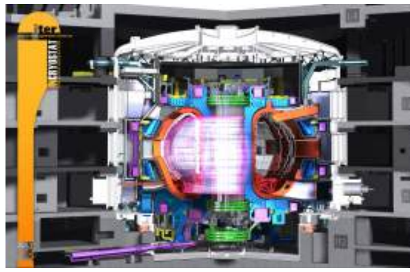
Efficient Computational Algorithms for Magnetic Equilibrium in a Fusion Reactor

Jiaxing Liang (Rice University)
Advisors: Matthias Heinkenschloss

Introduction

- Equilibrium Computation in magnetic confinement fusion devices Tokamaks.

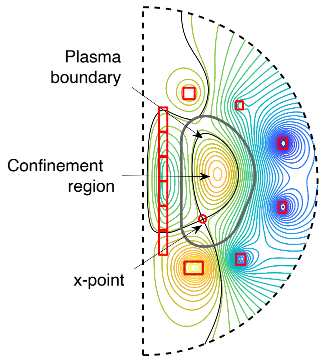


- **The Grad-Shafranov Equation:**

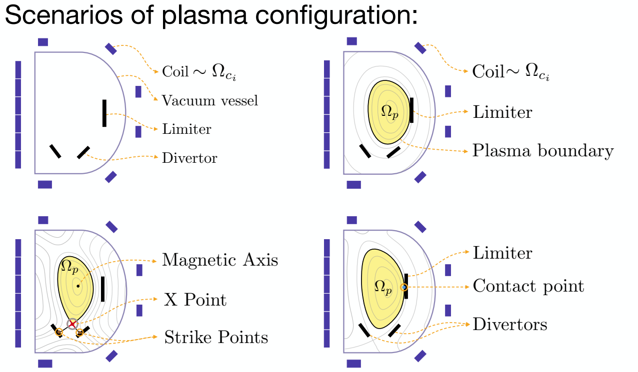
$$-\nabla\cdot\left(\frac{1}{\mu x}\nabla\psi\right)=\begin{cases} x\frac{d}{d\psi}p(\psi)+\frac{1}{2\mu x}\frac{d}{d\psi}g^2(\psi) & \text{in } \Omega_p(\psi)\\ I_i/S_i & \text{in } \Omega_{C_i}\\ 0 & \text{elsewhere}\end{cases}$$

$$\psi(0,y)=0; \qquad \lim_{\|(x,y)\|\to\infty}\psi(x,y)=0.$$

ψ: poloidal flux function, flow in poloidal direction with constant angle in the cylindrical coordinate.
p: hydrodynamic pressure, $p(\psi)$
I_i: current intensity in the coils
$\Omega_p(\psi)$: confinement region within plasma boundary
μ: magnetic permeability
g: toroidal field function, $g(\psi)$
S_i: cross section area of the coils
Ω_{c_i}: regions where the currents locate



- Semilinear elliptic, **free boundary problem.**
- **Uncertainty Quantification:** Current intensities of 12 external coils.
- Scenarios of plasma configuration:



Limited configuration: no separatrix - contact point - limiter prevents plasma from touching the vacuum vessel physically.
Diverted configuration: last closed line before limiter (separatrix) - x point & strike points - divertor removes impurities.

Objective

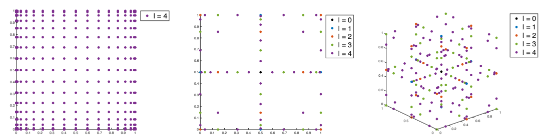
- **Quantify the impact of variations** in the current intensity on key physical quantities, like x-point & strike point locations etc..
- **Approximate the expectation** of the equilibrium configuration with sampling methods like Monte Carlo & **enhance efficiency** with surrogate computations.

Surrogate function

- **Build surrogate function** with sparse grid stochastic collocation to replace the nonlinear solve.

Idea: sets of current $\{\xi^{(k)}, k=1,\dots,n_{sc}\}$ in the parameter space are chosen in a special way. For each ξ, we evaluate the nonlinear equation to get the direct solutions $\{\psi_h^{(k)}\}$. An interpolation (surrogate) is then constructed to mimic the original nonlinear function with $\widehat{\psi_h}(\cdot,\xi)=\sum_k \psi_h^{(k)}(\cdot)L_{\xi^{(k)}}(\xi)$

Reference: V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. Advances in Computational Mathematics, 12:273–288, 2000.



