

Measures of quasisymmetry for stellarators

E. Rodríguez ^{1,2,†}, E.J. Paul ^{1,2} and A. Bhattacharjee^{1,2}

¹Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08543, USA

²Princeton Plasma Physics Laboratory, Princeton, NJ 08540, USA

(Received 27 September 2021; revised 13 December 2021; accepted 15 December 2021)

Quasisymmetric stellarators are an attractive class of optimised magnetic confinement configurations. The property of quasisymmetry (QS) is in practice limited to be approximate, and thus the construction requires measures that quantify the deviation from the exact property. In this paper we study three measure candidates used in the literature, placing the focus on their origin and a comparison of their forms. The analysis shows clearly the lack of universality in these measures. As these metrics do not directly correspond to any physical property (except when exactly quasisymmetric), optimisation should employ additional physical metrics for guidance. It is suggested that close to QS minima, one should treat QS metrics through inequality constraints so that additional physics metrics dominate optimisation. The impact of different quasisymmetric measures on optimisation is presented through an example, for which the standard metric that weights the asymmetric Fourier modes of the field magnitude appears to perform best.

Key words: fusion plasma, plasma devices

1. Introduction

The stellarator concept (Spitzer 1958) has seen a revival in the last few decades as an option to achieve controlled thermonuclear fusion. Stellarators are three-dimensional magnetic field configurations which confine hot plasmas within. The three-dimensionality of these devices provides the necessary freedom to avoid some of the limiting features from which axisymmetric tokamaks suffer: most notably, current-driven instabilities. However, this freedom can come at a price: the three-dimensionality of the magnetic field might lead to a loss of confinement that continuous symmetry confers upon a tokamak.

Quasisymmetry (QS) (Boozer 1983; Nührenberg & Zille 1988; Tessarotto *et al.* 1996; Helander 2014; Burby, Kallinikos & MacKay 2020; Rodríguez, Helander & Bhattacharjee 2020) is a property of the field that attempts to preserve in three-dimensional stellarators the good neoclassical properties (Mynick 2006) of the tokamak. The promise of this property has led to a search of designs that bear it. Although theoretical work suggests that in an ideal, static equilibrium with isotropic pressure one may not construct exact QS solutions (Garren & Boozer 1991; Landreman & Sengupta 2018; Rodríguez & Bhattacharjee 2021*b*), one may design such configurations approximately. In practice there have been two main approaches to finding approximately quasisymmetric equilibria. One is to directly construct approximate solutions by employing an asymptotic expansion near

† Email address for correspondence: eduardor@princeton.edu

the magnetic axis (Garren & Boozer 1991; Landreman & Sengupta 2018, 2019; Rodríguez & Bhattacharjee 2021c) or near axisymmetry (Plunk & Helander 2018). The other, and most commonly employed, approach is to treat the construction of QS configurations as part of an optimisation problem (Bader *et al.* 2019; Henneberg *et al.* 2019).

To proceed with optimisation it is necessary to construct cost functions that penalise deviations from QS. This has been done in a number of contexts and forms (Bader *et al.* 2019; Henneberg *et al.* 2019; Paul *et al.* 2020). However, at the present time there appears to be a lack of comparative studies of the different forms of cost functions used. The purpose of this paper is to re-evaluate carefully the origin of some of these commonly used forms and to investigate their differences, significance and content.

The paper is organised as follows. In § 2 three different measures of QS are considered. From these basics, the most apparent formal aspects of the metrics are discussed. In § 3 the measures are formally compared on the same footing. The connection between physical properties associated with QS and the form of the metrics is then explored. Section 4 presents a more practical perspective on the measures by looking at the universality of QS measures (how they compare across different near-QS designs) and the importance of the cost functions for optimisation in practice. We close the paper with some concluding remarks and suggestions for future optimisation.

2. Measures of QS

Let us start by defining the concept of QS from the fundamental perspective of single-particle dynamics. The class of magnetic fields that grants an approximate conserved momentum to the gyrocentre dynamics of charged particles is called quasisymmetric. In particular, we focus our attention on weakly quasisymmetric magnetic fields as recently defined (Rodríguez *et al.* 2020; Constantin, Drivas & Ginsberg 2021). Formally, a magnetic field \mathbf{B} is weakly quasisymmetric if and only if there exists a symmetry vector field \mathbf{u} such that (Rodríguez *et al.* 2020)

$$\mathbf{u} \cdot \nabla \mathbf{B} = 0, \quad (2.1)$$

$$\mathbf{B} \times \mathbf{u} = \nabla \Phi, \quad (2.2)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (2.3)$$

where Φ labels magnetic flux surfaces that we take to be nested (Rodríguez & Bhattacharjee 2021a). We are assuming the electrostatic potential in the problem to be symmetric, in the sense that $\mathbf{u} \cdot \nabla \Lambda = 0$ for Λ the electrostatic potential. This symmetry requirement preserves the approximate conservation of the momentum characterising a QS configuration.

Although this definition has the benefit of directly emerging from single-particle considerations, it is difficult to use it practically to assess whether a field is quasisymmetric or not. Doing so would require solving for the field \mathbf{u} . Instead, we would like to have a condition that only involves magnetic field quantities directly. This can be achieved in a straightforward way by constructing \mathbf{u} such that it satisfies (2.1) and (2.2): $\mathbf{u} = \nabla \Phi \times \nabla \mathbf{B} / (\mathbf{B} \cdot \nabla \mathbf{B})$. At the singular points where $\mathbf{B} \cdot \nabla \mathbf{B} = 0$ this construction breaks down unless $\nabla \Phi \times \nabla \mathbf{B} \cdot \mathbf{B} = 0$. (Note that this follows from the conditions (2.1) and (2.2).) The vanishing of $\nabla \Phi \times \nabla \mathbf{B} \cdot \mathbf{B}$ when $\mathbf{B} \cdot \nabla \mathbf{B} = 0$ is referred to as pseudosymmetry (Mikhailov *et al.* 2002; Skovoroda 2005), and it guarantees that all of the contours of constant $|\mathbf{B}|$ on a flux surface are linked to the torus in the same way, avoiding local extrema on the surfaces (and thus providing well-behaved streamlines of \mathbf{u}). Breaking pseudosymmetry leads to a diverging orbit width (Rodríguez *et al.* 2020) (i.e. loss of

deeply trapped and barely trapped particles) at these $|\mathbf{B}|$ extrema. In a QS solution, the construction of \mathbf{u} is then well defined because QS implies pseudosymmetry.

With the given construction of \mathbf{u} , we may substitute it into (2.3) to obtain

$$\frac{\nabla\Phi \times \nabla B \cdot \nabla(\mathbf{B} \cdot \nabla B)}{(\mathbf{B} \cdot \nabla B)^2} = 0. \quad (2.4)$$

This yields the triple vector formulation of QS. Namely, if we define the scalar quantity

$$f_T = \nabla\psi \times \nabla B \cdot \nabla(\mathbf{B} \cdot \nabla B), \quad (2.5)$$

the magnetic field is quasisymmetric if and only if $f_T = 0$. (By continuity, this includes pseudosymmetry.)

For the purpose of this paper, all magnetic configurations have well-behaved nested flux surfaces and are non-isodynamic. By isodynamic we mean a field whose magnitude $|\mathbf{B}|$ is a flux function (Bernardin, Moses & Tataronis 1986). With these provisions, f_T can be evaluated for non-quasisymmetric configurations and serves as a measure of QS. We have here used ψ , the toroidal magnetic flux, as a label of flux surfaces rather than the arbitrary flux function $\Phi(\psi)$. This is one of many possible choices, and one could, for instance, replace $\nabla\psi$ with $\hat{\mathbf{n}}$ (the normal unit vector to the flux surface) in evaluating (2.5) on a magnetic surface within a region of continuously nested surfaces.

The measure f_T is a coordinate-independent, scalar measure of the departure from QS. It is a local quantity, although it also has some global properties. For instance, $\langle f_T \rangle = 0$, where $\langle \dots \rangle$ is the flux-surface average.

We have seen how (2.5) results from the fundamental definition of QS in (2.1)–(2.3). This grants f_T the special feature that its application is not tied to a particular form of equilibrium. This could be relevant for constructing cost functions in so-called one-shot optimisation schemes (Akçelik *et al.* 2006; Dekeyser, Reiter & Baelmans 2014), i.e. a formulation in which finding equilibria is part of the optimisation problem. Although QS may be explored in the presence of forces more general than isotropic plasma pressure, p , it is standard to adopt the magnetohydrostatic assumption $\mathbf{j} \times \mathbf{B} = \nabla p$, where $\mathbf{j} = \nabla \times \mathbf{B}$ is the current density. Following standard practice, we shall do so for the remainder of this work.

A key consequence of this form of equilibrium is that $\mathbf{j} \cdot \nabla\psi = 0$. Thus, there exist Boozer coordinates (Boozer 1981; Helander 2014) $\{\psi, \theta, \phi\}$. In these straight-field line coordinates the covariant form of the magnetic field is $\mathbf{B} = I(\psi)\nabla\theta + G(\psi)\nabla\phi + B_\psi\nabla\psi$ and the coordinate Jacobian $\mathcal{J} = (G + \iota I)/B^2$, where ι is the rotational transform. Equation (2.5) may then be written explicitly in the form

$$f_T = [\partial_\theta B \partial_\phi - \partial_\phi B \partial_\theta] (\partial_\phi + \iota \partial_\theta) B. \quad (2.6)$$

The QS condition can be rewritten in the form $(\partial_\theta B)(\partial_\phi + \iota \partial_\theta)(\partial_\phi B / \partial_\theta B) = 0$, from which it follows that $\partial_\phi B / \partial_\theta B$ must be a flux function. This implies $|\mathbf{B}|$ has an explicit symmetry and the conventional definition of QS then follows (Boozer 1983; Rodríguez, Sengupta & Bhattacharjee 2021). We refer to it as the Boozer formulation of QS: a magnetic field is quasisymmetric if and only if the magnetic field magnitude can be written as a function $B = B(\psi, M\theta - N\phi)$, where $N, M \in \mathbb{Z}$. It is convenient to define the symmetry helicity $\tilde{\alpha} = N/M$ and a helical angle $\chi = \theta - \tilde{\alpha}\phi$. (We will not consider quasipoloidal symmetry, corresponding to $M = 0$, although the extension should be straightforward.) There is no unique way to construct a scalar measure that quantifies how close a given magnetic field

is to QS in the presented formulation. It is, however, customary to define the minimal measure

$$f_B = \sum_{\substack{n,m \\ n \neq \tilde{\alpha}m}} |B_{nm}|^2. \quad (2.7)$$

Here B_{nm} stands for the Fourier components of the magnetic field magnitude B in Boozer coordinates, and the sum is over the non-symmetric components. All the mode components are weighted equally in this form (although this could be modified). In practice, only the lower-order modes significantly contribute to f_B , a result that follows from the smoothness of $|B|$: an s times differentiable function in L^1 on a torus will have Fourier coefficients that generally scale like $\hat{f}(m) \sim 1/|m|^s$ (Grafakos 2008).

