Sequential Linear Programming With Adaptive Linearization Error Limits for All-Time Feasibility

Leko, Dorijan; Vašak, Mario

Source / Izvornik: **IEEE Control Systems Letters (L-CSS), 2024, 8, 3051 - 3056**

Journal article, Published version
Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

https://doi.org/10.1109/LCSYS.2024.3519547

Sequential Linear Programming With Adaptive Linearization Error Limits for All-Time Feasibility

Dorijan Leko, *Graduate Student Member, IEEE*, and Mario Vašak, *Senior Member, IEEE*

***Abstract*—This letter presents an enhanced Trust Region Method (TRM) for Sequential Linear Programming (SLP) designed to improve the initial feasible solution to a constrained nonlinear programming problem while maintaining the interim solutions feasibility throughout the SLP iterations. The method employs a polytopic sub-approximation of the feasible region, defined around the interim solution as a level set based on variable limits for the linearization error. This polytopic feasible region is established by using a trust region that ensures that maximum limits of the linearization errors are respected. The method adaptively adjusts the size of the feasible region during iterations to achieve convergence to a local optimum by employing variable linearization error limits. Local convergence is attained by reducing the size of the trust radius. A case study illustrates the effectiveness of the proposed method, which is compared to the benchmark TRM that uses heuristic limits on the permissible changes in manipulated variables.**

***Index Terms*—Optimization algorithms, optimization, computational methods.**

I. Introduction

Sequential Linear Programming (SLP) is an optimization technique designed to approximately solve nonlinear programming (NLP) problems [1]. This iterative method relies on local linearizations at interim solution points, typically achieved through first-order Taylor series expansion, to identify (local) optimal solutions. The linear approximation effectively captures the behavior of the original nonlinear problem in the vicinity of the interim solution. Thus, SLP requires well-defined moving limits at these points to maintain interim solution feasibility throughout the iterations and to ensure convergence of the solution.

There are two main classes of methods for navigating the manipulated variables through the SLP iterations. The first class, line search methods [2], focuses on finding the best direction before determining the permissible step size. The second class, central to this letter, involves Trust Region Methods (TRMs) [3], where the trust region of moving limits is established first, followed by identifying the best solution within that region. The size and shape of the trust region can significantly impact the feasibility of the solutions generated during the iteration process. Trust region methods can incorporate feasibility checks to ensure that proposed solutions adhere to constraints, which is crucial for practical applications. However, these feasibility checks can require numerous iterations, which may significantly increase the solving time [4].

In practice, developing an efficient algorithm that addresses a wide range of NLP problems is still challenging. Thus, ensuring feasibility in TRM across different applications has a significant focus of research in recent years [5]. Notable examples include a feasible SLP algorithm for time-optimal control problems proposed in [6], an Anderson-accelerated feasible SLP algorithm in [7], and an almost feasible SLP algorithm [8]. Additionally, a contextual optimization method based on sample-based trust region dynamics is presented in [9], along with trust-region inverse reinforcement learning [10] and a TRM for data-driven iterative learning control [11]. Nevertheless, the benchmark TRM applicable to a wide class of nonlinear problems relies on adaptive heuristic limits for permissible changes in manipulated variables [12].

The benchmark trust region subproblem is formulated as follows:

$$\begin{aligned} \min_{\mathbf{x}\in\mathcal{F}} \quad & c(\mathbf{x}^*) + \nabla c(\mathbf{x}^*)^\top (\mathbf{x}-\mathbf{x}^*) \\ \text{s.t.} \quad & \|\mathbf{x}-\mathbf{x}^*\| \le \delta. \end{aligned} \tag{1}$$

Here, the vector $\mathbf{x} \in \mathbb{R}^n$ represents the decision vector within the feasible set $\mathcal{F} \subset \mathbb{R}^n$. The vector $\mathbf{x}^* \in \mathbb{R}^n$ denotes the operating point, and $c : \mathbb{R}^n \to \mathbb{R}$ is a continuous differentiable function. The parameter $\delta \in \mathbb{R}_+$ represents the trust radius. The operator ∇ denotes the gradient, whereas $\|\cdot\|$ indicates the Euclidean norm. The trust radius is determined heuristically based on performance in the previous iteration. If the objective function is successfully minimized and the feasibility of the original optimization problem is maintained, the trust radius is increased by a specified value for the next iteration. Conversely, if these conditions are not satisfied, the trust radius is decreased, and the calculation of the trust region

Received 16 September 2024; revised 17 November 2024; accepted 2 December 2024. Date of publication 18 December 2024; date of current version 30 December 2024. This work was supported by the European Union from the European Regional Development Fund through the Operative Programme Competitiveness and Cohesion 2014–2020 for Croatia within the Project 'Dynamic Predictive Health Protection of an Electric Vehicle Battery' (EVBattPredtect) under Contract KK.01.1.1.07.0029 and the Project 'Machine Learning-Based Model of Battery Cell Aging for Online Battery Pack Health Protection' (MLBattProt) under Project NPOO.C3.2.R3-I1.04.0263. Recommended by Senior Editor S. Olaru. *(Corresponding author: Dorijan Leko.)*

The authors are with the Department of Control and Computer Engineering, University of Zagreb Faculty of Electrical Engineering and Computing, 10000 Zagreb, Croatia (e-mail: dorijan.leko@fer.unizg.hr; mario.vasak@fer.unizg.hr).

Digital Object Identifier 10.1109/LCSYS.2024.3519547

subproblem is repeated. The parameters such as initial trust radius and navigation factors between iterations are typically set heuristically, often informed by the scale of the problem or domain knowledge.

