Bayesian methods for magnetic and mechanical optimization of superconducting magnets for fusion

Sam Packman[1,2*], Nicolò Riva[2,3] and Pablo Rodriguez-Fernandez[2]

[1*]Mathematics Department, Massachusetts Institute of Technology, Massachusetts Ave., Cambridge, 02139, MA, USA.
[2]Plasma Science and Fusion Center, Massachusetts Institute of Technology, Massachusetts Ave., Cambridge, 02139, MA, USA.
[3]Proxima Fusion, Ainmillerstrasse 27A, Munich, 80801, Germany.

*Corresponding author(s). E-mail(s): sampack@mit.edu;
Contributing authors: nicoriva@mit.edu; pablorf@mit.edu;

Abstract

Stellarators as compact fusion power sources have incredible potential to help combat climate change. However, the task of making that a reality faces many challenges. This work uses Bayesian optimization, (BO) which is a method that is well suited to black-box optimizations, to address the complicated optimization problem inherent by stellarator design. In particular it focuses on the mechanical optimization necessary to withstand the Lorentz forces generated by the magnetic coils. This work leverages surrogate models that are constructed to integrate as much information as possible from the available data points, significantly reducing the number of required model evaluations. It showcases the efficacy of Bayesian optimization as a versatile tool for enhancing both magneto-static and mechanical properties within stellarator winding packs. Employing a suite of Bayesian optimization algorithms, we iteratively refine 2D and 3D models of solenoid and stellarator configurations, and demonstrate a 15% increase in optimization speed using multi-fidelity Bayesian optimization. For fusion technology to progresses from experimental stages to commercial viability, precise and efficient design methodologies will be essential. By emphasizing its modularity and transferability, our approach lays the foundation for streamlining optimization processes, facilitating the integration of fusion power into a sustainable energy infrastructure.

Keywords: Bayesian Optimization, High-Temperature Superconductor, Optimization, Solenoid, Stellarator, Critical Current

1 Introduction

Fusion technology has the potential to provide clean and cheap electric power, and will be an important tool in fighting climate change. Stellarators are a type of magnetic confinement fusion reactor that provides an alternative to axisymmetric tokamak designs. They use a non-axisymmetric coil configuration to produce a twisted magnetic field that confines the plasma. This provides advantages such as disruption-free and steady-state operation, [13][48] making stellarators attractive for power generation. Moreover, steady state operation reduces control systems issues and mitigates mechanical problems due to load cycling [28]. Because stellarators are not axisymmetric, there is a massive design space of magnetic field and winding pack configurations. As the finite element calculations required to evaluate possible designs are computationally expensive, an efficient optimization algorithm that can find global optimums is essential. Bayesian optimization, which is a surrogate model-based optimization method that uses uncertainty estimates to balance exploration and local optimization, is perfectly suited for this task.

Bayesian Optimization is formulated to minimize the number of evaluations of expensive black-box functions that are required to optimize some given metric of interest. It operates by fitting a surrogate model to a the evaluated data points, then optimizing an acquisition function to choose the next points to evaluate [9]. In this work, Gaussian processes were used as the surrogate model. Gaussian processes model a probability distribution over a space of functions by treating the value of each point in the domain as a Gaussian random variable and modeling the covariance between each point with a predetermined covariance kernel. By using a Gaussian process to model the data, the acquisition function can account for the uncertainty in the model, allowing an efficient balance of exploration and fine-scale optimization.

This work focuses on two specific optimization problems: magneto-static optimization of solenoid designs and mechanical optimization of stellarator winding packs. Solenoids were chosen as a starting point because they are better understood than stellarator designs, and so the optimal designs could be referenced against known methods and equations [64]. Additionally, this work is applicable to the optimization of central solenoids for spherical tokamak designs, where the central solenoid shape needs to be optimized due to space constraint [63]. In addition, solenoid design shares key similarities with stellarators. They require similar magnetostatic and mechanical simulations that generally use either finite element or Green's function based methods [64]. Additionally, the objective functions for both solenoids and stellarators are concerned with magnetic field properties and mechanical constraints. For this work the two objectives for the solenoid model were the deviation from a target (uniform) magnetic field, and the critical current of the solenoid. For the stellarator model the volume of steel was chosen as the objective to be minimized, and it was constrained by a maximum stress on the superconducting cables past which the critical current would degrade.

Many solenoids in fusion and other fields employ superconducting cables [33][71], providing a good comparison for superconducting cables in stellarator design. However, there are key differences which need to be accounted for. Stellarator design requires simulation of many different types of physics not present in solenoids, including magneto-hydrodynamics and radiation defect simulations [72][26]. Many of these physics require simulations at a very high resolution, and they are difficult to couple. Additionally, stellarator designs may require tighter bending radii than solenoids, which have a uniform bending radius and so will generally have a larger minimum bending radius (for comparably sized coils). As the critical current of superconductors can be decreased when they are bent into tight winding radii, this is an important constraint to account for [52]. These factors are not fully addressed in this work, and it is the topic of future research.

As well as the physical characteristics, the designing processes of solenoids and stellarators shares key challenges. The parameters for both are primarily geometric, making an efficient parametrization of the design space essential. Simulating both requires finite element solvers which are computationally intensive. Additionally, it is nontrivial to calculate gradient data from finite element solvers, limiting many effective gradient-based optimization methods. This research focuses on relatively small models, that require seconds to minutes to evaluate. Even with these evaluation times, optimization algorithms requiring more than several hundred model evaluations quickly become slow to compute. As two goals of this research are to provide a workflow that can be effectively iterated in conjunction with other modeling and optimization tools, and to scale to much larger models, it is crucial to maximize the information that the optimization algorithm gains from each model evaluation. The diagram shown in Figure 1 is a simplified and one-way coupling version of the full workflow, and outlines the different physics that must be computed for stellarator design. As there are many physics constraints, the workflow requires many back and forth steps, as solutions to some steps may not be compatible with further steps. BO allows the optimization to avoid starting over, because it can integrate the whole dataset into its Gaussian process, efficiently leveraging each simulation.

This work demonstrates that Bayesian Optimization is an effective algorithm for fast optimization of finite element models, allowing optimization in fewer iterations. While Bayesian optimization and related surrogate modeling methods have been successfully adopted to model tasks such as neural network hyper-parameter tuning, systems optimizations, and A/B testing [8], and even fusion plasma modeling [54][53] it has been adopted more slowly in stellarator design. In particular, multi-objective and multi-fidelity algorithms, which are recent additions to the BO literature, [67][21][32][16][17] have not been fully adopted for the design and optimization of the stellarators themselves. Previous research has separately addressed both multi-objective and multi-fidelity optimization [62][12] in fusion research, but research leveraging multi-fidelity multi-objective BO algorithms [32] in fusion is limited. This research addresses that gap by exploring several ways that multi-fidelity multi-objective Bayesian optimization can be incorporated into workflows for stellarator coil design.

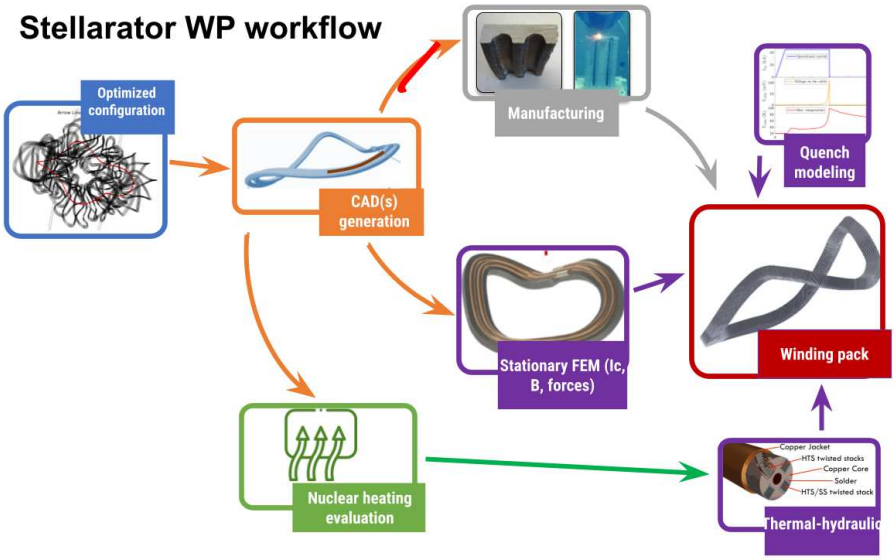


Figure 1: Simplified and one-way coupling version of the workflow envisioned for stellarator winding pack design

We apply Bayesian optimization to the solenoid and a stellarator models. Solenoids were investigated as a proof-of-concept, as they are better understood than solenoids, and the optimization objective was magnetic field regularity. Stellarator optimization was conducted to minimize the volume of the coil casing

2 Background

Magnetic coil design methods are based around several techniques, mostly variations of finite element methods and inverse methods [64]. They often leverage spatial symmetries to simplify calculations [64]. When possible, analytic and semi-analytic methods are used to calculate magnetic fields and forces [64]. Unfortunately, stellarator designs are not amenable to these methods, owing to their complicated geometries.

Current design methods for stellarators use tools such as the VMEC code for physics calculations, which is a variational method [29]. Several different codes are used for optimization, including the SIMSOPT [34], STELLOPT [35], and ROSE [19] codes. The design of stellarators is generally split into two phases: the magnetic field geometry design, and then the corresponding coil design [47][34]. However, there are are other methods that seek to incorporate both steps into a single process [58][24]. These algorithms generally focus on the magnetic field optimization, and don't address the later stages of design. These later stages include winding pack design, cooling system, and vacuum chamber design [51]. As these all have different functions: the winding pack must reinforce the coils, the cooling system must regulate temperature at cryogenic levels, and the vacuum chamber must be airtight and resistant to the plasma, they are generally modeled with separate codes [51]. However, an integrated

design workflow is essential because the system is space-constrained and highly interconnected. The focus of this work is the winding pack design. Winding packs must be engineered to compensate for the extremely high Lorentz forces present in the coil [50]. These Lorentz force can reach hundreds of megapascals, and can reduce the performance of the superconductors and deform the whole winding pack [50]. Preventing this requires structural reinforcement in the form of structural components (namely metal coil casings) [50][55]. However, due to the size constraints, these components must be made as small as possible, necessitating optimization.

The optimization step in stellarator codes use a variety of methods, such as genetic algorithms [41], metaheuristic algorithms [14], quasi-Newton methods [24], and the Levenberg-Marquardt algorithm [18][47]. One difficulty is that many of the most efficient optimization algorithms require gradient information, which requires closely integrating finite element computations with the optimization procedure [20]. There are few examples of Bayesian Optimization in the literature for magnetic coil design, despite it being popular in other areas of physics and engineering. There is work using other forms of surrogate modeling, such as using neural networks to approximate the finite element model [69]. However, these methods have challenges modeling uncertainty and noise in the way Gaussian processes can, and furthermore require far more datapoints to train. For instance, in Ultra-fast Surrogate Model for Magnetic Field Computation of a Superconducting Magnet Using Multi-layer Artificial Neural Networks the lowest number of data points used for training neural networks was 750 [69]. Bayesian optimization, on the other hand, can produce valuable results with as few as 20 data points, as is demonstrated in this research. Given this gap, as well as the unanswered questions about mechanical design optimization, this work shows that BO can address a key knowledge gap in stellarator design.

