AIMS Mathematics, 11(10): 33673–33726.
DOI: 10.3934/math.20261317
Received: 08 August 2026
Revised: 22 September 2026
Accepted: 05 October 2026
Published: 15 October 2026

https://www.aimspress.com/journal/Math

Research article

A safeguarded nonmonotone Barzilai–Borwein smoothing Newton method for weighted complementarity problems: convergence analysis and numerical validation

Teng Wu and Xiangjing Liu*

School of Science, Xi'an Technological University, No. 2 Xuefuzhonglu Road, Weiyang District, Xi'an, Shaanxi 710021, China

* **Correspondence:** Email: liuxiangjing504@163.com; Tel: +8617789103274.

Abstract: We developed a nonmonotone smoothing Newton method for weighted complementarity problems (WCPs) in $\mathbb{R}^n$. After full-step rejection, a safeguarded Barzilai–Borwein (BB) factor with step-length memory initialized backtracking. Under the structural rank–monotonicity condition and a bounded initial level set, we proved that the residual tends to zero, and every accumulation point solves the WCP; uniqueness gives whole-sequence convergence. Nonsingularity of every matrix in the generalized Jacobian at an accumulation point, together with a locally Lipschitz derivative of the original equation block, yielded quadratic convergence and eventual full-step acceptance. For symmetric positive-definite linear problems, we verified level-set boundedness, uniqueness, and local regularity without strict complementarity. On 150 paired nonconvex instances, rejected-trial counts were significantly lower than those of a one-step comparator in five of six settings after Holm correction, whereas iteration-count differences were not significant. An ablation study supported the BB initializer's benefit without establishing a general timing or success-rate advantage. Further tests included an independent recent one-step comparator, monotone linear tests, sparse positive-weight problems with up to 102400 complementarity pairs, and three zero-weight partial differential equation (PDE) benchmarks with grid-refinement and independent reference checks. On nonconvex tests, success depends on the fixed complementarity parameter and initialization; failures at larger parameter values persist from the original signed starts. Some signed-start Bratu failures persisted under a larger budget. These results support reduced post-rejection line-search work, not uniform robustness across problem classes.

Keywords: weighted complementarity problem; complementarity function; Barzilai–Borwein method; nonmonotone smoothing Newton algorithm; global residual convergence
Mathematics Subject Classification: 65K05, 90C33

Note: The table abbreviations (see Tanle 1) are defined below. Unless stated otherwise, the prefix A denotes an arithmetic mean over successful runs in the stated group, using whole-run totals for work counts. Failed runs enter the success rate (SR) but not these means.

Table 1. The abbreviations defined.

Abbreviation	The meaning
SR	Success rate: The percentage of attempted runs satisfying both stopping criteria above. "Successes" gives successful/attempted run counts.
AIT	Average iteration count: The mean number of outer Newton iterations completed with an accepted update.
ATime	Average elapsed time: The mean wall-clock solver time in seconds, not process CPU time.
ANF	Average number of function evaluations: The mean count of smoothing-residual mapping evaluations, including initialization, trial points, and re-evaluation at accepted points. Each vector evaluation counts once, not once per component.
ALS	Average number of linear solves: The mean count of reduced Newton systems solved, not of matrix factorizations.
ARED	Average number of rejected trials: The mean total rejected-trial count per successful run, including a rejected full-step trial and all subsequently rejected backtracking trials. Accepted trials are excluded; all methods use this convention.
ABB	Average number of BB-initialized searches: The mean count of full-step-rejected iterations on which a finite positive raw Barzilai–Borwein factor is available and the safeguarded BB initializer enters the search. It does not count accepted BB trial steps.
AKKT	Average final KKT residual: The mean terminal common Karush–Kuhn–Tucker residual R_{KKT} defined above, not the smoothing residual or the error relative to an exact solution.

1. Introduction

Complementarity problems (CPs) provide a standard framework for modeling mutually exclusive equilibrium conditions arising in mathematical programming and equilibrium theory. The standard CP is formulated as

$$x \geq 0, \quad F(x) \geq 0, \quad x^{\top} F(x) = 0,$$

which requires at least one member of each pair $(x_i, F_i(x))$ to be zero [1]. Linear programming, quadratic programming, Nash equilibrium problems, and traffic network equilibrium problems can be formulated as complementarity problems, which highlights the broad relevance of CPs in optimization theory and equilibrium modeling.

As a natural generalization of the standard CP, Potra introduced weighted complementarity problems (WCPs), focusing particularly on linear weighted complementarity problems (LWCPs) [1]. An LWCP has the form

$$x \geq 0, \quad s \geq 0, \quad Px + Qs + Ry = a, \quad x \circ s = w, \tag{1.1}$$

where $P, Q \in \mathbb{R}^{(n+m)\times n}$, $R \in \mathbb{R}^{(n+m)\times m}$, $a \in \mathbb{R}^{n+m}$, and the weight vector $w \in \mathbb{R}^n_+$ are given; the unknowns are $x, s \in \mathbb{R}^n$, and $y \in \mathbb{R}^m$. Potra proposed two interior-point algorithms and analyzed their time complexity. Later, researchers further developed full-Newton and infeasible interior-point variants for special or monotone LWCPs [2–4].

Potra's interior-point developments focused primarily on LWCPs, whereas general nonlinear WCPs have also attracted growing interest. In this paper, we consider WCPs in $\mathbb{R}^n$, which are typically formulated as

$$x \geq 0, \quad s \geq 0, \quad F(x, s, y) = 0, \quad x \circ s = w, \tag{1.2}$$

where $x, s, w \in \mathbb{R}^n$, $w \geq 0$, $y \in \mathbb{R}^m$, and $F\colon \mathbb{R}^{2n+m} \to \mathbb{R}^{n+m}$ is continuously differentiable. Here, "$\circ$" denotes the Hadamard product, that is, $(x \circ s)_i = x_i s_i$. When $w = 0$, the WCP reduces to a mixed nonlinear complementarity problem in $\mathbb{R}^n$.

Many algorithms have been developed for WCPs. Two important classes are interior-point and smoothing Newton methods. Under appropriate conditions, smoothing Newton methods can attain superlinear or quadratic local convergence [5]. Zhang developed a smoothing Newton algorithm for the weighted linear case [6]. Tang proposed a nonmonotone smoothing variant for large-scale LWCPs [7]. Tang et al. studied a smoothing Newton method for nonlinear complementarity problems [8]. For general WCPs, Tang and Zhang developed a nonmonotone smoothing Newton algorithm [9]. Tang et al. subsequently proposed an accelerated smoothing Newton method [10]. Tang and Zhou studied WCP solvability and a smoothing Newton algorithm under a local error bound [11].

The nonmonotone smoothing Newton methods in [9, 12] use line searches to globalize their Newton steps. Barzilai–Borwein (BB) step sizes use consecutive iterate and gradient differences to encode local secant information [13, 14].

BB-type steps have been incorporated into projection methods for variational inequalities and modified gradient methods for quadratic minimization [15, 16]. In these settings, BB information scales a projected or gradient-based step, whose construction differs from the smoothing Newton direction considered here.

In complementarity-related work, Li [17] developed a smoothing nonmonotone BB gradient method for stochastic linear complementarity problems through a box-constrained minimization reformulation. Ma et al. introduced stabilized BB step sizes for stochastic ordinal embedding [18]. In multiobjective optimization, Chen et al. used BB information to scale the objective gradients in the direction-finding subproblem [19]. Xu proposed parameterized BB step sizes based on interpolated least squares for unconstrained optimization [20]. These methods concern different problem classes and direction constructions; they provide conceptual context but are not interchangeable numerical baselines for the deterministic WCP smoothing Newton method studied here.

The question considered here is more specific: Can a scalar inverse-secant estimate reduce backtracking work without changing the smoothing Newton direction or its eventual full-step regime? The BB formula and the use of safeguards are established ideas. Our contribution is their integration into the post-rejection line search for the present WCP reformulation, together with the associated well-definedness and convergence analysis and an assessment of line-search work. Table 2 distinguishes this role from a BB-based smoothing gradient method and an existing smoothing Newton method.

Table 2. Roles of BB information and Newton steps in representative methods. The entries describe mechanisms, not a numerical ranking.

Method	Problem and direction	Step-selection mechanism
Li [17]	Stochastic linear complementarity; gradient reformulation	BB scaling in a nonmonotone smoothing gradient method.
Tang & Zhang [9]	WCP; smoothing Newton direction	Nonmonotone reference update and geometric backtracking.
Present method	WCP; smoothing Newton direction	Full step first; safeguarded BB initialization only after rejection.

The proposed method uses consecutive iterate and residual differences to obtain a scalar inverse-secant estimate. A safeguarded version of this estimate, combined with the previously accepted step length, initializes the search after full-step rejection. The Newton direction is unchanged. The main contributions are as follows:

- A single-parameter smoothing reformulation is constructed for WCPs in $\mathbb{R}^n$, and the nonsingularity of the Jacobian of the smoothing mapping is proved under an explicit condition combining rank and monotonicity.

- A safeguarded nonmonotone BB Newton (NBB) algorithm is proposed. The full Newton step is always tested first, and the truncated BB factor is used only as a post-rejection trial factor.

- Under the structural rank–monotonicity assumption, we establish well-definedness and prove global convergence when the initial merit-function level set is bounded: The residual tends to zero and every accumulation point is a solution. A unique zero in the level set gives convergence of the whole sequence. Generalized-Jacobian regularity and a locally Lipschitz derivative F' give the local quadratic rate. We verify these conditions for symmetric positive-definite linear models. The BB factor does not affect the eventual full-step phase.

The remainder of this paper is organized as follows: In Section 2, we introduce the notation and regularity concepts used in $\mathbb{R}^n$. In Section 3, we construct the smoothing reformulation, present the safeguarded NBB algorithm, and establish its well-definedness. In Section 4, we develop the global and local convergence analysis. In Section 5, we report the numerical experiments. In Section 6, we give concluding remarks and future directions.

2. Preliminaries

In this section, we introduce the notation and basic concepts used for WCPs in $\mathbb{R}^n$. The componentwise nonnegative orthant is denoted by $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x_i \geq 0,\ i = 1, \ldots, n\}$. For any $x, y \in \mathbb{R}^n$, the Hadamard product is defined by

$$x \circ y := (x_1 y_1, x_2 y_2, \ldots, x_n y_n)^\top.$$

For real vectors of the same dimension, we use the Euclidean inner product

$$\langle u, v\rangle := u^\top v = \sum_i u_i v_i$$

and its induced norm $\|u\| = \sqrt{\langle u, u\rangle}$. We next recall several concepts used in the convergence analysis [15, 21, 22].

Definition 2.1. A mapping $\Phi\colon \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz continuous if, for every $\bar{x} \in \mathbb{R}^n$, there are a neighborhood U of $\bar{x}$ and a constant $L_U > 0$, such that

$$\|\Phi(x) - \Phi(y)\| \le L_U \|x - y\|, \qquad x, y \in U. \tag{2.1}$$

If there exists a constant $L > 0$, such that

$$\|\Phi(x) - \Phi(y)\| \le L\|x - y\|, \qquad x, y \in \mathbb{R}^n,$$

then Φ is globally Lipschitz continuous.

Definition 2.2. Let $F\colon \Omega \to \mathbb{R}^n$, where $\Omega \subseteq \mathbb{R}^n$. The mapping F is monotone on Ω if

$$(u - v)^\top [F(u) - F(v)] \ge 0, \qquad u, v \in \Omega. \tag{2.2}$$

Definition 2.3. Let $\Phi\colon \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitz continuous, and let $\partial\Phi(x)$ denote its Clarke generalized Jacobian at x.

(1) Suppose that Φ is directionally differentiable at x. If, for every sufficiently small increment Δx and every $V \in \partial\Phi(x + \Delta x)$,

$$\Phi(x + \Delta x) - \Phi(x) - V\Delta x = o(\|\Delta x\|),$$

then Φ is semismooth at x.

(2) If Φ is semismooth at x and there exists $p \in (0, +\infty)$ such that, for every sufficiently small Δx and every $V \in \partial\Phi(x + \Delta x)$,

$$\Phi(x + \Delta x) - \Phi(x) - V\Delta x = O\Big(\|\Delta x\|^{1+p}\Big),$$

then Φ is p-order semismooth at x.

In particular, when $p = 1$,

$$\Phi(x + \Delta x) - \Phi(x) - V\Delta x = O(\|\Delta x\|^2),$$

and Φ is said to be strongly semismooth at x. If the corresponding property holds at every point of $\mathbb{R}^n$, then Φ is called semismooth, p-order semismooth, or strongly semismooth on $\mathbb{R}^n$, respectively.

Vectors are columns; we abbreviate a block column $(u_1^\top, \dots, u_q^\top)^\top$ as $(u_1, \dots, u_q)$. Vector inequalities, squares, square roots, and maxima are understood componentwise. We write I for the identity matrix, $\mathrm{diag}(x)$ for the diagonal matrix with diagonal x, e for the all-ones vector, and $x_- := \max\{-x, 0\}$ for the negative part. For a set S, $\mathrm{dist}(u, S) := \inf_{v\in S} \|u - v\|$. The symbols $O(\cdot)$ and $o(\cdot)$ have their usual asymptotic meanings.

3. A nonmonotone NBB algorithm and its well-definedness

In this section, we develop a nonmonotone smoothing Newton method with a safeguarded BB trial factor for solving the WCP (1.2). We first construct a single-parameter smoothing function that reformulates the WCP as a system of smoothing equations. We then present the algorithm and establish its well-definedness. Its convergence properties are analyzed in Section 4.

3.1. A parametric class of complementarity functions

Motivated by the Fischer–Burmeister and natural-residual constructions, we consider the following componentwise function for the WCP (1.2). For $\tau \in [0, 2)$, we define

$$\phi(x, s) = (x + s) - \sqrt{[(x - s)^2 + 2\tau(x \circ s)] + (4 - 2\tau)w}. \tag{3.1}$$

We first verify that the square root is defined componentwise.

Lemma 3.1. *For any $\tau \in [0, 2)$, $w \in \mathbb{R}^n_+$, and $x, s \in \mathbb{R}^n$, the function ϕ in* (3.1) *is well defined.*

Proof. For any $\tau \in [0, 2)$ and $w \in \mathbb{R}^n_+$, we have

$$[(x - s)^2 + 2\tau(x \circ s)] + (4 - 2\tau)w = [(x + (\tau - 1)s)^2 + \tau(2 - \tau)s^2] + (4 - 2\tau)w.$$

Each term on the right-hand side is componentwise nonnegative. Hence, the square root in (3.1) exists, and ϕ is well defined. □

Theorem 3.1. *Let ϕ be defined by* (3.1) *with $\tau \in [0, 2)$ and $w \in \mathbb{R}^n_+$. Then, for every $x, s \in \mathbb{R}^n$,*

$$\phi(x, s) = 0 \quad \Longleftrightarrow \quad x \geq 0,\ s \geq 0,\ x \circ s = w.$$

Proof. If $\phi(x, s) = 0$, then $x_i + s_i \geq 0$ for every i. Squaring the defining equality gives

$$(4 - 2\tau)(x_i s_i - w_i) = 0.$$

Since $\tau < 2$, we have $x_i s_i = w_i \geq 0$. A nonnegative sum and a nonnegative product imply $x_i, s_i \geq 0$. Conversely, if $x_i, s_i \geq 0$ and $x_i s_i = w_i$, the radicand in (3.1) equals $(x_i + s_i)^2$, whose nonnegative square root is $x_i + s_i$. Thus, $\phi_i(x, s) = 0$ for every i. □

The preceding equivalence shows that $\phi(x, s)$ is a complementarity function for the WCP.

3.2. Smoothing reformulation of the WCP

We fix the complementarity-family parameter $\tau \in [0, 2)$. This fixed parameter selects a member of the family, whereas the smoothing parameter μ varies during the algorithm. Because $\phi(x, s)$ is not differentiable everywhere, we introduce $\mu \in \mathbb{R}_+$ and define the single-parameter smoothing function

$$\varphi(\mu, x, s) = (1 + \mu)(x + s) - \sqrt{(1 + \mu)^2[(x - s)^2 + 2\tau(x \circ s)] + (4 - 2\tau)w + 2\mu^2 e}. \tag{3.2}$$

We write $z = (\mu, x, s, y) \in \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$. For the WCP in (1.2), we define

$$H : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \times \mathbb{R}^{m+n} \times \mathbb{R}^n$$

by

$$H(z) = \begin{pmatrix} \mu \\ F(x, s, y) \\ \varphi(\mu, x, s) \end{pmatrix}. \tag{3.3}$$

Because $\varphi(0, x, s) = \phi(x, s)$, Theorem 3.1 implies that $H(z) = 0$ if and only if $\mu = 0$ and (x, s, y) solves the WCP.

For $\mu > 0$, the mapping H is continuously differentiable, with Jacobian

$$H'(z) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & F'_x(x, s, y) & F'_s(x, s, y) & F'_y(x, s, y) \\ \varphi'_\mu(\mu, x, s) & \varphi'_x(\mu, x, s) & \varphi'_s(\mu, x, s) & 0 \end{bmatrix}.$$

For later use, we define

$$c(\mu, x, s) = \sqrt{(1 + \mu)^2[(x - s)^2 + 2\tau(x \circ s)] + (4 - 2\tau)w + 2\mu^2 e}. \tag{3.4}$$

The derivative blocks are

$$\varphi'_\mu(\mu, x, s) = (x + s) - \operatorname{diag}(c)^{-1}\left[(1 + \mu)((x - s)^2 + 2\tau(x \circ s)) + 2\mu e\right], \tag{3.5}$$

$$\varphi'_x = (1 + \mu)I - (1 + \mu)^2 \operatorname{diag}(c)^{-1} \operatorname{diag}(x - (1 - \tau)s), \tag{3.6}$$

$$\varphi'_s = (1 + \mu)I - (1 + \mu)^2 \operatorname{diag}(c)^{-1} \operatorname{diag}(s - (1 - \tau)x). \tag{3.7}$$

Following Tang and Zhang [9], we adopt the following structural assumption for the smoothing Newton analysis.

Assumption 3.1. At every point (x, s, y) in a region containing the (x, s, y)-components of all generated iterates and all their accumulation points, $\operatorname{rank}(F'_y(x, s, y)) = m$. Moreover, for every $(\Delta x, \Delta s, \Delta y) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$ satisfying $F'(x, s, y)(\Delta x, \Delta s, \Delta y) = 0$, we have $\langle \Delta x, \Delta s \rangle \geq 0$.

Remark 3.1. (1) Assumption 3.1 is a standard structural condition in smoothing Newton analyses of WCPs. For linear or equality-constrained models, it can be checked on the null space of the linearized equation; full rank of the constraint block together with monotonicity of the reduced complementarity mapping gives a common sufficient condition.

(2) This assumption is used to prove the nonsingularity of $H'(z)$ for $\mu > 0$ and hence the well-definedness of each Newton direction. It supports the global convergence analysis but does not replace the additional local regularity conditions required for quadratic convergence in Section 4.2. We retain this assumption in the present analysis. Establishing comparable results under weaker, verifiable monotonicity or rank conditions remains a topic for further research.

Example 3.1. For $m = 0$ and $F(x, s) = Bx + q - s$, with B symmetric positive semidefinite, a null direction satisfies $\Delta s = B\Delta x$. Hence, $\langle \Delta x, \Delta s \rangle = \Delta x^\top B \Delta x \geq 0$, and the rank condition is vacuous. This covers the positive-definite finite-difference operators in the obstacle and Reynolds examples below.

For the equality-constrained mapping $F(x, s, y) = (Ax - b, \nabla f(x) - s - A^\top y)$, suppose A has full row rank and $d^\top \nabla^2 f(x) d \geq 0$ for every $d \in \ker A$ throughout the region under consideration. A null direction satisfies $A\Delta x = 0$ and $\Delta s = \nabla^2 f(x)\Delta x - A^\top \Delta y$, so

$$\langle \Delta x, \Delta s \rangle = \Delta x^\top \nabla^2 f(x) \Delta x \geq 0.$$

Additionally, $F'_y = (0, -A^\top)$ has full column rank. Thus, Assumption 3.1 holds. This is a condition on the reduced Hessian, not a requirement that the full Hessian be positive definite.

The nonconvex tests in Section 5.2 are not asserted to satisfy the structural assumption globally.