This letter presents an enhancement of the benchmark TRM (1) that ensures the feasibility of the interim solution in each SLP iteration without requiring heuristic trust region parameters. The methodology employs variable linearization error limits to construct feasible regions around the interim solution by adaptively adjusting the size of the trust region that ensures that limits of the linearization errors are respected. The trust region is defined as a level set of the linearization error limit, evaluated using the Lagrange remainder [13] within a predefined domain. Consequently, the method requires that all functions are twice continuously differentiable, with their second derivatives bounded over the specified domain. Throughout the iterative process, convergence of the local solution is achieved by adaptively adjusting the trust radius.

The contributions of this letter are as follows:

- Enhanced benchmark TRM to ensure solution feasibility in each SLP iteration without compromising the solving complexity;
- Introduced two offline procedures for polytopic sub-approximation of the feasible region around the interim solution, each suited to different problem sizes;
- Developed procedure for dynamically adjusting the size of the trust radius by adaptively modifying variable linearization error limits with the intention to generate large feasible regions and thus facilitate convergence to the SLP solution.

The proposed TRM consists of both offline and online components. The offline part involves determining the polytopic sub-approximation of the trust region as a level set of the linearization error limit around the interim solution. The online part includes solving iterative subproblems, which encompass navigating the trust radius between iterations and solving the linear program of the subproblem.

This letter is organized as follows. Section II presents the problem setup and objectives of the proposed TRM. Section III outlines the procedures for polytopic sub-approximation of the trust region. Section IV discusses the navigation of linearization error limits through SLP iterations to achieve convergence. The scalability of the proposed method is analyzed in Section V. Finally, Section VI applies the proposed methodology to solve a constrained NLP problem and compares its benefits with the benchmark TRM that uses heuristic limits on permitted changes of manipulated variables.

II. Problem Setup

This section introduces the problem setup and outlines the objectives of the enhanced Trust Region Method (TRM) for solving constrained nonlinear programming (NLP) problems:

$$\min_{\mathbf{x}\in\mathcal{X}} \quad g_0(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{g}(\mathbf{x}) \le 0, \tag{2}$$

where the vector $\mathbf{x} \in \mathbb{R}^n$ represents the decision variables bounded by the set $\mathcal{X} \subset \mathbb{R}^n$. The cost function $g_0 : \mathbb{R}^n \to \mathbb{R}$ and constraint function $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^p$ are twice continuously differentiable. The feasible region $\mathcal{F} \subset \mathbb{R}^n$ is defined by the nonlinear constraints as follows:

$$\mathcal{F} = \{\mathbf{x} \in \mathcal{X} : \mathbf{g}(\mathbf{x}) \le 0\}. \tag{3}$$

The objective is to enhance the benchmark TRM (1) by ensuring feasibility at each SLP iteration through the consideration of linearization error limits, while maintaining a comparable solving complexity for the NLP problems defined in (2). To establish feasibility around the interim solution, it is essential to introduce the definitions of the trust region and the indented half-space, as outlined on the generic constraint function below.

Let a generic scalar function $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable (i.e., f is of class C^2) on a closed convex set $\mathcal{X} \subset \mathbb{R}^n$. Suppose there exists $\mathbf{x} \in \mathcal{X}$ such that $f(\mathbf{x}) \le 0$. Additionally, the function is linearized at a feasible operating point $\mathbf{x}^* \in \mathcal{X}$, where $f(\mathbf{x}^*) \le 0$ as follows:

$$f_{\mathrm{L}}(\mathbf{x}) = f(\mathbf{x}^*) + \mathbf{J}_f^\top(\mathbf{x} - \mathbf{x}^*), \quad \mathbf{J}_f = \nabla f(\mathbf{x}^*), \tag{4}$$

where $\mathbf{J}_f \in \mathbb{R}^n$ represents the Jacobian of the function f at the operating point $\mathbf{x}^*$. Furthermore, let the linearization error limit be represented by a positive real number, i.e., $R_{\mathrm{L}} \in \mathbb{R}_+$.

Definition 1: The *trust region* $\mathcal{T} \subset \mathbb{R}^n$ of a generic function f is defined as a level set of the linearization error limit around an operating point, characterized by the following condition:

$$\mathcal{T}(f, \mathbf{x}^*, R_{\mathrm{L}}) = \{\mathbf{x} \in \mathcal{X} \mid |f(\mathbf{x}) - f_{\mathrm{L}}(\mathbf{x})| \le R_{\mathrm{L}}\}.$$

Definition 2: The *indented half-space* $\mathcal{I} \subset \mathbb{R}^n$ of a generic function f is defined as a half-space in which the linearized generic function f_{L} satisfies the constraint $f(\mathbf{x}) \le 0$ under the following condition:

$$\mathcal{I}(f, \mathbf{x}^*, R_{\mathrm{L}}) = \{\mathbf{x} \in \mathcal{X} \mid f(\mathbf{x}) = f_{\mathrm{L}}(\mathbf{x}) + R_{\mathrm{L}} \le 0\}.$$

The intersection of the trust region and the indented half-space constructs a feasible region around an operating point, as stated in the following proposition.

Proposition 1: For any feasible operating point $\mathbf{x}^* \in \mathcal{X}$ such that $f(\mathbf{x}^*) \le 0$, it is guaranteed that the constraint $f(\mathbf{x}) \le 0$ is satisfied for all

$$\mathbf{x} \in \mathcal{T}(f, \mathbf{x}^*, R_{\mathrm{L}}) \cap \mathcal{I}(f, \mathbf{x}^*, R_{\mathrm{L}}), \quad R_{\mathrm{L}} \in [0, -f(\mathbf{x}^*)].$$

Proof: Let $\mathbf{x}^* \in \mathcal{X}$ be any feasible operating point, thus $f(\mathbf{x}^*) \le 0$. The linearized constraint function at this point is given by (4) as $f_{\mathrm{L}}(\mathbf{x}) = f(\mathbf{x}^*) + \mathbf{J}_f^\top(\mathbf{x} - \mathbf{x}^*)$.

