SAND2026-18922B

Encyclopedia of Optimization manuscript No.
(will be inserted by the editor)

Trust-region methods with inexact and adaptive computations

Robert J. Baraldi · Drew P. Kouri · Denis Ridzal

the date of receipt and acceptance should be inserted later

Abstract Keywords Nonconvex Optimization · Trust Region · Optimal Control · Inexactness · PDE-Constrained Optimization · Nonsmooth Optimization

Mathematics Subject Classification (2020) 49M15, 49M37, 65K05, 65K10, 90C06, 90C30

1 Introduction

Optimization problems involving large-scale, nonlinear simulations are fraught with computational challenges stemming from memory and arithmetic limitations. For problems constrained by partial differential equations (PDEs), simply evaluating the objective function requires the solution of the governing PDEs, which in turn necessitates the use of discretizations along with iterative linear and nonlinear solvers. These features lead to various forms of inexactness when evaluating the objective function, the constraints, their derivatives, and other derived quantities.

Drew P. Kouri
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87125, USA
Tel.: 1-505-845-8127
Fax.: 1-630-252-5986
E-mail: dpkouri@sandia.gov

Robert J. Baraldi
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87125, USA
E-mail: rjbaral@sandia.gov

Denis Ridzal
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87125, USA
E-mail: dridzal@sandia.gov

Consider the general optimization problem

$$\min_{x \in X} f(x) + \phi(x) \qquad \text{subject to} \qquad c(x) = 0, \tag{1}$$

where X is a Hilbert space, $f : X \to \mathbb{R}$ is a smooth function, $\phi : X \to (-\infty, +\infty]$ is proper, closed and convex, but potentially nonsmooth, and $c : X \to Y$ is a smooth constraint map with Y a Hilbert space. As previously discussed, one often must discretize (1) when X or Y is infinite dimensional, or if the evaluations of f, ϕ, or c require, e.g., the evaluation of integrals or solutions of differential equations. For example, in simulation-constrained optimization, the smooth objective function typically has the composite form

$$f(x) = J(S(x), x),$$

where $S(x)$ is the solution to a system of PDEs, ordinary differential equations (ODEs), or differential algebraic equations (DAEs). Given the functional nature of S, one typically computes an approximation using finite elements for PDEs [34] or time-stepping for ODEs and DAEs [12, 27]. The discretized system of differential equations is then solved using iterative linear and nonlinear solvers [6, 21]. In stochastic optimization, the evaluation of f and ϕ typically requires the computation of high-dimensional integrals, which can be estimated using Monte Carlo sampling or numerical quadrature [7, 18, 23, 24, 25]. Additionally, the objective function in dynamic programming depends on the trajectory of an often large-scale dynamical system, which can be prohibitive to store for derivative computations. Inexactness can stem from measures taken to overcome this memory burden, such as the use of reduced-order models [5, 17, 33, 35] or lossy compression [19, 20, 30].

Algorithms for solving (1) must rigorously control these approximations to guarantee convergence to a stationary point. Trust-region methods are popular choices for solving large-scale, nonconvex and multidisciplinary optimization problems because they provide strong global convergence properties in the presence of inexactness. Pioneering papers by Moré [26, 28], Toint [31], and Carter [9, 10, 11] first articulated the use of inexact gradients and objective function evaluations in trust-region methods. Later, Dennis, Alexandrov, Lewis and Torczon analyzed trust-region algorithms as general frameworks for managing approximations throughout the optimization iteration [1, 2, 3, 14, 15]. These early methods were foundational for recent work on adaptive approximations, where the trust-region framework manages the accuracy of the objective and constraint evaluations, gradient computations, Jacobian applications, and linear system solves.

