Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Over-Relaxed Projected-Forward Iterations for Cocoercive Variational Inequalities: Active-Face Spectral Tuning

Olaoluwa Jeremiah Omidire

olaoluwa.omidire@uniosun.edu.ng

Osun State University https://orcid.org/0000-0003-1321-684X

K. R. Tijani

Osun State University

B. T. Ishola

Obafemi Awolowo University

M. O. Olatinwo

Obafemi Awolowo University

F. S. Adeyinka

Bowen University

D. R. Ariyo

University of Ilesa

Research Article

Keywords: Variational inequality, projected-forward method, over-relaxation, active-set identification, cocoercive operator, spectral rate

Posted Date: August 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10685329/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Over-Relaxed Projected-Forward Iterations for Cocoercive Variational Inequalities: Active-Face Spectral Tuning

O. J. Omidire[1,*] K. R. Tijani[1] B. T. Ishola[2]
M. O. Olatinwo[2] F. S. Adeyinka[3] R. D. Ariyo[4]

[1]Department of Mathematical Sciences, Osun State University, Osogbo, Nigeria
[2]Department of Mathematics, Obafemi Awolowo University, Ile-Ife, Nigeria
[3]Department of Mathematics and Statistics, Bowen University, Iwo, Nigeria
[4]Department of Mathematics, University of Ilesa, Ilesa, Nigeria
[*]Corresponding author: olaoluwa.omidire@uniosun.edu.ng; omidireolaoluwa@gmail.com

Abstract

We study outer relaxation of the projected-forward iteration for cocoercive variational inequalities. The admissible over-relaxation interval follows from standard forward–backward averagedness, so the issue addressed here is parameter selection rather than global admissibility. For box constraints, strict complementarity and local differentiability imply finite identification of the active face. After identification, the relaxed map has a block triangular derivative; when the free–free Jacobian is symmetric positive definite, its local spectral radius is the maximum of the active factor $|1-\lambda|$ and the free-face factors $|1-\lambda\gamma\theta_j|$. This yields the closed-form spectral-radius minimizer $\lambda^*_{\mathrm{loc}} = 2/(a_\gamma + b_\gamma)$ and the exact criterion $a_\gamma + b_\gamma < 2$ for genuine over-relaxation. Joint tuning shows that an unrelaxed member is locally optimal when the forward step can be freely retuned, which identifies fixed or globally constrained forward steps as the natural regime for outer relaxation. We propose a pattern-reset safeguarded selector that clips spectral estimates to the global averagedness interval. In a dense 120-dimensional benchmark family where high-curvature directions are active, the locally preferred forward step violates the global step bound, while the adaptive relaxation reduces iteration counts by about 33% across three conditioning levels. A nonlinear 80-dimensional test confirms the same active-face mechanism and also shows the advantage of retuning the forward step when that option is globally admissible.

Keywords: variational inequality; projected-forward method; over-relaxation; active-set identification; cocoercive operator; spectral rate.
2020 Mathematics Subject Classification: 47H05, 47J20, 47J25, 49J40, 65K15, 90C25.

1 Introduction

Variational inequalities provide a common framework for constrained optimization, complementarity, equilibrium and nonlinear operator equations. Their classical development goes back to Hartman and Stampacchia [1] and Lions and Stampacchia [2]. Given a nonempty closed convex set X in a real Hilbert space $\mathcal{H}$ and an operator $f : \mathcal{H} \to \mathcal{H}$, we consider

$$\text{find } p \in X \text{ such that } \langle f(p), x - p\rangle \geq 0 \quad \text{for every } x \in X. \tag{1}$$

Its solution set is denoted by $\mathrm{VI}(X, f)$. The metric projection gives

$$p \in \mathrm{VI}(X, f) \quad \Longleftrightarrow \quad p = P_X(p - \gamma f(p)), \qquad \gamma > 0, \tag{2}$$

so the natural fixed-point map is

$$G_\gamma = P_X(I - \gamma f). \tag{3}$$

Equivalently, this is forward–backward splitting for $0 \in f(x) + N_X(x)$ because the resolvent of the normal cone is P_X.

The global relaxation theory of (3) is well established. Averagedness calculus gives an explicit interval of outer relaxation parameters for which

$$x_{n+1} = x_n + \lambda_n(G_\gamma x_n - x_n) \tag{4}$$

converges; variable-metric and relaxed forward–backward frameworks also permit relaxation above one [17, 21, 22]. Global contraction estimates under strong monotonicity are likewise classical [9, 24]. These results determine values of λ that are safe, but they do not determine which safe value is asymptotically fastest after the projection has identified the active constraints.

Activity identification provides the geometric structure needed for that question. Forward–backward methods identify active manifolds in finite time under partial smoothness and then enter a local linear regime [20]. For box-constrained optimization, active-set and spectral projected-gradient methods exploit curvature restricted to inactive variables [12–14, 30, 32]. These methods primarily tune the *forward steplength*. The present paper instead fixes the inner forward step γ and isolates the effect of the distinct outer relaxation λ in (4). The distinction is structural because, when the projection is active,

$$(1 - \lambda)x + \lambda P_X(x - \gamma f(x)) \neq P_X(x - \lambda\gamma f(x))$$

in general.

Optimal parameter selection for splitting methods has also been studied through worst-case performance estimation. In particular, Ryu, Taylor, Bergeling and Giselsson [31] derive tight contraction factors and select splitting parameters through semidefinite performance-estimation problems. Our result is different in scope: it is a post-identification, face-specific calculation for the projected-forward map, and the box geometry reduces the parameter problem to a closed-form scalar minimax rule. Thus the contribution is not a new global forward–backward interval or a general replacement for performance-estimation techniques; it is an analytic rule that exposes how active-coordinate damping and free-face curvature jointly determine the best scalar outer relaxation for a fixed γ.

The main contributions are as follows.

(i) For strongly monotone cocoercive operators, we combine the standard sharp forward contraction factor with averagedness to obtain a global linear certificate throughout the full admissible outer-relaxation window. This provides the global baseline for the local analysis.

(ii) For box constraints, strict complementarity yields finite stabilization of the projection face. For a C^1 operator we derive the post-identification derivative; in the affine case it is the exact eventual error recurrence.

(iii) When the free–free Jacobian is symmetric positive definite, we obtain the exact local spectral radius and its unique scalar minimizer. We state the result explicitly as *spectral-radius* or *asymptotic-root-factor* tuning, rather than as a one-step norm contraction claim.

(iv) We characterize the complete jointly optimal (γ, λ) family. This shows that $\lambda = 1$ can already be locally optimal when γ is freely retuned, and therefore identifies fixed, conservative or globally constrained forward steps as the setting in which outer relaxation has a distinct role.

(v) We propose a globally safeguarded adaptive selector with pattern reset. Every tentative spectral choice is clipped to the averagedness interval, so premature face declarations remain globally safe. The numerical section tests the adaptive algorithm itself, compares it with retuning γ, and includes a conditioning stress family in which the locally preferred forward step is globally inadmissible.

The update (4) is a Krasnosel'skiĭ–Mann/relaxed forward–backward step. Similar convex combinations also appear under Schaefer-type terminology in fixed-point literature [3–5, 23]; we use the standard forward–backward language because it makes the relation with contemporary optimization theory direct.

