# Evolution strategies for robust optimization

Kruisselbrink, J.W.

**Author**: Kruisselbrink, Johannes Willem
**Title**: Evolution strategies for robust optimization
**Date**: 2012-05-10

# Chapter 9

# Conclusion

In this work we have presented a definition and a framework of robust optimization that extends the definition of classical optimization and provided practical guidelines for approaching such problems. For two particular scenarios of robust optimization, namely the problem of optimization of noisy objective functions and the problem of finding robust optima, we have studied how Evolution Strategies should be adapted in order to tackle such problems. For these two scenarios we have provided a systematic overview of existing methods and pointed out yet unexplored directions of algorithmic improvement. The algorithmic contributions presented in this work aim to fill these blanks. The empirical comparisons presented in this work provide a refined benchmark set for algorithmic comparison for the two scenarios of robust optimization.

Section 9.1 provides a summary of this work and discusses the conclusions that can be drawn from this research and Section 9.2 closes with an outlook on possible future directions of research.

## 9.1 Summary

Real-world optimization problems often involve various types of uncertainties and noise emerging in different parts of the optimization problem. When not accounting for these matters, optimization may fail, or may yield solutions that are optimal in the classical strict notion of optimality, but fail in practice. Robust optimization is the practice of optimization that accounts for uncertainties and/or noise in the system or (simulation) model. That is, it considers all types of noise and uncertainties that emerge within the system or (simulation) model, but it does not include uncertainties in the formulation of the goals and constraints.

The goal of robust optimization is twofold: 1) to find optimal solutions despite uncertainties and noise in the optimization model, and 2) to find optimal solutions that are robust with respect to uncertainties and noise, and therewith useful in practice. Dealing with robust optimization problems requires the integration of the notion of robustness in the specification of solution quality. That is, effective objective and constraint functions are needed that incorporate the notion of robustness or robust quality. This notion changes the original goal of optimization, because robustness and solution quality are often conflicting objectives.

The different sources and types of uncertainty and noise causes a combinatorial explosion of different robust optimization scenarios. However, some scenarios occur more often than others. In this work, two particular robust optimization scenarios are considered: 1) optimization of noisy objective functions, and 2) finding robust optima. These two scenarios frequently emerge in different forms in real-world optimization settings. It is studied how two Evolution Strategy instances, namely the $(5/2_{DI}, 35)$-$\sigma$SA-ES and the CMA-ES, perform in their canonical form on these two scenarios and how they should be adapted to make them more robust.

Optimization of noisy objectives requires a measure for optimization that includes an account for noise, i.e., an effective objective function. When considering the expected quality of candidate solutions as such a measure, Evolution Strategies are fairly insensitive when the noise is relatively small. However, if a higher convergence accuracy is required, additional measures should be taken.

Implicit and explicit averaging provide a straightforward way to increase the convergence accuracy. Implicit averaging refers to the practice of increasing the population size and explicit averaging refers to assessing the quality of candidate solutions by taking the average over multiple evaluations. However, these two techniques require the a priori specification of a sample size or population size and still yield a limited convergence accuracy.

Adaptive averaging techniques are extensions of static noise handling techniques that aim to automatically adapt the evaluation intensity within the process of evolution. These techniques consist of two components: 1) an uncertainty quantification mechanism that measures the effects of noise on the selection operator, and 2) an uncertainty treatment mechanism that involves a static noise handling scheme (such as explicit averaging) and a way to adapt the evaluation intensity (e.g., the sample size) based on the uncertainty quantification.

In this work we consider adaptive averaging schemes that are based on explicit averaging as noise handling method. For a simple quadratic model with Gaussian additive noise it is shown that for an optimally tuned adaptive averaging strategy the resampling effort grows cubically with the inverse distance to the optimum. To achieve a linear convergence rate over the generations, it is thus necessary to at least exponentially increase the resampling effort.

An empirical study shows that an uncertainty measure that is based on rank-differences that emerge when splitting up the evaluation in two rounds is the most promising method for quantifying the uncertainty. Moreover, it is shown that adaptive averaging schemes can yield results comparable to well tuned static noise handling schemes. That is, except for one scenario; a well-tuned implicit averaging scheme for the CMA-ES outperforms all other methods. Being less sensitive to parameter settings, adaptive averaging techniques provide a good alternative to implicit and explicit averaging techniques.

When aiming to find robust optima given anticipated input uncertainty, a number of different effective objective functions can be constructed. Among these, the expected solution quality under consideration of its possible perturbations is a common measure. The expected objective function can be seen as an integral transform of the original objective function. The difficulty of this scenario lies in the fact that precise evaluation of the effective objective function is impossible, hence methods are needed that can approximate the robust quality of candidate solutions. Two types of robust optima can be distinguished: 1) shifted robust optima and 2) emergent robust optima. These types yield two distinct challenges.

The simplest approach for finding robust optima is to do nothing at all, but to rely on the inherent capabilities of Evolution Strategies to target the more robust peaks. This myopic approach is supported by the observation that Evolutionary Algorithms already have an inherent tendency to converge to the more robust parts of the search space. However, it fails when the robust optimizer is a shifted version of the original optimizer.

When actively targeting for robustness, Monte-Carlo integration methods can be used to approximate the expected objective function values for candidate solutions. It is pointed out that doing so yields approximations of the expected objective function value of candidate solutions that makes the problem of finding robust optima very similar to optimization of noisy objective functions. However, in this scenario the noise in the objective function is due to approximation errors and not an inherent part of the system. The limitation of Monte-Carlo integration methods is that they are limited in approximation accuracy and therefore limit the accuracy with which robust optima can be targeted.

