Figures of merit for stellarators near the magnetic axis

Matt Landreman [iD],†

Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, MD 20742, USA

A new paradigm for rapid stellarator configuration design has been recently demonstrated, in which the shapes of quasisymmetric or omnigenous flux surfaces are computed directly using an expansion in small distance from the magnetic axis. To further develop this approach, here we derive several other quantities of interest that can be rapidly computed from this near-axis expansion. First, the ∇B and $\nabla \nabla B$ tensors are computed, which can be used for direct derivative-based optimization of electromagnetic coil shapes to achieve the desired magnetic configuration. Moreover, if the norm of these tensors is large compared with the field strength for a given magnetic field, the field must have a short length scale, suggesting it may be hard to produce with coils that are suitably far away. Second, we evaluate the minor radius at which the flux surface shapes would become singular, providing a lower bound on the achievable aspect ratio. This bound is also shown to be related to an equilibrium beta limit. Finally, for configurations that are constructed to achieve a desired magnetic field strength to first order in the expansion, we compute the error field that arises due to second-order terms.

Key words: fusion plasma, plasma confinement

---

1. Introduction

The geometry of a stellarator's magnetic field needs to be designed to achieve good confinement, magnetohydrodynamic (MHD) stability and other desirable properties. In recent stellarator experiments such as HSX and W7-X, this design was performed by optimizing an objective function computed from numerical MHD equilibrium solutions. This approach suffers from the fact that the optimization algorithm can get trapped in local minima of the objective function, and relatively little insight is gained into the space of solutions. To address these issues, we have recently developed a complementary approach based on a high-aspect-ratio approximation (Landreman & Sengupta 2018; Landreman 2019; Landreman & Sengupta 2019; Landreman, Sengupta & Plunk 2019; Plunk, Landreman & Helander 2019; Jorge, Sengupta & Landreman 2020*a*,*b*). In this complementary approach, the three-dimensional shape of magnetic surfaces are directly constructed to achieve a desired magnetic field strength in Boozer coordinates. This 'near-axis construction' method is based on an expansion devised by Garren & Boozer (1991*a*,*b*). The reduced equations can be solved many orders of magnitude faster than

† Email address for correspondence: mattland@umd.edu

the three-dimensional MHD equilibrium solution required at each iteration of traditional stellarator optimization.

Past work on the near-axis construction has focused on neoclassical confinement, through the properties of quasisymmetry and omnigenity (Landreman & Sengupta 2019; Landreman *et al.* 2019; Plunk *et al.* 2019). Recently it has been shown that within this approach, magnetic well and Mercier stability can be computed (Landreman & Jorge 2020). One can also evaluate the geometric quantities entering the gyrokinetic equation for microinstabilities and turbulence (Jorge & Landreman 2021). In the present paper, we show how a number of other useful quantities can be computed directly from a solution of the near-axis equations. In future work, these figures of merit could be targeted during optimization within the space of near-axis solutions. Such an optimization would have the advantage that the objective function could be evaluated orders of magnitude faster than in traditional stellarator optimization based on full three-dimensional equilibrium calculations. The figures of merit derived here could also be applied to high-resolution scans over parameter space, which are made possible by the speed of the near-axis approach. Specifically, during such a scan over near-axis configuration parameters, the figures of merit here could be used to identify regions of parameter space that are uninteresting due to the need for close electromagnetic coils, unreasonably high aspect ratio, or large errors in quasisymmetry. Attention could then be focused on the remaining regions of parameter space.

The first quantities we will evaluate, in § 3, are the tensors ∇B and $\nabla \nabla B$ along the magnetic axis, where B is the magnetic field. These tensors are useful for two reasons. First, electromagnetic coils can be designed to produce a field matching these tensors, which (for vacuum fields) means they produce the desired stellarator configuration (Giuliani *et al.* 2020). Second, these tensors encode all the scale lengths in the magnetic field (up to the order of interest), reflecting how far away the coils may be from the axis. Configurations with short scale lengths in B are likely to require very close coils, which is undesirable. Next, in § 4, we compute the maximum minor radius for which the constructed surface shapes are smooth and nested. This critical minor radius is equivalent to a minimum aspect ratio, beyond which the near-axis construction does not give physical surface shapes. Configurations for which this minimum aspect ratio is large can then be rejected. We also show in § 4.4 that this limit on the aspect ratio is related to an equilibrium limit on β, the ratio of plasma to magnetic pressure. Finally, in § 5, we consider the case in which the surface shapes are constructed to give some desired pattern of B in Boozer coordinates to first order (in inverse aspect ratio). The 'error field' associated with the B at next order is computed. This information could enable the parameters of the near-axis equations to be optimized so that this error field is small.

All numerical calculations shown in this paper were carried out using the code in Landreman (2020*a*), with data available in Landreman (2020*b*).

2. Notation

We employ the 'inverse expansion' of Garren & Boozer (1991*a*,*b*), while using identical notation to Landreman & Sengupta (2019), hereafter denoted LS. Several key definitions are repeated here for convenience. In Boozer coordinates, the magnetic field has the forms

$$
\begin{aligned}
\boldsymbol{B} &= \nabla \psi \times \nabla \theta + \iota \nabla \varphi \times \nabla \psi \\
&= \beta \nabla \psi + I \nabla \theta + G \nabla \varphi,
\end{aligned}
\tag{2.1}
$$

where θ and φ are the poloidal and toroidal Boozer angles, $2\pi\psi$ is the toroidal flux, I and G are constant on ψ surfaces and $\iota(\psi)$ is the rotational transform. To facilitate

calculations with quasihelical symmetry, it is convenient to introduce a helical angle $\vartheta = \theta - N\varphi$, where N is a constant integer that can be set to zero if not considering quasihelical symmetry. Then

$$\boldsymbol{B} = \nabla\psi \times \nabla\vartheta + \iota_N \nabla\varphi \times \nabla\psi \qquad (2.2)$$

$$= \beta\nabla\psi + I\nabla\vartheta + (G + NI)\nabla\varphi, \qquad (2.3)$$

where $\iota_N = \iota - N$.

The position vector $\boldsymbol{r}$ at an arbitrary point can be written as

$$\boldsymbol{r}(r, \vartheta, \varphi) = \boldsymbol{r}_0(\varphi) + X(r, \vartheta, \varphi)\boldsymbol{n}(\varphi) + Y(r, \vartheta, \varphi)\boldsymbol{b}(\varphi) + Z(r, \vartheta, \varphi)\boldsymbol{t}(\varphi), \qquad (2.4)$$

where $\boldsymbol{r}_0(\varphi)$ is the position vector along the magnetic axis. The Frenet–Serret unit vectors of the axis $(\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b})$ satisfy

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\ell}\frac{\mathrm{d}\boldsymbol{r}_0}{\mathrm{d}\varphi} = \boldsymbol{t}, \quad \frac{\mathrm{d}\varphi}{\mathrm{d}\ell}\frac{\mathrm{d}\boldsymbol{t}}{\mathrm{d}\varphi} = \kappa\boldsymbol{n}, \quad \frac{\mathrm{d}\varphi}{\mathrm{d}\ell}\frac{\mathrm{d}\boldsymbol{n}}{\mathrm{d}\varphi} = -\kappa\boldsymbol{t} + \tau\boldsymbol{b}, \quad \frac{\mathrm{d}\varphi}{\mathrm{d}\ell}\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}\varphi} = -\tau\boldsymbol{n} \qquad (2.5a\text{--}d)$$

and $\boldsymbol{t} \times \boldsymbol{n} = \boldsymbol{b}$. Here, $\kappa(\varphi)$ is the axis curvature, $\tau(\varphi)$ is the axis torsion and ℓ is the arclength along the axis.

The expansion about the magnetic axis is developed by writing

$$X(r, \vartheta, \varphi) = rX_1(\vartheta, \varphi) + r^2 X_2(\vartheta, \varphi) + r^3 X_3(\vartheta, \varphi) + \cdots, \qquad (2.6)$$

where the effective minor radius coordinate r is defined via $2\pi\psi = \pi r^2 \bar{B}$, and where the constant $\bar{B}$ is a reference field strength. The quantities Y and Z in (2.4) are expanded analogously. All scale lengths in the system except r are ordered as comparable to a large length scale $\mathcal{R}$, from which we define the expansion parameter $\epsilon = r/\mathcal{R}$. Then (2.6) represents an expansion in $\epsilon \ll 1$. We expand B and β in a similar way to (2.6) but with an $O(\epsilon^0)$ term as follows:

$$B(r, \vartheta, \varphi) = B_0(\varphi) + rB_1(\vartheta, \varphi) + r^2 B_2(\vartheta, \varphi) + r^3 B_3(\vartheta, \varphi) + \cdots. \qquad (2.7)$$

The radial profiles $G(r)$, $I(r)$, $p(r)$ and $\iota_N(r)$ have expansions with only even powers of r, since they must be even in r:

$$G(r) = G_0 + r^2 G_2 + r^4 G_4 + \cdots. \qquad (2.8)$$

In the analogous expansion of $I(r)$, $I_0 = 0$ since $I(r)$ is proportional to the toroidal current inside the surface r. Considering that physical quantities should be analytic at the magnetic axis, the expansion coefficients have the form

$$\left.\begin{aligned}
X_1(\vartheta, \varphi) &= X_{1s}(\varphi)\sin(\vartheta) + X_{1c}(\varphi)\cos(\vartheta), \\
X_2(\vartheta, \varphi) &= X_{20}(\varphi) + X_{2s}(\varphi)\sin(2\vartheta) + X_{2c}(\varphi)\cos(2\vartheta), \\
X_3(\vartheta, \varphi) &= X_{3s3}(\varphi)\sin(3\vartheta) + X_{3s1}(\varphi)\sin(\vartheta) \\
&\quad + X_{3c3}(\varphi)\cos(3\vartheta) + X_{3c1}(\varphi)\cos(\vartheta).
\end{aligned}\right\} \qquad (2.9)$$

(A detailed discussion is given in appendix A of Landreman & Sengupta (2018).) The quantities Y, Z, B and β have expansion coefficients of the same form as (2.9).

Next, we apply the dual relations, $\nabla r = (\partial\boldsymbol{r}/\partial\vartheta \times \partial\boldsymbol{r}/\partial\varphi)/\sqrt{g}$ and cyclic permutations, where $\sqrt{g} = (\nabla r \cdot \nabla\vartheta \times \nabla\varphi)^{-1} = (G + \iota I)r\bar{B}/B^2$ is the Jacobian of the

(r, ϑ, φ) coordinates. Derivatives of the position vector (2.4) are substituted into the dual relations to obtain the vectors ∇r, $\nabla \vartheta$ and $\nabla \varphi$. These gradient vectors are then substituted into (2.2) = (2.3) and $(\nabla \times \boldsymbol{B}) \times \boldsymbol{B} = \mu_0 \nabla p$. The condition $B(r, \vartheta, \varphi) = B(r, \vartheta)$ can also be imposed if one desires quasisymmetry. Equations are thereby obtained at each order in $\epsilon \ll 1$. These equations can be found in the appendix of Garren & Boozer (1991*a*) and appendix A of LS.

We note two signs that appear in the analysis: $s_G = \text{sgn}(G) = \pm 1$ and $s_\psi = \text{sgn}(\psi) = \text{sgn}(\bar{B}) = \pm 1$. As discussed in appendix A.1 of Landreman & Jorge (2020), these signs are associated with the choice of φ and θ: each sign can be flipped by reversing the direction in which these angles increase. If one is considering quasisymmetry, the normalizing field can be taken to be $\bar{B} = s_\psi B_0$.

With the notation and expansion now defined, we proceed to derive the new figures of merit that can be computed from these expansion coefficients.

3. The ∇B and $\nabla \nabla B$ tensors

There are several reasons why it is valuable to evaluate the tensors ∇B and $\nabla \nabla B$ on the magnetic axis in terms of the near-axis expansion coefficients. First, consistency between these tensors and the magnetic field from the Biot–Savart formula can be achieved by numerical optimization, enabling direct optimization of coil shapes for quasisymmetry. Demonstrations of this method for $O(\epsilon)$ quasisymmetry are presented in Giuliani *et al.* (2020), in which the formula for ∇B derived here is used. It may be possible to optimize coils for $O(\epsilon^2)$ quasisymmetry (requiring $\nabla \nabla B$) in future work.

Second, ∇B and $\nabla \nabla B$ provide information about how feasible it is to generate a given near-axis stellarator solution using distant magnets. This is because ∇B encodes all the (inverse) scale lengths in the magnetic field that can be known from a $O(\epsilon^1)$ solution, and $\nabla \nabla B$ encodes all the additional (inverse) scale lengths in the magnetic field that can be known from a $O(\epsilon^2)$ solution. It is unlikely that magnets can be much farther from the axis than any of these scale lengths. This is because in magnetostatics the field from a small-scale current structure decays rapidly as one moves away from the current; for example in slab geometry a steady sheet current $\propto \sin kx$ with wavenumber k on the $z = 0$ plane produces a vacuum field $\propto \exp(-|kz|)$ that is exponentially small beyond distances $|z| > 1/|k|$. Note that these scale lengths associated with ∇B and $\nabla \nabla B$ are not directly related to scale lengths in magnetic surface shapes: surface shapes near an X-point have a radius of curvature that shrinks to zero, but the scale lengths in $\boldsymbol{B}$ remain non-zero on a separatrix. We therefore expect scale lengths derived from ∇B and $\nabla \nabla B$ to be better indicators of the distance to a coil than scale lengths in the surface shape. 'Worst-case' scale lengths can be constructed by summing the squares of the matrix elements in these tensors, as in

$$L_{\nabla B} = B \sqrt{2/||\nabla B||^2}, \tag{3.1}$$

$$L_{\nabla \nabla B} = \sqrt{4B/||\nabla \nabla B||}, \tag{3.2}$$

for the (squared) Frobenius norm $||\nabla B||^2 = \sum_{i,j=1}^{3} (\nabla B)_{i,j}^2$ and $||\nabla \nabla B||^2 = \sum_{i,j,k=1}^{3} (\nabla \nabla B)_{i,j,k}^2$. These expressions are motivated at the end of this section by considering an infinite straight wire, with $L_{\nabla B}$ and $L_{\nabla \nabla B}$ both constructed to give the distance from the wire. Since it is important in stellarators to maximize the plasma-to-coil distance, we expect that configurations with small $L_{\nabla B}$ or $L_{\nabla \nabla B}$ can be excluded as impractical. Both this application and the application in Giuliani *et al.* (2020) motivate a calculation of ∇B and $\nabla \nabla B$ in terms of the Garren–Boozer expansion.

