Article

Not peer-reviewed version

An Improved Column Generation Algorithm Based on Minimum-Norm Multipliers

Dingfang Su , Jie Tao * , Jiaxu Huang , Erzhan Gao

Posted Date: 11 February 2026

doi: 10.20944/preprints202602.0853.v1

Keywords: nonlinear programming; column generation; minimum norm multiplier; shadow price; unit commitment

doi:10.20944/preprints202602.0853.v1

Article

An Improved Column Generation Algorithm Based on Minimum-Norm Multipliers

Dingfang Su [1], Jie Tao [1,2,*], Jiaxu Huang [1] and Erzhan Gao [1]

[1] Business School, University of Shanghai for Science and Technology, No. 516 Jun Gong Road, Shanghai 200093, China

[2] School of Intelligent Emergency Management, University of Shanghai for Science and Technology, Shanghai 200093, China

* Correspondence: taojie@usst.edu.cn;

Abstract

Column generation is a fundamental technique for solving large-scale combinatorial optimization problems such as unit commitment and vehicle routing, yet its performance is often limited by dual oscillation. This study explores the intrinsic cause of this phenomenon from the perspective of shadow price theory and demonstrates that dual oscillation arises from the lack of marginal interpretability of Lagrange multipliers when multiple dual solutions coexist. To address this issue, an improved column generation framework is proposed in which traditional multipliers are replaced with minimum- norm multipliers that possess clear economic meaning and act as directional shadow prices. A generalized pricing subproblem is formulated, and partial minimum-norm multipliers are obtained through convex quadratic optimization to guide column generation. Numerical experiments on a simplified single-period unit commitment case show that the proposed approach eliminates invalid column generation and achieves speedy convergence to the optimal solution within only two iterations. The results indicate that the stabilization method enhances the consistency of dual variables and provides a more robust foundation for the theoretical and practical development of column generation algorithms.

Keywords: nonlinear programming; column generation; minimum-norm multiplier; shadow price; unit commitment

---

1. Introduction

Column Generation (CG) represents a classical paradigm for solving large-scale linear and integer programming problems. Its theoretical foundation can be traced back to the Dantzig–Wolfe decomposition principle proposed in 1960 [1]. The method iteratively decomposes the original problem into a Restricted Master Problem (RMP) and one or more Pricing Subproblems (PSP). During each iteration, CG introduces new columns obtained by solving the PSP, which are subsequently added to the master problem for re-optimization. The iterative process continues until no further valid columns can be produced, that is, until all variable coefficients satisfy the optimality conditions.

The fundamental advantage of CG lies in its ability to circumvent the computational intractability of explicitly enumerating all feasible columns. The method is particularly suited to large-scale combinatorial optimization problems with decomposable structures and has been successfully applied in areas such as cutting stock, crew scheduling, transportation planning, and vehicle routing problem(VRP) [2]. For instance, Furini et al. [3] developed a CG-based heuristic for two-stage multi-size cutting-stock problems, where dynamic-programming heuristics or integer formulations are used to solve PSPs. A restricted master problem is constructed with a subset of generated columns, and the corresponding upper bound is obtained through solver-based rounding techniques. In the field of crew scheduling, Janacek et al. [4] proposed a periodic crew-scheduling framework formulated via Dantzig–Wolfe decomposition. CG iteratively reconstructs feasible crew

 doi:10.20944/preprints202602.0853.v1

assignments while avoiding excessive integer constraints, enabling efficient computation on real-world railway instances. Similarly, An et al. [5] introduced an improved strategy for large-scale crew scheduling in which a pre-compiled set of promising shifts expedites the CG process, leading to significant computational acceleration. Within transportation and routing optimization, Archetti et al. [6] presented a branch and price and cut approach for the split-delivery VRP, dynamically generating routes and achieving new best solutions for multiple benchmark instances. Faiz et al. [7] considered vehicle-routing and scheduling problems with time windows and developed a path-based mixed integer programming formulation that incorporates CG to efficiently handle large-scale routing tasks. In addition, Alfandari et al. [8] demonstrated the effectiveness of a hybrid CG strategy for large-size covering integer programs in real-world transportation planning, showing accelerated convergence and improved integer-solution quality. Beyond classical scheduling and routing applications, Muts et al. [9] extended the CG paradigm to decentralized energy system planning. Their decomposition-based framework enables parallel resolution of nonlinear subproblems and achieves high-quality solutions for models with thousands of variables, further confirming the versatility of CG in modern optimization contexts. These studies collectively illustrate that CG is a cornerstone methodology in operations research—offering efficient decomposition, adaptive scalability, and broad applicability across diverse large-scale integer and mixed-integer optimization problems.

Despite its broad applicability, conventional CG suffers from a major practical obstacle—dual oscillation. In the early iterations, the RMP contains only a limited number of columns, resulting in unstable dual solutions that may deviate significantly from the optimal dual solution [10]. Such instability leads PSP to generate inefficient columns that temporarily improve the objective but contribute little to convergence, thereby increasing the required number of iterations [11]. This slow convergence is particularly severe in large-scale settings, restricting the algorithm's efficiency. To address this issue, various dual stabilization techniques have been proposed. Their central idea is to modify the dual output from the RMP so that the PSP receives dual information closer to optimality, enabling the generation of more effective columns and accelerating convergence. Common stabilization techniques include dual price smoothing, piecewise linear penalty functions, and artificial column insertion. These approaches aim to balance dual optimality with stability, suppressing oscillations while maintaining convergence.

Dual smoothing methods, such as the weighted averaging approach of Wentges [12], mitigate oscillations by averaging historical dual solutions. Although straightforward and computationally efficient, their performance depends strongly on smoothing parameters; excessive smoothing may obscure the optimal direction. Briant et al. [13] compared the stabilization mechanisms of the Bundle and Kelley methods, noting that the former adapts trust region step sizes to suppress oscillations but requires heuristic tuning and incurs higher per-iteration costs. The piecewise linear penalty function approach restricts the dual feasible region by adding penalty terms related to dual variables, as in the stabilization center framework by Du Merle [14] and its dynamic trust region extension. Ben Amor [11] further incorporated dual optimality inequalities, improving robustness using prior knowledge. The artificial column method, encompassing techniques such as the Big-M approach [15], ensures feasibility but exhibits strong parameter sensitivity and may distort dual interpretations. Artificial columns based on dual optimality inequalities [16] or pre-generated column pools also enhance stability, though their success depends on domain knowledge and parameter tuning. While these techniques achieve notable success, each suffers from limitations: parameter sensitivity, complex implementation, and potential numerical instability. Thus, obtaining high-quality dual information in a robust and efficient manner remains a pivotal challenge in CG.

Mathematically, the dual solution obtained from the RMP corresponds to the Lagrange multiplier of the restricted problem, and PSP generate valid columns by analyzing these multipliers. Consequently, dual oscillation is intrinsically related to multiplier selection. In resource allocation problems, multipliers can be interpreted as shadow prices that represent the marginal value of resources. However, when constraint qualifications are violated, multipliers may lose this interpretability. If multipliers are nonunique, they may fail to represent shadow prices accurately, leading to instability and erroneous CG [17,18]. Bertsekas and Ozdaglar showed that when the Linear Independence Constraint Qualification (LICQ) does not hold, multiple multipliers can emerge, and not all possess the meaning of shadow prices. In nonlinear programming, shadow prices may even fail to exist [19], implying that multipliers obtained from numerical solvers such as the simplex method might not reflect true marginal resource values and can oscillate dramatically across iterations.