Unlike f_T , the helicity of the symmetry $\tilde{\alpha}$ is required for the evaluation of f_B . Such knowledge is needed to avoid the symmetric modes of $|B|$ in the summation of (2.7). In a sense, therefore, f_B measures deviations from a particular form of QS. The possibility of enforcing a particular helicity can be of practical importance when optimising for QS configurations.

This measure is also different from f_T in two other important respects. First, f_B is a single flux-surface scalar instead of a local measure. Second, as the Fourier mode resolution of $|B|$ needed for (2.7) is in Boozer coordinates, an evaluation of f_B requires explicit knowledge of these coordinates. This prevents application of f_B to scenarios requiring $\mathbf{j} \cdot \nabla \psi \neq 0$, unlike f_T . The need to compute Boozer coordinates and Fourier-resolve B also imposes an additional numerical burden (e.g. using existing codes such as BOOZXFORM) (Sanchez *et al.* 2000; Paul *et al.* 2020). While computing the Boozer coordinate transformation is typically not a major numerical bottleneck, avoiding the transformation may be convenient for some applications.

While f_T does not require any particular coordinate system (in this case Boozer coordinates), it also does not allow specification of the desired helicity, $\tilde{\alpha}$. It is thus of interest to explore alternative measures of QS that incorporate both aspects. We now obtain what we call the two-term formulation of QS (Helander 2014; Paul *et al.* 2020).

We begin from (2.1)–(2.3) again. Defining the inner product $C = \mathbf{B} \cdot \mathbf{u} / \Phi'$ and using (2.1) and (2.2), we can write $C\mathbf{B} \cdot \nabla B = \mathbf{B} \cdot \nabla \psi \times \nabla B$. This expression holds even when $\mathbf{B} \cdot \nabla B = 0$. To involve C directly in the remaining (2.3), we rewrite it in the form (Rodríguez *et al.* 2021)

$$\mathbf{B} \cdot \nabla (\mathbf{B} \cdot \mathbf{u}) = \mathbf{j} \cdot \nabla \Phi. \quad (2.8)$$

Under the assumption of $\mathbf{j} \cdot \nabla \psi = 0$, it follows that $C = C(\psi)$ must be a flux function. We can then write

$$f_C = \mathbf{B} \cdot \nabla \psi \times \nabla B - C(\psi)\mathbf{B} \cdot \nabla B, \quad (2.9)$$

which must vanish for some flux function C for a QS field. This is the two-term formulation of QS. To avoid having to perform a search for C every time (2.9) is evaluated, we need some form for the flux function C in terms of more recognisable quantities. To do so, we consider the limit of a QS configuration and adopt Boozer coordinates. In that case, and for $\partial_x B \neq 0$ (otherwise $f_C = 0$ for any value of C),

$$C = \frac{\mathbf{B} \cdot \nabla \psi \times \nabla B}{\mathbf{B} \cdot \nabla B} = \frac{G + \tilde{\alpha}I}{\iota - \tilde{\alpha}}. \quad (2.10)$$

(This flux function C is often given the symbol F in the literature; Helander 2014; Paul *et al.* 2020). Now C is expressed in terms of physical quantities: the Boozer currents

	Helicity	Boozer	$\mathbf{j} \cdot \nabla \psi = 0$	Info
f_B	I	Y	Y	G
f_C	I	N	Y	L
f_T	O	N	N	L

TABLE 1. Basic properties of QS measures. Summary of the basic formal features of the QS measures introduced in § 2. These include: whether the helicity needs to be specified as an input (I) or not (O), whether Boozer coordinates are required to evaluate it, whether $\mathbf{j} \cdot \nabla \psi = 0$ is a necessary assumption and whether the measure is local (L) or global (G).

I and G and the rotational transform ι . The currents may be calculated without resorting to Boozer coordinates directly (see equation (16) in Helander (2014)). Any flux coordinate system will do, under the assumption of well-defined, nested flux surfaces.

It is clear that the helicity of the symmetry $\tilde{\alpha}$ needs to be specified, just as one must do to evaluate f_B . The function C for which it is needed has important physical meaning (Rodríguez *et al.* 2020), as it is a direct measure of the banana width of trapped particles, Λ , in the QS limit. In more detail, $\Lambda |\nabla \psi| \sim \rho_{\parallel} C$, where ρ_{\parallel} is the Larmor radius associated with the parallel velocity. The resonance at the surface $\iota = \tilde{\alpha}$ leads to $C \rightarrow \infty$, unless $\mathbf{B} \cdot \nabla \psi \times \nabla B = 0$ on the surface. This latter condition forces the surface to be isodynamic, which also prevents the appearance of a current singularity and the potential opening of a magnetic island (Rodríguez & Bhattacharjee 2021a). In the present work, we avoid such surfaces altogether.

In summary, we have a measure f_C that is local like f_T , for which the helicity of the symmetry must be prescribed as in f_B , but for which no Boozer coordinates are needed. This makes this formulation amenable to gradient-based optimisation techniques such as adjoint methods (Paul *et al.* 2020).

This constitutes the derivation of the three measures of QS— f_B , f_T and f_C —that will occupy us in this paper. In table 1 we summarise the main formal and construction differences of the three forms.

3. Comparison of cost functions

All three forms introduced in the previous section can be interpreted as measures of QS: they all share the basic feature that they vanish only for a configuration that is exactly quasisymmetric. However, each form treats deviations from QS differently. The focus of this section is to evaluate how these measures compare with each other formally, and to connect them to some of the physical properties usually associated with QS.

3.1. Formal relation

It is convenient to measure deviations from QS in terms of the asymmetric components of the magnetic field magnitude. Thus, to understand the forms of f_T and f_C , it is informative to learn how the different Fourier modes of $|\mathbf{B}|$ are weighted in each of these metrics. Then a comparison with f_B may be done on an equal footing.

Let us start with the definition of f_C and write the magnitude of the magnetic field as a Fourier series:

$$B = \sum_{n,m} B_{nm} e^{i(m\theta - n\phi)}, \quad (3.1)$$

where the angular Boozer coordinates θ and ϕ are used for the Fourier resolution of the field. Using the construction of C in (2.10) we may write, after some algebra,

$$\begin{aligned}
 f_C &= \frac{B^2}{\tilde{\alpha} - \iota} (\partial_\phi + \tilde{\alpha} \partial_\theta) B \\
 &= i \frac{B^2}{\iota - \tilde{\alpha}} \sum_{n/m \neq \tilde{\alpha}} (n - m\tilde{\alpha}) B_{nm} e^{i(m\theta - n\phi)}.
 \end{aligned}
 \tag{3.2}$$

The first line shows that f_C/B^2 , although a local measure, has some global properties like f_T does. In particular, the integral along closed streamlines of $\mathbf{Q} = \mathcal{J} \nabla \psi \times \nabla \chi$ must vanish, that is, $\oint (f_C/B^2) dl/|\mathbf{Q}| = 0$. The function f_C/B^2 is also clearly linear in the Fourier content of B , but interestingly, not all modes are weighted equally. The modes ‘furthest’ from the symmetry direction are weighted more heavily than those close to it.

Although we have successfully written f_C in terms of the Fourier content of B , its form is still far from that of f_B . To make a closer comparison with f_B we need to construct a single flux-surface scalar from f_C . We take (3.2) and write it as

$$B_{nm} = -i \frac{\iota - \tilde{\alpha}}{n - m\tilde{\alpha}} \left(\frac{f_C}{B^2} \right)_{nm},
 \tag{3.3}$$

where $(\dots)_{nm}$ represents the (n, m) Fourier component of the function in parentheses in Boozer coordinates. Using Parseval’s theorem, we write

$$\sum_{n,m} (n - m\tilde{\alpha})^2 |B_{nm}|^2 = \frac{(\iota - \tilde{\alpha})^2}{(2\pi)^2} \int d\theta d\phi \left(\frac{f_C}{B^2} \right)^2,
 \tag{3.4}$$

which can be rewritten as

$$\overline{\left(\frac{f_C}{B^2} \right)^2} = \sum_{n,m} \left(\frac{n - m\tilde{\alpha}}{\iota - \tilde{\alpha}} \right)^2 |B_{nm}|^2,
 \tag{3.5}$$

where $\overline{(\dots)}$ denotes the $(0, 0)$ Fourier component of the expression in parentheses. Equation (3.5) is now in a form close to f_B . The scalarised version of f_C , in this case normalised to B^2 and averaged over the surface, has a form similar to f_B in that it involves a sum over the non-symmetric Fourier modes of B . The main difference comes through the different weighting of the modes. The measure f_C can be thought of as a modification of f_B in which asymmetric components of the magnetic field magnitude furthest from QS are more heavily penalised.

The difference in weights implies that f_B and f_C are actually not monotonic with respect to one another. Changing the energy content of the Fourier modes may lead to f_B increasing but f_C decreasing (and vice versa). Having said that, there also is a class of changes for which both metrics respond in the same way. Given that all terms in the summation of (2.7) and (3.5) are positive, reducing any individual mode $|B_{nm}|$ will, for instance, lead to the reduction of both metrics. This observation has important implications for the universality of QS measures and optimisation, to be explored in the next sections.

Let us now see how f_T fares in comparison. When expressed in terms of the Fourier content of the magnetic field strength, the nonlinear character of the triple product

formulation is clear. We explicitly write

$$(\mathcal{J}^2 f_T)_{nm} = i \sum_{k,l} (nl - k)(km - nl) B_{n-k,m-l} B_{k,l}, \quad (3.6)$$

where \mathcal{J} is the Jacobian of Boozer coordinates and k, l are integers. The nonlinearity appears in the form of a convolution that mixes modes in a non-trivial way. We also have one higher derivative of the field strength compared to f_B or f_C . Scalarising this expression by computing an average like in the case of f_C does not help in bringing this form any closer to f_B . The convolution couples modes together, making it unclear whether reducing a single asymmetric mode will reduce f_T as it did for f_C and f_B .

We may try to gain some additional understanding of how f_T is affected by changes in the mode content by considering a system that lies close to QS. Consider a small deviation \tilde{B} in the field magnitude from a dominantly QS field. Then a linearisation of f_T takes the form

$$(\mathcal{J}^2 f_T)_{nm} = i(\tilde{\alpha}m - n)(\nu m - n) \sum_l l B_{\tilde{\alpha}l,l} \tilde{B}_{n-\tilde{\alpha}l,m-l}. \quad (3.7)$$

It is clear from this expression that decreasing \tilde{B} decreases the magnitude of f_T . However, even upon linearisation, it is not obvious that reducing a particular Fourier harmonic will guarantee a lower value of f_T .

