WPSA General Meeting - Integrated Data Analysis and Validation

R. Fischer[1], A. Bock[1], S.S. Denk[2], D. Stieglitz[1]
[1]Max Planck Institute for Plasma Physics, Garching, Germany
[2]Plasma Science and Fusion Center, MIT, Cambridge, USA

IDA for Nuclear Fusion

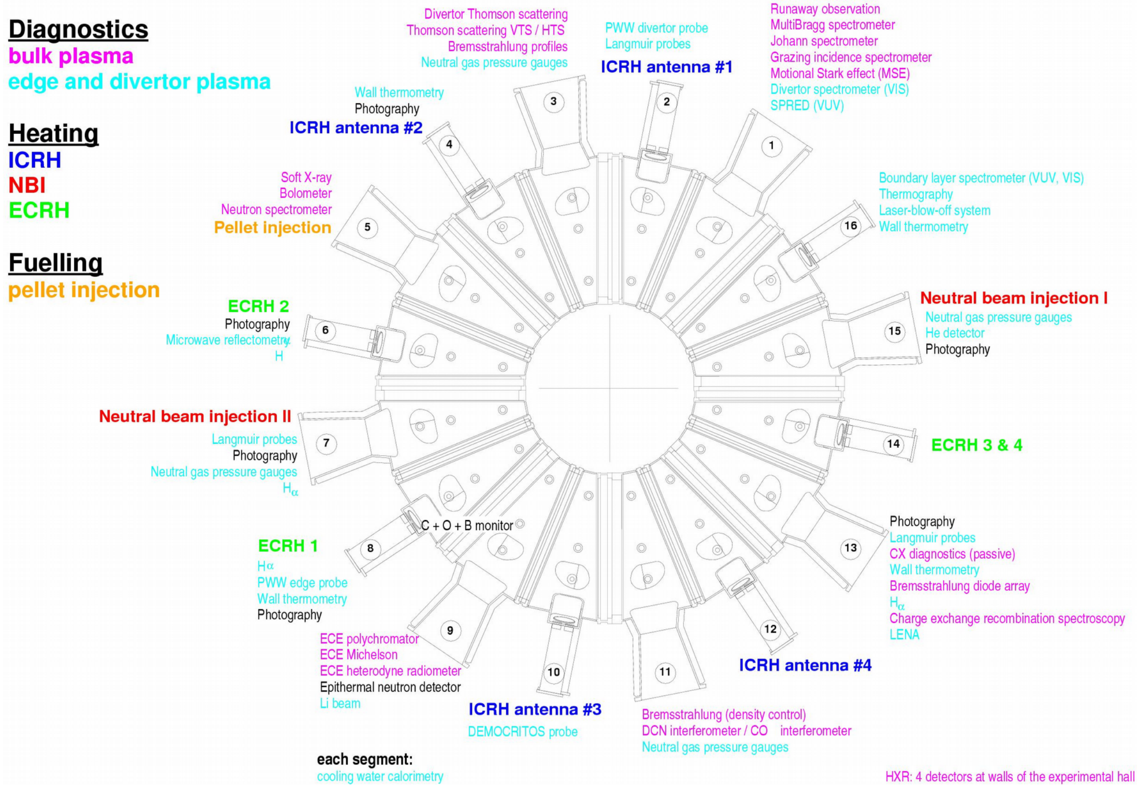
Scientific exploitation of fusion devices crucially depends on the reliability of measurements

Different measurement techniques for the same quantities → redundant and complementary data