Bayesian optimization originated in the 1970s with the work of Jonas Mockus, [42] but it has recently become heavily used with applications to neural network hyperparameter tuning [56][27] and systems optimization [8]. While the original method is a fairly simple optimization loop, with a single Gaussian process as the surrogate model and an acquisition function based on the Gaussian process, recent work has sought to apply Bayesian optimization to larger models and integrate secondary sources of information. This includes high output models [39], multi-fidelity optimization [37][66] and multi-objective optimization [60]. These are all potentially valuable for stellarator design, and this work focuses on the use of multi-fidelity optimization. Multi-fidelity optimization uses low fidelity models that approximate the objective model to provide information for the optimization process. To do this, it uses a Gaussian process to jointly model the dependence of the objective function on both the parameters and the fidelity. It then uses a knowledge based acquisition function to decide which parameters and fidelities to query [66][67].

One of the constraints of some multi-fidelity optimization methods is that they assume that the fidelities are unbiased. However, this is not always the case, which became apparent in this work with solenoid modeling. Methods that model the different fidelities with a single Gaussian process [67][11] implicitly assume that the different fidelities are unbiased. This assumption comes from the fact that at unknown fidelities, the only information a Gaussian process has is the objective function values from

the nearest known fidelities. In the absence of a priori information to set the mean function, the Gaussian process mean will be approximately the low fidelity value. It is possible to implement different mean functions to correct this, which is done in [45]. Poloczek proposes modeling the bias of each fidelity with its own Gaussian process, which is then added as a correction term to the original Gaussian process. As this samples only the difference however, it assumes that the correlation coefficient of the different fidelities is 1. This was found to be inaccurate in this work, and we propose a modification to the statistical model that uses a full linear regression on the different fidelities instead of modeling the difference. We applied this to a simplified version of the warm-start optimization method [46] that does not require the knowledge based acquisition functions used in other multi-fidelity work. Warm start Bayesian optimization is a method where the model is pre-trained on a lower fidelity data set, and then the initial set of parameters is chosen by an acquisition function on that model [46].

3 Computational Methods

The workflow separates the modeling and the optimization steps. The finite element solver used in this work was COMSOL Multiphysics version 6.1.

3.1 Model Geometries

3.1.1 Solenoid Model Geometry

A solenoid magnet, as shown in Figure 2a, was used as a proof-of-concept for the benefits of Bayesian Optimization. This work studied both an axially symmetric and a 3D helical geometry for the solenoid models. The cross-sections of the coil were rectangular 1 cm by 2 cm, the radius was initially 10 cm, and the spacing between the layers was initially 6 cm. The solenoid pancake models had 6 layers. The control points were the radial and vertical positions of each coil, and the rotation of the coil cross section in the $\theta = 0$ plane. The top 3 layers were reflected across the $z = 0$ plane, so in total there were 9 parameters: $r_1, r_2, r_3, z_1, z_2, z_3, \theta_1, \theta_2, \theta_3$.

The helical model, as shown in Figure 2b, was based off of an interpolation of the pancake model. The goal of the interpolation was for a cross section of the coil in the $r > 0$, $\theta = 0$ half plane to be identical to the cross section of the pancake model, with an interpolation to connect the rectangle-cross sections. Therefore it had the same 9 parameters as the pancake model. Smooth spline interpolation functions were used to separately interpolate the r, z, and angle parameter ϕ of the coil. The helix was defined in Cartesian coordinates by the parametric curve

$$\mathbf{x} = \langle r(t)\cos(t), r(t)\sin(t), z(t)\rangle \tag{1}$$

where $r(t)$ and $z(t)$ were the interpolation functions, with domains scaled so that a difference of 2π in the argument corresponds to the distance between two interpolation points. Then

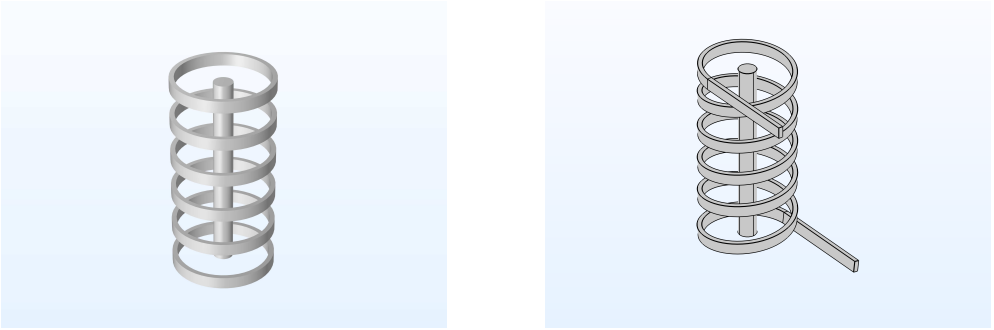
$$r(\pi) = r_1,\ r(3\pi) = r_2,\ r(5\pi) = r_3 \tag{2}$$

and the parameters were similarly reflected over the r-axis so that

$$r(-\pi) = r_1,\ r(-3\pi) = r_2,\ r(-5\pi) = r_3 \tag{3}$$

$z(t)$ was defined similarly but reflected around 0 as an odd function rather than an even function.

The twisting of the coil to achieve the correct angles was done by sweeping the cross section along the parametric curve, with a twist of $\phi(t)$ at each point along the curve. The $\phi(t)$ values were interpolated in the same way as the $z(t)$ values, to make $\phi(t)$ an odd function. Note that twist compensation was used to avoid introducing extra twists as artifacts of the sweep construction. Twist compensation rotates the swept face by an amount equal to the integral of the torsion along the curve to compensate for the rotation that would otherwise be added [59].



(a) (b)

Figure 2: (a) Default pancake solenoid design. The cylinder at the center is the critical magnetic field region (b) Default helical solenoid configuration.

3.1.2 Stellarator Model Geometry

The second Bayesian Optimization covered is mechanical optimization of a stellarator coil, shown in Figure 4. The stellarator coil optimized in this work was a winding pack with 48 turns. The magnetic field data was based on the work of Backmeyer et al., in which a domain decomposition method was used to compute the magnetic field generated by the entire stellarator [7]. The stellarator model was the HELIAS-3 model computed by Andreeva et al [4][10][38]. The current in the model was assumed to be uniform across the cables, and have a value set from Andreeva et al's critical current calculations.The model was scaled down spatially by 1/3 to achieve higher Lorentz forces . We note that this scaling makes the design unphysical. However, the change was made to demonstrate the success of the method when the mechanical challenges are exacerbated. Future work would use the correct scale and would optimize the magnetic field in tandem with mechanical constraints.

The mechanical simulations were done in COMSOL with a linear elastic model. This mechanical modeling method is effective for designing fusion reactor components

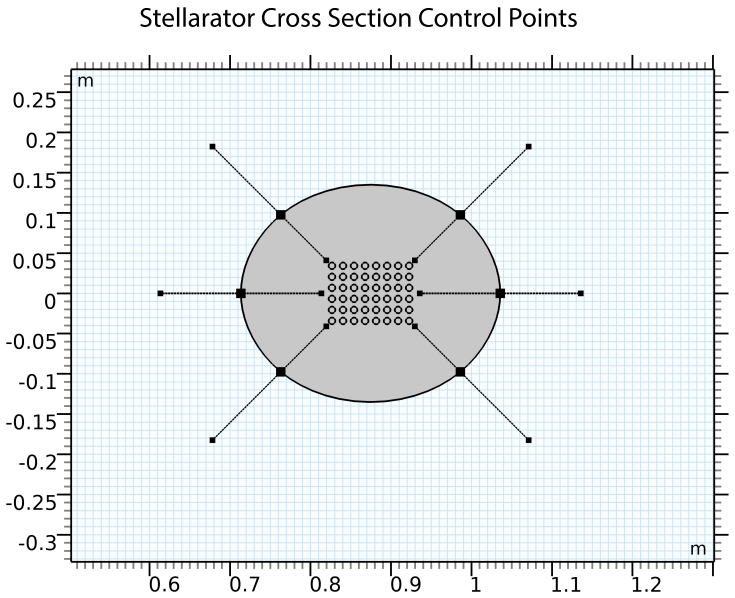


Figure 3: Intermediate size coil casing cross section with control points shown as six bolded points around the outside of the coil casing. The 6 line segments are the possible positions for each corresponding point. The fixed points in the model are the upper and lower right control points.

[1][30], and so the simplification does not produce much loss in accuracy. The coil casing was simulated as high strength alloy steel and the cable as copper. The coil casing was parameterized with control points placed both along its length and around its radius.

At each cross section of the stellarator, there were 6 independent control points. Note that the central axis of the stellarator, which corresponds to the center point in the axisymmetric model, is fixed. This is done to avoid coupling the optimization to the magnetic field. The control points could be moved along the predefined line segments shown in Figure 3 according to the free parameters. As there are six points which each have one degree of freedom to move along a line, there are six degrees of freedom for each cross section. With the control points, the boundary of the cross section was constructed using a smooth spline. Note that the use of smooth splines instead of line segments in this construction precludes the rectangular cross section that was used as a default. We made this decision because prior testing with line segments resulted in jagged and irregular shapes. The cross sections were lofted together to create a single 3D geometry. As there were 20 cross sections optimized for the full model, the 3D stellarator has 120 degrees of freedom. The physics was then simulated with a linear elastic mechanical simulation. The constrained regions of the model were two lines around the outside of the stellarator, which were constrained with a spring foundation. These lines are coincident with the fixed points on the cross sections as they are

embedded in the model. This is necessary because the net force on the winding pack pushes it outwards from the center of the stellarator. In a full design the stellarator would be anchored by a larger superstructure, but for the purposes of this optimization the simplification to two lines is sufficient.

