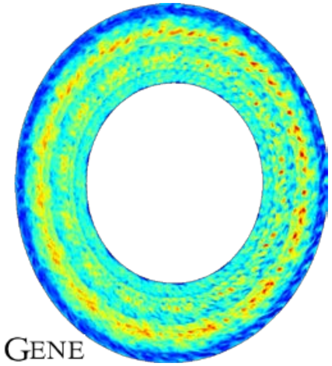
Circumventing the conflicting constraints of speed and accuracy for tokamak turbulence modelling

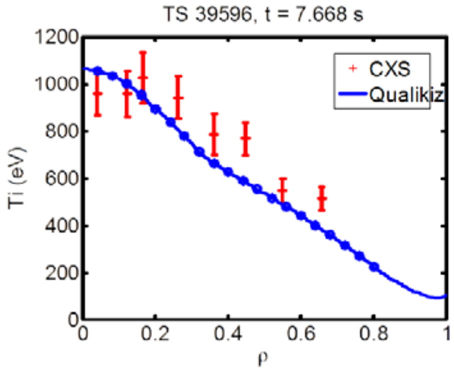
J. Citrin[1,2], J.F. Artaud[2], C. Bourdelle[2], S. Breton[2], Y. Camenen[3], F. Felici[4], J. Garcia[2], F. Imbeaux[2], P. Mantica[5], J. Redondo[2] and JET Contributors[6]

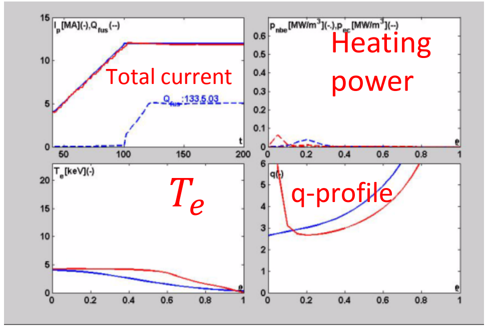
1 FOM Institute DIFFER: Dutch Institute for Fundamental Energy Research, Eindhoven, The Netherlands
2 CEA, IRFM, F-13108 Saint Paul Lez Durance, France
3 Aix-Marseille Université, CNRS, Marseille, France
4 Eindhoven University of Technology, Department of Mechanical Engineering, Eindhoven, The Netherlands
5 Istituto di Fisica del Plasma 'P. Caldirola', Associazione Euratom-ENEA-CNR, Milano, Italy
6 Eurofusion Consortium, Culham Science Centre, Abingdon, OX14 3DB, UK

Pathway towards unprecedented model tractability while remaining first-principle-based

Bridging 12 orders of magnitude in calculation speed







Local nonlinear gyrokinetics
10^8 CPUh for 1s JET-scale profile evolution

Reduced quasilinear model.
10^2 CPU hours for 1s JET-scale profile evolution

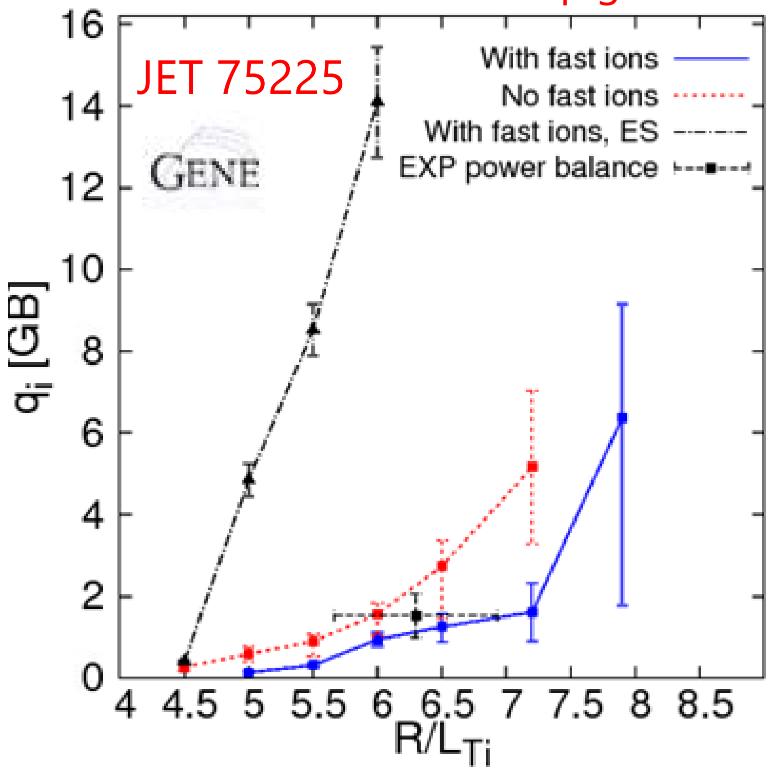
Realtime capability.
Neural network emulation technique.
Faster than realtime!

