Direct multi-scale coupling of a transport code to gyrokinetic turbulence codes

Michael Barnes

University of Oxford

Culham Centre for Fusion Energy

In collaboration with W. Dorland, G. Hammett, F. Jenko, T. Goerler, and I. Abel

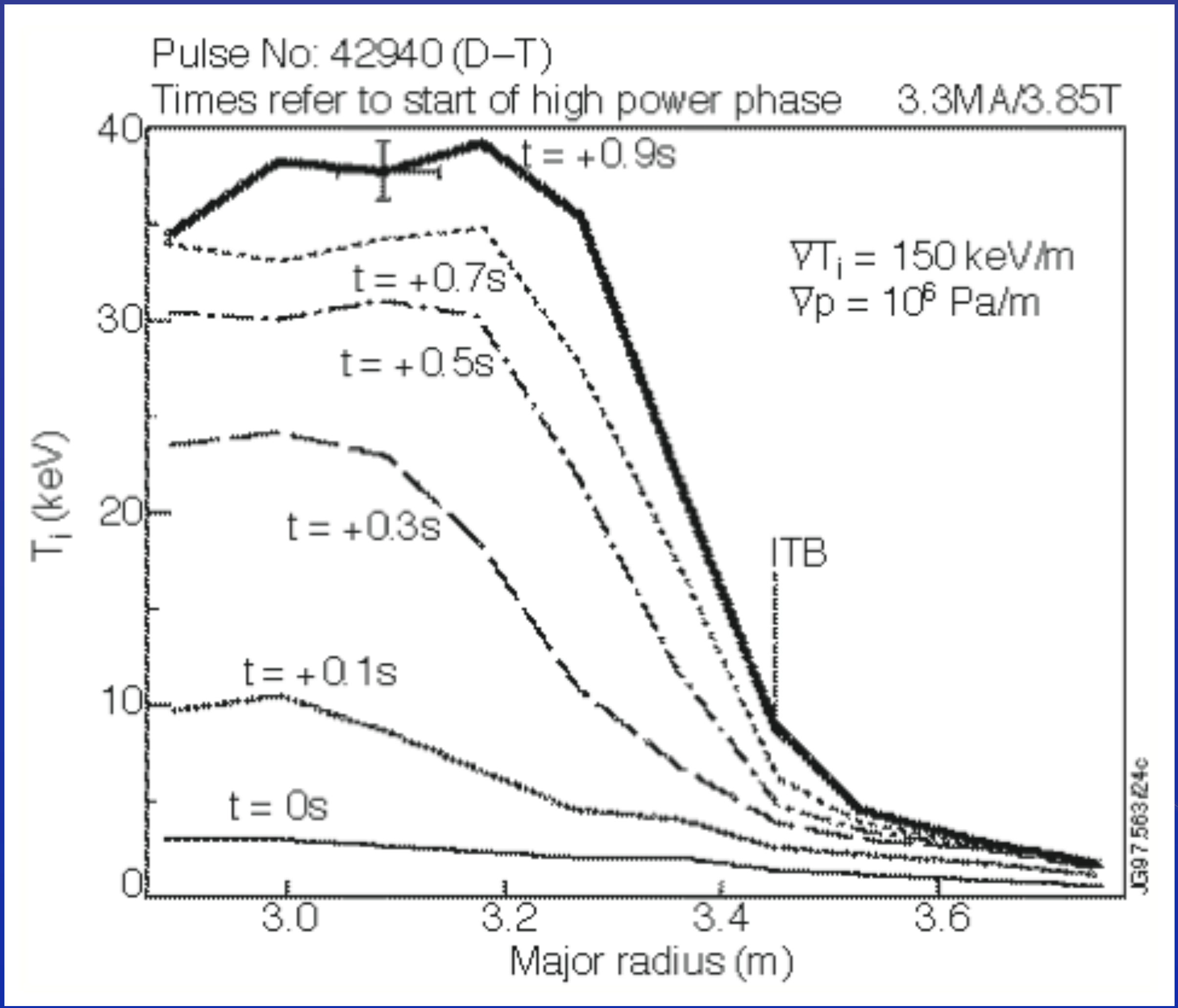
Overview

- Motivation
- Theoretical framework
- Numerical approach
- Example simulation results
- Future directions

Overview

- Motivation
- Theoretical framework
- Numerical approach
- Example simulation results
- Future directions

Objective

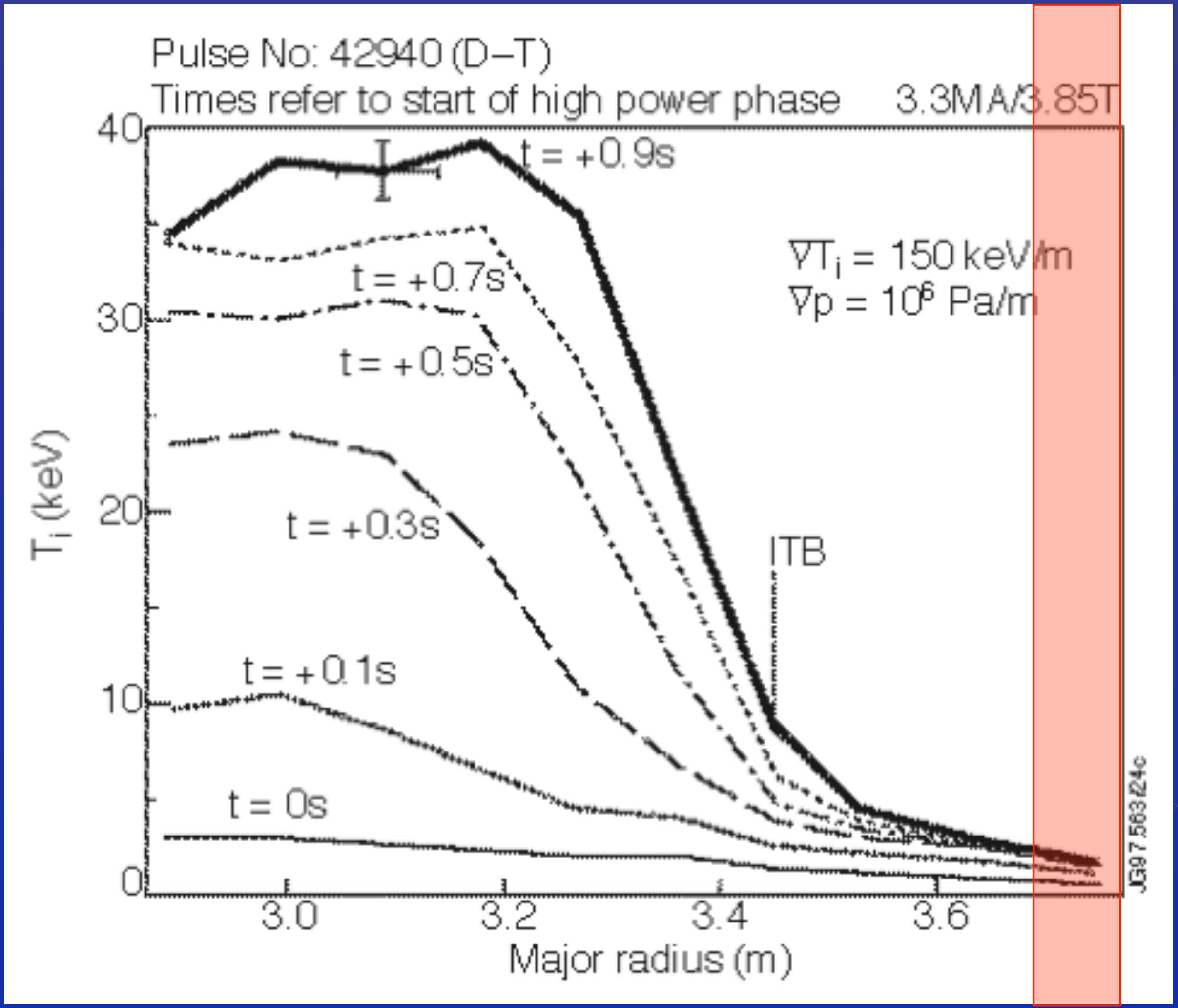


Connor et al. (2004)

Core:
multi-physics,
multi-scale

Edge:
multi-physics,
multi-scale

Objective



Connor et al. (2004)

Core:
multi-physics,
multi-scale

- **kinetic turbulence**
- neoclassical
- sources
- magnetic equilibrium
- MHD

Scale separation in ITER

$$\frac{\partial n}{\partial t} + \nabla \cdot \mathbf{\Gamma} = S_n \qquad \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{Q} + ... = S_p$$

Physics	Perpendicular spatial scale	Temporal scale
Turbulence from ETG modes	$k_\perp^{-1}$ ~ 0.005 – 0.05 cm	ω_* ~ 0.5 - 5.0 MHz
Turbulence from ITG modes	$k_\perp^{-1}$ ~ 0.3 - 3.0 cm	ω_* ~ 10 - 100 kHz
Transport barriers	Measurements suggest width ~ 1 - 10 cm	100 ms or more in core?
Discharge evolution	Profile scales ~ 100 cm	Energy confinement time ~ 2 - 4 s

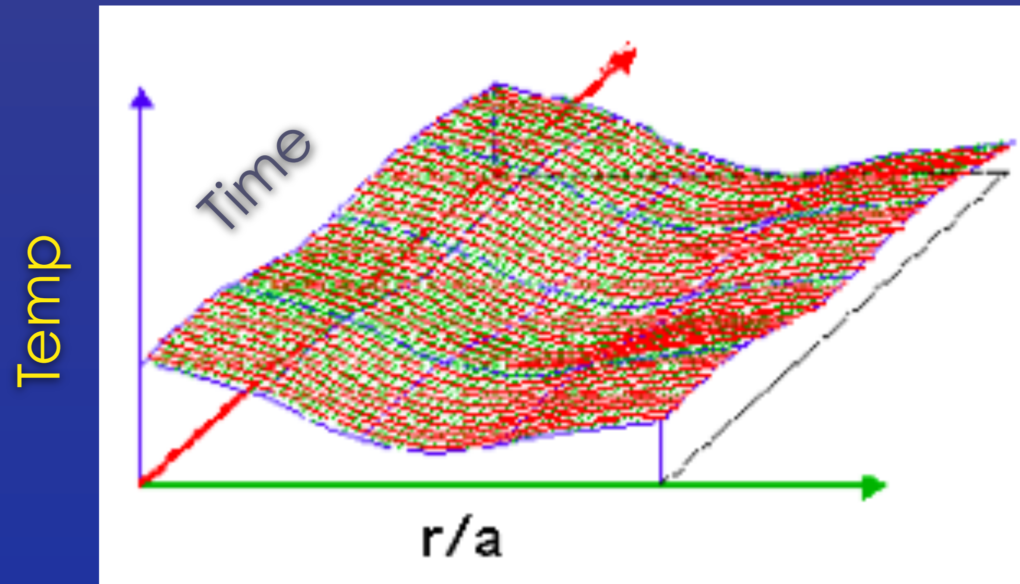
Direct simulation cost

- Grid spacings in space (3D), velocity (3D) and time:

$$\Delta x \sim 0.001\ cm, \quad L_x \sim 100\ cm$$

$$\Delta v \sim 0.1\ v_{th}, \quad L_v \sim v_{th}$$

$$\Delta t \sim 10^{-7}\ s, \quad L_t \sim 1\ s$$



- Grid points required:

$$(L_x/\Delta x)^3 \times (L_v/\Delta v)^3 \times (L_t/\Delta t) \sim 10^{25}$$

- Factor of ~10^9 more than largest fluid turbulence calculations
- Direct simulation not possible; need physics guidance

Improved simulation cost

- Field-aligned coordinates take advantage of $k_{||} \ll k_{\perp}$: savings of ~1000
- Eliminate gyro-angle variable: savings of ~10
- Total saving of ~10^4
- Factor of ~10^5 more than largest fluid turbulence calculations
- Simulation still not possible; need multiscale approach

Major Theoretical & Algorithmic Speedups

Slide from G.W. Hammett

relative to simplest brute force, fully resolved, algorithm, for ITER $1/\rho_* = a/\rho \sim 700$

- Nonlinear gyrokinetic equation
  - ion polarization shielding eliminates plasma freq. $\omega_{pe}/\Omega_{ci} \sim m_i/m_e$ — $\times 10^3$
  - ion polarization eliminates ρ_e & Debye scales $(\rho_i/\rho_e)^3$ — $\times 10^5$
  - average over fast ion gyration, $\Omega_{ci} / \omega_* \sim 1/\rho_*$ — $\times 10^3$
- Continuum or δf PIC, reduces noise, $(f_0/\delta f)^2 \sim 1/\rho_*^2$ — $\times 10^6$
- Field-aligned coordinates (nonlinear extension of ballooning coord.)

$$\Delta_{||} / (\Delta_\perp q R / a) \sim a / (q R \rho_*) \qquad \times 70$$

- Flux-tube / Toroidal annulus wedge, ↓ simulation volume
  - $k_\theta \rho_i = 0, 0.05, 0.1, \ldots, 1.0$
    $n = 0, \; 15, \; 30, \ldots, \; 300$ (i.e., 1/15 of toroidal direction) — x15
  - $L_r \sim a/5 \sim 140\, \rho \sim 10$ correlation lengths — x5
- High-order / spectral algorithms in 5-D, $2^5 \times 2$ — x64
- Implicit electrons — x5-50
- Total combined speedup of all algorithms — $\times 10^{23}$
- Massively parallel computers (Moore's law 1982-2007) — $\times 10^5$

Overview

- Motivation
- Theoretical framework
- Numerical approach
- Example simulation results
- Future directions

Gyrokinetic multiscale assumptions

$$\frac{\partial f}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \frac{\partial f}{\partial \mathbf{R}} + \frac{d\mu}{dt}\frac{\partial f}{\partial \mu} + \frac{dE}{dt}\frac{\partial f}{\partial E} + \frac{d\vartheta}{dt}\frac{\partial f}{\partial \vartheta} = C[f]$$

Gyrokinetic multiscale assumptions

$$\frac{\partial f}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \frac{\partial f}{\partial \mathbf{R}} + \frac{d\mu}{dt}\frac{\partial f}{\partial \mu} + \frac{dE}{dt}\frac{\partial f}{\partial E} + \frac{d\vartheta}{dt}\frac{\partial f}{\partial \vartheta} = C[f]$$

