

A model-based, multichannel, real-time capable sawtooth crash detector

H. van den Brand^{1,2}, M. R. de Baar^{1,2}, M. van Berkel³, T. C. Blanken², F. Felici², E. Westerhof¹, M. Willensdorfer⁴, the ASDEX Upgrade team⁴ & the EUROfusion MST1 team⁵

¹ FOM Institute DIFFER - Dutch Institute for Fundamental Energy Research, P.O. Box 6336, 5600 HH Eindhoven, The Netherlands

² Control Systems Technology Group, Department of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

³ Department of Fundamental Electricity and Instrumentation, Vrije Universiteit Brussel (VUB), Pleinlaan 2, 1050 Brussels, Belgium

⁴ Max Planck Institut für Plasmaphysik, Boltzmannstraße 2, 85748 Garching bei München, Germany

⁵ See <http://www.euro-fusionscipub.org/mst1>

E-mail: H.vandenbrand@diffier.nl

Abstract. Control of the time between sawtooth crashes, necessary for ITER and DEMO, requires real-time detection of the moment of the sawtooth crash. In this paper, estimation of sawtooth crash times is demonstrated using the model-based Interacting Multiple Model (IMM) estimator, based on simplified models for the sawtooth crash. In contrast to previous detectors, this detector uses the spatial extent of the sawtooth crash as detection characteristic. The IMM estimator is tuned and applied to multiple ECE channels at once. A model for the sawtooth crash is introduced, which is used in the IMM algorithm. The IMM algorithm is applied to 7 datasets from the ASDEX Upgrade tokamak. Five crash models with different mixing radii are used. All sawtooth crashes that have been identified beforehand by visual inspection of the data, are detected by the algorithm. A few additional detections are made, which upon closer inspection are seen to be sawtooth crashes, which show a partial reconnection. A closer inspection of the detected normal crashes shows that about 42 % are not well fitted by any of the full reconnection models and show some characteristics of a partial reconnection. In some case, the measurement time is during the sawtooth crashes, which also results in an incorrect estimate of the mixing radius. For data provided at a sampling rate of 1 kHz, the run time of the IMM estimator is below 1 ms, thereby fulfilling real-time requirements.

Keywords: tokamak, sawtooth, detector, interacting multiple model estimator

PACS numbers: 52.55.Fa, 52.55.Tn

Submitted to: *Plasma Phys. Control. Fusion*

1. Introduction

In tokamak plasmas, internal kink modes can be driven unstable, leading to reconnection of flux surfaces and mixing of energy and particles in the plasma core [1, 2, 3]. This phenomenon, known as the sawtooth instability, has been studied for decades, resulting in an increased understanding of the nature and interactions of this instability [4, 5].

The growth of the internal kink mode is followed by a fast magnetic reconnection at $q = 1$, i.e. the sawtooth crash. During this sawtooth crash, core plasma pressure decreases. A redistribution of this core energy results in an increase of plasma pressure outside the core up to the mixing radius ρ_{mix} . At the inversion radius ρ_{inv} , for which $\rho_{inv} < \rho_{mix}$, the pressure difference is zero. In between subsequent sawtooth crashes, plasma pressure increases for $\rho < \rho_{inv}$ and decreases for $\rho_{inv} < \rho < \rho_{mix}$. As a result, the plasma pressure as a function of time shows a shape that is similar to the jagged teeth of a saw.

The sawtooth crash is known to trigger other magnetic instabilities such as edge localized modes [6] and neoclassical tearing modes (NTMs) [7, 8]. NTMs could trigger a disruption and should therefore be prevented [9, 10]. By limiting the time between two consecutive sawtooth crashes, the probability of triggering an NTM is decreased [11, 12, 13]. Additionally, preemptive electron cyclotron heating could be used to reduce the triggering probability of an NTM during the sawtooth crash [14]. The application of such strategies requires real-time knowledge of when the sawtooth crash occurs. Additionally, sawtooth crashes can help avoid core impurity accumulation, which could be controlled using a real-time detector for the crash time [15].

The time between two consecutive sawtooth crashes, also known as the sawtooth period, is altered by coupling high power waves into the plasma. Electron cyclotron current drive (ECCD) is efficient in modifying the local magnetic shear in the vicinity of the $q = 1$ rational surface [16, 17]. Ion cyclotron resonance heating (ICRH) is used to modify fast ion orbits, with the view to negate the stabilizing effect of the fast ion population [18]. Therefore, both ECCD and ICRH could be used for sawtooth control.

Control of the sawtooth crashes is foreseen for ITER and DEMO operation [19, 20, 21]. Most algorithms for sawtooth control rely on controlling the time between sawtooth crashes. The time between sawtooth crashes can be controlled without feedback, e.g. [22], or with feedback based on the time difference between consecutive crashes [17, 23, 24]. In both cases, a real-time detector for the sawtooth crash time is required to, for the former, assess if the control action has the desired result and, for the latter, to provide an input signal to the feedback controller.

The sawtooth crash time can be determined by the fast increase or decrease of the plasma pressure. Several detectors exist for detection of the sawtooth crash time, which rely on a single channel of temperature or density measurements as a function of time. For instance, the wavelet-based algorithm by Van Berkel *et al.* presents a detector optimized for steps in the input signals [25]. Gude *et al.* detects sawtooth crashes in a single channel of soft X-ray measurements and combines multiple single-channel

crash detections to separate actual crashes from measurement noise with automated algorithms [26]. A few methods based on edge detectors (band-pass and difference filters) are applied in real-time sawtooth control experiments [17, 27, 28]. Isei *et al* demonstrate a sawtooth detection algorithm based on a neural network [29] and Weisen *et al* determine the inversion radius using an SVD decomposition[30].

Most algorithms use data from multiple measurement channels at once. For instance, for the real-time detectors an average of the measurements is processed by the sawtooth detector, while Van Berkel *et al.* and Gude *et al.* combine the channels after applying edge detectors. However, the detection algorithms do not explicitly take into account the redistribution of plasma pressure due to the sawtooth crash. A sawtooth detector should detect if temperature is redistributed instead of detecting step-like deviations of a single channel. Therefore, in this paper, a model-based detection method for sawtooth crashes is presented. Hence, multiple measurement channels are used simultaneously for the detection of a sawtooth crash based on its pressure redistribution. The detection algorithm is designed to be real-time capable, for which it should be causal and fast. For non-real-time algorithms, in contrast, one could use non-causal methods and optimize for detection of sawtooth features, without compromising on the computational time.

In this paper, an interacting multiple model (IMM) estimator [31] is used for model-based detection of the sawtooth crash time. This allows for a detection of the sawtooth crash based on the redistribution of plasma pressure. At every time step, predictions of the measured quantity of multiple channels as a function of the spatial coordinate are made. These predictions are made using multiple, physics-based models, each with a different mixing radius. The likelihood that the measurements originate from a crash at a given mixing radius is assessed by comparing the measurements with the predictions from the multiple models. Based on these likelihoods, it is determined if a sawtooth crash occurred.

In the next section, the crash detection algorithm is introduced. In section 3, the crash detection algorithm is applied to electron cyclotron emission (ECE) data from the ASDEX Upgrade tokamak. Section 4 discusses the model and results on the measurement data. Concluding remarks are found in section 5.

2. Crash detection algorithm

The crash detection is based on the change of plasma pressure due to the sawtooth crash. This change is visible in multiple channels of temperature or density diagnostics. The measurements are a result of the underlying temperature and density profiles, which are assumed to be a function of the flux coordinate ρ which ranges from 0 at the plasma core to 1 at the plasma edge. At every time instant k , the measurements are compared to an estimation of the measurements. Using spatial discretization, an estimated temperature or density profile is represented using a finite number of basis functions. Spatial discretization is introduced in subsection 2.1.

A sawtooth crash results in a drop of temperature and density in the core of the plasma and a rise further outward. Between consecutive sawtooth crashes, the temperature and density change on a much slower diffusive time scale. This observation of a different temperature and density evolution with and without a sawtooth crash is used in subsection 2.2 to construct models for the different situations. A model is constructed for the temperature profile following a sawtooth crash.

Based on temperature measurements, the temperature profile evolution is predicted using multiple models, presented in subsection 2.2. For each model, the estimated temperature is compared with measurements \mathbf{y}_k to assess which model is the most likely. The results of the estimated temperatures are used as a starting point for the next time step and the interaction between different models is specified. All these steps are taken into account by the interacting multiple model (IMM) estimator, which is presented in subsection 2.3. Parameters for the IMM estimator, that are independent of the datasets, are selected in subsection 2.4.

2.1. Spatial discretization

In this subsection, spatial discretization of the temperature profile $T(\rho)$ at time instant k is considered. The same techniques are applicable to other profiles. Due to the magnetic topology of the tokamak, the temperature T is assumed to be a function of the flux coordinate ρ . Furthermore, the temperature profile $T(\rho)$ is assumed to be a smooth function of ρ . In this paper, an estimate of the temperature profile $\hat{T}_k(\rho)$ at time k is based on a fixed set of N_x basis functions $\Lambda_i(\rho)$ (where $i = 1, \dots, N_x$) and is given by

$$\hat{T}_k(\rho) = \sum_{i=1}^{N_x} \Lambda_i(\rho) x_{i,k} = \mathbf{C}(\rho) \mathbf{x}_k \quad (1)$$

where $x_{i,k}$ are the coefficients at time k . Therefore, the vector $\mathbf{x}_k = [x_{1,k} \dots x_{N_x,k}]^T$ contains all the information about the estimated temperature profile $\hat{T}_k(\rho)$. For a fixed position ρ , $\mathbf{C}(\rho)$ is a constant row vector with N_x elements.

For time instant k , a set of N_y temperature measurements at locations $\rho_{1,k}$ to $\rho_{N_y,k}$ are contained in a column vector \mathbf{y}_k . Estimated temperature measurements $\hat{\mathbf{y}}_k$ are given by

$$\hat{\mathbf{y}}_k = \begin{bmatrix} \hat{T}_k(\rho_{1,k}) \\ \vdots \\ \hat{T}_k(\rho_{N_y,k}) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N_x} \Lambda_i(\rho_{1,k}) x_{i,k} \\ \vdots \\ \sum_{i=1}^{N_x} \Lambda_i(\rho_{N_y,k}) x_{i,k} \end{bmatrix} = \mathbf{C}_k \mathbf{x}_k \quad (2)$$

where \mathbf{C}_k is a time dependent matrix of dimensions $N_y \times N_x$, depending on the time-varying measurement positions $\rho_{j,k}$. If \mathbf{C}_k has a rank of N_x , a least-squares estimate of \mathbf{x}_k can be used to determine the best fit of the coefficients \mathbf{x}_k to the measurements \mathbf{y}_k .

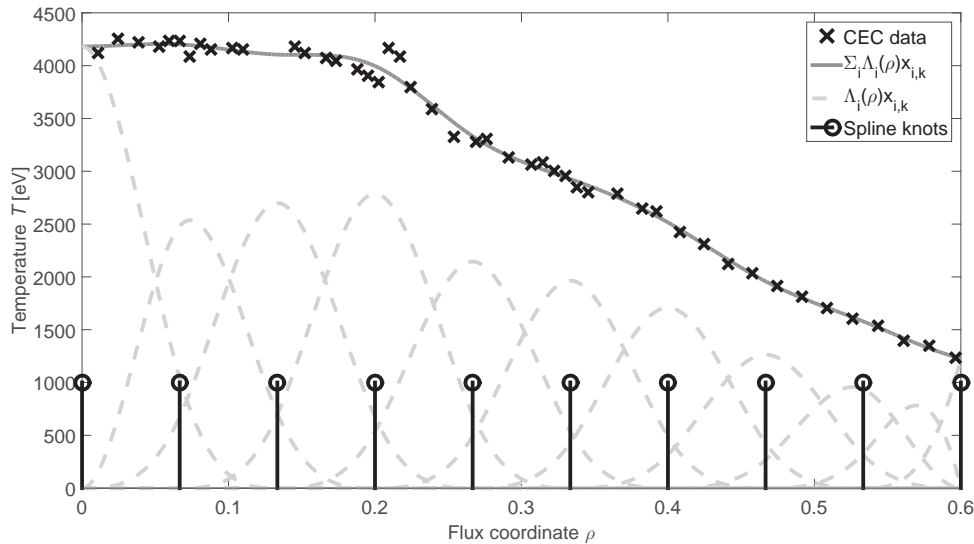


Figure 1. Cubic B-spline approximation of ECE data of ASDEX Upgrade discharge 30538 at 4.2 s, shown in black crosses. The approximation is shown in dark grey and the 11 basis functions multiplied by their associated coefficients, are shown in a light grey dashed line. The splines are generated using a set of 10 equidistant knots, indicated by black vertical lines. The derivative of the spline interpolation is forced to zero at $\rho = 0$.