2 Preliminaries and exact fixed-point structure

Throughout, $\mathcal{H}$ is a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$, and $X \subset \mathcal{H}$ is nonempty, closed and convex.

Definition 2.1. *An operator $f : \mathcal{H} \to \mathcal{H}$ is*

(i) monotone if $\langle f(x) - f(y), x - y\rangle \geq 0$ for all $x, y \in \mathcal{H}$;

(ii) μ-strongly monotone, $\mu > 0$, if

$$\langle f(x) - f(y), x - y\rangle \geq \mu\|x - y\|^2;$$

(iii) α-inverse strongly monotone (or α-cocoercive), $\alpha > 0$, if

$$\langle f(x) - f(y), x - y\rangle \geq \alpha\|f(x) - f(y)\|^2. \tag{5}$$

Cocoercivity implies $1/\alpha$-Lipschitz continuity. The metric projection $P_X : \mathcal{H} \to X$ is firmly nonexpansive and satisfies

$$\langle u - P_X u, y - P_X u\rangle \leq 0, \qquad u \in \mathcal{H},\ y \in X. \tag{6}$$

Proposition 2.2 (Projection characterization)**.** *For every $\gamma > 0$,*

$$\mathrm{VI}(X, f) = \mathrm{Fix}(G_\gamma), \qquad G_\gamma = P_X(I - \gamma f). \tag{7}$$

Proof. Let $p \in X$ and apply (6) to $u = p - \gamma f(p)$. The identity $p = P_X u$ is equivalent to

$$\langle -\gamma f(p), y - p\rangle \leq 0 \qquad (y \in X),$$

which is exactly (1) because $\gamma > 0$. □

The preceding proposition also explains why a direct fixed-point iteration of f is generally unrelated to the variational inequality: $\mathrm{Fix}(f)$ need not equal $\mathrm{VI}(X, f)$. The projection is not a cosmetic modification; it is the resolvent step that enforces the normal-cone part of the inclusion.

Definition 2.3. *A map $T : \mathcal{H} \to \mathcal{H}$ is η-averaged, with $0 < \eta < 1$, if $T = (1 - \eta)I + \eta R$ for some nonexpansive R. Equivalently,*

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 - \frac{1 - \eta}{\eta}\|(I - T)x - (I - T)y\|^2. \tag{8}$$

Firmly nonexpansive maps are $1/2$-averaged. The composition formula of Combettes and Yamada [17] will be used repeatedly.

3 Known global over-relaxation baseline

This section records the standard global facts needed later. They are included to make the paper self-contained and to make clear that the new rate-selection results in Section 5 are not claims about discovery of the classical relaxation interval.

Proposition 3.1 (Averaged projected-forward map [17])**.** *Suppose f is α-cocoercive and $0 < \gamma < 2\alpha$. Then $I - \gamma f$ is $\gamma/(2\alpha)$-averaged and*

$$G_\gamma = P_X(I - \gamma f) \tag{9}$$

is η_γ-averaged with

$$\eta_\gamma = \frac{2\alpha}{4\alpha - \gamma} \in \left(\frac{1}{2}, 1\right). \tag{10}$$

Consequently,

$$\frac{1}{\eta_\gamma} = 2 - \frac{\gamma}{2\alpha} > 1. \tag{11}$$

Proof. The map $I - \gamma f$ is $\gamma/(2\alpha)$-averaged by (5). Since P_X is $1/2$-averaged, the composition formula

$$\eta = \frac{\eta_1 + \eta_2 - 2\eta_1\eta_2}{1 - \eta_1\eta_2}$$

with $\eta_1 = 1/2$ and $\eta_2 = \gamma/(2\alpha)$ gives (10). □

For constant γ, define

$$z_n = G_\gamma x_n, \qquad x_{n+1} = x_n + \lambda_n(z_n - x_n). \tag{12}$$

If $\lambda_n > 1$, x_{n+1} can leave X, whereas $z_n \in X$ remains feasible. The "shadow" point z_n is therefore the natural feasible output when transient infeasibility of the outer iterate is undesirable.

Theorem 3.2 (Standard Fejér descent and convergence)**.** *Assume* $\mathrm{VI}(X, f) \neq \emptyset$, $0 < \gamma < 2\alpha$, *and*

$$0 < \underline{\lambda} \leq \lambda_n \leq \overline{\lambda} < \frac{1}{\eta_\gamma}. \tag{13}$$

Then, for every $p \in \mathrm{VI}(X, f)$,

$$\|x_{n+1} - p\|^2 \leq \|x_n - p\|^2 - \lambda_n \left(\frac{1}{\eta_\gamma} - \lambda_n\right) \|x_n - G_\gamma x_n\|^2. \tag{14}$$

Hence $\|x_n - G_\gamma x_n\| \to 0$, *and*

$$\min_{0 \leq k \leq N} \|x_k - G_\gamma x_k\| \leq \frac{\|x_0 - p\|}{\sqrt{(N+1)\delta}}, \qquad \delta = \underline{\lambda}\left(\frac{1}{\eta_\gamma} - \overline{\lambda}\right) > 0. \tag{15}$$

Moreover, x_n *converges weakly to a point of* $\mathrm{VI}(X, f)$*; in finite-dimensional spaces the convergence is strong. The shadow sequence* z_n *has the same limit.*

Proof. Apply (8) to G_γ and use

$$\|(1-t)a + tb\|^2 = (1-t)\|a\|^2 + t\|b\|^2 - t(1-t)\|a - b\|^2$$

with $a = x_n - p$, $b = G_\gamma x_n - p$ and $t = \lambda_n$. This yields (14). Summation gives (15) and asymptotic regularity. Weak convergence follows from demiclosedness of $I - G_\gamma$ and Opial's argument [6, 7]; finite-dimensional weak and strong convergence coincide. Finally, $z_n - x_n \to 0$. □

Remark 3.3. *The interval* (11) *is a standard forward–backward relaxation interval. Combettes and Yamada explicitly developed extended forward–backward relaxation using improved averagedness constants [17], and later variable-metric analyses retain relaxation parameters above one [21]. Thus, the role of Theorem 3.2 in this paper is certification, not novelty.*

4 Strong monotonicity: a global contraction baseline

Assume in this section that f is both α-cocoercive and μ-strongly monotone. The simultaneous assumptions imply $0 < \mu\alpha \leq 1$.