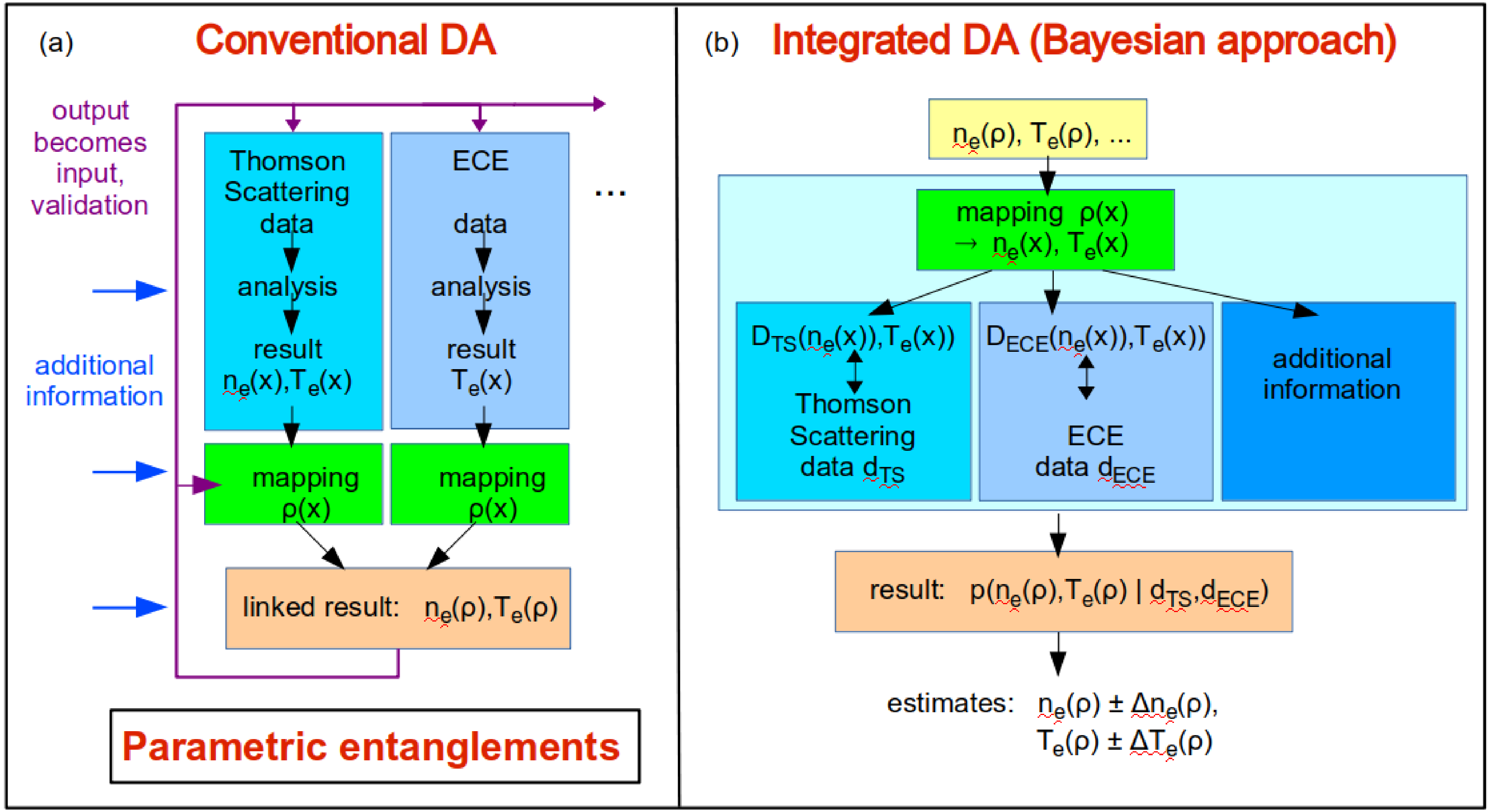
Coherent combination of measurements from different diagnostics

Goal:

- replace combination of results from individual diagnostics
- with combination of measured data
  → one-step analysis of pooled data



Conventional vs. Integrated Data Analysis



Conventional vs. Integrated Data Analysis (cont.)

Drawbacks of conventional data analysis: iterative

- (self-)consistent results? (cumbersome; do they exist?)
- difficult to be automated (huge amount of data from steady state devices: ITER, ...)
- information propagation? (Single estimates as input for analysis of other diagnostics?)
- error propagation? (frequently neglected: underestimation of the uncertainty)
- data and result validation? (How to deal with inconsistencies?)
- often backward inversion techniques (noise fitting? numerical stability?)
- result: estimates and error bars (sufficient? non-linear dependencies?)

