Single-stage gradient-based stellarator coil design: stochastic optimization

Florian Wechsung1,∗, Andrew Giuliani1, Matt Landreman2, Antoine Cerfon1 and Georg Stadler1

1 Courant Institute of Mathematical Sciences, New York University, NY, United States of America
2 University of Maryland-College Park, MD, United States of America

E-mail: wechsung@nyu.edu, giuliani@cims.nyu.edu, mattland@umd.edu, cerfon@cims.nyu.edu and stadler@cims.nyu.edu

Received 22 June 2021, revised 11 October 2021
Accepted for publication 15 December 2021
Published 10 May 2022

Abstract

We extend the single-stage stellarator coil design approach for quasi-symmetry on axis from (Giuliani et al 2020) to additionally take into account coil manufacturing errors. By modeling coil errors independently from the coil discretization, we have the flexibility to consider realistic forms of coil errors. The corresponding stochastic optimization problems are formulated using a risk-neutral approach and risk-averse approaches. We present an efficient, gradient-based descent algorithm which relies on analytical derivatives to solve these problems. In a comprehensive numerical study, we compare the coil designs resulting from deterministic and risk-neutral stochastic optimization and find that the risk-neutral formulation results in more robust configurations and reduces the number of local minima of the optimization problem. We also compare deterministic and risk-neutral approaches in terms of quasi-symmetry on and away from the magnetic axis, and in terms of the confinement of particles released close to the axis. Finally, we show that for the optimization problems we consider, a risk-averse objective using the conditional value-at-risk leads to results which are similar to the risk-neutral objective.

Keywords: stellarator, coil optimization, stochastic optimization, risk-averse optimization, optimization for quasi-symmetry

((Some figures may appear in colour only in the online journal)

1. Introduction

The design, manufacturing, and assembly of the primary coil system for stellarators are among the most technically challenging aspects of the construction of stellarators [17, 31, 41], and represent a large fraction of the total construction cost. Some of these challenges are intrinsic to the nature of stellarators: non-axisymmetric coil systems are expected to be more complex than the axisymmetric coil systems of tokamaks [11, 41]. However, some of these challenges are directly linked to the design optimization process, and the optima which have been selected for construction. Specifically, the emphasis has historically been given to optimization metrics corresponding to plasma performance, and less attention was given to the engineering requirements and constraints to achieve such performance. Stellarator designs considered optimal from a physics point of view could only be realized with complex and expensive coil configurations. The difficulties encountered in the construction of large scale stellarators [17, 31, 41] has recently triggered a renewed research effort toward the development of tools enabling the design of simpler and more efficient coil systems that still lead to strong plasma performance [5, 16, 21, 22, 32].

Beside the engineering complexity of a given design, the lack of robustness of the physics performance in the presence of coil manufacturing and assembly errors is also a strong driver for the cost of a machine, since it requires tight tolerances at every step of the manufacturing and assembly process. Both challenges are related but not identical: a relatively
simple coil system whose performance degrades strongly with manufacturing and alignment errors may not be as desirable as a more complex coil system with more robust performance. Work has therefore also been lately devoted to the design of efficient numerical methods for the evaluation of the sensitivity of error fields to coil perturbations [45, 46], and the sensitivity of physical quantities to error fields [13, 23]. These methods can be included in deterministic stellarator optimization codes, and serve to narrow the search to configurations with lower sensitivity or with sensitivity with respect to perturbations that are more easily controlled.

A complementary approach to deal with the challenge of strong sensitivity and tight tolerances is to account for errors during the optimization process, via stochastic optimization. These errors may either be engineering errors or possibly errors due to the limitations of the physics models used. In stochastic optimization, the objective is a function both of the controls and of the random model errors and hence is itself a random variable. This is precisely the approach we adopt in this article, with the randomness modeling errors of the location and shape of the coils. We observe that while stochastic optimization is not yet a widely applied method for the design of magnetic confinement devices, it was used in the design of the CNT stellarator [19, 33]. CNT consists of four circular coils, namely two interlocking (IL) coils and two poloidal field (PF) coils, and the CNT optimization problem involves the angle between the IL coils, as well as the current ratio between the IL and the PF coils. Coil errors involving tilts and shifts of these coils were considered and the goal was to optimize the average volume with good flux surfaces. More recently, for the problem of designing coils corresponding to a desired plasma boundary, Lobsien et al. [28, 29] demonstrated that stochastic optimization leads to simpler and better performing coils.

In this manuscript we present a stochastic version of the single-stage coil design framework recently introduced in [14]. The goal of the single stage framework is to optimize coils directly for the properties of the magnetic field they generate. This is in contrast to the classical two-stage method, which consists of first finding a desirable equilibrium and then searching for coils which generate a magnetic field that is consistent with its boundary.

As the purpose of our work is mainly to introduce a new paradigm for stochastic coil design optimization, the physics basis for the design is simple: we consider a vacuum field, and optimize directly for a target value of the rotational transform and for quasi-symmetry on the magnetic axis. The main contribution of this work is a formulation that enables us to consider more general and realistic coil errors than considered thus far, to implement gradient based optimisation algorithms relying on analytic derivatives, and to compare the performance of different forms of stochastic optimization. We emphasize that the approach to model coil errors presented here can be applied directly to other stellarator coil design problems and does not rely on the particular framework of [14] considered in this work. Furthermore, to make the article self-contained and more easily readable, we give a brief but complete summary of the objective function in section 4.

