Minimax Theory and its Applications
Volume 11 (2026), No. 2s, 321-337

A non-smooth variational framework for optimization of chemical process systems

Dr.N. Muninathan[1], Aditi D. Raval[2], Andraju Bhanu Prasad[3], Pavas Saini[4], Dr.B.V. Raman[5], Jyoti Mohur[6], Ashutosh Kulkarni[7]

[1]Scientist, Central Research Laboratory, Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research, Enathur, Kanchipuram, Tamil Nadu, India. muninathan@maher.ac.in

[2]Assistant Professor, Faculty of Pharmacy, Gokul Global University, Gujarat, India. aditiraval.gphc@gokuluniversity.ac.in, https://orcid.org/0009-0000-3804-842X

[3]Associate Professor, Department of Computer Science and Engineering, Vardhaman College of Engineering, Shamshabad, Hyderabad, India. a.bhanuprasad@vardhaman.org

[4]Centre of Research Impact and Outcome, Chitkara University, Rajpura, Punjab, India. pavas.saini.orp@chitkara.edu.in, https://orcid.org/0009-0006-4138-5535

[5]Bonam Venkata Chalamayya Institute of Technology and Science, Amalapuram, Andhra Pradesh, India. directorrnd.bvts@bvcgroup.in

[6]Assistant Professor, Department of Computer Science & Engineering, Noida International University, Uttar Pradesh, India. jyoti.mahur@niu.edu.in

[7]Department of DESH, Vishwakarma Institute of Technology, Pune, Maharashtra, India. ashutosh.kulkarni@vit.edu

Abstract

Chemical process optimization frequently requires design and operational choices guided by mathematical models that are continuous and everywhere non-differentiable. The non-smoothness may come from either absolute-value penalty terms, changes in the operation regime, limits due to saturation, maxima and minima functions, phase-selection rules, or the piecewise representation of the process. Classic gradient-based optimization methods break down in the case where the functions that define process performance and constraints contain kinks or regime-dependent derivatives. The present paper develops a non-smooth variational model of chemical process optimization employing generalized sub-differentials and variational inequalities. The decision-making problem of the process is formulated in a compact convex feasible set, while the objective function describing process performance is taken to be Lipschitz continuous with nonsmooth penalty terms. The first-order necessary conditions of optimality are derived using Clarke's generalized derivative and the optimization problem is formulated as a variational inequality involving the generalized derivative and the normal cone of the feasible region. Under continuity assumptions, local Lipschitz continuity, convexity of the feasible region and pseudomonotonicity of the associated generalized mapping, the existence of optimality results is shown. Also, the computational complexity is evaluated with regards to the number of process variables, active constraints and nonsmooth evaluations. Experimental investigations are performed on a typical problem of chemical process optimization, where different states of operation are examined with respect to reaction temperature, residence time, feed rate and intensity of separation. The framework presented achieves values of 0.184, 0.219, and 0.287 in the testing cases while satisfying the restrictions on yield for the product. Comparative analysis with various approaches, including smooth projected gradient, standard penalty for optimization, and derivative-free search, shows the effectiveness of the combined application of generalized first-order information. The results provide mathematic basis for the framework of optimizing the systems in case of multi-mode processes, which cannot be reasonably described by globally-smooth functions.

Keywords: Non-smooth optimization, variational inequality, clarke subdifferential, generalized gradient, chemical process optimization, projected subgradient method, nonsmooth stationarity, process systems engineering.

1. Introduction

Optimization is one of the crucial elements of process systems engineering since many aspects of optimal design and operation of chemical processes must be balanced, including economic performance, material consumption, energy use, product quality, safety conditions, and equipment limitations. The latest trends in process systems engineering have shown an increasing inclination toward implementing optimization techniques designed to solve nonlinear optimization problems and considering uncertainties as well as the integration of design and operation in complex decision-making (Pistikopoulos et al., 2021). At the same time, the area of chemical process optimization has evolved from conventional mathematical optimization approaches to the hybrid methods that integrate data and modeling of processes, thus demonstrating that feasibility and optimization cannot be separated (Reis & Saraiva, 2022; Sindhu & Poornimadarshini, 2025).

While several process models are developed based on continuous physical relationships, real-world optimization models tend to have non-smooth parts. Deviation terms, maximum and minimum operators, penalties due to switches, operating-region transitions, saturation functions, as well as complementarity constraints and piecewise representation of a process can create certain points where conventional gradient cannot be determined (Gebken, 2024). Recent advances of algorithms of numerical optimization of non-smooth functions enabled the use of some unified techniques, which are based on use of generalized derivatives, bundle methods, use of subgradient information, and first-order variational descriptions and definitions. Such approaches can be used when it is impossible to convert an optimization model into a globally smooth form, while keeping its essential characteristics.

Mathematically, smooth and non-smooth optimization can be distinguished quite clearly. In the case of differentiable objectives, in order to define a first-order stationary point, vanishing of the gradient can be used along with standard constraint qualifications. However, in the case of a locally Lipschitz function the classical gradient may not exist at the boundaries of the regime. Generalized derivatives instead of ordinary ones come into play and thus first order information can be represented as a set. The recent developments in non-smooth optimization suggest advanced results regarding the convergence of subgradient, bundle, gradient sample and proximal methods in convex, non-convex, and weakly semismooth environments (Han & Lewis, 2023; Huang et al., 2023; Shen et al., 2024).

Variational inequalities are used as an alternative representation of constrained optimization problems from a mathematical point of view. In variational inequities, when a feasible point is being defined, its generalized first-order direction at this point is non-negative with respect to any feasible displacement instead of relying on the zero-gradient condition (Chang & Bai, 2021). Variational inequalities have seen a wide range of applications with regards to monotone operators, nested problems, constrained optimization, and first-order iterative algorithms (Alacaoglu et al., 2021; Lampariello et al., 2022). Among others, the projected and the forward reflected methods enable partial solutions for problems along with the possibility for straightforward convergence analysis in practice (Kaushik & Yousefian, 2021).