To proceed, we first evaluate the vector $\boldsymbol{B}$ near the axis. This can be done using (2.2) in the form

$$\boldsymbol{B} = \frac{B^2}{G + \iota I} \left[\left(\Lambda + \iota_N \frac{\partial Z}{\partial \vartheta} \right) \boldsymbol{t} + \left(\Xi + \iota_N \frac{\partial X}{\partial \vartheta} \right) \boldsymbol{n} + \left(\Upsilon + \iota_N \frac{\partial Y}{\partial \vartheta} \right) \boldsymbol{b} \right], \qquad (3.3)$$

where

$$\Lambda = \ell' + \frac{\partial Z}{\partial \varphi} - X\ell'\kappa, \quad \Xi = \frac{\partial X}{\partial \varphi} - Y\ell'\tau + Z\ell'\kappa, \quad \Upsilon = \frac{\partial Y}{\partial \varphi} + X\ell'\tau, \qquad (3.4a\text{--}c)$$

following notation from Garren & Boozer (1991*b*), with $\ell' = |\partial \boldsymbol{r}_0/\partial \varphi|$. Notice $X \sim Y \sim \epsilon^1$, $Z \sim \epsilon^2$, $\Lambda \sim \epsilon^0$ and $\Xi \sim \Upsilon \sim \epsilon^1$. Writing $\boldsymbol{B} = \boldsymbol{B}_0 + r\boldsymbol{B}_1 + r^2 \boldsymbol{B}_2 + \cdots$, and noting $\ell' = |G_0|/B_0$, the leading term in (3.3) is

$$\boldsymbol{B}_0 = s_G B_0 \boldsymbol{t}. \qquad (3.5)$$

The terms of next order in (3.3) give

$$\boldsymbol{B}_1 = \frac{B_0^2}{G_0} \left[X_1 \ell'\kappa \boldsymbol{t} + \left(\frac{\partial X_1}{\partial \varphi} - Y_1 \ell'\tau + \iota_{N0} \frac{\partial X_1}{\partial \vartheta} \right) \boldsymbol{n} + \left(\frac{\partial Y_1}{\partial \varphi} + X_1 \ell'\tau + \iota_{N0} \frac{\partial Y_1}{\partial \vartheta} \right) \boldsymbol{b} \right], \qquad (3.6)$$

where $B_1 = \kappa B_0 X_1$ has been employed.

We next use the dual relations to write

$$\nabla \boldsymbol{B} = (\nabla r) \frac{\partial \boldsymbol{B}}{\partial r} + (\nabla \vartheta) \frac{\partial \boldsymbol{B}}{\partial \vartheta} + (\nabla \varphi) \frac{\partial \boldsymbol{B}}{\partial \varphi}$$

$$= \frac{B^2}{(G + \iota I) r \bar{B}} \left[\left(\frac{\partial X}{\partial \vartheta} \boldsymbol{n} + \frac{\partial Y}{\partial \vartheta} \boldsymbol{b} + \frac{\partial Z}{\partial \vartheta} \boldsymbol{t} \right) \times (\Lambda \boldsymbol{t} + \Xi \boldsymbol{n} + \Upsilon \boldsymbol{b}) \frac{\partial \boldsymbol{B}}{\partial r} \right.$$

$$+ (\Lambda \boldsymbol{t} + \Xi \boldsymbol{n} + \Upsilon \boldsymbol{b}) \times \left(\frac{\partial X}{\partial r} \boldsymbol{n} + \frac{\partial Y}{\partial r} \boldsymbol{b} + \frac{\partial Z}{\partial r} \boldsymbol{t} \right) \frac{\partial \boldsymbol{B}}{\partial \vartheta}$$

$$\left. + \left(\frac{\partial X}{\partial r} \boldsymbol{n} + \frac{\partial Y}{\partial r} \boldsymbol{b} + \frac{\partial Z}{\partial r} \boldsymbol{t} \right) \times \left(\frac{\partial X}{\partial \vartheta} \boldsymbol{n} + \frac{\partial Y}{\partial \vartheta} \boldsymbol{b} + \frac{\partial Z}{\partial \vartheta} \boldsymbol{t} \right) \frac{\partial \boldsymbol{B}}{\partial \varphi} \right]. \qquad (3.7)$$

The terms of $O(\epsilon^0)$ are

$$\nabla \boldsymbol{B} \approx \frac{B_0^2}{G_0 \bar{B}} \left[\ell' \left(-\frac{\partial X_1}{\partial \vartheta} \boldsymbol{b} + \frac{\partial Y_1}{\partial \vartheta} \boldsymbol{n} \right) \boldsymbol{B}_1 + \ell'(X_1 \boldsymbol{b} - Y_1 \boldsymbol{n}) \frac{\partial \boldsymbol{B}_1}{\partial \vartheta} \right.$$

$$\left. + \left(X_1 \frac{\partial Y_1}{\partial \vartheta} - Y_1 \frac{\partial X_1}{\partial \vartheta} \right) \boldsymbol{t} \frac{\mathrm{d}\boldsymbol{B}_0}{\mathrm{d}\varphi} \right], \qquad (3.8)$$

where in the last term, (3.5) gives $\mathrm{d}\boldsymbol{B}_0/\mathrm{d}\varphi = s_G B_0' \boldsymbol{t} + s_G B_0 \ell'\kappa \boldsymbol{n}$. We substitute in (3.6), noting that for any functions

$$P(\vartheta) = P_s \sin\vartheta + P_c \cos\vartheta, \quad Q(\vartheta) = Q_s \sin\vartheta + Q_c \cos\vartheta, \qquad (3.9a,b)$$

we have

$$P \frac{\partial Q}{\partial \vartheta} - Q \frac{\partial P}{\partial \vartheta} = P_c Q_s - P_s Q_c, \quad PQ + \frac{\partial P}{\partial \vartheta} \frac{\partial Q}{\partial \vartheta} = P_s Q_s + P_c Q_c \qquad (3.10a,b)$$

and $\partial^2 P/\partial \vartheta^2 = -P$. We also note (A21) in LS, $X_{1c}Y_{1s} - X_{1s}Y_{1c} = s_G \bar{B}/B_0$ (the condition that the first-order flux surface enclose the proper toroidal flux at each φ). We thereby

obtain the final result for a general configuration as follows:

$$
\nabla B \approx \frac{B_0^2}{\bar{B}\ell'} \left[\left(X_{1c}'Y_{1s} - X_{1s}'Y_{1c} + \iota_{N0}[X_{1s}Y_{1s} + X_{1c}Y_{1c}] \right) \boldsymbol{nn} \right.
$$

$$
+ \left(X_{1c}X_{1s}' - X_{1s}X_{1c}' - \frac{s_G\bar{B}\ell'\tau}{B_0} - \iota_{N0}[X_{1s}^2 + X_{1c}^2] \right) \boldsymbol{bn}
$$

$$
+ \left(Y_{1c}'Y_{1s} - Y_{1s}'Y_{1c} + \frac{s_G\bar{B}\ell'\tau}{B_0} + \iota_{N0}[Y_{1s}^2 + Y_{1c}^2] \right) \boldsymbol{nb}
$$

$$
\left. + \left(X_{1c}Y_{1s}' - X_{1s}Y_{1c}' - \iota_{N0}[X_{1s}Y_{1s} + X_{1c}Y_{1c}] \right) \boldsymbol{bb} \right]
$$

$$
+ s_G B_0 \kappa \left(\boldsymbol{tn} + \boldsymbol{nt} \right) + \frac{s_G B_0'}{\ell'} \boldsymbol{tt}. \tag{3.11}
$$

Using (A23) in LS, (which expresses how on-axis rotational transform is driven by rotating elongation, axis torsion and toroidal current), it can be seen that the $\boldsymbol{nb}$ component and the $\boldsymbol{bn}$ component become equal when $I_2 = 0$, giving the expected symmetry of ∇B for a vacuum field.

In the case of quasisymmetry, (3.11) simplifies due to $B_0' = 0$ and $X_{1s} = 0$, and we also have $\bar{B} = s_\psi B_0$. Thus, the ∇B tensor for quasisymmetry is

$$
\nabla B \approx \frac{s_\psi B_0}{\ell'} \left[\left(X_{1c}'Y_{1s} + \iota_{N0}X_{1c}Y_{1c} \right) \boldsymbol{nn} + \left(-s_G s_\psi \ell' \tau - \iota_{N0}X_{1c}^2 \right) \boldsymbol{bn} \right.
$$

$$
+ \left(Y_{1c}'Y_{1s} - Y_{1s}'Y_{1c} + s_G s_\psi \ell' \tau + \iota_{N0}[Y_{1s}^2 + Y_{1c}^2] \right) \boldsymbol{nb}
$$

$$
\left. + \left(X_{1c}Y_{1s}' - \iota_{N0}X_{1c}Y_{1c} \right) \boldsymbol{bb} \right] + s_G B_0 \kappa \left(\boldsymbol{tn} + \boldsymbol{nt} \right). \tag{3.12}
$$

Analogously to (3.7)–(3.8), we can evaluate the $\nabla\nabla B$ tensor as well. To leading order,

$$
\nabla\nabla B \approx \frac{B_0^2}{G_0 \bar{B}} \left[\ell' \left(-\frac{\partial X_1}{\partial \vartheta} \boldsymbol{b} + \frac{\partial Y_1}{\partial \vartheta} \boldsymbol{n} \right) (\nabla B)_1 + \ell'(X_1 \boldsymbol{b} - Y_1 \boldsymbol{n}) \frac{\partial (\nabla B)_1}{\partial \vartheta} \right.
$$

$$
\left. + \left(X_1 \frac{\partial Y_1}{\partial \vartheta} - Y_1 \frac{\partial X_1}{\partial \vartheta} \right) \boldsymbol{t} \frac{\partial (\nabla B)_0}{\partial \varphi} \right]. \tag{3.13}
$$

The result is too lengthy to write here, but it can be found in a Mathematica notebook in Landreman (2020*a*). The result depends on X_2, Y_2 and Z_2.

For reference, we note the values of ∇B and $\nabla\nabla B$ in the limit of an axisymmetric and purely toroidal vacuum field, which is the field of an infinite straight vertical wire at major radius $R = 0$. This limit can be used for verification and to derive figures of merit such as (3.1)–(3.2), the equivalent distance to a coil. The field is $\boldsymbol{B} = s_G\mu_0 I\boldsymbol{t}/(2\pi R)$, where $I > 0$ is the current in the wire, R is the distance from the wire and $\boldsymbol{t} = \boldsymbol{e}_\phi$ is the unit vector in the direction of the standard cylindrical angle ϕ. Note also $\boldsymbol{n} = -\boldsymbol{e}_R$. Then

$$
\nabla B = (\nabla R)\frac{\partial B}{\partial R} + (\nabla \phi)\frac{\partial B}{\partial \phi} = \frac{s_G\mu_0 I}{2\pi R^2}(\boldsymbol{nt} + \boldsymbol{tn}), \tag{3.14}
$$

consistent with the appropriate limit of (3.12) ($d/d\varphi = 0$, $\tau = 0$, $\iota_{N0} = 0$). Forming the squared Frobenius norm of this expression and solving for R gives $R = L_{\nabla B}$ where $L_{\nabla B}$ is
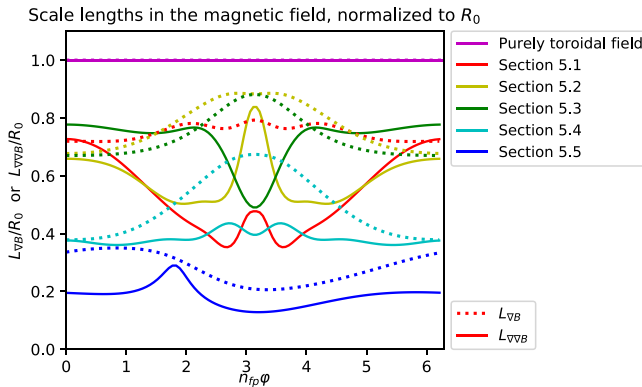
FIGURE 1. Scale lengths in the magnetic field, indicating measures of the effective distance to an electromagnetic coil, evaluated for the five examples in Landreman & Sengupta (2019). Section numbers in the legend refer to this earlier paper. Both of the scale lengths (3.1) and (3.2) are shown.

the figure of merit (3.1). Similarly,

$$\nabla\nabla B = (\nabla R)\frac{\partial(\nabla B)}{\partial R} + (\nabla\phi)\frac{\partial(\nabla B)}{\partial\phi} = \frac{s_G\mu_0 I}{\pi R^3}(\boldsymbol{tnn} + \boldsymbol{ntn} + \boldsymbol{nnt} - \boldsymbol{ttt}). \tag{3.15}$$

Squaring and solving for R gives $R = L_{\nabla\nabla B}$ where $L_{\nabla\nabla B}$ is the expression in (3.2).

For this case of a purely toroidal $B \propto 1/R$ field, the same B can be produced in an axisymmetric toroidal domain by a poloidal sheet current on the domain boundary. Hence the currents can be moved closer to an evaluation point at some fixed $R > 0$ without changing $L_{\nabla B}$ or $L_{\nabla\nabla B}$. This example illustrates that for a given field, there is no unique distance to a coil. The best we can hope to compute from a given B is a maximum distance to a current, reflecting the distance to a singularity in the exterior (Greene 1965). In this axisymmetic example, where there must be current through the donut hole of the toroidal region where $B \propto 1/R$, $L_{\nabla B}$ and $L_{\nabla\nabla B}$ give the correct maximum possible distance to a current.

Figure 1 shows the scale lengths $L_{\nabla B}$ and $L_{\nabla\nabla B}$ for the five $O(\epsilon^2)$ quasisymmetric configurations discussed in LS. Colours and legend text in the figure refer to section numbers in LS, each representing a different stellarator configuration. The scale lengths are normalized to the major radius to yield dimensionless figures of merit. By design, this normalized measure is 1 for a purely toroidal field, and small values are undesirable. The examples from sections 5.1–5.3 are quasi-axisymmetric, while the examples from sections 5.4 and 5.5 are quasihelically symmetric. The example of section 5.5 lacks stellarator symmetry, while the other examples are stellarator symmetric. Therefore the two curves for this example lack reflection symmetry about $n_{fp}\varphi = \pi$, whereas the curves for the stellarator-symmetric examples do have this reflection symmetry. The example of section 5.5 has quite strong shaping, as is evident in figure 17 of LS. Therefore it is not surprising that this example has significantly smaller values for the normalized scale lengths compared with the other examples. This initial evidence is at least suggestive that $L_{\nabla B}$ and $L_{\nabla\nabla B}$ may be useful as proxies for the complexity of a magnetic configuration. Specifically, noting that $L_{\nabla B}$ and $L_{\nabla\nabla B}$ are somewhat independent, and that the shortest scale length dominates, we conjecture that small values of $\min(L_{\nabla B}, L_{\nabla\nabla B})$ are correlated with high coil complexity. Further study involving coil optimization is needed to verify this conjecture. Note that $L_{\nabla B}$ and $L_{\nabla\nabla B}$ can be evaluated in under 1 ms for reasonable
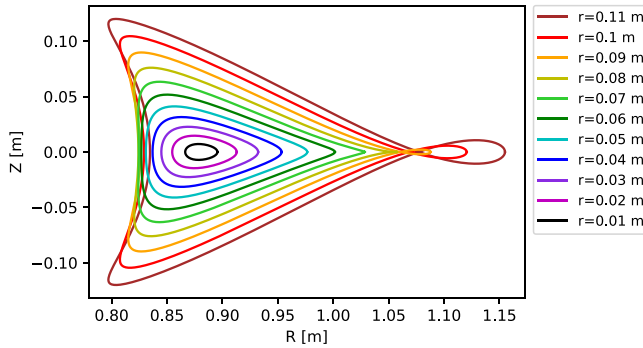
FIGURE 2. Several unphysical behaviours can arise in the constructed flux surface shapes for sufficiently large minor radius r. In the example here, showing the $\phi = 0$ cross-section of the configuration of §4.3, each surface intersects itself on the large-R side when $r > 0.085$ m, and the surfaces are not properly nested on the small-R side for $r > 0.077$ m.

resolutions, roughly four orders of magnitude faster than even the fastest calculations of coil shapes with current–potential methods like REGCOIL (Landreman 2017).