Previous studies introduced special types of multipliers exhibiting shadow price-like properties. Akgul [20] distinguished "buy" and "sell" multipliers—representing, respectively, the producer's maximum buying price and minimum selling price for resource units under optimal use. Gauvin [21] extended this concept to nonlinear contexts, though practical computation remains challenging under general conditions. Bertsekas and Ozdaglar [14] further proposed "informational multipliers," describing partial shadow price information, and showed that Minimum-norm multipliers constitute a particular class of such informational multipliers. Tao and Gao [22] advanced this framework by defining directional shadow prices and proving that under lower semicontinuity of the optimal solution set, the Minimum-norm multiplier coincides with the shadow price. Moreover, compared with alternative multipliers, the Minimum-norm multiplier exhibits superior stability and robustness with respect to data perturbations, underscoring its potential for real-world applications.

Motivated by these observations, this paper develops an improved column generation framework based on minimum-norm multipliers. The main contributions are summarized as follows:

1. Theoretical contribution. It is shown that when multiple Lagrange multipliers exist, not all of them preserve the economic interpretation of shadow prices. The minimum-norm multiplier is proved to characterize the steepest-ascent direction of shadow prices and to provide the maximal lower bound on the objective improvement rate over all feasible directions.
2. Methodological contribution. To overcome the tendency of traditional PSPs to generate inactive columns under multiple-multiplier settings, a generalized pricing framework is proposed. A convex quadratic formulation is used to compute partial minimum-norm multipliers, together with a column selection strategy that guarantees objective improvement.
3. Empirical contribution: Through a one-dimensional cutting stock example, the existence and impact of multiple multipliers are explicitly illustrated. A single-period Unit Commitment problem further demonstrates that the proposed method avoids inactive columns and significantly accelerates convergence compared with the traditional CG approach.

Structure of the Paper

The remainder of this paper is organized as follows. Section 2 introduces a class of resource-constrained nonlinear optimization models and the associated value function framework. Section 3 analyzes the origin of nonunique Lagrange multipliers, revisits classical and extended shadow-price concepts, and highlights the limitations of single-resource marginal analysis. Section 4 establishes the theoretical properties of minimum-norm multipliers and their interpretation as directional shadow prices. Section 5 illustrates multiplier non-uniqueness in column generation using a cutting stock example. Section 6 presents the generalized pricing framework and validates the proposed method on a unit commitment problem. Section 7 concludes the paper and outlines directions for future research.

2. Basic Model

This study considers a class of nonlinear economic system optimization problems with resource constraints, which frequently arise in practical scenarios such as resource allocation and economic scheduling. The basic mathematical formulation is expressed as follows:

 doi:10.20944/preprints202602.0853.v1

$$\begin{aligned} \max_x \quad & f(x) \\ s.t. \quad & g_i(x) \le y_i, i = 1, \ldots, p \\ & g_i(x) = y_i, i = p+1, \ldots, q \end{aligned} \tag{1}$$

In model (1), both f and g_i ($i = 1, 2, \ldots, q$) are continuous mappings from the n-dimensional Euclidean space $\mathbb{R}^n$ to the real number field $\mathbb{R}$. Specifically, the objective function $f: \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, representing the overall utility of the economic system (e.g., total profit or operational efficiency). Each function g_i describes an individual resource supply constraint, illustrating the relationship between decision variables and resource consumption.

The vector $x \in \mathbb{R}^n$ denotes the decision variable, corresponding to the operational scheme of the system. The vector $y \in \mathbb{R}^q$ represents the resource constraint vector, characterizing the total available resources of the system. Its value fluctuates within a small neighborhood around the nominal resource level $\bar{y}$. The finiteness of resource availability y serves as a fundamental bottleneck in system optimization—small perturbations in y can exert non-negligible effects on the overall utility of the economic system.

For a given resource constraint vector y, the feasible solution set $M(y)$ consists of all decision vectors x that satisfy the constraints of model (1). Formally, $M(y) = \{x \in \mathbb{R}^n \mid g_i(x) \le y_i, i = 1, \ldots, p; g_i(x) = y_i, i = p+1, \ldots, q\}$. This set characterizes all operationally feasible configurations of the system under the specific resource supply level y. The value function $v(y)$ of model (1) represents the maximum achievable utility of the economic system for a fixed resource level y, defined by $v(y) = \max\{f(x) \mid x \in M(y)\}$. As an indicator of system performance, $v(y)$ reflects the optimal utility level corresponding to different resource endowments. Correspondingly, the solution set $S(y)$ of model (1) contains all decision schemes attaining the optimal utility : $S(y) = \{x \in M(y) \mid f(x) = v(y)\}$. When $S(y)$ includes a single element, the optimal solution is unique; otherwise, multiple optimal solutions exist—a phenomenon closely related to the subsequent theoretical analysis of multiple multipliers.

3. Shadow Price Theory Under Multiple Multipliers

In nonlinear optimization theory, Lagrange multipliers naturally correspond to the shadow prices of resources. Under the nominal resource supply level $\bar{y}$, let $\bar{x} \in S(\bar{y})$ denote a local optimal solution of model (1). A vector $\bar{\lambda} \in \mathbb{R}^q$ is defined as a Lagrange multiplier of model (1) at $\bar{x}$ if it satisfies the following Karush–Kuhn–Tucker (KKT) conditions:

$$\begin{gathered} \nabla_x L(\bar{x}, \bar{\lambda}) = 0 \\ \bar{\lambda}_i g_i(\bar{x}) = 0, i = 1, \ldots, p \\ \bar{\lambda}_i \ge 0, i = 1, \ldots, p \end{gathered}$$

Here, the Lagrangian function is defined by $L(x, \lambda) = f(x) + \sum_{i=1}^{q} \lambda_i g_i(x)$. For subsequent analysis, denote by $\Lambda(\bar{y})$ the set of all Lagrange multipliers corresponding to the optimal solutions $x \in S(\bar{y})$, and let $\lambda_{\min}$ represent the multiplier with the minimum Euclidean norm in this set.

3.1. Classical Definition of Shadow Prices

 doi:10.20944/preprints202602.0853.v1

Classical shadow prices are defined based on the marginal change of the value function. The shadow price p_i of the i-th resource equals the partial derivative of $v(y)$ with respect to the resource supply $\bar{y}$, evaluated at $y=\bar{y}$:

$$p_i = \left.\frac{\partial v(y)}{\partial y_i}\right|_{y=\bar{y}} = \lim_{\Delta y_i \to 0} \frac{(\bar{y}+y_i e_i)-v(\bar{y})}{\Delta y_i}, i=1,\dots,q$$

where $e_i \in \mathbb{R}^q$ denotes the unit vector with one in the i-th position and zeros elsewhere. This derivative quantifies the marginal rate of change in the system's optimal utility due to infinitesimal resource perturbations, and thus serves as an indicator of resource scarcity. However, classical shadow prices require that the value function $v(y)$ be differentiable. For non-convex optimization problems or those violating regularity conditions, $v(y)$ is often non-differentiable, making classical shadow prices inapplicable.

3.2. Extended Shadow Prices

To overcome this limitation, Gauvin et al. introduced the concepts of buy and sell shadow prices, which characterize marginal profit changes using Dini directional derivatives without relying on differentiability. Let $v(y+te_i)-v(\bar{y})]/t \ge M_i$ and m_i denote the market buy and sell prices of resource i, respectively ($i=1,\dots,q$). If the marginal utility gain from acquiring an additional tt units of resource i exceeds the market buy price (that is,[$v(\bar{y}+te_i)-v(\bar{y})]/t \ge M_i$), the resource is worth purchasing. The corresponding buy shadow price is defined by $p_i^+(\bar{y}) = \liminf_{t\to 0^+} \frac{v(\bar{y}+te_i)-v(\bar{y})}{t}$. Conversely, if the marginal utility loss from selling tt units of resource ii is smaller than the market sell price (that is, $v(\bar{y})-v(\bar{y}-te_i)]/t \ge m_i$), the resource is worth selling. The corresponding sell shadow price is defined as $p_i^-(\bar{y}) = \limsup_{t\to 0^+} \frac{v(\bar{y})-v(\bar{y}-te_i)}{t}$. By removing the differentiability requirement of the value function, buy and sell shadow prices substantially broaden the theoretical applicability of shadow pricing, allowing analysis in non-convex and irregular optimization contexts.