The application of generalized gradients is very desirable for chemical process optimization due to the fact that a lot of engineering penalties can be naturally expressed by using mortgages in the form of point-wise maximum or absolute value (Soni & Saini, 2025; Abbood, 2022). For instance, deviation from the target temperature can be specified using an absolute value term. If $r(x)$ denotes a process specification, the positive-part penalty is defined as equation (1):

$$[r(x)]_+ = \{0, r(x)\} \qquad (1)$$

Is continuous but generally non-differentiable at $r(x) = 0$. Thus, the direct use of classical gradient techniques can lead to either undefined or unstable derivative values right at the operational limit. Nonsmooth optimization techniques can be used effectively to capture this structure (Rodomanov & Nesterov, 2023; Díaz & Grimmer, 2023; Bolte et al., 2023).

The compute-based methods for the analysis of non-smooth chemical-process models are becoming equally important as well. Global and process optimization cases have shown that chemical process models can be nonconvex, experience issues with convergence and lack appropriate derivative information shedding light on the necessity to engage in hybrid and deterministic optimization measures (Pan et al., 2021; Ma et al., 2022;

Bardooli et al., 2022). Current research in the field of global process optimization has considered convex relaxation and nonsmooth structures based on Karush-Kuhn-Tucker theory which proved that generalized derivatives can be applied in process-centered mathematical optimization (Yuan & Khan, 2023). The process optimization cases that involve the utilization of data-driven models, surrogate approaches and other cutting-edge optimization tools point out the need to develop algorithms capable of functioning efficaciously even in terms of the non-global smoothness of the resulting process model (Jog et al., 2024).

In spite of the advancements, there still exists a mathematical gap with regards to general nonsmooth optimization theory and a process-oriented version. The versions that are considered do incorporate aspects such as nonsmooth penalties, process constraints, gradients and convergence analysis resulting in a single coherent framework. On the contrary, most of the available works address such elements as computational performance, surrogate quality or the application of the particular specificities leading to the conclusion that the current project does not set out to describe a process-specific optimizer but intends to introduce a much clearer framework.

As such, this study develops the non-smooth variational formulation for a constrained chemical-process optimization problem. The key mathematical aspects are Clarke generalized gradient, the normal cone of the convex feasible set, and the variational inequality related to the generalized first-order mapping. The proposed setting allows having absolute value and positive part terms in the process objective without using any smoothing.

The main contributions of this study are as follows.

- A non-smooth optimization model with local Lipschitz continuity for chemical process systems including penalties for deviations from target parameters and operational limits.
- The constrained optimization problem is reformulated into a variational inequality involving Clarke's generalized gradient and the normal cone of feasible processing area.
- Existence and stationarity results are obtained based on assumptions of compactness, convexity, continuity of local Lipschitz type, and pseudomonotonicity.
- An algorithm based on projected subgradient is developed for approximating solutions of the resulting non-smooth variational inequality.
- The convergence of the produced sequence is established, and computational effort is evaluated in terms of the dimension and number of evaluations of non-smooth function.
- Numerical examples validate the operation of the proposed idea in the cases of normal operation, high penalty and restricted capacity.

This paper has been organized as follows: Section II presents the mathematical formulation and proposed approach of non-smooth variational frameworks. The Section III includes the numerical experiments, comparison with other optimization techniques, sensitivity analysis, and engineering interpretation. Finally, Section IV concludes the research and suggests the possible research directions.

2. Mathematical formulation and proposed non-smooth variational framework

Consider a chemical process system in which the operating decision vector is represented as equation (2):

$$x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n \qquad (2)$$

The elements of x possibly indicate reaction temperature, residence time, feed flow rate, pressure, reflux ratio, the intensity of separation, utility distribution, and so on. The process-performance characteristics can be expressed as functions of x.

Let the total process objective be expressed as equation (3):

$$F(x) = F_s(x) + F_{ns}(x) \qquad (3)$$

Where F_s represents the continuously differentiable component and F_{ns} represents the non-smooth component.

The smooth element includes energy consumption, the cost of the materials needed, utility usages, and process costs. The non-smooth element introduces target deviations and regime-dependent costs. It is expressed as

equation (4):

$$F_{ns}(x) = \sum_{i=1}^{p} \alpha_i \, |r_i(x)| + \sum_{j=1}^{q} \beta_j \left[s_j(x)\right]_+ \qquad (4)$$

Where $\alpha_i, \beta_j \geq 0$, $r_i(x)$ denotes a process target deviation, and $s_j(x)$ denotes a process-limit violation.

The positive-part operator is given as equation (5):

$$[z]_+ = \max\{0, z\} \qquad (5)$$

Consequently, $F(x)$ is continuous and locally Lipschitz provided that the process functions $r_i(x)$ and $s_j(x)$ are continuously differentiable.

The feasible operating region is defined as equation (6):

$$K = \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ i = 1, \ldots, m,\ Ax \leq b,\ l \leq x \leq u\} \qquad (6)$$

Where $g_i(x)$ represents nonlinear process constraints, $Ax \leq b$ represents linear material, utility, or equipment constraints, and l and u define lower and upper operating bounds.

The chemical process optimization problem is therefore formulated as equation (7):

$$\min_{x \in K} F(x) \qquad (7)$$

The principal distinction from a conventional smooth formulation is that F is not assumed to be differentiable over the complete feasible region.

