SOBRAPO

TRUST-REGION-BASED METHODS FOR NONLINEAR PROGRAMMING: RECENT ADVANCES AND PERSPECTIVES

Sandra A. Santos

ABSTRACT. The aim of this text is to highlight recent advances of trust-region-based methods for nonlinear programming and to put them into perspective. An algorithmic framework provides a ground with the main ideas of these methods and the related notation. Specific approaches concerned with handling the trust-region subproblem are recalled, particularly for the large scale setting. Recent contributions encompassing the trust-region globalization technique for nonlinear programming are reviewed, including nonmonotone acceptance criteria for unconstrained minimization; the adaptive adjustment of the trust-region radius; the merging of the trust-region step into a line search scheme, and the usage of the trust-region elements within derivative-free optimization algorithms.

Keywords: trust-region methods, global convergence, nonlinear programming.

1 INTRODUCTION

In the numerical solution of nonlinear optimization problems, usually by iterative schemes, it is desirable to reach convergence to stationary points starting from an arbitrary approximation, what defines the so-called *global convergence*. Trust-region methods, originally devised for unconstrained optimization, are robust globalization strategies that rest upon a model (usually quadratic) for the objective function around the current iterate and a measure for the agreement between the model and the original function. Their robustness may be connected with the regularization effect of minimizing (quadratic) models over regions of predetermined size.

A thorough reference to the subject is Conn, Gould & Toint's book [16], published in 2000, that includes an extensive annotated bibliography, known at that time. As pointed out in the introduction of such a book (§ 1.2), the term *trust region* seems to have been coined by Dennis, in a course that he taught shortly after he heard Powell talk about his technique for solving nonlinear equations [60] (see also [61]). The first official appearance of the term *trust region* seems to be in Dennis [22]. Nevertheless, it took a while before such a terminology spread among the

Department of Applied Mathematics, University of Campinas, Campinas, SP, Brazil. E-mail: sandra@ime.unicamp.br

community, being the survey of Moré [54] highly influential, not to mention the papers on convergence and algorithms of Powell [62, 63] and Moré & Sorensen [56]. Related nomenclature are *restricted step method*, adopted by Fletcher [30] and *confidence region*, employed by Toint [83].

The books of Dennis & Schnabel [24], and of Nocedal & Wright [57] are also relevant references on the trust-region scenario. The former is a classic textbook about unconstrained minimization and nonlinear equations that presents not only the details of the *optimal hooked step* (cf. Hebden [42]) but also describes the *dog-leg* [61] and the *double-dog leg* [23] strategies for approximately solving the trust-region subproblem, together with a discussion about the initial choice and the updating of the trust-region radius. The latter, a contemporary textbook on general nonlinear programming, goes further into *nearly exact solutions* to the subproblems, including Steihaug's approach [81] (see also Toint [83]) and the description of the so-called *hard case*. The presented trust-region algorithm for unconstrained minimization is accompanied by the global convergence analysis under *fraction of Cauchy decrease*. Besides, the text tackles scaling, non-Euclidean trust regions and the implementation of the Levenberg-Marquardt strategy upon the trust-region perspective (see also Moré [53] and references therein). It is worth mentioning that the Levenberg-Marquardt strategy for the nonlinear least-squares problem may be considered as a precursor of the trust-region method for unconstrained minimization. When it comes to constrained optimization, Nocedal & Wright have also analyzed the trust-region sequential quadratic programming (SQP) approach, presenting a practical trust-region SQP algorithm.

The purpose of this survey is to highlight recent advances in the area and to put them into perspective. It is organized as follows. To provide a ground with the main ideas and the related notation, an algorithmic framework is recalled in Section 2. Specific approaches concerned with handling the trust-region subproblem are discussed in Section 3. Recent contributions encompassing trust-region globalization for nonlinear programming are reviewed in Section 4, including, among others, nonmonotone acceptance criteria for unconstrained minimization and the usage of the trust-region philosophy within derivative-free optimization algorithms. Finally, a brief conclusion is given in Section 5.

2   THE ALGORITHMIC FRAMEWORK

The trust-region strategy for minimization on an arbitrary domain is outlined next, following the presentation of Martínez & Santos [50]. The problem upon consideration is

$$\text{minimize } f(x) \text{ subject to } x \in \mathcal{D}, \tag{1}$$

where $\mathcal{D} \subset \mathbb{R}^n$ is an arbitrary closed set and f is continuously differentiable in an open set that contains $\mathcal{D}$. The gradient is denoted by $g := \nabla f$.

Let $\| \cdot \|$ be an arbitrary norm on $\mathbb{R}^n$ and the associated induced matrix norm. Let the parameters σ_1, σ_2, θ, $\Delta_{\min}$, M, γ be such that $0 < \sigma_1 < \sigma_2 < 1$, $\theta \in (0, 1)$, $\Delta_{\min} > 0$, $M > 0$ and $\gamma \in (0, 1]$. Initially, an arbitrary feasible point $x_0 \in \mathcal{D}$ is known, a symmetric matrix $B_0 \in \mathbb{R}^{n \times n}$ such that $\|B_0\| \leq M$ is given, together with an initial radius $\Delta^0 \geq \Delta_{\min}$. At the k-th iteration, the first trust-region radius tried is denoted by Δ^k, whereas the trust-region radius finally accepted

is denoted by Δ_k. Given $x_k \in \mathcal{D}$, $g_k := g(x_k)$, $B_k = B_k^T \in \mathbb{R}^{n \times n}$ such that $\|B_k\| \leq M$ and $\Delta^k \geq \Delta_{\min}$, the steps for obtaining Δ_k and x_{k+1} are given next:

Algorithm 1. *Trust-region for minimization on an arbitrary domain*

Step 0. Set $\Delta \leftarrow \Delta^k$.

