EUROFUSION WPMST1-PR(15) 14314

A S Jacobsen et al.

Inversion methods for fast-ion velocity-space tomography in fusion plasmas

Preprint of Paper to be submitted for publication in
Plasma Physics and Controlled Fusion

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Inversion methods for fast-ion velocity-space tomography in fusion plasmas

A. S. Jacobsen[1], L. Stagner[2], M. Salewski[1], B. Geiger[3], W. W. Heidbrink[2], S. B. Korsholm[1], F. Leipold[1], S. K. Nielsen[1], J. Rasmussen[1], M. Stejner[1], H. Thomsen[4], M. Weiland[3], and the ASDEX Upgrade team[3]

[1] Technical University of Denmark, Department of Physics, DK-2800 Kgs. Lyngby, Denmark

[2] University of California Irvine, Irvine, CA, USA

[3] Max-Planck-Institute for Plasma Physics, Boltzmannstr. 2, 85748 Garching, Germany

[4] Max-Planck-Institute for Plasma Physics, Wendelsteinstrasse 1, 17491 Greifswald, Germany

E-mail: ajsen@fysik.dtu.dk

E-mail: lstagner@uci.edu

Abstract. Velocity-space tomography has been used to infer 2D fast-ion velocity distribution functions. Here we compare the performance of five different inversion methods: Truncated singular value decomposition, maximum entropy, minimum Fisher information and zeroth- and first-order Tikhonov regularization. The inversion methods are applied on fast-ion D_α [1,2] measurements taken just before and just after a sawtooth crash in the ASDEX Upgrade tokamak as well as on synthetic measurements from different test distributions. We find that the methods regularizing by penalizing steep gradients perform best. Furthermore, most of the methods agree that passing as well as trapped ions are ejected from the plasma center and that this ejection depletes the passing ion population more strongly than the trapped ion population. We assess the uncertainty of the calculated inversions taking into account photon noise, uncertainties in the forward model as well as uncertainties introduced by the regularization.

1. Introduction

Traditional fast-ion diagnostics and analysis procedures provide only incomplete information about the 2D fast-ion velocity distribution function. Using velocity-space tomography it is possible to combine data from several measurements to infer the 2D fast-ion velocity distribution function [3–10]. With this approach it should even be possible to combine measurements from different diagnostics which is beneficial as they are always sensitive to different regions of velocity-space [6]. This velocity-space sensitivity is quantified by velocity-space sensitivity functions, also called weight

functions, which have been developed for fast-ion D_α spectroscopy (FIDA) [11, 12], collective Thomson scattering (CTS) [4], fast-ion loss detectors (FILD) [13], neutron emission spectrometry (NES) [14, 15] and gamma-ray spectroscopy (GRS) [16]. The weight functions, w, relate a measurement, s, to the fast-ion velocity distribution function, f:

$$s = \int\int w f \, \mathrm{d}E \, \mathrm{d}p \,, \tag{1}$$

where E is the fast-ion energy and p is the pitch defined as $p = \frac{v_\parallel}{v}$. Here $v_\parallel$ is the ion velocity parallel to the magnetic field and v is the ion speed. p is defined positive in the co-current direction. By discretizing equation (1) and normalizing by the uncertainties [6], a linear system of equations is obtained:

$$S = WF \,, \tag{2}$$

where S and F are vectors of length m and n, respectively, and W is an $m \times n$ matrix dubbed the transfer matrix. Calculating F from equation (2) is a mathematically ill-posed inverse problem. Nevertheless, it is possible to obtain sensible solutions to such an ill-posed inverse problem by regularizing it. Many different inversion methods have been developed and applied in many scientific fields. Here we compare five inversion methods to measure fast-ion velocity distribution functions by tomographic inversion: truncated singular value decomposition, maximum entropy, zeroth- and first-order Tikhonov regularization and minimum Fisher information. These methods have been compared for position-space tomography in fusion plasmas [17, 18]. Here we test these methods for velocity-space tomography. In our comparison we use a transfer matrix describing FIDA measurements taken simultaneously in five views at ASDEX Upgrade in discharge #31557. This diagnostic set-up (i.e. this transfer matrix) is used, firstly, with synthetic measurements to calculate inversions for known velocity distribution functions and, secondly, with real five-view FIDA measurements taken just before and just after a sawtooth crash. The synthetic measurements enable us to quantify the performance of the different methods for assumed measurement uncertainties since the true solution is known. The real measurements allow us to investigate the redistribution of fast ions due to a sawtooth crash resolved in 2D velocity space for ion energies above about 20 keV.

The paper is organised as follows. The FIDA diagnostic set-up is described in section 2. Section 3 explains the different inversion methods. Section 4 presents how to estimate the uncertainties of the inversions. In section 5 we quantify the performance of the inversion methods by inverting synthetic data based on known test functions. In section 6 the methods are used to investigate the effect of a sawtooth crash on the central fast-ion population. The results are discussed in section 7 and conclusions are summarized in section 8.

2. ASDEX Upgrade FIDA system

A FIDA diagnostic set-up measures Doppler-shifted deuterium Balmer-alpha light from the plasma. It is extensively used to diagnose fast ions at ASDEX Upgrade [19, 20], DIII-D [21, 22], NSTX [23], MAST [24, 25] and LHD [26]. The newly upgraded set-up at ASDEX Upgrade now consists of five different views, each with several lines of sight measuring at different radial locations. We use one line of sight from each view, each intersecting the beam path of neutral beam injector (NBI) Q3 in the plasma centre. Each view has a different angle between its line of sight and the magnetic field. Thereby, they probe different regions in velocity space as described by their weight functions [12]. In the plasma centre, the respective angles are 14°, 73°, 103°, 133° and 153°. A description of the updated FIDA system is found in [10].

3. Inversion methods

3.1. Singular value decomposition

Truncated singular value decomposition (SVD) is the method used previously to calculate velocity-space tomographies in fusion plasmas [5–9]. The $m \times n$ transfer matrix, W, can as any matrix be written as the product of three matrices:

$$W = U\Sigma V^T \,, \tag{3}$$

where the columns of the $m \times m$ matrix U are the eigenvectors of the matrix WW^T and the columns of the $n \times n$ matrix V are the eigenvectors of the matrix W^TW [27]. U as well as V are orthogonal matrices. Σ is a diagonal $m \times n$ rectangular matrix. The values in the diagonal are called the singular values. They are the square roots of the non-zero eigenvalues of both WW^T and W^TW [28]. The values in the diagonal of Σ are ordered in a decreasing manner.

Given a linear system of equations

$$WF = S \,, \tag{4}$$

the solution, F_{SVD}, is found as

$$F_{SVD} = W^+ S \,, \tag{5}$$

where W^+ is called the *pseudoinverse* of W. W^+ can be calculated using the SVD factorization:

$$W^+ = V\Sigma^+ U^T \,, \tag{6}$$

where Σ^+ is a rectangular diagonal matrix with the reciprocals of the diagonal elements of Σ on the diagonal. Writing the SVD as a sum, one gets

$$WF = \sum_{j=1}^{r} u_j \sigma_j \left(v_j^T F\right) \,, \tag{7}$$

where r is the number of non-zero singular values, u_j and v_j are the j'th columns of U and V, respectively, and σ_j is the j'th singular value. v_j^T indicates the transpose of the j'th column of V. F_{SVD} can likewise be expressed as a sum:

$$F_{SVD} = \sum_{j=1}^{r} \frac{v_j \left(u_j^T S\right)}{\sigma_j} \,. \tag{8}$$

Experimental data always contains some form of noise. Here we define

$$S = S_{exact} + e \,, \tag{9}$$

where S_{exact} is the idealized measurement without noise and e is the noise. Inserting equation (9) in equation (8) we get

$$F_{SVD} = \sum_{j=1}^{r} \frac{v_j \left(u_j^T S_{exact}\right)}{\sigma_j} + \sum_{j=1}^{r} \frac{v_j \left(u_j^T e\right)}{\sigma_j} = F_{exact} + \sum_{j=1}^{r} \frac{v_j \left(u_j^T e\right)}{\sigma_j} \,, \tag{10}$$

where F_{exact} is the exact solution we seek and the last sum describes the effect of the noise. For very small singular values, the SVD solution can be completely dominated by the noise. To reduce its influence, a possibility is to truncate the sum after k terms. However, this makes it impossible to reconstruct F_{exact} completely. This method is called truncated SVD. Truncated SVD introduces the problem of choosing the optimum truncation level, k. Several criteria exist for choosing k. Here we use the L-curve criterion to choose k [29].