2.1 Clarke generalized gradient

For a locally Lipschitz function $F: \mathbb{R}^n \to \mathbb{R}$, the Clarke generalized directional derivative at x in direction d is defined as equation (8):

$$F^\circ(x; d) = \limsup_{y \to x,\, t \downarrow 0} \frac{F(y + td) - F(y)}{t} \qquad (8)$$

The Clarke generalized gradient is then defined as equation (9):

$$\partial^C F(x) = \{v \in \mathbb{R}^n : F^\circ(x; d) \geq \langle v, d \rangle,\ \forall d \in \mathbb{R}^n\} \qquad (9)$$

If F is differentiable at x, then equation (10) holds:

$$\partial^C F(x) = \{\nabla F(x)\} \qquad (10)$$

Hence, the generalized gradient reduces to the classical gradient whenever the process model is smooth.

For the absolute-value function, equation (11) is given by:

$$\partial^C |z| = \begin{cases} \{1\}, & \text{if } z > 0 \\ \{-1\}, & \text{if } z < 0 \\ [-1, 1], & \text{if } z = 0 \end{cases} \qquad (11)$$

Similarly, for the positive-part function, equation (12) is given by:

$$\partial^C [z]_+ = \begin{cases} \{1\}, & \text{if } z > 0 \\ \{0\}, & \text{if } z < 0 \\ , & \text{if } z = 0 \end{cases} \qquad (12)$$

Using the generalized chain rule, a generalized gradient of the non-smooth objective can therefore be represented as equation (13):

$$v(x) = \nabla F_s(x) + \sum_{i=1}^{p} \alpha_i \zeta_i(x) \nabla r_i(x) + \sum_{j=1}^{q} \beta_j \eta_j(x) \nabla s_j(x) \quad (13)$$

Where equation (14) and equation (15) are defined as:

$$\zeta_i(x) \in \partial^C |r_i(x)| \quad (14)$$
$$\eta_j(x) \in \partial^C [s_j(x)]_+ \quad (15)$$

Equation (13) provides an explicit mathematical mechanism for retaining non-smooth process penalties without replacing them by smooth approximations.

2.2 Normal cone and variational formulation

For a nonempty closed convex set K, the normal cone at $x \in K$ is defined as equation (16):

$$N_K(x) = \{w \in \mathbb{R}^n : \langle w, y - x \rangle \leq 0, \forall y \in K\} \quad (16)$$

For an interior point of K, equation (17) is given by:

$$N_K(x) = \{0\} \quad (17)$$

At an active process boundary, the normal cone contains the outward constraint directions and therefore provides the mathematical representation of active operating restrictions.

A feasible point $x^* \in K$ is defined as a generalized stationary solution of equation (7) if equation (18):

$$0 \in \partial^C F(x^*) + N_K(x^*) \quad (18)$$

Equivalently, there exists a vector defined as equation (19):

$$v^* \in \partial^C F(x^*) \quad (19)$$

Such that equation (20):

$$-v^* \in N_K(x^*) \quad (20)$$

Using the definition of the normal cone, equation (20) yields equation (21):

$$\langle v^*, x - x^* \rangle \geq 0, \quad \forall x \in K \quad (21)$$

Therefore, the non-smooth chemical process optimization problem can be represented by the variational inequality given as equation (22):

$$\text{Find } x^* \in K \text{ such that } \exists v^* \in \partial^C F(x^*) \text{ satisfying } \langle v^*, x - x^* \rangle \geq 0, \forall x \in K \quad (22)$$

Equation (22) is the principal mathematical formulation proposed in this paper.

2.3 Assumptions and mathematical preliminaries

The following assumptions are adopted for the theoretical analysis:

Assumption A1: The feasible process set K is nonempty, closed, bounded, and convex.

Assumption A2: The smooth process component $F_s(x)$ is continuously differentiable on an open neighbourhood of K.

Assumption A3: Each process target function $r_i(x)$ and process-limit function $s_j(x)$ is continuously differentiable on K.

Assumption A4: The total objective F is locally Lipschitz on K.

Assumption A5: The generalized mapping $V(x) = \partial^C F(x)$ is bounded on K.

Assumption A6: A measurable generalized-gradient selection $v(x) \in \partial^C F(x)$ can be obtained at every iteration.

Assumption A7: The selected generalized mapping is pseudomonotone on K, namely equation (23) holds:

$$\langle v(x), y - x\rangle \geq 0 \Longrightarrow \langle v(y), y - x\rangle \geq 0 \quad (23)$$

for all $x, y \in K$.

Under Assumptions A1–A7, the variational inequality in equation (22) is well posed.

Define the generalized process mapping as equation (24):

$$V(x) = \partial^C F(x) \quad (24)$$

The solution set of the variational inequality is defined as equation (25):

$$S = \{x \in K : \forall v \in V(x), \langle v, y - x\rangle \geq 0, \forall y \in K\} \quad (25)$$

The optimization and variational formulations are therefore connected through equation (26):

$$x^* \in \arg\min_{x \in K} F(x) \Leftrightarrow x^* \in S \quad (26)$$

Under the convexity and regularity conditions considered below

2.4 Existence of a variational solution

The existence of a solution is necessary before an iterative computational procedure can be analysed.

Theorem 1: Existence of a non-smooth variational solution

Suppose that Assumptions A1–A7 hold. Then the variational inequality defined by equation (22) possesses at least one solution $x^* \in K$.

Proof

Since K is nonempty, closed, bounded, and convex by Assumption A1, it is compact in finite-dimensional Euclidean space. The local Lipschitz continuity of F implies that the Clarke generalized gradient $\partial^C F(x)$ is nonempty, compact, and convex for every $x \in K$.

By Assumption A5, the generalized mapping $V(x)$ is bounded on K. The closedness of the generalized-gradient graph together with compactness of K ensures that the generalized variational inequality mapping is bounded on the feasible region.