- Turbulent fluctuations are low amplitude:

$$f = F + \delta f \qquad\qquad \delta f \sim \epsilon f$$

Gyrokinetic multiscale assumptions

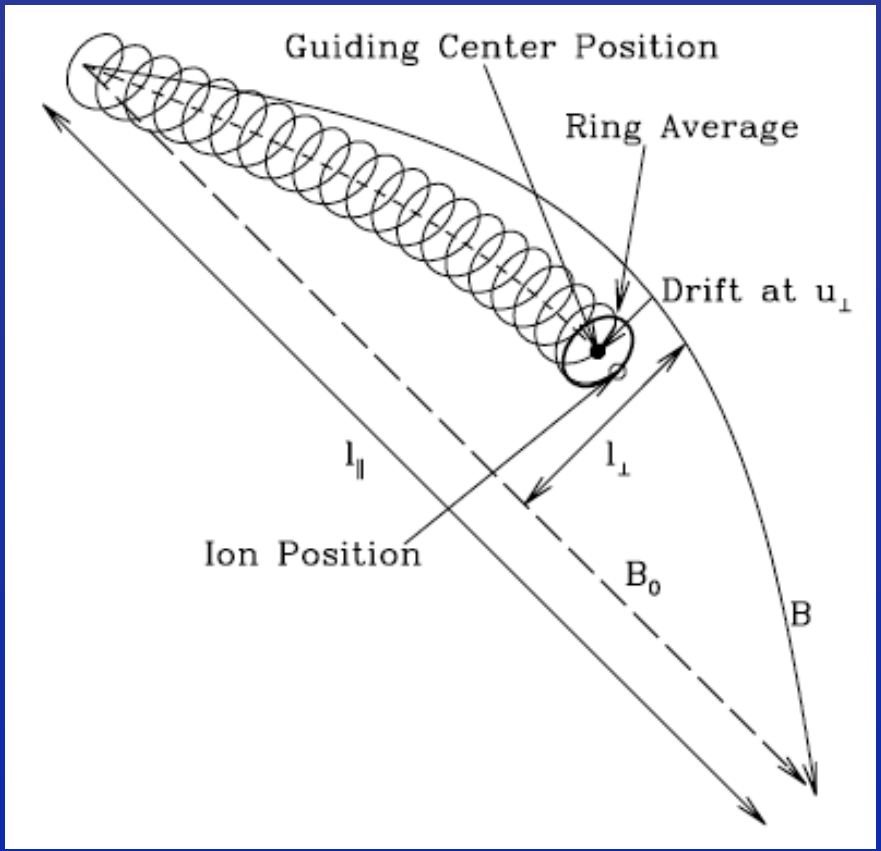
$$\frac{\partial f}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \frac{\partial f}{\partial \mathbf{R}} + \frac{d\mu}{dt}\frac{\partial f}{\partial \mu} + \frac{dE}{dt}\frac{\partial f}{\partial E} + \frac{d\vartheta}{dt}\frac{\partial f}{\partial \vartheta} = C[f]$$

- Turbulent fluctuations are low amplitude:

$$f = F + \delta f \qquad\qquad \delta f \sim \epsilon f$$

- Separation of time scales:

$$\frac{\partial_t \delta f}{\delta f} \sim \omega \sim \epsilon \Omega$$



Gyrokinetic multiscale assumptions

$$\frac{\partial f}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \frac{\partial f}{\partial \mathbf{R}} + \frac{d\mu}{dt}\frac{\partial f}{\partial \mu} + \frac{dE}{dt}\frac{\partial f}{\partial E} + \frac{d\vartheta}{dt}\frac{\partial f}{\partial \vartheta} = C[f]$$

- Turbulent fluctuations are low amplitude:

$$f = F + \delta f \qquad\qquad \delta f \sim \epsilon f$$

- Separation of time scales:

$$\frac{\partial_t \delta f}{\delta f} \sim \omega \sim \epsilon \Omega \qquad\qquad \frac{\partial_t F}{F} \sim \tau^{-1} \sim \epsilon^2 \omega$$

Gyrokinetic multiscale assumptions

$$\frac{\partial f}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \frac{\partial f}{\partial \mathbf{R}} + \frac{d\mu}{dt}\frac{\partial f}{\partial \mu} + \frac{dE}{dt}\frac{\partial f}{\partial E} + \frac{d\vartheta}{dt}\frac{\partial f}{\partial \vartheta} = C[f]$$

- Turbulent fluctuations are low amplitude:

$$f = F + \delta f \qquad\qquad \delta f \sim \epsilon f$$

- Separation of time scales:

$$\frac{\partial_t \delta f}{\delta f} \sim \omega \sim \epsilon \Omega \qquad\qquad \frac{\partial_t F}{F} \sim \tau^{-1} \sim \epsilon^2 \omega$$

- Separation of space scales:

$$\nabla F \sim F/L, \quad \nabla_{\|}\delta f \sim \delta f/L, \quad \nabla_{\perp}\delta f \sim \delta f/\rho$$

Gyrokinetic multiscale assumptions

$$\frac{\partial f}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \frac{\partial f}{\partial \mathbf{R}} + \frac{d\mu}{dt}\frac{\partial f}{\partial \mu} + \frac{dE}{dt}\frac{\partial f}{\partial E} + \frac{d\vartheta}{dt}\frac{\partial f}{\partial \vartheta} = C[f]$$

- Turbulent fluctuations are low amplitude:

$$f = F + \delta f \qquad\qquad \delta f \sim \epsilon f$$

- Separation of time scales:

$$\frac{\partial_t \delta f}{\delta f} \sim \omega \sim \epsilon\Omega \qquad\qquad \frac{\partial_t F}{F} \sim \tau^{-1} \sim \epsilon^2 \omega$$

- Separation of space scales:

$$\nabla F \sim F/L, \quad \nabla_{||}\delta f \sim \delta f/L, \quad \nabla_{\perp}\delta f \sim \delta f/\rho$$

- "Smooth" velocity space:

$$\epsilon \lesssim \nu/\omega \lesssim 1 \Rightarrow \sqrt{\epsilon} \lesssim \delta v/v_{th} \lesssim 1$$

Gyrokinetic multiscale assumptions

$$\frac{\partial f}{\partial t} + \frac{d\mathbf{R}}{dt}\cdot\frac{\partial f}{\partial \mathbf{R}} + \frac{d\mu}{dt}\frac{\partial f}{\partial \mu} + \frac{dE}{dt}\frac{\partial f}{\partial E} + \frac{d\vartheta}{dt}\frac{\partial f}{\partial \vartheta} = C[f]$$

- Turbulent fluctuations are low amplitude:

$$f = F + \delta f \qquad\qquad \delta f \sim \epsilon f$$

- Separation of time scales:

$$\frac{\partial_t \delta f}{\delta f} \sim \omega \sim \epsilon\Omega \qquad\qquad \frac{\partial_t F}{F} \sim \tau^{-1} \sim \epsilon^2\omega$$

- Separation of space scales:

$$\nabla F \sim F/L, \quad \nabla_{||}\delta f \sim \delta f/L, \quad \nabla_{\perp}\delta f \sim \delta f/\rho$$

- "Smooth" velocity space:

$$\epsilon \lesssim \nu/\omega \lesssim 1 \Rightarrow \sqrt{\epsilon} \lesssim \delta v/v_{th} \lesssim 1$$

- Sub-sonic drifts: $v_D \sim \epsilon v_{th}$

Key results: turbulence and transport

$$f = F_0 + h + ... \qquad\qquad F_0 = F_M(\mathbf{R}) \exp\left(-\frac{q\Phi}{T}\right)$$

Key results: turbulence and transport

$$f = F_0 + h + \ldots \qquad\qquad F_0 = F_M(\mathbf{R}) \exp\left(-\frac{q\Phi}{T}\right)$$

Gyrokinetic equation for turbulence:

$$\partial h/\partial t + v_{\|}\hat{\mathbf{b}}\cdot\nabla h + \langle \mathbf{v}_\chi \rangle_{\mathbf{R}} \cdot \nabla (F_0 + h) + \mathbf{v}_{\mathbf{B}} \cdot \nabla h = \frac{qF_0}{T_0}\frac{\partial \langle \chi \rangle_{\mathbf{R}}}{\partial t} + \langle C[h] \rangle_{\mathbf{R}}$$

Key results: turbulence and transport

$$f = F_0 + h + \ldots \qquad F_0 = F_M(\mathbf{R}) \exp\left(-\frac{q\Phi}{T}\right)$$

Gyrokinetic equation for turbulence:

$$\partial h/\partial t + v_{\|}\hat{\mathbf{b}}\cdot\nabla h + \langle \mathbf{v}_\chi \rangle_{\mathbf{R}} \cdot \nabla(F_0 + h) + \mathbf{v_B}\cdot\nabla h = \frac{qF_0}{T_0}\frac{\partial \langle \chi \rangle_{\mathbf{R}}}{\partial t} + \langle C[h] \rangle_{\mathbf{R}}$$

Moment equations for equilibrium evolution:

$$\begin{aligned}
\frac{\partial n_s}{\partial t} &= -\frac{1}{V'}\frac{\partial}{\partial \psi}\left(V'\left\langle \mathbf{\Gamma}_s \cdot \nabla\psi \right\rangle\right) + S_n \\
\frac{3}{2}\frac{\partial n_s T_s}{\partial t} &= -\frac{1}{V'}\frac{\partial}{\partial \psi}\left(V'\left\langle \mathbf{Q}_s \cdot \nabla\psi \right\rangle\right) \\
&+ T_s\left(\frac{\partial \ln n_s}{\partial \psi} - \frac{3}{2}\frac{\partial \ln T_s}{\partial \psi}\right)\left\langle \mathbf{\Gamma}_s \cdot \nabla\psi \right\rangle + \frac{\partial \ln T_s}{\partial \psi}\left\langle \mathbf{Q}_s \cdot \nabla\psi \right\rangle \\
&- \left\langle \int d^3v \frac{h_s T_s}{F_{0s}} \left\langle C[h_s] \right\rangle_{\mathbf{R}} \right\rangle + n_s \nu_\epsilon^{su}\left(T_u - T_s\right) + S_p
\end{aligned}$$

Sugama (1997)

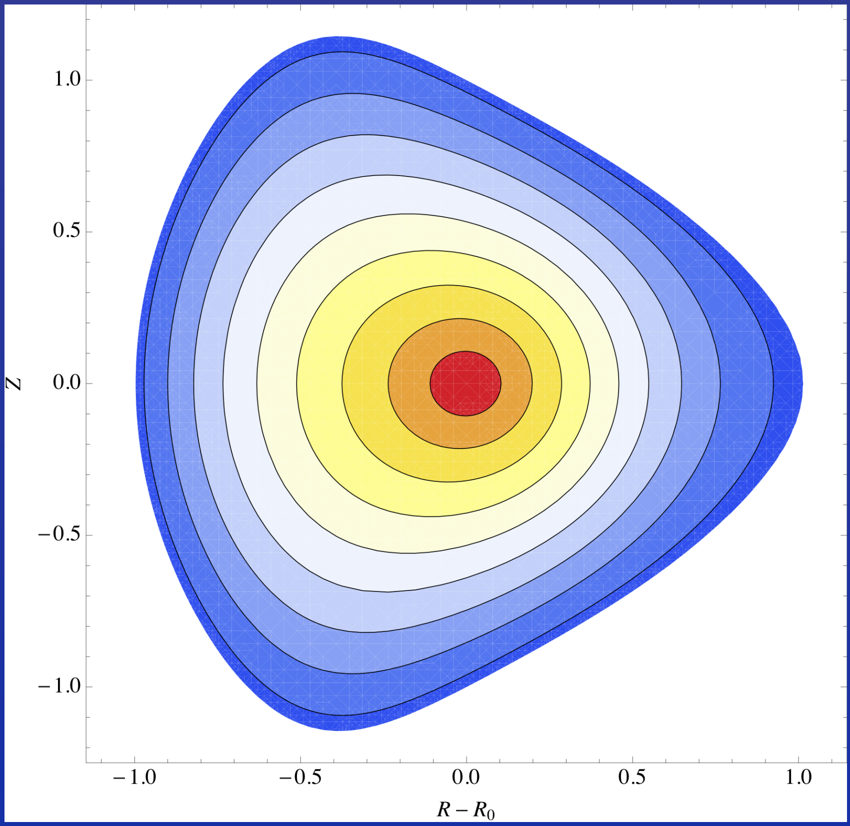
Space-time averages

- Flux surface average
- Intermediate space average:

$$\rho \ll \Delta_\psi \ll L$$

- Intermediate time average:

$$\tau_t \ll \Delta_\tau \ll \tau_E$$



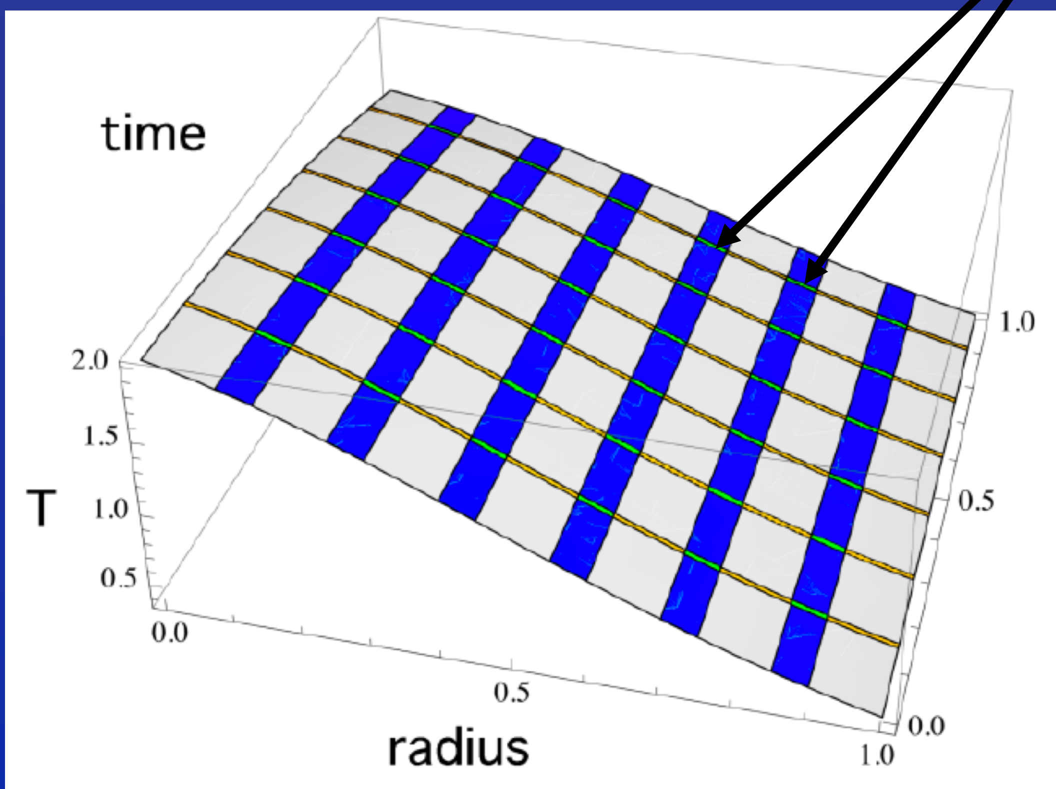
Overview

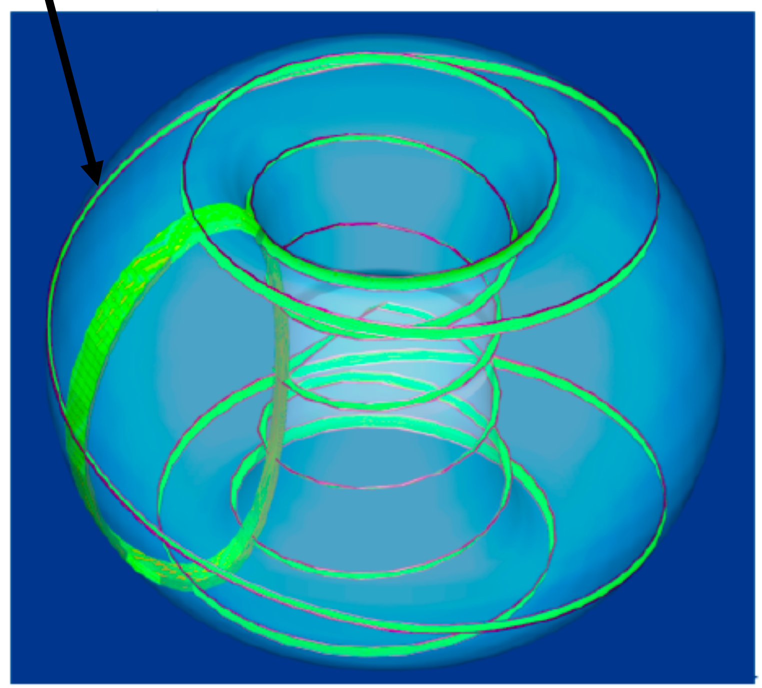
- Motivation
- Theoretical framework
- Numerical approach
- Example simulation results
- Future directions

Multiscale grid

- Turbulent fluctuations calculated in small regions of fine space-time grid embedded in “coarse” grid (for mean quantities)

Flux tube simulation domain





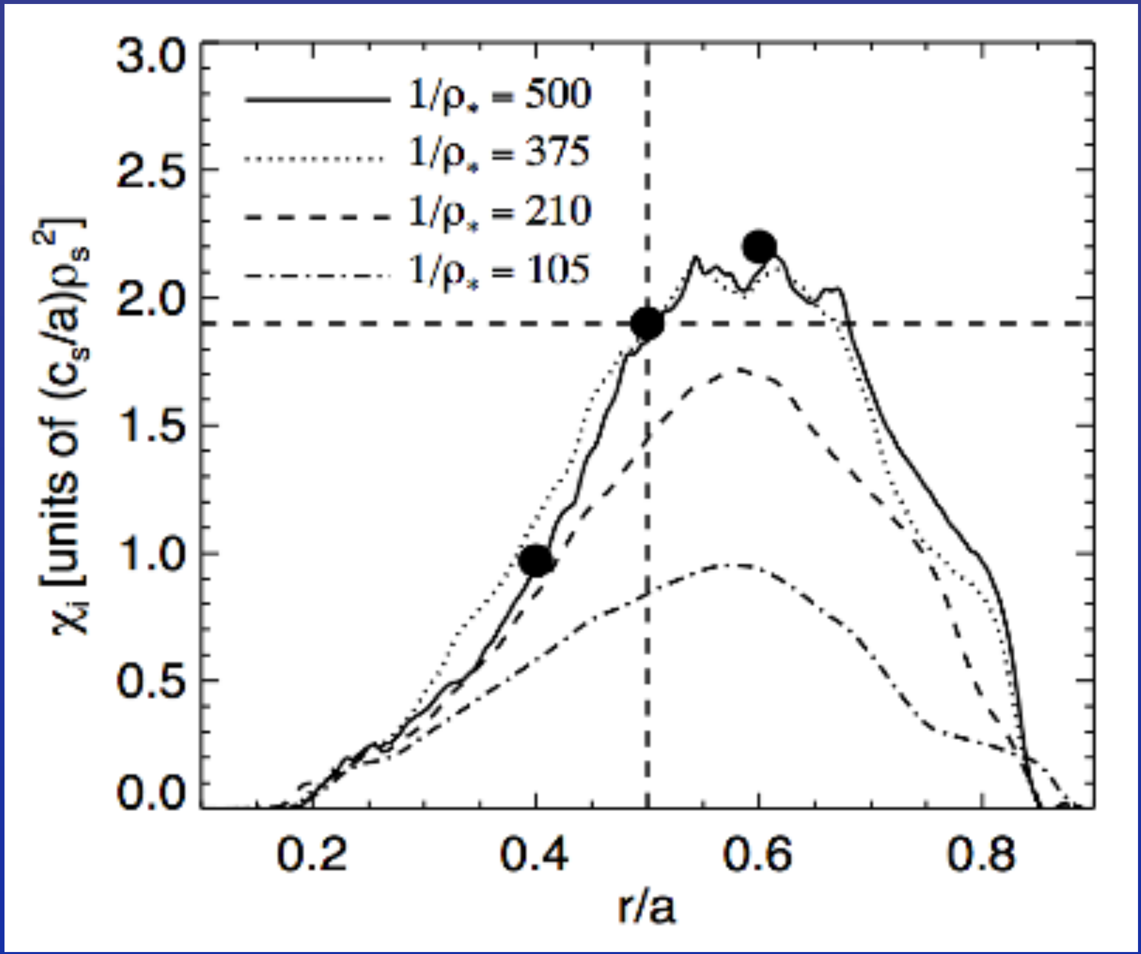
Flux tube assumptions

- Macroscopic quantities (density, flow, temperature, etc. constant across simulation domain)
- Gradient scale lengths of macroscopic quantities constant across simulation domain
  - Total gradient NOT constant (corrugations possible due to fluctuation + equilibrium gradients)
- In addition to delta-f assumption that equilibrium quantities constant in time over simulation
- => No important meso-scale physics

Validity of flux tube approximation

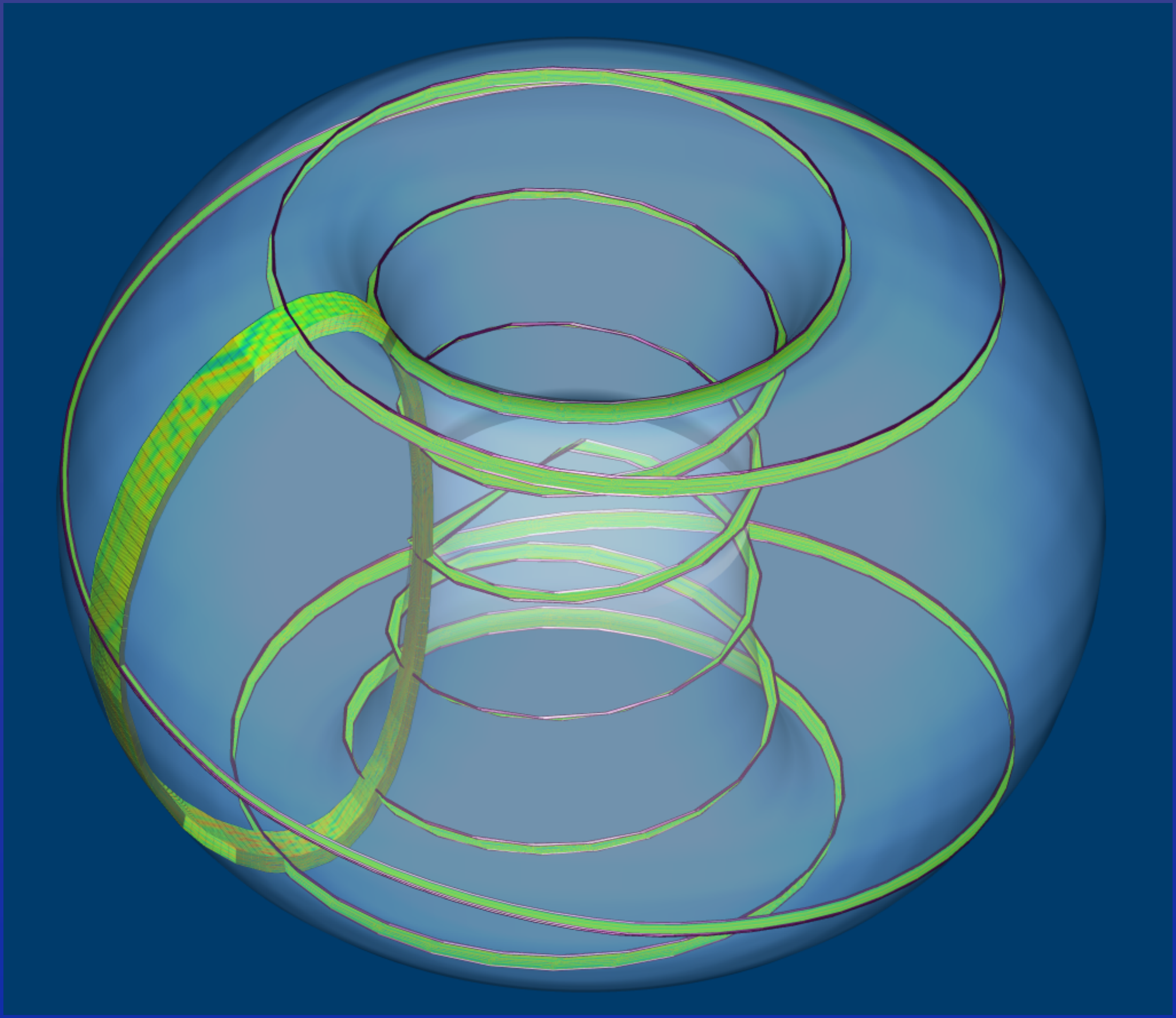
- Lines represent global simulations from GYRO
- Dots represent local (flux tube) simulations from GS2
- Excellent agreement for

$$\rho_* \ll 1$$



*J. Candy, R.E. Waltz and W. Dorland, The local limit of global gyrokinetic simulations, Phys. Plasmas **11** (2004) L25.

Flux tubes minimize flux surface grid points



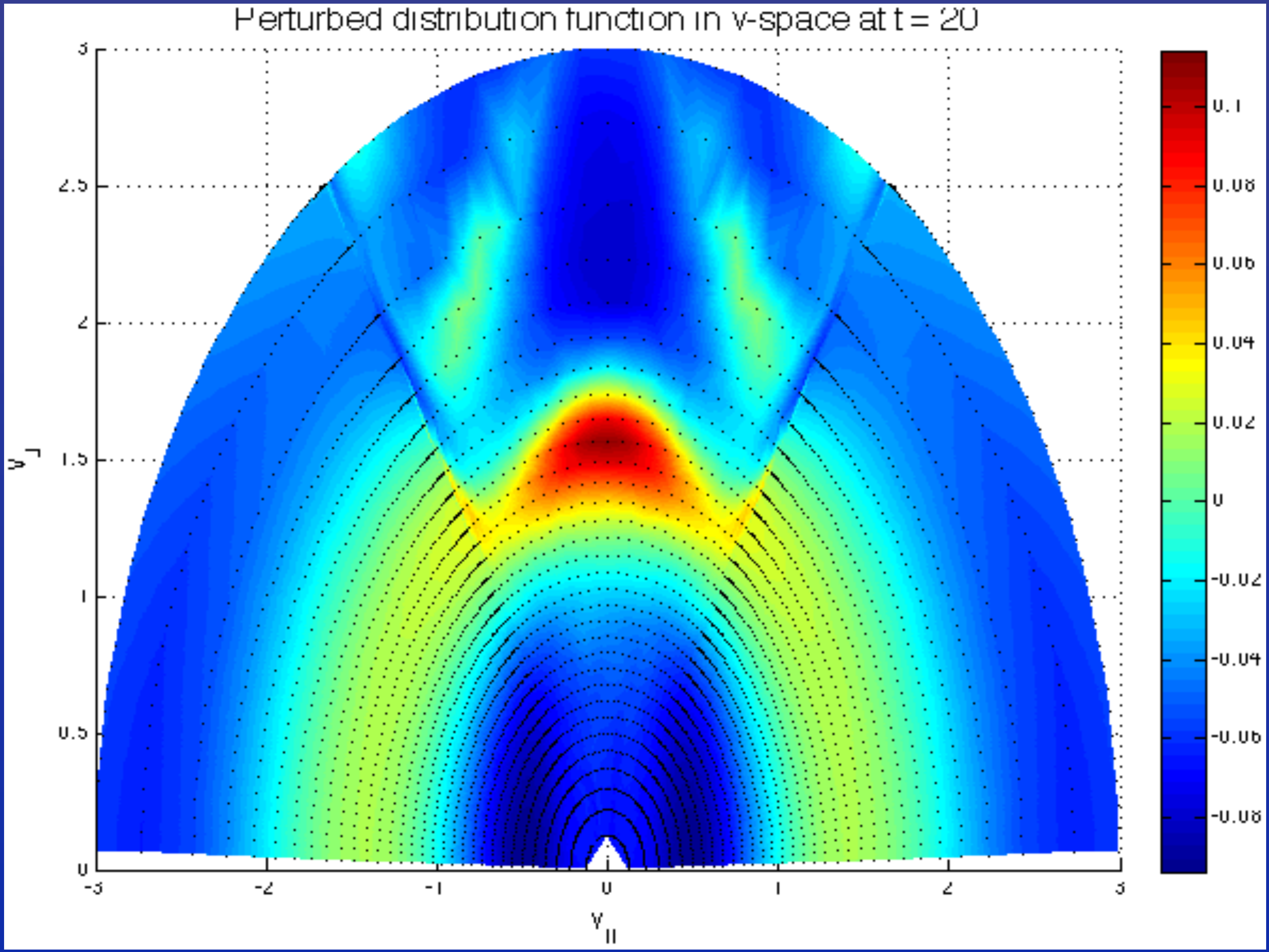
Simulation volume reduced by factor of ~10s