4. Minimum aspect ratio without intersecting surfaces

When a stellarator configuration is constructed to $O(\epsilon^2)$ using the near-axis expansion, a common problem is the following. The triangularity of the constructed boundary surface grows with r, as does the shift of the surface centroid compared with the magnetic axis. For r beyond a critical value r_c, the surface can begin to self-intersect, or the surfaces may no longer be nested, as shown in figure 2. These intersections and overlaps of the surfaces are not physically realizable, but rather are an indication that the small-r expansion has broken down. These problems with the surface geometry put a lower limit on the aspect ratio at which boundary surfaces can be constructed by the near-axis method. Therefore, it is useful to be able to calculate r_c.

In effect, r_c is a convenient summary of how rapidly the coefficients of the r expansion increase with the order in the expansion. The (X_j, Y_j, Z_j) coefficients tend to increase with j, since in the Garren–Boozer equations the order-j terms depend on the $d/d\varphi$ derivatives of the order-$(j-1)$ terms. This increase with j limits the radius of convergence of the expansion in r. When r_c is small, the coefficients increase rapidly with order so the expansion is accurate only for small r. We wish to find solutions of the near-axis equations for which the coefficients do not increase rapidly with order, so the expansion is accurate for relatively large r. Hence, we seek solutions with large r_c.

For quasisymmetric configurations, this critical value r_c for singularity is also useful as a measure of the accuracy of quasisymmetry. This is because the construction achieves quasisymmetry through $O(\epsilon^2)$ but not at $O(\epsilon^3)$, so quasisymmetry is not accurately obtained at values of r for which the $O(\epsilon^3)$ terms matter. One can expect the quantities $\{B_j, X_j, Y_j, Z_j\}$ at each order j to all be roughly comparable in magnitude since they are related by the equations at each order of the expansion, e.g. (A32)–(A48) of LS. The $O(\epsilon^3)$ terms in the shape evidently matter when $r \sim r_c$, for then the $O(\epsilon^2)$ shape is unphysical. Therefore, configurations with large r_c can be expected to have good quasisymmetry throughout a larger volume than configurations with small r_c. Motivated by these reasons above, in this section we seek a method to calculate r_c.

Considering the type of singularity on the large-R side of figure 2, the sharp edge in the surface shape is reminiscent of the X-point in a diverted tokamak. It may therefore be

possible to take advantage of this kind of singularity to design a stellarator divertor. We will not attempt to pursue this possibility here.

4.1. *Problem formulation*

The condition for self-intersection or overlap of the flux surfaces, i.e. constant-r surfaces, is $\sqrt{g} = 0$, where $\sqrt{g}$ is the Jacobian,

$$\sqrt{g} = \frac{\partial \boldsymbol{r}}{\partial r} \times \frac{\partial \boldsymbol{r}}{\partial \vartheta} \cdot \frac{\partial \boldsymbol{r}}{\partial \varphi}. \tag{4.1}$$

We define r_c as the minimum positive value of r such that $\sqrt{g} = 0$. Since this definition has the form of a constrained optimization problem, we introduce a Lagrange multiplier λ, and seek stationary points of the Lagrangian $\mathcal{L}(\lambda, r, \vartheta, \varphi) = r + \lambda\sqrt{g}$. Variation with respect to λ recovers the constraint $\sqrt{g} = 0$. Variation of $\mathcal{L}$ with respect to the spatial coordinates yields

$$\lambda \partial\sqrt{g}/\partial r = -1, \tag{4.2}$$

$$\partial\sqrt{g}/\partial\vartheta = 0, \tag{4.3}$$

$$\partial\sqrt{g}/\partial\varphi = 0. \tag{4.4}$$

Equation (4.2) effectively determines λ, and will not be needed. In practice, since the shape coefficients $\{X, Y, Z\}$ are available on a grid in φ, it is convenient to replace (4.4) with minimization over the same φ grid, as follows. We define $\hat{r}_c(\varphi)$ as the solution of $\sqrt{g} = 0$ and (4.3) at given φ, we can evaluate $\hat{r}_c(\varphi)$ on the numerical grid in φ, and we then identify $r_c = \min_\varphi \hat{r}_c(\varphi)$. The key task then is to compute $\hat{r}_c(\varphi)$.

To this end, we substitute the position vector (2.4) and expansions (2.6) and (2.9) into (4.1), using (2.5*a–d*). If the expansion for $\{X, Y, Z\}$ is truncated after $\{X_1, Y_1, Z_1\}$, corresponding to the $O(\epsilon)$ construction, then $\sqrt{g} = (1 - rX_1\kappa)rs_G\bar{B}\ell'/B_0$, and the solution of $\sqrt{g} = 0$ and $\partial\sqrt{g}/\partial\vartheta = 0$ gives $\hat{r}_c(\varphi) = 1/(\kappa\sqrt{X_{1s}^2 + X_{1c}^2})$. For the rest of §4 we will consider the more complicated problem of the $O(\epsilon^2)$ construction, in which the expansion for $\{X, Y, Z\}$ is truncated after $\{X_2, Y_2, Z_2\}$. In this case, the product in (4.1) involving three copies of the position vector includes terms scaling as r^1 through r^5,

$$\sqrt{g} = r\sum_{j=0}^{4} r^j g_j(\vartheta, \varphi) \tag{4.5}$$

where $g_0 = g_0(\varphi)$ is independent of ϑ,

$$g_1 = g_{1s}(\varphi)\sin\vartheta + g_{1c}(\varphi)\cos\vartheta, \tag{4.6}$$

$$g_2 = g_{20}(\varphi) + g_{2s}(\varphi)\sin 2\vartheta + g_{2c}(\varphi)\cos 2\vartheta, \text{ etc.}, \tag{4.7}$$

analogous to (2.6). We find $g_0 = (X_{1c}Y_{1s} - X_{1s}Y_{1c})\ell' = s_G\bar{B}\ell'/B_0$ and $g_{1s,c} = -2s_G\bar{B}\ell' B_{1s,c}/B_0^2$. These expressions can be derived using (A21) and (A32)–(A33) in LS, which express equality of the $\boldsymbol{t}$ components of the covariant and contravariant representations of $\boldsymbol{B}$. These results for g_0 and g_1 can also be derived by expanding $(G + \iota I)r\bar{B}/B^2$, which is equal to $\sqrt{g}$ if all orders in the expansion are retained. However, since the expansion is truncated after $\{X_2, Y_2, Z_2\}$, g_j does not equal the corresponding term in the expansion of $(G + \iota I)r\bar{B}/B^2$ for $j \geq 2$. Explicit expressions for g_2 are given in appendix A.

At each φ, the equations $\sqrt{g} = 0$ and $\partial\sqrt{g}/\partial\vartheta$ then give two equations for the two unknowns ϑ and $r = \hat{r}_c$. In principle this nonlinear system could be solved numerically using Newton's method. As with any such nonlinear problem, it is hard to ensure that all possible solutions are found, since the numerical solution depends on an initial guess. However, we are primarily interested in the smallest positive solution for $\hat{r}_c$, which is the root that sets the limit on minimum aspect ratio. In the next subsection we show how this smallest positive root can be found more robustly using an asymptotic approach. The approximate root location computed in the next subsection can serve as a good initial guess for Newton iteration if a more accurate value of r_c is desired.

4.2. *Robust solution*

Since we seek solutions of $\sqrt{g} = 0$ and $\partial\sqrt{g}/\partial\vartheta = 0$ for small r, we keep only the leading three orders in $\sqrt{g}$: $\sqrt{g} \approx r(g_0 + rg_1 + r^2 g_2)$. Then (4.3) can be written as

$$r = \frac{g_{1c} \sin\vartheta - g_{1s} \cos\vartheta}{2(g_{2s} \cos 2\vartheta - g_{2c} \sin 2\vartheta)}. \tag{4.8}$$

Using this result to eliminate r in $\sqrt{g} = 0$, we obtain

$$K_0 + K_{2s} \sin 2\vartheta + K_{2c} \cos 2\vartheta + K_{4s} \sin 4\vartheta + K_{4c} \cos 4\vartheta = 0, \tag{4.9}$$

where

$$\left.\begin{aligned}
K_0 &= 2g_{20}(g_{1c}^2 + g_{1s}^2) + 8g_0(g_{2c}^2 + g_{2s}^2) + 3g_{2c}(g_{1s}^2 - g_{1c}^2) - 6g_{1c}g_{1s}g_{2s}, \\
K_{2s} &= 2g_{2s}(g_{1c}^2 + g_{1s}^2) - 4g_{1s}g_{1c}g_{20}, \\
K_{2c} &= 2g_{20}(g_{1s}^2 - g_{1c}^2) + 2g_{2c}(g_{1s}^2 + g_{1c}^2), \\
K_{4s} &= g_{2s}(g_{1c}^2 - g_{1s}^2) + 2g_{1c}g_{1s}g_{2c} - 16g_0g_{2c}g_{2s}, \\
K_{4c} &= g_{2c}(g_{1c}^2 - g_{1s}^2) + 8g_0(g_{2s}^2 - g_{2c}^2) - 2g_{1s}g_{1c}g_{2s}.
\end{aligned}\right\} \tag{4.10}$$

We now convert (4.9) to an equation for the roots of a polynomial, since all roots of such equations can be computed robustly using the companion matrix method described below. Introducing $w = \sin 2\vartheta$, then $\cos 2\vartheta = \varpi\sqrt{1 - w^2}$ where $\varpi = \pm 1$. Also, $\sin 4\vartheta = 2w\varpi\sqrt{1 - w^2}$ and $\cos 4\vartheta = 1 - 2w^2$. Using these results to eliminate the trigonometric functions, (4.9) can be written as

$$P_4 w^4 + P_3 w^3 + P_2 w^2 + P_1 w + P_0 = 0 \tag{4.11}$$

where

$$\left.\begin{aligned}
P_4 &= 4K_{4c}^2 + 4K_{4s}^2, \\
P_3 &= 4K_{4s}K_{2c} - 4K_{4c}K_{2s}, \\
P_2 &= K_{2s}^2 + K_{2c}^2 - 4K_0K_{4c} - 4K_{4c}^2 - 4K_{4s}^2, \\
P_1 &= 2K_0K_{2s} + 2K_{4c}K_{2s} - 4K_{4s}K_{2c}, \\
P_0 &= (K_0 + K_{4c})^2 - K_{2c}^2.
\end{aligned}\right\} \tag{4.12}$$

The solutions w of (4.11) can be computed robustly by finding the eigenvalues of the companion matrix

$$M = \begin{pmatrix} -P_3/P_4 & -P_2/P_4 & -P_1/P_4 & -P_0/P_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{4.13}$$

a matrix constructed such that its characteristic polynomial $\det|M - wI|$ (for identity matrix I) is proportional to the original polynomial (4.11).

Then for each real solution w, ϖ can in principle be computed from (4.9) in the form

$$\varpi = -\frac{K_0 + K_{2s}w + K_{4c}(1 - 2w^2)}{(2wK_{4s} + K_{2c})\sqrt{1 - w^2}}, \tag{4.14}$$

where the square of the right-hand side is 1 due to (4.11). Thus, precisely one of the two choices for ϖ is consistent with each real root w. Due to possible precision loss in the expression (4.14) in finite-precision arithmetic, it is more convenient in practice to select ϖ as the element of $\{-1, 1\}$ that minimizes the residual in (4.9). With $\cos 2\vartheta = \varpi\sqrt{1 - w^2}$ now known, $\cos \vartheta$ can be computed from $\varsigma\sqrt{(1 + \cos 2\vartheta)/2}$, with $\varsigma = \pm 1$. Note that $\sin \vartheta$ can be computed from $(\sin 2\vartheta)/(2\cos \vartheta)$. The two choices of ς correspond to a pair of solutions of $\{\sqrt{g} = 0, \ \partial\sqrt{g}/\partial\vartheta = 0\}$ in which r differs by a factor of -1, i.e. the two choices represent two representations $(\vartheta_0, \ r_0)$ and $(\vartheta_0 + \pi, \ -r_0)$ of the same physical point. It is no loss of generality to consider only the solution with positive r. One can compute r from (4.8), or from $\sqrt{g} = 0$ if the denominator of (4.8) vanishes. The smallest positive solution for r derived from the various roots of (4.11) is then $\hat{r}_c$.

By this method, the nonlinear equations for $\hat{r}_c$ can be solved without any need for an initial guess. The resulting value of $\hat{r}_c$ can either be used directly as a good approximation for the exact solution (by which we mean a solution of $\{\sqrt{g} = 0, \ \partial\sqrt{g}/\partial\vartheta = 0\}$ in which g_4 and g_5 are retained in $\sqrt{g}$), or it can be used as an initial guess for Newton iteration to find the exact solution.

4.3. *Example*

Figures 2 and 3 demonstrate the methods of this section for an example quasi-axisymmetric configuration. The configuration is constructed using the $O(\epsilon^2)$ equations, with axis shape $R[\text{m}] = 1 - 0.12\cos(2\phi)$ and $Z[\text{m}] = 0.12\sin(2\phi)$. The other input parameters for the construction were $\bar{\eta} = -0.7 \text{ m}^{-1}$, $\sigma(0) = 0$, $I_2 = 0$, $p_2 = 0$, $B_{2s} = 0$ and $B_{2c} = -0.5 \text{ T/m}^2$. This configuration was chosen since multiple types of singularities can be seen in the $\phi = 0$ cross-section, as shown in figure 2. On the large-R side, each magnetic surface with sufficiently large r crosses through itself, and on the small-R side, the surfaces are not simply nested beyond a critical r.

Figure 3 shows the value of $\hat{r}_c$ computed using the method of § 4.2 for this example. The solution found is the one at the small-R side, occurring at $\hat{r}_c = 0.0762$ m in the $\varphi = 0$ plane, a slightly smaller value of r than the singularity at the large-R side. At each grid point in φ, this solution is used as an initial guess for a solution of the full problem from § 4.1 using Newton's method. Newton's method converges to machine precision at all grid points, and the solution is plotted on the same graph. In this case the Newton solution barely differs from the robust approximate solution, with a solution $\hat{r}_c = 0.0767$ m in the $\varphi = 0$ plane. For other magnetic configurations in which the minimum aspect ratio is much lower, significant differences can arise between the robust and Newton solutions for
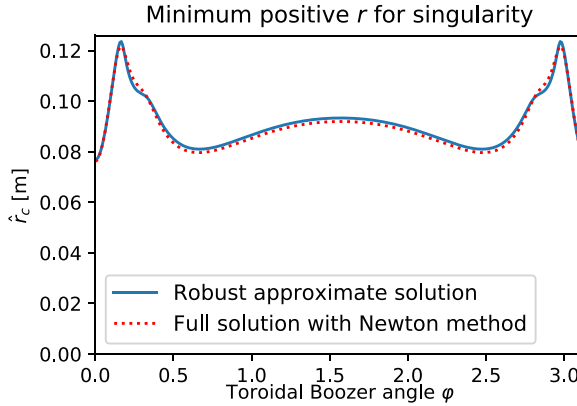
FIGURE 3. Effective minor radius r at which the magnetic surfaces intersect, for the example of § 4.3 and figure 2. The blue solid curve is computed by the method of § 4.2.

r_c. Nevertheless, the robust method is found to be extremely accurate (as in the example shown) when the minimum aspect ratio is large, making it useful for screening out such configurations.

