Design stellarator coils by a modified Newton method using FOCUS

Caoxiang Zhu, Stuart R. Hudson

December 4, 2017

Abstract

To find the optimal coils for stellarators, nonlinear optimization algorithms are applied in existing coil design codes. However, none of these codes have used the information from the second-order derivatives. In this paper, we present a modified Newton method in the recently developed code FOCUS. The Hessian matrix is calculated with analytically derived equations. Its inverse is approximated by a modified Cholesky factorization and applied in the iterative scheme of a classical Newton method. Using this method, FOCUS is able to recover the W7-X modular coils starting from a simple initial guess. Results demonstrate significant advantages in convergence and robustness.

1 Introduction

Unlike tokamaks, stellarators rely on the external current-carrying coils to produce the rotational-transform required for confining plasmas. As a consequence, stellarators are inherently stable and free of disruptions, which is very attractive for future fusion reactors. However, due to the three-dimensional nature, the coils in stellarators are much more complicated.

There are two main approaches to designing stellarator coils. The first one is to approximate the external coils with a surface current potential on a predefined toroidal "winding surface" surrounding the plasma [?, ?, ?]. The mathematical problem reduces to a least squares minimization problem, and can be solved linearly. This method is fast and elegant, but it lacks direct control on the final coil shapes. To incorporate engineering constraints, other approaches that optimize the geometry of the coil filaments directly were developed [?, ?, ?]. These employ nonlinear optimization algorithms to find the coil shapes that satisfy both physics requirements and engineering constraints, as encapsulated by a "cost-function". The minimization algorithms used in these codes require either only function values, such as the Brent's method [?], or function values together with numerically approximated gradients, like the Levenberg-Marquardt method [?]. Although Hessian-based minimization algorithms are generally more powerful, none of these codes, as yet, have employed any Hessian-based minimization algorithms.

Recently, a new coil design algorithm, named FOCUS, has been implemented [?]. FOCUS describe the coil filaments as arbitrary, closed one-dimensional curves embedded in three-dimensional space. The "physics" and "engineering" cost functions were carefully chosen to be both relevant, and differentiable. The previous version of FOCUS calculated only the first derivatives and gradient-based minimization algorithms were employed. Newton's method is potentially a much faster algorithm, and it has been applied in many codes in the field of plasma physics [?, ?, ?, ?]. In this paper, we will introduce a modified Newton method using the analytically calculated Hessian matrix (comprised of the second derivatives) in FOCUS.

Section II gives a overview of the FOCUS code, and details of the modified Newton method are presented in Section III. In Section IV, an example of optimizing the modular coils of the W7-X stellarator from a simple circular coils initialization is illustrated, and section V gives some conclusions and comments.

2 Overview of the FOCUS code

Coils are approximated by single, closed filaments embedded in three-dimensional space. FOCUS presently uses Fourier series to represent coils in the Cartesian coordinates,

$$x(t) = X_{c,0} + \sum_{n=1}^{N_F} \left[X_{c,n} \cos(nt) + X_{s,n} \sin(nt) \right], \tag{1}$$

where the parameter t varys between $[0, 2\pi]$, and similarly for $y(t)$ and $z(t)$. Each coil is fully determined by $3 \times (2N_F + 1)$ Fourier coefficients, which constitute the independent degrees-of-freedom. The current through each coil can also be considered as a free-parameter.

All the coil parameters are allowed to vary to minimize a so-called "target function" or "cost function" consisting of multiple objective functions

$$\chi^2(\mathbf{X}) = \sum_j w_j \left(\frac{f_j(\mathbf{X}) - f_{j,o}}{f_{j,o}} \right)^2, \tag{2}$$

where $\mathbf{X}$ describes the degrees-of-freedom in the coil geometry and currents, $f_j(\mathbf{X})$ is the j^{th} objective function, and $f_{j,o}$ denotes the desired "target" value, and w_j is a user-prescribed weight. Several objective functions have been implemented in FOCUS, subject to physical requirements and engineering constraints.

The most important requirement of the external coils is that they produce the desired magnetic field. The mathematical theory of vacuum fields show that the magnetic field inside a certain domain is uniquely determined ~~(up to a scalar multiple)~~ by the normal magnetic field at the domain boundary and a prescribed loop integral. The first objective functional is thus

$$f_B(\mathbf{X}) \equiv \int_S \frac{1}{2} \left(\mathbf{B_V} \cdot \mathbf{n} - T_{Bn} \right)^2 \mathrm{d}s. \tag{3}$$

Here, T_{Bn} is the target normal field distribution on an arbitrary predefined boundary S, i.e T_{Bn} is the magnetic field produced by the plasma, and $\mathbf{B_V}$ is the total magnetic field generated by external coils. Normally, the smaller f_B is, the better the coils produce the required field (but so-called resonant field distributions can be particularly problematic, even though they are small).

