

A non-Convex Optimization Approach for Process Parameter Control Using Variational Analysis

Danish Kundra¹, B. Selva Babu², Dr. Sameer Singh³, Jyoti Mohur⁴, Ashutosh Kulkarni⁵,
Dr. Chetanaba G. Rajput⁶, R. Arivukkodi⁷

¹Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India. danish.kundra.orp@chitkara.edu.in,
<https://orcid.org/0009-0004-1739-8219>

²Assistant Professor - Gr. II, Department of Mechanical Engineering, Aarupadai Veedu Institute of Technology, Vinayaka Mission's Research Foundation (DU),
Tamil Nadu, India. selvababu@avit.ac.in, <https://orcid.org/0000-0003-4172-7882>

³Rungta College of Engineering and Technology, Rungta International Skills University, Bhilai, Chhattisgarh, India. sameer.singh@rungta.org

⁴Assistant Professor, Department of Computer Science & Engineering, Noida International University, Greater Noida, Uttar Pradesh, India. jyoti.mahur@niu.edu.in

⁵Department of DESH, Vishwakarma Institute of Technology, Pune, Maharashtra, India.
ashutosh.kulkarni@vit.edu

⁶Assistant Professor, Faculty of Arts, Gokul Global University, Sidhpur, Gujarat, India.
ad.admin@gokuluniversity.ac.in, <https://orcid.org/0009-0006-2894-1234>

⁷Assistant Professor, Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India.
arivukodir@maher.ac.in

Abstract

The problem of controlling process parameters in case of non-convex processes represents a mathematically difficult one since the process operator can ignore global monotonicity while the feasible domain of the process parameters may have a non-convex structure. In this paper, an approach to solving such problems based on variational optimization is developed. The process-control problem is formulated as a non-convex variational inequality and equivalently represented through the inclusion $0 \in F(x^*) + N_K(x^*)$. To eliminate the ambiguity of the projection into a non-convex set, the feasible set is supposed to be locally r -prox-regular, thus guaranteeing a unique metric projection along the optimization path. The weak monotonicity constant for the function is calculated from the symmetric Jacobian, hence the initial regularization coefficient is defined. The proposed method introduces $F_\mu(x) = F(x) + \mu(x - x_0)$ and progressively reduces μ toward zero. Lyapunov descent is used to prove that a certain type of resummability is attained and that the accumulation points converge to the state of variational stationarity, under the described local regularity assumptions. The results are confirmed by numerical tests, conducted for a nonlinear three-parameter process model, where a convergence for different starting points is achieved with the variational gap being less than 10^{-6} . The locally minimal stationary states are separated from other stationary points based on the analysis of the reduced Hessian. The unregularized projected gradient stress benchmark shows improved stability.

Keywords: Non-convex optimization, variational inequality, prox-regularity, hypomonotonicity, process parameter control, regularization.

1. Introduction

Non-convex optimization offers a natural way to mathematically formulate process parameter control in cases where the correlation between the operating parameters and the process performance is non-linear, coupled, and

multi-feasible. In real-life process systems, interactivity among parameters will generate multiple local stationary points, whereas physical constraints on operation might give rise to feasible sets that do not satisfy global convexity assumptions. These two issues make the classical convex optimization approach infeasible, and hence the necessity of mathematical models to address the problem of non-convexity directly (Jain & Kar, 2017; Danilova et al., 2022).

Variational analysis offers a solid mathematical basis for this class of optimization problems via concepts such as tangent cone, normal cone, variational inequalities, and generalized stationarity conditions. The variational inequality approach, in particular, allows for formulating optimization and equilibrium problems in terms of an operator defined over a feasible set (El Idrissi et al., 2024). In addition, normal cone formulations offer a compact way to represent first-order optimality conditions and enable solving constrained problems without having to convert them into unconstrained formulations (Sherjonov et al., 2025). Nonetheless, most of the classic variational optimization theory is developed under convexity and monotonicity assumptions, making it less suitable for optimization of processes with non-convex objective functions and non-convex feasible regions (Bassily et al., 2021).

There is therefore a considerable gap in the development of variational optimization approaches that would take into account non-convex feasible sets, non-monotone operators, and numerically stable solution algorithms (Alhazmi, 2026; Csetnek, 2020). Conventional gradient-based methods might converge to local stationary points without giving a variational formulation of the problem. At the same time, standard variational inequality approaches are often based on the monotonicity and convexity assumptions. Weak monotonicity is a less restrictive assumption compared to the former and allows for a bounded level of non-monotonicity (Häberle et al., 2020). Nonetheless, an appropriate regularization and projection technique is needed to take advantage of this property in the context of a non-convex process control problem (Wei & Xie, 2023; Bianchi & Jakubowicz, 2012; Wang et al., 2023). Regularization offers a way to enhance the mathematical stability of ill-posed or difficult optimization problems, but it is yet to be incorporated efficiently in the non-convex variational process optimization setting (Grammatico et al., 2015; Liu et al., 2019; Bianchi et al., 2022; Anantrasirichai et al., 2020).

This gap is being addressed in this study by formulating the process parameter control problem as a variational inclusion

$$0 \in F(x^*) + N_K(x^*),$$

where $x \in \mathbb{R}^n$ denotes the process-parameter vector, K represents the potentially non-convex feasible region, F is the process-performance operator, and $N_K(x^*)$ is the normal cone of K . A regularized operator is used to handle weak non-monotonicity, and then an iterative projection scheme is applied to find the solution of stationary solutions. The suggested formulation is studied through variational gap functions and stationarity residuals, and convergence analysis of the algorithm (Shirke & Udayakumar, 2019). Engineering process model is used just to illustrate mathematically the suggested formulation.

Contributions of this work can be outlined in four points. Firstly, a non-convex variational formulation is constructed for process parameter optimization under constraints. Secondly, a concept of weakly monotone operators is introduced as a relaxation of the classical monotonicity condition. Thirdly, a regularized projected variational iteration is formulated for finding stationary solutions in the non-convex feasible set. Fourthly, a study on convergence to variationally stationary points is conducted under certain regularity and boundedness conditions and step-size selection rules. Hence, the primary contribution is mathematical, and the process parameter control problem is considered as an engineering test bench.

