Application of an enhanced control-variate scheme

for particle-in-cell simulations

with and without collisions

R. Kleiber, R. Hatzky, A. Könies
K. Kauffmann, P. Helander, M. Borchardt

Content

- Introduction
- Particle discretisation using δf method
- Including collisions (two-weight scheme)
- Examples for usage of an enhanced control variate
  - Improving the two-weight scheme
  - Number conservation in δf PIC
- Adaptive control variate for electromagnetic calculations
- Electromagnetic calculation using the EUTERPE code

Introduction

- Micro instabilities and turbulence in magnetised plasmas
  $\Rightarrow$ gyrokinetic theory (5D space $\{\vec{R}, v_{\|}, \mu\}$)

- Special interest in stellarators: - 3D geometry
  - Neoclassical effects
  - Radial electric field

- Numerical simulation by PIC methods: EUTERPE code

General formulation

- **(Gyro) kinetic equation** $(z = \{\vec{R}, v_\parallel, \mu\})$**:**

$$\frac{\mathrm{d}}{\mathrm{d}t} f(z) = \frac{\partial f}{\partial t} + v \frac{\partial f}{\partial z} = \begin{cases} 0 \\ C(f,f) \end{cases}$$

$v = v^0 + v^1(\Phi)$, v^0 **only determined by magnetic configuration**
$C(f,f)$**: collision term**

- **Field equation:**

$$\mathcal{L}(\Phi) = \int A(z - z') f(z') dz' = E(A) \qquad (\mathcal{L} \text{ : linear operator})$$

- **Particle-in-cell: density binning:** $n(z) = E(\Lambda)$
Λ**: binning function (e.g.** $\frac{1}{2\Delta} I_{[z-\Delta, z+\Delta]}$**, or B-spline)**

Classical δf method

- **Problem with full-f:** for small variations on a large background (e.g. for linearised equations) most particles are wasted for representation of the background $\Rightarrow$ **huge noise**
- Splitting of f into background f_0 and fluctuation δf [Kotschenreuther 88]:

$$f(z,t) = f_0(z) + \delta f(z,t)$$

$$\frac{\mathrm{d}f}{\mathrm{d}t} = 0 \quad \Rightarrow \quad \frac{\partial}{\partial t}\delta f + (v^0 + v^1)\frac{\partial}{\partial z}\delta f = -(v^0 + v^1)\frac{\partial f_0}{\partial z} =: S$$

$$\Rightarrow \text{characteristics: } \dot{z} = v^0 + v^1, \quad \dot{\delta f} = S$$

- Introduce N marker with distribution function g evolving according to $\frac{\mathrm{d}g}{\mathrm{d}t} = 0$
- For each marker p ($w = \frac{\delta f}{g}$):

$$\Rightarrow \quad \dot{z}_p = v^0 + v^1, \quad \dot{w}_p = \frac{1}{g(z_p)} S(z_p) \quad \text{with} \quad g(z_p(t)) = g(z_p(0))$$

δf method with control variate

δf can be regarded as a simple control variate (CV) method [Aydemir 94]

$$E(A) = \int A f dz' = \int A(f - f_0) dz' + \int A f_0 dz'$$

$$E(A) = \int A \frac{\delta f}{g} g \, dz' + A_0$$

$$\delta A = E(A) - A_0 = \frac{1}{N} \sum_{p=1}^{N} w_p A(z_p) + \epsilon \qquad \epsilon = \frac{\sqrt{V(A \frac{\delta f}{g})}}{\sqrt{N}}$$

$\Rightarrow$ - Distribution of (numerical) markers decoupled from distribution of (physical) particles
- CV: no equation for $\dot{w}_p$ needed
- Order of magnitude better statistics than full-f, if $\delta f \ll f$
- GK PIC simulations in linear regime only possible with δf method

Including collisions

Collisional simulations

- Some instabilities (e.g. trapped electron mode) react strongly on collisions
- Zonal flows damped by collisions $\rightarrow$ back-reaction on turbulence
- Stellarators have strong collisional (neoclassical) transport
  This may interact with instabilities and turbulence
- Collisions can provide numerical dissipation needed for nonlinear simulations

---

Basic kinetic equation:

$$\frac{\mathrm{d}}{\mathrm{d}t} f(z) = \frac{\partial f}{\partial t} + v \frac{\partial f}{\partial z} = C(f, f)$$

$C(f, f)$ linearised collision operator where $C(f, f_0)$ has Fokker-Planck form

Basic theorem

The Kolmogorov forward (Fokker-Planck) equation

$$\frac{\partial}{\partial t} f = -\frac{\partial}{\partial z^i}\left[\mu^i f\right] + \frac{1}{2}\frac{\partial^2}{\partial z^i \partial z^j}\left[(DD^T)^{ij} f\right]$$

(μ^i, D^{ij}: drift coefficients, diffusion matrix)

$$\Updownarrow$$

f is given by the distribution function of the stochastic (Itoh) process

$$\mathrm{d}z^i = \mu^i \mathrm{d}t + D^{ij}\mathrm{d}W^j$$

(W^i: Wiener process)

δf method

Equation to solve:

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial z} = C(f, f)$$

use $f = f_0 + \delta f$

$$\Rightarrow \quad \mathcal{P}\delta f = C(f_0, f_0) + C(f_0, \delta f) - \mathcal{S}$$

with

$$\mathcal{P}\delta f = \frac{\partial \delta f}{\partial t} + \frac{\partial (v\delta f)}{\partial z} - C(\delta f, f_0) \quad \text{and} \quad \mathcal{S} = v\frac{\partial f_0}{\partial z}$$

- $C(\delta f, f_0)$ **has Fokker-Planck form**
- $C(\delta f, \delta f)$ **neglected,** $C(f_0, \delta f)$ **usually replaced by model [Xu, Rosenbluth 90]**

$\Rightarrow$ **Equation has the form of an inhomogeneous FP equation**

Extended phase space formalism

- Extend phase space by one [Chen, White 97] or two [Brunner, Valeo, Krommes 99, Wang et al. 99] weight dimensions
- Postulate extended distribution function $F(z, w_1, w_2)$ evolving according to:

$$\mathcal{P}F + \frac{\partial}{\partial w_1}(A_1 F) + \frac{\partial}{\partial w_2}(A_2 F) = 0$$