We find that both the stochastic and deterministic formulations result in different designs depending on the initialization of the optimization algorithm, which indicates the existence of multiple local minima. However, the variability of the designs corresponding to the stochastic problem is substantially reduced as compared to the designs obtained from the deterministic optimization problem. The coil systems we obtain for different initial conditions are much more similar to one another when using stochastic optimization than with deterministic optimization. We compare the performance of the obtained configurations in the presence of coil errors by evaluating the level of quasi symmetry near and away from the axis, the rotational transform on axis as well as particle loss fraction. We observe that the configurations found by stochastic optimization outperform those obtained from the deterministic formulation. Furthermore, the different minima obtained from stochastic minimization all perform very similarly, which suggests that the optimization algorithm does not get trapped in poor local minima. For all these reasons, our work demonstrates the strong potential of stochastic optimization for stellarator design, and motivates its application to more detailed reactor design studies.

The structure of the article is as follows. In section 2, we present our mathematical description of the coils, and explain how we model random coil errors. In section 3, we provide a brief summary of stochastic optimization, with a description of several variants which are relevant to stellarator design. We then review in section 4 the direct coil design paradigm that we modify for stochastic optimization, first introduced in [14]. We present our main numerical results in section 5, and summarize our work in section 6, where we also suggest directions for future work.

2. Modeling coil perturbations

2.1. Physical representation of coil perturbations

For this work, we make the common assumption in coil design that we can represent the coils as current-carrying filaments, i.e., we neglect the non-zero thickness of the coils, and simply model coils as curves in space. A coil is then described by a periodic function \( \Gamma : [0, 2\pi) \rightarrow \mathbb{R}^3 \). A standard approach to discretizing such coils is given by a truncated Fourier expansion, that is the \( j \)th coordinate of the \( i \)th coil \( \Gamma^{(i)} \) is given by

\[
\Gamma^{(i)}(\theta) = c_{i,0}^{(i)} + \sum_{j=1}^{n_{\text{coil}}} c_{i,j}^{(i)} \sin(j \theta) + \sum_{j=1}^{n_{\text{coil}}} c_{i,j}^{(i)} \cos(j \theta). \tag{1}
\]

We collect the degrees of freedom for coil \( i \) in the vector \( \epsilon^{(i)} \in \mathbb{R}^{3(2n_{\text{coil}}+1)} \). This approach is also used in the FOCUS code [47]. We note that this formulation allows the coils to move freely in space, as compared to coil optimization codes that restrict the coils to lie on the so-called winding surface. This latter approach is employed in the ONSET [9], COILOPT [40], COILOPT++ [12] codes.

A straightforward way of modeling errors is to perturb the vector containing the degrees of freedom, \( \epsilon^{(i)} = \epsilon^{(i)} + \varepsilon \) where \( \varepsilon \) is a vector of independent random variables. In the context
of coil optimization, this approach has been used in [28–30] which builds on ONSET and uses splines to represent coils. There the spline anchor points were perturbed by independent, centered Gaussian random variables with small variance. This approach does not require any modification of the objective function implementation and hence a deterministic code can be extended toward stochastic optimization with little effort.

Most coil design codes assume rotational and stellarator symmetry of the coils [8]. This assumption significantly reduces the dimensionality of the optimization space and is used in most optimization studies [3, 10, 15, 47]. For example, the National Compact Stellarator Experiment consists of 24 coils that are obtained by applying symmetries to four base coils, which were optimized for. However, coil errors are not guaranteed to satisfy these same symmetries. Using the same parametrization for coil errors as used for the coils imposes this symmetry on the coil errors and is hence inherently limiting.

Additionally, when Fourier modes are used for coils and coil errors, these errors affect coils globally and repeat themselves along the coils, with more repetitions with increasing Fourier mode numbers. Using splines to describe the coil geometry as well as manufacturing errors has the advantage of allowing the description of local manufacturing errors. However, if the number of spline anchor points changes, the characteristics of the manufacturing error changes as well. Thus, when the description of the coils and the errors are the same, changing the coil discretization also means changing the type of errors considered and convergence to a limit when refining the coil description is unclear.

For this reason, we separate the discretization of the coils and the modeling of perturbations by considering additive perturbations modeled by Gaussian processes [35]. To the best of our knowledge, this is a novel approach for stellarator coil optimization applications. However, similar approaches have been used in other areas. For instance, such an approach has been used in the context of airfoil optimization in two dimensions [7, 27, 43].

We consider here the case of errors that are independent for each coil, but we will outline at the end of the section how that approach can be extended to allow for errors that partially satisfy symmetries.

2.2. Randomly perturbed coils via stochastic processes

We model the perturbations of the coil by random, periodic functions \( \Xi : [0, 2\pi] \rightarrow \mathbb{R}^3 \), and denote the perturbed coil by \( \tilde{C}(\theta) = C(\theta) + \Xi(\theta) \). We choose to model the components \((\Xi_1, \Xi_2, \Xi_3)\) of \( \Xi \) as independent centered Gaussian processes.

We briefly recall the definition and some basic properties of Gaussian processes. A random function \( \Xi \) is a centered Gaussian process if for any fixed set \( \{\theta_1, \ldots, \theta_n\} \) the random vector \( (\Xi(\theta_1), \ldots, \Xi(\theta_n)) \) follows a multi-variate normal distribution with mean zero. The function \( \text{Cov}(\Theta, \Theta) = \text{Cov}(\Xi(\theta), \Xi(\theta')) \) is referred to as the covariance function. Sampling a Gaussian process at points \( \{\theta_1, \ldots, \theta_n\} \) is then as straightforward as drawing a Gaussian vector with mean zero and covariance matrix \( \{C(\theta_i, \theta_j)\}_{i,j} \).