Probabilistic combination of different diagnostics (IDA)

✔ uses only forward modeling (complete set of parameters → modeling of measured data)
✔ additional physical information easily to be integrated
✔ systematic effects (inconsistency) → describe with (nuisance) parameters; identify and resolve inconsistencies
✔ unified error interpretation → Bayesian Probability Theory
✔ result: probability distribution of parameters of interest incl. all dependencies

IDA offers a unified way
of combining data (information) from various experiments (sources)
to obtain improved results

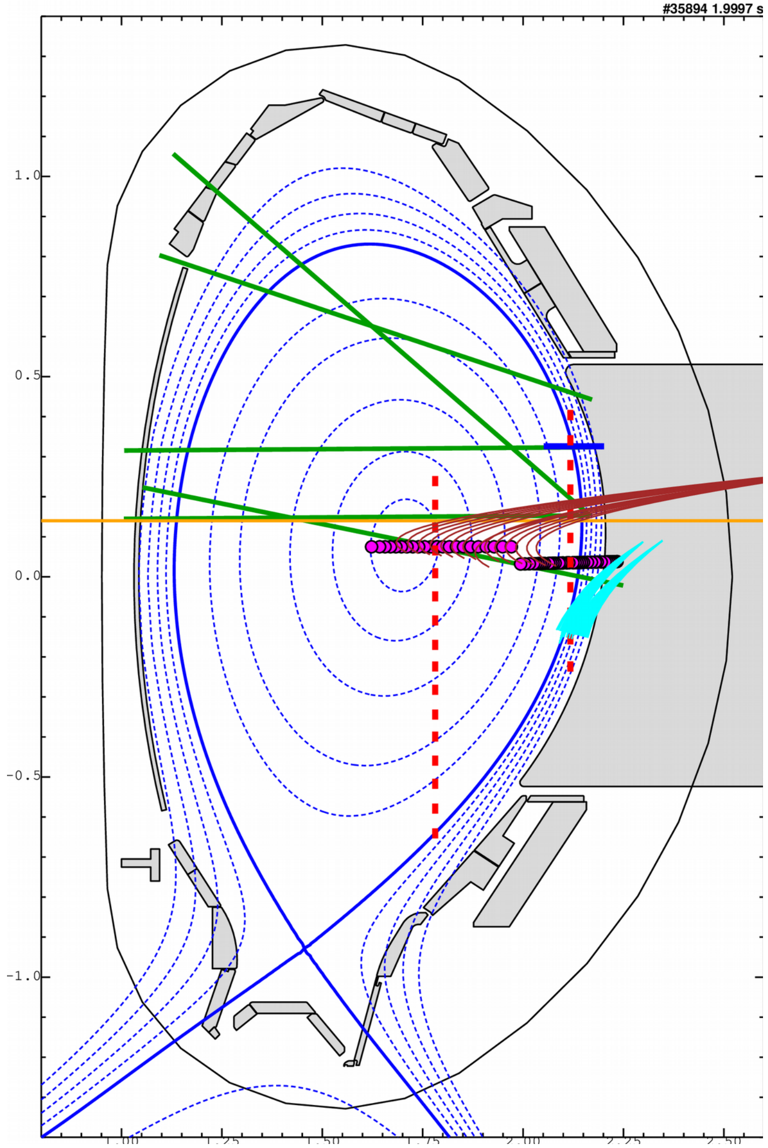
IDA at ASDEX Upgrade

multi-diagnostic profile reconstruction: n_e , T_e

- Lithium beam impact excitation spectroscopy (LIB)
  collisional radiative model → $n_e\ (T_e)$
- Interferometry measurements (DCN) → n_e
- Electron cyclotron emission (ECE)
  ECRad: Electron cyclotron radiation transport → $T_e\ (n_e)$
- Thomson scattering (TS) → n_e , T_e
- Reflectometry → n_e
- Beam emission spectroscopy → $n_e\ (Z_{eff})$
- Thermal Helium beam spectroscopy → n_e , T_e

- Equilibrium reconstructions for diagnostics mapping
  (*IDE*: kinetic Grad-Shafranov solution coupled with current diffusion)

A lot of dependencies and uncertainties:
We need a probabilistic approach!