- Basic theorem applies $\Rightarrow$ trajectory equations

$$\frac{\mathrm{d}z}{\mathrm{d}t} = v + \mu + D\frac{\mathrm{d}W}{\mathrm{d}t}$$
$$\frac{\mathrm{d}w_{(1,2)}}{\mathrm{d}t} = A_{(1,2)}$$

- Determine $A_{(1,2)}$ from consistency with δf equation

Extended phase space formalism

- Moments of F:

$$g = \int F \mathrm{d}w_1 \mathrm{d}w_2, \quad u = \int w_1 F \mathrm{d}w_1 \mathrm{d}w_2, \quad h = \int w_2 F \mathrm{d}w_1 \mathrm{d}w_2$$

- Evolution equations for moments

$$\mathcal{P}g = 0, \quad \mathcal{P}u - \int A_1 F \mathrm{d}w_1 \mathrm{d}w_2 = 0, \quad \mathcal{P}h - \int A_2 F \mathrm{d}w_1 \mathrm{d}w_2 = 0$$

- Identification of u with δf
- First ansatz: $A_1 = \frac{1}{g}\left[C(f_0, f_0) + C(f_0, \delta f) - \mathcal{S}\right]$ and $A_2 = 0$

One weight scheme: $$\frac{\mathrm{d}w_1}{\mathrm{d}t} = \frac{1}{g}\left[C(f_0, f_0) + C(f_0, \delta f) - \mathcal{S}\right]$$

Problem: Contrary to collisionless case the marker distribution function $g(t)$ can not be calculated from $g(0)$
Also g can not be calculated accurately by binning of markers

Extended phase space formalism

Second ansatz:

$$A_1 = \frac{w_2}{h}\left[C(f_0, f_0) + C(f_0, \delta f) - \mathcal{S}\right] \qquad A_2 = -\frac{w_2}{h}\left[C(h, f_0) - \frac{\mathrm{d}f_0}{\mathrm{d}t}\right]$$

$$\Rightarrow \frac{\mathrm{d}h}{\mathrm{d}t} = \frac{\mathrm{d}f_0}{\mathrm{d}t}$$

using $h = f_0 \rightarrow$ **Two-weight scheme**

$$\frac{\mathrm{d}w_1}{\mathrm{d}t} = \left[C(f_0, \delta f) + C(f_0, f_0) - \mathcal{S}\right]\frac{w_2}{f_0}$$

$$\frac{\mathrm{d}w_2}{\mathrm{d}t} = -\left[C(f_0, f_0) - \mathcal{S}\right]\frac{w_2}{f_0}$$

Interpretation: $w_1 = \dfrac{\delta f}{g}, \; w_2 = \dfrac{f_0}{g}$

Behaviour of the two-weight scheme

Ornstein-Uhlenbeck process

Diffusion with drift:

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial v}(\mu v f) + \frac{1}{2}D^2\frac{\partial^2}{\partial v^2}f \qquad \text{with} \quad f(0) = \frac{a}{\sqrt{\pi}}\mathrm{e}^{-a^2v^2}$$

Solution (decaying Maxwellian):

$$f = \frac{1}{\sqrt{\pi}h(t)}e^{-\frac{v^2}{h(t)^2}}, \quad h(t)^2 = \frac{D^2}{\mu}\left[1 - e^{-2\mu t}\left(1 - \frac{\mu}{a^2D^2}\right)\right]$$

Solution tends to a stationary state for $t \to \infty$:

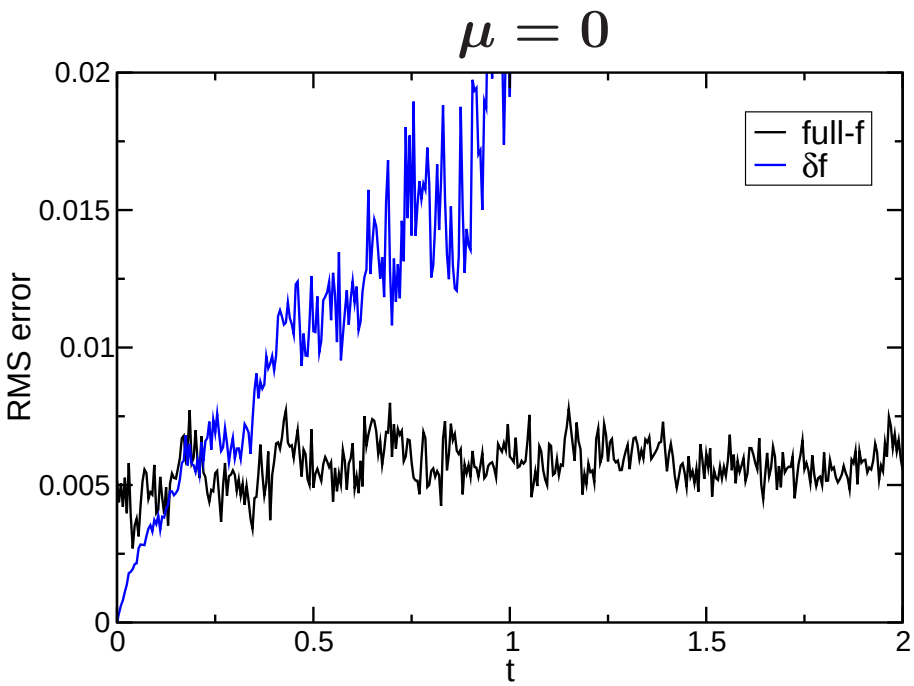
$$f_s = \sqrt{\frac{\mu}{D^2\pi}}\mathrm{e}^{-\frac{\mu}{D^2}v^2}$$

