Reflector Antenna Optimization using One-Sided Least-Squares

Anders Eltved[1], Oscar Borries[2], Martin S. Andersen[1]
[1]DTU Compute, Technical University of Denmark, Kongens Lyngby, Denmark, {aelt,mskan}@dtu.dk
[2]TICRA, Copenhagen, Denmark, ob@ticra.com

***Abstract*—Numerical optimization is commonly used for the design of reflector antennas that produce a contoured beam. The optimization problems that arise in this context are generally nonlinear and difficult to solve to global optimality. Different problem formulations lead to different properties and challenges. We demonstrate that the popular minimax approach can have many local minima and that a local optimization method with a poor initialization may result in a bad design. As an alternative to the minimax approach, we propose a one-sided least-squares formulation. We outline a trust-region method for finding a local minimum, and we illustrate the merits of the new approach with some numerical examples. In particular, our preliminary results indicate that our method is often less sensitive to the initial design than the minimax method. Finally, we investigate the use of the one-sided least-squares model as a means to find an initial guess for the minimax approach, and our results show that it sometimes leads to a better local minimum, and hence improves the design.**

I. Introduction

Numerical optimization plays an important role in the design of reflector antennas for modern communication satellites. For contoured beams, optimization is used to produce the desired coverage by means of surface shaping. A common design goal is to produce a system that meets a set of specified goals in the coverage, e.g., ensuring some level of performance in the entire coverage region.

The performance of a reflector system can be simulated by applying Physical Optics (PO) which is sometimes augmented with the Physical Theory of Diffraction (PTD) to take the effects of diffraction from the reflector edges into account. With PO/PTD, it is possible to optimize the antenna design by evaluating the gain, the cross-polar performance, and other relevant quantities across the coverage region as the optimization variables are modified. The objective is to design a system that meets some specified goals at given points in the coverage region. To this end, some measure of system performance is needed, i.e., the objective function of the minimization problem must be chosen. A common way to measure the performance of an antenna system is by its worst-case performance, i.e., the performance at the point in the coverage region where the gain is furthest from the specified goal. This leads to a so-called minimax problem which seeks to optimize the worst-case performance; see e.g. [1], [2], [3]. It is a nonlinear optimization problem which means that it is generally too computationally expensive to solve it to global optimality. Thus, in practice, local optimization is used to find a local minimum. This approach often depends on the antenna designer providing a suitable initial design as a starting point for the local optimization. Unfortunately, the quality of the design obtained via the minimax approach is sensitive to the initialization, and for some systems it can be hard to come up with a good initial design.

To address the sensitivity of the minimax approach to the starting point, we propose a different measure of the performance of a particular design. The result is a one-sided nonlinear least-squares (one-sided LS) problem in which all points in the coverage that do not attain their goal enter into the cost function. In other words, the cost function takes all points where the goal is not met into account instead of only the worst point. The resulting problem is still a nonlinear optimization problem, so we will be satisfied with a local minumum. Even so, our results indicate that in practice, the modified problem is much less sensitive to the initial design when compared to the minimax approach. We also explore the use of the one-sided LS problem as a means to find an initial design for the minimax approach. In all of our test cases, this strategy resulted in improved worst-case performance.

The outline of the paper is as follows. In Section II, we formulate the design optimization problem and present a simple case to demonstrate the differences between the minimax approach and the proposed approach that is based on a one-sided nonlinear LS formulation. In Section III, we describe how to locally optimize the one-sided LS problem and how it can be used to initialize the minimax algorithm. We present some numerical results and a short discussion in Section IV, and Section V contains final remarks.

II. Methodology

A. Problem Formulation

We start by defining m functions $f_1, \ldots, f_m$ where $f_i : \mathbb{R}^n \to \mathbb{R}$ represents the performance associated with the ith position in the coverage region as a function of a vector of n design variables. For each of the m positions, we have a given performance goal g_i, and we introduce m scalar weights $w_1, \ldots, w_m$ that may be used to emphasize important goals and de-emphasize less important goals. Furthermore, we define a residual function

$$r_i(x) = w_i(g_i - f_i(x)) \quad (1)$$

for each of the m positions, where x denotes the vector of design variables to be optimized. The elements of x may represent the parameters in a shaped surface, the position or excitation of feeds and arrays, etc.