Proposition 4.1 (Sharp forward contraction factor)**.** *For* $0 < \gamma < 2\alpha$, *the forward map* $I - \gamma f$, *and hence* G_γ, *is Lipschitz continuous with constant*

$$\widehat{\rho}_\gamma = \sqrt{1 - \frac{\mu}{\alpha}\gamma(2\alpha - \gamma)} \in [0, 1). \tag{16}$$

In particular, $\mathrm{VI}(X, f)$ *is a singleton.*

Proof. Set $d = x - y$ and $g = f(x) - f(y)$. By cocoercivity, $\|g\|^2 \le \langle g, d\rangle/\alpha$. Hence

$$\begin{aligned}\|d - \gamma g\|^2 &= \|d\|^2 - 2\gamma\langle g, d\rangle + \gamma^2\|g\|^2 \\ &\le \|d\|^2 - \left(2\gamma - \frac{\gamma^2}{\alpha}\right)\langle g, d\rangle \\ &\le \left[1 - \frac{\mu}{\alpha}\gamma(2\alpha - \gamma)\right]\|d\|^2.\end{aligned}$$

Nonexpansiveness of P_X transfers the same bound to G_γ. The factor (16) is the standard cocoercivity-sensitive forward contraction bound; see, for example, Briceño-Arias and Pustelnik [24]. Banach's theorem gives uniqueness of the fixed point and Proposition 2.2 identifies it with the variational-inequality solution. □

Remark 4.2 (Why the sharper estimate matters)**.** *A weaker estimate is obtained if strong monotonicity is first used only to deduce $\|f(x) - f(y)\| \ge \mu\|x - y\|$ and that lower bound is inserted after a cocoercive descent inequality. The resulting factor contains μ^2. Proposition 4.1 instead uses cocoercivity to eliminate* $\|g\|^2$ before *strong monotonicity is applied and produces the sharper coefficient μ/α. The distinction is substantial when $\mu\alpha \ll 1$.*

The next proposition combines contractivity and averagedness. Its purpose is to give a single global rate certificate valid throughout the standard over-relaxation window, against which the active-face rate can be compared.

Proposition 4.3 (Global relaxed contraction certificate)**.** *Let $T : \mathcal{H} \to \mathcal{H}$ be both η-averaged and ρ-contractive, where $1/2 \le \eta < 1$ and $0 \le \rho < 1$. Define*

$$R_\lambda = (1 - \lambda)I + \lambda T.$$

Then:

(i) for $0 < \lambda \le 1$,

$$\|R_\lambda x - R_\lambda y\| \le [1 - \lambda(1 - \rho)]\|x - y\|; \tag{17}$$

(ii) for $1 \le \lambda < 1/\eta$,

$$\|R_\lambda x - R_\lambda y\|^2 \le q_\eta(\lambda, \rho)^2\|x - y\|^2, \tag{18}$$

where

$$q_\eta(\lambda, \rho)^2 = 1 - \frac{\lambda(1 - \eta\lambda)}{1 - \eta}(1 - \rho^2) < 1. \tag{19}$$

Proof. Part (i) follows from the triangle inequality. For part (ii), put $d = x - y$ and $e = Tx - Ty$. The identity

$$\|(1 - \lambda)d + \lambda e\|^2 = (1 - \lambda)\|d\|^2 + \lambda\|e\|^2 + \lambda(\lambda - 1)\|d - e\|^2$$

and averagedness imply

$$\|d - e\|^2 \le \frac{\eta}{1 - \eta}(\|d\|^2 - \|e\|^2).$$

Substitution gives

$$\|R_\lambda x - R_\lambda y\|^2 \le \frac{1 - \eta - \lambda + \eta\lambda^2}{1 - \eta}\|d\|^2 + \frac{\lambda(1 - \eta\lambda)}{1 - \eta}\|e\|^2.$$

Since $\lambda < 1/\eta$, the coefficient of $\|e\|^2$ is positive; use $\|e\| \le \rho\|d\|$ and simplify. For $\eta \ge 1/2$, the multiplier $\lambda(1-\eta\lambda)/(1-\eta)$ lies in $(0, 1]$ on $[1, 1/\eta)$, so the right-hand side of (19) belongs to $[\rho^2, 1)$. □

Applying Proposition 4.3 with $T = G_\gamma$, $\eta = \eta_\gamma$ and $\rho = \widehat{\rho}_\gamma$ gives a global Q-linear guarantee for every fixed

$$0 < \lambda < 2 - \frac{\gamma}{2\alpha}.$$

The bound deliberately does not claim that over-relaxation is globally faster. In fact, the worst-case certificate in (19) deteriorates as λ approaches the upper endpoint. The practical acceleration observed in projected problems is therefore a local geometric phenomenon, which motivates the next section.

5 Active-face spectral tuning for box constraints

We now work in $\mathbb{R}^m$ with a box

$$X = [\ell, u] = \{x \in \mathbb{R}^m : \ell_i \le x_i \le u_i,\ i = 1, \dots, m\}. \tag{20}$$

Let p be the unique solution of (1). Define the lower-active, upper-active and free sets

$$\begin{aligned} \mathcal{A}_\ell &= \{i : p_i = \ell_i\}, \\ \mathcal{A}_u &= \{i : p_i = u_i\}, \\ \mathcal{F} &= \{i : \ell_i < p_i < u_i\}, \end{aligned}$$

and $\mathcal{A} = \mathcal{A}_\ell \cup \mathcal{A}_u$.

Definition 5.1 (Strict complementarity)**.** *The solution p satisfies strict complementarity if*

$$f_i(p) > 0 \quad (i \in \mathcal{A}_\ell), \qquad f_i(p) < 0 \quad (i \in \mathcal{A}_u), \qquad f_i(p) = 0 \quad (i \in \mathcal{F}). \tag{21}$$

For a box, (21) has an immediate geometric consequence. At a lower-active coordinate, $p_i - \gamma f_i(p) < \ell_i$; at an upper-active coordinate, $p_i - \gamma f_i(p) > u_i$; and at a free coordinate, $p_i - \gamma f_i(p) = p_i$ lies strictly inside the interval. These inequalities persist in a neighborhood of p.

5.1 Finite stabilization of the projection face

Theorem 5.2 (Finite face identification)**.** *Suppose f is continuous, the iteration* (12) *converges to p, and p satisfies strict complementarity. Then there exists N such that for every $n \ge N$, the projected point*

$$z_n = P_X(x_n - \gamma f(x_n))$$

has exactly the active/free pattern of p:

$$\begin{aligned} (z_n)_i &= \ell_i, & i &\in \mathcal{A}_\ell, \\ (z_n)_i &= u_i, & i &\in \mathcal{A}_u, \\ \ell_i < (z_n)_i &< u_i, & i &\in \mathcal{F}. \end{aligned}$$

Proof. Set $y(x) = x - \gamma f(x)$. For $i \in \mathcal{A}_\ell$, strict complementarity gives $y_i(p) < \ell_i$. Continuity of y implies $y_i(x) < \ell_i$ throughout a neighborhood of p, hence $(P_X y(x))_i = \ell_i$. The upper-active case is analogous. If $i \in \mathcal{F}$, then $y_i(p) = p_i \in (\ell_i, u_i)$, so $y_i(x)$ remains interior near p and the projection is locally inactive in that coordinate. Since $x_n \to p$, all these relations hold simultaneously for sufficiently large n. □

Remark 5.3. *Finite identification for forward–backward methods is known under far more general partly smooth structures [20]. Theorem 5.2 is a direct box-specific proof included because it exposes the exact derivative used below without invoking the machinery of partial smoothness.*

5.2 Local derivative and exact affine recurrence

Assume now that f is continuously differentiable near p and write

$$J = \nabla f(p).$$

Order the coordinates as $(\mathcal{A}, \mathcal{F})$ and partition J accordingly. In particular, $J_{\mathcal{F}\mathcal{F}}$ denotes the principal Jacobian submatrix formed from the rows and columns corresponding to the free coordinates, i.e., the variables that lie strictly inside their box bounds at p. Let

$$R_{\gamma,\lambda}(x) = (1 - \lambda)x + \lambda P_X(x - \gamma f(x)). \tag{22}$$