3.2. Tikhonov regularization

The inverse problem posed in equation (2) can be formulated as a least squares problem, i.e. find the solution F which minimizes the norm of the residual:

$$\text{minimize} \left\{ ||WF - S||^2 \right\} \,. \tag{11}$$

Well-posed problems can be solved using the normal equations:

$$F = \left(W^T W\right)^{-1} W^T S \,. \tag{12}$$

However, for ill-posed problems a small change in S can have a significant impact on F. In Tikhonov regularization, the ill-posed least squares problem is replaced by a nearby well-posed least squares problem

$$\text{minimize} \left\{ ||WF - S||^2 + \alpha ||LF||^2 \right\} \,, \tag{13}$$

where L is a regularization matrix of size $n \times n$ and α is a non-negative number determining the weight of the regularization term. As for truncated SVD, we determine the value of α using the L-curve method [30]. The Tikhonov solution, F_α, becomes

$$F_\alpha = \left(W^T W + \alpha L^T L\right)^{-1} W^T S \,. \tag{14}$$

The choice of regularization matrix determines the nature of the regularization. Common choices of L penalize the magnitude of f or its derivative to different orders.

Therefore, Tikhonov regularization is also sometimes called linear regularization. The simplest regularization matrix is

$$L = I\,, \tag{15}$$

where I is the $n \times n$ identity matrix so that $L^T L = I$ and equation (14) becomes

$$F_\alpha = \left(W^T W + \alpha I\right)^{-1} W^T S\,. \tag{16}$$

This penalizes large absolute values of f and is called 0'th order regularization.

1'st order regularization penalizes large gradients. In 2D velocity space $\left(v_\parallel, v_\perp\right)$, the penalty operator is

$$L^T L = \nabla_{v_\parallel}^T \nabla_{v_\parallel} + \nabla_{v_\perp}^T \nabla_{v_\perp}\,. \tag{17}$$

Here $\nabla_{v_\parallel}$ and $\nabla_{v_\perp}$ are matrix representations of finite difference operators. In (E, p)-coordinates, the velocity-space gradient is

$$\nabla F = \sqrt{2mE}\,(\nabla_E F)\,\hat{e}_E + \sqrt{\frac{m}{2E}}\sqrt{1-p^2}\,(\nabla_p F)\,\hat{e}_p\,. \tag{18}$$

The derivation of equation 18 is included in Appendix A. Hence in (E, p)-coordinates the penalty operator becomes

$$L^T L = 2mE\,\nabla_E^T \nabla_E + \frac{m}{2E}\left(1-p^2\right)\,\nabla_p^T \nabla_p\,. \tag{19}$$

3.3. Minimum Fisher information regularization

The principle of minimum Fisher information has been used to compute inversions in soft X-ray tomography in tokamak plasmas [17]. In reference [17] the minimum Fisher information principle is effectively built in as a Tikhonov penalty function. It can therefore be seen as a variant or extension of the general Tikhonov regularisation method. The minimum Fisher information method penalizes large gradients divided by the function values. The normalization with the distribution itself means that the smoothing effect is strongest where the distribution has low values.

The minimum Fisher information method is implemented as an iterative algorithm [17]. First a solution $F^{(1)}$ is found using Tikhonov regularization with a first-order linear penalty function. In the subsequent iterations, the penalty function in $\left(v_\parallel, v_\perp\right)$-coordinates becomes

$$L^T L = \nabla_{v_\parallel}^T M^{(n)} \nabla_{v_\parallel} + \nabla_{v_\perp} M^{(n)} \nabla_{v_\perp}\,, \tag{20}$$

where

$$M_{i,j}^{(n)} = \frac{1}{F_i^{(n-1)}}\,\delta_{i,j} \qquad \text{if} \qquad F_i^{(n-1)} > 0 \tag{21}$$

$$M_{i,j}^{(n)} = M_{max}^{(n)}\,\delta_{i,j} \qquad \text{if} \qquad F_i^{(n-1)} \leq 0\,. \tag{22}$$

$M_{max}^{(n)}$ is the largest $M^{(n)}$ for $F_i > 0$. In (E, p)-coordinates the penalty function becomes

$$L^T L = 2mE \nabla_E^T M^{(n)} \nabla_E + \frac{m}{2E}\left(1-p^2\right)\nabla_p^T M^{(n)} \nabla_p\,. \tag{23}$$

In each iteration, the corresponding Tikhonov solution with the appropriate minimum Fisher information penalty function is found. We find that the solution converges after only a few iterations.

Maximum entropy regularization

The last inversion method we have implemented is maximum entropy regularization. In the case of maximum entropy regularization it is assumed that the object that is to be reconstructed from data is positive $f(E,p) \geq 0$. The specific formulation of maximum entropy regularization can be found in references [17,31]. Maximum entropy regularization can be formulated as a minimization problem of the form

$$\text{minimize} \left\{ \frac{1}{2}||WF - S||^2 + \alpha H \right\} . \tag{24}$$

where α is a free parameter controlling the strength of our assumptions similar to the free parameter introduced in Tikhonov and minimum Fisher information regularization. We determine the optimal value of α using the L-curve method [30]. H is the Shannon information entropy given by

$$H = -\sum_{i=1}^{N} (F_i - m_i - F_i \ln(F_i/m_i)) . \tag{25}$$

The entropy H is minimized when $F_i = m_i$. Thus m_i is called the default model as it is the value F_i will take when there is no information or data influencing it. While the default model is usually set to be constant in phase-space to prevent biasing of the solution, we may choose to set the default model to be given by a theoretical model. Thus, the process of inference becomes the process of updating the theoretical model to be consistent under new observations [32]. For this work, the default model is set to be constant. The solution of this minimization problem, called the maximum entropy solution, is found using a general non-linear optimization library [33, 34], implemented in a Bayesian framework.

4. Uncertainties in tomographic inversions

It is possible to estimate the uncertainties in a tomography. Here we consider uncertainties due to noise in the measurements, uncertainties due to uncertainties in the forward model as well as uncertainties imposed by the regularization methods themselves.

The forward model is given by the FIDA weight functions. These are calculated based on profiles of several nuisance parameters. The weight functions are most sensitive to the ion temperature and drift velocity, the electron temperature and density and the effective charge Z_{eff}. Hence we consider the impact of these parameters on the tomographic inversion results for the different regularization methods. The uncertainties in the bulk plasma parameters lead to uncertainties in the weight functions, δW. Assuming a Gaussian error distribution of the bulk plasma parameters, we calculate the uncertainty in the forward model (i.e. in the weight functions) by sampling a population of weight functions calculated while varying one nuisance plasma parameter at a time and keeping the other parameters fixed. The total variance of the weight function is

then obtained by summing up the variances obtained from each plasma parameter. The corresponding error, e_m, from the forward model error is

$$e_m = \delta W F_{true}\,. \tag{26}$$

e_m depends on the (often unknown) true distribution function. However, if an estimate of F_{true} can be obtained, e_m can be estimated. The combined uncertainty due to uncertainty in the forward model and measurement uncertainty is then

$$e_{data} = \sqrt{e_{ph}^2 + e_m^2}\,. \tag{27}$$

where e_{ph} is the photon noise. In this work we approximate F_{true} by the sum of a Maxwellian for the bulk ions and a TRANSP/NUBEAM [35] simulation to estimate the NBI ions. The Maxwellian is calculated using measured ion temperature and density and we then use e_{data} to normalize the measurements and transfer matrix.

The covariance matrix of the tomography due to uncertainty in the spectra, C_e^F, is [6]

$$C_e^F = W^\dagger C_e^S \left(W^\dagger\right)^T\,, \tag{28}$$

where C_e^S is the covariance matrix of the uncertainty in the spectra and $W^\dagger$ denotes the regularized inverse. For uncorrelated uncertainties the uncertainty in the tomography is then given by the diagonal elements of the covariance matrix. For SVD, $W^\dagger = W^+$, the pseudoinverse of W. For Tikhonov and minimum Fisher information regularization, $W^\dagger = \left(W^T W + \alpha L^T L\right)^{-1} W^T$. The covariance matrix of the maximum entropy solution is given by [36]

$$C_e^F = \left(\left(W^T W\right)_{ij} + \frac{\alpha}{F_i}\delta_{ij}\right)^{-1}_{ij}\,. \tag{29}$$

As $W^\dagger$ depends on the regularization, C_e^F depends both on the choice of regularization method and the level of regularization. In fact, the purpose of the regularization is to decrease the effect of noise in the spectra. The price one pays by regularizing the problem is to introduce a regularization or smoothing error. This error makes it impossible to recreate the true distribution, even if the measurements were noise free. The regularization error, e_{reg}, is defined as

$$\mathrm{e}_{reg} = \left(I - W^\dagger W\right) F_{true}\,, \tag{30}$$

where I is the $n \times n$ identity matrix. e_{reg}, as e_m, depends on F_{true}. However, if an estimate for the true distribution function such as a simulated velocity distribution function is known, e_{reg} can be estimated. Here we use a TRANSP/NUBEAM simulation as our best estimate of F_{true}.

We define a measure of the complete uncertainty of the tomographic inversion as:

$$e_{tot} = \sqrt{\left(\mathrm{diag}\left(C_e^F\right)\right)^2 + e_{reg}^2}\,. \tag{31}$$

However, it should be noted that the regularization error has a different character than the data uncertainty which represents an assumed Gaussian distribution of values in many realizations of the experiment. On the contrary, the regularization error

characterizes how well the inversion method reconstructs a known test function. We nevertheless combine these uncertainties as a measure of the total uncertainty of the tomographic inversion.