However, these extended shadow prices capture only the individual marginal effects of single resources. They neglect multi-resource interactions, such as complementarity or substitutability among different resources, which may lead to suboptimal system-wide allocation decisions and reduced economic efficiency. This limitation motivates the development of more general multiplier-based extensions—discussed in the next section—capable of representing joint resource effects via the concept of minimum-norm multipliers.

4. Implications of the Minimum-Norm Multiplier

In the resource-constrained economic system optimization model (1), Lagrange multipliers λ_i are typically given a clear economic interpretation—representing the marginal value of the i-th constrained resource. However, when Lagrange multipliers are non-unique, which often occurs near inequality boundaries or degenerate points, resource decisions based on a single multiplier may obscure potential improvements in efficiency and introduce ambiguity into economic meaning.

4.1. Minimum-Norm Multiplier Construction and Economic Interpretation

To overcome multiplier ambiguity, we construct the Minimum-norm multiplier problem in the neighborhood of a local optimal solution $\bar{x} \in S(\bar{y})$, assuming that the Mangasarian–Fromovitz Constraint Qualification (MFCQ) holds at $\bar{x}$. The objective is to obtain a unique and stable measure of marginal value, formalized as the following convex optimization problem:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 February 2026 doi:10.20944/preprints202602.0853.v1

$$\begin{aligned} \min_{\lambda} \quad & \frac{1}{2}\|(\lambda)\|^2 \\ s.t. \quad & \nabla f(\bar{x}) - \sum_{i=1}^{q} \lambda_i^T \nabla g_i(\bar{x}) = 0 \\ & \lambda_i g_i(\bar{x}) = 0, i = 1,\dots,p \\ & \lambda_i \geq 0, i = 1,\dots,p \end{aligned} \tag{2}$$

Under the MFCQ, the multiplier set $\Lambda(\bar{x},\bar{y})$ is bounded, closed, and convex; therefore, (2) always admits an optimal solution. Let $\lambda_{\min} \in \Lambda(\bar{x},\bar{y})$ denote the optimal solution. Economically, $\lambda_{\min}$ represents the vector with the smallest total penalty intensity on constraint violations—hence the "core resource support price" of the system. From the complementary slackness condition, for inactive inequality constraints ($g_i(\bar{x}) < 0$), we have $\lambda_i = 0$. Such constraints can therefore be omitted in subsequent analysis, and we assume $g_i(\bar{x}) = 0$, $i = 1,\dots,p$.

Introducing a dual variable $d \in \mathbb{R}^n$, the dual form of the primal problem (2) can be expressed as:

$$\max_{d\in\mathbb{R}^n} \nabla f(\bar{x})^T d - \frac{1}{2}\left[\sum_{i}^{q}\left(\left(\nabla g_i(\bar{x})^T d\right)^+\right)^2\right] \tag{3}$$

where $(\cdot)^+$ denotes the positive part function. By strong duality, the dual problem (3) and the primal problem (2) share the same optimal value. Let $\bar{d} \in \mathbb{R}^n$ denote the dual optimal solution. Then, optimality yields the correspondence:

$$(\lambda_{\min})_i = \left(\nabla g_i(x)^T d\right)^+, i = 1,\dots,q$$

where $(\lambda_{\min})_i$ denotes the i-th component of the Minimum-norm multiplier, and the essential equality $\nabla f(\bar{x})^T \bar{d} = \|\Lambda^{\min}\|^2$. Thus, the dual problem can be equivalently rewritten as:

$$\begin{aligned} \max_{d\in\mathbb{R}^n} \quad & \nabla f(\bar{x})^T d \\ s.t. \quad & \left(\left\langle \nabla g_i(\bar{x}), \bar{d} \right\rangle\right) = (\lambda_{\min})_i, i = 1,..,q \end{aligned} \tag{4}$$

The dual optimal direction $\bar{d}$ has geometric and economic significance: it characterizes the resource adjustment direction that maximizes the growth rate of the system's marginal utility $v(\bar{y}+d)$ around the nominal resource supply $\bar{y}$. Meanwhile, $\lambda_{\min}$ represents the marginal utility improvement rate per unit resource increment along this direction.

Based on the signs of the components of $\lambda_{\min}$, we can divide resources into two mutually exclusive sets: marginally sensitive resources $I^+ = \{i | (\lambda_{\min})_i > 0\}$ and marginally insensitive resources $I^0 = \{i | (\lambda_{\min})_i = 0\}$ (it is obvious that $I^+ \cup I^0 = \{1,\dots,q\}$).

- For $i \in I^+$: $g_i(\bar{x}+\bar{\lambda}t) = g_i(\bar{x}) + t\langle \nabla g_i(\bar{x}), \lambda \rangle + o(t) = g_i(\bar{x}) + (\lambda_{\min})_i t + o(t), i = 1,\dots,q$. Increasing the supply of such resources directly enhances the system value. Moreover, purchasing resources in the proportion of the components of $\lambda_{\min}$ maximizes the improvement in the system's marginal utility (i.e., $\langle \nabla f(\bar{x}), \bar{\mathrm{d}} \rangle = \|\lambda_{\min}\|^2$).
- For $i \in I^0, g_i(\bar{x}+\bar{\lambda}t) \leq g_i(\bar{x}), i = 1,\dots,q$, Increasing the supply of such resources has no impact on the system's marginal utility, indicating that these resources are already in a state of sufficiency or just meeting the demand.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 February 2026 doi:10.20944/preprints202602.0853.v1

The above analysis reveals the intuitive economic implications of the Minimum-norm multiplier, but the rigorous theoretical link between $\lambda_{\min}$ and the directional derivative of the value function remains to be established. This gap is addressed by the following theorem and its proof.

4.2. The Relationship Between the Minimum-Norm Multiplier and the Directional Derivative of the Value Function

We now establish the theoretical link between $\lambda_{\min}$ and the directional derivative of the value function.

Theorem 1. *Let the nominal resource supply* $\bar{y}$ *lie in the interior of the domain of the optimal solution mapping* $S(\cdot)$*, let* $\bar{x} \in S(\bar{y})$*, and assume* $\lambda_{\min}(\bar{x}, \bar{y}) \neq 0$*. If the MFCQ holds at* $\bar{x}$ *and the feasible set mapping* $M(\cdot)$ *satisfies the Aubin property at* $(\bar{x}, \bar{y})$*, then the directional derivative of the value function along the direction* $d^* = \frac{\lambda_{\min}(\bar{x},\bar{y})}{\|\lambda_{\min}(\bar{x},\bar{y})\|}$ *exists and satisfies:* $v'(\bar{y}, d^*) = \|\lambda_{\min}(\bar{x}, \bar{y})\|$.

Proof of Theorem 1. By the Aubin property, the value function v is locally Lipschitz at $\bar{y}$ [23], and its associated multiplier set satisfies $\Lambda(\bar{x}, \bar{y}) \subseteq \partial_c v(\bar{y})$, where $\partial_c v(\bar{y})$ denotes the Clarke subdifferential of v at $\bar{y}$. Since $\lambda_{\min} \in \Lambda(\bar{x}, \bar{y}) \subseteq \Lambda(\bar{x}, \bar{y})$, we have $\lambda_{\min} \in \Lambda(\bar{x}, \bar{y}) \subseteq \partial_c v(\bar{y})$. Furthermore, under MFCQ and Robinson's regularity [24], the mapping $\Lambda(\cdot,\cdot)$ is upper semicontinuous at $(\bar{x}, \bar{y})$.

