Tomography of the Fast-Ion Velocity Distribution in Tokamak Plasmas with a Physics-Based Prior

Per Christian Hansen
Technical University of Denmark

Main ingredients

1. A bit of physics
2. The inverse problem
3. Revisiting the standard-form transformation

Joint work with

← Mirko Salewski
Assoc. Prof., DTU Physics

← Bo Simmendefeldt
PhD student , DTU Physics

Thanks to Michiel E. Hochstenbach, TU/e and Martin S. Andersen, DTU Compute

Please don't ask me about the physics!

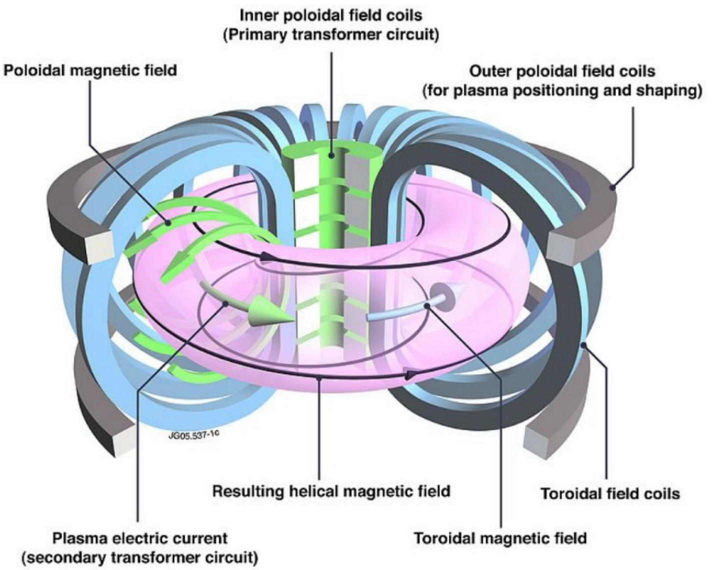
DTU Compute
Department of Applied Mathematics and Computer Science

The Tokamak Reactor

A **tokamak** uses a magnetic field to confine *plasma*, with a temperature of several million °C, in the shape of a torus.

The plasma typically consists of *deuterium* (D, a hydrogen isotope).

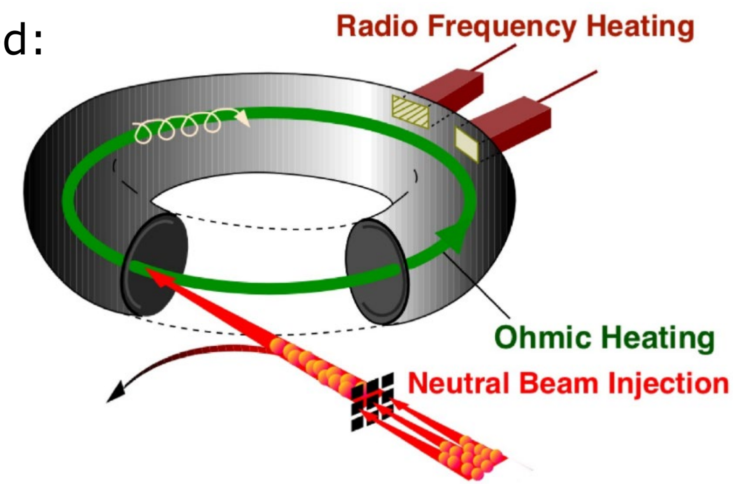
The energy produced through *fusion* is absorbed as heat in the walls. This heat produces *electricity* by way of turbines and generators.



Images: www.euro-fusion.org

To reach fusion conditions the plasma is heated:

- by launching radio waves into the plasma,
- by Ohmic heating (the plasma behaves as a current),
- by injecting neutral particles (i.e., particles with no electric charge) into the plasma.