With the definition of the residual functions given in (1), the antenna design problem can be expressed as a multi-objective optimization problem

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & (r_1(x), \ldots, r_m(x)) \\ \text{subject to} & Cx + d \geq 0 \end{array}$$

where the elementwise inequality $Cx+d \geq 0$ (with $C \in \mathbb{R}^{p \times n}$ and $d \in \mathbb{R}^p$) represents p constraints on the design variables (e.g., restricting the maximum curvature of the produced surface such that the reflector can be manufactured in practice). We obtain a single-objective problem by introducing a cost function $r_0 : \mathbb{R}^m \to \mathbb{R}$

$$\begin{array}{lll} \underset{u \in \mathbb{R}^m, x \in \mathbb{R}^n}{\text{minimize}} & r_0(u) & \\ \text{subject to} & u_i \geq r_i(x), \quad i = 1, \ldots, m & \\ & Cx + d \geq 0, & \end{array} \tag{2}$$

where u is a vector of m auxiliary variables. The minimax problem is a special case of (2) in which the cost function is defined as

$$r_0(u) = \max_i u_i. \tag{3}$$

This function is convex, but it is not everywhere differentiable. As we will illustrate later in the paper, the resulting problem may have many local minima if any of the residual functions are non-convex. As an alternative to the max objective (3), we propose a one-sided least-squares objective of the form

$$r_0(u) = \sum_{i=1}^{m} \max\{0, u_i\}^2. \tag{4}$$

Recall that a negative residual function implies that the associated goal is met, and hence only positive residuals are penalized. Unlike the max objective function (3), the one-sided LS objective is continuously differentiable, and as we demonstrate next, it may overcome some of the difficulties with the max function.

B. The Importance of Initialization

To compare the two cost functions (3) and (4) and to demonstrate the challenges with the minimax approach, we present a simple case where the design objective is to position a feed. The case is based on a simple offset reflector antenna system where the surface has been shaped to provide a good contoured beam for a feed positioned at the focal point of the system. We then move the feed away from the focal point, allowing us to test whether algorithms designed for the two different formulations can move the feed back to the focal point. The position of the feed may be described by the variables (x, y, z) (i.e., $n = 3$) with box constraints $\|(x, y, z)\|_\infty \leq 41.6955\lambda$ (corresponding to $p = 6$ inequality constraints) where λ denotes the wavelength. We have $m = 897$ residual functions.

The focal point is the optimal position of the feed in the minimax formulation since the reflector has been shaped for this position. The focal point is

$$(x_s, y_s, z_s) = (-3.5944, 5.2119, -3.7741)\lambda, \tag{5}$$

where we note that the symmetry plane of the system is $y = 5.2119\lambda$. Thus, for now, we only consider movement in the xz-plane. We will refer to the focal point (5) as the minimax optimum, since this is the point of interest for the optimization.

To compare the two formulations, we now focus on the cost functions restricted to the xz-plane and centered at the position (5). Figure 1 includes a contour plot for both the one-sided LS cost function (4) and the max cost function (3). The one-sided LS cost function is well-behaved in the sense that there appears to be a single minimum whereas the max cost function has multiple local minima with widely different function values. Comparing the contours of the two functions, it is evident that a local optimization algorithm is more likely to terminate in a poor local minimum with the minimax approach if the initial guess is not sufficiently close to the global minimum or a "good" local minimum. This case illustrates that the minimax formulation can be sensitive to the initial guess, and it also suggests that the one-sided LS formulation may be less sensitive to initialization.

III. Implementation

We now turn to the implementation of a trust-region method for local minimization of the one-sided LS problem (problem (2) with the cost function (4)), followed by a description of an initialization strategy for the minimax algorithm based on the one-sided LS algorithm.