The measure f_T is, in this form, distinct from the previous metrics. However, there is actually a natural relation between the triple vector form and f_C through a simple magnetic differential equation (see [appendix A](#)). This is reasonable given the common origin of the metrics.

3.2. Physical relations

We have explored three forms for QS in this work. These constructions describe deviations from QS differently. To understand the potential physical meaning of these differences, we now explore physical phenomena often associated with QS configurations and their relation to the cost functions.

3.2.1. Single-particle dynamics

The concept of QS as presented in this work is built on single-particle dynamics. In §2, beginning with the most fundamental definition, we constructed three measures of QS. Thus, it is natural to expect that all these metrics should have some relation to single-particle dynamics.

Let us start from the definition of the momentum: $\bar{p} = -\Phi + v_{\parallel} \mathbf{b} \cdot \mathbf{u}$, where v_{\parallel} is the parallel velocity of the particle under consideration (Rodríguez *et al.* 2020). For a QS configuration where \mathbf{u} satisfies (2.1)–(2.3), this momentum is approximately conserved to $O(\rho/L)$, where ρ is the particle Larmor radius and L a characteristic length of the system. When the symmetry is broken, we may ask how this momentum evolves in time. To describe that evolution we need to define the vector field \mathbf{u} . We can always choose this field in a way that (2.1) and (2.2) are satisfied by construction. To do so we assume the existence of flux surfaces as well as pseudosymmetry (Mikhailov *et al.* 2002; Skovoroda 2005). The latter guarantees that the contours of constant $|\mathbf{B}|$ are linked to the torus avoiding local extrema on the surfaces. These well-behaved contours are vital to the notion of a continuous symmetry, and necessary to prevent $\mathbf{b} \cdot \mathbf{u} \rightarrow \infty$. Once \mathbf{u} has been constructed, one evaluates the dynamics of \bar{p} using the leading-order form of the equations of motion from the Littlejohn Lagrangian (Littlejohn 1983; Rodríguez *et al.* 2020). We may thus

write, to $O(\rho/L)$,

$$\frac{d}{dt}\bar{p} = \frac{v_{\parallel}^2}{B^2} [\mathbf{B} \cdot \nabla(\mathbf{B} \cdot \mathbf{u}) - \mathbf{j} \cdot \nabla\psi] = - \left(\frac{v_{\parallel}}{\mathbf{B} \cdot \nabla B} \right)^2 f_T. \tag{3.8}$$

This shows that f_T constitutes a measure of the conservation of \bar{p} . Of course, even if $f_T \neq 0$ the above could result in conservation of \bar{p} in the bounce-averaged sense. This averaged conservation could be seen as the difference between QS and omnigenity (Landreman & Catto 2012; Helander 2014). It is also important to note from (3.8) that the cost function f_T appears with a factor of $1/(\mathbf{B} \cdot \nabla B)^2$. This factor penalises deviations from QS close to the extrema of $|\mathbf{B}|$ along field lines. The special importance of the regions close to the extrema is reminiscent of the requirements for the confinement of barely and deeply trapped particles in omnigenity. The presence of this resonance makes the combination $f_T/(\mathbf{B} \cdot \nabla B)^2$ numerically ill-behaved as a modified cost function. A practical implementation could perhaps be achieved by regularising the combination, as has been employed in optimisation for pseudosymmetry (Mikhailov *et al.* 2002). Although attractive from the single-particle perspective, and even if the resonance is amended, the measure would still exhibit sharp gradients due to the heavy weight near extrema of $|\mathbf{B}|$ along field lines.

The argument related to the dynamics of \bar{p} might appear to some extent artificial, as \bar{p} really only gains special dynamical meaning in the limit of QS. We may alternatively look at a more physically relevant property: the net drift of magnetised particles off flux surfaces. It is convenient to express the off-surface bounce-averaged drift of particles, $\Delta\psi$, as $\partial_{\alpha}\mathcal{J}_{\parallel}$, where $\mathcal{J}_{\parallel} = \oint v_{\parallel} dl$ is the second adiabatic invariant (calculated taking the integral along the field line and between bouncing points) and α is the usual magnetic field-line label (Helander 2014). To evaluate this derivative of \mathcal{J}_{\parallel} , consider two nearby lying magnetic field lines on the same magnetic surface at α and $\alpha + \delta\alpha$ (see figure 1). Take the turning points that define the integral for \mathcal{J}_{\parallel} to be at a field magnitude $|\mathbf{B}| = B_t$. We assume that these bounce points exist on both of these field lines (as we are looking at the \mathcal{J}_{\parallel} associated with a class of particles defined by its bounce points). Then taking the line integral along the contour in figure 1, we compute

$$\left(\int_{C_{\alpha}} + \int_{C_{\delta\alpha}} + \int_{C_{B+}} + \int_{C_{B-}} \right) v_{\parallel} \mathbf{b} \cdot d\mathbf{l} = \delta\alpha \int \nabla\psi \cdot \nabla \times (v_{\parallel} \mathbf{b}) \frac{dl}{B}. \tag{3.9}$$

To obtain the equality, we apply the Stokes theorem and take the limit of small $\delta\alpha$, taking the latter integral with respect to length along a field line, l . Assuming that $\mathbf{j} \cdot \nabla\psi = 0$ and noting that the integrals along C_{B+} and C_{B-} vanish as $v_{\parallel} = 0$ wherever $|\mathbf{B}| = B_t$, we may rewrite

$$\partial_{\alpha}\mathcal{J}_{\parallel} = - \int \frac{\nabla\psi \times \nabla B \cdot \mathbf{B}}{\mathbf{B} \cdot \nabla B} \frac{1}{B^2} \left(\frac{E}{v_{\parallel}} + \frac{v_{\parallel}}{2} \right) d\mathbf{B}, \tag{3.10}$$

where B is being used to parameterise the integral along the field line. To be precise, we should split the integral along the field line into concatenated pieces joint at $|\mathbf{B}|$ extrema. The vanishing of $\mathbf{B} \cdot \nabla B$ at the stitching points (and the consequent divergence of the integrand) is generally not problematic as one can consider switching back to integrating with respect to l arbitrarily close to them, with that small piece of the integral constituting a vanishingly small contribution. We drop this distinction for simplicity.

One may write the kinematic expression $(E/v_{\parallel} + v_{\parallel}/2)/B^2 = \mathbf{B} \cdot \nabla(v_{\parallel}/B)/(\mathbf{B} \cdot \nabla B)$. As a result, any flux function multiplying this kinematic expression will lead to a vanishing integral. Thus, the first fraction in the integrand can be written as $f_C/(\mathbf{B} \cdot \nabla B)$. This

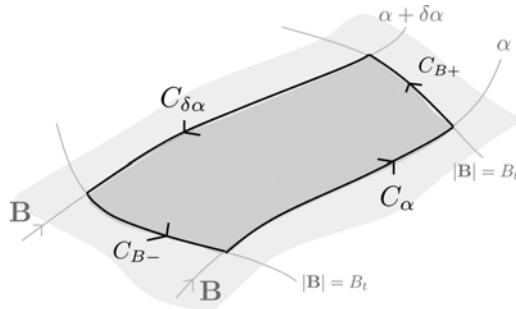


FIGURE 1. Schematic of the loop integral for $\partial_\alpha \mathcal{J}_\parallel$. Schematic of the path integral followed to compute the field line dependence of the second adiabatic invariant \mathcal{J}_\parallel . The contours match field lines and contours of constant $|\mathbf{B}| = B_t$.

shows that f_C has a direct relation to the field-line dependence of the second adiabatic invariant. The complicated nature of the kinematic factor, however, obscures this relation. Proceeding with integration by parts and using (A2):

$$\partial_\alpha \mathcal{J}_\parallel = \frac{\nabla \psi \times \nabla \mathbf{B} \cdot \mathbf{B}}{\mathbf{B} \cdot \nabla \mathbf{B}} \frac{v_\parallel}{B} \Big|_{v_\parallel=0} + \int \frac{f_T}{(\mathbf{B} \cdot \nabla \mathbf{B})^2} v_\parallel dl. \quad (3.11)$$

Thus, we have constructed a simple relation between the bounce-averaged, off-surface drift $\Delta \psi$ and the QS measure f_T . The measure comes once again normalised by a factor $(\mathbf{B} \cdot \nabla \mathbf{B})^2$. However, this f_T term is not the only term in (3.11). For the expression in (3.11) to be finite (or vanish as we want for QS), the boundary term in (3.11) should not diverge, which requires $\nabla \psi \times \nabla \mathbf{B} \cdot \mathbf{B} = 0$ wherever $\mathbf{B} \cdot \nabla \mathbf{B} = 0$. This condition is precisely the condition of pseudosymmetry introduced earlier (Mikhailov *et al.* 2002; Skovoroda 2005). Note that this pseudosymmetry condition appears separate from the f_T integral in (3.11). Thus, generally, even when $f_T \neq 0$ as in omnigenic configurations, the condition should be satisfied if $\partial_\alpha \mathcal{J}_\parallel$ vanishes (and deeply trapped and barely trapped particles are confined). As discussed in § 2, requiring $f_T = 0$ everywhere includes this condition. Guaranteeing that the boundary term vanishes also ensures the existence of bouncing points B_t on nearby field lines, in agreement with the assumptions made in evaluating $\partial_\alpha \mathcal{J}_\parallel$.

The discussion above shows that both f_T and f_C can be regarded as natural measures of QS in the single-particle perspective. However, when the configuration fails to be pseudosymmetric, the interpretation of f_T is obscured.

The measure f_C offers some additional particle dynamics insight. As we mentioned when constructing f_C in § 2, the function C in f_C is a measure of the banana width of bouncing particles in the QS limit. The expression in (3.2) inherits the scaling $\propto 1/\bar{\iota}$, where $\bar{\iota} = \iota - \tilde{\alpha}$ from C . Hence, f_C also has scaling related to the orbit width and the bootstrap current, which is proportional to the banana width (Helander 2014).

Although the metrics bear some important physical content in their relation to the behaviour of single-particle dynamics, the mapping is not one-to-one. More sophisticated measures are required for an accurate description of the dynamics of single particles. Especially important are alpha particles, for which complicated metrics (see Bader *et al.* 2019) or expensive simulations are needed.

3.2.2. Neoclassical transport

Let us now consider the important physics of neoclassical transport (also related to enhanced neoclassical viscosity). One of the distinguishing features of QS is to grant a low

level of neoclassical transport to the magnetic configuration (Boozer 1983). In particular, it prevents the detrimental neoclassical $1/\nu$ transport (we do not consider turbulent transport here, which, however, has been seen to be a dominant form of transport in many optimised stellarators once neoclassical transport has been sufficiently reduced) that tends to spoil confinement in stellarators. Of course, as we deviate from QS, this property will degrade. We would like to understand how this degradation is captured (if at all) by the different metrics of QS presented in the paper.