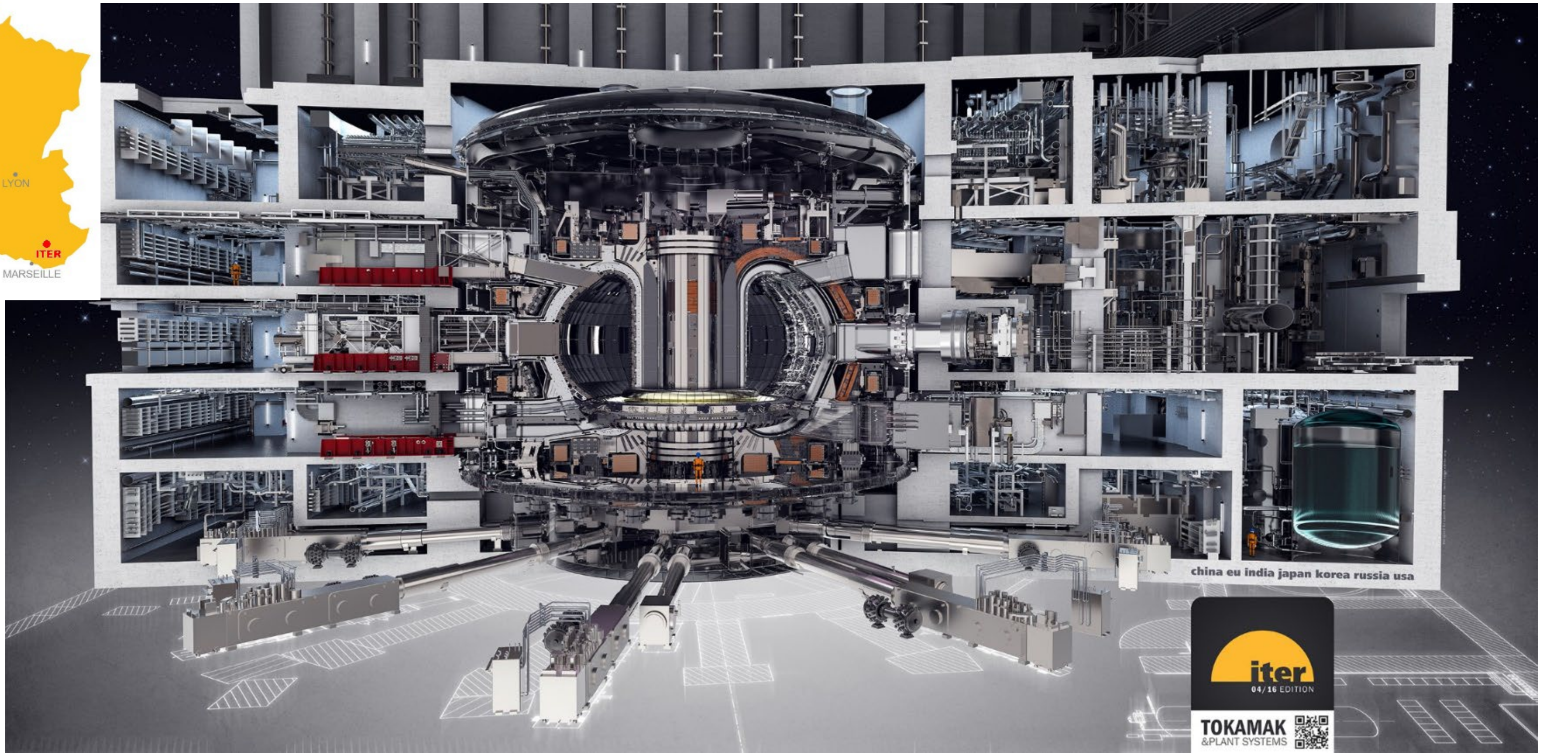


ITER

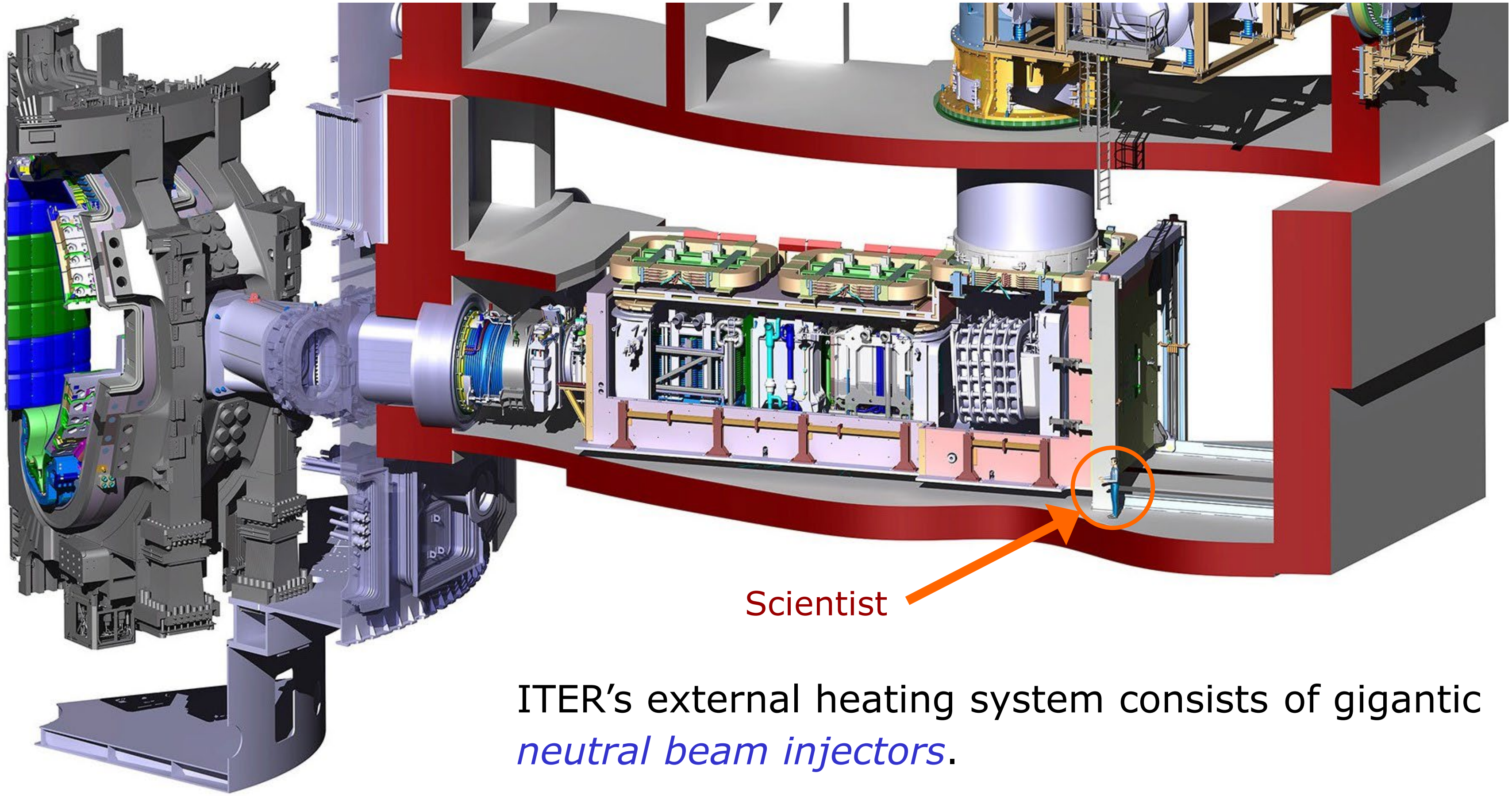
International Thermonuclear Experimental Reactor