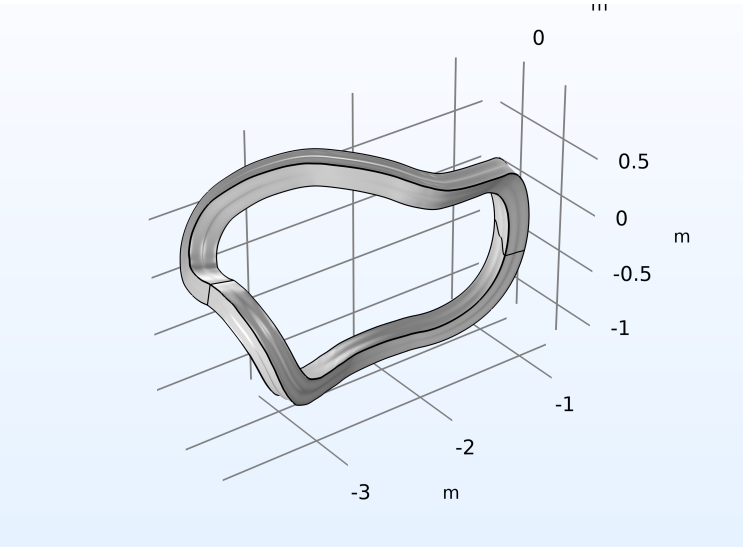


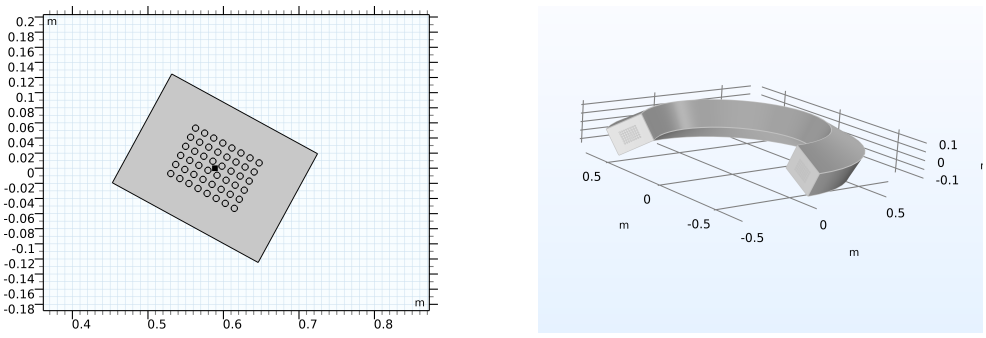
Figure 4: Default winding pack design for 3D stellarator model

The stellarator model was optimized by splitting it into many 2D cross sections. A default cross section and the revolved geometry are shown in Figures 5a and Figure 5b Each 2D cross section was treated as a separate model, with the simplifying assumption of being axially symmetric. The physical assumption is that the changes in stresses in the axial direction were negligible, so the cross section could be treated as a plane strain model. The 2D models had the same field as the 3D model imported, as well as equivalent geometric parameters. The fixed points were slightly different, as an axially symmetric model is inherently constrained in the radial direction. The points were fixed only in the vertical direction, which is necessary because there is a nonzero vertical force.

3.2 Finite Element Modeling

3.2.1 Solenoid Magnetic Field Modeling

The physics of the system was modeled with an A-formulation of Maxwell's equations, which formulates the EM field in terms of magnetic and electric potentials [25][70]. The current was assumed to be homogeneous across the coil. The critical current at each point was calculated from from a table of current density values [31] depending on field strength and the angle of the field vector from the normal vector to the tape [25][52][61]. The engineering current density was taken to be the same as J_c, which



(a) (b)

Figure 5: (a) Default cross section of stellarator model (b) Full revolved geometry of cross section

makes the assumption that the calculated geometry is the REBCO tape itself and that a finished design would be embedded in a larger structure. The finite element calculation evaluates the ratio of the magnetic field to the current at a reference current, then calculates the maximum allowable current everywhere. The initial magneto-static calculation uses the linear equation

$$\nabla \times B_0 = \mu J_0 \tag{4}$$

where B is the magnetic field. As this was formulated in the A formulations, it was handled internally as

$$\nabla \times \nabla \times A_0 = \mu J_0 \tag{5}$$

As the current is assumed to be uniform, then at the critical current value $J = \alpha J_0$ and $B = \alpha B_0$ for some scaling parameter α. Then from the initial solutions B_0 and J_0 the maximum current at any point can be calculated as the solution to the equation

$$\alpha J_0 = J_c(\alpha B_0) \tag{6}$$

The J_c values were interpolated from a dataset of J_c values for REBCO coils at 77 K [31], that give the dependence of critical current on the magnetic field angle and strength. REBCO tape has the property that the critical current depends on both the angle of the tape relative to the magnetic field as well as the field strength, unlike low temperature superconductors REBCO coils are commonly constructed with many layered tapes [2], so modeling approaches must account for possible effects caused by the stacks of tapes. Fortunately, for low frequency applications the tapes can be modeled as a homogenous bulk while retaining accuracy [61]. As this work was primarily focused with time-independent models, accuracy at low frequencies is sufficient.

The solutions to 6 were precomputed, so α could be be directly looked up from a table. As the critical current cannot be exceeded anywhere, the critical current of the whole solenoid is the minimum of the critical currents everywhere on the coil. Then the factor α_{coil} for the whole solenoid is the minimum of the alpha values over every

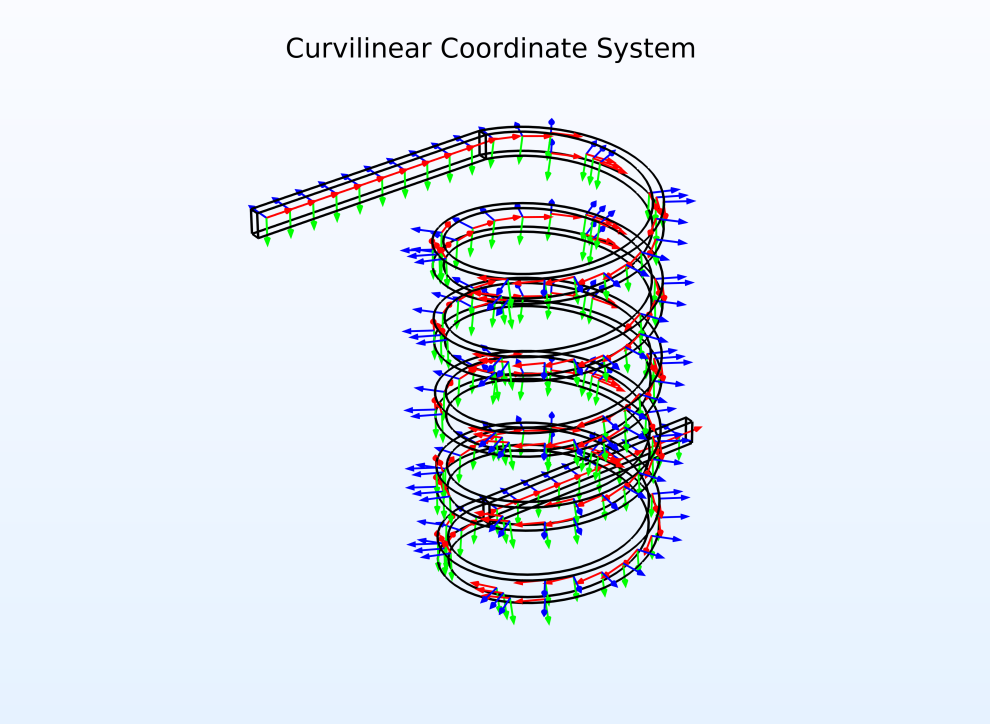


Figure 6: Curvilinear coordinate system of helical model. Red vectors are direction of current, blue are perpendicular to tapes, and green are parallel to tapes.

point on the solenoid. This allowed calculating the total current and therefore the magnetic field strength within the critical region.

The angular dependence of J_c was modeled using a curvilinear coordinate system, which used a diffusion method to solve the wave equation within the coil and find local coordinates tangent and normal to the coil [15].

After using the curvilinear coordinate system as in Figure 6 to calculate the angle between the field and the tape, the critical current density was calculated using a linear interpolation from the J_c data set [31]. The current was then calculated from an estimated cross section of the tape in the coils. This was calculated by multiplying the actual cross section of the coils with a constant between 0 and 1 representing the fraction of the coil consisting of REBCO tape. This simplification was based on work calculating the critical current by integrating the critical current of the tape over the cross section of the coil [52][61][25].

3.2.2 Mechanical Modeling

A mechanical elastic mechanical model was used [5]. The materials used were the COMSOL 6.1 built-in copper and high strength steel, as the mechanical properties of the HTS itself on the whole structure were considered to be negligible. This may impact

the final results, as the stress on REBCO tape may be significantly different than the stress on the copper in these simulations. Additionally, the assumption of elasticity ignores the magneto-plasticity of copper, which is significant in certain conditions [68]. This is another area of further work in refining the models used in this research.

3.3 Bayesian Optimization

The Bayesian optimization component of this work utilized the BoTorch python library [9] to perform each of the necessary steps of Bayesian optimization, from model training to acquisition function optimization. Bayesian optimization, illustrated in Figure 7 works by first fitting a Gaussian process to a few selected data points used for the initial training. The formal definition of a Gaussian process is a collection of random variables indexed by a domain such that every finite collection is multivariate Gaussian. For a Gaussian process Y and a parameter space X, for each $x \in X$ then $Y(x)$ will be a normally distributed random variable. The mean is defined by a mean function

$$\mathbb{E}\left[Y(X)\right] = \mu(x) \tag{7}$$

which in this case is taken to be constant a priori. The covariance between two random variables $Y(x)$ and $Y(x')$ is defined by a covariance kernel

$$\mathrm{Cov}[Y(x), Y(x')] = k(x, x') \tag{8}$$

for a covariance kernel $k : X \times X \to \mathbb{R}$. In this work the covariance is taken a priori to be the Matérn kernel [40][23].

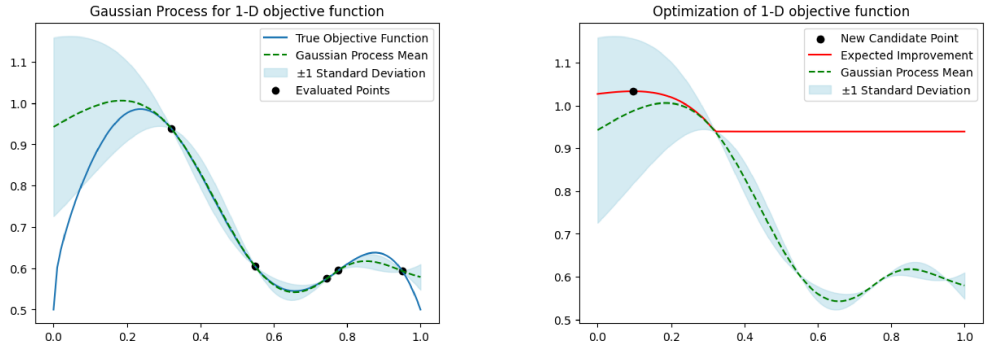
More intuitively, the covariance between any two of the variables determines how close to each other they are likely to be. The Gaussian processes used in Bayesian optimization generally have a covariance that is higher for variables that are spatially closer together, which in essence predicts that if a point is close to a measured data point in the domain, it likely will have a similar measured value as well. This allows the Gaussian Process to interpolate the known data and provide estimates of error for the interpolation.

Once the Gaussian process has been fitted to the observed data points using Bayes' rule, an acquisition function is created, which evaluates the potential of any new point in the search space. The acquisition models the trade off between exploration of unknown regions and improvement in the known best value, in the search for the global maximum. The acquisition function used in this work is the hypervolume log Expected Improvement [17]. This method applied the log expected improvement acquisition function [9][3] and to optimize the hypervolume of the observed space. The hypervolume of a set of observed points Y is defined as

$$HV(Y) = \mu(\{x : \exists y \in Y, r \preccurlyeq x \preccurlyeq y\}) \tag{9}$$

where μ is the Lebesgue measure, r is a pre-chosen reference point and $x \preccurlyeq y$ means that x is dominated by y, so that for every index i, $x_i \leq y_i$.

