Multifidelity methods for assessing energy loss in tokamak fusion reactors

Frederick Law

May 13, 2020

1. Introduction

Harnessing fusion energy, in which hydrogen atoms fuse in hot plasma to release energy, is a crucial step in improving global energy infrastructure. Not only is fusion energy extraordinarily clean, but it is far more sustainable and energetic than traditional energy sources. As a comparison, one metric ton of fusion energy material would provide a million times more energy than one metric ton of oil. Yet to reap these benefits, physicists and engineers must first design reactors to hold, or confine, the fusion plasma efficiently enough so that energy is not lost as the reaction occurs.

One popular design for fusion reactors is the tokamak, which uses powerful magnetic fields to spin the plasma around in a toroidal, donut-like shape (Fig.1). The fusion reaction inside the spinning plasma emits byproducts called *alpha particles* in all directions. These alpha particles have wildly different trajectories depending on the direction they are emitted. As a result, some alpha particles may escape the plasma which can lead to significant energy losses. These energy losses are exactly what we want to reduce in order to develop efficient fusion reactors. However, there is an inherent uncertainty here in even quantifying energy loss, since each alpha particle gets emitted in a random direction. Factoring in this uncertainty in emission direction is critical for assessing energy loss and thereby informing reactor design. When tokamak operators turn their reactor on, will they keep enough energy to successfully attain fusion?

We assess energy loss in tokamaks using the tools of *uncertainty quantification* (UQ), where we model the uncertainty in the emission direction of alpha particles by a probability distribution and we analyze the resulting statistics of their trajectories. To study these statistics, we typically need *many trajectories* using different samples of emission direction. However, numerical simulation of even a single trajectory is expensive since the dynamics for alpha particles have two distinct time scales. As a result, a prohibitively small time step is needed to resolve a fast time scale motion. Since numerically simulating individual trajectories is expensive, simulating many trajectories using different emission direction samples becomes rapidly intractable.

Our goal is to develop methods to limit the overall number of expensive particle trajectories that we need to simulate. Specifically, in section 3 we discuss how a *multifidelity* approach can reduce the number of expensive trajectories required by utilizing other cheaper models which are well correlated to the alpha particle dynamics.

2. Equations of motion and a naive method for estimating statistics

First we provide some details on our mathematical model for the dynamics of alpha particles. Specifically, we highlight some features of the model which make estimating statistics particularly challenging. Then we illustrate how these challenges manifest by applying a natural approach estimating statistics of our alpha particle model.

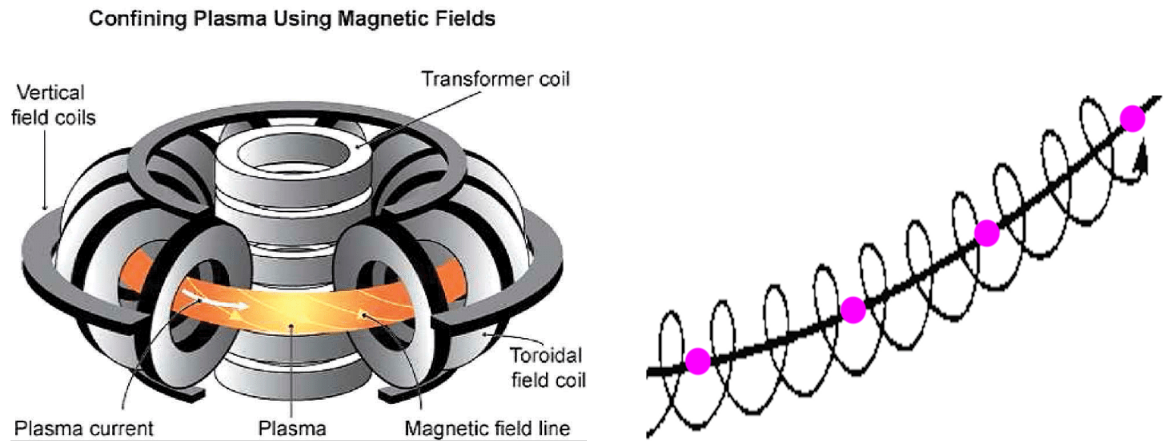

Figure 1: **Left:** Schematic of a tokamak reactor device. Various coils induce a toroidal magnetic field which spins the fusion plasma around the device. Inside of this plasma, reactions occur which emit the alpha particles as byproducts. **Right:** Zoomed in view of alpha particle trajectories inside of tokamaks. Their trajectories consist of a rapid oscillation around the slower moving center which we want to track (indicated by the pink dots).

