EVOLUTIONARY AND DETERMINISTIC METHODS FOR DESIGN, OPTIMIZATION AND CONTROL
C. Poloni, D. Quagliarella, J. Périaux, N. Gauger and K. Giannakoglou (Eds.)

GAUSSIAN ADAPTATION FOR ROBUST DESIGN CENTERING

Christian L. Müller, Ivo F. Sbalzarini
Institute of Theoretical Computer Science
and Swiss Institute of Bioinformatics
ETH Zurich
CH–8092 Zurich, Switzerland
Email: {christian.mueller, ivos}@inf.ethz.ch
Web page: http://www.mosaic.ethz.ch

Abstract. In real-world applications, engineers are often faced with the task of finding nominal design parameters that guarantee proper operation of a system under uncertainty. The problem of design centering considers finding a parameter set that has maximum distance from the borders of the feasible region in parameter space. We present the application of the stochastic process of Gaussian Adaptation to robust design centering and highlight some of its properties. We argue that Gaussian Adaptation provides information about both the most robust nominal parameter set and the volume of the feasible region of parameter space, a measure of problem robustness.

Key words: design centering, Gaussian Adaptation, maximum entropy, convex volume estimation, STS.

1 INTRODUCTION

Design centering addresses the problem of finding "good" nominal operation points for systems in constrained, uncertain environments[1]. Given the (potentially high-dimensional) real-valued parameter space $\mathbb{R}^n$ the goal is to find a vector of design parameters $\mathbf{x} \in \mathbb{R}^n$ that fulfills two requirements: First, the parameters satisfy the specifications imposed by the engineer, i.e., some objective (or criterion) function $f(\mathbf{x}) = f(\mathbf{y}(\mathbf{x}))$ applied to the system output $\mathbf{y}(\mathbf{x})$. Second, the parameter values should be maximally robust with respect to intrinsic uncertainties during operation. Ideally, one would also like to have information about the expected robustness of the resulting system, i.e., the volume of the "feasible region" $\mathcal{A} \subset \mathbb{R}^n$ of parameter vectors that fulfill the design criteria.

Here, we review the stochastic process of Gaussian Adaptation (GaA) and present its application to design centering under uncertainty. We show that GaA is a promising candidate for robust design centering and present some preliminary work as well as a software implementation. GaA has originally been developed in the context of electrical network design. There, the key goal is to find nominal values of, e.g., resistances and capacities in an analog network that render the circuit's response robust with respect to intrinsic random variations of the components and environmental changes during operation of the electrical device. In the late 1960's Gregor Kjellström, at the time an engineer at the Ericsson Telephone Company, realized that with increasing network complexity classical optimizers such as conjugate gradients perform poorly, especially when analytical gradients are not readily available or when the objective function is multimodal. He suggested to search the space of valid parameter settings with stochastic methods that only rely on evaluations of the objective function. Starting from an exploration method that can be considered an adaptive random walk through design space[2], he refined his algorithm to what he called Gaussian Adaptation[3]. In the following decades, the algorithm has been largely ignored by the optimization and systems design communities.

We have recently revisited and reintroduced GaA in the context of black-box optimization and sampling[4,5]. We contributed several ideas to the basic GaA scheme. First, we showed that the foundation of the algorithm can be derived from Jaynes' Maximum Entropy principle[4,6]. Second, we provided suitable standard parameter settings, initialization and boundary handling schemes, as well as an effective restart strategy[4]. These enhancements render GaA a ready-to-use, parameter-free black-box algorithm. Restart GaA has been tested on the IEEE CEC 2005 benchmark test suite. Its performance ranks GaA among the top black-box optimizers ever tested on this benchmark. Third, we have constructed an adaptive-proposal MCMC sampler based on GaA, called Metropolis-GaA[5]. Together with the extensions presented here, this renders GaA a unifying framework for design centering, black-box optimization, and adaptive MCMC sampling. A MATLAB toolbox implementing GaA for all of these three applications is available from the web page of the authors.