The expected improvement is defined as the logarithm of the expected improvement acquisition function, which is defined as



(a) (b)

Figure 7: (a) Gaussian Process fitted to known evaluation points from objective function and (b) Optimization of acquisition function for 1-D objective function. Note that the acquisition function is shifted to align with the best value so far

$$EI(x) = \mathbb{E}[\max(f(x), y^*)] \tag{10}$$

where y^* is the best evaluated value so far. The improvement of a measurement is how much that measurement improves over the known values, so taking the expected value of that given the Gaussian process gives EI. The logarithm is taken to adapt it to LEI for numerical stability [3]. LEI balances between the search and optimization by incentivising improvement rather than high expected value for new evaluations. When it predicts improvements with high certainty near previously evaluated parameters, then it will improve upon the current best optimal value. However, once the predicted improvement in explored regions is small, the expected improvement in unexplored regions will be higher, causing the algorithm to move towards the unknown regions which have a low probability of improvement but larger possible values, giving larger expected improvement. The hypervolume expected improvement is then

$$EHVI(x) = \mathbb{E}[HV(Y \cup \{f(x)\})] \tag{11}$$

The maximum can be omitted, as the hypervolume of a set is never decreased by adding another point.

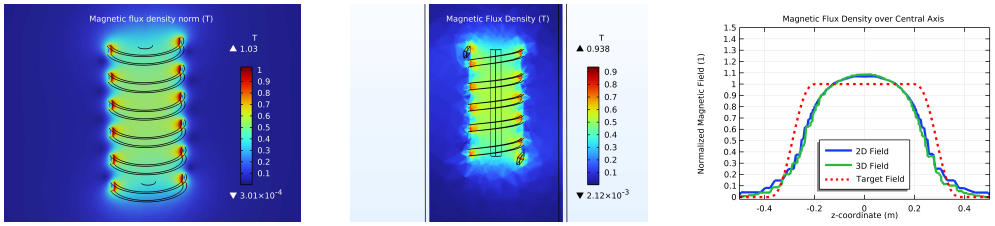
The acquisition function is then optimized. This can be done with many numerical optimization methods, but for this research the algorithm used was based on greedy approaches to differential methods. Those methods have been found to work well with the Monte-Carlo integral based implementations of many acquisition functions, especially for the expected improvement acquisition function [65]. After the acquisition function is optimized, either over a single point in the search space or as a collection of points, then the new points are evaluated and added to the data set. The process then repeats.

3.4 Solenoid Model Optimization

The objective functions for the solenoid models were the field regularity and the magnetic field strength. This was measured as the deviation from the desired field in a cylinder contained within the solenoid (solenoid critical region) and the average magnetic field within the critical region. The desired field in this work was a field aligned in the vertical direction, with strength equal to the average vertical component of the field. The way that the deviation from the desired field was measured was not with a maximum value, but an L^p norm. This change was done to avoid numerical errors from the mesh. L^p norms are calculated via an integral over the whole space, and have the property that for larger p, they approximate the maximum value. The formula for the L^p norm is

$$||B_{err}||_{L^p} = (\mathbb{E}[|B_{err}(x)|^p])^{1/p} \tag{12}$$

where $B_{err}(x)$ is the relative error in the magnetic field at point x. An expected value was used instead of just an integral over the whole space because it more closely approximates the maximum value (as p approaches infinity the $1/p$ exponent would remove any constant factor from the integral, but relatively slowly). In this work the value used for p was 10. Illustrations of the default geometries and the magnetic fields are shown in Figure 8.



(a) (b) (c)

Figure 8: Magnetic field norm for pancake (a) and helical (b) solenoid model with default configuration. (c) Normalized magnetic flux density of default coil configurations compared to optimal flux geometry, showing that the default solenoid models did not conform well to the desired magnetic field. Note that the field is only compared to the target field in the $|z| \leq 0.2\,\text{m}$ region, the rest of the field is a smooth interpolation and is shown only for illustrative purposes.

The Bayesian Optimization trials used 50 iterations for the solenoid models, and used only 1 model evaluation per iteration. We note that the use of multiple model evaluations per iteration is particularly advantageous when used in conjunction with parallelized hardware. This will be explored in future work on more complicated geometries.

The next step after demonstrating the feasibility of Bayesian optimization for magneto-static optimization problems was to use a variant called multi-fidelity Bayesian Optimization [67]. Multi-fidelity Bayesian optimization is a label for a class

of algorithms that seek to use data from lower fidelity analogues of the model of interest to accelerate the optimization process. They generally either model or assume a correlation between the models. They additionally need to use a time weighted acquisition function that evaluates the trade-off between model runtime and improvement [67] In the case of this research, the axially symmetric model took about 3 seconds per evaluation, while the helical model took 1 to 2 minutes. This large difference in evaluation time suggested that the multi-fidelity algorithm would be advantageous.

However, one of the deficiencies of many current implementations of multi-fidelity BO is that they assume that both low and high-fidelity versions of the model have the same mean. The failure of multi-fidelity optimization between the axially symmetric and the helical solenoid models showed that this condition was not met. To correct this, a linear regression on the pancake vs. helical model data was used, and then a GP modeled the deviation from the linear regression. As the mean of the deviation from a linear regression is 0, this corrected for that problem. This method was then simplified by turning the optimization into a 2-step process. The 2D model was optimized first, and then a set of initialisation parameters were evaluated for both the 2D and the 3D models. These parameters were chosen using the upper confidence bound acquisition function [6][57]. In this work 20 parameter choices were evaluated for the initial data set, as 20 is around twice the dimension of the parameter space. This was found in this work to be an adequate number of evaluations for initialization of the Bayesian optimization algorithm. Then a linear regression was done on the 20 parameters for which both the 2D and 3D models were evaluated. The multi-fidelity GP was then fit, using the original parameters as inputs and the deviations of each data point from the linear regression as the output. This GP was then used to produce posterior distributions for the 3D model at each point where the 2D model was evaluated. Those distributions were then treated as new evaluations and added to the 3D model data set, with the posterior mean as the value and the posterior variance for the dataset variance.

With the extra data, the 3D model was then optimized through a standard Bayesian optimization loop, with the exception that each time the 3D model was evaluated, the 2D model was evaluated at the same parameters and the data points were entered into the multi-fidelity GP. This was done because the extra 2D model evaluations are very cheap compared to the 3D model evaluations, and the extra data points allowed refinement of the multi-fidelity GP, and consequently better inference of the 3D model from the 2D model.

3.4.1 Multi-Objective Multi-Fidelity (MOMF) Optimization

In addition to the warm start optimization, the MOMF method described in [32] was compared to the hypervolume expected improvement optimization. In this method the fidelity parameter is treated as an additional objective to be jointly optimized along with the other objectives. The acquisition function is then weighted by the cost of the function evaluation. The acquisition function is then

$$MFEHVI(x, s) = \frac{\mathbb{E}[HV(Y \cup \{(f(x), s)\})]}{cost(s)} \tag{13}$$

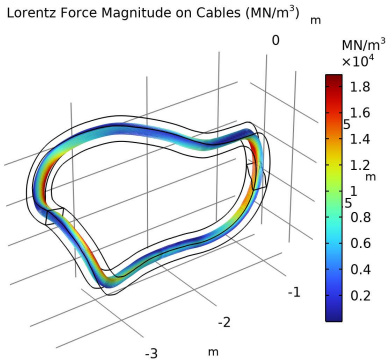


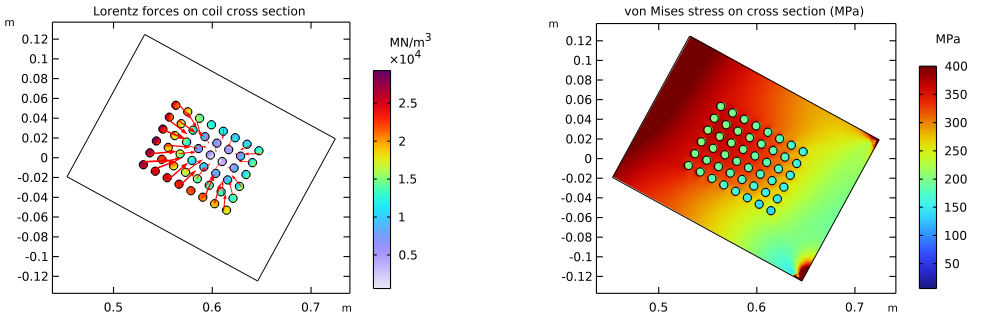
Figure 9: Lorentz forces on cables

where s is the fidelity parameter and Y is the set of all tuples $(y_1, ..., y_k, s_i)$ of objective function and fidelity values. In this work the 3D model took about 20 times longer to evaluate than the 2D model, so the cost factor was 20. As there were only two fidelities, s could take values 0 and 1, so $cost(0) = 1$ and $cost(1) = 20$.

3.5 Stellarator Model Optimization

As Figure 9 shows, the Lorentz forces on the winding pack modeled in this work are very high, up to $1.8 \times 10^{13}\,\mathrm{N/m^3}$, demonstrating the necessity of optimization. As directly optimizing the model would be difficult due to its complexity and high number of geometry parameters, the optimization was done in two steps. In the first step, the stellarator was sampled at 20 cross sections. Each cross section was treated as an axially symmetric model to speed up computation. Then each axially symmetric cross section was modeled using the local Lorentz forces. Note that instead of directly importing Lorentz force data into the cross-section model, magnetic field data was used, and the forces were computed within the 2D model. The 2D models were then optimized via a constrained Bayesian optimization algorithm, in which the constraint was von mises forces on the REBCO coil and the objective was low steel volume. Von Mises stress was chosen as a constraint because the dependence of REBCO critical current on mechanical stress is well studied [2] [43]. Experiments have found that the critical current of REBCO begins to sharply decline with stresses around several hundred megapascals, with specific numbers varying by experiment and type of strain [2][43]. In this work we used a cutoff value of 250 MPa of von Mises stress as the upper limit allowed in optimizations. However, as von Mises stress is more suited for isotropic and ductile materials, this is a simplification to the model. In reality each circular coil would be composed of many tapes each wound around a core, likely similar to a VIPER cable [52]. Fully modeling a VIPER cable would require computing the orientation of every REBCO tape, which was infeasible for this work. Further work aims to more accurately model the coils and understand how they can be homogenized. For this work however, the cables were treated as fixed parts of the geometry, so that the magnetic field did not change with respect to the optimization parameters. The

objective function for minimization was chosen to be the steel volume because a smaller volume occupied by the coil casing will allow more space for other components of a fusion reactor. A more sophisticated objective might account for the precise geometries of the coil casing and the adjacent components, but that was decided to be beyond the scope of this work. The stresses on one of the cross sections and on the full coil casing and shown in Figures 10 and 11



(a) (b)

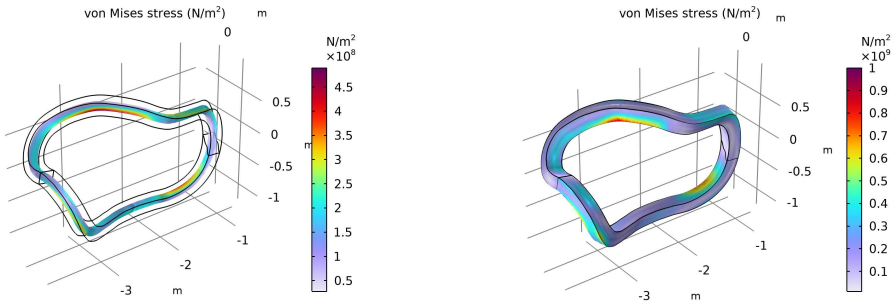
Figure 10: (a) Lorentz forces for default winding pack cross section. The colors show magnitude and the arrows show direction of the force. (b) Von Mises stress on default cross section. The two corners on the right hand side are fixed points, so they have high stresses.