If there are no plasma currents and the plasma boundary is required to be a flux surfaces, then T_{Bn} is zero. As a consequence, trivial solutions, like when $I_i \to 0 \ \forall i$, $f_B \to 0$, exist. An objective function is defined as

$$f_\Psi(\mathbf{X}) \equiv \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{2} \left(\frac{\Psi_\zeta \ - \ \Psi_o}{\Psi_o} \right)^2 \mathrm{d}\zeta, \tag{4}$$

to produce the target toroidal flux and avoid such trivial solutions.

If only the magnetic field is considered, coils would prefer to be further away from the plasma to reduce the ripple, or to form wiggles to better match the plasma shape. By constructing a penalty on the coil length, we can prevent coils from getting arbitrarily long and forming unrealistic wiggles. The length penalty can be realized either by an exponential form

$$f_L(\mathbf{X}) = \frac{1}{N_C} \sum_{i=1}^{N_C} \frac{e^{L_i}}{e^{L_{i,o}}}, \tag{5}$$

or by a quadratic form

$$f_L(\mathbf{X}) = \frac{1}{N_C} \sum_{i=1}^{N_C} \frac{1}{2} \frac{(L_i - Li, o)^2}{L_{i,o}^2} \ . \tag{6}$$

In Eq.(5) and Eq.(6), N_C is the total number of coils and $L_i(\mathbf{X})$ is the length of i-th coil, while $L_{i,o}$ is a user-specified normalization or the target length.

The above objective functions are easily differentiable functions of the coil geometry, and previously FOCUS used the steepest descent method (and a nonlinear conjugate gradient method [**?**]) to optimize the coil parameters.

3 Modified Newton method for coil optimization

3.1 Analytically calculated Hessian matrix

The second derivatives of the target function with respect to the coil parameters can also be calculated analytically, albeit with a little extra work. For instance, an arbitrary term, $\partial^2 f_B / \partial X_m^2$, is calculated as,

$$\frac{\partial^2 f_B}{\partial X_m^2} = \int_S \left(\frac{\partial \mathbf{B}_V}{\partial X_m} \cdot \mathbf{n} \right)^2 + (\mathbf{B}_V \cdot \mathbf{n} - T_{Bn}) \left(\frac{\partial^2 \mathbf{B}_V}{\partial X_m^2} \cdot \mathbf{n} \right) \ , \tag{7}$$

where X_m is the m-th variable in the vector of degrees of freedom $\mathbf{X}$. The magnetic field $\mathbf{B}_V$ produced by external coils is computed by the Biot-Savart law,

$$\mathbf{B}_V = \frac{\mu_0}{4\pi} \sum_{i=1}^{N_C} I_i \int_{C_i} \frac{\mathrm{d}\mathbf{l}_i \times \mathbf{r}}{r^3}. \tag{8}$$

The displacement vector $\mathbf{r} = \mathbf{x}_0 - \mathbf{x}_i$ is from the evaluation point to the source point on the i-th coil and $\mathrm{d}\mathbf{l}_i = \dot{\mathbf{x}}_i \mathrm{d}t$. $\mathbf{B}_V$ is a functional of the coil geometries $\mathbf{x}(\mathbf{X})$. To calculate the derivatives of $\mathbf{B}_V$ with respect to coil Fourier coefficients, functional derivatives can be applied. For illustration, the first and second derivatives of $\mathbf{B}_V$ with respect to X_m are

$$\frac{\partial \mathbf{B}_V}{\partial X_m} = \int_0^{2\pi} \frac{\delta \mathbf{B}_V}{\delta x} \frac{\partial x}{\partial X_m} + \frac{\delta \mathbf{B}_V}{\delta y} \frac{\partial y}{\partial X_m} + \frac{\delta \mathbf{B}_V}{\delta z} \frac{\partial z}{\partial X_m} \mathrm{d}t \ , \tag{9}$$

$$\frac{\partial^2 \mathbf{B}_V}{\partial X_m^2} = \int_0^{2\pi} \frac{\delta^2 \mathbf{B}_V}{\delta x^2} \left(\frac{\partial x}{\partial X_m} \right)^2 + \frac{\delta^2 \mathbf{B}_V}{\delta x \delta y} \frac{\partial y}{\partial X_m} \frac{\partial x}{\partial X_m} + \frac{\delta^2 \mathbf{B}_V}{\delta x \delta z} \frac{\partial z}{\partial X_m} \frac{\partial x}{\partial X_m} \mathrm{d}t \ , \tag{10}$$