This paper is organized as follows. Mathematical background, including non-convex feasible set, process objective functional, variational inequality, tangent and normal cones, and variational gap function, is introduced in Section 2. Non-convex variational formulation employing weak monotonicity and regularization as well as the corresponding stationarity condition is formulated in Section 3. A regularized projected variational approach is proposed and analyzed in Section 4. Numerical analysis, including convergence, sensitivity, and stationary solution computation for a nonlinear engineering process parameter model, is provided in Section 5. Conclusions and mathematical extensions are summarized in Section 6.

2. Mathematical preliminaries and problem formulation

This section sets out the mathematical setting necessary to describe the problem of process parameters control as a constrained non-convex optimization problem (Luedke & Kingdone, 2025; Suvarna & Bharadwaj, 2024). This description employs a finite-dimensional space of parameters, a possibly non-convex feasible set, a nonlinear

functional describing the performance of the process, and the variational inequality form. Definitions introduced in this section are prerequisites to developing the regularized variational optimization approach in following sections.

2.1 Process parameter space and feasible region

Let the process parameter vector be represented by

$$x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n,$$

where x_i denotes an adjustable process parameter. Depending on the engineering process, these parameters may represent temperature, pressure, feed rate, rotational speed, residence time, or other operating variables.

The admissible parameter domain is defined as

$$K = \{x \in \mathbb{R}^n : g_j(x) \leq 0, j = 1, \dots, m\}, \quad (1)$$

where $g_j: \mathbb{R}^n \rightarrow \mathbb{R}$ represents a nonlinear process constraint. Simple parameter bounds can additionally be expressed as

$$l_i \leq x_i \leq u_i, \quad i = 1, \dots, n.$$

The feasible set K is assumed to be nonempty and closed but is not required to be convex. Consequently, there may exist $x, y \in K$ and $\lambda \in (0,1)$ such that

$$\lambda x + (1 - \lambda)y \notin K.$$

This assumption distinguishes the proposed formulation from conventional convex process optimization.

2.2 Process performance functional

Let

$$\Phi: K \rightarrow \mathbb{R}$$

denote the process-performance functional. It represents the combined effect of process quality, energy consumption, operating cost, deviation from desired operating conditions, or other engineering objectives.

A general weighted formulation can be expressed as

$$\Phi(x) = \sum_{r=1}^q w_r \Phi_r(x), \quad w_r \geq 0, \quad \sum_{r=1}^q w_r = 1. \quad (2)$$

Here, $\Phi_r(x)$ denotes an individual process objective and w_r represents its relative contribution.

The resulting constrained optimization problem is

$$\boxed{\min_{x \in K} \Phi(x)}. \quad (3)$$

The objective is considered non-convex when there exist $x, y \in K$ and $\lambda \in (0,1)$ such that

$$\Phi(\lambda x + (1 - \lambda)y) > \lambda \Phi(x) + (1 - \lambda)\Phi(y).$$

For a twice differentiable objective, global convexity is not guaranteed when

$$\nabla^2 \Phi(x) \geq 0$$

does not hold throughout K . Therefore, the optimization problem in (3) may contain multiple local stationary points.

2.3 Process operator

For the variational formulation, define an operator

$$F: K \rightarrow \mathbb{R}^n.$$

When the process functional is differentiable, the natural choice is

$$F(x) = \nabla \Phi(x).$$

Thus,

$$F(x) = \begin{bmatrix} \frac{\partial \Phi}{\partial x_1} \\ \frac{\partial \Phi}{\partial x_2} \\ \vdots \\ \frac{\partial \Phi}{\partial x_n} \end{bmatrix}.$$

The operator formulation is more general than directly treating the problem through the objective function because it can accommodate process interactions that may not have a simple scalar-potential representation.

For $x, y \in K$, a classical monotone operator satisfies

$$\langle F(x) - F(y), x - y \rangle \geq 0.$$

The proposed non-convex formulation does not require this condition globally.

2.4 Weak monotonicity

To characterize controlled non-monotonicity, F is assumed to satisfy the weak monotonicity condition

$$\langle F(x) - F(y), x - y \rangle \geq -\nu \|x - y\|^2, \quad (4)$$

for some $\nu \geq 0$ and all x, y within the region of analysis.

When $\nu = 0$, (4) reduces to the conventional monotonicity condition. For $\nu > 0$, the operator may be non-monotone, but its departure from monotonicity is bounded by ν . This assumption is appropriate for nonlinear process models in which interacting parameters can produce locally adverse curvature or multiple stationary regions.

2.5 Variational inequality formulation

The finite-dimensional variational inequality associated with the process parameter problem is defined as follows.

Definition 1. Find $x^* \in K$ such that

$$\langle F(x^*), x - x^* \rangle \geq 0, \quad \forall x \in K. \quad (5)$$

The solution is denoted by

$$x^* \in \text{VI}(F, K).$$

If $F = \nabla \Phi$, (5) represents the first-order variational characterization of a constrained stationary point. For a convex feasible region and convex differentiable objective, this condition can characterize a global minimum. In the present study, however, K and Φ may be non-convex. Consequently, the variational inequality is interpreted primarily as a stationarity condition rather than automatically as a global optimality condition.

2.6 Tangent cone

For a feasible point $x \in K$, the tangent cone describes feasible first-order directions. It is defined by

$$T_K(x) = \left\{ d \in \mathbb{R}^n : \exists t_k \downarrow 0, x_k \in K, \frac{x_k - x}{t_k} \rightarrow d \right\}.$$

A vector $d \in T_K(x)$ therefore represents a direction along which the process parameters can be perturbed while remaining asymptotically feasible.

A necessary first-order condition for a local minimum x^* is

$$\boxed{\langle \nabla \Phi(x^*), d \rangle \geq 0, \quad \forall d \in T_K(x^*)}. \quad (6)$$

Equation (6) provides the local variational interpretation of process parameter stationarity.