A. One-Sided Least-Squares Algorithm

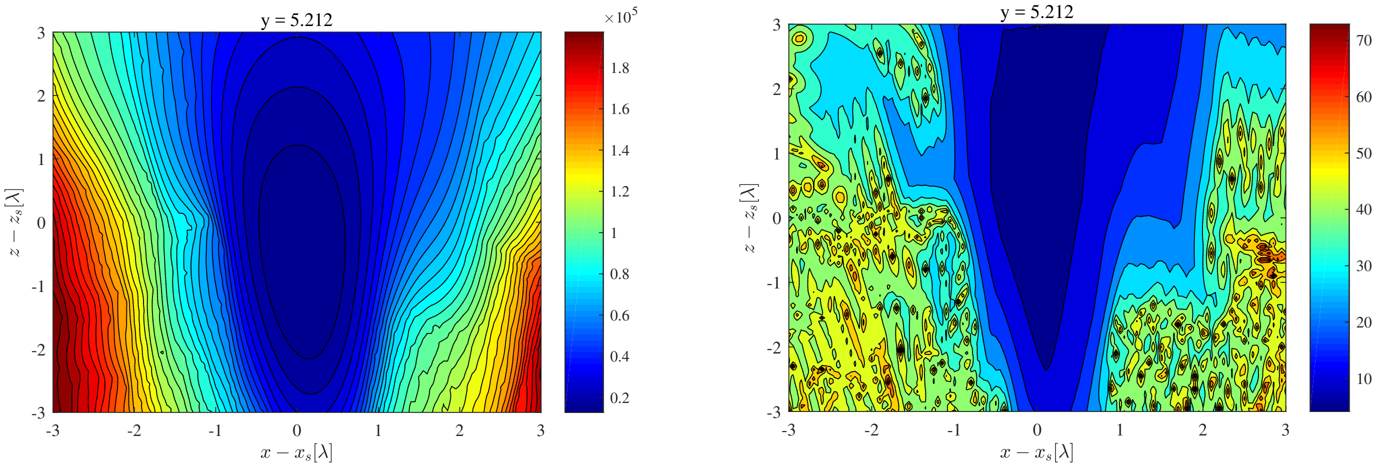
The one-sided least-squares problem is a special case of a general nonlinear least-squares problem of the form

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & \|R(x)\|_2^2 \\ \text{subject to} & Cx + d \geq 0, \end{array} \tag{6}$$

where $R : \mathbb{R}^n \to \mathbb{R}^m$ denotes a vector-valued function with elements $R_i(x) = \max\{0, r_i(x)\}$. The function R is non-differentiable (because of the max-function) and the problem has constraints, so we cannot use the usual nonlinear least-squares methods such as the Gauss-Newton or Levenberg-Marquardt [4], [5] method. Using the epigraph formulation $u_i \geq R_i(x),\ i = 1, \ldots, m$, we arrive at the equivalent problem

$$\begin{array}{lll} \underset{x \in \mathbb{R}^n, u \in \mathbb{R}^m}{\text{minimize}} & \|u\|_2^2 & \\ \text{subject to} & Cx + d \geq 0 & \\ & u_i \geq r_i(x), & i = 1, \ldots, m \\ & u_i \geq 0, & i = 1, \ldots, m. \end{array} \tag{7}$$

The problem has a convex quadratic objective function, but the inequality constraints $u_i \geq r_i(x)$ may be non-convex. Solvers such as IPOPT [6] and SNOPT [7] can be used to optimize (7), but these solvers require second-order derivatives which



(a) $f(x) = \sum_{i=1}^{m} \max(0, r_i(x))^2$ (b) $f(x) = \max_{i=1,\ldots,m}\ r_i(x)$

Fig. 1. Contours of (a) one-sided LS objective and (b) worst-case (minimax) objective.

unfortunately are not available in our application. Furthermore, function evaluations are often quite costly, so we wish to avoid finite-difference approximations of second-order information.

To minimize (7) locally, we propose to use a trust-region method [8]. This requires a model of the problem that serves as a surrogate within a trust-region. At the kth iteration, we obtain a convex model by linearizing the residual functions $r_i(x)$ around x^k, resulting in the trust-region problem

$$\begin{array}{ll} \underset{\Delta x \in \mathbb{R}^n, u \in \mathbb{R}^m}{\text{minimize}} & \|u\|_2^2 \\ \text{subject to} & C\Delta x + d^k \geq 0 \\ & u \geq r^k + J^k \Delta x \\ & u \geq 0 \\ & \delta^k \geq \|\Delta x\|_\infty , \end{array} \tag{8}$$

where $\delta^k > 0$ is the trust-region radius, and

$$\begin{aligned} d^k &= Cx^k + d, \quad r^k = r(x^k), \\ J^k &= \left[\nabla r_1(x^k) \ \ldots \ \nabla r_m(x^k)\right]^T . \end{aligned} \tag{9}$$