4.4. *Tokamak equilibrium limit*

In both tokamaks and stellarators, an equilibrium beta limit exists associated with an X-point approaching and reaching the plasma boundary (Mukhovatov & Shafranov 1971; Freidberg 2014). As the method of §§ 4.1 and 4.2 finds singularities in the flux surface geometry, we might expect that this method can calculate this equilibrium beta limit. Here we show that this is indeed true for a high-aspect-ratio tokamak. This correspondence gives added meaning and value to the figure of merit r_c in §§ 4.1 and 4.2.

For this calculation we use the 'high-β tokamak' ordering, (6.92) of Freidberg (2014): $\epsilon \ll 1$ where $\epsilon = a/R$, $\mu_0 p_0/B_0^2 \sim \epsilon$ and $\iota \sim 1$. Since axisymmetry is considered here, the Garren–Boozer equations in this case are given in appendix B. We consider a pressure profile $p(r) = p_0 + r^2 p_2$ such that $p(a) = 0$ at the plasma boundary $r = a$, so $p_2 = -p_0/a^2$. Then using (B 1), it follows that

$$\mu_0|p_2| \gg I_2^2. \tag{4.15}$$

We limit our attention to up–down symmetric geometry, so $\sigma = 0$. The near-axis solutions have a degree of freedom B_{2c} reflecting the freedom in triangularity, and for simplicity we will limit attention to the case of zero triangularity, which can be expressed as

$$(X - \Delta)^2 + Y^2 b = c, \tag{4.16}$$

for some flux functions $\Delta(r)$, $b(r)$ and $c(r)$. This equation is independent of the near-axis expansion. To demand consistency between (4.16) and the near-axis equations, we take $\Delta(r)$, $b(r)$ and $c(r)$ to be polynomials and equate terms at each order. The $r^2 \cos^2 \vartheta$ terms give the leading (constant) term of b to be X_{1c}^2/Y_{1s}^2. Then the $r^3 \cos 3\vartheta$ terms give

$$Y_{1s}X_{2c} - X_{1c}Y_{2s} = 0. \tag{4.17}$$

This result, with (B 11) and (B 6), determines B_{2c}.

Using these results and the equations of appendix B, and considering the major radius of the magnetic axis to be R_0, the first few terms in the expansion of $\sqrt{g}$ are $g_0 = s_G s_\psi R_0$, $g_{1s} = 0$, $g_{1c} = -2 s_G s_\psi \bar{\eta} R_0$,

$$g_{20} \approx \frac{s_G s_\psi \mu_0^2 p_2^2 (1 + \bar{\eta}^4 R_0^4)^4}{I_2^4 \bar{\eta}^2 R_0^3 (3 + \bar{\eta}^4 R_0^4)^2}, \tag{4.18}$$

$g_{2s} = 0$ and $g_{2c} \approx -2 g_{20}$, where (4.15) has been applied. The largest terms in (4.9) and (4.10) are then $K_0 \approx 8 g_0 g_{2c}^2$ and $K_{4c} \approx -K_0$, so $K_0(1 - \cos 4\vartheta) = 0$, implying $\vartheta = j\pi/2$ for integers j. The solutions for odd j do not yield real solutions for r. For the even-j solutions, $\partial\sqrt{g}/\partial\vartheta = 0$ is satisfied automatically. The $\vartheta = \pi$ solution is just the $\vartheta = 0$ solution with $r_c \to -r_c$, so it is sufficient to consider $\vartheta = 0$. Solving $0 = \sqrt{g} \propto g_0 + r_c g_{1c} + r_c^2 (g_{20} + g_{2c})$ for r_c then gives the positive solution to be

$$r_c \approx \sqrt{\frac{g_0}{g_{20}}} \approx \frac{I_2^2 \bar{\eta} R_0^2 (3 + \bar{\eta}^4 R_0^4)}{\mu_0 |p_2| (1 + \bar{\eta}^4 R_0^4)^2} = \frac{\iota_0^2 B_0^2 (3 + \bar{\eta}^4 R_0^4)}{4 \mu_0 |p_2| \bar{\eta}^3 R_0^4}, \tag{4.19}$$

where the last equality follows from (B 1). If all terms through g_4 are retained in $\sqrt{g}$ instead of just terms through g_2, a root of $\sqrt{g} = 0$ and $\partial\sqrt{g}/\partial\vartheta$ exists at $\vartheta = 0$ which is identical to (4.19) to leading order in $I_2^2/|\mu_0 p_2| \ll 1$. Therefore the truncation at g_2 used for the robust solution is a good approximation.

The plasma minor radius a must be within the singularity: $a < r_c$. This inequality can be written using (4.19) in the form of a β-limit as follows:

$$\frac{2\mu_0 \bar{p}}{B_0^2} < E(\bar{\eta} R_0) \iota_0^2 \frac{a}{R_0} \sim \iota_0^2 \frac{a}{R_0}, \tag{4.20}$$

where $\bar{p} = (2/a^2) \int_0^a p r \, dr = p_0/2$ is an average pressure and

$$E(\bar{\eta} R_0) = \frac{3 + \bar{\eta}^4 R_0^4}{4 \bar{\eta}^3 R_0^3} \tag{4.21}$$

is a coefficient of order unity that depends on the elongation κ_e. As discussed in appendix B, $\kappa_e = |Y_1(\vartheta = \pi/2)/X_1(\vartheta = 0)| = |Y_{1s}/X_{1c}| = 1/(\bar{\eta}^2 R_0^2)$. For surfaces with circular cross-section, $E \to 1$.

For comparison, the tokamak equilibrium beta limit can be found in (236) of Mukhovatov & Shafranov (1971) or § 6.5 of Freidberg (2014). These previous results can be expressed as

$$\left\langle \frac{2\mu_0 p}{B^2} \right\rangle \lesssim \iota^2 \frac{a}{R_0}, \tag{4.22}$$

where $\langle \rangle$ represents an average over the plasma and ι is a characteristic rotational transform. The correspondence between (4.20) and (4.22) is apparent. An exact match of the coefficient (4.21) to earlier results like Freidberg (2014) is complicated because the earlier calculations include radially varying elongation and magnetic shear, which do not appear until higher orders in the Garren–Boozer expansion than the order we consider. The boundary separatrix in the earlier tokamak calculations cannot truly be represented by the shape (2.6) and (2.9) through $O(r^2)$ used here, so we would not necessarily expect the numerical coefficient to match exactly. Nevertheless the basic scaling $\beta \sim \iota^2 a/R$ is in clear agreement.

Based on this successful comparison, the singularity calculation in §§ 4.1 and 4.2 could perhaps be used to find stellarator configurations with a suitably high equilibrium beta limit, as follows. Given a desired on-axis pressure p_0 and desired minor radius a, p_2 could be set equal to $-p_0/a^2$. Then other parameters of the Garren–Boozer model (axis shape, $\bar{\eta}$, etc.) could be found numerically such that $r_c > a$. Any resulting configuration satisfying this inequality should have an acceptably high equilibrium beta limit (at least of the sort considered here, associated with separatrices). It should be acknowledged that stellarators also are thought to have an equilibrium beta limit associated with increasing stochasticity of the magnetic field (Loizu *et al.* 2017), which is not computed by the method here.

Near the equilibrium β limit, the shift of the geometric centre of the boundary flux surfaces relative to the magnetic axis (Shafranov shift) becomes a substantial fraction of the plasma minor radius. It is noteworthy that this Shafranov shift is accurately represented in the Garren–Boozer equations, as shown in detail for axisymmetry in appendix B. Therefore, using the Garren–Boozer equations, it may also be possible to assess the equilibrium β limit in a general stellarator without using the r_c metric of §§ 4.1 and 4.2 by instead examining the shift of the surface centroids relative to the axis. While such an approach would only detect the kind of singularity at the left of figure 2 and not the kind at the right, it may provide more insight than the procedure of §§ 4.1 and 4.2. This alternative approach is left for future work.

5. Error field for the first-order construction

Not only can the near-axis analysis be used to construct boundary shapes that give quasisymmetry to a certain order; the expansions can also provide formulae for the size of the quasisymmetry-breaking variation in B at the next order. If the magnitude of this symmetry-breaking is large, then the volume of good quasisymmetry is evidently quite limited, whereas configurations with small symmetry-breaking terms can be expected to have a larger volume of good quasisymmetry. Therefore, we expect it will be useful to derive formulae for these symmetry-breaking terms.

In this section, we will consider the $O(\epsilon^1)$ quasisymmetry construction, in which the cross-section of the boundary flux surface (in the plane perpendicular to the magnetic axis) is a perfect ellipse. We will then compute the $O(\epsilon^2)$ variation in B, which will generally break the quasisymmetry. The analysis will closely follow the method of § 3 and appendix B of LS for computing the field strength that is realized inside a constructed boundary. However, in LS, the surface shape coefficients (X_j, Y_j, Z_j) were constructed through $j = 2$, and some $j = 3$ terms were included, whereas here the boundary shape includes only $j = 1$ terms. The calculation in LS showed what the $j = 2$ and $j = 3$ terms must be to achieve a desired B_2, whereas the calculation here shows what B_2 is achieved if the $j \geq 2$ terms in the boundary shape are not included. These two calculations, while related, are sufficiently different that the latter requires the separate derivation here.

5.1. *Expansion*

We now review the expansion developed in § 3 and appendix B of LS. Throughout § 5 we do not assume quasisymmetry. We suppose a solution of the Garren–Boozer equations for X_1 and Y_1 is fixed. From this solution, a finite-aspect-ratio surface is constructed by setting r to a finite value a. The higher-order terms X_j, Y_j and Z_j for $j > 1$ are not considered when generating this surface. Then, $\boldsymbol{J} \times \boldsymbol{B} = \nabla p$ is solved inside this boundary without making a high-aspect-ratio expansion, with the boundary condition that $\boldsymbol{B}$ has no component normal to the boundary surface. This step could be done by running the VMEC code in fixed-boundary mode using the constructed surface. The result is a configuration that is similar to but not identical to the original near-axis solution: we integrated outward

from the axis only approximately, but then solved for the equilibrium inside the boundary surface exactly. As a result the final axis shape will be slightly different from the initial one. The resulting finite-aspect-ratio configuration also satisfies the Garren–Boozer equations, but with slightly altered expansion coefficients, denoted with tildes, e.g. $\tilde{X}_2$. The original 'non-tilde' expansion represents a single configuration we would like to achieve, whereas the tilde expansion represents a family of configurations parameterized by a. Thus, while the position vector in the non-tilde configuration is given by (2.4), the position vector in a tilde configuration is

$$\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{r}}_0(a, \tilde{\varphi}) + \tilde{X}(a, r, \tilde{\vartheta}, \tilde{\varphi})\tilde{\boldsymbol{n}}(a, \tilde{\varphi}) + \tilde{Y}(a, r, \tilde{\vartheta}, \tilde{\varphi})\tilde{\boldsymbol{b}}(a, \tilde{\varphi}) + \tilde{Z}(a, r, \tilde{\vartheta}, \tilde{\varphi})\tilde{\boldsymbol{t}}(a, \tilde{\varphi}). \quad (5.1)$$

The Boozer angles $\tilde{\vartheta}$ and $\tilde{\varphi}$ in the constructed configuration are allowed to differ slightly from the original Boozer angles ϑ and φ, and the differences are given by single-valued quantities t and f as follows:

$$\tilde{\vartheta}(a, \vartheta, \varphi) = \vartheta + t(a, \vartheta, \varphi), \quad \tilde{\varphi}(a, \vartheta, \varphi) = \varphi + f(a, \vartheta, \varphi). \quad (5.2a,b)$$

These expressions define $(\tilde{\vartheta}, \tilde{\varphi})$ only up to an additive constant, and it is convenient to eliminate this freedom using the following constraints:

$$\int_0^{2\pi} d\vartheta \int_0^{2\pi} d\varphi\, t(a, \vartheta, \varphi) = 0, \quad \int_0^{2\pi} d\vartheta \int_0^{2\pi} d\varphi\, p(a, \vartheta, \varphi) = 0. \quad (5.3a,b)$$

Similar to (2.6), the $\tilde{X}$, $\tilde{Y}$ and $\tilde{Z}$ coefficients each have expansions of the form

$$\tilde{X}(a, r, \tilde{\vartheta}, \tilde{\varphi}) = \sum_{j=1}^{\infty} r^j \tilde{X}_j(a, \tilde{\vartheta}, \tilde{\varphi}), \quad (5.4)$$

while $\tilde{B}$ and $\tilde{\beta}$ have analogous expansions that include a $j = 0$ term, e.g.

$$\tilde{B}(a, r, \tilde{\vartheta}, \tilde{\varphi}) = \sum_{j=0}^{\infty} r^j \tilde{B}_j(a, \tilde{\vartheta}, \tilde{\varphi}). \quad (5.5)$$

For each quantity in the tilde configurations, the a dependence has a Taylor series denoted by superscripts in parentheses,

$$\tilde{\boldsymbol{r}}_0(a, \tilde{\varphi}) = \sum_{k=0}^{\infty} a^k \tilde{\boldsymbol{r}}_0^{(k)}(\tilde{\varphi}). \quad (5.6)$$

Analogous expansions exist for $\tilde{\boldsymbol{n}}$, $\tilde{\boldsymbol{b}}$ and $\tilde{\boldsymbol{t}}$. Similarly,

$$\tilde{X}_j(a, \tilde{\vartheta}, \tilde{\varphi}) = \sum_{k=0}^{\infty} a^k \tilde{X}_j^{(k)}(\tilde{\vartheta}, \tilde{\varphi}), \quad (5.7)$$

and analogous expansions exist for $\tilde{Y}_j$, $\tilde{Z}_j$, $\tilde{B}_j$ and $\tilde{\beta}_j$. The analogous expansion for the angle differences is

$$t(a, \vartheta, \varphi) = \sum_{k=0}^{\infty} a^k t^{(k)}(\vartheta, \varphi), \quad f(a, \vartheta, \varphi) = \sum_{k=0}^{\infty} a^k f^{(k)}(\vartheta, \varphi). \quad (5.8a,b)$$

The profiles $I(r)$ and $p(r)$ are considered to be the same in the tilde and non-tilde configurations, since these profiles are typically inputs to an MHD equilibrium calculation.

Therefore in the finite-minor-radius calculation they can be matched exactly to the non-tilde profiles. However, the profiles $G(r)$ and $\iota(r)$ may generally differ in the tilde configurations, so we expand

$$\tilde{G}(a, r) = \sum_{j=0}^{\infty} r^{2j} \tilde{G}_{2j}(a), \quad \tilde{\iota}(a, r) = \sum_{j=0}^{\infty} r^{2j} \tilde{\iota}_{2j}(a), \tag{5.9a,b}$$

where

$$\tilde{G}_j(a) = \sum_{k=0}^{\infty} a^k \tilde{G}_j^{(k)}, \quad \tilde{\iota}_j(a) = \sum_{k=0}^{\infty} a^k \tilde{\iota}_j^{(a)}. \tag{5.10a,b}$$

We emphasize that subscripts refer to an expansion in distance from the axis of a given configuration, while superscripts indicate a separate expansion in the value of minor radius substituted into the original non-tilde expansion.