Heinkenschloss and Vicente [22] developed an inexact composite-step trust-region sequential quadratic programming (SQP) method, where inexactness stems from iterative linear system solves or approximations of derivatives, and accuracy is tuned based on the feasibility and optimality of iterates. This work, motivated by optimal control, optimal design and parameter identification problems, applies to the case in which the optimization variables admit a natural splitting into state and control variables, i.e., basic and non-basic variables. In [34], Ulbrich and Ziems adapted and specialized the work in [22] to PDE-constrained optimization by enabling adaptive mesh refinement. Heinkenschloss and Ridzal [21] generalized [22] to handle problem formulations where the state-control variable splitting may not be applicable or desired, and opened the door to handling inexact solutions of Karush-Kuhn-Tucker (KKT) systems in composite-step SQP methods. In [32], Walther also considered problem formulations without the splitting restriction, and developed an SQP approach that does not require exact evaluations of the constraint Jacobian. The approach in [32] is best suited for optimization problems of moderate size, with constraint Jacobians that possess special structure, e.g., dense Jacobians.

In [24, 25], Kouri et al. introduced a trust-region algorithm for smooth unconstrained optimization that manages inexactness in the objective function evaluation and gradient. Largely motivated by [9, 10, 11, 22, 34], Kouri et al. alleviated the dependence of the traditional inexactness criteria on constants tied to other algorithmic parameters, enabling the use of general error indicators that are accurate up to unknown or uncomputable constants. They applied their algorithm to adaptive numerical quadrature for stochastic programming using anisotropic dimension-adaptive sparse grids. Garreis and Ulbrich [18] built on this work, by allowing for convex constraints on x with inexact projections. Further applications and extensions of [24, 25] include [30] which considers adaptive compression for dynamic problems, [33, 35] which apply trust regions to adaptive numerical integration and reduced-order modeling, and [6] which extends the algorithms in [18, 29, 31] to handle more general nonsmooth objective functions as in (1).

These existing trust-region methods have also seen use as subproblem solvers within penalty frameworks such as augmented Lagrangian. Notably, the method of Heinkenschloss and Ridzal [21], which leverage iterative and inexact linear system solves is used to solve equality-constrained subproblems within the ALESQP framework [4].

2 Inexact Trust-Region Methods

This article highlights the treatment of inexact computations in trust-region methods through brief reviews of the algorithms presented in [6] and [21]. The algorithm in [6] is applicable to (1) without the equality constraint $c(x) = 0$. In particular, this algorithm is applicable for smooth unconstrained, convex-constrained and nonsmooth regularized problems. At each iteration, the trust-region method generates a trial iterate x_k^+ as an approximate solution to the trust-region subproblem

$$\min_{x \in X} \; f_k(x) + \phi(x) \qquad \text{subject to} \qquad \|x - x_k\| \le \Delta_k, \tag{2}$$

where $x_k \in X$ is the current iterate, $f_k : X \to \mathbb{R}$ is a local model of f around x_k, such as a quadratic Taylor model, and $\Delta_k > 0$ is the trust-region radius. The trial iterate x_k^+ must be feasible with respect to the trust-region constraint and must satisfy the fraction of Cauchy decrease condition, cf. [6, 13]. For example, when f_k is twice Fréchet differentiable and the sequence of Hessians is bounded, i.e., there exists $\kappa_{\text{bnd}} \in (0, 1)$ such $\|\nabla^2 f_k(x_k)\| \le (1 - \kappa_{\text{bnd}})/\kappa_{\text{bnd}}$, we require that x_k^+ satisfies

$$\|x_k^+ - x_k\| \le \kappa_{\text{rad}} \Delta_k$$

and

$$(f_k(x_k) - \phi(x_k)) - (f_k(x_k^+) - \phi(x_k^+)) \ge \kappa_{\text{fcd}} h_k \min\{\kappa_{\text{bnd}} h_k, \Delta_k\},$$

where κ_{rad} and κ_{fcd} are positive constants and

$$h_k := \tfrac{1}{r} \|x_k - \operatorname{prox}_{r\phi}(x_k - r \nabla f_k(x_k))\|,$$

for fixed $r > 0$. Here, $\mathrm{prox}_{r\phi}$ denotes the usual proximity operator [8, chapter 24].