The problem (8) is a linearly constrained convex quadratic program (since the ∞-norm is linearly representable), and we can express it in standard form as

$$\begin{array}{ll} \underset{y \in \mathbb{R}^{m+n}}{\text{minimize}} & \frac{1}{2} y^T H y \\ \text{subject to} & A^T y \geq b \end{array} \tag{10}$$

with

$$\begin{aligned} y &= \begin{bmatrix} u \\ \Delta x \end{bmatrix}, \ H = \begin{bmatrix} 2I_{m\times m} & 0_{m\times n} \\ 0_{n\times m} & 0_{n\times n} \end{bmatrix}, \\ A^T &= \begin{bmatrix} 0_{p\times m} & C \\ I_{m\times m} & -J^k \\ I_{m\times m} & 0_{m\times n} \\ 0_{n\times m} & -I_{n\times n} \\ 0_{n\times m} & I_{n\times n} \end{bmatrix}, \ b = \begin{bmatrix} -d^k \\ r^k \\ 0_m \\ -\bar{\delta}^k \\ -\bar{\delta}^k \end{bmatrix} \in \begin{bmatrix} \mathbb{R}^p \\ \mathbb{R}^m \\ \mathbb{R}^m \\ \mathbb{R}^n \\ \mathbb{R}^n \end{bmatrix}. \end{aligned} \tag{11}$$

We use an interior-point method (IPM) to solve (10) which yields a step Δx. IPMs are iterative methods, and the main computation in each iteration is the solution of the linear system

$$(H + ADA^T)\Delta y = v, \tag{12}$$

where D is a diagonal matrix and v is a vector. The matrix on the left-hand side of (12) is of order $n+m$, so factorizing this matrix costs $O((n+m)^3)$. However, under the assumption that $n < m$ (i.e., we have fewer variables than functions), we can use the approach described in [9] which takes advantage of the structure in H, A, and b. With this approach, the complexity of one interior-point iteration becomes $O(n^2(m+p))$. In other words, the cost grows linearly with the number of residual functions.

B. Initialization Strategy

As demonstrated with the feed positioning case in Section II-B, the quality of a local minimum obtained with the minimax formulation can be more sensitive to the initial guess than with the one-sided LS formulation. However, the two formulations yield different solutions in general, and the solutions obtained with the minimax approach may be preferred over those obtained with the one-sided LS approach. In the following, we assume that a minimax solution is of interest, and we wish to investigate the use of a (locally optimal) one-sided LS solution as initialization for the minimax algorithm. We begin by optimizing the one-sided LS cost function starting from the provided initial guess, and then the (locally optimal) solution is used as initial guess for the minimax algorithm.

To avoid many costly function evaluations, we also consider initialization with a crude, approximate local solution to the one-sided LS problem. Solving the one-sided LS problem accurately may require a lot of function evaluations, but high accuracy may not be necessary to obtain a good initial guess for the minimax algorithm. We will assume that the user chooses a maximum number of iterations, and half of these

TABLE I

COMPARING COST FUNCTIONS USING THE DIFFERENT FORMULATIONS AND ALGORITHMS FOR FEED POSITIONING. THE LEFT COLUMN IS THE MINIMAX COST FUNCTION AND HAS UNIT dB. THE RIGHT COLUMN IS THE ONE-SIDED LS COST FUNCTION AND HAS UNIT dB^2.

	$\max_i r_i(x)$	$\sum_i \max\{0, r_i(x)\}^2$
Minimax algorithm	35.790	332,020
One-sided LS algorithm	5.517	12,772
Minimax optimum	4.151	13,493

will be assigned to computing an initial guess using the one-sided LS algorithm, and the other half will be assigned to the minimax algorithm. For example, if the user specifies a maximum of 500 function calls, we will stop the one-sided LS algorithm after 250 iterations (or earlier if the stopping criteria are met), and then we run the minimax algorithm with no more than 250 iterations.

We conclude this section by mentioning that our implementation of the one-sided LS algorithm may be improved by removing constraints in the trust-region subproblem (8) that will remain inactive within a given trust-region. This would most likely reduce the computation time, but since we assume that function evaluations are costly, the added complexity of eliminating constraints may not be worthwhile.

IV. NUMERICAL RESULTS

We begin by revisiting the case described in Section II-B before we move on to a more complicated case of surface shaping of a reflector with two feeds. Note that we have chosen not to report the computational time of the algorithms used to produce the results since this is insignificant compared to the time it takes to evaluate the objective function.