At $r = a$, the boundary shape represented by the non-tilde and tilde expansions coincides exactly, since this equivalence defines the tilde configuration. The equation representing this equivalence is, using (2.4) and (5.1),

$$\boldsymbol{r}_0(\varphi) + X(a, \vartheta, \varphi)\boldsymbol{n}(\varphi) + Y(a, \vartheta, \varphi)\boldsymbol{b}(\varphi) + Z(a, \vartheta, \varphi)\boldsymbol{t}(\varphi)$$
$$= \tilde{\boldsymbol{r}}_0(a, \tilde{\varphi}) + \tilde{X}(a, a, \tilde{\vartheta}, \tilde{\varphi})\tilde{\boldsymbol{n}}(a, \tilde{\varphi}) + \tilde{Y}(a, a, \tilde{\vartheta}, \tilde{\varphi})\tilde{\boldsymbol{b}}(a, \tilde{\varphi}) + \tilde{Z}(a, a, \tilde{\vartheta}, \tilde{\varphi})\tilde{\boldsymbol{t}}(a, \tilde{\varphi}),$$
$$\tag{5.11}$$

and it plays a central role in the analysis.

To the order of interest, the field strength in the constructed (i.e. tilde) configurations is

$$\tilde{B}(a, r, \tilde{\vartheta}, \tilde{\varphi}) = \tilde{B}_0^{(0)}(\tilde{\varphi}) + r\tilde{B}_1^{(0)}(\tilde{\vartheta}, \tilde{\varphi}) + a\tilde{B}_0^{(1)}(\tilde{\varphi})$$
$$+ r^2 \tilde{B}_2^{(0)}(\tilde{\vartheta}, \tilde{\varphi}) + ra\tilde{B}_1^{(1)}(\tilde{\vartheta}, \tilde{\varphi}) + a^2 \tilde{B}_0^{(2)}(\tilde{\varphi}) + O(\epsilon^3). \tag{5.12}$$

In appendix B.2 of LS, it was shown that for the first-order construction, $\tilde{B}_0^{(0)}(\varphi) = B_0(\varphi)$, $\tilde{B}_1^{(0)}(\vartheta, \varphi) = B_1(\vartheta, \varphi)$ and $\tilde{B}_0^{(1)}(\varphi) = 0$. Computation of the remaining terms is shown in detail in appendix C. In §C.1 we will compute $\tilde{B}_2^{(0)}$ and $\tilde{B}_1^{(1)}$, and in subsection C.2 we will compute $\tilde{B}_0^{(2)}$. This will complete the calculation of all terms in the second-order field strength (5.12) in terms of non-tilde quantities, which are considered known.

5.2. *Summary of results*

We now summarize the practical consequences of appendix C. When a magnetic surface is constructed to achieve a desired $B(\vartheta, \varphi)$ to $O(\epsilon^1)$, the $O(\epsilon^2)$ 'error field' terms in B are given by $r^2 \tilde{B}_2^{(0)}(\tilde{\vartheta}, \tilde{\varphi}) + a^2 \tilde{B}_0^{(2)}(\tilde{\varphi})$, since $\tilde{B}_1^{(1)}$ vanishes. To compute these terms we first evaluate $\tilde{Z}_2^{(0)}$ from (A27)–(A31) of LS. Then we solve a linear system given by (A41)–(A42) (with tildes) in LS, together with (C 9)–(C 12), with $\tilde{X}_{20}^{(0)}$ and $\tilde{Y}_{20}^{(0)}$ as the unknowns. Then $\tilde{B}_2^{(0)}(\tilde{\vartheta}, \tilde{\varphi}) = \tilde{B}_{20}^{(0)}(\tilde{\varphi}) + \tilde{B}_{2s}^{(0)}(\tilde{\varphi}) \sin 2\tilde{\vartheta} + \tilde{B}_{2c}^{(0)}(\tilde{\varphi}) \cos 2\tilde{\vartheta}$ is evaluated from (A34)–(A40) (adding tildes and $^{(0)}$ superscripts to the subscript-2 quantities). Finally, $\tilde{B}_0^{(2)}$ is found from (C 24)–(C 30). For the special case of quasisymmetry, (C 9)–(C 12) simplify as shown in § 3.

5.3. *Numerical verification*

We now demonstrate the methods of this section for a concrete example, the quasi-axisymmetric configuration of § 5.1 in Landreman *et al.* (2019), shown in figures 2
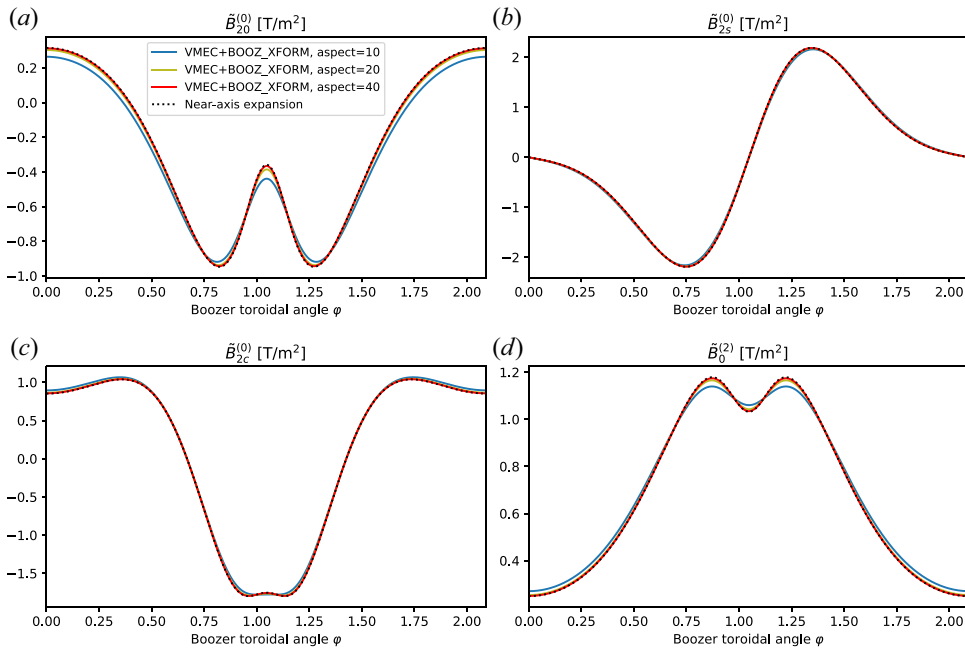
FIGURE 4. The leading quasisymmetry-breaking terms in B for the constructed quasi-axisymmetric example of §5.1 of Landreman *et al.* (2019). Excellent agreement is evident between the values of these terms computed by the near-axis analysis (dotted) compared with the values from finite-aspect-ratio equilibria (solid curves). The latter are computed with the VMEC and BOOZ_XFORM codes at a range of aspect ratios.

and 3 of that reference. This configuration is defined by the axis shape $R = 1 + 0.045 \cos(3\phi)$ and $Z = -0.045 \sin(3\phi)$ in units of metres, where (R, ϕ, Z) are standard cylindrical coordinates, and $\bar{\eta} = -0.9 \text{ m}^{-1}$, along with $I_2 = 0$ and $p_2 = 0$. For the $O(\epsilon)$ construction, the leading-order terms in B that cause departures from quasisymmetry are $\tilde{B}_{20}^{(0)}$, $\tilde{B}_{2s}^{(0)}$, $\tilde{B}_{2c}^{(0)}$ and $\tilde{B}_{0}^{(2)}$. These four quantities are computed using the near-axis analysis of this section and plotted in figure 4 as dotted curves. Then, toroidal magnetic surfaces are constructed for several different choices of the boundary minor radius a: 1/10, 1/20 and 1/40 metres, giving effective aspect ratios of 10, 20 and 40. The VMEC code (Hirshman & Whitson 1983) is used to solve for the MHD equilibrium inside the constructed boundaries without making the near-axis expansion, and the VMEC results are converted to Boozer coordinates using the BOOZ_XFORM code (Sanchez *et al.* 2000). From these results, we can identify $\tilde{B}_{2s}^{(0)}$ and $\tilde{B}_{2c}^{(0)}$ as the magnitudes of the $\sin 2\theta$ and $\cos 2\theta$ modes at the boundary (scaled by $1/a^2$), and identify $\tilde{B}_{0}^{(2)}$ as the magnitude of B on the magnetic axis after subtracting off $B_0 = 1$ T and scaling by $1/a^2$. We can also identify $\tilde{B}_{2}^{(0)}$ as the difference between the θ-independent Fourier mode at the boundary and axis, scaled by $1/a^2$. These values of $\tilde{B}_{20}^{(0)}$, $\tilde{B}_{2s}^{(0)}$, $\tilde{B}_{2c}^{(0)}$ and $\tilde{B}_{0}^{(2)}$ extracted from the finite-aspect-ratio VMEC solutions are displayed in figure 4 as solid curves. It can be seen that there is excellent agreement between the quantities derived from finite-aspect-ratio VMEC solutions and the near-axis analysis, and the agreement improves as a decreases.

Figure 5 provides another perspective on the same quasi-axisymmetric example. Panel (*a*) shows the total field strength at the constructed $a = 1/10$ boundary, as computed by finite-aspect-ratio computations with VMEC and BOOZ_XFORM.
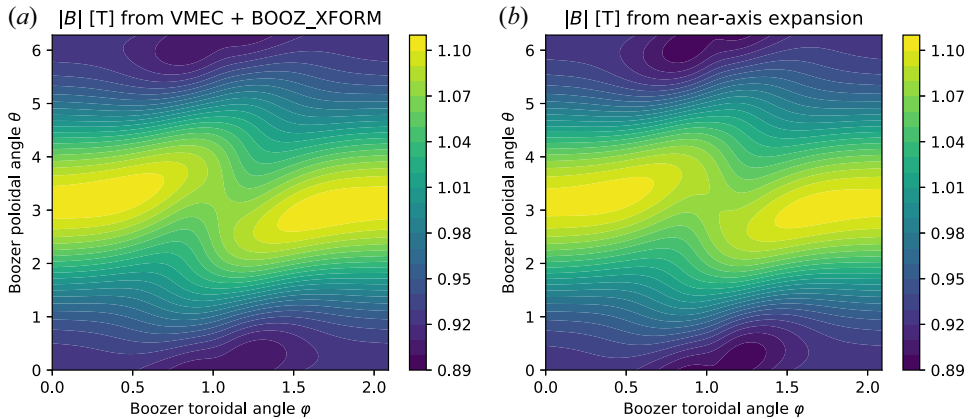
FIGURE 5. Total magnetic field strength on the boundary of the constructed quasi-axisymmetric example of § 5.1 of Landreman *et al.* (2019) for aspect ratio 10. Panel (*a*) is the result from a finite-aspect-ratio MHD equilibrium solution inside the constructed boundary. Panel (*b*) is the result of the near-axis analysis, considering both the $O(\epsilon)$ quasisymmetric component and the $O(\epsilon^2)$ non-quasisymmetric component. The two methods evidently agree well.

Panel (*b*) shows the boundary field strength predicted for this case by the near-axis expansion, (5.12), including both the $O(\epsilon)$ quasisymmetric component and the $O(\epsilon^2)$ quasisymmetry-breaking terms. Good agreement is apparent.

6. Discussion and conclusions

In this paper we have derived several diagnostic quantities that can be used to assess stellarator configurations. These diagnostics can all be computed directly from a solution of the near-axis equilibrium equations of Garren & Boozer (1991*a*). The first diagnostics are the scale lengths $L_{\nabla B}$ and $L_{\nabla \nabla B}$. These quantities could be maximized, or one could require that these quantities be above some threshold, since a short scale length probably implies that an electromagnetic coil must be close to the magnetic axis. Instead, coils should be far from the magnetic axis in order to have sufficient space for a vacuum vessel and blanket, and to reduce ripple in B. The next diagnostic derived was r_c, the maximum minor radius at which the flux surfaces are smooth and nested. This quantity should be maximized, or one should require this quantity to be above some threshold, since otherwise a near-axis configuration may be limited to very high aspect ratio. Moreover, we showed r_c is related to an equilibrium β limit, so ensuring r_c is sufficiently large at the desired pressure should ensure that this β limit is not too severe. The last diagnostic derived here was the field strength at $O(r^2)$ for a configuration constructed to have a desired B at $O(r^1)$. This 'error field' could be minimized using numerical optimization while varying the axis shape and other model parameters to ensure that deviations from quasisymmetry or omnigenity are not too great.

These diagnostics can all be computed by algebraic manipulation of the solution of the near-axis equilibrium equations, without requiring a finite-aspect-ratio numerical equilibrium from a code such as VMEC. Therefore these diagnostics can all be computed within the time scale of a few milliseconds on which the near-axis equilibrium equations are solved (Landreman & Sengupta 2019; Landreman *et al.* 2019), orders of magnitude faster than the objective function evaluations of traditional stellarator optimization. In future work, we intend to apply these diagnostics to both optimization and to scans over

the parameter space of the near-axis equations. Due to the speed by which the near-axis equations and these diagnostics can be evaluated, it is feasible to conduct wide and high-resolution scans over parameter space.

Another application of a result of this paper is single-stage optimization of coil shapes for quasisymmetry. This technique, explained in Giuliani *et al.* (2020), depends fundamentally on the ∇B tensor derived here, (3.12). There are many possible generalizations of this work, such as extending the method to $O(r^2)$ quasisymmetry using the $\nabla \nabla B$ tensor from § 3, and including other figures of merit in the optimization beyond those in Giuliani *et al.* (2020).

Besides the applications mentioned above, several other avenues for future work are evident. First, it would be useful to verify the accuracy of $L_{\nabla B}$ or $\min(L_{\nabla B}, L_{\nabla \nabla B})$ as surrogates for coil complexity by comparison with coil designs. For example, a database of stellarator equilibria could be gathered, and coils could be computed for each configuration using a code such as REGCOIL (Landreman 2017) or FOCUS (Zhu *et al.* 2018). The correlation could then be checked between the complexity of these coils versus $L_{\nabla B}$ or $\min(L_{\nabla B}, L_{\nabla \nabla B})$, depending on the order of expansion used. If the correlation is reasonably good, it may then be useful to compute the $L_{\nabla B}$ or $\min(L_{\nabla B}, L_{\nabla \nabla B})$ diagnostics not only for near-axis equilibrium solutions but also for finite-aspect-ratio MHD solutions, from codes such as VMEC. One could maximize $L_{\nabla B}$ and $L_{\nabla \nabla B}$ in traditional optimization of the plasma boundary shape to ensure that the equilibria obtained do not require very close coils. A second opportunity for future work could be to repeat the method of § 5 at next order, computing the $O(r^3)$ field strength for a configuration that was constructed to have a desirable B through $O(r^2)$. To extend the calculation to next order in this way, the algebra will be quite complicated, but it may be possible to make progress with assistance from symbolic algebra software. Finally, using the knowledge of flux surface singularities from § 4, it may be possible to put a singularity at a desired location for a divertor.

Editor Per Helander thanks the referees for their advice in evaluating this article.

Declaration of interests

The author reports no conflict of interest.