Step 1. Compute a global solution $s_k^Q(\Delta)$ of

$$\begin{aligned} \text{minimize} \quad & Q_k(s) := \tfrac{M}{2} s^T s + g_k^T s \\ \text{subject to} \quad & x_k + s \in \mathcal{D}, \|s\| \leq \Delta. \end{aligned} \tag{2}$$

If $Q_k(s_k^Q(\Delta)) = 0$, stop.

Step 2. Compute $\bar{s}_k(\Delta)$ such that

$$\begin{aligned} \Phi_k(\bar{s}_k(\Delta)) &\leq \gamma \, Q_k(s_k^Q(\Delta)) \\ x_k + \bar{s}_k(\Delta) &\in \mathcal{D}, \|\bar{s}_k(\Delta)\| \leq \Delta, \end{aligned} \tag{3}$$

where $\Phi_k(s) := \tfrac{1}{2} s^T B_k s + g_k^T s$ for all $s \in \mathbb{R}^n$.

Step 3. If

$$f(x_k + \bar{s}_k(\Delta)) \leq f(x_k) + \theta \Phi_k(\bar{s}_k(\Delta)) \tag{4}$$

then define $s_k = \bar{s}_k(\Delta)$, $x_{k+1} = x_k + s_k$, $\Delta_k = \Delta$. Choose $\Delta^{k+1} \geq \Delta_{\min}$, $B_{k+1} = B_{k+1}^T$ such that $\|B_{k+1}\| \leq M$ and return.

Otherwise, replace Δ by $\Delta_{\text{new}} \in [\sigma_1 \|\bar{s}_k(\Delta)\|, \sigma_2 \Delta]$ and go to Step 1.

Some remarks about the Algorithm 1 are in order:

i. If the Algorithm 1 stops at *Step 1*, so that $Q_k(s_k^Q(\Delta)) = 0$, then x_k is stationary for problem (1) (cf. [50]).

ii. In *Step 2*, the quadratic model is required to decrease a fraction of the minimum of the auxiliary subproblem (2). The solution of the auxiliary subproblem plays the role of the classical Cauchy point, with a practical advantage of being computed possibly with a single projection step, whereas computing a Cauchy point approximation may require more than one projection on the feasible region.

iii. The fixed parameter $\Delta_{\min} > 0$ imposes a lower bound for the first trust-region radius tried at each iteration. As a result, larger steps are tried far from the solution, and artificially small trial steps inherited from previous iterations are eliminated.

iv. The acceptance condition (4) of *Step 3* is an Armijo-like alternative presentation of the usual ratio between the *actual* and the *predicted* reductions:

$$\frac{\text{Ared}_k}{\text{Pred}_k} := \frac{f(x_k) - f(x_k + \bar{s}_k(\Delta))}{\Phi_k(0) - \Phi_k(\bar{s}_k(\Delta))} \geq \theta. \tag{5}$$

It is worth mentioning that the current value of such a ratio may be used in more sophisticated schemes for updating the trust-region radius.

v. Although the subproblems

$$
\begin{aligned}
\text{minimize} \quad & \Phi_k(s) \\
\text{subject to} \quad & x_k + s \in \mathcal{D}, \|s\| \leq \Delta.
\end{aligned}
\tag{6}
$$

might be difficult, as linear approximations of the set $\mathcal{D}$ are not employed, particular cases are solvable, which include Euclidean balls; spheres; complements of Euclidean balls; intersection of the aforementioned sets, among others. Thus, the algorithm is not restricted to a convex domain. For further details, see [50].

vi. By choosing $\| \cdot \|$ as the ℓ_∞ norm, in case de domain $\mathcal{D}$ is a polytope, the subproblem (6) is solvable, with polynomial complexity for an ε-approximate solution, for arbitrary quadratics (cf. Vavasis [84]).

2.1   Convergence results and further comments

The uniform boundedness assumption $\|B_k\| \leq M$ is a common hypothesis for the trust-region algorithms to reach global convergence ([54, 78]). A weaker condition was used by Powell [63], namely $\|B_k\| \leq M + \widetilde{M}k$, with $k = 1, 2, 3 \ldots$. Together with a step s_k that provides simple decrease $f(x_k + s_k) < f(x_k)$, Powell proved global convergence in the sense of $\liminf_{k \to \infty} \|g_k\| = 0$ for the unconstrained case ($\mathcal{D} = \mathbb{R}^n$).

Imposing the sufficient decrease condition (4) allows to strengthen the result, so that, for a continuously differentiable function f with Lipschitz continuous gradient, if $\|B_k\| \leq M$, $\mathcal{D} = \mathbb{R}^n$ and f is bounded below in the level set $\{x \in \mathbb{R}^n \mid f(x) \leq f(x_0)\}$ then $\lim_{k \to \infty} \|g_k\| = 0$ (see, e.g. Theorem 4.8 of Nocedal & Wright [57]).

No doubt, one of the advantages of trust-region methods, as compared with line search methods, is that the matrix B_k is allowed to be indefinite (cf. Nocedal & Yuan [58]). In Algorithm 1 this provides more freedom to form the quadratic model defined by $\Phi_k(\cdot)$, from quasi-Newton approximations to the true Hessian if the function f is twice continuously differentiable. In the latter case, convergence to points that satisfy the second order necessary optimality conditions may be obtained for the unconstrained problem (cf. Shultz, Schnabel & Byrd [78]).

In terms of local convergence, at the early iterations, when x_k may be far from the solution $x_\star$, the values of Δ_k may be small and may prevent a full (quasi) Newton step from being taken. However, at later iterations in which x_k is closer to $x_\star$, it is hoped that there will be greater trust in the model. Then Δ_k can be made sufficiently large so that full (quasi) Newton steps are acceptable and a (superlinear) quadratic convergence rate is achievable.

Under an assumption of weak regularity, defined in terms of feasible arcs, it has been proved that the Algorithm 1 is well defined and globally convergent to a stationary point, also defined in terms of feasible arcs emanating from it (see, resp. Theorems 2.3 and 3.2 of [50]). A specific algorithm for handling the Euclidean ball domain was presented in [50] as well, together with an extensive set of numerical examples.