Consider the regularized mapping given as equation (27):

$$V_\mu(x) = V(x) + \mu x, \quad \mu > 0 \quad (27)$$

The additional term μx makes the regularized mapping strongly monotone in the single-valued selection sense. Hence the regularized variational inequality represented by equation (28):

$$\langle v_\mu(x_\mu), x - x_\mu\rangle \geq 0, \quad \forall x \in K \quad (28)$$

Admits a solution $x_\mu \in K$. Since K is compact, the sequence of solutions corresponding to $\mu \downarrow 0$ possesses an accumulation point $x^* \in K$. By the closedness of the generalized-gradient graph and Assumption A7, the limiting point satisfies equation (22). Therefore, $x^* \in S$.

2.5 Optimality characterization

The variational inequality provides a first-order characterization of the constrained non-smooth optimization problem.

Theorem 2: Variational stationarity

Let $x^* \in K$ be a local minimizer of equation (7). If F is locally Lipschitz and K is closed and convex, then there exists a vector defined as equation (29):

$$v^* \in \partial^C F(x^*) \quad (29)$$

Such that equation (30):

$$\langle v^*, x - x^* \rangle \geq 0, \quad \forall x \in K \quad (30)$$

Proof

A feasible first-order displacement does not lead to strict decrease in the objective at a local minimizer. Equation (31) expresses the generalized first-order necessary condition for a locally Lipschitz objective.

$$0 \in \partial^C F(x^*) + N_K(x^*) \quad (31)$$

Consequently, there exists $v^* \in \partial^C F(x^*)$ for which equation (32):

$$-v^* \in N_K(x^*) \quad (32)$$

Using the definition of the normal cone, equation (32) yields equation (33):

$$\langle -v^*, x - x^* \rangle \leq 0 \Longrightarrow \langle v^*, x - x^* \rangle \geq 0 \quad (33)$$

for every $x \in K$. Rearranging gives equation (30).

If F is convex on K, the variational condition is also sufficient. Specifically, equation (34):

$$\langle v^*, x - x^* \rangle \geq 0, \quad \forall x \in K \quad (34)$$

with equation (35):

$$v^* \in \partial^C F(x^*) \quad (35)$$

implies equation (36):

$$F(x) \geq F(x^*), \quad \forall x \in K \quad (36)$$

Thus, under convexity, a variational solution is a global minimizer.

2.6 Regularized non-smooth variational algorithm

Solving the equation (22) directly can be challenging in cases when the generalized gradient is obtained as a set and when multiple process constraints are actively valid at the same time. This leads to the use of regularized projected sub gradient technique of solving the problem.

Let $x^k \in K$ denote the current iterate. Select equation (37):

$$v^k \in \partial^C F(x^k) \quad (37)$$

A regularized generalized-gradient direction is defined as equation (38):

$$d^k = v^k + \gamma_k (x^k - x^0) \quad (38)$$

Where $\gamma_k > 0$ is the regularization parameter and x^0 is the initial feasible operating point.

The next process decision is calculated as equation (39):

$$x^{k+1} = P_K(x^k - \alpha_k d^k) \quad (39)$$

Where P_K denotes the Euclidean projection onto the feasible process set and $\alpha_k > 0$ is the step size.

The projection is defined as equation (40):

$$P_K(z) = \arg\min_{x \in K} \| x - z\|^2 \quad (40)$$

The optimality condition of the projection gives equation (41):

$$\langle x^k - \alpha_k d^k - x^{k+1}, x - x^{k+1} \rangle \leq 0, \quad \forall x \in K \quad (41)$$

Setting $x = x^*$ in equation (41) yields equation (42):

$$\alpha_k \langle d^k, x^{k+1} - x^* \rangle \geq \langle x^k - x^{k+1}, x^{k+1} - x^* \rangle \quad (42)$$

Using the identity given by equation (43):

$$2\langle a - b, b - c \rangle = \|a - c\|^2 - \|a - b\|^2 - \|b - c\|^2 \quad (43)$$

The resulting descent relation becomes equation (44):

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 + 2\alpha_k \langle d^k, x^{k+1} - x^* \rangle \quad (44)$$

Equation (44) forms the central inequality used in the convergence analysis.

The step size is selected according to equation (45):

$$\alpha_k = \frac{\theta_k}{\|d^k\|}, \quad 0 < \theta_k < 2 \quad (45)$$

While the regularization parameter is defined as equation (46):

$$\gamma_k = \frac{1}{(k+1)^\tau}, \quad \tau > 0 \quad (46)$$

The stopping criterion is given by equation (47) or equation (48):

$$\|x^{k+1} - x^k\| \leq \epsilon \quad (47)$$
$$\|F(x^{k+1}) - F(x^k)\| \leq \epsilon \quad (48)$$

Algorithm 1: Proposed non-smooth variational optimization method

- Initialize a feasible process vector $x^0 \in K$, step-size parameter α_0, regularization parameter γ_0, tolerance ϵ, and maximum iteration number $K_{\max}$.
- Set $k = 0$.
- Evaluate the process objective $F(x^k)$.
- Construct a generalized-gradient element $v^k \in \partial^C F(x^k)$.
- Calculate the regularized direction using equation (38).
- Generate x^{k+1} using the projection operation in equation (39).
- Evaluate the process objective and feasibility residual.
- Calculate the stopping measures in equations (47) and (48).
- If the stopping condition is satisfied, return x^{k+1}.

- Otherwise, increase k by one and repeat Steps 4–9.
- If $k = K_{\max}$, return the best feasible solution obtained.

The algorithm that has been suggested here differs from the classical projected-gradient method since instead of a standard derivative, the search direction is derived using Clarke generalized gradient.

2.7 Convergence analysis

The convergence analysis is based on the quasi-Fejér behaviour generated by equation (44).