Left to right: Full tensor grid 2d, level 4. Chebyshev sparse grids for 2d and 3d from level 0 to level 4.

Compared to full tensor grid, **Sparse Grid Stochastic Collocation** requires **fewer grids** but with **satisfying accuracy**.

Theorem: Interpolation error bound in L_∞ norm

$$\|f-\mathscr{S}(f)\|_\infty=\mathcal{O}\left(N^{-k}\cdot|\log N|^{(k+2)(d-1)+1}\right)$$

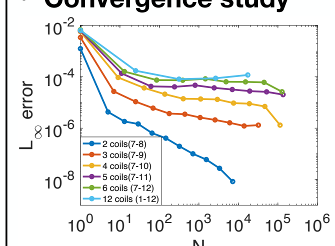
N: # of collocation nodes, d: dimension, k: measure of smoothness of ψ wrt ξ.

Three Approaches

- **Monte Carlo + Surrogate**
- **Multilevel Monte Carlo + Nonlinear solve**
- **Multilevel Monte Carlo + Surrogate**

Monte Carlo + Surrogate

- **Convergence study**



- Fewer coils lead to faster error decay.
- Most errors decrease except for perturbing currents in all 12 coils. Reason: as d increases, need large N to observe error decay.

- **Efficiency study**

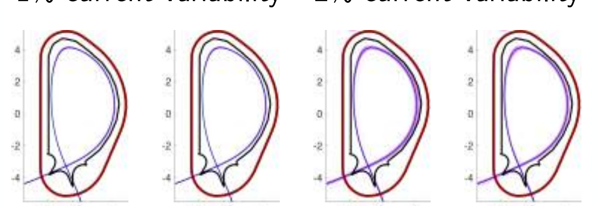
	Level 2	Level 3
1% Noise	Factor of 33	Factor of 9
2% Noise	Factor of 33	Factor of 7

		Surrogate Level 2	Surrogate Level 3	Surrogate Level 4	Direct Solver
1% Noise	# Samples	5100	5000	5000	5200
	Time per Sample (s.)	0.16	0.59	3.03	5.22
	Total Time (s.)	831	2944	15141	27166
2% Noise	# Samples	20000	25300	31100	19100
	Time per Sample (s.)	0.16	0.59	3.04	5.69
	Total Time (s.)	3264	14911	94602	108620
		1 hour	4.1 hours	26 hours	30 hours

- Cost is way low for surrogate than nonlinear solve.

- **Accuracy study**

1% current variability 2% current variability



Surrogate level 3 Direct solution Surrogate level 3 Direct solution

- Surrogate evaluations and direct solutions behave similarly for separatix, x points, and strike points.
- Surrogate mimics the non-linear function with a minor accuracy loss.

	Evaluation of surrogate Sample A (25300 realizations)		Direct solver Sample B (19100 realizations)	
Quantity of interest	Mean	Variance	Mean	Variance
x point	(5.14,-3.29)	(2.89e-04,1.48e-03)	(5.14,-3.29)	(2.56e-04,1.31e-03)
magnetic axis	(6.41,0.61)	(1.53e-03,1.14e-03)	(6.41,0.61)	(1.19e-03,6.42e-04)
strike points	(4.16,-3.71) (5.56,-4.22)	(2.42e-03,2.38e-03) (5.18e-08,3.20e-03)	(4.16,-3.71) (5.56,-4.21)	(8.59e-04,2.09e-03) (5.06e-08,3.13e-03)
inverse aspect ratio	0.32	5.53e-06	0.32	4.85e-06
elongation	1.86	3.00e-04	1.86	3.12e-04
upper triangularity	0.43	2.87e-04	0.43	2.65e-04
lower triangularity	0.53	2.37e-04	0.53	2.15e-04

- Means and variances are comparable for both methods.
- Surrogate requires more samples than nonlinear solve.