Step 1. Establishing that $\lambda_{\min}$ is the Minimum-Norm Element in $\partial_c v(\bar{y})$

Suppose there exists $\xi \in \partial_c v(\bar{y})$ such that $\|\xi\| < \|\lambda_{\min}\|$. Then, by the definition of the Clarke subdifferential, there exists sequence $\{(x_k, y_k)\}_{k=1}^{\infty} \to (x_k, y_k)$ such that $x_k \in S(y_k)$ and $\{\nabla v(y_k)\}_{k=1}^{\infty} \to \xi$. Since MFCQ holds in a neighborhood of (x_k, y_k), the Envelope theorem implies

$$\nabla v(y_k) \in \Lambda(x_k, y_k)$$

By upper semicontinuity of $\Lambda(\cdot,\cdot)$, we obtain $\xi \in \Lambda(\bar{x}, \bar{y})$, contradicting the minimality of $\lambda_{\min}$. Hence, for every $\xi \in \partial_c v(\bar{y})$, $\|\xi\| \geq \|\lambda_{\min}\|$ and therefore $\lambda_{\min}$ is the unique element of smallest norm in $\partial_c v(\bar{y})$.

Step 2. Computing the Clarke Directional Derivative of v at $\bar{y}$ along d^*

The Clarke directional derivative of v at $\bar{y}$ in direction d^* is given by

$$v_c'(\bar{y}; d^*) = \limsup_{\substack{y \to \bar{y} \\ t \to 0^+}} \frac{v(y + td^*) - v(y)}{t}$$

Since the Clarke subdifferential $\partial_c(\bar{y})$ is a compact convex set, the Clarke directional derivative is equal to the support function of this set along direction d^*:

$$v_c'(\bar{y}; d^*) = \max_{\xi \in \partial_c v(\bar{y})} \xi^T d^* \tag{5}$$

From Rockafellar (1970, Theorem 27.4) [25], the minimal-norm element $\lambda_{\min}$ satisfies the variational inequality $(\xi - \lambda_{\min})^T \lambda_{\min} \geq 0, \forall \xi \in \partial_c v(\bar{y})$. Substituting $d^* = \frac{\lambda_{\min}}{\|\lambda_{\min}\|}$, we have

doi:10.20944/preprints202602.0853.v1

$$\xi^T d^* = \frac{\xi^T \lambda_{\min}}{\|\lambda_{\min}\|} = \frac{(\xi - \lambda_{\min})^T \lambda_{\min} + \|\lambda_{\min}\|^2}{\|\lambda_{\min}\|} \geq \|\lambda_{\min}\|$$

Equality holds only if $\xi = \lambda_{\min}$. Hence, the maximum of the support function (5) is uniquely attained at $\lambda_{\min}$, and $v_c(\bar{y}; d^*) = \|\lambda_{\min}\|$.

Step 3. Proving the Existence of the Directional Derivative and Its Equivalence to the Clarke Derivative

According to nonsmooth analysis theory, if the support function of the Clarke subdifferential has a unique maximizer along a given direction, then the (standard) directional derivative exists and coincides with the Clarke directional derivative. Since $\lambda_{\min}$ is the unique maximum point of (5), we conclude that the directional derivative of v at $\bar{y}$ along d^* exists, and

$$v'(\bar{y}; d^*) = v_c'(\bar{y}; d^*) = \|\lambda_{\min}\|$$

This completes the proof. □

The theorem establishes a fundamental link between the minimum-norm multiplier and the directional behavior of the value function in non-smooth optimization settings. It demonstrates that the direction of the minimum-norm multiplier corresponds to the steepest-ascent direction of shadow prices, thereby providing a rigorous economic interpretation of the multiplier as a directional shadow price. This theoretical insight forms the cornerstone of the improved CG framework developed in this study.

5. The Multiplier Multiplicity Phenomenon in Column Generation Algorithms

In CG algorithms, dual multipliers often play the role of shadow prices and guide the PSP to identify columns that can improve the objective value of the master problem.
However, when the dual problem is degenerate, multiple optimal multipliers may exist, and distinct multipliers can produce different subproblem optima.

More critically, not all optimal multipliers generate effective columns: even if a newly generated column has a negative reduced cost, it may fail to enter the basis, preventing improvement and delaying convergence.Therefore, designing a proper multiplier selection strategy is crucial to the efficiency of CG procedures. The cutting stock problem provides a simple and intuitive setting to illustrate this phenomenon. In what follows, a one-dimensional cutting stock problem is used to demonstrate how multiple optimal dual multipliers lead to distinct algorithmic behaviors in CG.

5.1. Problem Description

Consider a one-dimensional cutting stock problem. A batch of steel pipes of length 18 m must be cut to satisfy three orders: 20 pieces of 3 m, 20 pieces of 6 m, and 18 pieces of 7 m. The goal is to determine cutting patterns that satisfy all orders while minimizing the number of raw pipes used.

Let $y_i, i = 1, \ldots, N$ denote the number of raw pipes used with the i-th cutting pattern, and let $\left(\alpha_1^{[i]}, \alpha_2^{[i]}, \alpha_3^{[i]}\right), i = 1, \ldots, N$ represent the numbers of 3 m, 6 m, and 7 m pieces produced by pattern i. The integer programming model is

$$\begin{aligned} \min \quad & \sum_{i=1}^{N} y_i \\ s.t. \quad & \sum_{i=1}^{N} y_i \cdot (\alpha_1^{[i]}, \alpha_2^{[i]}, \alpha_3^{[i]})^T \geq (20, 20, 18)^T \\ & y_i \geq 0, i = 1, \ldots, N \end{aligned}$$

Because the number of feasible cutting patterns grows exponentially with the problem scale, explicit enumeration is computationally prohibitive. The CG algorithm addresses this issue by dynamically generating columns (cutting patterns) that have the potential to improve the objective value.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 February 2026 doi:10.20944/preprints202602.0853.v1

5.2. Problem Analysis

CG algorithm begins with a RMP containing a limited number of initial patterns. Assume the RMP starts with one pattern $\left(\alpha_1^{[i]},\alpha_2^{[i]},\alpha_3^{[i]}\right)=(1,1,1)$, meaning each 18 m pipe yields one 3 m, one 6 m, and one 7 m piece (2 m waste). The RMP is then

$$\begin{aligned} \min\quad & y_1 \\ s.t.\quad & y_1 \ge 20 \\ & y_1 \ge 20 \\ & y_1 \ge 18 \\ & y_1 \ge 0 \end{aligned} \tag{6}$$

By linear-programming duality, the dual problem is

$$\begin{aligned} \max\quad & 20\lambda_1+20\lambda_2+18\lambda_3 \\ s.t.\quad & \lambda_1+\lambda_2+\lambda_3 \le 1 \\ & \lambda_1,\lambda_2,\lambda_3 \ge 0 \end{aligned}$$

At the primal optimum $y_1^*=20$, the first two constraints of (6) are tight and the third is slack; thus $\lambda_3=0$ .Strong duality requires equality of objective values at optimum: $20\lambda_1 + 20\lambda_2=20$. Combining this with the active-constraint condition $\lambda_1+\lambda_2 \le 1$ yields the entire optimal dual solution set:

$$\Lambda\left\{(\lambda_1,\lambda_2,0,0)\mid \lambda_1+\lambda_2=1,\lambda_1\ge 0,\lambda_2\ge 0\right\} \tag{7}$$

Equation (7) indicates infinitely many dual optima forming a one-dimensional line segment. The non-uniqueness arises from degeneracy in the primal problem: several constraints are active simultaneously at the optimum.

The PSP seeks new cutting patterns with negative reduced costs using the current dual multipliers $(\lambda_1,\lambda_2,\lambda_3)$:

$$\begin{aligned} \min\quad & 1-\left(\lambda_1\alpha_1+\lambda_2\alpha_2+\lambda_3\alpha_3\right) \\ s.t.\quad & 3\alpha_1+6\alpha_2+7\alpha_3 \le 18 \\ & \alpha_1,\alpha_2,\alpha_3\in Z_+ \end{aligned} \tag{8}$$

A negative reduced cost indicates a promising column. However, when the dual optimum is not unique, different multipliers in Λ lead to different subproblem optima, affecting convergence behavior.