In this work, we make the common assumption that the covariance is stationary, i.e., it is only a function of \( \theta - \theta' \). Thus, one can write \( C(\theta, \theta') = K(\theta - \theta') \) for some function \( K \). The regularity of the random functions \( \Xi \) is directly linked to the regularity of \( K \); in this work we consider a classical squared exponential covariance function,

\[
k(d) = \sigma^2 \exp\left(-\frac{d^2}{2\ell^2}\right),
\]

which results in rather smooth perturbations. Here \( \sigma > 0 \) controls the overall magnitude of the perturbations, and \( \ell > 0 \) is a measure for its length scale. For our problem, we cannot use the covariance function \( k \) directly, because we need to guarantee periodicity of \( \Xi \). To address this minor difficulty, we rely on the fact that any covariance function \( k \) can be made periodic on \( [0, 2\pi] \) by defining

\[
\tilde{k}(d) = \sum_{j \in \mathbb{Z}} k(d + j2\pi),
\]
Gaussian and satisfy, for just a few terms, since \( k \) has exponential decay. Finally, since our optimization algorithms require the knowledge of derivatives, an additional property of Gaussian processes is relevant to our work. By linearity (see [34, section 9.4], [1, section 2.2]) the derivatives of \( \Xi \) are also Gaussian and satisfy, for \( \theta, \bar{\theta} \):

\[
\text{Cov}(\Xi'(\theta), \Xi'(\bar{\theta})) = \partial_\theta C(\theta, \bar{\theta}) = \hat{\Sigma}(\theta - \bar{\theta})
\]

\[
\text{Cov}(\Xi(\theta), \Xi'(\bar{\theta})) = \partial_\theta \partial_\theta C(\theta, \bar{\theta}) = -\hat{\bar{\Sigma}}(\theta - \bar{\theta})
\]

where the prime denotes derivative with respect to its argument. Thus, we can draw samples \( \{\Xi(\theta_1), \ldots, \Xi(\theta_n)\} \) by drawing a Gaussian vector with covariance matrix

\[
\Sigma = \begin{bmatrix}
\hat{\Sigma}(\theta_1 - \bar{\theta}_1) & \ldots & \hat{\Sigma}(\theta_1 - \bar{\theta}_n) \\
\ldots & \ldots & \ldots \\
\hat{\Sigma}(\theta_n - \bar{\theta}_1) & \ldots & \hat{\Sigma}(\theta_n - \bar{\theta}_n)
\end{bmatrix}
\]

The standard approach to drawing such samples is to compute a matrix square root \( \Sigma = LL^T \) (e.g. via a Cholesky decomposition) and to draw a standard Gaussian vector \( \zeta \). It is then straightforward to check that \( \text{Cov}(\zeta \bar{\Xi}) = \Sigma \). We show several random function draws from this Gaussian process as well as a perturbed coil in figure 1.

Finally, we remark that while here we only consider the case of independent perturbations for each coil, it is straightforward to include perturbations that repeat across coils. The former are also referred to as statistical errors and the latter as systematic errors [2, 37]. To illustrate this, let us assume that coil \( j \) is simply a repetition of coil \( i \), that is \( \Gamma_j = \Gamma_i \), for some rotation matrix \( \Gamma \). Now consider an error of the form \( \Xi = \Xi_{\text{stat}} + \Xi_{\text{sys}} \), then one may choose \( \Xi_{\text{sys}} = L\Xi_{\text{sys}} \) to enforce symmetry on only part of the error. We note that different Gaussian process parameters could be used to model \( \Xi_{\text{stat}} \) and \( \Xi_{\text{sys}} \) to represent different magnitudes and length scales for statistical and systematic error.

3. Stochastic and risk-averse optimization

In this section, we briefly review the mathematical formulations of the different forms of stochastic optimization one may favor, depending on the coil optimization design goals and the level of risk one is willing to tolerate.

Let \( f(\{\Gamma_i\}_{i=1}^n, q) \) be some quantity of interest which one wants to minimize, and which depends on the coil geometry \( \{\Gamma_i\}_{i=1}^n \), as well as other quantities \( q \in \mathbb{R}^q \) (e.g. coil currents). We define

\[
g(\{\Gamma_i\}_{i=1}^n, q, \{\Xi(\theta)\}_{i=1}^n) = f(\{\Gamma_i + \Xi(\theta)\}_{i=1}^n, q).
\]

For notational brevity we write \( x = (\{\Gamma_i\}_{i=1}^n, q) \) and \( \zeta = \{\Xi(\theta)\}_{i=1}^n \), i.e., the variables to be optimized are contained in \( x \) and the randomness is contained in \( \zeta \). For fixed \( x \), \( g(x, \zeta) \) is now a random variable, which needs to be scalarized in order to perform optimization. Risk-neutral, risk-averse, and robust stochastic optimization formulations are all obtained by different ways of scalarising \( g(x, \zeta) \), and thus all take the distribution of manufacturing errors into account. While other stochastic optimization formulations exist (e.g., [39]), we only describe these three below, since they are the most common formulations, and well suited to stellarator optimization.

Before we do so, we observe that the deterministic optimization problem of minimizing \( f \) is equivalent to minimizing

\[
\min_x g(x, 0),
\]

i.e., it is assumed that no coil error is present. This is the approach traditionally taken in stellarator optimization, sometimes followed by perturbation tests at the optimal design [14, 37, 44, 47] to evaluate the sensitivity of the objective with respect to coil errors. In contrast, the approaches discussed next take the distribution of manufacturing errors into account during the optimization process.

3.1. Types of optimization under uncertainty

3.1.1. Risk-neutral stochastic optimization. Risk-neutral stochastic optimization corresponds to the situation in which one wants to find a solution \( x \) that performs optimally with respect to the mean of the realisations. We thus obtain the optimization problem

\[
\min_x \mathbb{E}[g(x, \zeta)],
\]

where \( \mathbb{E}[-] \) denotes expectation over the distribution of the perturbations \( \zeta \). In this article, we will mainly focus on this approach, and thus sometimes simply refer to it as stochastic optimization. However, for certain stellarator design problems and certain objective functions, it can be desirable to explicitly avoid poor objective values for some realizations of \( \zeta \). This can be achieved using risk-averse or robust formulations, which we summarize next.