Theorem 3.2. *If Assumption 3.1 holds, then $H'(z)$ is nonsingular at every point $z = (\mu, x, s, y)$ covered by that assumption with $\mu > 0$.*

Proof. Suppose $H'(z)\Delta z = 0$. The first row gives $\Delta\mu = 0$, so the remaining equations are

$$\begin{cases} F'_x\Delta x + F'_s\Delta s + F'_y\Delta y = 0, \\ \varphi'_x\Delta x + \varphi'_s\Delta s = 0. \end{cases}$$

For each component, $\mu > 0$, $w_i \geq 0$, and $\tau \in [0, 2)$ imply

$$c_i^2 = (1+\mu)^2[(x_i - (1-\tau)s_i)^2 + \tau(2-\tau)s_i^2] + (4-2\tau)w_i + 2\mu^2 > (1+\mu)^2(x_i - (1-\tau)s_i)^2,$$

and the analogous strict inequality holds with x_i and s_i interchanged. Equations (3.6) and (3.7) therefore show that diagonal matrices φ'_x and φ'_s have strictly positive entries. Consequently,

$$\langle \Delta x, \Delta s\rangle = -\sum_{i=1}^{n} \frac{(\varphi'_x)_{ii}}{(\varphi'_s)_{ii}}(\Delta x_i)^2 \leq 0.$$

Assumption 3.1 gives the reverse inequality, forcing $\Delta x = 0$ and hence $\Delta s = 0$. The equation block then reduces to $F'_y\Delta y = 0$. Full column rank gives $\Delta y = 0$, so $H'(z)$ has a trivial null space and is nonsingular. □

3.3. Safeguarded nonmonotone NBB algorithm

Based on the smoothing reformulation developed above, we now give the formal statement of the proposed method. The method combines a smoothing Newton direction with a nonmonotone line search whose post-rejection trial factor uses safeguarded BB information and the previously accepted step length. We refer to this method as the safeguarded nonmonotone NBB algorithm throughout the remainder of the paper. Its complete procedure is stated in Algorithm 3.1.

We first explain the positive forcing term in the Newton equation. Since the first component of $H(z)$ is μ, the ordinary Newton equation $H'(z^k)\Delta z^k = -H(z^k)$ gives $\Delta\mu^k = -\mu_k$. Thus, a full step would set $\mu_{k+1} = 0$, although the remaining nonlinear equations need not yet be solved. At that boundary, H can be nonsmooth and the positive-smoothing nonsingularity result of Theorem 3.2 no longer applies. A damped ordinary step with $\alpha_k < 1$ would keep $\mu_{k+1} > 0$, but merely forbidding full steps does not provide the eventual full-step mechanism used in our local convergence proof.

Instead, we use $h = (1, 0, \ldots, 0)^\top$ and the positive scalar $\beta_k = \gamma \min\{1, C_k\}$, where $C_k \geq M(z^k) > 0$ at a nonterminal iteration. The perturbed equation can be written as

$$H(z^k) + H'(z^k)\Delta z^k = \beta_k h.$$

Thus, only the first component of the right-hand side is perturbed: The linearized smoothing parameter is assigned the positive target β_k, while the linearized F and φ blocks still have zero targets. Its first row gives

$$\Delta\mu^k = -\mu_k + \beta_k, \qquad \mu_{k+1} = (1-\alpha_k)\mu_k + \alpha_k\beta_k > 0 \quad \text{for} \quad 0 < \alpha_k \leq 1.$$

In particular, a full step gives $\mu_{k+1} = \beta_k > 0$. Together with Assumption 3.1, this preserves the region where the next classical Newton system is guaranteed to be well defined. Theorem 4.1 also proves $\beta_k \le \mu_k$, so this positive target does not increase μ_k.

The target is not a fixed positive floor: Lemma 4.2 gives $\beta_k \le \gamma(C_0 + 1)\|H(z^k)\|^2/2$ for $k \ge 1$. Consequently it vanishes with the residual and does not replace the limiting equation $H(z) = 0$ by a different problem. Under the local regularity assumptions, the resulting direction perturbation is second order, as explained after that lemma, and hence is compatible with local quadratic convergence. This is a right-hand-side forcing term, not a regularization of $H'(z^k)$; it cannot make a singular Jacobian invertible. Although only the first right-hand-side component is changed, the coupled system can change all components of Δz^k. We do not assert that an unperturbed or a separately designed nonsmooth Newton method must fail.

For the BB factor, let $d_k = z^k - z^{k-1}$ and $q_k = H(z^k) - H(z^{k-1})$. If $q_k \neq 0$, the scalar inverse-secant least-squares problem has the solution

$$\operatorname*{argmin}_{p\in\mathbb{R}} \|d_k - pq_k\|^2 = \frac{d_k^\top q_k}{q_k^\top q_k} = P_{bb}^k.$$

Indeed, differentiating the squared norm gives $2pq_k^\top q_k - 2d_k^\top q_k$. Residual differences are used because the underlying problem is the equation $H = 0$, and these residuals are already available. Gradient differences for M would instead define a different scaling strategy for the merit function. The scalar pI is only a motivation for the trial factor; it does not replace the Newton matrix. The unconstrained least-squares value need not be positive, which motivates the fallback and safeguard rules below. Variable and residual orderings are fixed throughout, since the scalar factor is not invariant under an independent permutation of the residual entries.

Algorithm 3.1. (NBB) **Initialization.** Choose $z^0 = (\mu_0, x^0, s^0, y^0) \in \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$ with $\mu_0 > 0$. Select $\lambda \in (0, 1)$, $\gamma \in (0, 1)$ with $\gamma \le \mu_0$, and $\sigma \in (0, 1)$ satisfying

$$\sigma(2 - \sqrt{2}\gamma) < 1.$$

Choose safeguard parameters $0 < p_{\min} \le p_{\max}$, and define

$$M(z) = \frac{1}{2}\|H(z)\|^2, \qquad C_0 = M(z^0), \qquad \alpha_{-1} = 1, \qquad k := 0.$$

Step 1. If $\|H(z^k)\| = 0$, stop. Otherwise, set

$$\beta_k = \gamma \min\{1, C_k\}. \tag{3.8}$$

Step 2. Let $h = (1, 0, \ldots, 0)^\top$ be the first unit vector with the same dimension as $H(z)$. Compute the search direction $\Delta z^k = (\Delta\mu^k, \Delta x^k, \Delta s^k, \Delta y^k)$ from

$$H'(z^k)\Delta z^k = -H(z^k) + \beta_k h. \tag{3.9}$$

Step 3. If $k = 0$, set $p_0 = p_{\max}$. For $k \ge 1$, set $p_k = p_{\max}$ when $H(z^k) = H(z^{k-1})$. Otherwise, compute the raw BB factor

$$P_{bb}^k = \frac{(z^k - z^{k-1})^\top (H(z^k) - H(z^{k-1}))}{(H(z^k) - H(z^{k-1}))^\top (H(z^k) - H(z^{k-1}))}. \tag{3.10}$$

If the computed factor is nonpositive or nonfinite, set $p_k = p_{\max}$. Otherwise, set

$$p_k = \min\{p_{\max}, \max\{p_{\min}, P_{bb}^k\}\}. \tag{3.11}$$

Step 4. First test the full Newton step. If

$$M(z^k + \Delta z^k) \le [1 - \sigma(2 - \sqrt{2}\gamma)]C_k, \tag{3.12}$$

set $\alpha_k = 1$. Otherwise define the post-rejection trial factor

$$\rho_k = \max\{\min\{1, p_k\}, \min\{p_{\max}, \alpha_{k-1}\}\}. \tag{3.13}$$

If $\rho_k = 1$, replace ρ_k by λ. Then find $\alpha_k = \rho_k \lambda^i$, where i is the smallest nonnegative integer satisfying

$$M(z^k + \alpha_k \Delta z^k) \le [1 - \sigma(2 - \sqrt{2}\gamma)\alpha_k]C_k. \tag{3.14}$$

Step 5. Set $z^{k+1} = z^k + \alpha_k \Delta z^k$ and update

$$C_{k+1} = \frac{(C_k + 1)M(z^{k+1})}{M(z^{k+1}) + 1}. \tag{3.15}$$

Set $k := k + 1$ and go to Step 1.

Remark 3.2. The full Newton step is tested before any BB trial factor is used. If it is rejected, the safeguarded value $p_k \in [p_{\min}, p_{\max}]$, together with the previously accepted step length, initializes the backtracking sequence. After the possible replacement of $\rho_k = 1$ by λ,

$$\underline{\rho} := \min\{\lambda, p_{\min}\} \le \rho_k < 1.$$

Thus, the BB information affects only globalization and neither changes the Newton direction nor prevents eventual full-step acceptance. The positive lower bound on ρ_k is used in the proof of Theorem 4.2 to identify a rejected predecessor when accepted steps tend to zero. It is not a uniform lower bound on $\alpha_k = \rho_k \lambda^i$. The upper safeguard limits the initial trial scale; it does not constrain the separately tested full Newton step.

Remark 3.3. For the convergence analysis, if Step 1 terminates at an index K, we set $z^k = z^K$ for all $k > K$. In the nonterminating case, Step 5 uses the rational reference update of Tang and Zhang [9]; whenever $M(z^{k+1}) \le C_k$,

$$M(z^{k+1}) \le C_{k+1} \le C_k.$$

This relation provides the decreasing line-search reference used in the convergence analysis.

3.4. Well-definedness

Well-definedness here concerns existence and uniqueness in exact arithmetic. A nearly singular but invertible Jacobian still defines a unique direction, although solving the system in finite precision may yield an inaccurate direction. The BB safeguard does not improve the conditioning of that linear system. If the Jacobian is singular, the result does not apply; replacing the system by a regularized or pseudoinverse solve would require a separate algorithmic analysis.

Theorem 3.3. *Suppose Assumption 3.1 holds. Then NBB is well defined, and every generated iterate satisfies*

$$\mu_k > 0, \qquad M(z^k) \le C_k.$$

Proof. We proceed by induction. The initialization gives $\mu_0 > 0$ and $M(z^0) = C_0$. Suppose $\mu_k > 0$ and $M(z^k) \le C_k$ at a nonterminal iteration. Then $C_k > 0$ and $\beta_k > 0$. Theorem 3.2 ensures that (3.9) has a unique solution. Its first row gives $\Delta\mu^k = -\mu_k + \beta_k$, so every trial with $0 < \alpha \le 1$ satisfies

$$\mu_k + \alpha\Delta\mu^k = (1-\alpha)\mu_k + \alpha\beta_k > 0.$$

To establish finite termination of the line search, differentiate M and use the Newton equation:

$$M'(z^k)\Delta z^k = H(z^k)^\top[-H(z^k) + \beta_k h] = -2M(z^k) + \mu_k\beta_k.$$

Since $\mu_k \le \sqrt{2C_k}$ and $\beta_k = \gamma\min\{1, C_k\} \le \gamma\sqrt{C_k}$, we have $\mu_k\beta_k \le \sqrt{2}\gamma C_k$. For $0 < \alpha \le 1/2$, Taylor expansion therefore gives

$$\begin{aligned} M(z^k + \alpha\Delta z^k) &= (1-2\alpha)M(z^k) + \alpha\mu_k\beta_k + o(\alpha) \\ &\le [1 - (2 - \sqrt{2}\gamma)\alpha]C_k + o(\alpha). \end{aligned}$$

Because $C_k > 0$, $2 - \sqrt{2}\gamma > 0$, and $\sigma < 1$, the remainder is at most $(1-\sigma)(2-\sqrt{2}\gamma)\alpha C_k$ for all sufficiently small positive α. Thus, (3.14) holds for such trials. If the full step is rejected, the geometric sequence $\rho_k\lambda^i$ tends to zero, so backtracking terminates finitely. In either case, $\mu_{k+1} > 0$ and

$$M(z^{k+1}) \le [1 - \sigma(2 - \sqrt{2}\gamma)\alpha_k]C_k \le C_k. \tag{3.16}$$

Writing $M_{k+1} = M(z^{k+1})$, the update in (3.15) satisfies

$$C_{k+1} - M_{k+1} = \frac{M_{k+1}(C_k - M_{k+1})}{M_{k+1} + 1} \ge 0.$$

Hence, $M(z^{k+1}) \le C_{k+1}$, and the induction is complete. □

4. Convergence analysis

In this section, we analyze the convergence properties of NBB. We first prove global residual convergence under boundedness of the initial merit-function level set and state the additional uniqueness condition for convergence of the whole sequence. We then prove local quadratic convergence under additional regularity assumptions.

4.1. Global convergence

Theorem 4.1. *Suppose Assumption 3.1 holds. Let $\{z^k\}$ be the sequence generated by NBB. For every index k for which z^{k+1} is generated,*

$$C_{k+1} \le C_k, \qquad \mu_k \ge \beta_k, \qquad \mu_{k+1} \le \mu_k.$$

Proof. We write $M_{k+1} = M(z^{k+1})$. If the full Newton step is accepted, then $\alpha_k = 1$, and the full-step condition (3.12) gives

$$M_{k+1} \le [1 - \sigma(2 - \sqrt{2}\gamma)]C_k \le C_k.$$

If the full Newton step is rejected, the accepted backtracking step satisfies (3.14), and hence

$$M_{k+1} \le [1 - \sigma(2 - \sqrt{2}\gamma)\alpha_k]C_k \le C_k.$$

Thus, in either case, $M_{k+1} \le C_k$. Therefore,

$$C_k - C_{k+1} = \frac{C_k - M_{k+1}}{M_{k+1} + 1} \ge 0,$$

which proves $C_{k+1} \le C_k$. For $k = 0$, the condition $\gamma \le \mu_0$ implies

$$\mu_0 \ge \gamma \min\{1, C_0\} = \beta_0.$$

If $\mu_k \ge \beta_k$, then

$$\mu_{k+1} = (1 - \alpha_k)\mu_k + \alpha_k\beta_k \ge \beta_k.$$

Since $C_{k+1} \le C_k$, we have $\beta_{k+1} \le \beta_k$, and hence $\mu_{k+1} \ge \beta_{k+1}$. Finally, $\beta_k \le \mu_k$ gives

$$\mu_{k+1} = (1 - \alpha_k)\mu_k + \alpha_k\beta_k \le \mu_k.$$

This proves the theorem. □

Theorem 4.2. *(Global convergence on a bounded initial level set) Suppose Assumption 3.1 holds and the initial level set*

$$\mathcal{L}_0 = \{z = (\mu, x, s, y) \in \mathbb{R}^{1+2n+m} : \ \mu \ge 0, \ M(z) \le C_0\}, \quad C_0 = M(z^0)$$

is bounded. Then the sequence generated by NBB is bounded and has at least one accumulation point. Moreover,

$$\lim_{k\to\infty} \|H(z^k)\| = 0, \qquad \lim_{k\to\infty} \mu_k = 0,$$

and every accumulation point $z^* = (0, x^*, s^*, y^*)$ satisfies $H(z^*) = 0$; hence, (x^*, s^*, y^*) solves the WCP.

Proof. Finite termination already gives a zero of H, so consider an infinite generated sequence. Theorems 3.3 and 4.1 give

$$M(z^k) \le C_k \le C_0 \quad \text{for all} \quad k.$$

Thus, every iterate belongs to $\mathcal{L}_0$. Since H is continuous for $\mu \ge 0$, this level set is closed; by the assumed boundedness, it is compact. Consequently the generated sequence is bounded and has an accumulation point in $\mathcal{L}_0$. This is where the bounded-level-set assumption enters the argument; monotonicity of C_k alone does not imply boundedness of the iterates.

By Theorem 4.1, $\{C_k\}$ is monotonically nonincreasing and bounded below by zero. Hence,

$$C^* = \lim_{k\to\infty} C_k \ge 0$$

exists. If $C^* = 0$, then $M(z^k) \le C_k$ implies

$$\lim_{k\to\infty} M(z^k) = 0, \qquad \lim_{k\to\infty} \|H(z^k)\| = 0.$$

The continuity of H then gives $H(z^*) = 0$ along any convergent subsequence.

We now exclude the case $C^* > 0$. Solving the reference-value update (3.15) for the new merit value gives

$$M(z^{k+1}) = \frac{C_{k+1}}{C_k + 1 - C_{k+1}}.$$

Because $C_k \to C^*$ and $C_{k+1} \to C^*$, the denominator tends to 1. Consequently,

$$\lim_{k\to\infty} M(z^{k+1}) = C^*.$$

Because $C^* > 0$, we have $C_k > 0$ for all sufficiently large k. Dividing the accepted-step estimate (3.16) by C_k gives

$$0 \le \sigma(2 - \sqrt{2}\gamma)\alpha_k \le 1 - \frac{M(z^{k+1})}{C_k} \longrightarrow 0.$$

Thus, $\alpha_k \to 0$. Hence, for all sufficiently large k, the full Newton step is rejected, and the accepted step is generated by the post-rejection backtracking sequence. The starting factor of this sequence satisfies $\rho_k \ge \underline{\rho} > 0$, while $\alpha_k \to 0$. At least one backtracking reduction therefore occurs eventually, so the trial immediately preceding the accepted step is rejected.

For these sufficiently large indices, we let $\widehat{\alpha}_k = \alpha_k/\lambda$ denote the last rejected trial step. We choose a subsequence, still denoted by $\{z^k\}$, such that $z^k \to z^*$. Since the preceding limit also implies $M(z^k) \to C^*$, continuity gives

$$M(z^*) = C^*.$$

Moreover,

$$\beta_k \to \beta^* = \gamma \min\{1, C^*\} > 0, \qquad \mu_k \ge \beta_k.$$

Thus, the subsequential limit satisfies $\mu^* \ge \beta^* > 0$. Therefore, $H'(z^*)$ is nonsingular by Theorem 3.2. Since H' is continuous for positive smoothing parameters, $H'(z^k)$ remains nonsingular and $H'(z^k)^{-1} \to H'(z^*)^{-1}$ along this subsequence. By (3.9), the Newton direction can be written as

$$\Delta z^k = H'(z^k)^{-1}\{-H(z^k) + \beta_k h\}.$$

Consequently, Δz^k converges along the same subsequence to the vector Δz^* satisfying the limiting Newton equation

$$H'(z^*)\Delta z^* = -H(z^*) + \beta^* h,$$

and, in particular, $\{\Delta z^k\}$ is bounded on this subsequence. Since $\widehat{\alpha}_k = \alpha_k/\lambda \to 0$, we also have

$$z^k + \widehat{\alpha}_k \Delta z^k \to z^*.$$

For the last rejected trial step,

$$M(z^k + \widehat{\alpha}_k \Delta z^k) > [1 - \sigma(2 - \sqrt{2}\gamma)\widehat{\alpha}_k] C_k.$$

Since $M(z^k) \le C_k$,

$$\frac{M(z^k + \widehat{\alpha}_k \Delta z^k) - M(z^k)}{\widehat{\alpha}_k} > -\sigma(2 - \sqrt{2}\gamma)C_k.$$

Because $\mu^* > 0$, M is continuously differentiable in a neighborhood of z^*. The limits $z^k \to z^*$, $\Delta z^k \to \Delta z^*$, and $\widehat{\alpha}_k \to 0$ permit passage to the directional-derivative limit in the preceding quotient. We therefore obtain

$$M'(z^*)\Delta z^* \ge -\sigma(2 - \sqrt{2}\gamma)C^*.$$

On the other hand, the limiting Newton equation gives

$$\begin{aligned} M'(z^*)\Delta z^* &= H(z^*)^\top[-H(z^*) + \beta^* h] \\ &= -\|H(z^*)\|^2 + \mu^*\beta^*. \end{aligned}$$

Since $M(z^*) = C^*$, $\mu^* \le \sqrt{2C^*}$, and $\beta^* \le \gamma\sqrt{C^*}$, we obtain

$$M'(z^*)\Delta z^* \le -(2 - \sqrt{2}\gamma)C^*.$$

Thus,

$$-(2 - \sqrt{2}\gamma)C^* \ge M'(z^*)\Delta z^* \ge -\sigma(2 - \sqrt{2}\gamma)C^*,$$

which contradicts $\sigma \in (0, 1)$ and $C^* > 0$. Hence, $C^* = 0$ and $\|H(z^k)\| \to 0$. Since $0 \le \mu_k \le \|H(z^k)\|$, we also have $\mu_k \to 0$. Continuity of H shows that every accumulation point is a zero of H, which is equivalent to a WCP solution by Section 3.2. □

Accumulation-point extension. After obtaining an accumulation point, the residual argument above does not otherwise use boundedness of $\mathcal{L}_0$. Thus, under Assumption 3.1, $\|H(z^k)\| \to 0$ and every accumulation point is a zero of H whenever the generated sequence has at least one accumulation point, even without assuming that the initial level set is bounded. We retain this weaker extension for the local analysis while using the explicit bounded-level-set condition as the main global convergence hypothesis.

