Strongly Convergent Inertial Reflected-Forward-backward Algorithm for Solving Monotone Inclusion Problems

Chibueze Christian Okeke

School of Mathematics, University of the Witwatersrand, South Africa.

Presentation @ Mathematical Optimization and Control for Industrial and Societal Challenges in Sub-Saharan Africa.

September 7, 2026

Introduction

Overview

Optimization is one of the fundamental mathematical frameworks for modelling and solving decision-making and computational problems in which one seeks to determine an optimal element from a prescribed feasible set. In its most general form, an optimization problem can be written as

$$\min_{x \in C}; \phi(x),$$

where (C) is a nonempty subset of an appropriate space and (ϕ) is an objective function whose minimizers represent the desired solutions

Introduction

Introduction

Optimization problems arise naturally in a wide range of areas, including engineering design, economics, finance, signal and image processing, machine learning, data science, optimal control, inverse problems and medical imaging. In many of these applications, the objective function may contain several components with different mathematical structures. For example, one component may be smooth and differentiable, whereas another may be nonsmooth but possess a structure that can be efficiently exploited through a proximal operator.



Common Optimization Problems

Some important optimization problems in Hilbert spaces include:

1. Convex optimization;
2. Variational inequalities;
3. Monotone inclusion problems;
4. Fixed-point problems;
5. Split feasibility and variational problems;
6. Constrained optimization;
7. Nonsmooth optimization.

Description of Problems

Monotone Inclusion Problem

We investigate the monotone inclusion problem, which involves searching for a point z in the solution set $(G+\gamma)^{-1}(\mathbf{0})$ such that

$$\mathbf{0} \in (G+\gamma)z, \tag{1}$$

where $G : \mathcal{H} \to 2^{\mathcal{H}}$ and $\gamma : \mathcal{H} \to \mathcal{H}$ are maximal monotone operator and monotone and L-Lipschitz continuous operator, respectively; and $\mathcal{H}$ is a real Hilbert



Description of Problems

Description of Problems

The broad applicability of this framework has generated considerable interest in developing efficient iterative algorithms for approximating solutions of monotone inclusion problems. The principal challenge is to construct algorithms that simultaneously have strong convergence guarantees, low computational cost, weak assumptions on the operators, and good numerical performance for large-scale applications.

Proximal point algorithm

Proximal point algorithm

One of the foundational approaches for solving monotone inclusion problems is the proximal point algorithm (PPA) introduced by Rockafellar. For the inclusion

$$0 \in Tz, \qquad T = G + \gamma,$$

the proximal point iteration is

$$z_{n+1} = J_{\delta_n T} z_n,$$

where

$$J_{\delta_n T} = (I + \delta_n T)^{-1}$$

is the resolvent of T, and $(\delta_n > 0)$ is a regularization parameter.

Advantages and Drawback of PPA

Advantages and Drawback of PPA

The major advantage of PPA is its robust convergence theory for maximal monotone operators. It provides a conceptually simple way of transforming the problem of finding a zero of T into a sequence of regularized subproblems. However, its principal drawback is computational: evaluating the resolvent of the entire sum $(G + \gamma)$ may be difficult or even computationally prohibitive. In large-scale problems, solving the implicit subproblem at every iteration can therefore make PPA expensive.

Foward-Backward-Splitting

A natural splitting strategy is the forward-backward splitting method (FBSM). Instead of evaluating

$$J_{\delta(G+\gamma)},$$

the method evaluates the two operators separately according to

$$J_{\delta G}\left(z_n - \delta\gamma z_n\right).$$

Here,

$$J_{\delta G} = (I + \delta G)^{-1}$$

is the resolvent of G. Thus, each iteration consists of:

Advantages of FBSM

The principal advantages of FBSM are:

- it separates the two operators rather than treating $(G + \gamma)$ as one complicated operator;
- it requires only one forward evaluation and one backward evaluation per iteration;
- the backward step can often be computed efficiently when the resolvent of (G) has a simple form;
- it provides a unified framework for many proximal-gradient-type optimization algorithms.

Limitations of FBSM:
Despite its attractive structure, classical FBSM generally requires a stronger assumption on (γ), typically cocoercivity.

Forward-backward-forward splitting

To overcome the restrictive cocoercivity requirement, Tseng introduced the forward-backward-forward splitting method (FBF). The method takes the form

$$J_{\delta G}(z_n - \delta\gamma z_n),$$

followed by

$$\widehat{z}_n - \delta\left(\gamma\widehat{z}_n - \gamma z_n\right).$$

Unlike classical forward-backward splitting, Tseng's method can operate under the weaker assumption that γ is monotone and Lipschitz continuous. This is a significant theoretical advantage because it removes the need for cocoercivity.

Advantages of FBF

The principal advantage is therefore its broader applicability:

- it accommodates monotone and Lipschitz continuous operators;
- it avoids the restrictive cocoercivity assumption;
- it retains a splitting structure;
- it has an established weak convergence theory under appropriate assumptions.

Limitation of FBF:
The major computational disadvantage is that the forward operator γ is evaluated twice per iteration.