To study this, we introduce first the classical stellarator concept of helical ripple. We define helical ripple as $\epsilon_h \sim B_{nm}/B$, a measure of the symmetry-breaking magnetic field magnitude. This quantity is clearly linked to f_B in (2.7). When a single, small asymmetric mode on top of a QS field is considered, in the $1/\nu$ and $\sqrt{\nu}$ regimes, this ripple leads to an enhancement of neoclassical losses that scales roughly as $D \sim \epsilon_h^{3/2}$ (Ho & Kulsrud 1987; Nemov *et al.* 1999; Helander & Simakov 2008). Therefore, there is a direct relationship between ϵ_h and transport. In this single-mode classical stellarator scenario f_B is the ripple (unlike in the more general case). It is then clear that f_B does incorporate some direct information about transport.

This classical view on transport has the drawback of treating all ‘ripples’ on the same footing. However, a more detailed approach shows that different modes contribute distinctly (Calvo *et al.* 2014). Transport in the $1/\nu$ regime is driven by bouncing particles near minima of $|\mathbf{B}|$ along field lines. The enhancement of transport can then come from the introduction of significant additional wells in which new bouncing particles will live. However, this is not the only mechanism that can enhance transport. As particles precess following constant \mathcal{J}_\parallel surfaces, a significant difference in the $|\mathbf{B}|$ profile between field lines on the same flux surface will lead to an enhancement in the drifting of bouncing particles off the surface. (Additional effects such as collisional detrapping may also occur.) These phenomena are quantified in Calvo *et al.* (2014). Small deviations from QS with helical ripple ϵ_h that do not affect the $|\mathbf{B}|$ profile significantly lead to a transport enhancement $\langle \mathbf{\Gamma} \cdot \nabla \psi \rangle \sim \epsilon_h^2$. (The same scaling is anticipated for flow damping and generally deviations from intrinsic ambipolarity.) Now, if the deviation is due to a mode with a large gradient along \mathbf{B} , then the changes lead to an enhancement by a factor $O(\epsilon_h)$. When the deviations lead to significant differences between field lines, then the transport changes significantly, not amenable to a perturbative treatment. (Aspects of this scenario can be understood from an analysis of the second adiabatic invariant, studied in the previous section.)

In this more detailed picture it is clear that different modes of $|\mathbf{B}|$ contribute differently to transport. The enhancement is significant when the gradients along the field line are large in the sense $|\mathbf{b}_0 \cdot \nabla B_1| \sim |\mathbf{b}_0 \cdot \nabla B_0|$, with B_1 the asymmetric field strength. In a more illuminating form, $|\mathbf{b}_0 \cdot \nabla B_1|/|\mathbf{b}_0 \cdot \nabla B_0| \sim [(n - im)/(M(\tilde{\alpha} - i))](\epsilon_h/\epsilon_{QS})$, where we have taken ϵ_{QS} to represent the helical ripple associated with the symmetric part of B with symmetric poloidal mode number M . The weight in the numerator indicates the predominance of modes with large gradients along the field line, while the denominator serves as a comparative measure for the the gradient along the field line due to the symmetric field.

This difference between modes is clearly missed by the construction of f_B . However, we have shown that the angle-averaged $(f_C/B^2)^2$ provides a weighted version of f_B in which higher-order modes are increasingly penalised $\propto (n - m\tilde{\alpha})^2$. Although this unequal weight provides a distinction between modes, the weight is not of the form $(n - im)$ as in the transport criterion. It instead weights higher-order modes according to their similarity with respect to the symmetry direction. Thus, f_C only

qualitatively matches the high-order mode weighting of the transport scaling. The weight that appears in the definition of f_C can be interpreted as a consequence of the smoothness of the function. The quasisymmetric modes are excluded from the sum through a weight that can be written as a differential operator: the weight must be zero at $n/m = \tilde{\alpha}$ and smoothly increasing away from it. It thus appears natural that the weight in f_C is centred about the symmetry direction. This requirement is avoided in f_B , where the symmetric discrete modes are avoided. To more accurately treat the impact on transport, one could imagine a modified version of f_B in which the mode weights are given by $(n - im)$ and the symmetric components are excluded from the summation. It would, however, still require Boozer coordinates (unlike f_C). Interestingly, the weight $(m - n)$ appears in the linearised version of f_T in § 3.1.

So far we have focused on small, isolated deviations from QS. In general, however, transport calculations are significantly more complex. The presence of multiple ripples will lead to intricate changes in the trapping and passing particle behaviour. There exist more sophisticated metrics that attempt to describe these changes in more detail. A commonly used example is the $1/\nu$ transport regime effective ripple, $\epsilon_{\text{eff}}^{3/2}$ (Nemov *et al.* 1999; Nemov, Kasilov & Kernbichler 2014). The complexity of this measure is not reflected in either f_B or f_C (see the following sections). Thus, although f_B and f_C do contain relevant information regarding transport, they cannot be taken as accurate measures of it.

3.2.3. Current singularities and islands

Although flux surfaces are assumed throughout our analysis, we may also relate the QS metrics to the potential presence of current singularities and magnetic islands (Tessarotto, Johnson & Zheng 1995; Rodríguez & Bhattacharjee 2021a). Fourier modes of $|B|$ that resonate with the finite rotational transform can lead to Pfirsch–Schlüter (PS) current singularities. These finite- β currents are associated with the potential opening of magnetic islands at the location of the singularities (Reiman & Boozer 1984; Rodríguez & Bhattacharjee 2021a). The PS current singularity arises due to the periodicity constraint on the magnetic differential equation,

$$\mathbf{B} \cdot \nabla \left(\frac{j_{\parallel}}{B} + \frac{p' C}{B^2} \right) = -2p' \frac{f_C}{B^3}, \quad (3.12)$$

for the parallel current j_{\parallel} . The periodicity condition required for solvability is

$$p' \oint \frac{f_C}{B^4} dl = 0, \quad (3.13)$$

where l measures length along a field line. We note that f_C appears explicitly. This condition will be satisfied in QS, thus precluding the existence of PS singularities (except possibly at the conflictive $\iota = \tilde{\alpha}$ surface; see § 2). Note that solvability may still be obtained away from QS if $p' = 0$ or the line integral vanishes.

One may generally show that in magnetohydrostatics, the resonant radial magnetic field due to the singularities is $B_{\rho nm} \propto (1/B^2)_{nm}$ (in simplified geometry), and thus $W_i \propto \sqrt{(1/B^2)_{nm}}$ (Reiman & Boozer 1984; Bhattacharjee *et al.* 1995; Rodríguez & Bhattacharjee 2021a). A non-QS field will potentially have contributions at many rational surfaces, as driven by the asymmetric modes. In that sense, any sum over the asymmetric modes of B , such as f_B or f_C , will include some of this information. Resonances only occur when the rotational transform matching the asymmetric mode helicity exists within the plasma, and thus a more faithful metric would only account for the resonant asymmetric modes. Due to the generally asymmetric geometry of configurations, resonant

fields arising from PS currents may appear at other rational surfaces through toroidal coupling (Rodríguez & Bhattacharjee 2021a). This is of course only a partial account of the relevant island physics, focusing on the potentially deleterious seeding of islands. The appearance of islands will in practice depend on a plethora of additional effects beyond the simple magnetohydrostatic equilibria considerations, including, for instance, the healing effects of flows or additional kinetic effects (Fitzpatrick 1993; Hegna 2011, 2012).

4. Quasisymmetry in practice: universality and cost functions

The discussion so far has explored the origin of the three different measures of QS— f_B , f_C and f_T —and discussed their formal properties and related physical content. In this section we explore some more practical aspects of these measures. First, we address the question ‘Which configuration is more quasisymmetric?’ and discuss the universality of QS. At a second stage, we explore the implications of these different formulations for the optimisation of stellarators through numerical examples.

4.1. Universality of QS

The goal of the metrics f_B , f_C and f_T is to quantify the deviation from QS. In this sense, all these forms would appear to be equally valid candidates. Given the variety of choice, we would like to understand to what extent it is meaningful to state that a certain field configuration is more quasisymmetric than another as measured by these metrics.

As demonstrated in the comparison of the previous section, QS metrics are not generally monotonic with respect to each other. For example, decreasing f_B does not imply decreasing f_C . One could, for instance, exchange the energies of a larger low-order mode to a high-order one without changing f_B but modifying f_C . The nonlinearity of f_T makes the lack of monotonicity somewhat more marked. This lack of monotonicity makes it difficult to determine magnetic configurations which are ‘closer’ to QS, or for that matter, if the question is meaningful. As different measures of QS yield different answers, any comparison must be defined with respect to a particular standard. The notion thus lacks universality: the quantification of closeness to QS is in some sense arbitrary, and does not necessarily have a clear physical meaning attached to it.

Let us consider some practical examples. In order to use the metrics of QS quantitatively, we first need to normalise and scalarise them. Only then will we be able to make fair comparisons across different devices. Perhaps unsurprisingly, we will see that this procedure is not unique, emphasising the lack of universality.

Let us start with the metric f_B , which measures the Fourier content of B in Boozer coordinates. It is customary to normalise it to the total magnetic field energy. So we define

$$\hat{f}_B = \sqrt{\frac{1}{B^2} f_B}. \quad (4.1)$$

Doing so ensures $\hat{f}_B \in [0, 1)$. This finite range of values makes f_B convenient, although the requirement $B > 0$ excludes the value of 1. Although this normalisation seems the most natural, it is not unique. For instance, we could decide to normalise f_B to (only) the symmetric part of the field excluding the flux-averaged part, or we could normalise by the flux-surface average of B^2 , $\langle B^2 \rangle$. For the purpose of this discussion, we take the natural and more conventional form (4.1).

The choice of a ‘natural’ normalisation is unclear for f_C and f_T . Fundamentally, our normalisation should give a dimensionless measure. This principle is, of course, not

specific enough. To constrain the construction of the measure further, we try making their Fourier forms look closest to f_B . This is a convenient choice. Let us start with f_C , which dimensionally scales as $f_C \sim B^3$. In addition to this dimensional factor, motivated by the form of (3.2), we also choose to normalise by $1/(\tilde{\alpha} - \iota)$ to bring the measure closer to the form of f_B . One could argue for adding additional dimensionless factors such as the inverse aspect ratio, ϵ , which characterises the magnitude of $CB \cdot \nabla B$ for $\chi = \theta$, but for simplicity we do not do so here. In view of the normalisation of f_B , we normalise and scalarise f_C as follows:

$$\hat{f}_C = \sqrt{\frac{\langle f_C^2 \rangle}{\langle B^2 \rangle^3} |\tilde{\alpha} - \iota|^2}. \quad (4.2)$$

This leads to a form that is close to but not identical to (3.2).