1. For all $\mathbf{x} \in \mathcal{T}(f, \mathbf{x}^*, R_{\mathrm{L}}) \cap \mathcal{I}(f, \mathbf{x}^*, R_{\mathrm{L}})$, the following inequalities must hold:

$$|f(\mathbf{x}) - f_{\mathrm{L}}(\mathbf{x})| \le R_{\mathrm{L}} \quad \text{and} \quad f_{\mathrm{L}}(\mathbf{x}) + R_{\mathrm{L}} \le 0.$$

2. From these inequalities, $f(\mathbf{x})$ can be expressed as:

$$f_{\mathrm{L}}(\mathbf{x}) - R_{\mathrm{L}} \le f(\mathbf{x}) \le f_{\mathrm{L}}(\mathbf{x}) + R_{\mathrm{L}} \le 0.$$

3. Since $\mathbf{x}^*$ is a feasible operating point with $f(\mathbf{x}^*) \le 0$, using the inequalities above and (4), it follows that:

$$f(\mathbf{x}) \le 0 \text{ for all } 0 \le R_{\mathrm{L}} \le -f(\mathbf{x}^*).$$

Thus, for any feasible operating point $\mathbf{x}^*$ and linearization error limit $R_{\mathrm{L}} \in [0, -f(\mathbf{x}^*)]$, it is guaranteed that the constraint

$f(\mathbf{x}) \leq 0$ holds for all $\mathbf{x}$ in the intersection of the trust region and the indented half-space. ■

The condition that ensures feasibility (Proposition 1) is extended to the case of multiple constraints defined by p generic constraint functions, such as $f_1 \leq 0, \ \ldots, \ f_p \leq 0$, which share the same properties as f. Thus, the intersected feasible region $\mathcal{F}_\cap \subset \mathbb{R}^n$ is determined as the level set of p linearization error limits around the interim solution:

$$\mathcal{F}_\cap\left(\mathbf{x}^*, R_{\mathrm{L}1}, \ldots, R_{\mathrm{L}p}\right) = \bigcap_{i=1}^{p} \mathcal{F}_i\left(f_i, \mathbf{x}^*, R_{\mathrm{L}i}\right), \tag{5}$$

where $\mathcal{F}_i \subset \mathbb{R}^n$ represents the feasible region for the f_i constraint function, defined as the intersection of its trust region $\mathcal{T}_i$ and indented half-space $\mathcal{I}_i$: $\mathcal{F}_i(f_i, \mathbf{x}^*, R_{\mathrm{L}i}) = \mathcal{T}_i(f_i, \mathbf{x}^*, R_{\mathrm{L}i}) \cap \mathcal{I}_i(f_i, \mathbf{x}^*, R_{\mathrm{L}i})$.

III. Sub-Approximation of the Trust Region

This section describes two offline procedures for the polytopic sub-approximation of the feasible region around an interim solution, based on the intersection of the trust region and the indented half-space (Proposition 1). Both procedures leverage the characterization of linearization error via the Lagrange remainder to establish the polytopic sub-approximation of the trust region.

The linearization of the generic function f is achieved by employing the first-order Taylor series and its Lagrange remainder, as stated in the following theorem.

Theorem 1 [13]: The first-order Taylor series and its Lagrange remainder of the function f for any $\mathbf{x} \in \mathcal{X}$ around the operating point $\mathbf{x}^* \in \mathcal{X}$ is given by:

$$f(\mathbf{x}) = f\left(\mathbf{x}^*\right) + \nabla f\left(\mathbf{x}^*\right)^\top \left(\mathbf{x} - \mathbf{x}^*\right) + R(\mathbf{x}),$$

where the remainder $R : \mathbb{R}^n \rightarrow \mathbb{R}$ is given in Lagrange's form by

$$R(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x} - \mathbf{x}^*\right)^\top \nabla^2 f(\boldsymbol{\xi}(\mathbf{x}))\left(\mathbf{x} - \mathbf{x}^*\right), \tag{6}$$

for some $\boldsymbol{\xi}(\mathbf{x}) \in \{\mathbf{x}^* + c(\mathbf{x} - \mathbf{x}^*) \mid c \in [0, 1]\}$.

The absolute value of the Lagrange remainder can be over-approximated when all second-order partial derivatives of the function f are bounded on a specific domain, as stated in the following corollary.

Corollary 1 [13]: If all second-order partial derivatives, that form the Hessian matrix of the function f are bounded by a bounding-range matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ such that $|\nabla^2 f(\boldsymbol{\xi}(\mathbf{x}))| \leq \mathbf{M}$ for all $\boldsymbol{\xi}(\mathbf{x}) \in \{\mathbf{x}^* + c(\mathbf{x} - \mathbf{x}^*) \mid \mathbf{x} \in \mathcal{X}, \ c \in [0, 1]\}$, then the Lagrange remainder can be over-approximated by $|R(\mathbf{x})| \leq R_{\mathrm{m}}(\mathbf{x})$,

$$R_{\mathrm{m}}(\mathbf{x}) = \frac{1}{2}|\mathbf{x} - \mathbf{x}^*|^\top \mathbf{M} |\mathbf{x} - \mathbf{x}^*|, \tag{7}$$

where the absolute value of a vector is applied element-wise.

