Commun. Math. Res.
doi: 10.4208/cmr.2026-0004

Vol. **x**, No. x, pp. 1-15
x 2026

Transfer of Improvement Strategies Between DRS and ADMM: A Unified Classification Framework

Shuting Liu*
School of Mathematics, Jilin University, Changchun 130012, P.R. China.

Received 5 January 2026; Accepted 16 March 2026

Abstract. Douglas-Rachford splitting (DRS) and the alternating direction method of multipliers (ADMM) are two fundamental first-order methods for structured convex optimization. Although derived from different viewpoints, ADMM can be interpreted as the application of DRS to the dual problem. Based on this structural equivalence, this paper studies how algorithmic improvement strategies can be transferred between the two methods. We classify transferable strategies into three categories: exact operator-level transfer, parameter-driven transfer, and heuristic transfer. Representative examples including relaxation, metric scaling, adaptive parameter updates, and residual balancing are discussed to illustrate the different levels of transferability. This perspective provides a systematic way to understand the relationship between DRS and ADMM and clarifies how algorithmic ideas developed for one method may inform the design of variants of the other, offering a unified framework that both explains existing variants and guides the design of new ones.

AMS subject classifications: 65D18, 65K10

Key words: Douglas-Rachford splitting, alternating direction method of multipliers, monotone operator splitting, convex optimization, algorithmic equivalence, relaxation techniques, adaptive step sizes.

1 Introduction

In large-scale convex optimization, image processing, machine learning, and distributed computation, many practical problems can be formulated as composite

*Corresponding author. *Email address:* `lst23@mails.jlu.edu.cn` (S. Liu)

optimization models with separable structures, such as objective functions composed of several simple terms or variables coupled through linear constraints. For this class of problems, first-order operator splitting methods have become fundamental tools in numerical optimization and computational mathematics, due to their low per-iteration cost, ease of parallelization, and scalability to large problem sizes. From the viewpoint of monotone operator theory, these models can often be equivalently cast as finding a zero of the sum of two maximal monotone operators, which provides a unified framework for algorithm design and convergence analysis [8].

Among various splitting methods, the Douglas-Rachford splitting method and the alternating direction method of multipliers are two particularly influential algorithms. DRS originates from the early work of Douglas and Rachford on heat conduction problems, and was later formalized by Lions and Mercier as a general method for solving monotone inclusions [5, 8]. ADMM, on the other hand, is typically derived from the augmented Lagrangian framework and primal-dual decomposition, and is especially effective for separable convex optimization problems with linear equality constraints. Since the seminal survey by Boyd *et al.* [2], ADMM has become a cornerstone algorithm in distributed optimization and statistical learning. Although their derivations differ substantially, DRS and ADMM are often applied to problems with similar structures and exhibit closely related numerical behavior.

A fundamental connection between these two methods was established by Eckstein and Bertsekas. For the composite convex optimization problem

$$\min_x f(x)+g(Mx), \tag{1.1}$$

under suitable regularity assumptions, they showed that ADMM can be interpreted as applying Douglas-Rachford splitting in the dual space to the maximal monotone operators induced by f^* and g^* [6]. This equivalence embeds ADMM, originally derived from the augmented Lagrangian method, into the broader framework of monotone operator splitting and fixed-point iterations. As a consequence, it not only offers new insights into the convergence analysis of ADMM, but also provides a systematic mechanism for transferring algorithmic modifications and parameter strategies between the two methods.

Recent research has moved beyond establishing the formal equivalence in the convex setting, and increasingly focuses on exploiting this connection to unify parameter selection strategies and extend convergence theory. One active direction concerns relaxation and adaptive parameter schemes. It has been shown that relaxed DRS iterations correspond naturally to over-relaxed variants of ADMM, and in certain simplified settings, the relaxation parameters can even be chosen identically [6]. Moreover, non-stationary step-size strategies developed for DRS have been successfully transferred to ADMM, leading to adaptive variants with improved practical performance [10].

Another important direction lies in convergence analysis beyond the classical convex framework. Although originally designed for convex problems, DRS and ADMM have been widely applied to structured nonconvex models in practice. Recent studies show that, by introducing tools such as the Douglas-Rachford envelope, one can analyze the convergence behavior of both algorithms within a unified framework. In particular, the convergence of ADMM for certain nonconvex problems can be established via its primal equivalence to DRS, significantly simplifying classical arguments [11].