Funding

Appendix A. Coefficients in the Jacobian

The g_2 terms in the Jacobian $\sqrt{g}$ in §§ 4.1 and 4.2 are

$$
\begin{aligned}
g_{20} &= 2\kappa\ell' X_{20}(X_{1s}Y_{1c} - X_{1c}Y_{1s}) + X_{2c}(-\kappa\ell'(X_{1c}Y_{1s} + X_{1s}Y_{1c}) + 4\ell'Y_{2s}) \\
&\quad + X_{2s}(\kappa\ell'(X_{1c}Y_{1c} - X_{1s}Y_{1s}) - 4\ell'Y_{2c}) + 2\kappa\ell'X_{1c}X_{1s}Y_{2c} + \kappa\ell'Y_{2s}\left(X_{1s}^2 - X_{1c}^2\right) \\
&\quad - Z_{20}(X_{1c}Y_{1s} - X_{1s}Y_{1c})' + Z'_{20}(X_{1c}Y_{1s} - X_{1s}Y_{1c}) \\
&\quad + Z_{2s}\left(-\ell'\tau V_3 + Y_{1c}X'_{1c} - X_{1c}Y'_{1c} - Y_{1s}X'_{1s} + X_{1s}Y'_{1s}\right) \\
&\quad + Z_{2c}\left(\ell'\tau V_2 - Y_{1s}X'_{1c} + X_{1c}Y'_{1s} - Y_{1c}X'_{1s} + X_{1s}Y'_{1c}\right),
\end{aligned}
\tag{A 1}
$$

$$g_{2s} = X_{20}(\kappa\ell'(X_{1c}Y_{1c} - X_{1s}Y_{1s}) - 4\ell'Y_{2c}) + X_{2c}(-\kappa\ell'(X_{1c}Y_{1c} + X_{1s}Y_{1s}) + 4\ell'Y_{20})$$
$$+ 2\kappa\ell'X_{2s}(X_{1s}Y_{1c} - X_{1c}Y_{1s}) + \kappa\ell'Y_{20}\left(X_{1s}^2 - X_{1c}^2\right) + \kappa\ell'Y_{2c}\left(X_{1c}^2 + X_{1s}^2\right)$$
$$+ Z_{20}\left(-\ell'\tau V_3 + Y_{1c}X_{1c}' - X_{1c}Y_{1c}' - Y_{1s}X_{1s}' + X_{1s}Y_{1s}'\right)$$
$$- Z_{2s}(X_{1c}Y_{1s} - X_{1s}Y_{1c})' + Z_{2s}'(X_{1s}Y_{1c} - X_{1c}Y_{1s})$$
$$+ Z_{2c}\left(\ell'\tau V_1 - Y_{1c}X_{1c}' + X_{1c}Y_{1c}' - Y_{1s}X_{1s}' + X_{1s}Y_{1s}'\right), \tag{A 2}$$

$$g_{2c} = X_{20}(-\kappa\ell'(X_{1c}Y_{1s} + X_{1s}Y_{1c}) + 4\ell'Y_{2s}) + 2\kappa\ell'X_{2c}(X_{1s}Y_{1c} - X_{1c}Y_{1s})$$
$$+ X_{2s}(\kappa\ell'(X_{1c}Y_{1c} + X_{1s}Y_{1s}) - 4\ell'Y_{20}) + 2\kappa\ell'X_{1c}X_{1s}Y_{20} + \kappa\ell'Y_{2s}\left(-X_{1c}^2 - X_{1s}^2\right)$$
$$+ Z_{20}\left(\ell'\tau V_2 - Y_{1s}X_{1c}' + X_{1c}Y_{1s}' - Y_{1c}X_{1s}' + X_{1s}Y_{1c}'\right)$$
$$+ Z_{2s}\left(-\ell'\tau V_1 + Y_{1c}X_{1c}' - X_{1c}Y_{1c}' + Y_{1s}X_{1s}' - X_{1s}Y_{1s}'\right)$$
$$- Z_{2c}(X_{1c}Y_{1s} - X_{1s}Y_{1c})' + Z_{2c}'(X_{1c}Y_{1s} - X_{1s}Y_{1c}), \tag{A 3}$$

where, as in Garren & Boozer (1991*b*), $V_1 = X_{1s}^2 + X_{1c}^2 + Y_{1s}^2 + Y_{1c}^2$, $V_2 = 2(X_{1s}X_{1c} + Y_{1s}Y_{1c})$ and $V_3 = X_{1c}^2 - X_{1s}^2 + Y_{1c}^2 - Y_{1s}^2$.

Appendix B. Garren–Boozer equations in axisymmetry

In this appendix we show how Garren and Boozer's near-axis equilibrium equations reduce in the case of axisymmetry. These results are used in § 4.4. Garren and Boozer's equations for general non-axisymmetric geometry are given in the appendix of Garren & Boozer (1991*a*) and in appendix B of LS; here we will refer to equation numbers in the latter reference. In axisymmetry, with the magnetic axis at major radius R_0, then $\ell' = R_0$, $G_0 = s_G B_0 R_0$, $\kappa = 1/R_0$ and $\tau = 0$. Along the magnetic axis, $\varphi = \phi$ where ϕ is the standard toroidal angle. Equation (A26) of LS becomes

$$\iota_0 = \frac{2s_G R_0^3 \bar{\eta}^2 I_2}{B_0(\bar{\eta}^4 R_0^4 + 1 + \sigma^2)}, \tag{B 1}$$

where σ is now a constant parameter. The shape of the flux surfaces to $O(r)$ is given by

$$X_{1c} = \bar{\eta}R_0, \quad Y_{1s} = \frac{s_G s_\psi}{\bar{\eta}R_0}, \quad Y_{1c} = \frac{s_G s_\psi \sigma}{\bar{\eta}R_0}. \tag{B 2a–c}$$

The corresponding elongation $\kappa_e(\phi)$ of the surfaces can be computed using (B.4) of Landreman & Sengupta (2018) as follows:

$$\kappa_e = \left[V_1 + \sqrt{V_1^2 - 4q^2}\right]/(2|q|), \tag{B 3}$$

where $V_1 = X_{1s}^2 + X_{1c}^2 + Y_{1s}^2 + Y_{1c}^2$ and $q = X_{1s}Y_{1c} - X_{1c}Y_{1s} = -s_G s_\psi$. The result in axisymmetry is

$$\kappa_e = \frac{1}{2}\left[\bar{\eta}^2 R_0^2 + \frac{1}{\bar{\eta}^2 R_0^2}(1 + \sigma^2) + \sqrt{\bar{\eta}^4 R_0^4 - 2 + 2\sigma^2 + \frac{1}{\bar{\eta}^4 R_0^4}(1 + \sigma^2)^2}\right]. \tag{B 4}$$

For stellarator symmetry ($\sigma = 0$), this expression simplifies to $\kappa_e = \max(\bar{\eta}^2 R_0^2, \bar{\eta}^{-2} R_0^{-2})$. Under the further assumption of $\bar{\eta} = \pm 1/R_0$, so the elongation is 1,

then $|\iota_0| = R_0 I_2/B_0$, which matches the well known formula for a high-aspect-ratio tokamak $|\iota| \approx (R_0/r)(B_\theta/B_0)$, with B_θ the poloidal field computed from Ampere's Law.

Returning to the case of general σ and $\bar{\eta}$, LS eq (A42) reduces to

$$C\iota_0 \left[-\frac{4\mu_0 p_2 \sigma}{B_0^2} + \bar{\eta}^2(\sigma - 4s_G s_\psi R_0 Y_{2c}) \right.$$
$$\left. -\frac{2}{B_0}\left(B_{2s}[1 - 3\bar{\eta}^4 R_0^4 - 3\sigma^2] + 2B_{20}\sigma + 4B_{2c}\sigma\right) \right] = 0, \qquad (B\,5)$$

where C is a non-zero constant. Therefore, as long as $\iota_0 \neq 0$,

$$Y_{2c} = s_G s_\psi \left[-\frac{\mu_0 p_2 \sigma}{B_0^2 \bar{\eta}^2 R_0} + \frac{\sigma}{4R_0} - \frac{B_{2s}[1 - 3\bar{\eta}^4 R_0^4 - 3\sigma^2] + 2B_{20}\sigma + 4B_{2c}\sigma}{2B_0 \bar{\eta}^2 R_0} \right]. \qquad (B\,6)$$

Then LS eq (A41) gives, upon substitution of (B 6),

$$\frac{B_{20}}{B_0} = -\frac{\mu_0 p_2}{B_0^2} + \frac{1}{3 - \bar{\eta}^4 R_0^4 + 3\sigma^2}\left[-\frac{\mu_0 p_2 F^2}{2I_2^2 R_0^2} + 3(\bar{\eta}^4 R_0^4 - 1 - 3\sigma^2)\frac{B_{2c}}{B_0} \right.$$
$$\left. + 6\sigma(\bar{\eta}^4 R_0^4 + \sigma^2)\frac{B_{2s}}{B_0} + \frac{\bar{\eta}^2}{2}(7 - 2\bar{\eta}^4 R_0^4 + 4\sigma^2) + \frac{4I_2^2 \bar{\eta}^6 R_0^6 (F - 3)}{B_0^2 F^2} \right], \qquad (B\,7)$$

where $F = \bar{\eta}^4 R_0^4 + \sigma^2 + 1$. The rest of the coefficients needed to describe the flux surface shapes to $O(r^2)$ are $Z_{20} = 0$,

$$Z_{2s} = \frac{s_G I_2 (F - 2)}{2B_0 F}, \qquad Z_{2c} = -\frac{s_G I_2 \sigma}{B_0 F}, \qquad (B\,8)$$

$$X_{20} = \frac{I_2^2 \bar{\eta}^2 R_0^3}{B_0^2 F} - \frac{\bar{\eta}^2 R_0}{2} + \frac{\mu_0 p_2 R_0}{B_0^2} + \frac{B_{20}}{B_0}R_0, \qquad (B\,9)$$

$$X_{2s} = \frac{2I_2^2 \bar{\eta}^2 R_0^3 \sigma}{B_0^2 F^2} + \frac{B_{2s}}{B_0}R_0, \qquad (B\,10)$$

$$X_{2c} = \frac{I_2^2 \bar{\eta}^2 R_0^3 (F - 2)}{B_0^2 F^2} - \frac{\bar{\eta}^2 R_0}{2} + \frac{B_{2c}}{B_0}R_0, \qquad (B\,11)$$

$$Y_{20} = Y_{2c} + \frac{s_G s_\psi}{\bar{\eta}^2 R_0}\left[\frac{\mu_0 p_2 \sigma}{B_0^2} - \frac{B_{2s} + (B_{2c} - B_{20})\sigma}{B_0} \right], \qquad (B\,12)$$

$$Y_{2s} = -\frac{2s_G s_\psi I_2^2 \bar{\eta}^4 R_0^5}{B_0^2 F^2} + \frac{s_G s_\psi}{2R_0} - \frac{s_G s_\psi}{\bar{\eta}^2 R_0}\left[\frac{\mu_0 p_2}{B_0^2} + \frac{B_{20} + B_{2c} - B_{2s}\sigma}{B_0} \right]. \qquad (B\,13)$$

We can thus consider the equilibrium to be parameterized by the constants R_0, B_0, I_2, $\bar{\eta}$, σ, p_2, B_{2c} and B_{2s}; then B_{20} is determined by (B 7), and the surface shapes are given by (B 2a–c), (B 6) and (B 8)–(B 13). The fact that $Z_2 \neq 0$ in general reflects the fact that the Boozer toroidal angle is generally only identical to the cylindrical coordinate angle ϕ on the axis.

We now demonstrate the above results are consistent with the textbook result for the Shafranov shift in a tokamak with circular cross-section. First, to consider surfaces that are circles to $O(r)$, we set $\sigma = 0$ and $\bar{\eta} = 1/R_0$, so $F = 2$. Then to consider surfaces that remain circular to $O(r^2)$, we plug (2.6) and (2.9) (and the analogous expansion of Y)

into the equation for a shifted circle, $(X - \Delta)^2 + Y^2 = r^2$, assuming $\Delta = r^2\Delta_2 + O(r^3)$. Collecting terms with shared θ dependence, one finds $\Delta_2 = X_{20} + X_{2c}$ and $X_{2c} = s_G s_\psi Y_{2s}$. Then (B 7)–(B 13) give the distance between the axis and the geometric centre of the surface at radius r to be

$$\Delta = r^2 \left(\frac{1}{8R_0} - \frac{\mu_0 p_2}{2I_2^2 R_0} \right) = r^2 \left(\frac{1}{8R_0} - \frac{\mu_0 p_2 R_0}{2\iota_0^2 B_0^2} \right). \tag{B 14}$$

For comparison, analytic reduction of the Grad-Shafranov equation for shifted-circle geometry gives ((150) in Hazeltine & Meiss (1992) or (3.6.7) in Wesson (2004))

$$\frac{\mathrm{d}}{\mathrm{d}r} \left(rB_\theta^2 \frac{\mathrm{d}\Delta}{\mathrm{d}r} \right) = \frac{r}{R_0} \left(B_\theta^2 - 2\mu_0 r \frac{\mathrm{d}p}{\mathrm{d}r} \right), \tag{B 15}$$

where $B_\theta = \iota_0 B_0 r / R_0$. Substitution of $p \approx p_0 + r^2 p_2$ and $\Delta \approx r^2\Delta_2$ then gives the same near-axis shift (B 14) as the Garren–Boozer equations.

Appendix C. Details of the error field calculation

In this appendix, details are given of the calculation leading to the result in § 5.2. We first note that by examining (5.11) in appendix B.2 of LS, it was shown that for the first-order construction, $f^{(0)} = 0$, $f^{(1)} = 0$, $t^{(0)} = 0$, $\tilde{r}_0^{(0)}(\varphi) = r_0(\varphi)$, $\tilde{t}^{(0)}(\varphi) = t(\varphi)$, $\tilde{n}^{(0)}(\varphi) = n(\varphi)$, $\tilde{b}^{(0)}(\varphi) = b(\varphi)$, $\tilde{r}_0^{(1)} = 0$, $\tilde{t}^{(1)} = 0$, $\tilde{n}^{(1)} = 0$, $\tilde{b}^{(1)} = 0$, $\tilde{X}_1^{(0)}(\vartheta, \varphi) = X_1(\vartheta, \varphi)$, $\tilde{Y}_1^{(0)}(\vartheta, \varphi) = Y_1(\vartheta, \varphi)$. $\tilde{G}_0^{(0)} = G_0$, $\tilde{G}_0^{(1)} = 0$ and $\tilde{\iota}_0^{(0)} = \iota_0$.