A sub-approximation of the trust region around an operating point, denoted by $\hat{\mathcal{T}} \subset \mathbb{R}^n$, is characterized by ensuring that the linearization error does not exceed its limit. This is achieved by employing an over-approximation of the Lagrange remainder, as stated in the following corollary.

Corollary 2: If all second-order partial derivatives of the function f are bounded by a bounding-range matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ so that $|\nabla^2 f(\boldsymbol{\xi}(\mathbf{x}))| \leq \mathbf{M}$ for all $\boldsymbol{\xi}(\mathbf{x}) \in \{\mathbf{x}^* + c(\mathbf{x} - \mathbf{x}^*) \mid \mathbf{x} \in \mathcal{X}, \ c \in [0, 1]\}$, then the sub-approximation of the trust region $\hat{\mathcal{T}}(f, \mathbf{x}^*, R_{\mathrm{L}}) \subset \mathcal{T}(f, \mathbf{x}^*, R_{\mathrm{L}})$ can be characterized as a level set of the linearization error limit around the feasible operating point, as follows:

$$\hat{\mathcal{T}}\left(f, \mathbf{x}^*, R_{\mathrm{L}}\right) = \left\{\mathbf{x} : \frac{1}{2}|\mathbf{x} - \mathbf{x}^*|^\top \mathbf{M} |\mathbf{x} - \mathbf{x}^*| \leq R_{\mathrm{L}}\right\}.$$

Proof: The trust region around the operating point $\mathbf{x}^* \in \mathcal{X}$ is defined by the condition that the linearization error is bounded by a limit R_{L}:

$$|R(\mathbf{x})| \leq R_{\mathrm{m}}(\mathbf{x}) \leq R_{\mathrm{L}}, \quad \mathbf{x} \in \mathcal{X}.$$

By using the over-approximation of the Lagrange remainder given in (7), where $|\nabla^2 f(\boldsymbol{\xi}(\mathbf{x}))| \leq \mathbf{M}$ for all $\boldsymbol{\xi}(\mathbf{x}) \in \{\mathbf{x}^* + c(\mathbf{x} - \mathbf{x}^*) \mid \mathbf{x} \in \mathcal{X}, \ c \in [0, 1]\}$, it is straightforward to characterize the region where the linearization error does not exceed its limit, thereby defining $\hat{\mathcal{T}}(f, \mathbf{x}^*, R_{\mathrm{L}})$ as the sub-approximation of the trust region $\mathcal{T}(f, \mathbf{x}^*, R_{\mathrm{L}})$. ■

The quadratic form of the trust region $\hat{\mathcal{T}}$ is symmetric and can be either convex or non-convex, depending on the eigenvalues of the bounding-range matrix $\mathbf{M}$. Due to the symmetry of the quadratic form stemming from the absolute value applied, it is enough to find the approximation of the quadratic form in the first orthant (all elements of $\mathbf{x} - \mathbf{x}^*$ are non-negative in it) and map it to the other orthants.

A potentially non-convex set is sub-approximated by a convex set $\mathcal{T}_{\mathrm{c}}(f, \mathbf{x}^*, R_{\mathrm{L}}) \subset \mathbb{R}^n$ such that $\mathcal{T}_{\mathrm{c}} \subseteq \hat{\mathcal{T}}$. This is achieved by eliminating negative eigenvalues in the orthonormal basis:

$$\mathcal{T}_{\mathrm{c}} = \left\{\mathbf{x} : \frac{1}{2}\left(\mathbf{x} - \mathbf{x}^*\right)^\top \mathbf{M}_{\mathrm{c}}\left(\mathbf{x} - \mathbf{x}^*\right) \leq R_{\mathrm{L}}\right\}, \tag{8}$$

where $\mathbf{M}_{\mathrm{c}} \in \mathbb{R}^{n \times n}$ is a positive semi-definite matrix defined as $\mathbf{M}_{\mathrm{c}} = 0.5\mathbf{V}(\mathbf{D} + |\mathbf{D}|)\mathbf{V}^\top$. Here, the diagonal matrix $\mathbf{D} = \mathbf{V}^\top \mathbf{M} \mathbf{V}$ is obtained through the Gram-Schmidt procedure [14] to find an orthonormal basis in $\mathbb{R}^n$ by the orthogonal matrix $\mathbf{V} \in \mathbb{R}^{n \times n}$.

A. Polytopic Sub-Approximation of the Trust Region

The quadratic form of the trust region sub-approximation is not suitable for use in SLP. Therefore, the following two propositions outline procedures for constructing a polytopic sub-approximation of the trust region as a level set of the linearization error limit around an operating point. The first procedure offers a polytopic sub-approximation using a hyper-rectangle constructed by $2n$ hyperplanes in $\mathbb{R}^n$, whereas the second procedure provides a polytopic sub-approximation using a "gem" shape constructed by 2^n hyperplanes.

A hyper-rectangle is defined by its center (the operating point $\mathbf{x}^* \in \mathbb{R}^n$) and half-lengths along each dimension $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{R}^n$, with $h_i \geq 0$ for $i = 1, 2, \ldots, n$.