2.7 Normal cone

The normal cone associated with K at $x \in K$ is defined as

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

The relationship between the tangent and normal cones can be expressed through the polar operation as

$$N_K(x) = T_K(x)^\circ.$$

Using the normal cone, the variational inequality in (5) can be equivalently represented by the inclusion

$$\boxed{0 \in F(x^*) + N_K(x^*)}. \quad (7)$$

For the differentiable objective case,

$$\boxed{0 \in \nabla \Phi(x^*) + N_K(x^*)}. \quad (8)$$

Equation (8) represents the primary first order stationarity relation used in the following developments. In the non-convex case, it defines the variational stationary point without guaranteeing the latter to be the global minimum point.

2.8 Variational gap function

To quantify the deviation of an arbitrary feasible point from a variational solution, define the gap function as

$$\boxed{G(x) = \sup_{y \in K} \{ \langle F(x), x - y \rangle \}}. \quad (9)$$

For a variational solution x^* , the gap satisfies

$$G(x^*) = 0.$$

Therefore, during numerical optimization, the quantity $G(x_k)$ can be used to monitor the approach of the generated sequence toward variational stationarity.

For numerical implementation, a normalized gap can additionally be defined as

$$\boxed{G_N(x_k) = \frac{G(x_k)}{\max\{I, \|F(x_k)\|\}}}. \quad (10)$$

The normalized form reduces the influence of the magnitude of the operator and provides a more consistent convergence indicator across different parameter scales.

2.9 Regularized operator

Because the process operator may be weakly monotone rather than monotone, introduce the regularized operator

$$F_\mu(x) = F(x) + \mu(x - x_0), \quad \mu > 0, \quad (11)$$

where $x_0 \in K$ is an initial feasible point.

For $x, y \in K$,

Hence, whenever

$$\mu > \nu,$$

the regularized operator possesses a strong monotonicity property with modulus

$$\mu - \nu > 0 \quad (12)$$

This property forms the mathematical basis for stabilizing the iterative procedure developed in Section 4.

2.10 Regularized problem formulation

The regularized variational inequality is defined as finding $x_\mu \in K$ such that

$$\langle F(x_\mu) + \mu(x_\mu - x_0), x - x_\mu \rangle \geq 0, \quad \forall x \in K. \quad (13)$$

Equivalently, using the normal cone,

$$0 \in F(x_\mu) + \mu(x_\mu - x_0) + N_K(x_\mu). \quad (14)$$

The corresponding regularized gap function is

$$G_\mu(x) = \sup_{y \in K} \{ \langle F(x) + \mu(x - x_0), x - y \rangle \}. \quad (15)$$

A solution x_μ of the regularized variational inequality satisfies

$$G_\mu(x_\mu) = 0.$$

As the regularization parameter decreases,

$$\lim_{\mu \rightarrow 0^+} F_\mu(x) = F(x). \quad (16)$$

Consequently, under the required continuity and stability conditions, the regularized inclusion in (14) approaches the original non-convex variational condition

$$0 \in F(x^*) + N_K(x^*). \quad (17)$$

Therefore, the mathematical problem studied in this investigation is a solution of a series of regularized variational problems with gradual decrease in the regularization coefficient such that the obtained solution converges to a stationary point of the initial non-convex process optimization problem.

In further numerical analysis, the corresponding stationary criterion may be controlled using the residual

$$\mathcal{R}_k = \max \left\{ G_{\mu_k}(x_k), \frac{\|x_{k+1} - x_k\|}{\max\{1, \|x_k\|\}} \right\}. \quad (18)$$

$\mathcal{R}_k$ of a certain sufficiently small value would imply simultaneous decrease of the variational gap and the sequential parameter shift. Thus, equation (18) will serve as the main stopping rule for the suggested non-convex variational optimization technique discussed in Section 4.

3. Non-convex variational formulation

Parameter control of the process is stated as a non-convex variational inequality where both the performance process operator and the set of admissible parameters may have non-convex behavior. As projection onto an arbitrary non-convex set is always multi-valued, some extra geometrical conditions must be imposed in order to obtain a uniquely defined iteration scheme. Hence, it is assumed that the feasible region is r_K -prox-regular in the vicinity of the generated trajectory.

3.1 Local variational inequality

Let

$$K \subset \mathbb{R}^n$$

be a nonempty closed feasible parameter set and let

$$F: K \rightarrow \mathbb{R}^n$$

denote the process operator. A point $x^* \in K$ is a local variational solution if there exists $\rho > 0$ such that

$$\langle F(x^*), x - x^* \rangle \geq 0, \quad \forall x \in K \cap B(x^*, \rho). \quad (19)$$

For a differentiable process-performance functional,

$$F(x) = \nabla \Phi(x),$$

and therefore

$$\langle \nabla \Phi(x^*), x - x^* \rangle \geq 0. \quad (20)$$

Because K and Φ are not assumed to be convex, this condition represents local variational stationarity rather than global optimality.

3.2 Prox-regular feasible region

The non-convex projection issue is addressed by imposing a local geometric assumption on K .

Assumption 1. The feasible region K is r_K -prox-regular along the trajectory generated by the proposed method. That is, there exists $r_K > 0$ such that for every trajectory point $x \in K$, the metric projection

$$P_K(z) = \arg \min_{y \in K} |z - y| \quad (21)$$

is a singleton for all

$$z \in U_{r_K}(K) = \{z: \text{dist}(z, K) < r_K\}.$$

Under this assumption,

$$P_K: U_{r_K}(K) \rightarrow K$$

is single-valued and locally Lipschitz continuous.

This is less restrictive than global convexity and allows for non-convex sets to have sufficiently smooth boundaries. This also removes the arbitrariness involved in choosing an arbitrary element from disconnected projections when analyzing convergence.

3.3 Normal-cone stationarity

For a prox-regular set, the tangent and normal cone construction remains applicable locally. The normal cone is

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

under the regularity conditions considered here.