As an example, electron cyclotron emission (ECE) measurements are approximated using cubic B-splines, which are piecewise cubic polynomials [32]. It is assumed that ECE is well localized (which holds for optically thick plasmas [33]), that ECE is emitted from the cold resonance positions, that the electrons are in thermal equilibrium, and that all measurement locations are positioned on a horizontal line through the plasma center. The magnetic field strength is approximated by $B(R) = B_0 R_0 / R$ with R the major radius of the plasma, B_0 the magnetic field strength at the magnetic axis, and R_0 the major radius of the magnetic axis. N_y measurements are considered and, with the assumptions listed above, the locations ρ_j are determined using

$$\rho_j = \left\| \left\| \frac{n_j e B_0 R_0}{a m_e 2\pi f_j} - \frac{R_0}{a} \right\| \right\| \quad (3)$$

in which n_j is the harmonic number, e is the elementary charge, a is the minor radius of the plasma, m_e is the electron mass, and f_j is the frequency of ECE that is collected in measurement channel j . Figure 1 shows the approximation of ECE data using a cubic B-spline with $N_x = 11$ basis functions.

2.2. Sawtooth crash model

In the previous subsection, a discretization of the temperature profiles is introduced such that a smooth temperature profile at time instant k can be described using the vector \mathbf{x}_k . The continuous temperature evolution is evaluated at times kt_s , with k a

time index number and t_s a fixed sample time. As a result, the temperature profile evolution is described by \mathbf{x}_k for $k = 1, 2, 3, \dots$, where the fixed sample time is omitted. A prediction of the temperature profile at $k + 1$ is made using the coefficients \mathbf{x}_k at time k , such that

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k, \quad (4)$$

and the measurements are predicted by $\hat{\mathbf{y}}_{k+1} = \mathbf{C}_k\mathbf{x}_{k+1}$. The matrix \mathbf{A} is chosen to reflect the expected evolution of the temperature. In absence of a sawtooth crash, the predicted temperature profile is taken equal to the previous temperature profile, such that

$$\mathbf{x}_{k+1} = \mathbf{x}_k = \mathbf{A}^{(\text{NoCrash})}\mathbf{x}_k := \mathbf{I}_{N_x}\mathbf{x}_k, \quad (5)$$

with \mathbf{I}_{N_x} an $N_x \times N_x$ identity matrix. The temperature profile is assumed constant in time due to a much shorter sampling time compared to the time scale of temperature evolution.

During a sawtooth crash, a fast reconnection results in a flattening of the temperature profile between $\rho = 0$ and $\rho = \rho_{mix}$. Therefore, the matrix $\mathbf{A}^{(\text{NoCrash})}$ cannot be used to predict the temperature evolution in this case. Instead, it is assumed that directly after the sawtooth crash the thermal energy in a volume within $\rho = \rho_{mix}$ is redistributed evenly over the volume. Density up to $\rho = \rho_{mix}$ is taken constant. Following this assumption, only the temperature evolution has to be considered for sawtooth detection. The volume contained within $\rho = \rho_{mix}$ is calculated using $V = \int_0^{\rho_{mix}} \pi \rho d\rho$. It is assumed that the volume can be approximated as being based on a circular cross section of which the shape remains unaltered, i.e. no changes in κ and δ . The flat temperature is calculated by dividing the energy inside $\rho = \rho_{mix}$ by the volume

$$T_{k,flat}(\rho_{mix}) = \frac{\pi}{V} \int_0^{\rho_{mix}} T_k(\rho) \rho d\rho = \mathbf{V}^{(\text{Crash})}(\rho_{mix})\mathbf{x}_k, \quad (6)$$

where in the last step it is used that the profile $T_k(\rho)$ is fully determined by the coefficients \mathbf{x}_k and each basis function is integrated exactly. For a given set of N_x basis functions $\mathbf{V}^{(\text{Crash})}(\rho_{mix})$ is a constant row vector with N_x elements. The row vector $\mathbf{V}^{(\text{Crash})}(\rho_{mix})$ is applied for positions $\rho \leq \rho_{mix}$. This results in the matrix $\mathbf{A}^{(\text{Crash})}(\rho_{mix})$, as defined in (A.1) of Appendix A, that determines the coefficients \mathbf{x}_{k+1} , based on \mathbf{x}_k , just after a sawtooth crash with ρ_{mix} . The number of knots, used in the spline approximation, limits the variety of mixing radii that can be accurately portrayed.

The temperature profile at time instant $k + 1$ can be predicted using

$$\mathbf{x}_{k+1} = \begin{cases} \mathbf{A}^{(\text{NoCrash})}\mathbf{x}_k & \text{No crash} \\ \mathbf{A}^{(\text{Crash})}(\rho_{mix})\mathbf{x}_k & \text{Crash at } \rho_{mix} \end{cases} \quad (7)$$

with the matrices $\mathbf{A}^{(\text{NoCrash})}$ and $\mathbf{A}^{(\text{Crash})}(\rho_{mix})$ as defined in (5) and (A.1), respectively.

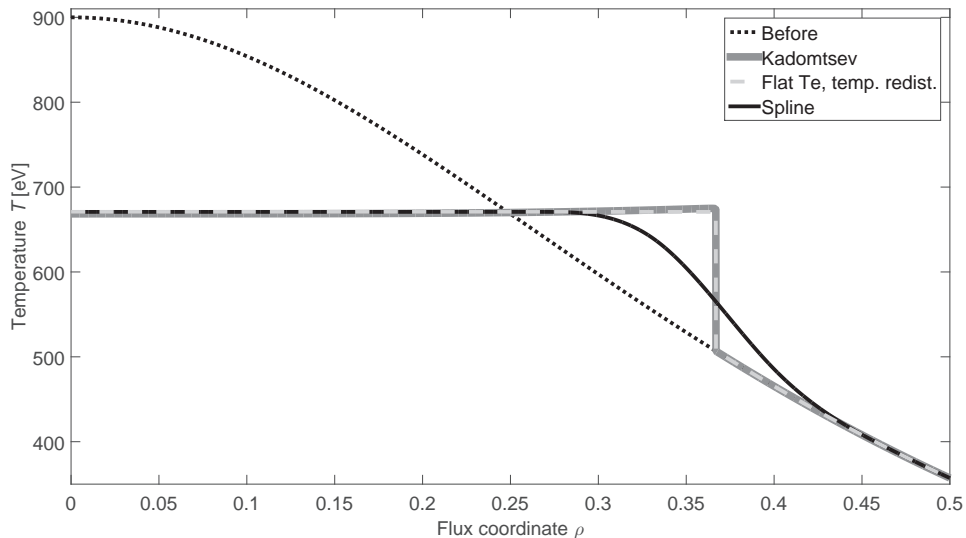


Figure 2. Temperature as a function of the flux coordinate ρ before (dotted black line) and after a sawtooth crash following three different models. The result of Kadomtsev’s full reconnection model is shown in dark grey and is overlapped for the largest part by the model assuming a flat redistribution of the temperature up to the mixing radius (dashed light-grey line). A spline-based crash model is shown with a solid black line. The full reconnection model uses the parameters $T_0 = 900$ eV, $q_a = 4$, $a = 0.5$ m, $R_0 = 1.65$ m, and $B_0 = 2.3$ T to determine the temperature and safety factor profile in (8).

The assumption that after a sawtooth crash a flat temperature profile is achieved is verified by comparing with the full reconnection model of Kadomtsev [34, 35]. For this comparison, the safety factor profile and temperature profile given by

$$\begin{aligned} q(\rho) &= q_0 + (q_a - q_0) \rho^2 \\ T(\rho) &= T_0 (1 + q_a \rho^2)^{-4/3} \end{aligned} \quad (8)$$

are used, where q_a is the safety factor at the edge of the plasma, $q_0 = \frac{q_a}{1+q_a}$, and T_0 is the core temperature. The full reconnection model is applied following the procedure in [35]. The temperature after the crash at position $\rho_c^2 = \rho_2^2 - \rho_1^2$ is the result of reconnecting flux tubes at positions ρ_1 and ρ_2 , which have the same helical flux, but are on opposite sides of the $q = 1$ surface. Taking into account the volume of the reconnecting flux tubes, the temperature after reconnection is calculated.

Figure 2 shows the temperature profile given by (8) with a black dashed line. Three different sawtooth crash models are applied to the temperature profile. The full reconnection model is shown as a dark grey line. This crash model results in a temperature profile that shows a higher temperature at $\rho = \rho_{mix}$ compared to the temperature at $\rho = 0$. At the mixing radius there is an abrupt step after which the original temperature profile is followed.

A flat redistribution of temperature following (6) up to the mixing radius is depicted as a dashed light-grey line. This model is a good approximation of the Kadomtsev model,

while using the mixing radius ρ_{mix} as its only parameter. A limited set of 11 basis functions is also used to describe the temperature profile. With these basis functions, the crash is modelled by applying the crash matrix in (A.1) to the coefficients \mathbf{x}_k , such that the post crash temperatures are given by $C_k A^{(\text{Crash})}(\rho_{mix}) \mathbf{x}_k$, with C_k defined as in (2). The result is shown with the black line in figure 2. As a result of smoothness of the basis functions, the step at the mixing radius cannot be represented in the crash model based on the spline. However, the step in the post-crash temperature at ρ_{mix} is not observed in experiments. The high temperature gradient drives diffusive processes and, thereby, results in a post-crash temperature profile that is smoothed, similar to the spline-based temperature profile. In this paper, the smoothed spline-based temperature profile is a result of the smoothness and limited number of the underlying splines and does not take the diffusive processes into account.

2.3. Interacting multiple model estimator

Given measurements \mathbf{y}_k at time k , it needs to be determined if a sawtooth crash occurred or not. To do so, the measurements are compared with predicted temperature profiles based on various hypotheses for the occurrence of a crash. Multiple hypotheses for a crash are used, each with different mixing radii. The predicted temperature profile is determined by previous measurements and the expected time evolution of the temperature profile, following (7).

The true temperature profile and the associated coefficients \mathbf{x}_k are unknown, but information about the profile is available through the measurements \mathbf{y}_k . If the temperature evolution follows a single model A, the measurements \mathbf{y}_k can be described by a process

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k \end{aligned} \quad (9)$$

where \mathbf{w}_k is an unknown process noise with N_x elements and \mathbf{v}_k is an unknown measurement noise with N_y elements. The process noise is assumed to be normally distributed with a known process noise covariance $\mathbb{E}(\mathbf{w}_k\mathbf{w}_k^T) = \mathbf{Q}_k$, where $\mathbb{E}(\cdot)$ is the expected value. The measurement noise is assumed to be normally distributed with a known measurement noise covariance $\mathbb{E}(\mathbf{v}_k\mathbf{v}_k^T) = \mathbf{R}_k$. Both process and measurement noise are assumed to be zero mean with a time-dependent noise covariance matrix.