This algorithm works by modeling both the objective and the constraint with a Gaussian Process, then using an acquisition function based on a constraint weighted objective of the two GPs [36][22]. This essentially used a posterior transform on the data so that the acquisition function would only consider parameters where the constraints are satisfied (or predicted to be satisfied). After optimizing the cross sections, we interpolated the control points along each cross section into a CAD for the full 3D coil casing. This CAD was then run in a full 3D mechanical simulation.

4 Results

4.1 Solenoid Model Results

In the solenoid optimization, the initial field quality (measured as normalized deviation from the target homogenous field) was 13.7 % for the pancake model and 16.3 % for the helical model. The average field strength was 0.5 T for the pancake model and 0.48 T for the helical model. After the multi-objective optimization, for both models, BO was used to choose 10 parameters to minimize the field deviation while ensuring that the average field was at least 0.5 T. The optimized field quality was then 3.7 % for the pancake model and 5.9 % for the helical model. Figure 13 show the default and optimal geometries of the models, as well as the deviation of their magnetic fields from the desired uniform field.



(a) (b)

Figure 11: Stresses on default stellarator model. (a) shows stresses on the coils and (b) shows stresses on the coil casing.

	Default Parameters	Optimized for pancake model	Optimized for helical model
r_1 (cm)	10	8.6	8.0
r_2 (cm)	10	8.4	7.6
r_3 (cm)	10	7.1	6.0
z_1 (cm)	4	3.9	4.1
z_2 (cm)	12	12.32	12.0
z_3 (cm)	20	20.1	20.1
θ_1 (deg)	0	3.8	-3.0
θ_2 (deg)	0	-7.3	-3.0
θ_3 (deg)	0	4.9	-15.5
Pancake model deviation (%)	13.7	3.7	4.2
Helical model deviation (%)	16.3	7.6	5.9
Pancake model average field (T)	0.50	0.50	0.51
Helical model average field (T)	0.48	0.49	0.50

Figure 12: Solenoid geometry parameters for default and optimized configurations, as well as objective functions (normalized deviation from target field and average field strength). Note that all parameter and output values are rounded, so may produce slightly different results when rerun on model.

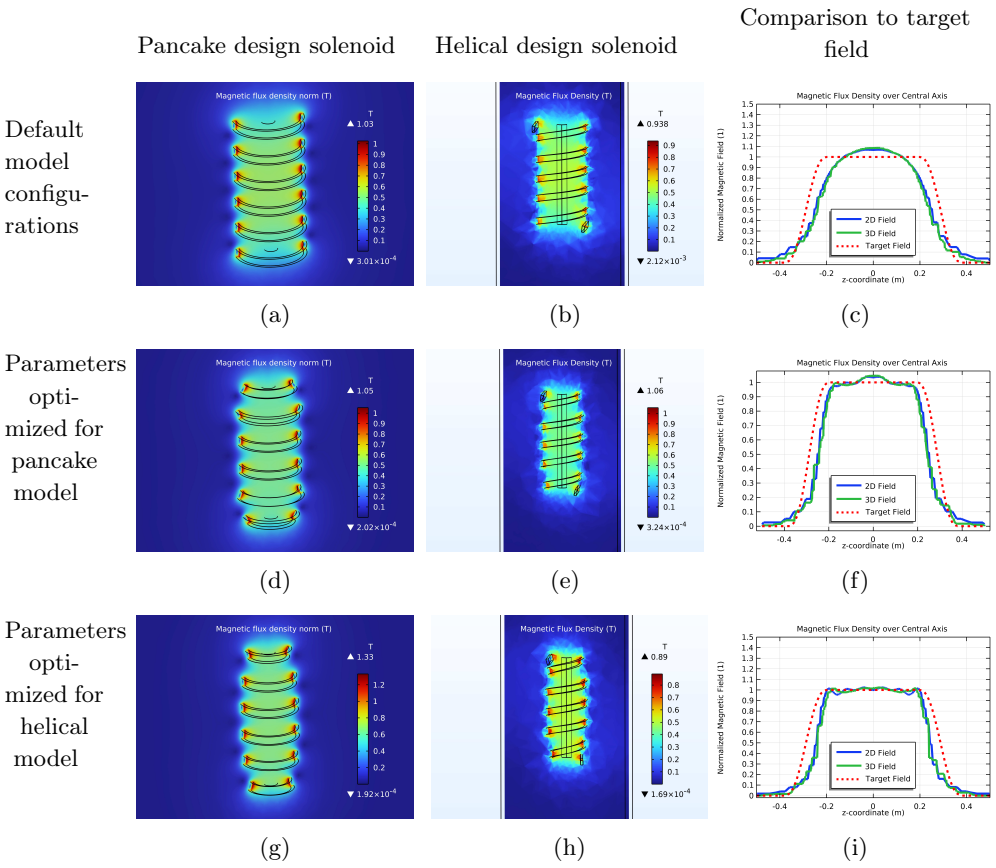
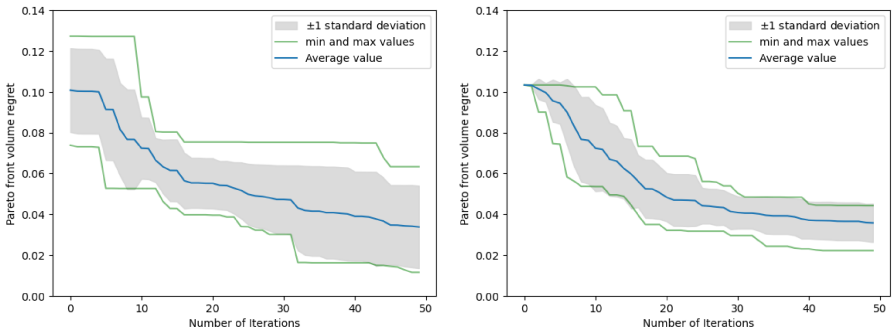


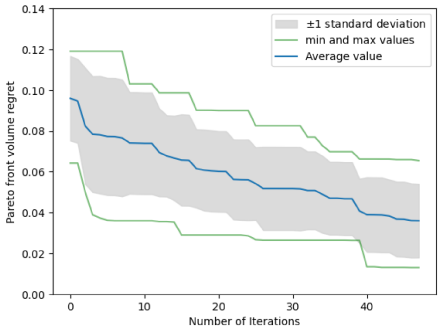
Figure 13: Default and optimized (both for pancake and helical model) solenoids. For the graphs, note that the critical region in which the field was measured is between $z = -0.2\,\mathrm{m}$ and $z = 0.2\,\mathrm{m}$

The results below were obtained with 10 model evaluations for the initial training set and the optimization loop was run for 50 additional model evaluations. The regret was calculated by computing the hypervolume for all of the recorded observations (1253 over several dozen optimization trials) and treating that value (1.138) as the true hypervolume of the Pareto front. The regret was then the difference between the Pareto front volume during optimization and the estimated true Pareto front.



(a) (b)

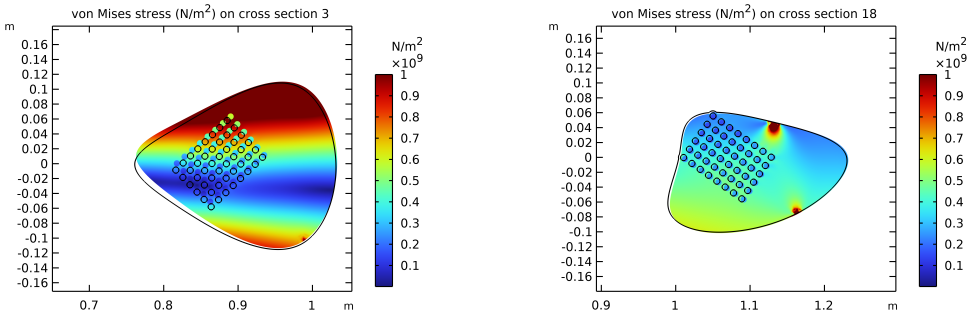
Figure 14: Model optimization withouth (a) and with (b) warm start for BO algorithm over 7 trials



(a) Model optimization with MF-EHVI algorithm in [32]

4.2 Stellarator Results

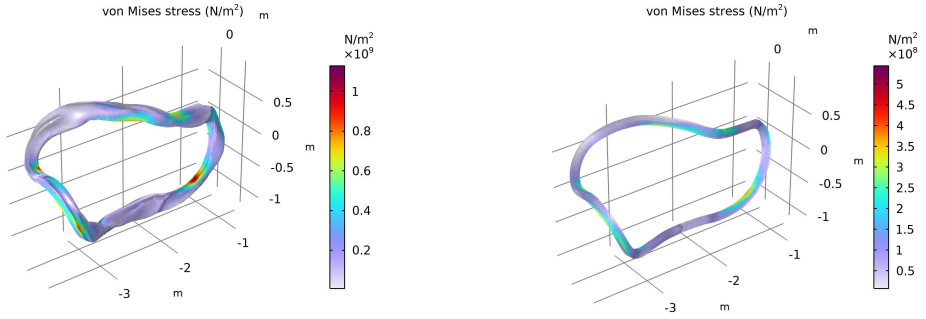
Von Mises stresses were calculated for 20 evenly interspersed cross sections of the default stellarator model. The mean steel volume of the axially symmetric models was 0.22 m^3 and the mean stress level was 286 MPa. The mean steel volume of the optimized axially symmetric models was 0.24m^3 and the mean von Mises stresses were 280 MPa. Two of the optimized cross sections are shown in Figure 16.



(a) (b)

Figure 16: Optimized cross sections 3 (a) and 18 (b). The two points with high stresses in the upper and lower right of both models are the fixed points.

After optimizing the models, the full stellarator was generated by interpolating the cross sections. The default cross section had 0.289 m^3 of steel, and stresses up to 266 MPa. The optimized model had stresses up to 242 MPa and 0.270 m^3 of steel. Below is the optimized stellarator model.



(a) (b)

Figure 17: (a) Stresses on the whole optimized winding pack. (b) Stresses on the coils

5 Discussion

5.1 Quality of optimization results

The performance of the Bayesian optimization algorithm in finding optimum values was promising. To establish that it was finding values close to the true Pareto front, the Pareto front was estimated from all values over all of the optimizations. As the initial regret was around 0.1 and decreased to around 0.03 on average, the optimization

was moderately successful. However, it is not surprising that the true Pareto front is larger than a Pareto front composed of less than 50 points, as 50 points is not enough to exactly contain the whole volume dominated by the Pareto front.

As another test of the optimization, we compared the values to standards for solenoids used in actual detectors. The pancake model had 3.7 % error and the helical model had 5.9% error. While this is slightly higher than standards for actual detectors of around 0.1% or lower [44], this result is promising because of the small parameter space used, as well as the constraint on field strength that affected the field regularity. This shows that the optimal values found in this work were likely close to the true optimum values. The comparison to real solenoids also suggests that a larger design space with a more complex geometry could yield directly applicable results with the same optimization techniques [44].