The variational inequality can therefore be expressed as

$$0 \in F(x^*) + N_K(x^*). \quad (22)$$

For $F = \nabla\Phi$,

$$0 \in \nabla\Phi(x^*) + N_K(x^*). \quad (23)$$

Equation (23) is the principal first-order stationarity condition used throughout the proposed method.

3.4 Local weak monotonicity

The weak-monotonicity constant introduced previously is treated as a local quantity, rather than an unrestricted global constant.

Let $D \subset K$ be a bounded neighborhood containing the optimization trajectory. Assume that there exists

$$\nu_D \geq 0$$

such that

$$\langle F(x) - F(y), x - y \rangle \geq -\nu_D |x - y|^2, \quad \forall x, y \in D. \quad (24)$$

Thus, ν_D is a local hypomonotonicity bound over the region actually explored by the algorithm.

For a continuously differentiable operator, a practical estimate can be obtained from the smallest eigenvalue of the symmetric Jacobian,

$$J_s(x) = \frac{J_F(x) + J_F(x)^T}{2}.$$

An empirical bound over sampled points $x^{(q)} \in D$ is

$$\hat{\nu}_D = \max_q \left[-\lambda_{\min} \left(J_s(x^{(q)}) \right) \right]_+. \quad (25)$$

A safety factor $\kappa_\nu > 1$ is then introduced and

$$\mu_0 = \kappa_\nu \hat{\nu}_D \quad (26)$$

is selected, with the practical requirement

$$\mu_0 > \nu_D.$$

This provides a reproducible method for selecting the initial regularization rather than treating ν as an unspecified tuning constant.

3.5 Regularized operator

Define

$$F_{\mu_k}(x) = F(x) + \mu_k(x - x_0). \quad (27)$$

Using (24),

Hence, during the phase satisfying

$$\mu_k > \nu_D, \quad (28)$$

the regularized operator is strongly monotone with modulus

$$m_k = \mu_k - \nu_D > 0. \quad (29)$$

3.6 Regularized variational problem

The regularized problem is

$$\langle F_{\mu_k}(x_{\mu_k}), x - x_{\mu_k} \rangle \geq 0, \quad \forall x \in K. \quad (30)$$

Equivalently,

$$0 \in F(x_{\mu_k}) + \mu_k(x_{\mu_k} - x_0) + N_K(x_{\mu_k}). \quad (31)$$

The regularization is gradually reduced so that the regularized stationary problem approaches the original inclusion (22).

3.7 Variational gap function

The regularized gap is defined as

$$G_{\mu_k}(x) = \sup_{y \in K} \{ \langle F_{\mu_k}(x), x - y \rangle \}. \quad (32)$$

A regularized variational solution satisfies

$$G_{\mu_k}(x_{\mu_k}) = 0. \quad (33)$$

The gap is therefore used together with the update residual to establish numerical stationarity.

3.8 First- and second-order stationarity

First-order stationarity does not distinguish between a local minimum, saddle point, or local maximum in a non-convex problem. Therefore, after obtaining a stationary solution x^* , a second-order screening procedure is applied.

Let $A(x^*)$ denote the matrix of active constraints and let

$$Z \in \mathbb{R}^{n \times d}$$

contain an orthonormal basis for the tangent subspace satisfying

$$A(x^*)Z = 0. \quad (34)$$

The reduced Hessian is

$$H_R(x^*) = Z^T \nabla^2 \Phi(x^*) Z. \quad (35)$$

A candidate stationary point is classified as a second-order local minimizer when

$$\lambda_{\min}(H_R(x^*)) > 0. \quad (36)$$

This post-processing step prevents a first-order stationary saddle point from being incorrectly reported as an optimal process configuration.

4. Proposed non-convex variational optimization method

The developed algorithm employs a regularized variational operator along with a unique local metric projection (Hadri et al., 2021). Prox-regularity, discussed in Section 3, guarantees the well-posedness of the projection operator in the neighborhood that includes the generated sequence.

4.1 Prox-regular projected variational iteration

The proposed update is

$$x_{k+1} = P_K[x_k - \alpha_k F_{\mu_k}(x_k)]. \quad (37)$$

Using (27),

$$x_{k+1} = P_K[x_k - \alpha_k \{F(x_k) + \mu_k(x_k - x_0)\}]. \quad (38)$$

Because K is assumed to be r_K -prox-regular, the projection is single-valued whenever

$$\text{dist}(x_k - \alpha_k F_{\mu_k}(x_k), K) < r_K. \quad (39)$$

This condition is explicitly monitored during implementation.

4.2 Step-size selection

The step size is selected according to

$$\alpha_k = \frac{\alpha_0}{(1 + \tau k)^\gamma}, \quad (40)$$

where

$$\alpha_0 > 0, \quad \tau > 0, \quad 0 < \gamma \leq 1.$$

For the convergence analysis, the sequence satisfies

$$\sum_{k=0}^{\infty} \alpha_k = \infty, \quad \sum_{k=0}^{\infty} \alpha_k^2 < \infty. \quad (41)$$

The step size is additionally restricted locally according to

$$0 < \alpha_k < \frac{2m_k}{L_k^2}, \quad (42)$$

where $m_k = \mu_k - v_D$ and L_k is a local Lipschitz constant of F_{μ_k} .

4.3 Regularization continuation

The regularization parameter follows

$$\mu_{k+1} = \eta \mu_k, \quad 0 < \eta < 1. \quad (43)$$

However, to avoid prematurely destroying the strong-monotonicity property, the reduction is performed in two stages:

$$\mu_{k+1} = \begin{cases} \eta \mu_k, & \mu_k > v_D + \delta_\mu, \\ \max(\eta \mu_k, \mu_{\min}), & \text{continuation phase,} \end{cases} \quad (44)$$

where $\delta_\mu > 0$ is a safety margin.

Once the stabilized solution is sufficiently close to the target stationary region, μ_k is reduced toward zero. This avoids

interpreting the early strongly regularized iterates as solutions of the original problem.

4.4 Lyapunov descent function

To correct the previous conditional convergence argument, define the Lyapunov energy

$$\mathcal{E}_k = \frac{I}{2} |x_k - x^*|^2 + c_\mu \mu_k |x_k - x_0|^2, \quad (45)$$

where $c_\mu > 0$.