For the system (9), the optimal estimator for the coefficients \mathbf{x}_k is the Kalman filter [36, 37]. The Kalman filter uses the matrix A to make an *a priori* estimate $\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1}$ at time k based on the estimate $\hat{\mathbf{x}}_{k-1|k-1}$ from the time step $k-1$. An *a posteriori* estimate $\hat{\mathbf{x}}_{k|k}$ is determined based on the difference between the measurements \mathbf{y}_k and the predicted measurements $\mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}$. The second index of the *a priori* estimate $\hat{\mathbf{x}}_{k|k-1}$ and *a posteriori* estimate $\hat{\mathbf{x}}_{k|k}$ indicates that all measurements up to, respectively, $k-1$ and k are taken into account. For a small measurement noise the *a posteriori* estimate $\hat{\mathbf{x}}_{k|k}$ is determined for the most part based on the measurements \mathbf{y}_k . In contrast, if the measurement noise is large and the process noise is small, the *a*

posteriori estimate $\hat{\mathbf{x}}_{k|k}$ is mostly determined based on the *a priori* estimate $\hat{\mathbf{x}}_{k|k-1}$. A filter gain determines the optimal trade-off between relying on measurements and the *a priori* estimate, based on the process and measurement noise covariances \mathbf{Q}_k and \mathbf{R}_k [36], such that the expected value of the difference between the *a posteriori* estimate $\hat{\mathbf{x}}_{k|k}$ and the actual coefficients \mathbf{x}_k is minimized.

A plasma with sawtooth crashes cannot be described using the model (9), as the matrix \mathbf{A} differs for time steps with and without a sawtooth crash. Instead, multiple hypotheses are formulated for the state the plasma is in. For each hypothesis i a separate model $\mathbf{A}^{(i)}$ is considered. *A priori* estimates $\hat{\mathbf{x}}_{k|k-1}^{(i)}$ and *a posteriori* estimates $\hat{\mathbf{x}}_{k|k}^{(i)}$ are determined for each hypothesis i using a single step of a Kalman filter. The probability $\mu_{k|k}^{(i)}$ that hypothesis i holds is evaluated by comparing predictions of the measurements $\mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}^{(i)}$ with the measurements \mathbf{y}_k . Based on the probabilities $\mu_{k|k}^{(i)}$ found for time step k and *a priori* expectations for the probability of the hypotheses in the next time step (contained in a matrix $\mathbf{\Pi}$), a *mixed* estimate $\bar{\mathbf{x}}_{k|k}^{(i)}$ (note the bar instead of the hat) is determined that serves as the starting point for the next step with the Kalman filter for hypothesis i . All these features are contained in the interacting multiple model (IMM) estimator [31].

For the detection of a sawtooth crash, N_h hypotheses are considered with models $\mathbf{A}^{(i)}$. The first hypothesis ($i = 1$) is that no sawtooth crash has occurred and for this hypothesis $\mathbf{A}^{(1)} = \mathbf{A}^{(\text{NoCrash})}$, given by (5). For the other hypotheses, a crash is expected at a mixing radius $\rho_{mix}^{(i)}$ and therefore $\mathbf{A}^{(i)} = \mathbf{A}^{(\text{Crash})}(\rho_{mix}^{(i)})$ given by (A.1). Based on the probabilities $\mu_{k|k}^{(i)}$, it is determined if a crash occurred between time step $k - 1$ and time step k . An *a priori* transition matrix $\mathbf{\Pi}$ determines which hypotheses are expected to be true in time step $k + 1$ given probabilities $\mu_{k|k}^{(i)}$ for all hypotheses i . The matrix $\mathbf{\Pi}$ is chosen to reflect that one does not expect a sawtooth crash to occur at time step $k + 1$ following a crash at any location at time step k and that the chance at finding no crash at any time step is higher than the chance of finding a sawtooth crash.

The IMM algorithm is started with an initial distribution of probabilities for the hypotheses and an initial estimate of the profile. With both values present, one step of the IMM algorithm from $k - 1$ to k consists of the following steps:

- **Mixing step:** Based on the probabilities $\mu_{k-1|k-1}^{(i)}$ and the *a priori* transition matrix $\mathbf{\Pi}$, the *mixed* estimates $\bar{\mathbf{x}}_{k-1|k-1}^{(i)}$ and an associated covariance matrix $\bar{\mathbf{P}}_{k-1|k-1}^{(i)} = \mathbb{E} \left(\bar{\mathbf{x}}_{k-1|k-1}^{(i)} (\bar{\mathbf{x}}_{k-1|k-1}^{(i)})^T \right)$ are determined. An element π_{ij} (from row i and column j of matrix $\mathbf{\Pi}$) contains the *a priori* probability that hypothesis i is true given that hypothesis j was true in the previous step. As part of the *mixing* step, *mixing* weights $\mu_{k|k-1}^{j|i}$ are determined which give the contribution of the *a posteriori* estimate $\bar{\mathbf{x}}_{k-1|k-1}^{(j)}$ to the *mixed* estimate $\bar{\mathbf{x}}_{k-1|k-1}^{(i)}$. *A priori* probabilities $\mu_{k|k-1}^{(i)}$ are calculated, which determine whether hypothesis i holds for the plasma at time k based on measurements up to time instant $k - 1$.
- **Prediction step:** The *a priori* estimate $\hat{\mathbf{x}}_{k|k-1}^{(i)}$ and an associated covariance matrix

Algorithm 1 One step of the interacting multiple model estimator [31]**Mixing step for all hypotheses**

$$\begin{aligned}
\text{A priori hypothesis probability} & \mu_{k|k-1}^{(i)} &= \sum_j \pi_{ij} \mu_{k-1|k-1}^{(j)} \\
\text{Mixing weights} & \mu_{k|k-1}^{j|i} &= \pi_{ij} \mu_{k-1|k-1}^{(j)} / \mu_{k|k-1}^{(i)} \\
\text{Mixed estimate} & \bar{\mathbf{x}}_{k-1|k-1}^{(i)} &= \sum_j \bar{\mathbf{x}}_{k-1|k-1}^{(j)} \mu_{k|k-1}^{j|i} \\
\text{Mixed process covariance} & \bar{\mathbf{P}}_{k-1|k-1}^{(i)} &= \sum_j \left[\mathbf{P}_{k-1|k-1}^{(j)} + \right. \\
& & \left. \left(\bar{\mathbf{x}}_{k-1|k-1}^{(i)} - \bar{\mathbf{x}}_{k-1|k-1}^{(j)} \right) \left(\bar{\mathbf{x}}_{k-1|k-1}^{(i)} - \bar{\mathbf{x}}_{k-1|k-1}^{(j)} \right)^T \right] \mu_{k|k-1}^{j|i}
\end{aligned}$$

Prediction step for all hypotheses

$$\begin{aligned}
\text{A priori estimate} & \hat{\mathbf{x}}_{k|k-1}^{(i)} &= \mathbf{A}^{(i)} \bar{\mathbf{x}}_{k-1|k-1}^{(i)} \\
\text{A priori process covariance} & \mathbf{P}_{k|k-1}^{(i)} &= \mathbf{A}^{(i)} \bar{\mathbf{P}}_{k-1|k-1}^{(i)} \left(\mathbf{A}^{(i)} \right)^T + \mathbf{Q}_k
\end{aligned}$$

Update step for all hypotheses

$$\begin{aligned}
\text{Measurement residual} & \mathbf{z}_k^{(i)} &= \mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}^{(i)} \\
\text{Residual covariance} & \mathbf{S}_k^{(i)} &= \mathbf{C}_k \mathbf{P}_{k|k-1}^{(i)} \mathbf{C}_k^T + \mathbf{R}_k \\
\text{Filter gain} & \mathbf{K}_k^{(i)} &= \mathbf{P}_{k|k-1}^{(i)} \mathbf{C}_k^T \left(\mathbf{S}_k^{(i)} \right)^{-1} \\
\text{A posteriori estimate} & \hat{\mathbf{x}}_{k|k}^{(i)} &= \hat{\mathbf{x}}_{k|k-1}^{(i)} + \mathbf{K}_k^{(i)} \mathbf{z}_k^{(i)} \\
\text{A posteriori process covariance} & \mathbf{P}_{k|k}^{(i)} &= \mathbf{P}_{k|k-1}^{(i)} - \mathbf{K}_k^{(i)} \mathbf{S}_k^{(i)} \left(\mathbf{K}_k^{(i)} \right)^T
\end{aligned}$$

Probability update for all hypotheses

$$\begin{aligned}
\text{Hypothesis likelihood} & L_k^{(i)} &= (2\pi)^{-\frac{N_x}{2}} \left| \mathbf{S}_k^{(i)} \right|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \left(\mathbf{z}_k^{(i)} \right)^T \left(\mathbf{S}_k^{(i)} \right)^{-1} \mathbf{z}_k^{(i)} \right] \\
\text{A posteriori hypothesis probability} & \mu_{k|k}^{(i)} &= \frac{\mu_{k|k-1}^{(i)} L_k^{(i)}}{\sum_j \mu_{k|k-1}^{(j)} L_k^{(j)}}
\end{aligned}$$

$\mathbf{P}_{k|k-1}^{(i)} = \mathbb{E} \left(\hat{\mathbf{x}}_{k|k-1}^{(i)} \left(\hat{\mathbf{x}}_{k|k-1}^{(i)} \right)^T \right)$ are calculated for hypotheses $i = 1, \dots, N_h$.

- **Update step:** The measurement residual $\mathbf{z}_k^{(i)}$ and an associated covariance $\mathbf{S}_k^{(i)} = \mathbb{E} \left(\mathbf{z}_k^{(i)} \left(\mathbf{z}_k^{(i)} \right)^T \right)$ are determined. A filter gain $\mathbf{K}_k^{(i)}$ is calculated and used to determine the *a posteriori* estimate $\hat{\mathbf{x}}_{k|k}^{(i)}$ and an associated covariance $\mathbf{P}_{k|k}^{(i)} = \mathbb{E} \left(\hat{\mathbf{x}}_{k|k}^{(i)} \left(\hat{\mathbf{x}}_{k|k}^{(i)} \right)^T \right)$.
- **Probability update:** The likelihood $L_k^{(i)}$ of measuring \mathbf{y}_k assuming hypothesis i holds is evaluated, using a multivariate normal distribution with zero mean and covariance $\mathbf{S}_k^{(i)}$, and, based on the likelihoods and *a priori* probabilities $\mu_{k|k-1}^{(i)}$, the *a posteriori* probabilities $\mu_{k|k}^{(i)}$ are evaluated.

The steps described above are listed in algorithm 1 for time step k with all equations. Detection of a crash is done by considering the *a posteriori* hypothesis probabilities and comparing the values to a threshold. To run the IMM estimator the following parameters need to be chosen:

- (i) Number of hypotheses N_h
- (ii) The *a priori* transition matrix $\mathbf{\Pi}$
- (iii) Initial hypothesis probabilities $\mu_{1|1}^{(i)}$ for all hypotheses i
- (iv) Initial estimates $\hat{\mathbf{x}}_{1|1}^{(i)}$ for all hypotheses i
- (v) Initial process covariances $\mathbf{P}_{1|1}^{(i)}$ for all hypotheses i
- (vi) Prediction matrices $\mathbf{A}^{(i)}$ for all hypotheses i , with the associated mixing radii

- (vii) The process noise covariance \mathbf{Q}_k
- (viii) The measurement noise covariance \mathbf{R}_k
- (ix) The output matrix \mathbf{C}_k
- (x) A threshold on the probabilities $\mu_{k|k}^{(i)}$ to determine if a sawtooth crash occurred

Choices for items (i) to (vi) and item (x) are introduced in the next subsection. Items (vii) to (ix) are selected in subsection 3.2 after the datasets are introduced.

2.4. Selection of dataset-independent IMM parameters

In the previous subsection, the IMM estimator is introduced and an overview is given of the parameters that need to be set for application of the estimator. The parameters that do not depend on the datasets are chosen in this subsection.