Despite these advances, existing results on the DRS-ADMM equivalence remain scattered across different algorithmic variants and analytical settings. The purpose of this paper is to provide a systematic and integrated perspective from the viewpoint of operator splitting theory. Taking the equivalence between DRS and ADMM as the main thread, we clarify how relaxation techniques and adaptive step-size strategies developed for DRS can be systematically transferred to ADMM, and analyze the resulting algorithms using unified convergence tools. Rather than proposing new algorithmic structures, this work aims to offer a coherent equivalence-based framework that serves as a solid theoretical reference for the design and analysis of efficient ADMM variants.

2 Preliminaries

2.1 Douglas-Rachford splitting

The Douglas-Rachford splitting method is a first-order algorithm for solving optimization problems with separable structures and constraints. It is designed to solve the following monotone inclusion problem:

$$0 \in Ax + Bx, \tag{2.1}$$

or equivalently, to find

$$x^\star \in \operatorname{zer}(A+B) := \{x \in \mathbb{R}^n : 0 \in Ax + Bx\}.$$

Here, $A, B : \mathbb{R}^n \to \mathbb{R}^n$ are maximal monotone operators. When $A = \partial F$ and $B = \partial G$, where F and G are closed, proper, and convex functions, (2.1) coincides with the first-order optimality condition of the convex optimization problem

$$\min_{x \in \mathbb{R}^n} F(x) + G(x).$$

Definition 2.1 (Maximal Monotone Operator). *Let $\mathbb{R}^n$ be an n-dimensional real vector space. A set-valued operator $A:\mathbb{R}^n \to \mathbb{R}^n$ is said to be monotone if for any $(x,u),(y,v) \in \operatorname{gra} A$,*

$$\langle x-y, u-v\rangle \geq 0,$$

where

$$\operatorname{gra} A := \big\{ (x,u) \in \mathbb{R}^n \times \mathbb{R}^n \,|\, u \in A(x) \big\}$$

denotes the graph of A. If there exists no monotone operator B such that $\operatorname{gra} A \subsetneq \operatorname{gra} B$, then A is called a maximal monotone operator [1].

Definition 2.2 (Resolvent [4]). *Let $A:\mathbb{R}^n \to \mathbb{R}^n$ be a maximal monotone operator and let $\lambda > 0$. The resolvent of A is defined by*

$$J_{\lambda A} := (I+\lambda A)^{-1}.$$

In finite-dimensional spaces, the resolvent $J_{\lambda A}$ is single-valued, everywhere defined, and nonexpansive.

When A is the subdifferential operator ∂f of a closed, proper, and convex function $f:\mathbb{R}^n \to (-\infty,+\infty]$, the resolvent coincides with the proximal operator, namely,

$$J_{\lambda\partial f} = \operatorname{prox}_{\lambda f}.$$

For completeness, we recall the classical Douglas-Rachford splitting method in Algorithm 1.

Algorithm 1 Douglas-Rachford Splitting

Initialization: Choose $\lambda > 0$ and $z^0 \in \mathbb{R}^n$, and set $k=0$.
Iteratively update u^k, v^k, z^{k+1} according to the following scheme:

$$\begin{aligned} u^k &= (I+\lambda B)^{-1}(z^k), \\ v^k &= (I+\lambda A)^{-1}\big(2u^k - z^k\big), \\ z^{k+1} &= z^k + v^k - u^k. \end{aligned}$$

Terminate if a prescribed stopping criterion is satisfied; otherwise set $k \leftarrow k+1$.

2.2 Basic form of the alternating direction method of multipliers

The alternating direction method of multipliers is a first-order algorithm for solving convex optimization problems with separable objective functions and linear constraints. It is applicable to problems of the form [3]

$$\begin{aligned} &\min_{x\in\mathbb{R}^n, w\in\mathbb{R}^m} f(x)+g(w) \\ &\text{s.t.} \quad Mx-w=0, \end{aligned} \tag{2.2}$$

where $f:\mathbb{R}^n\to(-\infty,+\infty]$ and $g:\mathbb{R}^m\to(-\infty,+\infty]$ are proper closed convex functions, and $M\in\mathbb{R}^{m\times n}$ is a given linear operator.