- Nonlinear predictions must be validated by experiments
- Reduced model must be verified by nonlinear, and validated by experiment
- Recent focus: neural network emulation of quasilinear transport models. Realtime capable. Powerful tool for experiment prediction and optimization

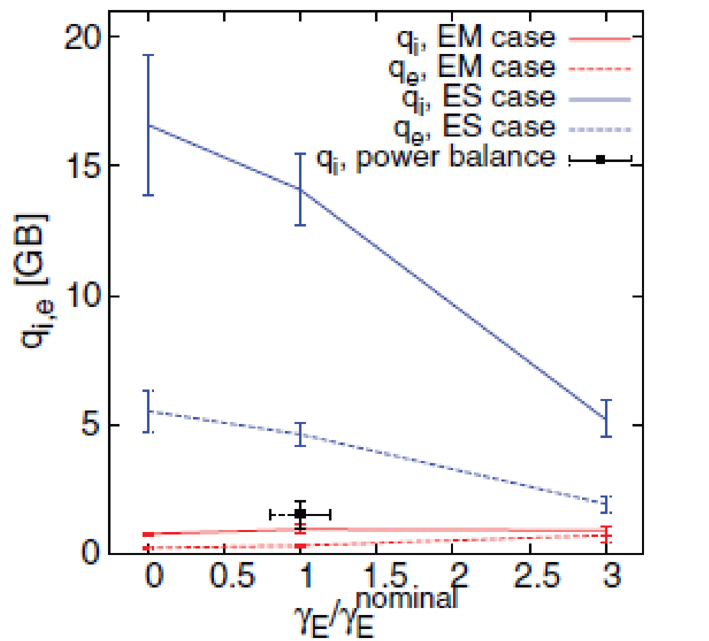
Increasing "routine" validation of nonlinear simulation

Continuous exploration of new regimes due to affordable computation → increasing agreement. Example: T_i peaking by electromagnetic stabilization of ITG in certain regimes. Previously thought to be due to rotation