The matrices $\mathbf{A}^{(i)}$ and the output matrix \mathbf{C}_k can only be properly selected if a choice is made as to what the estimates $\hat{\mathbf{x}}_{k|k}^{(i)}$ represent. As indicated in subsection 2.1, the temperature $T(\rho)$ is a smooth function of a flux coordinate ρ and can be described using a set of basis functions $\Lambda_i(\rho)$. For the detection of sawtooth crashes, an equidistant grid on ρ is chosen. The values of ρ are determined on 10 locations ranging from $\rho = 0$ to $\rho = 0.6$, and therefore the locations

$$\rho_l = \frac{0.6}{10 - 1}(l - 1) \quad \text{with } l = 1 \cdots 10 \quad (10)$$

are used. These locations are shown in figure 1. The number of spline locations should be chosen such that typical temperature profiles before and after the crash can be accurately represented. A Neumann boundary condition, $\left. \frac{dT}{d\rho} \right|_{\rho=0} = 0$, is used for the core, which results in 11 cubic B-spline basis functions that represent the temperature profile [32]. The basis functions both lead to a reduction of elements of the matrices that need to be evaluated and enforce that the temperature profile is a smooth function of ρ . The estimates $\hat{\mathbf{x}}_{k|k}^{(i)}$ therefore consists of $N_x = 11$ spline coefficients. The matrix \mathbf{C}_k depends on the locations of the measurements and the choice of this matrix is therefore made in section 3.2.

Six hypotheses are considered ($N_h = 6$), of which the first hypothesis considers the situation that no crash has occurred and the other five hypotheses all consider a crash happening at different ρ_{mix} . For hypothesis $i = 1$, the model $\mathbf{A}^{(1)} = \mathbf{A}^{(\text{NoCrash})}$, given by (5), is used. For hypotheses 2 to 6, a crash is expected to occur. In subsection 2.2, a model is presented for a sawtooth crash given a temperature profile in terms of basis functions. A key parameter of this model, given in (A.1), is the mixing radius ρ_{mix} . Variation in ρ_{mix} is accounted for by selecting

$$\begin{bmatrix} \rho_{mix}^{(2)} \\ \rho_{mix}^{(3)} \\ \rho_{mix}^{(4)} \\ \rho_{mix}^{(5)} \\ \rho_{mix}^{(6)} \end{bmatrix} = \begin{bmatrix} 0.30 \\ 0.35 \\ 0.40 \\ 0.45 \\ 0.50 \end{bmatrix}. \quad (11)$$

For this choice of mixing radii, the distance between spline knots should also be on the order of 0.05, which is the case with $N_x = 11$ spline basis functions.

At every time instant k , where the probability of the first hypothesis is large (no crash) for the previous time step $k - 1$, one would expect that a crash could occur. A crash chance of $p_c = 0.1$ is used and all mixing radii are assumed equally likely. Therefore, given that hypothesis one was true in the previous time step, hypotheses 2 to 5 all would have an *a priori* probability of $\frac{p_c}{N_h - 1} = 0.02$. The *a priori* probability that a sawtooth crash will not occur in the current time step would be $1 - p_c = 0.9$. After each sawtooth crash, one would expect another sawtooth crash not to follow immediately. This results in the following *a priori* transition matrix

$$\Pi = \begin{bmatrix} 0.9 & 1 & 1 & 1 & 1 & 1 \\ 0.02 & 0 & 0 & 0 & 0 & 0 \\ 0.02 & 0 & 0 & 0 & 0 & 0 \\ 0.02 & 0 & 0 & 0 & 0 & 0 \\ 0.02 & 0 & 0 & 0 & 0 & 0 \\ 0.02 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (12)$$

in which an element π_{ij} (from row i and column j) contains the *a priori* probability that hypothesis i is true given that hypothesis j was true in the previous step. A sawtooth crash is assumed to have occurred when the probability of a crash at an arbitrary mixing radius is higher than 50 %. Therefore, a sawtooth crash is detected at time k if the condition

$$\sum_{i=2}^{N_h} \mu_k^{(i)} > 0.5 \quad (13)$$

is satisfied. The detection criterium takes into account how well the measurements are related to the crash models. This differs from other multichannel algorithm which either put a detection threshold on every single channel and combine data afterwards (e.g. [25]) or use an averaged measurement channel in which the relation with the spatial distribution is not accounted for [28].

Finally, the initial values for the IMM estimator are determined. At the start of the IMM estimator, the assumption is made that a crash has not occurred. As a result the probability $\mu_1^{(1)} = 1$ for the no-crash hypothesis and for the other hypotheses $i = 2, \dots, N_h$ it follows that $\mu_1^{(i)} = 0$. From $\mu_1^{(1)} = 1$, it follows that only the initial condition for hypothesis one is relevant. The initial estimate $\hat{\mathbf{x}}_{1|1}^{(1)}$ is determined using a least squares fit of the measurements \mathbf{y}_1 at the first time step, such that

$$\hat{\mathbf{x}}_{1|1}^{(1)} = (\mathbf{C}_k^T \mathbf{C}_k)^{-1} \mathbf{C}_k^T \mathbf{y}_1. \quad (14)$$

The initial process covariance $\mathbf{P}_{1|1}^{(1)}$ is chosen the same as the process covariance matrix \mathbf{Q}_k , which is introduced in subsection 3.2.

To summarize, the following parameters are used for the IMM estimator:

- (i) Number of hypotheses $N_h = 6$

- (ii) The *a priori* transition matrix Π given by (12)
- (iii) Initial hypothesis probabilities $\mu_1^{(1)} = 1$ and for $i = 2, \dots, 6$ the probabilities are $\mu_1^{(i)} = 0$
- (iv) Initial estimate $\hat{\mathbf{x}}_{1|1}^{(1)}$ following (14)
- (v) Initial process covariance $P_{1|1}^{(1)} = Q_k$, with Q_k introduced in subsection 3.2
- (vi) Prediction matrices are $A^{(1)} = A^{(\text{NoCrash})}$, given by (5), and $A^{(i)} = A^{(\text{Crash})}(\rho_{mix}^{(i)})$ using (A.1) and (11) for hypotheses $i = 2, \dots, 6$.
- (vii) A sawtooth crash is detected at time instant k if condition (13) is satisfied.

3. Application to measurement data

In this section, the algorithm presented in section 2 is applied to ECE data from the ASDEX Upgrade tokamak. In subsection 3.1, the datasets that will be used are introduced briefly along with the type of crashes they exhibit. The IMM estimator parameters are chosen in subsection 3.2. The IMM estimator is applied and the detected sawtooth crash times are presented in subsection 3.3. The sawtooth crash models are compared with the measurement data in subsection 3.4. Details about the real-time capabilities are given in subsection 3.5.

3.1. ECE datasets

At ASDEX Upgrade, a 60-channel radiometer is used for measuring ECE [38]. The 60-channel heterodyne radiometer collects only ECE in second harmonic X-mode. The ECE measurements present a spatially localized electron temperature, which in the remainder of this paper is referred to as the temperature T . In this paper, ECE measurements at a sample rate of 1 MHz are used and these measurements are filtered and downsampled. A set of five discharges (30538, 30543, 30550, 30552, and 31105) are chosen which are mostly performed for feedforward sawtooth control [39, 26]. By looking at the ECE channels, time ranges are selected with different sawtooth characteristics. For discharges, 30538 and 30543 two sets of time ranges are selected. The time ranges are selected based solely on the crashes that are observed in the ECE measurements. The datasets are given in table 1, which includes a description of the variation in sawteeth that is visible in the ECE signals.

The time resolution provided by the ECE signal at a sampling rate of 1 MHz is higher than what is required for the detection of sawtooth crashes. The ECE signal is therefore filtered (using a low pass filter) and downsampled, such that the required number of computations for the IMM estimator is reduced. Figure 3 shows the ECE data at 1 MHz and filtered and downsampled versions at 100 kHz, 10 kHz, 1 kHz, and 100 Hz. The 100 Hz signal (white line) shows that this averaged temperature does not capture the typical sawtooth crash and is therefore not suitable. At all other sampling rates the typical crash structure is visible. Therefore, a sampling rate of 1 kHz is chosen to result

Table 1. Datasets that were used for IMM estimator tests.

Dataset	Discharge number	Start time t_0 [s]	End time t_1 [s]	Magnetic field strength B_0 [T]	Type of sawteeth
1	30538	4.0	4.3	2.593	Irregular sawteeth with intermediate temperature drop
2	30538	4.5	4.7	2.593	Irregular sawteeth
3	30543	1.4	1.8	2.609	Regular sawteeth
4	30543	3.3	3.6	2.609	Small sawteeth followed by larger sawteeth
5	30550	4.4	5.1	2.610	Long sawteeth with intermediate temperature drop
6	30552	2.8	3.2	2.611	Irregular sawteeth
7	31105	3.8	4.2	2.505	Short precursor oscillation followed by longer sawteeth

in the smallest amount of data points to which the IMM algorithm needs to be applied. An overview of the ECE data for all datasets at 1 kHz is given in figure 4. By looking at time traces of multiple ECE channels, a *human observer* can easily identify the times at which core temperature drops and temperature further out increases consistently. The algorithm needs to make a similar decision based on only the difference between the current and the next time step. The times at which a sawtooth crash is *observed*, by a human observer, are indicated using black vertical lines. The *observed* crash times are also used to evaluate the detection results in the remainder of this section.

3.2. Dataset dependent IMM parameters

In subsection 2.3, the IMM estimator is introduced and most parameters for the IMM estimator are selected in subsection 2.4. In this subsection, the parameters C_k , R_k , and Q_k are selected.

The matrix C_k determines the mapping of the spline coefficients to estimates of the measurements $\hat{\mathbf{y}}_k$. This mapping could be chosen to be time-dependent and include the spatial dependence of the emission region. In this paper, a fixed matrix $C_k = C$ is assumed for every dataset. The matrix C is constructed by calculating the measurement locations ρ_j , using (3), for every ECE channel with frequency f_j . The values $n_j = 2$, $R_0 = 1.65$ m, $a = 0.5$ m, and B_0 from table 1 are used. Use of (3) implies the following approximations:

- (i) all channels measure an optically thick plasma (well-localized ECE),
- (ii) the ECE resonance is given by the cold plasma approximation,
- (iii) the ECE results from electrons in thermal equilibrium,
- (iv) measurements are made in the midplane,

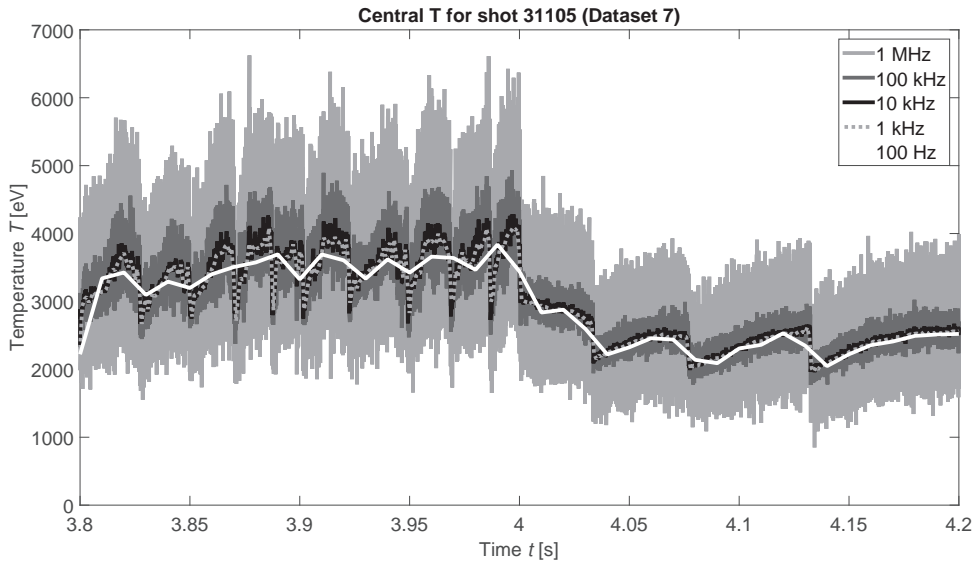


Figure 3. ECE measurements of a core measurement channel and downsampled data for ASDEX Upgrade discharge 31105 (dataset 7). 1 MHz data is shown in light grey, 100 kHz data in dark grey, 10 kHz data in black, 1 kHz data as a dotted grey line, and 100 Hz data as a white line. An 8th order Chebyshev Type-I IIR filter is applied before downsampling.