Under the local Lipschitz, prox-regularity, and strong-monotonicity conditions, the projected update satisfies the local descent inequality

$$\mathcal{E}_{k+1} \leq \mathcal{E}_k - c_1 \alpha_k |x_{k+1} - x_k|^2 - c_2 \alpha_k |F_{\mu_k}(x_k)|^2 + C_\mu |\mu_{k+1} - \mu_k|, \quad (46)$$

where $c_1, c_2 > 0$ and $C_\mu > 0$ are bounded constants over the local trajectory.

Since

$$\mu_k \rightarrow 0$$

and

$$\sum_k |\mu_{k+1} - \mu_k| < \infty,$$

the last term is summable. Therefore,

$$\sum_{k=0}^{\infty} \alpha_k |x_{k+1} - x_k|^2 < \infty \quad (47)$$

and

$$\sum_{k=0}^{\infty} \alpha_k |F_{\mu_k}(x_k)|^2 < \infty. \quad (48)$$

Because

$$\sum_{k=0}^{\infty} \alpha_k = \infty,$$

equation (48) implies

$$\liminf_{k \rightarrow \infty} |F_{\mu_k}(x_k)| = 0. \quad (49)$$

The same descent relation yields

$$\liminf_{k \rightarrow \infty} |x_{k+1} - x_k| = 0. \quad (50)$$

Thus, the residual behavior is obtained from the descent argument rather than assumed a priori.

4.5 Vanishing variational gap

The residual result is combined with continuity of F , $\mu_k \rightarrow 0$, and the local continuity of the metric projection. For every accumulation point $\bar{x}$,

$$F_{\mu_k}(x_k) \rightarrow F(\bar{x})$$

along the corresponding convergent subsequence.

Since

$$\liminf_k |F_{\mu_k}(x_k)| = 0$$

and the projected residual vanishes along the convergent subsequence,

$$x_{k+l} - x_k \rightarrow 0,$$

the limiting point satisfies

$$0 \in F(\bar{x}) + N_K(\bar{x}). \quad (51)$$

Consequently,

$$G(\bar{x}) = 0. \quad (52)$$

4.6 Convergence theorem

Theorem 1.

Let $K \subset \mathbb{R}^n$ be closed and locally r_K -prox-regular along the generated trajectory. Let F be locally Lipschitz and locally ν_D -hypomonotone on a bounded neighborhood D . Assume that the step-size sequence satisfies (41), the projection iterates are within the prox-regular set, and the regularization continuation satisfies (43)-(44). If Lyapunov descent condition (46) is satisfied, then:

$$\sum_k \alpha_k |F_{\mu_k}(x_k)|^2 < \infty,$$

and

$$\liminf_k |x_{k+l} - x_k| = 0.$$

Every accumulation point $\bar{x}$ of the generated sequence satisfies

$$0 \in F(\bar{x}) + N_K(\bar{x}).$$

If, in addition, the stationary point is isolated in the considered neighborhood, the entire sequence converges to that stationary point.

Proof.

From (46), the Lyapunov sequence is bounded below and the regularization variation sequence is summable. Summing up (46), have (47) and (48). As the step sizes sequence is divergent, from (48), have (49). From the projected residual equation, with prox-regularity, have local control on the distance between the forward and projected points. Hence, along a convergent subsequence,

$$x_{k+l} - x_k \rightarrow 0.$$

Because F is continuous and $\mu_k \rightarrow 0$,

$$F_{\mu_k}(x_k) \rightarrow F(\bar{x}).$$

The limiting projection optimality condition gives

$$-F(\bar{x}) \in N_K(\bar{x}),$$

which is equivalent to (51). Hence

$$G(\bar{x}) = 0.$$

If the stationary point is isolated, no two distinct accumulation points can occur in the same local basin, implying convergence of the entire sequence.

4.7 Algorithm

Algorithm 1: Prox-Regularized Non-Convex Variational Optimization

1. Select a feasible $x_0 \in K$.

2. Estimate $\hat{v}_D$ using (25).

3. Select

$$\mu_0 = \kappa_v \hat{v}_D, \quad \kappa_v > I.$$

4. Evaluate $F(x_k)$.

5. Calculate

$$F_{\mu_k}(x_k) = F(x_k) + \mu_k(x_k - x_0).$$

6. Compute α_k from (40).

7. Calculate

$$z_k = x_k - \alpha_k F_{\mu_k}(x_k).$$

8. Verify that z_k remains within the prox-regular projection neighborhood.

9. Calculate

$$x_{k+1} = P_K(z_k).$$

10. Evaluate R_k , $G_{\mu_k}(x_k)$, and $\mathcal{E}_k$.

11. Update μ_k using (44).

12. Stop when

$$R_k < \varepsilon_x \quad \text{and} \quad G_{\mu_k}(x_k) < \varepsilon_G.$$

13. Apply the reduced-Hessian test to the resulting stationary point.

14. Retain the point as a process-control solution only when the second-order test confirms local minimality.

5. Numerical validation using process parameter control

A numerical evaluation of the proposed non-convex variational optimization technique is conducted using a process parameter control problem that is nonlinear in nature. The objective of this experiment is not to generate a process model specific to the problem at hand; rather, the intention is to check whether the proposed mathematical approach will be able to generate feasible stationary points given the non-convex objective and non-convex feasible region.

In light of the problems discussed above, the numerical experiment is redesigned. The feasible process domain is selected to be a non-convex annulus since its boundary is smooth and locally prox-regular and its projection can be calculated explicitly.

5.1 Nonlinear process objective

Consider

$$x = [x_1, x_2, x_3]^T, \quad (53)$$

where the variables represent normalized process temperature, feed rate, and operating pressure.