3.1.2. Risk-averse stochastic optimization:C VaR. A measure that focuses on the tail of the distribution is the conditional value-at-risk, or CVaR. The CVaR of a random variable \( Z \) is defined as the expected value given that the random variable falls into the \( \alpha \)-quantile of its distribution, i.e.,

\[
\text{CVaR}_\alpha[Z] = \mathbb{E}[Z | Z > \text{CDF}^{-1}_Z(\alpha)],
\]

where \( \text{CDF}_Z \) denotes the cumulative distribution function and \( \alpha \in [0, 1] \). The difference between the risk-neutral formulation and CVaR is illustrated with an example probability density function in figure 2. This example highlights the fact that the CVaR only depends on the tail of the distribution. One reason for the popularity of CVaR over other risk-averse measures is its convexity as well as the following equivalent formulation (cf [36, theorem 1])

\[
\text{CVaR}_\alpha[Z] = \inf_{t \in \mathbb{R}} \left[ t + \frac{1}{1 - \alpha} \mathbb{E}[(Z - t)^+] \right],
\]
where \( s^+ : \max(0, s) \), which is a convenient formulation for computational purposes. To cope with the non-differentiable nature of the max-function, one relies in practice on a smooth approximation \( h : \mathbb{R} \to [0, \infty) \) with \( h_{\varepsilon} \to (\cdot)^+ \) as \( \varepsilon \to 0 \). The specific form of \( h_{\varepsilon} \) used in this work is

\[
h_{\varepsilon}(x) = \begin{cases} x & \text{if } x \geq \varepsilon / 2, \\
(\varepsilon / 2)^3 - \frac{(x + \varepsilon / 2)^3 - x^3}{2\varepsilon^3} & \text{if } -\varepsilon / 2 < x < \varepsilon / 2, \\
0 & \text{otherwise}.
\end{cases}
\]

We thus obtain the optimization problem

\[
\min_x \left[ t + \frac{1}{1 - \alpha} \mathbb{E} \left[ h_{\varepsilon}(g(x, \zeta) - t) \right] \right]
\]

for a small \( \varepsilon \). Risk-averse formulations based on CVaR are successfully used in the insurance and finance industries and in engineering \([18, 20, 36]\) for instance.

### 3.1.3. Robust stochastic optimization

To completely control the probability of poor outcomes, one can optimize the worst possible scenario, i.e.,

\[
\min_x \max_{\zeta} g(x, \zeta).
\]

This formulation is typically combined with a model for randomness that results in perturbations that are almost surely bounded. We will observe in section 5 that the difference between risk-neutral and risk-averse coil designs is minor for the stellarator optimization problem we consider. Thus, we do not explore robust optimization further in this article, as it can be viewed as an extreme version of risk-averse optimization.

### 3.2. Sample average approximation

The expected value in (8) and (12) can typically not be computed analytically but has to be approximated numerically. For that purpose, we use the sample average approximation, i.e., we draw \( N_{\text{MC}} \) independent realisations of \( \zeta_k \) and approximate

\[
\mathbb{E}[g(x, \zeta)] \approx \frac{1}{N_{\text{MC}}} \sum_{k=1}^{N_{\text{MC}}} g(x, \zeta_k),
\]

\[
\mathbb{E}[h_{\varepsilon}(g(x, \zeta) - t)] \approx \frac{1}{N_{\text{MC}}} \sum_{k=1}^{N_{\text{MC}}} h_{\varepsilon}(g(x, \zeta_k) - t),
\]

for the risk-neutral and risk-averse formulation respectively. As \( N_{\text{MC}} \to \infty \), the random space approximation error is of the typical Monte Carlo order \( O(N_{\text{MC}}^{-1/2}) \). Since the samples \( \zeta_k \) are kept fixed throughout the optimization, (14) results in a deterministic optimization problem with \( N_{\text{MC}} \) terms. Note that by linearity the sample average approximation of the gradients is exactly equal to the gradient of the sample average approximation.

Our numerical tests in section 5 focus on a comparison between deterministic, risk-neutral and CVaR risk-averse stochastic designs. Moreover, we study the role of the Monte Carlo sample size \( N_{\text{MC}} \) for approximating the distribution.

### 4. Direct coil design for quasi-symmetry in vacuum fields

In [14] a new formulation was presented to directly design coils generating vacuum magnetic fields which are quasi-symmetric to high accuracy in a region close to the magnetic axis. We briefly recall the basic structure of the objective that was developed there and then show the corresponding stochastic and risk-averse formulations.

Given a so-called expansion axis \( \mathbf{\Gamma}_a \) and real parameter \( \eta \), it was shown in \([24, 25]\) how to construct a magnetic field \( \mathbf{B}_{\text{QS}} \) that is quasi-symmetric near the axis and how to compute its rotational transform \( \mathbf{L} \). Calling \( \mathbf{B}_{\text{col}} \) the magnetic fields induced by the coils \( \{ \mathbf{\Gamma}_c^{(i)} \} \), the approach of [14] is then to find coils so that \( \mathbf{B}_{\text{QS}} \approx \mathbf{B}_{\text{col}} \). Grouping the coefficients that describe the expansion axis and the real parameter \( \eta \) in a vector \( \mathbf{a} \), and grouping the coefficients that describe the coils and their currents in a vector \( \mathbf{c} \), the objective is given by