Corollary 4.1. *(Whole-sequence convergence under uniqueness) Under the assumptions of Theorem 4.2, if H has a unique zero $\bar{z}$ in $\mathcal{L}_0$, then the whole sequence converges to $\bar{z}$.*

Proof. Theorem 4.2 shows that the sequence is bounded and every accumulation point is a zero of H in $\mathcal{L}_0$. By uniqueness, all such points coincide with $\bar{z}$. If the sequence does not converge to $\bar{z}$, then a subsequence stays outside some fixed neighborhood of $\bar{z}$. Boundedness then gives an accumulation point different from $\bar{z}$, a contradiction. Therefore, $z^k \to \bar{z}$. □

Proposition 4.1. *Let $m = 0$ and $F(x, s) = Bx + q - s$, where B is symmetric positive definite. For fixed $w \ge 0$ and $\tau \in [0, 2)$, every level set $\{z : \mu \ge 0,\ M(z) \le C\}$ is bounded.*

Proof. Suppose otherwise. Since $\mu^2 \le 2C$, an unbounded sequence in the level set has a subsequence for which $\mu \to \bar{\mu} \ge 0$ and $r = \|(x, s)\| \to \infty$, while $(x/r, s/r) \to (u, v)$ with $\|(u, v)\| = 1$. Dividing the bounded equation block by r gives $Bu = v$. Dividing the bounded smoothing block by r, passing to the limit, and cancelling $1 + \bar{\mu} > 0$ gives, componentwise,

$$u_i + v_i = \sqrt{(u_i - v_i)^2 + 2\tau u_i v_i}.$$

Squaring this identity yields $(4 - 2\tau)u_i v_i = 0$. Together with $u_i + v_i \ge 0$, this implies $u_i, v_i \ge 0$ and $u_i v_i = 0$. Consequently, $0 = u^\top v = u^\top Bu$, so $u = 0$ and $v = 0$, contradicting the normalization. □

This proposition supplies a concrete bounded-level-set condition for the fixed-grid linear obstacle and Reynolds models in Section 5.3. It is not asserted for the random nonconvex mappings or for general equality-constrained models with multipliers. In particular, monotonicity alone is not used as a substitute for level-set boundedness.

4.2. Local quadratic convergence

The local nonsingularity assumption below is stronger than the positive-smoothing conclusion of Theorem 3.2. It requires every matrix in $\partial H(z^*)$ to be nonsingular, providing a uniform inverse bound and a local error bound. If each component is nondegenerate (for example, if $w_i > 0$ or, when $w_i = 0$, strict complementarity holds), the norm expression is nonzero at the solution, and this condition reduces to ordinary Jacobian nonsingularity. Smoothness or strict complementarity alone does not guarantee that nonsingularity. At a component with $x_i^* = s_i^* = w_i = 0$, the generalized Jacobian is set-valued, and the all-elements condition is a substantive sufficient regularity requirement. It is used for the local rate, not as an additional assumption of the global residual theorem.

The quadratic smoothing term $2\mu^2 e$ enables the square-root expression in φ to be written as the Euclidean norm of a smooth mapping. Consequently, the same formula defines a locally Lipschitz extension of H to a neighborhood of $\mu = 0$, including negative values of μ. This extension is used when referring to the Clarke generalized Jacobian at $\mu = 0$. The rational reference-value update then makes the forcing term second order in the current residual.

Lemma 4.1. *If F' is locally Lipschitz continuous near (x^*, s^*, y^*), then H is strongly semismooth near $z^* = (0, x^*, s^*, y^*)$.*

Proof. For each component, we use

$$(x_i - s_i)^2 + 2\tau x_i s_i = \big(x_i + (\tau - 1)s_i\big)^2 + \tau(2 - \tau)s_i^2$$

and define

$$G_i(\mu, x_i, s_i) = \begin{pmatrix} (1 + \mu)\big(x_i + (\tau - 1)s_i\big) \\ \sqrt{\tau(2 - \tau)}(1 + \mu)s_i \\ \sqrt{4 - 2\tau}\,\sqrt{w_i} \\ \sqrt{2}\,\mu \end{pmatrix}.$$

The square-root term in φ_i is exactly $\|G_i(\mu, x_i, s_i)\|$. The Euclidean norm is strongly semismooth. Away from the origin it has a locally Lipschitz derivative, whereas at the origin, its semismooth residual vanishes. Since G_i is twice continuously differentiable and strong semismoothness is preserved under smooth composition, each φ_i is strongly semismooth. Moreover, a continuously differentiable mapping with a locally Lipschitz derivative is strongly semismooth. The same conclusion therefore holds for F and hence for H. □

Remark 4.1. The quadratic smoothing term also covers degenerate complementarity components. Indeed, if $w_i = x_i = s_i = 0$, then the locally extended component is

$$\varphi_i(\mu, 0, 0) = -\sqrt{2}\,|\mu|,$$

which is locally Lipschitz and strongly semismooth at $\mu = 0$.

Lemma 4.2. *Let $M_k = M(z^k)$. For every $k \geq 1$,*

$$M_k \leq C_k \leq (C_0 + 1)M_k, \qquad \beta_k \leq \frac{\gamma(C_0 + 1)}{2}\|H(z^k)\|^2.$$

Proof. The lower bound follows from Theorem 3.3. By Theorem 4.1, $C_{k-1} \leq C_0$, and (3.15) gives

$$C_k = \frac{(C_{k-1} + 1)M_k}{M_k + 1} \leq (C_0 + 1)M_k.$$

Since $\min\{1, C_k\} \leq C_k$ and $M_k = \frac{1}{2}\|H(z^k)\|^2$, the estimate for β_k follows. □

To make the role of this estimate explicit, let d_N^k solve the ordinary Newton equation $H'(z^k)d_N^k = -H(z^k)$ at the same positive-smoothing iterate. Subtraction from (3.9) gives

$$\Delta z^k - d_N^k = \beta_k H'(z^k)^{-1}h.$$

Under the local assumptions of Proposition 4.2, the inverse is uniformly bounded near the solution. Since $\|h\| = 1$, Lemma 4.2 and local Lipschitz continuity of H imply

$$\|\Delta z^k - d_N^k\| = O(\|H(z^k)\|^2) = O(\|z^k - z^*\|^2).$$

The positive-smoothing correction therefore has the same order as the local Newton error, rather than a first-order or fixed-size bias. The proposition and theorem below establish the full-step acceptance and quadratic convergence under their stated assumptions.

Proposition 4.2. *Let $z^* = (0, x^*, s^*, y^*)$ be a zero of H. Suppose every $V \in \partial H(z^*)$ is nonsingular and F' is locally Lipschitz continuous near (x^*, s^*, y^*). Let*

$$\widetilde{z}^{k+1} = z^k + \Delta z^k$$

be the full Newton trial point generated by Step 2. There exist a neighborhood U of z^* and constants $c_N, c_H > 0$ such that every generated iterate $z^k \in U$ satisfies

$$\|\widetilde{z}^{k+1} - z^*\| \leq c_N(\|z^k - z^*\|^2 + \beta_k)$$

and

$$\|H(\widetilde{z}^{k+1})\| \leq c_H(\|z^k - z^*\|^2 + \beta_k).$$

Proof. By Lemma 4.1, H is strongly semismooth near z^*. Because $\mu_k > 0$, the classical Jacobian satisfies $H'(z^k) \in \partial H(z^k)$. Upper semicontinuity of the Clarke generalized Jacobian, together with nonsingularity of every element of $\partial H(z^*)$, implies that $\|H'(z^k)^{-1}\|$ is uniformly bounded for all generated z^k sufficiently close to z^*. The Newton system (3.9) gives

$$z^k + \Delta z^k - z^* = z^k - z^* + H'(z^k)^{-1}[-H(z^k) + \beta_k h].$$

Equivalently, using $H(z^*) = 0$,

$$\widetilde{z}^{k+1} - z^* = H'(z^k)^{-1}\{H'(z^k)(z^k - z^*) - [H(z^k) - H(z^*)] + \beta_k h\}.$$

After shrinking the neighborhood if necessary, strong semismoothness gives a uniform constant $c_s > 0$ such that

$$\|H(z^k) - H(z^*) - H'(z^k)(z^k - z^*)\| \le c_s\|z^k - z^*\|^2,$$

for every generated z^k in that neighborhood. These estimates, together with the uniform boundedness of $H'(z^k)^{-1}$ and the fixed vector h, give the first bound with a constant independent of k.

It remains to ensure that the full trial point stays in a common Lipschitz neighborhood of z^*. For $k \ge 1$, Lemma 4.2 and local Lipschitz continuity of H give

$$\beta_k \le \frac{\gamma(C_0 + 1)}{2}\|H(z^k)\|^2 \le c_\beta\|z^k - z^*\|^2$$

for some constant $c_\beta > 0$ whenever z^k is sufficiently close to z^*; for $k = 0$, the same conclusion follows directly from $C_0 = M(z^0)$. The first estimate therefore implies that $\overline{z}^{k+1}$ remains in the chosen local neighborhood after U is reduced if necessary. Since H is locally Lipschitz continuous there and $H(z^*) = 0$, the second estimate follows from

$$\|H(\overline{z}^{k+1})\| \le L\|\overline{z}^{k+1} - z^*\|$$

for some local Lipschitz constant $L > 0$. □

Theorem 4.3. *Suppose Assumption 3.1 holds, and suppose $z^* = (0, x^*, s^*, y^*)$ is an accumulation point of the sequence generated by NBB. If every $V \in \partial H(z^*)$ is nonsingular and F' is locally Lipschitz continuous near (x^*, s^*, y^*), then $z^k \to z^*$ and*

$$\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2).$$

Proof. By the accumulation-point extension following Theorem 4.2, $H(z^*) = 0$; boundedness of the initial level set is not an additional hypothesis of this local theorem. Lemma 4.2 gives, for all $k \ge 1$,

$$\beta_k \le \frac{\gamma(C_0 + 1)}{2}\|H(z^k)\|^2.$$

The local Lipschitz continuity of H therefore implies $\beta_k = O(\|z^k - z^*\|^2)$ whenever z^k is sufficiently close to z^*. Proposition 4.2 therefore yields

$$\|\overline{z}^{k+1} - z^*\| = O(\|z^k - z^*\|^2).$$

Moreover,

$$M(\overline{z}^{k+1}) = \frac{1}{2}\|H(\overline{z}^{k+1})\|^2 = O(\|z^k - z^*\|^4).$$

The nonsingularity of all elements of $\partial H(z^*)$ also gives a local error bound [22]: For all z^k sufficiently close to z^*,

$$c_1\|z^k - z^*\| \le \|H(z^k)\| \le c_2\|z^k - z^*\|$$

with positive constants c_1, c_2. Hence,

$$C_k \ge M(z^k) \ge \frac{1}{2}c_1^2\|z^k - z^*\|^2.$$

Since $M(\overline{z}^{k+1}) = O(\|z^k - z^*\|^4)$ and $1 - \sigma(2 - \sqrt{2}\gamma) > 0$, the full Newton trial satisfies the full-step test (3.12) whenever z^k is sufficiently close to z^*. Because NBB tests the full Newton step before using the safeguarded BB factor, $\alpha_k = 1$ is then accepted. Hence, $z^{k+1} = z^k + \Delta z^k$, and

$$\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2).$$

More explicitly, we choose a closed ball $B_\delta(z^*) = \{z : \|z - z^*\| \le \delta\}$ and a constant $c > 0$ such that the preceding full-step acceptance and the estimate

$$\|z^{k+1} - z^*\| \le c\|z^k - z^*\|^2$$

hold whenever a generated iterate z^k belongs to $B_\delta(z^*)$. By reducing δ if necessary so that $c\delta \le 1/2$, any such iterate satisfies

$$\|z^{k+1} - z^*\| \le c\delta\|z^k - z^*\| \le \tfrac{1}{2}\|z^k - z^*\|.$$

It therefore remains in the same local region. Since z^* is an accumulation point, some sufficiently late iterate enters this ball. Induction shows that every subsequent iterate remains there and converges to z^*, while the preceding quadratic estimate gives the asserted rate. □

Corollary 4.2. *Under the assumptions of Theorem 4.3, there exists an integer K such that $\alpha_k = 1$ for all $k \ge K$. Consequently, the safeguarded BB trial factor and the step-length memory do not affect the local tail of the sequence.*

Proof. The proof of Theorem 4.3 shows that the full Newton step satisfies (3.12) for all sufficiently large k. Since Step 4 tests the full Newton step before using the safeguarded BB factor, NBB accepts $\alpha_k = 1$ from that point onward. □

The following result verifies the local regularity condition for a concrete class, including degenerate complementarity components.

Corollary 4.3. *(Positive-definite linear problems) Let $m = 0$ and $F(x, s) = Bx + q - s$, where B is symmetric positive definite. For every fixed $w \ge 0$ and $\tau \in [0, 2)$, the WCP has a unique solution. From any initialization satisfying Algorithm 3.1, the generated sequence converges to the corresponding zero of H, eventually accepts full Newton steps, and converges quadratically. Strict complementarity is not required.*

Proof. Example 3.1 and Proposition 4.1 establish the structural assumption and boundedness of the initial level set. Theorem 4.2 therefore yields at least one solution. If (x, s) and $(\widehat{x}, \widehat{s})$ are two solutions, then componentwise nonnegativity and $x_i s_i = \widehat{x_i}\widehat{s_i} = w_i$ imply

$$(x_i - \widehat{x_i})(s_i - \widehat{s_i}) = 2w_i - x_i\widehat{s_i} - \widehat{x_i}s_i \le 0,$$

where the last inequality follows from the arithmetic–geometric mean inequality. Since $s - \widehat{s} = B(x - \widehat{x})$, positive definiteness forces $x = \widehat{x}$ and $s = \widehat{s}$. Thus, the solution is unique, and Corollary 4.1 gives convergence of the whole sequence to it.

It remains to verify nonsingularity of every $V \in \partial H(z^*)$. Each such matrix has the block form

$$V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & B & -I \\ a & D_x & D_s \end{bmatrix},$$

where D_x and D_s are diagonal. If $x_i^* + s_i^* > 0$, direct differentiation gives

$$(D_x)_{ii} = \frac{(2-\tau)s_i^*}{x_i^* + s_i^*}, \qquad (D_s)_{ii} = \frac{(2-\tau)x_i^*}{x_i^* + s_i^*}.$$

Both entries are nonnegative and their sum is $2 - \tau > 0$. If $x_i^* = s_i^* = 0$, then $w_i = 0$. In the norm representation used in Lemma 4.1, the derivatives with respect to x_i and s_i at this point are

$$t_x = (1, 0, 0, 0)^\top, \qquad t_s = (\tau - 1, \sqrt{\tau(2-\tau)}, 0, 0)^\top.$$

The norm-subgradient chain rule represents each pair of generalized derivative entries as $1 - v_i^\top t_x$ and $1 - v_i^\top t_s$ for some $\|v_i\| \le 1$. Since $\|t_x\| = \|t_s\| = 1$ and $\|t_x + t_s\| = \sqrt{2\tau}$, these entries satisfy

$$(D_x)_{ii} \ge 0, \qquad (D_s)_{ii} \ge 0, \qquad (D_x)_{ii} + (D_s)_{ii} \ge 2 - \sqrt{2\tau} > 0.$$

These inequalities also hold for convex combinations of limiting Jacobians. If $V\Delta z = 0$, then $\Delta\mu = 0$, $\Delta s = B\Delta x$, and $D_x\Delta x + D_s\Delta s = 0$. The last relation and the nonnegative coefficients with positive sum imply $\Delta x_i \Delta s_i \le 0$ for every i. Hence,

$$0 \ge \Delta x^\top \Delta s = \Delta x^\top B \Delta x,$$

which gives $\Delta x = \Delta s = 0$. Thus, V is nonsingular. Since F' is constant, Theorem 4.3 and Corollary 4.2 prove the rate and full-step conclusions. □

This corollary applies to the fixed-grid linear obstacle, Reynolds, and positive-weight models below. It does not cover the nonlinear Bratu or random nonconvex tests. The general local theorem still requires generalized-Jacobian nonsingularity; the corollary supplies a verifiable sufficient model class rather than removing that requirement for arbitrary WCPs.

5. Numerical experiments

In this section, we report numerical experiments for NBB. We first compare one-step methods on fixed nonconvex instances, isolate the BB mechanism by ablation, and examine parameter sensitivity. In Sections 5.2.5 and 5.2.6, we add a recent independent one-step comparator, monotone linear tests, and tests at nonzero values of τ. In Section 5.3, we study three zero-weight physical models, including failure diagnostics, nested meshes, and an independent discrete contact-set check, followed by a separate analytical-solution benchmark. In Section 5.4, we test larger sparse problems, including positive weights.

5.1. Experimental settings

NBB solves one Newton system per outer iteration, while its safeguarded BB update requires only vector inner products and scalar truncation. We denote the real-coordinate nonmonotone smoothing Newton algorithm by Rn-NSNA and the fixed-trial smoothing Newton method by FT-SNM. Rn-NSNA provides a literature-based one-step comparator, and the controlled FT-SNM ablation isolates the BB trial factor within the same smoothing equation and globalization framework. We denote our real-coordinate implementation of the smoothing Newton method under a local error bound by Rn-SNEB.

In Section 5.2.5, we compare this independent one-step method, retaining its own smoothing and monotone acceptance rules. Linear-system and mapping-evaluation counts are reported separately from outer iterations.

All experiments are implemented in MATLAB (matrix laboratory), version R2023b. They run on a computer with an advanced micro devices (AMD) Ryzen 5 6600H processor at 3.30 gigahertz (GHz). The machine has 16 gigabytes (GB) of random-access memory (RAM). Unless otherwise stated, the tolerance for the method-specific smoothing residual is 10^{-8}, and the maximum number of iterations is 220. The line search permits at most 92 trial evaluations per outer iteration, including the initial full-step trial. In Rn-SNEB, the full trial is reused if its second acceptance test is needed at the same point. Fresh runs of NBB and Rn-NSNA on the original 150 instances reproduce their previous success flags and work counts. A run is counted as successful only if its method-specific smoothing residual satisfies the stopping tolerance and its common Karush–Kuhn–Tucker (KKT) residual

$$R_{\mathrm{KKT}} = \max\{\|Ax - b\|_\infty, \|\nabla f(x) - s - A^\top y\|_\infty, \|x \circ s - w\|_\infty, \|x_-\|_\infty, \|s_-\|_\infty\}$$

is at most 10^{-7}. This residual is evaluated on the original complementarity variables rather than on a method-specific smoothed equation. It therefore provides a common accuracy measure for all methods and variants. Wall-clock time is included only as a supplementary indicator because linear-algebra timings depend on the local MATLAB environment. The historical raw-data fields named `cpu` or `cpu_time` are recorded with `tic/toc`; they measure elapsed solver time, excluding model construction, not process central processing unit (CPU) time. In Section 5.4, we report repeated complete-call timings, construction times, and separately measured process CPU times. Less hardware-dependent work measures are the numbers of Newton systems, residual evaluations, and rejected line-search trial steps.

The original Bratu and Reynolds experiments use a 200-iteration budget and at most 82 line-search trials per iteration. The later nested-grid and analytical-solution studies use the default 220-iteration and 92-trial budgets; extended failure diagnostics are identified separately.

All random instances in the one-step comparison use fixed MATLAB random streams. In the tests with different dimensions, the problem seed is

$$920000 + 1000n + j$$

and the starting-vector seed is

$$925000 + 1000n + j,$$

where j is the instance index. In the tests with different initialization scales, the corresponding seeds are

$$960000 + 1000r + j \quad \text{and} \quad 965000 + 1000r + j.$$

MATLAB's backslash operator solves the reduced Newton systems. The raw instance-level comma-separated values (CSV) files record the seed, success flag, iteration count, residual evaluations, Newton linear solves, rejected trial steps, number of iterations with an admissible positive BB factor, and final residual. Unless otherwise stated, the main tables average successful runs only. Failure counts and paired win rates are reported in the surrounding text. Together with the tolerances, metrics, and averaging convention, the seeds and records specify the protocol used to generate the tables.

Implementation checks are included in the supplementary code. Across 48 combinations of parameter value, sign pattern, and equality-constraint dimension, the largest relative central-difference Jacobian discrepancy is 1.46×10^{-10}, and the largest relative reduced/full Newton-direction discrepancy is 1.41×10^{-15}. Separate tests check known roots, domain boundaries, the scalar unconstrained case, and exact trial-count accounting when the search budget is exhausted. These finite tests check implementation consistency; they do not replace the convergence assumptions or constitute a proof of correctness for every input.

The literature baseline, denoted by Rn-NSNA, is a coordinatewise $\mathbb{R}^n$ specialization of the one-step nonmonotone smoothing Newton framework of Tang and Zhang [9]. It solves one Newton system per iteration and uses the published rational reference-value update with geometric backtracking. We also implement a fixed-trial control, FT-SNM, which uses the same smoothing equation, reduced Newton system, forcing term, and stopping rules as NBB but disables the BB factor and accepted-step memory. Thus, FT-SNM starts the post-rejection sequence at λ, after the full step has been tested. Across all 150 paired instances reported below, FT-SNM and Rn-NSNA have identical success flags, iteration counts, residual-evaluation counts, and rejected-trial counts. This empirical identity does not make the two formulations mathematically identical. To avoid duplicate rows, the main comparison tables report Rn-NSNA. The FT-SNM records are retained in the raw data, and the fixed-trial mechanism is examined directly in the ablation study.

NBB, Rn-NSNA, and FT-SNM use $\mu_0 = 10^{-4}$, $\gamma = 10^{-5}$, $\sigma = 0.2$, and $\tau = 0$, together with the common tolerances and reduced linear solver stated above. The geometric contraction value is 0.5: It is denoted by λ in NBB and FT-SNM and by δ in Rn-NSNA. NBB additionally uses $p_{\min} = 0.25$ and $p_{\max} = 0.5$; these safeguards do not enter Rn-NSNA or FT-SNM. The choice $\tau = 0$ is the natural-residual member of the family in Section 3.2. The main paired comparisons retain this value to isolate the line-search mechanism; additional nonzero values are tested in Section 5.2.6. No uniform robustness claim is made over $\tau \in [0, 2)$.