2 DESIGN CENTERING WITH GAUSSIAN ADAPTATION

We present the application of GaA to design-centering problems under uncertainty. Assume that the engineer of an electrical circuit can vary the values of design parameters and can decide whether each parameter set fulfills a specified criterion, or not. How can one describe the set $\mathcal{A} \subset \mathbb{R}^n$ of acceptable parameter vectors in a general and compact manner? Based on Shannon's information theory, Kjellström derived that under the assumption of finite mean $\mathbf{m} \in \mathbb{R}^n$ and covariance $\mathbf{C} \in \mathbb{R}^{n \times n}$ of the samples, a Gaussian distribution may be used to characterize $\mathcal{A}$ optimally[3]. In doing so, although not explicitly stated in the original publication, Kjellström applied the Maximum Entropy principle developed by Jaynes in 1957[6]. There, Jaynes states that the maximum-entropy solution "is the least biased estimate possible on the given information; i.e., it is maximally noncommittal with regard to missing information."[6] In the case of given mean and covariance information, the Gaussian distribution maximizes the entropy $\mathcal{H}$, and hence is the preferred choice to describe the region of acceptable points. The entropy of a multivariate Gaussian distribution is

$$\mathcal{H}(\mathcal{N}) = \ln\left(\sqrt{(2\pi \mathrm{e})^n \det(\mathbf{C})}\right) , \tag{1}$$

where $\mathbf{C}$ is the covariance matrix and e is Euler's number. In order to obtain the most informative characterization of the region $\mathcal{A}$, Kjellström envisioned an iterative sampling strategy with a Gaussian distribution that satisfies the following criteria: (i) The probability of finding a feasible design parameter vector should be fixed to a predefined hitting probability $P < 1$, and (ii) the spread of the samples as quantified by their entropy should be maximized. As Eq. 1 shows, this can be achieved by maximizing the determinant of the covariance matrix under the constraint of the fixed hitting probability.

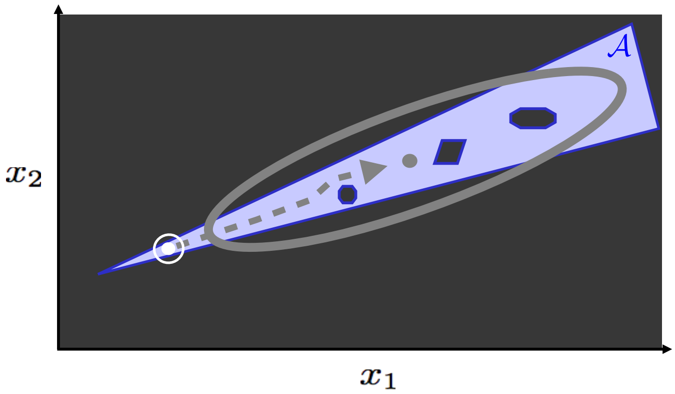

Figure 1: Illustration of Gaussian Adaptation. The light-blue, non-convex area depicts the feasible region $\mathcal{A}$ in a 2D design-parameter space. Both the left (white) and right (gray) dots and ellipsoids represent the means and covariances of two Gaussian distributions with the same hitting probability. Starting from the white distribution, GaA converges to the gray one by moving away from the left corner toward the center of $\mathcal{A}$, adapting the distribution to the shape of $\mathcal{A}$ and circumventing holes in the feasible region.