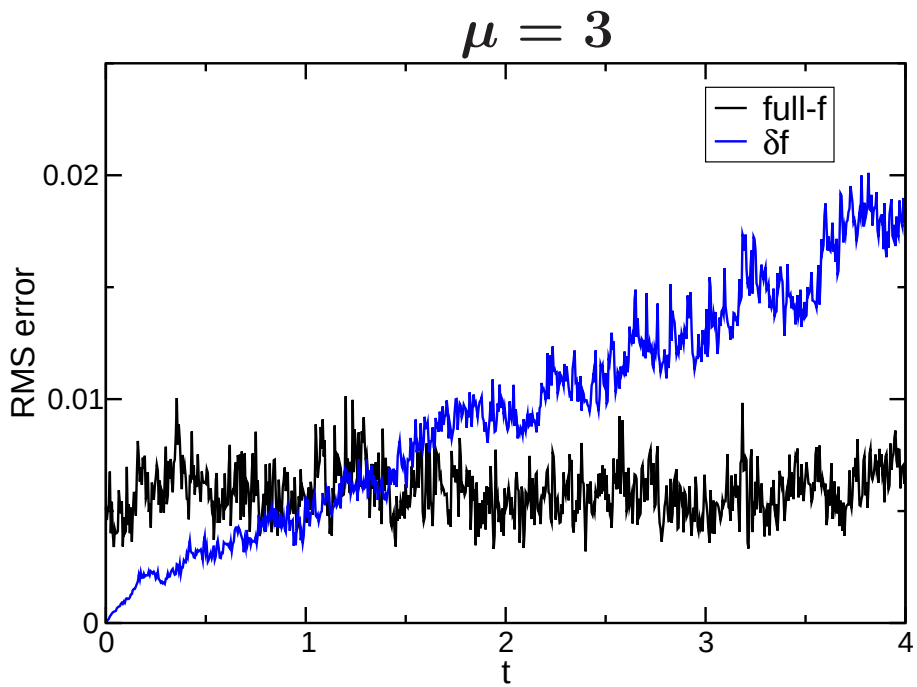
No time evolution if $\mu = a^2D^2$

Ornstein-Uhlenbeck process, numerical solution

Error between analytical and numerical solution

Choose $f_0 = f(t=0)$ for δf splitting ($a = 2, D = 1$)





⇒ - δf behaviour: error smaller if f_0 is closer to solution

- Error grows with time, even if stationary solution is reached

Ornstein-Uhlenbeck process, numerical solution

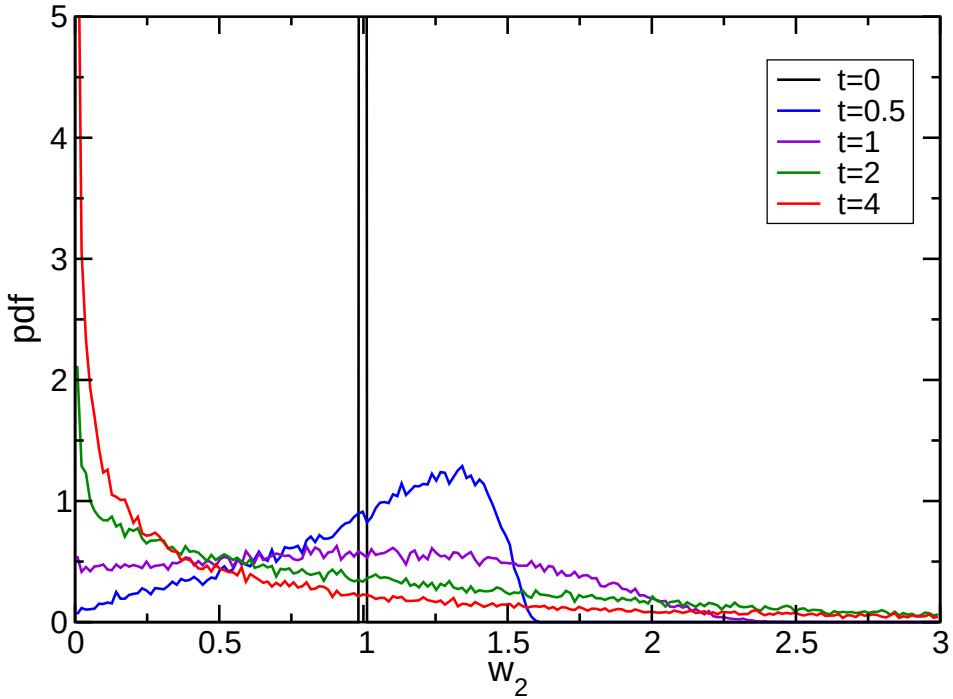
$\Rightarrow$ **Error of the two-weight scheme grows with time (weight spreading)**

Evolution equation for w_2:

$$\dot{w}_{2,p} = G(v_p) w_{2,p}$$
$$\dot{v}_p = -\mu v_p + D \frac{\mathrm{d}W}{\mathrm{d}t}$$

$$G = (\mu - a^2 D^2)(2a^2 v_p^2 - 1)$$

$E(G) \leq 0$ **for large** t
$\Rightarrow$ $w_{2,p}$ **evolves towards fixed point at zero**



By definition of w_2: $E(w_2)_v = f_0(v)$
$\Rightarrow$ **Variance must increase with time**

Methods for treating collisions

Avoid weight spreading:

- Add artificial terms ($\dot{w}_p = \eta(w_p - W)$) to the weight evolution equations $\rightarrow$ restore the weights to their mean value W in a grid cell [Brunner 99]

- Coarse graining: set weights in a grid cell equal to their mean value [Chen 07]

Other methods to include collisions:

- Add Krook operator ($\dot{w}_p = -\nu w_p$) [McMillan 08]

- Solve collision operator directly on a grid (e.g. by finite difference method) [Broemstrup 08]

Using an enhanced control variate

Noise reduction by variance reduction: control variate (CV)

CV method used in δf approach ($\int f dz = \int (f - f_0) dz + \int f_0 dz$)
is only a special case of the general CV method

Given a stochastic process X, need to calculate $\mathrm{E}[X]$
Error proportional to variance $V(X)$ should be small

- CV: assume process Y with known $\mathrm{E}[Y] = y$
- Construct new process $\tilde{X}(t) = X - \alpha(Y - y)$
  $\Rightarrow$ expectation value unchanged: $\mathrm{E}[\tilde{X}] = \mathrm{E}[X]$
- Use freedom in α to minimise $\mathrm{V}[\tilde{X}]$

Noise reduction by variance reduction: control variate (CV)