2.1. Equations of motion for alpha particles in tokamaks

We consider the dynamics of an emitted alpha particle in the magnetic field of the tokamak reactor, where we neglect the electric and magnetic fields generated by the fusion plasma itself. The equations of motion are then give by the Lorentz force

$$m\ddot{\mathbf{x}}(t) = (+2e) \cdot \mathbf{v}(t) \times \mathbf{B}, \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{R}^3$ is the position vector of the alpha particle, $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$ is its velocity, and m is its mass. The reactor's magnetic field $\mathbf{B}$ is a vector field which depends on position, but not time. Alpha particles emitted in the tokamak follow the equations of motion (1) which cause the particles to orbit the tokamak device both toroidally (long way around the donut) and poloidally (short way around the donut through the hole). These dynamics are inherently multiscale in time, with a *small, rapid gyration whose period is the fastest time scale* in the system. However, the alpha particle tends to drift on a much slower time scale, and it is the position of the particle on the slow time scale that we are interested in, see Fig. 1.

For any initial position $\mathbf{x}(0)$ and initial velocity $\mathbf{v}(0)$ we can solve (1) to any time by solving (1). However, due to the uncertainty in the emission direction, there is uncertainty in $\mathbf{v}(0)$. We model this uncertainty by setting $\mathbf{v}(0)$ to be distributed uniform in all directions. The magnitude $||\mathbf{v}(0)||_2$ is determined by the energy released in the fusion reaction and is a fixed constant.

To emphasize the special role that direction plays over magnitude, we define the *emission direction* by the initial velocity normalized by its magnitude: $\boldsymbol{\omega} = \mathbf{v}(0)/||\mathbf{v}(0)||_2$. So $\boldsymbol{\omega}$ is a random variable distributed uniformly on the unit sphere in $\mathbb{R}^3$. Thus, for an initial position $\mathbf{x}(0)$ and random emission direction $\boldsymbol{\omega}$, we let $\mathbf{x}(t;\boldsymbol{\omega})$ denote the solution to (1) at time t with emission direction $\boldsymbol{\omega}$.

While $\mathbf{x}(t;\boldsymbol{\omega})$ is the alpha particle's location at time t based on uncertainty in emission direction, we want to measure how close the particle is to escaping the plasma since this is what leads to energy losses. To do this we use a certain flux function Ψ which maps positions inside the tokamak to $\mathbb{R}$, where the sublevel set $\{\mathbf{x} : \Psi(\mathbf{x}) \leq 0\}$ represents the plasma, and the level set $\{\mathbf{x} : \Psi(\mathbf{x}) = 0\}$ is outermost edge of the plasma. Thus we are interested in tracking the variable

$$Q_t(\boldsymbol{\omega}) := \Psi(\mathbf{x}(t;\boldsymbol{\omega})), \tag{2}$$

and we say that the alpha particle with emission direction $\boldsymbol{\omega}$ escapes if $Q_t(\boldsymbol{\omega}) \geq 0$ for any t.

Since we are interested in slow time scale statistics, we let $t_1 < t_2 < \cdots < t_{N_{\mathrm{MAX}}}$ denote the *time points of interest*, assuming $t_{j+1} - t_j$ is much larger than the period of rapid gyration for all j. Between

t_j and t_{j+1} the particle has done many small, rapid oscillations, but has moved little in its orbit around the reactor. For shorthand, we denote $Q^{(n)} = Q_{t_n}$, so

$$Q(\boldsymbol{\omega}) = \left(Q^{(1)}(\boldsymbol{\omega}), Q^{(2)}(\boldsymbol{\omega}), \ldots, Q^{(N_{\mathrm{MAX}})}(\boldsymbol{\omega})\right) \in \mathbb{R}^{N_{\mathrm{MAX}}}. \tag{3}$$

is our statistical quantity of interest arising from the uncertain emission direction $\boldsymbol{\omega}$. We will refer to the model (3) as the *full-orbit* model.