5.3. Comparative Analysis of Different Multiplier Selections

First, select the optimal dual solution $(\lambda_1,\lambda_2,\lambda_3)=(1,0,0)$. In this case, the PSP (8) is simplified as:

$$\begin{aligned} \min\quad & 1-\alpha_1 \\ s.t.\quad & 3\alpha_1+6\alpha_2+7\alpha_3 \le 18 \\ & \alpha_1,\alpha_2,\alpha_3\in Z_+ \end{aligned}$$

This subproblem is equivalent to maximizing α_1 under the capacity constraint. The optimal solution is $(\alpha_1,\alpha_2,\alpha_3)=(6,0,0)$, with a corresponding reduced cost of $1-6=-5<0$. After adding this column to the master problem and resolving it, the optimal solution is $y_1=20, y_2=0$, and the objective function value remains 20, indicating no improvement.

Next, select $(\lambda_1,\lambda_2,\lambda_3)=(0,1,0)$. The corresponding optimal solution is $(\alpha_1,\alpha_2,\alpha_3)=(0,3,0)$, with a reduced cost of $-2<0$. However, adding this column also fails to improve the objective function value. The reason is that both the new column $(6,0,0)$ and $(0,3,0)$ only perform well in meeting the demand for a single specification and cannot satisfy the requirements of other specifications at all. Therefore, even though the new columns have negative reduced costs, they fail to enter the basis due to the degeneracy and constraint structure of the master problem, and thus cannot improve the objective function.

In contrast, select the multiplier $(\lambda_1,\lambda_2)=(1/3,2/3)$. At this point, the subproblem (8) is simplified as:

$$\begin{aligned} \min \quad & 1-\frac{1}{3}\alpha_1-\frac{2}{3}\alpha_2 \\ s.t. \quad & 3\alpha_1+6\alpha_2+7\alpha_3\le 20 \\ & \alpha_1,\alpha_2,\alpha_3\in Z_+ \end{aligned}$$

There are multiple optimal solutions, namely $(\alpha_1,\alpha_2,\alpha_3)=(0,3,0),(2,2,0),(4,1,0),\ (6,0,0)$. Although the reduced cost of all these solutions is −1, adding these new columns to the master problem separately shows that only the pattern $(\alpha_1,\alpha_2,\alpha_3)=(2,2,0)$ improves the objective function. The updated master problem is:

$$\begin{aligned} \min \quad & y_1+y_2 \\ s.t. \quad & y_1+2\cdot y_2\ge 20 \\ & y_1+2\cdot y_2\ge 20 \\ & y_1+0\cdot y_2\ge 18 \\ & y_1\ge 0, y_2\ge 0, \text{integer} \end{aligned}$$

The optimal solution is $y^*=19,\ \ y_1=18, y_2=1$, which means that only 19 raw steel pipes are needed to meet all delivery requirements.

Different multiplier selections lead to different sequences of columns generated in subsequent iterations, thereby affecting the overall convergence efficiency. This phenomenon indicates that in the case of multiple multipliers, the choice of multipliers not only affects the improvement potential of a single iteration but also determines whether the algorithm can quickly approach the optimal solution. Therefore, designing a reasonable multiplier selection strategy is essential for enhancing the performance of CG algorithms.

6. Application of the Improved Column Generation Algorithm to the UC Problem

In CG algorithms, dual multipliers play the role of shadow prices that guide the PSP to improve the master problem.When the dual problem is degenerate, multiple optimal multipliers may exist. Using different multipliers can lead to distinct subproblem solutions, and some generated columns may remain inactive with non-positive reduced costs.Therefore, the selection of multipliers is crucial to avoid inactive CG and to accelerate convergence.

This section develops an improved CG method that employs the minimum-norm multiplier introduced in Section 4.By embedding the minimum-norm criterion into the dual update of the RMP, the method stabilizes the iterative process without altering the decomposition structure.The approach is validated on the UC problem where each generating unit's feasible schedule represents a potential column.

6.1. Multiple Multipliers in the UC Problem

A standard UC problem can be represented by the mixed-integer optimization model

 doi:10.20944/preprints202602.0853.v1

$$
\begin{aligned}
\min_{x} \quad & \sum_{g\in G} c_g(x_g) \\
st. \quad & \sum_{g\in G} A_g x_g = \alpha \\
& x_g \in \bar{X}_g, g \in G
\end{aligned}
\tag{9}
$$

where G denotes the set of generating units, x_g includes both discrete variables (unit on/off status) and continuous variables (power output), and the independent physical constraints for each unit are expressed as $D_g x_g \le d_g$, with $D_g \in \mathbb{R}^{l\times(n+m)}$ and $d_g \in \mathbb{R}^l$. Here l represents the number of constraints for unit g, including power limits, ramp rates, minimum up/down times, and logical limits. The cost function $c_g(x_g):\mathbb{R}^{n+m} \to \mathbb{R}$ comprises start-up, no-load, and marginal generation cost coefficients. $A_g \in \mathbb{R}^{T\times(n+m)}$ is the global constraint matrix (for example, power-balance equations), and $\alpha \in \mathbb{R}^T$ denotes the demand vector. A constant load is assumed for simplicity.

Because the UC problem is high-dimensional and highly coupled across time periods, direct optimization is difficult. To improve tractability, the Dantzig–Wolfe decomposition divides the original problem (9) into a global master problem and independent unit-level subproblems. The CG method proceeds iteratively: if the current solution satisfies the optimality conditions, the procedure stops; otherwise, new feasible scheduling schemes (columns) are added into the RMP to gradually reach the optimum.

Each feasible schedule x_g for unit g can be represented as a convex combination of the extreme points of its feasible region, i.e., $x_g = \sum_{n=1}^{N} z_g^n x_g^n, \sum_{n=1}^{N} z_g^n = 1, z_g^n \ge 0, \forall g \in G,$ $n = 1,\dots,N$. where $\{x_g^1, x_g^2, \dots, x_g^N\}$ denotes the set of extreme points of $\bar{X}_g$. Substituting the convex-combination coefficients z_g^n into (9) yields the master linear program:

$$
\begin{aligned}
\min_{z\in\mathbb{R}^N} \quad & \sum_{g\in G}\sum_{n=1}^{N} c_g(x_g)\cdot z_g^n \\
st. \quad & \sum_{g\in G}\sum_{N=1}^{N} A_g\left(x_g^n \cdot z_g^n\right) = \alpha \\
& \sum_{n=1}^{N} z_g^n = 1, \forall g \in G \\
& z_g^n \ge 0, \forall g \in G, n = 1,\dots,N
\end{aligned}
\tag{10}
$$