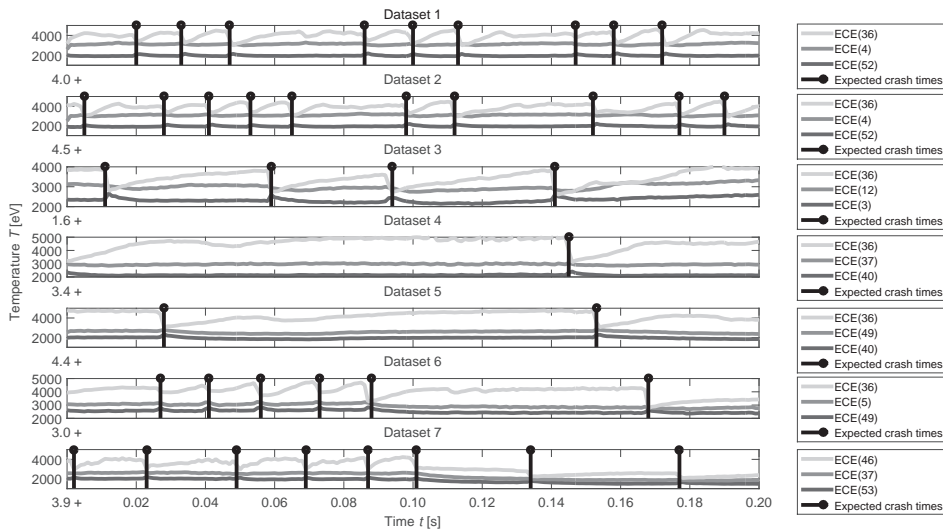


Figure 4. Three ECE measurement channels are shown for the datasets in table 1. A time ranges of 0.2 s is shown with a channel inside the inversion radius (light grey), a channel near the inversion radius (middle grey), and a channel outside the inversion radius and inside the mixing radius (dark grey). Times at which sawtooth crashes are *observed* are shown with black vertical lines.

- (v) the magnetic field strength scales inversely with major radius R , and
- (vi) the magnetic equilibrium is constant with fixed magnetic axis at R_0 .

Matrix C is calculated with (2) and the locations ρ_j . Only channels with $\rho_j \leq 0.6$ are taken into account to reduce the number of data points, while retaining coverage of $\rho \leq \rho_{mix}$. All approximations could be removed by using a proper time-varying matrix C_k . For example, an ECE measurement channel collects radiation in a frequency range, which can be represented by a weighted integral over part of the temperature profile. This could be represented exactly in the matrix C . A drawback is the additional computational time that is required for the construction of the matrix C_k at every time instant k . The computational effort increases when the approximations are replaced by more complex models, e.g. [38].

The noise covariance matrix R_k is determined based on the measurements. The mean-square difference between the measurements and a ten-point moving average is determined. This showed that the noise depends on the average temperature. Therefore, the mean-square differences are divided by the square for the ten-point moving average. This resulted in an average normalized mean-square difference for every channel of every dataset. An average is first calculated for all channels of a specific dataset, followed by an average over all datasets. The average of all channels for all datasets of the normalized mean-square difference is 0.001. In the IMM estimator the mean temperature is approximated by the estimated measurements $C\hat{\mathbf{x}}_{k|k-1}^{(i)}$. Therefore, the measurement covariance is given by

$$R_k = 0.001C\hat{\mathbf{x}}_{k|k-1}^{(i)}. \quad (15)$$

It is known that for this choice the resulting estimator is not optimal [40], but this could be compensated using an estimated measurement covariance [41].

The process noise covariance Q_k is used as a tuning parameter, as the relation between R_k and Q_k determines for what part predictions and measurements are followed. A time-independent process noise covariance matrix $Q_k = Q = Q_0\bar{Q}$ is chosen, where Q_0 is a tuning parameter that is chosen such that the average temperature is followed, while measurement noise is suppressed. For temperature fluctuations $\Delta T(\rho_l)$ and $\Delta T(\rho_m)$ one would expect that

$$\mathbb{E}(\Delta T(\rho_l)\Delta T(\rho_m)) = g(|\rho_l - \rho_m|) \quad (16)$$

as heat sources and losses are not assumed to be localized on a single ρ . In this paper, the function g is set to

$$g(x) = \exp\left(-\frac{x^2}{w_\rho^2}\right) \quad (17)$$

with $w_\rho = 0.1$. A process noise covariance matrix \underline{Q} is constructed for the spline knots given by (10) and the location $\rho_{N_x} = \frac{\rho_{N_x-1} + \rho_{N_x-2}}{2}$ using (17). As a result \underline{Q} is a symmetric matrix with ones on the diagonal and values below on on off-diagonal terms. Using a

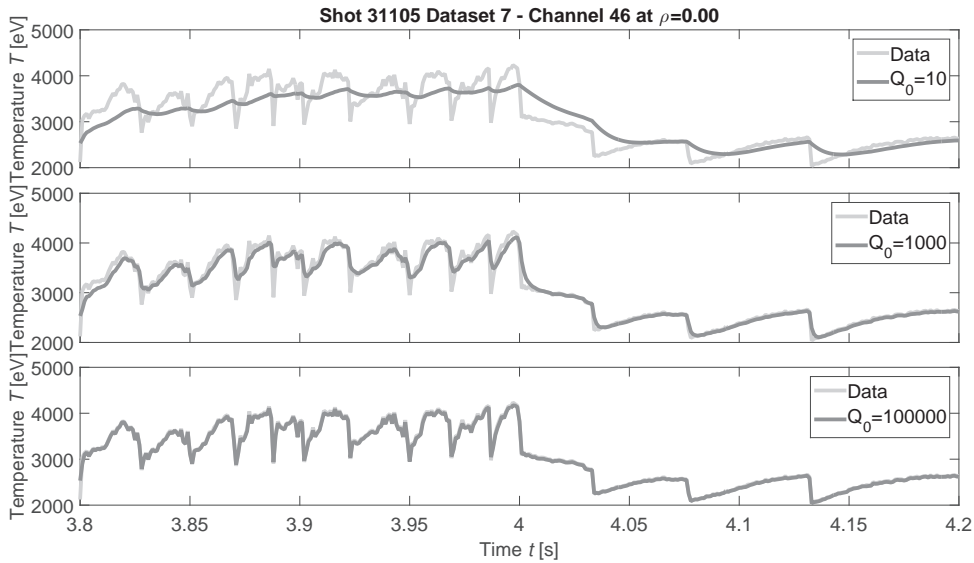


Figure 5. Central channel of ASDEX Upgrade discharge 31105 (Dataset 7) with estimated temperature. The estimated temperatures are obtained for the no crash hypothesis only, for which $A = I_{N_x}$. R_k is defined as (15) and the process noise covariance as in (18), with $Q_0 = 10$, $Q_0 = 1000$, and $Q_0 = 100000$ in the top, middle, and bottom plot, respectively.

matrix \underline{C} , based on (2), for mapping between spline knots and spline coefficients, the process noise in terms of the spline coefficients is given by

$$Q_k = Q = Q_0 \underline{C}^{-1} Q (\underline{C}^T)^{-1}. \quad (18)$$

Temperature estimates for three different values of Q_0 are shown in figure 5. The bottom plot shows that the estimated temperature follows the core temperature closely. For this value $Q_0 = 100000$, the no-crash hypothesis would converge instantaneously to the measurement and therefore the no-crash hypothesis would be just as valid as any of the crash hypotheses. The convergence of the temperature would need to be slower such that applying a sawtooth crash model results in an estimate that better describes the measurements. The top plot shows that the estimated temperature converges slowly to the measurements. This results in a larger deviation between the *a priori* estimates of the temperature and the measurements and would make both the no-crash and sawtooth crash models unlikely. A value of $Q_0 = 1000$ is used for the IMM estimator. This value shows the slow evolution of temperature is tracked, while at the same time not tracking the sawtooth crash. As a result, the sawtooth crash model is *observed* to be most likely at times of the sawtooth crashes.

To summarize, the following dataset-based parameters are used for the IMM estimator:

- (i) The output matrix $C_k = C$, determined by (2) for ρ_j given by (3)
- (ii) The measurement noise covariance R_k , given by (15)
- (iii) The process noise covariance $Q_k = Q$ given by (18) with $Q_0 = 1000$

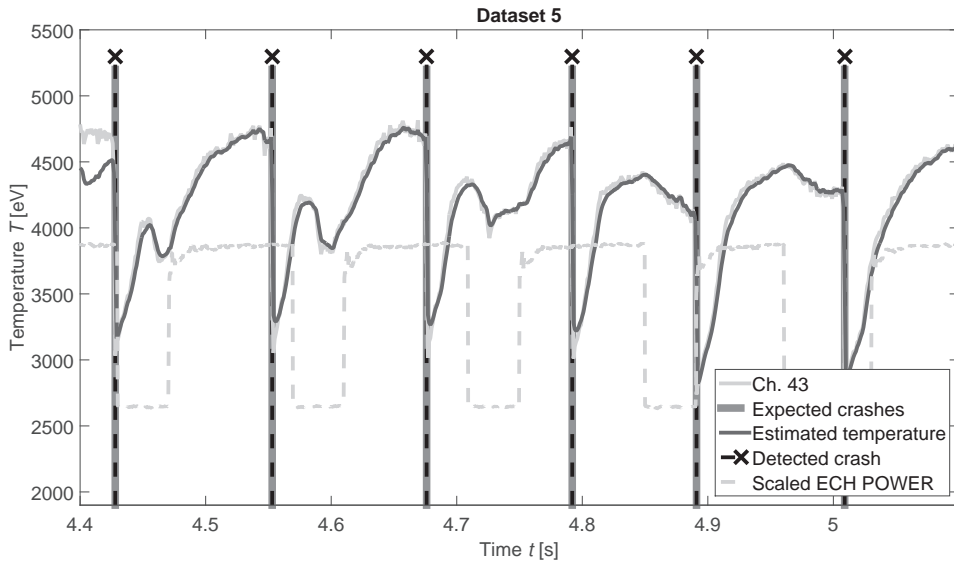


Figure 6. Detected sawtooth crashes for dataset 5. The core temperature channel is shown in light grey and the estimated temperature, based on the no-crash hypothesis, is shown in dark grey. Expected times of sawtooth crashes are indicated by dark grey vertical bars. The detected sawtooth crashes are shown by black dashed lines topped with crosses. The total ECH power in watt (divided by a factor 1000 and with an offset of 2000) is shown with a dashed light grey line.

3.3. Detection of sawtooth crash times

The IMM estimator is applied to the seven datasets listed in table 1. In applying the IMM estimator, false detections due to the initialization are discarded and detections within 5 ms of a previous detection are also discarded. Checks for these conditions could easily be incorporated in the algorithm, but are omitted to make the most use of the data generated by the IMM estimator.

The result of the IMM estimator for dataset 5 is shown in figure 6. In this figure, the temperature of a measurement channel (light grey line) is shown together with the estimated temperature for the no-crash hypothesis ($i = 1$) in dark grey. Based on the time evolution of all measurement channels, the sawtooth crash detections are *observed* at several times, which are shown with thick dark grey vertical bars in figure 6. The sawtooth crashes that are detected by the IMM estimator are shown with black dashed vertical lines topped by a cross. Every *observed* crash is found by the IMM estimator. Note that although only one ECE measurement channel is shown, all measurement channels are used in the IMM estimator. The detection is not hampered by the ECRH which varied during this time window. The total ECH power is shown with a dashed light grey line.