Theorem 5.4 (Post-identification linearization)**.** *Under the assumptions of Theorem 5.2, if f is C^1 near p, then $R_{\gamma,\lambda}$ is differentiable in a neighborhood of p in which the projection face is fixed, and*

$$DR_{\gamma,\lambda}(p) = M_{\gamma,\lambda} = \begin{pmatrix} (1-\lambda)I_{\mathcal{A}} & 0 \\ -\lambda\gamma J_{\mathcal{F}\mathcal{A}} & I_{\mathcal{F}} - \lambda\gamma J_{\mathcal{F}\mathcal{F}} \end{pmatrix}. \tag{23}$$

Consequently,

$$x_{n+1} - p = M_{\gamma,\lambda}(x_n - p) + o(\|x_n - p\|). \tag{24}$$

If $f(x) = Ax - b$ is affine, then (24) *is exact for all sufficiently large n, with $J = A$.*

Proof. After face identification, the active components of $G_\gamma(x)$ are locally constant at the corresponding bounds, so their derivative is zero. On the free coordinates the projection is the identity, hence

$$D(G_\gamma)_{\mathcal{F}}(p) = \begin{pmatrix} -\gamma J_{\mathcal{F}\mathcal{A}} & I_{\mathcal{F}} - \gamma J_{\mathcal{F}\mathcal{F}} \end{pmatrix}.$$

Combining this derivative with (22) gives (23). The differentiability expansion yields (24). For affine f, the post-identification map is affine, so no remainder term is present. □

The block triangular form is the key point: active components and free components contribute distinct scalar slopes to the local spectrum.

Corollary 5.5 (Local spectral factor)**.** *Suppose, in addition, that $J_{\mathcal{F}\mathcal{F}}$ is symmetric positive definite, with eigenvalues*

$$0 < m_F = \theta_1 \le \cdots \le \theta_s = L_F.$$

Then

$$\operatorname{spec}(M_{\gamma,\lambda}) = \{1 - \lambda : \mathcal{A} \neq \varnothing\} \cup \{1 - \lambda\gamma\theta_j : j = 1, \ldots, s\}, \tag{25}$$

and the governing local spectral factor is

$$r_{\mathrm{loc}}(\gamma,\lambda) = \rho(M_{\gamma,\lambda}) = \max\left\{ |1-\lambda| \ (\mathcal{A} \neq \varnothing), \max_j |1 - \lambda\gamma\theta_j| \right\}. \tag{26}$$

If $r_{\mathrm{loc}}(\gamma,\lambda) < 1$, then for every $q \in (r_{\mathrm{loc}}, 1)$ there are a neighborhood of p, an index N and a constant C such that

$$\|x_n - p\| \le C q^{n-N} \|x_N - p\|, \qquad n \ge N. \tag{27}$$

For the affine problem, r_{loc} is the worst-case asymptotic root factor of the exact post-identification recurrence; a particular initial error can exhibit a smaller observed factor if the dominant spectral modes are absent.

Proof. The spectrum of a block triangular matrix is the union of the spectra of the diagonal blocks. Since $J_{\mathcal{F}\mathcal{F}}$ is symmetric positive definite, the eigenvalues of the free block are exactly $1 - \lambda\gamma\theta_j$. The local linear convergence statement follows from the standard linearization theorem for fixed-point iterations whenever the derivative has spectral radius below one. In the affine case the iteration is eventually linear with matrix $M_{\gamma,\lambda}$. □

Remark 5.6 (Spectral radius versus one-step norm contraction)**.** *Even when $J_{\mathcal{F}\mathcal{F}}$ is symmetric positive definite, the full matrix $M_{\gamma,\lambda}$ need not be normal because $J_{\mathcal{F}\mathcal{A}}$ can be nonzero. Hence r_{loc} is the asymptotic root factor, not necessarily an exact Euclidean one-step contraction factor. Repeated dominant eigenvalues can also create polynomial transients. This is why Corollary 5.5 states the local bound for every $q > r_{\mathrm{loc}}$ and why the numerical comparisons distinguish asymptotic spectral tuning from finite-tolerance iteration counts.*

5.3 Closed-form spectral-radius-optimal outer relaxation

Define the positive slope set

$$\mathcal{S}_\gamma = \big(\{1\} \text{ if } \mathcal{A} \neq \varnothing\big) \cup \{\gamma\theta_j : j = 1, \dots, s\}. \tag{28}$$

Let

$$a_\gamma = \min \mathcal{S}_\gamma, \qquad b_\gamma = \max \mathcal{S}_\gamma. \tag{29}$$

Then (26) is simply

$$r_{\mathrm{loc}}(\gamma, \lambda) = \max_{s \in \mathcal{S}_\gamma} |1 - \lambda s|.$$

Because the maximum of $|1 - \lambda s|$ over a positive interval occurs at an endpoint, the scalar minimax problem has a closed form.

Theorem 5.7 (Locally spectral-radius-optimal scalar relaxation)**.** *Assume $\mathcal{S}_\gamma \neq \varnothing$. Over $\lambda > 0$, the local spectral factor* (26) *is minimized uniquely at*

$$\boxed{\lambda^*_{\mathrm{loc}}(\gamma) = \frac{2}{a_\gamma + b_\gamma}} \tag{30}$$

with optimal factor

$$\boxed{r^*_{\mathrm{loc}}(\gamma) = \frac{b_\gamma - a_\gamma}{a_\gamma + b_\gamma}.} \tag{31}$$

Moreover,

$$\lambda^*_{\mathrm{loc}} > 1 \quad \Longleftrightarrow \quad a_\gamma + b_\gamma < 2. \tag{32}$$

Proof. For $a = a_\gamma$ and $b = b_\gamma$, $0 < a \leq b$, and

$$r_{\mathrm{loc}}(\gamma, \lambda) = \max\{|1 - \lambda a|, |1 - \lambda b|\}.$$

At the minimizer, the endpoint errors must equioscillate; otherwise one can change λ slightly to reduce the larger endpoint without increasing the smaller to the same level. Thus

$$1 - \lambda a = -(1 - \lambda b),$$

which gives (30). Substitution yields (31). Finally, $2/(a + b) > 1$ if and only if $a + b < 2$. □

Theorem 5.7 gives a precise interpretation of over-relaxation: the parameter is beneficial only in identifiable spectral regimes, and the claim concerns the local spectral radius.

Corollary 5.8 (Conservative forward steps force local over-relaxation)**.** *Suppose both $\mathcal{A}$ and $\mathcal{F}$ are nonempty and*

$$0 < \gamma < \frac{1}{L_F}. \tag{33}$$

Then

$$a_\gamma = \gamma m_F, \qquad b_\gamma = 1, \tag{34}$$

and hence

$$\lambda^*_{\mathrm{loc}} = \frac{2}{1 + \gamma m_F} > 1, \qquad r^*_{\mathrm{loc}} = \frac{1 - \gamma m_F}{1 + \gamma m_F}. \tag{35}$$

Thus genuine over-relaxation is locally spectral-radius-optimal whenever the forward step is conservative relative to the curvature of the identified free face.