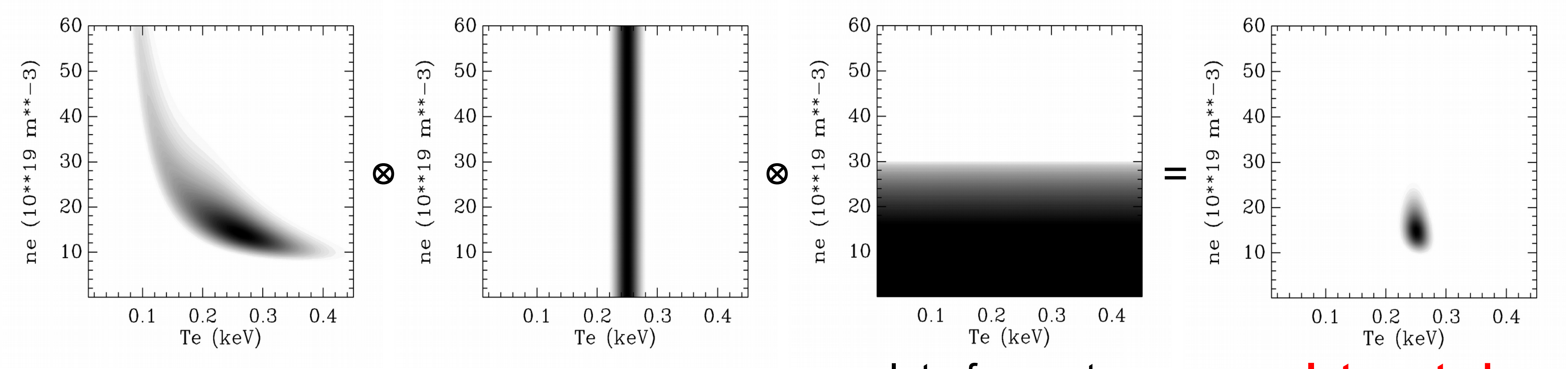


IDA: Probabilistic approach

Combination of a *set* of diagnostics

- n_e … electron density
- T_e … electron temperature

R. Fischer, A. Dinklage, E. Pasch, PPCF 2002, PPCF 2003



Thomson Scattering

Soft-X-ray

Interferometer Operation

Integrated Result

Probabilistic framework

IDA for Nuclear Fusion

Bring together different diagnostics/diagnosticians/theoreticians with redundant/complementary/modeling data

- Probabilistic modeling of individual diagnostics (forward models, likelihoods for all kind of uncertainties)
- Probabilistic combination of different diagnostics (multiply pdfs, unified error analysis, error propagation)
- Probabilistic combination with prior / modeling information

- Redundant data:
  - more reliable results by larger (meta-) data set → reduction of estimation uncertainties
  - detect and resolve data inconsistencies (reliable/consistent diagnostics) using standardized error/uncertainty treatment

- Complementary data:
  - resolve parametric entanglement
  - resolve complex error propagation (non-Gaussian)
  - synergistic effects (exploiting full probabilistic correlation)
  - automatic *in-situ* and *in-vivo* calibration (transient effects, degradation, …)
  - advanced data analysis technique

    → improvements in modeling (ECE) and diagnostics hardware (LIB)

Further Applications of IDA

W7-AS: n_e , T_e : TS, interferometry, soft X-ray

R. Fischer et al., PPCF, 45, 1095-1111 (2003)

W7-X: non-Maxwellian electron energy distribution function: visible emission spectrum

n_e , $T_{e/i}$, impurity densities, flows: TS, X-ray imaging

n_e , T_e : TS, interferometry, helium beam

Z_{eff} : bremsstrahlung spectra

D. Dodt et al., J. Phys. D: Appl. Phys., 41:205207, 2008.
A. Langenberg et al., RSI, 90(6), 063505 (2019)
S. Kwak et al., arXiv:2103.07582, 2021
S. Kwak et al., RSI, 92:043505 (2021)

MST RFP: T_e : TS, soft X-ray

Z_{eff} : soft X-ray, CXRS

L. M. Reusch et al., RSI, 85:11D844, 2014.
M.E. Galante et al., NF, 55:123016, 2015.

TJ-II: n_e , T_e : TS, interferometry, reflectometry, Helium beam

B. Ph. van Milligen, et al., RSI 82, 073503 (2011)

JET: n_e : LIB

n_e , T_e : LIDAR, interferometry

fast-ion distributions : velocity-space tomography of fast-ion D-alpha spectroscopy, collective TS, gamma-ray and neutron emission spectrometry, and neutral particle analyzers.

D. Dodt, et al., P-2.148, EPS 2009
O. Ford, et al., P-2.150, EPS 2009
M. Salewski et al., FST 74:23–36, 2018.

ITPA Diagnostic TG: SWG IDAV

Integrated Data Analysis and Validation (IDAV) specialist working group

- founded: 01.01.2020
- Rainer Fischer (chair, IPP Garching, Germany)
- Keisuke Fujii (co-chair, Kyoto University, Japan)
- Simon Pinches (co-chair, IO)
- Topic:
  - ITER provides many measurements coming from multiple diagnostics systems and this information will be used for machine control and safety as well as physics studies.
  - SWG working to develop self-consistent data validation procedures (profiles, error bar estimates…) for key ITER measurement parameters (density, ion and electron temperature, current and q-profile, Zeff, wall condition, etc.)
  - integrating all measurements to optimize information available for ITER operations, control and safety.
- started developing a modular IDA tool box for present and next generation fusion devices using a modern language

IDA Basic Implementation for ITER, JT-60SA, ...