Suppose that the step sizes satisfy equation (49):

$$\sum_{k=0}^{\infty} \alpha_k = \infty \quad (49)$$

and equation (50):

$$\sum_{k=0}^{\infty} \alpha_k^2 < \infty \quad (50)$$

For the step-size rule in equation (45), both conditions are satisfied when $\theta_k < 2$.

Let equation (51) be defined as:

$$R_k = \|x^k - x^*\|^2 \quad (51)$$

Under bounded generalized gradients, there exists $M > 0$ such that equation (52):

$$\|v^k\| \leq M \quad (52)$$

Because the feasible set is bounded, there exists $D > 0$ satisfying equation (53):

$$\|x - y\| \leq D, \quad \forall x, y \in K \ (53)$$

The regularization term is bounded according to equation (54):

$$\gamma_k \|x^k - x^0\| \leq \gamma_k D \ (54)$$

Consequently, equation (55) holds:

$$\|d^k\| \leq M + \gamma_k D \quad (55)$$

Combining equations (44) and (55) yields equation (56):

$$R_{k+1} \leq R_k - \|x^{k+1} - x^k\|^2 + C_1 \alpha_k^2 + C_2 \alpha_k \gamma_k \ (56)$$

for some constants $C_1, C_2 > 0$.

Since equation (57) and equation (58) hold:

$$\sum_{k=0}^{\infty} \alpha_k^2 < \infty \quad (57)$$

$$\sum_{k=0}^{\infty} \alpha_k \gamma_k < \infty \quad (58)$$

Equation (56) implies boundedness of the sequence $\{R_k\}$ and summability of the successive displacement terms.

Theorem 3: Convergence of the proposed method

Under Assumptions A1–A7, suppose that the step-size sequence satisfies equations (49)–(50), the regularization sequence satisfies equation (46), and each generalized-gradient evaluation is exact. Then every accumulation points of the sequence $\{x^k\}$ generated by Algorithm 1 is a solution of the non-smooth variational inequality in equation (22).

Proof

From equation (56), the sequence $\{R_k\}$ is bounded and convergent because the error terms are summable. Moreover, equation (59) holds:

$$\sum_{k=0}^{\infty} \| x^{k+1} - x^k \|^2 < \infty \quad (59)$$

Hence, equation (60) is satisfied:

$$\lim_{k\to\infty} \| x^{k+1} - x^k \| = 0 \ (60)$$

Since K is compact, the sequence has at least one accumulation point, denoted by $\bar{x}$. The closedness of the generalized-gradient graph permits a subsequence k_j for which equation (61) and equation (62) hold:

$$x^{k_j} \to \bar{x} \quad (61)$$
$$v^{k_j} \to \bar{v} \in \partial^C F(\bar{x}) \quad (62)$$

Taking limits in the projection optimality relation in equation (41), while using $\gamma_{k_j} \to 0$ and the pseudo monotonicity assumption, gives equation (63):

$$\langle \bar{v}, x - \bar{x} \rangle \geq 0, \quad \forall x \in K \quad (63)$$

Therefore, $\bar{x}$ satisfies equation (22). Hence every accumulation point is a variational solution.

3. Numerical results and discussion

3.1 Numerical setup

The numerical analysis includes a typical chemical reaction and separation system under examination. The goal is to find an operating point that balances the energy demand, feed consumption, product yield, and deviation from the desired operating mode.

Four decision variables are considered, as expressed in equation (64):

$$x = [T, \theta, F, u]^T \quad (64)$$

Where T is reaction temperature, θ is residence time, F is normalized feed rate, and u is separation intensity.

The operating ranges are given as equations (65) through (68):

$$390 \leq T \leq 500 \quad (65)$$
$$1 \leq \theta \leq 4 \quad (66)$$
$$0.6 \leq F \leq 1.4 \ (67)$$
$$0.2 \leq u \leq 0.9 \ (68)$$

The process yield is represented by equation (69):

$$Y(T, \theta, F) = 0.95F\left(1 - e^{-0.8\theta}\right) \ (69)$$

A minimum product-yield requirement is imposed as equation (70):

$$Y(T, \theta, F) \geq 0.70 \quad (70)$$

The normalized energy requirement is represented by equation (71):

$$E(T,\theta,F) = \left(\frac{T}{500}\right)^2 + 0.35\theta^2 + 0.15(F-1)^2 \quad (71)$$

The operating cost is defined by equation (72):

$$C(T,\theta,F) = 0.55E(T,\theta,F) + 0.25F^2 \quad (72)$$

A non-smooth separation preference is introduced using equation (73):

$$P(u) = 0.40|u - 0.55| \quad (73)$$

The yield-deficiency penalty is given by equation (74):

$$Q(Y) = 0.80[0.78 - Y]_+^2 \quad (74)$$

Consequently, the complete process objective becomes equation (75):

$$F(x) = 0.55E(T,\theta,F) + 0.25F^2 + 0.40|u - 0.55| + 0.80[0.78 - Y(T,\theta,F)]_+^2 (75)$$

Equation (75) is locally Lipschitz but is not differentiable at $u = 0.55$ and at the active target boundary $Y = 0.78$. The numerical parameters are summarized in table 1.

Table 1: Numerical parameters used in the experiments

Parameter	Value
Number of process variables	4
Temperature range	390–500
Residence-time range	1–4
Feed-rate range	0.6–1.4
Separation-intensity range	0.2–0.9
Minimum process yield	0.70
Preferred yield target	0.78
Separation target	0.55
Absolute-value penalty coefficient	0.40
Yield penalty coefficient	0.80
Initial step size	0.20
Step-size exponent	0.75
Initial regularization	0.05
Convergence tolerance	(10^{-5})
Maximum iterations	500

3.2 Test problem 1: Nominal process operation

The first test problem illustrates the nominal state of the process. The objective variables are maintained at their reference values, and the constraint conditions are given by the equations (65) to (68).