A. Feed Positioning

Recall that for the case described in Section II-B, the focal point coincides with the minimax optimum. We then moved the feed out of position to see if the optimization algorithm can bring it back. We now compare the performance of the minimax and one-sided LS algorithms using the same initial guess.

$$(x, y, z) = (0, 0, 0). \tag{13}$$

Note that the contours in Figure 1 are centered at the focal point (5), and the initial guess is outside the top right corner at another depth (a different y) than the slice shown in the contour plots. With this initial guess, the minimax algorithm stops after 5 iterations, and the one-sided LS algorithm stops after 38 iterations.

Table I shows the value of the cost functions in the different formulations at the obtained local optimums as well as the value at the known minimax optimum. We see from the table that the minimax algorithm does not perform well with this initial guess; the maximal residual is 31.64 dB higher than at the focal point. On the other hand, the one-sided LS algorithm yields a solution of comparable magnitude to the minimax optimum; the largest residual is 1.37 dB from the largest residual value at the minimax optimum. However, the sum of the squared positive residuals, which is what the one-sided LS formulation seeks to minimize, is lower than at the minimax optimum. This suggests that a comparison based on the numbers in Table I is insufficient to show the qualitative difference between the two approaches. Therefore, we consider all residuals by plotting a histogram of these in Figure 2. The histograms show a clear difference in the solutions. The value of the worst residual is lower for the minimax optimum, but many residuals take a value close to this. For the one-sided LS solution, some residuals have better performance (lower value) at the expense of others.

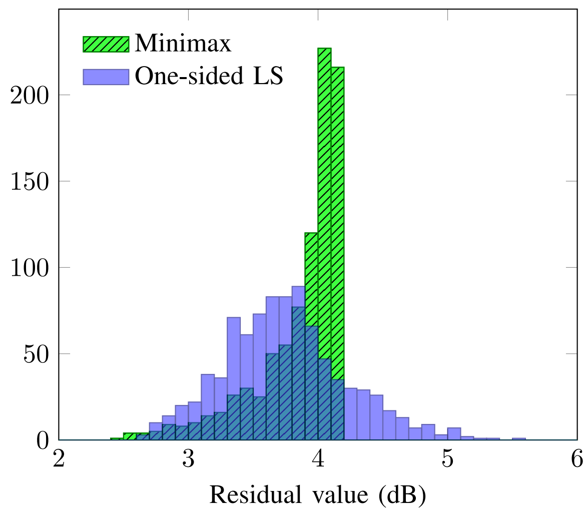


Fig. 2. Histogram of residuals at optimum in the two different formulations. Smaller residual values (left) are more desirable.

We now investigate our initialization strategy from Section III-B, i.e., we use the output from the one-sided LS algorithm as initialization for the minimax algorithm. With this approach, the minimax algorithm used 5 iterations (in addition to 38 iterations for the one-sided LS algorithm) and reached the minimax optimum that is listed in the last row of Table I and shown in green in Figure 2. In other words, the one-sided LS algorithm took us from the initial position to the position represented by the blue histogram, and the minimax algorithm took us from the point represented by the blue histogram to the point represented by the green histogram.

The histograms in Figure 2 are qualitatively different; with the minimax solution we can guarantee a certain minimum gain for all points in the coverage, while the one-sided LS solution has a better average gain for the points in the coverage.

In this particular case, all residuals are positive which suggests that we could have used a standard nonlinear least-squares algorithm instead of the one-sided LS algorithm. However, if we were to reduce some or all of the goals by, say, 4 dB (which corresponds to subtracting 4 from all residuals), the one-sided LS formulation could still be used with a desired behavior, but a standard nonlinear LS formulation would penalize negative residuals. Moreover, if there exists a design for which all residuals are negative, the one-sided LS cost function (4) and its gradient become zero which means that we have reached a global minimum. However, we could then increase the goals to aim for a better design, or we could use the one-sided LS solution as an initial guess in the minimax algorithm.

B. Surface Shaping

As a second case, we consider the radiation from a reflector illuminated by two feeds, intended to produce two separate hemisphere beams. The case is illustrated in Figure 3. We note that this is a so-called super-coverage antenna, intended to be able to provide dual-hemisphere coverage from a position either over the Atlantic Ocean or over the Pacific Ocean.

The surface of the reflector is discretized using 355 spline variables, and since we also optimize the xy-position of the feeds in the focal plane, the total number of variables in this case is $n = 359$. The coverage region is discretized using $m = 658$ residuals. To ensure that the reflector can be manufactured, we include $p = 2512$ constraints such that the surface curvature is at most $0.1\,\mathrm{m}$.