Ion heat flux vs ion temp gradient



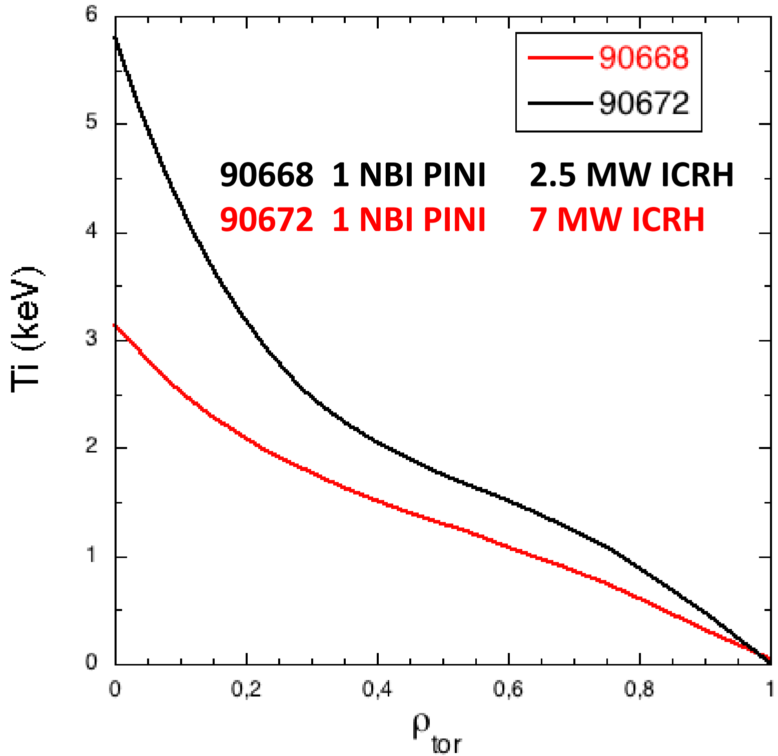
Weak impact of rotation in EM cases

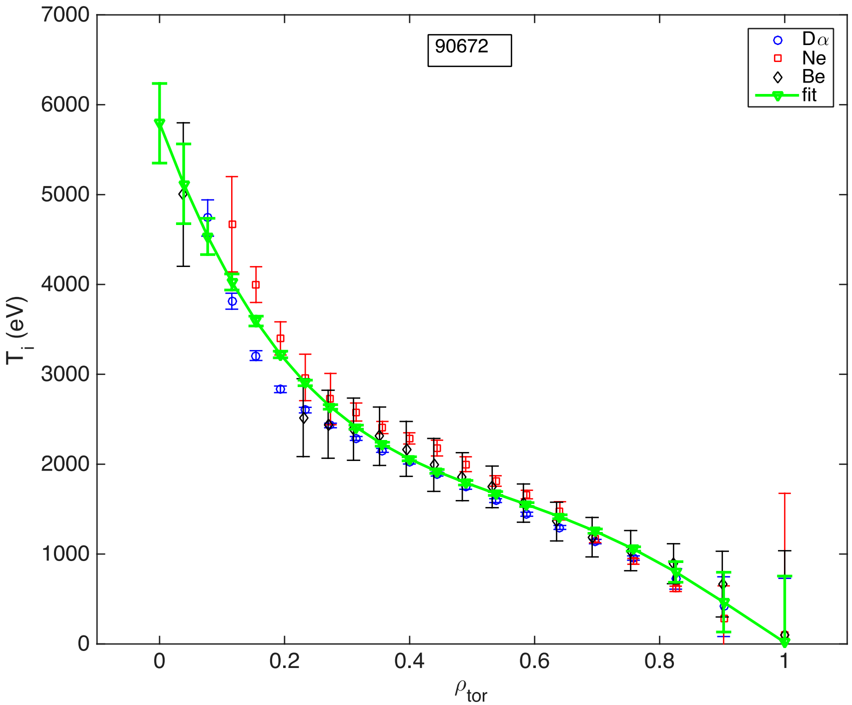


Mantica PRL 2011, Citrin PRL 2013, PPCF 2015. Garcia Nucl Fusion 2015 (JET C-wall)
Doerk accepted PPCF 2016 (JET-ILW, see next talk!)
Bravenec accepted PPCF 2016 (successful GYRO-GENE benchmark)

Predictions from nonlinear simulations has guided experimental design

Recent JET experiments designed to test prediction of strong EM-stabilization with fast ions also at low rotation





Strong peaking seen at high ICRH power and low rotation, consistent with prediction! Detailed GK simulations planned. See P. Mantica P.1.8 (this meeting)

Quasilinear modelling a significant acceleration compared to nonlinear

Fast reduced transport model QuaLiKiz: quasilinear gyrokinetic ITG/TEM/ETG heat, particle, and momentum turbulent transport [Bourdelle PPCF 2016]

10 CPUs per flux, $\times 10^6$ faster than nonlinear [Citrin EPS 2016, Varenna 2016]

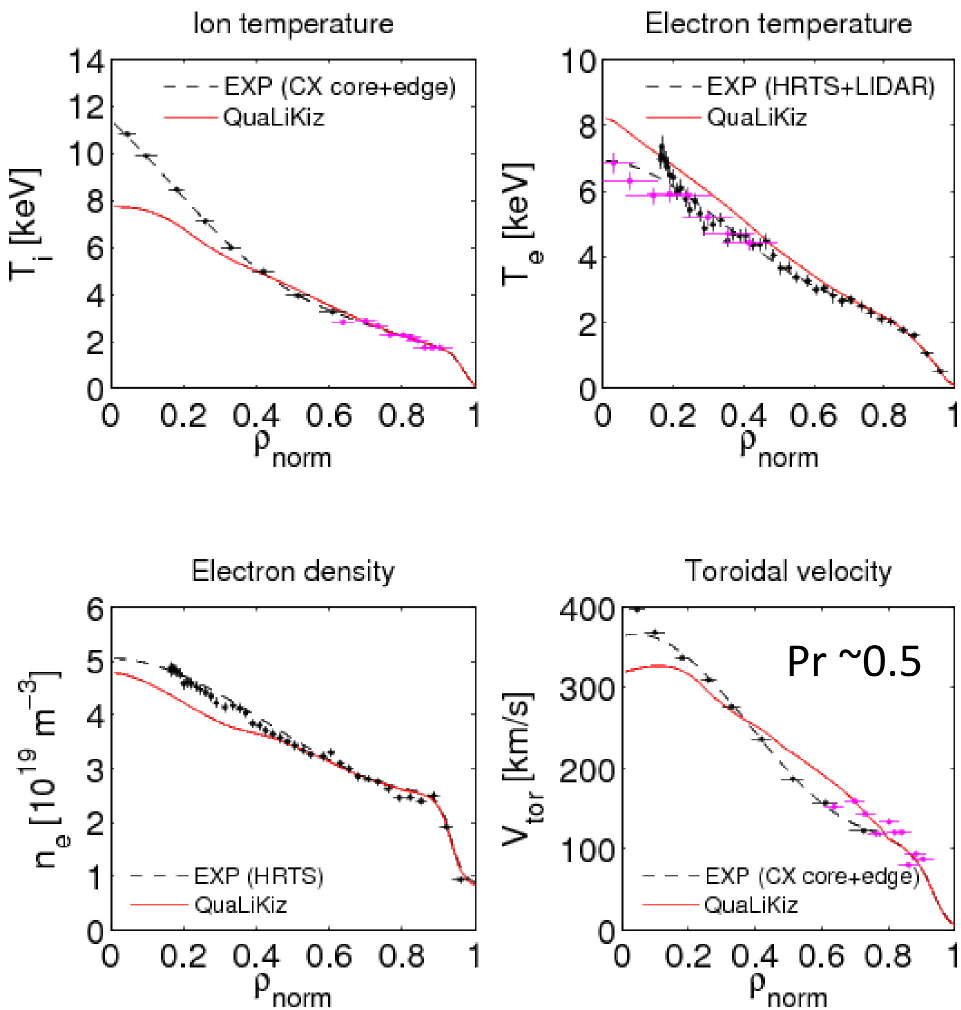
QuaLiKiz assumptions

- Ballooned Gaussian eigenfunction ansatz
- Shifted circle ($s - \alpha$) geometry
- Electrostatic only (nonlinear EM-stabilization effects to be added to nonlinear saturation rule)
- Collisions only with Krook operator for trapped electrons
- No $E \times B$ stabilization included for $\rho_{norm} < 0.5$ (leads to better agreement with nonlinear, likely due to underpredicted PVG destabilization)