- α minimised by $\alpha_* = \dfrac{\mathrm{Cov}(X,Y)}{\mathrm{V}(Y)}$
- New process $\tilde{X}(t) = X - \alpha_*(Y - y)$ has variance

$$\mathrm{V}(\tilde{X}) = V(X)\left[1 - \mathrm{Corr}(X,Y)^2\right]$$

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{V}(X)\mathrm{V}(Y)}} \qquad -1 \leq \mathrm{Corr}(X,Y) \leq 1$$

$\Rightarrow$ **Variance reduced if X and Y are correlated**

- Classical δf is a limiting case:
  Optimal α could be calculated but if $\delta f \ll f_0 \Rightarrow \alpha_* \approx 1$

Application to collisions

- Binned distribution function needed : $\hat{f}_i = E(\Lambda_i \frac{f}{g})$ (Λ_i: binning function)
- Choose $\Lambda_i \frac{f_0}{g}$ as CV
- New process with free parameter α_i in each bin:

$$\tilde{X} = \Lambda_i \frac{f}{g} - \alpha_i \left(\Lambda_i \frac{f_0}{g} - F_{0,i} \right) \qquad F_{0,i} = \int \Lambda_i f_0 dv = f_0(v_i)$$

$$\Rightarrow \quad \hat{f}_i = E(\tilde{X}) \approx \frac{1}{N} \sum_p \Lambda_i \left(c_p - \alpha_i w_{2,p} \right) + \alpha_i F_{0,i}$$

with $c_p = w_{1,p} + w_{2,p}$

$w_{1,p}, w_{2,p}$ evolving according to the two-weight equations

Application to collisions

$$\Rightarrow \quad \alpha_i = \frac{\mathrm{Cov}(c, w_2)}{\mathrm{V}(w_2)} \quad \text{in each bin}$$

Calculation of Cov and V by standard estimators

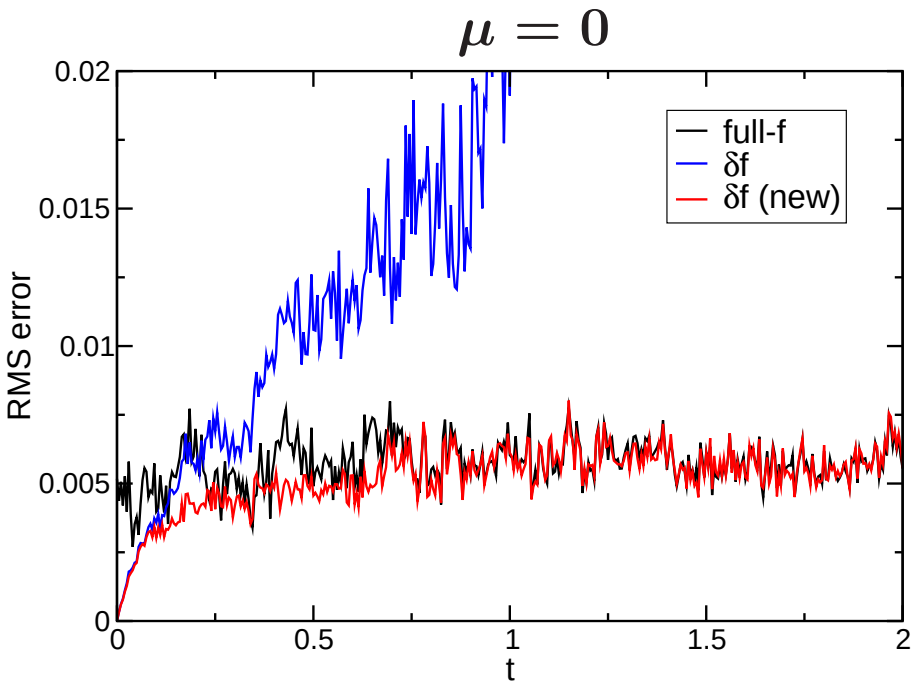
- α_i interpolates (in each bin) between δf ($\alpha_i = 1$) and full-f ($\alpha_i = 0$) scheme

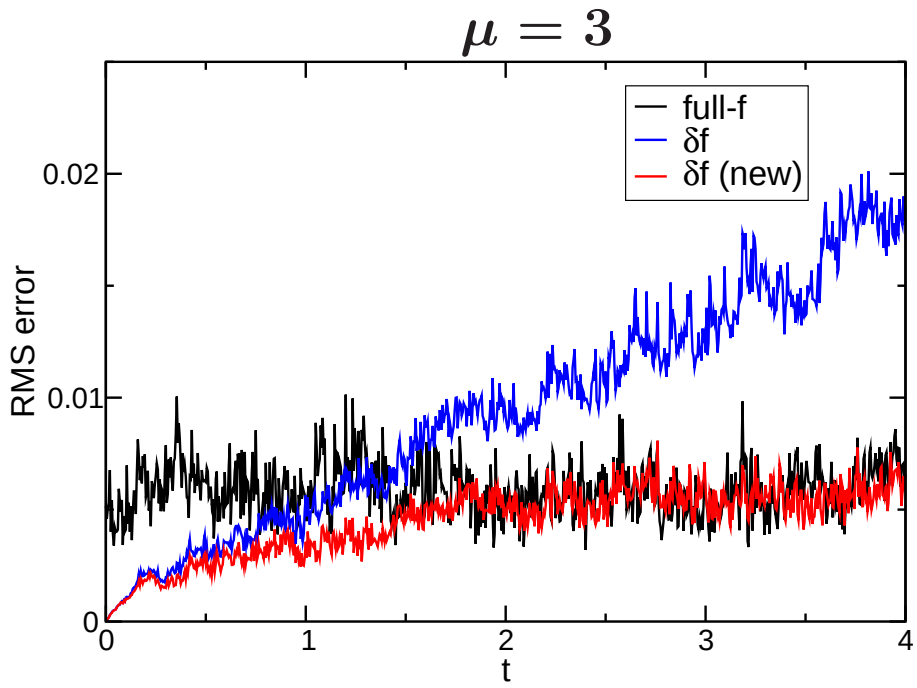
- Statistics may be too bad for good estimate of α_i
  e.g. not enough particles in a bin
  $\rightarrow$ Necessary to introduce cutoff function for α_i guaranteeing $\alpha_i \in [0, 1]$