Figure 7 shows the result of the IMM estimator applied to dataset 1. Core temperature and simulated temperature using the no-crash hypothesis are shown in light and dark grey solid lines, respectively. Expected times of sawtooth crashes are

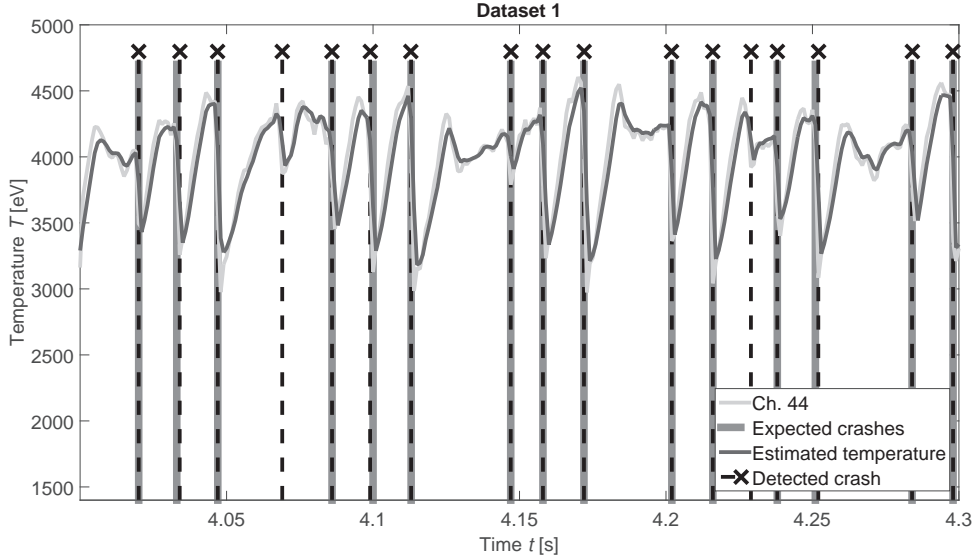


Figure 7. Detected sawtooth crashes for dataset 1. The core temperature channel is shown in light grey and the estimated temperature, based on the no-crash hypothesis, is shown in dark grey. Expected times of sawtooth crashes are indicated by dark grey vertical bars. The detected sawtooth crashes are shown by black dashed lines topped with crosses.

shown with dark grey vertical bars and detected sawtooth crashes are shown with dashed vertical black lines topped by a cross. The ECH power is not shown as it is constant during the considered time interval. Figure 7 shows that the IMM estimator detects all the *observed* sawtooth crashes (introduced in subsection 3.1). On top of that, sawtooth crashes are also detected at 4.069 s and 4.229 s. In the following subsection the detections that do not correspond with *observed* sawtooth crashes are considered.

Datasets 2, 3, 4, 6, and 7 all showed detection of all *observed* sawtooth crashes without showing any additional detections, not corresponding with *observed* crashes.

3.4. Comparison of the crash model with measurements

In this subsection, the crash model is compared with the measurements. Dataset 1 is used as an example. For this dataset, the sawtooth crash at 4.02 s result in a temperature redistribution that is comparable to the model, as is depicted in figure 8. The figure shows measurements just before and just after the sawtooth crash with circles and crosses, respectively. The *a priori* temperature profiles, according to the no-crash model and the crash model, are shown in light and dark grey, respectively. The post-crash temperature measurements show a strong correspondence with the *a priori* crash profile and the estimated mixing radius is close to the radial location where the measurements show mixing ends. Nevertheless, it should be noted that the step in temperature at ρ_{mix} , based on which the crash model is developed, is not found in the measurements.

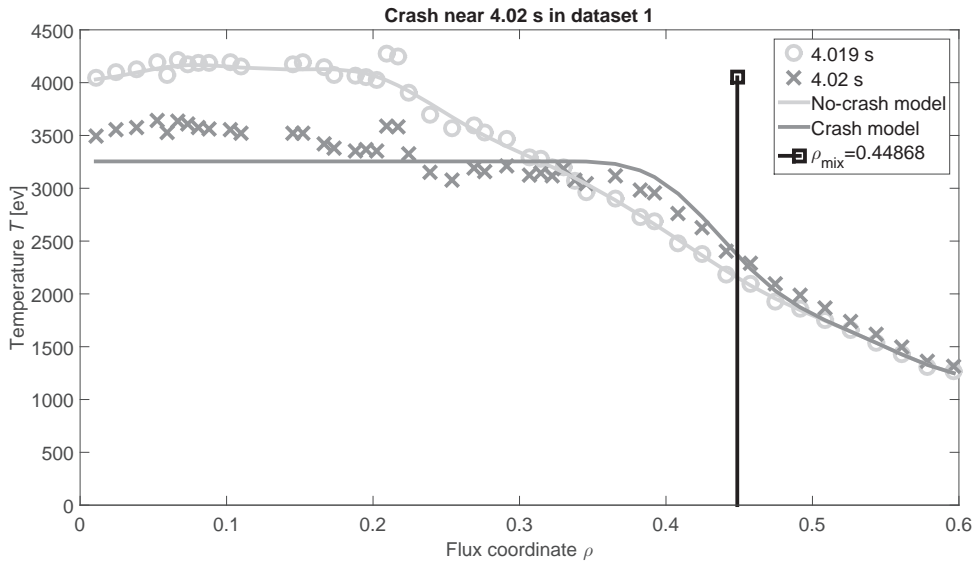


Figure 8. Radial plot of the temperature for the detected sawtooth crash around 4.02 s for dataset 1. Measurements before and after the crash are shown with circles and crosses, respectively. The result of the no-crash model and a crash model with $\rho_{mix} = 0.45$ are shown with light and dark grey lines, respectively. The mixing radius is indicated by a vertical black line.

Figure 9 shows the radial temperature profile for another *observed* sawtooth crash. In this case, there is much more difference between the crash model and the temperature measurement at 4.047 s, just after the sawtooth crash. This difference is caused, because the measurement at 4.047 s is made during the crash, while full flattening of the core temperature is only reached at 4.048 s. This problem arises, because a single time point is used to determine if a crash occurred. If multiple time steps are considered, the time step with the largest temperature difference could be selected. However, such procedures will inherently increase the computational time.

Figure 10 shows one of the two detections not corresponding to an *observed* crash. In this case, the temperature measurements show a decrease of core temperature. However, the flat temperature only extends between $\rho = 0.24$ and $\rho = 0.33$. Temperatures at $\rho < 0.24$ are also decreased, but do not show the same flattening. This indicates that a partial reconnection has occurred. An incomplete temperature flattening is also visible for the detected crash at 4.147 s. In both cases, soft X-ray measurements shows post-cursors which indicates the presence after the sawtooth crash.

The difference between the crash model and the measurements is assessed for all detected sawtooth crashes. About 60 sawtooth crashes are detected in the 7 datasets. For 42 % of the detected sawtooth crashes, the match between the measurements and the crash model is similar to what is visible in figure 8. For 16 % of the detected sawtooth crashes, the measurements for $0 < \rho < \rho_{inv}$ exhibit a difference with the crash model, that is similar to the difference in figure 8, while for $\rho > \rho_{inv}$ the predicted temperature is higher than the measured values. For the remaining 42 % of the detected sawtooth

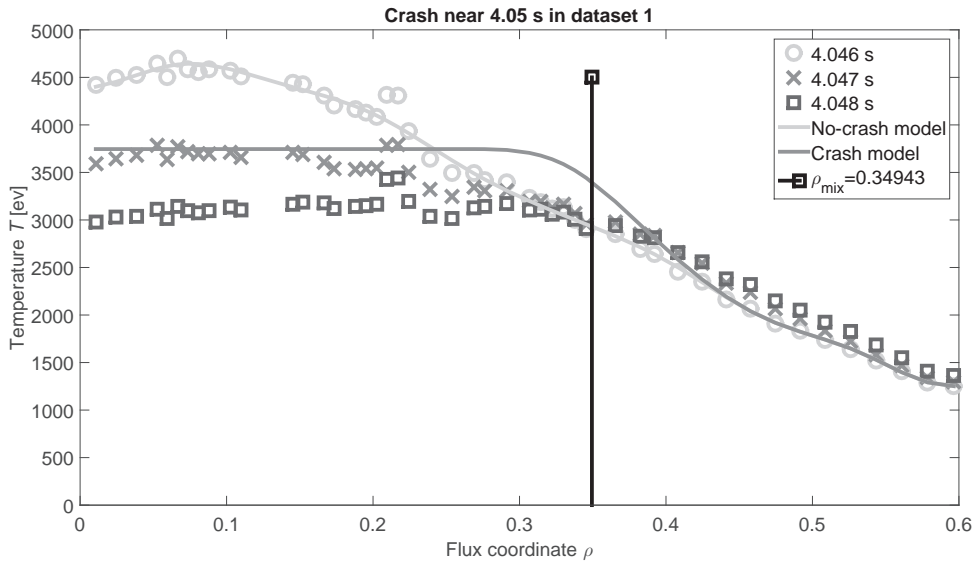


Figure 9. Radial plot of the temperature for the detected sawtooth crash around 4.05 s for dataset 1. Measurements before the crash are shown with circles. Two measurements after the crash are shown with crosses and squares. The result of the no-crash model and a crash model with $\rho_{mix} = 0.35$ are shown with light and dark grey lines, respectively. The mixing radius is indicated by a vertical black line.

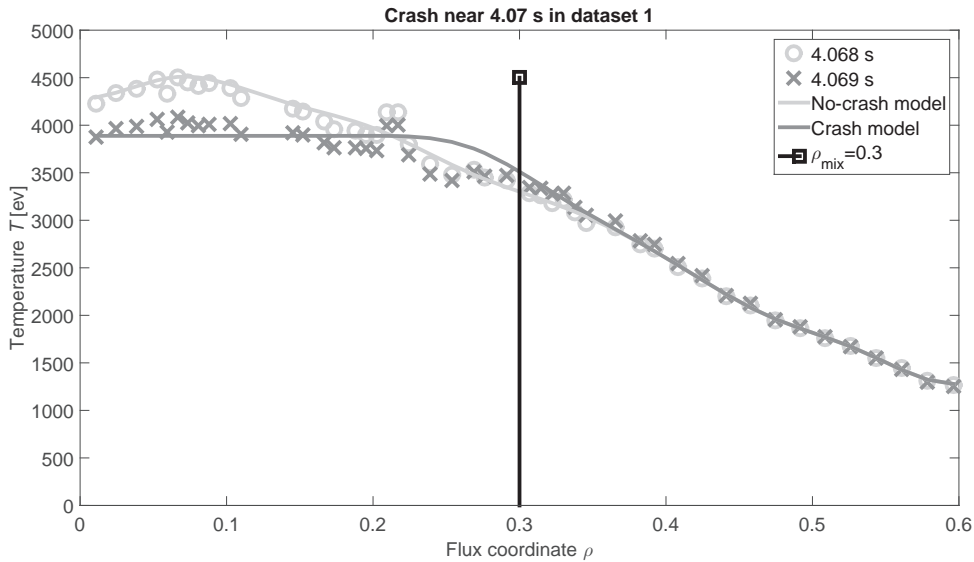


Figure 10. Radial plot of the temperature for the detected sawtooth crash around 4.07 s for dataset 1. Measurements before and after the crash are shown with circles and crosses, respectively. The result of the no-crash model and a crash model with $\rho_{mix} = 0.30$ are shown with light and dark grey lines, respectively. The mixing radius is indicated by a vertical black line.

crashes, there is a significant difference between the crash model and the measurements. For most of these detected crashes, a partial reconnection is observed similar to figure 10. Therefore, although the crash model is suitable to determine if a crash occurred, the implemented crash model is not suitable to accurately determine the mixing radius. In post-shot analysis additional sawtooth crash models could be considered, such that features like the inversion and mixing radius could be determined more accurately. The measurements before and after a detected sawtooth crash can be compared using a more refined spline grid or could be compared against additional models to determine the mixing radius ρ_{mix} in more detail. This process could run in parallel to the detection of sawtooth crash times. Regardless of the mismatch between the sawtooth crash in the model and in the measurements, all *observed* sawtooth crashes are detected and only three detections are made without a corresponding *observed* sawtooth crash. This is achieved with the same IMM estimator parameters for all seven datasets, as given in subsection 2.4 and subsection 3.2.