where $\delta \mathbf{B}_V/\delta x$ and $\delta^2 \mathbf{B}_V/\delta x^2$ are the first and second functional derivatives, respectively. For a small change on the geometry of the i-th coil, written as $\delta \mathbf{x}_i$, the first and second variation of the vacuum magnetic field are

$$\delta \mathbf{B_V} = \frac{\mu_0}{4\pi} I_i \int_0^{2\pi} \left[\frac{3\mathbf{r}\cdot\mathbf{x}_i'}{r^5} \mathbf{r}\times\delta\mathbf{x}_i \;+\; \frac{2}{r^3}\delta\mathbf{x}_i\times\mathbf{x}_i' \;+\; \frac{3\mathbf{r}\cdot\delta\mathbf{x}_i}{r^5}\mathbf{x}_i'\times\mathbf{r} \right] \mathrm{d}t, \tag{11}$$

$$\begin{aligned} \delta^2 \mathbf{B}_V = \frac{\mu_0}{4\pi} I_i \int_0^{2\pi} & \frac{3\left[\mathbf{r}\cdot\delta^2\mathbf{x} + 5(\mathbf{r}\cdot\delta\mathbf{x})^2 - \delta\mathbf{x}\cdot\delta\mathbf{x}\right]}{r^5}(\mathbf{x}'\times\mathbf{r}) \\ & + \frac{3\left[\mathbf{r}\cdot\delta\mathbf{x}' + 5(\mathbf{r}\cdot\mathbf{x}')(\mathbf{r}\cdot\delta\mathbf{x}) - \delta\mathbf{x}\cdot\mathbf{x}'\right]}{r^5}(\mathbf{r}\times\delta\mathbf{x}) \\ & + \frac{9\mathbf{r}\cdot\delta\mathbf{x}}{r^5}(\delta\mathbf{x}\times\mathbf{x}') + \frac{3\mathbf{r}\cdot\delta\mathbf{x}}{r^5}(\delta\mathbf{x}'\times\mathbf{r}) \\ & + \frac{3\mathbf{r}\cdot\mathbf{x}}{r^5}(\mathbf{r}\times\delta^2\mathbf{x}) + \frac{2}{r^3}(\delta^2\mathbf{x}\times\mathbf{x}') + \frac{2}{r^3}(\delta\mathbf{x}\times\delta\mathbf{x}'). \end{aligned} \tag{12}$$

In the Cartesian coordinates, the functional derivatives are now written as,

$$\frac{\delta B_V^x}{\delta x} = \frac{3\Delta x\left(\Delta z y' - \Delta y z'\right)}{r^5}\ , \tag{13}$$

$$\frac{\delta^2 B^x{}_V}{\delta x^2} = \frac{3\left(5\Delta x^2 - r^2\right)\left(\Delta z y' - \Delta y z'\right)}{r^7}\ , \tag{14}$$

where $\mathbf{r} = (\Delta x, \Delta y, \Delta z)$, $y' = \mathrm{d}y/\mathrm{d}t$ and $z' = \mathrm{d}z/\mathrm{d}t$. Only the x components are listed here and other components can be calculated likewise.

Inserting Eq.(8), Eq.(9), Eq.(10), Eq.(13) and Eq.(14) into Eq.(7), we can calculate $\partial^2 f_B/\partial X_m^2$ fast and accurately. Similarly, the whole Hessian matrix, $\mathbf{H}$, can be computed.

3.2 Modified Newton method

The iterative scheme of a classical Newton method is [**?**]

$$\mathbf{X}_{k+1} = \mathbf{X}_k - \mathbf{H}_k^{-1}\cdot\mathbf{g}_k\ . \tag{15}$$

Here, $\mathbf{H}_k^{-1}$ is the inverse of Hessian and $\mathbf{g}_k$ is the gradient at the $k-th$ iteration.

As yet, it is not guaranteed that the Hessian has a well-defined inverse. There are two problems that may be encountered. The first is purely numerical. The representation for the coils is not unique in the sense that the Fourier harmonics of the coils can change, but with no change in the coil geometry. This amounts to a tangential, reparameterization of the coils. The second problem is "physical", or geometrical. The physical solution may not be unique. Consider for example designing a set of coils for an axisymmetric tokamak. Because of the rotational symmetry, the geometry of the coils can be rotated without affecting the magnitude of the objective functions.