The augmented Lagrangian function associated with problem (2.2) is defined as

$$\mathcal{L}_\lambda(x,w,p)=f(x)+g(w)+\langle p,Mx-w\rangle+\frac{\lambda}{2}\|Mx-w\|^2, \tag{2.3}$$

where $p\in\mathbb{R}^m$ denotes the Lagrange multiplier and $\lambda>0$ is the penalty parameter.

ADMM proceeds by alternately minimizing the augmented Lagrangian $\mathcal{L}_\lambda(x,w,p)$ with respect to the primal variables x and w, followed by an update of the dual variable p. Under convexity and suitable regularity conditions, the sequence generated by ADMM converges to a primal-dual solution of problem (2.2).

From the viewpoint of operator splitting, ADMM can be interpreted as an implementation of Douglas-Rachford splitting in an appropriately defined variable space. This equivalence not only provides a unified perspective for the convergence analysis of ADMM, but also lays the theoretical foundation for transferring algorithmic improvement strategies between ADMM and Douglas-Rachford splitting. For completeness, we present the classical alternating direction method of multipliers in Algorithm 2.

Algorithm 2 Alternating Direction Method of Multipliers

Iteratively update $x^{k+1}, w^{k+1}, p^{k+1}$ according to the following scheme:

$$\begin{aligned} x^{k+1}&=\operatorname*{argmin}_x\left\{f(x)+\langle p^k,Mx\rangle+\frac{\lambda}{2}\|Mx-w^k\|^2\right\}, \\ w^{k+1}&=\operatorname*{argmin}_w\left\{g(w)-\langle p^k,w\rangle+\frac{\lambda}{2}\|Mx^{k+1}-w\|^2\right\}, \\ p^{k+1}&=p^k+\lambda\big(Mx^{k+1}-w^{k+1}\big). \end{aligned}$$

Terminate if a prescribed stopping criterion is satisfied; otherwise set $k\leftarrow k+1$.

Theorem 2.1 (Subgradient Inversion of Conjugate Functions [9])**.** *Let $f:\mathbb{R}^n\to(-\infty,+\infty]$ be a proper, closed, and convex function, and let f^* denote its convex conju-*

gate. Then for any $x \in \operatorname{dom} f$ and $y \in \operatorname{dom} f^*$, the following equivalence holds:

$$y \in \partial f(x) \iff x \in \partial f^*(y).$$

Proof. Since f is a proper closed convex function, we have $f^{**} = f$. By definition of the conjugate function,

$$f^*(y) = \sup_u \{ \langle y,u \rangle - f(u) \}.$$

If $y \in \partial f(x)$, then by the Fenchel-Young inequality and the equality condition of subgradients, x attains the supremum in the above expression, and hence

$$\langle y,x \rangle - f(x) = f^*(y).$$

Using the definition of the biconjugate,

$$f^{**}(x) = \sup_v \{ \langle x,v \rangle - f^*(v) \},$$

and combining with the above identity, we obtain

$$f(x) = f^{**}(x) = \langle x,y \rangle - f^*(y).$$

This shows that y attains the supremum in the definition of $f^{**}(x)$, and therefore satisfies the optimality condition

$$x \in \partial f^*(y).$$

The converse implication $x \in \partial f^*(y) \Rightarrow y \in \partial f(x)$ can be derived in a completely analogous manner. $\square$

3 Douglas-Rachford splitting in the dual space and its equivalence to ADMM

This section demonstrates that, under an appropriate construction, applying the Douglas-Rachford splitting method to the dual problem yields an iteration that is equivalent to the augmented Lagrangian alternating direction method of multipliers applied to the primal problem. The following derivation closely follows the notation and framework of Eckstein-Bertsekas [6]. We make the following standard assumptions.

Let $f: \mathbb{R}^n \to (-\infty, +\infty]$ and $g: \mathbb{R}^m \to (-\infty, +\infty]$ be proper closed convex functions. Assume that the constraint $Mx = w$ satisfies a standard constraint qualification so that strong duality holds. Moreover, the proximal subproblems associated with f and g are assumed to be well-defined.