- Dynamics of the system is not affected, only the calculation of expectation values is improved

- More complex systems: expectation values couple back to the dynamics

Ornstein-Uhlenbeck process, numerical solution





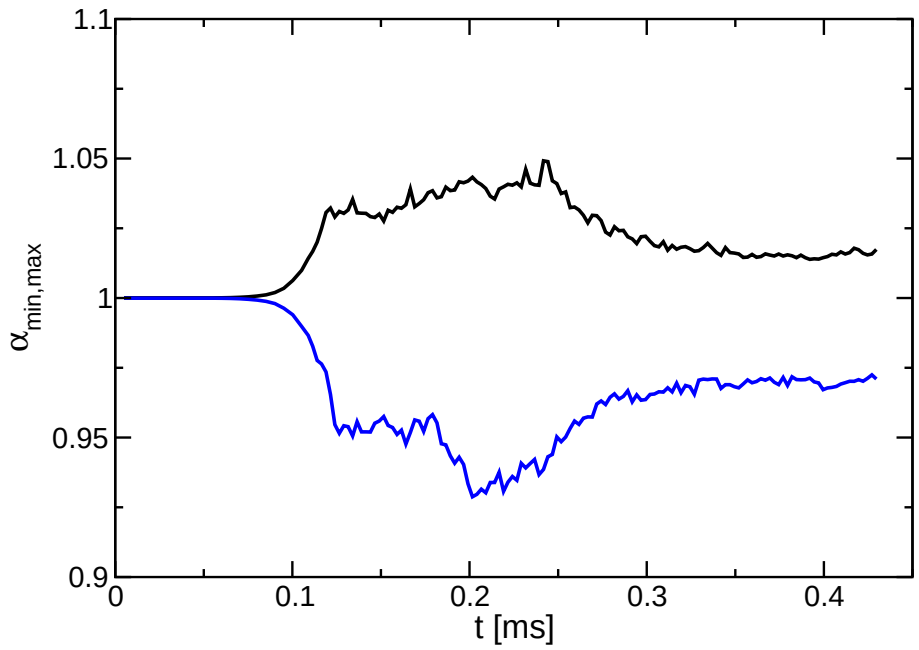
- **Smaller times: Small error of δf scheme**
- **Larger times: Error bounded by that of the full-f scheme**

Use of CV for diagnostics

Example: Collisionless nonlinear gyrokinetic turbulence in a cylinder

Quantify quality of the CV by

$$\alpha_i = \frac{\mathrm{Cov}(\Lambda_i \frac{f}{g}, \Lambda_i \frac{f_0}{g})}{\mathrm{V}(\Lambda_i \frac{f_0}{g})}$$



Linear phase: $\alpha_i \approx 1$

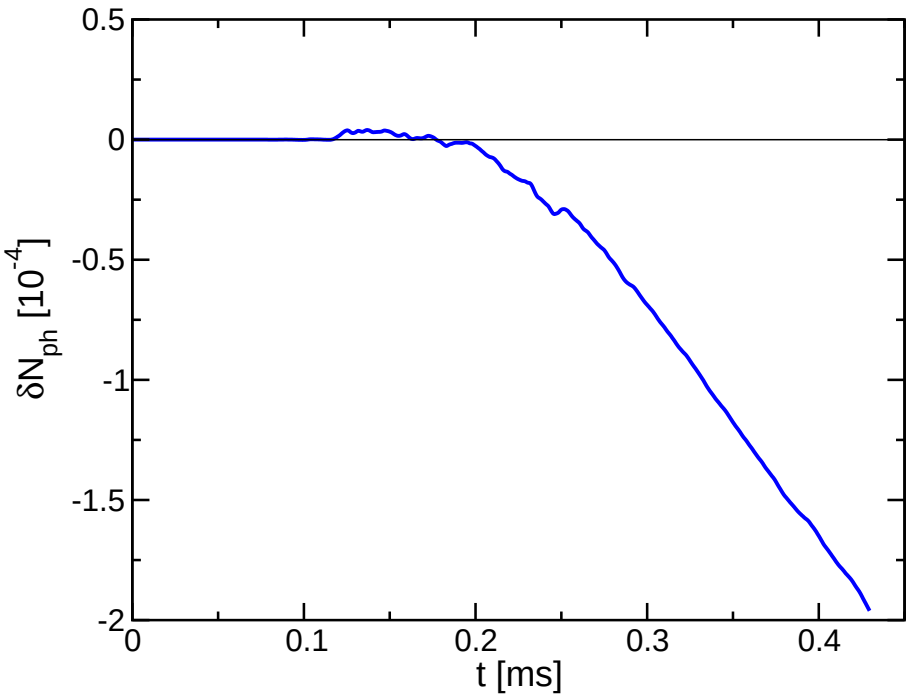
Nonlinear phase: α_i **starts deviating from one but CV is still sufficiently good**
$\rightarrow$ **Back coupling to density calculation brings no advantage**