GS2 features

- V-space variables: energy and magnetic moment
- Realistic magnetic geometry
  - Numerical equilibrium from experiment
  - Miller local equilibrium
  - S-alpha model
- Multiple kinetic species
- Model Fokker-Planck collision operator
- Implicit treatment of linear physics
- Includes background flow shear

GS2 v-space

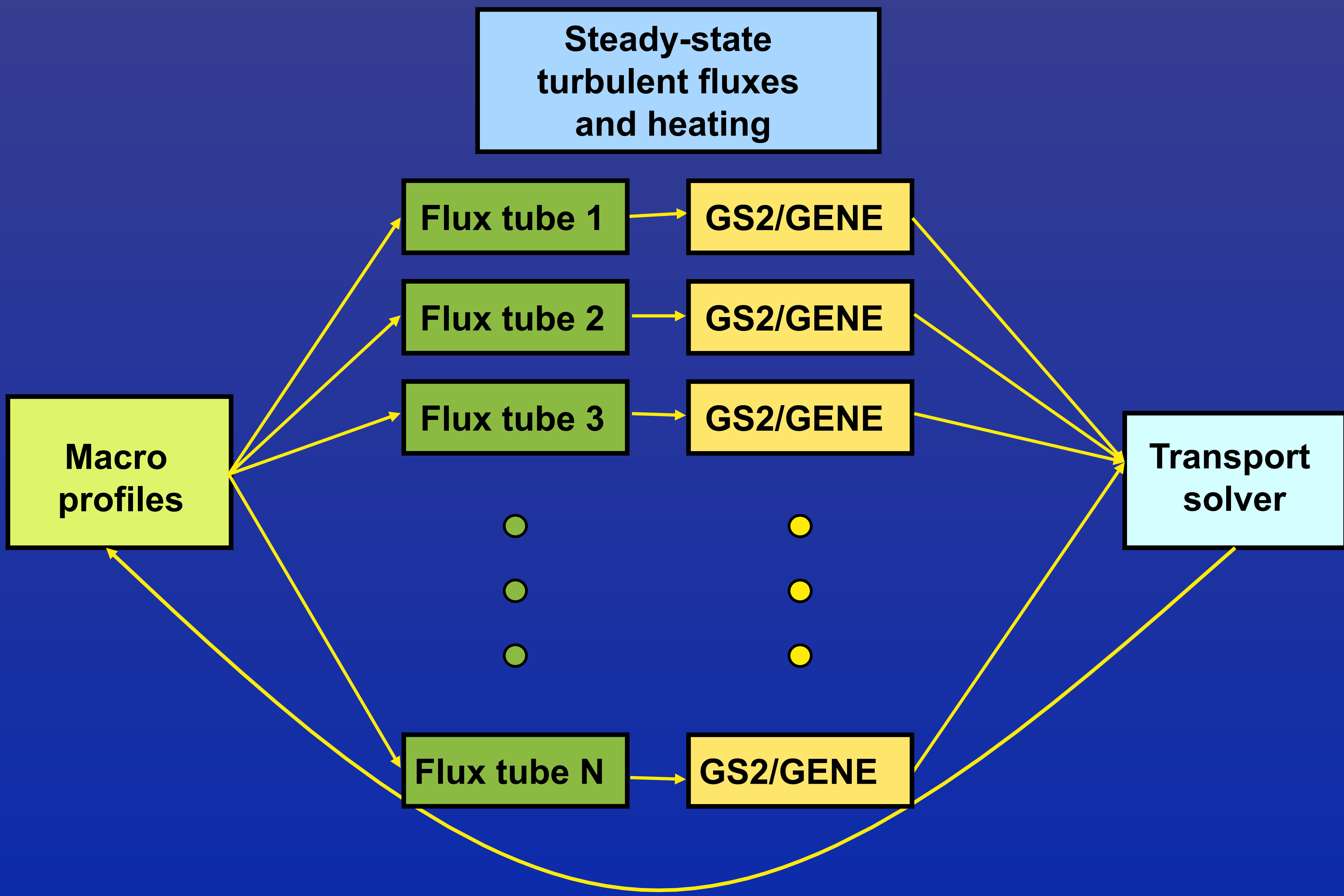
$$\frac{\langle \delta f \rangle_{\mathbf{R}}}{F_M}$$



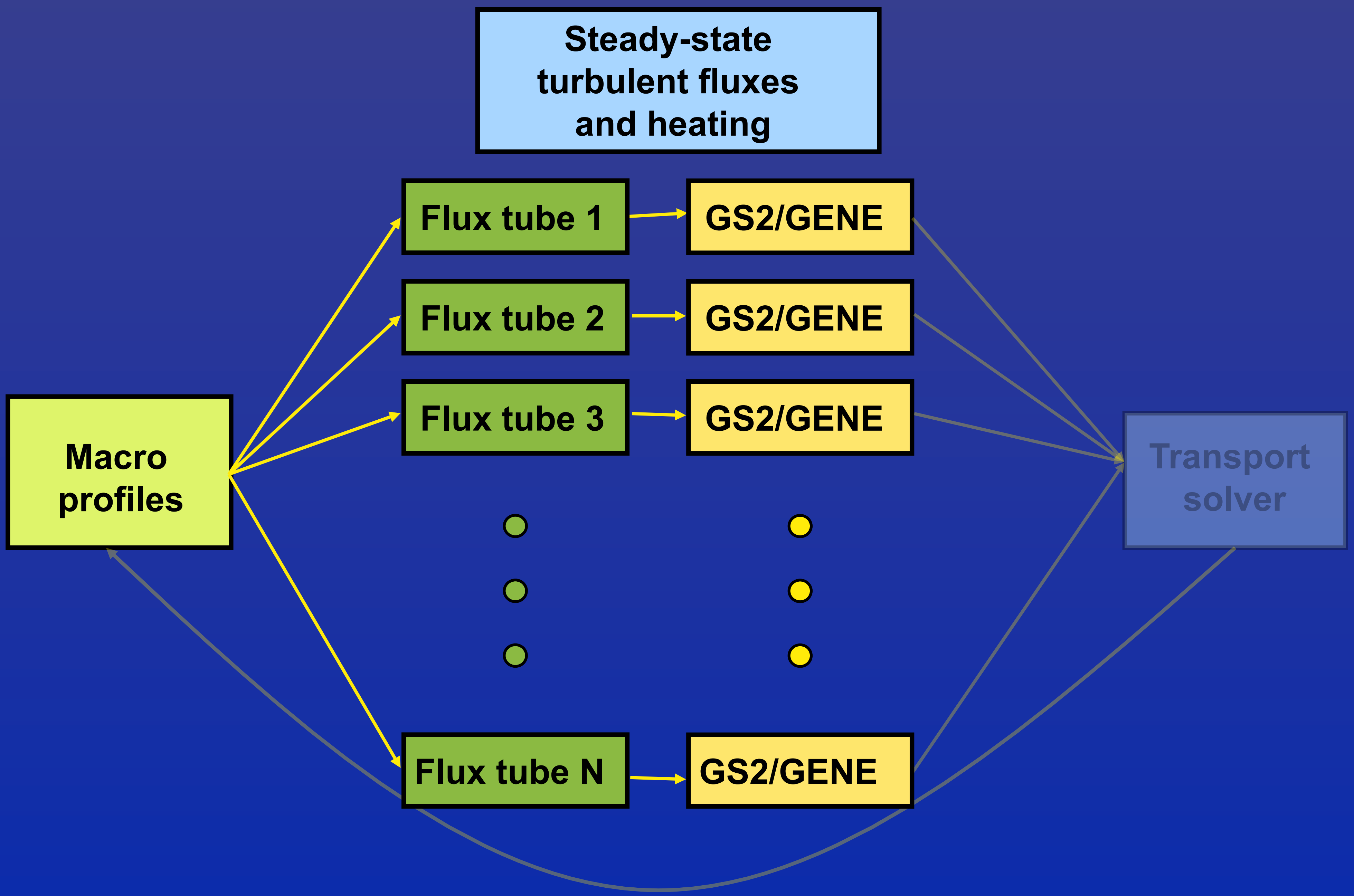
GS2 features

- V-space variables: energy and magnetic moment
- Realistic magnetic geometry
  - Numerical equilibrium from experiment
  - Miller local equilibrium
  - S-alpha model
- Multiple kinetic species
- Model Fokker-Planck collision operator
- Implicit treatment of linear physics
- Includes background flow shear