In order to a generate a single sample of our full-orbit model $Q(\boldsymbol{\omega})$, we start by sampling an emission direction $\boldsymbol{\omega}$. Then we solve (1) using $\boldsymbol{\omega}$ to times $t_1, t_2, \ldots, t_{N_{\mathrm{MAX}}}$. This is done by an expensive numerical integration where we must use a time step smaller than the fastest time scale, the period of rapid oscillation. Then we evaluate Ψ at $x(t_i; \boldsymbol{\omega})$ for $i = 1, \ldots, N_{\mathrm{MAX}}$ which gives $Q(\boldsymbol{\omega})$.

2.2. A first attempt for estimating statistics: Monte Carlo

Now that our mathematical model for the alpha particle dynamics is set up, we want to estimate statistics of $Q(\boldsymbol{\omega})$, namely the mean and variance. This would enable tokamak operators to estimate what proportion of alpha particles might escape the plasma. For simplicity, here we only focus on estimating the mean $\mathbb{E}Q^{(k)}(\boldsymbol{\omega})$ at the times t_k for $k = 1, \ldots, N_{\mathrm{MAX}}$.

A natural approach to estimate $\mathbb{E}Q(\boldsymbol{\omega})$ would be to just generate a very large number of i.i.d. samples of $Q(\boldsymbol{\omega})$ and then take their average. This is called the *Monte Carlo* (MC) method and the exact procedure is as follows:

Monte Carlo (MC)

1. Draw N i.i.d. samples $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_N$.
2. For each $i = 1, \ldots, N$, generate $Q(\boldsymbol{\omega}_i)$ using expensive numerical integration to solve (1).
3. For $k = 1, \ldots, N_{\mathrm{MAX}}$, estimate $\mathbb{E}Q^{(k)}(\boldsymbol{\omega})$, i.e. the mean at time point t_k by

$$\widehat{Q}^{(k)}_{\mathrm{MC},N} := \frac{1}{N} \sum_{i=1}^{N} Q^{(k)}(\boldsymbol{\omega}_i). \tag{4}$$

For notation, we always use a hat as in $\widehat{Q}^{(k)}_{\mathrm{MC},N}$ to denote an estimator of $\mathbb{E}Q^{(k)}(\boldsymbol{\omega})$. For each time t_k, this is an unbiased estimator for $\mathbb{E}Q^{(k)}(\boldsymbol{\omega})$ with variance $\mathrm{Var}(Q^{(k)}(\boldsymbol{\omega}))/N$. Based on empirical measurements, $\mathrm{Var}(Q^{(k)}(\boldsymbol{\omega}))$ is large for all times t_k, which means a massive N will be necessary for an accurate estimator.

Moreover, since each $Q(\boldsymbol{\omega}_i)$ is expensive to simulate due to the multiscale nature of the dynamics, using a large number of trajectories N is intractable. Note that the variance of the MC estimator decays like $1/N$. Since accuracy of estimators scales like the square root of variance, a ten-fold increase in accuracy requires a hundred-fold increase in N. In concrete terms, this means one more digit of accuracy requires 100 times more samples. Two more digits of accuracy requires 10,000 times more samples. Since our full-orbit model $Q(\boldsymbol{\omega})$ is already expensive to simulate, using more and more samples is out of the question. So we instead look for *variance reduction* techniques to shrink $\mathrm{Var}(Q^{(k)})$, which would allow for the same accuracy to be achieved for less computational cost.

3. Using multiple models simultaneously

First we use a simple example to motivate how introducing other random variables can yield variance reduction. We then extend this idea by introducing other models alongside the full-orbit model for variance reduction.

3.1. *Adding correlated random variables for variance reduction*

As a quick diversion, let X be any random variable and suppose we want to estimate $\mathbb{E}X$. Suppose that we already have some unbiased estimator $\widehat{X}$.