Proposition 2: The maximal volume hyper-rectangle $\mathcal{P}_{\mathrm{t}}(f, \mathbf{x}^*, R_{\mathrm{L}}) \subset \mathbb{R}^n$ inscribed within the trust region such that $\mathcal{P}_{\mathrm{t}} \subseteq \mathcal{T}_{\mathrm{c}}$, can be expressed as a level set of the linearization error limit around the feasible operating point, as follows:

$$\mathcal{P}_{\mathrm{t}} = \left\{\mathbf{x} : \mathbf{x}^* - \mathbf{h}(R_{\mathrm{L}}) \leq \mathbf{x} \leq \mathbf{x}^* + \mathbf{h}(R_{\mathrm{L}})\right\}, \tag{9}$$

where $\mathbf{h}(R_L)$ represents the half-lengths as functions of the linearization error limit. The half-lengths are obtained by maximizing the volume inside the convex trust region $\mathcal{T}_c$, as follows:

$$\max_{\mathbf{h}(R_L)} \quad 2^n \prod_{i=1}^{n} h_i(R_L)$$
$$\text{s.t.} \quad \mathbf{h}^T(R_L)\mathbf{M}_c\mathbf{h}(R_L) - 2R_L = 0. \tag{10}$$

Proof: To find the maximal volume hyper-rectangle inscribed within the ellipsoid of the convex trust region defined by (8), the following optimization problem is set up:

$$\max_{\mathbf{h}(R_L)} \quad 2^n \prod_{i=1}^{n} h_i(R_L)$$
$$\text{s.t.} \quad \frac{1}{2}\left(\mathbf{x}-\mathbf{x}^*\right)^\top \mathbf{M}_c\left(\mathbf{x}-\mathbf{x}^*\right) \le R_L,$$
$$\mathbf{x}^* - \mathbf{h}(R_L) \le \mathbf{x} \le \mathbf{x}^* + \mathbf{h}(R_L),$$
$$h_i(R_L) \ge 0, \quad i = 1, 2, \ldots, n. \tag{11}$$

To ensure the hyper-rectangle fits within the ellipsoid and its volume is maximized, the vertices must satisfy the ellipsoid equality, i.e., $\mathbf{h}^T(R_L)\mathbf{M}_c\mathbf{h}(R_L) = 2R_L$. Thus, the optimization problem can be reformulated to (10). ■

The optimization problem (10) can be effectively solved using Lagrange multipliers [15], allowing the half-lengths to be explicitly obtained as functions of the linearization error limit.

Another possible procedure for sub-approximating the trust region using a "gem" shape formed by 2^n half-spaces is introduced below.

Proposition 3: The polytopic sub-approximation of the trust region $\mathcal{P}_t(f, \mathbf{x}^*, R_L) \subset \mathbb{R}^n$, such that $\mathcal{P}_t \subseteq \mathcal{T}_c$, can be obtained using a "gem" shape as a level set of the linearization error limit around the feasible operating point:

$$\mathcal{P}_t = \left\{\mathbf{x} : \operatorname{diag}\left(\mathbf{M} + \mathbf{V}|\mathbf{D}|\mathbf{V}^\top\right)^{\frac{1}{2}} |\mathbf{x} - \mathbf{x}^*| \le 2R_L^{\frac{1}{2}}\right\}, \tag{12}$$

where $\mathbf{D} = \mathbf{V}^\top\mathbf{M}\mathbf{V}$ is a diagonal matrix. The operator $\operatorname{diag}(\cdot)$ returns a row vector of the matrix diagonal, whereas the square root and absolute value operators are applied element-wise.

Proof: To obtain a polytopic sub-approximation of $\mathcal{T}_c$ as a "gem" shape, each orthant is approximated by a half-space as a level set of the linearization error limit around the operating point. The vertices correspond to the maximum possible displacements around the operating point along the coordinate axes, depending on the linearization error limit: $|\Delta x_{\max,j}| = \sqrt{\frac{2R_L}{M_{cj}}}$, $j = 1, \ldots, n$, where M_{cj} is the j^{th} diagonal element of the matrix $\mathbf{M}_c$. Thus, the polytopic level set $\mathcal{P}_t$ can be expressed as $\sum_{j=1}^{n} |\frac{x_j - x_j^*}{\Delta x_{\max,j}}| \le 1$, from where the final form (12) follows. ■

IV. Adaptive Limits and Convergence

This section describes the navigation of the adaptive linearization error limits through SLP iterations, focusing on adjusting the size of the trust radius and the distance between the operating point and the indented half-space.

The trust radius $r_T \in \mathbb{R}_+$ is defined as the shortest distance from the operating point to the boundary of the polytopic sub-approximation of the trust region, as described in (9) or (12). It can be expressed as a function of the linearization error limit in the following form:

$$r_T(R_L) = \sqrt{\frac{R_L}{\alpha(\mathbf{M}_c)}}, \tag{13}$$

where the constant $\alpha(\mathbf{M}_c) \in \mathbb{R}_+$ depends on the matrix $\mathbf{M}_c$. For a polytopic sub-approximation using the hyper-rectangle in (9), $\alpha(\mathbf{M}_c)$ is determined using Lagrange multipliers. For the polytopic sub-approximation represented in (12), $\alpha(\mathbf{M}_c) = \frac{1}{4}\sum_{j=1}^{n} M_{cj}$, where M_{cj} is the j^{th} diagonal element of the matrix $\mathbf{M}_c$.

The minimal distance $d \in \mathbb{R}_+$ between the feasible operating point and the indented half-space can be expressed as a function of the linearization error limit and the operating point:

$$d\left(\mathbf{x}^*, R_L\right) = \frac{|f(\mathbf{x}^*)| - R_L}{\|\mathbf{J}_f\|}, \quad \mathbf{J}_f = \nabla f\left(\mathbf{x}^*\right). \tag{14}$$

The relationship between the trust radius and the minimum distance to the indented half-space is governed by the linearization error limit, as stated in the following proposition.