Validation within JETTO-SANCO integrated modelling

Agreement excellent in ALL channels for ρ>0.5

First ever 4-channel QuaLiKiz integrated modelling simulation. ~100 CPUh



JET 75225. C-wall hybrid scenario (1.7MA/2T). Time window 6 - 7s

C impurity in SANCO → D and C modelled separately

Boundary condition at $\rho = 0.8$

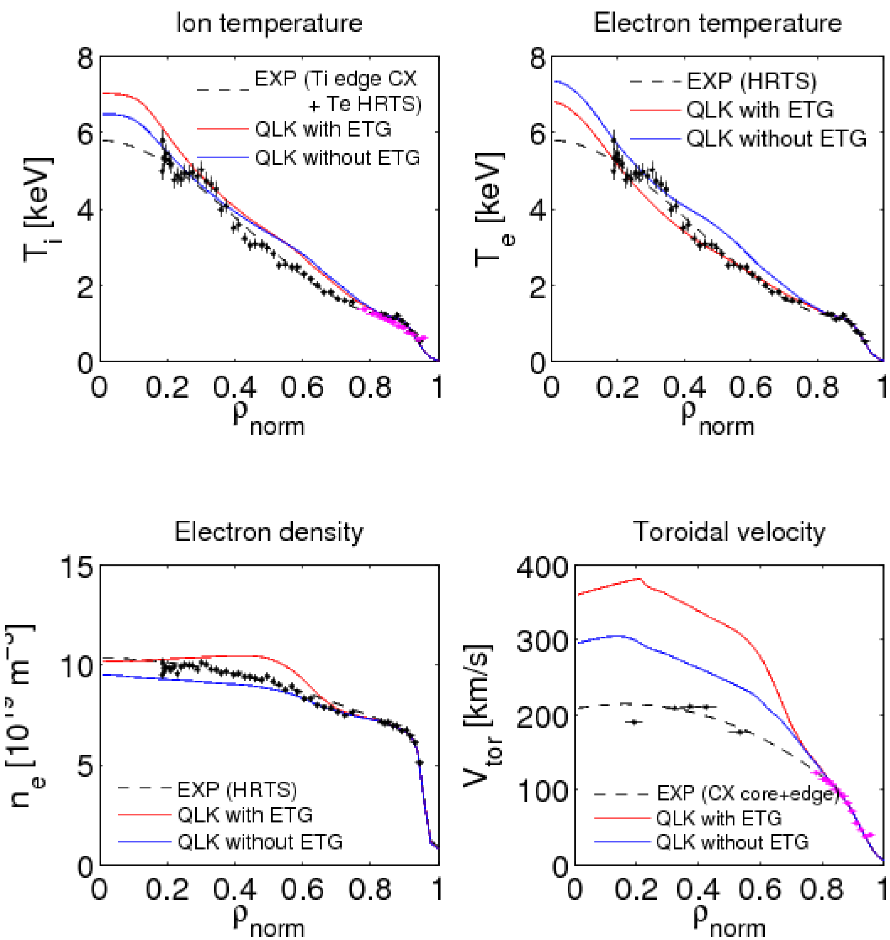
Includes rotation ($\rho > 0.5$) and momentum transport!

Agreement excellent in all channels for $\rho > 0.5$

For $\rho < 0.5$, Ti underprediction due to lack of EM effects in QLK, important for hybrid scenarios

QuaLiKiz validated also by comparison to a JET-ILW baseline scenario

Comparison with and without ETG-scales
Time window averaged between 10-10.5s



ILW baseline scenario
JET 87412 (3.5MA/3.35T)

Good agreement in
All channels apart from V_{tor}

- Boundary condition at $\rho = 0.85$
- Stable for $\rho < 0.2$. No sawtooth model
- Assuming core measurements T_i=T_e due to poor core CX
- NTV torque due to NTMs flatten profile? Quality of core CX for V_{tor}?

Ongoing work and further validations underway, including with W-transport (S. Breton, C. Bourdelle, F. Casson)

Neural networks can provide a further speedup in turbulence modelling

$10\ CPUs$ is fast, but we can go much further!