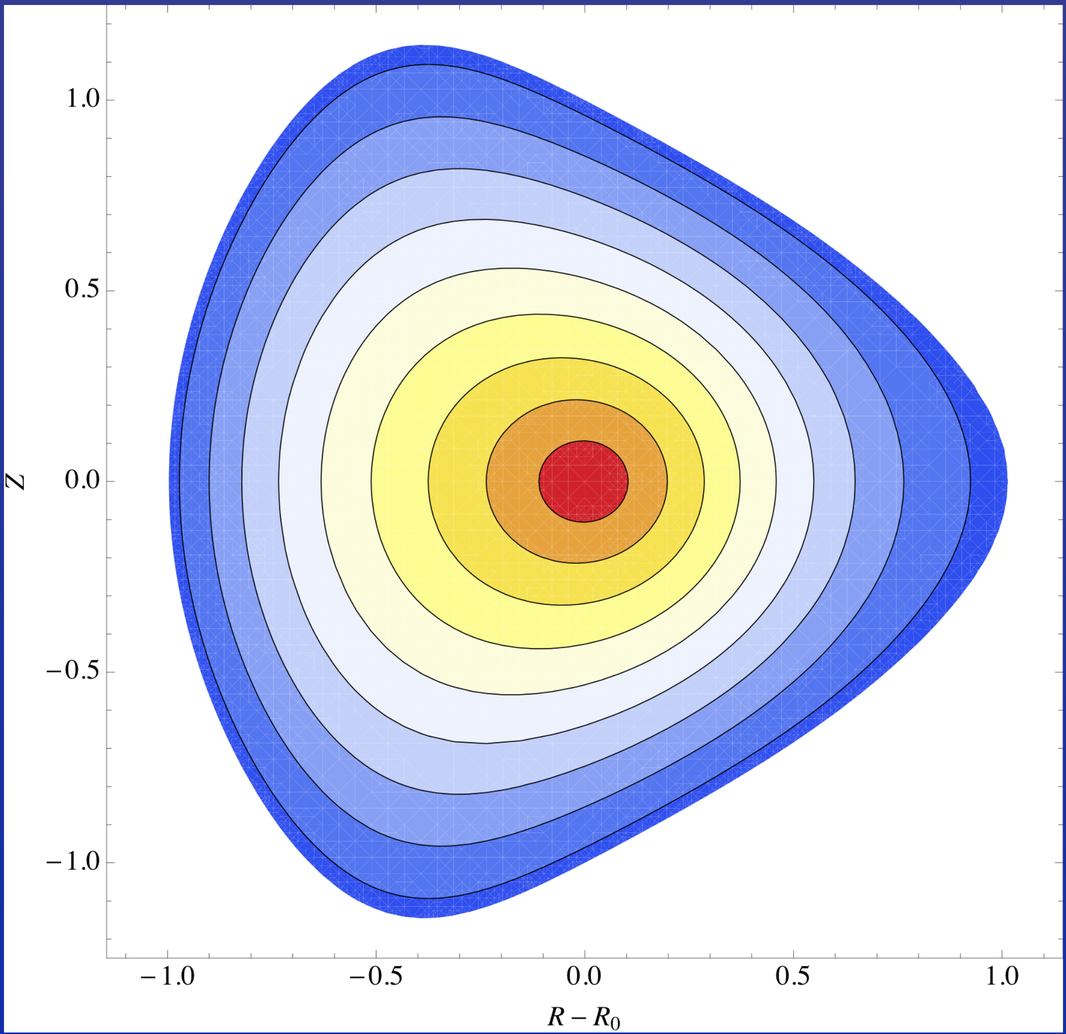
Trinity schematic

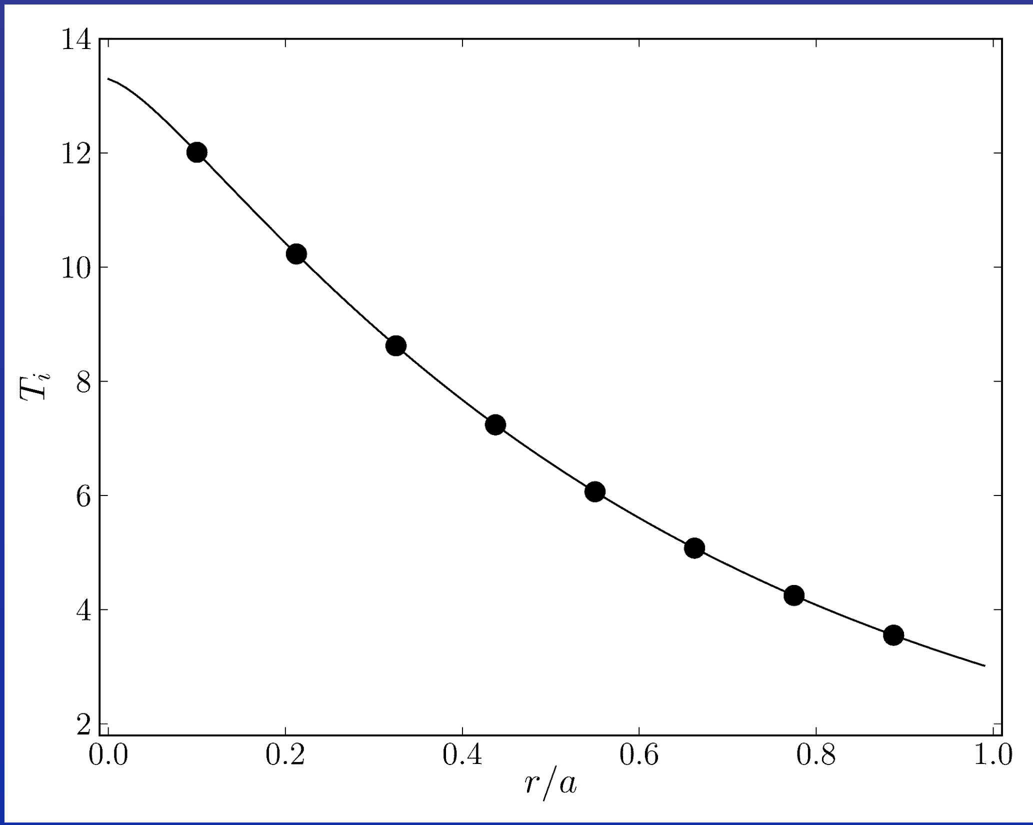


Trinity schematic

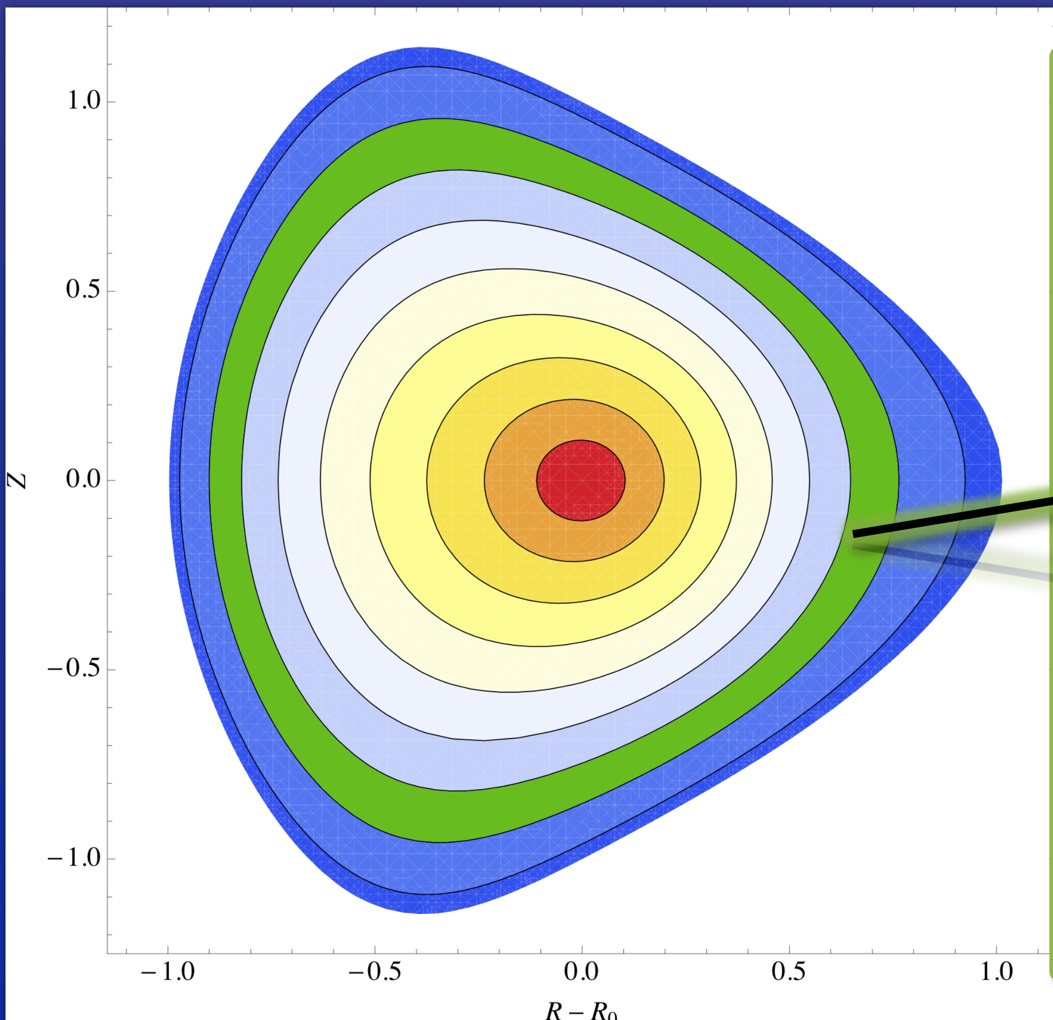


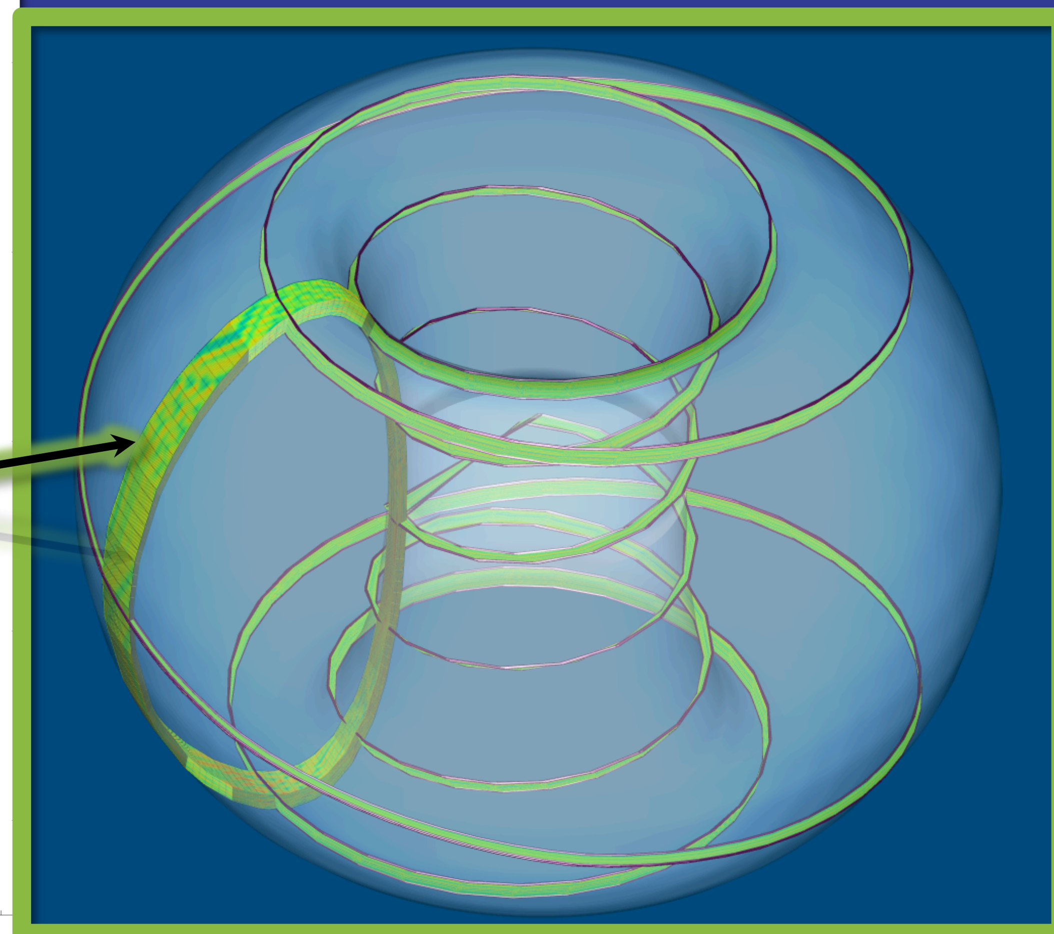
Sampling profile with flux tubes



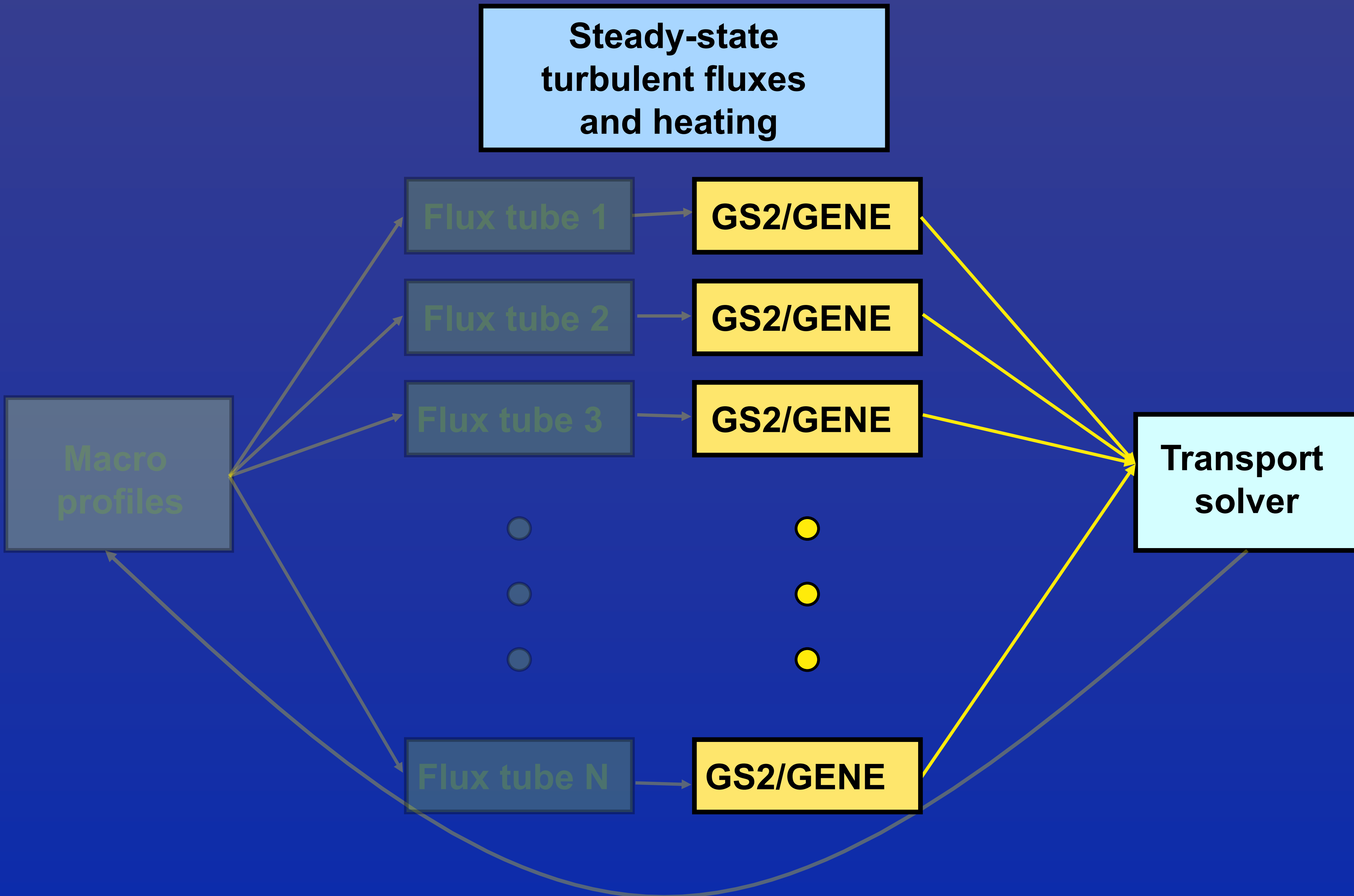


Sampling profile with flux tubes





Trinity schematic



Transport equations in GK

Moment equations for evolution of mean quantities:

$$
\begin{aligned}
\frac{\partial n_s}{\partial t} &= -\frac{1}{V'}\frac{\partial}{\partial \psi}\left(V'\left\langle \mathbf{\Gamma}_s \cdot \nabla\psi\right\rangle\right) + S_n \\
\frac{3}{2}\frac{\partial n_s T_s}{\partial t} &= -\frac{1}{V'}\frac{\partial}{\partial \psi}\left(V'\left\langle \mathbf{Q}_s \cdot \nabla\psi\right\rangle\right) \\
&+ T_s\left(\frac{\partial \ln n_s}{\partial \psi} - \frac{3}{2}\frac{\partial \ln T_s}{\partial \psi}\right)\left\langle \mathbf{\Gamma}_s \cdot \nabla\psi\right\rangle + \frac{\partial \ln T_s}{\partial \psi}\left\langle \mathbf{Q}_s \cdot \nabla\psi\right\rangle \\
&- \left\langle \int d^3v \frac{h_s T_s}{F_{0s}}\left\langle C[h_s]\right\rangle_{\mathbf{R}}\right\rangle + n_s \nu_\epsilon^{su}\left(T_u - T_s\right) + S_p
\end{aligned}
$$

$$
\frac{\partial L}{\partial t} = -\frac{1}{V'}\frac{\partial}{\partial \psi}\left(V'\sum_s \left\langle \pi_s \right\rangle\right) + S_L
$$

...depend on fluctuations

Sugama (1997)

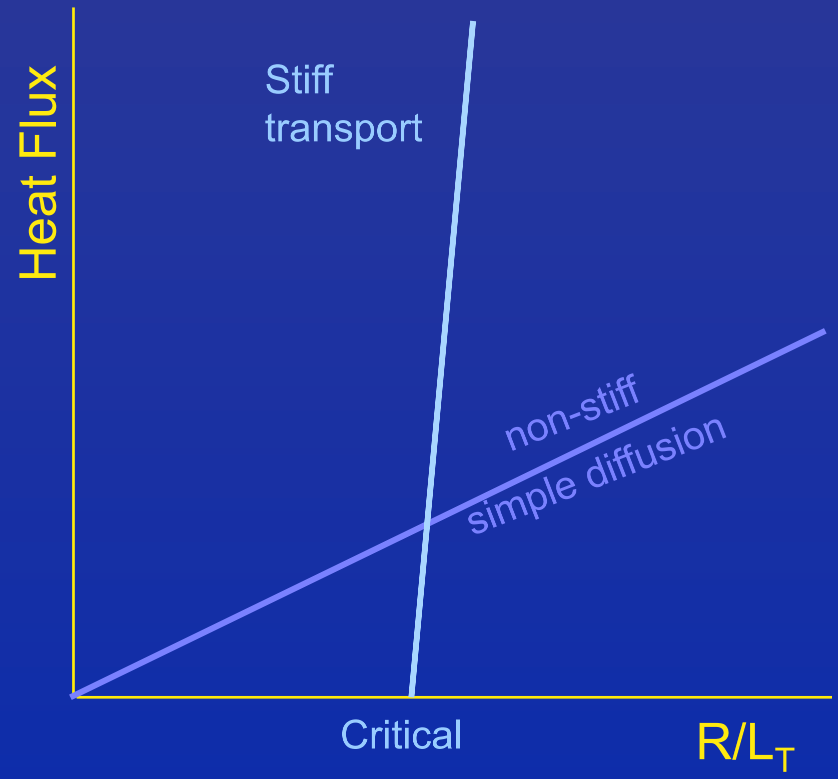
Trinity transport solver

- Transport equations are stiff, nonlinear PDEs:

$$\frac{3}{2}\frac{\partial p_s}{\partial t} = -\frac{1}{V'}\frac{\partial}{\partial \psi}\left(V' \left\langle \mathbf{Q}_s \cdot \nabla \psi \right\rangle\right) + \ldots$$

$$\mathbf{Q}_s = \mathbf{Q}_s(n(\psi,t), T(\psi,t); \psi, t)$$

Implicit treatment needed for stiffness



Trinity transport solver

$$\frac{\partial n}{\partial t} = H(r)\frac{\partial}{\partial r}G[n(r,t),T(r,t);r,t]$$

- General (single-step or multi-step) time discretization:

$$\frac{n^{m+1}-n^{m}}{\Delta\tau} = \alpha\left[H\frac{\partial G}{\partial r}\right]^{m+1} + (1-\alpha)\left[H\frac{\partial G}{\partial r}\right]^{m}$$

- 2nd order centered difference in radial coordinate (equally spaced grid):

$$\frac{\partial G}{\partial r} = \frac{G_{j+1/2}-G_{j-1/2}}{\Delta r}$$

Trinity transport solver

- Implicit treatment via Newton's Method) allows for time steps ~0.1 seconds (vs. turbulence sim time ~0.001 seconds)
- Challenge: requires computation of quantities like

$$\Gamma_j^{m+1} \approx \Gamma_j^m + \left(\mathbf{y}^{m+1} - \mathbf{y}^m\right) \left.\frac{\partial \Gamma_j}{\partial \mathbf{y}}\right|_{\mathbf{y}^m} \qquad \mathbf{y} = \left[\{n_k\}, \{p_{i_k}\}, \{p_{e_k}\}\right]^T$$

- Local approximation: $$\frac{\partial \Gamma_j}{\partial n_k} = \frac{\partial \Gamma_j}{\partial n_j} + \frac{\partial \Gamma_j}{\partial (R/L_n)_j} \frac{\partial (R/L_n)_j}{\partial n_k}$$
- Simplifying assumption: normalized fluxes depend primarily on gradient scale lengths

*S.C. Jardin, G. Bateman, G.W. Hammett, and L.P. Ku, On 1D diffusion problems with a gradient-dependent diffusion coefficient, J. Comp. Phys. **227**, 8769 (2008).