Algorithm 1: GaA for design centering

Input: $\mathbf{m}^{(0)}$, $\mathbf{C}^{(0)}$, $r^{(0)}$, K, and c_T
Result: $\mathbf{m}^{(K)}$, $\mathbf{C}^{(K)}$, $r^{(K)}$, and P_emp
$s = 0$
for $g = 1, 2, \ldots, K$ **do**
  1. Sample $\mathbf{x}^{(g)} \sim \mathcal{N}\big(\mathbf{m}^{(g-1)}, r^{(g-1)\,2}\mathbf{C}^{(g-1)}\big)$
  2. Evaluate $f(\mathbf{x}^{(g)})$.
  3. **if** $f(\mathbf{x}^{(g)}) < c_\mathrm{T}$ **then**
    $s = s + 1$
    $r^{(g)} = f_\mathrm{e} r^{(g-1)}$
    $\mathbf{m}^{(g)} = (1 - \frac{1}{N_\mathrm{m}})\mathbf{m}^{(g-1)} + \frac{1}{N_\mathrm{m}}\Delta\mathbf{x}^{(g)}$
    $\mathbf{C}^{(g)} = (1 - \frac{1}{N_\mathrm{C}})\mathbf{C}^{(g-1)} + (\frac{1}{N_\mathrm{C}})\Delta\mathbf{x}^{(g)}\Delta\mathbf{x}^{(g)\mathrm{T}}$
    with $\Delta\mathbf{x}^{(g)} = \mathbf{x}^{(g)} - \mathbf{m}^{(g-1)}$.
    Normalize $\mathbf{C}^{(g)}$ such that $\det(\mathbf{C}^{(g)}) = 1$.
  else
    $r^{(g)} = f_\mathrm{c} r^{(g-1)}$
$P_\mathrm{emp} = s/K$

If the system (or the parameters) has to fulfill a predefined, static design criterion, the iterative sampler should push the mean of the distribution toward the *center* of the feasible design space $\mathcal{A}$. Simultaneously, it should adapt the orientation and scale of the covariance matrix to the shape of $\mathcal{A}$ under the constraint of the fixed hitting probability. The final mean can then be used as the nominal design parameter vector, providing maximal robustness against uncertainties in the parameters and in the specified constraints. Figure 1 illustrates this process. Mathematically, the key ingredients for achieving such an adaptation are rank-one updates of the covariance matrix based on sampled feasible directions, as well as proper expansion and contraction of the covariance upon sample acceptance and rejection, respectively. The expansion and contraction factors $f_e > 1$ and $f_c < 1$ are determined by the desired hitting probability and are used to impose the hitting probability as previously described[5]. The overall procedure of GaA for design centering is given in Algorithm 1. The algorithm takes as input the initial mean $\mathbf{m}^{(0)}$, covariance matrix $\mathbf{C}^{(0)}$, and step size $r^{(0)}$, as well as an acceptance threshold c_T (see below) and the total number of iterations K to be performed. It returns the adapted mean $\mathbf{m}^{(K)}$, covariance matrix $\mathbf{C}^{(K)}$, step size $r^{(K)}$, and the actually achieved empirical hitting probability P_{emp}. For additional algorithmic details we refer to the respective publications[3,4]. A discussion of the parameters of the algorithm and their recommended standard settings is also available in the literature[5].

If the objective function $f(\mathbf{x})$ yields real values, the feasible region $\mathcal{A}$ can be defined as the set of all points in parameter space where the objective-function value is below a certain threshold c_T, hence $\mathcal{A} = \{\mathbf{x} : f(\mathbf{x}) < c_T\}$. Repeating the design centering process for decreasing acceptance thresholds c_T then provides a sequence of robust solutions for increasingly stringent design criteria, as illustrated in Fig. 2. For a given threshold c_T, GaA iteratively adapts a Gaussian distribution to the largest region of parameter space where the objective-function values of Gaussian sample points are expected to be below the threshold c_T with probability P. This statistical definition of the feasible region $\mathcal{A}$ allows that both the objective function and the feasible region may be non-convex, although no performance guarantees can be given then.