3.1 Primal and dual problems

Consider the convex optimization problem

$$\min_{x\in\mathbb{R}^n} f(x)+g(Mx), \tag{3.1}$$

where $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ and $g:\mathbb{R}^m\to\mathbb{R}\cup\{+\infty\}$ are proper closed convex functions, and $M\in\mathbb{R}^{m\times n}$ is a given matrix. By introducing an auxiliary variable $w\in\mathbb{R}^m$, problem (3.1) can be rewritten in an explicitly constrained form

$$\begin{aligned} \min_{x,w}\ & f(x)+g(w), \\ \text{s.t.}\ & Mx=w. \end{aligned} \tag{3.2}$$

Introducing the Lagrange multiplier $p\in\mathbb{R}^m$ for the constraint $Mx=w$, the Lagrangian function is given by

$$L(x,w,p)=f(x)+g(w)+\langle p,Mx-w\rangle.$$

Taking the dual of (3.2) yields the dual problem

$$\max_{p\in\mathbb{R}^m} -\left[f^*(-M^\top p)+g^*(p)\right], \tag{3.3}$$

where f^* and g^* denote the convex conjugates of f and g, respectively.

Define

$$F(p):=f^*(-M^\top p),\quad G(p):=g^*(p).$$

Then problem (3.3) is equivalent to

$$\min_{p\in\mathbb{R}^m} F(p)+G(p),$$

whose optimality condition can be written as $0\in\partial F(p)+\partial G(p)$. Let

$$A:=\partial F=\partial\big(f^*\circ(-M^\top)\big),\quad B:=\partial G=\partial g^*.$$

Then any dual solution $p^\star$ satisfies $0\in Ap^\star+Bp^\star$, i.e., $p^\star\in\operatorname{zer}(A+B)$.

3.2 Variable definitions and correspondence

We first introduce the key correspondence between primal and dual variables. Let z^k denote the DRS iteration variable, and define

$$p^k:=J_{\lambda B}(z^k),\quad w^k:=\frac{1}{\lambda}(z^k-p^k). \tag{3.4}$$

Intuitively, p^k can be interpreted as an approximation of the dual multiplier associated with the constraint $Mx=w$, while w^k represents a primal variable that approximately satisfies this constraint. This definition implies the relation

$$z^k = p^k + \lambda w^k. \tag{3.5}$$

3.3 Expanding the DRS iteration and linking to the primal problem

We now expand the three-step DRS iteration

$$\begin{cases} u^k = J_{\lambda B}(z^k), \\ v^k = J_{\lambda A}(2u^k - z^k), \\ z^{k+1} = z^k + v^k - u^k, \end{cases} \tag{3.6}$$

and, using the definitions in (3.4), rewrite it in terms of the variables (x,w,p).

Step 1. Handling $u^k = J_{\lambda B}(z^k)$. By the definition of the resolvent and the fact that $B = \partial g^*$,

$$u^k = J_{\lambda B}(z^k)$$

is equivalent to

$$z^k - u^k \in \lambda \partial g^*(u^k).$$

Substituting $u^k = p^k$ and using (3.5), we obtain $\lambda w^k \in \lambda \partial g^*(p^k)$, i.e., $w^k \in \partial g^*(p^k)$. By the conjugate subgradient inverse relation (Proposition 2.1), this is equivalent to

$$p^k \in \partial g(w^k). \tag{3.7}$$

This relation forms the basis for constructing the subsequent w-subproblem.

Step 2. Handling $v^k = J_{\lambda A}(2u^k - z^k)$. First note that

$$2u^k - z^k = 2p^k - (p^k + \lambda w^k) = p^k - \lambda w^k.$$

Since $A = \partial(f^* \circ (-M^\top))$, by the definition of the resolvent,

$$v^k = J_{\lambda A}(p^k - \lambda w^k) \iff p^k - \lambda w^k - v^k \in \lambda \partial\big(f^* \circ (-M^\top)\big)(v^k).$$

Using the chain rule

$$\partial\big(f^* \circ (-M^\top)\big)(p) = -M \partial f^*(-M^\top p),$$

and introducing a variable x^{k+1} such that $x^{k+1} \in \partial f^*(-M^T v^k)$, the above inclusion can be rewritten as

$$0 \in -Mx^{k+1} + \frac{1}{\lambda}\big(v^k - (p^k - \lambda w^k)\big).$$

Solving for v^k yields

$$v^k = p^k - \lambda w^k + \lambda M x^{k+1}. \tag{3.8}$$

Applying again the conjugate subgradient inverse relation to $-M^\top v^k \in \partial f(x^{k+1})$, we obtain $0 \in \partial f(x^{k+1}) + M^\top v^k$. Substituting (3.8) and completing the square, one verifies that x^{k+1} is equivalently given by

$$x^{k+1} \in \arg\min_x \left\{ f(x) + \langle p^k, Mx \rangle + \frac{\lambda}{2}\|Mx - w^k\|^2 \right\}. \tag{3.9}$$