Given an acceptable trial iterate x_k^+, the trust-region algorithm accepts or rejects x_k^+ based on the ratio of *actual* and *predicted* reduction

$$\rho_k^* := \frac{\mathrm{ared}_k}{\mathrm{pred}_k},$$

where

$$\mathrm{ared}_k := (f(x_k) + \phi(x_k)) - (f(x_k^+) + \phi(x_k^+))$$

and

$$\mathrm{pred}_k := (f_k(x_k) + \phi(x_k)) - (f_k(x_k^+) + \phi(x_k^+)).$$

Here, ared_k is the reduction of the objective function achieved by x_k^+ relative to x_k and pred_k is the reduction of the model. In many applications of practical importance, the objective function cannot be computed accurately [13, 18, 25, 23, 35]. Instead, ared_k is replaced by an approximation denoted cred_k—the *computed reduction.* The algorithm instead decides whether or not to accept x_k^+ based on the ratio of computed and predicted reduction

$$\rho_k := \frac{\mathrm{cred}_k}{\mathrm{pred}_k}. \tag{3}$$

If $\rho_k \ge \eta_1$, we accept the trial iterate $x_{k+1} = x_k^+$. Otherwise, we reject it, $x_{k+1} = x_k$, resulting in a null step. The trust-region algorithm then increases the radius Δ_k if $\rho_k \ge \eta_2$ and reduces Δ_k if $\rho_k < \eta_1$. The algorithmic parameters $0 < \eta_1 < \eta_2 < 1$ are user-specified with common values $\eta_1 = 0.05$ and $\eta_2 = 0.9$. To ensure that cred_k is a sufficiently accurate approximation of ared_k, we require satisfaction of the following condition: *There exists a constant* $\kappa_{\mathrm{obj}} \ge 0$ *such that*

$$|\mathrm{ared}_k - \mathrm{cred}_k| \le \kappa_{\mathrm{obj}}[\eta \min\{\mathrm{pred}_k, \theta_k\}]^\zeta \quad \forall\, k, \tag{4}$$

where ζ*,* η*, and* θ_k *are (user-specified) positive real numbers that satisfy*

$$\zeta > 1, \qquad 0 < \eta < \min\{\eta_1, (1-\eta_2)\}, \qquad \text{and} \qquad \lim_{k\to\infty} \theta_k = 0.$$

Similarly, the gradient $\nabla f(x_k)$ often cannot be computed exactly. Instead, we require that the model gradient $\nabla f_k(x_k)$ satisfies: *There exists a constant* $\kappa_{\mathrm{grad}} \ge 0$ *such that*

$$\|\nabla f_k(x_k) - \nabla f(x_k)\| \le \kappa_{\mathrm{grad}} \min\{h_k,\ \Delta_k\}. \tag{5}$$

The key feature of (4) and (5) is their dependence on the arbitrary constants κ_{obj} and κ_{grad}, which enables the use of general error indicators that are accurate up to unknown constants. The trust-region algorithm from [6] is stated in Algorithm 1.

The method presented in [21] assumes that $\phi \equiv 0$ and instead focuses on the constraint $c(x) = 0$ in (1). In this setting, the trust-region subproblem includes a linearization of $c(x) = 0$ about x_k, taking the form

$$\min_{s\in X}\ \ell_k(s) \qquad \text{subject to} \qquad c'(x_k)s + c(x_k) = 0,\ \ \|s\| \le \Delta_k, \tag{6}$$

where ℓ_k is a model of the Lagrangian $\ell(x,\lambda) := f(x) + \langle \lambda, c(x)\rangle$, e.g.,

$$\ell_k(s) := \ell(x_k, \lambda_k) + \langle \nabla_x \ell(x_k,\lambda_k), s\rangle + \frac{1}{2}\langle B_k s, s\rangle,$$

Algorithm 1 Trust-region Algorithm with Inexact Computations

Require: Initial guess $x_1 \in \operatorname{dom}\phi$, initial radius $\Delta_1 > 0$, $0 < \eta_1 < \eta_2 < 1$, and $0 < \gamma_1 \leq \gamma_2 < 1 \leq \gamma_3$, tolerance $\tau > 0$.

for $k = 1, 2, \ldots$ **do**

Convergence Check: If $h_k < \tau$, then terminate.