Planned for late 2025 (?)

China, EU, India, Japan, Russia, South Korea, USA



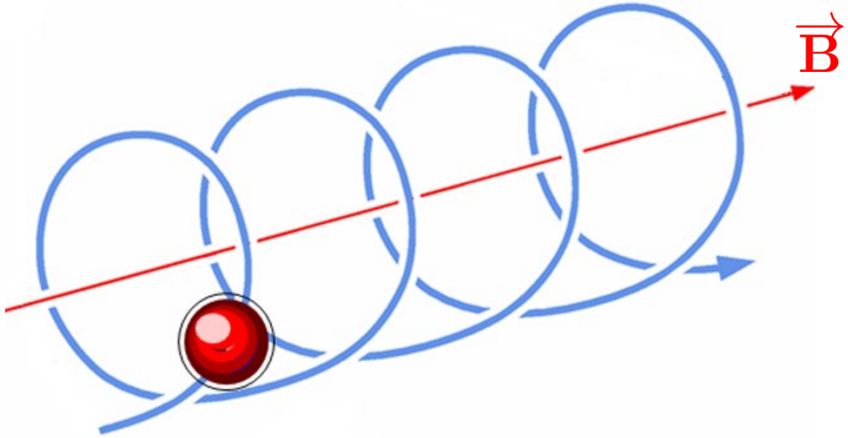
Neutral-Beam Injector (NBI)



ITER’s external heating system consists of gigantic *neutral beam injectors*.

They shoot high-energy electrically neutral atoms, “neutrals,” into the plasma – thus transferring energy to the plasma particles.

Distribution of Plasma Velocities



Locally, the magnetic field is uniform and each *deuterium ion* in the plasma moves in a helix along the magnetic field **B**:

- $v_{||}$ denotes the velocity along the field.
- $v_{\perp}$ denotes the tangential speed (the radial speed is zero).

We want to know the probability distribution $f(v_{||}, v_{\perp})$.

Towards the Inverse Problem

When fast plasma ions pass through the neutral beam, they undergo a reaction that creates a newly-born fast neutral and a photon.

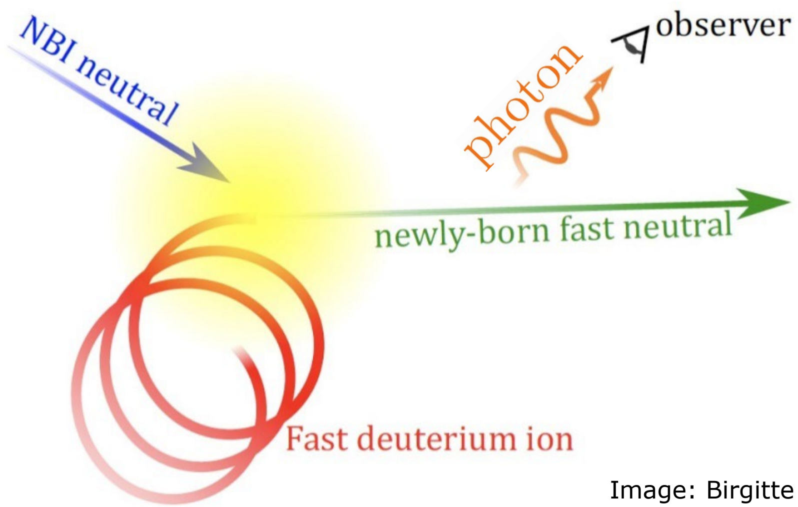


Image: Birgitte Madsen

The photons are *Doppler-shifted* depending on the velocity of the fast ion.

The inverse problem:
By measuring the Doppler-shift, we can infer about the *velocity of the fast ion*.

The Inverse Problem, Part I

On a spectral detector, we measure the *intensity per wavelength* $I(\lambda, \phi)$ of the photons, as a function of their wavelength λ.

This intensity also depends on the angle ϕ between the magnetic field $\mathbf{B}$ and the line-of-sight $\mathbf{u}$ to the photon detector.