C.1. *Helical modes*

We begin with the $O(\epsilon^2)$ terms in (5.11), (B 12) in LS, copied here:

$$\begin{aligned}
X_2(\vartheta, \varphi)\boldsymbol{n} &+ Y_2(\vartheta, \varphi)\boldsymbol{b} + Z_2(\vartheta, \varphi)\boldsymbol{t} \\
&= \tilde{\boldsymbol{r}}_0^{(2)} + f^{(1)}\tilde{\boldsymbol{r}}_0^{(1)\prime} + f^{(2)}(\vartheta, \varphi)\boldsymbol{r}_0' + \tfrac{1}{2}f^{(1)2}\boldsymbol{r}_0'' \\
&\quad + \tilde{X}_2^{(0)}(\vartheta, \varphi)\boldsymbol{n} + \tilde{X}_1^{(1)}(\vartheta, \varphi)\boldsymbol{n} + X_1(\vartheta, \varphi)\tilde{\boldsymbol{n}}^{(1)} + f^{(1)}X_1(\vartheta, \varphi)\boldsymbol{n}' \\
&\quad + t^{(1)}(\vartheta, \varphi)\boldsymbol{n}\partial_1 X_1(\vartheta, \varphi) + f^{(1)}\boldsymbol{n}\partial_2 X_1(\vartheta, \varphi) \\
&\quad + \tilde{Y}_2^{(0)}(\vartheta, \varphi)\boldsymbol{b} + \tilde{Y}_1^{(1)}(\vartheta, \varphi)\boldsymbol{b} + Y_1(\vartheta, \varphi)\tilde{\boldsymbol{b}}^{(1)} + f^{(1)}Y_1(\vartheta, \varphi)\boldsymbol{b}' \\
&\quad + t^{(1)}(\vartheta, \varphi)\boldsymbol{b}\partial_1 Y_1(\vartheta, \varphi) + f^{(1)}\boldsymbol{b}\partial_2 Y_1(\vartheta, \varphi) + \tilde{Z}_2^{(0)}(\vartheta, \varphi)\boldsymbol{t}, \tag{C 1}
\end{aligned}$$

where ∂_1 and ∂_2 indicate partial derivatives with respect to the first or second argument. To simplify notation, the argument is not shown for functions of only φ. We assume the construction is done only through $O(\epsilon^1)$, so the left-hand side of (C 1) is zero. As mentioned above, in appendix B of LS, it was proved that $\tilde{\boldsymbol{r}}_0^{(1)} = 0$, $\tilde{\boldsymbol{n}}^{(1)} = 0$, $\tilde{\boldsymbol{b}}^{(1)} = 0$ and $f^{(1)} = 0$ for the first-order construction. Then (C 1) reduces to

$$\begin{aligned}
0 &= \tilde{\boldsymbol{r}}_0^{(2)} + f^{(2)}(\vartheta, \varphi)\boldsymbol{r}_0' + \tilde{X}_2^{(0)}(\vartheta, \varphi)\boldsymbol{n} + \tilde{X}_1^{(1)}(\vartheta, \varphi)\boldsymbol{n} + t^{(1)}(\vartheta, \varphi)\boldsymbol{n}\partial_1 X_1(\vartheta, \varphi) \\
&\quad + \tilde{Y}_2^{(0)}(\vartheta, \varphi)\boldsymbol{b} + \tilde{Y}_1^{(1)}(\vartheta, \varphi)\boldsymbol{b} + t^{(1)}(\vartheta, \varphi)\boldsymbol{b}\partial_1 Y_1(\vartheta, \varphi) + \tilde{Z}_2^{(0)}(\vartheta, \varphi)\boldsymbol{t}. \tag{C 2}
\end{aligned}$$

The $\boldsymbol{t}$ component is

$$0 = \boldsymbol{t} \cdot \tilde{\boldsymbol{r}}_0^{(2)} + \ell' f^{(2)}(\vartheta, \varphi) + \tilde{Z}_2^{(0)}(\vartheta, \varphi). \tag{C 3}$$

The last quantity, $\tilde{Z}_2^{(0)}(\vartheta, \varphi)$, is known since it is a unique function of X_1 and Y_1, (A27)–(A29) of LS. Therefore,

$$f^{(2)}(\vartheta, \varphi) = \bar{f}^{(2)}(\varphi) - \left[\tilde{Z}_{2s}^{(0)}(\varphi)\sin 2\vartheta + \tilde{Z}_{2c}^{(0)}(\varphi)\cos 2\vartheta\right]/\ell', \tag{C 4}$$

where $\bar{f}^{(2)}(\varphi)$ has not yet been determined.

The argument in the paragraph of LS containing (B24)–(B26) applies here, except with $X_{2s} = X_{2c} = Y_{2s} = Y_{2c} = 0$, so

$$t^{(1)}(\vartheta, \varphi) = \bar{t}^{(1)} + \frac{2}{X_{1s}^2 + X_{1c}^2}\left[\left(X_{1c}\tilde{X}_{2s}^{(0)} - X_{1s}\tilde{X}_{2c}^{(0)}\right)\cos\vartheta - \left(X_{1s}\tilde{X}_{2s}^{(0)} + X_{1c}\tilde{X}_{2c}^{(0)}\right)\sin\vartheta\right] \tag{C 5}$$

and

$$t^{(1)}(\vartheta, \varphi) = \bar{t}^{(1)} + \frac{2}{Y_{1s}^2 + Y_{1c}^2}\left[\left(Y_{1c}\tilde{Y}_{2s}^{(0)} - Y_{1s}\tilde{Y}_{2c}^{(0)}\right)\cos\vartheta - \left(Y_{1s}\tilde{Y}_{2s}^{(0)} + Y_{1c}\tilde{Y}_{2c}^{(0)}\right)\sin\vartheta\right], \tag{C 6}$$

where $\bar{t}^{(1)}(\varphi)$ has yet to be determined. Equating the $\sin\vartheta$ and $\cos\vartheta$ modes of (C 5) and (C 6), we find

$$\frac{\tilde{X}_{2s}^{(0)}X_{1s} + \tilde{X}_{2c}^{(0)}X_{1c}}{X_{1s}^2 + X_{1c}^2} = \frac{\tilde{Y}_{2s}^{(0)}Y_{1s} + \tilde{Y}_{2c}^{(0)}Y_{1c}}{Y_{1s}^2 + Y_{1c}^2}, \tag{C 7}$$

$$\frac{\tilde{X}_{2c}^{(0)}X_{1s} - \tilde{X}_{2s}^{(0)}X_{1c}}{X_{1s}^2 + X_{1c}^2} = \frac{\tilde{Y}_{2c}^{(0)}Y_{1s} - \tilde{Y}_{2s}^{(0)}Y_{1c}}{Y_{1s}^2 + Y_{1c}^2}. \tag{C 8}$$

Equations (C 7) and (C 8) provide two linear equations constraining the subscript-2 shape coefficients $\tilde{X}_{20}^{(0)}$, $\tilde{X}_{2s}^{(0)}$, $\tilde{X}_{2c}^{(0)}$, $\tilde{Y}_{20}^{(0)}$, $\tilde{Y}_{2s}^{(0)}$ and $\tilde{Y}_{2c}^{(0)}$. These six quantities are also constrained by (A32), (A33), (A41) and (A42) of LS. ((A32)–(A33) express the equality at second order of the t components of the covariant and contravariant B representations, while (A41)–(A42) are the second-order version of the relation between on-axis transform, axis torsion, rotating elongation and toroidal current.) Thus we have six linear equations for six unknowns, so we can solve the linear system. Four of the six equations are algebraic rather than differential, so the system can be reduced to a system of two equations and two unknowns for faster numerical solution. This reduction can be accomplished by considering $\tilde{X}_{20}^{(0)}$ and $\tilde{Y}_{20}^{(0)}$ as the unknowns, solving (C 7) and (C 8) with LS equations (A32) and (A33) to obtain

$$\tilde{X}_{2s}^{(0)} = \frac{1}{2b}\left[(X_{1s}Y_{1s} - X_{1c}Y_{1c})\tilde{X}_{20}^{(0)} + (X_{1c}^2 - X_{1s}^2)\tilde{Y}_{20}^{(0)} + \frac{s_G\bar{B}X_{1s}X_{1c}\kappa}{B_0}\right], \tag{C 9}$$

$$\tilde{X}_{2c}^{(0)} = \frac{1}{2b}\left[(X_{1s}Y_{1c} + X_{1c}Y_{1s})\tilde{X}_{20}^{(0)} - 2X_{1c}X_{1s}\tilde{Y}_{20}^{(0)} + \frac{s_G\bar{B}(X_{1c}^2 - X_{1s}^2)\kappa}{2B_0}\right], \tag{C 10}$$

$$\tilde{Y}_{2s}^{(0)} = \frac{1}{2b}\left[(Y_{1s}^2 - Y_{1c}^2)\tilde{X}_{20}^{(0)} + (X_{1c}Y_{1c} - X_{1s}Y_{1s})\tilde{Y}_{20}^{(0)} + \frac{s_G\bar{B}(X_{1s}Y_{1c} + X_{1c}Y_{1s})\kappa}{2B_0}\right], \tag{C 11}$$

$$\tilde{Y}_{2c}^{(0)} = \frac{1}{2b}\left[2Y_{1s}Y_{1c}\tilde{X}_{20}^{(0)} - (X_{1s}Y_{1c} + X_{1c}Y_{1s})\tilde{Y}_{20}^{(0)} + \frac{s_G\bar{B}(X_{1c}Y_{1c} - X_{1s}Y_{1s})\kappa}{2B_0}\right], \tag{C 12}$$

where $b = X_{1s}Y_{1c} - X_{1c}Y_{1s} = -s_G\bar{B}/B_0$. With $\tilde{X}_{20}^{(0)}$, $\tilde{X}_{2s}^{(0)}$ and $\tilde{X}_{2c}^{(0)}$ now known, we can evaluate the tilde version of (A34)-(A36) in LS (which express equality of B^2 to the square

of the contravariant representation) to find $\tilde{B}_{20}^{(0)}$, $\tilde{B}_{2s}^{(0)}$ and $\tilde{B}_{2c}^{(0)}$. We have thus determined $\tilde{B}_2^{(0)}$, the first term on the bottom line of (5.12).

Given (C 5) and (C 6), and noting that the $\sin\vartheta$ and $\cos\vartheta$ modes of $t^{(1)}$ do not contribute to (B15)–(B17) of LS, (B29) holds after substituting $t^{(1)} \to \bar{t}^{(1)}$. Then (B30) and the argument that follows it implies $\bar{t}^{(1)} = 0$, so $\tilde{X}_{1s}^{(1)} = \tilde{X}_{1c}^{(1)} = \tilde{Y}_{1s}^{(1)} = \tilde{Y}_{1c}^{(1)} = 0$ and $\tilde{B}_1^{(1)} = 0$. This determines the middle term on the bottom line of (5.12).

Now that $\tilde{X}_2^{(0)}$, $\tilde{Y}_2^{(0)}$ and $t^{(1)}$ are known, the $\boldsymbol{n}$ and $\boldsymbol{t}$ components of (C 2) give $\tilde{\boldsymbol{r}}_0^{(2)} \cdot \boldsymbol{n}$ and $\tilde{\boldsymbol{r}}_0^{(2)} \cdot \boldsymbol{b}$. However, these components of $\tilde{\boldsymbol{r}}_0^{(2)}$ are not needed for the remainder of the calculation here, so these results will not be displayed here.

C.2. *Mirror term*

It remains to compute the quantity $\tilde{B}_0^{(2)}(\tilde{\varphi})$ in (5.12), which represents a 'mirror term' in the field strength. To compute this term, we proceed to examine the $O(\epsilon^3)$ terms in (5.11). Compared with (B32) in LS, there are two differences in the present context. First, $f^{(2)}$ now depends on ϑ, as we found in (C 4). Second, $t^{(1)}$ is non-zero, given by (C 5) or (C 6) with $\bar{t}^{(1)} = 0$. Thus, we have

$$
\begin{aligned}
X_3(\vartheta,\varphi)\boldsymbol{n} + Y_3(\vartheta,\varphi)\boldsymbol{b} + Z_3(\vartheta,\varphi)\boldsymbol{t} &= \tilde{\boldsymbol{r}}_0^{(3)} + f^{(3)}(\vartheta,\varphi)\boldsymbol{r}_0' + \boldsymbol{t}\tilde{Z}_3^{(0)}(\vartheta,\varphi) + \boldsymbol{t}\tilde{Z}_2^{(1)}(\vartheta,\varphi) \\
&\quad + \boldsymbol{n}\tilde{X}_3^{(0)}(\vartheta,\varphi) + \tilde{\boldsymbol{n}}^{(2)}X_1(\vartheta,\varphi) + \boldsymbol{n}\tilde{X}_1^{(2)}(\vartheta,\varphi) + \boldsymbol{n}\tilde{X}_2^{(1)}(\vartheta,\varphi) \\
&\quad + \boldsymbol{b}\tilde{Y}_3^{(0)}(\vartheta,\varphi) + \tilde{\boldsymbol{b}}^{(2)}Y_1(\vartheta,\varphi) + \boldsymbol{b}\tilde{Y}_1^{(2)}(\vartheta,\varphi) + \boldsymbol{b}\tilde{Y}_2^{(1)}(\vartheta,\varphi) \\
&\quad + f^{(2)}(\vartheta,\varphi)Y_1(\vartheta,\varphi)\boldsymbol{b}' + f^{(2)}(\vartheta,\varphi)X_1(\vartheta,\varphi)\boldsymbol{n}' \\
&\quad + \boldsymbol{n}f^{(2)}(\vartheta,\varphi)\partial_2 X_1(\vartheta,\varphi) + \boldsymbol{b}f^{(2)}(\vartheta,\varphi)\partial_2 Y_1(\vartheta,\varphi) \\
&\quad + t^{(2)}(\vartheta,\varphi)\boldsymbol{n}\partial_1 X_1(\vartheta,\varphi) + t^{(2)}(\vartheta,\varphi)\boldsymbol{b}\partial_1 Y_1(\vartheta,\varphi) \\
&\quad + t^{(1)}(\vartheta,\varphi)\left[\boldsymbol{n}\partial_1 \tilde{X}_2^{(0)}(\vartheta,\varphi) + \boldsymbol{b}\partial_1 \tilde{Y}_2^{(0)}(\vartheta,\varphi) + \boldsymbol{t}\partial_1 \tilde{Z}_2^{(0)}(\vartheta,\varphi)\right] \\
&\quad + \tfrac{1}{2}\left[t^{(1)}(\vartheta,\varphi)\right]^2 \left[\boldsymbol{n}\partial_1^2 X_1(\vartheta,\varphi) + \boldsymbol{b}\partial_1^2 Y_1(\vartheta,\varphi)\right]. \quad (C\,13)
\end{aligned}
$$