5. Tomographies using synthetic measurements

In this section we calculate tomographies using synthetic data obtained using equation (1) and known distribution functions. Inversions of synthetic spectra calculated from known distribution functions enable us to compare the performance of the inversion methods by quantitative figures of merit as we know the true solution.

5.1. Figures of merit

We define two different figures of merit. The first is the total sum of the absolute difference between the true solution and the tomography normalized by the sum of the true solution,

$$\frac{\sum_i |F_{tomography,i} - F_{true,i}|}{\sum_i F_{true,i}}\,, \tag{32}$$

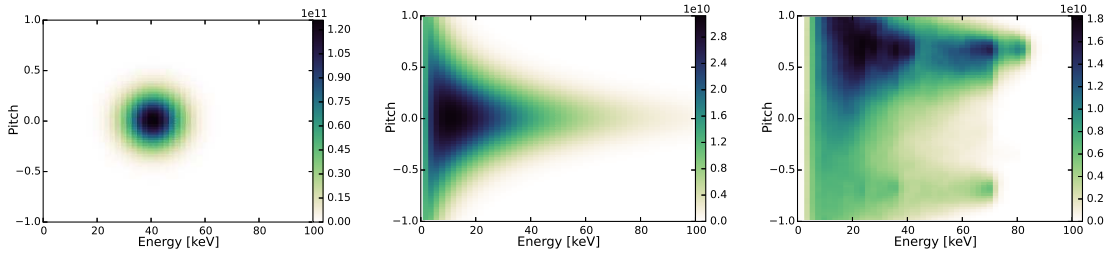
For this figure of merit the smallest values correspond to the best performance of the inversion method. The second figure of merit is the ratio of the inferred fast-ion density to the true fast-ion density which is calculated as the integral of the tomography normalized by the integral of the true distribution,

$$\frac{\int\int F_{tomography}\,\mathrm{d}E\mathrm{d}p}{\int\int F_{true}\,\mathrm{d}E\mathrm{d}p}\,. \tag{33}$$

Optimally, this figure of merit should be one. However, in this study we reconstruct the velocity-space above about 20 keV which is realistic for the ASDEX Upgrade case. As the bi-Maxwellian and NBI distributions contain ions with energies below 20 keV, the optimal value is below one for these two distributions.

5.2. Test velocity-distribution functions

Three different velocity functions will be investigated in this analysis. A Gaussian distribution, a bi-Maxwellian distribution and a simulated NBI-distribution from TRANSP/NUBEAM. The three distributions are shown in figure 1. We choose these three distribution functions as they pose different challenges to the inversion methods. The Gaussian distribution is highly localized and requires good resolution properties of the method. A Gaussian blob may further represent a source of fast particles typical for the peaks at the injection energies for neutral beam heating. The bi-Maxwellian is a wide function covering the entire pitch range. Here the challenge is to recreate the large-scale undulation. Lastly, we study a distribution function typical for neutral beam injection as simulated by TRANSP. This is an important test case as it should be very similar to the distribution functions in experiments with NBI heating. The challenge here is the structure complexity on both small and large spatial scales.



(a) Gaussian. (b) Bi-Maxwellian. (c) NBI.

Figure 1. Test velocity distributions functions as a function of energy and pitch of the ions. The functions are given in units of [ions/keV/cm^3]

5.3. Modelling of measurement noise

The photon noise of FIDA light scales approximately with the square-root of the signal. However, in the absence of FIDA light the photon noise is dominated by bremsstrahlung setting a lower limit on the noise level. These two effects are modelled as

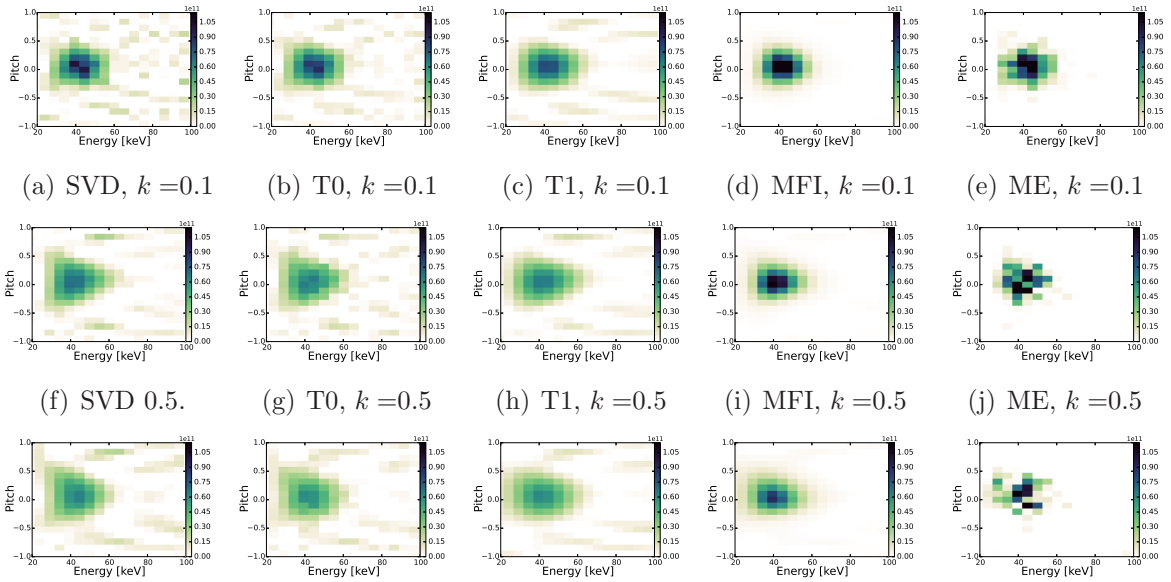
$$S_{noisy} = S_{exact} + k \left\langle \sqrt{S_{exact}} \right\rangle \mathcal{N}\left(0, \max\left(e_{min}, \sqrt{S_{exact}}\right)\right) , \tag{34}$$

where S_{noisy} is the noisy spectrum, S_{exact} is the exact noise-free spectrum, e_{min} is the bremsstrahlung level and k is a scaling constant that allows us to vary the noise level. $\mathcal{N}\left(0, \max\left(e_{min}, \sqrt{S_{exact}}\right)\right)$ denotes a Gaussian distribution with a mean of zero and a standard deviation of the maximum of $\sqrt{S_{exact}}$ and the bremsstrahlung level e_{min}. Varying the noise level allows us to investigate how robust the methods are against noise.

5.4. Inversion results

Figure 2 shows tomographies of the Gaussian blob shown in figure 1(a) calculated with the different methods for various noise levels. All methods reconstruct the position of the Gaussian blob well. The characteristic widths of the Gaussians are approximately right but tend to be slightly larger than in the original test distribution. Measurement noise enhances this trend. We further observe the appearance of jitter in the inversions throughout velocity space. The minimum Fisher information regularization stands out from the other methods in that it resembles the original function the most. This suggests superior resolution performance of the method. The minimum Fisher information regularization is furthermore most robust against measurement noise. Of the other methods, the maximum entropy regularization performs best. It is not as smooth, but it is able to localize the large values of the Gaussian distribution very well in velocity-space.

Figure 3 shows the reconstructions of the bi-Maxwellian distribution function. The large-scale shape of the distribution is reproduced well by all five inversion methods. The pitch angle symmetry with respect to $p = 0$ is reproduced well and the larger perpendicular temperature compared with the parallel temperature is reflected well in



(a) SVD, k =0.1 (b) T0, k =0.1 (c) T1, k =0.1 (d) MFI, k =0.1 (e) ME, k =0.1

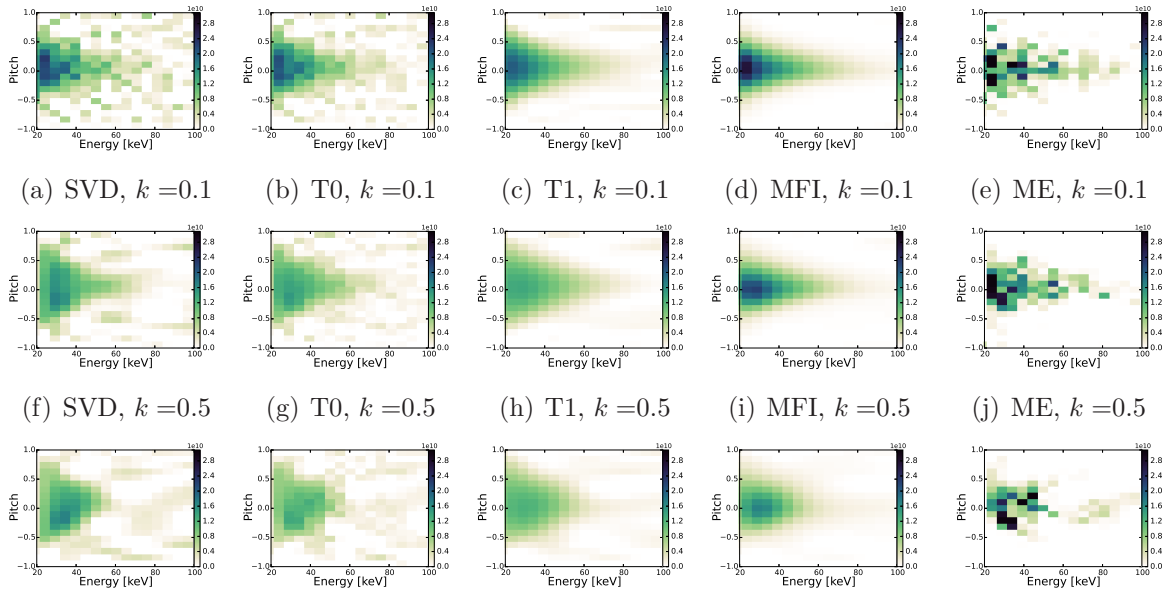
(f) SVD 0.5. (g) T0, k =0.5 (h) T1, k =0.5 (i) MFI, k =0.5 (j) ME, k =0.5