1. Obtain a good quasilinear model validated vs nonlinear simulations and experiments

2. Use quasilinear model to create datasets of turbulent flux calculations. Include all tokamak parameters of interest (e.g. based on experiments). Feasible with 10^7 CPUh scale HPC projects (currently ‘routine’)

3. Define ‘training sets’ from the database to construct nonlinear regressions of the data

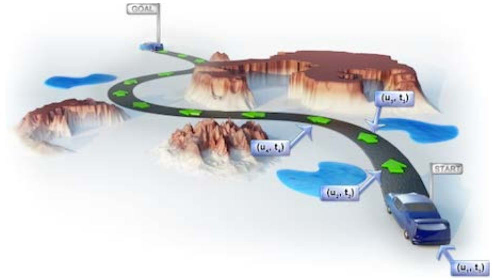
4. Use the fitted nonlinear regression as the ‘transport model’

The nonlinear regression technique we’ve been exploring is Multilayer Perceptron Neural Networks

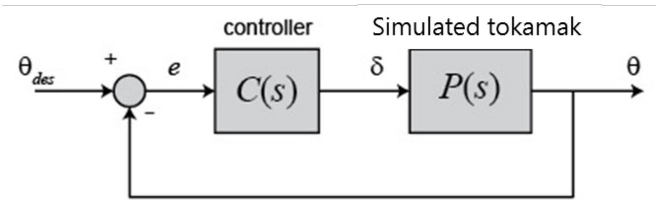
Similar work as NN fit to DIIID power balance database [Meneghini PoP 2014], NN fit to TGLF and EPED subsets [Staebler ITPA March 2016, Meneghini PoP 2016]

Neural network technique opens up wide applications for scenario prediction and control

“Very fast” tokamak simulator

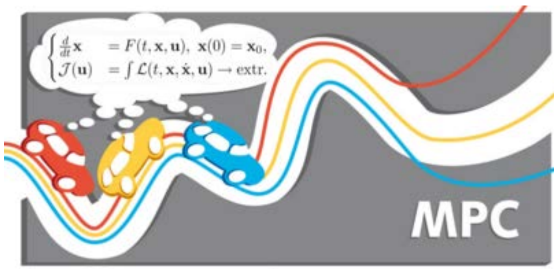


- Offline trajectory optimization
- Controller design
- Controller validation



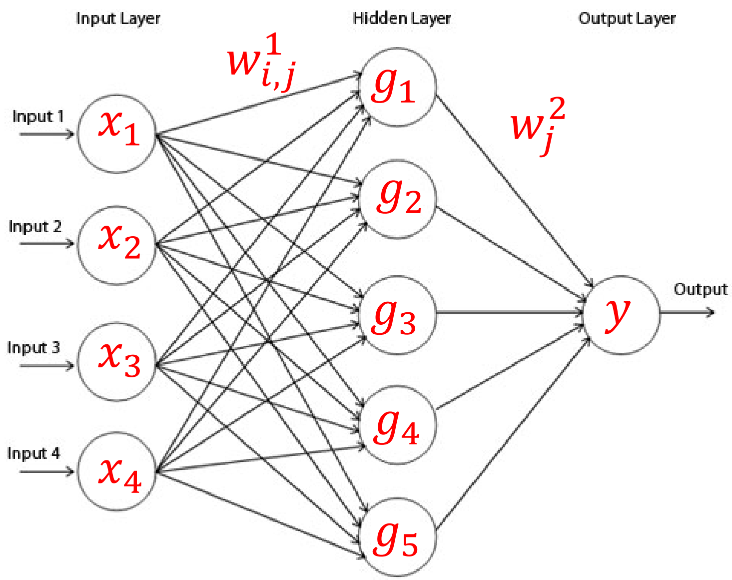
Realtime tokamak simulator

- Discharge supervision and monitoring (e.g. disruption mitigation)
- Online trajectory optimization faster-than-real-time (model-based predictive control)



Demystification of neural networks: just a tunable, general, nonlinear mapping

Multilayer perceptron neural network (simple topology)



x: Inputs: e.g. T_i/T_e, q, $\hat{s}$, R/L_{ti}
y: Output: e.g. ion heat flux
$w^{1,2}$: free weights for optimization

$$y = \sum w_j^2 g_j \left(\sum w_{i,j}^1 x_i \right)$$

With, e.g. $g(x) = \frac{2}{1+e^{-x}} - 1$

Optimize weights by minimizing: $\sum_N (t_N - y_N)^2 + \lambda \sum \left(w_{ij}\right)^2$

t_N are target values, known from, e.g. QuaLiKiz runs
λ is the regularization factor. Critical for avoiding overfitting

Neural network QuaLiKiz – a realtime first-principle-based transport model

Cover parameter space with the mapping

Local plasma parameters → QuaLiKiz → Heat/particle fluxes

1. Compile knowledge of mapping: code output
2. Train the free weights in the "neural network" to then emulate the mapping

Universal continuous function approximator (basic literature, Bishop 1995, Haykin 1999)