The normalisation of f_T is inherently more complex. The triple vector product introduces explicit length scales which must be non-dimensionalised. Once again, we look at the form of f_T in (3.5). It is then natural to normalise and scalarise f_T using

$$\hat{f}_T = \sqrt{\langle f_T^2 \rangle \frac{R^4}{\langle B^2 \rangle^4}}. \quad (4.3)$$

In obtaining this normalisation we have used the scaling $\mathcal{J} \sim R/B$, where R is the major radius. Once again, one could argue that the choice is, to some extent, arbitrary. It is interesting to note that the combination $f_T/(B \cdot \nabla B)^2$, recurrent in the description of single-particle dynamics in the previous section, is by construction dimensionless. Thus, it appears to be a natural choice. However, we do not consider it here due to the added complication of the resonant denominators and the associated steep gradients.

Now that we have chosen specific normalisations for the QS metrics, we may compare their values in stellarator equilibria. We consider two scenarios. First, we focus on the comparison of the metrics in configurations that have been designed to be approximately quasisymmetric. Then we study how these measures differ when the system has not necessarily been optimised to be quasisymmetric.

We present in figure 2 a comparison of the metrics in six quasisymmetric configurations:¹ GAR (Garabedian 2008; Garabedian & McFadden 2009), HSX (Anderson *et al.* 1995), NCSX (Zarnstorff *et al.* 2001), ESTELL (Drevlak *et al.* 2013), QHS48 (Ku & Boozer 2010) and ARIESCS (Najmabadi *et al.* 2008). Within a given device, the overall correlation between cost functions is good. All measures qualitatively show a similar behaviour: QS improves as we move from the boundary inwards. However, in regions closest to the axis, this correspondence is not as strong (especially in configurations such as GAR and QHS48). However, the same cost function correspondence does not seem to hold if we make comparisons between different configurations.

To quantify this lack of ordering across devices, we compute the Spearman correlation coefficient, which considers correlation only taking order into account. We show this in figure 3. The high correlation between cost functions within a particular device (which in all cases was ~ 0.9 – 1) clearly does not hold across devices. Although \hat{f}_C and \hat{f}_T retain a high degree of correlation (~ 0.9), the conventional \hat{f}_B measure seems largely uncorrelated to the others. (Of course, in this inter-configuration comparison the normalisation factors chosen are particularly important.)

¹https://github.com/landreman/vmec_equilibria.git.

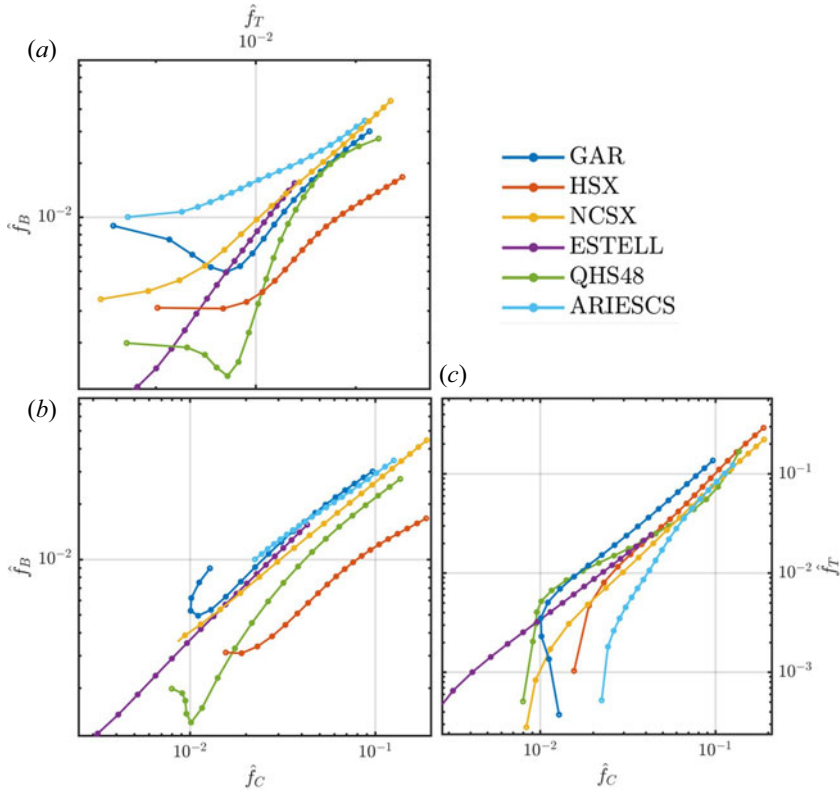


FIGURE 2. Comparison of defined cost functions in quasisymmetric designs. Plots in logarithmic scale including a comparison of the cost functions \hat{f}_B , \hat{f}_C and \hat{f}_T for a number of quasisymmetric devices. The scatter points for each device indicate values at a number of equally spaced magnetic flux surfaces joined by lines in order from the centre to the edge of the configurations (the rightmost points represent the edge).

The loss of correlation across machines emphasises the main point on universality from a practical perspective. Depending on the form of cost function taken, one could say either A or B is more quasisymmetric. The statement has only a relative meaning at best for these optimised configurations. Without a physical basis to the measures it is difficult to argue in favour of one definition over the other. For a more meaningful statement, one should instead compare these functions through some physical property. Ultimately we are interested in finding configurations that are close to QS because we seek some of the physical properties that QS confers upon magnetic configurations. In [figure 3](#) we compare ϵ_{eff} (which also contains some arbitrary normalisation factors in its definition; Nemov *et al.* 1999) with the three cost functions for the devices of [figure 2](#). Other physics measures such as alpha particle confinement could also be suited for such comparison (Bader *et al.* 2019). There exists some correlation between the transport and the normalised values of f_C and f_T . Especially remarkable is the seemingly little correlation between \hat{f}_B , which is widely used for optimisation, and ϵ_{eff} . Overall, for these ‘quasisymmetric’ configurations, the QS cost functions appear not to be good indicators of their transport levels. Similar behaviour has been observed in previous studies of quasisymmetric configurations (Martin & Landreman 2020).

In summary, when comparing quasisymmetric configurations, there is no universal measure of QS. The comparison depends on the metric used, and these differ from

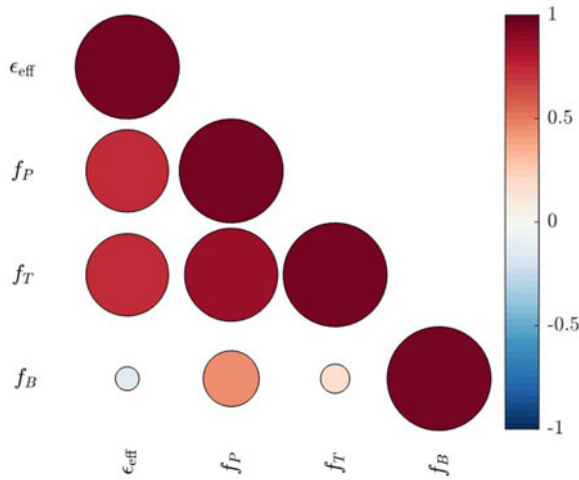


FIGURE 3. Correlation between edge values of cost functions and ϵ_{eff} of different quasisymmetric configurations. Diagram showing the Spearman correlation between the edge values of the different cost functions and ϵ_{eff} of the configurations in figure 2. The colour represents the coefficient (as does the diameter of the coloured circles).

each another. To make a comparison physically meaningful, especially close to optimised configurations, we are in need of resorting to direct physical measures of the configuration (e.g. neoclassical transport or alpha particle confinement).

4.2. Optimising for QS

The differences in the formal structure of the cost functions in conjunction with the difference in behaviour close to QS suggest that the different measures of QS will behave quite differently as optimisation cost functions.

From a formal perspective, we would expect to find f_B and f_C to perform similarly, while the nonlinear f_T to differ significantly. We now present some examples to illustrate the differences in the optimisation. We consider an example optimisation problem in which QS is targeted throughout the plasma volume. The optimisation parameters correspond to the Fourier modes describing the outermost surface:

$$\left. \begin{aligned} R &= \sum_{nm} R_{nm} \cos(m\vartheta - nN_p\phi_c), \\ Z &= \sum_{nm} Z_{nm} \sin(m\vartheta - nN_p\phi_c), \end{aligned} \right\} \quad (4.4)$$

where N_p is the number of field periods, ϕ_c is the cylindrical angle and ϑ is a general poloidal angle. The configuration optimised is based on a quasihelicallysymmetric configuration² with $N_p = 4$ field periods. We investigate two aspects of the problem. First, we consider slices of parameter space with the purpose of observing the differences in QS measures in practice beyond the region close to the minima. This should give a partial idea of how the problem changes as a result of changing the form of the optimisation cost function. Then, we perform various optimisations using STELLOPT (Lazerson *et al.*

²The basis of the optimisation is the 34DOF_varyBoundary_targetQuasisymmetry scenario in https://github.com/landreman/stellopt_scenarios.

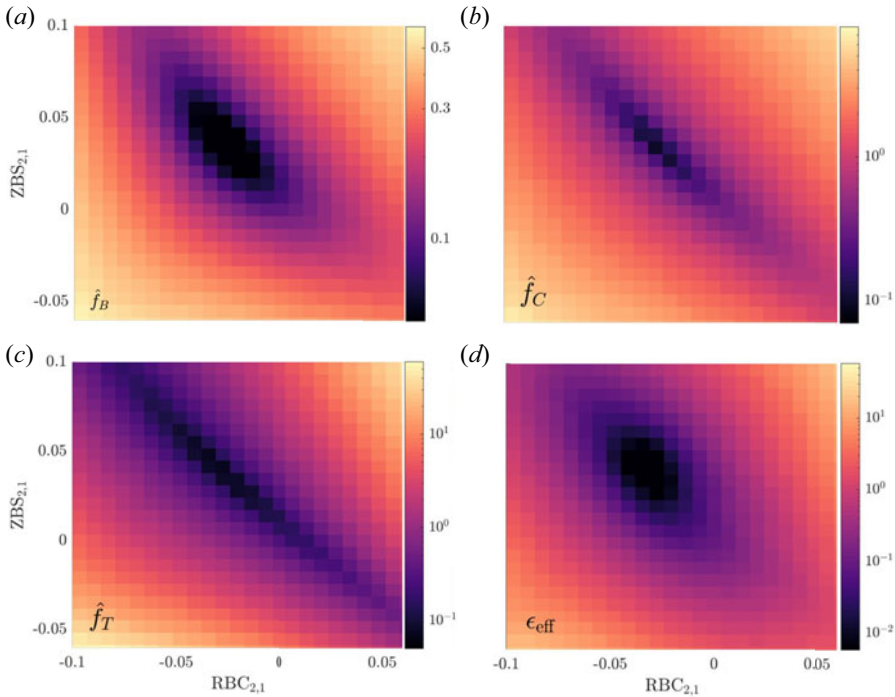


FIGURE 4. Parameter space spanned by $n = 2, m = 1$ surface mode for $\hat{f}_B, \hat{f}_C, \hat{f}_T$ and ϵ_{eff} . The plots show, clockwise, the cost functions $\hat{f}_B, \hat{f}_C, \epsilon_{\text{eff}}$ and \hat{f}_T as a function of the $(2, 1)$ surface modes. The parameter scan is performed using the map option of STELLOPT around the stellarator design in the footnote of p. 28. The minima for each of these are found (clockwise) at: $(-0.031, 0.031), (-0.031, 0.031), (-0.037, 0.045)$ and $(-0.037, 0.037)$.