(k) SVD, k =0.9 (l) T0, k =0.9 (m) T1, k =0.9 (n) MFI, k =0.9 (o) ME, k =0.9

Figure 2. Tomographies of the Gaussian blob from figure 1(a) based on synthetic measurements using various inversion methods and noise levels. The inversion methods are SVD, 0'th and 1'st order Tikhonov (T0 and T1, respectively), minimum Fisher information (MFI) and maximum entropy (ME). The noise level k is defined in equation (34).

the large fast-ion densities for pitches close to zero. For large noise the SVD, the 0'th order Tikhonov and the maximum entropy become erroneously biased towards negative pitches, but this bias is small and the general trends remain apparent. The 1'st order Tikhonov and the minimum Fisher information reproduce the symmetry and the large fast-ion densities for pitches close to zero particularly well. Additionally, penalizing large gradients produces smooth functions which resemble the smooth bi-Maxellian function well. Further, the jitter appearing throughout velocity space for larger noise levels is significantly smaller.

Figure 4 shows reconstructions of the NBI distribution function for various noise levels and inversion methods. This fast-ion distribution function is typical for neutral beam injection with two co-current beams with injection energies at 80 keV and 70 keV and one counter-current beam with an injection energy of 70 keV. Therefore, this distribution function is a more difficult test case than previously presented reconstructions of distribution functions which are more typical for a single NBI beam. The overall shape of the NBI distribution function with bias towards positive pitches is well reproduced by all five inversion methods. The protrusion at pitches of about 0.7 originates from the co-current beam injection, and the weaker protrusion at pitches of -0.7 from the counter-current beam injection. All reconstructions show that the full energy beam injection peak for co-current injection (positive pitch) is at larger energies than that for counter-current injection (negative pitch). As observed for the Gaussian blob and the bi-Maxwellian, the 1'st order Tikhonov and the minimum Fisher



(a) SVD, k =0.1 (b) T0, k =0.1 (c) T1, k =0.1 (d) MFI, k =0.1 (e) ME, k =0.1

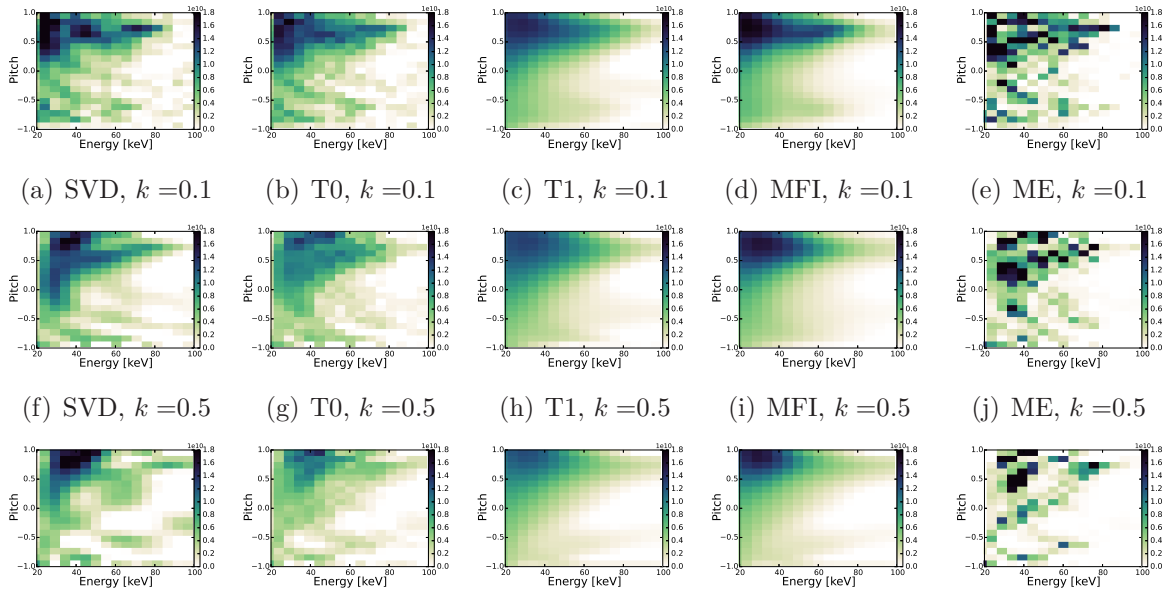
(f) SVD, k =0.5 (g) T0, k =0.5 (h) T1, k =0.5 (i) MFI, k =0.5 (j) ME, k =0.5

(k) SVD, k =0.9 (l) T0, k =0.9 (m) T1, k =0.9 (n) MFI, k =0.9 (o) ME, k =0.9

Figure 3. Tomographies of the bi-Maxwellian from figure 1(b) based on synthetic measurements using various inversion methods and noise levels. The inversion methods are SVD, 0'th and 1'st order Tikhonov (T0 and T1, respectively), minimum Fisher information (MFI) and maximum entropy (ME). The noise level k is defined in equation (34).

information regularization result in smooth tomographies with very little jitter compared with the SVD, 0'th order Tikhonov and maximum entropy. The small amplitudes of the jitter makes the overall shape of the function with protrusions at positive and negative pitches stand out most clearly for the 1'st order Tikhonov and the minimum Fisher information regularization. The local maxima due to the beam injection peaks at full, half and third energies can be seen in the SVD, 0'th order Tikhonov and minimum Fisher information regularization in the case of low noise ($k = 0.1$). For larger noise levels, none of the methods are able to resolve more than one peak.

Figure 5 shows the behaviour of the performance parameters as a function of noise level for the tomographies of the three test functions. Figures 5(a), 5(c) and 5(e) show the normalized absolute difference. The difference increases for larger noise levels for all inversion methods and test distributions. Furthermore the first order Tikhonov and minimum Fisher information regularization methods perform best, except for the Gaussian distribution where the maximum entropy method performs well. A similar behaviour is seen in the density ratios in figures 5(b), 5(d) and 5(f). The methods which minimize gradients reconstruct the densities most accurately. The dashed lines represent the optimal ratio, which for the bi-Maxwellian and NBI distributions are below one since the test distributions has values below an energy of 20 keV, but our reconstructions are limited to energies above 20 keV due to the limited weight function coverage. The ratios in 5(d) and 5(f) all have a distinct negative slope. Thus, for very



(a) SVD, k =0.1 (b) T0, k =0.1 (c) T1, k =0.1 (d) MFI, k =0.1 (e) ME, k =0.1

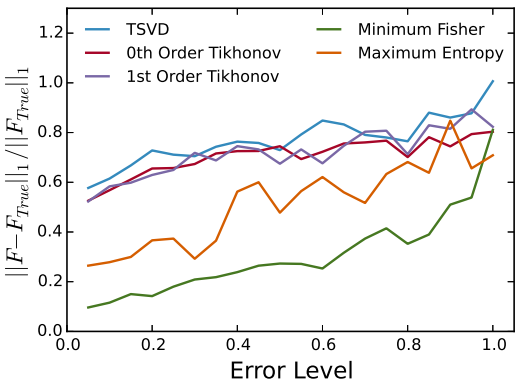
(f) SVD, k =0.5 (g) T0, k =0.5 (h) T1, k =0.5 (i) MFI, k =0.5 (j) ME, k =0.5

(k) SVD, k =0.9 (l) T0, k =0.9 (m) T1, k =0.9 (n) MFI, k =0.9 (o) ME, k =0.9

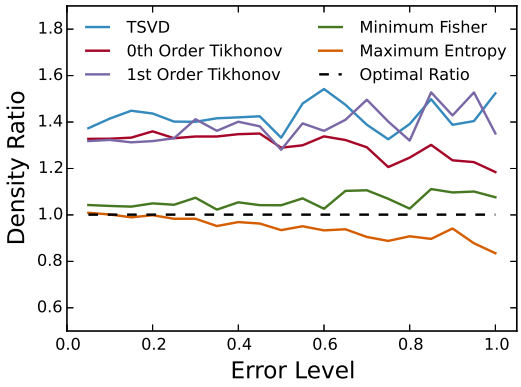
Figure 4. Tomographies of the beam distribution from figure 1(c) based on synthetic measurements using various inversion methods and noise levels. The inversion methods are SVD, 0'th and 1'st order Tikhonov (T0 and T1, respectively), minimum Fisher information (MFI) and maximum entropy (ME). The noise level k is defined in equation (34).

large noise levels in the measurements the absolute values of a density obtained from a reconstruction might be unreliable.