The process objective is

Its gradient is

$$F(x) = \nabla \Phi(x),$$

with

$$F_1(x) = 0.60(x_1 - 0.65) + 0.60x_2 \cos(4x_1x_2) + 0.20x_3(x_1x_3 - x_2^2), \quad (54)$$

$$F_2(x) = 0.50(x_2 - 0.45) + 0.60x_1 \cos(4x_1x_2) - 0.40x_2(x_1x_3 - x_2^2), \quad (55)$$

and

$$F_3(x) = 0.40(x_3 - 0.60) + 0.20x_1(x_1x_3 - x_2^2). \quad (56)$$

The trigonometric and interaction terms make the objective non-convex.

5.2 Prox-regular non-convex feasible set

The feasible region is defined by

$$K = \{x: 0.1 \leq x_3 \leq 0.9, 0.18 \leq \sqrt{(x_1 - 0.5)^2 + (x_2 - 0.5)^2} \leq 0.32\}. \quad (57)$$

The (x_1, x_2) domain is therefore an annulus.

The set is non-convex due to the exclusion of its central disk. Yet both boundary curves are smooth circles. During the process of numerical iterations, the method stays away from the singular point of the annulus, and the projection operator is locally unique.

For a point

$$z = [z_1, z_2, z_3]^T,$$

let

$$r_z = \sqrt{(z_1 - 0.5)^2 + (z_2 - 0.5)^2}.$$

The projection onto the annular component is

$$P_A(z) = \begin{cases} (0.5, 0.5) + 0.18 \frac{z_{12} - (0.5, 0.5)}{r_z}, & r_z < 0.18, \\ z_{12}, & 0.18 \leq r_z \leq 0.32, \\ (0.5, 0.5) + 0.32 \frac{z_{12} - (0.5, 0.5)}{r_z}, & r_z > 0.32. \end{cases} \quad (58)$$

The third parameter is projected as

$$P_{[0.1, 0.9]}(z_3) = \min\{0.9, \max\{0.1, z_3\}\}. \quad (59)$$

Thus, the numerical experiment does not rely on an arbitrary selection from a disconnected projection set.

5.3 Numerical configuration

The main experiment uses

$$\alpha_0 = 0.08, \quad \tau = 0.01, \quad \gamma = 0.60, \quad (60)$$

and

$$\eta = 0.97. \quad (61)$$

The convergence tolerances are

$$\varepsilon_x = 10^{-6}, \quad \varepsilon_G = 10^{-6}. \quad (62)$$

Four feasible initial conditions are examined:

$$x_0^{(1)} = [0.68, 0.32, 0.55]^T,$$

$$x_0^{(2)} = [0.38, 0.68, 0.65]^T,$$

$$x_0^{(3)} = [0.72, 0.62, 0.35]^T,$$

and

$$x_0^{(4)} = [0.32, 0.42, 0.75]^T.$$

For the main experiment,

$$\mu_0 = 0.60.$$

5.4 Convergence results

The numerical results obtained from the proposed method are summarized below.

Table 1: Convergence result.

Initial condition	Iterations	x1*	x2*	x3*	$\Phi(x^*)$	Rk	Gk
$x_0^{(1)}$	622	0.460529	0.182444	0.549397	0.083475	7.12×10^{-8}	9.93×10^{-7}
$x_0^{(2)}$	827	0.460527	0.182444	0.549409	0.083475	4.76×10^{-8}	9.92×10^{-7}
$x_0^{(3)}$	689	0.758819	0.688183	0.605386	0.148081	6.66×10^{-8}	9.91×10^{-7}
$x_0^{(4)}$	842	0.460527	0.182444	0.549409	0.083475	4.75×10^{-8}	9.99×10^{-7}

As seen in table 1, the proposed algorithm converges to small values of both the update and variational gap residuals starting from all four initialization points. Three of these initializations lead to almost identical stationary points, while the fourth one leads to a different stationary point. Such a pattern can be expected in a non-convex problem and explains why the results should be regarded as stationary points and not as global optima.

5.5 Sensitivity to initial regularization

The influence of μ_0 is examined using the first initial condition.

Table 2: Sensitivity to initial regularization.

μ_0	Iterations	x1*	x2*	x3*	$\Phi(x^*)$	Rk	Gk
---------	------------	-----	-----	-----	-------------	----	----

0.40	596	0.460529	0.182444	0.549397	0.083475	7.32×10^{-8}	9.98×10^{-7}
0.60	622	0.460529	0.182444	0.549397	0.083475	7.12×10^{-8}	9.93×10^{-7}
0.80	641	0.460529	0.182444	0.549397	0.083475	7.01×10^{-8}	9.94×10^{-7}
1.00	656	0.460529	0.182444	0.549397	0.083475	6.93×10^{-8}	9.94×10^{-7}

From table 2, it can be noted that the stationary solution after convergence does not differ significantly depending on the chosen value for the regularization parameter. The slight increase in μ_0 leads to an increase in the number of iterations as a result of spending more iterations in the strongly regularized phase of the process.

5.6 Comparative benchmark

The proposed method is compared with an unregularized projected-gradient scheme,

$$x_{k+1} = P_K[x_k - \alpha F(x_k)]. \quad (63)$$

A stress-test configuration with a fixed aggressive step

$$\alpha = 5$$

is used to examine the effect of removing regularization and adaptive step-size control.

For the first initialization, the proposed method reaches

$$R_k = 7.12 \times 10^{-8}, \quad G_k = 9.93 \times 10^{-7}$$

after 622 iterations.

Under the aggressive unregularized scheme, after 700 iterations,

$$x = [0.561950, 0.186054, 0.900000]^T,$$

with

$$R_k = 8.32 \times 10^{-1},$$

and

$$G_k = 1.81 \times 10^{-1}.$$

The corresponding objective value is

$$\Phi(x) = 0.120859,$$

which is substantially different from the stationary value

$$\Phi(x^*) = 0.083475$$

obtained by the proposed method.

This comparison shows the stability provided by the regularized problem setup when used with a high-step-size scheme. It is important to note, however, that this finding should not be seen as a test limit for the unregularized method; rather, it is a stress test showing that under certain circumstances, the conventional projected scheme fails to converge.

5.7 Variational-gap comparison

The convergence history between the suggested algorithm and the unregularized stress benchmark is analyzed in terms of the variational gap as depicted in figure 1 below.