Many techniques have been proposed to deal with such problems. Among them, one of the most practical methods is a modified Cholesky factorization [**?**]. For an arbitrary real symmetric matrix, like the Hessian matrix $\mathbf{H}$ in our problem, it would perform a standard Cholesky decomposition on the new matrix $\mathbf{H}^* = \mathbf{H} + \mathbf{E}$, in which $\mathbf{E}$ is a carefully chosen diagonal matrix such that $\mathbf{H}^*$ is positive definite and well-conditioned ($\mathbf{E} = 0$ if $\mathbf{H}$ is already positive definite). Now the inverse of the Hessian is approximated with $(\mathbf{H}^*)^{-1}$,

$$(\mathbf{H}^*)^{\mathbf{-1}} = \mathbf{P^T}\ (\mathbf{L^{-1}})^{\mathbf{T}}\ \mathbf{L^{-1}P}\ . \tag{16}$$

$\mathbf{L}$ is the lower triangular matrix from the standard Cholesky factorization and $\mathbf{P}$ denotes the permutation matrix.

To ensure a descent direction, a step length α_k along the new search direction $\mathbf{p}_k = -\,(\mathbf{H}^*)^{-1}\mathbf{g}_k$ is selected to satisfy the strong Wolfe condition [?, ?],

$$\chi^2(\mathbf{X}_{k+1}) \leq \chi^2(\mathbf{X}_k) + c_1\alpha_k\mathbf{g}_k^T\mathbf{p}_k\ , \tag{17}$$

$$|\mathbf{g}_{k+1}^T\mathbf{p}_k| \leq c_2|\mathbf{g}_k^T\mathbf{p}_k|\ , \tag{18}$$

where parameters c_1 and c_2 satisfy $0 < c_1 < c_2 < 1$. Finally, the actual iterative scheme applied in the code is

$$\mathbf{X}_{k+1} = \mathbf{X}_k - \alpha_k(\mathbf{H}_k^*)^{-1}\mathbf{g}_k\ , \tag{19}$$

and we call this the modified Newton method in FOCUS.

4 Application for designing W7-X modular coils

For demonstration, we use the modified Newton method to recover the design the W7-X [?] modular coils. The W7-X "standard configuration" [?] was optimized to have a rotational-transform ($\iota\!\!\!-$) equal to 1.0 at the last closed flux surface (LCFS), with a 5/5 island chain outside the LCFS. This configuration is achived by using 50 modular (nonplanar) coils arranged in five field periods. In each period, the modular coils are stellarator symmetric [?], and consequently there are only five unique modular coils. The modular coils were obtained from an extension of the NESCOIL code [?], which systematically adjusted the Fourier coefficients of the winding surface to satisfy multiple criteria [?].

To use FOCUS, a boundary surface, S, that is close to the LCFS, together with T_{Bn} on S as calculated from the actual coils, is provided. Note that these "come as a pair", so to speak. If a slightly different boundary is provided, a slightly different T_{Bn} would result, but the magnetic field inside the boundaries will be identical. The "coils" used in this calculation are the filamentary models of the actually built coils, represented in Fourier series with $N_F = 8$. The modified Newton method in FOCUS requires a suitable initial guess for the coil geometry, and this is obtained by placing 50 circular coils at equal toroidal intervals surrounding the boundary. The radius of the initial coils is 1.25m. The shape of the selected boundary, S, and the initial coils are shown in Fig.1. The differences between the initial normal field B_n, which is produced the circular coils, and the target distribution T_{Bn} are illustrated on the boundary. The maximum value of $(B_n - T_{Bn})$ is 1.30T.

Besides the constraint on the normal magnetic field, a length penalty in the quadratic form, as decribed in Eq.(6), is also applied. The target length of each coil, $L_{i,o}$, is set to the length of the actual coil that is at the closest toroidal angle. Weights for f_B and f_L are chosen to satisfy $w_B = 200/f_{B0}$ and $w_L = 0.1/f_{L0}$, where f_{B0} and f_{L0} are the initial values of f_B and f_L respectively. Not only do Newton methods require an initial guess, they require a good initial guess; if a bad initial guess is provided, Newton methods can be very unstable. The mathematical problem of