The last two rows are new compared with (B32) in LS. We take the $\boldsymbol{n}$ component, noting $\boldsymbol{n}\cdot\boldsymbol{n}' = 0$ and $\boldsymbol{n}\cdot\tilde{\boldsymbol{n}}^{(2)} = 0$, since the latter is the $O(\epsilon^2)$ term in $|\tilde{\boldsymbol{n}}|^2 = 1$. The $\sin\vartheta$ and $\cos\vartheta$ modes of the result are similar to (B33) of LS, but with a few extra terms,

$$
\begin{aligned}
\tilde{X}_{1s}^{(2)} &= X_{3s1} - \tilde{X}_{3s1}^{(0)} - \boldsymbol{n}\cdot\tilde{\boldsymbol{b}}^{(2)}Y_{1s} - X_{1s}t_{sc}^{(2)} + X_{1c}t_{ss}^{(2)} - T_{Xs} \\
&\quad + \left(\frac{Y_{1c}\tilde{Z}_{2s}^{(0)} - Y_{1s}\tilde{Z}_{2c}^{(0)}}{2\ell'} - \bar{f}^{(2)}Y_{1s}\right)\boldsymbol{n}\cdot\boldsymbol{b}' - \left(\frac{\tilde{Z}_{2c}^{(0)}}{2\ell'} + \bar{f}^{(2)}\right)X_{1s}' + \frac{\tilde{Z}_{2s}^{(0)}}{2\ell'}X_{1c}', \quad (C\,14)
\end{aligned}
$$

$$
\begin{aligned}
\tilde{X}_{1c}^{(2)} &= X_{3c1} - \tilde{X}_{3c1}^{(0)} - \boldsymbol{n}\cdot\tilde{\boldsymbol{b}}^{(2)}Y_{1c} - X_{1s}t_{cc}^{(2)} + X_{1c}t_{sc}^{(2)} - T_{Xc} \\
&\quad + \left(\frac{Y_{1c}\tilde{Z}_{2c}^{(0)} + Y_{1s}\tilde{Z}_{2s}^{(0)}}{2\ell'} - \bar{f}^{(2)}Y_{1c}\right)\boldsymbol{n}\cdot\boldsymbol{b}' + \left(\frac{\tilde{Z}_{2c}^{(0)}}{2\ell'} - \bar{f}^{(2)}\right)X_{1c}' + \frac{\tilde{Z}_{2s}^{(0)}}{2\ell'}X_{1s}', \quad (C\,15)
\end{aligned}
$$

where (C 4) has been used,

$$
t_{ss}^{(2)}(\varphi) = \pi^{-1}\int_0^{2\pi} \mathrm{d}\vartheta\, t^{(2)}(\vartheta,\varphi)\sin^2\vartheta, \quad (C\,16)
$$

$$
t_{sc}^{(2)}(\varphi) = \pi^{-1}\int_0^{2\pi} \mathrm{d}\vartheta\, t^{(2)}(\vartheta,\varphi)\sin\vartheta\cos\vartheta, \quad (C\,17)
$$

$$t_{cc}^{(2)}(\varphi) = \pi^{-1} \int_0^{2\pi} d\vartheta \, t^{(2)}(\vartheta, \varphi) \cos^2 \vartheta, \tag{C 18}$$

$$T_{Xs} = \frac{1}{\pi} \int_0^{2\pi} d\vartheta \, \sin \vartheta \left(t^{(1)}(\vartheta, \varphi) \partial_1 \tilde{X}_2^{(0)}(\vartheta, \varphi) + \frac{1}{2} \left[t^{(1)}(\vartheta, \varphi) \right]^2 \partial_1^2 X_1(\vartheta, \varphi) \right) \tag{C 19}$$

and T_{Xc} is defined by (C 19) with $\sin \to \cos$. Similarly, taking the $\boldsymbol{b}$ component of (C 13), the $\sin \vartheta$ and $\cos \vartheta$ modes give

$$\tilde{Y}_{1s}^{(2)} = Y_{3s1} - \tilde{Y}_{3s1}^{(0)} - \boldsymbol{b} \cdot \tilde{\boldsymbol{n}}^{(2)} X_{1s} - Y_{1s} t_{sc}^{(2)} + Y_{1c} t_{ss}^{(2)} - T_{Ys}$$
$$+ \left(\frac{X_{1c} \tilde{Z}_{2s}^{(0)} - X_{1s} \tilde{Z}_{2c}^{(0)}}{2\ell'} - \bar{f}^{(2)} X_{1s} \right) \boldsymbol{b} \cdot \boldsymbol{n}' - \left(\frac{\tilde{Z}_{2c}^{(0)}}{2\ell'} + \bar{f}^{(2)} \right) Y_{1s}' + \frac{\tilde{Z}_{2s}^{(0)}}{2\ell'} Y_{1c}', \tag{C 20}$$

$$\tilde{Y}_{1c}^{(2)} = Y_{3c1} - \tilde{Y}_{3c1}^{(0)} - \boldsymbol{b} \cdot \tilde{\boldsymbol{n}}^{(2)} X_{1c} - Y_{1s} t_{cc}^{(2)} + Y_{1c} t_{sc}^{(2)} - T_{Yc}$$
$$+ \left(\frac{X_{1c} \tilde{Z}_{2c}^{(0)} + X_{1s} \tilde{Z}_{2s}^{(0)}}{2\ell'} - \bar{f}^{(2)} X_{1c} \right) \boldsymbol{b} \cdot \boldsymbol{n}' + \left(\frac{\tilde{Z}_{2c}^{(0)}}{2\ell'} - \bar{f}^{(2)} \right) Y_{1c}' + \frac{\tilde{Z}_{2s}^{(0)}}{2\ell'} Y_{1s}', \tag{C 21}$$

where

$$T_{Ys} = \frac{1}{\pi} \int_0^{2\pi} d\vartheta \, \sin \vartheta \left(t^{(1)}(\vartheta, \varphi) \partial_1 \tilde{Y}_2^{(0)}(\vartheta, \varphi) + \frac{1}{2} \left[t^{(1)}(\vartheta, \varphi) \right]^2 \partial_1^2 Y_1(\vartheta, \varphi) \right) \tag{C 22}$$

and T_{Yc} is defined by (C 22) with $\sin \to \cos$.

We now consider the $O(\epsilon^2)$ terms of (A21) in LS, expressing the condition that the $O(r^1 a^2)$ flux surfaces enclose the proper toroidal flux:

$$X_{1c} \tilde{Y}_{1s}^{(2)} + \tilde{X}_{1c}^{(2)} Y_{1s} - X_{1s} \tilde{Y}_{1c}^{(2)} - \tilde{X}_{1s}^{(2)} Y_{1c} = -s_G \bar{B} \tilde{B}_0^{(2)} / B_0^2. \tag{C 23}$$

Substituting (C 14), (C 15), (C 20) and (C 21) into this expression, terms involving $\tilde{\boldsymbol{n}}^{(2)}$, $\tilde{\boldsymbol{b}}^{(2)}$ and $t^{(2)}$ cancel to leave

$$\tilde{B}_0^{(2)} = \hat{B} B_0 - \bar{f}^{(2)} B_0', \tag{C 24}$$

where

$$\hat{B} = (X_{1s} Y_{3c1} + Y_{1c} X_{3s1} - X_{1c} Y_{3s1} - Y_{1s} X_{3c1} - \tilde{q}) s_G B_0 / \bar{B}, \tag{C 25}$$

$$\tilde{q} = \tilde{Q}^{(0)} + (A/2) + T, \tag{C 26}$$

$$\tilde{Q}^{(0)}(\varphi) = \frac{(G_2 + I_2 N) \bar{B} \ell'}{2 G_0^2} + 2(\tilde{X}_{2c}^{(0)} \tilde{Y}_{2s}^{(0)} - \tilde{X}_{2s}^{(0)} \tilde{Y}_{2c}^{(0)}) + \frac{\bar{B}}{2 G_0} (\ell' \tilde{X}_{20}^{(0)} \kappa - \tilde{Z}_{20}^{(0)\prime})$$
$$+ \frac{I_2}{4 G_0} (-\ell' \tau V_1 + Y_{1c} X_{1c}' - X_{1c} Y_{1c}' + Y_{1s} X_{1s}' - X_{1s} Y_{1s}')$$
$$+ \frac{\beta_0 \bar{B}}{4 G_0} (X_{1s} Y_{1c}' + Y_{1c} X_{1s}' - X_{1c} Y_{1s}' - Y_{1s} X_{1c}'), \tag{C 27}$$

$$A = -2\tau \tilde{Z}_{2c}^{(0)} (X_{1s} X_{1c} + Y_{1s} Y_{1c}) - \tau \tilde{Z}_{2s}^{(0)} (X_{1s}^2 - X_{1c}^2 - Y_{1c}^2 + Y_{1s}^2)$$
$$+ \frac{\tilde{Z}_{2s}^{(0)}}{\ell'} (X_{1c} Y_{1c}' + Y_{1s} X_{1s}' - X_{1s} Y_{1s}' - Y_{1c} X_{1c}')$$
$$- \frac{\tilde{Z}_{2c}^{(0)}}{\ell'} (X_{1c} Y_{1s}' - Y_{1s} X_{1c}' + X_{1s} Y_{1c}' - Y_{1c} X_{1s}') \tag{C 28}$$

and $T = X_{1s}T_{Yc} + Y_{1c}T_{Xs} - X_{1c}T_{Ys} - Y_{1s}T_{Xc}$. To obtain these results we have used $\boldsymbol{b} \cdot \boldsymbol{n}' = -\boldsymbol{n} \cdot \boldsymbol{b}' = \tau\ell'$, the tilde version of (A49) of LS (in which the $\boldsymbol{t}$ components of the covariant and contravariant $\boldsymbol{B}$ representations are equated at third order), and the $d/d\varphi$ derivative of $X_{1c}Y_{1s} - X_{1s}Y_{1c} = s_G\bar{B}/B_0$. Using (C 5), (C 7) and (C 8) to eliminate $\tilde{Y}_2^{(0)}$, we find

$$T = 2\left[(\tilde{X}_{2s}^{(0)})^2 + (\tilde{X}_{2c}^{(0)})^2\right]\frac{X_{1s}Y_{1c} - X_{1c}Y_{1s}}{X_{1s}^2 + X_{1c}^2}. \tag{C 29}$$

Expressions (C 24), (C 25) and (C 27) mirror (3.10)–(3.12) of LS, where $\tilde{B}_0^{(2)}$ was computed for the $O(\epsilon^2)$ construction (in contrast to the $O(\epsilon)$ construction here). Just as described at the end of Appendix B of LS, $\bar{f}^{(2)}$ is determined by

$$\bar{f}^{(2)}(\varphi) = \left(\int_0^{\varphi} d\bar{\varphi}\,\hat{B}(\bar{\varphi})\right) + \left(\frac{1}{2} - \frac{\varphi}{2\pi}\right)\left(\int_0^{2\pi} d\bar{\varphi}\,\hat{B}(\bar{\varphi})\right) - \frac{1}{2\pi}\int_0^{2\pi} d\bar{\bar{\varphi}}\int_0^{\bar{\bar{\varphi}}} d\bar{\varphi}\,\hat{B}(\bar{\varphi}). \tag{C 30}$$

At this point we have fully determined $\tilde{B}_0^{(2)}$ in terms of known quantities. Normally, the $O(\epsilon)$ construction would be carried out with $X_{3s1} = X_{3c1} = Y_{3s1} = Y_{3c1} = 0$, so these terms would be absent in (C 25). Alternatively, non-zero values of X_{3s1}, X_{3c1}, Y_{3s1} and Y_{3c1} could be chosen in order to make $\tilde{B}_0^{(2)}$ vanish, using (3.14)–(3.15) of LS with $Q \to \tilde{q}$.

C.3. *Quasisymmetry*

In quasisymmetry, $X_{1s} = 0$ and $\bar{B} = s_\psi B_0$, so (C 9)–(C 12) reduce to

$$\tilde{X}_{2s}^{(0)} = \frac{Y_{1c}}{2Y_{1s}}\tilde{X}_{20}^{(0)} - \frac{X_{1c}}{2Y_{1s}}\tilde{Y}_{20}^{(0)}, \tag{C 31}$$

$$\tilde{X}_{2c}^{(0)} = -\frac{1}{2}\tilde{X}_{20}^{(0)} - \frac{s_G s_\psi X_{1c}\kappa}{4Y_{1s}}, \tag{C 32}$$

$$\tilde{Y}_{2s}^{(0)} = \frac{Y_{1c}^2 - Y_{1s}^2}{2X_{1c}Y_{1s}}\tilde{X}_{20}^{(0)} - \frac{Y_{1c}}{2Y_{1s}}\tilde{Y}_{20}^{(0)} - \frac{s_G s_\psi \kappa}{4}, \tag{C 33}$$

$$\tilde{Y}_{2c}^{(0)} = -\frac{Y_{1c}}{X_{1c}}\tilde{X}_{20}^{(0)} + \frac{1}{2}\tilde{Y}_{20}^{(0)} - \frac{s_G s_\psi Y_{1c}\kappa}{4Y_{1s}}. \tag{C 34}$$

REFERENCES

FREIDBERG, J. P. 2014 *Ideal MHD*. Cambridge University Press.

GARREN, D. A. & BOOZER, A. H. 1991*a* Existence of quasihelically symmetric stellarators. *Phys. Fluids* B **3**, 2822.

GARREN, D. A. & BOOZER, A. H. 1991*b* Magnetic field strength of toroidal plasma equilibria. *Phys. Fluids* B **3**, 2805.

GIULIANI, A., WECHSUNG, F., CERFON, A., STADLER, G. & LANDREMAN, M. 2020 Single-stage gradient-based stellarator coil design: optimization for near-axis quasi-symmetry. arXiv:2010.02033.

GREENE, J. M. 1965 Calculation of the singularities of a magnetic field. *Phys. Fluids* **8**, 704.

HAZELTINE, R. D. & MEISS, J. D. 1992 *Plasma Confinement*. Addison-Wesley.

HIRSHMAN, S. P. & WHITSON, J. C. 1983 Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria. *Phys. Fluids* **26**, 3553.

JORGE, R. & LANDREMAN, M. 2021 The use of near-axis magnetic fields for stellarator turbulence simulations. *Plasma Phys. Control. Fusion* **63**, 014001.

JORGE, R., SENGUPTA, W. & LANDREMAN, M. 2020*a* Construction of quasisymmetric stellarators using a direct coordinate approach. *Nucl. Fusion* **60**, 076021.

JORGE, R., SENGUPTA, W. & LANDREMAN, M. 2020b Near-axis expansion of stellarator equilibrium at arbitrary order in the distance to the axis. *J. Plasma Phys.* **86**, 905860106.

LANDREMAN, M. 2017 An improved current potential method for fast computation of stellarator coil shapes. *Nucl. Fusion* **57**, 046003.

LANDREMAN, M. 2019 Optimized quasisymmetric stellarators are consistent with the Garren-Boozer construction. *Plasma Phys. Control. Fusion* **61**, 075001.

LANDREMAN, M. 2020a Dataset on Zenodo. http://doi.org/10.5281/zenodo.4011733.

LANDREMAN, M. 2020b Dataset on Zenodo. http://doi.org/10.5281/zenodo.4294769.

LANDREMAN, M. & JORGE, R. 2020 Mercier stability of stellarators near the magnetic axis. *J. Plasma Phys.* **86**, 905860510.

LANDREMAN, M. & SENGUPTA, W. 2018 Direct construction of optimized stellarator shapes. I. Theory in cylindrical coordinates. *J. Plasma Phys.* **84**, 905840616.

LANDREMAN, M. & SENGUPTA, W. 2019 Constructing stellarators with quasisymmetry to high order. *J. Plasma Phys.* **85**, 905850608.

LANDREMAN, M., SENGUPTA, W. & PLUNK, G. G. 2019 Direct construction of optimized stellarator shapes. II. Numerical quasisymmetric solutions. *J. Plasma Phys.* **85**, 905850103.

LOIZU, J., HUDSON, S. R., NÜHRENBERG, C., GEIGER, J. & HELANDER, P. 2017 Equilibrium β-limits in classical stellarators. *J. Plasma Phys.* **83**, 715830601.

MUKHOVATOV, V. S. & SHAFRANOV, V. D. 1971 Plasma equilibrium in a tokamak. *Nucl. Fusion* **11**, 605.

PLUNK, G. G., LANDREMAN, M. & HELANDER, P. 2019 Direct construction of optimized stellarator shapes. III. Omnigenity near the magnetic axis. *J. Plasma Phys.* **85**, 905850602.

SANCHEZ, R., HIRSHMAN, S. P., WARE, A. S., BERRY, L. A. & SPONG, D. A. 2000 Ballooning stability optimization of low-aspect-ratio stellarators. *Plasma Phys. Control. Fusion* **42**, 641.

WESSON, J. 2004 *Tokamaks*, 3rd edn. Oxford University Press.

ZHU, C., HUDSON, S. H., SONG, Y. & WAN, Y. 2018 New method to design stellarator coils without the winding surface. *Nucl. Fusion* **58**, 016008.