Remark 5.9 (Intermediate and large forward steps)**.** *If $1/L_F \le \gamma \le 1/m_F$, then the slope 1 lies inside $[\gamma m_F, \gamma L_F]$ and*

$$\lambda^*_{\mathrm{loc}} = \frac{2}{\gamma(m_F + L_F)}, \qquad r^*_{\mathrm{loc}} = \frac{L_F - m_F}{L_F + m_F}.$$

Within this interval, $\lambda^*_{\mathrm{loc}} > 1$ precisely when $\gamma < 2/(m_F + L_F)$, equals one at $\gamma = 2/(m_F + L_F)$, and is smaller than one afterwards. Thus the same problem can favor over-relaxation, no relaxation, or under-relaxation depending on the chosen forward step.

Proposition 5.10 (Joint local spectral tuning of the forward and outer parameters)**.** *Suppose both $\mathcal{A}$ and $\mathcal{F}$ are nonempty and $J_{\mathcal{F}\mathcal{F}}$ is symmetric positive definite with extreme eigenvalues m_F and L_F. If $\gamma > 0$ and $\lambda > 0$ are both free at the local level, then*

$$\inf_{\gamma>0,\,\lambda>0} r_{\mathrm{loc}}(\gamma, \lambda) = \frac{L_F - m_F}{L_F + m_F}. \tag{36}$$

The complete family attaining this value is

$$\lambda\gamma = \frac{2}{m_F + L_F}, \qquad \frac{2m_F}{m_F + L_F} \le \lambda \le \frac{2L_F}{m_F + L_F}. \tag{37}$$

In particular,

$$\lambda = 1, \qquad \gamma = \frac{2}{m_F + L_F} \tag{38}$$

is jointly spectral-radius-optimal whenever that forward step is otherwise admissible.

Proof. Set $\tau = \lambda\gamma$. The contribution of the free block is

$$\max_{\theta \in [m_F, L_F]} |1 - \tau\theta|,$$

whose minimum over $\tau > 0$ is the classical two-endpoint minimax value $(L_F - m_F)/(L_F + m_F)$, attained uniquely at $\tau = 2/(m_F + L_F)$. Therefore the full factor cannot be smaller than this value. At this τ, the free contribution equals $r_F = (L_F - m_F)/(L_F + m_F)$. The active contribution satisfies $|1 - \lambda| \le r_F$ exactly when $1 - r_F \le \lambda \le 1 + r_F$, which is precisely the interval in (37). Hence every pair in (37) attains the lower bound, and no other pair can do so. □

Remark 5.11 (What outer relaxation contributes)**.** *Proposition 5.10 prevents an overstatement of the role of λ. If the forward step can be retuned freely after the face is known, the unrelaxed choice* (38) *already attains the best local scalar spectral radius. Outer relaxation is most useful when γ is fixed in advance, kept conservative for global reasons, or is constrained by information available before the active face is identified. In that setting λ changes the effective free-face step $\lambda\gamma$ while simultaneously damping the active error through the distinct factor $1 - \lambda$.*

5.4 A globally safeguarded spectral selector

Let $\Lambda_\gamma = 1/\eta_\gamma$ be the standard global relaxation cap. Fix $\varepsilon > 0$ and a lower safeguard $\lambda_{\min} > 0$ such that $\lambda_{\min} < \Lambda_\gamma - \varepsilon$. With exact spectral extremes, define

$$\lambda_{\mathrm{sf}} = \Pi_{[\lambda_{\min},\, \Lambda_\gamma - \varepsilon]}\left(\frac{2}{a_\gamma + b_\gamma}\right), \tag{39}$$

where $\Pi_{[c,d]}$ denotes scalar projection onto $[c, d]$. Because $\lambda \mapsto r_{\mathrm{loc}}(\gamma, \lambda)$ is convex and has the unique unconstrained minimizer in Theorem 5.7, (39) is also the exact minimizer of the local spectral radius over the safeguarded interval.

Algorithm 1 Pattern-reset safeguarded active-face relaxation

Require: x_0, $\gamma \in (0, 2\alpha)$, $K \geq 2$, $\varepsilon > 0$, $\lambda_{\min} > 0$, activity tolerance $\varepsilon_{\mathrm{act}} > 0$.

$\Lambda_\gamma \leftarrow 2 - \gamma/(2\alpha)$; set $\lambda_0 \leftarrow 1$ and no face is locked.
for $n = 0, 1, 2, \ldots$ **do**
  $z_n \leftarrow P_X(x_n - \gamma f(x_n))$ and classify each coordinate as lower-active, upper-active or free using $\varepsilon_{\mathrm{act}}$.
  if a locked pattern exists and the current pattern differs from it **then**
    unlock the pattern, reset the stability counter, and set $\lambda_n \leftarrow 1$.
  end if
  if no pattern is locked and the current pattern has persisted for K iterations **then**
    estimate $\widehat{m}_F, \widehat{L}_F$ from the current free–free Jacobian.
    form $\widehat{a}_\gamma, \widehat{b}_\gamma$ from the active slope 1 and the free slopes $\gamma\widehat{m}_F, \gamma\widehat{L}_F$.
    $\lambda_n \leftarrow \Pi_{[\lambda_{\min},\, \Lambda_\gamma - \varepsilon]}\Big(2/(\widehat{a}_\gamma + \widehat{b}_\gamma)\Big)$ and lock the pattern.
  else if no pattern is locked **then**
    $\lambda_n \leftarrow 1$.
  end if
  $x_{n+1} \leftarrow x_n + \lambda_n(z_n - x_n)$.
end for

Corollary 5.12 (When the unconstrained local minimizer is globally admissible)**.** *The value in* (30) *lies strictly inside the standard global averagedness interval if and only if*

$$a_\gamma + b_\gamma > 2\eta_\gamma = \frac{4\alpha}{4\alpha - \gamma}. \tag{40}$$

If equality holds, the unconstrained minimizer is the excluded global endpoint and must be clipped.

Proof. Because all quantities are positive, $2/(a_\gamma + b_\gamma) < 1/\eta_\gamma$ is equivalent to $a_\gamma + b_\gamma > 2\eta_\gamma$. □

For implementation, the true active face and the true spectral extremes are not known in advance. We therefore use a pattern-reset rule with an activity tolerance $\varepsilon_{\mathrm{act}} > 0$. For a projected point z_n, coordinate i is declared lower-active if $(z_n)_i \leq \ell_i + \varepsilon_{\mathrm{act}}$, upper-active if $(z_n)_i \geq u_i - \varepsilon_{\mathrm{act}}$, and free otherwise. A face is tentatively accepted only after this tolerance-based pattern has remained unchanged for K consecutive iterations. The free-face extremes are then computed or estimated at the current iterate, the spectral value is clipped as in (39), and that value is held while the pattern remains unchanged. If the pattern changes, the method immediately returns to $\lambda = 1$ and restarts the stability counter.