5.2 Comparison of pancake and helical solenoid models

The pancake and helical models were designed to be as close as possible to each other while still retaining the fundamentally distinct geometries. Despite this, there were still significant differences between the models. While they are reasonably well correlated for most random parameters, the current tended to be higher in the pancake model. This is likely because coil passed through more angles in the helical model, rather than having a finite set of angles, which produced a lower critical current in the minimization over the entire coil.

Because of these effects, the two models could not be considered as different fidelity versions of each other. Additionally, the models are not only biased, but have regression slope not equal to 1. Across a large portion of parameters the helical model displayed an error between 50% and 100% larger than the corresponding error for the 2D model. While multi-objective Bayesian optimization algorithms have been described [67][45] they do not account for both of these effects. This work uses a linear regression to train a Gaussian process, modeling the difference between the fidelities in a more sophisticated way.

5.3 Multi-fidelity optimization

As Figure 14 and 15a show, the warm start optimization and the MF-EHVI optimizations have very similar final regret values to the optimization without a warm start. However, there are two notable differences. The warm start optimization has less variance than both the MF-EHVI and the default optimization across all numbers of iterations. There is an especially large difference in variance between the warm start and the default at the end of the optimization, where the warm start optimization has a standard deviation in regret of 9.4×10^{-3} and the default optimization has a standard deviation of 2.0×10^{-2}. The other important difference between the different optimization procedures was the optimization speed. The warm start optimization reached a regret of 0.04 within 34 optimization iterations, while both the MF-EHVI and default algorithms requires 40 iterations. This is a speed improvement of 15%.

Some of the similarity betweeen the different optimization methods may be due to the relatively low number of low fidelity evaluations used - the warm start optimization

used 60 and the MF-EHVI generally used around 30. This suggests that careful tuning of the cost-weighted acquisition function may produce better results. Additionally, the multi-fidelity hypervolume knowledge gradient acquisition function [16] may do better than MF-EHVI at balancing the exploration and optimization in multi-fidelity optimization.

5.4 Stellarator Model Optimization

The results for the stellarator cross sections were not a huge improvement, as the average steel volume (in the revolved geometry) of the cross sections increased slightly while the average stresses decreased by 6 MPa. However, Figure 18 showing a scatterplot of the volume and stress values demonstrates an area of improvement.

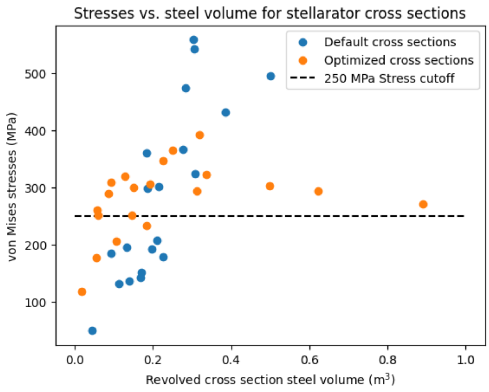


Figure 18: Scatter plot of steel volumes and stress levels on axially symmetric cross section models.

While the optimized 3D model was unable to keep most of the stress levels below the threshold of 250 MPa on the superconducting cables, it did significantly lower the stresses on the cross sections with the highest stresses. The maximum stress level across all of the cross sections decreased from 559 MPa to 392 MPa. This suggests that more work will be successful in optimizing the cross sections better, perhaps with a better choice of geometry parameterization or a different Bayesian optimization algorithm.

For the full stellarator coil seen in Figures 17a and 17b, both the stresses and the steel volume decreased slightly. The differences in stresses between the cross section model and the full 3D model is likely due to the way the models were constrained. The axially symmetric models were fixed in the z direction, and the full model had a spring foundation constraint, which might reduce the stresses. Nevertheless, the improvement in the full model shows that the Bayesian optimization was successful in its goal of reducing both the stress levels and the steel volume.

However, one concern with the optimized designs from this work is the strange geometry, which is apparent in Figure 17a. Rapid variation in steel thickness along the length of winding pack suggests that it may be suboptimal. This is likely true, as

the optimization procedure had limited ability to correct for twists in the coil. Further work will investigate ways to better optimize all of the parameters in the whole coil simultaneously.

Another issue is whether the design could be manufactured. Fabricating such a highly irregular and non-planar shape that contains fragile and heat sensitive coils will be difficult. However, additive manufacturing for stellarator coils is rapidly advancing [49], and is a promising method for future coil designs.

Beyond fabrication, there are other infeasible aspects of this design, such as the absence of a superstructure to hold the winding pack, as well as geometric constraints to accomate other components of the stellarator. Because stellarators require so many feasibility constraints, we decided that it would be beyond the scope of this work to enforce them all. In a full workflow, design constraints would also be modeled with a Gaussian process and optimized via the same constrained Bayesian optimization algorithm used for stellarator design.

6 Conclusion

This work demonstrates the efficacy of Bayesian optimization as a viable tool for engineering stellarator winding pack design. The improvements from 16.3% to 5.9% in magnetic field error for the helical solenoid model and the combined 24 MPa (9.0% relative decrease) and $0.019\,\mathrm{m}^3$ (6.6% relative decrease) reductions in stress and steel volume for the stellarator show that Bayesian optimization is an effective optimization tool. Additionally, the examination of multi-fidelity multi-objective optimization shows that further work is necessary to understand which algorithms work best for optimizing finite element models.

While many simplifications were made in the models in this work, they demonstrate a workflow that can be extended to more complex and interconnected systems. The many constraints, including mechanical stresses, quench stability, magnetic field geometry, and manufacturability, will all need to be balanced in a successful design. As these constraints and objective often require different simulations and codes to compute, the flexible framework that Bayesian optimization provides is invaluable. This research demonstrates that multi-task models are well suited to the task of correlating and interpolating many different requirements, then judiciously choosing each model evaluation to achieve the task as quickly as possible. Further work building comprehensive workflows with Bayesian optimization will enable the possibility of a faster timeline for fusion reactor development, and the eventual deployment of a fusion power plant.

7 Acknowledgements

Authors would like to thank Rachel Shulman (Massachusetts Institute of Technology), Dr. Michael Short (Massachusetts Institute of Technology), Tyler Ea (Massachusetts Institute of Technology), Dr. Luisa Chiesa (Tufts University), and Dr. Robert Granetz (Massachusetts Institute of Technology) for the support and helpful discussions. We would also like to thank the Plasma Science and Fusion Center MIT and the Fusion

for Undergraduate Scholars program with its donor Mr. Burns for sponsoring nuclear research.

References

[1] Alexander B. Alekseev. "Chapter 12 - Mechanics of Magnetic Fusion Reactors". In: *Fundamentals of Magnetic Thermonuclear Reactor Design*. Ed. by Vasilij Glukhikh, Oleg Filatov, and Boris Kolbasov. Woodhead Publishing Series in Energy. Woodhead Publishing, Jan. 1, 2018, pp. 351–386. ISBN: 978-0-08-102470-6. DOI: 10.1016/B978-0-08-102470-6.00012-3. URL: https://www.sciencedirect.com/science/article/pii/B9780081024706000123 (visited on 03/04/2024).

[2] N. C. Allen, L. Chiesa, and M. Takayasu. "Structural modeling of HTS tapes and cables". In: *Cryogenics*. Chats on Applied Superconductivity 2015 University of Bologna, Italy, 14-16 September 2015 80 (Dec. 1, 2016), pp. 405–418. ISSN: 0011-2275. DOI: 10.1016/j.cryogenics.2016.02.002. URL: https://www.sciencedirect.com/science/article/pii/S0011227516000266 (visited on 01/26/2024).

[3] Sebastian Ament et al. *Unexpected Improvements to Expected Improvement for Bayesian Optimization*. Oct. 31, 2023. DOI: 10.48550/arXiv.2310.20708. arXiv: 2310.20708[cs,math,stat]. URL: http://arxiv.org/abs/2310.20708 (visited on 01/17/2024).

[4] Tamara Andreeva et al. "The Helias Reactor Concept: Comparative Analysis of Different Field Period Configurations". In: *Fusion Science and Technology* 46.2 (Sept. 1, 2004). Publisher: Taylor & Francis _eprint: https://doi.org/10.13182/FST04-A579, pp. 395–400. ISSN: 1536-1055. DOI: 10.13182/FST04-A579. URL: https://doi.org/10.13182/FST04-A579 (visited on 03/03/2024).

[5] Hedy Attouch, Giuseppe Buttazzo, and Gérard Michaille. "Chapter 15: Variational problems with a lack of coercivity". In: *Variational Analysis in Sobolev and BV Spaces*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, Oct. 2, 2014, pp. 599–642. ISBN: 978-1-61197-347-1. DOI: 10.1137/1.9781611973488.ch15. URL: https://epubs.siam.org/doi/10.1137/1.9781611973488.ch15 (visited on 03/16/2023).

[6] Peter Auer. "Using Confidence Bounds for Exploitation-Exploration Trade-offs." In: *Journal of Machine Learning Research* 3 (Jan. 1, 2002), pp. 397–422. DOI: 10.1162/153244303321897663.

[7] Merle Backmeyer et al. *Modeling non-planar coils in a full-scale stellarator*. Publication Title: arXiv e-prints ADS Bibcode: 2023arXiv231000945B. Oct. 1, 2023. DOI: 10.48550/arXiv.2310.00945. URL: https://ui.adsabs.harvard.edu/abs/2023arXiv231000945B (visited on 01/26/2024).

[8] Tianyi Bai et al. *Transfer Learning for Bayesian Optimization: A Survey*. Feb. 12, 2023. arXiv: 2302.05927[cs]. URL: http://arxiv.org/abs/2302.05927 (visited on 10/27/2023).

[9] Maximilian Balandat et al. "BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization". In: *Advances in Neural Information Processing Systems*. Vol. 33. Curran Associates, Inc., 2020, pp. 21524–21538. URL: https://proceedings.neurips.cc/paper/2020/hash/f5b1b89d98b7286673128a5fb112cb9a-Abstract.html (visited on 10/26/2023).

[10] C. D. Beidler et al. "The Helias reactor HSR4/18". In: *Nuclear Fusion* 41.12 (Dec. 2001), p. 1759. ISSN: 0029-5515. DOI: 10.1088/0029-5515/41/12/303. URL: https://dx.doi.org/10.1088/0029-5515/41/12/303 (visited on 03/03/2024).

[11] Edwin V Bonilla, Kian Chai, and Christopher Williams. "Multi-task Gaussian Process Prediction". In: *Advances in Neural Information Processing Systems*. Vol. 20. Curran Associates, Inc., 2007. URL: https://papers.nips.cc/paper_files/paper/2007/hash/66368270ffd51418ec58bd793f2d9b1b-Abstract.html (visited on 01/29/2024).

[12] Theodore Brown et al. "Multi-objective Bayesian optimization for design of Pareto-optimal current drive profiles in STEP". In: *IEEE Transactions on Plasma Science* 52.9 (Sept. 2024), pp. 3904–3909. ISSN: 0093-3813, 1939-9375. DOI: 10.1109/TPS.2024.3382775. arXiv: 2310.02669[physics]. URL: http://arxiv.org/abs/2310.02669 (visited on 12/30/2024).