Model Selection: Choose a model f_k that satisfies (5).

Step Computation: Compute $x_k^+ \in X$ that *approximately* solves (2).

Computed Reduction: Compute cred_k that satisfies (4).

Step Acceptance and Radius Update:

i: Compute ρ_k as in (3).

ii: If $\rho_k < \eta_1$, set $x_{k+1} \leftarrow x_k$ and choose $\Delta_{k+1} \in [\gamma_1 \Delta_k, \gamma_2 \Delta_k]$.
Otherwise, set $x_{k+1} \leftarrow x_k^+$ and choose Δ_{k+1} such that

$$\begin{cases} \Delta_{k+1} \in [\gamma_2 \Delta_k, \Delta_k], & \text{if } \rho_k \in [\eta_1, \eta_2); \\ \Delta_{k+1} \in [\Delta_k, \gamma_3 \Delta_k], & \text{if } \rho_k \geq \eta_2. \end{cases}$$

end for

with an approximation $B_k = B(x_k, \lambda_k)$ of the Hessian of the Lagrangian. To handle the potential inconsistency of the constraints in (6), the step s is split into the so-called tangential and quasi-normal components, motivated by the Byrd-Omojokun composite-step approach [13, section 15.4.2]. The quasi-normal step aims to reduce linearized infeasibility by approximately solving

$$\min_{n \in X} \; \|c'(x_k)n + c(x_k)\|^2 \qquad \text{subject to} \qquad \|n\| \leq \zeta \Delta_k,$$

where $\zeta \in (0, 1)$. Under the assumptions of [21], n_k must satisfy the condition

$$\|n_k\| \leq \kappa_{\mathrm{qn}} \|c(x_k)\|, \tag{7}$$

where $\kappa_{\mathrm{qn}} > 0$ is independent of k, and the fraction of Cauchy decrease condition

$$\|c(x_k)\|^2 - \|c'(x_k)n_k + c(x_k)\|^2 \geq \kappa_{\mathrm{fcd}} \|c(x_k)\| \min\left\{\kappa_{\mathrm{bnd}} \|c(x_k)\|, \Delta_k\right\}, \tag{8}$$

where $\kappa_{\mathrm{fcd}}, \kappa_{\mathrm{bnd}} > 0$ are independent of k. These conditions were derived in [16]. The purpose of the tangential step t_k is to reduce the model ℓ_k while maintaining the linearized feasibility obtained by n_k. This is achieved via the subproblem

$$\min_{t \in X} \; \ell_k(n_k + t) \qquad \text{subject to} \qquad c'(x_k)t = 0, \;\; \|n_k + t\| \leq \Delta_k. \tag{9}$$

To compute the tangential step t_k, the linear constraint $c'(x_k)t = 0$ in (9) is eliminated using a null-space representation of $c'(x_k)$, denoted $W_k : Z \to X$ with the property

$$\mathrm{Range}(W_k) = \mathrm{Null}(c'(x_k)).$$

Here, Z is a Hilbert space. The application of the operator W_k often necessitates the iterative, hence inexact, solution of very large linear systems. Additionally, in PDE-constrained optimization, the application of W_k requires the solution of systems involving the linearized PDEs, where the operator $c'(x_k)$ may not be formed explicitly, making iterative linear system solvers the only option for applying W_k. In summary, managing the inexactness in the application of W_k is the principal challenge in solving (9).

Inexactness in W_k gives rise to the linear operator $\widetilde{W}_k$, which approximates W_k. After introducing the quantity $g_k \coloneqq \nabla_x \ell(x_k, \lambda_k) + B_k n_k$, it is shown in [21] that the tangential-step subproblem (9) may be replaced with the subproblem

$$\min_{\widetilde{t} \in X} \ \frac{1}{2}\langle B_k \widetilde{t}, \widetilde{t}\rangle + \langle \widetilde{W}_k g_k, \widetilde{t}\rangle \quad \text{subject to} \quad \widetilde{t} \in \operatorname{Range}(\widetilde{W}_k),\ \|n_k + \widetilde{t}\| \le \Delta_k, \tag{10}$$