The measurements are related to the velocity distribution $f(v_\parallel, v_\perp)$ via the integral

$$I(\lambda, \phi) = \int_0^\infty \int_{-\infty}^\infty k(\lambda, \phi; v_\parallel, v_\perp)\, f(v_\parallel, v_\perp)\, \mathrm{d}v_\parallel\, \mathrm{d}v_\perp \ ,$$

in which the kernel is given by

$$\begin{aligned} k(\lambda, \phi; v_\parallel, v_\perp) &= R(v_\parallel, v_\perp)\, \pi(\lambda \,|\, \phi, v_\parallel, v_\perp) \\ R(v_\parallel, v_\perp) &= \text{total intensity for any wavelength} \\ \pi(\lambda \,|\, \phi, v_\parallel, v_\perp) &= \text{probability density function for } \lambda \end{aligned}$$

For simplicity, we consider R a constant – leading to a simpler “proxi model.”

The Inverse Problem, Part II

Recall there is a Dopple shift, which we can write

$$\lambda - \lambda_{\mathrm{D}} = \frac{u\,\lambda_{\mathrm{D}}}{c} \qquad \Leftrightarrow \qquad u = c\,\frac{\lambda - \lambda_{\mathrm{D}}}{\lambda_{\mathrm{D}}}$$

where $\lambda_{\mathrm{D}} = 656.1\,\mathrm{nm}$, $c =$ speed of light, and $u =$ the ion's velocity component along the line-of-sight $\mathbf{u}$ to the detector. We now switch to the formulation

$$I(u,\phi) = \int_0^{\infty}\int_{-\infty}^{\infty} k(u,\phi;v_{\|},v_{\perp})\,f(v_{\|},v_{\perp})\,\mathrm{d}v_{\|}\,\mathrm{d}v_{\perp} \ ,$$

with

$$k(u,\phi;v_{\|},v_{\perp}) = R(v_{\|},v_{\perp})\,\pi(u\,|\,\phi,v_{\|},v_{\perp}) \ .$$

Again we assume $R(v_{\|},v_{\perp})$ is a constant, and

$$\pi(u\,|\,\phi,v_{\|},v_{\perp}) = \frac{1}{\pi\,v_{\perp}}\left(1-\left(\frac{u - v_{\|}\,\cos\phi}{v_{\perp}\,\sin\phi}\right)^2\right)^{-1/2} .$$

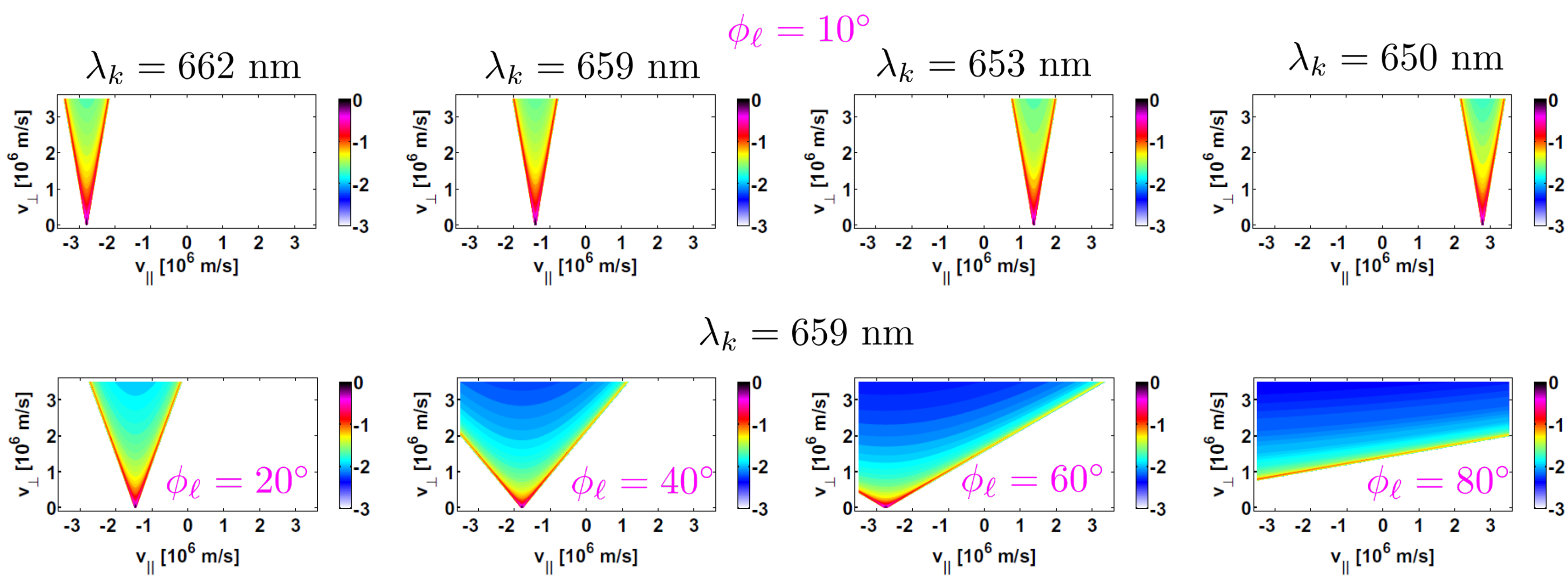
See, e.g., Salewski et al. (2014) for a derivation of this.