Now let $\rho(\cdot,\cdot)$ denote the correlation coefficient between two random variables, and suppose we have another variable Y which is well correlated with X, e.g. $\rho(X,Y) = 0.995$. Additionally, suppose we know $\mathbb{E}Y$ exactly and that we have an unbiased estimator $\widehat{Y}$ for $\mathbb{E}Y$. Then

$$\widehat{X}_* := \widehat{X} + c\left[\mathbb{E}Y - \widehat{Y}\right] \tag{5}$$

is also an unbiased estimator for $\mathbb{E}X$ for any coefficient c. Moreover if the coefficient c is correctly chosen, it can lead to a large variance reduction: $\mathrm{Var}\left(\widehat{X}_*\right) = \left(1 - \rho(X,Y)^2\right)\mathrm{Var}\left(\widehat{X}\right) \approx \mathrm{Var}\left(\widehat{X}\right)/100$. Hence, the more correlated Y is with X, the larger variance reduction we achieve. In this ideal case, the optimal c can be found analytically by minimizing the variance of (5) with respect to c.

3.2. *Combining the full-orbit model with other low-fidelity models*

Motivated by the simple example in section 3.1, we now seek to provide a similar variance reduction for the full-orbit model through the use of multifidelity methods. These methods speed up estimating statistics of high-fidelity models (expensive, very accurate, physics-based, X in section 3.1) by using low-fidelity models (cheap, inaccurate, any source, Y in section 3.1). As in section 3.1, the hope is that the low-fidelity model is correlated enough to the high-fidelity model to provide variance reduction. An example of a high-fidelity, low-fidelity combination common in elliptic PDE is a fine and coarse mesh discretization of the same elliptic operator.

For us, the high-fidelity model is our expensive full-orbit model to track alpha particles, but there are many possible low-fidelity models that can be used. We consider two low-fidelity models:

- An existing physics-based *drift-orbit* model which essentially takes the equations of motion (1) and averages out the fast gyration. This yields an ODE in 4 variables which no longer has fast time-scale phenomenon to resolve, and hence numerical integration is cheap.
- Piecewise linear interpolation. Since $\boldsymbol{\omega}$ lies on the unit sphere in $\mathbb{R}^3$, we can use polar and azimuthal angles to parametrize the unit sphere. With this, $Q_t(\boldsymbol{\omega})$ can be viewed as a mapping from $[-\pi/2, -\pi/2] \times [-\pi, \pi] \times [0, t_{N_{\mathrm{MAX}}}] \to \mathbb{R}$. Thus we can use full-orbit model simulations on a mesh in $[-\pi/2, -\pi/2] \times [-\pi, \pi]$ to fit a piecewise linear interpolant.

Let $q(\boldsymbol{\omega}) \in \mathbb{R}^{N_{\mathrm{MAX}}}$ denote a low-fidelity model at the same time points as the full-orbit model. Without loss of generality, assume any notation involving the full-orbit model $Q(\boldsymbol{\omega})$ also now holds for the low-fidelity model $q(\boldsymbol{\omega})$ but with the models changed. Then using our full-orbit model $Q(\boldsymbol{\omega})$ and a low-fidelity model $q(\boldsymbol{\omega})$, the multifidelity estimator (MF) is given by:

Multifidelity estimator (MF)

1. Draw N i.i.d. samples $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_N$. Then draw $M - N$ more i.i.d. samples $\boldsymbol{\omega}_{N+1}, \ldots, \boldsymbol{\omega}_M$.
2. For each $i = 1, \ldots, N$, generate $Q(\boldsymbol{\omega}_i)$ using expensive numerical integration to solve (1). For each $j = 1, \ldots, M$, generate low-fidelity model evaluations $q(\boldsymbol{\omega}_j)$.
3. For $k = 1, \ldots, N_{\mathrm{MAX}}$, estimate $\mathbb{E}Q^{(k)}(\boldsymbol{\omega})$, i.e. the mean at time point t_k by

$$\widehat{Q}^{(k)}_{\mathrm{MF},N} := \widehat{Q}^{(k)}_{\mathrm{MC},N} + c^{(k)}\left[\widehat{q}^{(k)}_{\mathrm{MC},M} - \widehat{q}^{(k)}_{\mathrm{MC},N}\right]. \tag{6}$$

where $c^{(k)}$ are constant coefficients.