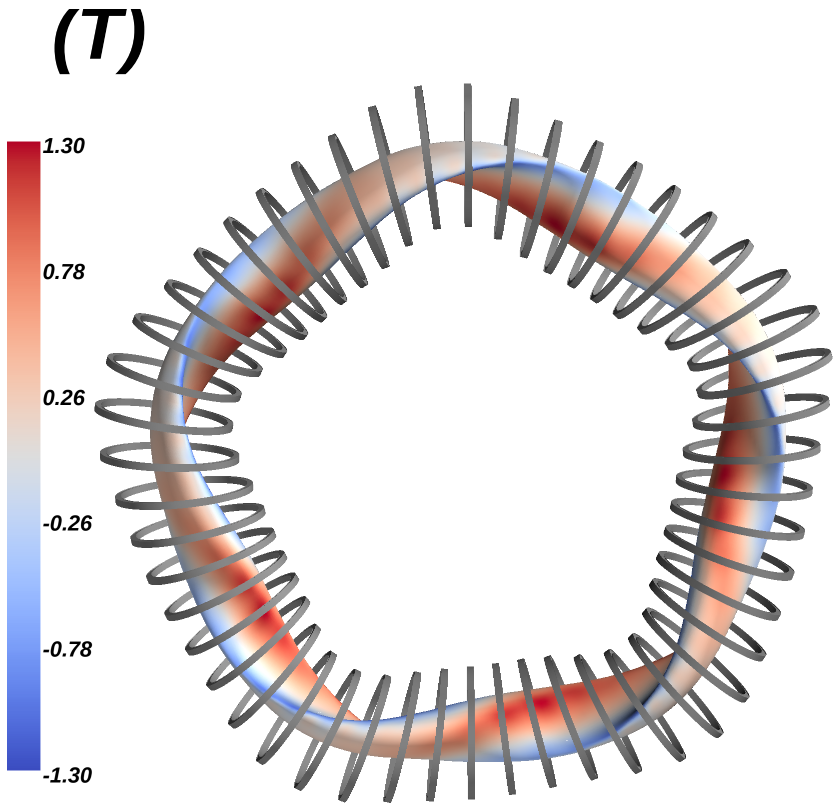

Figure 1: The input plasma boundary and initial coils for FOCUS. The colors on the boundary indicate the magnitude of $(B_n - T_{Bn})$. The maximum value is about 1.30 T. The single-filamentary coils are plotted with finite width for better viewing, same hereafter.

coil design is very nonlinear, and this fact combined with the very free coil representation used by FOCUS suggests that it is advisable to begin with a more stable method. So, for this calculation the optimization starts with 50 iterations of the nonlinear conjugate gradient method (CG). After a closer estimate is obtained, 50 iterations of the modified Newton method (MN) are employed.

Fig.2 shows the evolutions of χ^2, f_B and f_L with respect to the accumulated iterations. In the first 10 iterations, the objective functions are decreased rapidly by CG. Then the rate of descent slows down until the beginning of MN. Applying the MN method makes the objective functions keep decreasing again, at a considerable rate. Quantitatively, f_B is reduced from 1.71×10^1 to 7.02×10^{-3} by CG, and finally to 2.24×10^{-6} after MN. Similarly, f_L is 2.91×10^{-3} at the beginning, 9.15×10^{-5} after CG and 1.13×10^{-8} at the end of MN. It should be noted that the evolution curves are not completely flat at the end of MN, which means the objective functions would be reduced more if the modified Newton method was continued. However, the objective functions are already small enough and there wouldn't be significant reductions. (In reality, the coils have finite thickness, and this places a limit on the how accurately it is required to design filamentary coils.)