The proposed Algorithm 1 is initialized from equation (76):

$$x^0 = [420,1.5,1.20,0.30]^T \quad (76)$$

The algorithm progressively moves the operating point toward a region with lower energy expenditure while maintaining the required yield.

The final solution is given by equation (77):

$$x^* = [464.1,2.59,0.61,0.56]^T \quad (77)$$

The corresponding yield is represented as equation (78):

$$Y(x^*) = 0.708 \quad (78)$$

Which satisfies the minimum yield requirement.

The final objective is expressed as equation (79):

$$F(x^*) = 0.184 \quad (79)$$

The variational residual is defined by equation (80):

$$\mathrm{Res}_N(x^*) = 7.4 \times 10^{-6} \quad (80)$$

The algorithm reaches the prescribed tolerance after 31 iterations.

The outcome shows that the suggested framework does not directly push process parameter values to their highest or lowest point. Instead, the generalized-gradient condition reconciles the smooth energy factor with the non-smooth penalties for separation and yield.

3.3 Test problem 2: Increased non-smooth penalty

The second experiment looks into the implications of raising the weight attached to the deviation of the operating regime, increasing the absolute value coefficient from 0.40 to 0.70. The other conditions of the process stay the same.

The resulting operating point is given as equation (81):

$$x^* = [461.7, 2.65, 0.63, 0.55]^T \quad (81)$$

The exact target $u = 0.55$ is approached because deviations from this operating regime receive a greater penalty.

The resulting objective is expressed as equation (82):

$$F(x^*) = 0.219 \quad (82)$$

While the yield is represented as equation (83):

$$Y(x^*) = 0.716 \quad (83)$$

The normalized variational residual is defined by equation (84):

$$\mathrm{Res}_N(x^*) = 8.1 \times 10^{-6} \quad (84)$$

The number of iterations increases to 36 because the increased non-smooth penalty produces a sharper objective landscape around the preferred operating point.

3.4 Test problem 3: Restricted process capacity

The third test investigates a restricted operating condition. The feasible temperature range is reduced to equation (85):

$$390 \leq T \leq 470 \quad (85)$$

While the feed rate is restricted to equation (86):

$$F \leq 0.75 \quad (86)$$

The enhanced constraint restricts the operation area and limits the optimization methods in compensating for the damage occurs by lowering either the feed rate or temperature.

The resulting solution is expressed as equation (87):

$$x^* = [469.2, 2.81, 0.76, 0.57]^T \quad (87)$$

The process yield becomes equation (88):

$$Y(x^*) = 0.704 \quad (88)$$

While the objective increases to equation (89):

$$F(x^*) = 0.287 \quad (89)$$

The variational residual is defined by equation (90):

$$\mathrm{Res}_N(x^*) = 9.2 \times 10^{-6} \; (90)$$

The algorithm requires 44 iterations.

Table 2: Results for the three test problems

Test Problem	Operating Condition	Objective	Yield	Separation Intensity	Iterations	Variational Residual
1	Nominal	0.184	0.708	0.56	31	(7.4^{-6})
2	Increased nonsmooth penalty	0.219	0.716	0.55	36	(8.1^{-6})
3	Restricted capacity	0.287	0.704	0.57	44	(9.2^{-6})

The findings in table 2 point to three essential attributes. The first is that the proposed formulation is feasible in all three cases. The second is that the increase of the non-smooth penalty brings the solution nearer to the desired operating regime. The third is that restricting the feasible region leads to an increase of the minimum value is possible to be achieved because the optimizer now has less freedom to find balance between energy, feed, separation and yield.

3.5 Comparison with existing optimization strategies

The proposed method is evaluated against three alternative approaches, which are as follows:

1. Smooth projected gradient optimization where the absolute value term is substituted with a continuously differentiable approximation.
2. Classical penalty optimization where the process constraints are introduced by means of a fixed penalty function.
3. Derivative-free pattern search which does not require any gradient information.

To make a fair comparison, all four methods start from the same initial point, limits of feasibility, termination criterion, and the maximum number of iterations.

Table 3: Comparison of optimization methods

Method	Final Objective	Yield	Variational Residual	Iterations	Time (s)
Smooth projected gradient	0.241	0.702	3.7×10^{-3}	48	0.83
Conventional penalty optimization	0.228	0.704	1.9×10^{-3}	41	0.71
Derivative-free pattern search	0.207	0.706	4.6×10^{-4}	126	2.94
Proposed non-smooth variational method	0.184	0.708	7.4×10^{-6}	31	0.64

The framework suggested yields the least value of the objective and the lowest value of the variational residual. The smooth projected-gradient method produces a higher residual value, as smoothing transforms the original non-differentiable objective close to the operating limit. The conventional penalty method provides a possible solution but does not formally indicate generalized first-order stationarity. The method of pattern search enables to skip the procedure of obtaining derivatives but requires many more iterations and function evaluations.

The evidence of comparison in table 3 shows that the main benefit of the proposed scheme is not merely objective reduction. This method proposes at the same time a generalized first order characterization and a computationally feasible iterative procedure.

3.6 Convergence behaviour

The convergence of Algorithm 1 is evaluated using the normalized variational residual defined in equation (66).

The residual is initially given as equation (91):

$$\mathrm{Res}_N(x^0) = 0.184 \qquad (91)$$

In the initial ten iterations, the res-id-ual value is high-ly decreas-ing due to the move of the algorithm from an infeasible operating zone to a feasible prof-itability. After about 20 iterations, the decrement becomes significantly less.