Forward-reflected-backward splitting

A significant development in this direction is the forward-reflected-backward splitting method (FRBS) proposed by Malitsky and Tam. Its iteration is

$$J_{\delta_n G}\left(z_n - \delta_n \gamma z_n - \delta_{n-1}(\gamma z_n - \gamma z_{n-1})\right).$$

The key feature is the use of information from the previous iteration to approximate the additional forward evaluation that appears in Tseng's method. Consequently, the method requires only one current evaluation of the forward operator per iteration.

Advantages of FRBS

FRBS has several important advantages:

- it requires only one forward evaluation of γ per iteration;
- it works under monotonicity and Lipschitz continuity rather than cocoercivity;
- it retains a simple forward-backward splitting structure;
- it can recover or relate to several well-known methods, including the projected reflected gradient method, proximal point algorithm and Popov's algorithm.

Limitation of FRBS:
The main limitation, in the context of the present work, is that the established convergence result is weak convergence.



Reflected-forward-backward splitting

Another important development is the reflected-forward-backward (RFB) splitting method. When the forward operator is affine, the structure of the forward-reflected-backward method coincides with the reflected-forward-backward method

$$\widehat{z}_n = 2z_n - z_{n-1},$$

followed by

$$J_{\delta G}\left(z_n - \delta\gamma\widehat{z}_n\right).$$

The manuscript identifies this as the reflected-forward-backward method considered in the work of Cevher and Vũ.

Advantages of RFB

The reflected approach has several attractive features:

- only one forward evaluation of the single-valued operator is required;
- only one backward/resolvent evaluation is required;
- it can be applied under monotonicity and Lipschitz continuity;
- its computational structure is simple and suitable for large-scale problems.

Limitation of RFB:
Despite these advantages, the convergence obtained by the existing reflected-forward-backward schemes considered in the manuscript is not the strong convergence sought in the present work.

Inertial Term

Polyak introduced the heavy-ball dynamical system

$$\ddot{x}(t) + \gamma\dot{x}(t) + \nabla f(x(t)) = 0, \qquad \gamma > 0.$$

A discretization leads to

$$x_{n+1} = x_n + \beta(x_n - x_{n-1}) - \alpha\nabla f(x_n).$$

The term

$$x_n - x_{n-1}$$

is called the **momentum term**.



Inertial Extrapolation

The inertial method can be written as

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1}), \\ x_{n+1} = y_n - \alpha\nabla f(x_n). \end{cases}$$

Here,

$$\beta_n(x_n - x_{n-1})$$

is the inertial or momentum component.
The parameter β_n controls the amount of extrapolation.

Our Contributions

Strong convergence

Strong convergence is generally a more demanding property and is often more desirable when the actual distance between the computed iterate and the target solution is important.

Our Contribution

We introduce and study a strongly convergent one-step inertial-type reflected-forward-backward method. The introduced method incorporates a novel inertial technique that enhances the convergence property of the recursive sequence produced by the introduced method. Notably, the method requires only one forward evaluation of the single-valued operator and one backward evaluation of the set-valued operator per iteration, a feature not present in related methods.:

Our Contributions

Algorithm

Strongly Convergent Inertial RFB Splitting Algorithm

For arbitrary points $z_{-1}, z_0 \in \mathcal{H}$, let the sequence $\{z_t\}$ be iteratively produced by

$$\begin{aligned} e_t &= (1-\sigma_t)(z_t + \vartheta_t(z_t - z_{t-1})), \\ \hat{z}_t &= 2z_t - z_{t-1}, \\ z_{t+1} &= \mathcal{J}_{\lambda G}(e_t - \lambda\gamma\hat{z}_t). \end{aligned}$$

Thus, the proposed method combines inertial extrapolation and reflection while maintaining one backward evaluation of G and one forward evaluation of γ at each iteration.



Some Remarks

If $\vartheta_t = \sigma_t = 0$, the proposed algorithm recovers the reflected-forward-backward method studied in Cevher. Also, if G is a normal cone of C ($G = N_C$), that is, searching for a point $z \in C : \mathbf{0} \in \gamma z + N_C(z)$, and $\vartheta_t = \sigma_t = 0$, then our method recovers the projected reflected gradient method Malitsky. In contrast to these methods, the proposed algorithm produces recursive sequence that converges faster and strongly to the unique solution of our problem.

Our Contributions

Remark

The principal theoretical contribution is the establishment of strong convergence of the generated sequence to the selected solution of the monotone inclusion problem under the stated assumptions. This is important because the proposed method combines strong convergence with the computational advantage of requiring only one evaluation of each component operator per iteration.



THANK YOU

References



R. D. Beer, A dynamical systems perspective on agent environment interaction, *Artificial Intelligence* **72** (1995) 173-215.



P. L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.,* **16**(6) (1979) 964–979.



B. Martinet, Brève communication. Régularisation d'inéquations variationnelles par approximations successives *Rev Fr Inform Rech Opér.,***4** (1970) 154–158. https://doi.org/10.1051/m2an/197004R301541



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