which motivates the quadratic model $\widetilde{q}_k(t) \coloneqq \frac{1}{2}\langle B_k t, t\rangle + \langle \widetilde{W}_k g_k, t\rangle$. The quantity $\widetilde{W}_k g_k$, called the inexact reduced gradient, must satisfy

$$\|\widetilde{W}_k g_k - W_k g_k\| \le \kappa_{\text{grad}} \min\left\{\|\widetilde{W}_k g_k\|, \Delta_k\right\}, \tag{11}$$

for some $\kappa_{\text{grad}} > 0$ independent of k. We also require the fraction of Cauchy decrease condition for $\widetilde{t}_k$ and the quadratic model $\widetilde{q}_k$, specifically

$$\widetilde{q}_k(0) - \widetilde{q}_k(\widetilde{t}_k) \ge \kappa_{\text{fcd}} \|\widetilde{W}_k g_k\| \min\left\{\kappa_{\text{bnd}}\|\widetilde{W}_k g_k\|, \Delta_k\right\}, \tag{12}$$

where $\kappa_{\text{fcd}}, \kappa_{\text{bnd}} > 0$ are independent of k.

Progress of the composite-step algorithm is assessed with the help of the augmented Lagrangian merit function

$$\mathcal{L}(x, \lambda; \rho) \coloneqq \ell(x, \lambda) + \rho\|c(x)\|^2.$$

Let λ_{k+1} be a Lagrange multiplier estimate for the trial point $x_k + s_k$, where $s_k = n_k + t_k$. The global convergence theory for the composite-step trust-region method requires that the sequence $\{\lambda_k\}_{k\in\mathbb{N}}$ be bounded, which is easily satisfied. Without inexactness in W_k, the progress can be evaluated based on the ratio of the actual reduction

$$\operatorname{ared}(s_k; \rho_k) = \mathcal{L}(x_k, \lambda_k; \rho_k) - \mathcal{L}(x_k + s_k, \lambda_{k+1}; \rho_k)$$

and the predicted reduction

$$\begin{aligned}\operatorname{pred}(s_k; \rho_k) = \mathcal{L}(x_k, \lambda_k; \rho_k) - \Big[&\ell(x_k, \lambda_k) + \langle \nabla_x \ell(x_k, \lambda_k), s_k\rangle + \frac{1}{2}\langle B_k s_k, s_k\rangle \\ &+ \langle \lambda_{k+1} - \lambda_k, c'(x_k)s_k + c(x_k)\rangle + \rho_k\|c'(x_k)s_k + c(x_k)\|^2\Big].\end{aligned}$$

Additionally, without inexactness, it is correct to identify t_k with $\widetilde{t}_k$, the approximate solution of the tangential subproblem (10), and set $s_k = n_k + t_k = n_k + \widetilde{t}_k$. With inexactness in W_k, i.e., with $\widetilde{W}_k$ replacing W_k, the computed tangential step t_k is no longer in the null space of $c'(x_k)$, resulting in the following observation

$$c'(x_k)s_k = c'(x_k)(n_k + t_k) = c'(x_k)n_k + r_k^t \ne c'(x_k)n_k,$$

where $r_k^t \coloneqq c'(x_k)t_k$. To restore a sufficent amount of linearized constraint feasibility, the tangential step t_k is redefined, and computed by post-processing $\widetilde{t}_k$. Subsequently, the *predicted reduction with inexactness* is introduced,

$$\widehat{\operatorname{pred}}(s_k; \rho_k) \coloneqq \operatorname{pred}(n_k, \widetilde{t}_k; \rho_k) + \operatorname{rpred}(r_k^t; \rho_k),$$