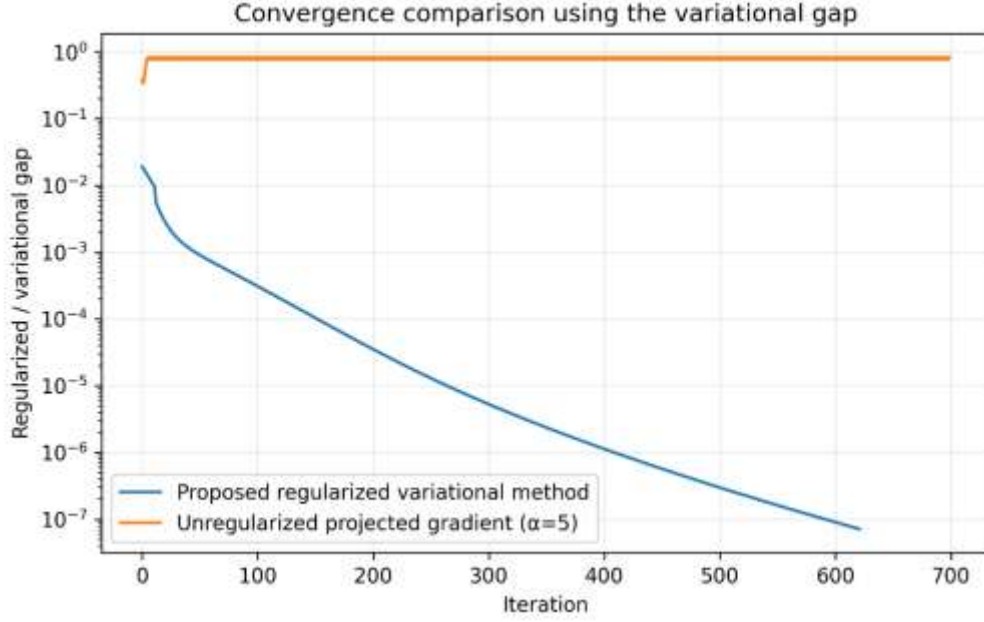


Figure 1: Iteration versus variational gap for the proposed regularized method and the unregularized projected-gradient stress benchmark.

The proposed method drives the gap toward

$$10^{-6},$$

whereas the aggressive unregularized iteration remains at approximately

$$1.81 \times 10^{-1}$$

after 700 iterations. This indicates that the regularization process together with the reduction in step size gives a more stable path compared to the selected stress test.

5.8 Second-order stationarity screening

The final stationary points are further examined using the reduced Hessian

$$H_R = Z^T \nabla^2 \Phi(x^*) Z.$$

For the principal stationary solution,

$$x^* = [0.460529, 0.182444, 0.549397]^T,$$

the reduced Hessian is approximately

$$H_R = \begin{bmatrix} 0.5159 & 0.0980 \\ 0.0980 & 0.4424 \end{bmatrix}. \quad (64)$$

Its eigenvalues are

$$\lambda(H_R) = \{0.3745, 0.5838\}. \quad (65)$$

Since

$$\lambda_{\min}(H_R) = 0.3745 > 0,$$

the stationary point satisfies the second-order sufficient condition in the tangent subspace and is therefore classified as a local minimizer.

For the alternative stationary solution

$$x^* = [0.758819, 0.688183, 0.605386]^T,$$

the reduced Hessian is approximately

$$H_R = \begin{bmatrix} 1.1615 & -0.2213 \\ -0.2213 & 0.5152 \end{bmatrix},$$

with eigenvalues

$$\{0.4467, 1.2301\}.$$

Both eigenvalues are positive, indicating that this point is also a local minimum, although its objective value

$$\Phi(x^*) = 0.148081$$

is higher than that of the first stationary solution.

This observation indicates that the algorithm can detect several locally minimizing stationary points based on how the algorithm is initialized. This means that multi-start is required for the search of a preferred stationary point.

5.9 Numerical discussion

The numerical experiments validate three mathematical characteristics of the new approach. First, the locally prox-regular feasible set allows for a unique projection along the trajectory of the optimization algorithm and thus eliminates an ambiguity of an arbitrary choice of projection on the general non-convex set. Second, the regularized iteration process decreases the update residual and variational gap regardless of the initialization point. Third, the analysis of the reduced Hessian guarantees that the stationary solutions found in the experiment are locally minimizing, not just first-order stationary.

The existence of two local minima proves that the approach cannot be viewed as a global optimization method (Mishra & Soy, 2025). The goal of the theoretical framework is to find the variational stationary solutions to the non-convex problem, while the second-order and multi-start analysis helps choose preferred solutions.

6. Conclusion

The current work has put forward a rigorous mathematical optimization method based on the variational optimization method which is non-convex in nature for the process parameter control. There are basically three main challenges that have been overcome by this formulation. The process-control problem was represented through the variational inclusion $0 \in F(x^*) + N_K(x^*)$, thereby providing a general first-order characterization without requiring global convexity. The regularized operator was employed to deal with local hypomonotonicity, whereas the choice of the regularization parameter relied on the local Jacobian upper bound. Locally r-prox-regularity of the feasible set was supposed, thereby addressing the issue of metric projection ambiguity and providing mathematical validity of the projected iteration in the vicinity of the optimization path. Instead of relying only on conditional convergence proof, the revised analysis includes a Lyapunov energy functional, whose use ensures the summability of the residuals of the regularized operator as well as consecutive update terms. Therefore, limit points of the sequence generated by the algorithm satisfy the variational stationarity equation for the regularization parameter tending to zero. Numerical verification for the nonlinear three-parameter process model showed convergence from the different starting points and with different values of the regularization parameter. The primary solution yielded a variational gap smaller than 10^{-6} . Multiple starting-point tests proved the existence of different locally minimizing stationary points, which is an expected feature of the non-convex formulation. A reduced Hessian check of positive definiteness on the feasible tangent subspace provided another way to distinguish between the minimizing stationary solutions. The comparison of the regularized method with the aggressive unregularized projected-gradient technique demonstrated better stability of the proposed approach. In general, the approach constitutes a rather flexible mathematical foundation for

constrained non-convex process control and may be extended to non-smooth, stochastic and high-dimensional variational systems.