The master problem (10) is a linear programming problem with decision variables $z_g^n(g \in G, n = 1,\dots,N)$. However, since the number of extreme points in the master problem grows exponentially with the number of units G, enumeration is infeasible, making direct solution impractical. CG avoids full enumeration by dynamically adding effective columns z_g^n as needed. The corresponding RMP takes the form:

$$
\begin{aligned}
\min_{z\in\mathbb{R}^N} \quad & \sum_{g\in G, n\in N} c_g(x_g)\cdot z_g^n \\
st. \quad & \sum_{g\in G, n\in N} A_g\left(x_g^n \cdot z_g^n\right) = \alpha \\
& \sum_{n\in N} z_g^n = 1, \forall g \in G \\
& z_g^n \ge 0, \forall g \in G, n \in \hat{N} \\
& z_g^n = 0, \forall g \in G, n \in \left(\hat{N}\right)^c
\end{aligned}
\tag{11}
$$

where $\hat{N}$ is a subset of the set N, and $\left(\hat{N}\right)^c$ is its complement. Since the initial RMP (11) contains only a small number of columns, it is easy to solve. Let $\hat{z}\in\mathbb{R}^N$ be the optimal solution to RMP (10), and $\hat{\mu}\in\mathbb{R}$、$\hat{\pi}_g\in\mathbb{R}$、$\hat{\lambda}_g^n\in\mathbb{R}^N$ be the Lagrange multipliers corresponding to the constraints $\sum_{g\in G,n\in N}A_g\left(x_g^n\cdot z_g^n\right)=\alpha$, $\sum_{n\in N}z_g^n=1,\forall g\in G$, $z_g^n\geq 0,\forall g\in G,n\in\hat{N}$ and $z_g^n=0,\forall g\in G,n\in\left(\hat{N}\right)^c$, respectively. The optimality conditions of the RMP (11) are as follows:

$$\begin{cases}\hat{\lambda}_g^n=c_g(x_g^n)+\left\langle\hat{\mu},A_g x_g^n\right\rangle+\hat{\pi}_g,\forall g\in G,n\in N\\ \hat{\lambda}_g^n\geq 0,g\in G,n\in\hat{N}\end{cases}\tag{12}$$

where λ_g is the reduced cost of the decision variable z_g^n in the current basis of the RMP (11). And the optimality conditions of model (10) can be easily derived as:

$$\begin{cases}\hat{\lambda}_g^n=c_g(x_g^n)+\left\langle\hat{\mu},A_g x_g^n\right\rangle+\hat{\pi}_g,g\in G,n\in N\\ \hat{\lambda}_g^n\geq 0,g\in G,n\in N\end{cases}\tag{13}$$

Since (12) ⊂ (13), if every $\hat{\lambda}_g^n\geq 0$ also holds for all $n\in N$, then the RMP solution is globally optimal. Otherwise, columns with negative reduced cost must be generated via the PSP:

$$\bar{\lambda}=\min_{x_g\in\bar{X}_g,n\in N}\left\{c_g(x_g^n)+\left\langle\hat{\mu},A_g x_g^n\right\rangle+\hat{\pi}_g\right\}\tag{14}$$

If $\bar{\lambda}\geq 0$ for all g,n the current solution is optimal; if $\bar{\lambda}<0$, new columns are added. However, when the dual multipliers are non-unique, different $\left(\hat{\mu},\hat{\pi}_g\right)$ may generate the same or inactive columns, slowing convergence.

To illustrate the limitations of the traditional PSP in this scenario, consider a simple single-period UC problem with a system load of 100 MW, including two units G1 and G2. The power output range of G1 is [40, 80] MW with an energy cost of 50 \$/MWh, and the power output range of G2 is [0, 50] MW with an energy cost of 40 \$/MWh. The model reads

$$\begin{aligned}\min\quad & 50a_1+40a_2\\ st.\quad & a_1+a_2=100\\ & 40\leq a_1\leq 80\\ & 0\leq a_2\leq 50\end{aligned}\tag{15}$$

Assume initial columns $z_1^{(1)}:a_1^{(1)}=80$, $z_2^{(1)}:a_2^{(1)}=20$. where $a_1^{(1)}$ and $a_2^{(1)}$ denote the power outputs of unit 1 and unit 2 under the first scheme, respectively. The corresponding RMP is

$$\begin{aligned}\min_z\quad & 4000z_1^{(1)}+800z_2^{(1)}\\ st.\quad & 80z_1^{(1)}+20z_2^{(1)}=100 && \to\mu^{(1)}\\ & z_1^{(1)}=1,z_2^{(1)}=1 && \to\pi_1,\pi_2\\ & z_1^{(1)}\geq 0,z_2^{(1)}\geq 0 && \to\lambda_1^{(1)},\lambda_2^{(1)}\end{aligned}\tag{16}$$

Obviously, the optimal solution of the RMP (16) is $z_1^{(1)}=1,z_2^{(1)}=1$ with an optimal value of 4800. Let $\left(\mu^{(1)},\pi_1,\pi_2,\lambda_1^{(1)},\lambda_2^{(1)}\right)$ be the multipliers corresponding to the five constraints in the RMP (16). The associated multipliers satisfy:

$$\begin{cases}80\mu^{(1)}+\pi_1-\lambda_1^{(1)}=4000\\ 20\mu^{(1)}+\pi_2-\lambda_2^{(1)}=800\\ \lambda_1^{(1)}\geq 0,\lambda_2^{(1)}\geq 0\end{cases}\tag{17}$$

 doi:10.20944/preprints202602.0853.v1

Clearly, these multipliers are not unique, and arbitrary selection may yield inactive columns. Hence, a multiplier-selection strategy based on the minimum-norm multiplier is proposed.

6.2. Improved CG Method Based on Minimum-Norm Multipliers

In the case of multiple multipliers, the traditional PSP is prone to dual oscillation, which leads to an increased number of redundant iterations. To address this issue, this paper proposes a stabilization improvement solely for the "multiplier selection" step, without altering the original problem decomposition structure.

We note that the traditional PSP generates new columns by judging the non-negativity of the minimum reduced cost $\hat{\lambda}_g^n$ (corresponding to the decision variable z_g^n in the master problem (10)) within the feasible region. In fact, $\hat{\lambda}_g^n$ only corresponds to one component in the Lagrange multiplier set $\Lambda(x_g^n), g \in G, n \in N$. As analyzed in Section 3, the minimum l_2-norm multiplier represents the "optimal" increment of resources—meaning that increasing the resource supply level in this way enables the system to achieve the maximum marginal benefit. In other words, if a component of the Minimum-norm multiplier is large, the increase in the corresponding resource will exert a more significant positive impact on the overall system utility. Based on this property, we design the following generalized PSP to calculate partial Minimum-norm multipliers in the multiplier set.

$$\begin{aligned}
\min_{\lambda \in \mathbb{R}} \quad & \frac{1}{2}\left\|\hat{\lambda}\right\|^2 \\
st. \quad & \hat{\lambda}_g^n = c_g\left(x_g^n\right) + \left\langle \hat{\mu}, A_g x_g^n \right\rangle + \hat{\pi}_g, g \in G, n \in N \\
& \hat{\lambda}_g^n = 0, g \in G, n \in \hat{N} \\
& \hat{\lambda}_g^n \in \mathbb{R}, \forall g \in G, n \in \left(\hat{N}\right)^c \\
& \hat{\pi}_g, \hat{\mu} \in \mathbb{R}, \forall g \in G, n \in N
\end{aligned} \tag{18}$$

We refer to model (18) as the generalized PSP, in which $\mu \in \mathbb{R}, \pi_g \in \mathbb{R}, \lambda_g \in \mathbb{R}^N$ represent the Lagrange multipliers corresponding to the constraints in the RMP (11). The equality constraints in model (18) are directly derived from the optimality conditions (11). It is important to note that the optimal solution $\hat{\lambda}^*$ of model (18) does not denote the global minimum-norm multiplier over the entire multiplier set, i.e., $\underset{\lambda \in \Lambda}{min} \| \lambda \|$.

Rather, $\hat{\lambda}^*$ corresponds to the component λ_g^n with the smallest norm within the local Lagrange-multiplier set $\Lambda\left(\mu, \pi_g, \lambda_g^n\right) \in \mathbb{R}$ associated with the decision variable z_g^n in the RMP (11). Therefore, $\hat{\lambda}^*$ is regarded as a partial minimum-norm multiplier in the sense of the multiplier decomposition described above.

Model (18) is a convex quadratic program, which can be solved efficiently by standard quadratic-programming techniques. Its optimal solution, denoted $\hat{\lambda}^*$, serves as a stable and effective metric for identifying valuable scheduling schemes, thereby establishing the foundation of the improved CG algorithm.