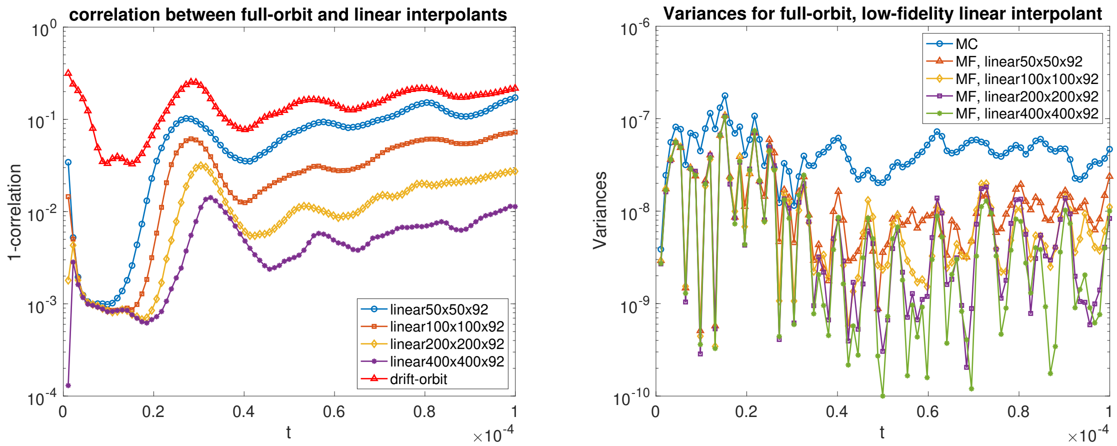

Figure 2: **Left:** Correlation of low-fidelity models to full-orbit model. Drift-orbit does the worst, and the piecewise linear interpolant improves for finer meshes. **Right:** Variances for MF using piecewise linear interpolant as the low-fidelity. In all cases, MF beats MC for the computational time. In both plots, lower is better.

Existing results in the literature [2] give us an optimal choice of $c^{(k)}$ and M for a fixed computational time. Let w be the ratio of the cost for a single high-fidelity trajectory $Q(\boldsymbol{\omega})$ to the cost for a single low-fidelity trajectory $q(\boldsymbol{\omega})$. Then using the optimal $c^{(k)}$ and M, the accuracy of MF is

$$\mathrm{Var}(\widehat{Q}^{(k)}_{\mathrm{MF},N}) = A^{(k)}\,\mathrm{Var}(\widehat{Q}^{(k)}_{\mathrm{MC},N}). \tag{7}$$

where $A^{(k)}$ is a term which goes to 0 rapidly as the correlation between the full-orbit and low-fidelity model goes to 1, and depends inversely on w. Hence, a good low-fidelity model $q((\boldsymbol{\omega})$ should be both cheap and well correlated to the full-orbit model to achieve a large variance reduction.

In Fig. 2, we see that the drift-orbit model has relatively poor correlation to the high-fidelity model. The piecewise linear interpolants were built on $L \times L$ meshes with $L = 50, 100, 200, 400$ and $N_{\mathrm{MAX}} = 92$. In all cases they had stronger correlation to the full-orbit model than the drift orbit. Drift-orbit also lost in in speed, with w around 350, but the interpolants had w near 400,000. We see that MF using any of the piecewise linear interpolants provides smaller variance for the same computational time compared to MC. All interpolation meshes get 10 times variance reduction at some time points, with the finest mesh achieving almost 100 times reduction at certain time points.

4. Conclusions

We used multifidelity methods to estimate statistics of alpha particles emitted in tokamak reactors, and with low-fidelity interpolation models we were able to provide a noticeable improvement over the Monte Carlo approach. Such multifidelity methods serve as a tool for tokamak researchers to efficiently assess energy losses due to escaping alpha particles. And in doing so, these methods can help inform tokamak design in order to minimize energy losses which would be a major step towards clean, sustainable fusion energy.

Future work includes using correlation between time points of interest t_k and t_{k+1} to improve estimation. Some work in this exists in the optimization literature under the name *information reuse*. Another direction is training machine learning models such neural nets to learn correlation to high-fidelity models, specifically to be then used in a multifidelity estimator.

References:

[1] L. Ng and K. Willcox. "Multifidelity approaches for optimization under uncertainty". Int. J. Numer. Meth. Engng **100**, 746 (2014).

[2] B. Peherstorfer, K. Willcox, and M. Gunzburger. "Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization". SIAM Review, **60**, 550 (2018).