Trinity transport solver

- Calculating flux derivative approximations:
  - at every radial grid point, simultaneously calculate $\Gamma_j[(R/L_n)_j^m]$ and $\Gamma_j[(R/L_n)_j^m + \delta]$ using 2 different flux tubes
  - use 2-point finite differences:

$$\frac{\partial \Gamma_j}{\partial (R/L_n)_j} \approx \frac{\Gamma_j[(R/L_n)_j^m] - \Gamma_j[(R/L_n)_j^m + \delta]}{\delta}$$

  - possible because flux tubes independent (do not communicate during calculation)
  - perfect parallelization (almost)

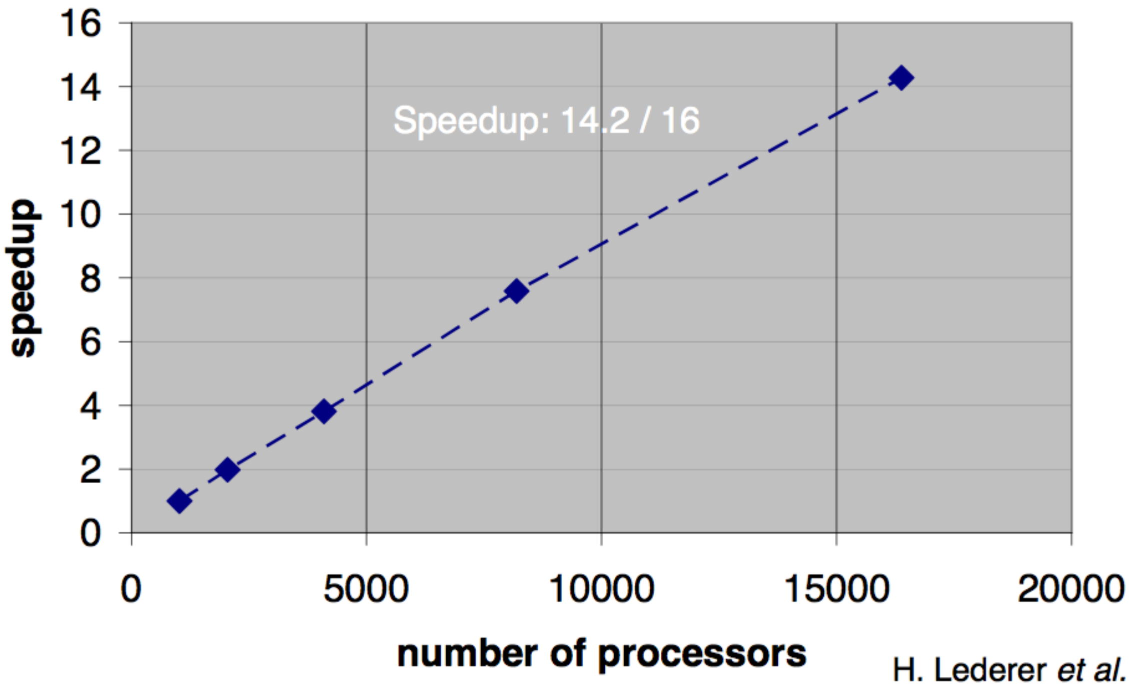
Trinity transport solver

- Nonlinear turbulence simulation runs until fluxes converged
- Turbulence for new transport time step initialized to saturated state from previous transport time step – faster convergence
- Option to use model fluxes (IFS-PPPL, quasilinear, etc.)
- Sources specified analytically or taken from experiment

Trinity scaling

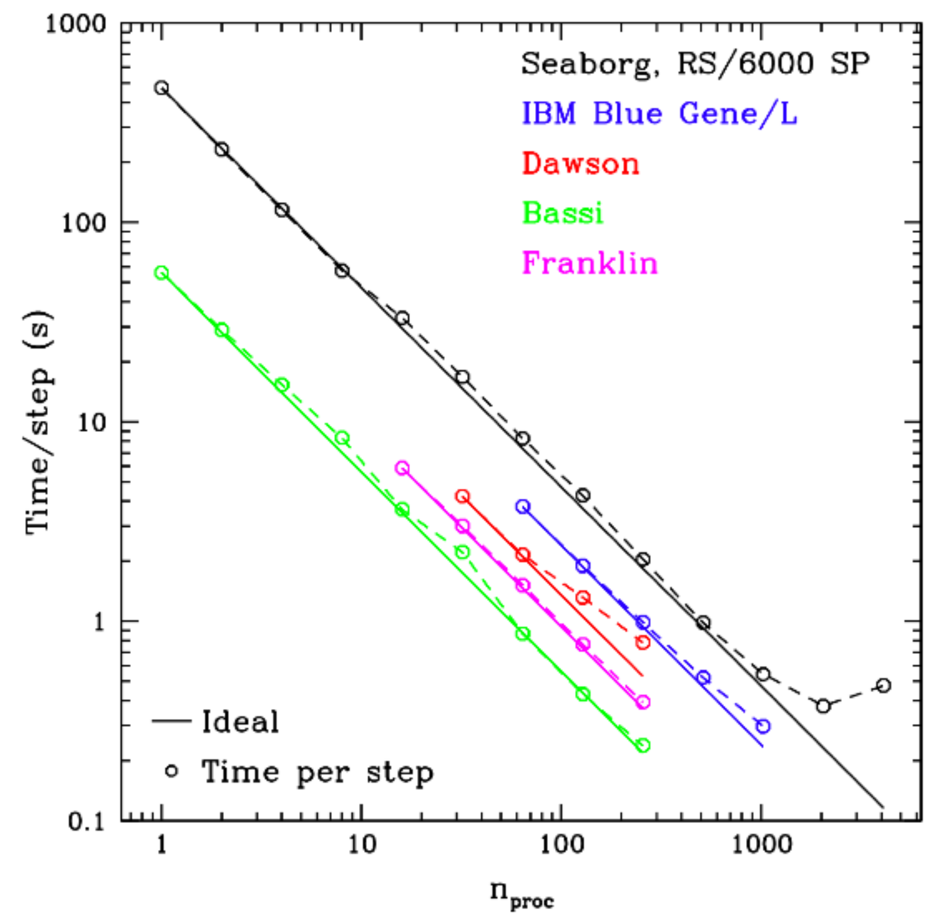
- Example calculation with 10 radial grid points:
  - evolve density, toroidal angular momentum, and electron/ion pressures
  - simultaneously calculate fluxes for equilibrium profile and for 4 separate profiles (one for each perturbed gradient scale length)
  - total of 50 flux tube simulations running simultaneously
  - ~2000-4000 processors per flux tube => scaling to over 100,000 processors with high efficiency
- Adding radial grid points, multiple species, electron space-time scales, and other physics increases weak scaling by up to 10^4

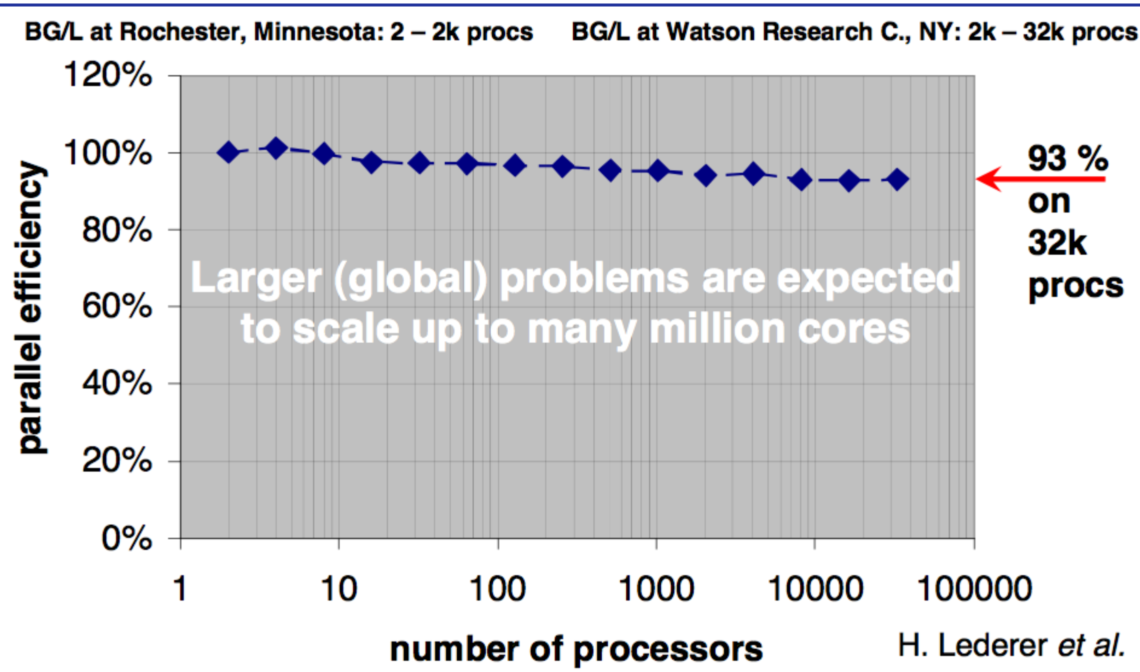
Flux tube scaling



GENE strong scaling

GS2 strong scaling





GENE weak scaling

Boundary conditions

- Various initialization options:
  - Analytic specification
  - Experimental profiles
  - Numerical profiles (from IFS-PPPL, etc.)
- Fix density and temperature at outer edge of simulation domain
  - Predict performance as a function of pedestal height
- Vanishing fluxes at magnetic axis:

$$\psi \rightarrow 0 : \quad V'Q = V'\Gamma = 0$$

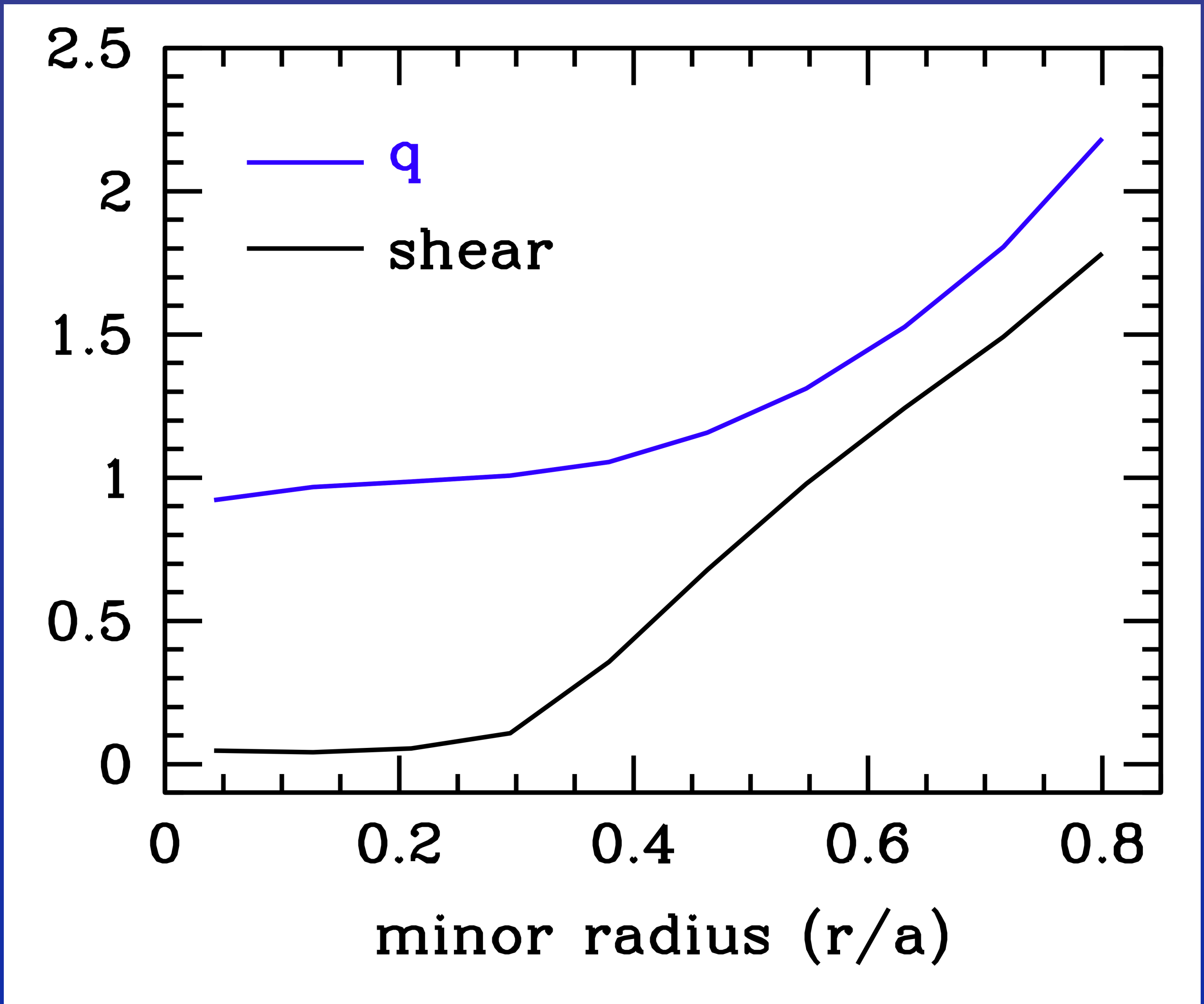
Multi-scale simulation savings

- Statistical periodicity in toroidal direction takes advantage of $k_\perp^{-1} \ll L_\theta$: volume savings factor of ~10-100
- Exploitation of scale separation between turbulence and equilibrium evolution: time savings factor of ~100
- Total saving of ~10^4 + extreme parallelizability: simulation possible on current machines