Proposition 4: The linearization error limit $R_L \in (0, |f(\mathbf{x}^*)|)$ is defined as a function of the feasible operating point, such that the condition $d(\mathbf{x}^*, R_L) = r_T(R_L)$ is satisfied, as follows:

$$R_L\left(\mathbf{x}^*\right) = |f\left(\mathbf{x}^*\right)| + a - \sqrt{a^2 + 2a|f(\mathbf{x}^*)|}, \quad a = \frac{\|\mathbf{J}_f\|^2}{2\alpha(\mathbf{M}_c)}. \tag{15}$$

Proof: The relationship for $R_L(\mathbf{x}^*)$ is defined to maximize the length of the trust radius, which is achieved when the condition $d(\mathbf{x}^*, R_L) = r_T(R_L)$ holds, serving as a solution to the quadratic equation.

To prove that the linearization error limit R_L is correctly defined, it is necessary to demonstrate that R_L takes values in the interval $(0, |f(\mathbf{x}^*)|)$ for all possible values of a. Firstly, R_L is a monotonically decreasing function of a because $R_L' < 0$ for all $a > 0$. Secondly, its behavior is analyzed on its domain:

1. As $a \to 0$, R_L approaches $|f(\mathbf{x}^*)|$.
2. By using L'Hospital rule, it follows that as $a \to \infty$, R_L approaches 0.

Thus, it is established that R_L varies continuously between these bounds. Therefore, for any possible value of a, it follows that:

$$R_L \in \left(0, |f(\mathbf{x}^*)|\right). \quad ■$$

A. Convergence of the TRM

The objective is to ensure convergence in minimizing the cost function to the local optimum, as stated in the following proposition.

Proposition 5: The convergence of the minimization of the cost function is achieved by the following conditions:

$$\nabla g_0\left(\mathbf{x}^*\right)^\top\left(\mathbf{x}-\mathbf{x}^*\right) + R_{L0}\left(\mathbf{x}^*\right) \le 0, \quad \mathbf{x} \in \mathcal{P}_t\left(g_0, \mathbf{x}^*, R_{L0}\right),$$

where the linearization error limit on the cost function is adaptively determined by $R_{\mathrm{L0}}(\mathbf{x}^*) = \frac{\|\nabla g_0(\mathbf{x}^*)\|^2}{\alpha(\mathbf{M}_\mathrm{c})}$.

Proof: To ensure the convergence of the minimization of the cost function, its progress through the iterations must satisfy the following condition:

$$g_0(\mathbf{x}^*) + \nabla g_0(\mathbf{x}^*)^\top(\mathbf{x} - \mathbf{x}^*) + R_{\mathrm{L0}}(\mathbf{x}^*) \le C_{i-1}, \quad (16)$$

where $C_{i-1} = g_0(\mathbf{x}^*)$ represents the interim value of the cost function from the previous iteration $i-1$, while ensuring that the linearization error remains within its limit, i.e., $\mathbf{x} \in \mathcal{P}_\mathrm{t}(g_0, \mathbf{x}^*, R_{\mathrm{L0}})$. The trust radius of the objective function and the distance between the operating point and the constraint (16) are determined as follows:

$$r_\mathrm{T}(R_{\mathrm{L0}}) = \sqrt{\frac{R_{\mathrm{L0}}}{\alpha(\mathbf{M}_\mathrm{c})}}, \quad d(\mathbf{x}^*, R_{\mathrm{L0}}) = \frac{R_{\mathrm{L0}}}{\|\nabla g_0(\mathbf{x}^*)\|}. \quad (17)$$

For $d(\mathbf{x}^*, R_{\mathrm{L0}}) = r_\mathrm{T}(R_{\mathrm{L0}})$, it follows that $R_{\mathrm{L0}}(\mathbf{x}^*) = \frac{\|\nabla g_0(\mathbf{x}^*)\|^2}{\alpha(\mathbf{M}_\mathrm{c})}$. Thus, when the gradient of the objective function is equal to zero, convergence of the solution is achieved, indicating that a local minimum has been found. ■

V. Algorithm Scalability

This section presents the overview and scalability analysis of the proposed enhanced TRMs. The NLP problem (2) is addressed as outlined below:

$$\begin{aligned}
\min_{\mathbf{x}_i \in \mathcal{X}} \quad & g_0(\mathbf{x}_{i-1}^*) + \nabla g_0(\mathbf{x}_{i-1}^*)^\top(\mathbf{x}_i - \mathbf{x}_{i-1}^*) + R_{\mathrm{L0}}(\mathbf{x}_{i-1}^*) \\
\text{s.t.} \quad & \nabla g_0(\mathbf{x}_{i-1}^*)^\top(\mathbf{x}_i - \mathbf{x}_{i-1}^*) + R_{\mathrm{L0}}(\mathbf{x}_{i-1}^*) \le 0, \\
& \mathbf{x}_i \in \mathcal{I}_\cap = \bigcap_{j=1}^{p} \mathcal{I}(g_j, \mathbf{x}_{i-1}^*, R_{\mathrm{Lj}}), \\
& \mathbf{x}_i \in \mathcal{T}_\cap = \bigcap_{j=0}^{p} \mathcal{P}_\mathrm{t}(g_j, \mathbf{x}_{i-1}^*, R_{\mathrm{Lj}}),
\end{aligned} \quad (18)$$

where $\mathbf{x}_0^* \in \mathcal{F}$ is the initial feasible operating point and $i = \{1, 2, \ldots\}$ denotes the iteration index. Linearization error limits are determined based on interim solution so that R_{L0} is determined according to Prop. 5, while $R_{\mathrm{L}j}$ is determined according to Prop. 4. In case the solution approached the constraint defined with the function f_j such that $f_j(\mathbf{x}_{i-1}^*) + f_\varepsilon \ge 0$, where $f_\varepsilon \in \mathbb{R}_+$ is a numerical tolerance, one can try to minimize the criterion along this function with $R_{\mathrm{L}j}(\mathbf{x}_{i-1}^*) := -f_j(\mathbf{x}_{i-1}^*)$. Otherwise, the local convergence is achieved.