\[ \min_{\mathbf{a}, \mathbf{c}} \left[ \sum_i \| \mathbf{a} \|^2 + \sum_{i,j} \| \mathbf{c}_i \| \right] \]
\[ J(c, a) = \frac{1}{2} \int_{\Gamma_a} ||B_{\text{coils}}(c) - B_{\text{QS}}(a)||^2 \, dl \\
+ \frac{1}{2} \int_{\Gamma_a} ||\nabla B_{\text{coils}}(c) - \nabla B_{\text{QS}}(a)||^2 \, dl \\
+ \frac{1}{2} \left( \frac{(a - l_{0,a})^2}{l_{0,a}^2} \right) + R_q(a) + R_c(c), \]

where \( l_{0,a} \) is a target rotational transform, and \( R_q \) and \( R_c \) contain various regularizations for the expansion axis and coils respectively. The regularization terms include penalty functions for the length of the axis and the length of the coils, the curvature of the coils, and the distance between coils. In [14] it was shown that this formulation leads to an efficient method to design from scratch coils producing nearly quasi-symmetric vacuum magnetic configurations, and to improve the quasi-symmetry properties of existing designs.

Random perturbations of the coils can be taken into account by considering the objective

\[ J'(c, a, \zeta) = \frac{1}{2} \int_{\Gamma_a} ||B_{\text{coils}}(c, \zeta) - B_{\text{QS}}(a)||^2 \, dl \\
+ \frac{1}{2} \int_{\Gamma_a} ||\nabla B_{\text{coils}}(c, \zeta) - \nabla B_{\text{QS}}(a)||^2 \, dl \\
+ \frac{1}{2} \left( \frac{(a - l_{0,a})^2}{l_{0,a}^2} \right) + R_q(a) + R_c(c), \]

where \( B_{\text{coils}}(c, \zeta) \) corresponds to the magnetic field produced by the perturbed coils \( \{\Gamma^{(i)} + \Xi^{(i)}\} \). We emphasize here that the field \( B_{\text{QS}} \) and the rotational transform are independent from the random variable \( \zeta \), and hence we only have to recompute \( B_{\text{coils}}(c, \zeta) \) and \( \nabla B_{\text{coils}}(c, \zeta) \) for different samples \( \zeta \). We observe that as it is stated here, the optimization problem can lead to numerical difficulties, because vanishing derivatives of \( \Gamma^{(i)} \) result in a non differentiable curve length objective, and because of the large nullspace of the objective, which is partially due to the fact that different parametrizations give the same physical curve. In order to address these numerical difficulties, we add the following regularization term, in addition to the axis length, coil length, coil curvature, and coil distance terms already included in [14],

\[ R_{\text{arc}}(c) = \sum_{i=1}^{n_a} \int_{(0,2\pi)} (||\Gamma^{(i)}(\theta)|| - l^{(i)})^2 \, d\theta, \]

where \( l^{(i)} = \frac{\pi}{n_a} \) is the value that would correspond to a constant-arclength parametrization of a circle with coil length \( l^{(i)} \).

5. Numerical results for NCSX-like example

5.1. Implementation and setup

We implement this optimization in the open source PyPlasmaOpt package available under https://github.com/florianwechsung/PyPlasmaOpt. PyPlasmaOpt is a Python library that relies on the geometric objects and the Biot Savart implementation of the SIMSOPT stellarator optimization package http://github.com/hiddenSymmetries/SIMSOPT. The implementation is parallelized across samples using MPI, and the Biot Savart computation is accelerated using SIMD instructions as well as OpenMP. A function and gradient evaluation of a typical configuration (18 coils, 120 quadrature points per coil) with 1024 samples takes less than half of a second on a machine with two 24 Core Intel Xeon Platinum 8268 processors.

To solve the optimization problems in (7), (8), and (12), we use the L-BFGS implementation in SciPy [42]. For the smoothed risk-averse objective, we use the risk-neutral minimizer as initial guess, then solve the optimization problem, then reduce the smoothing parameter \( \epsilon \), and then solve the problem again, using the previous solution as initial guess. This is repeated until \( \epsilon = 10^{-5} \).

We study a configuration that is inspired by the National Compact Stellarator Experiment (NCSX). NCSX consists of three distinct modular coils, which results in 18 coils after applying three fold rotational symmetry and stellarator
We study this problem in particular as the NCSX project was canceled due to increasing costs because of, among other reasons, the requirement of tight engineering tolerances on the coils. To model the distribution of manufacturing errors, we choose a length scale of $l = 0.4\pi$ and a standard deviations of either $\sigma = 10^{-2}$ or $\sigma = 3 \times 10^{-3}$ in the kernel (2) defining the Gaussian process. For each of the coils, these values correspond to manufacturing errors of a few centimeters for $\sigma = 10^{-2}$ and several millimeters for $\sigma = 3 \times 10^{-3}$.

When running either the deterministic or the stochastic version of the optimization algorithm with different initial guesses, we observe that multiple local minima exist. For this reason, in the following sections we usually show results for several different minima, that were obtained by starting the optimizer from eight different initial guesses. These initial guesses were obtained by randomly perturbing the Fourier coefficients describing the coils using independent normal random variables with standard deviation of 0.01. One of the obtained configurations together with a range of magnetic surfaces is shown in figure 3. Next, in sections 5.2 and 5.3 we mainly focus on the role the number of samples $N_{\text{MC}}$ in the sample average approximation (14) plays in approximating the distribution of coil perturbations, and on the sensitivity of both the deterministic and the stochastic minimizers on the initial guesses for the optimization algorithm.