3.5. Real-time capabilities of the IMM estimator for sawtooth crash detection

The IMM estimator is implemented as a function, which relies on ECE input data, frequencies of the ECE channels, and the IMM parameters. The code ran on a laptop (Intel i7-2630QM, 2 GHz) within Mathworks Matlab 2015a. The run time for processing a dataset using the IMM estimator is determined and divided by the number of time steps at 1 kHz in the data. This results in an average run time of less than 5 ms for an average step of every dataset. This run time is limited due to the recursive nature of the IMM algorithm. With compilation of the Matlab code (mex), the run time could be reduced to below 3 ms for every time step at 1 kHz. However, as a result of compilation, data transfer between Matlab and the compiled code is a significant part of the run time.

The IMM estimator is also run on a dedicated computational node (Intel i7-5930K, 3.5 GHz). On this computational node, the IMM estimator function is compiled together with the script that calls the function for the seven datasets. As a result, there is no need for data transfer between Matlab and the compiled code and it is possible to time only the compiled IMM estimator function. The compiled (mex) code is executed in Matlab and runs within 1 ms for every time step at 1 kHz. A stand-alone executable is also compiled. However, this showed slower run time than the compiled code in Matlab.

Parallelization was not attempted on either the laptop or the computational node. However, it should be realized that algorithm 1 is suitable for this. Each of the six models could potentially perform the prediction step, the update step, and the calculation of the likelihood in a separate thread. On both computers, the IMM estimator is run within a non-real-time operating system. As a result, interrupts are also expected to result in a longer run time than what is minimally achievable. Therefore, it is expected that the IMM estimator with the parameters as presented in this paper could process ECE data at 1 kHz in real-time.

4. Discussion

An IMM estimator is designed and used for sawtooth crash detection. Below the detection model and the use of spline basis functions are discussed. Additionally, the influence of the ECE data on the detection is also addressed. A few considerations for real-time implementation are also given. The complexity of the presented detector, in comparison with real-time detectors, is discussed in the last paragraph.

The IMM estimator uses several models for the crash and a model in case of no crash. In subsection 3.4, the crash model is compared with measurement data. The used crash models result in a good distinction between a crash and no crash. However, the crash model only provided a good approximation of the measurements for 42 % of the detected sawtooth crashes. The reliability of the determined mixing radius could be increased by better crash models. The current models are derived from a physics-based model, but alternatively the models could also be derived using data of sawtooth crashes. The number of models could also be increased, but incorporating additional models would increase the run time of the algorithm. The run time constraint could be relaxed by picking only a subset of all sawtooth crash models that correspond to the type of sawtooth crashes that are expected based on the safety factor profile of the programmed discharge. Another solution would be to rely on a limited set of models to determine the crash time and use a different method to determine the inversion or mixing radius. One could, for instance, use a set of pre-generated sawtooth crashes and use this to train a neural network [42, 43]. The response of the IMM estimator to different models could also be varied by changing the measure for the likelihood of the measurements for hypotheses $i = 1, \dots, N_h$ in algorithm 1. Analysis of the residual distribution could indicate whether a model, different from the multivariate normal distribution, would be more suitable for the likelihood.

In absence of a crash, a constant temperature is assumed. In this case, more extensive models are available. A real-time profile estimator, e.g. RAPTOR, could provide a more accurate *a priori* prediction of the temperature which also takes into account the influence of heating actuators such as ECRH and NBI [44]. In this case the *a priori* estimate and covariance estimates in algorithm 1 are replaced by estimates from a profile estimator. In particular, the RAPTOR code contains the Porcelli model to calculate the temperature profile after a sawtooth crash, which could be used instead of the crash model of the IMM estimator [45]. The RAPTOR code runs within 10 ms at ASDEX Upgrade (a reduction of run time to 5 ms is also feasible). Therefore, for several time steps of the 1 kHz ECE data, the predicted temperature profiles are the same or based on a linear approximation of the same predicted temperature profiles. Due to the energy confinement time of 30 ms, inside the $q = 1$ surface, for a standard H-mode discharge of ASDEX Upgrade, the estimated temperature profiles provide a valid model during the 10 ms (or less) that the models are used.

Spline basis functions in terms of the flux coordinate ρ are used. As a result of the limited number of basis functions the estimated temperature profile is smooth.

As is illustrated in subsection 2.2, this makes it impossible to completely capture the temperature difference between $\rho < \rho_{mix}$ and $\rho > \rho_{mix}$. A possible solution would be inserting additional knots. However, in this case there is a risk of overfitting the temperature measurements. For example, a case with 17 knots showed a smooth temperature profile that oscillated in between the measurement locations, such that the measurements were reproduced exactly, but the estimated temperature profile is non-physical. A way to deal with overfitting could be the inclusion of a monotonicity constraint similar to what is used in integrated data analysis [46].

The results shown in this paper include a limited dataset for the detection of crashes. This dataset could be extended to include all types of discharges in which the IMM estimator should be applied. The results of the IMM estimator can be compared to those of an offline detector, which should provide better detection capabilities as such a detector need not be causal nor fast. The location of the measurement ECE channels should also be considered to best match the region in which the sawtooth crash is expected. In addition, the estimator relies on a redistribution of temperature, which would require that the measured ECE signal values could be interpreted as a reliable temperature measurement. This requires the used ECE frequencies to be optically thick and not in cut-off and necessitates an accurate temperature calibration. Based on density and temperature, optical thickness and lack of cut-off can be verified. The effect of an inaccurate calibration is, however, not straightforward to identify.

The algorithm with five full reconnection hypotheses and one no-crash hypothesis is able to process ECE data at 1 kHz in real-time. Implementation of the algorithm in real-time would require that the same set of parameters is used for all discharges, or that the parameters are changed automatically. In subsection 3.3, all *observed* crashes are detected with a single set of algorithm parameters. To achieve this result, multiple models for the mixing radius ρ_{mix} are used. In a real-time implementation the IMM algorithm should also deal with a sudden loss of one of the radiometer channels. To avoid having to recalculate the output matrix C , the measurement covariance $R_{k,l}$ of a failing channel l could be increased such as to minimize the influence of this channel on the *a posteriori* estimate. In this case care should be taken that the matrix $S_k^{(i)}$ retains a full rank. This implies that a sufficient coverage of the flux coordinate grid at which the estimator models are evaluated should be maintained.

Compared to real-time detectors, such as presented in [17, 27, 28], the presented detector contains more computational steps and requires setting more parameters. Such an increase of complexity should be justified by providing a better solution for the detection of sawtooth crashes. This complexity is introduced primarily to provide a sensible way to incorporate a number of measurement channels in the detection process. In [25], the possibility of detecting a false positive on a single measurement channel is noted as a reason to combine multiple channels. Multiple papers have combined several channels, or the detection results thereof, in an adhoc manner. In contrast, this paper presents a detector which uses the results of all channels and compares temperature profiles based on a model of the expected temperature redistribution. Using multiple

models, a difference could also be made in terms of the mixing radius. For the parameters used in this paper, the detected mixing radius is not accurate. However, by increasing or changing the basis functions and by a different set of detection models, detection capabilities for providing the mixing radius in real-time can be improved, at the expense of a large computational time. If the added complexity of the presented detector is required depends on the quality of the signals available and the requested performance in terms of speed and accuracy of a sawtooth detector.

5. Conclusions

In this paper, an IMM estimator is used for the detection of sawtooth crashes. The estimator evaluates the probability of a crash by comparing measurements with the expected evolution of temperature. The temperature profile, following a sawtooth crash, is expected to have a flat temperature up to the mixing radius ρ_{mix} and no change in the temperature profile is expected in absence of a crash. Multiple mixing radii are considered.

The IMM estimator is applied to 7 datasets based on 5 ASDEX Upgrade discharges. Filtered and downsampled ECE measurements at 1 kHz are used. The same set of design parameters for the IMM estimator is used for all 7 datasets. The resulting times of detected crashes are presented and correspond well with the *observed* crash times. All *observed* sawtooth crashes are detected and additionally a few crash detections are found not corresponding with *observed* crashes.

The temperature profiles, predicted by the crash models, are compared with the post-crash ECE measurements. A valid mixing radius is only found in 42 % of the cases, while in all cases the correct crash times are determined. An example is shown of a case in which a sawtooth is detected based on a measurement during the sawtooth crash, which also resulted in an incorrect determination of the mixing radius. The IMM estimator can be compared to offline sawtooth detection methods to assess if all or most sawtooth crashes are properly detected.

Acknowledgments

This project was carried out with financial support from NWO. The 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.

Appendix A. Crash matrix

Based on the flat temperature $T_{k,flat}$, a post-crash temperature profile is calculated in terms of the coefficients \mathbf{x}_k . The matrix $A^{(Crash)}(\rho_{mix})$ is defined assuming a cubic B-

spline grid with equidistant knots starting at $\rho = 0$ and forcing the derivative at $\rho = 0$ to be zero. A similar matrix $A^{(\text{Crash})}(\rho_{mix})$ could be defined for a different set of basis functions.

For B-splines it holds that $\sum_{i=1}^{N_x} \Lambda_i(\rho) = 1$ for all ρ . As a result of the constant sum, a flat temperature $T_{k,flat}$ is achieved if all coefficients \mathbf{x}_k are equal and set to $T_{k,flat}$. However, in case of a sawtooth crash, the temperature profile only needs to be flat up to $\rho = \rho_{mix}$ and should be unaffected for $\rho > \rho_{mix}$.

For a given mixing radius ρ_{mix} , it is ensured that the highest basis function is set to $T_{k,flat}(\rho_{mix})$ for all knots with $\rho \leq \rho_{mix}$. For an equidistant grid, the number of spline coefficients to set to $T_{k,flat}(\rho_{mix})$ is determined using $n_m = \lfloor \frac{\rho_{mix}}{\Delta\rho} \rfloor + 1$, where $\Delta\rho$ is the distance between the knots and $\lfloor a \rfloor$ is the largest integer smaller than or equal to a . Using the value of n_m , the crash matrix to calculate the post-crash spline coefficients is given by

$$A^{(\text{Crash})}(\rho_{mix}) = \begin{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \mathbf{V}^{(\text{Crash})}(\rho_{mix}) \\ \mathbf{0}_{(N_x - n_m) \times n_m} \quad I_{N_x - n_m} \end{bmatrix} \quad (\text{A.1})$$

in which $\mathbf{0}_{(N_x - n_m) \times n_m}$ is a matrix of zeros with $N_x - n_m$ rows and n_m columns.