[13] R. Burhenn et al. "Diagnostics Development for Steady State Operation of the Stellarator Wendelstein 7-X". In: *Contributions to Plasma Physics* 51.2 (2011). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ctpp.201000053, pp. 271–278. ISSN: 1521-3986. DOI: 10.1002/ctpp.201000053. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ctpp.201000053 (visited on 01/26/2024).

[14] F. Castejón et al. "Stellarator optimization under several criteria using metaheuristics". In: *Plasma Physics and Controlled Fusion* 55.1 (Dec. 2012). Publisher: IOP Publishing, p. 014003. ISSN: 0741-3335. DOI: 10.1088/0741-3335/55/1/014003. URL: https://dx.doi.org/10.1088/0741-3335/55/1/014003 (visited on 01/30/2024).

[15] COMSOL. *Defining Curvilinear Coordinates for Anisotropic Materials*. COMSOL. URL: https://www.comsol.com/blogs/defining-curvilinear-coordinates-anisotropic-materials/ (visited on 01/30/2024).

[16] Sam Daulton, Maximilian Balandat, and Eytan Bakshy. "Hypervolume Knowledge Gradient: A Lookahead Approach for Multi-Objective Bayesian Optimization with Partial Information". In: *Proceedings of the 40th International Conference on Machine Learning*. International Conference on Machine Learning. ISSN: 2640-3498. PMLR, July 3, 2023, pp. 7167–7204. URL: https://proceedings.mlr.press/v202/daulton23a.html (visited on 12/30/2024).

[17] Samuel Daulton, Maximilian Balandat, and Eytan Bakshy. *Parallel Bayesian Optimization of Multiple Noisy Objectives with Expected Hypervolume Improvement*. Oct. 26, 2021. DOI: 10.48550/arXiv.2105.08195. arXiv: 2105.08195[cs]. URL: http://arxiv.org/abs/2105.08195 (visited on 12/30/2024).

[18] Peter Deuflhard. "Systems of Equations: Global Newton Methods". In: *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*. Ed. by Peter Deuflhard. Springer Series in Computational Mathematics. Berlin, Heidelberg: Springer, 2011, pp. 109–172. ISBN: 978-3-642-23899-4. DOI: 10.1007/978-3-642-23899-4_3. URL: https://doi.org/10.1007/978-3-642-23899-4_3 (visited on 03/03/2024).

[19] M. Drevlak et al. "Optimisation of stellarator equilibria with ROSE". In: *Nuclear Fusion* 59.1 (Nov. 2018). Publisher: IOP Publishing, p. 016010. ISSN: 0029-5515.

DOI: 10.1088/1741-4326/aaed50. URL: https://dx.doi.org/10.1088/1741-4326/aaed50 (visited on 10/12/2023).

[20] Daniel Dudt et al. "The DESC Stellarator Code Suite Part III: Quasi-symmetry optimization". In: *Journal of Plasma Physics* 89.2 (Apr. 2023), p. 955890201. ISSN: 0022-3778, 1469-7807. DOI: 10.1017/S0022377823000235. arXiv: 2204.00078[physics]. URL: http://arxiv.org/abs/2204.00078 (visited on 10/16/2023).

[21] Zahra Zanjani Foumani et al. "Multi-Fidelity Cost-Aware Bayesian Optimization". In: *Computer Methods in Applied Mechanics and Engineering* 407 (Mar. 2023), p. 115937. ISSN: 00457825. DOI: 10.1016/j.cma.2023.115937. arXiv: 2211.02732[stat]. URL: http://arxiv.org/abs/2211.02732 (visited on 12/14/2024).

[22] Jacob Gardner et al. "Bayesian Optimization with Inequality Constraints". In: *Proceedings of the 31st International Conference on Machine Learning*. International Conference on Machine Learning. ISSN: 1938-7228. PMLR, June 18, 2014, pp. 937–945. URL: https://proceedings.mlr.press/v32/gardner14.html (visited on 03/04/2024).

[23] Marc G. Genton. "Classes of kernels for machine learning: a statistics perspective". In: *The Journal of Machine Learning Research* 2 (Mar. 1, 2002), pp. 299–312. ISSN: 1532-4435.

[24] Andrew Giuliani et al. *Single-stage gradient-based stellarator coil design: Optimization for near-axis quasi-symmetry*. Mar. 15, 2022. DOI: 10.48550/arXiv.2010.02033. arXiv: 2010.02033[physics]. URL: http://arxiv.org/abs/2010.02033 (visited on 01/30/2024).

[25] F. Gömöry and B. Klincok. "Self-field critical current of a conductor with an elliptical cross-section". In: *Superconductor Science Technology* 19 (Aug. 1, 2006). ADS Bibcode: 2006SuScT..19..732G, pp. 732–737. ISSN: 0953-2048. DOI: 10.1088/0953-2048/19/8/007. URL: https://ui.adsabs.harvard.edu/abs/2006SuScT..19..732G (visited on 01/26/2024).

[26] L.A. El-Guebaly. "Neutronics analysis for the stellarator power plant study SPPS". In: *Proceedings of 16th International Symposium on Fusion Engineering*. Proceedings of 16th International Symposium on Fusion Engineering. Vol. 2. Sept. 1995, 1162–1165 vol.2. DOI: 10.1109/FUSION.1995.534432. URL: https://ieeexplore.ieee.org/document/534432 (visited on 01/26/2024).

[27] Huong Ha et al. "Bayesian Optimization with Unknown Search Space". In: *Advances in Neural Information Processing Systems*. Vol. 32. Curran Associates, Inc., 2019. URL: https://papers.nips.cc/paper_files/paper/2019/hash/ccf0304d099baecfbe7ff6844e1f6d91-Abstract.html (visited on 01/29/2024).

[28] H.-J. Hartfuss, R. König, and A. Werner. "Diagnostics for steady state plasmas". In: *Plasma Physics and Controlled Fusion* 48.10 (Sept. 2006), R83. ISSN: 0741-3335. DOI: 10.1088/0741-3335/48/10/R01. URL: https://dx.doi.org/10.1088/0741-3335/48/10/R01 (visited on 01/26/2024).

[29] S. P. Hirshman and J. C. Whitson. "Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria". In: *The Physics of Fluids* 26.12 (Dec. 1, 1983), pp. 3553–3568. ISSN: 0031-9171. DOI: 10.1063/1.864116. URL: https://doi.org/10.1063/1.864116 (visited on 03/03/2024).

[30] Yue Huang et al. "Thermo-structural design of the European DEMO water-cooled blanket with a multiscale-multiphysics framework". In: *Fusion Engineering and Design* 135 (Oct. 1, 2018), pp. 31–41. ISSN: 0920-3796. DOI: 10.1016/j.fusengdes.2018.07.007. URL: https://www.sciencedirect.com/science/article/pii/S0920379618305623 (visited on 03/04/2024).

[31] Robinson Research Institute. *Robinson HTS Wire Critical Current Database*. URL: https://htsdb.wimbush.eu/ (visited on 01/26/2024).

[32] Faran Irshad, Stefan Karsch, and Andreas Döpp. "Leveraging Trust for Joint Multi-Objective and Multi-Fidelity Optimization". In: *Machine Learning: Science and Technology* 5.1 (Mar. 1, 2024), p. 015056. ISSN: 2632-2153. DOI: 10.1088/2632-2153/ad35a4. arXiv: 2112.13901[cs]. URL: http://arxiv.org/abs/2112.13901 (visited on 12/30/2024).

[33] D. C. Van Der Laan et al. "A CORC® cable insert solenoid: the first high-temperature superconducting insert magnet tested at currents exceeding 4 kA in 14 T background magnetic field". In: *Superconductor Science and Technology* 33.5 (Apr. 2020). Publisher: IOP Publishing, 05LT03. ISSN: 0953-2048. DOI: 10.1088/1361-6668/ab7fbe. URL: https://dx.doi.org/10.1088/1361-6668/ab7fbe (visited on 01/26/2024).

[34] Matt Landreman et al. "SIMSOPT: A flexible framework for stellarator optimization". In: *Journal of Open Source Software* 6.65 (Sept. 10, 2021), p. 3525. ISSN: 2475-9066. DOI: 10.21105/joss.03525. URL: https://joss.theoj.org/papers/10.21105/joss.03525 (visited on 10/16/2023).

[35] Samuel Lazerson et al. *STELLOPT*. STELLOPT. Princeton Plasma Physics Laboratory (PPPL), Princeton, NJ (United States), May 7, 2020. DOI: 10.11578/dc.20180627.6. URL: https://www.osti.gov/biblio/1617636 (visited on 10/16/2023).

[36] Benjamin Letham et al. *Constrained Bayesian Optimization with Noisy Experiments*. June 26, 2018. DOI: 10.48550/arXiv.1706.07094. arXiv: 1706.07094[cs,stat]. URL: http://arxiv.org/abs/1706.07094 (visited on 03/04/2024).

[37] Shibo Li et al. *Multi-Fidelity Bayesian Optimization via Deep Neural Networks*. Dec. 10, 2020. DOI: 10.48550/arXiv.2007.03117. arXiv: 2007.03117[cs,stat]. URL: http://arxiv.org/abs/2007.03117 (visited on 05/24/2023).

[38] J. Lion et al. "A general stellarator version of the systems code PROCESS". In: *Nuclear Fusion* 61.12 (Oct. 2021). Publisher: IOP Publishing, p. 126021. ISSN: 0029-5515. DOI: 10.1088/1741-4326/ac2dbf. URL: https://dx.doi.org/10.1088/1741-4326/ac2dbf (visited on 03/03/2024).

[39] Wesley J. Maddox et al. *Bayesian Optimization with High-Dimensional Outputs*. Oct. 28, 2021. DOI: 10.48550/arXiv.2106.12997. arXiv: 2106.12997[cs,stat]. URL: http://arxiv.org/abs/2106.12997 (visited on 01/17/2024).

[40] Bertil Matérn. "Some particular models". In: *Spatial Variation*. Ed. by Bertil Matérn. Lecture Notes in Statistics. New York, NY: Springer, 1986, pp. 27–51. ISBN: 978-1-4615-7892-5. DOI: 10.1007/978-1-4615-7892-5_3. URL: https://doi.org/10.1007/978-1-4615-7892-5_3 (visited on 03/04/2024).

[41] William Miner et al. "Use of a genetic algorithm for compact stellarator coil design". In: *Nuclear Fusion - NUCL FUSION* 41 (Sept. 1, 2001). DOI: 10.1088/0029-5515/41/9/307.

[42] J. Močkus. "On bayesian methods for seeking the extremum". In: *Optimization Techniques IFIP Technical Conference Novosibirsk, July 1–7, 1974*. Ed. by G. I. Marchuk. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1975, pp. 400–404. ISBN: 978-3-540-37497-8. DOI: 10.1007/3-540-07165-2_55.

[43] Yoh Nagasaki et al. "Axial Compressive Stress Dependence of Critical Current of REBCO Double-Pancake Coil". In: *IEEE Transactions on Applied Superconductivity* 31.5 (Aug. 2021). Conference Name: IEEE Transactions on Applied Superconductivity, pp. 1–5. ISSN: 1558-2515. DOI: 10.1109/TASC.2021.3070119. URL: https://ieeexplore.ieee.org/document/9392320 (visited on 01/26/2024).