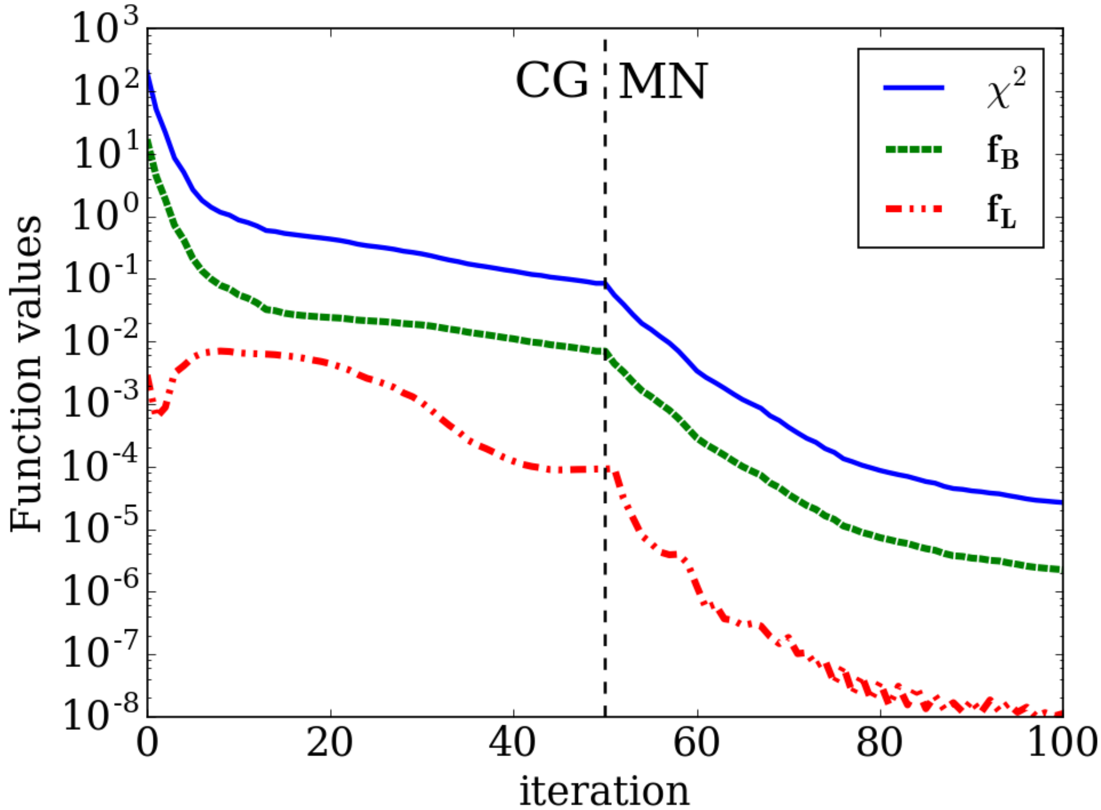

Figure 2: Convergence curves of the objective functions over iterations. The first 50 iterations are using the CG method and the latter is MN.

Fig.3 illustrates the geometries of the optimized coils resulting from the FOCUS calculation and the actual W7-X coils. The two coils sets almost perfectly coincide with each other.

In addition to computing the magnitude of the objective functionals and their gradients with respect to the coil geometry, there are many criteria that can be used for evaluating the quality of the magnetic field produced by the "FOCUS coils". We compare the Poincaré plots of the flux surfaces and the rotational-transform profiles. At the bean-shaped cross-section, 64 points that are linearly interpolated between $(R_l, Z_l) = (5.95, 0.00)$ and $(R_u, Z_u) = (6.30, 0.00)$ are chosen as starting points for fieldline tracing calculations. The intersection points are computed by following the fieldlines starting from the starting points with 2000 periods/iterations. As shown in Fig.4, inside the boundary, the flux surfaces produced by the two coils sets are almost identical. A quantitative comparison on the rotational-transform profiles is carried out, as shown in Fig.5. The average relative deviation of the rotational-transforms on the flux surfaces inside the boundary is 0.028% , and the maximum deviation is 0.071%. The rotational-transform profiles outside the

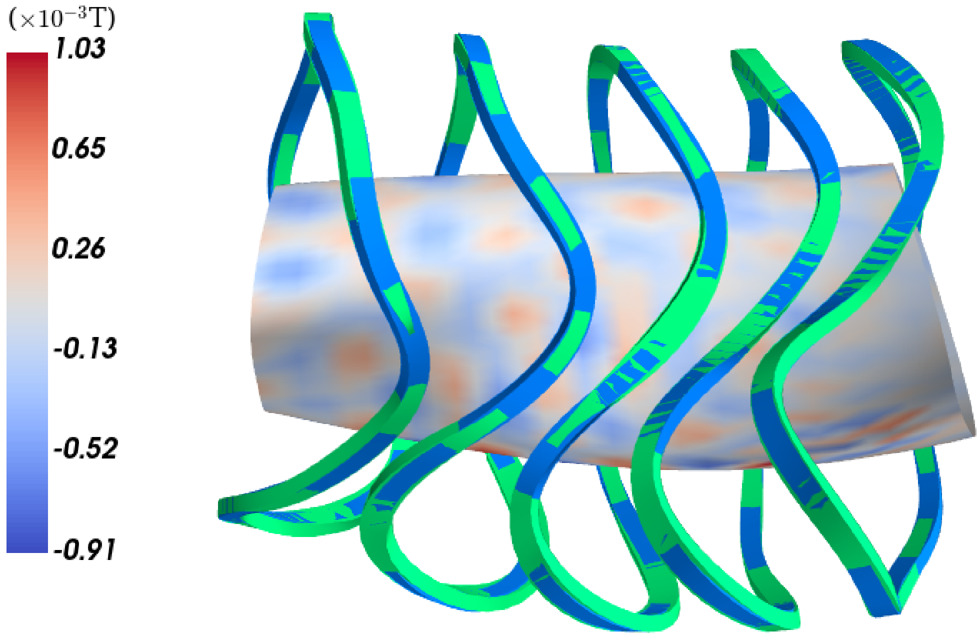

Figure 3: Comparison of FOCUS optimized coils (blue) and the actual coils (green). Only half period is plotted here. The colors on the boundary indicate the distribution of $(B_n - T_{Bn})$. The maximum value of normal field error is now reduced to 1.03×10^{-3}T.