2020) with implementations of the three different forms of the cost functions. Let us start with the exploration of parameter space. Figure 4 presents the evaluation of the QS metrics and ϵ_{eff} as a function of two parameters characterising the surface (the $n = 2, m = 1$ modes of the boundary (4.4)), and Figure 5 the magnitude of the associated gradient. The metrics are evaluated at six equally spaced magnetic flux surfaces, and their contributions are added together.

The objective landscape varies depending on the metrics used. Qualitatively, though, all the measures show a similar behaviour. The metrics share a region of reduced cost, which seem to roughly coincide. In a global sense, the metrics seem not to disagree with each other. This might appear as a surprise given the observations of the QS configurations and some of the formal discussion above. We should, however, remember that there exist certain changes that affect all metrics favourably. Away from a QS configuration it thus seems that there are directions in which one may reduce many of these modes such that all metrics decrease. This is supported by the parameter space representations in figure 4. However, note that the similarity in global trends does not preclude differences in local trends.

First, the relative variation of the cost functions in space is significantly different. Gradients seem to be relatively strong for the highly nonlinear measures \hat{f}_T and ϵ_{eff} , while the other two metrics behave more smoothly. This distinction in the gradients also seems to leave a region close to the minima much shallower for \hat{f}_T and ϵ_{eff} , comparatively, which may lead to difficulties in accessing the exact minima. Second, beyond these differences

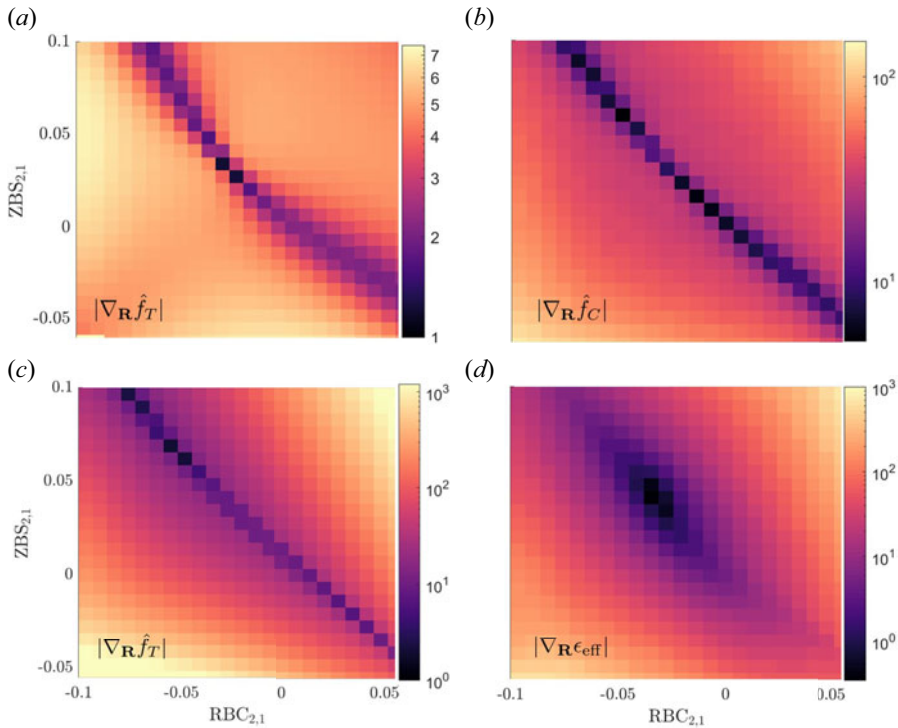


FIGURE 5. Gradients in parameter space spanned by $n = 2, m = 1$ surface mode for $\hat{f}_B, \hat{f}_C, \hat{f}_T$ and ϵ_{eff} . The plots show, clockwise, the parameter space gradients of $\hat{f}_B, \hat{f}_C, \epsilon_{\text{eff}}$ and \hat{f}_T as a function of the $(2, 1)$ surface modes. Scales of the gradients are significantly different between measures.

in magnitude, there are also differences in the shape and location of the minima. The region near a minimum can be more or less elongated depending on the measure, and, importantly, the minimiser is located at different values. In this particular case, minimisers vary by up to $\sim 20\% - 50\%$ in the boundary coefficients between different measures (although \hat{f}_B and \hat{f}_C have, in this case, roughly the same minimiser). In brief, although far from a minimum the various metrics may exhibit a qualitatively similar behaviour, close to the minimum there are significant differences. These latter differences are consistent with the discussion on the lack of universality in § 4.1.

Of course, this illustration of the optimisation space is only one slice of a multidimensional manifold. The large number of degrees of freedom means that the parameter space slices will change as other parameters do. In general, this makes optimisation a highly complex process in which the interaction between different dimensions may amplify the differences between metrics. In addition, the algorithm employed to move around in optimisation space will also affect the result of a given optimisation. A discussion on how the optimisation is likely to perform is therefore extremely difficult, and we instead present an illustrative example. With that being said, from our discussion we expect the optimisation of \hat{f}_B and \hat{f}_C to be similar, while the nonlinearity of \hat{f}_T to change the optimisation significantly. We hypothesise here that there is a rough relation between changes in boundary shape and the Fourier content of B .

We present the resulting cross-sections and values associated with the various metrics from optimisations in which QS is targeted in a volume using the different forms of

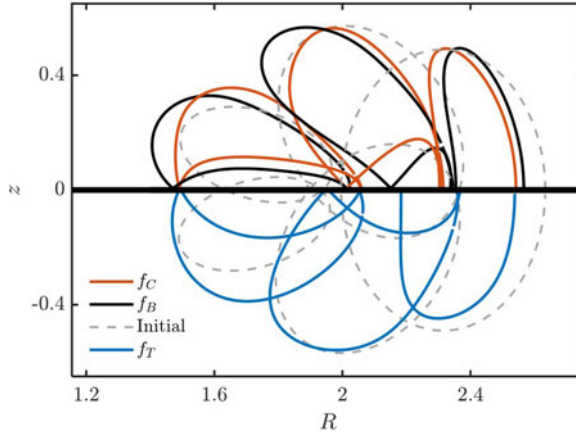


FIGURE 6. Comparison of cross-sections of optimisation results for different cost functions. Plot showing the cross-section of the stellarators as obtained from the optimisation using different forms of the cost function. The optimisation stops when the optimiser is unable to make further improvement (the characteristic change in the parameters during an iteration is 10^{-4} – 10^{-3}). The broken line represents the starting point, while the black, blue and red lines represent the f_B, f_T and f_C optimisation results, respectively. The cross-sections are only presented in one half of the plane but the missing plane can be reconstructed symmetrically.

Opt.	\hat{f}_B	\hat{f}_C	\hat{f}_T	ϵ_{eff}	Asp.
Start	0.54	5.3	24	66	6.3
f_B	3.3×10^{-3}	2.4×10^{-2}	3.0×10^{-2}	5.1×10^{-5}	7.1
f_C	2.3×10^{-2}	7.6×10^{-2}	0.50	3.9×10^{-3}	6.8
f_T	6.2×10^{-2}	1.9×10^{-1}	4.8×10^{-2}	3.6×10^{-2}	6.0

TABLE 2. Averaged metric values for optimised stellarators. Summary of the average value of the metric over the volume of the optimised stellarators with respect to the different cost functions.

QS metrics as cost functions in figure 6 and table 2. We start from a configuration that is deliberately chosen to be some distance away in parameter space from the QS configuration found in the footnote of p. 28. The optimisation uses a BFGS gradient-based method (as part of the suite MANGO³) in which there is a single scalarised cost function with 32 degrees of freedom in the form of Fourier components (from 0 through 3, keeping the major radius R_{00} fixed) defining the surface as in (4.4). The process does not penalise any additional property such as the commonly targeted rotational transform or aspect ratio. By studying a simplified problem, we hope to gain more insight into the differences between the various QS metrics. The optimisation is performed for several values of the finite-difference step size to ensure convergence.

Comparing cross-sections in figure 6 and the values in table 2 we observe that the resulting optima are quite different from one another. This emphasises the points discussed in this paper: that the precise form of the cost function has important qualitative implications. The optimisation using the standard form of \hat{f}_B seems, in this case, to perform

³<https://github.com/hiddenSymmetries/mango>.

best in allowing the configuration to reach a state of minimal transport and QS measures. The optimisation by \hat{f}_C is not far away, and a comparison of the cross sections shows similar shaping features. Interestingly, optimising for \hat{f}_C does not seem to yield the smallest value for \hat{f}_C (which \hat{f}_B optimisation does). This suggests that the optimisation for \hat{f}_C gets stuck in some effective local minimum.

In light of the comparison made in § 3, and with the hypothesis that the surface variations are roughly related to variations in B , we can think of optimisation as a process that rearranges energy in the space of Fourier modes of $|B|$. The relation is generally complex and nonlinear, but thinking in these terms may still shed some light on the differences in optimisation. In this rearrangement of energy in Fourier space, \hat{f}_B contains a large space of degenerate interchanges. Energy can be freely exchanged between asymmetric modes of B , regardless of their mode numbers. In that sense, reducing \hat{f}_B means, necessarily, reducing some asymmetric mode. On the other hand, the cost function \hat{f}_C distinguishes between modes. Different weights at different mode numbers lift part of the degeneracy in the exchange of mode energy that occurs for \hat{f}_B . This distinction effectively excludes certain regions of the mode space from being used for the restacking of the mode energy. In addition to decreasing the magnitude of an asymmetric mode, moving energy to modes closer to the symmetry direction will decrease \hat{f}_C . This could be a mixed blessing. Having this additional mechanism could be a way to avoid certain local minima by adding a new direction of descent. However, it could also present additional local minima if the optimisation prematurely reduces modes very close to the symmetry direction.