Similar to when dealing with noisy objective functions, adaptive noise handling strategies can also be used for finding robust optima. Using this approach has the same advantages as with noisy optimization, namely that it does not suffer from convergence accuracy limitations and it does not require the a priori specification of a sample size.

Another branch of approaches is formed by archiving and metamodeling approaches. These approaches store previously evaluated candidate solutions and use these for estimating the objective function values for newly generated candidate solutions. This makes them especially useful when objective function evaluations are (computationally) expensive. An archive maintenance approach is reviewed that incorporates an advanced scheme to update the archive and to make sure that it is well usable for the obtaining reliable approximations for newly generated candidate solutions. Besides this, a Kriging metamodeling approach is considered and tested that uses the archive not directly, but builds a model from which effective objective function estimates are obtained.

The idea of using niching approaches for the goal of finding robust optima has the alleged advantage that the search focuses on more regions of the search space (which in particular for emergent robust optima looks promising). A straightforward implementation of a standard niching strategy shows, however, that using this idea directly introduces more problems than it solves. For these kinds of purposes, niching strategies are required that can deal with noisy objective functions.

Last but not least, a method to boost accuracy when aiming to find robust optima is to exploit the overlap of the regions of uncertainty (or $\eta$-neighborhoods) of candidate solutions and base the evaluation on how pairs of solutions compare rather than aiming to obtain precise approximations of the effective objective function.

The results of an empirical comparative study show that an adaptive averaging strategy using Latin Hypercube Sampling and using the same disturbances for all individuals in the population for evaluation of the effective objective function is the most promising approach. Compared to the myopic approach and optimally tuned resampling, it yields better performance across the set of test problems. The archive based evaluation approach and the metamodel based approach yield a good performance on specific test problems.

## 9.2 Outlook

There are many other scenarios that have not been discussed in detail, but which also fall within the scope of robust optimization. Although the observations for the two scenarios considered in this work can be used to a great extent in other scenarios as well, particular dynamics remain to be studied. In particular the scenario of optimization under uncontrollable perturbations of the design variables (see, e.g., [BOS03, SBO04, BS06b]) and the scenario of optimization given uncertainties in the environmental parameters form two interesting classes. Besides this, both scenarios considered in this work can be researched in more depth. For instance, by considering different types of noise distributions or uncertainties and different types of effective objective function measures.

Extending this research to multi-objective optimization is another direction that is useful for many real-world optimization problems. Approaches that have been proposed to find robust optima for multi-objective optimization are presented in, e.g., [JS03, DG06, GA05, LAA05, LOL05, BA06, PL09, Bad10, SRS11]. Linking the observations and findings of this thesis to these studies forms a challenging project. Along the same line of thought, it would interesting to extend this work to robust optimization with constraints.

In this work, the focus was on Evolution Strategies, and in particular the $(5/2_{DI}, 35)$-$\sigma$SA-ES and the CMA-ES. Extending this research in the direction of other Evolution Strategy variants or other Evolutionary Algorithms is another logical next step. Besides that, it would be interesting to see to what extent the methods for robust optimization that have been presented in this work can be used within other types of optimization algorithms.

Besides extending this research to different types of robust optimization scenarios, some approaches are well worth to be investigated in more detail, also from the algorithmic perspective.

In Section 5.4.6, an alternative uncertainty quantification approach is proposed. The empirical results of Chapter 6 show that this approach yields comparable results to the uncertainty quantification measure of Hansen et al. [HNGK09] (see Section 5.4.5). This measure counts the rank-inversions for the uncertainty quantification and compares them with known statistics on purely random orderings. Compared to the uncertainty quantification of Hansen et al. [HNGK09], it has a more stable statistical basis, it is simpler to implement, and yields comparable results. Investigating this uncertainty quantification in more depth is therefore important.

The empirical results of Chapter 6 show that implicit averaging is a very efficient way of countering the effects of noise for the CMA-ES. Studying adaptive averaging techniques that use implicit averaging as noise treatment scheme is therefore a worthwhile object of study.

Maintaining an archive of previously evaluated solutions for the purpose of reusing them for the evaluation of other candidate solutions can enhance the efficiency of optimization when function evaluations are costly. A promising approach for doing so is presented in Section 7.2.7. This approach aims to fill the archive such that it is better spread in the region of interest (i.e., the region around candidate solutions). Also combining this approach in combination with other metamodeling techniques, as shown in Section 7.2.8, is worthwhile being studied.

Also niching for finding robust optima (discussed in Section 7.2.9) is an issue worthwhile of further investigation. Although the straightforward utilization of a standard niching technique does not function well, the idea of focusing the optimization on multiple parts of the search space is interesting. Assuming that the noise in the robustness approximations of the candidate solutions is the main cause of the failure of the niching approach in this context, niching for noisy objective functions is an interesting line of study.

A particularly promising idea is to exploit the overlap in the regions of uncertainty when aiming to find robust optima, presented in Section 7.2.10. For this only a proof of concept is given. Extending this method for efficient comparison among multiple candidate solutions could potentially make this approach very suitable for finding robust optima. However, this remains to be studied.

Finally, this work has provided a benchmark setup for empirical comparison. Based on test problems from literature, two benchmark sets have been defined that are small, representative test sets for 1) optimization of noisy objective functions with Gaussian additive noise, and 2) optimization for finding robust optima given uniform input noise. The test problems are described in detail in Appendices A and B. Although the empirical testing in this thesis is limited to 10-dimensional search spaces, the set of benchmarks is generalizable for arbitrary dimensions. The results of the experiments presented in Chapter 6 and Chapter 8 respectively can be used for comparison of new algorithmic methods. Extending the empirical study to higher dimensional problems is a relatively simple, but is interesting future step.