boundary are also coincident, with a relatively high precision, and the 5/5 islands are reproduced.

5 Summary and discussions

In this paper, a modified Newton method has been proposed for designing stellarator coils. To exploit the power of Newton methods, we have overcome the following difficulties: (i) By employing an explicit coil representation and constructing differentiable cost functions, the FOCUS code can compute the first and second derivatives analytically. (ii) A modified Cholesky factorization is applied to quickly invert the Hessian matrix, which is not necessarily invertible. (iii) Before using the modified Newton method, we have available the previously implemented steepest descent and/or the conjugate gradient method to obtain a sufficient initial guess. The modified Newton method has been applied to design modular coils for W7-X. We started from an arbitrary initial guess of circular coils and only provided the plasma boundary, the target B_n on it and the target lengths. By using a combination of CG and MN methods, FOCUS successfully and rapidly recovered the actual coils with remarkable precisions. Moreover, it shows that the modified Newton method is of great robustness, considering that initial circular coils are relatively "bad" guesses and the informations needed are poor. Neither any winding surfaces nor other engineering constraints except the length penalty on the coils were given.

More constraints, such as the spectral condensation on the Fourier coefficients [?], could be implemented to avoid singularities in the Hessian matrix so that the exact inverse could be solved by a direct algebra technique. Global minimization algorithms [?, ?, ?] could also be used to find the global minimum.

6 Acknowledgments

This work was supported by the Chinese Scholarship Council (CSC) with No. 201506340040 and U.S. Department of Energy with under DE-AC02-09CH11466. The authors gracefully acknowledge the fruitful discussions with S. Lazerson, J. Breslau, D. Gates, N. Pomphrey from PPPL and M. Landreman from the University of Maryland. Special appreciations are expressed to J. Geiger and J. Loizu from IPP for providing the W7-X data. The views and opinions expressed herein do not necessarily reflect those of the Max Planck Institute for Plasma Physics. The author C.Z. would also like to thank the PPPL Theory Department for hosting his visiting with this work.