Multilevel Monte Carlo + Nonlinear solve

- **Monte Carlo**
  - Estimator: $A_{MC}:=\mathbb{E}(u_h)=\frac{1}{N}\sum_{i=1}^{N}u_h^{(i)}$
  - Relative MSE: $\bar{\mathcal{E}}^2_{A_{MC}}=\underbrace{\frac{\|\mathbb{E}(u)-\mathbb{E}(u_h)\|_Z^2}{\|\mathbb{E}(u)\|^2_{L^2(W,Z)}}}_{\text{discretization error } \bar{\mathcal{E}}_{\text{Bias}}}+\underbrace{\frac{\mathbb{V}(u_h)}{N\|\mathbb{E}(u)\|^2_{L^2(W,Z)}}}_{\text{statistical error } \bar{\mathcal{E}}_{\text{Stat}}}=\bar{\mathcal{E}}^2_{\text{Bias}}(h)+\bar{\mathcal{E}}^2_{\text{Stat}}(N)$
  - Error splitting: $\begin{cases}\bar{\mathcal{E}}^2_{\text{DisErr}}\le(1-\theta)\epsilon^2 & \to \quad M=M(\epsilon),\\ \bar{\mathcal{E}}^2_{\text{Stat}}\le\theta\epsilon^2 & \to \quad N=N(\epsilon).\end{cases}$
  - Sampling cost: $C_{MC}=\mathcal{O}\left(\epsilon^{-2-\gamma/\alpha}\right)$
- **Multilevel Monte Carlo**
  - Estimator: $A_{MLMC}:=\mathbb{E}(u_L)=\sum_{\ell=0}^{L}\mathbb{E}(Y_\ell)=\frac{1}{N_0}\sum_{i=1}^{N_0}u_0^{(i)}+\sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{i=1}^{N_\ell}\left(u_\ell^{(i)}-u_{\ell-1}^{(i)}\right)$
  - Relative MSE: $\bar{\mathcal{E}}^2_{A_{MLMC}}=\underbrace{\frac{\|\mathbb{E}(u)-\mathbb{E}(u_L)\|_Z^2}{\|\mathbb{E}(u)\|^2_{L^2(W,Z)}}}_{\text{discretization error } \bar{\mathcal{E}}_{\text{Bias}}}+\underbrace{\sum_{\ell=0}^{L}\frac{\mathbb{V}(u_\ell-u_{\ell-1})}{N_\ell\|\mathbb{E}(u)\|^2_{L^2(W,Z)}}}_{\text{statistical error } \bar{\mathcal{E}}_{\text{Stat}}}=\bar{\mathcal{E}}^2_{\text{Bias}}(h_L)+\bar{\mathcal{E}}^2_{\text{Stat}}(N_\ell,L)$
  - Error splitting: $\begin{cases}\bar{\mathcal{E}}^2_{\text{DisErr}}\le(1-\theta)\epsilon^2 & \to \quad M_L,\\ \bar{\mathcal{E}}^2_{\text{Stat}}\le\theta\epsilon^2 & \to \quad N_\ell.\end{cases}$
  - Sampling cost:

Theorem

Suppose there exist positive constants α,β,γ such that $\alpha\ge\frac{1}{2}\min(\beta,\gamma)$,

(i) $\|\mathbb{E}(u-u_\ell)\|_Z=\mathcal{O}\left(M_\ell^{-\alpha}\right)$,
(ii) $\mathbb{V}(Y_\ell)=\mathcal{O}\left(M_\ell^{-\beta}\right)$,
(iii) $C_\ell=\mathcal{O}\left(M_\ell^{\gamma}\right)$.

Then for any $\epsilon<1$ small enough, there exists level L and sample size N_ℓ for which the multilevel estimator A_{MLMC} has a relative MSE with

$$\frac{\|\mathbb{E}(u)-A_{MLMC}(u_L)\|^2_{L^2(W,Z)}}{\|\mathbb{E}(u)\|^2_{L^2(W,Z)}}<\epsilon^2,$$

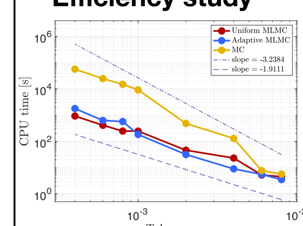
and the total computation cost C_ϵ with bound

$$C_{MLMC}=\begin{cases}\mathcal{O}\left(\epsilon^{-2}\right) & \beta>\gamma,\\ \mathcal{O}\left(\epsilon^{-2}(\log\epsilon)^2\right) & \beta=\gamma,\\ \mathcal{O}\left(\epsilon^{-2-\frac{(\gamma-\beta)}{\alpha}}\right) & 0<\beta<\gamma.\end{cases}$$