Step 3. Updating $z^{k+1} = z^k + v^k - u^k$. Substituting (3.5), (3.8), and $u^k = p^k$ into the update gives

$$\begin{aligned} z^{k+1} &= (p^k + \lambda w^k) + (p^k - \lambda w^k + \lambda M x^{k+1}) - p^k \\ &= p^k + \lambda M x^{k+1}. \end{aligned}$$

Using the definition $z^{k+1} = p^{k+1} + \lambda w^{k+1}$, we obtain

$$p^{k+1} + \lambda w^{k+1} = p^k + \lambda M x^{k+1}. \tag{3.10}$$

To determine w^{k+1}, we consider an update consistent with (3.7) while incorporating the new information from x^{k+1}. Specifically, consider [9]

$$w^{k+1} \in \arg\min_w \left\{ g(w) - \langle p^k, w \rangle + \frac{\lambda}{2}\|Mx^{k+1} - w\|^2 \right\}. \tag{3.11}$$

Its first-order optimality condition is

$$0 \in \partial g(w^{k+1}) - p^k + \lambda(w^{k+1} - Mx^{k+1}), \tag{3.12}$$

or equivalently,

$$p^k \in \partial g(w^{k+1}) + \lambda(Mx^{k+1} - w^{k+1}).$$

Combining (3.10) and (3.12) yields the multiplier update

$$p^{k+1} = p^k + \lambda(Mx^{k+1} - w^{k+1}). \tag{3.13}$$

3.3.1 Algorithm summary and equivalence statement

Collecting (3.9), (3.11), and (3.13), we arrive at the following three-step iteration:

$$\begin{cases} x^{k+1} \in \operatorname{argmin}_x \left\{ f(x) + \langle p^k, Mx \rangle + \dfrac{\lambda}{2} \|Mx - w^k\|^2 \right\}, \\ w^{k+1} \in \operatorname{argmin}_w \left\{ g(w) - \langle p^k, w \rangle + \dfrac{\lambda}{2} \|Mx^{k+1} - w\|^2 \right\}, \\ p^{k+1} = p^k + \lambda \left(Mx^{k+1} - w^{k+1} \right). \end{cases} \tag{3.14}$$

This is precisely the standard form of the augmented Lagrangian alternating direction method of multipliers. Since the above derivation is reversible, we obtain the following equivalence theorem.

Theorem 3.1 (Equivalence of DRS and ADMM)**.** *Consider problem* (3.1)*, where f and g are proper closed convex functions. Define the dual operators $A = \partial(f^* \circ (-M^\top))$ and $B = \partial g^*$. For any initial dual variable z^0, generate $\{z^k\}$ by the DRS iteration* (3.6) *and define $\{p^k, w^k\}$ by* (3.4).

If $\{x^{k+1}\}$ is defined by (3.9) *and $\{w^{k+1}, p^{k+1}\}$ are defined by* (3.11) *and* (3.13)*, respectively, then the sequence $\{(x^k, w^k, p^k)\}$ coincides with the sequence generated by the ADMM iteration* (3.14) *starting from the same initial (w^0, p^0). The converse also holds.*

Proof. The sufficiency follows directly from the step-by-step derivation above. The necessity can be shown by reversing the arguments: starting from the ADMM iteration (3.14) and defining $z^k = p^k + \lambda w^k$, one can verify that $\{z^k\}$ satisfies the DRS iteration (3.6). □

4 A classification of transferable strategies under the DRS-ADMM equivalence

4.1 Exact operator-level transfer

We first characterize the strongest form of strategy transfer, namely the transfer at the operator level.

Definition 4.1 (Exact Operator-Level Transfer)**.** *Suppose that the DRS iteration can be written in the fixed-point form*

$$z^{k+1} = T_\gamma(z^k),$$

where

$$T_\gamma = I + J_{\gamma A}(2J_{\gamma B} - I) - J_{\gamma B}$$

is the Douglas-Rachford operator. If, under the DRS-ADMM equivalence mapping, the corresponding ADMM iteration can be written as a coordinate representation of the same operator family, and the parameters admit an explicit correspondence between the two algorithms, then the strategy is said to belong to exact operator-level transfer.

Below we present two typical examples.