- Optimal weights found by optimization based on 'training set' of known mapping
  Extensive literature and libraries. Can use off-the-shelf techniques
- Can be used as a nonlinear regression technique to emulate transport models
- Trained network output in <ms, ~6 orders of magnitude less than original calc.
  circumvents conflicting constraints of model accuracy and tractability
- Provides an analytical formula with analytical derivatives.
  Critical for trajectory optimization applications and implicit timestep solvers

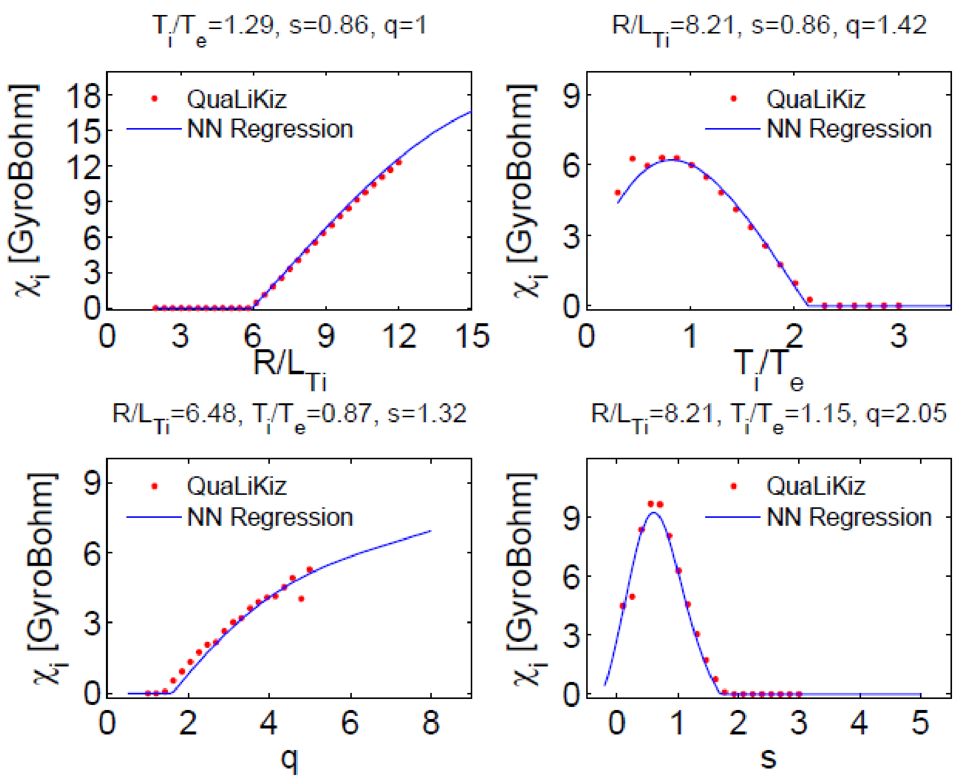
A proof-of-principle NN transport model has been developed

Neural network fit for QuaLiKiz output. ITG regime

(S.Breton MSc, J. Redondo MSc ; Citrin, Breton *et al.*, Nucl. Fusion Lett. 2015)

5D input training set for ~50,000 fluxes

$q = 1-5\,;\ \hat{s} = 0.1-3\,;\ \frac{T_i}{T_e} = 0.3-3\,;\ \frac{R}{L_{Ti}} = 2-12\,;\ k_\theta\rho_s = 0.05-0.8$ (ion scales)



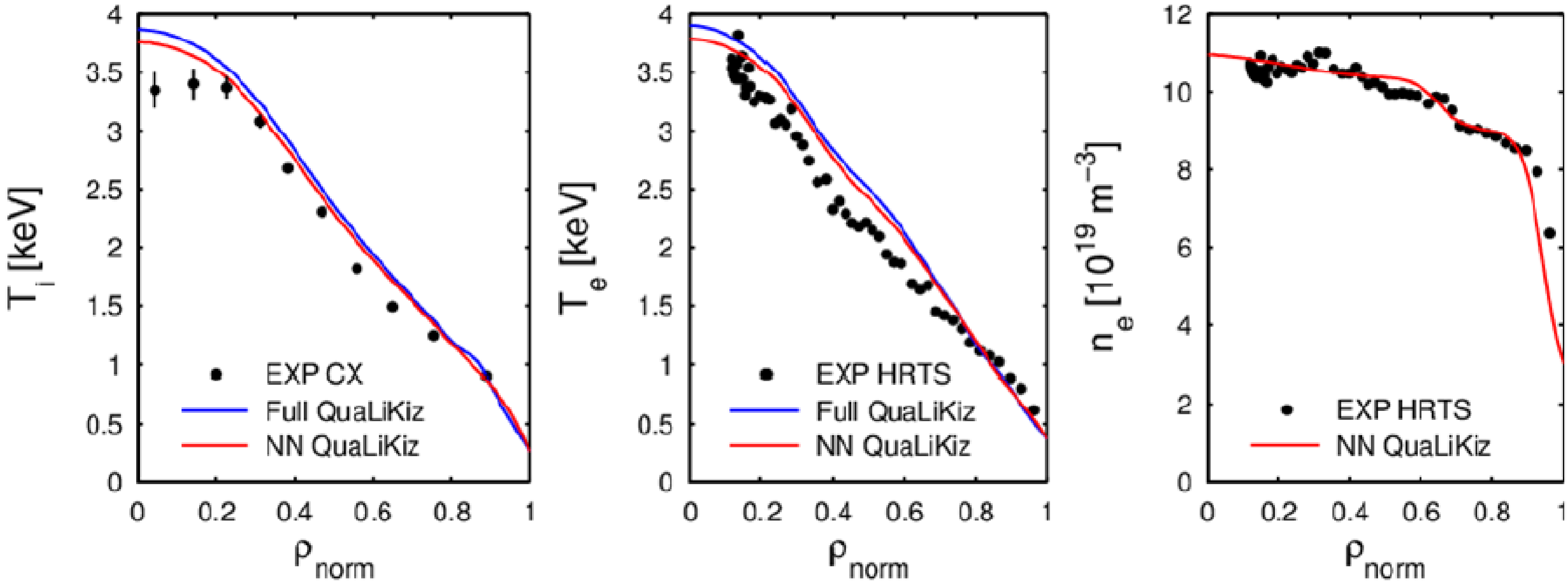
Parameter scans of NN ion heat conductivity vs original QuaLiKiz results