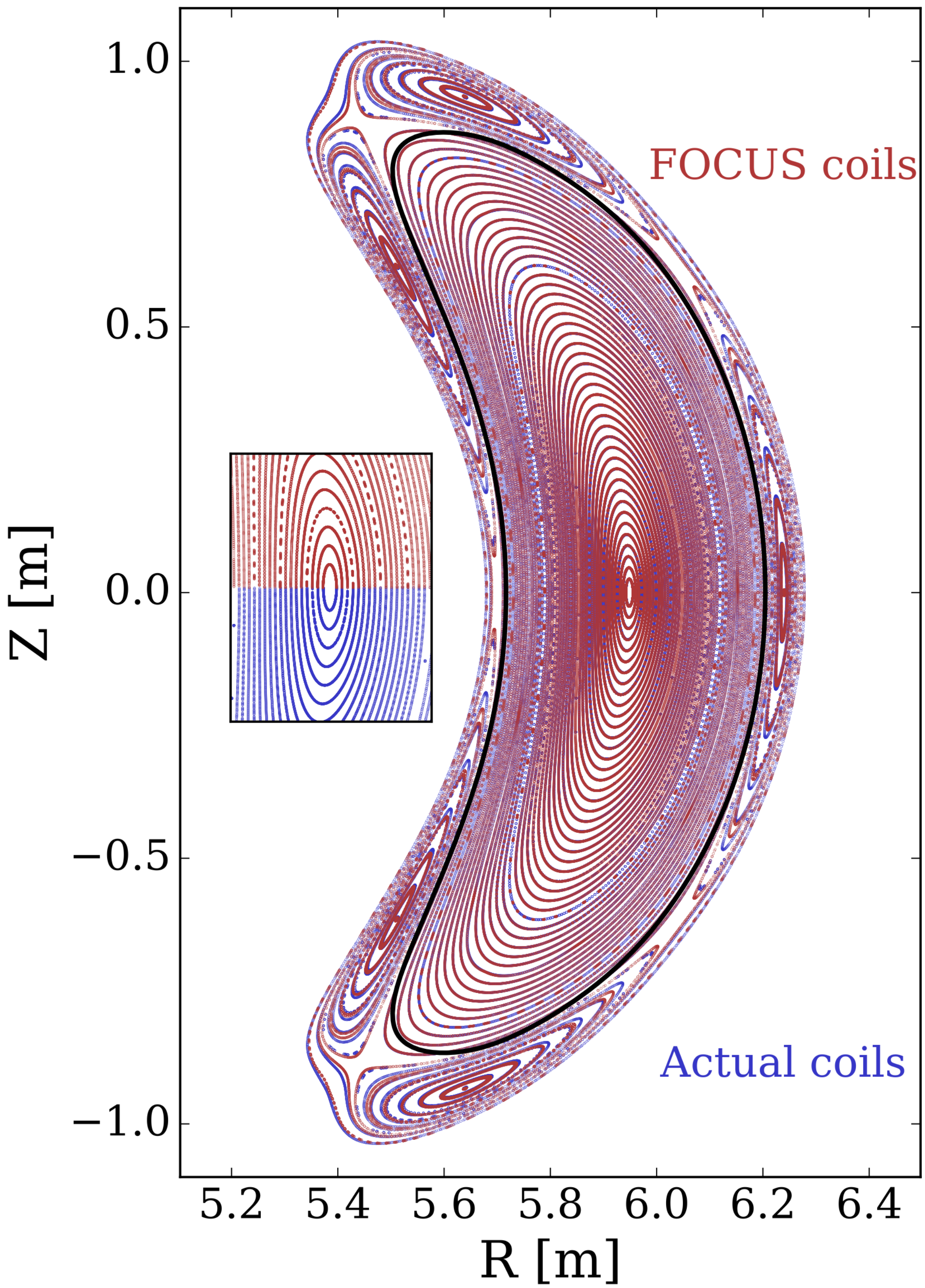

Figure 4: Poincaré plots of the flux surfaces at the bean-shaped cross-section. The data from the magnetic field generated by the FOCUS optimized coils (FOCUS coils) is in red while the actual coils in blue. The solid line (black) is the input plasma boundary. In the zoomed subfigure, the upper half is the flux surfaces from the FOCUS coils and the lower half is from the actual coils.

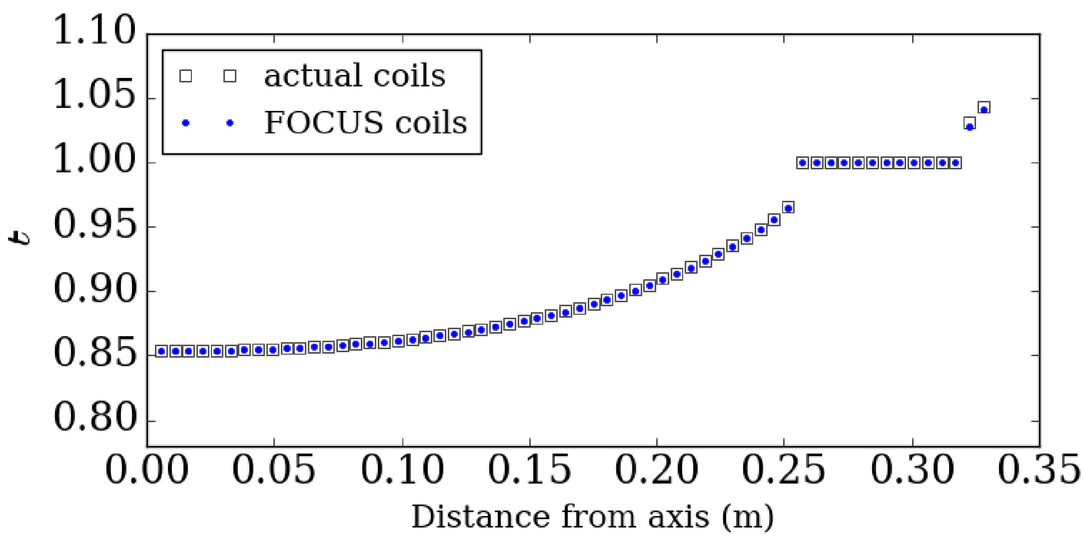

Figure 5: Rotational-transforms on the flux surfaces of the vacuum field produced by the FOCUS coils and the actual coils. The abscissa is the distance from the magnetic axis.