The initial guess for the feed positions is determined by the spacing between the two intended beams, while the initial guess for the reflector surface is determined by defocusing the reflector until an area the size of the coverage is sufficiently illuminated.

We optimize by applying the initialization strategy, starting with 250 iterations with the one-sided LS algorithm, followed by 250 iterations with the minimax algorithm. We end up with a maximum residual of 7.444 dB. As a reference, if we just perform 500 iterations with the minimax algorithm, we end up with a significantly higher maximum residual of 8.557 dB.

The solution based on the initialization strategy is illustrated in Figure 4. Clearly, the beams are not as precisely shaped as for standard single-coverage shaped beams, due to the intended super-coverage of the antenna.

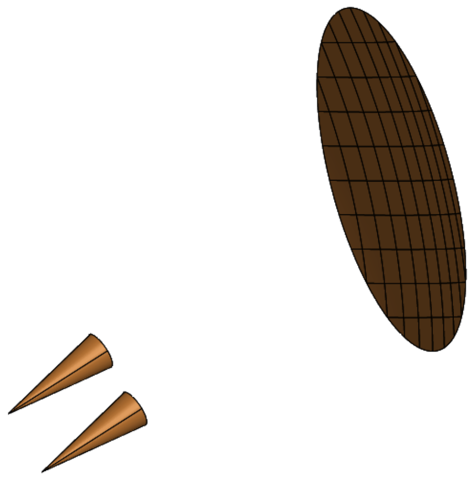


Fig. 3. Scenario 2.

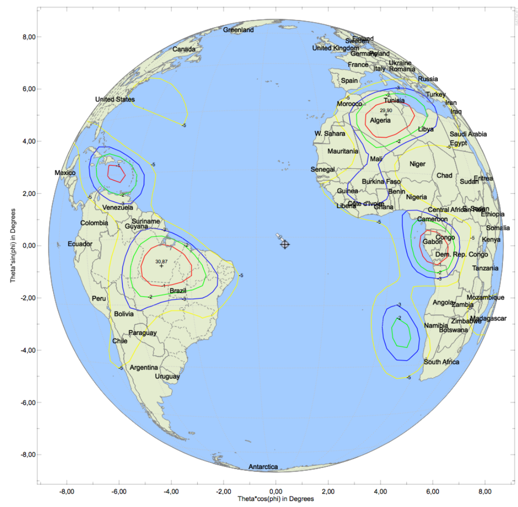


Fig. 4. Illustration of the dual hemisphere coverage.

V. Conclusions

We have proposed a new approach to reflector antenna design based on one-sided least-squares, and we have demonstrated its usefulness as (i) an alternative to worst-case optimization and (ii) as an initialization method for worst-case optimization. The problem can be optimized locally using a trust-region method, and compared to the minimax algorithm, our one-sided LS algorithm appears to be less sensitive to the initial guess, easing the demand on the end-user to provide a good initial design. Moreover, we have demonstrated that the use of the one-sided LS formulation for finding an initial design for the minimax algorithm may result in improved designs as well as fewer iterations.

References

[1] V. Demyanov and V. Molozemov, *Introduction to Minimax*. Wiley: New York, 1974.

[2] J. Hald, *Two stage algorithms for linearly constrained minimax optimization*. PhD thesis, Technical University of Denmark, 1984.

[3] K. Madsen and H. Schjær-Jacobsen, “Linearly constrained minimax optimization,” *Mathematical Programming*, vol. 14, no. 1, 1978.

[4] K. Levenberg, “A method for the solution of certain non-linear problems in least square,” *Quarterly of applied mathematics*, 1944.

[5] D. W. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[6] A. Wächter and L. T. Biegler, “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming,” *Mathematical Programming*, vol. 106, pp. 25–57, Mar 2006.

[7] P. E. Gill, W. Murray, and M. A. Saunders, “SNOPT: An SQP algorithm for large-scale constrained optimization,” *SIAM Rev.*, vol. 47, 2005.

[8] A. Conn, N. Gould, and P. Toint, *Trust Region Methods*. Society for Industrial and Applied Mathematics, 2000.

[9] S. Sra, S. Nowozin, and S. Wright, “Interior-point methods for large-scale cone programming,” in *Optimization for Machine Learning*, ch. 3, MIT Press, 2012.