The Inverse Problem, Part III

Note that each measurement, for a given pair of velocity u_k and measurement angle ϕ_ℓ, can be written as

$$I(u_k,\phi_\ell) = \int_{-1}^{1}\int_{0}^{\infty} K(u_k,\phi_\ell;v_\|,v_\perp)\, f(v_\|,v_\perp)\,\mathrm{d}v_\|\,\mathrm{d}v_\perp\ .$$

Here, $K(u_k,\phi_\ell;v_\|,v_\perp)$ is called the *weight function* for measurement (k,ℓ).



Discretization

The discretized problem takes the usual form $\boxed{A\,x = b}$.

- We discretize the unknown function $f(v_{\|}, f_{\perp})$ on a grid, and stack the values in the vector x of unknowns.
- For each measurement angle ϕ we measure $I(u, \phi)$ for discrete values of u, and we stack all these measurements of I in the vector b.
- The coefficient matrix (system or weight matrix) A contains discrete values of the kernel $k(u, \phi; v_{\|}, v_{\perp})$ – each row represents one weight function.

Typical sizes of the ingredients:

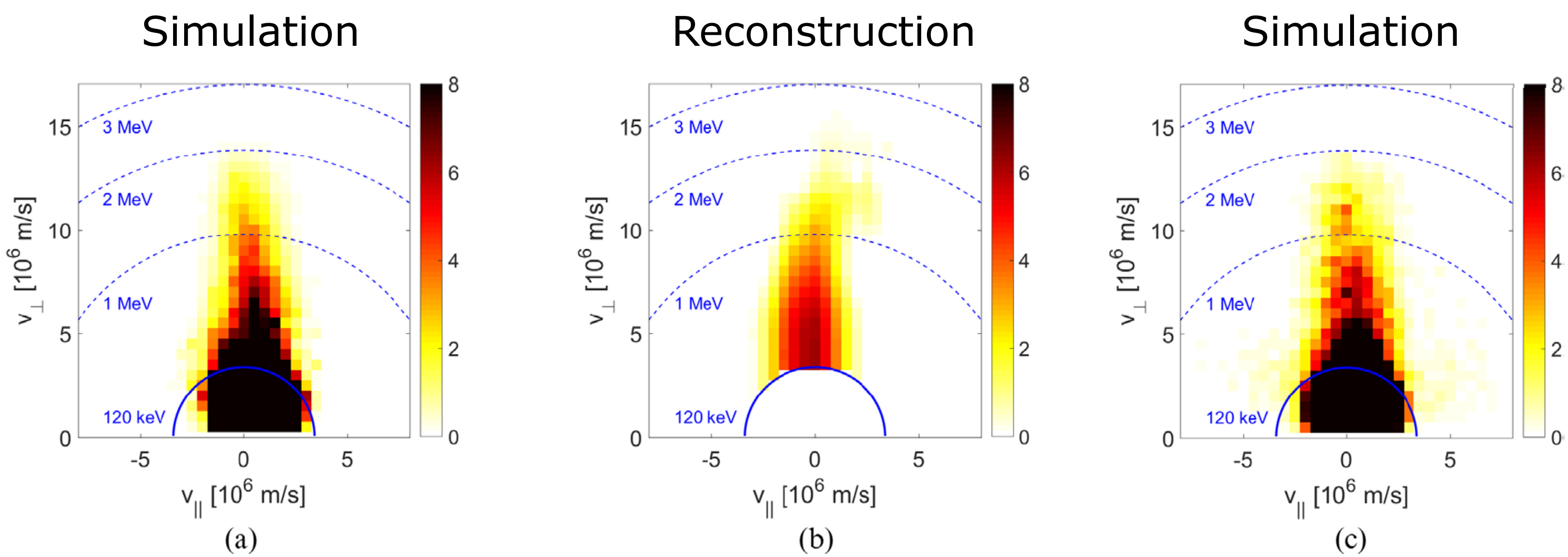
- 100 values of u and 5 detectors (each with its own ϕ) $\rightarrow$ 500 data elements.
- $(v_{\|}, v_{\perp})$ discretised, say, in a 20×20 grid $\rightarrow$ 400 unknowns.
- The matrix A is therefore $m \times n$ with $m \approx 500$ and $n \approx 400$.

Hence we deal with a small-scale problem.

Reconstructions

Plain vanialla Tikhonov regularization with non-negativity constraints:

$$\min_{x} \|A\,x - b\|_2 + \alpha^2 \|x\|_2^2 \quad \text{s.t.} \quad x \geq 0 \ ,$$



From Salewski et al. (2017).