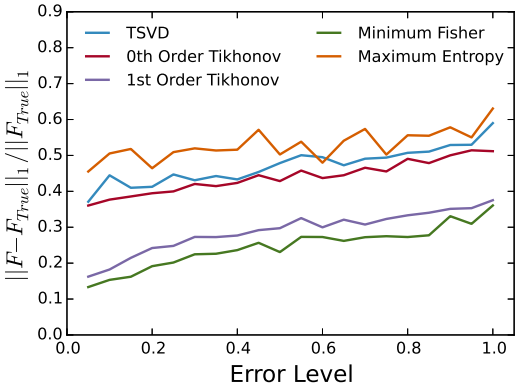
The uncertainties of the tomographies of the beam distribution as defined in section 4 are shown in figure 6 for a noise level of $k = 0.5$ in equation (34). The simulated photon noise is propagated to show the uncertainties in the tomography as described in equation (28). Here we disregard the model uncertainty for simplicity. They uncertainties are shown in figures 6(a) - 6(e). The uncertainty based on the noise in the data is almost independent of energy and largest for pitch values close to ± 1 for all but the maximum entropy regularization. Compared with the values of the tomographies in figure 4, the uncertainties are about one order of magnitude smaller, and smallest for first order Tikhonov and minimum Fisher information regularization. Figures 6(f) - 6(j) show the regularization error, e_{reg}. Positive values denote regions where too few ions are placed. Negative values denote regions where too many ions are placed. The beam peaks are seen in the regularization errors, especially for first order Tikhonov and minimum Fisher information regularization as these two are only able to resolve the spiky nature of the peaks for low noise levels. Adding the uncertainty contributions in quadrature gives e_{tot} which are shown in figures 6(k) - 6(o). It is seen that for all but ME, the main contribution to the uncertainty is the regularization error.



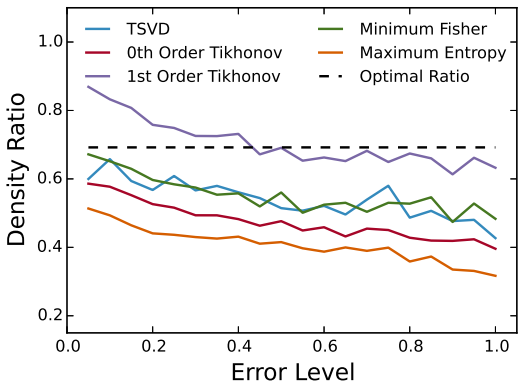
(a) Normalized absolute difference, Gaussian



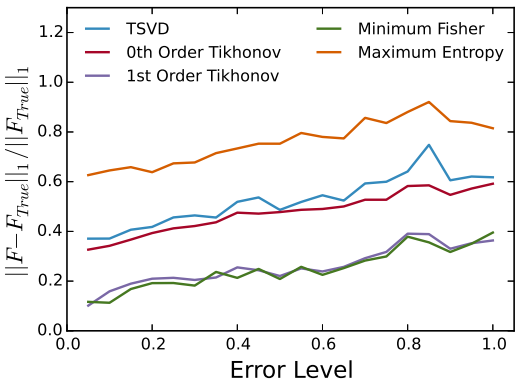
(b) Density ratio, Gaussian



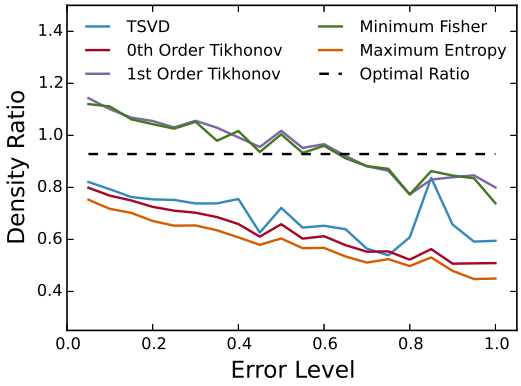
(c) Normalized absolute difference, bi-Maxwellian



(d) Density ratio, bi-Maxwellian

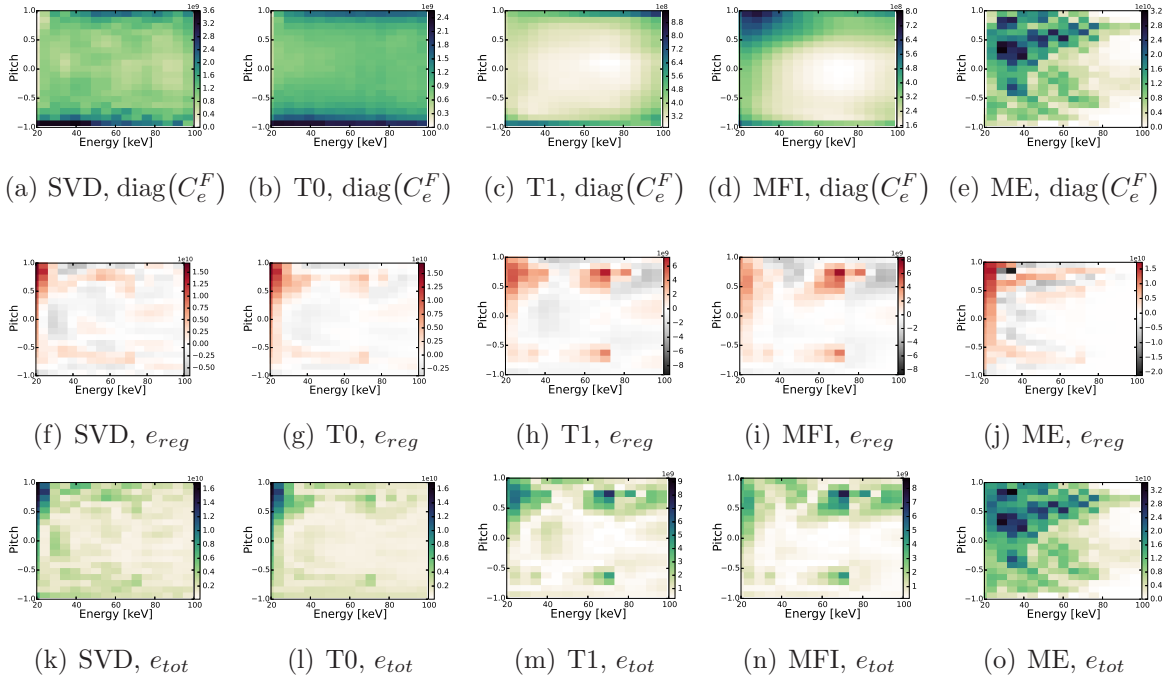


(e) Normalized absolute difference, NBI



(f) Density ratio, NBI

Figure 5. Performance parameters of the reconstructions of the test distributions.



(a) SVD, $\mathrm{diag}(C_e^F)$ (b) T0, $\mathrm{diag}(C_e^F)$ (c) T1, $\mathrm{diag}(C_e^F)$ (d) MFI, $\mathrm{diag}(C_e^F)$ (e) ME, $\mathrm{diag}(C_e^F)$

(f) SVD, e_{reg} (g) T0, e_{reg} (h) T1, e_{reg} (i) MFI, e_{reg} (j) ME, e_{reg}

(k) SVD, e_{tot} (l) T0, e_{tot} (m) T1, e_{tot} (n) MFI, e_{tot} (o) ME, e_{tot}

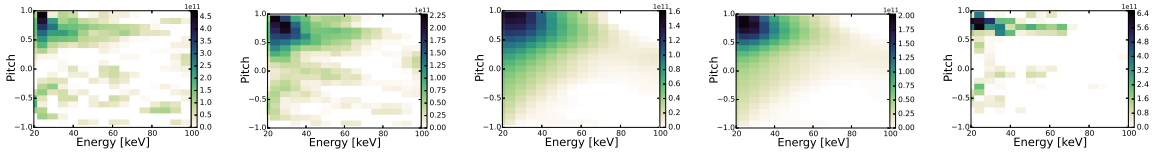
Figure 6. Uncertainties for the tomographies of the beam distribution. The inversion methods are SVD, 0'th and 1'st order Tikhonov (T0 and T1, respectively), minimum Fisher information (MFI) and maximum entropy (ME). All uncertainties are calculated for a noise level of $k = 0.5$.