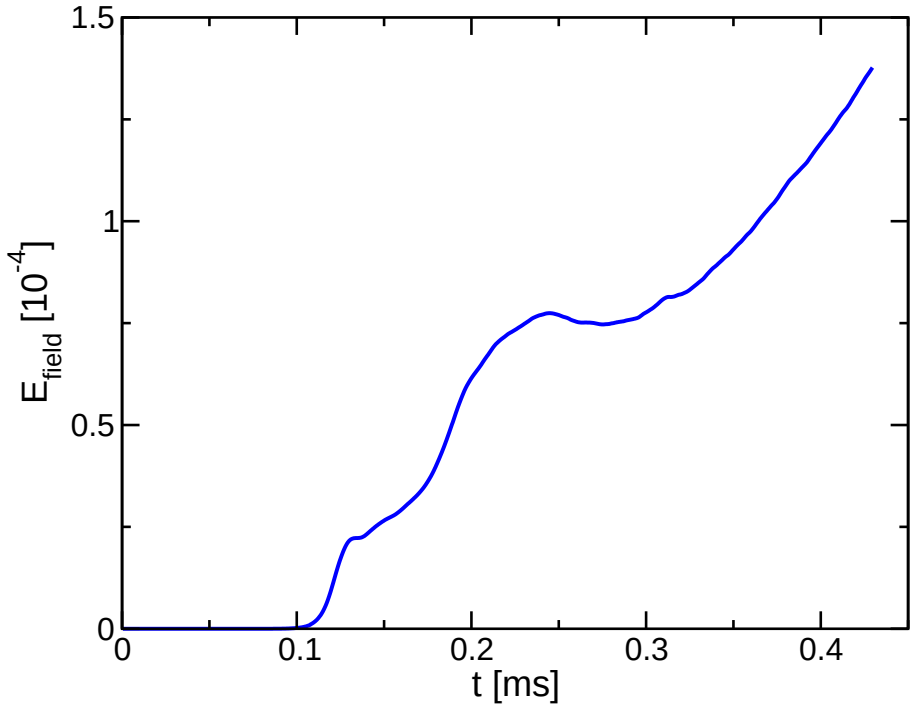
Number conservation in δf PIC

Number N_{ph} of physical particles in volume should be conserved:

$$\delta N_{\mathrm{ph}} = E(w) = 0$$

Problems for δf simulations:
- $\delta N_{\mathrm{ph}} \neq 0$
- Growth of field energy ($\int n \Phi dV$) at later times





CV as a constraint: Number conservation in δf PIC

- **Interpretation of CV as a constraint [Szechtman, Glynn 2001]**
- **Construct CV which ensures number conservation:**
  $E(w) = 0$ **is required** $\rightarrow$ **use** w **as CV**

$$\tilde{X} = \Lambda_i w - \frac{\mathrm{Cov}(\Lambda_i w, w)}{\mathrm{V}(w)} w$$

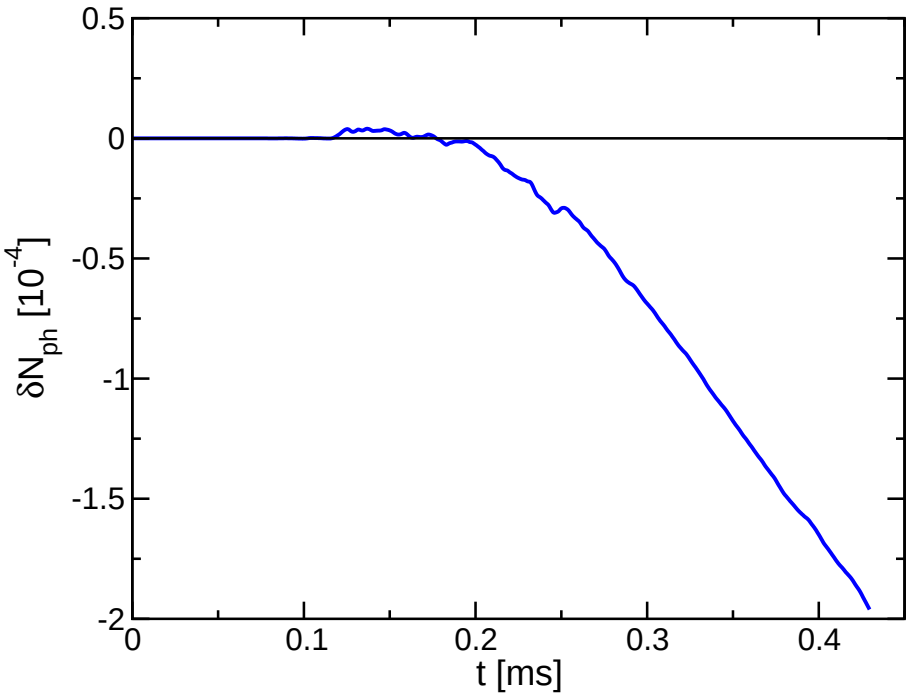
$$E(\tilde{X}) = E\left(\Lambda_i w \left[1 - \frac{w - E(w)}{E(w^2) - E(w)^2} E(w)\right]\right)$$

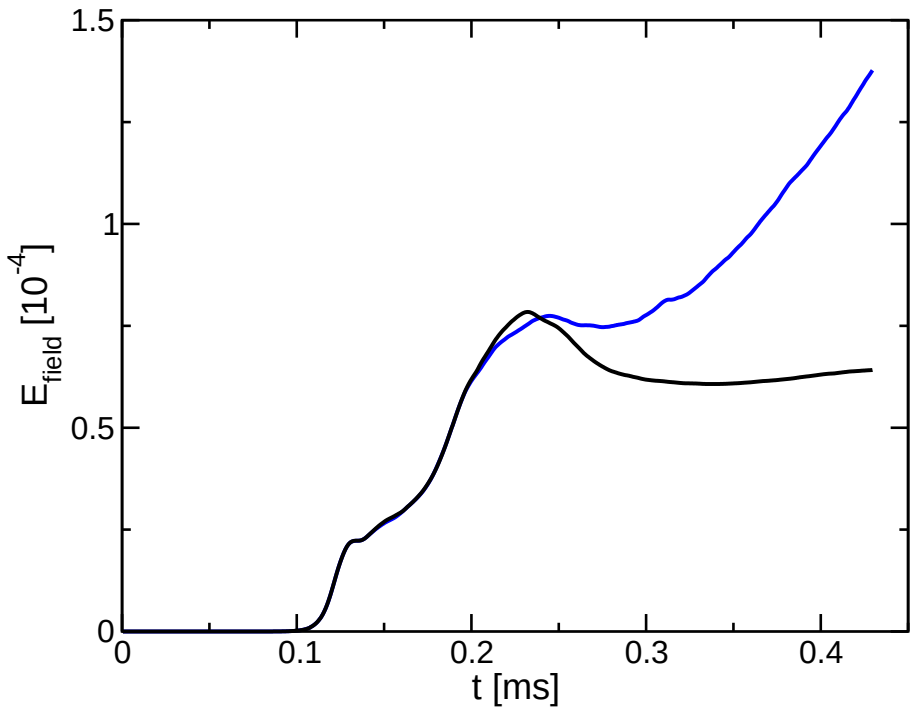
$\Rightarrow$ **Introduce modified weight**

$$\tilde{w} = w \frac{E(w^2) - wE(w)}{E(w^2) - E(w)^2}$$

CV as a constraint: Number conservation in δf PIC

By construction the particle number is conserved: $E(\tilde{w}) = 0$





$\Rightarrow$ **Usage of $\tilde{w}$ also improves behaviour of the field energy**