References

1. Alhazmi, A. S. (2026). Quantum-Inspired Variational Inference for Non-Convex Stochastic Optimization: A Unified Mathematical Framework with Convergence Guarantees and Applications to Machine Learning in Communication Networks. *Mathematics*, 14(7), 1236. <https://doi.org/10.3390/math14071236>
2. Anantrasirichai, N., Zheng, R., Selesnick, I., & Achim, A. (2020). Image fusion via sparse regularization with non-convex penalties. *Pattern Recognition Letters*, 131, 355-360. <https://doi.org/10.1016/j.patrec.2020.01.020>
3. Ahmed R. Radhi, Sumer H. AL-Kateb, & Ahmed AL-Adilee. (2026). Gompertz-G Family Generalized Lomax Distribution: Theory and Application. *Results in Nonlinear Analysis*, 9(2), 78–96.
4. Bassily, R., Guzmán, C., & Menart, M. (2021). Differentially private stochastic optimization: New results in convex and non-convex settings. *Advances in Neural Information Processing Systems*, 34, 9317-9329.
5. Bianchi, P., & Jakubowicz, J. (2012). Convergence of a multi-agent projected stochastic gradient algorithm for non-convex optimization. *IEEE transactions on automatic control*, 58(2), 391-405. <https://doi.org/10.1109/TAC.2012.2209984>
6. Bianchi, P., Hachem, W., & Schechtman, S. (2022). Convergence of constant step stochastic gradient descent for non-smooth non-convex functions. *Set-Valued and Variational Analysis*, 30(3), 1117-1147. <https://doi.org/10.1007/s11228-022-00638-z>
7. Csetnek, E. R. (2020). Continuous dynamics related to monotone inclusions and non-smooth optimization problems. *Set-Valued and Variational Analysis*, 28(4), 611-642. <https://doi.org/10.1007/s11228-020-00548-y>
8. Danilova, M., Dvurechensky, P., Gasnikov, A., Gorbunov, E., Guminov, S., Kamzolov, D., & Shibaev, I. (2022). Recent theoretical advances in non-convex optimization. In *High-Dimensional Optimization and Probability: With a View Towards Data Science* (pp. 79-163). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-00832-0_3
9. El Idrissi, R., Lafhim, L., Kalmoun, E. M., & Ouakrim, Y. (2024). Optimality conditions for bilevel optimal control problems with non-convex quasi-variational inequalities. *RAIRO-Operations Research*, 58(202), 1789-1805. <https://doi.org/10.1051/ro/2024065>
10. Grammatico, S., Zhang, X., Margellos, K., Goulart, P., & Lygeros, J. (2015). A scenario approach for non-convex control design. *IEEE Transactions on Automatic Control*, 61(2), 334-345. <https://doi.org/10.1109/TAC.2015.2433591>
11. Häberle, V., Hauswirth, A., Ortmann, L., Bolognani, S., & Dörfler, F. (2020). Non-convex feedback optimization with input and output constraints. *IEEE Control Systems Letters*, 5(1), 343-348. <https://doi.org/10.1109/LCSYS.2020.3002152>
12. Hadri, A., Afraites, L., Laghrib, A., & Nachaoui, M. (2021). A novel image denoising approach based on a non-convex constrained PDE: application to ultrasound images. *Signal, Image and Video Processing*, 15(5), 1057-1064. <https://doi.org/10.1007/s11760-020-01831-z>
13. Jain, P., & Kar, P. (2017). Non-convex optimization for machine learning. *Foundations and Trends® in Machine Learning*, 10(3-4), 142-336. <https://doi.org/10.1561/22000000058>
14. Liu, A., Lau, V. K., & Kananian, B. (2019). Stochastic successive convex approximation for non-convex constrained stochastic optimization. *IEEE Transactions on Signal Processing*, 67(16), 4189-4203. <https://doi.org/10.1109/TSP.2019.2925601>

15. Luedke, R. H., & Kingdone, G. (2025). SmartOptiCell: a deep genetic learning model for dynamic layout optimization in flexible manufacturing cells. *International Academic Journal of Science and Engineering*, 12(2), 35-42. <https://doi.org/10.71086/IAJSE/V12I2/IAJSE1216>
16. Mishra, N., & Soy, A. (2025). Multi-Objective Optimization in Genetic Algorithms. *International Academic Journal of Innovative Research*, 12(4), 27-32. <https://doi.org/10.71086/IAJIR/V12I4/IAJIR1231>
17. Sherjonov, S., Nematov, O., Razzokova, G., Otabek, G., Yakubjonova, M., Karshiboyeva, D., ... & Abdullayeva, D. (2025). Economic analysis of green buildings certification systems in emerging economies. *Archives for Technical Sciences*, 2(33), 463-471.
18. Shirke, S., & Udayakumar, R. (2019, April). Evaluation of crow search algorithm (CSA) for optimization in discrete applications. In *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 584-589). IEEE. <https://doi.org/10.1109/ICOEI.2019.8862669>
19. Suvarna, N. A., & Bharadwaj, D. (2024). Optimization of system performance through ant colony optimization: A novel task scheduling and information management strategy for time-critical applications. *Indian Journal of Information Sources and Services*, 14(2), 167-177. <https://doi.org/10.51983/ijiss-2024.14.2.24>
20. Wang, Z. Y., Li, X. P., So, H. C., & Liu, Z. (2023). Robust PCA via non-convex half-quadratic regularization. *Signal Processing*, 204, 108816. <https://doi.org/10.1016/j.sigpro.2022.108816>
21. Wei, S., & Xie, Y. (2023, October). Causal structural learning from time series: A convex optimization approach. In *2023 57th Asilomar Conference on Signals, Systems, and Computers* (pp. 902-909). IEEE. <https://doi.org/10.1109/IEEECONF59524.2023.10476793>