Proposition 5.13 (Global safety of the adaptive selector)**.** *Suppose f is α-cocoercive,* $\mathrm{VI}(X, f) \neq \emptyset$, $0 < \gamma < 2\alpha$, *and every relaxation produced by the adaptive rule belongs to* $[\lambda_{\min}, \Lambda_\gamma - \varepsilon]$. *Then the iterates converge globally as in Theorem 3.2. Consequently, a premature or inaccurate face declaration cannot destroy convergence as long as the selected relaxation is clipped to this interval. Under strong monotonicity the limit is the unique solution. If, in addition, the assumptions of Corollary 5.5 and strict complementarity hold, the true face is eventually identified. Once the true face is retained and the exact free-face extremes are used, the asymptotic root factor is $r_{\mathrm{loc}}(\gamma, \lambda_{\mathrm{sf}})$; if no clipping occurs, this equals the minimum in Theorem 5.7.*

Proof. The clipping rule gives uniform bounds $0 < \lambda_{\min} \leq \lambda_n \leq \Lambda_\gamma - \varepsilon < 1/\eta_\gamma$, so Theorem 3.2 applies independently of how the tentative patterns are generated. Strong monotonicity gives uniqueness. Finite identification follows from Theorem 5.2; the final local statement follows from Corollary 5.5 and Theorem 5.7. □

The spectral calculation is deliberately performed only when a pattern is locked, rather than at every iteration. For a small or moderate free face, the extreme eigenvalues can be computed directly. For large

problems they can be estimated by Lanczos or power-type procedures using Jacobian–vector products, so the full free Jacobian need not be formed. Inexact spectral estimates affect local tuning quality but not the global convergence certificate because the final relaxation remains clipped to the averagedness interval.

6 Numerical experiments

All experiments use the fixed-point residual stopping rule

$$\|x_n - G_\gamma x_n\| < 10^{-8}. \tag{41}$$

The adaptive runs use $K = 5$, $\lambda_{\min} = 0.25$, $\varepsilon = 10^{-3}$ and $\varepsilon_{\mathrm{act}} = 10^{-12}$. Matrix constructions, random seeds, adaptive face events, tables and figures are reproduced by the supplementary Python script. Exact eigendecompositions are used in these moderate-dimensional tests so that parameter-selection error is not confounded with the theory; the practical large-scale alternatives are described after Algorithm 1.

6.1 Two-dimensional sanity check

Let

$$X = [0,1]^2, \qquad f(x) = Ax - b, \qquad A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \qquad b = \begin{pmatrix} 1.2 \\ 0.2 \end{pmatrix}.$$

The solution is $p = (1, 0.1)^T$, with one active and one free coordinate. For $\gamma = 0.1$, $m_F = L_F = 2$ and Theorem 5.7 gives $\lambda^*_{\mathrm{loc}} = 5/3$. Starting from $x_0 = (0,1)^T$, the unrelaxed method requires 75 iterations, whereas the fixed spectral choice requires 44. Algorithm 1 also requires 44 iterations; it locks the final one-active/one-free face at iteration 15 and then uses $\lambda = 5/3$. If γ is allowed to be retuned, Proposition 5.10 gives $\gamma = 2/(m_F + L_F) = 0.5$ with $\lambda = 1$, which reaches the tolerance in 3 iterations. This elementary case already illustrates both sides of the theory: outer relaxation accelerates a fixed conservative forward step, but it does not replace forward-step tuning when that tuning is available.

6.2 Dense benchmark with a globally constrained forward step

The main affine experiment is designed to isolate the setting in which outer relaxation has a distinct purpose. We use 120 variables, ordered with 60 active coordinates followed by 60 free coordinates, and construct

$$A = \begin{pmatrix} A_A & C^T \\ C & A_F \end{pmatrix}, \qquad A_A = Q_A \operatorname{diag}(\xi_i^A) Q_A^T, \qquad A_F = Q_F \operatorname{diag}(\xi_i^F) Q_F^T, \tag{42}$$

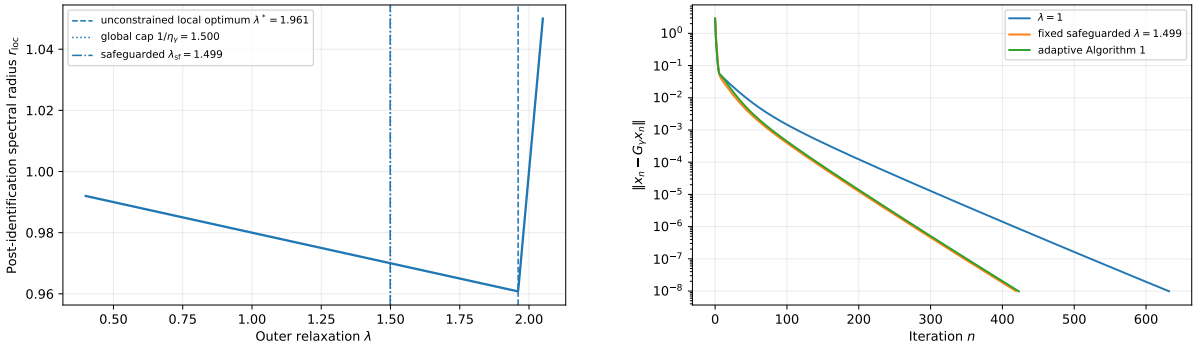
where ξ_i^A are equally spaced in $[0.4M, M]$, ξ_i^F are equally spaced in $[1, 3]$, and $C = 0.15R/\sqrt{60}$. The orthogonal factors and R are generated deterministically with NumPy seed 20260811. Thirty active coordinates are prescribed at each bound and the remaining 60 solution coordinates are equally spaced in $[0.15, 0.85]$; the right-hand side is then chosen so that the active residuals are 0.2 and -0.2, giving strict complementarity.

The free block has $m_F = 1$ and $L_F = 3$, so the locally optimal unrelaxed forward step would be $\gamma_F = 2/(m_F + L_F) = 0.5$. In contrast, the global Lipschitz constant is approximately M, and the globally admissible interval is $0 < \gamma < 2/L$. We deliberately take $\gamma = 1/L$, for which the global outer-relaxation cap is $\Lambda_\gamma = 1.5$. Thus γ_F is globally inadmissible for all three conditioning levels below. This is precisely the regime in which changing λ can safely compensate for a forward step that must remain globally conservative.

For all three cases the unconstrained local spectral minimizer lies beyond the global cap and is therefore clipped to $\lambda_{\mathrm{sf}} = 1.499$. Algorithm 1 uses the same safeguarded value once the final face is locked. Its iteration reductions are 33.3%, 32.9% and 33.1%, respectively. The small difference between the fixed safeguarded run and the adaptive run is the price of identifying the face rather than supplying it in advance.

Table 1: Dense 120-dimensional benchmark family. The adaptive reduction is relative to the fixed-γ unrelaxed iteration.

M	L	$2/L$	λ^*_{loc}	$\lambda = 1$	fixed λ_{sf}	adaptive	final lock
10	10.0028	0.1999	1.8182	135	87	90	8
25	25.0010	0.0800	1.9231	328	217	220	9
50	50.0005	0.0400	1.9608	632	419	423	10



(a) Local spectral radius for $M = 50$.

(b) Residual histories for $M = 50$.

Figure 1: Globally constrained affine benchmark. The unconstrained local minimizer is outside the global relaxation interval, so the safeguarded endpoint is the best globally certified scalar outer relaxation.