where

$$\begin{aligned}\mathrm{pred}(n_k,\widetilde{t}_k;\rho_k) \coloneqq &-\langle \widetilde{W}_k g_k, \widetilde{t}_k\rangle - \frac{1}{2}\langle B_k\widetilde{t}_k,\widetilde{t}_k\rangle\\ &-\langle \nabla_x \ell(x_k,\lambda_k), n_k\rangle - \frac{1}{2}\langle B_k n_k, n_k\rangle\\ &-\langle \lambda_{k+1}-\lambda_k, c'(x_k)n_k + c(x_k)\rangle\\ &+\rho_k\left(\|c(x_k)\|^2 - \|c'(x_k)n_k + c(x_k)\|^2\right),\end{aligned}$$

and

$$\mathrm{rpred}(r_k^t;\rho_k) \coloneqq -\langle \lambda_{k+1}-\lambda_k, r_k^t\rangle - \rho_k\|r_k^t\|^2 - 2\rho_k\langle r_k^t, c'(x_k)n_k + c(x_k)\rangle.$$

To ensure global convergence, a penalty parameter ρ_k is first computed to satisfy

$$\mathrm{pred}(n_k,\widetilde{t}_k;\rho_k) \geq \frac{\rho_k}{2}\left(\|c(x_k)\|^2 - \|c'(x_k)n_k + c(x_k)\|^2\right).$$

Then, an approximate null-space projection is applied to $\widetilde{t}_k$, to generate the tangential step t_k satisfying

$$|\mathrm{rpred}(r_k^t;\rho_k)| \leq \eta_0\,\mathrm{pred}(n_k,\widetilde{t}_k;\rho_k), \tag{13}$$

where $\eta_0 \in (0, 1-\eta_1)$, and $\eta_1 \in (0,1)$ is the smallest acceptable ratio of the actual and predicted reduction. Finally, the tangential step t_k must satisfy two additional conditions, which ensure that it remain sufficiently close to the exact projection of $\widetilde{t}_k$ and that $\widetilde{t}_k$ remain bounded by s_k, given by

$$\|t_k - W_k\widetilde{t}_k\| \leq \kappa_{\mathrm{proj}}\Delta_k\min\{\Delta_k,\|s_k\|\} \qquad \text{and} \qquad \|\widetilde{t}_k\| \leq \kappa_{\mathrm{bnd}}\|s_k\|, \tag{14}$$

for some $\kappa_{\mathrm{proj}}, \kappa_{\mathrm{bnd}} > 0$ independent of k. The general framework for a composite-step trust-region method with inexact constraint null-space computations is included as Algorithm 2. Specific implementations are provided in [21]. The algorithmic framework developed in the early work [22] can be thought of as a specialization of Algorithm 2 to the case where the optimization variables x can be naturally split into basic and nonbasic variables, as is common in optimal control.

3 Open Problems

The two algorithms described here apply to subclasses of the more general problem (1). To our knowledge, there are no inexact trust-region methods that can directly handle (1), nor is there a trust-region algorithm for the more general problem

$$\min_{x\in X}\; f(x) + \psi(g(x)), \tag{15}$$

where $\psi : Y \to (-\infty,+\infty]$ is nonsmooth and $g : X \to Y$ is smooth. Note that (15) encapsulates (1) as well as other optimization problems arising in nonlinear programming and risk-averse optimization. Finally, a more thorough treatment of inexact projections and proximity operators is needed. In [18], Garreis and Ulbrich provided a starting point for this line of research. This is an important direction since many projections and proximity operators cannot be computed analytically, especially when X is an arbitrary Hilbert space, not simply $\mathbb{R}^n$ or L^2.

Algorithm 2 Composite-step Trust-region Algorithm with Inexact Computations

Require: Initial guess x_0, initial trust-region radius Δ_0, constants $0 < \alpha_1, \eta_1 < 1$, $0 < \eta_0 < 1 - \eta_1$, $\rho_{-1} \geq 1$, $\bar{\rho} > 0$, and tolerance $\tau > 0$, initial Lagrange multiplier estimate λ_0.

for $k = 1, 2, \dots$ **do**

Convergence Check: If $\|\nabla_x \ell(x_k, \lambda_k)\| < \tau$ and $\|c(x_k)\| < \tau$, then terminate.