The number of outer Newton iterations (IT), number of function evaluations (NF), and number of rejected trial steps (RED) are individual-run counts. Their successful-run means are AIT, ANF, and ARED. Paired and dispersion tables summarize these counts as specified in their captions.

The random benchmark construction follows numerical-testing practice in smoothing Newton studies for WCPs and related complementarity problems. Dense random matrices, nonlinear componentwise perturbations, and infeasible or noninterior starting points test globalization behavior [9, 10, 12]. The dimensions and initialization scales are chosen to activate the line search and expose the effect of the safeguarded BB trial factor.

The supplementary reproducibility archive contains the MATLAB implementations, seeded instance generators, original and added instance-level CSV files, paired statistical summaries, saved field/trajectory data, and scripts for the added tables and figures. Its README gives the exact entry points and settings. It also includes derivative and reduced/full-system verification records. The archive permits reconstruction of the comparisons from unrounded instance-level records.

5.2. Test problems and comparative experiments

The two benchmark groups below are nonconvex stress tests. We do not assume that they satisfy Assumption 3.1 along every generated trajectory, and empirical success on them is not a verification

of the global theorem. They use the equality-constrained WCP mapping

$$F(x,s,y)=\begin{pmatrix} Ax-b \\ \nabla f(x)-s-A^{\top}y \end{pmatrix}, \qquad x\geq 0, \quad s\geq 0, \quad x\circ s=w.$$

This is also the fixed row order in the implementation:

$$z=(\mu,x,s,y) \quad \text{and} \quad H(z)=(\mu,Ax-b,\nabla f(x)-s-A^{\top}y,\varphi).$$

An independent permutation of the residual rows would preserve the Newton step but could change the BB inner product, so the same order is used in all comparisons. The nonlinear objective defining the stationarity block is

$$\begin{aligned} f(x) &= \frac{1}{2}x^{\top}Qx+(c+P(x))^{\top}x, \\ Q &= \frac{n}{4}\frac{N^{\top}N}{\|N^{\top}N\|_2}, \\ P_i(x) &= x_i^2+\sin x_i+\cos x_i+1. \end{aligned}$$

The gradient and Hessian used in the Newton equations are

$$\begin{aligned} \nabla f(x) &= Qx+c+P(x)+x\circ P'(x), \\ \nabla^2 f(x) &= Q+\operatorname{diag}(2P'(x)+x\circ P''(x)), \end{aligned}$$

where

$$P_i'(x)=2x_i+\cos x_i-\sin x_i \quad \text{and} \quad P_i''(x)=2-\sin x_i-\cos x_i.$$

Here,

$$N=\operatorname{randn}(n,n), \quad c=\operatorname{rand}(n,1), \quad w=\operatorname{rand}(n,1), \quad A=\operatorname{randn}(m,n), \quad \text{and} \quad b=A\hat{x}$$

for $\hat{x}=\operatorname{rand}(n,1)$. Thus, Q is positive semidefinite after scaling, whereas the diagonal nonlinear term can make $\nabla^2 f(x)$ indefinite. The choice $b=A\hat{x}$ ensures feasibility of the equality constraints. The initial points are mixed-sign noninterior points

$$x^0=r\ \operatorname{randn}(n,1), \qquad s^0=r\ \operatorname{randn}(n,1), \qquad y^0=r\ \operatorname{randn}(m,1),$$

which are deliberately outside the usual positive-start local regime. The Gaussian vectors are not normalized, so r is an initialization scale, not a prescribed Euclidean radius or distance to a solution. Scaling a fixed signed vector by $r>0$ also scales its distance to the nonnegative orthant. These tests therefore emphasize globalization rather than only local Newton behavior near a solution.

5.2.1. Tests with different problem dimensions

The first group fixes $r=20$ and tests $(n,m)=(100,50),(150,75)$, and $(200,100)$. For each dimension, 20 independent instances are generated. This experiment examines whether the BB globalization remains useful as the number of complementarity variables and the number of linear equality constraints increase. Table 3 gives the results.

Table 3. Comparison of NBB and Rn-NSNA for nonconvex nonlinear WCPs of different dimensions at $r = 20$ (20 instances per dimension; averages over successful runs). NBB denotes the proposed nonmonotone NBB algorithm. ABB is the average number of BB-initialized searches after full-step rejection, not accepted BB trials.

n	Method	SR (%)	AIT	ANF	ARED	ABB	AKKT
100	NBB	100	23.90	76.50	27.70	5.55	8.45×10^{-10}
	Rn-NSNA	100	26.50	91.20	37.20	0	1.18×10^{-9}
150	NBB	100	34.90	121.95	51.15	11.25	1.59×10^{-9}
	Rn-NSNA	100	42.05	161.25	76.15	0	6.40×10^{-10}
200	NBB	90	78.78	320.50	161.94	33.44	1.42×10^{-9}
	Rn-NSNA	85	89.00	394.35	215.35	0	5.38×10^{-10}

Both methods solve all instances for $n = 100$ and $n = 150$. For $n = 200$, NBB fails on two instances and Rn-NSNA fails on three. Of the 20 paired instances, both methods solve 15, NBB alone solves three, and Rn-NSNA alone solves two. A two-sided exact McNemar test on these five discordant outcomes gives $p = 1.000$, so the one-instance difference is descriptive rather than evidence of a systematic success-rate improvement. Relative to Rn-NSNA, the successful-run means for NBB are lower by 9.8%, 17.0%, and 11.5% in iteration count; 16.1%, 24.4%, and 18.7% in residual evaluations; and 25.5%, 32.8%, and 24.8% in rejected trial steps for $n = 100, 150, 200$, respectively. Because one Newton system is solved per iteration, ALS equals AIT and is omitted from the table. Elapsed-time measurements are retained in the raw files but are not used as the primary comparison metric.

On jointly successful instances, NBB uses fewer residual evaluations on 80.0%, 75.0%, and 66.7% of the runs and fewer rejected trial steps on 80.0%, 80.0%, and 66.7% for $n = 100, 150, 200$, respectively. The final common residuals have the same order of magnitude. All instance-level counts are retained in the raw files, from which standard deviations can be computed. The increase in ABB with dimension indicates that an admissible positive BB factor is available on more post-rejection iterations as the line search becomes harder. These observations motivate the paired statistical assessment in Section 5.2.2.

5.2.2. Tests with different initialization scales

The second group fixes $(n, m) = (200, 100)$ and varies the initialization scale r. Thirty independent instances are generated for each of $r = 20, 50, 100$. The seeds depend on r, so the three groups have the same dimensions but different problem instances and starting-vector directions; only the methods within each group are paired. Consequently, cross-group differences cannot be attributed solely to the change in r. All three groups are reported, including $r = 20$, where the methods have different failure sets. Table 4 gives the results; the column abbreviations have the same meanings as in Table 3.

Table 4. Comparison of NBB and Rn-NSNA for nonconvex nonlinear WCPs at different initialization scales, with $(n, m) = (200, 100)$ (30 instances per scale; averages over successful runs). ABB counts positive BB availability after full-step rejection, not accepted BB trials.

r	Method	SR (%)	AIT	ANF	ARED	ABB	AKKT
20	NBB	90.00	68.67	272.22	133.89	27.85	9.72×10^{-10}
	Rn-NSNA	93.33	94.96	416.04	225.11	0	1.10×10^{-9}
50	NBB	100	44.90	164.63	73.83	15.07	3.40×10^{-10}
	Rn-NSNA	100	60.03	242.10	121.03	0	5.93×10^{-10}
100	NBB	100	38.37	135.77	58.03	11.83	7.46×10^{-10}
	Rn-NSNA	100	41.37	155.13	71.40	0	1.01×10^{-9}

At $r = 20$, NBB solves 27 instances and Rn-NSNA solves 28. Both methods solve 25 instances, NBB alone solves two, and Rn-NSNA alone solves three. The two-sided exact McNemar test again gives $p = 1.000$. Among successful runs, NBB has a 27.7% lower mean iteration count, 34.6% fewer residual evaluations, and 40.5% fewer rejected trial steps. At $r = 50$, the corresponding reductions are 25.2%, 32.0%, and 39.0%; at $r = 100$, they are 7.3%, 12.5%, and 18.7%. NBB uses fewer residual evaluations on 76.0%, 73.3%, and 73.3% of jointly successful runs for $r = 20, 50, 100$, and fewer rejected trial steps on 76.0%, 76.7%, and 76.7%, respectively. Thus, the most consistent reduction occurs in rejected trial steps, the metric most directly affected by the post-rejection BB factor.

Because successful-run means can be influenced by a small number of long runs, Table 5 supplements Tables 3 and 4 with medians and interquartile ranges. Each entry has the form median $[Q_1, Q_3]$ and is computed using the same successful-run subsets as the main tables.

Table 5. Distributional summary of work counts. Entries are median $[Q_1, Q_3]$ over successful runs.

Setting	Method	IT	NF	RED
$n = 100$	NBB	24 [20, 26.5]	77 [58, 92.5]	26 [17, 38.5]
	Rn-NSNA	24.5 [19, 29.5]	79 [56.25, 110]	30 [17.5, 49]
$n = 150$	NBB	31.5 [24.75, 42]	108.5 [78.75, 147.75]	43.5 [28.75, 64.25]
	Rn-NSNA	40 [27.5, 46.75]	149.5 [94.75, 186.75]	68 [42.25, 95.25]
$n = 200$	NBB	56.5 [42.5, 99.5]	244 [157.75, 397.75]	130 [72.25, 198.75]
	Rn-NSNA	72 [44, 106]	311 [172, 427]	166 [86, 233]
$r = 20$	NBB	55 [33, 93.5]	214 [124, 378.5]	107 [57, 187.5]
	Rn-NSNA	76.5 [53, 129.25]	325 [216.75, 592]	172.5 [109.75, 332.5]
$r = 50$	NBB	39 [31.75, 50.5]	145 [114.25, 185.75]	68 [47.75, 81.75]
	Rn-NSNA	44 [36.5, 73]	173 [132.5, 294]	86 [57.75, 151]
$r = 100$	NBB	34.5 [27.25, 44.5]	120.5 [90, 170.5]	50.5 [33, 76.25]
	Rn-NSNA	34.5 [27, 46.75]	120.5 [90.5, 184]	50.5 [36.25, 94.5]

The marginal median rejected-trial count is lower for NBB in five settings and equal at $r = 100$. Median residual evaluations are also lower in five settings and equal at $r = 100$, while median iteration counts are lower in five settings and equal at $r = 100$. The interquartile ranges widen for the hardest $n = 200$ and $r = 20$ tests, explaining why means alone are not a sufficient summary. These descriptive

distributional results are consistent with, but do not replace, the paired inference below.

Paired statistical assessment. The preceding summaries describe successful-run distributions but do not account for pairing. We therefore apply a two-sided exact sign test to each collection of jointly successful paired instances. This nonparametric test uses the direction of each nonzero paired difference and does not require a normality assumption [23]. For notational brevity in this paragraph, NSNA denotes the Rn-NSNA comparator. For any reported work metric, define

$$\Delta(\cdot) = \mathrm{NBB}(\cdot) - \mathrm{NSNA}(\cdot), \tag{5.1}$$

where the two metric values are obtained on the same problem instance. Thus, $\Delta(\cdot) < 0$ favors NBB, $\Delta(\cdot) > 0$ favors NSNA, and $\Delta(\cdot) = 0$ is a tie. The entries denoted by Δ in Table 6 are the medians of these instance-wise paired differences.

For a fixed setting and metric, let d_i denote the paired difference (5.1) on instance i, omit ties, and define $n_- = \#\{i : d_i < 0\}$, $n_+ = \#\{i : d_i > 0\}$, $N_s = n_- + n_+$. Under the null hypothesis that neither algorithm is systematically favored, a negative or positive sign has probability 0.5, so the number of signs in either specified direction follows $X \sim \mathrm{Binomial}(N_s, 0.5)$. The two-sided exact p-value is therefore

$$p = \min\left\{1,\ 2\sum_{j=0}^{\min\{n_-,n_+\}} \binom{N_s}{j}\left(\frac{1}{2}\right)^{N_s}\right\}. \tag{5.2}$$

This p-value is the probability, under the null hypothesis, of observing a sign imbalance at least as extreme as the one obtained. A smaller p-value indicates stronger incompatibility with the null hypothesis, whereas a large p-value indicates insufficient evidence against it and does not prove that the two algorithms are equivalent [24]. We use 0.05 as the nominal significance level. Because six setting-wise tests are conducted within each work-metric family, Holm correction is applied to control the family-wise error rate [25]. Table 6 reports the unadjusted exact p-values, and the adjusted conclusions are stated below the table.

Table 6. Paired exact sign tests on jointly successful instances. Entries are median differences $\Delta(\cdot)$ with unadjusted p-values in parentheses.

Setting	Pairs	Δ(IT) (p)	Δ(NF) (p)	Δ(RED) (p)
$n = 100$	20	0 (0.754)	−4.5 (0.0118)	−3.5 (0.00443)
$n = 150$	20	0 (0.424)	−7.5 (0.0414)	−7.5 (0.0118)
$n = 200$	15	−2 (0.581)	−23 (0.302)	−19 (0.302)
$r = 20$	25	−10 (0.115)	−51 (0.0146)	−31 (0.0146)
$r = 50$	30	−10 (0.108)	−51.5 (0.0161)	−31.5 (0.00232)
$r = 100$	30	0 (0.503)	−8 (0.0161)	−6 (0.00522)

For example, consider RED in the $n = 100$ row of Table 6. Here, IT is the number of outer Newton iterations, NF is the number of function evaluations, and RED is the number of rejected trial steps. Their successful-run averages are AIT, ANF, and ARED. Among the 20 jointly successful pairs, NBB has fewer rejected trial steps on 16 instances, NSNA has fewer on three instances, and one pair is tied.

After the tie is omitted, $n_- = 16$, $n_+ = 3$, and $N_s = 19$. Substitution into (5.2) gives

$$p = 2\sum_{j=0}^{3}\binom{19}{j}2^{-19} = 0.004425 \approx 0.00443.$$

The accompanying value $\Delta(\mathrm{RED}) = -3.5$ means that the median of the 20 paired differences is -3.5; in other words, NBB requires 3.5 fewer rejected trial steps in the median paired comparison. The small unadjusted p-value shows that the observed 16-to-3 sign imbalance would be unlikely under equal sign probabilities; its Holm-adjusted value is $0.0221 < 0.05$, so the difference remains statistically significant after correction. By contrast, the same row has $\Delta(\mathrm{IT}) = 0$ and $p = 0.754$, which provides no evidence of a systematic difference in iteration count.

The iteration-count differences are not significant in any setting. Residual-evaluation differences have unadjusted $p < 0.05$ in five of six settings, but none remains significant after the within-metric Holm correction. For rejected trial steps, five of the six comparisons remain significant after correction at the 0.05 level, with adjusted p-values between 0.0139 and 0.0355; the $n = 200$ comparison is not significant (adjusted $p = 0.302$). Success-rate tests give $p = 1.000$, and wall-clock-time differences are not consistent across settings. The evidence therefore supports a narrow conclusion. Within these tests, the safeguarded BB mechanism reduces line-search rejections in most settings and often lowers residual-evaluation work. It does not establish a general improvement in success rate, Newton iteration count, or wall-clock time.

5.2.3. Ablation study of the BB globalization mechanism

To separate the effects of the safeguarded BB trial factor, step-length memory, and the safeguard interval, we perform an ablation study on the same nonconvex WCP family with $(n, m) = (80, 40)$ and $r = 30$. Thirty independent instances are generated with problem seeds $970000 + j$ and starting-vector seeds $975000 + j$, where $j = 1, \ldots, 30$. This experiment addresses a narrower question than the Rn-NSNA comparison. It examines which component of the trial-factor update reduces line-search cost after full-step rejection. The four variants are

NBB	the proposed safeguarded BB method,
NBB-M0	NBB without the accepted-step memory,
FIXED	a fixed post-rejection trial factor with no BB factor,
NBB-U	an unclipped positive BB factor with the same line search.

The full Newton step is tested first in all variants; the differences concern only the initial trial factor used after this full step is rejected. Table 7 reports averages over successful runs.

Table 7. Ablation of the BB globalization mechanism on 30 nonconvex WCP instances with $(n, m) = (80, 40)$ and $r = 30$. ANF is the average number of function evaluations; ARED is the average number of rejected trial steps.

Variant	SR (%)	AIT	ANF	ARED	ABB	AKKT
NBB	100	20.73	61.43	18.97	3.73	4.10×10^{-10}
NBB-M0	100	20.70	61.37	18.97	3.47	3.60×10^{-10}
FIXED	100	20.83	65.73	23.07	0	4.10×10^{-10}
NBB-U	100	22.67	69.00	22.67	5.17	7.21×10^{-11}

All variants solve all 30 instances, so the ablation compares computational work without confounding from failed runs. Because the four variants use the same Newton system and stopping test, ANF and ARED reflect changes induced by the line-search trial factor rather than changes in the local model. Compared with FIXED, NBB has nearly the same average iteration count but 6.5% fewer residual evaluations and 17.8% fewer rejected trial steps. In the paired exact sign tests, NBB has lower NF and RED values than FIXED on 23 instances and higher values on six, with one tie. After Holm correction across the three NBB-versus-variant contrasts within each metric, the NF and RED comparisons with FIXED remain significant (adjusted $p = 0.00695$). NBB-M0 is nearly indistinguishable from NBB, and neither its NF nor RED contrast is significant, indicating that accepted-step memory is not the dominant source of improvement on this set. NBB-U has larger successful-run means than NBB, but the paired differences between these two variants are not significant. The ablation therefore supports a specific, limited conclusion: The safeguarded BB initializer reduces line-search work relative to the fixed initializer on these instances, while the separate effects of memory and clipping are less decisive.

5.2.4. Parameter sensitivity and convergence profiles

We next test whether the observed behavior depends strongly on a single parameter choice. The sensitivity study uses

$$(n, m) = (200, 100), \quad r = 50,$$

and 12 independent instances with problem seeds $980000 + j$ and starting-vector seeds $985000 + j$, where $j = 1, \ldots, 12$. The base setting is

$$(p_{\min}, p_{\max}) = (0.25, 0.50), \quad \lambda = 0.5, \quad \text{and} \quad \gamma = 10^{-5}.$$

The interval $(p_{\min}, p_{\max})$ controls truncation of the BB factor. Parameter λ is the backtracking contraction factor, and γ controls the forcing term in the smoothing-parameter update. Table 8 perturbs these three components separately.

Table 8. Parameter sensitivity of NBB on 12 nonconvex WCP instances with $(n, m) = (200, 100)$ and $r = 50$.

Setting	$(p_{\min}, p_{\max})$	λ	γ	SR (%)	AIT	ANF	ARED	AKKT
baseline	$(0.25, 0.50)$	0.5	10^{-5}	100	46.17	175.25	81.92	5.40×10^{-10}
interval A	$(0.10, 0.40)$	0.5	10^{-5}	100	45.25	159.50	68.00	1.40×10^{-10}
interval B	$(0.25, 0.80)$	0.5	10^{-5}	100	41.42	157.08	73.25	3.89×10^{-10}
interval C	$(0.40, 0.80)$	0.5	10^{-5}	100	48.17	202.50	105.17	6.62×10^{-10}
$\lambda = 0.4$	$(0.25, 0.50)$	0.4	10^{-5}	100	40.08	141.67	60.50	6.36×10^{-10}
$\lambda = 0.7$	$(0.25, 0.50)$	0.7	10^{-5}	100	45.08	201.58	110.42	4.43×10^{-10}
$\gamma = 10^{-6}$	$(0.25, 0.50)$	0.5	10^{-6}	100	41.08	154.33	71.17	1.32×10^{-9}
$\gamma = 10^{-4}$	$(0.25, 0.50)$	0.5	10^{-4}	100	44.50	163.33	73.33	5.96×10^{-14}

All tested settings solve all 12 instances. The average iteration count ranges from 40.08 to 48.17, and the mean final common residual is below 1.4×10^{-9} for every setting. Across the 96 individual

runs, the largest final common residual is 8.56×10^{-9}, still below the common 10^{-7} threshold. The larger trial factors associated with $(0.40, 0.80)$ and $\lambda = 0.7$ are accompanied by more backtracking. No failure occurs under any tested single-component perturbation, although the work counts remain sensitive to the initial trial scale.