[44] Eugene Perepelkin et al. "Optimization of the Magnetic Field Homogeneity Area for Solenoid Type Magnets". In: *EPJ Web of Conferences* 173 (2018). Ed. by Gh. Adam et al., p. 03018. ISSN: 2100-014X. DOI: 10.1051/epjconf/201817303018. URL: https://www.epj-conferences.org/10.1051/epjconf/201817303018 (visited on 10/26/2023).

[45] Matthias Poloczek, Jialei Wang, and Peter Frazier. "Multi-Information Source Optimization". In: *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc., 2017. URL: https://papers.nips.cc/paper_files/paper/2017/hash/df1f1d20ee86704251795841e6a9405a-Abstract.html (visited on 01/29/2024).

[46] Matthias Poloczek, Jialei Wang, and Peter I. Frazier. *Warm Starting Bayesian Optimization*. Aug. 11, 2016. DOI: 10.48550/arXiv.1608.03585. arXiv: 1608.03585[cs,stat]. URL: http://arxiv.org/abs/1608.03585 (visited on 01/29/2024).

[47] N. Pomphrey et al. "Innovations in compact stellarator coil design". In: *Nuclear Fusion* 41.3 (Mar. 2001), p. 339. ISSN: 0029-5515. DOI: 10.1088/0029-5515/41/3/312. URL: https://dx.doi.org/10.1088/0029-5515/41/3/312 (visited on 10/16/2023).

[48] Jinggang Qin et al. "New design of cable-in-conduit conductor for application in future fusion reactors". In: *Superconductor Science and Technology* 30.11 (Oct. 2017). Publisher: IOP Publishing, p. 115012. ISSN: 0953-2048. DOI: 10.1088/1361-6668/aa8900. URL: https://dx.doi.org/10.1088/1361-6668/aa8900 (visited on 01/26/2024).

[49] V. Queral et al. "Evaluation of metal additive manufacturing for high-field modular-stellarator radial plates and conductors". In: *Nuclear Materials and Energy* 30 (Mar. 1, 2022), p. 101149. ISSN: 2352-1791. DOI: 10.1016/j.nme.2022.101149. URL: https://www.sciencedirect.com/science/article/pii/S2352179122000369 (visited on 01/06/2025).

[50] Vicente Queral et al. "High-Field Ignition-Capable Stellarator i-ASTER: Initial Structural Evaluation". In: *IEEE Transactions on Plasma Science* 48.6 (June 2020), pp. 1842–1848. ISSN: 0093-3813, 1939-9375. DOI: 10.1109/TPS.2020.2972408. URL: https://ieeexplore.ieee.org/document/9007027/ (visited on 10/16/2023).

[51] A. R. Raffray et al. “Engineering challenges in designing an attractive compact stellarator power plant”. In: *Fusion Engineering and Design*. Proceedings of the 24th Symposium on Fusion Technology 82.15 (Oct. 1, 2007), pp. 2696–2704. ISSN: 0920-3796. DOI: 10.1016/j.fusengdes.2007.05.023. URL: https://www.sciencedirect.com/science/article/pii/S0920379607002499 (visited on 10/16/2023).

[52] Nicolò Riva et al. “Development of the first non-planar REBCO stellarator coil using VIPER cable”. In: *Superconductor Science and Technology* 36 (Aug. 6, 2023). DOI: 10.1088/1361-6668/aced9d.

[53] P. Rodriguez-Fernandez, N. T. Howard, and J. Candy. “Nonlinear gyrokinetic predictions of SPARC burning plasma profiles enabled by surrogate modeling”. In: *Nuclear Fusion* 62.7 (May 2022). Publisher: IOP Publishing, p. 076036. ISSN: 0029-5515. DOI: 10.1088/1741-4326/ac64b2. URL: https://dx.doi.org/10.1088/1741-4326/ac64b2 (visited on 01/26/2024).

[54] P. Rodriguez-Fernandez et al. *Enhancing predictive capabilities in fusion burning plasmas through surrogate-based optimization in core transport solvers*. Dec. 19, 2023. DOI: 10.48550/arXiv.2312.12610. arXiv: 2312.12610[physics]. URL: http://arxiv.org/abs/2312.12610 (visited on 01/26/2024).

[55] A. Shimizu et al. “Recent developments in engineering design for the quasi-axisymmetric stellarator CFQS”. In: *Nuclear Fusion* 62.1 (Dec. 2021). Publisher: IOP Publishing, p. 016010. ISSN: 0029-5515. DOI: 10.1088/1741-4326/ac369a. URL: https://dx.doi.org/10.1088/1741-4326/ac369a (visited on 10/16/2023).

[56] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. “Practical Bayesian Optimization of Machine Learning Algorithms”. In: *Advances in Neural Information Processing Systems*. Vol. 25. Curran Associates, Inc., 2012. URL: https://papers.nips.cc/paper_files/paper/2012/hash/05311655a15b75fab86956663e1819cd-Abstract.html (visited on 01/29/2024).

[57] Niranjan Srinivas et al. “Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design”. In: *IEEE Transactions on Information Theory* 58.5 (May 2012), pp. 3250–3265. ISSN: 0018-9448, 1557-9654. DOI: 10.1109/TIT.2011.2182033. arXiv: 0912.3995[cs]. URL: http://arxiv.org/abs/0912.3995 (visited on 03/04/2024).

[58] Dennis J. Strickler, Lee A. Berry, and Steven P. Hirshman. “Designing Coils for Compact Stellarators”. In: *Fusion Science and Technology* 41.2 (Mar. 1, 2002). Publisher: Taylor & Francis _eprint: https://doi.org/10.13182/FST02-A206, pp. 107–115. ISSN: 1536-1055. DOI: 10.13182/FST02-A206. URL: https://doi.org/10.13182/FST02-A206 (visited on 10/16/2023).

[59] *Sweep*. URL: https://doc.comsol.com/5.5/doc/com.comsol.help.comsol/comsol_ref_geometry.14.085.html (visited on 03/04/2024).

[60] Ben Tu et al. “Joint Entropy Search for Multi-Objective Bayesian Optimization”. In: *Advances in Neural Information Processing Systems* 35 (Dec. 6, 2022), pp. 9922–9938. URL: https://proceedings.neurips.cc/paper_files/paper/2022/hash/4086fe59dc3584708468fba0e459f6a7-Abstract-Conference.html (visited on 01/20/2024).

[61] Carlos Roberto Vargas-Llanos et al. "3D homogenization of the T-A formulation for the analysis of coils with complex geometries". In: *Superconductor Science and Technology* 35.12 (Oct. 2022). Publisher: IOP Publishing, p. 124001. ISSN: 0953-2048. DOI: 10.1088/1361-6668/ac9932. URL: https://dx.doi.org/10.1088/1361-6668/ac9932 (visited on 01/18/2024).

[62] N. N. Vazirani et al. "Coupling multi-fidelity xRAGE with machine learning for graded inner shell design optimization in double shell capsules". In: *Physics of Plasmas* 30.6 (June 1, 2023), p. 062704. ISSN: 1070-664X, 1089-7674. DOI: 10.1063/5.0129565. URL: https://pubs.aip.org/pop/article/30/6/062704/2894857/Coupling-multi-fidelity-xRAGE-with-machine (visited on 12/15/2024).

[63] P. Wąchal et al. "Development and mechanical investigation of central solenoid structure for COMPASS-U tokamak". In: *AIP Conference Proceedings* 2239.1 (May 22, 2020), p. 020047. ISSN: 0094-243X. DOI: 10.1063/5.0008077. URL: https://doi.org/10.1063/5.0008077 (visited on 01/26/2024).

[64] Qiuliang Wang. *Practical Design of Magnetostatic Structure Using Numerical Simulation | Wiley*. Wiley.com. URL: https://www.wiley.com/en-us/Practical+Design+of+Magnetostatic+Structure+Using+Numerical+Simulation-p-9781118398142 (visited on 03/14/2023).

[65] James T. Wilson, Frank Hutter, and Marc Peter Deisenroth. *Maximizing acquisition functions for Bayesian optimization*. Dec. 2, 2018. DOI: 10.48550/arXiv.1805.10196. arXiv: 1805.10196[cs,stat]. URL: http://arxiv.org/abs/1805.10196 (visited on 01/17/2024).

[66] Di Wu and G. Gary Wang. "Knowledge-Assisted Optimization for Large-Scale Design Problems: A Review and Proposition". In: *Journal of Mechanical Design* 142.1 (Oct. 3, 2019). ISSN: 1050-0472. DOI: 10.1115/1.4044525. URL: https://doi.org/10.1115/1.4044525 (visited on 04/30/2023).

[67] Jian Wu et al. *Practical Multi-fidelity Bayesian Optimization for Hyperparameter Tuning*. Mar. 11, 2019. arXiv: 1903.04703[cs,math,stat]. URL: http://arxiv.org/abs/1903.04703 (visited on 06/28/2023).

[68] Yueqing Yang et al. "Significant mechanical softening of copper under coupled electric and magnetic stimuli". In: *Scripta Materialia* 231 (July 1, 2023), p. 115438. ISSN: 1359-6462. DOI: 10.1016/j.scriptamat.2023.115438. URL: https://www.sciencedirect.com/science/article/pii/S1359646223001628 (visited on 01/26/2024).

[69] Mohammad Yazdani-Asrami, Alireza Sadeghi, and Wenjuan Song. "Ultra-fast Surrogate Model for Magnetic Field Computation of a Superconducting Magnet Using Multi-layer Artificial Neural Networks". In: *Journal of Superconductivity and Novel Magnetism* 36.2 (Feb. 1, 2023), pp. 575–586. ISSN: 1557-1947. DOI: 10.1007/s10948-022-06479-z. URL: https://doi.org/10.1007/s10948-022-06479-z (visited on 03/14/2023).

[70] Huiming Zhang, Min Zhang, and Weijia Yuan. "An efficient 3D finite element method model based on the T–A formulation for superconducting coated conductors". In: *Superconductor Science and Technology* 30.2 (Dec. 2016). Publisher: IOP Publishing, p. 024005. ISSN: 0953-2048. DOI: 10.1088/1361-6668/30/2/

024005. URL: https://dx.doi.org/10.1088/1361-6668/30/2/024005 (visited on 01/26/2024).

[71] S. J. Zheng et al. "Quench analysis and experiment of CiADS linac superconducting solenoid". In: *IEEE Transactions on Applied Superconductivity* (2024). Conference Name: IEEE Transactions on Applied Superconductivity, pp. 1–6. ISSN: 1558-2515. DOI: 10.1109/TASC.2024.3352519. URL: https://ieeexplore.ieee.org/document/10387740 (visited on 01/26/2024).

[72] Yao Zhou et al. "Approach to nonlinear magnetohydrodynamic simulations in stellarator geometry". In: *Nuclear Fusion* 61.8 (July 2021). Publisher: IOP Publishing, p. 086015. ISSN: 0029-5515. DOI: 10.1088/1741-4326/ac0b35. URL: https://dx.doi.org/10.1088/1741-4326/ac0b35 (visited on 01/26/2024).