3   TRUST-REGION SUBPROBLEMS

As previously detailed, the exact solution of the trust-region subproblem (TRS) is not required for reaching convergence of the main algorithm. Moreover, the subproblem does not need to be solved to high precision and the step s_k may satisfy $\|s_k\| \leq \xi \Delta_k$ for some constant $\xi \geq 1$. As expected, the cost of obtaining a solution for the TRS turns to be more significative as the problem dimension increases.

A consolidated strategy for computing an approximate solution for the TRS based on Cholesky factorizations was proposed by Moré & Sorensen [55]. Theoretical properties of the TRS were presented by Gay [33] and Sorensen [79]. Further, Ben-Tal & Teboulle [9] have deepened the analysis of the tremendous amount of structure of such a problem. Taking into account the modern sparse linear algebra subroutines available, the ideas of Gay & Moré-Sorensen were revisited by Gould, Robinson & Thorne [37]. The resulting software is freely available as the packages TRS and RQS, as part of the GALAHAD optimization library [36].

In case a factorization turns to be too expensive or not affordable, and just matrix-vector products are at hand, the inexact conjugate-gradient-like strategy of Steihaug-Toint [81, 83] might be an alternative. To overcome the likely premature stopping of such a strategy whenever negative curvature is present, Gould, Lucidi, Roma & Toint [35] have used the Lanczos method to solve further the subproblem in case the boundary is encountered, defining the GLTR (generalized Lanczos trust-region) algorithm.

An extension of the strategy of Steihaug & Toint has been recently proposed by Erway, Gill & Griffin [26], in which a solution of the TRS may be calculated to any prescribed accuracy. A controlling parameter allows the user to take advantage of the tradeoff between the overall number of function evaluations and matrix-vector products associated with the underlying trust-region method. An improvement upon Steihaug-Toint has been suggested by Erway & Gill [25], where the trust-region norm is defined independently of the employed preconditioner. Numerical experiments corroborate the efficiency of the proposed improvement in terms of function evaluations, as compared with Steihaug's algorithm and with the GLTR of [35].

In the survey [59], Palagi addresses other possibilities for the numerical solution of the large-scale TRS: the parametric eigenvalue reformulation-based strategy of Sorensen [80]; the semidefinite programming approach of Rendl & Wolkowicz [70]; the exact penalty function based algorithm of Lucidi, Palagi & Roma [46], and the DC (difference of convex functions) based algorithm of Pham Dinh Tao & Le Thi Hoai An [82].

Along the parametric eigenvalue reformulation-based philosophy for the large-scale TRS is the work of Rojas, Santos & Sorensen [72], a matrix-free method that improves upon Sorensen's original idea [80] by encompassing both the easy and the hard case in a unified and superlinearly convergent interpolating scheme. The implementation of the LSTRS method, that stands for large-scale trust-region subproblems, is presented and described in [73].

Another recent matrix-free method for the large-scale TRS has been proposed by Apostolopoulou, Sotiropoulos & Pintelas [5]. By assuming that the matrix B_k is updated by the limited

memory L-BFGS formula with a correction of low rank (one or, at most, two), they have obtained analytic formulas for the eigenvalues of the involved matrices. Based on such formulas, they have constructed a positive definite matrix with analytically computable inverse, without any factorization. Moreover, in the hard case, the inverse power method may be used to compute the demanded direction of negative curvature. Numerical comparative experiments illustrate the efficiency of the approach in terms of the obtained accuracy, small running time and negligible needed amount of memory.

In a following work, Apostolopoulou, Sotiropoulos, Botsaris & Pintelas have extended the previous results of [5], where the authors have assumed the initial B_0 to be a scalar matrix defined by the Barzilai & Borwein spectral parameter. In [4] the authors have studied the eigenstructure of minimal-memory BFGS matrices with the usage of any non-zero real number as the initial scaling factor. Likewise, analytic expressions are derived, factorizations are avoided, and an algorithm that solely rests upon inner products and vector summations is obtained, with an extremely favorable numerical performance when compared with the GLTR algorithm of Gould, Lucidi, Roma & Toint [35].

Along the same perspective is the recent work of Erway & Marcia [27], that addresses the solution of linear systems of the form $(B + \sigma I)x = c$, being $B \in \mathbb{R}^{n \times n}$ a limited-memory BFGS matrix (with m updates, $m \ll n$) and σ a positive constant. A recursive formula is devised upon simple conditions on B_0 and σ, with complexity $m^2 n$. Experiments with $m = 5$ and n from 10^3 up to 10^7 illustrate the performance of the proposed formula comparatively with the `matlab` direct *backslash* command (with $n \leq 2 \times 10^4$) and the built-in conjugate-gradient routine `pcg.m`.

4   NONLINEAR PROGRAMMING

In this section, several current trust-region-based methods are reviewed, separated in three classes, namely: unconstrained minimization; constrained minimization and other problems, and derivative-free optimization.

4.1   Unconstrained minimization

Concerning the convergence of unconstrained minimization algorithms, following the pioneering work of Powell [62, 63], the general analysis presented by Shultz, Schnabel & Byrd [78] extends it and practically encompasses the main aspects, addressing first and second order necessary optimality conditions under fraction of optimal decrease and fraction of Cauchy decrease. Although the authors have focused a class of trust-region-based algorithms, their scheme is sufficiently broad to include line search algorithms as well.

After such a thorough analysis, completed by the systematization presented in the comprehensive book of Conn, Gould & Toint [16], one could think that the unconstrained minimization scenario upon the trust-region perspective had been exhausted. Nevertheless, two aspects have generated several contributions in the last few years.

First, the attempt to devise an automatic adjustment of the trust-region radius, as examined by Sartenaer [75], has generated the *adaptive trust-region methods*, being Zhang, Zhang, &

Liao [90] one of the first references in this context. Shi & Guo [76] have also proposed an algorithm that automatically adjusts the trust-region radius at each iteration.