Relating changes in the mode content to changes in the boundary, we expect these two metrics to facilitate different ways of modifying the boundary. It is not clear which metric will lead to improved optimisation. We have here presented a single numerical exercise in which \hat{f}_B appeared to perform better, but the performance of these metrics will likely depend on the particulars of the optimisation problem as well as other aspects such as the optimisation algorithm used.

The optimisation of \hat{f}_T in the presented example exhibits the worst performance of all. The cross-sections show large differences with respect to the other configurations, suggesting that optimisation has stopped at a distinct local minimum. That the measure \hat{f}_T is lower in the \hat{f}_B -optimised case serves as evidence that the optimisation of \hat{f}_T terminated at a different local minimum. The presence of nonlinearities in f_T is perhaps responsible for this additional minimum. In the light of the Fourier space energy picture, \hat{f}_T not only depends on the magnitude and the location of the mode energy, but also on the relative location of modes. This nonlinear coupling complicates the problem, potentially leading to additional local minima. Furthermore, the nonlinearity of \hat{f}_T results in an objective landscape which appears flatter near the minimum (see figure 4), which may make difficult reaching an accurate optimum. Although this qualitative picture fits with observations made for this optimisation example, we remark that the behaviour is likely to vary between problems. An important additional point of distinction of the \hat{f}_T optimisation with respect to the others is, as emphasised in § 2, that the form of symmetry of the configuration is not specified. Hence, the optimisation must locally decide upon the type of QS towards which to evolve. If close to a configuration with good QS, optimisation will naturally bring the system towards the corresponding helicity. However, generally it is difficult to predict given an initial configuration the form of QS towards which optimisation will tend.

In the example presented, the aspect ratio of the configuration does not seem to change significantly between and along optimisations (values lie within $\sim 15\%$ of the initial point). This seems counterintuitive, as from a near-axis perspective one would expect a larger aspect ratio to allow for a more quasisymmetric solution. This follows from the possibility of satisfying QS conditions sufficiently close to the magnetic axis, as the relevant expansion parameter for near axis scales as $\sqrt{\psi/B}/R$ (Landreman & Sengupta 2019). Since we are fixing the total flux, decreasing the minor radius of the toroidal configuration can increase the differential flux over which the QS error is small. In practice, however, the optimiser has difficulty in finding these large-aspect-ratio minima. The difficulty in changing the aspect ratio without sacrificing QS along the way and the choice of parameterisation (Henneberg, Helander & Drevlak 2021) (4.4) might provide an explanation for this behaviour.

What conclusions can we thus draw from this illustrative example? First, the form of the cost function must be carefully considered. Here we only examined QS and observed that the optimisation result varies significantly depending on the form of the cost function. Similar behaviour will likely affect the optimisation of other physical measures. Second, we see in this example that the ‘simplest’ forms (in this case, f_B and f_C) seemed to provide a better global performance. (By simpler we mean a lower number of field derivatives and a simpler dependence on the Fourier modes.) Increased complexity seems to make optimisation harder. However, it is difficult to make general statements based only on this example, and there might be exceptions not treated in this paper.

As discussed in the previous section, close to local QS minima, improvements in QS metrics do not always lead to improvements in physical properties. Thus insisting on making marginal improvements on \hat{f}_B (and for that matter, any of the other QS measures) does not seem to be a useful exercise. Close to a QS minimum it does seem appropriate to prioritise optimisation with respect to more physical measures. Optimising stellarators based solely on physical properties has been shown to be successful in the literature (Hirshman *et al.* 1998, 1999; Spong *et al.* 2000). Here we propose a hybrid approach between optimising for QS alone and optimising for physical metrics (without QS). The QS optimisation should bring the search close to a region of configurations with QS-like properties following a simpler form of cost function. Then for refinement, it will follow the more detailed physical cost function optimisation. Perhaps insisting on QS at all times may stand its ground if an absolute QS zero could be reached (a point with clear physical meaning), such as in the optimisation for QS on a single magnetic surface. But even then, stellarator design usually requires additional physical considerations.

We seem then to need to alternate between an optimisation dominated by the QS metrics and an optimisation guided by other physics metrics. A way forward is to impose QS as an inequality constraint, so that physical metrics can dominate optimisation when QS is ‘small enough’. Of course there is the unavoidable difficulty of choosing the value of a QS metric to indicate ‘sufficient QS.’ Picking a fixed value will modify the shape of optimisation space in a non-trivial way. Instead of choosing a fixed value, a more dynamic alternative would be to turn QS optimisation on and off guided by the relative changes or gradient norms of the QS metrics. One might choose to not optimise for QS at all, instead initialising the search from an already fairly quasisymmetric configuration. This would place the optimisation in the right basin, but lead to refinements in it based on physically relevant features. Possible initial seeds could be constructed from near-axis solutions.

As explored, the non-universality of the QS metrics close to minima seems to require amendment by introducing additional physics metrics. At the same time, detailed physical calculations are likely to lead to complicated parameter spaces including multiple local

minima compared with those of the simpler QS metrics. Using QS metrics in conjunction with a more complex physical metric may provide the best of both worlds. This could shed some light on recent results in Bader *et al.* (2019). In this work, optimisation of QS through f_B together with a more complex physical proxy for α particle losses seems to perform better than each of the pieces separately. We hypothesise that f_B allows for a smoother global navigation (avoiding some potential local minima), while the physics metric helps in finding a more physically meaningful final minimum. Using both metrics in an appropriately balanced way seems to lead to a convenient compromise in which benefits of both aspects are brought together.

Stellarator design often requires multi-objective optimisation. Given the nonlinearity of the objectives of interest, adding additional terms to the cost function can change the problem dramatically. Thus, the differences observed above for the QS metrics will only be more complex in the presence of other physics objectives. Those differences may be amplified or reduced depending on their relation to the behaviour of the other metrics. Such an analysis will be necessary to reason how to best compose QS with these other metrics. In these complex scenarios, the idea of initially seeding optimisation with an approximately QS configuration gains strength, as it would make the problem simplest. Further study is required to verify these proposed optimisation strategies.

5. Conclusion

In this paper we have explored three main measures of QS: the standard Fourier form, the two-term formulation and the triple product form. We performed an analytical comparison of their origin and mathematical structure. Particularly important are their differences in weighting the asymmetric modes of B in Boozer coordinates. While the standard Fourier form treats all modes equally, the two-term form weights them differently, and the triple vector formulation involves a mode convolution.

The differences in these valid QS metrics show that there is no universal measure of QS. While the measures are all rigorously equivalent when QS is satisfied exactly (assuming $\mathbf{j} \cdot \nabla \psi = 0$), there is no unique measure that quantifies deviations from QS. This lack of universality makes the comparison between configurations particularly difficult near approximately QS solutions, and some physically relevant property (such as neoclassical transport through ϵ_{eff} or alpha particle confinement) would be needed for a more meaningful comparison. Near those local minima that are not exactly QS, optimising QS cost functions for marginal improvements generally lacks physical meaning. Away from the minima, the global structures of the QS metrics are qualitatively similar, providing more physical meaning to the metrics.

We also show that the mathematical form of the measure affects the optimisation significantly via an example. In the particular example presented, the complexity of a cost function seems to lead to ending in alternative local minima.

In light of the above, we suggest enforcing approximate QS either through an inequality constraint (so that optimisation close to minima is dominated by some direct physics measure) or through an appropriate initial configuration. Additional attention should also be paid to the form of physics or engineering cost functions.

Acknowledgements

The authors would like to acknowledge Daniel Dudt for fruitful discussion.

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

Funding

This research is primarily supported by a grant from the Simons Foundation/SFARI (560651, A.B.) and DoE contract no. DE-AC02-09CH11466. E.R. was also supported by the Charlotte Elizabeth Procter Fellowship at Princeton University. E.P. was also supported by the Presidential Postdoctoral Research Fellowship at Princeton University.

Declaration of interests

The authors report no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Appendix A. Magnetic differential relation between f_C and f_T

Although the triple vector product form of QS seems distinct from the other forms given its nonlinear character, we can establish a simple relation between f_T and f_C . To obtain this relation it is most convenient to use the same tools that we used in the derivations of the different forms of QS in § 2. Away from QS, one may still construct the symmetry field as $\tilde{\mathbf{u}} = \nabla\Phi \times \nabla B / (\mathbf{B} \cdot \nabla B)$, where we assume $\mathbf{B} \cdot \nabla B \neq 0$. At extremal points of the magnetic field magnitude along \mathbf{B} , this construction fails. We have seen the role of the extrema, for example, in the discussion of single-particle dynamics (§ 3). Requiring pseudosymmetry takes care of these, as it guarantees $\mathbf{B} \cdot \tilde{\mathbf{u}}$ to be well-behaved. Constructing this vector field $\tilde{\mathbf{u}}$ in this way is equivalent to enforcing (2.1) and (2.2). With these as well as $\nabla \cdot \mathbf{B} = 0$:

$$\mathbf{B} \cdot \nabla (\mathbf{B} \cdot \tilde{\mathbf{u}}) = B^2 \nabla \cdot \tilde{\mathbf{u}}, \quad (\text{A } 1)$$

using $\mathbf{j} \cdot \nabla \psi = 0$ and appropriate vector identities to rewrite the left-hand side. Now, from $\mathbf{B} \cdot \tilde{\mathbf{u}} / \Phi' = C + f_C / (\mathbf{B} \cdot \nabla B)$, it then follows that f_C and f_T are related to each other through a magnetic differential equation:

$$\mathbf{B} \cdot \nabla \left(\frac{f_C}{\mathbf{B} \cdot \nabla B} \right) = - \frac{f_T}{(\mathbf{B} \cdot \nabla B)^2}. \quad (\text{A } 2)$$

This magnetic differential equation shows that f_T does, in fact, contain a higher-order derivative in comparison with f_C .

REFERENCES

- AKÇELİK, V., BIROS, G., GHATTAS, O., HILL, J., KEYES, D. & VAN BLOEMEN WAANDERS, B. 2006 Parallel algorithms for PDE-constrained optimization. In *Parallel Processing for Scientific Computing* (eds M.A. Heroux, P. Raghavan & H.D. Simon), pp. 291–322. SIAM.
- ANDERSON, F.S.B., ALMAGRI, A.F., ANDERSON, D.T., MATTHEWS, P.G., TALMADGE, J.N. & SHOHET, J.L. 1995 The Helically Symmetric Experiment (HSX) goals, design and status. *Fusion Technol.* **27** (3 T), 273–277.
- BADER, A., DREVLAK, M., ANDERSON, D.T., FABER, B.J., HEGNA, C.C., LIKIN, K.M., SCHMITT, J.C. & TALMADGE, J.N. 2019 Stellarator equilibria with reactor relevant energetic particle losses. *J. Plasma Phys.* **85** (5), 905850508.
- BERNARDIN, M.P., MOSES, R.W. & TATARONIS, J.A. 1986 Isodynamical (omnigenous) equilibrium in symmetrically confined plasma configurations. *Phys. Fluids* **29** (8), 2605–2611.
- BHATTACHARJEE, A., HAYASHI, T., HEGNA, C.C., NAKAJIMA, N. & SATO, T. 1995 Theory of pressure-induced islands and self-healing in three-dimensional toroidal magnetohydrodynamic equilibria. *Phys. Plasmas* **2** (3), 883–888.