Figure 1 shows an illustrative trajectory for a reproducibly generated instance with $(n, m) = (200, 100)$ and $r = 50$. The problem and starting-vector seeds are 990001 and 995001, respectively. The upper panel plots the smoothing residual, while the lower panel records the accepted step length and safeguarded BB trial factor. NBB reaches

$$\|H(z^k)\|_2 = 1.04 \times 10^{-13}$$

in 52 iterations, whereas the fixed-trial variant reaches 2.79×10^{-9} in 65 iterations. NBB uses small safeguarded trial factors away from the solution and eventually accepts full Newton steps. This single trajectory illustrates the mechanism and is not used in the paired inference. Its full-step tail agrees with Corollary 4.2: BB globalization acts in the nonlocal phase, while full Newton steps govern the local tail.

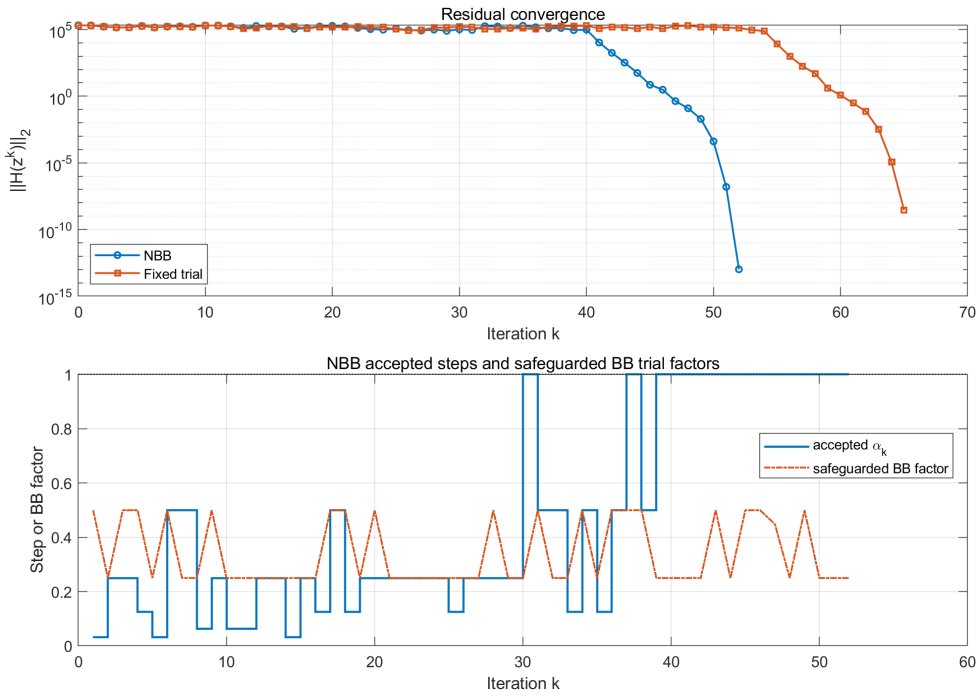


Figure 1. Convergence profile for the illustrative nonconvex WCP instance. The upper panel compares residual decay for NBB and the fixed-trial variant. The lower panel shows the accepted step length and safeguarded BB trial factor used by NBB.

5.2.5. Comparison with a recent one-step smoothing Newton method

We implement Algorithm 3.1 of Tang and Zhou [11], whose article appeared online in 2023 and in a 2025 journal issue. We call our Euclidean-coordinate implementation Rn-smoothing Newton under an error bound (SNEB); this is not the authors' original software. It computes one Newton direction

per iteration, without a second correction. Its local quadratic convergence theorem requires the rank-monotonicity assumption, a local error bound, local Lipschitz continuity, a quadratic linearization remainder, and a uniformly bounded inverse on a neighborhood with positive smoothing parameters, together with an accumulation point. An error bound alone is not the complete hypothesis.

With

$$H_{\mathrm{E}}(z) = (\mu, F(x, s, y), \psi_{\mathrm{E}}(\mu, x, s)),$$

the method uses

$$\psi_{\mathrm{E}}(\mu, x, s) = x + s - \sqrt{x^2 + s^2 + 2w + 4\mu^2 e}, \qquad \Psi_{\mathrm{E}}(z) = \|H_{\mathrm{E}}(z)\|^2,$$

where squares and square roots are componentwise. The direction satisfies

$$H'_{\mathrm{E}}(z)\Delta z = -H_{\mathrm{E}}(z) + \gamma_{\mathrm{E}}\Psi_{\mathrm{E}}(z)e_1.$$

The full step is accepted if its residual norm is at most τ_{E} times the current norm. Otherwise, it tests $\alpha = \delta^l$, $l = 0, 1, \ldots$, in

$$\Psi_{\mathrm{E}}(z + \alpha\Delta z) \le (1 - \lambda_1\alpha^2)\Psi_{\mathrm{E}}(z) - \lambda_2\|\alpha\Delta z\|^2.$$

We retain the published numerical choices

$$\mu_0 = 10^{-3}, \;\; \delta = \tau_{\mathrm{E}} = 0.5, \;\; \lambda_1 = \lambda_2 = 10^{-5},$$

and the fixed

$$\gamma_{\mathrm{E}} = \mu_0/(\Psi_{\mathrm{E}}(z^0) + 1).$$

Here, τ_{E} is a full-step acceptance factor, not our complementarity-family parameter τ. The already evaluated full trial is reused in the fallback test; it is not counted twice. The residual tolerance is tightened from the source's 10^{-6} to 10^{-8} for our comparison.

The same 150 nonconvex problems and initial x, s, y vectors are used for all three methods. Their distinct smoothing parameters and forcing rules are retained, including $\mu_0 = 10^{-4}$ for NBB and Rn-NSNA. All use the common KKT threshold, 220-iteration budget, and 92-trial cap. The 300 NBB/Rn-NSNA reruns reproduce their previous success flags and work counts. For Rn-SNEB, four small Jacobian checks have a maximum relative finite-difference error of 8.52×10^{-11}; full and reduced directions agree to 7.28×10^{-16}. Twenty independently computed dense first steps also verify the fallback test, including acceptance at $\alpha = 1$ after the initial norm test fails. Final native and common residuals are independently recomputed from saved fields.

Table 9a reports the nonconvex stress tests. Rn-SNEB succeeds on none of these 150 instances: 135 reach the iteration budget and 15 exhaust the trial budget. These tests lie outside its monotonicity guarantee as well as ours. There is consequently no common-success subset for a three-method work comparison; dashes must not be interpreted as zero work or infinite running time. The supplement retains every terminal residual and failure status. No claim of general inferiority of Rn-SNEB follows from this experiment.

Table 9. Independent one-step comparison with Rn-SNEB. Work means include successful runs only.

Setting	Method	Successes	AIT	ALS	ANF	ARED
(a) Original nonconvex signed-start suite						
$n = 100$	NBB	20/20	23.90	23.90	76.50	27.70
$n = 100$	Rn-NSNA	20/20	26.50	26.50	91.20	37.20
$n = 100$	Rn-SNEB	0/20	–	–	–	–
$n = 150$	NBB	20/20	34.90	34.90	121.95	51.15
$n = 150$	Rn-NSNA	20/20	42.05	42.05	161.25	76.15
$n = 150$	Rn-SNEB	0/20	–	–	–	–
$n = 200$	NBB	18/20	78.78	78.78	320.50	161.94
$n = 200$	Rn-NSNA	17/20	89.00	89.00	394.35	215.35
$n = 200$	Rn-SNEB	0/20	–	–	–	–
$r = 20$	NBB	27/30	68.67	68.67	272.22	133.89
$r = 20$	Rn-NSNA	28/30	94.96	94.96	416.04	225.11
$r = 20$	Rn-SNEB	0/30	–	–	–	–
$r = 50$	NBB	30/30	44.90	44.90	164.63	73.83
$r = 50$	Rn-NSNA	30/30	60.03	60.03	242.10	121.03
$r = 50$	Rn-SNEB	0/30	–	–	–	–
$r = 100$	NBB	30/30	38.37	38.37	135.77	58.03
$r = 100$	Rn-NSNA	30/30	41.37	41.37	155.13	71.40
$r = 100$	Rn-SNEB	0/30	–	–	–	–
(b) Monotone linear construction of Tang–Zhou, Example 6.2						
$n = 100$	NBB	6/6	4.83	4.83	11.50	0.83
$n = 100$	Rn-NSNA	6/6	4.83	4.83	11.50	0.83
$n = 100$	Rn-SNEB	6/6	5.50	5.50	12.67	0.67
$n = 500$	NBB	6/6	5.67	5.67	14.33	2.00
$n = 500$	Rn-NSNA	6/6	5.50	5.50	14.00	2.00
$n = 500$	Rn-SNEB	6/6	5.83	5.83	13.83	1.17
$n = 1000$	NBB	6/6	5.67	5.67	14.50	2.17
$n = 1000$	Rn-NSNA	6/6	5.50	5.50	14.17	2.17
$n = 1000$	Rn-SNEB	6/6	5.67	5.67	13.67	1.33

Note: Panel (a) uses the original 150 instances: The first three groups have $r = 20$, and the last three have $n = 200$. Panel (b) uses three generated problems and two starts per dimension (18 matched instances). All methods use a 220-iteration budget, 92 trial points per iteration, native residual tolerance 10^{-8}, and common KKT tolerance 10^{-7}. Method-specific smoothing and forcing parameters are retained. Dashes mean no successful run, not zero cost. ALS counts reduced linear

solves.

To also compare within a monotone class, we use the construction of [11, Example 6.2]:

$$F(x, s) = Bx - q - s,$$

where B consists of four normalized Gram blocks, $q = -\operatorname{rand}(n, 1)$, and $w = \operatorname{rand}(n, 1)$.

At $n = 100, 500, 1000$, three generated problems and the two published starting-point types give 18 matched instances. The source provides no seeds, so these are reproducible new instances of its construction, not a reproduction of its exact tables. All methods succeed on all 18 instances; Table 9b therefore compares identical successful subsets. Across these instances, the mean iteration/evaluation/rejection counts are 5.39/13.44/1.67 for NBB, 5.28/13.22/1.67 for Rn-NSNA, and 5.67/13.39/1.06 for Rn-SNEB. Thus, NBB has no uniform work advantage, and Rn-SNEB rejects fewer trials here. Three additional sparse positive-weight checks at $n = 100, 400, 1600$ also succeed for all methods and are retained in the supplement. All methods in this subsection use one reduced linear solve per attempted Newton iteration; no timing ranking is inferred from these tests.

5.2.6. Further parameter sensitivity tests

The complementarity-family parameter $\tau \in [0, 2)$ is fixed within each run and is distinct from the iteratively updated smoothing parameter μ. Theorem 3.1 gives the same WCP solution set for every admissible τ; nonzero values are also allowed in Theorem 4.2 under its structural and bounded-level-set assumptions. These statements do not guarantee convergence within a fixed budget for arbitrary nonconvex starts. We test $\tau \in \{0, 10^{-8}, 10^{-6}, 10^{-4}, 10^{-3}, 0.01, 0.1, 0.25, 0.5, 1, 1.5, 1.9\}$ on the same 12 nonconvex weighted instances, keeping the signed starts at initialization scale $r = 50$, other parameters, and the 220-iteration budget fixed. All 12 runs succeed at zero and at each of the four tested positive values up to 10^{-3}. Every run at each of the seven tested values from 0.01 to 1.9 reaches the iteration limit. Table 10a summarizes all 144 runs. A zero success count means that the stated tolerances are not reached, not that the solver returns no iterate or that the WCP has no solution.

Independent implementation checks reproduce the success flags and work counts of all 144 runs. At 96 initial/final states, an independently evaluated sum-of-squares/rationalized residual agrees with the implemented mapping to 1.34×10^{-14} relative to $\max\{1, \|H\|\}$. Blockwise central-difference derivative discrepancies are at most 5.72×10^{-8}, and reduced versus directly assembled full-Jacobian directions differ relatively by at most 1.52×10^{-11}, with normwise backward errors below 1.70×10^{-17}. Full-system runs on the first instance retain the same success/failure pattern at all twelve values, although work counts change on these sensitive trajectories. These checks reveal no missing τ term, derivative-sign error, or reduced-system inconsistency; finite checks are not a proof of numerical reliability at every iterate.

Increasing only the original-start iteration budget to 2200 leaves all 84 previously unsuccessful runs unsuccessful. Their final $\|H\|$ values range from 6.05×10^4 to 1.88×10^5, C_{2200}/C_0 lies between 0.99999960 and 0.99999990, and μ remains at 10^{-5}. Thus, these are large-residual trajectories, not solutions rejected solely by an overly strict final tolerance. The structural hypothesis is also violated at every original start: For an orthonormal basis Z of $\ker A$, the minimum eigenvalues of $Z^\top \nabla^2 f(x^0) Z$ range from -695.0 to -426.7. The global theorem therefore does not cover these starts. The large retained reference values document weak progress of the nonmonotone globalization

on this stress set; they do not identify a unique failure cause or prove nonconvergence with an unlimited budget. Separate initialization diagnostics use the same problems, not replacement instances. For $\tau = 0, 0.5, 1, 1.5, 1.9$, replacing x^0, s^0 by their absolute values at $r = 50$ gives success counts of $12, 0, 0, 0, 0$ out of 12. Scaling the original Gaussian directions for all three variables to initialization scale $r = 1$, without using any solution information, instead gives $11, 7, 3, 0, 0$ successes at those five values. Near-singular-system warnings occur in three unsuccessful $\tau = 1.9$ runs with $r = 1$ and are retained in the audit log. Finally, perturbing each computed $\tau = 0$ solution by fixed-seed noise of scale 10^{-3} gives 12/12 successes at every tested $\tau \in \{0.01, 0.1, 0.25, 0.5, 1, 1.5, 1.9\}$, each in two iterations, with maximum

$$R_{\mathrm{KKT}} = 2.80 \times 10^{-10};$$

see Table 10b. This last check uses a computed solution and is explicitly a local-usability diagnostic, not an independent-start performance comparison or an analytical solution test. None of these outcomes replaces the original-start results.

Table 10. Fixed-τ sensitivity and follow-up diagnostics on the same 12 nonconvex weighted problems.

τ	Successes	AIT	ANF	ARED	AKKT
(a) Original signed starts, 220-iteration budget					
0	12/12	46.17	175.25	81.92	5.40×10^{-10}
10^{-8}	12/12	37.50	140.25	64.25	3.32×10^{-9}
10^{-6}	12/12	49.50	186.33	86.33	2.61×10^{-9}
10^{-4}	12/12	34.00	115.25	46.25	2.73×10^{-10}
10^{-3}	12/12	45.08	164.92	73.75	1.01×10^{-9}
0.01	0/12	–	–	–	–
0.1	0/12	–	–	–	–
0.25	0/12	–	–	–	–
0.5	0/12	–	–	–	–
1	0/12	–	–	–	–
1.5	0/12	–	–	–	–
1.9	0/12	–	–	–	–

(b) Larger budget and separate near-solution diagnostics

	Original starts, 2200 iterations		Near-solution starts		
τ	Successes	Final $\|\|H\|\|$ range	Successes	AIT	Max. R_{KKT}
0.01	0/12	$[9.26 \times 10^4, 1.85 \times 10^5]$	12/12	2.00	4.22×10^{-11}
0.1	0/12	$[7.69 \times 10^4, 1.82 \times 10^5]$	12/12	2.00	4.39×10^{-11}
0.25	0/12	$[8.18 \times 10^4, 1.88 \times 10^5]$	12/12	2.00	4.67×10^{-11}
0.5	0/12	$[9.74 \times 10^4, 1.76 \times 10^5]$	12/12	2.00	5.15×10^{-11}
1	0/12	$[1.17 \times 10^5, 1.71 \times 10^5]$	12/12	2.00	8.22×10^{-11}
1.5	0/12	$[8.35 \times 10^4, 1.71 \times 10^5]$	12/12	2.00	1.72×10^{-10}
1.9	0/12	$[6.05 \times 10^4, 1.56 \times 10^5]$	12/12	2.00	2.80×10^{-10}

Note: Panel (a) retains every original result; means use successful runs only, and dashes do not mean that the solver was not run. In panel (b), only the budget changes for the original-start runs. The near-solution runs instead perturb the computed $\tau = 0$ solution by fixed-seed noise of scale 10^{-3} and retain the 220-iteration cap. They test local usability, not success from an independent start, and do not replace panel (a). All runs require $\|H\| \le 10^{-8}$ and $R_{\mathrm{KKT}} \le 10^{-7}$.

The monotone checks independently confirm that larger nonzero values are usable. On the 20×20 active-obstacle grid, the five values $0, 0.5, 1, 1.5, 1.9$ all succeed with iteration counts $8, 10, 10, 12, 22$ and rejected-trial counts $0, 0, 0, 1, 23$, respectively. On the monotone positive-weight linear model with $n = 400$ defined in Section 5.4, the same five values succeed in $3, 4, 4, 5, 5$ iterations without rejected trials, and the common residual is at most 2.30×10^{-9}. The primary comparison uses $\tau = 0$. These results establish neither a sharp parameter threshold nor uniform optimality of zero; changing τ changes the smoothing geometry and Newton trajectory even though the limiting WCP solution set is unchanged. Improving robustness from difficult nonconvex starts or choosing τ adaptively requires a separate algorithmic study. We examine the forcing and contraction parameters in the two blocks on the same 12 instances, with $\tau = 0$ and $\sigma = 0.2$. The first block fixes

$$\mu_0 = 10^{-4}$$

and tests all nine combinations of

$$\gamma \in \{10^{-6}, 10^{-5}, 10^{-4}\} \quad \text{and} \quad \lambda \in \{0.4, 0.5, 0.7\}.$$

All 108 runs succeed. Their mean iteration counts range from 34.00 to 48.17, and their mean rejected-trial counts range from 45.08 to 117.17.

The second block tests all six values

$$\gamma \in \{10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 0.1\}$$

with the same three contraction factors. Since Algorithm 3.1 requires $\gamma \le \mu_0$, we set a common $\mu_0 = 0.1$ in this block to accommodate the larger forcing coefficients. The condition

$$\sigma(2 - \sqrt{2}\gamma) < 1$$

holds throughout this grid. The problem instances, signed (x^0, s^0, y^0) starts, other parameters, and computational budgets remain unchanged. Fixing μ_0 within each block permits within-block assessment of γ and λ; differences between the blocks also reflect the changed initial smoothing value.

All 216 runs in the second block succeed. Their mean iteration counts range from 37.75 to 59.17, and their mean rejected-trial counts range from 50.67 to 156.17. In particular, $\gamma = 0.1$ gives 12/12 successes at each $\lambda = 0.4, 0.5, 0.7$, with mean iteration counts 42.17, 44.58, and 46.42 and mean rejected-trial counts 55.08, 67.83, and 97.50, respectively. Table 11 reports both blocks. These results support the usability of larger γ on this test set, but establish neither uniform parameter insensitivity nor a general advantage of $\gamma = 0.1$. The primary comparisons use $\mu_0 = 10^{-4}$ and $\gamma = 10^{-5}$; their parameters are not selected using these sensitivity results.

Table 11. Joint forcing and contraction study at two initial smoothing values on the same 12 instances.

γ	λ	Successes	AIT	ANF	ARED
(a) Original block: $\mu_0 = 10^{-4}$					
10^{-6}	0.4	12/12	35.92	125.25	52.42
10^{-6}	0.5	12/12	41.08	154.33	71.17
10^{-6}	0.7	12/12	44.25	191.83	102.33
10^{-5}	0.4	12/12	40.08	141.67	60.50
10^{-5}	0.5	12/12	46.17	175.25	81.92
10^{-5}	0.7	12/12	45.08	201.58	110.42
10^{-4}	0.4	12/12	34.00	114.08	45.08
10^{-4}	0.5	12/12	44.50	163.33	73.33
10^{-4}	0.7	12/12	48.17	214.50	117.17
(b) Expanded block: $\mu_0 = 0.1$					
10^{-6}	0.4	12/12	40.42	141.08	59.25
10^{-6}	0.5	12/12	57.00	217.50	102.50
10^{-6}	0.7	12/12	48.17	217.33	120.00
10^{-5}	0.4	12/12	39.08	133.33	54.17
10^{-5}	0.5	12/12	46.33	172.00	78.33
10^{-5}	0.7	12/12	43.58	207.58	119.42
10^{-4}	0.4	12/12	37.75	127.17	50.67
10^{-4}	0.5	12/12	45.83	170.17	77.50
10^{-4}	0.7	12/12	59.17	275.50	156.17
10^{-3}	0.4	12/12	43.33	146.75	59.08
10^{-3}	0.5	12/12	42.42	153.75	67.92
10^{-3}	0.7	12/12	56.92	259.92	145.08
10^{-2}	0.4	12/12	43.92	149.25	60.42
10^{-2}	0.5	12/12	43.17	156.08	68.75
10^{-2}	0.7	12/12	47.25	205.00	109.50
0.1	0.4	12/12	42.17	140.42	55.08
0.1	0.5	12/12	44.58	158.00	67.83
0.1	0.7	12/12	46.42	191.33	97.50

Note: Each row uses the same 12 problems and signed (x^0, s^0, y^0) starts, with $\tau = 0$, $\sigma = 0.2$, and 220 iterations at most. The initial smoothing value is fixed within each block to satisfy $\gamma \leq \mu_0$. Means include successful runs only. Cross-block differences also reflect the change in μ_0; neither block is used to retune the primary comparisons.