Second, adopting a nonmonotone criterion for acceptance of the step has generated *nonmonotone trust-region methods*. It was motivated by the popularity of nonmonotone line search techniques for the solution of unconstrained optimization problems by Newton's method. Condition (4), or equivalently (5), is relaxed by replacing the objective value at the kth-iteration by the maximum of the function values at the current and the previous p iterations, for a given positive integer p. This strategy has been proposed by Deng, Xiao & Zhou [21], and it is based on Grippo, Lampariello & Lucidi's ideas [39]. Subsequent related contributions are the works of Mo, Liu & Yan [52], Gu & Mu [41] and Liu & Ma [45]; nevertheless, these three references have employed the average function value of the latest $p + 1$ iterations instead of the maximum, following the approach of Zhang & Hager [88].

Combining the aforementioned two features, several *nonmonotone adaptive trust-region methods* have been devised for unconstrained minimization. Among them, and chronologically, we should mention Zhang & Zhang [89], Fu & Sun [32], Shi & Wang [77] and Ahookhosh & Amini [2], among others.

Following Nocedal & Yuan's ideas [58], some authors have adopted a combination of trust-region and line search ideas. This is the case of Shi & Wang [77], Ahookhosh, Amini & Peyghami [1] and Liu & Ma [45]. It is worth mentioning that the quadratic models adopted by Liu & Ma are convex, what is not mandatory in trust-region methods. The matrices B_k are diagonal, so it is possible to solve exactly the trust-region subproblems. Moreover, the obtained directions are always of descent at the current iterate x_k. As a consequence of the adopted line search, the global convergence result obtained is weaker than the usual global convergence of trust-region methods: instead of reaching stationarity at all limit points, the authors have just established that there exists a stationary limit point of the generated sequence.

Subspace properties of trust-region methods were analyzed and assessed by Wang & Yuan [86]. Filter trust-region algorithms for unconstrained optimization have been first addressed by Gould, Sainvitu & Toint [38], based on the multidimensional filter devised by Gould, Leyffer & Toint [34] for solving (possibly) large-scale systems of nonlinear equations and nonlinear least-squares problems. More recently, Fatemi & Mahdavi-Amiri [29] improved upon Gould, Sainvitu & Toint's ideas.

Last but not least, in the latest textbook of Griva, Nash & Sofer [40] the trust-region methods are described in the basics for unconstrained optimization, in parallel with line search methods, as a strategy for guaranteeing convergence.

4.2 Constrained minimization and additional problems

When it comes to *bound-constrained optimization*, there are a few recent contributions to be reviewed. The first is the filter-trust-region method of Sainvitu & Toint [74], in which the authors have extended the technique of [38] by means of a gradient-projection method.

The second is the trust-region affine method of Wang [85], where two trial steps are computed. The primary is obtained by solving an appropriate quadratic model in an ellipsoidal region defined by an affine scaling technique that depends on both the distance of the current iterate to the boundary and the trust-region radius. For establishing convergence and avoiding iterations trapped around nonstationary points, an auxiliary step is defined along an approximate projected gradient. The trial step to generate the next iterate is chosen as the one that produces more reduction of the quadratic model, so that the limit points of this algorithm are not bounded away from stationarity.

The third is the DC (difference of convex) trust-region method of Le Thi Hoai An et al. [3]. The authors have used the DC framework in the solution of nonlinear optimization problems within convex domains using trust regions. More specifically, DC local models for the quadratic model of the objective function were used to compute the trust-region step, and a primal-dual subgradient method was applied to the solution of the associated trust-region subproblems. Exact second-order derivatives of the objective function turned out to be essential, theoretically and practically. Moreover, the applicability of the approach rests upon projections on the feasible set being affordable. Bound constrained minimization was used to illustrate and validate the proposed idea, in a thorough set of computational tests.

Lukšan, Matonoha & Vlček [47] have solved the ℓ_1 *optimization problem* by means of a sequence of parametrized trust-region interior point-barrier methods, and the sequence of solutions thus obtained is shown to converge to the solution of the original problem.

For *bound-constrained nonlinear systems*, Bellavia, Macconi & Morini [8] have devised an affine scaling trust-region approach. Concerning *general nonlinear programming*, the global convergence of a trust-region SQP-filter algorithm has been addressed by Fletcher, Gould, Leyffer, Toint & Wächter [31] and also by Maciel & Mendonça [48].

Kanzow & Petra [43] have reformulated the *mixed complementarity problem* as a bound-constrained nonlinear least-squares problem with zero residual. On the basis of this reformulation, a trust-region method for the solution of mixed complementarity problems is considered. This trust-region method combines a projected Levenberg-Marquardt step to guarantee local fast convergence under suitable assumptions, affine scaling matrices which are used to improve the global convergence properties, and a multidimensional filter technique to accept the full step more frequently.

For *unconstrained multiobjective problems*, Carrizo, Lotito & Maciel [12] have proposed a trust-region quasi-Newton algorithm, based on the BFGS updates for scalar optimization. Comparative results with the usage of exact Hessians have shown a clear advantage for the BFGS approximation, when it comes to the total number of demanded functional evaluations.

4.3 Derivative-free optimization

Among the strategies employed in the derivative-free optimization scenario, the trust-region algorithms rest upon linear or quadratic approximations to the objective function, which are based only on the objective function values at sample points. These local surrogate models are the core

of the interpolation-based derivative-free optimization methods, reviewed by Karasözen in the survey [44], also described in the book of Conn, Scheinberg & Vicente [20] and pointed out in the survey of Rios & Sahinidis [71].