Note that regularization allows reasonable extrapolation.

Extrapolation not recommended, but encouraging for robustness in sparse datasets

Neural Network QuaLiKiz validated by JET discharge modelling

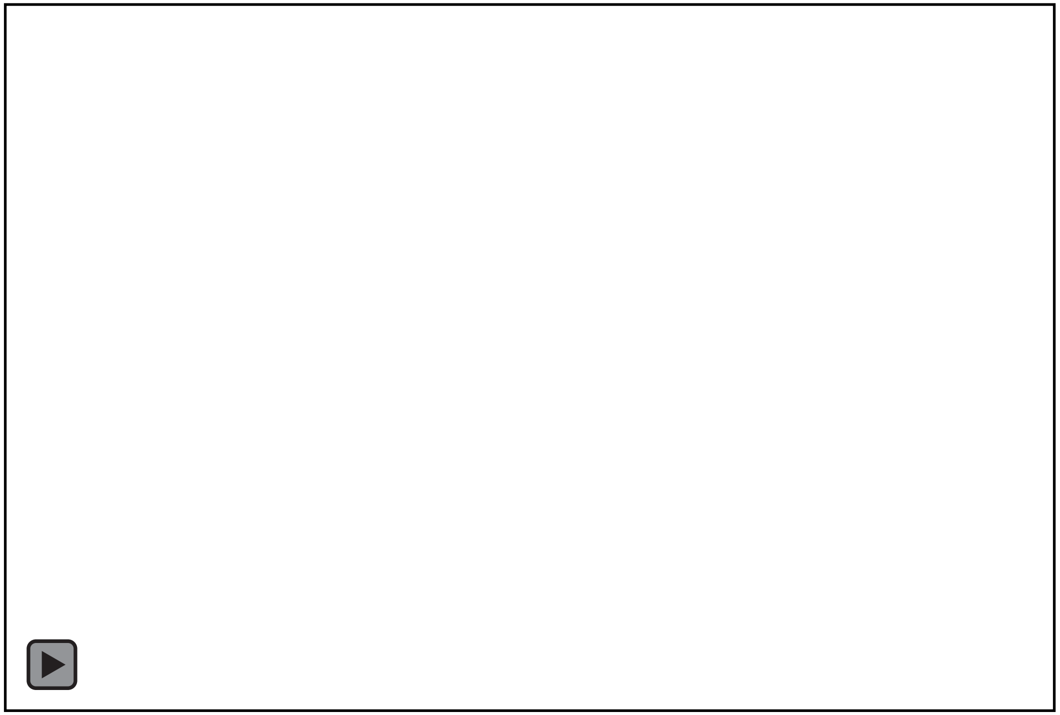
CRONOS/QLKNN simulation of flat top in JET 73342 standard H-mode.
Original QLK simulation in Baiocchi PPCF 2015. Boundary condition at $\rho = 0.88$



- Neural network successfully reproduces QuaLiKiz results $\times 10^6$ faster! ~1ms for a flux
- Already works well on JET ITG dominated case in flux driven integrated modelling

Faster than realtime predictions in the RAPTOR control-oriented simulator

- QuaLiKiz neural network applied in ultra-fast RAPTOR simulator (F. Felici et al 2012)
- Extrapolation to ITER hybrid scenario. QuaLiKiz NN compared with GLF23 modelling (Citrin NF 2010). GLF23 gyrofluid quasilinear model valid in ITG regime
- GLF23/CRONOS took 1 week for this calculation (on 1 CPU)
  QLKANN/RAPTOR took 8s to calculate 300 ITER seconds! Faster than realtime!