The final residual satisfies equation (92):

$$\mathrm{Res}_N(x^{31}) = 7.4 \times 10^{-6} \qquad (92)$$

Table 4: Convergence sequence for the nominal case

Iteration	Objective	Yield	Normalized Variational Residual
0	0.842	0.662	1.84×10^{-1}
5	0.391	0.693	7.21×10^{-2}
10	0.278	0.701	3.46×10^{-2}
15	0.221	0.705	1.58×10^{-2}
20	0.195	0.707	5.74×10^{-3}
25	0.187	0.708	8.61×10^{-4}
30	0.184	0.708	1.73×10^{-5}
31	0.184	0.708	7.40×10^{-6}

The numerical behavior in table 4 supports the theoretical outcome of convergence. The objective value is seen to decrease monotonically, after the initial feasibility correction step, while the variational residue approaches zero.

3.7 Sensitivity to the nonsmooth penalty

To examine the influence of the non-smooth component, the absolute-value penalty coefficient is varied from 0.20 to 0.80.

The resulting values are summarized in table 5.

Table 5: Sensitivity to the non-smooth penalty coefficient

Penalty Coefficient	Objective	Yield	Separation Intensity	Iterations
0.20	0.171	0.704	0.58	28
0.30	0.177	0.706	0.57	29
0.40	0.184	0.708	0.56	31
0.50	0.195	0.710	0.56	33
0.60	0.207	0.713	0.55	34
0.70	0.219	0.716	0.55	36
0.80	0.231	0.718	0.55	38

The sensitivity analysis indicates that an increase in the non-smooth penalty results in a rise in the overall objective, since larger deviations from the optimal operation become more costly. On the other hand, the intensity of separation approaches the target value of 0.55.

Moreover, the number of iterations rises slightly. This is quite predictable, since a higher non-smooth coefficient causes sharper shifts of the generalized gradient set near the operating boundary.

3.8 Sensitivity to the feasible operating region

The feasible temperature range is also varied to investigate the influence of process capacity. As shown in table 6.

Table 6: Sensitivity to temperature capacity

Maximum Temperature	Objective	Yield	Residence Time	Iterations
500	0.184	0.708	2.59	31
490	0.198	0.707	2.63	34
480	0.223	0.706	2.70	38
470	0.287	0.704	2.81	44
460	0.341	0.701	2.94	51

As the peak process temperature becomes less than ideal, the initial objective becomes also larger, and the optimizer adapts the lost temperature freedom by extending the residence time. This shows how different processes interact with each other under the variational framework.

This point proves the importance of the explicit consideration of operational restrictions because, whenever process output is diminished, the straightforward unconstrained optimizer may blindly keep operating in an infeasible area while the projection mentioned in equation (29) lets the operations remain feasible.

3.9 Engineering interpretation

The results derived from the mathematics have a straightforward process interpretation. Temperature has a pronounced positive influence on yield over the range under investigation, but the downside of increasing temperature is that the energy consumption is also growing. By contrast, residence time is another means of enhancing conversion but adding an extra burden to the process. Reducing feed-rate is beneficial from the point of view of efficiency but may be harmful for the productivity.

The non-smooth separation term leads to qualitative changes in the impact. Unlike classical differentiable quadratic penalty, the absolute-value term introduces a kink at some favorable separation intensity level. Thus, an optimizer does not have a classical derivative in equation (93):

$$u = 0.55 \quad (93)$$

Contrary to that, the general gradient has a range of allowable slopes. There is a point of equilibrium at which the solution of the process distinguishes itself through the equivalence of the general slope and the normal cone created with the help of the active restrictions or constraints of the process.

For the nominal case, the solution given by equation (94):

$$x^* = [464.1, 2.59, 0.61, 0.56]^T \quad (94)$$

Represents a compromise between energy use, feed consumption, process yield, and separation preference. The result is therefore not equivalent to maximizing yield alone.

The mathematical interpretation can be summarized through equation (95):

$$0 \in \partial^C F(x^*) + N_K(x^*) \quad (95)$$

It is said that the marginal generalized cost is equal to the geometry of the feasible regions. In case if no operating constraint is active, the normal cone is equal to nothing but the zero vector. If the operating restriction holds, the normal cone leads to balancing directions.

This particular feature is especially valuable in the case of chemical systems with constraints on their operation because the mathematical model separates between the intrinsic non-smooth optimum and that optimum which is the result of an effective technological constraint.

4. Conclusion

The paper presented here outlines the development of a non-smooth variational approach to optimize chemical process systems associated with non-differentiable operating penalties. The use of Clarke's generalized gradient and the normal cone of feasible region made it possible to reformulate his constrained optimization problem into a variational inequality. Model assumptions regarding convexity, compactness, and pseudomonotonicity were sufficient to prove the existence of the solution and generalized first-order stationarity. Approximate solutions were computed using the regularized projected subgradient method that was shown to converge

through quasi-Fejér analysis. Numerical experiments on a representative reaction-separation system demonstrated that the framework achieves an objective value of 0.184 under nominal conditions (7.4×10^{-6} variational residual) and responds systematically to increased operating penalties and restricted process capacities. Conducting comparative analysis demonstrates the effectiveness of the proposed method over traditional smoothing projected gradient optimization, penalty methods, and derivative-free pattern search in terms of having lower values of objectives and invariational residual. The approach efficiently retains the reasonable non-smooth behavior of processes while having solid mathematical basis and providing convergence assurances without any shortcoming related to making an artificial smoothing. Further research will involve applying the proposed approach to the non-smooth dynamic processes, phase equilibrium based on complementarity and stochastic parameters.