Adjustable CV method: electromagnetic simulations

- Ampère's law (dimensionless units)

$$-\frac{1}{\beta}\nabla_\perp^2 A_\parallel + \frac{\beta}{\rho_e^2} A_\parallel = j_{\parallel i} + j_{\parallel e}$$

- $\frac{1}{\beta}\nabla_\perp^2 A_\parallel \ll \frac{\beta}{\rho_e^2} A_\parallel$ (especially for MHD modes, since $k_\perp$ is small)
- Different numerical representations: LHS: matrix, RHS: particles
- $j_{\parallel e}$ contains adiabatic part $j_{\parallel ad} = \int \frac{q_e^2}{T} f_0 v_\parallel^2 A_\parallel \; d^3v$
  which cancels $\frac{\beta}{\rho_e^2} A_\parallel$ exactly

$\Rightarrow$ **Cancellation problem:** numerical cancellation of two large terms
$\Rightarrow$ Adjustable CV method [Hatzky, Könies, Mishchenko 07]

Adjustable CV method

- Ampère's law (only electron part): $\mathcal{L}A_\parallel = j_{\parallel e} = q_e \int v_\parallel f_e \, d^3v$
- Use $f_{\mathrm{ad}} = \frac{q_e}{T} f_0 v_\parallel A_\parallel$ as CV (with still unknown $A_\parallel$):

$$\mathcal{L}A_\parallel = q_e \int v_\parallel (f_e - f_{\mathrm{ad}}) \, d^3v + \frac{q_e^2 n}{m_e} A_\parallel = j_{\parallel e} - \hat{\mathcal{S}} A_\parallel + \mathcal{S} A_\parallel$$

- same quantity represented by $\hat{\mathcal{S}}$ (particles) and $\mathcal{S}$ (grid)
- $(\mathcal{L} - \mathcal{S} + \hat{\mathcal{S}}) A_\parallel = j_{\parallel e} \quad \Rightarrow \quad \left[1 - (\mathcal{S} - \hat{\mathcal{S}})\mathcal{L}^{-1}\right] \mathcal{L}A_\parallel = j_{\parallel e}$
- Using Neumann series $\Rightarrow \mathcal{L}A_\parallel = \sum_{i=0}^{\infty} \left[(\mathcal{S} - \hat{\mathcal{S}})\mathcal{L}^{-1}\right]^i j_{\parallel e}$

$\Rightarrow$ Iterative method for determining the CV
Strong reduction in the number of necessary markers

EUTERPE code

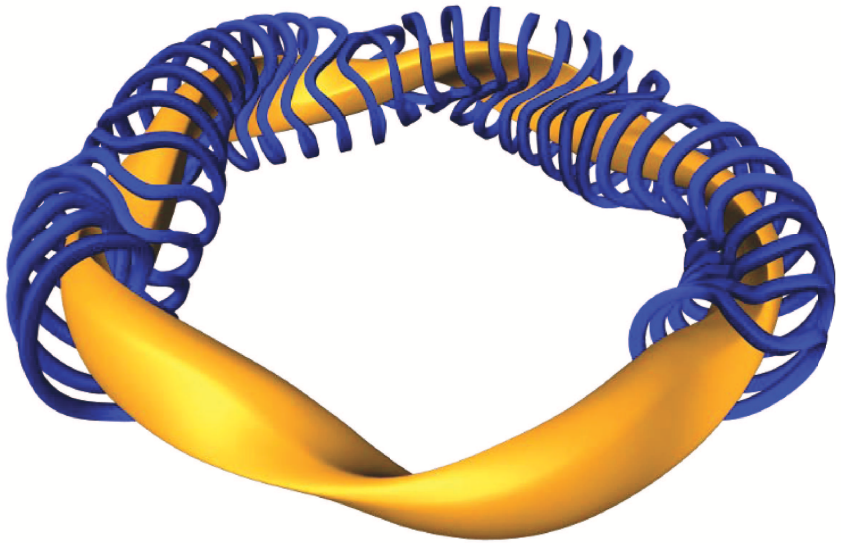
- Simulates the gyrokinetic equation (δf PIC scheme)
- Global simulation domain: full flux-surface, full radius
- 3D stellarator equilibria (from VMEC)
- Multiple kinetic species (e.g. ions, electrons, fast ions/impurities)
- Electromagnetic, adjustable CV method
- Nonlinear terms included
- 4th, 5th order Runge-Kutta with adaptive time step
- Pitch angle collision operator