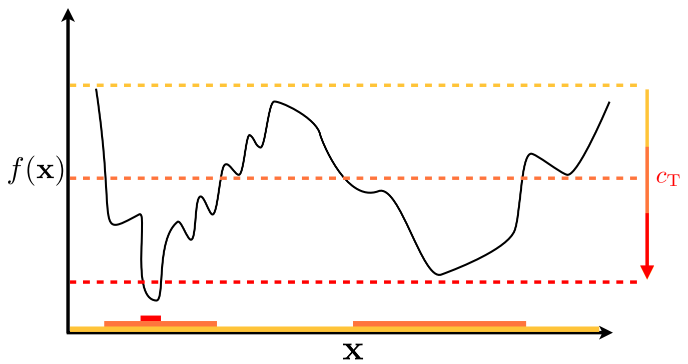

Figure 2: Illustration of lowering the objective-function value threshold c_T in GaA. The dashed lines represent decreasing values of the acceptance threshold c_T. The bars along the $\mathbf{x}$-axis show in the same color the corresponding feasible region of parameter space.

3 FEASIBLE-VOLUME ESTIMATION AS A MEASURE OF ROBUSTNESS

GaA heuristically finds the maximum-volume covariance matrix that fits the feasible region $\mathcal{A}$ with a predefined hitting probability P. This can not only provide the most robust parameter vector as the mean of the adapted distribution, but also a measure for the overall robustness as quantified by the final covariance matrix. In order to estimate a uniform ellipsoidal volume of the feasible region of parameter space, we first compute the eigen-decomposition of $\mathbf{\Sigma} = r^{(K)} r^{(K)} \mathbf{C}^{(K)}$, resulting in $\mathbf{\Sigma} = \mathbf{B}^{\mathrm{T}} \mathbf{\Lambda} \mathbf{B}$, where $\mathbf{B}$ is the orthonormal matrix of the n eigenvectors and $\mathbf{\Lambda}$ the diagonal matrix of the eigenvalues $\lambda_1, \ldots, \lambda_i, \ldots, \lambda_n$. Next, we rescale each λ_i by the factor $c_n = \mathrm{inv}\,\chi^2_n(P_{\mathrm{emp}})$, i.e., the n-dimensional inverse χ^2-distribution with parameter P_{emp}. An approximate ellipsoidal volume estimate $\mathrm{vol}(\hat{\mathcal{E}}_n)$ can then be computed as:

$$\mathrm{vol}(\hat{\mathcal{E}}_n) = \prod_{i=1}^{n} \left(\sqrt{\lambda_i c_n} \right) \mathrm{vol}(\mathcal{S}_n)\,, \tag{2}$$

where $\mathrm{vol}(\mathcal{S}_n) = \pi^{\frac{n}{2}} / \Gamma(\frac{n}{2} + 1)$ is the volume of the n-dimensional unit hyper-sphere. The rescaling converts the probabilistic multivariate Gaussian description into an ellipsoidal description with uniform density. If the region $\mathcal{A}$ is convex, it is thus likely that GaA can provide good volume estimates even in high dimensions. The approximate hyper-ellipsoid can be interpreted as a statistical analog of the *maximum-volume inscribed ellipsoid* in a convex set. These ellipsoids are known to have good approximation properties for polyhedral bodies (see Ref.[7], p. 414, for further information).

In addition, GaA might be a viable method for approximately solving the problem of inner-ellipsoidal approximation of the convex hull of a finite set, a problem that is known to be NP-hard[8]. This problem classically occurs whenever a body (here the feasible region of parameter space) is given by an oracle that can decide for a given sample whether it is inside the body or not (here the objective function with threshold c_{T}). We have investigated this topic only recently, but we expect that estimating the volume of a convex body given by a membership oracle or a finite set of feasible sample points is closely related to robust design centering. The best currently known convex-volume algorithms are without exception randomized schemes, such as the Ball-walk and Hit-and-run samplers[9].