6. Tomographies of a measured sawtooth crash

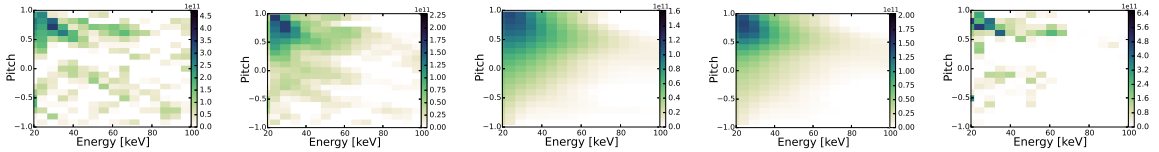
A sawtooth crash is a periodic plasma instability which occurs when the central safety factor drops below one. It changes the magnetic field topology and has been observed to redistribute particles and energy from the center of the plasma. It has furthermore been observed on several machines that passing fast ions are redistributed more strongly compared to trapped ions [9,37,38]. Figure 7 shows tomographies based on experimental data. They are calculated using the different methods on FIDA spectra measured just before and after a sawtooth crash in ASDEX Upgrade discharge 31557 at 2.25 s. The corresponding total measures of uncertainty as described in equation (31) are shown in figure 8. Common for all regularization methods, the fast ion density drops significantly during the sawtooth crash. To further investigate the velocity-space dependence of the change in the fast-ion distribution function, we calculate the relative change:

$$(\Delta F)_{rel} = \frac{F_{after} - F_{before}}{F_{before}}. \tag{35}$$

The relative change is calculated for every regularization method and plotted in figures 9(a)-9(e). The velocity-space dependence of the relative change is especially clear in the 1'st order Tikhonov and the minimum Fisher information figures as the amount of jitter in these tomographies are significantly less pronounced compared to the other methods. Both 1'st order Tikhonov and minimum Fisher information suggest that the passing ions are redistributed more compared with the passing ions. Similar trends were

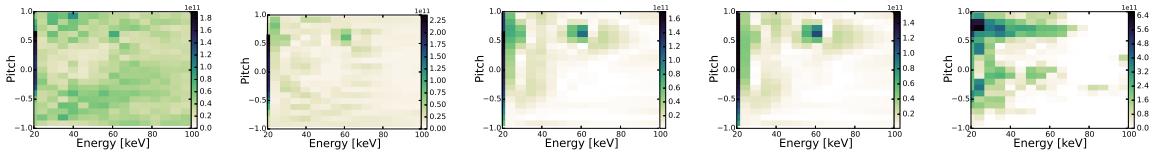


(a) SVD before. (b) T0 before. (c) T1 before. (d) MFI before. (e) ME before.

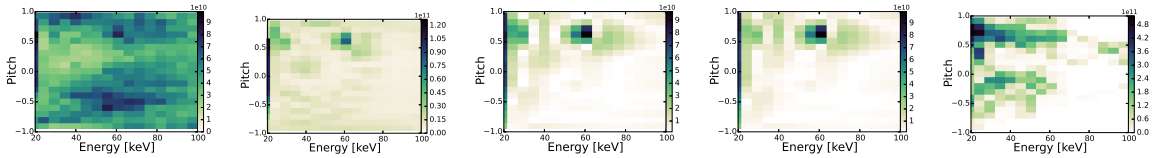


(f) SVD after. (g) T0 after. (h) T1 after. (i) MFI after. (j) ME after.

Figure 7. Tomographies before and after a sawtooth crash calculated using the different regularization methods.

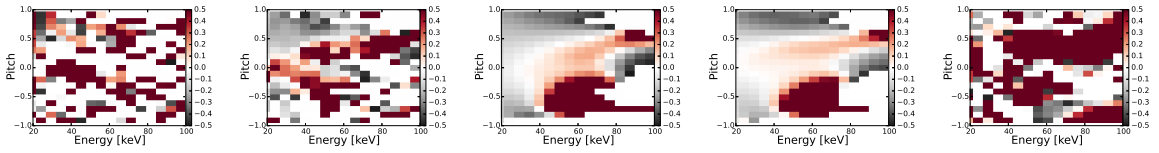


(a) SVD, e_{tot} before. (b) T0, e_{tot} before. (c) T1, e_{tot} before. (d) MFI, e_{tot} before. (e) ME, e_{tot} before.

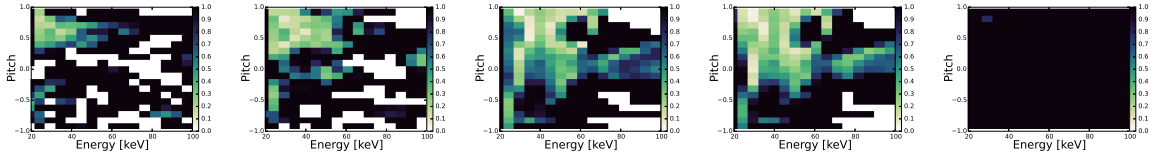


(f) SVD, e_{tot} after. (g) T0, e_{tot} after. (h) T1, e_{tot} after. (i) MFI, e_{tot} after. (j) ME, e_{tot} after.

Figure 8. Measures of uncertainties using the different regularization methods calculated using equation (31).



(a) SVD $(\Delta_F)_{rel}$ (b) T0 $(\Delta_F)_{rel}$ (c) T1 $(\Delta_F)_{rel}$ (d) MFI $(\Delta_F)_{rel}$ (e) ME $(\Delta_F)_{rel}$



(f) SVD $\sigma_{(\Delta_F)_{rel}}$ (g) T0 $\sigma_{(\Delta_F)_{rel}}$ (h) T1 $\sigma_{(\Delta_F)_{rel}}$ (i) MFI $\sigma_{(\Delta_F)_{rel}}$ (j) ME $\sigma_{(\Delta_F)_{rel}}$

Figure 9. Relative change of the fast-ion velocity-space distribution function.

observed previously using singular value decomposition [9] and a variant of a first-order Tikhonov [10] where different regularization levels were chosen rather than set by the L-curve method. For SVD the L-curve method does not show the stronger redistribution of passing particles for this crash. Figures 9(f)-9(j) show the uncertainties of the relative change as a function of energy and pitch calculated using the total errors shown in figure 8. The minimum Fisher information regularization and the first-order Tikhonov regularization show these trends with confidence even when the uncertainty includes the regularization uncertainty.

7. Discussion

Both the regularization uncertainty and the forward model uncertainty require an estimate of the true fast-ion velocity distribution function. For actual measurements where F_{true} is unknown, the total uncertainty, e_{tot}, will be biased by the estimate of F_{true}. However, the uncertainty will have approximately the correct magnitude if a good estimate of F_{true} can be calculated. Furthermore it is useful to gain insight in to whether the uncertainties obtained with a given regularization strength is dominated by the uncertainties in the measurements or the regularization itself.

It is seen that when the noise level is not too large, the first order Tikhonov and minimum Fisher information regularization methods can reconstruct the overall shape of the true distribution function very well. However, they lack in capability to resolve very fine and detailed features. It is seen that the absolute values of a derived quantity such as the fast-ion density depends on the noise level in the data. However, we find that the ratio of such quantities to less extend depends on the specific noise level and regularization strength.

8. Conclusion

We have compared the performance of five different regularization methods for velocity-space tomography. In order to estimate the credibility of the presented analysis, uncertainties of the tomographies are defined and calculated taking into account the photon noise, uncertainties in the forward model as well as uncertainty introduced by the regularization methods themselves. It is found that for the regularization level used here, the uncertainty introduced by the regularization methods are the major contribution. The performance is tested using synthetic data calculated using a realistic transfer-matrix from the five-view FIDA-system at ASDEX Upgrade. It is found that the regularization methods which penalize large gradients perform best for realistic test functions. Furthermore, the various methods are applied to actual FIDA measurements obtained in ASDEX Upgrade discharge #31557 just before and just after a sawtooth crash. Using velocity-space tomography it is possible to investigate the velocity-space dependence of the fast-ion redistribution. We find that sawtooth crashes at ASDEX Upgrade affect passing ions more than trapped ions.

Acknowledgement

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

[1] W W Heidbrink, K H Burrell, Y Luo, N a Pablant, and E Ruskov. Hydrogenic fast-ion diagnostic using Balmer-alpha light. *Plasma Physics and Controlled Fusion*, 46(12):1855–1875, December 2004.

[2] W W Heidbrink. Fast-ion Dα measurements of the fast-ion distribution (invited). *The Review of scientific instruments*, 81(10):10D727, October 2010.

[3] J. Egedal and H. Bindslev. Reconstruction of gyrotropic phase-space distributions from one-dimensional projections. *Physics of Plasmas*, 11(5):2191, 2004.

[4] M. Salewski, S.K. Nielsen, H. Bindslev, V. Furtula, N.N. Gorelenkov, S.B. Korsholm, F. Leipold, F Meo, P.K. Michelsen, D. Moseev, and M. Stejner. On velocity space interrogation regions of fast-ion collective Thomson scattering at ITER. *Nuclear Fusion*, 51(8):083014, August 2011.