- **Adaptive grids—Error estimator**

$$\eta_{K,\ell}(\omega):=h_K^2\left\|\nabla\cdot\left(\frac{1}{\mu r}\nabla u_\ell(\omega)\right)-f(u_\ell(\omega))\right\|_K+h_K^{3/2}\left\|\left[\left[\frac{1}{\mu r}\nabla u_\ell(\omega)\cdot\boldsymbol{n}\right]\right]\right\|_{\partial K\setminus\partial\Omega}$$

- **Efficiency study**



ϵ	MC-FE Time	Uniform MLMC-FE Time	Uniform MLMC-FE Speedup	Adaptive MLMC-FE Time	Adaptive MLMC-FE Speedup
8×10^{-3}	5.67e+00	4.52e+00	1.3	3.50e+00	1.6
6×10^{-3}	7.69e+00	5.25e+00	1.5	5.51e+00	1.4
4×10^{-3}	1.30e+02	2.32e+01	5.6	9.01e+00	14.4
2×10^{-3}	4.83e+02	4.62e+01	10.5	3.16e+01	15.3
10^{-3}	9.22e+03	2.47e+02	37.3	1.84e+02	50.1
8×10^{-4}	1.50e+04	2.48e+02	60.5	5.73e+02	26.2
6×10^{-4}	2.48e+04	4.13e+02	60.0	6.30e+02	39.4
4×10^{-4}	5.68e+04	9.29e+02	61.1	1.79e+03	31.7
2×10^{-4}	9.62e+05*	4.21e+03	228.5	1.21e+04	79.6

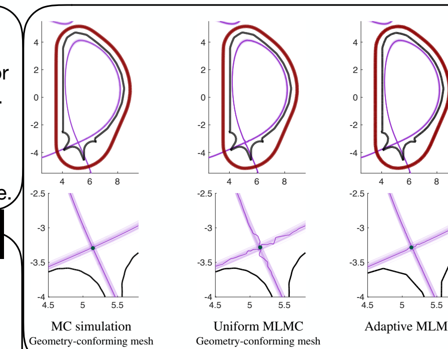
- For efficiency, uniform MLMC> Adaptive MLMC>MC.

- **Accuracy study**

	MC-FE	Uniform MLMC-FE	Adaptive MLMC-FE
x point	(5.14,-3.29)	(5.14,-3.29)	(5.14,-3.28)
magnetic axis	(6.41,0.61)	(6.44,0.56)	(6.46,0.54)
strike points	(4.16,-3.71) (5.56,-4.22)	(4.16,-3.71) (5.56,-4.22)	(4.16,-3.71) (5.56,-4.21)
inverse aspect ratio	0.32	0.32	0.32
elongation	1.86	1.87	1.86
upper triangularity	0.43	0.43	0.43
lower triangularity	0.53	0.53	0.53

- Means and variances are comparable for three approaches.

Acknowledgment: Work supported by the US Department of Energy and the US National Science Foundation.



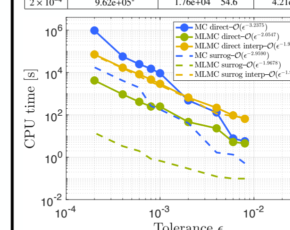
MC simulation Geometry-conforming mesh Uniform MLMC Geometry-conforming mesh Adaptive MLMC

- Uniform MLMC-Large distortion around x-point due to extrapolation on geometry-conforming mesh.
- Adaptive MLMC-better representation of geometric quantities.

Multilevel Monte Carlo + Surrogate

- **Three types of surrogate**
  - Single-level sparse grid + single-level spatial grid
  - Single-level sparse grid + multilevel spatial grid
  - Multilevel in both sparse & spatial grids
- **Two sampling methods**
  - Monte Carlo
  - Multilevel Monte Carlo
- **Efficiency study**