As a proof of concept we present here a numerical benchmark for the quality of approximation of the volumes of anisotropic, axis-aligned, shifted ellipsoids using GaA. We consider ellipsoids $\mathcal{E}_n = \{\mathbf{x} \,:\, (\mathbf{x} - \mathbf{c})^{\mathrm{T}} \mathbf{A}_n (\mathbf{x} - \mathbf{c}) < 1\}$ with $\mathbf{A}_n = \mathrm{diag}(1, \ldots, n)$ an n-dimensional diagonal matrix and the ellipsoid centers at $\mathbf{c} = [0.5, \ldots, 0.5]^{\mathrm{T}}$. The exact volumes of the ellipsoids are:

$$\mathrm{vol}(\mathcal{E}_n) = \det(\mathbf{A}_n^{-1/2}) \mathrm{vol}(\mathcal{S}_n)\,. \tag{3}$$

We numerically estimate $\mathrm{vol}(\mathcal{E}_n)$ for $n = 2, 5, 10, 15, 20$ using GaA with $N_{\mathrm{m}} = 10n$, $N_{\mathrm{C}} = 10n^2$, and $P = 1/\mathrm{e}$. The derivation of the corresponding factors f_{c} and f_{e} is as previously described[5]. Each element of $\mathbf{m}^{(0)}$ is chosen uniformly at random in $[0.45, 0.55]$. The initial covariance is the identity.

GaA generates $K = 1000n^2$ samples in each of 10 independent design-centering runs per tested value of n. In order to estimate the volumes of the above ellipsoids, we use the objective function $f(\mathbf{x}) = (\mathbf{x} - \mathbf{c})^{\mathrm{T}} \mathbf{A}_n (\mathbf{x} - \mathbf{c})$ and $c_{\mathrm{T}} = 1$. The results are summarized in Fig. 3. Figure 3**a** shows the true $\mathrm{vol}(\mathcal{E}_n)$ (blue dots) and the GaA

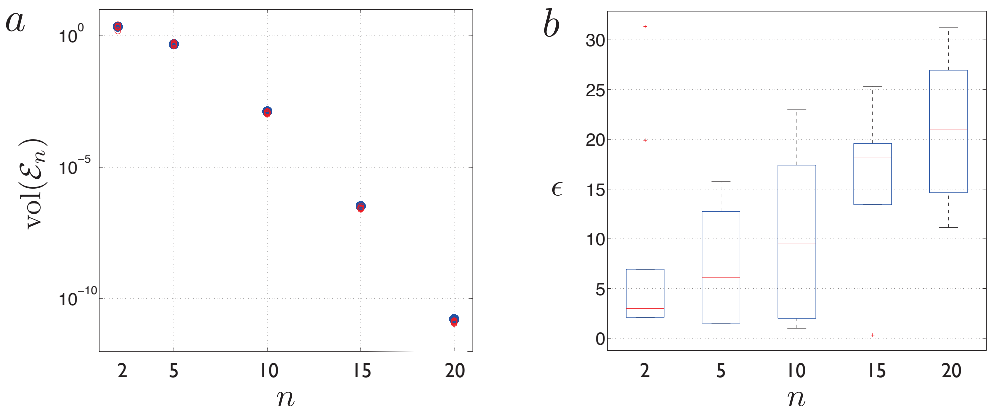

Figure 3: **a**: True ellipsoid volumes (blue dots) and approximate GaA volume estimates (red circles) from 10 independent runs per tested parameter-space dimensionality n. **b**: Box plot of the relative error ϵ of the volume estimation for increasing problem dimensionality.

approximations vol($\hat{\mathcal{E}}_n$) from each of the 10 runs (red circles) per dimensionality n. Figure 3**b** shows a box plot of the relative error

$$\epsilon = \frac{|\mathrm{vol}(\mathcal{E}_n) - \mathrm{vol}(\hat{\mathcal{E}}_n)|}{\mathrm{vol}(\mathcal{E}_n)} \cdot 100\,. \tag{4}$$

As expected, the approximation accuracy decreases with increasing parameter-space dimensionality n. Nevertheless, GaA is able to estimate all volumes within $\epsilon \approx 20\%$ and with very small absolute errors (see Fig. 3). All computational results were generated with a publicly available MATLAB toolbox that we developed (see Fig. 4 for a screenshot).