Step Computation:

i: Compute a quasi-normal step n_k satisfying (7) and (8).

ii: Compute $\widetilde{t}_k$ satisfying (11) and (12).

Step Acceptance and Radius Update:

i: Compute a new Lagrange multiplier estimate λ_{k+1}.

ii: Update the penalty parameter: If

$$\operatorname{pred}(n_k, \widetilde{t}_k; \rho_{k-1}) \geq \frac{\rho_{k-1}}{2} \left(\|c(x_k)\|^2 - \|c'(x_k) n_k + c(x_k)\|^2 \right),$$

then set $\rho_k \leftarrow \rho_{k-1}$; otherwise, set

$$\rho_k \leftarrow \frac{-2 \operatorname{pred}(n_k, \widetilde{t}_k; \rho_{k-1})}{\|c(x_k)\|^2 - \|c'(x_k) n_k + c(x_k)\|^2} + 2\rho_{k-1} + \bar{\rho}.$$

iii: Compute the tangential step t_k satisfying the conditions (13) and (14).

iv: Compute the trial step $s_k \leftarrow n_k + t_k$.

v: Compute $\theta_k \leftarrow \operatorname{ared}(s_k; \rho_k) / \operatorname{pred}(n_k, \widetilde{t}_k; \rho_k)$.

vi: If $\theta_k \geq \eta_1$, set $x_{k+1} \leftarrow x_k + s_k$, and choose $\Delta_{k+1} \geq \Delta_k$.
Otherwise, set $x_{k+1} \leftarrow x_k$, $\lambda_{k+1} \leftarrow \lambda_k$, and $\Delta_{k+1} \leftarrow \alpha_1 \|s_k\|$.

end for

References

1. Alexandrov N (1996) Robustness properties of a trust-region framework for managing approximation models in engineering optimization. In: AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, pp 1056–1059
2. Alexandrov N (1998) A trust-region framework for managing approximations in constrained optimization problems. In: First ISSMO/NASA Internet Conference on Approximation and Fast Reanalysis Techniques in Engineering Optimization
3. Alexandrov N, Dennis J, Lewis R, Torczon V (1998) A trust-region framework for managing the use of approximation models in optimization. Structural Optimization 15:16–23
4. Antil H, Kouri DP, Ridzal D (2023) ALESQP: An augmented Lagrangian equality-constrained SQP method for optimization with general constraints. SIAM Journal of Optimization 33(1):237–266
5. Arian E, Fahl M, Sachs EW (2000) Trust-region proper orthogonal decomposition for flow control. 2000-2025, Institute for Computer Applications in Science and Engineering, NASA Langley
6. Baraldi RJ, Kouri DP (2022) A proximal trust-region method for nonsmooth optimization with inexact function and gradient evaluations. Mathematical Programming 201(1):1–40
7. Bastin F, Cirillo C, Toint PL (2006) An adaptive Monte Carlo algorithm for computing mixed logit estimators. Computational Management Science 3:55–79