Powell's algorithm COBYLA (constrained optimization by linear approximations) [64] was created for addressing nonlinearly constrained optimization problems with a few variables and it is easy to use, according to the author. Each iteration forms linear approximations to the objective and constraint functions by interpolation at the vertices of a simplex, and a trust-region bound restricts each change to the variables. Thus, a new vector of variables is calculated, which may replace one of the current vertices, either to improve the shape of the simplex or because it is the best vector that has been found so far, according to a merit function that gives attention to the greatest constraint violation. COBYLA was followed by UOBYQA (unconstrained optimization by quadratic approximation) [65], NEWUOA [66, 67] and BOBYQA [68]. The latter is an efficient algorithm that handles bound-constrained minimization and does not have a companion convergence theory.

Also based on models is the so-called DFO (derivative free optimization) method of Conn, Scheinberg & Toint [17], for unconstrained minimization. It encompasses techniques that ensure *geometric quality* of the considered models, based upon Lagrangian polynomials. Indeed, for the model to be well-defined, the interpolation points must be *poised*, meaning that they must be compatible with the interpolation conditions imposed on them. The convergence of this approach is presented in [14], using Newton's fundamental polynomials, an alternative to Lagrange functions.

Marazzi & Nocedal [49] have also proposed (linear and quadratic)-model-based algorithms for unconstrained derivative-free optimization, whose convergence is ensured by trust regions. The geometric quality of the model is controlled by means of a taboo region for the potentially degenerate points, that are avoided by imposing an aditional constraint, wedged-shape in the linear case.

A numerical study was presented by Fasano, Morales & Nocedal [28] concerning unconstrained derivative-free optimization, aimed to investigate the effect of dispensing with the geometry phase altogether. To their surprise, although ill-conditioning had been observed, a self-correcting mechanism seemed to be present, so that no failure was observed. To remove a point from the current set, Moré-Sorensen [55] algorithm was used. Very competitive comparative results with DFO [17, 14] and NEWUOA [66, 67] were shown. Scheinberg & Toint [17] have analyzed further the intrinsic self-correcting mechanism of combining trust regions and interpolating models for unconstrained derivative-free optimization.

An algorithm for least-squares minimization upon the derivative-free perspective was proposed by Zhang, Conn & Scheinberg [87], taking advantage of the intrinsic structure of the problem, but following the features of Conn, Scheinberg & Vicente [18] and of Powel [66] for practical efficiency.

Conn, Scheinberg & Vicente [19] have broaden the theoretical analysis of global convergence of trust-region algorithms to first- and second-order critical points. They have considered a class of

methods based on the minimization of quadratic or linear models, with results that do not depend on the sampling techniques to generate the sets of interpolant points. Important issues addressed include global convergence when acceptance of iterates is based on simple decrease of the objective function; trust-region radius maintenance at the criticality step, and global convergence for second-order critical points.

An opportune discussion concerning the differences between the description of an algorithm for practical use and its description for developing convergence theory is given by Powell [69]. Moreover, in such a work he presents the global convergence of an algorithm for unconstrained derivative-free optimization in $\mathbb{R}^n$ under the assumption that the models interpolate the objective function in $n + 1$ points, what ensures unicity of the models in the linear case.

Bandeira, Scheinberg & Vicente [6] have analyzed the sparse recovery of models for functions with sparse Hessians, in unconstrained minimization. Probabilistic models for the unconstrained case have been handled by Bandeira, Scheinberg & Vicente [7].

When it comes to model-based general constrained derivative-free optimization, two algorithms stand out: the DFO of Conn, Scheinberg & Toint [15] and Berghen & Bersini's CONDOR [10], an extension of Powell's UOBYQA. Both had been designed for small dimensional problems and high-computing-load objective functions. DFO uses linear or quadratic models to guide the search, in contrast to UOBYQA and CONDOR, thus requiring less function evaluations to build the local models. The authors of CONDOR, however, based on their experimental results, have surprisingly discovered that their code used less function evaluations than DFO to reach an optimum point, despite the fact that the cost to build a local model is higher. The heuristic used inside UOBYQA (and also inside CONDOR) has shown to be relevant to reduce the number of function evaluations in the presence of noisy and high computing load objective functions. A primary aim of Berghen & Bersini was to provide an updated version and a more accessible description of such a heuristic. It is worth mentioning that for both DFO and CONDOR, the performance improves in case the gradients of the constraints are available.

In Conejo et al. [13], the authors have established the global convergence of a trust-region-based algorithm developed for convex-constrained derivative-free optimization under usual assumptions. Although the problem upon consideration is assumed to be smooth, only the derivatives of the constraints are available.

Problems with smooth constraints (not necessarily convex) and derivative-free objective function were also tackled by Bueno, Friedlander, Martínez & Sobral [11] within the inexact restoration approach, that performed favourably in terms of robustness in comparison with COBYLA and three other benchmarks. For problems with thin domains, defined by computationally inexpensive but highly nonlinear functions, Martínez & Sobral [51] have proposed the algorithm SKINNY, that splits the main iteration into a restoration step, where infeasibility is decreased without evaluating the objective function, followed by the derivative-free minimization on a relaxed feasible set. In the presented comparative numerical experiments, SKINNY were able to solve more problems than DFO [15].

5   FINAL REMARKS

The trust-region-based methods constitute a relatively recent research area among the nonlinear programming community. The fact that over half of the referred contributions of this survey date from 2000 onwards corroborates the potential strength that the focused subject has being reached, with a promising even greater impact in the near future. To name a few, nonconvex constrained optimization; robust optimization and noisy optimization are challenging branches of increasing interest and under current investigation, in which the trust-region framework may offer valuable ingredients to produce globally convergent and robust algorithms.

ACKNOWLEDGMENTS

REFERENCES

[1] AHOOKHOSH M, AMINI K & PEYGHAMI MR. 2012. A nonmonotone trust-region line search method for large-scale unconstrained optimization. *Appl. Math. Model.*, **36**(1): 478–487.

[2] AMINI K & AHOOKHOSH M. 2011. Combination adaptive trust region method by non-monotone strategy for unconstrained nonlinear programming. *Asia-Pac. J. Oper. Res.*, **28**(5): 585–600.