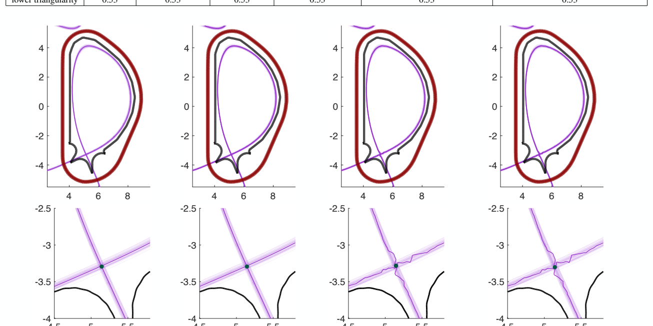
ϵ	MC-FE Direct solver Time	MC-FE Surrogate Time	MC-FE Surrogate Speedup	MLMC-FE Direct solver Time	MLMC-FE Direct solver Speedup	MLMC-FE Surrogate Time	MLMC-FE Surrogate Speedup	MLMC-FE DS interp Time	MLMC-FE DS interp Speedup	MLMC-FE Surrogate interp Time	MLMC-FE Surrogate interp Speedup
8×10^{-3}	5.67e+00	5.16e-01	11.0	4.52e+00	1.3	9.78e-02	5.8e+01	6.49e+01	0.08	7.17e+01	0.08
6×10^{-3}	7.69e+00	1.33e+00	5.8	5.25e+00	1.5	9.81e-02	7.8e+01	9.49e+01	0.08	8.84e+01	0.09
4×10^{-3}	1.30e+02	1.67e+00	78.1	2.32e+01	5.6	1.23e-01	1.1e+03	2.16e+02	0.6	1.14e+02	1.1
2×10^{-3}	4.83e+02	4.07e+01	11.9	4.62e+01	10.5	2.91e-01	1.7e+03	6.56e+02	0.7	6.16e+02	0.8
10^{-3}	9.22e+03	1.80e+02	51.3	2.47e+02	37.3	6.77e-01	1.4e+04	2.92e+03	3.2	2.46e+03	3.8
8×10^{-4}	1.50e+04	2.60e+02	57.8	2.48e+02	60.5	8.36e-01	1.8e+04	4.63e+03	3.3	3.53e+03	4.3
6×10^{-4}	2.48e+04	1.93e+03	12.9	4.13e+02	60.0	1.97e+00	1.3e+04	8.26e+03	3.0	6.63e+03	3.8
4×10^{-4}	5.68e+04	4.12e+03	13.8	9.29e+02	61.1	3.33e+00	1.7e+04	1.69e+04	3.4	1.62e+04	3.5
2×10^{-4}	9.62e+05*	1.76e+04	54.6	4.21e+03	228.5	1.43e+01	6.7e+04	7.13e+04	13.5	6.38e+04	15.1



- For efficiency, MLMC+surrogate>MLMC+nonlinear solve>MC+surrogate>MC+nonlinear solve, when tolerance is small.
- For accuracy, we interpolate to a common grid that encloses all coarser meshes, but this incurs a significant interpolation cost.

- **Accuracy study**

	MC-FE DS	MC-FE Surrogate	MLMC-FE DS	MLMC-FE Surrogate	MLMC-FE DS(Interp 2 common)	MLMC-FE Surrogate(Interp 2 common)
x point	(5.14,-3.29)	(5.14,-3.29)	(5.14,-3.29)	(5.14, -3.30)	(5.14,-3.29)	(5.14,-3.29)
magnetic axis	(6.41,0.61)	(6.41,0.61)	(6.44,0.56)	(6.44, 0.56)	(6.41,0.61)	(6.41,0.61)
strike points	(4.16,-3.71) (5.56,-4.22)	(4.16,-3.71) (5.56,-4.22)	(4.16,-3.71) (5.56,-4.22)	(4.16, -3.71) (5.56, -4.22)	(4.16,-3.71) (5.56,-4.22)	(4.16,-3.71) (5.56,-4.22)
inverse aspect ratio	0.32	0.32	0.32	0.32	0.32	0.32
elongation	1.86	1.86	1.87	1.87	1.86	1.86
upper triangularity	0.43	0.43	0.43	0.46	0.43	0.43
lower triangularity	0.53	0.53	0.53	0.53	0.53	0.53



MC-FE + Direct Solve MC-FE + Surrogate MLMC-FE + Direct Solve MLMC-FE + Surrogate

- Surrogate-based sampling has similar behavior as the sampling+nonlinear solve.

Reference

- Surrogate approximation of the Grad-Shafranov free boundary problem via stochastic collocation on sparse grids, with Howard Elman and Tonatiuh Sánchez-Vizuet. (Journal of Computational Physics, 448 (2022), paper No. 110699.)
- Multilevel Monte Carlo methods for the Grad-Shafranov free boundary problem, with Howard Elman and Tonatiuh Sánchez-Vizuet. (Computer Physics Communications, 298 (2024), paper No. 109099.)