Integrated Data Analysis: basic implementation in python

- open source license
- working implementation on https://git.iter.org
  - being completely modular
  - to be compatible with any fusion device (ITER:IMAS, ...)
  - diagnostics: Thomson scattering, ECE and interferometry, ...
  - likelihoods (data uncertainty): Gaussian, Cauchy (outlier robust), …
  - multi-fidelity forward models / synthetic diagnostics
    - ECE: $T_{rad} = T_e$ vs radiation transport modeling $T_{rad}(T_e, n_e)$
    - real-time vs offline analysis
  - flexible parameterisation of, e.g., profiles: splines, GPR, ...
  - priors: smoothness, positivity, physical modeling, ...
  - results and their uncertainties:
    - MAP solution (probability maximum and width)
    - MCMC sampling methods (explore full probability space)

```
.
├── run_ida_sample.py
└── lib
    ├── core
    │   ├── likelihoods
    │   ├── optimizers
    │   └── IDA
    └── diagnostics
        ├── base
        ├── aug
        │   ├── ECE
        │   ├── INTF
        │   └── TS
        ├── iter
        └── jt60sa
```

IDA for JT-60SA: Status, plan and goal

Present status:

1) The adaption of a generic IDA code package to JT-60SA diagnostics has been specified, starting with commissioning diagnostics.
2) Continuation of IDA(ITER) code development (ITER IMAS, diagnostics forward modeling)

Plan (to be agreed):

1) start with commissioning diagnostics (PO-1):
   - interferometry → n_e
   - soft-X ray → $T_e(n_e,Z_{eff})$
   - visible spectroscopy → $Z_{eff}(n_e,T_e)$
2) augment with PO-2 synthetic diagnostics:
   - Thomson scattering → n_e,T_e
   - ECE → $T_e(n_e)$

Benefits of IDA approach:

1) same IDA as for ITER
2) mutual development for various devices
3) mutual development for similar diagnostics

4) diagnostics inter-dependencies resolved
5) probabilistic parameter space exploration (MCMC)
   → to characterize diagnostics to be commissioned
6) unified uncertainty quantification of data and parameters
7) addtl. information: positivity (n_e, T_e), Z_{eff}>=1, modeling, ...
8) (nuisance) parameters, e.g. calibration and uncertainty
9) easily to be augmented with further diagnostics...

Goal: IDA for physics exploitation in 2025

IDA for JT-60SA: Commissioning diagnostics

1) n_e : 2-colour CO_2 laser interferometry $\int n_e \, dl$

and polarimeter $\sim \lambda^2 \int n_e B_{t,\parallel} \, dl$

2) $T_e(n_e, Z_{eff})$: Line-integrated soft-X ray bremsstrahlung intensity $\sim \int n_e^2 \sqrt{T_e} Z_{eff} \, dl$

3) $Z_{eff}(n_e, T_e)$: Visible spectroscopy, bremsstrahlung intensity $\sim \int g_{ff}(Z_{eff}, T_e, \lambda) \frac{n_e^2 Z_{eff}}{\sqrt{T_e} \lambda^2} \exp(hc/\lambda T_e) \, dl$

Parametric entanglements

IDA for JT-60SA: Technicalities

IDA ingredients (QST support needed):

1) measured data and an exploration and quantification of statistical and systematic errors (calibration)
2) physical (forward) model for the data/diagnostic
3) outlier robust method optionally (likelihood)
4) prior/modeling information

IT:

1) Platform: EURO*fusion* Gateway cluster or QST server at Naka
2) preferred data-dictionary and access-layer? (IMAS is already implemented)
3) python infrastructure

Proposed collaboration:

1) Contact persons: H. Tojo (QST), K. Fujii (Kyoto Univ.), R. Fischer (IPP), D. Stieglitz (IPP), G. Falchetto (CEA)
   (1st meeting 10 Dez 2020)
2) Manning (IPP): 12pm/year (postdoc);
3) Visits: short visit 1-2 weeks in the initial phase, longer visit possible to be discussed at a later stage