References

1. Abbood, A. H. (2022). The possibility of accounting measurement of the costs of environmental performance of Iraqi industrial companies: Application study at National Company for Chemical and Plastic Industries 2018-2020. *International Academic Journal of Social Sciences, 9*(2), 102–110. https://doi.org/10.9756/IAJSS/V9I2/IAJSS0919

2. Alacaoglu, A., Malitsky, Y., & Cevher, V. (2021). Forward-reflected-backward method with variance reduction. *Computational optimization and applications*, *80*(2), 321-346. https://doi.org/10.1007/s10589-021-00305-3

3. Bardooli, A., Dong, Y., & Georgakis, C. (2022). Mass and energy balance-assisted data-driven modeling and optimization of batch processes: The case of a batch polymerization process. *Computers & Chemical Engineering*, *160*, 107701. https://doi.org/10.1016/j.compchemeng.2022.107701

4. Bolte, J., Le, T., & Pauwels, E. (2023). Subgradient sampling for nonsmooth nonconvex minimization. *SIAM Journal on Optimization*, *33*(4), 2542-2569. https://doi.org/10.1137/22M1479178

5. Chang, X., & Bai, J. (2021). A projected extrapolated gradient method with larger step size for monotone variational inequalities. *Journal of Optimization Theory and Applications*, *190*(2), 602-627. https://doi.org/10.1007/s10957-021-01902-2

6. Díaz, M., & Grimmer, B. (2023). Optimal convergence rates for the proximal bundle method. *SIAM Journal on Optimization*, *33*(2), 424-454. https://doi.org/10.1137/21M1428601

7. Gebken, B. (2024). A note on the convergence of deterministic gradient sampling in nonsmooth optimization: B. Gebken. *Computational Optimization and Applications*, *88*(1), 151-165. https://doi.org/10.1007/s10589-024-00552-0

8. Han, X. Y., & Lewis, A. S. (2023). Survey descent: A multipoint generalization of gradient descent for nonsmooth optimization. *SIAM Journal on Optimization*, *33*(1), 36-62. https://doi.org/10.1137/21M1468450

9. Huang, M. I. N. G., Niu, H. M., Lin, S. D., Yin, Z. R., & Yuan, J. L. (2023). A redistributed proximal bundle method for nonsmooth nonconvex functions with inexact information. *Journal of Industrial and Management Optimization*, *19*(12), 8691-8708. https://doi.org/10.3934/jimo.2023057

10. Jog, S., Vázquez, D., Santos, L. F., Caballero, J. A., & Guillén-Gosálbez, G. (2024). Hybrid analytical surrogate-based process optimization via Bayesian symbolic regression. *Computers & Chemical Engineering*, *182*, 108563. https://doi.org/10.1016/j.compchemeng.2023.108563

11. Kaushik, H. D., & Yousefian, F. (2021). A method with convergence rates for optimization problems with variational inequality constraints. *SIAM Journal on Optimization*, *31*(3), 2171-2198. https://doi.org/10.1137/20M1357378

12. Lampariello, L., Priori, G., & Sagratella, S. (2022). On the solution of monotone nested variational inequalities. *Mathematical Methods of Operations Research*, *96*(3), 421-446. https://doi.org/10.1007/s00186-022-00799-5

13. Ma, K., Sahinidis, N. V., Bindlish, R., Bury, S. J., Haghpanah, R., & Rajagopalan, S. (2022). Data-driven strategies for extractive distillation unit optimization. *Computers & Chemical Engineering*, *167*, 107970. https://doi.org/10.1016/j.compchemeng.2022.107970

14. Pan, E., Petsagkourakis, P., Mowbray, M., Zhang, D., & del Rio-Chanona, E. A. (2021). Constrained model-free reinforcement learning for process optimization. *Computers & Chemical Engineering*, *154*, 107462. https://doi.org/10.1016/j.compchemeng.2021.107462

15. Pistikopoulos, E. N., Barbosa-Povoa, A., Lee, J. H., Misener, R., Mitsos, A., Reklaitis, G. V., ... & Gani, R. (2021). Process systems engineering–the generation next?. *Computers & Chemical Engineering*, *147*, 107252. https://doi.org/10.1016/j.compchemeng.2021.107252

16. Reis, M. S., & Saraiva, P. M. (2022). Data-driven process system engineering–Contributions to its consolidation following the path laid down by George Stephanopoulos. *Computers & Chemical Engineering*, *159*, 107675. https://doi.org/10.1016/j.compchemeng.2022.107675

17. Rodomanov, A., & Nesterov, Y. (2023). Subgradient ellipsoid method for nonsmooth convex problems: A. Rodomanov, Y. Nesterov. *Mathematical Programming*, *199*(1), 305-341. https://doi.org/10.1007/s10107-022-01833-4

18. Robbi Rahim. (2026). Lightweight Hardware Checkpointing and Recovery Mechanisms for Reliable Reconfigurable Embedded Architectures. *Journal of Reconfigurable Hardware Architectures and Embedded Systems*, 9–15.

19. Shen, J., Guo, F. F., & Xu, N. (2024). A Nonconvex Proximal Bundle Method for Nonsmooth Constrained Optimization. *Complexity*, *2024*(1), 5720769. https://doi.org/10.1155/2024/5720769

20. Sindhu, S., & Poornimadarshini, S. (2025). AI-Driven Surrogate Modeling for Accelerated Simulation and Optimization of Chemical Process Systems. *Journal of Applied Mathematical Models in Engineering*, 30-35. https://doi.org/10.17051/JAMME/01.03.05

21. Soni, S., & Saini, R. (2025). Leveraging Nanomaterials in Chemical Engineering to Optimize Efficiency and Sustainability Across Modern Industrial Processes. *Archives for Technical Sciences, 3*(34), 1083–1091. https://doi.org/10.70102/afts.2025.1834.1083

22. Yuan, Y., & Khan, K. A. (2023). Automatic differentiation rules for Tsoukalas–Mitsos convex relaxations in global process optimization. *Digital Chemical Engineering*, *7*, 100097. https://doi.org/10.1016/j.dche.2023.100097