : \lambda^k = \Gamma^k : \mathbb{C}^k : (e^k - \varepsilon^{k+\frac{1}{2}}).$$

Inserting the latter identity into the formula (B5) yields

$$e^k - \varepsilon^{k+\frac{1}{2}} = \Gamma^k : \lambda^k + (\text{Id} - \Gamma^k : \mathbb{C}^k - \langle \cdot \rangle_Q) : e^k + \mathbb{P} : \langle e^k \rangle_Q - \bar{\varepsilon} - \mathbb{D}^k : (\bar{\sigma} - \mathbb{Q} : \langle \lambda^k \rangle_Q).$$

Upon taking the (squared) \mathbb{C}^k -norm, the Helmholtz decomposition in the form (B2) implies the formula (B1).