[5] M. Salewski, B. Geiger, S.K. Nielsen, H. Bindslev, M. García-Muñoz, W.W. Heidbrink, S.B. Korsholm, F. Leipold, F. Meo, P.K. Michelsen, D. Moseev, M. Stejner, and G. Tardini. Tomography of fast-ion velocity-space distributions from synthetic CTS and FIDA measurements. *Nuclear Fusion*, 52(10):103008, October 2012.

[6] M. Salewski, B. Geiger, S.K. Nielsen, H. Bindslev, M. García-Muñoz, W.W. Heidbrink, S.B. Korsholm, F. Leipold, J. Madsen, F. Meo, P.K. Michelsen, D. Moseev, M. Stejner, and G. Tardini. Combination of fast-ion diagnostics in velocity-space tomographies. *Nuclear Fusion*, 53(6):063019, June 2013.

[7] M. Salewski, B. Geiger, a.S. Jacobsen, M. García-Muñoz, W.W. Heidbrink, S.B. Korsholm, F. Leipold, J. Madsen, D. Moseev, S.K. Nielsen, J. Rasmussen, M. Stejner, G. Tardini, and M. Weiland. Measurement of a 2D fast-ion velocity distribution function by tomographic inversion of fast-ion D-alpha spectra. *Nuclear Fusion*, 54(2):023005, February 2014.

[8] M Salewski, B Geiger, W W Heidbrink, A S Jacobsen, S B Korsholm, F Leipold, J Madsen, D Moseev, S K Nielsen, J Rasmussen, L Stagner, D Steeghs, M Stejner, G Tardini, and M Weiland. Doppler tomography in fusion plasmas and astrophysics. *Plasma Physics and Controlled Fusion*, 57(1):014021, January 2015.

[9] B. Geiger, M. Weiland, A.S. Jacobsen, D. Rittich, R. Dux, R. Fischer, C. Hopf, M. Maraschek, R.M. McDermot, S.K. Nielsen, T. Odstrcil, M. Reich, F. Ryter, M. Salewski, P.A. Schneider, and G. Tardini. Fast-ion transport and neutral beam current drive in ASDEX Upgrade. *Nuclear Fusion*, 55:083001, 2015.

[10] M. Weiland, B. Geiger, A.S. Jacobsen, M. Reich, M. Salewski, and T. Odstrcil. Velocity-space resolved fast-ion measurements from five FIDA views at the tokamak ASDEX Upgrade. submitted.

[11] W W Heidbrink, Y Luo, K H Burrell, R W Harvey, R I Pinsker, and E Ruskov. Measurements of fast-ion acceleration at cyclotron harmonics using Balmer-alpha spectroscopy. *Plasma Physics and Controlled Fusion*, 49(9):1457–1475, September 2007.

[12] M Salewski, B Geiger, D Moseev, W W Heidbrink, a S Jacobsen, S B Korsholm, F Leipold, J Madsen, S K Nielsen, J Rasmussen, M Stejner, and M Weiland. On velocity-space sensitivity of fast-ion D-alpha spectroscopy. *Plasma Physics and Controlled Fusion*, 56(10):105005, October 2014.

[13] W W Heidbrink, J R Ferron, C T Holcomb, M a Van Zeeland, X Chen, C M Collins, a Garofalo,

X Gong, B a Grierson, M Podestà, L Stagner, and Y Zhu. Confinement degradation by Alfvén-eigenmode induced fast-ion transport in steady-state scenario discharges. *Plasma Physics and Controlled Fusion*, 56(9):095030, 2014.

[14] a. S. Jacobsen, M. Salewski, J. Eriksson, G. Ericsson, a. Hjalmarsson, S. B. Korsholm, F. Leipold, S. K. Nielsen, J. Rasmussen, and M. Stejner. Velocity-space sensitivity of the time-of-flight neutron spectrometer at JETa). *Review of Scientific Instruments*, 85(11):11E103, November 2014.

[15] A.S. Jacobsen, M. Salewski, J. Eriksson, G. Ericsson, S.B. Korsholm, F. Leipold, S.K. Nielsen, J. Rasmussen, and M. Stejner. Velocity-space sensitivity of neutron spectrometry measurements. *Nuclear Fusion*, 55(5):053013, May 2015.

[16] M. Salewski, M. Nocente, G. Gorini, A.S. Jacobsen, V. Kiptily, S. Korsholm, F. Leipold, J. Madsen, D. Moseev, S.K. Nielsen, J. Rasmussen, and M. Stejner. Velocity-space observation regions of high-resolution two-step reaction gamma-ray spectroscopy. submitted.

[17] M Anton, H Weisen, M J Dutch, W von der Linden, F Buhlmann, R Chavan, B Marletaz, P Marmillod, and P Paris. X-ray tomography on the TCV tokamak. *Plasma Physics and Controlled Fusion*, 38(11):1849–1878, November 1996.

[18] T. Craciunescu, G. Bonheure, V. Kiptily, a. Murari, I. Tiseanu, and V. Zoita. A comparison of four reconstruction methods for JET neutron and gamma tomography. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 605(3):374–383, July 2009.

[19] B Geiger, R Dux, R M McDermott, S Potzel, M Reich, F Ryter, M Weiland, D Wünderlich, and M Garcia-Munoz. Multi-view fast-ion D-alpha spectroscopy diagnostic at ASDEX Upgrade. *The Review of scientific instruments*, 84(11):113502, November 2013.

[20] B. Geiger, M. Garcia-Munoz, R. Dux, F. Ryter, G. Tardini, L. Barrera Orte, I.G.J. Classen, E. Fable, R. Fischer, V. Igochine, and R.M. McDermott. Fast-ion transport in the presence of magnetic reconnection induced by sawtooth oscillations in ASDEX Upgrade. *Nuclear Fusion*, 54(2):022005, February 2014.

[21] W.W. Heidbrink, M.a. Van Zeeland, B.a. Grierson, C.M. Muscatello, J.M. Park, C.C. Petty, R. Prater, and Y.B. Zhu. Initial measurements of the DIII-D off-axis neutral beams. *Nuclear Fusion*, 52(9):094005, September 2012.

[22] D. C. Pace, M. E. Austin, E. M. Bass, R. V. Budny, W. W. Heidbrink, J. C. Hillesheim, C. T. Holcomb, M. Gorelenkova, B. a. Grierson, D. C. McCune, G. R. McKee, C. M. Muscatello, J. M. Park, C. C. Petty, T. L. Rhodes, G. M. Staebler, T. Suzuki, M. a. Van Zeeland, R. E. Waltz, G. Wang, a. E. White, Z. Yan, X. Yuan, and Y. B. Zhu. Energetic ion transport by microturbulence is insignificant in tokamaks. *Physics of Plasmas*, 20(5):056108, 2013.

[23] a Bortolon, W W Heidbrink, and M Podestà. A tangentially viewing fast ion D-alpha diagnostic for NSTX. *The Review of scientific instruments*, 81(10):10D728, October 2010.

[24] C a Michael, N Conway, B Crowley, O Jones, W W Heidbrink, S Pinches, E Braeken, R Akers, C Challis, M Turnyanskiy, a Patel, D Muir, R Gaffka, and S Bailey. Dual view FIDA measurements on MAST. *Plasma Physics and Controlled Fusion*, 55(9):095007, September 2013.

[25] O M Jones, C a Michael, K G McClements, N J Conway, B Crowley, R J Akers, R J Lake, and S D Pinches. Fast-ion deuterium alpha spectroscopic observations of the effects of fishbones in the Mega-Ampere Spherical Tokamak. *Plasma Physics and Controlled Fusion*, 55(8):085009, August 2013.

[26] T. Ito, M. Osakabe, K. Ida, M. Yoshinuma, M. Kobayashi, M. Goto, S. Murakami, M. Isobe, S. Kobayashi, K. Toi, K. Ogawa, Y. Takeiri, and S. Okamura. Fast ion charge exchange spectroscopy adapted for tangential viewing geometry in LHD. *Review of Scientific Instruments*, 81(10):1–5, 2010.

[27] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical recipies: The art of scientific computing*. Cambridge university press, 2007.

[28] G. Strang. *Linear algebra and its applications*. Thomson Learning, 1988.

[29] Per Christian Hansen, Toke Koldborg Jensen, and Giuseppe Rodriguez. An adaptive pruning algorithm for the discrete L-curve criterion. *Journal of Computational and Applied Mathematics*, 198(2):483–492, 2007.

[30] P.C. Hansen. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. SIAM, 1998.

[31] SF Gull and J Skilling. Maximum entropy method in image processing. *Communications, Radar and Signal Processing, IEE Proceedings F*, 131(6):646–659, 1984.

[32] Adom Giffin and Ariel Caticha. Updating probabilities with data and moments. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, volume 954, pages 74–84, 2007.