5.3. Physical-model benchmarks

The preceding tests isolate algorithmic behavior on controlled dense instances. We now examine three complementarity models arising from continuum mechanics. In all three applications, $w = 0$ and there are no equality constraints. The WCP therefore reduces to the nonlinear complementarity

problem

$$q \geq 0, \qquad F(q) \geq 0, \qquad q \circ F(q) = 0.$$

This reduced problem fits the present formulation when we set $x = q$ and use the WCP equation block $\mathcal{F}(q, s) = F(q) - s = 0$. Sparse five-point finite differences are used. Consistently with the implementation, the reported common residual is

$$\max\{\|F(q) - s\|_\infty, \|q \circ s\|_\infty, \|q_-\|_\infty, \|s_-\|_\infty\}.$$

When the equation block is satisfied exactly, eliminating $s = F(q)$ gives the familiar nonlinear complementarity problem (NCP) residual based on $q \circ F(q)$, q_-, and $F(q)_-$. The implemented form also records the finite-precision equation-block error and simultaneously measures feasibility of the gap or pressure, feasibility of the multiplier or residual force, and complementarity. The physical-model tables use SR, IT, and NF as defined in Section 5.1. LSys denotes the number of Newton linear-system solves, whose successful-run mean is ALS. BTR denotes the rejected-trial count, with the same counting rule as RED; its successful-run mean is ARED. The combined heading IT/NF/LSys lists three separate counts in that order, not their ratio. Time (s) denotes elapsed solver time, with any repeated-run averaging stated alongside the table. The main physical-model tables report NBB alone. In Section 5.3.4, we add an independent quadratic-programming reference for the contact set, rather than a timing comparison with a specialized physical solver. The purpose is to check whether NBB solves these discrete models to a small complementarity residual and whether selected field quantities stabilize under mesh refinement. The first obstacle example has a nonempty active contact set, the obstacle Bratu problem tests a nonlinear nonmonotone membrane operator, and the Reynolds model displays complementarity and free-boundary behavior through cavitation.

These zero-weight examples illustrate the NCP special case and do not alone validate the nonzero-weight extension. A separate sparse positive-weight test is given in Section 5.4. The active-obstacle and Reynolds mappings also connect directly to the global theory. Their Jacobians are the symmetric positive-definite finite-difference matrices L_N and K_h, respectively. Since $m = 0$, the rank condition on $\mathcal{F}'_y$ is vacuous. Moreover, every direction in the null space of the linearized equation block satisfies $\Delta s = F'(q)\Delta q$, and hence $\Delta q^\top \Delta s = \Delta q^\top F'(q)\Delta q \geq 0$. This verifies the monotonicity part of Assumption 3.1; Proposition 4.1 also verifies boundedness of their fixed-grid level sets. The nonlinear Bratu Jacobian need not be positive semidefinite, so that benchmark remains a stress test outside this guaranteed monotone subclass.

5.3.1. Active obstacle problem with a nonempty contact set

We first consider a linear obstacle problem [26–28] with an active contact region. We denote the membrane displacement and the obstacle by u and ψ, respectively, and define the gap variable by $q = u - \psi$. The continuous complementarity conditions are discretized in the form

$$q \geq 0, \qquad L_N(q + \psi) - \ell \geq 0, \qquad q \circ (L_N(q + \psi) - \ell) = 0,$$

where L_N is the five-point discretization of $-\Delta$ with zero boundary data. We choose

$$\psi(\xi, \eta) = 0.1 \sin(\pi\xi) \sin(\pi\eta), \qquad \ell \equiv 1,$$

on $\Omega = (0, 1)^2$. This obstacle is positive in the interior and zero on the boundary; the chosen obstacle–load pair produces the nonempty active contact zone reported below. The starting point is $q^0 = s^0 = \mathbf{1}$. Table 12 reports averages over five repeated runs.

Table 12. NBB results for the active obstacle problem with a nonempty contact set.

N	n	SR (%)	IT/NF/LSys	BTR	R_{KKT}	Contact (%)	$u_{\min}$	$u_{\max}$
20	400	100	8/17/8	0	4.04×10^{-14}	19.00	0.004258	0.099442
30	900	100	10/21/10	0	4.83×10^{-13}	17.33	0.002212	0.099743
40	1600	100	13/27/13	0	1.78×10^{-13}	16.50	0.001371	0.099853

NBB succeeds on all grids and keeps R_{KKT} below 5×10^{-13}. The active contact fraction, computed as the proportion of nodes with $q_i \leq 10^{-7}$, remains between 16.50% and 19.00%. The maximum displacement approaches the obstacle amplitude 0.1 under refinement. BTR is zero in this example, meaning that every accepted update uses the full Newton trial without a rejected line-search evaluation. The small residuals indicate that the computed gap, multiplier, and complementarity product are simultaneously accurate.

Figure 2 shows a central contact region where the membrane coincides with the positive obstacle, while values near the outer boundary approach the prescribed zero displacement. The computed contact contour has no visually apparent oscillations on this grid. Together with the contact fractions in Table 12, this provides a qualitative consistency check, not a proof of active-set convergence.

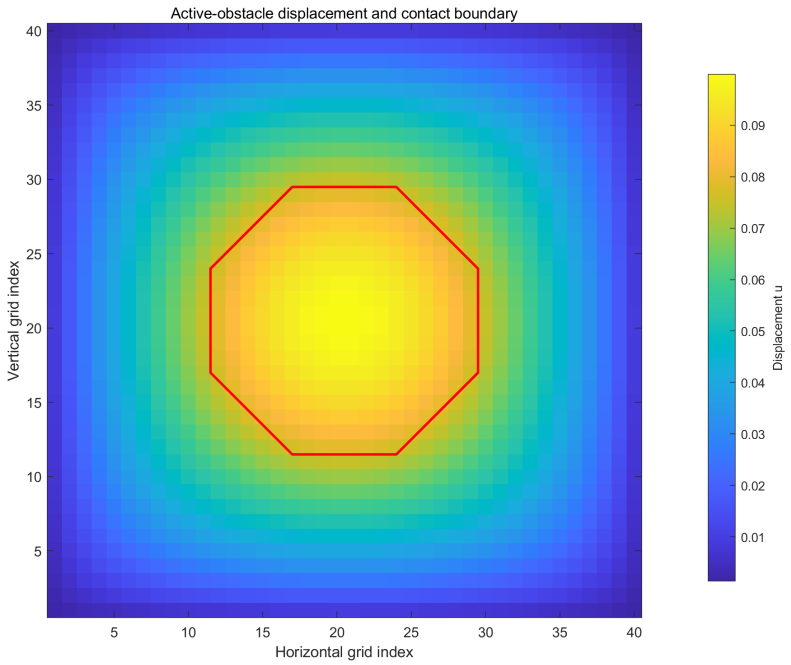


Figure 2. Computed displacement field for the active obstacle problem on the 40×40 interior grid. Color represents the displacement u, and the red contour marks the boundary of the active contact region.

5.3.2. Obstacle Bratu problem for an elastic membrane

We consider an elastic membrane over

$$\Omega = (0, 1)^2$$

that is constrained by an obstacle ψ and subjected to a Bratu-type load. The exponential source follows the classical Bratu/Liouville–Bratu–Gelfand family of nonlinear elliptic problems [29, 30]. The unilateral obstacle constraint follows the variational-inequality formulation [26, 28]. The equilibrium conditions can be written as

$$-\Delta u + \kappa e^{-u} \geq 0, \qquad u - \psi \geq 0, \qquad (-\Delta u + \kappa e^{-u})(u - \psi) = 0, \qquad u|_{\partial\Omega} = 0.$$

We set $v = u - \psi$, denote by L_N the five-point approximation of $-\Delta$ on an $N \times N$ interior grid, and use the test data $\psi = -4$ and $\kappa = 1$, where κ is the Bratu load parameter. The discrete mapping is

$$F_N(v) = L_N(v + \psi) + \kappa \exp(-v - \psi), \qquad n = N^2.$$

This mapping is nonlinear and nonmonotone in general. Table 13 reports NBB results obtained from the starting point $v^0 = s^0 = \mathbf{1}$. Each timing is averaged over five runs. The columns $u_{\min}$ and $u_{\max}$ are included to check whether the computed membrane shape stabilizes as the grid is refined.

Table 13. NBB results for the obstacle Bratu problem under grid refinement.

N	n	IT	NF	LSys	Time (s)	R_{KKT}	$u_{\min}$	$u_{\max}$
20	400	5	13	5	0.0176	5.00×10^{-11}	-0.077663	-0.004092
30	900	5	12	5	0.0552	5.00×10^{-11}	-0.077900	-0.002137
40	1600	5	11	5	0.0248	5.00×10^{-11}	-0.077986	-0.001328

NBB succeeds on every grid and keeps R_{KKT} at or below 5.0×10^{-11}. The iteration and linear-system counts remain equal to five as n increases from 400 to 1600. The timing is not monotone across these short runs and is reported only as a supplementary implementation-dependent quantity. The minimum displacement changes from -0.077663 to -0.077986; the relative change between the two finest grids is approximately 0.11%. Moreover, $u_{\max}$ approaches the zero Dirichlet boundary value. These observations provide a mesh-consistency check that is independent of the stopping residual.

For the data $\psi = -4$ and $\kappa = 1$, the computed gap $v = u - \psi$ remains positive, and the contact multiplier is zero to numerical precision. The obstacle is therefore inactive at equilibrium: This benchmark tests the nonlinear nonmonotone Bratu operator rather than an active contact zone. Figure 3 shows the resulting smooth, centrally depressed displacement profile, consistent with the zero boundary condition and the displacement extrema in Table 13.

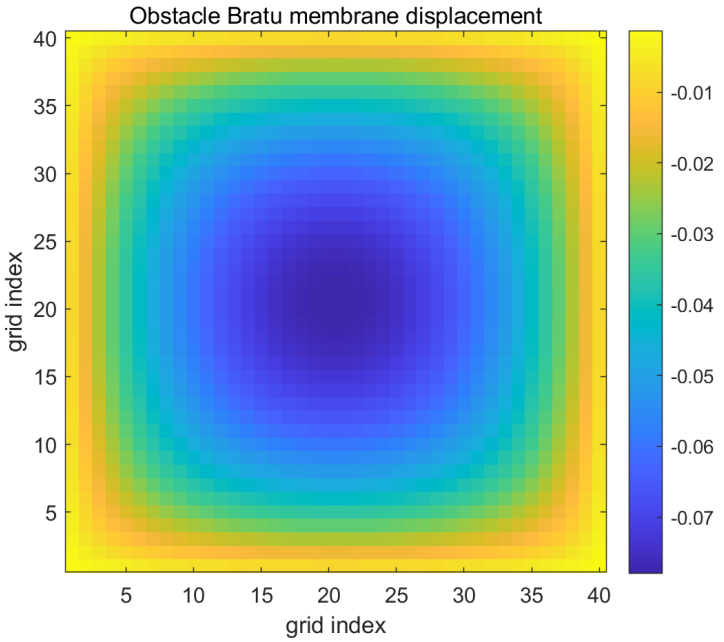


Figure 3. Computed displacement field for the obstacle Bratu membrane on the 40×40 interior grid. The color scale represents the membrane displacement u; the most negative displacement occurs near the center.

To test sensitivity to initialization, we fix the same discrete Bratu problem on the 30×30 interior grid ($n = 900$) and use 20 signed starting pairs at each scale:

$$v_j^0 = r\xi_j, \qquad s_j^0 = r\eta_j, \qquad j = 1, \ldots, 20.$$

Here, $\xi_j, \eta_j \in \mathbb{R}^{900}$ are independent standard normal vectors generated with seeds $830000 + j$. The same 20 vector pairs are reused at $r = 0.5, 1, 2, 5$. These vectors are not normalized: r is an initialization scale, not a physical radius or the Euclidean distance to a solution. In particular, $r = 2$ doubles each component of the corresponding starting pair at $r = 1$. The problem, mesh, and solver parameters are unchanged; only the initial values vary.

A run is counted as successful if it satisfies $\|H(z^k)\|_2 \le 10^{-8}$ and $R_{\mathrm{KKT}} \le 10^{-7}$ within 200 iterations. Each line search permits at most 82 trials, including the full-step trial. For each scale, the empirical success rate is

$$\mathrm{SR} = 100\,\frac{N_{\mathrm{succ}}}{20}\%,$$

where N_{succ} is the number of successful starting pairs, not the number of solved grid nodes. Thus, the entry 25% at $r = 2$ means that 5 of the 20 runs meet the stopping criteria, while the remaining 15 reach the iteration limit. It is neither a solution-accuracy percentage nor an estimate established for arbitrary initial values. Table 14 reports the counts and percentages. Its work and timing means include successful runs only; in particular, AIT = 13.00 at $r = 2$ averages the five successes and excludes the fifteen 200-iteration failures.

Table 14. Sensitivity of NBB to the initialization scale for one fixed obstacle Bratu problem on the 30×30 interior grid. Each row uses 20 starting pairs; work and timing means include successful runs only.

r	Successes	SR (%)	AIT	ATime (s)	ANF	ALS	ARED	ABB
0.5	20/20	100	17.65	0.0425	37.30	17.65	1.00	0.00
1	20/20	100	18.40	0.0414	37.80	18.40	0.00	0.00
2	5/20	25	13.00	0.0298	28.00	13.00	1.00	0.00
5	20/20	100	32.80	0.0754	106.50	32.80	39.90	12.30

Every successful run converges to the same membrane field as the grid-refinement experiment, with an average common residual between 3.50×10^{-11} and 5.00×10^{-11}. Success is not monotone in the initialization scale. For example, the starting pair generated with seed 830002 succeeds in 18 iterations at $r = 1$ and 32 iterations at $r = 5$, but fails within 200 iterations at $r = 2$. Scaling a signed start changes the trajectory through the nonlinear, nonmonotone Bratu mapping; a larger scale is not necessarily harder. At $r = 5$, an admissible positive BB factor is available on 12.3 iterations and NBB records 39.9 rejected trial steps on average. These tests demonstrate sensitivity to initialization, not uniform robustness over all starts or a universal 25% success probability for Bratu problems.

Failure diagnostics. We rerun all 80 signed-start cases while recording accepted steps, the smoothing parameter, the rational reference value, an estimate of the reciprocal condition number of the reduced Newton matrix, and the normwise backward error of the full linear equation. All success flags, termination statuses, and work counts are reproduced. The five successes at $r = 2$ correspond to $j = 1, 3, 4, 6, 14$ and require $12, 14, 13, 11, 15$ iterations, respectively. The other 15 runs reach 200 iterations with $\|H\|_2$ between 4.89×10^3 and 2.03×10^5, far above the 10^{-8} threshold. None succeeds when rerun from the same initial pair with the budget increased to 800 iterations. Thus, these are genuine failures of the specified finite-budget test, not rounding effects at the stopping threshold. They do not establish nonconvergence with an unlimited budget.

The reference and smoothing updates explain an obstruction along these failed trajectories. With $M_{k+1} = M(z^{k+1})$, the implemented rational update satisfies

$$C_k - C_{k+1} = \frac{C_k - M_{k+1}}{M_{k+1} + 1}.$$

Consequently, accepted steps with large merit values can reduce the reference only slightly relative to its size. At iteration 200, the 15 failed runs retain between 99.9815% and 99.99998% of their initial reference values. In the 200- and 800-iteration records, their reference values remain above 1 at every step, so $\beta_k = \gamma = 10^{-5}$. The first Newton equation then gives

$$\mu_{k+1} = (1 - \alpha_k)\mu_k + \alpha_k\gamma.$$

Since $\mu_0 \geq \gamma$ and $0 < \alpha_k \leq 1$, these steps keep $\mu_k \geq \gamma$, and therefore $\|H(z^k)\|_2 \geq \mu_k \geq 10^{-5} > 10^{-8}$. Thus, the stopping criterion cannot be met while the reference remains in this regime. This identifies an observed obstruction; it does not prove that the reference must stay above 1 indefinitely or that the forcing term is the sole cause of the large equation residuals.

Across the failed 200-iteration trajectories, the smallest estimated reciprocal condition number is 2.33×10^{-9}. The largest normalized linear backward error,

$$\frac{\|H'(z^k)\Delta z^k - (-H(z^k) + \beta_k h)\|_\infty}{\|H'(z^k)\|_\infty \|\Delta z^k\|_\infty + \| - H(z^k) + \beta_k h\|_\infty},$$

is 2.04×10^{-19}. Small backward errors do not guarantee small forward errors for ill-conditioned systems, but none of these failures is reported as a singular-solve exception. Figure 4 shows the residual excursions and slow reference reduction. Independent checks reproduce all 80 original-budget and 15 extended-budget outcomes. Directional finite differences verify the Bratu Jacobian at all four scales for seed 830002, with a maximum relative discrepancy of 6.49×10^{-9}. These checks support the reported observations but do not establish a unique cause or characterize all initial values that will succeed. The nonlinear Bratu problem is not asserted to satisfy Assumption 3.1 along these trajectories, so the global convergence theorem does not supply a guarantee for these starts. BB safeguarding does not regularize the Newton matrix.

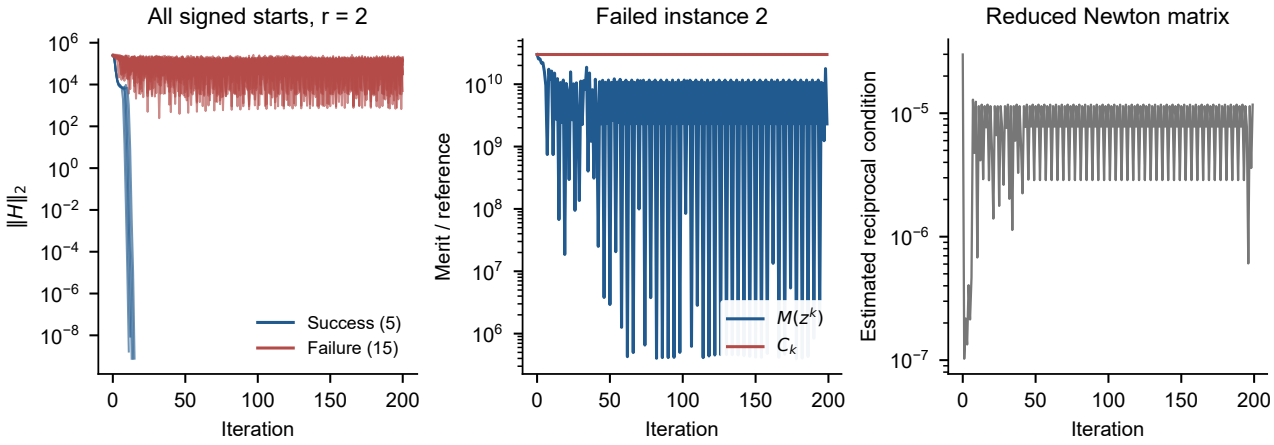


Figure 4. Signed-start Bratu diagnostics at $r = 2$. Left: All 20 residual trajectories, including the five successes and 15 failures. Middle: The merit and reference values for failed instance 2. Right: The reciprocal-condition estimate for that instance. Diagnostics stop at the original 200-iteration budget; extended-budget outcomes are reported in the text.

5.3.3. Reynolds cavitation in a finite journal bearing

Hydrodynamic lubrication with Swift–Stieber cavitation provides another inequality-governed complementarity model [31–33]. Complementarity formulations have also been developed for mass-conserving cavitation models [34]. In this experiment, we use the classical, simpler Swift–Stieber conditions. The experiment should be interpreted as a numerical model benchmark rather than a complete cavitation model. On $\Omega = (0, 2\pi/3) \times (0, 1)$, we impose zero pressure on the boundary and use the complementarity conditions

$$\mathcal{K}p = -\frac{\partial}{\partial\theta}\left(d^3\frac{\partial p}{\partial\theta}\right) - \frac{1}{4}\frac{\partial}{\partial y}\left(d^3\frac{\partial p}{\partial y}\right), \qquad p \geq 0, \quad \mathcal{K}p - g \geq 0, \quad p(\mathcal{K}p - g) = 0,$$

where $d(\theta) = 1 - \varepsilon\cos(\theta - \theta_0)$, $g = -6\partial d/\partial\theta$, and the circumferential phase is $\theta_0 = 0.548388$. The minus-cosine convention fixes only the angular origin at the minimum film thickness. It is equivalent

to the usual plus-cosine form after a phase shift. In the positive-pressure region, the Reynolds equation is active. In the cavitated region, the pressure is zero, and the residual inequality is active. Flux coefficients are evaluated at cell faces, giving the sparse linear complementarity mapping $F_h(p) = K_h p - g_h$ with $n = n_\theta n_y$. We set $\varepsilon = 0.9$ on all three grids. Table 15 reports the grid-refinement results.

Table 15. NBB results for Reynolds cavitation under grid refinement at $\varepsilon = 0.9$. Cav. denotes the cavitation fraction.