6.3 Nonlinear 80-dimensional box VI and comparison with retuning γ

Consider

$$f(x) = Ax + \tau \tanh(x) - b, \qquad \tau = 0.5, \tag{43}$$

with componentwise tanh and a deterministic dense symmetric positive definite matrix A with eigenvalues equally spaced in $[1, 2]$ (NumPy seed 20260810). The prescribed solution has 20 lower-active, 40 free and 20 upper-active coordinates, with active residuals 0.2 and -0.2. Here

$$f = \nabla\Phi, \qquad \Phi(x) = \frac{1}{2}x^T Ax + \tau \sum_{i=1}^{80} \log(\cosh x_i) - b^T x,$$

and

$$\nabla^2\Phi(x) = \nabla f(x) = A + \tau\, \mathrm{diag}(\mathrm{sech}^2 x_i).$$

Because $A \succeq I$ and $0 \le \mathrm{sech}^2 x_i \le 1$, the function Φ is strongly convex and its gradient is globally L-Lipschitz with $L \le 2 + \tau = 2.5$. The Baillon–Haddad theorem therefore applies to the convex differentiable function Φ and gives that $f = \nabla\Phi$ is $1/2.5$-cocoercive [19]. With $\gamma = 0.1$, the global relaxation cap is 1.875. At the solution,

$$m_F = 1.5258223354, \qquad L_F = 2.3108508164,$$

so $\lambda^*_{\mathrm{loc}} = 1.7352341046$. The jointly optimal unrelaxed step is $\gamma_F = 2/(m_F + L_F) = 0.5212849573$, which is globally admissible because $\gamma_F < 2/2.5 = 0.8$.

The adaptive method locks the final 40-active/40-free pattern at iteration 13 and selects $\lambda = 1.734909$, using current-face estimates $\widehat{m}_F = 1.527981$ and $\widehat{L}_F = 2.308991$. It therefore reaches the tolerance in 55 iterations, compared with 93 for the fixed-γ unrelaxed method. The value near the global cap is much slower, confirming that global admissibility does not imply local speed. On the other hand, the globally admissible retuned forward step reaches the tolerance in 12 iterations. This is not a contradiction of the outer-relaxation result: it is the numerical counterpart of Proposition 5.10, and it makes the intended scope of the method explicit.

Table 2: Nonlinear 80-dimensional box VI.

Configuration	γ	λ	Iterations
Fixed γ, unrelaxed	0.100000	1.000000	93
Fixed γ, spectral minimizer	0.100000	1.735234	56
Adaptive Algorithm 1	0.100000	adaptive	55
Near global cap	0.100000	1.865000	115
Retuned γ, unrelaxed	0.521285	1.000000	12

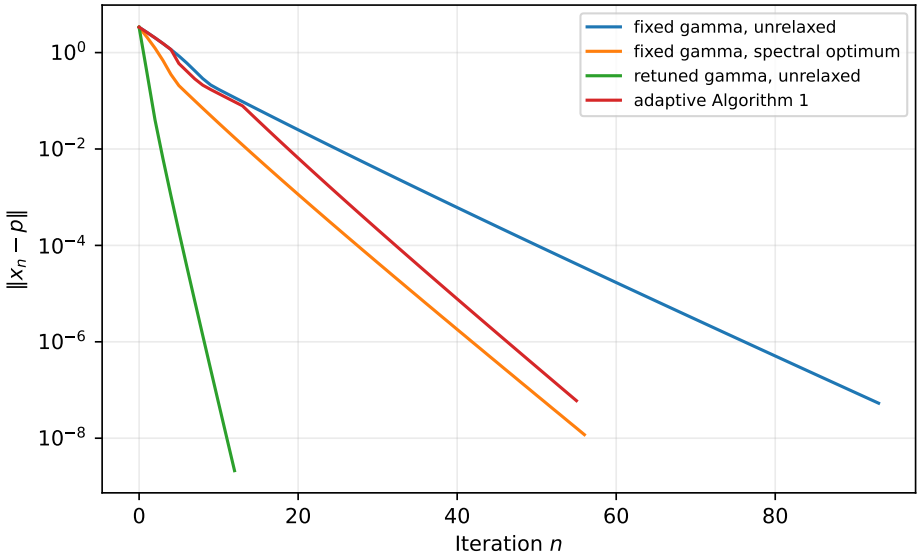


Figure 2: Error histories for the nonlinear problem. Outer relaxation accelerates the fixed conservative step, whereas direct forward-step retuning is preferable when the locally optimal forward step is itself globally admissible.

6.4 Interpretation of the computational evidence

The experiments establish four points relevant to the theory. First, the adaptive selector itself reproduces the fixed spectral choice without assuming the final face in advance. Second, clipping makes tentative or imperfect face declarations harmless for global convergence. Third, the dense stress family demonstrates a regime in which the locally preferred forward step is incompatible with the global forward bound, so outer relaxation provides a genuine globally certified acceleration mechanism. Fourth, the nonlinear experiment shows the complementary regime: when γ can be safely retuned, an unrelaxed member of the jointly optimal family can outperform fixed-γ over-relaxation. The numerical evidence therefore supports a conditional, mechanism-based use of over-relaxation rather than a universal acceleration claim.

7 Discussion and relation to existing theory

The contribution is best understood by separating three levels of theory. At the global level, the averagedness coefficient and the interval $0 < \lambda < 1/\eta_\gamma$ are standard consequences of forward–backward calculus [17, 21, 22]; the contraction estimate under strong monotonicity is also a known baseline [9, 24]. We use these facts to certify every adaptive relaxation, not as novelty claims.

At the parameter-selection level, performance-estimation methods can derive tight contraction factors and optimize parameters for broad operator-splitting classes [31]. Our calculation is narrower but more explicit: once a box face is identified, the projection derivative is a diagonal active mask and the post-identification matrix separates the active slope 1 from the free slopes $\gamma\theta_j$. The resulting one-dimensional

minimax problem has the closed form in Theorem 5.7. No semidefinite performance-estimation problem is required because the active-face geometry has already reduced the local operator to a finite set of scalar slopes.

At the box-optimization level, spectral projected-gradient and active-set methods use curvature restricted to inactive variables to tune the forward step or switch to a reduced-space method [12–14, 30, 32]. The present selector does not compete with those methods as a general-purpose optimizer. Its role is different: it keeps the globally chosen forward map G_γ fixed and tunes the outer relaxation. Proposition 5.10 precisely quantifies this difference. If γ can be retuned to $2/(m_F + L_F)$ without violating other requirements, $\lambda = 1$ is already locally optimal. If that forward step is fixed, costly to change, or incompatible with the global bound, λ provides a separate control variable while preserving the original forward map and its global certificate.

The closed-form analysis requires a box and a symmetric positive definite free–free Jacobian. The full post-identification matrix can still be nonnormal through the active-to-free coupling, so spectral-radius tuning predicts the asymptotic root factor rather than all finite transients. For nonsymmetric free Jacobians, eigenvalues alone may be insufficient because nonnormality can dominate transient behavior. For general convex sets, the derivative of the projection is governed by tangent geometry instead of a diagonal mask. Extending the same parameter-selection principle to polyhedral and partly smooth manifolds is a natural next step, but it requires additional geometric structure and is not claimed here.