A Convenient Reformulation

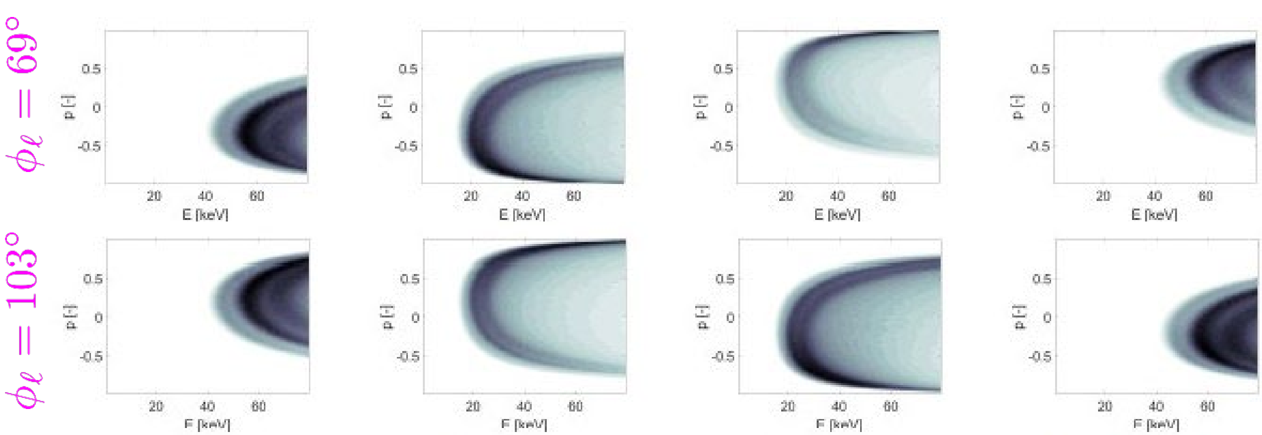
Instead of working with the velocity distribution in the form $f(v_{\|}, v_{\perp})$, the physicists often prefer a *transformation of variables* to energy E and pitch p:

$$E = {}^1\!/_2\, m_\mathrm{D}\, v^2 \ , \qquad p = v_{\|}/v$$

$$v = \sqrt{v_{\|}^2 + v_{\perp}^2} \ , \qquad m_\mathrm{D} = \text{mass of deuterium ion}$$

and instead work with the distribution function $F(E,p)$:

$$I(u,\phi) = \int_{-1}^{1} \int_{0}^{\infty} K(u,\phi; E, p)\, F(E,p)\, \mathrm{d}E\, \mathrm{d}p \ .$$



The weight functions $K(u_k, \phi_\ell; E, p)$ for four different Doppler shifts $-5\,\mathrm{nm}, -3\,\mathrm{nm}, 3\,\mathrm{nm}, 5\,\mathrm{nm}$

Salewski et al. (2018).

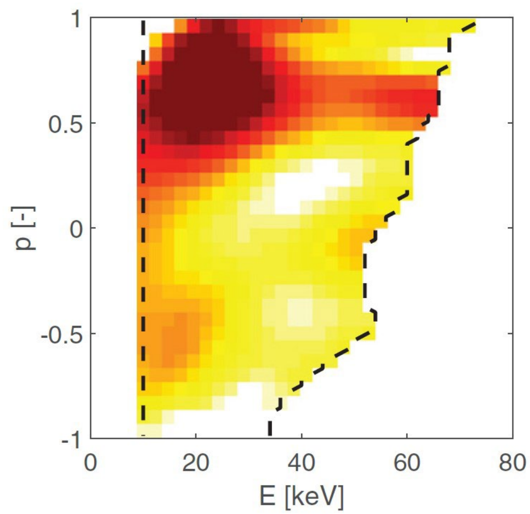
Utilizing Knowledge about the Physics

Tikhonov regularization in general form, with constraint:

$$\min_x \|A\,x - b\|_2 + \alpha^2 \|L\,x\|_2^2 \qquad \text{s.t.} \qquad x \geq 0 \ ,$$

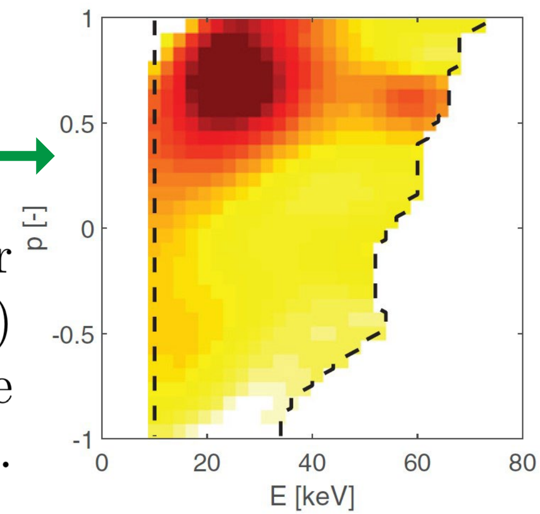
where $\alpha = \alpha(E, p)$ may depend on energy and pitch; Salewski et al. (2016).

Prior	Benefit	Risk
$x \geq 0$	Improves solution	$x < 0$ could diagnose data error
$L \sim 1.$ deriv	Gives smooth solutions	Misses spikes and ridges
$\alpha(E, p)$	Accounts for NBI peaks	Might introduce spurious peaks.



Constant α.

$\alpha = \alpha(E, p)$ with smaller values (less regularization) at known locations of the NBI peaks in (E, p)-space.