[33] Andreas Wächter and Lorenz T Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, 106(1):25–57, 2006.

[34] Miles Lubin and Iain Dunning. Computing in operations research using julia. *INFORMS Journal on Computing*, 27(2):238–248, 2015.

[35] Alexei Pankin, Douglas McCune, Robert Andre, Glenn Bateman, and Arnold Kritz. The tokamak Monte Carlo fast ion module NUBEAM in the National Transport Code Collaboration library. *Computer Physics Communications*, 159(3):157–184, June 2004.

[36] W. von den Linden, V. Dose, and U. von Toussaint. *Bayesian probability theory: Applications in the physical sciences*. Cambridge university press, 2014.

[37] S.K. Nielsen, M. Salewski, H. Bindslev, a. Bürger, V. Furtula, M. Kantor, S.B. Korsholm, H.R. Koslowski, a. Krämer-Flecken, F. Leipold, F. Meo, P.K. Michelsen, D. Moseev, J.W. Oosterbeek, M. Stejner, and E. Westerhof. Dynamics of fast ions during sawtooth oscillations in the TEXTOR tokamak measured by collective Thomson scattering. *Nuclear Fusion*, 51(6):063014, June 2011.

[38] C M Muscatello, W W Heidbrink, Ya I Kolesnichenko, V V Lutsenko, M a Van Zeeland, and Yu V Yakovenko. Velocity-space studies of fast-ion transport at a sawtooth crash in neutral-beam heated plasmas. *Plasma Physics and Controlled Fusion*, 54(2):025006, February 2012.

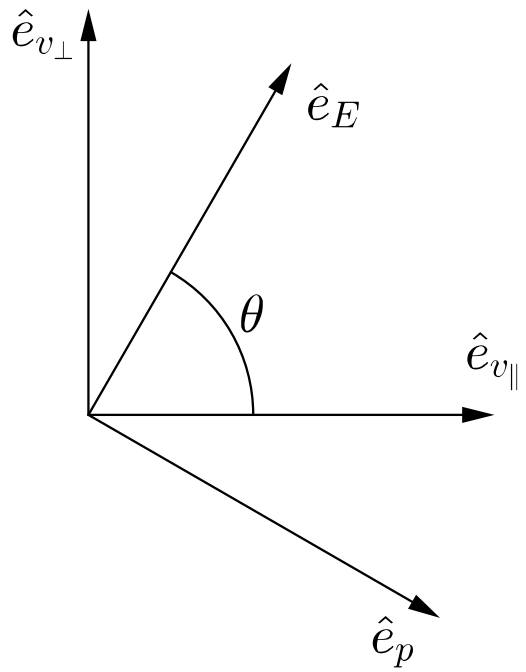


Figure A1. The relations between the unit vectors $\hat{e}_{v_\parallel}, \hat{e}_{v_\perp}, \hat{e}_E$ and $\hat{e}_p$.

Appendix A. Derivation of velocity-gradient in (E, p)-coordinates

To calculate the velocity-space gradient in (E, p)-coordinates, the gradient is transformed from $\left(v_\parallel, v_\perp\right)$-coordinates to (E, p)-coordinates. 1'st order regularization in $\left(v_\parallel, v_\perp\right)$-coordinates can be achieved by setting

$$L^T L = \nabla_{v_\parallel}^T \nabla_{v_\parallel} + \nabla_{v_\perp}^T \nabla_{v_\perp} \,, \tag{A.1}$$

where $\nabla_{v_\parallel}$ and $\nabla_{v_\perp}$ are finite difference matrix representations of the first-order differential operators. These have to be transformed to (E,p)-coordinates. It is apparent that the velocity-space gradient in (E, p)-coordinates has similarities to the real-space gradient in polar coordinates. The relations between the unit vectors $\hat{e}_{v_\parallel}, \hat{e}_{v_\perp}, \hat{e}_E$ and $\hat{e}_p$ are illustrated graphically in figure A1. The velocity-space gradient of f is

$$\nabla f = \left(\nabla_{v_\parallel} f\right) \hat{e}_{v_\parallel} + \left(\nabla_{v_\perp} f\right) \hat{e}_{v_\perp} = a \left(\nabla_E f\right) \hat{e}_E + b \left(\nabla_p f\right) \hat{e}_p \,, \tag{A.2}$$

where a and b are unknowns that must be calculated from the Jacobian. ∇_E and ∇_p are velocity-gradients along $\hat{e}_E$ and $\hat{e}_p$ respectively. Writing $\hat{e}_E$ and $\hat{e}_p$ as functions of

$\hat{e}_{v_\parallel}$ and $\hat{e}_{v_\perp}$ gives

$$\hat{e}_E = \cos(\theta)\,\hat{e}_{v_\parallel} + \sin(\theta)\,\hat{e}_{v_\perp} = p\,\hat{e}_{v_\parallel} + \sqrt{1-p^2}\,\hat{e}_{v_\perp}\,, \tag{A.3}$$

$$\hat{e}_p = \sin(\theta)\,\hat{e}_{v_\parallel} - \cos(\theta)\,\hat{e}_{v_\perp} = \sqrt{1-p^2}\,\hat{e}_{v_\parallel} - p\,\hat{e}_{v_\perp}\,, \tag{A.4}$$

where the relation $p = \cos(\theta)$ has been used. The gradient in energy is now found by dotting equation (A.2) with $\hat{e}_E$:

$$\begin{aligned}(\nabla f)\cdot\hat{e}_E \Rightarrow \left(\nabla_{v_\parallel} f\right) p + \left(\nabla_{v_\perp} f\right)\sqrt{1-p^2} &= a\,(\nabla_E f) \\ = a\left(\left(\nabla_{v_\parallel} f\right)\frac{\partial v_\parallel}{\partial E} + \left(\nabla_{v_\perp} f\right)\frac{\partial v_\perp}{\partial E}\right)&,\end{aligned} \tag{A.5}$$

To calculate the partial derivatives, the relations between $v_\parallel, v_\perp, E$ and p are needed:

$$v_\parallel = p\sqrt{\frac{2E}{m}} \tag{A.6}$$

$$v_\perp = \sqrt{1-p^2}\sqrt{\frac{2E}{m}}\,. \tag{A.7}$$

The partial derivatives are:

$$\frac{\partial v_\parallel}{\partial E} = \frac{p}{\sqrt{2mE}} \tag{A.8}$$

$$\frac{\partial v_\perp}{\partial E} = \frac{\sqrt{1-p^2}}{\sqrt{2mE}}\,. \tag{A.9}$$

Inserting equations (A.8) and (A.9) in equation (A.5) gives

$$\left(\nabla_{v_\parallel} f\right) p + \left(\nabla_{v_\perp} f\right)\sqrt{1-p^2} = a\left(\left(\nabla_{v_\parallel} f\right)\frac{p}{\sqrt{2mE}} + \left(\nabla_{v_\perp} f\right)\frac{\sqrt{1-p^2}}{\sqrt{2mE}}\right)\,. \tag{A.10}$$

Equation (A.10) is fulfilled for

$$a = \sqrt{2mE}\,. \tag{A.11}$$

Similarly, b can be found by dotting equation (A.2) with $\hat{e}_p$:

$$\begin{aligned}(\nabla f)\cdot\hat{e}_p \Rightarrow \left(\nabla_{v_\parallel} f\right)\sqrt{1-p^2} - \left(\nabla_{v_\perp} f\right) p &= b\,(\nabla_p f) \\ = b\left(\left(\nabla_{v_\parallel} f\right)\frac{\partial v_\parallel}{\partial p} + \left(\nabla_{v_\perp} f\right)\frac{\partial v_\perp}{\partial p}\right)&\,.\end{aligned} \tag{A.12}$$

The partial derivatives are:

$$\frac{\partial v_\parallel}{\partial p} = \sqrt{\frac{2E}{m}}\,, \tag{A.13}$$

$$\frac{\partial v_\perp}{\partial p} = -\frac{p}{\sqrt{1-p^2}}\sqrt{\frac{2E}{m}}\,. \tag{A.14}$$

Inserting equations (A.13) and (A.14) in equation (A.12) gives

$$\left(\nabla_{v_\parallel} f\right)\sqrt{1-p^2} - \left(\nabla_{v_\perp} f\right) p = b\left(\left(\nabla_{v_\parallel} f\right)\sqrt{\frac{2E}{m}} - \left(\nabla_{v_\perp} f\right)\frac{p}{\sqrt{1-p^2}}\sqrt{\frac{2E}{m}}\right). \tag{A.15}$$

Equation (A.15) is fulfilled for

$$b = \sqrt{\frac{m}{2E}}\sqrt{1-p^2}\,. \tag{A.16}$$

Thus, the velocity-space gradient in energy-pitch coordinates becomes

$$\nabla f = \sqrt{2mE}\left(\nabla_E f\right)\,\hat{e}_E + \sqrt{\frac{m}{2E}}\sqrt{1-p^2}\left(\nabla_p f\right)\,\hat{e}_p\,. \tag{A.17}$$