The primary computational cost arises from solving the subproblem at each iteration by linear programming, with time efficiency depending on the problem size. The size of the TRM subproblem defined in (18) can be analyzed in relation to the number of decision variables n and the number of linear constraints N_c. Since the two procedures for polytopic sub-approximation of the trust region (propositions 2 and 3) exhibit different complexities, the number of constraints in the linear programs solved through the SLP iterations varies significantly with respect to the approach used.

TABLE I
Total Number of Linear Constraints for TRM Subproblems

SLP Methodology	Number of constraints N_c
Benchmark TRM	$2n + p$
Enhanced TRM (Prop 3.1)	$2n + p + 1$
Enhanced TRM (Prop 3.2)	$2n + (p+1)(2^n + 1)$

When the trust regions are sub-approximated by hyper-rectangles (according to Prop. 2), the intersected trust region $\mathcal{T}_\cap$ can be simplified as follows:

$$\begin{aligned}
\mathcal{T}_\cap &= \{\mathbf{x} : \mathbf{x}^* - \mathbf{h}_{\min} \le \mathbf{x} \le \mathbf{x}^* + \mathbf{h}_{\min}\}, \\
\mathbf{h}_{\min} &= \min\left(\mathbf{h}_0(R_{\mathrm{L0}}), \mathbf{h}_1(R_{\mathrm{L1}}), \ldots, \mathbf{h}_p(R_{\mathrm{L}p})\right),
\end{aligned} \quad (19)$$

where the operator $\min(\cdot)$ returns the minimal half-lengths along each dimension, identifying the n minimum values from $p+1$ vectors. Thus, the first approach (Prop. 2) provides a minimal representation of the intersected trust region using $2n$ half-spaces at each iteration, same as benchmark TRM. In contrast, the second approach (Prop. 3) does not efficiently yield a minimal half-space representation of the intersected trust region, resulting in a total of $(p+1)2^n$ half-spaces. Table I summarizes the concerned linear program size, i.e., the total number of linear constraints in it, comparing the proposed enhanced TRMs with the benchmark TRM. The presented numbers reflect the worst-case scenario, where the cost and all constraints are strictly nonlinear functions.

Consequently, the enhanced TRM based on Prop. 2 exhibits a comparable time complexity to the benchmark TRM, while additionally ensures feasibility at each iteration. Since the SLP with the enhanced TRM will never result in an interim infeasible solution and since it uses large feasible regions by selecting linearization errors according to Prop. 4, it is expected to converge faster to the SLP solution compared to the SLP with benchmark TRM. Although the second enhanced TRM provides better sub-approximation of the trust region and thus larger feasible regions, the interim LP complexity is compromised as the dimension n and number of constraints p in the original NLP increase. Therefore, its benefits are emphasized for problems with low p and n.

VI. Case Study

This section presents two examples of solving NLP problems using the proposed enhanced TRMs and comparing them to the benchmark TRM (1).

1) The First Example: The following constrained NLP problem is considered:

$$\begin{aligned}
\min_{x_1, x_2} \quad & -x_1^4 x_2 \\
\text{s.t.} \quad & x_1 \ge 0, \ x_2 \ge 0, \\
& g_1(x_1, x_2) = x_1^2 + x_2^2 - 25 \le 0, \\
& g_2(x_1, x_2) = 4e^{0.08(x_1+1)} + x_2 - 7 \le 0.
\end{aligned}$$

Both the proposed enhanced TRMs and the benchmark TRM were applied to solve this NLP problem for 100 randomly generated initial feasible points, $\mathbf{x}_{0k}^* \in \mathcal{F}$, $k =$

TABLE II
PERFORMANCE METRICS TABLE FOR THE FIRST EXAMPLE

TRM	N_c	$\bar{n}_i$	$\bar{t}$ [ms]
Benchmark	6	74	22.8
Enhanced (Prop 3.1)	7	15	4.7
Enhanced (Prop 3.2)	19	11	3.2

$1, \dots, 100$ where the optimal solution $\mathbf{x}^* = (4.8, 6)$ is calculated.

All three methods successfully minimized the cost function at $\mathbf{x}^* = (4.8, 0.6)$, achieving a cost value of $g_0(\mathbf{x}^*) = -318.5$, although with notable differences in performance. Performance metrics, including the number of constraints N_c in the SLP subproblem, the average number of iterations $\bar{n}_i$ and average total computation times $\bar{t}$, are summarized in Table II. The computation time reflects the entire online procedure; for the proposed enhanced TRMs, it includes the time taken to solve all linear programs in the subproblems and to calculate adaptive linearization error limits until convergence is achieved. The computation time of the benchmark TRM additionally includes the time required to find the feasible region at each SLP iteration. All computations were conducted using MATLAB with IBM ILOG CPLEX on a computer equipped with an AMD Ryzen 7 PRO 4750U processor, running at a clock speed of 1.7 GHz.

The results align with expectations. The enhanced TRM featuring a "gem"-shaped trust region demonstrated the best performance for this example. This can be attributed to the small size of the NLP problem, where all three methods exhibit similar time complexity. The "gem"-shaped trust region allows for larger volume trust regions, facilitating faster convergence. In contrast, the benchmark TRM encountered numerous infeasible subproblems while determining the feasible region at each iteration. Its initial trust region is set to 10% of the maximum range of the feasible region, with a changing factor of 10% applied to adjust the trust region for feasibility between SLP iterations.