[3] AN LTH, NGAI HV, TAO PD, VAZ AIF & VICENTE LN. 2014. Globally convergent DC trust-region methods. *J. Glob. Optim.*, **59**: 209–225.

[4] APOSTOLOPOULOU MS, SOTIROPOULOS DG, BOTSARIS CA & PINTELAS P. 2011. A practical method for solving large-scale TRS. *Optim. Lett.*, **5**(2): 207–227.

[5] APOSTOLOPOULOU MS, SOTIROPOULOS DG & PINTELAS P. 2008. Solving the quadratic trust-region subproblem in a low-memory BFGS framework. *Optim. Methods Softw.*, **23**(5): 651–674.

[6] BANDEIRA AS, SCHEINBERG K & VICENTE LN. 2012. Computation of sparse low degree interpolating polynomials and their application to derivative-free optimization. *Math. Program.*, **134**(1, Ser. B): 223–257.

[7] BANDEIRA AS, SCHEINBERG K & VICENTE LN. 2013. Convergence of trust-region methods based on probabilistic models, arXiv.org, `http://arxiv.org/abs/1304.2808`, Submitted 09-April-2013.

[8] BELLAVIA S, MACCONI M & MORINI B. 2003. An affine scaling trust-region approach to bound-constrained nonlinear systems. *Appl. Numer. Math.*, **44**(3): 257–280.

[9] BEN-TAL A & AND TEBOULLE M. 1996. Hidden convexity in some nonconvex quadratically constrained quadratic programming. *Math. Program.*, **72**(1, Ser. A): 51–63.

[10] BERGHEN FV & BERSINI H. 2005. CONDOR, a new parallel, constrained extension of powells UOBYQA algorithm: Experimental results and comparison with the DFO algorithm. *J. Comput. Appl. Math.*, **181**(1): 1457–175.

[11] BUENO LF, FRIEDLANDER A, MARTÍNEZ JM & SOBRAL FNC. 2013. Inexact Restoration Method for Derivative-Free Optimization withSmooth Constraints. *SIAM J. Optim.*, **23**(2): 1189–1213.

[12] CARRIZO GA, LOTITO PA & MACIEL MC. 2013. Método Quasi-Newton para Optimización Multiobjetivo. In: G.L. Mura, D. Rubio, and E. Serrano, editors, MACI Vol. 4 (2013), pages 363–366. ASAMACI – Asociación Argentina de Matemática Aplicada, Computacional e Industrial, Argentina.

[13] CONEJO PD, KARAS EW, PEDROSO LG, RIBEIRO AA & SACHINE M. 2012 Global convergence of trust-region algorithms for constrained minimization without derivatives. *Appl. Math. Comput.*, **220**(1): 324–330.

[14] CONN AR, SCHEINBERG K & TOINT PHL. 1997. On the convergence of derivative-free methods for unconstrained optimization. In: Approximation Theory and Optimization (Cambridge, 1996), pages 83–108. Cambridge: Cambridge Univ. Press.

[15] CONN AR, SCHEINBERG K & TOINT PHL. 1998. A derivative free optimization algorithm in practice. In: Proceedings of 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization (St. Louis, MO).

[16] CONN AR, GOULD NIM & TOINT PHL. 2000. Trust-Region Methods. MPS/SIAM Series on Optimization. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

[17] CONN AR, SCHEINBERG K & TOINT PHL. 1997. Recent progress in unconstrained nonlinear optimization without derivatives. *Math. Program.*, **79**(1-3, Ser. B): 397–414.

[18] CONN AR, SCHEINBERG K & VICENTE LN. 2008. Geometry of interpolation sets in derivative free optimization. *Math. Program.*, **111**(1-2, Ser. B): 141–172.

[19] CONN AR, SCHEINBERG K & VICENTE LN. 2009. Global convergence of general derivative-free trust-region algorithms to first- and second-order critical points. *SIAM J. Optim.*, **20**(1): 387–415.

[20] CONN AR, SCHEINBERG K & VICENTE LN. 2009. Introduction to derivative-free optimization. Vol. 8 of MPS/SIAM Series on Optimization. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

[21] DENG NY, XIAO Y & ZHOU FJ. 1993. Nonmonotonic trust region algorithm. *J. Optim. Theory Appl.*, **76**(2): 259–285.

[22] DENNIS JR JE. 1978. A brief introduction to quasi-Newton methods. In: Numerical analysis (Proc. Sympos. Appl. Math., Atlanta, GA, 1978), Proc. Sympos. Appl. Math., XXII, pages 19–52. Providence, RI: American Mathematics Society (AMS).

[23] DENNIS JR JE & MEI HHW. 1979. Two new unconstrained optimization algorithms which use function and gradient values. *J. Optim. Theory Appl.*, **28**(4): 453–482.

[24] DENNIS JR JE & SCHNABEL RB. 1996. Numerical methods for unconstrained optimization and nonlinear equations. Vol. 16 of Classics in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM). Corrected reprint of the 1983 original.

[25] ERWAY JB & GILL PE. 2009. A subspace minimization method for the trust-region step. *SIAM J. Optim.*, **20**(3): 1439–1461.

[26] ERWAY JB, GILL PE & GRIFFIN JD. 2009. Iterative methods for finding a trust-region step. *SIAM J. Optim.*, **20**(2): 1110–1131.

[27] ERWAY JB & MARCIA RF. 2012. Limited-memory BFGS systems with diagonal updates. *Linear Algebra Appl.*, **437**(1): 333–344.

[28] FASANO G, MORALES JL & NOCEDAL J. 2009. On the geometry phase in model-based algorithms for derivative-free optimization. *Optim. Methods Softw.*, **24**(1): 145–154.

[29] FATEMI M & AND MAHDAVI-AMIRI N. 2012. A filter trust-region algorithm for unconstrained optimization with strong global convergence properties. *Comput. Optim. Appl.*, **52**(1): 239–266.

[30] FLETCHER R. 1987. Practical methods of optimization. Second edition. Chichester: John Wiley & Sons.