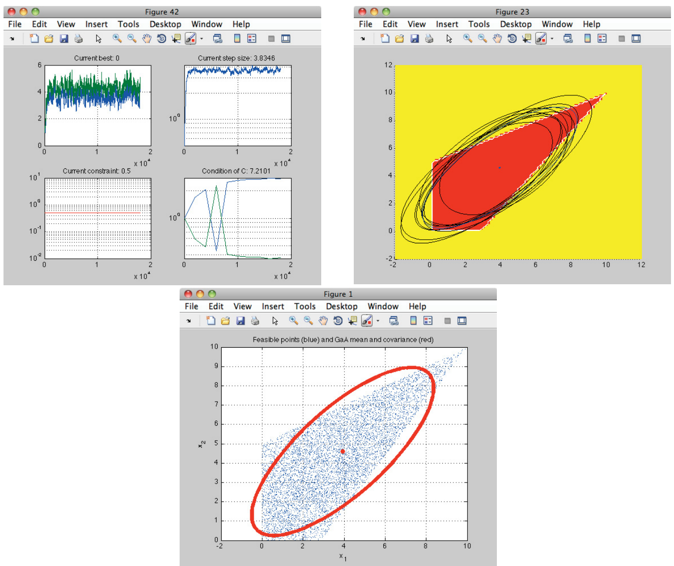

Figure 4: Screenshot of a GaA design-centering run using the present MATLAB toolbox. The upper-left panel monitors the progress of internal variables and the position of GaA's mean. The upper-right panel shows the sequence of covariance matrices used by GaA to iteratively approximate the feasible region (shown in red). The lower centered plot shows the collected feasible samples and the final ellipsoid approximation (in red).

4 CONCLUSIONS

We have detailed the use of Gaussian Adaptation (GaA) for sampling-based design centering under uncertain parameters and uncertain constraints. GaA iteratively learns a maximum-entropy approximation of the feasible region of a design-parameter space by a multivariate Gaussian. The mean of the final Gaussian has maximum distance from any border of the feasible region, hence providing a robust nominal set of design parameters. In addition, GaA can be used to estimate the volume of the feasible region below a given objective-function value threshold as approximated by the volume of the hyper-ellipsoid corresponding to the covariance matrix of the maximum-entropy Gaussian.

We have implemented GaA as a MATLAB toolbox that is freely available from the web site of the authors. The toolbox provides functions for robust design centering, black-box optimization, and adaptive-proposal MCMC sampling using GaA (see Fig. 4 for a screenshot). It also includes several test scripts that illustrate how GaA's strategy and control parameters can be set in the different scenarios. A performance-optimized Fortran90 version of the algorithm with a comfortable MATLAB mex interface will be released in the coming months.

REFERENCES

[1] Graeb, H. E. *Analog Design Centering and Sizing.* Springer, (2007).

[2] Kjellström, G. Network optimization by random variation of component values. *Ericsson Technics* **25**(3), 133–151 (1969).

[3] Kjellström, G. and Taxen, L. Stochastic optimization in system design. *IEEE Trans. Circ. and Syst.* **28**(7), 702–715 July (1981).

[4] Müller, C. L. and Sbalzarini, I. F. Gaussian Adaptation revisited — an entropic view on covariance matrix adaptation. In *Proc. EvoStar*, volume 6024 of *Lect. Notes Comput. Sci.*, 432–441. Springer, April (2010).

[5] Müller, C. L. and Sbalzarini, I. F. Gaussian Adaptation as a unifying framework for continuous black-box optimization and adaptive Monte Carlo sampling. In *Proc. IEEE Congress on Evolutionary Computation (CEC)*, 2594–2601. IEEE, July (2010).

[6] Jaynes, E. T. Information theory and statistical mechanics. *Phys. Rev.* **106**(4), 620–630 (1957).

[7] Boyd, S. and Vandenberghe, L. *Convex Optimization.* Cambridge University Press, March (2004).

[8] Nemirovski, A. Advances in convex optimization: conic programming. In *Proc. Intl. Congress of Mathematicians, Madrid, Spain*, (2007).

[9] Vempala, S. Geometric random walks: a survey. *MSRI Volume on Combinatorial and Computational Geometry* **52**, 577–616 (2005).