$n_\theta \times n_y$	n	IT/NF/LSys	Time (s)	R_{KKT}	max p	W_h	Cav. (%)
30×10	300	5/11/5	0.0051	2.48×10^{-10}	19.2297	6.0799	67.33
60×20	1200	8/17/8	0.0222	9.07×10^{-10}	19.4088	6.2031	67.00
90×30	2700	13/27/13	0.0951	6.69×10^{-14}	19.4248	6.2244	66.44

Wall-clock times are averages over five runs. Here, $W_h = \sum_{i,j} p_{ij} \Delta\theta \Delta y$ is a scalar dimensionless pressure integral, and the cavitation fraction is the percentage of interior nodes satisfying $p_{ij} \leq 10^{-7}$. The integral measures total pressure over the modeled sector; it is not the magnitude of the journal-bearing resultant load, which requires angular projections of pressure. The cavitation fraction describes the discrete zero-pressure area, not lubricant mass loss or a directly measured cavitation volume. Neither quantity is used as an efficiency score for NBB. NBB succeeds on all three grids and keeps $R_{\mathrm{KKT}} < 10^{-9}$. The pressure peak changes by approximately 0.93% from the 30×10 to the 60×20 grid and by 0.082% from the 60×20 to the 90×30 grid. The corresponding changes in W_h are 2.03% and 0.34%, showing consistent pointwise and integral quantities over the tested refinements.

Figure 5 shows positive pressure concentrated in a localized angular band, with zero pressure over most of the remaining domain. These regions correspond to the active Reynolds equation and the active pressure constraint, respectively. The small common residuals in Table 15 confirm that the computed fields satisfy the discrete Swift–Stieber complementarity conditions to the reported tolerance.

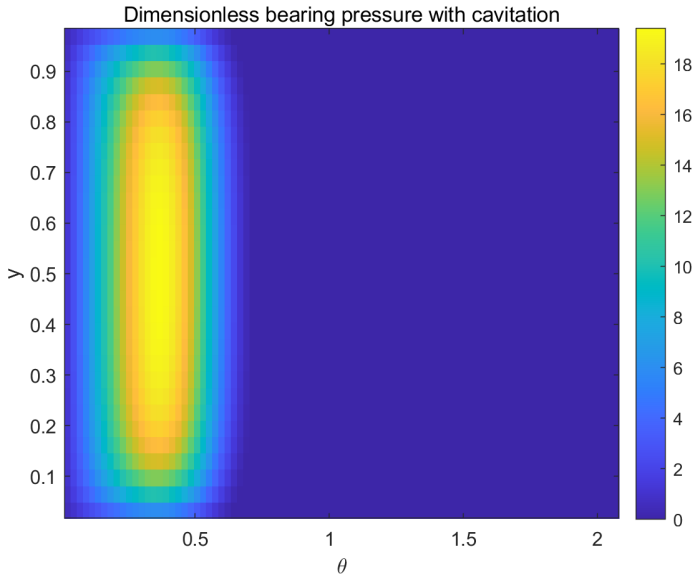


Figure 5. Dimensionless Reynolds pressure at $\varepsilon = 0.9$ on the 90×30 grid. Color represents the pressure p.

5.3.4. Mesh convergence and contact-set checks

We assess spatial refinement through physical outputs, field profiles, and successive-grid differences, separately from convergence of the nonlinear solver. Grid-refinement methodology distinguishes stabilization of a computed output from its error relative to a continuum solution [35]. In lubrication studies, changes in hydrodynamic pressure under mesh refinement are used to assess grid sensitivity [36, 37]. We adopt this reporting approach, not the geometries or multiphase models of those studies.

Each model uses five nested meshes with its physical parameters and boundary conditions fixed. For the square-domain models, the interior counts are $N = 15, 31, 63, 127, 255$, with $h = 1/(N + 1)$. For Reynolds, $n_y = 7, 15, 31, 63, 127$ and $n_\theta = 3(n_y + 1) - 1$; both coordinate spacings are halved at each refinement, and $h = 1/(n_y + 1)$.

To check that algebraic inaccuracy does not dominate the grid comparisons, we apply three active-set Newton corrections to each saved NBB field. At each correction, nodes satisfying $U - \psi \le G(U)$ are fixed to the obstacle, and the linearized equilibrium equation is solved on the remaining nodes. Here, G is the discrete physical operator and $\psi = 0$ for Reynolds. For Bratu, this check evaluates the equation directly in the displacement u, rather than repeatedly recovering it from the gap $v = u + 4$. The corrected fields satisfy

$$R_{\text{eq}} = \max\{\|(U - \psi)_-\|_\infty, \|G(U)_-\|_\infty, \|(U - \psi) \circ G(U)\|_\infty\} \le 2.27 \times 10^{-11}.$$

The largest change from a saved field is 8.51×10^{-10}. For each adjacent pair, the sum of the two correction norms divided by its infinity-norm grid difference is at most 4.45×10^{-7}. This is an a posteriori algebraic-accuracy check, not a rigorous continuum error bound or a competing-solver experiment. We use these residual-checked fields for the following spatial comparisons.

The original finest-grid Bratu run still has $\|H\|_2 = 1.14 \times 10^{-8}$ after 220 iterations and therefore fails the fixed 10^{-8} stopping test, despite $R_{\text{KKT}} = 1.05 \times 10^{-10}$. Its post-solve displacement correction is only 9.93×10^{-13}. The extra corrections do not change the recorded NBB status or work counts; 14 of the original 15 nested-grid NBB runs pass both original tests.

For the obstacle membrane, we monitor $V_h = \Delta x \Delta y \sum u_{ij}$, rather than the maximum displacement, which is pinned by the obstacle height. For Bratu we use the maximum downward deflection $d_h = -\min u_{ij}$. For Reynolds, we report the peak pressure $P_h = \max p_{ij}$ and the sector pressure integral $W_h = \Delta\theta \Delta y \sum p_{ij}$. The sums are tensor-product trapezoidal quadratures, including the prescribed zero boundary values. The quantities are dimensionless; W_h is not the resultant bearing load. The relative change assigned to a refined grid is

$$\delta Q_{h/2} = 100 \frac{|Q_{h/2} - Q_h|}{|Q_{h/2}|} \quad (\%).$$

Figure 6 and Table 16 show the approach to stable physical outputs. On the last refinement, the changes in V_h, d_h, P_h, and W_h are 0.0114%, 0.0034%, 0.0080%, and 0.0647%, respectively. These are relative changes between two grids, not percentages of error against an exact solution. The Reynolds peak pressure approaches its fine-grid value nonmonotonically; we retain that behavior rather than imposing a monotone fit or assigning a universal convergence order.

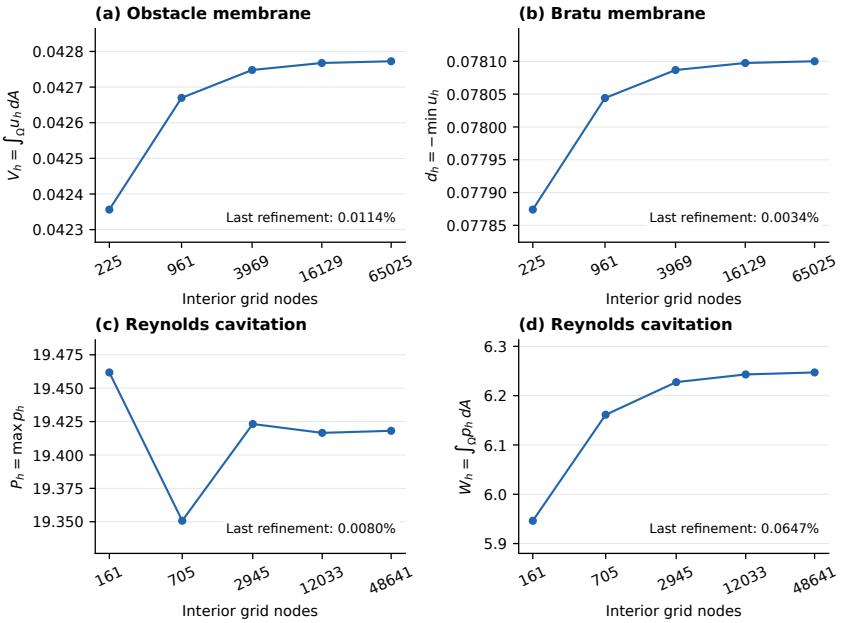


Figure 6. Physical-output grid-convergence study with five meshes per model. The abscissa is the number of interior nodes on a logarithmic scale. Panels show (a) the obstacle displacement integral, (b) the Bratu maximum downward deflection, (c) the Reynolds peak pressure, and (d) the sector pressure integral. Annotations give the last relative refinement change. Lines connect computed values only. The post-solve accuracy check and the original Bratu stopping-test exception are described in the text.

Table 16. Physical outputs under systematic grid refinement.

Model	Grid	Q_h	δQ (%)	W_h	δW (%)	R_{eq}
Obstacle	15	0.04235627	–	–	–	1.07×10^{-14}
Obstacle	31	0.04266975	0.7347	–	–	6.39×10^{-14}
Obstacle	63	0.04274792	0.1828	–	–	2.84×10^{-13}
Obstacle	127	0.04276762	0.0461	–	–	1.14×10^{-12}
Obstacle	255	0.04277249	0.0114	–	–	4.55×10^{-12}
Bratu	15	0.07787405	–	–	–	5.94×10^{-14}
Bratu	31	0.07804406	0.2178	–	–	2.53×10^{-13}
Bratu	63	0.07808677	0.0547	–	–	9.62×10^{-13}
Bratu	127	0.07809746	0.0137	–	–	4.50×10^{-12}
Bratu	255	0.07810013	0.0034	–	–	2.00×10^{-11}
Reynolds	23×7	19.46179436	–	5.94623677	–	1.05×10^{-13}
Reynolds	47×15	19.35072707	0.5740	6.16126067	3.4899	1.87×10^{-13}
Reynolds	95×31	19.42318664	0.3731	6.22739540	1.0620	1.38×10^{-12}
Reynolds	191×63	19.41654846	0.0342	6.24301075	0.2501	6.93×10^{-12}
Reynolds	383×127	19.41809927	0.0080	6.24705064	0.0647	2.26×10^{-11}

Note: $Q_h = V_h$ for the obstacle membrane, $Q_h = d_h$ for Bratu, and $Q_h = P_h$ for Reynolds. Square-grid entries denote $N \times N$ interior nodes. All quantities are dimensionless. Relative changes use the finer-grid value as denominator. These are

post-solve equilibrium checks, not additional successful NBB runs.

Let $I_h U_{h/2}$ be the restriction of the finer field to coincident coarse nodes. In addition to the midline profiles in Figure 7a–c, we compute

$$E_h = \left(\Delta x \Delta y \sum_{i,j} |U_h - I_h U_{h/2}|^2 \right)^{1/2}$$

and

$$E_{\infty,h} = \max_{i,j} |U_h - I_h U_{h/2}|.$$

These are self-convergence diagnostics, not errors against an exact continuum solution. The observed orders in Table 17 are approximately two for the obstacle and Bratu fields, while the Reynolds orders vary from 1.27 to 2.22. The two difference norms decrease on every tested refinement. Figure 7d–f reports these values; the h^2 guides for the two membrane models are not fitted bounds. We do not infer a mesh-independent iteration count, a continuum convergence theorem, or a calibrated physical-model error from these finite refinements.

Table 17. Successive-grid differences of the residual-checked physical fields.

Model	N	E_h	$E_{\infty,h}$	Order
Obstacle	15	5.55×10^{-5}	1.24×10^{-4}	2.08
Obstacle	31	1.31×10^{-5}	2.39×10^{-5}	1.93
Obstacle	63	3.42×10^{-6}	7.34×10^{-6}	2.13
Obstacle	127	7.83×10^{-7}	1.91×10^{-6}	–
Bratu	15	1.01×10^{-4}	1.70×10^{-4}	1.99
Bratu	31	2.53×10^{-5}	4.27×10^{-5}	2.00
Bratu	63	6.34×10^{-6}	1.07×10^{-5}	2.00
Bratu	127	1.59×10^{-6}	2.67×10^{-6}	–
Reynolds	7	9.09×10^{-2}	2.63×10^{-1}	1.27
Reynolds	15	3.76×10^{-2}	1.12×10^{-1}	2.22
Reynolds	31	8.05×10^{-3}	2.50×10^{-2}	1.87
Reynolds	63	2.21×10^{-3}	8.08×10^{-3}	–

Note: N is the coarse-grid interior count in each square-domain direction, or the axial count for Reynolds. Order is $\log_2(E_h/E_{h/2})$. Neither the differences nor the orders use an exact continuum solution.

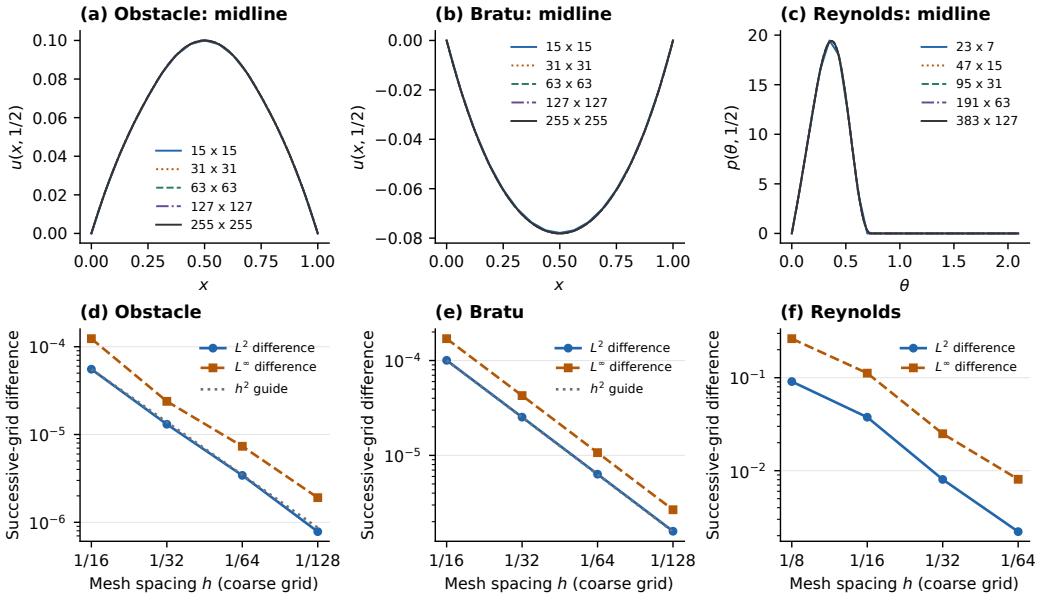


Figure 7. Field-level mesh-convergence checks. Top row: Displacement or pressure along the physical midline $y = 1/2$, with zero boundary values included; legend entries are interior grid dimensions. Bottom row: Discrete L^2 and infinity-norm differences between adjacent residual-checked grids, evaluated at coincident nodes. The bottom-row horizontal axes run toward finer meshes. The guides in (d) and (e) indicate an h^2 slope only. No exact continuum field is assumed.

For an independent discrete contact-set check, the obstacle problem is also solved as the strictly convex bound-constrained quadratic program

$$\min_{q \geq 0} \ \tfrac{1}{2} q^\top L_N q + (L_N \psi - \ell)^\top q$$

using MATLAB's `quadprog` interior-point-convex algorithm [38]. No NBB solution is supplied as its starting point.

At

$$N = 3,163,127,$$

the infinity-norm differences between the two gap vectors are

$$4.03 \times 10^{-8}, \quad 4.13 \times 10^{-6}, \quad \text{and} \quad 2.05 \times 10^{-7},$$

respectively; their contact classifications at threshold 10^{-7} disagree at $0, 8, 0$ nodes, respectively. The corresponding reference-solver KKT residuals are

$$1.85 \times 10^{-9}, \quad 2.15 \times 10^{-7}, \quad \text{and} \quad 1.24 \times 10^{-8}.$$

In particular, the $N = 63$ reference does not meet our 10^{-7} common threshold and its eight differing nodes should not be interpreted as exact classification errors of NBB. Figure 8 shows the agreement at $N = 127$. This is a cross-solver validation of a discrete contact set, not a comparison with an analytic free boundary.

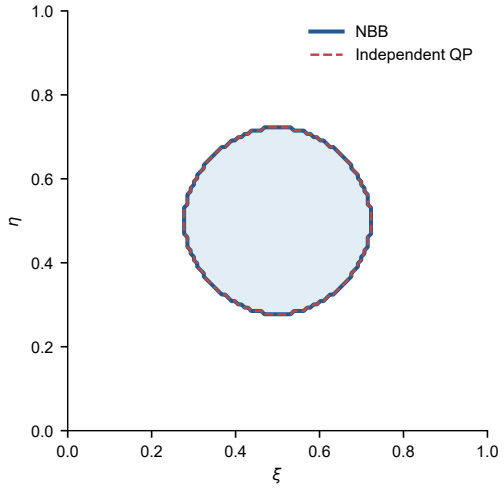


Figure 8. Auxiliary cross-solver check: Active-contact contours obtained with NBB and an independently solved quadratic program at $N = 127$, in physical coordinates. The two node classifications agree at the 10^{-7} gap threshold. These are discrete numerical boundaries, not the analytical-solution comparison, which is provided separately in Section 5.3.5.

5.3.5. Analytical solution and contact-boundary validation

The preceding cross-solver check compares two numerical solutions of the same discrete problem; it does not measure error against an analytical solution. To provide that distinct validation, we add Benchmark I from Harasim and Valdman [39], with the published contact radius $R = 0.7$. This is a separate obstacle benchmark, not an analytical solution claimed for the sinusoidal-obstacle, Bratu, or Reynolds models above.

On $\Omega = (-1, 1)^2$, set the obstacle to zero, write $r^2 = \xi^2 + \eta^2$, and prescribe

$$u^*(\xi, \eta) = \left(\max\{r^2 - R^2, 0\}\right)^2, \qquad u|_{\partial\Omega} = u^*|_{\partial\Omega},$$
$$f(\xi, \eta) = \begin{cases} -16r^2 + 8R^2, & r > R, \\ -8(R^4 + R^2) + 8R^2 r^2, & r \le R. \end{cases}$$

The continuum conditions hold almost everywhere: $u \ge 0$, $s = -\Delta u - f \ge 0$, and $us = 0$. Outside the contact disk, differentiating u^* gives $-\Delta u^* = f$. Inside, $u^* = 0$ and

$$s^* = 8(R^4 + R^2) - 8R^2 r^2 \ge 8R^2 > 0.$$

The displacement and its first derivatives match at $r = R$, so there is no interface contribution to the weak Laplacian. Thus, the analytical contact set is $r \le R$, and its free boundary is the circle $r = R$.

We discretize $-\Delta$ with the same five-point scheme, now with $h = 2/(N + 1)$ and the nonzero Dirichlet trace included on the right-hand side. The load is evaluated from its continuum formula above, not formed by applying the discrete operator to u^*; consequently, the sampled analytical field is not imposed as a discrete solution. NBB uses the unchanged standard settings and $x^0 = s^0 = e$, where x is the interior displacement since the obstacle is zero. All five grids $N = 15, 31, 63, 127, 255$ meet both stopping tests, without post-solve correction.

For the raw numerical displacement u_h, we report the actual solution errors

$$E^*_{\infty,h} = \max_{i,j} |u_{h,ij} - u^*_{ij}|, \qquad E^*_{2,h} = \left(h^2 \sum_{i,j} |u_{h,ij} - u^*_{ij}|^2 \right)^{1/2}.$$

To evaluate the contact boundary, classify nodes by

$$u_{h,ij} \le 10^{-7}$$

and let Γ_h be the piecewise-linear 0.5 contour of this binary nodal indicator. Define

$$D_{\Gamma,h} = \max_{z \in \Gamma_h} \big| \|z\|_2 - R \big|.$$

This measures the maximum radial deviation of the polygonal boundary from the exact circle; it is not an error in a fitted radius or an unqualified symmetric Hausdorff distance. Segment interiors as well as endpoints enter the calculation. Changing the classification threshold to 10^{-8} or 10^{-6} leaves every contact mask unchanged in these five tests.

Table 18. NBB errors relative to the analytical obstacle solution and its contact circle.

N	IT	R_{KKT}	$E^*_{\infty,h}$	$E^*_{2,h}$	$D_{\Gamma,h}$	$D_{\Gamma,h}/h$
15	8	5.68×10^{-14}	3.97×10^{-3}	1.89×10^{-3}	8.31×10^{-2}	0.665
31	9	2.84×10^{-13}	4.55×10^{-4}	3.03×10^{-4}	4.28×10^{-2}	0.685
63	15	2.36×10^{-12}	2.14×10^{-4}	8.03×10^{-5}	2.03×10^{-2}	0.649
127	25	6.37×10^{-12}	7.61×10^{-5}	2.19×10^{-5}	1.37×10^{-2}	0.875
255	51	2.91×10^{-11}	2.43×10^{-5}	5.78×10^{-6}	7.42×10^{-3}	0.950

Note: $h = 2/(N+1)$ and $R = 0.7$. All five runs satisfy both stopping thresholds; no post-solve correction is applied. $E^*_{\infty,h}$ and $E^*_{2,h}$ compare the computed displacement with the continuous analytical solution sampled at grid nodes. $D_{\Gamma,h}$ is the maximum radial deviation of the reconstructed numerical contact boundary from the exact circle, not a residual or a fitted radius error.