The procedure for solving the UC problem using the generalized PSP (18) can be summarized as follows:

1. Initialization: Select a subset $n \in \hat{N}$ from the set of all scheduling schemes N as the initial scheme set. Set the decision variables $z_g^n \geq 0, n \in \hat{N}$ for the selected schemes, while imposing $z_g^n = 0$ for the unselected schemes with $n \in \left(\hat{N}\right)^c$.
2. Feasible-Scheme Generation: Generate several sequences of scheduling-scheme points within each unit's feasible region, i.e., construct multiple feasible solutions z_g^n for all $g \in G$ and $n \in N$.

3. Partial Minimum-Norm-Multiplier Calculation: Compute the partial minimum-norm multiplier $\hat{\lambda}^*$ corresponding to each scheduling scheme using model (18), and identify those schemes z_g^n with the maximum values of $\hat{\lambda}^*$. These schemes contribute most significantly to improving the system's marginal benefit and should therefore be added to the RMP (11). Note that there may exist multiple schemes attaining the maximum $\hat{\lambda}^*$; all such schemes should be included in the RMP (11). An advantage of this strategy is that even if the constructed scheme set does not contain the exact optimal solution of the original UC problem, the global optimum can still be represented as a linear combination of these highest-valued schemes.
4. RMP Update and Pruning: Remove the scheduling schemes $z_g^n, n \in \hat{N}$ for which $\hat{\lambda}^*$, since they make no positive contribution to the objective function of the system. Re-solve the updated RMP.
5. Convergence Check: Verify whether the updated RMP's optimal solution coincides with the global optimum of the original master problem (10). If it does, output the result; otherwise, repeat the above steps iteratively to expand the scheme set of the RMP until convergence is achieved.

Notably, the convex quadratic programming nature of model (18) ensures high computational efficiency. In particular, when the original UC problem exhibits a complex structure and large scale, approximating the global optimum by solving a limited number of convex quadratic programs iteratively leads to a significantly reduced computational burden compared with directly tackling the original mixed-integer linear programming model.

To further illustrate the computational process of the improved algorithm, we continue with the UC problem defined in (15). Assume that there exist N feasible scheduling schemes in total; the specific power outputs of each scheme are omitted for brevity. The initial columns are selected as $z_1^{[1]} : a_1^{[1]} = 80; z_2^{[1]} : a_2^{[1]} = 20$, implying that the remaining $N-1$ schemes are not included in the initial RMP. Accordingly, set the decision variables $z_1^1 \geq 0,\ z_2^1 \geq 0,$ for these initial columns, while assigning $z_1^i = 0,\ z_2^i = 0,$ for all other schemes. The resulting RMP can be formulated as follows:

$$\begin{aligned}
\min \quad & 4000 z_1^{[1]} + 800 z_2^{[1]} + \cdots + 50 a_1^{[N]} z_1^{[N]} + 40 a_2^{[N]} z_2^{[N]} \\
st. \quad & 80 z_1^{[1]} + 20 z_2^{[1]} + \cdots + a_1^{[N]} z_1^{[N]} + \alpha_2^{[N]} z_2^{[N]} = 100 \\
& z_1^{[1]} + \cdots + z_1^{[N]} = 1 \\
& z_2^{[1]} + \cdots + z_2^{[N]} = 1 \\
& z_1^{[1]} \geq 0, z_2^{[1]} \geq 0 \\
& z_1^{[i]} = z_2^{[i]} = 0, i = 2, \ldots, N
\end{aligned}$$

Let μ, π_1, π_2 denote the Lagrange multipliers corresponding to the first three constraints of the RMP, and let $\lambda_i^{[j]}$ be the multipliers associated with $z_i^{[j]}, i = 1, 2, j = 1, \ldots, N$, where $\alpha_i^{[j]}$ represents the power output of the i-th unit under the j-th scheduling scheme. Through derivation, the multipliers satisfy the following optimality conditions:

$$\begin{cases}
4000 + 80\mu + \pi_1 - \lambda_1^{[1]} = 0 \\
800 + 20\mu + \pi_2 - \lambda_1^{[2]} = 0 \\
50\alpha_1^{[j]} + a_1^{[j]}\mu + \pi_1 - \lambda_1^{[j]} = 0, j = 2, \ldots, N \\
40\alpha_2^{[j]} + a_2^{[j]}\mu + \pi_2 - \lambda_2^{[j]} = 0, j = 2, \ldots, N \\
\lambda_1^{[1]} = \lambda_2^{[1]} = 0
\end{cases}$$

Based on these conditions, the generalized PSP can be formulated as:

$$\begin{aligned}
\min_{\lambda\in\mathbb{R}} \quad & \tfrac{1}{2}\sum_{j=1}^{N}(\lambda_1^{[j]})^2+(\lambda_2^{[j]})^2 \\
s.t. \quad & \lambda_1^{[1]}=\lambda_2^{[1]}=0 \\
& \lambda_1^{[j]}=50a_1^{[j]}+a_1^{[j]}\mu+\pi_1,\ j=2,\dots,N \\
& \lambda_2^{[j]}=40a_2^{[j]}+a_2^{[j]}\mu+\pi_2,\ j=2,\dots,N
\end{aligned} \tag{19}$$

Substituting $\pi_1=-(4000+80\mu)$ and $\pi_2=-(800+20\mu)$, which are derived from the first equalities $\lambda_1^{[1]}=\lambda_2^{[1]}=0$, transforms the model into a univariate convex optimization problem with respect to μ:

$$\min_{\mu\in\mathbb{R}}\frac{1}{2}\sum_{j=2}^{N}\left(\left(a_1^{[j]}-80\right)\mu+(50a_1^{[j]}-4000)\right)^2+\left(\left(a_2^{[j]}-20\right)\mu+\left(40a_2^{[j]}-800\right)\right)^2$$

To obtain the optimal multiplier value, the first-order derivative of the objective function with respect to μ is taken and set equal to zero ($\partial/\partial\mu=0$). Solving the resulting equation yields the optimal multiplier $\mu=45$.

New scheduling schemes are then generated by varying the power outputs of the two units in 10-MW increments. The corresponding reduced-cost vectors and partial minimum-norm values $\hat{\lambda}^j$ are summarized in Table 1.

Table 1. Reduced Costs and Partial Minimum-Norm Values of Generated Scheduling Schemes.

Scheduling Scheme $(a_1^{[j]},a_2^{[j]})$ **(MW)**	**Reduced-Cost Vector** $(\lambda_1^{[j]},\lambda_2^{[j]})$ **($)**	**Partial Minimum-Norm** $\hat{\lambda}^j$ **($)**
$(50,50)$	$(-150,-150)$	212.13
$(60,40)$	$(-100,-100)$	141.42
$(70,30)$	$(-50,-50)$	70.71
$(80,20)$	$(0,0)$	0

From Table 1, the scheduling scheme $(50,50)$ yields the maximum partial minimum-norm value $\hat{\lambda}^j=212.13$, indicating that it provides the most significant improvement to the system's marginal benefit. Therefore, this scheme is added to the RMP, leading to the following updated model:

$$\begin{aligned}
\min \quad & 4000z_1^{[1]}+800z_2^{[1]}+2500z_1^{[2]}+2000z_2^{[2]}\cdots+50a_1^{[N]}z_1^{[N]}+40a_2^{[N]}z_2^{[N]} \\
st. \quad & 80z_1^{[1]}+20z_2^{[1]}+50z_1^{[2]}+50z_2^{[2]}+\cdots+a_1^{[N]}z_1^{[N]}+\alpha_2^{[N]}z_2^{[N]}=100 \\
& z_1^{[1]}+\cdots+z_1^{[N]}=1 \\
& z_2^{[1]}+\cdots+z_2^{[N]}=1 \\
& z_1^{[i]}\ge 0, z_2^{[i]}\ge 0, i=1,2 \\
& z_1^{[i]}=z_2^{[i]}=0, i=3,\dots,N
\end{aligned}$$

The optimal solution of the updated RMP is $z_1^{[1]}=0.25, z_1^{[2]}=0.75,\ z_2^{[1]}=0, z_2^{[2]}=1,$ with a minimum total cost of $\$4500$, achieving the global optimum of the original UC problem.