BLUE – QLK Neural Network. RED – GLF23

Next steps: increased input dimensionality for more general neural network emulation

Ongoing work (Karel van de Plassche MSc, Eindhoven):

10D QuaLiKiz database including ITG/TEM/ETG. ~1.5MCPUh

$$k_\theta \rho_s, \frac{R}{L_{Ti}}, \frac{R}{L_{Te}}, \frac{R}{L_{ne}}, q, \hat{s}, \epsilon = \frac{r}{R}, \frac{T_i}{T_e}, \nu^*, Z_{eff}$$

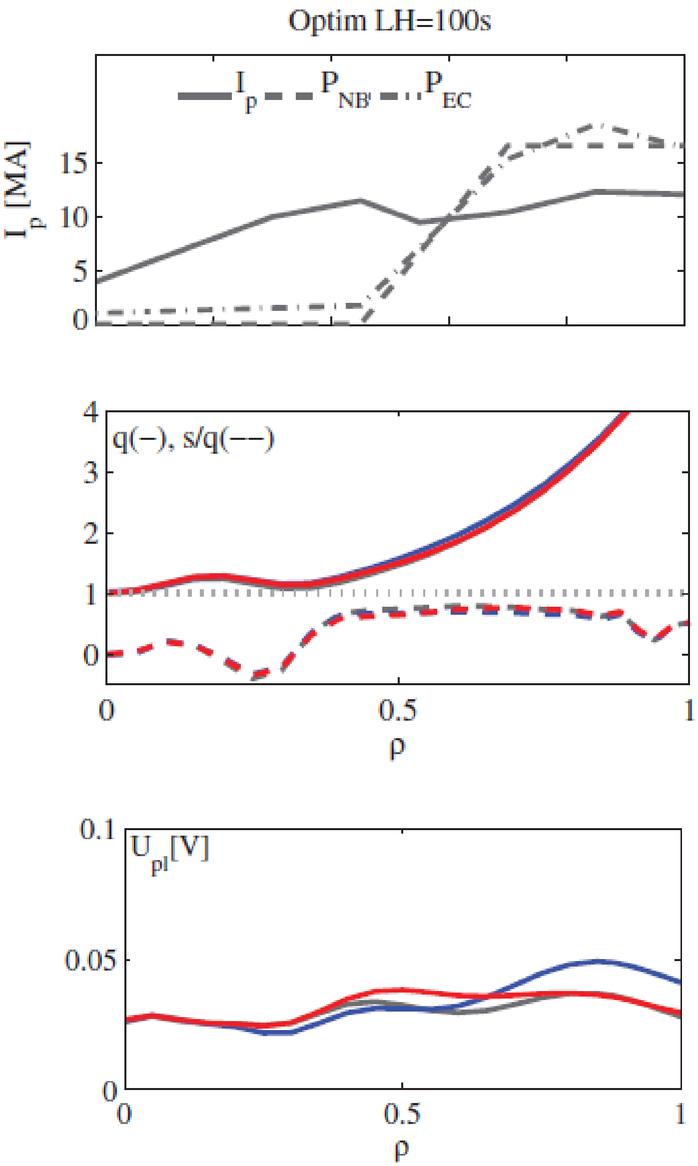
- Hypercube in experimentally relevant ranges
- Increased generality in neural network transport model

Upcoming work (Aaron Ho PhD, DIFFER, from Nov 2016)

~22D quasilinear database with GENE. ~10MCPUh

- Less approximations in linear solver (e.g. electromagnetic, shaping)
- Experimentally relevant subspace from multi-machine profile database
- Linear output database under construction to share with community

Potential applications for experimental design and optimization



Van Dongen, Felici PPCF 2014

- Automatic actuator trajectory optimization for ITER hybrid scenario within RAPTOR
- Sequential Quadratic Programming algorithm. Optimization of trajectory given cost functions (e.g. q-profile shape, flat loop voltage) and constraints (e.g. q>1)
- Algorithm “found” the current-overshoot as done at JET, without prior knowledge
- Needs many trajectories, simulation must be fast! Previously used Bohm-GyroBohm tuned to GLF23. Our NN transport model is a natural improvement

Similar work can be done to design current and future experiments

Summary

- Set out on a quest for simultaneous accuracy and tractability in turbulent transport models in tokamaks
- Key technique: neural network regression of a quasilinear reduced model, itself validated by nonlinear gyrokinetics
- Proof-of-principle 5D network created and applied in transport models. Validated by JET discharge in ITG regime. Can model faster than real-time
- Ongoing generalizations of the neural network model to higher input dimensions
- Technique not limited to turbulence modelling. Any code where full mapping in relevant parameter space in $<O(10^8)$ CPUh may be "neural networked". e.g. neural network EPED (Meneghini et al) Pathway to realtime tokamak simulator