4.1.1 Relaxation

The standard DRS iteration can be written in the following fixed-point form:

$$z^{k+1} = T_\gamma(z^k).$$

Introducing a relaxation parameter $\gamma \in (0,2)$ yields the relaxed DRS iteration

$$z^{k+1} = (1-\gamma)z^k + \gamma T_\gamma(z^k),$$

which is a typical Krasnosel'skiĭ-Mann iteration.

Under the variable transformation $z^k = p^k + \lambda w^k$, the corresponding ADMM iteration can be obtained, where

$$\tilde{w}^k = \gamma M x^{k+1} + (1-\gamma) w^k.$$

This leads to the following relaxed ADMM scheme:

$$x^{k+1} \in \operatorname*{argmin}_x \left\{ f(x) + \langle p^k, Mx \rangle + \frac{\lambda}{2} \|Mx - w^k\|^2 \right\}, \tag{4.1a}$$

$$\tilde{w}^k = \gamma M x^{k+1} + (1-\gamma) w^k, \tag{4.1b}$$

$$w^{k+1} \in \operatorname*{argmin}_w \left\{ g(w) - \langle p^k, w \rangle + \frac{\lambda}{2} \|\tilde{w}^k - w\|^2 \right\}, \tag{4.1c}$$

$$p^{k+1} = p^k + \lambda(Mx^{k+1} - w^{k+1}). \tag{4.1d}$$

Proposition 4.1 (Convergence)**.** *Let f and g be proper closed convex functions, and assume that the problem admits a primal-dual optimal solution. For any penalty parameter $\lambda > 0$ and relaxation parameter $\gamma \in (0,2)$, the relaxed ADMM iteration* (4.1) *generates a sequence that converges to a primal-dual optimal solution.*

This result follows from the averageness of the Douglas-Rachford operator together with the DRS-ADMM equivalence, see [1,6,8].

The relaxation applied to the DRS operator iteration is directly translated into the over-relaxation update in ADMM. The relaxation parameter γ has the same mathematical interpretation in both algorithms, namely forming a weighted combination between the current iterate and the operator update, thereby modifying the averaged property of the iteration mapping.

Since DRS and ADMM are strictly structurally equivalent at the operator level, the convergence theory of relaxed DRS can be directly transferred to the corresponding ADMM variant. Therefore, relaxation constitutes a typical example of exact operator-level transfer.

4.1.2 Metric scaling

Another modification that preserves the operator structure arises from metric scaling. Let M be a symmetric positive definite matrix and define the weighted inner product

$$\langle x,y\rangle_M = \langle Mx,y\rangle = \langle x,My\rangle.$$

This inner product changes the underlying metric structure of the space.

Under this weighted inner product, the resolvent of operator A becomes

$$J^M_{\gamma A} = (I+\gamma M^{-1}A)^{-1},$$

where the appearance of M^{-1} ensures that the operator is well-defined with respect to the weighted inner product (see [7]).

The resulting metric Douglas-Rachford operator is

$$T^M_\gamma = I + J^M_{\gamma A}\left(2J^M_{\gamma B} - I\right) - J^M_{\gamma B}.$$

The corresponding DRS iteration is

$$z^{k+1} = T^M_\gamma(z^k).$$

Since the resolvent $J^M_{\gamma A}$ defined under the weighted inner product remains nonexpansive, the operator T^M_γ is still averaged, and thus convergence of the iteration is preserved [7].

Under the variable mapping $z^k = p^k + \lambda w^k$ established in Section 3, this metric modification corresponds to a preconditioning of the augmented Lagrangian term in ADMM, leading to a preconditioned ADMM variant.

Because this strategy only changes the underlying inner-product structure while preserving the reflection composition structure of the Douglas-Rachford operator, the modification can be transferred to the ADMM framework via the equivalence mapping.

Therefore, metric scaling, similar to relaxation, constitutes another typical example of exact operator-level transfer.

4.2 Parameter-driven transfer

In many practical algorithms, the parameters of the method are not fixed but vary during the iteration process. In this case, the Douglas-Rachford iteration can be written as

$$z^{k+1} = T_{\gamma_k}(z^k),$$

where $\{\gamma_k\}$ is a sequence of stepsizes that changes along the iterations.

Unlike the fixed-operator case, this iteration is generated by a sequence of operators $\{T_{\gamma_k}\}$, and is therefore referred to as a nonstationary operator iteration.