5.2. Deterministic versus stochastic designs

Figure 4 shows minimizers that were obtained by solving the deterministic and the stochastic optimization problem for eight different initial guesses. We clearly see that for the...
5.3. Out of sample distribution at the minimizer

As we only optimized the mean of the objective for a finite number of perturbations, we have to check that this performance generalises to the full distribution of coil errors. To do this we draw a large number ($2^{18}$) of new samples, which are different from those used in the sample average approximation, and evaluate the objective at the minimizer for each of the samples. In the context of statistics and machine learning, this procedure is known as out-of-sample testing or cross validation. Figure 5 shows the resulting distribution for minimizers obtained from deterministic and stochastic optimization from eight different initial guesses. We see that the performance of the minimizers of the deterministic optimization problem varies strongly on perturbed coils. The minimizers found by stochastic optimization have lower objective value on average and for most perturbations. In other words, the stochastic designs perform significantly better than the deterministic designs. Additionally, different minimizers obtained as a result of different initializations of the algorithm perform vastly differently for the deterministic formulation, but very similarly for the stochastic formulations, in particular when $N_{MC} = 1024$ samples are used to estimate the expected value in (14) for the stochastic formulation.

5.4. Quasi-symmetry close to the axis

The objective is designed to ensure quasi-symmetry near the axis. To confirm that this is achieved and to investigate the magnetic field away from the axis, we compute magnetic flux surfaces $\mathbf{S}(\varphi, \theta)$ parametrized by Boozer angles $\varphi$ and $\theta$. We recall that a magnetic field is called quasi-axisymmetric if $|\mathbf{B}(\mathbf{S}(\varphi, \theta))|$ is a function of $\theta$ only. Figure 6 shows three surfaces computed for the best configuration obtained from stochastic optimization with $N_{MC} = 1024$ samples. We can see that for the surface closest to the axis, the field strength is indeed close to constant in $\varphi$. As we move away from the axis, this property is gradually lost.
Figure 7. Mean non-quasi-symmetry on a range of surfaces for eight different minimizers obtained from deterministic and stochastic ($N_{MC} = 1024$ samples) optimization. Here we use the surface area as a label for the surfaces. Larger surface areas correspond to surfaces farther from the axis.

Figure 8. Distribution of the rotational transform $\iota$ on axis for each of the eight minimizers found from deterministic and stochastic optimization. The distribution is approximated using 128 independently drawn perturbed coil sets. The dashed line indicated the target rotational transform.
To quantify this statement, we decompose $|B|$ into a quasi-symmetric and non-quasi-symmetric part by defining

$$
\|B\|_{QS}(\theta) = \int_{0,2\pi} \|B(S(\varphi, \theta))\| \|\partial_\varphi S \times \partial_\theta S\| d\varphi
$$

and

$$
\|B\|_{NonQS}(\varphi, \theta) = \|B(S(\varphi, \theta))\| - \|B\|_{QS}(\theta).
$$

We then measure the norm of the non-quasi-symmetric part and report $\left[ \frac{\int_{S} \|B\|_{NonQS}^2 dS}{\int_{S} \|B\|_{QS}^2 dS} \right]^{1/2}$ in figure 7. Since we only enforce quasi-symmetry near the axis, we expect this measure to be small for surfaces close to the axis, and to increase as we move away from the axis.

We perform the stochastic optimization for $\sigma_{opt} = 0.01$ and $\sigma_{opt} = 0.003$. Quasi-symmetry is then evaluated by drawing 20 new samples with standard deviation $\sigma_s$ and computing surfaces for the fields induced by the perturbed coils. The case of $\sigma_s < \sigma_{opt}$ can be viewed as the estimate for the perturbation size in the original optimization being pessimistic, or be due to improvements in the manufacturing process between design and construction of the coils. For comparison, we also compute surface and non-quasi-symmetry for the minimizers obtained from deterministic optimization (corresponding to $\sigma_{opt} = 0$). The results are displayed in figure 7.

Overall we observe that the configurations have very little non-quasi-symmetric contribution close to the axis, and that the non-quasi-symmetry then increases away from axis. As can be expected, the difference between stochastic and deterministic optimization is most significant for large coil perturbations (left plot, solid lines). Keeping those same configurations, but reducing the perturbation magnitude in the newly drawn samples, we can see that the configurations obtained from both types of optimization benefit from the added accuracy (left plot, dashed lines). Importantly, the designs from stochastic optimization systematically perform better than those from deterministic optimization.

For smaller perturbations (right) the behavior remains qualitatively the same, but the overall difference between stochastic and deterministic optimization is smaller. This is expected, as in the limit of $\sigma_{opt} \to 0$ the two optimization strategies become identical.

5.5. Rotational transform on axis

The objective includes a penalty that targets a certain rotational transform on the expansion axis. To investigate the impact of coil perturbation errors on rotational transform, we draw 128 sets of perturbed coils, compute the magnetic axis for the resulting magnetic fields, and then compute the rotational transform $\iota$ on axis. The resulting distribution of $\iota$ is shown in figure 8. As expected, for the perturbed coils the target rotational transform is not achieved exactly. In agreement with our previous results, we also observe that the different minimizers obtained from deterministic optimization vary more than those obtained from stochastic optimization.
Figure 10. Distribution of objective values for designs computed using risk-neutral stochastic formulation (blue) and CVaR risk-averse stochastic formulation (pink). \( N_{MC} = 1024 \) samples are used in the stochastic optimization, and the distributions are computed using 262 144 independent samples. Slightly different designs are obtained using eight initializations for the optimization.

5.6. Particle confinement

As a final measure of performance, we compute particle trajectories for perturbations of the configurations obtained from deterministic and stochastic optimization. We consider both 250 eV protons and 1 keV protons. For reference, we note that a 3.4 keV proton in our designs has approximately the same ratio of gyroradius to machine size as an energetic alpha particle in the ARIES-CS reactor. We draw 10 perturbed coil configurations for each minimizer obtained from deterministic and stochastic optimization, as well as for the initial NCSX-like configuration. We then spawn 1120 protons with random pitch angle on random points along the magnetic axis and follow them for 10 ms, using the guiding center approximation [4, 6, 26] without collisions. Particles are considered lost if they move more than 30 cm away from the axis. Figure 9 shows the average fraction of lost particles over time.