- BOOZER, A.H. 1981 Plasma equilibrium with rational magnetic surfaces. *Phys. Fluids* **24** (11), 1999–2003.
- BOOZER, A.H. 1983 Transport and isomorphic equilibria. *Phys. Fluids* **26** (2), 496–499.
- BURBY, J.W., KALLINIKOS, N. & MACKAY, R.S. 2020 Some mathematics for quasi-symmetry. *J. Math. Phys.* **61** (9), 093503.
- CALVO, I., PARRA, F.I., VELASCO, J.L. & ALONSO, J.A. 2014 Flow damping in stellarators close to quasisymmetry. *Plasma Phys. Control. Fusion* **57** (1), 014014.
- CONSTANTIN, P., DRIVAS, T.D. & GINSBERG, D. 2021 On quasisymmetric plasma equilibria sustained by small force. *J. Plasma Phys.* **87** (1), 905870111.
- DEKEYSER, W., REITER, D. & BAELEMAN, M. 2014 Automated divertor target design by adjoint shape sensitivity analysis and a one-shot method. *J. Comput. Phys.* **278**, 117–132.
- DREVLAK, M., BROCHARD, F., HELANDER, P., KISSLINGER, J., MIKHAILOV, M., NÜHRENBURG, C., NÜHRENBURG, J. & TURKIN, Y. 2013 ESTELL: a quasi-toroidally symmetric stellarator. *Contrib. Plasma Phys.* **53** (6), 459–468.
- FITZPATRICK, R. 1993 Interaction of tearing modes with external structures in cylindrical geometry (plasma). *Nucl. Fusion* **33** (7), 1049–1084.
- GARABEDIAN, P.R. 2008 Three-dimensional analysis of tokamaks and stellarators. *Proc. Natl Acad. Sci.* **105** (37), 13716–13719.
- GARABEDIAN, P.R. & MCFADDEN, G.B. 2009 Design of the DEMO fusion reactor following ITER. *J. Res. Natl Inst. Stand. Technol.* **114** (4), 229.
- GARREN, D.A. & BOOZER, A.H. 1991 Existence of quasihelically symmetric stellarators. *Phys. Fluids B* **3** (10), 2822–2834.
- GRAFAKOS, L. 2008 *Fourier Analysis on the Torus*, pp. 1–88. Springer.
- HEGNA, C.C. 2011 Kinetic shielding of magnetic islands in 3D equilibria. *Plasma Phys. Control. Fusion* **53** (2), 024003.
- HEGNA, C.C. 2012 Plasma flow healing of magnetic islands in stellarators. *Phys. Plasmas* **19** (5), 056101.
- HELANDER, P. 2014 Theory of plasma confinement in non-axisymmetric magnetic fields. *Rep. Prog. Phys.* **77** (8), 087001.
- HELANDER, P. & SIMAKOV, A.N. 2008 Intrinsic ambipolarity and rotation in stellarators. *Phys. Rev. Lett.* **101**, 145003.
- HENNEBERG, S., DREVLAK, M., NUHRENBURG, C., BEIDLER, C., TURKIN, Y., LOIZU, J. & HELANDER, P. 2019 Properties of a new quasi-axisymmetric configuration. *Nucl. Fusion* **59** (2), 026014.
- HENNEBERG, S., HELANDER, P. & DREVLAK, M. 2021 Representing the boundary of stellarator plasmas. *J. Plasma Phys.* **87** (5), 905870503.
- HIRSHMAN, S.P., SPONG, D.A., WHITSON, J.C., LYNCH, V.E., BATCHELOR, D.B., CARRERAS, B.A. & ROME, J.A. 1998 Transport optimization and MHD stability of a small aspect ratio toroidal hybrid stellarator. *Phys. Rev. Lett.* **80**, 528–531.
- HIRSHMAN, S.P., SPONG, D.A., WHITSON, J.C., NELSON, B., BATCHELOR, D.B., LYON, J.F., SANCHEZ, R., BROOKS, A., Y-FU, G., GOLDSTON, R.J., *et al.* 1999 Physics of compact stellarators. *Phys. Plasmas* **6** (5), 1858–1864.
- HO, D.D. & KULSRUD, R.M. 1987 Neoclassical transport in stellarators. *Phys. Fluids* **30** (2), 442–461.
- KU, L. & BOOZER, A. 2010 New classes of quasi-helically symmetric stellarators. *Nucl. Fusion* **51** (1), 013004.
- LANDREMAN, M. & CATTO, P.J. 2012 Omnigenity as generalized quasisymmetry. *Phys. Plasmas* **19** (5), 056103.
- LANDREMAN, M. & SENGUPTA, W. 2018 Direct construction of optimized stellarator shapes. Part 1. Theory in cylindrical coordinates. *J. Plasma Phys.* **84** (6), 905840616.
- LANDREMAN, M. & SENGUPTA, W. 2019 Constructing stellarators with quasisymmetry to high order. *J. Plasma Phys.* **85** (6), 815850601.
- LAZERSON, S., SCHMITT, J., ZHU, C., BRESLAU, J. & STELLOPT DEVELOPERS AND USDOE OFFICE OF SCIENCE 2020 STELLOPT, Version 2.7.5. <https://www.osti.gov/biblio/I617636>.
- LITTLEJOHN, R.G. 1983 Variational principles of guiding centre motion. *J. Plasma Phys.* **29** (1), 111–125.
- MARTIN, M.F. & LANDREMAN, M. 2020 Impurity temperature screening in stellarators close to quasisymmetry. *J. Plasma Phys.* **86** (3), 905860317.

- MIKHAILOV, M., SHAFRANOV, V., SUBBOTIN, A., ISAEV, M., HRENBURG, J.N., ZILLE, R. & COOPER, W. 2002 Improved α -particle confinement in stellarators with poloidally closed contours of the magnetic field strength. *Nucl. Fusion* **42** (11), L23–L26.
- MYNICK, H.E. 2006 Transport optimization in stellarators. *Phys. Plasmas* **13** (5), 058102.
- NAJMABADI, F., RAFFRAY, A., ABDEL-KHALIK, S., BROMBERG, L., CROSATTI, L., EL-GUEBALY, L., GARABEDIAN, P., GROSSMAN, A., HENDERSON, D., IBRAHIM, A., *et al.* 2008 The ARIES-CS compact stellarator fusion power plant. *Fusion Sci. Technol.* **54** (3), 655–672.
- NEMOV, V.V., KASILOV, S.V. & KERNBICHLER, W. 2014 Collisionless high energy particle losses in optimized stellarators calculated in real-space coordinates. *Phys. Plasmas* **21** (6), 062501.
- NEMOV, V.V., KASILOV, S.V., KERNBICHLER, W. & HEYN, M.F. 1999 Evaluation of $1/\nu$ neoclassical transport in stellarators. *Phys. Plasmas* **6** (12), 4622–4632.
- NÜHRENBERG, J. & ZILLE, R. 1988 Quasi-helically symmetric toroidal stellarators. *Phys. Lett. A* **129** (2), 113–117.
- PAUL, E.J., ANTONSEN, T., LANDREMAN, M. & COOPER, W.A. 2020 Adjoint approach to calculating shape gradients for three-dimensional magnetic confinement equilibria. Part 2. Applications. *J. Plasma Phys.* **86** (1), 905860103.
- PLUNK, G. & HELANDER, P. 2018 Quasi-axisymmetric magnetic fields: weakly non-axisymmetric case in a vacuum. *J. Plasma Phys.* **84** (2), 905840205.
- REIMAN, A. & BOOZER, A.H. 1984 Island formation and destruction of flux surfaces in three-dimensional MHD equilibria. *Phys. Fluids* **27** (10), 2446–2454.
- RODRÍGUEZ, E. & BHATTACHARJEE, A. 2021a Islands and current singularities in quasisymmetric toroidal plasmas. *Phys. Plasmas* **28** (9), 092506.
- RODRÍGUEZ, E. & BHATTACHARJEE, A. 2021b Solving the problem of overdetermination of quasisymmetric equilibrium solutions by near-axis expansions. I. Generalized force balance. *Phys. Plasmas* **28** (1), 012508.
- RODRÍGUEZ, E. & BHATTACHARJEE, A. 2021c Solving the problem of overdetermination of quasisymmetric equilibrium solutions by near-axis expansions. II. Circular axis stellarator solutions. *Phys. Plasmas* **28** (1), 012509.
- RODRÍGUEZ, E., HELANDER, P. & BHATTACHARJEE, A. 2020 Necessary and sufficient conditions for quasisymmetry. *Phys. Plasmas* **27** (6), 062501.
- RODRÍGUEZ, E., SENGUPTA, W. & BHATTACHARJEE, A. 2021 Generalized Boozer coordinates: a natural coordinate system for quasisymmetry. *Phys. Plasmas* **28** (9), 092510.
- 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** (6), 641–653.
- SKOVORODA, A.A. 2005 3D toroidal geometry of currentless magnetic configurations with improved confinement. *Plasma Phys. Control. Fusion* **47** (11), 1911–1924.
- SPLITZER, L. 1958 The stellarator concept. *Phys. Fluids* **1** (4), 253–264.
- SPONG, D., HIRSHMAN, S., WHITSON, J., BATCHELOR, D., SANCHEZ, R., CARRERAS, B., LYNCH, V., LYON, J., VALANJU, P., MINER, W., *et al.* 2000 Design studies of low aspect ratio quasi-omnigenous stellarators. *Nucl. Fusion* **40** (3Y), 563–567.
- TESSAROTTO, M., JOHNSON, J.L., WHITE, R.B. & ZHENG, L. 1996 Quasi-helical magnetohydrodynamic equilibria in the presence of flow. *Phys. Plasmas* **3** (7), 2653–2663.
- TESSAROTTO, M., JOHNSON, J.L. & ZHENG, L.J. 1995 Hamiltonian approach to the magnetostatic equilibrium problem. *Phys. Plasmas* **2** (12), 4499–4512.
- ZARNSTORFF, M., BERRY, L., BROOKS, A., FREDRICKSON, E., FU, G., HIRSHMAN, S., HUDSON, S., KU, L., LAZARUS, E., MIKKELSEN, D., *et al.* 2001 Physics of the compact advanced stellarator NCSX. *Plasma Phys. Control. Fusion* **43** (12A), A237.