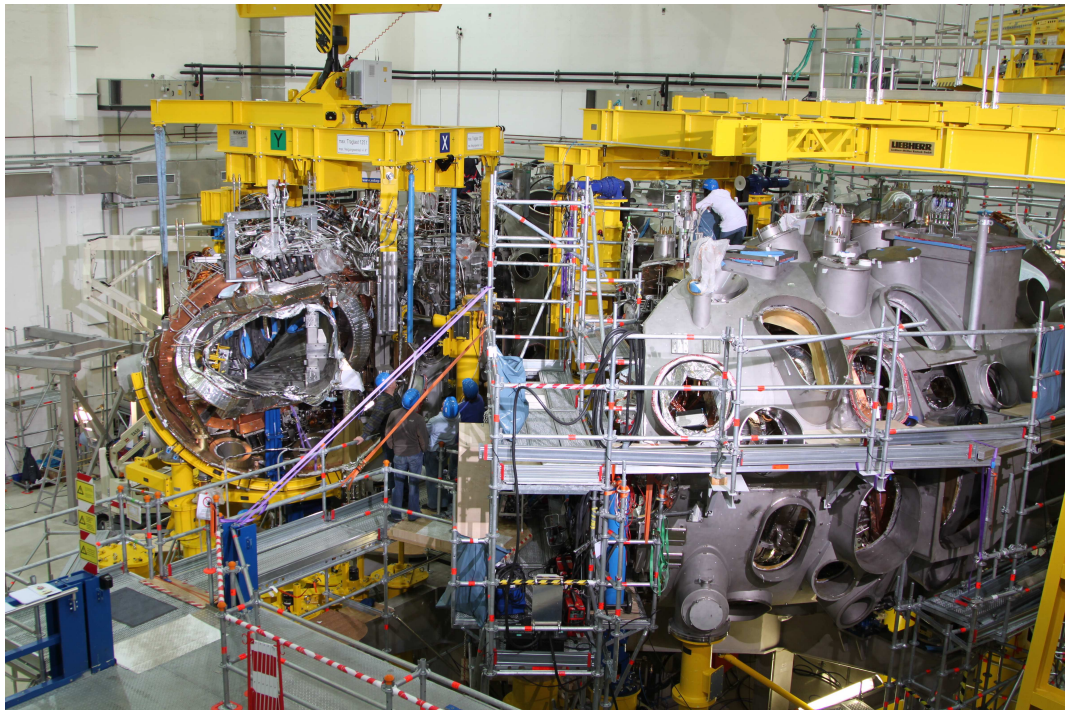
Simulation parameters

Stellarator simulations roughly $O(10)$ more expensive than tokamak simulations

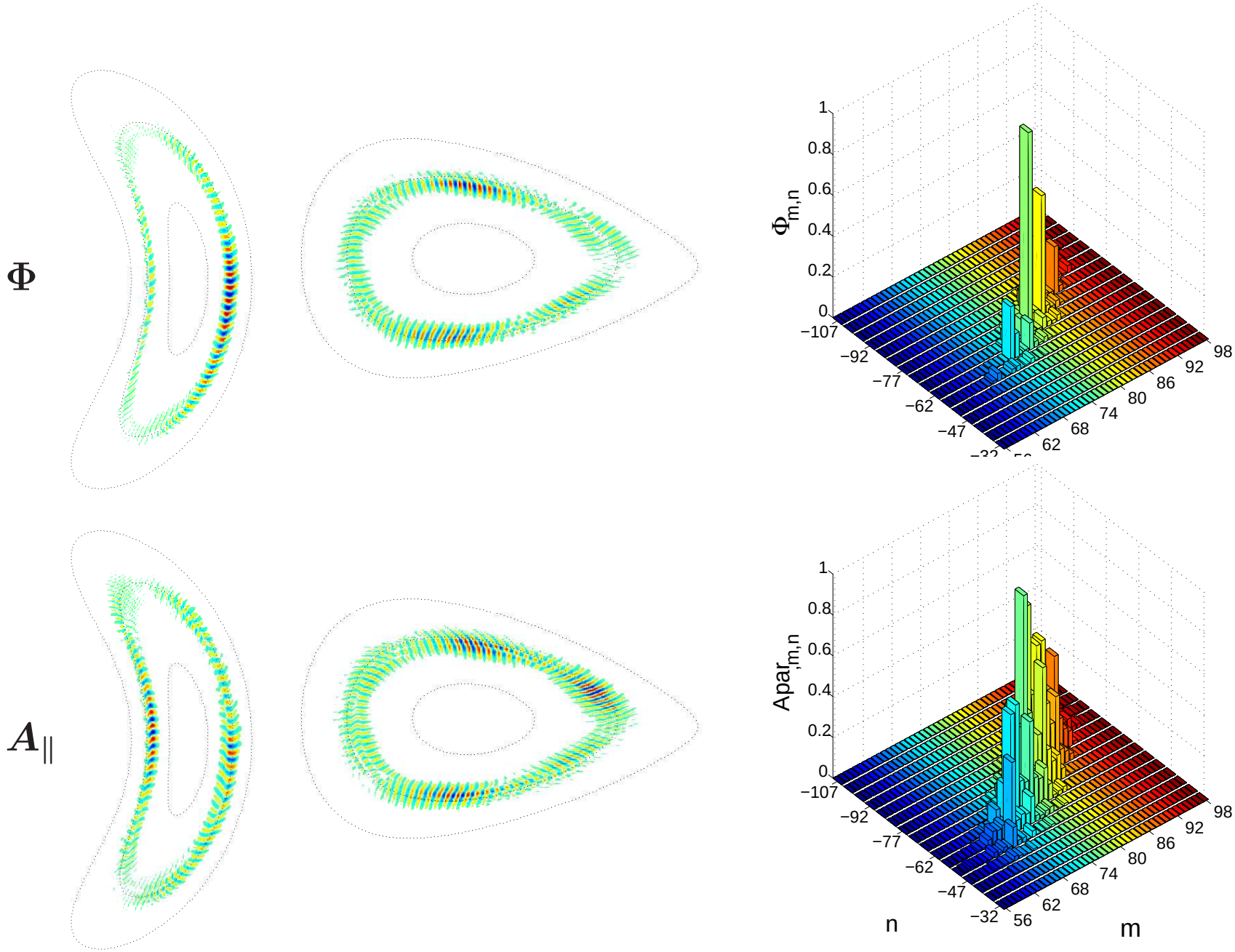
- **Equilibrium: Wendelstein 7-X**
- **Linear simulation of electromagnetic ITG with $\beta = 1\%$**
- **Fully kinetic electrons $\Rightarrow$ small time step**
- **Particle numbers: $N_\mathrm{i} = 32M$, $N_\mathrm{e} = 128M$**
- **Grid size (s, ϑ, φ): $100 \times 128 \times 128$**
- **High mode number: $M = 74$, $N = -62$**
- **Computing time: About 30d on 128 PEs**

Wendelstein 7-X stellarator





Under construction in Greifswald


Φ
$A_{\|}$
$\Phi_{m,n}$
$Apar_{,m,n}$
1
0.8
0.6
0.4
0.2
0
−107
−92
−77
−62
−47
−32
56
62
68
74
80
86
92
98
n
m

Conclusion

- Classical δf uses a control variate in a restricted sense

- Enhanced control variate gives more freedom:
  - Improvement of the collisional two-weight scheme
  - Constrain particle number in δf PIC

- Adjustable control variate:
  First global electromagnetic simulations for stellarators

$\Rightarrow$ - Application to more complex problems with collisions
- Energy conservation in δf PIC