2) The Second Example: The optimization problem for determining the State of Energy (SoE) for battery cell from [16] is considered:

$$\begin{aligned}
\max_{\bar{p}_0,\dots,\bar{p}_{N-1}} \quad & \pm \sum_{i=0}^{N-1} \bar{p}_i, \\
\text{s.t.} \quad & \mathbf{x}(0) = \mathbf{x}_0, \quad \mathbf{x}_{i+1} = \mathbf{f}(\mathbf{x}_i, \bar{p}_i), \\
& \bar{p}_i \in [P_{\min}, P_{\max}], \quad i_{\text{bat},i} \in [I_{\min}, I_{\max}], \\
& u_{\text{bat},i} \in [U_{\min}, U_{\max}], \quad u^{-}_{\text{bat},i} \in [U_{\min}, U_{\max}] \\
& i = 0, 1, \dots, N-1,
\end{aligned}$$

where $\mathbf{x} \in \mathbb{R}^3$ is the system state, $\bar{p}$ is the average power input, u_{bat} and i_{bat} are battery voltage and current, respectively.

The test was conducted with a prediction horizon of $N = 10$, indicating the use of 10 optimization variables, and involved 20 different initial states $\mathbf{x}_0$. Table III summarizes the performance metrics for this numerical experiment. Notably, the enhanced trust region method (TRM) utilizing a "hyper-rectangle" shape outperforms the others, as it swiftly identifies the minimum and maximum admissible trajectories for average powers on the prediction horizon. The second enhanced TRM with a "gem" shape achieves optimal results in 17 iterations, capitalizing on the largest trust region; however, it also entails a more complex subproblem with 110 constraints.

TABLE III
PERFORMANCE METRICS TABLE FOR THE SECOND EXAMPLE

TRM	N_c	$\bar{n}_i$	$\bar{t}$ [ms]
Benchmark	38	84	89.6
Enhanced (Prop 3.1)	38	21	22.1
Enhanced (Prop 3.2)	110	17	34

VII. CONCLUSION

In this letter, we propose an enhanced Trust Region Method (TRM) for Sequential Linear Programming (SLP). Subject to initial solution feasibility, it guarantees that all linear programs through the SLP iterations will result in a feasible solution for the original nonlinear program that is approximately solved by the SLP. The interim feasible regions through iterations are kept large by adaptively changing the linearization error limits used in the construction of trust regions, and this effectively reduces the number of LP iterations and the overall computation time needed to converge to the SLP solution.

REFERENCES

[1] F. Palacios-Gomez, L. Lasdon, and M. Engquist, "Nonlinear optimization by successive linear programming," *Manag. Sci.*, vol. 28, no. 10, pp. 1106–1120, 1982.

[2] J. Nocedal and S. J. Wright, "Line search methods," in *Numerical Optimization*, New York, NY, USA: Springer, 1999, pp. 34–63.

[3] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust Region Methods*. Philadelphia, PA, USA: SIAM, 2000.

[4] L. Nita and E. Carrigan, "SEQUOIA: A sequential algorithm providing feasibility guarantees for constrained optimization," in *Proc. 22nd IFAC World Congr.*, vol. 56, 2023, pp. 4840–4845.

[5] F. Messerer, K. Baumgartner, and M. Diehl, "Survey of sequential convex programming and generalized Gauss-Newton methods," *ESAIM: Proc. Surv.*, vol. 71, pp. 64–88, Sep. 2021.

[6] D. Kiessling et al., "A feasible sequential linear programming algorithm with application to time-optimal path planning problems," in *Proc. IEEE 61st Conf. Decision Control (CDC)*, 2022, pp. 1196–1203.

[7] D. Kiessling, P. Pas, A. Astudillo, P. Patrinos, and J. Swevers, "Anderson accelerated feasible sequential linear programming," in *Proc. 22nd IFAC World Congr.*, vol. 56, 2023, pp. 7436–7441.

[8] D. Kiessling, C. Vanaret, A. Astudillo, W. Decré, and J. Swevers, "An almost feasible sequential linear programming algorithm," in *Proc. Eur. Control Conf. (ECC)*, 2024, pp. 2328–2335.

[9] L. Sabug, L. Fagiano, and F. Ruiz, "Sample-based trust region dynamics in contextual global optimization," *IEEE Control Syst. Lett.*, vol. 8, pp. 1619–1624, 2024.

[10] K. Cao and L. Xie, "Trust-region inverse reinforcement learning," *IEEE Trans. Autom. Control*, vol. 69, no. 2, pp. 1037–1044, Feb. 2024.

[11] J. Wang, L. Hemelhof, I. Markovsky, and P. Patrinos, "A trust-region method for data-driven iterative learning control of nonlinear systems," *IEEE Control Syst. Lett.*, vol. 8, pp. 1847–1852, 2024.

[12] L. Grippo and M. Sciandrone, "Trust region methods," in *Introduction to Methods for Nonlinear Optimization*, vol. 152. Cham, Switzerland: Springer, 2023, pp. 297–324.

[13] M. Berz and G. Hoffstätter, "Computation and application of Taylor polynomials with interval remainder bounds," *Reliab. Comput.*, vol. 4, pp. 83–97, Feb. 1998.

[14] L. Beilina, E. Karchevskii, and M. Karchevskii, *Numerical Linear Algebra: Theory and Applications*. Cham, Switzerland: Springer, 2017.

[15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[16] D. Leko and M. Vašak, "Robust and computationally efficient estimation of state of power and maximum energy for lithium-ion batteries in electric vehicles," in *Proc. IEEE Conf. Control Technol. Appl. (CCTA)*, 2023, pp. 368–375.