[31] FLETCHER R, GOULD NIM, LEYFFER S, TOINT PhL & WÄCHTER A. 2002/03. Global convergence of a trust-region SQP-filter algorithm for general nonlinear programming. *SIAM J. Optim.*, **13**(3): 635–659.

[32] FU J & SUN W. 2005. Nonmonotone adaptive trust-region method for unconstrained optimization problems. *Appl. Math. Comput.*, **163**(1): 489–504.

[33] GAY DM. 1981. Computing optimal locally constrained steps. *SIAM J. Sci. Statist. Comput.*, **2**(2): 186–197.

[34] GOULD NIM, LEYFFER S & TOINT PhL. 2004. A multidimensional filter algorithm for nonlinear equations and nonlinear least-squares. *SIAM J. Optim.*, **15**(1): 17–38.

[35] GOULD NIM, LUCIDI S, ROMA M & TOINT PhL. 1999. Solving the trust-region subproblem using the Lanczos method. *SIAM J. Optim.*, **9**(2): 504–525.

[36] GOULD NIM, ORBAN D & TOINT PhL. 2003. GALAHAD, a library of thread-safe Fortran 90 packages for large-scale nonlinear optimization. *ACM Trans. Math. Software*, **29**(4): 353–372.

[37] GOULD NIM, ROBINSON DP & THORNE HS. 2010. On solving trust-region and other regularised subproblems in optimization. *Math. Program. Comput.*, **2**(1): 21–57.

[38] GOULD NIM, SAINVITU C & TOINT PhL. 2005. A filter-trust-region method for unconstrained optimization. *SIAM J. Optim.*, **16**(2): 341–357.

[39] GRIPPO L, LAMPARIELLO F & LUCIDI S. 1986. A nonmonotone line search technique for Newton's method. *SIAM J. Numer. Anal.*, **23**(4): 707–716.

[40] GRIVA I, NASH SG & SOFER A. 2009 Linear and Nonlinear Optimization. Second edition, Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

[41] GU N-Z & MO J-T. 2008. Incorporating nonmonotone strategies into the trust region method for unconstrained optimization. *Comput. Math. Appl.*, **55**(9): 2158–2172.

[42] HEBDEN MD. 1973. An algorithm for minimization using exact second derivatives. Technical Report T.P. 515, AERE Harwell Laboratory, Harwell, Oxfordshire, England.

[43] KANZOW C & PETRA S. 2007. Projected filter trust region methods for a semismooth least squares formulation of mixed complementarity problems. *Optim. Methods Softw.*, **22**(5): 713–735.

[44] KARASÖZEN B. 2007. Survey of trust-region derivative free optimization methods. *Journal of Industrial and Management Optimization*, **3**(2): 321–334.

[45] LIU J & MA C. 2013. A nonmonotone trust region method with new inexact line search for unconstrained optimization. *Numer. Algorithms*, **64**(1): 1–20.

[46] LUCIDI S, PALAGI L & ROMA M. 1998. On some properties of quadratic programs with a convex quadratic constraint. *SIAM J. Optim.*, **8**(1): 105–122.

[47] LUKŠAN L, MATONOHA C & VLČEK J. 2007. Trust-region interior-point method for large sparse l_1 optimization. *Optim. Methods Softw.*, **22**(5): 737–753.

[48] MACIEL MC& MENDONÇA MG. 2013. Trust-region-filter method for nonlinear optimization problems: Review and new proposal. In: G.L. Mura, D. Rubio, and E. Serrano, editors, MACI Vol. 4 (2013), pages 369–372. ASAMACI – Asociación Argentina de Matemática Aplicada, Computacional e Industrial, Argentina.

[49] MARAZZI M & NOCEDAL J. 2002. Wedge trust region methods for derivative free optimization. *Math. Program.*, **91**(2, Ser. A): 289–305.

[50] MARTÍNEZ JM & SANTOS SA. 1995. A trust-region strategy for minimization on arbitrary domains. *Math. Program.*, **68**(3, Ser. A): 267–301.

[51] MARTÍNEZ JM & SOBRAL FNC. 2013. Constrained derivative-free optimization on thin domains. *J. Global Optim.*, **56**(3): 1217–1232.

[52] MO J, LIU C & YAN S. 2007. A nonmonotone trust region method based on nonincreasing technique of weighted average of the successive function values. *J. Comput. Appl. Math.*, **209**(1): 97–108.

[53] MORÉ JJ. 1978. The Levenberg-Marquardt algorithm: implementation and theory. In: Numerical Analysis (Proc. 7th Biennial Conf., Univ. Dundee, Dundee, 1977), pages 105–116. Lecture Notes in Math., Vol. 630. Berlin: Springer.

[54] MORÉ JJ. 1983. Recent developments in algorithms and software for trust region methods. In: Mathematical Programming: the state of the art (Bonn, 1982), pages 258–287. Berlin: Springer.

[55] MORÉ JJ & SORENSEN DC. 1983. Computing a trust region step. *SIAM J. Sci. Statist. Comput.*, **4**(3): 553–572.

[56] MORÉ JJ & SORENSEN DC. 1984. Newton's method. In: Studies in Numerical Analysis, volume 24 of MAA Stud. Math., pages 29–82. Washington, DC: Math. Assoc. America.

[57] NOCEDAL J & WRIGHT SJ. 1999. Numerical Optimization. Springer Series in Operations Research. First edition. New York: Springer.

[58] NOCEDAL J & YUAN.Y-X. 1998. Combining trust region and line search techniques. In: Advances in Nonlinear Programming (Beijing, 1996), volume 14 of Appl. Optim., pages 153–175. Dordrecht: Kluwer Acad. Publ.

[59] PALAGI L. 2009. Large scale trust region problems. In: C.A. Floudas and P.M. Pardalos, editors, Encyclopedia of Optimization, pages 1822–1831. Springer.