8. Bauschke HH, Combettes PL (2011) Convex analysis and monotone operator theory in Hilbert spaces, vol 408. Springer
9. Carter RG (1989) Numerical optimization in Hilbert space using inexact function and gradient information. Tech. Rep. 89–45, ICASE, Langely, VA
10. Carter RG (1991) On the global convergence of trust region algorithms using inexact gradient information. SIAM Journal on Numerical Analysis 28(1):251–265
11. Carter RG (1993) Numerical experience with a class of algorithms for nonlinear optimization using inexact function and gradient information. SIAM Journal of Scientific Computing 14:368–388
12. Clever D, Lang J, Ulbrich S, Ziems C (2012) Generalized Multilevel SQP-methods for PDAE-constrained Optimization Based on Space-Time Adaptive PDAE Solvers, Springer Basel, Basel, pp 51–74. DOI 10.1007/978-3-0348-0133-1_4
13. Conn AR, Gould NIM, Toint PL (2000) Trust region methods. SIAM, Philadelphia, PA
14. Dennis J, El-Alem M, Maciel M (1997) A global convergence theory for general trust-region-based algorithms for equality constrained optimization. SIAM Journal of Optimization 7:177–207
15. Dennis JE, Li SB, Tapia R (1994) A unified approach to global convergence of trust region methods for nonsmooth optimization. Mathematical Programming 68:319–346
16. Dennis JE, El-Alem M, Maciel MC (1997) A global convergence theory for general trust-region-based algorithms for equality constrained optimization. SIAM Journal on Optimization 7(1):177–207, DOI 10.1137/S1052623492238881
17. Fahl M, Sachs EW (2003) Reduced order modelling approaches to pde-constrained optimization based on proper orthogonal decomposition. In: Large-scale PDE-constrained optimization, Springer, pp 268–280
18. Garreis S, Ulbrich M (2019) An inexact trust-region algorithm for constrained problems in Hilbert space and its application to the adaptive solution of optimal control problems with PDEs. Preprint, submitted, Technical University of Munich
19. Götschel S, Weiser M (2015) Lossy compression for PDE-constrained optimization: Adaptive error control. Computational Optimization and Applications 62:131–155
20. Götschel S, Chamakuri N, Kunisch K, Weiser M (2014) Lossy compression in optimal control of cardiac defibrillation. Journal of Scientific Computing 60:35–59
21. Heinkenschloss M, Ridzal D (2014) A matrix-free trust-region SQP method for equality constrained optimization. SIAM Journal on Optimization 24(3):1507–1541
22. Heinkenschloss M, Vicente LN (2001) Analysis of inexact trust-region SQP algorithms. SIAM Journal on Optimization 12(2):282–302
23. Kouri DP, Ridzal D (2018) Inexact trust-region methods for PDE-constrained optimization. In: Frontiers in PDE-Constrained Optimization, Springer, New York, NY, pp 83–121
24. Kouri DP, Heinkenschloss M, Ridzal D, van Bloemen Waanders BG (2013) A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty. SIAM Journal on Scientific Computing 35(4):A1847–

A1879

25. Kouri DP, Heinkenschloss M, Ridzal D, van Bloemen Waanders BG (2014) Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty. SIAM Journal on Scientific Computing 36(6):A3011–A3029
26. Lin CJ, Moré JJ (1999) Newton's method for large bound-constrained optimization problems. SIAM Journal on Optimization 9(4):1100–1127
27. Meidner D, Vexler B (2007) Adaptive space-time finite element methods for parabolic optimization problems. SIAM Journal on Control and Optimization 46(1):116–142
28. Moré JJ (1983) Recent developments in algorithms and software for trust region methods. In: Bachem A, Grötschel M, Korte B (eds) Mathematical Programming, The State of the Art, Springer-Verlag, Berlin, pp 258–287
29. Moré JJ (1988) Trust regions and projected gradients. In: System Modelling and Optimization, Springer, New York, NY, pp 1–13
30. Muthukumar R, Kouri DP, Udell M (2021) Randomized sketching algorithms for low-memory dynamic optimization. SIAM Journal on Optimization 31(2):1242–1275
31. Toint PL (1988) Global convergence of a class of trust-region methods for nonconvex minimization in Hilbert space. IMA Journal of Numerical Analysis 8(2):231–252, DOI 10.1093/imanum/8.2.231
32. Walther A (2008) A first-order convergence analysis of trust-region methods with inexact Jacobians. SIAM Journal on Optimization 19(1):307–325, DOI 10.1137/050634530
33. Zahr MJ, Carlberg KT, Kouri DP (2019) An efficient, globally convergent method for optimization under uncertainty using adaptive model reduction and sparse grids. SIAM/ASA Journal on Uncertainty Quantification 7(3):877–912
34. Ziems JC, Ulbrich S (2011) Adaptive multilevel inexact SQP methods for PDE-constrained optimization. SIAM Journal on Optimization 21(1):1–40, DOI 10.1137/080743160
35. Zou Z, Kouri DP, Aquino W (2022) A locally adapted reduced-basis method for solving risk-averse PDE-constrained optimization problems. SIAM/ASA Journal on Uncertainty Quantification 10(4):1629–1651