For both the lower energy protons and the higher energy protons considered here, we observe that the configurations from stochastic optimization have better confinement than the initial configuration and the configurations obtained from deterministic optimization. In addition, we see again that the different minimizers in the stochastic case all perform very similarly, whereas those from deterministic optimization have highly varying performance. For the protons with lower energy, nearly all optimized configurations outperform the initial configuration. However, for the protons with higher energy, this advantage is less clear. This suggests that quasi-symmetry only close to the axis is in general insufficient to guarantee good particle confinement, and motivates ongoing work on direct coil optimization enforcing quasi-symmetry away from the axis.

5.7. Risk-neutral vs risk-averse optimization

Finally, we compare the risk-neutral formulation (i.e. minimization of the expected value) with a risk-averse objective for an error distribution with \( \sigma = 10^{-2} \). We choose \( \alpha = 0.95 \) and minimize \( \text{CVaR}_\alpha(f) \), meaning that we minimize the expected value of the tail containing the 5% worst scenarios. In figure 10 we show the out-of-sample distribution of the objective evaluated at the minimizers. We can see that most of the distribution for the risk-neutral objective is closer to zero, while the tail is slightly thicker. However, the difference is insignificant. We attribute this to the quadratic penalty form of our objective: since all objective values are positive and the objective is squared, in order to control the mean large positive outliers have to be avoided.

6. Conclusion and future work

We have extended the direct stellarator design approach of [14] to include random coil errors. We emphasize that our formulation uses separate discretizations for the coils and their errors, which allows us to retain symmetries in the design space but considers perturbations that do not satisfy them.

We then studied and compared deterministic, risk-neutral, and risk-averse optimization for an NCSX-like example. We found that the deterministic problem admits a large number of distinct minimizers which perform quite differently. Including stochasticity reduces the number of different minimizers and results in minimizers that all perform nearly identical in terms of both objective value, as well as quasi-symmetry and rotational transform on and away from axis. Moving from a risk-neutral to a risk-averse formulation does not result in significantly different minimizing designs in our experiments: the tail of the distribution is reduced at the cost of somewhat worse average performance. Finally, while we are able to achieve quasi-symmetry near the axis, this property is lost away from axis. Current work is focused on including the non-quasi-symmetry measure presented in figure 7 on a range of surfaces to enforce quasi-symmetry away from axis.
Acknowledgments

This work was supported by a grant from the Simons Foundation (560651). A.G. is partially supported by a NSERC (Natural Sciences and Engineering Research Council of Canada) postdoctoral fellowship. In addition, this work was supported in part through the NYU IT High Performance Computing resources, and staff expertise.

Data availability statement

The code used to generate the numerical results is openly available under https://github.com/FlorianWechsung/PyPlasmaOpt/tree/tf/tree/paper-stochastic/examples/stochastic.

ORCID iDs

Florian Wechsung https://orcid.org/0000-0003-2195-6522
Andrew Giuliani https://orcid.org/0000-0002-4388-2782
Matt Landreman https://orcid.org/0000-0002-7233-577X
Antoine Cerfon https://orcid.org/0000-0003-3326-4785
Georg Stadler https://orcid.org/0000-0001-7762-6544