[60] POWELL MJD. 1970. A Fortran subroutine for solving systems of nonlinear algebraic equations. In: Numerical methods for nonlinear algebraic equations (Proc. Conf., Univ. Essex, Colchester, 1969), pages 115–161. London: Gordon and Breach.

[61] POWELL MJD. 1970. A new algorithm for unconstrained optimization. In: Nonlinear Programming (Proc. Sympos., Univ. of Wisconsin, Madison, Wis., 1970), pages 31–65. New York: Academic Press.

[62] POWELL MJD. 1974. Convergence properties of a class of minimization algorithms. In: Nonlinear Programming, 2 (Proc. Sympos. Special Interest Group on Math. Programming, Univ. Wisconsin, Madison, Wis., 1974), pages 1–27. New York: Academic Press.

[63] POWELL MJD. 1984. On the global convergence of trust region algorithms for unconstrained minimization. *Math. Program.*, **29**(3): 297–303.

[64] POWELL MJD. 1994. A direct search optimization method that models the objective and constraint functions by linear interpolation. In: Advances in optimization and numerical analysis (Oaxaca, 1992), volume 275 of Math. Appl., pages 51–67. Dordrecht: Kluwer Acad. Publ.

[65] POWELL MJD. 2002. UOBYQA: unconstrained optimization by quadratic approximation. *Math. Program.*, **92**(3, Ser. B): 555–582.

[66] POWELL MJD. 2006. The NEWUOA software for unconstrained optimization without derivatives. In: Large-scale nonlinear optimization, volume 83 of Nonconvex Optim. Appl., pages 255–297. New York: Springer.

[67] POWELL MJD. 2008. Developments of NEWUOA for minimization without derivatives. *IMA J. Numer. Anal.*, **28**(4): 649–664.

[68] POWELL MJD. 2009. The BOBYQA algorithm for bound constrained optimization without derivatives. Technical report DAMTP 2009/NA06, Cambridge. Available at: `http://www.damtp.cam.ac.uk/user/na/NA_papers/NA2009_06.pdf`.

[69] POWELL MJD. 2012. On the convergence of trust region algorithms for unconstrained minimization without derivatives. *Comput. Optim. Appl.*, **53**(2): 527–555.

[70] RENDL F & WOLKOWICZ H. 1997. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Math. Program.*, **77**(2, Ser. B): 273–299.

[71] RIOS LM & SAHINIDIS NV. 2013. Derivative-free optimization: a review of algorithms and comparison of software implementations. *J. Global Optim.*, **56**(3): 1247–1293.

[72] ROJAS M, SANTOS SA & SORENSEN DC. 2000/01. A new matrix-free algorithm for the large-scale trust-region subproblem. *SIAM J. Optim.*, **11**(3): 611–646.

[73] ROJAS M, SANTOS SA & SORENSEN DC. 2008. Algorithm 873: LSTRS: MATLAB software for large-scale trust-region subproblems and regularization. *ACM Trans. Math. Software*, **34**(2): Art. 11, 28.;

[74] SAINVITU C & TOINT PHL. 2007. A filter-trust-region method for simple-bound constrained optimization. *Optim. Methods Softw.*, **22**(5): 835–848.

[75] SARTENAER A. 1997. Automatic determination of an initial trust region in nonlinear programming. *SIAM J. Sci. Comput.*, **18**(6): 1788–1803.

[76] SHI Z-J & GUO J-H. 2008. A new trust region method for unconstrained optimization. *J. Comput. Appl. Math.*, **213**(2): 509–520.

[77] SHI Z-J & WANG S. 2011. Nonmonotone adaptive trust region method. *European J. Oper. Res.*, **208**(1): 28–36.

[78] SHULTZ GA, SCHNABEL RB & BYRD RH. 1985. A family of trust-region-based algorithms for unconstrained minimization with strong global convergence properties. *SIAM J. Numer. Anal.*, **22**(1): 47–67.

[79] SORENSEN DC. 1982. Newton's method with a model trust region modification. *SIAM J. Numer. Anal.*, **19**(2): 409–426.

[80] SORENSEN DC. 1997. Minimization of a large-scale quadratic function subject to a spherical constraint. *SIAM J. Optim.*, **7**(1): 141–161.

[81] STEIHAUG T. 1983. The conjugate gradient method and trust regions in large scale optimization. *SIAM J. Numer. Anal.*, **20**(3): 626–637.

[82] TAO PD & AN LTH. 1998. A D.C. optimization algorithm for solving the trust-region subproblem. *SIAM J. Optim.*, **8**(2): 476–505.

[83] TOINT PHL. 1981. Towards an efficient sparsity exploiting Newton method for minimization. In: DUFF IS, (editor), Sparse matrices and their uses, pages 57–88. London: Academic Press.

[84] VAVASIS SA. 1992. Approximation algorithms for indefinite quadratic programming. *Math. Program.*, **57**(2, Ser. B): 279–311.

[85] WANG X. 2013. A trust region affine scaling method for bound constrained optimization. *Acta Mathematica Sinica*, English Series, **29**(1): 159–182.

[86] WANG Z-H & YUAN Y-X. 2006. A subspace implementation of quasi-Newton trust region methods for unconstrained optimization. *Numer. Math.*, **104**(2): 241–269.

[87] ZHANG H, CONN AR & SCHEINBERG K. 2010. A derivative-free algorithm for least-squares minimization. *SIAM J. Optim.*, **20**(6): 3555–3576.

[88] ZHANG H & HAGER WW. 2004. A nonmonotone line search technique and its application to unconstrained optimization. *SIAM J. Optim.*, **14**(4): 1043–1056.

[89] ZHANG J-L & ZHANG X-S. 2003. A nonmonotone adaptive trust region method and its convergence. *Computers and Mathematics with Applications*, **45**(10–11): 1469–1477.

[90] ZHANG X, ZHANG J & LIAO L. 2002. An adaptive trust region method and its convergence. *Sci. China Ser. A*, **45**(5): 620–631.