Physics Prior via Slowing-Down Functions

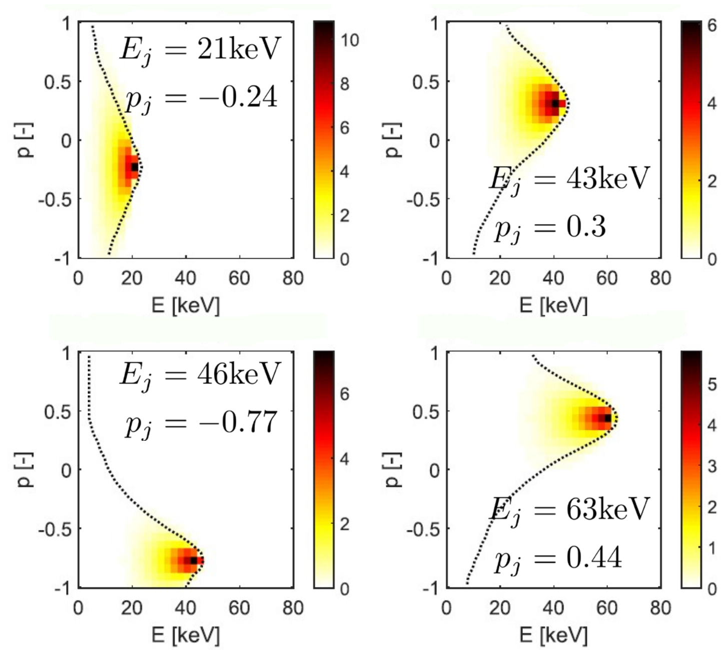
Instead of working with a standard "pixel basis" for $F(E,p)$, we can use suited basis functions $\psi_1, \psi_2, \ldots, \psi_{n_{\mathrm{sd}}}$ and write the reconstruction

$$F(E,p) = \textstyle\sum_{j=1}^{n_{\mathrm{sd}}} c_j \, \psi_j(E,p) \ .$$

The discretized and regularized problem for $\{c_j\}$ then takes the basic form

$$\min_c \| A \, \Psi \, c - b \|_2^2 + \alpha^2 \|c\|_2^2 \ , \qquad \Psi \in \mathbb{R}^{n \times n_{\mathrm{sd}}} \ , \qquad x = \Psi \, c \ ,$$

where the columns of Ψ are samples of the basis functions ψ_j = slowing down functions.



Each ψ_j is a distribution F excited by a δ-function $\delta(u_j, \phi_j)$ corresponding to (E_j, p_j), revealing how the plasma ions "slow down" due to collisions.

Hence, the basis functions ψ_j represent the physics of the plasma.

Interpretation of Basis with Physics Prior

The case $n_{\text{sd}} = n$

Ψ is square and we assume it has full rank. Recall $x = \Psi\, c \;\Leftrightarrow\; c = \Psi^{-1} x$:

$$\min_c \|A\,\Psi\,c - b\|_2^2 + \alpha^2 \|c\|_2^2 \qquad \Leftrightarrow \qquad \min_x \|A\,x - b\|_2^2 + \alpha^2 \|\Psi^{-1} x\|_2^2 \; .$$

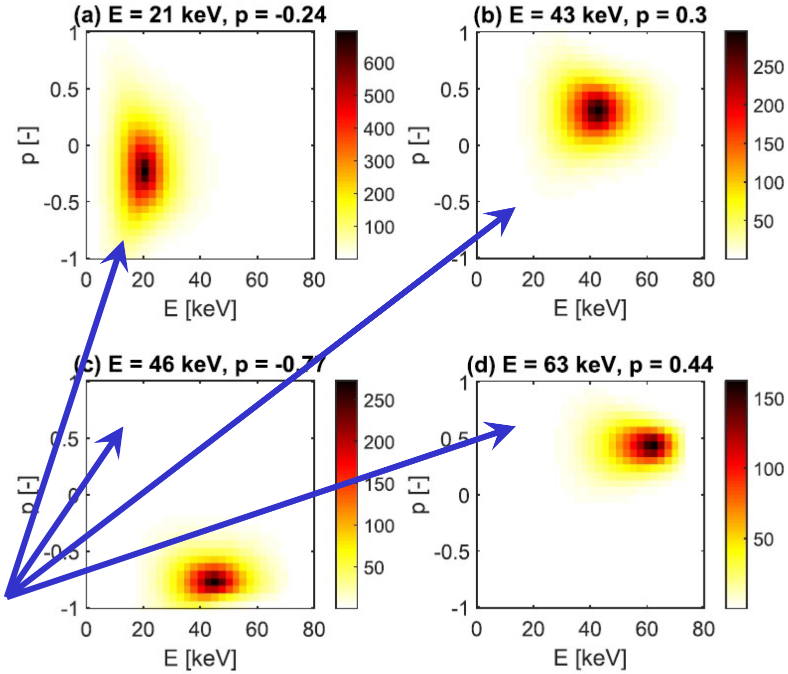
We can interpret the use of Ψ as a regularizer Ψ^{-1} for x in the reg. term.

In the Bayesian framework, the regularization term $\alpha^2 \|\Psi^{-1} x\|_2^2$ represents a Gaussian prior for x with covariance matrix

$$C_x = \left(\alpha^2 (\Psi^{-1})^T \Psi^{-1}\right)^{-1} = \alpha^{-2} \Psi\, \Psi^T \; .$$

This C_x ensures smoothness by correlating a given pixel to the pixels in its vicinity.

Rows of C_x reshaped to (E, p) domain



Interpretation, Part II

And now: the **standard-form transformation**

$$\min_x \|A\,x - b\|_2^2 + \alpha^2 \|L\,x\|_2^2 \qquad \Leftrightarrow \qquad \min_\xi \|(A\,L_A^\dagger)\,\xi - b\|_2^2 + \alpha^2 \|\xi\|_2^2$$

where $\xi = L\,x \;\Leftrightarrow\; x = L_A^\dagger \xi$, and $L_A^\dagger$ = A-weighted pseudoinverse of L.