References

[1] Adler R.J. 2010 The Geometry of Random Fields (Philadelphia, PA: SIAM)
[2] Andreeva T. et al 2014 Final assessment of Wendelstein 7-X magnetic field perturbations caused by construction asymmetries 25th IAEA Fusion Energy Conf. (St Petersburg, 13–18 October 2014) p 7 (https://www-pub.iaea.org/iaeameetings/46091/25th-fusion-energy-conference-feb-2014)
[3] Bader A., Drevlak M., Anderson D.T., Faber B.J., Hegna C.C., Likin K.M., Schmitt J.C. and Talmadge J.N. 2019 Stellarator equilibria with reactor relevant energetic particle losses J. Plasma Phys. 85 905880508
[4] Boozer A.H. 1980 Guiding center drift equations Phys. Fluids 23 904–8
[5] Carlton-Jones A., Paul E.J. and Dorland W. 2021 Computing the shape gradient of stellarator coil complexity with respect to the plasma boundary J. Plasma Phys. 87 905870222
[6] Cary J.R. and Brizard A.J. 2009 Hamiltonian theory of guiding-center motion Rev. Mod. Phys. 81 693–738
[7] Chen S. and Chen W. 2011 A new level-set based approach to shape and topology optimization under geometric uncertainty Struct. Multidiscip. Optim. 44 1–18
[8] Dewar R.L. and Hudson S.R. 1998 Stellarator symmetry Physica D 112 275–80
[9] Drevlak M. 1999 Optimization of heterogenous magnet systems Proc. 12th Int. Stellarator Workshop (Madison, USA)
[10] Drevlak M., Brochard F., Helander P., Kisslinger J., Mikhailov M., Nührenberg C., Nührenberg J. and Turkin Y. 2013 ESTELL: a quasi-toroidally symmetric stellarator Contrib. Plasma Phys. 53 459–68
[11] Gates D.A. et al 2018 Stellarator research opportunities: a report of the national stellarator coordinating committee J. Fusion Energy 37 51–94
[12] Gates D.A. et al 2017 Recent advances in stellarator optimization Nucl. Fusion 57 126064
[13] Geraldini A., Landreman M. and Paul E. 2021 An adjoint method for determining the sensitivity of island size to magnetic field variations J. Plasma Phys. 87 905870302
[14] Giuliani A. et al 2020 Single-stage gradient-based stellarator coil design: optimization for near-axis quasi-symmetry (arXiv:2010.02033)
[15] Henneberg S.A., Drevlak M., Nührenberg C., Beidler C.D., Turkin Y., Loizu J. and Helander P. 2019 Properties of a new quasi-axisymmetric configuration Nucl. Fusion 59 026014
[16] Hudson S.R., Zhu C., Pfefferle D. and Gunderson L. 2018 Differentiating the shape of stellarator coils with respect to the plasma boundary Phys. Lett. A 382 2732–7
[17] Klinger T. et al 2013 Towards assembly completion and preparation of experimental campaigns of Wendelstein 7-X in the perspective of a path to a stellarator fusion power plant Fusion Eng. Des. 88 461–5
[18] Kouri D.P. and Surowiec T.M. 2016 Risk-averse PDE-constrained optimization using the conditional value-at-risk SIAM J. Optim. 26 365–96
[19] Kremer J.P. et al 2003 The status of the design and construction of the Columbia non-neutral torus AIP Conf. Proc. 692 320–5
[20] Krokhmal P., Palmaquist J. and Uryasev S. 2002 Portfolio optimization with conditional value-at-risk objective and constraints J. Risk 4 43–68
[21] Landreman M. 2017 An improved current potential method for fast computation of stellarator coil shapes Nucl. Fusion 57 046003
[22] Landreman M. and Boozer A.H. 2016 Efficient magnetic fields for supporting toroidal plasmas Phys. Plasmas 23 032506
[23] Landreman M. and Paul E. 2018 Computing local sensitivity and tolerances for stellarator physics properties using shape gradients Nucl. Fusion 58 076023
[24] Landreman M. and Sengupta W. 2018 Direct construction of optimized stellarator shapes: I. Theory in cylindrical coordinates J. Plasma Phys. 84 905840616
[25] Landreman M., Sengupta W. and Plunk G.G. 2019 Direct construction of optimized stellarator shapes: II. Numerical quasisymmetric solutions J. Plasma Phys. 85 905885013
[26] Littlejohn R.G. 1983 Variational principles of guiding centre motion J. Plasma Phys. 29 111–25
[27] Liu D., Litvinenko A., Schillings C. and Schulz V. 2017 Quantification of airfoil geometry-induced aerodynamic uncertainties—comparison of approaches SIAM/ASA J. Uncertain. Quantification 5 334–52
[28] Lobkens J.-F. et al 2020 Improved performance of stellarator coil design optimization J. Plasma Phys. 86 815860202
[29] Lobkens J.-F. et al 2020 Physics analysis of results of stochastic and classic stellarator coil optimization Nucl. Fusion 60 046012
[30] Lobkens J.-F. et al 2018 Stellarator coil optimization towards higher engineering tolerances Nucl. Fusion 58 106013
[31] Neilson G.H., Gruber C.O., Harris J.H., Rej D.J., Simmons R.T. and Strykowski R.L. 2010 Lessons learned in risk management on NCSX IEEE Trans. Plasma Sci. 38 320–7
[32] Paul E.J., Landreman M., Bader A. and Dorland W. 2018 An adjoint method for gradient-based optimization of stellarator coil shapes Nucl. Fusion 58 076015
[33] Sunn Pedersen T. et al 2006 Experimental demonstration of a compact stellarator magnetic trap using four circular coils Phys. Plasmas 13 012302
[34] Rasmussen C.E. and Williams C.K.I. 2006 Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning) (Cambridge, MA: MIT Press)
[35] Rasmussen C.E. and Williams C.K.I. 2006 Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning) (Cambridge, MA: MIT Press) p 248
[36] Rockafellar R.T. and Uryasev S. 2000 Optimization of conditional value-at-risk J. Risk 2 21–41
[37] Rummler T. et al 2004 Accuracy of the construction of the superconducting coils for WENDELSTEIN 7-X IEEE Trans. Appl. Supercond. 14 1394–8
[38] Schölkopf B. and Smola A.J. 2002 Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning) (Cambridge, MA: MIT Press)

[39] Shapiro A., Dentcheva D. and Ruszczynski A. 2009 Lectures on Stochastic Programming: Modeling and Theory (Philadelphia, PA: SIAM)

[40] Strickler D.J., Berry L.A. and Hirshman S.P. 2002 Designing coils for compact stellarators Fusion Sci. Technol. 41 107–15

[41] Strykowsky R.L. et al 2009 Engineering cost schedule lessons learned on NCSX 23rd IEEE/NPSS Symp. on Fusion Engineering (San Diego, 1–5 June 2009) pp 1–4 (https://ieeexplore.ieee.org/xpl/conhome/5205182/proceeding)

[42] Virtanen P. et al 2020 SciPy 1.0: fundamental algorithms for scientific computing in Python Nat. Methods 17 261–72

[43] Wang Q. et al 2011 Conditional sampling and experiment design for quantifying manufacturing error of transonic airfoil 49th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition (Orlando, Florida, 04–07 January 2004) (American Institute of Aeronautics and Astronautics) (https://arc.aiaa.org/doi/book/10.2514/MASM1P)

[44] Zhu C., Hudson S.R., Song Y. and Wan Y. 2018 Designing stellarator coils by a modified Newton method using FOCUS Plasma Phys. Control. Fusion 60 065008

[45] Zhu C., Hudson S.R., Lazerson S.A., Song Y. and Wan Y. 2018 Hessian matrix approach for determining error field sensitivity to coil deviations Plasma Phys. Control. Fusion 60 054016

[46] Zhu C., Gates D.A., Hudson S.R., Liu H., Xu Y., Shimizu A. and Okamura S. 2019 Identification of important error fields in stellarators using the Hessian matrix method Nucl. Fusion 59 126007

[47] Zhu C., Hudson S.R., Song Y. and Wan Y. 2018 New method to design stellarator coils without the winding surface Nucl. Fusion 58 016008