Overview

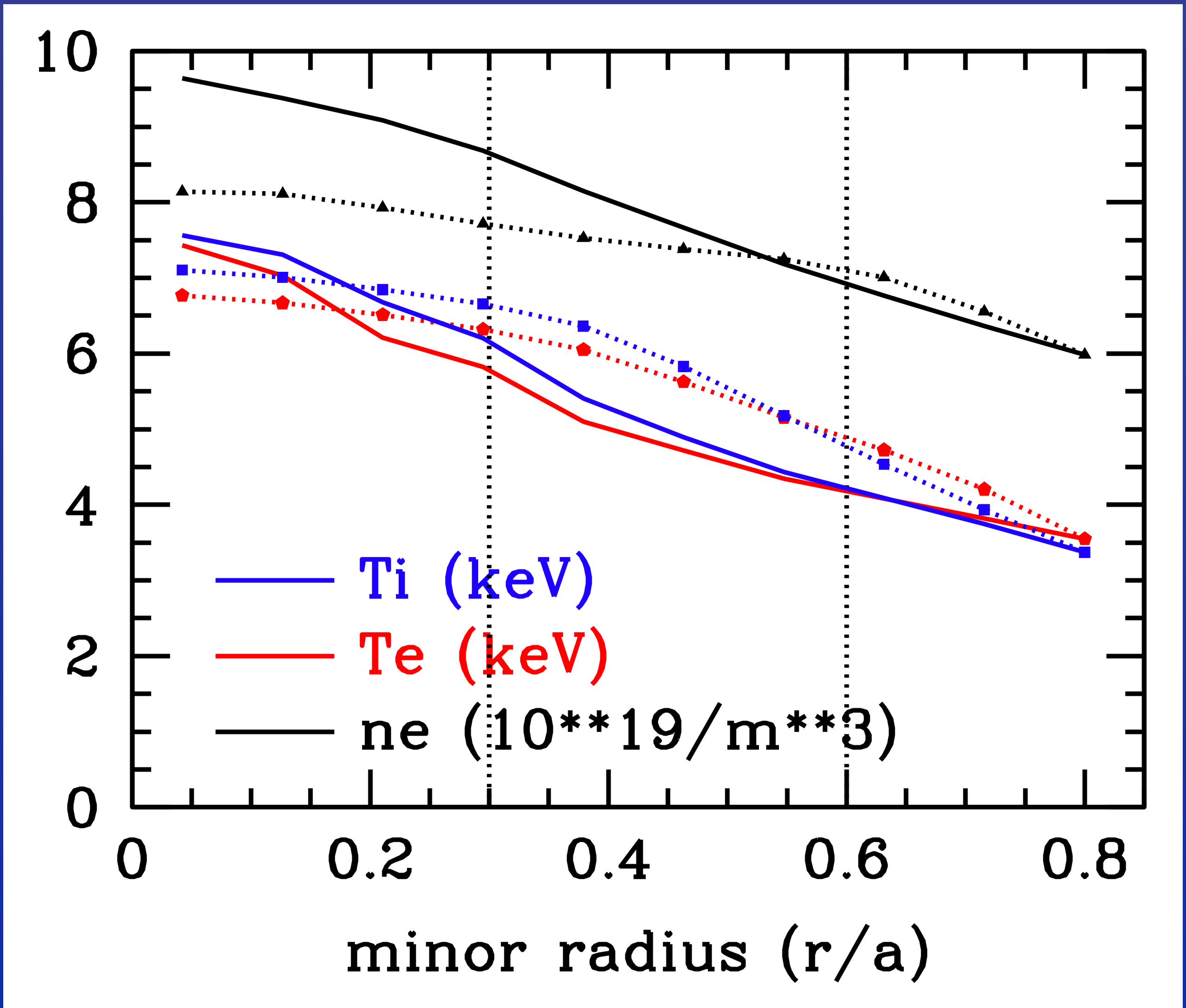
- Motivation
- Theoretical framework
- Numerical approach
- Example simulation results
- Future directions

JET shot #42982



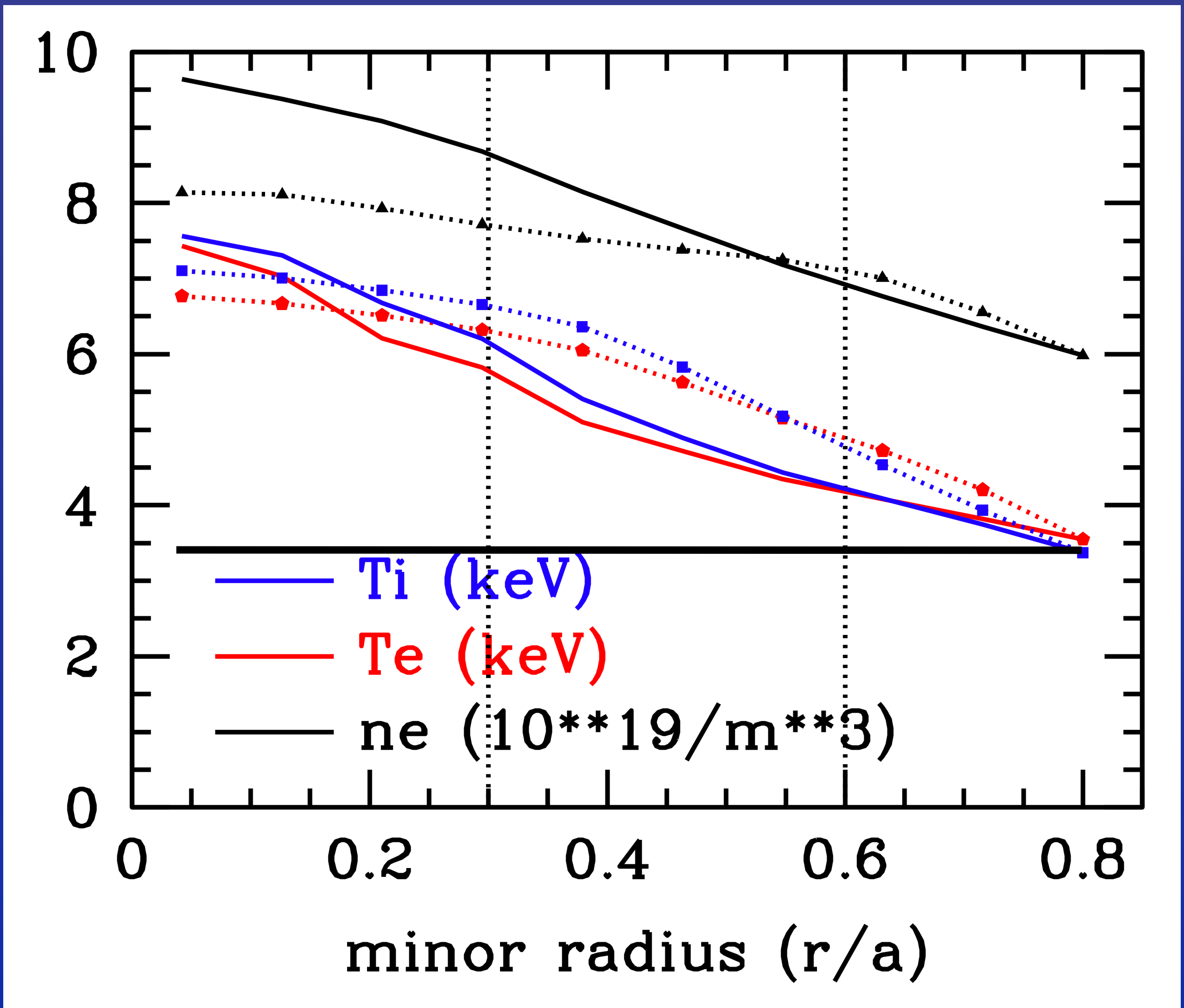
- ITER demo discharge
- H-mode D-T plasma, record fusion energy yield
- Miller local equilibrium model: q, shear, shaping
- B = 3.9 T on axis
- TRANSP fits to experimental data taken from ITER profile database

Evolving density profile



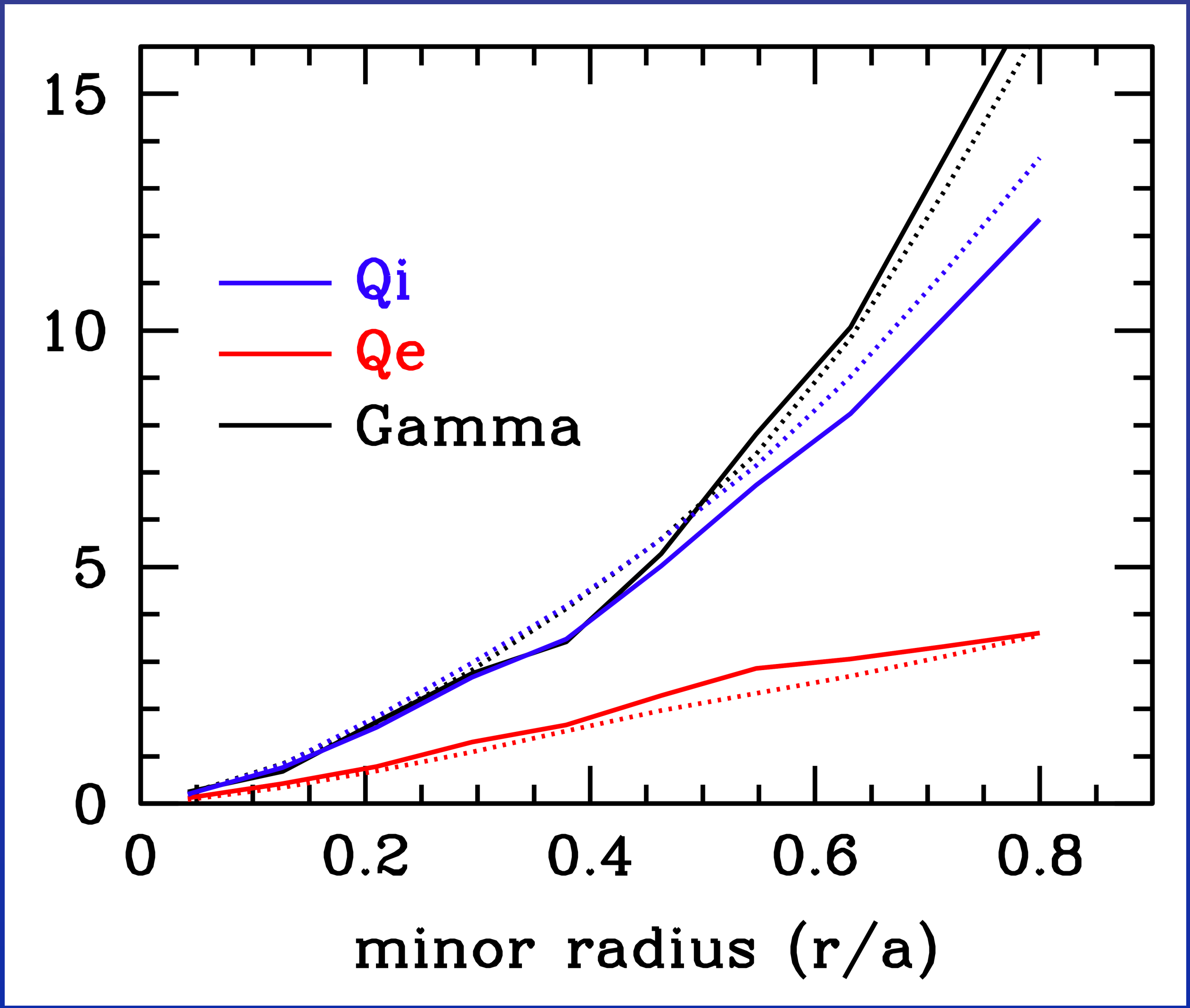
- 10 radial grid points
- Costs ~120k CPU hrs (<10 clock hrs)
- Dens and temp profiles agree within ~15% across device
- Energy off by 5%
- Incremental energy off by 15%
- Sources of discrepancy:
  - Large error bars
  - Flow shear absent