The numerical results verify that the proposed generalized PSP, which incorporates minimum-norm multiplier selection, effectively eliminates the generation of inactive or redundant columns that commonly occur in traditional CG algorithms. By prioritizing multipliers that contribute most to the marginal system benefit, the method stabilizes the dual iteration process and significantly accelerates convergence toward the global optimum. Although the verification is demonstrated using a simplified UC example, the same mechanism can be readily extended to large-scale UC models and other decomposition-based optimization frameworks, where it is expected to maintain both computational efficiency and convergence robustness.

doi:10.20944/preprints202602.0853.v1

7. Conclusions

This paper addresses the dual oscillation phenomenon in column generation algorithms for large-scale combinatorial optimization by examining its structural origin from the perspective of shadow-price theory and Lagrange multiplier nonuniqueness. Focusing on the role of multiple multipliers, an enhanced column generation scheme based on minimum-norm multiplier selection is developed.

The main conclusions and contributions can be summarized as follows:

1. It is shown that under multiple-multiplier conditions, the lack of a unique and economically meaningful marginal interpretation of Lagrange multipliers may cause conventional column generation to produce inactive or redundant columns, thereby degrading convergence behavior.
2. Among all admissible multipliers, the minimum-norm multiplier admits a clear variational and economic interpretation: it characterizes the steepest-ascent direction of shadow prices in the resource space and thus serves as a directional shadow price.
3. Leveraging this property, a generalized pricing framework is proposed in which traditional dual multipliers are replaced by partial minimum-norm multipliers, enabling principled multiplier selection in the presence of dual non-uniqueness.
4. By generating columns associated with the largest partial minimum-norm multipliers, the proposed strategy ensures effective improvement of the master-problem objective and enhances both the stability of pricing and the efficiency of the column generation process.
5. Numerical experiments on a single-period Unit Commitment problem confirm that the proposed method mitigates dual oscillation, avoids inactive columns, and accelerates convergence compared with the standard column generation approach.

Despite these encouraging results, several limitations merit further investigation. First, the current empirical validation is restricted to a Unit Commitment setting; extending the approach to broader classes of large-scale combinatorial and mixed-integer optimization problems is necessary to assess its generality. Second, the computation of partial minimum-norm multipliers introduces additional overhead; further research is needed to balance computational efficiency with solution quality.

Future research directions include the following:

1. Extension to more complex models. Apply the proposed framework to multi-period, multi-constraint Unit Commitment formulations and other large-scale mixed-integer programming problems to evaluate scalability and robustness.
2. Algorithmic acceleration. Integrate learning-based or hybrid optimization techniques to reduce the computational burden of computing partial minimum-norm multipliers and improve practical efficiency in large-scale settings.
3. Theoretical development. Further develop the theory of minimum-norm multipliers by investigating their economic and mathematical interpretations in dynamic, stochastic, and nonlinear combinatorial optimization contexts, thereby strengthening the theoretical foundation for stabilizing advanced decomposition algorithms.

Abbreviations

The following abbreviations are used in this manuscript:

CG	Column Generation
RMP	Restricted Master Problem
PSP	Pricing Subproblems
UC	Unit Commitment

References

1. Dantzig, G.B.; Wolfe, P. The Decomposition Principle for Linear Programs. Oper. Res. 1960, 8, 101–111.
2. Lübbecke, M.E.; Desrosiers, J. Selected Topics in Column Generation. Oper. Res. 2005, 53(6), 1007–1023.
3. Furini, F.; Malaguti, E.; Durán, R.M.; et al. A Column Generation Heuristic for the Two-Dimensional Two-Staged Guillotine Cutting Stock Problem with Multiple Stock Sizes. Eur. J. Oper. Res. 2012, 218(1), 251–260.
4. Janacek, J.; Kohani, M.; Koniorczyk, M.; Marton, P. Optimization of Periodic Crew Schedules with Application of Column Generation Method. Transp. Res. Part C: Emerg. Technol. 2017, 83, 165–178.
5. Chen, S.; Shen, Y. An Improved Column Generation Algorithm for Crew Scheduling Problems. J. Inf. Comput. Sci. 2013, 10(1), 175–183.
6. Archetti, C.; Bianchessi, N.; Speranza, M.G. A Column Generation Approach for the Split Delivery Vehicle Routing Problem. Networks 2011, 58(4), 241–254.
7. Faiz, T.I.; Vogiatzis, C.; Noor-E-Alam, M. A Column Generation Algorithm for Vehicle Scheduling and Routing Problems. Comput. Ind. Eng. 2019, 130, 222–236.
8. Alfandari, L.; Sadki, J.; Plateau, A.; Nagih, A. Hybrid Column Generation for Large-Size Covering Integer Programs: Application to Transportation Planning. Comput. Oper. Res. 2013, 40(8), 1938–1946.
9. Muts, P.; Bruche, S.; Nowak, I.; Wu, O.; Hendrix, E.M.; Tsatsaronis, G. A Column Generation Algorithm for Solving Energy System Planning Problems. Optim. Eng. 2023, 24(1), 317–351.
10. Vanderbeck, F. Implementing Mixed Integer Column Generation. In Column Generation; Springer US: Boston, MA, USA, 2005; pp. 331–358.
11. Ben Amor, H.; Desrosiers, J.; Valério de Carvalho, J.M. Dual-Optimal Inequalities for Stabilized Column Generation. Oper. Res. 2006, 54(3), 454–463.
12. Wentges, P. Weighted Dantzig-Wolfe Decomposition for Linear Mixed-Integer Programming. Int. Trans. Oper. Res. 1997, 4(2), 151–162.
13. Briant, O.; Lemaréchal, C.; Meurdesoif, P.; et al. Comparison of Bundle and Classical Column Generation. Math. Program. 2008, 113, 299–344.
14. Du Merle, O.; Villeneuve, D.; Desrosiers, J.; et al. Stabilized Column Generation. Discrete Math. 1999, 194(1-3), 229–237.
15. Valério de Carvalho, J.M. Exact Solution of Bin-Packing Problems Using Column Generation and Branch-and-Bound. Ann. Oper. Res. 1999, 86, 629–659.
16. Pessoa, A.; Sadykov, R.; Uchoa, E.; et al. A Generic Exact Solver for Vehicle Routing and Related Problems. Math. Program. 2020, 183, 483–523.
17. Tao, J.; Gao, Y. Identification and Calculation of Shadow Prices Under Multiple Lagrange Multipliers. J. Manag. Sci. China 2021, 24(3), 32–44. (in Chinese)
18. Cortez, K.L.; Rosenblueth, J.F. Normality and Uniqueness of Lagrange Multipliers. Discrete Contin. Dyn. Syst. Ser. A 2018, 38, 539-552.
19. Bertsekas, D.P.; Ozdaglar, A.E. Pseudonormality and a Lagrange Multiplier Theory for Constrained Optimization. J. Optim. Theory Appl. 2002, 114(2), 287–343.
20. Akgul, M. A Note on Shadow Prices in Linear Programming. J. Oper. Res. Soc. 1984, 35(5), 425–431.
21. Gauvin, J. Shadow Prices in Nonconvex Mathematical Programming. Math. Program. 1980, 19(1), 300–312.
22. Tao, J.; Gao, Y. Directional Shadow Price in Linearly Constrained Nonconvex Optimization Models. J. Glob. Optim. 2023, 88(2), 431–460.
23. Bonnans, J.F.; Shapiro, A. Perturbation Analysis of Optimization Problems; Springer Science & Business Media: Berlin, Germany, 2013.
24. Robinson, S.M. Strongly Regular Generalized Equations. Math. Oper. Res. 1980, 5(1), 43–62.
25. Rockafellar, R.T. Convex Analysis; Princeton University Press: Princeton, NJ, USA, 1970.