Both displacement error norms decrease on every refinement. At

$$N = 255, \quad E^*_{\infty,h} = 2.43 \times 10^{-5}, \quad \text{and} \quad E^*_{2,h} = 5.78 \times 10^{-6},$$

whereas the common algebraic residual is 2.91×10^{-11}. The different magnitudes illustrate why a small discrete residual is not a substitute for a continuum-solution error. The radial boundary deviation decreases from 8.31×10^{-2} to 7.42×10^{-3} and is less than one grid spacing on every tested mesh. Figure 9 displays the analytical/numerical displacement overlay, the pointwise error, and the comparison with the known circle. These are empirical accuracy checks for this benchmark, not a universal convergence-order theorem or validation against physical measurements.

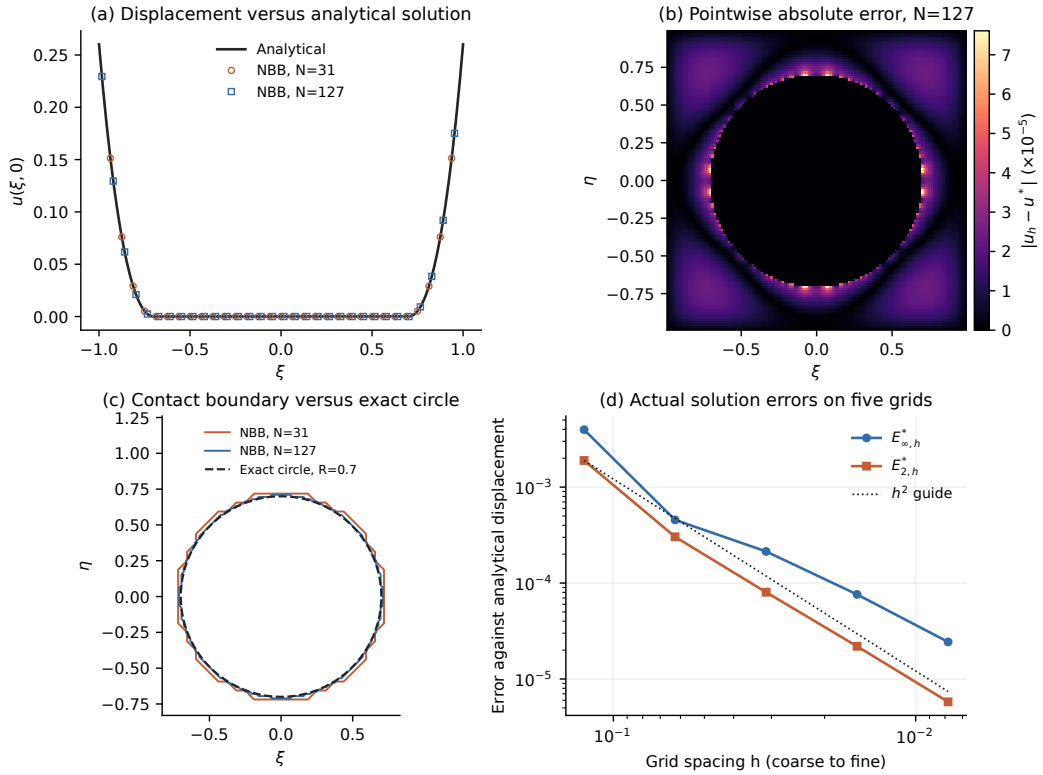


Figure 9. Comparison with the analytical obstacle solution of Harasim and Valdman [39]. (a) Analytical midline displacement and computed nodal values; markers on the $N = 127$ grid are subsampled for visibility. (b) Actual absolute nodal error at $N = 127$. (c) Numerical contact contours and the exact circle $r = 0.7$, using physical domain coordinates. (d) Displacement errors relative to the analytical field on five grids. The h^2 line is a slope guide anchored at the coarsest $E^*_{2,h}$, not a fitted bound or a claim of uniform second-order behavior. All quantities are nondimensional.

5.4. Large sparse weighted problems

To separate a genuine nonzero-weight test from the zero-weight physical examples, let B_N be the positive-definite five-point Dirichlet Laplacian on an $N \times N$ grid, and define

$$F(x, s) = B_N x + c - s, \qquad c = e - B_N e, \qquad w = e.$$

The known solution is $x^* = s^* = e$. We use $x^0 = s^0 = 0.5e$ and the standard NBB settings. This monotone model satisfies Assumption 3.1 and Proposition 4.1. It is a structured scalability check with a known solution, not a difficult random nonlinear instance. The active-obstacle model is also run on the same grids for comparison of sparse problem sizes, without claiming that the two problems have equal difficulty.

For each model and grid, we perform one warm-up followed by five measured runs, always rebuilding the model and restarting from the specified initial point. The six cases are run serially, with their order reversed on alternate repetitions. The environment is MATLAB R2023b on the AMD Ryzen 5 6600H machine described in Section 5.1; MATLAB reports a maximum of six computational threads. We separately measure model construction, the complete solver call, and their combined

elapsed time using `tic/toc`. The complete call includes initialization, iterations, and final residual and error reporting. MATLAB startup, file I/O, and independent post-run validation are outside these intervals. Process CPU time is separately measured with `cputime`; it accumulates processor work across threads and is not interchangeable with elapsed time. Table 19 reports solver-time medians and ranges. Construction, combined, and process CPU times are medians of their own samples, so their rounded entries need not add exactly.

Table 19. Sparse problem sizes, work, and repeated wall-clock timings.

Model	n	nnz(B)	IT	NF	BTR	Runs	R_{KKT}
Obstacle	10000	49600	23	49	2	5/5	1.14×10^{-12}
$w = e$	10000	49600	4	9	0	5/5	1.64×10^{-11}
Obstacle	40000	199200	39	82	3	5/5	5.00×10^{-12}
$w = e$	40000	199200	4	9	0	5/5	8.00×10^{-11}
Obstacle	102400	510720	56	116	3	5/5	1.30×10^{-11}
$w = e$	102400	510720	5	12	1	5/5	1.31×10^{-10}

Model	n	Solve: median [min, max]	Build	Total	CPU
Obstacle	10000	0.664 [0.648, 0.685]	0.001	0.665	4.109
$w = e$	10000	0.109 [0.107, 0.114]	0.002	0.110	0.656
Obstacle	40000	5.146 [5.101, 5.198]	0.003	5.149	30.781
$w = e$	40000	0.478 [0.471, 0.481]	0.005	0.483	2.891
Obstacle	102400	22.220 [22.103, 22.898]	0.008	22.227	133.203
$w = e$	102400	1.824 [1.785, 1.825]	0.015	1.841	11.000

Note: All times are in seconds, after one warm-up per case. Solve, Build, and Total are elapsed wall-clock times; CPU is separately measured process CPU time. Build, Total, and CPU are medians of five runs. Solve times cover the complete solver call; Total covers model construction and that call, excluding startup, file I/O, and independent validation. Each run starts afresh. IT, NF, and BTR are identical across repeats; one linear system is solved per iteration. R_{KKT} is the maximum over the five repeats. The full smoothing vector has $1 + 2n$ entries.

All thirty measured runs meet both stopping thresholds, and their iteration, residual-evaluation, linear-solve, and rejected-trial counts reproduce the original records. At $N = 320$, there are $n = 102400$ complementarity pairs and 204801 full smoothing variables, but only a reduced $n \times n$ sparse matrix is assembled and solved with MATLAB backslash. The full smoothing Jacobian is not formed. The Laplacian has $5N^2 - 4N = 510720$ nonzeros at this size, and diagonal scaling in the reduced system preserves this pattern. Thus, these tests do not involve a dense matrix with n^2 stored entries. Sparse factorization still introduces fill-in and additional storage; no linear memory or time bound is inferred from the input nonzero count.

The positive-weight problem takes five Newton solves at the largest size, with a median elapsed solver time of 1.824 s (range 1.785–1.825 s), whereas the obstacle problem takes 56 solves and 22.220 s (22.103–22.898 s). Their median process CPU times are 11.000 and 133.203 s, respectively. The short elapsed time of the positive-weight case is consistent with its sparse structure and small iteration count; the known solution is used only for error checking, not to initialize at the root or replace a Newton solve. Independent reconstruction of both model equations validates all six final fields saved

from the last repetition. The reduced directions also satisfy the full Newton equations, with a maximum normalized backward error of 4.64×10^{-16} in the six initial-state checks; small-grid full and reduced solves agree to relative tolerance 10^{-10}.

These measurements establish feasibility on the specified structured problems and hardware, not a general runtime guarantee for large WCPs. The increasing obstacle iteration counts do not support mesh-independent iteration complexity. The positive-weight case is deliberately simple and complements, rather than replaces, the nonconvex stress tests.

5.5. Summary of numerical findings

The paired comparisons and fixed-initializer ablation support a specific benefit: reduced rejected-trial work on the tested nonconvex problems. They do not establish a systematic reduction in outer Newton iterations or a general timing or success-rate advantage. The separate effects of accepted-step memory and clipping are not statistically established. The parameter tests show that larger forcing coefficients are usable on the tested instances, but neither parameter insensitivity nor exclusive usability of $\tau = 0$ follows. The nonconvex and signed-start Bratu failures demonstrate dependence on the problem and initialization; near-solution diagnostics do not replace these failed independent-start runs.

The physical tests provide complementary evidence: Successive-grid differences assess spatial stabilization, the independent quadratic-programming solve checks a discrete contact set, and the analytical obstacle benchmark measures error against a known continuum solution. These checks do not establish universal continuum error bounds or replace comparisons with specialized physical solvers. The finest Bratu run also shows sensitivity of the fixed aggregate stopping tolerance to problem size, and the Reynolds test is limited to Swift–Stieber cavitation. The random nonconvex tests lie outside the established global theory, while the sparse positive-weight tests assess feasibility on a deliberately simple structured class. Broader comparisons remain necessary.

6. Conclusions and outlook

We have developed a safeguarded nonmonotone NBB method in which BB information initializes backtracking only after rejection of the full Newton step. Under the structural rank–monotonicity assumption and boundedness of the initial merit-function level set, the method is well defined and globally convergent in the precise sense that its smoothing residual tends to zero and every accumulation point solves the WCP. A unique zero in that level set ensures convergence of the whole sequence. The estimate $C_k \leq (C_0 + 1)M(z^k)$ makes the forcing term second order; generalized-Jacobian regularity and local Lipschitz continuity of F' then yield eventual full steps and quadratic convergence. For symmetric positive-definite linear models, the level-set and local regularity conditions are verified, giving convergence to the unique solution without strict complementarity.

The numerical evidence supports the role of the BB factor as a post-rejection line-search initializer, rather than a modification of the Newton direction. Its benefit is problem dependent: NBB reduces rejected-trial work relative to Rn-NSNA on the paired nonconvex tests, whereas the independent one-step Rn-SNEB comparator fails on that stress suite but succeeds on all 18 monotone linear instances, with fewer mean rejected trials than NBB. No general timing or success-rate advantage is claimed.

Sparse positive-weight problems up to $n = 102400$ and discrete PDE checks extend the range of

tested models. The nonconvex parameter failures and persistent signed-start Bratu failures also identify limits of the present globalization strategy. In future work, researchers should examine adaptive selection of τ, alternative reference-value updates, and inexact sparse solves. For physical applications, comparisons with specialized solvers and mass-conserving cavitation models remain necessary beyond the discrete and analytical-benchmark validation provided here.

Author contributions

Teng Wu: conceptualization, methodology, software, formal analysis, investigation, data curation, visualization, and writing–original draft. Xiangjing Liu: conceptualization, methodology, validation, supervision, project administration, and writing–review and editing. All authors have read and agreed to the published version of the manuscript.

Use of generative artificial intelligence (AI) tools declaration

OpenAI's ChatGPT and Codex assisted with English-language editing, LaTeX formatting, literature organization, preliminary mathematical exposition, experimental-code drafting and checking, numerical-analysis scripts, figure preparation, and drafting of revision responses. The authors are responsible for independently checking the mathematical arguments, computations, references, interpretations, and final submitted content; generative-AI output is not a substitute for that verification.

Acknowledgments

This research was supported by the Scientific Research Program Funded by Education Department of Shaanxi Provincial Government (Program No. 23JK0487) and Natural Science Basic Research Program of Shaanxi (Program No. 2026JC-YBQN-0070).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. F. A. Potra, Weighted complementarity problems–a new paradigm for computing equilibria, *SIAM J. Optim.*, **22** (2012), 1634–1654. https://doi.org/10.1137/110837310
2. S. Asadi, Z. Darvay, G. Lesaja, N. Mahdavi-Amiri, F. A. Potra, A full-Newton step interior-point method for monotone weighted linear complementarity problems, *J. Optim. Theory Appl.*, **186** (2020), 864–878. https://doi.org/10.1007/s10957-020-01728-4
3. X. N. Chi, Z. P. Wan, Z. J. Hao, A full-modified-Newton step $O(n)$ infeasible interior-point method for the special weighted linear complementarity problem, *J. Ind. Manag. Optim.*, **18** (2022), 2579–2598. https://doi.org/10.3934/jimo.2021082

4. X. N. Chi, G. Q. Wang, A full-Newton step infeasible interior-point method for the special weighted linear complementarity problem, *J. Optim. Theory Appl.*, **190** (2021), 108–129. https://doi.org/10.1007/s10957-021-01873-4
5. L. Qi, D. Sun, G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, *Math. Program.*, **87** (2000), 1–35. https://doi.org/10.1007/s101079900127
6. J. Zhang, A smoothing Newton algorithm for weighted linear complementarity problem, *Optim. Lett.*, **10** (2016), 499–509. https://doi.org/10.1007/s11590-015-0877-4
7. J. Y. Tang, A variant nonmonotone smoothing algorithm with improved numerical results for large-scale LWCPs, *Comput. Appl. Math.*, **37** (2018), 3927–3936. https://doi.org/10.1007/s40314-017-0554-6
8. J. Tang, L. Dong, J. Zhou, L. Fang, A smoothing Newton method for nonlinear complementarity problems, *Comput. Appl. Math.*, **32** (2013), 107–118. https://doi.org/10.1007/s40314-013-0015-9
9. J. Tang, H. Zhang, A nonmonotone smoothing Newton algorithm for weighted complementarity problem, *J. Optim. Theory Appl.*, **189** (2021), 679–715. https://doi.org/10.1007/s10957-021-01839-6
10. J. Tang, J. Zhou, H. Zhang, An accelerated smoothing Newton method with cubic convergence for weighted complementarity problems, *J. Optim. Theory Appl.*, **196** (2023), 641–665. https://doi.org/10.1007/s10957-022-02152-6
11. J. Tang, J. Zhou, The solvability of weighted complementarity problems and a smoothing Newton algorithm under the local error bound, *Optimization*, **74** (2025), 811–841. https://doi.org/10.1080/02331934.2023.2269943
12. X. Liu, J. Zhang, J. Chen, A modified smoothing Newton method for solving weighted complementarity problems with a nonmonotone line search, *Eng. Lett.*, **31** (2023), 1684–1690.
13. J. Barzilai, J. M. Borwein, Two-point step size gradient methods, *IMA J. Numer. Anal.*, **8** (1988), 141–148. https://doi.org/10.1093/imanum/8.1.141
14. Y. H. Dai, A new analysis on the Barzilai–Borwein gradient method, *J. Oper. Res. Soc. China*, **1** (2013), 187–198. https://doi.org/10.1007/s40305-013-0007-x
15. H. He, D. Han, Z. Li, Some projection methods with the BB step sizes for variational inequalities, *J. Comput. Appl. Math.*, **236** (2012), 2590–2604. https://doi.org/10.1016/j.cam.2011.12.017
16. Y. Zheng, B. Zheng, A new modified Barzilai–Borwein gradient method for the quadratic minimization problem, *J. Optim. Theory Appl.*, **172** (2017), 179–186. https://doi.org/10.1007/s10957-016-1008-9
17. X. Li, Smoothing nonmonotone Barzilai–Borwein gradient method and its application to stochastic linear complementarity problems, *Math. Probl. Eng.*, **2015** (2015), 425351. https://doi.org/10.1155/2015/425351
18. K. Ma, J. Zeng, J. Xiong, Q. Xu, X. Cao, W. Liu, et al., Stochastic non-convex ordinal embedding with stabilized Barzilai–Borwein step size, *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, 2018, 3738–3745. https://doi.org/10.1609/aaai.v32i1.11599

19. J. Chen, L. Tang, X. Yang, A Barzilai–Borwein descent method for multiobjective optimization problems, *Eur. J. Oper. Res.*, **311** (2023), 196–209. https://doi.org/10.1016/j.ejor.2023.04.022
20. X. Xu, A parameterized Barzilai–Borwein method via interpolated least squares, *J. Ind. Manag. Optim.*, **21** (2025), 6246–6269. https://doi.org/10.3934/jimo.2025130
21. L. Qi, J. Sun, A nonsmooth version of Newton's method, *Math. Program.*, **58** (1993), 353–367. https://doi.org/10.1007/BF01581275
22. L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Math. Oper. Res.*, **18** (1993), 227–244. https://doi.org/10.1287/moor.18.1.227
23. W. J. Conover, *Practical Nonparametric Statistics*, 3 Eds., Wiley, 1999.
24. R. L. Wasserstein, N. A. Lazar, The ASA statement on *p*-values: context, process, and purpose, *Am. Stat.*, **70** (2016), 129–133. https://doi.org/10.1080/00031305.2016.1154108
25. S. Holm, A simple sequentially rejective multiple test procedure, *Scand. J. Stat.*, **6** (1979), 65–70.
26. G. Duvaut, J. L. Lions, *Inequalities in mechanics and physics*, Springer, 1976. https://doi.org/10.1007/978-3-642-66165-5
27. J. F. Rodrigues, *Obstacle problems in mathematical physics*, North-Holland Mathematics Studies, 1987. https://doi.org/10.1016/S0304-0208(08)X7010-0
28. R. Glowinski, J. L. Lions, R. Trémolières, *Numerical analysis of variational inequalities*, North-Holland, 1981.
29. G. Bratu, Sur les équations intégrales non linéaires, *Bull. Soc. Math. Fr.*, **42** (1914), 113–142. https://doi.org/10.24033/bsmf.943
30. D. D. Joseph, T. S. Lundgren, Quasilinear Dirichlet problems driven by positive sources, *Arch. Ration. Mech. Anal.*, **49** (1973), 241–269. https://doi.org/10.1007/BF00250508
31. H. W. Swift, The stability of lubricating films in journal bearings, *Minutes Proc. Inst. Civ. Eng.*, **233** (1932), 267–288. https://doi.org/10.1680/imotp.1932.13239
32. W. Stieber, *Das schwimmlager: hydrodynamische theorie des gleitlagers*, VDI-Verlag, 1933.
33. G. Bayada, M. Chambat, The transition between the Stokes equations and the Reynolds equation: a mathematical proof, *Appl. Math. Optim.*, **14** (1986), 73–93. https://doi.org/10.1007/BF01442229
34. M. Giacopini, M. T. Fowell, D. Dini, A. Strozzi, A mass-conserving complementarity formulation to study lubricant films in the presence of cavitation, *J. Tribol.*, **132** (2010), 041702. https://doi.org/10.1115/1.4002215
35. P. J. Roache, Perspective: a method for uniform reporting of grid refinement studies, *J. Fluids Eng.*, **116** (1994), 405–413. https://doi.org/10.1115/1.2910291
36. M. Muchammad, D. D. P. Wibowo, M. Tauviqirrahman, B. Setiyana, E. Yohana, J. Jamari, Investigation of the influence of turbulence to tribological performance on smooth and multistep journal bearing with hydrodynamics simulation, *Front. Mech. Eng.*, **8** (2022), 946074. https://doi.org/10.3389/fmech.2022.946074

37. M. Muchammad, M. Tauviqirrahman, Y. M. Rizki, I. Syafaat, F. Maharani, M. I. Ammarullah, Elastohydrodynamic analysis of multistep texture effects and partial surface roughness on the tribological performance of steel journal bearings, *Discov. Appl. Sci.*, **7** (2025), 301. https://doi.org/10.1007/s42452-025-06789-6

38. The MathWorks, Inc., *quadprog: quadratic programming*, Optimization toolbox documentation. Available from: https://www.mathworks.com/help/optim/ug/quadprog.html.

39. P. Harasim, J. Valdman, Verification of functional a posteriori error estimates for obstacle problem in 2D, *Kybernetika*, **50** (2014), 978–1002. https://doi.org/10.14736/kyb-2014-6-0978



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)