Evolving density profile



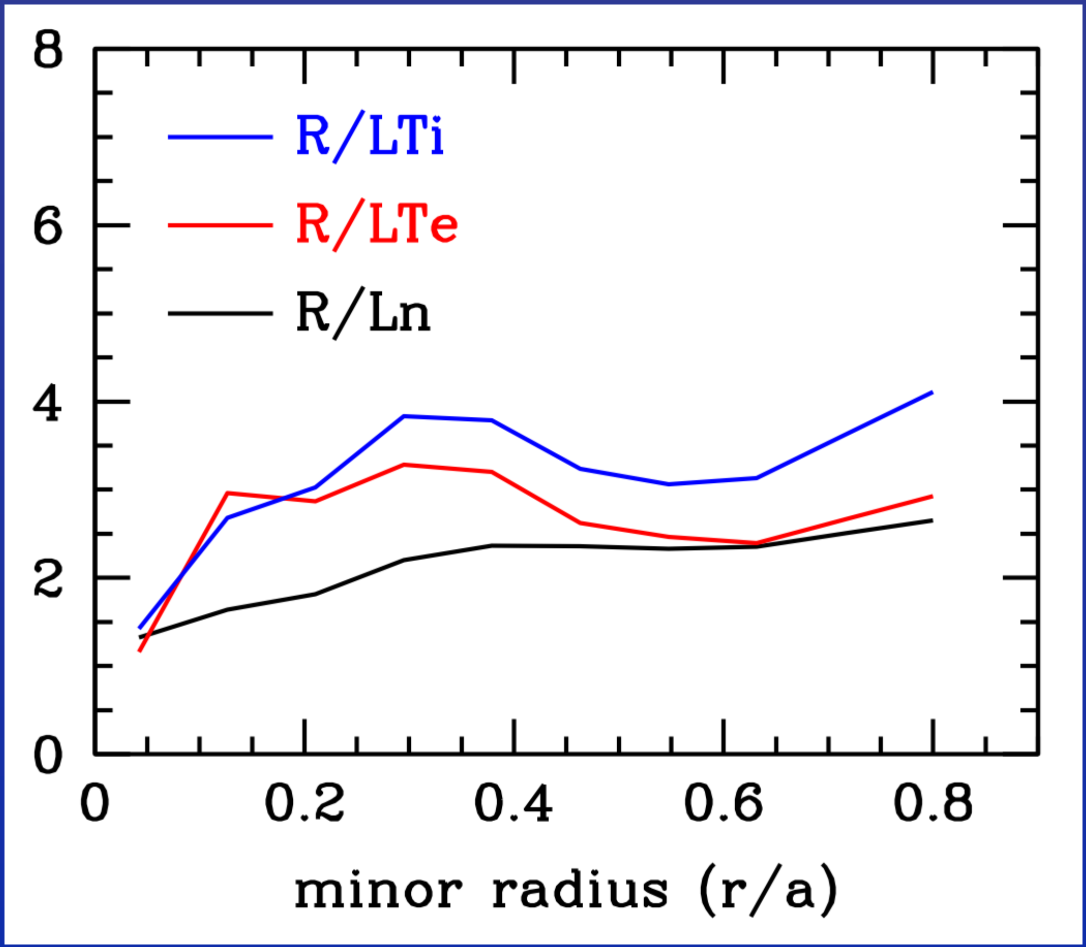
- 10 radial grid points
- Costs ~120k CPU hrs (<10 clock hrs)
- Dens and temp profiles agree within ~15% across device
- Energy off by 5%
- Incremental energy off by 15%
- Sources of discrepancy:
  - Large error bars
  - Flow shear absent

Power balance

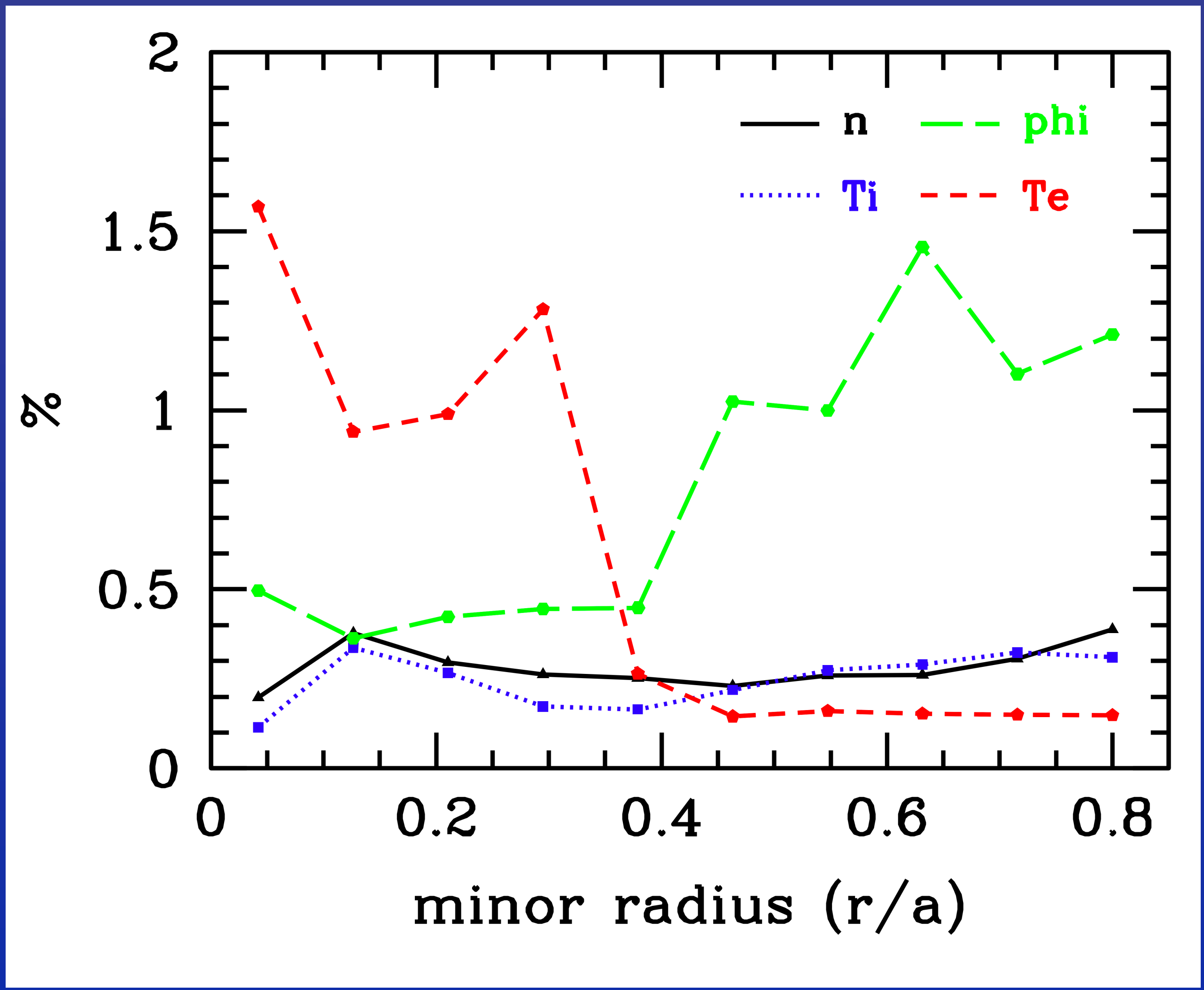


Profile stiffness

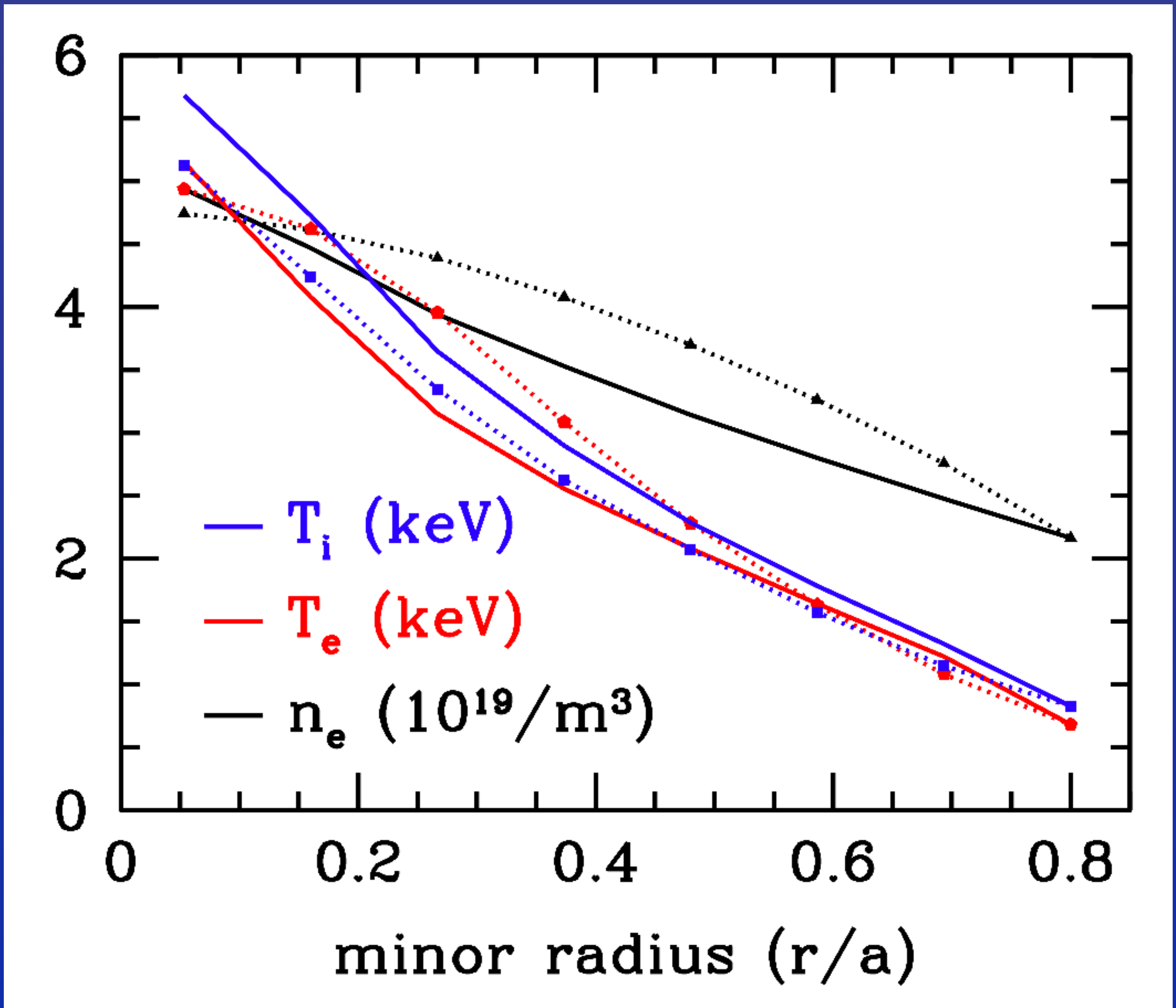
- ~ flat grad scale lengths indicative of stiffness (near critical gradient across most of minor radius)



Fluctuations

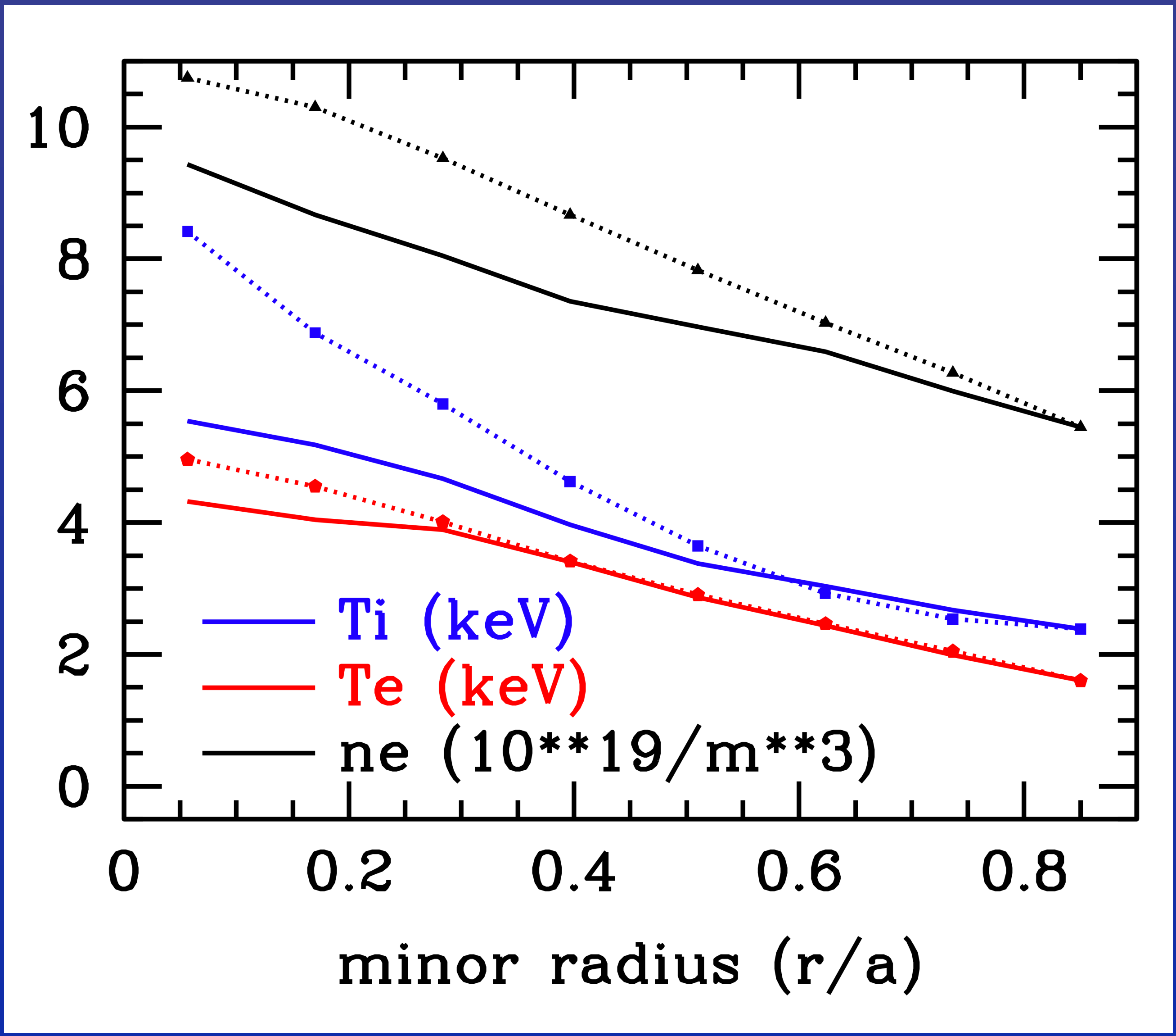


JET shot #19649



- L-mode discharge
- 8 radial grid points
- Costs ~25k CPU hrs (4 clock hrs)
- Flow shear absent

AUG shot #13151



- Fluxes calculated with GENE
- 8 radial grid points
- Costs ~400k CPU hrs (<24 clock hrs)
- Dens and electron temp profiles agree within ~10% across device
- Flow shear absent

Overview

- Motivation
- Theoretical framework
- Numerical approach
- Example simulation results
- Future directions

Future directions

- Summary: multi-scale approach saves factor of 10^3-10^4 and highly parallelizable
- Capture more physics
  - Magnetic equilibrium evolution
  - MHD stability
  - Momentum transport
  - Improved neoclassical model
- Improve convergence
  - Flux dependences on density, temperature, etc.
- Coupling to global gyrokinetic code (GENE)
  - Address meso-scale spatial structures
- Coupling to GPU-based gyrofluid code (GRYFFIN)
  - Entire calculation in minutes on several GPUs