The case on the previous slide ($n = n_{\mathrm{sd}}$) is a special case in which

$$L = \Psi^{-1} \quad \text{and} \quad L_A^\dagger = L^{-1} = \Psi\ .$$

The case $n_{\mathrm{sd}} > n$

The matrix Ψ is "obese" and we assume it has full row rank. Then

$$L = \Psi^\dagger \quad \text{because} \quad L_A^\dagger = L^\dagger = (\Psi^\dagger)^\dagger = \Psi$$

and

$$C_x = (\alpha^2 (\Psi^\dagger)^T \Psi^\dagger)^{-1} = \alpha^{-2} \Psi\,\Psi^T\ , \quad \text{similar to before.}$$

Interpretation, Part III

Ongoing work

The case $n_{\rm sd} < n$

Ψ is "tall & skinny" and we assume full row rank. How to interpret this case?

1 Can we determine a matrix L such that $L_A^\dagger = \Psi$? Nonlinear problem. ☹

→ **Detour!**

2 Write

$$\Psi = (\, Q \,,\, Q_0 \,) \begin{pmatrix} R \\ 0 \end{pmatrix} = Q\, R \,, \qquad \mathrm{range}(Q_0) = \mathrm{range}(\Psi)^\perp = \mathrm{null}(\Psi^T) \;.$$

Let $P = Q_0\, Q_0^T$ = orthogonal projector on $\mathrm{null}(\Psi^T)$. For a general x:

$$x = \Psi\, c + Q_0\, w \quad \Leftrightarrow \quad Q\, R\, c = x - Q_0\, w \quad \Leftrightarrow \quad c = R^{-1} Q^T x = \Psi^\dagger x \;.$$

We want $x \in \mathrm{range}(\Psi)$ and hence $Q_0\, w = P\, x = 0$, and we arrive at

$$\min_x \| A\, x - b \|_2^2 + \alpha^2 \| \Psi^\dagger x \|_2^2 \qquad \text{s.t.} \qquad P\, x = 0 \;.$$

and

$$C_x = (\alpha^2 (\Psi^\dagger)^T \Psi^\dagger)^{-1} = \alpha^{-2} \Psi\, \Psi^T \;, \quad \text{similar to before.}$$

The Covariance Matrix for a "Skinny" Basis

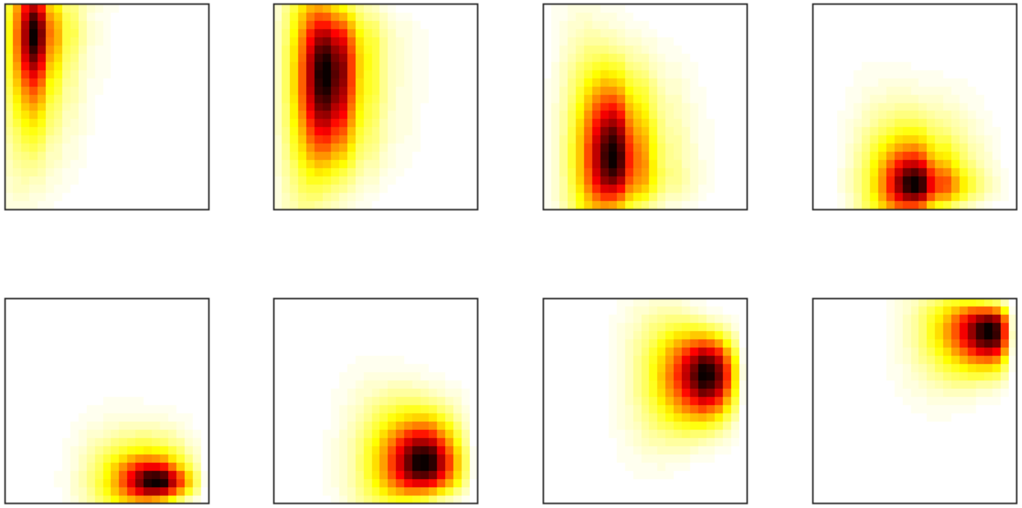
Recap: the reason for studying the standard-form transformation is **not** for computations, but in order to *interpret the role of the basis* Ψ via

- the regularization term $\alpha^2 \, \|\Psi^\dagger x\|_2^2$ (possibly with constraint $P\, x = 0$),
- the covariance matrix $C_x = \alpha^{-2} \Psi \, \Psi^T$.

We continue with the case where $\Psi \in \mathbb{R}^{625 \times 170}$ is "tall & skinny," i.e., we have fewer basis functions ψ_j than the number of unknowns (pixels in the image).

Rows of C_x reshaped to (E, p) domain; as before, they represent local averaging.

For lower energy E the correlation in pitch p increases. This confirms the intuition of the physicists.



Conclusions

- Tikhonov regularization is perhaps “old school” but it works well here.
- The slowing-down functions provide a set of basis vectors that represent the behavior of the ions in the plasma.
- Via a standard-form transformation paradigm we can regirously interpret the use of these basis functions as imposing local smoothing regularization.
- The specific smoothing that we observe confirms the intuition of the physicists.
- Next steps: further studies of the insight provided by this interpretation, and GSVD analysis of the pair $(A,\Psi^{\dagger})$, uncertainty quantification.