8 Conclusion

The global over-relaxation window for a projected-forward method is known; the new question addressed here is how to select the outer relaxation after the projection reveals the active geometry. Under strict complementarity, the active face of a box is identified finitely. For a C^1 operator with symmetric positive definite free–free Jacobian, the post-identification spectrum consists of the active factor $1 - \lambda$ and free factors $1 - \lambda\gamma\theta_j$. Their minimax balance yields the explicit spectral-radius minimizer $\lambda^*_{\mathrm{loc}} = 2/(a_\gamma + b_\gamma)$ and the criterion $a_\gamma + b_\gamma < 2$ for genuine over-relaxation.

The joint (γ, λ) analysis also delineates the limit of the result: if the forward step can be freely and safely retuned, an unrelaxed member is already locally optimal. Outer relaxation is therefore most useful when the forward map must remain fixed or globally conservative. The pattern-reset safeguarded selector implements this principle without sacrificing global convergence. In the dense benchmark family, the locally preferred forward step is globally inadmissible and adaptive outer relaxation reduces iteration counts by about one third; in the nonlinear example, retuning γ is globally admissible and is correspondingly superior. Together, the theory and experiments provide a precise criterion for when over-relaxation is useful rather than treating $\lambda > 1$ as a generic acceleration heuristic.

Statements and Declarations

Funding. No funding was received for conducting this study.
Competing interests. The authors have no relevant financial or non-financial interests to disclose.
Ethics approval and consent to participate. Not applicable.
Data and code availability. No external dataset was generated or analyzed. The deterministic Python script used to reproduce all numerical tables and figures is supplied as supplementary material.

References

[1] P. Hartman and G. Stampacchia, On some non-linear elliptic differential-functional equations, *Acta Math.* 115 (1966), 271–310.

[2] J.-L. Lions and G. Stampacchia, Variational inequalities, *Comm. Pure Appl. Math.* 20 (1967), 493–519.

[3] W. R. Mann, Mean value methods in iteration, *Proc. Amer. Math. Soc.* 4 (1953), 506–510.

[4] M. A. Krasnosel'skiĭ, Two remarks on the method of successive approximations, *Uspekhi Mat. Nauk* 10(63) (1955), 123–127.

[5] H. H. Schaefer, Über die Methode sukzessiver Approximationen, *Jahresber. Deutsch. Math.-Verein.* 59 (1957), 131–140.

[6] F. E. Browder and W. V. Petryshyn, Construction of fixed points of nonlinear mappings in Hilbert space, *J. Math. Anal. Appl.* 20 (1967), 197–228.

[7] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.* 73 (1967), 591–597.

[8] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.* 14 (1976), 877–898.

[9] H. G. Chen and R. T. Rockafellar, Convergence rates in forward–backward splitting, *SIAM J. Optim.* 7 (1997), 421–444.

[10] W. Takahashi and M. Toyoda, Weak convergence theorems for nonexpansive mappings and monotone mappings, *J. Optim. Theory Appl.* 118 (2003), 417–428.

[11] H. Iiduka and W. Takahashi, Strong convergence theorems for nonexpansive nonself-mappings and inverse-strongly-monotone mappings, *J. Convex Anal.* 11 (2004), 69–79.

[12] E. G. Birgin, J. M. Martínez and M. Raydan, Nonmonotone spectral projected gradient methods on convex sets, *SIAM J. Optim.* 10(4) (2000), 1196–1211.

[13] S. Crisci, V. Ruggiero and L. Zanni, Steplength selection in gradient projection methods for box-constrained quadratic programs, *Appl. Math. Comput.* 356 (2019), 312–327, `https://doi.org/10.1016/j.amc.2019.03.039`.

[14] W. W. Hager and H. Zhang, A new active set algorithm for box constrained optimization, *SIAM J. Optim.* 17(2) (2006), 526–557.

[15] Y. Censor, A. Gibali and S. Reich, The subgradient extragradient method for solving variational inequalities in Hilbert space, *J. Optim. Theory Appl.* 148 (2011), 318–335.

[16] H.-K. Xu, Averaged mappings and the gradient-projection algorithm, *J. Optim. Theory Appl.* 150 (2011), 360–378.

[17] P. L. Combettes and I. Yamada, Compositions and convex combinations of averaged nonexpansive operators, *J. Math. Anal. Appl.* 425(1) (2015), 55–70.

[18] Y. Malitsky, Projected reflected gradient methods for monotone variational inequalities, *SIAM J. Optim.* 25 (2015), 502–520.

[19] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed., Springer, Cham, 2017.

[20] J. Liang, J. Fadili and G. Peyré, Activity identification and local linear convergence of forward–backward-type methods, *SIAM J. Optim.* 27(1) (2017), 408–437.

[21] F. Cui, Y. Tang and C. Zhu, Convergence analysis of a variable metric forward–backward splitting algorithm with applications, *J. Inequal. Appl.* 2019 (2019), Article 141, 1–27.

[22] P. Cholamjiak, D. V. Hieu and Y. J. Cho, Relaxed forward–backward splitting methods for solving variational inclusions and applications, *J. Sci. Comput.* 88 (2021), Article 85.

[23] M. O. Olatinwo and O. J. Omidire, Some new convergence and stability results for Jungck generalized pseudo-contractive and Lipschitzian type operators using hybrid iterative techniques in the Hilbert space, *Rend. Circ. Mat. Palermo, II Ser.* 72 (2023), 1067–1086.

[24] L. M. Briceño-Arias and N. Pustelnik, Theoretical and numerical comparison of first order algorithms for cocoercive equations and smooth convex optimization, *Signal Process.* 206 (2023), 108900.

[25] R. I. Boţ and D.-K. Nguyen, Fast Krasnosel'skiĭ–Mann algorithm with a convergence rate of the fixed point iteration of $o(1/k)$, *SIAM J. Numer. Anal.* 61(6) (2023), 2813–2843.

[26] P. L. Combettes, The geometry of monotone operator splitting methods, *Acta Numerica* 33 (2024), 487–632.

[27] K. O. Aremu, M. I. Mona and M. Ibrahim, A modified self-adaptive inertial Tseng algorithm for solving a quasimonotone variational inequality and fixed point problems in real Hilbert space, *J. Anal.* 33 (2025), 319–340.

[28] G. Ogwo, C. Izuchukwu, O. Oyewole, Y. Shehu and J.-C. Yao, Projected reflected gradient method with momentum terms for variational inequalities, *Numer. Algebra Control Optim.* 17 (2026), 113–136.

[29] Y. Dong, A fully relaxed Tseng's splitting method, *Set-Valued Var. Anal.* (2026), online first, https://doi.org/10.1007/s11228-026-00790-w.

[30] E. G. Birgin and J. M. Martínez, Large-scale active-set box-constrained optimization method with spectral projected gradients, *Comput. Optim. Appl.* 23(1) (2002), 101–125, https://doi.org/10.1023/A:1019928808826.

[31] E. K. Ryu, A. B. Taylor, C. Bergeling and P. Giselsson, Operator splitting performance estimation: tight contraction factors and optimal parameter selection, *SIAM J. Optim.* 30(3) (2020), 2251–2271, https://doi.org/10.1137/19M1304854.

[32] S. Crisci, F. Porta, V. Ruggiero and L. Zanni, Hybrid limited memory gradient projection methods for box-constrained optimization problems, *Comput. Optim. Appl.* 84 (2023), 151–189, https://doi.org/10.1007/s10589-022-00409-4.