Under suitable conditions, for example when the stepsize sequence $\{\gamma_k\}$ satisfies the boundedness condition

$$0 < \underline{\gamma} \le \gamma_k \le \overline{\gamma} < \infty,$$

or an asymptotic stability condition such as $\gamma_k \to \gamma$, the iteration still converges, see [10].

Through the DRS-ADMM equivalence established in Section 3, the parameter sequence $\{\gamma_k\}$ corresponds to a sequence of penalty parameters $\{\lambda_k\}$ in ADMM. Therefore, adaptive-penalty ADMM can be interpreted as a nonstationary Douglas-Rachford iteration applied to the dual problem.

It is important to note that the convergence theory for nonstationary DRS requires the parameter sequence $\{\gamma_k\}$ to be predetermined, i.e., chosen before the algorithm starts, rather than adapted based on the iterates. Under such conditions, the corresponding ADMM with varying penalty parameters inherits the same convergence guarantees via the equivalence mapping [10].

4.3 Heuristic transfer

Besides the strategies with clear operator interpretations, there also exist parameter update rules based on empirical considerations. Such strategies dynamically adjust algorithm parameters according to residuals or other statistics observed during the iteration, rather than being derived from operator splitting theory.

A typical example is the residual balancing strategy commonly used in ADMM [2]. In ADMM, the primal and dual residuals are usually defined as

$$r^k = Mx^k - w^k, \quad s^k = \lambda M^\top (w^k - w^{k-1}),$$

where r^k measures the violation of the constraint $Mx = w$, while s^k reflects the change in the dual variables.

The residual balancing strategy updates the penalty parameter according to

$$\lambda_{k+1}=\begin{cases}\tau\lambda_k, & \|r^k\|>\mu\|s^k\|,\\ \lambda_k/\tau, & \|s^k\|>\mu\|r^k\|,\\ \lambda_k, & \text{otherwise},\end{cases}$$

where $\tau>1$ and $\mu>1$ are prescribed constants.

Such update rules are generally difficult to express as explicit operator iterations. Consequently, they cannot be rigorously derived through the DRS-ADMM structural equivalence.

However, since DRS and ADMM share similar structures in terms of primal and dual residuals, these residual-based parameter adjustment strategies can still be transplanted between the two algorithms in practice, often leading to significant numerical improvements.

Therefore, although these strategies do not admit a strict operator-level explanation, they can still be transferred empirically across algorithms.

Unlike the parameter-driven strategies discussed in Section 4.2, these heuristic update rules lack rigorous convergence guarantees within the operator splitting framework. The parameter adjustment depends on the current residuals, making the resulting parameter sequence state-dependent and thus falling outside the scope of nonstationary operator theory. Nevertheless, these rules are widely used in practice due to their robust numerical performance [2, 12].

These three types of strategy transfer illustrate different levels of structural relationships between DRS and ADMM, ranging from exact operator equivalence to heuristic cross-algorithm adaptations.

References

[1] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.

[2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Found. Trends Mach. Learn. 3 (2011), 1–122.

[3] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vis. 40 (2011), 120–145.

[4] D. Davis and W. Yin, *A three-operator splitting scheme and its convergence analysis*, Set-Valued Var. Anal. 25 (2017), 829–858.

[5] J. Douglas and H. H. Rachford, *On the numerical solution of heat conduction problems in two and three space variables*, Trans. Am. Math. Soc. 82 (1956), 421–439.

[6] J. Eckstein and D. P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Math. Program. 55 (1992), 293–318.

[7] P. Giselsson and S. Boyd, *Metric selection in Douglas-Rachford splitting and ADMM*, arXiv:1410.8479, 2014.

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

[9] H. Liu, J. Hu, Y. Li, and Z. Wen, *Optimization: Modeling, Algorithms and Theory* (in Chinese), Higher Education Press, 2020.

[10] D. A. Lorenz and Q. Tran-Dinh, *Non-stationary Douglas-Rachford and alternating direction method of multipliers*, Math. Program. 175 (2019), 369–410.

[11] A. Themelis and P. Patrinos, *Douglas-Rachford splitting and ADMM for nonconvex optimization: Tight convergence results*, SIAM J. Optim. 30 (2020), 149–181.

[12] B. Wohlberg, *ADMM penalty parameter selection by residual balancing*, arXiv:1704.06209, 2017.