References

- [1] S. von Goeler, W. Stodiek, and N. Sauthoff. Studies of Internal Disruptions and $m = 1$ Oscillations in Tokamak Discharges with Soft X-Ray Techniques. *Physical Review Letters*, 33(20):1201–1203, November 1974.
- [2] F. Porcelli, D. Boucher, and M. N. Rosenbluth. Model for the sawtooth period and amplitude. *Plasma Physics and Controlled Fusion*, 38(12):2163, 1996.
- [3] I. T. Chapman. Controlling sawtooth oscillations in tokamak plasmas. *Plasma Physics and Controlled Fusion*, 53(1):013001, 2011.
- [4] J. P. Graves. Influence of Asymmetric Energetic Ion Distributions on Sawtooth Stabilization. *Physical Review Letters*, 92(18):185003, May 2004.
- [5] KSTAR Team, G. S. Yun, H. K. Park, W. Lee, M. J. Choi, G. H. Choe, S. Park, Y. S. Bae, K. D. Lee, S. W. Yoon, Y. M. Jeon, C. W. Domier, N. C. Luhmann, B. Tobias, and A. J. H. Donné. Appearance and Dynamics of Helical Flux Tubes under Electron Cyclotron Resonance Heating in the Core of KSTAR Plasmas. *Physical Review Letters*, 109(14):145003, October 2012.
- [6] M. F. F. Nave, S. Ali-Arshad, B. Alper, B. Balet, H. J. De Blank, D. Borba, C. D. Challis, M. G. Von Hellermann, T. C. Hender, G. T. A. Huysmans, W. Kerner, G. J. Kramer, F. Porcelli, J. O'Rourke, L. Porte, G.J. Sadler, P. Smeulders, A. C. C. Sips, P. M. Stubberfield, D. Stork, R. Reichle, J. A. Wesson, and W. Zwingmann. MHD activity in JET hot ion H mode discharges. *Nuclear Fusion*, 35(4):409, 1995.
- [7] O. Sauter, E. Westerhof, M.-L. Mayoral, B. Alper, P. A. Belo, R. J. Buttery, A. Gondhalekar, T. Hellsten, T. C. Hender, D. F. Howell, T. Johnson, P. Lamalle, M. J. Mantsinen, F. Milani, M. F. F. Nave, F. Nguyen, A. L. Pecquet, S. D. Pinches, S. Podda, and J. Rapp. Control of Neoclassical Tearing Modes by Sawtooth Control. *Physical Review Letters*, 88(10):105001, February 2002.
- [8] G. P. Canal, B. P. Duval, F. Felici, T. P. Goodman, J. P. Graves, A. Pochelon, H. Reimerdes,

- O. Sauter, D. Testa, and the TCV Team. Fast seeding of NTMs by sawtooth crashes in TCV and their preemption using ECRH. *Nuclear Fusion*, 53(11):113026, 2013.
- [9] M. Maraschek. Control of neoclassical tearing modes. *Nuclear Fusion*, 52(7):074007, 2012.
- [10] R. J. La Haye. Neoclassical tearing modes and their control. *Physics of Plasmas (1994-present)*, 13(5):055501, 2006a.
- [11] I. T. Chapman, R. J. Buttery, S. Coda, S. Gerhardt, J. P. Graves, D. F. Howell, A. Isayama, R. J. La Haye, Y. Liu, P. Maget, M. Maraschek, S. Sabbagh, O. Sauter, the ASDEX Upgrade, DIII-D, HL-2A, JT-60U, MAST, NSTX, TCV, Tore Supra Teams, and JET-EFDA Contributors. Empirical scaling of sawtooth period for onset of neoclassical tearing modes. *Nuclear Fusion*, 50(10):102001, 2010.
- [12] D. Kim, T. P. Goodman, and O. Sauter. Real-time sawtooth control and neoclassical tearing mode preemption in ITER. *Physics of Plasmas (1994-present)*, 21(6):061503, June 2014.
- [13] E. Westerhof, O. Sauter, M.-L. Mayoral, D. F. Howell, M. J. Mantsinen, M. F. F. Nave, B. Alper, C. Angioni, P. Belo, R. J. Buttery, A. Gondhalekar, T. Hellsten, T. C. Hender, T. Johnson, P. Lamalle, M. E. Maraschek, K. G. McClements, F. Nguyen, A. L. Pécquet, S. Podda, J. Rapp, S. E. Sharapov, M. Zabiego, and contributors to the EFDA JET Workprogramme. Control of sawteeth and triggering of NTMs with ion cyclotron resonance frequency waves in JET. *Nuclear Fusion*, 42(11):1324, 2002.
- [14] F. Felici, T. P. Goodman, O. Sauter, G. Canal, S. Coda, B. P. Duval, J. X. Rossel, and the TCV Team. Integrated real-time control of MHD instabilities using multi-beam ECRH/ECCD systems on TCV. *Nuclear Fusion*, 52(7):074001, 2012.
- [15] M. F. F. Nave, J. Rapp, T. Bolzonella, R. Dux, M. J. Mantsinen, R. Budny, P. Dumortier, M. von Hellermann, S. Jachmich, H. R. Koslowski, G. Maddison, A. Messiaen, P. Monier-Garbet, J. Ongena, M. E. Puiatti, J. Strachan, G. Telesca, B. Unterberg, M. Valisa, P. de Vries, and contributors to the JET-EFDA Workprogramme. Role of sawtooth in avoiding impurity accumulation and maintaining good confinement in JET radiative mantle discharges. *Nuclear Fusion*, 43(10):1204, 2003.
- [16] V. Igochine, I. T. Chapman, V. Bobkov, S. Günter, M. Maraschek, D. Moseev, G. Pereversev, M. Reich, J. Stober, and the ASDEX Upgrade Team. Destabilization of fast particle stabilized sawteeth in ASDEX Upgrade with electron cyclotron current drive. *Plasma Physics and Controlled Fusion*, 53(2):022002, 2011.
- [17] M. Lennholm, L.-G. Eriksson, F. Turco, F. Bouquey, C. Darbos, R. Dumont, G. Giruzzi, M. Jung, R. Lambert, R. Magne, D. Molina, P. Moreau, F. Rimini, J.-L. Segui, S. Song, and E. Trainsel. Demonstration of Effective Control of Fast-Ion-Stabilized Sawteeth by Electron-Cyclotron Current Drive. *Physical Review Letters*, 102(11):115004, March 2009.
- [18] D. J. Campbell, D. F. H. Start, J. A. Wesson, D. V. Bartlett, V. P. Bhatnagar, M. Bures, J. G. Cordey, G. A. Cottrell, P. A. Dupperex, A. W. Edwards, C. D. Challis, C. Gormezano, C. W. Gowers, R. S. Granetz, J. H. Hammen, T. Hellsten, J. Jacquinet, E. Lazzaro, P. J. Lomas, N. Lopes Cardozo, P. Mantica, J. A. Snipes, D. Stork, P. E. Stott, P. R. Thomas, E. Thompson, K. Thomsen, and G. Tonetti. Stabilization of Sawteeth with Additional Heating in the JET Tokamak. *Physical Review Letters*, 60(21):2148–2151, May 1988.
- [19] M. Henderson, G. Saibene, C. Darbos, D. Farina, L. Figini, M. Gagliardi, F. Gandini, T. Gassmann, G. Hanson, A. Loarte, T. Omori, E. Poli, D. Purohit, and K. Takahashi. The targeted heating and current drive applications for the ITER electron cyclotron system. *Physics of Plasmas (1994-present)*, 22(2):021808, February 2015.
- [20] I. T. Chapman, R. J. La Haye, R. J. Buttery, W. W. Heidbrink, G. L. Jackson, C. M. Muscatello, C. C. Petty, R.I. Pinsker, B. J. Tobias, and F. Turco. Sawtooth control using electron cyclotron current drive in ITER demonstration plasmas in DIII-D. *Nuclear Fusion*, 52(6):063006, 2012.
- [21] W. Biel, M. de Baar, A. Dinklage, F. Felici, R. König, H. Meister, W. Treutterer, and R. Wenninger. DEMO diagnostics and burn control. *Fusion Engineering and Design*, 96–97:8–15, October 2015.
- [22] M. Lauret, F. Felici, G. Witvoet, T. P. Goodman, G. Vandersteen, O. Sauter, M. R. de Baar,

- and the TCV Team. Demonstration of sawtooth period locking with power modulation in TCV plasmas. *Nuclear Fusion*, 52(6):062002, 2012.
- [23] T. P. Goodman, F. Felici, O. Sauter, J. P. Graves, and the TCV Team. Sawtooth Pacing by Real-Time Auxiliary Power Control in a Tokamak Plasma. *Physical Review Letters*, 106(24):245002, June 2011.
- [24] M. Lennholm, T. Blackman, I. T. Chapman, L.-G. Eriksson, J. P. Graves, D. F. Howell, M. de Baar, G. Calabro, R. Dumont, M. Graham, S. Jachmich, M.-L. Mayoral, C. Sozzi, M. Stamp, M. Tsalas, P. de Vries, and JET EFDA Contributors. Feedback control of the sawtooth period through real time control of the ion cyclotron resonance frequency. *Nuclear Fusion*, 51(7):073032, 2011.
- [25] M. van Berkel, G. Witvoet, M. R. de Baar, P. W. J. M. Nuij, H. G. ter Morsche, and M. Steinbuch. Real-time wavelet detection of crashes in limit cycles of non-stationary fusion plasmas. *Fusion Engineering and Design*, 86(12):2908–2919, December 2011.
- [26] A. Gude, M. Maraschek, P. Eulenberg, V. Igochine, and ASDEX Upgrade Team. Automated sawtooth detection with multi-signal analysis. In *42nd EPS Conference on Plasma Physics*, page P1.122, 2015.
- [27] J. I. Paley, F. Felici, S. Coda, T. P. Goodman, F. Piras, and the TCV Team. Real time control of the sawtooth period using EC launchers. *Plasma Physics and Controlled Fusion*, 51(5):055010, 2009.
- [28] F. Felici, H. B. Le, J. I. Paley, B. P. Duval, S. Coda, J. M. Moret, A. Bortolon, L. Federspiel, T. P. Goodman, G. Hommen, A. Karpushov, F. Piras, A. Pitzschke, J. Romero, G. Sevillano, O. Sauter, and W. Vijvers. Development of real-time plasma analysis and control algorithms for the TCV tokamak using Simulink. *Fusion Engineering and Design*, 89(3):165–176, March 2014.
- [29] N. Isei, A. Isayama, S. Ishida, M. Sato, T. Oikawa, T. Fukuda, A. Nagashima, and N. Iwama. Electron cyclotron emission measurements in JT-60u. *Fusion Engineering and Design*, 53(1–4):213–220, January 2001.
- [30] H. Weisen, I. Furno, S. Alberti, C. Angioni, K. Appert, R. Behn, P. Blanchard, P. Bosshard, S. Coda, I. Condrea, A. Degeling, B. P. Duval, P. Gomez, T. P. Goodman, M. A. Henderson, F. Hofmann, J.-P. Hogge, B. Joye, J. B. Lister, X. Llobet, A. Manini, Y. Martin, A. Martynov, J.-M. Mayor, E. Minardi, J. Mlynar, J.-M. Moret, P. Nikkola, Z. A. Pietrzyk, R. A. Pitts, A. Pochelon, H. Reimerdes, J. H. Rommers, O. Sauter, E. Scavino, G. Tonetti, M. Q. Tran, and A. Zabolotsky. Shape dependence of sawtooth inversion radii and profile peaking factors in TCV L mode plasmas. *Nuclear Fusion*, 42(2):136, 2002.
- [31] X. R. Li and V. P. Jilkov. Survey of maneuvering target tracking. Part V. Multiple-model methods. *IEEE Transactions on Aerospace and Electronic Systems*, 41(4):1255–1321, October 2005.
- [32] M. T. Heath. *Scientific Computing*. McGraw-Hill, New York, second edition, 2005.
- [33] M. Bornatici, R. Cano, O. De Barbieri, and F. Engelmann. Electron cyclotron emission and absorption in fusion plasmas. *Nuclear Fusion*, 23(9):1153, September 1983.
- [34] B. B. Kadomtsev. *Soviet Journal of Plasma Physics*, 1:389, 1975.
- [35] G. Witvoet, M. R. de Baar, E. Westerhof, M. Steinbuch, and N. J. Doelman. Systematic design of a sawtooth period feedback controller using a Kadomtsev–Porcelli sawtooth model. *Nuclear Fusion*, 51(7):073024, July 2011.
- [36] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, March 1960.
- [37] B. D. O. Anderson and J. B. Moore. *Optimal filtering*. Dover publications, New York, second edition, 2005.
- [38] S. K. Rathgeber, L. Barrera, T. Eich, R. Fischer, B. Nold, W. Suttrop, M. Willensdorfer, E. Wolfrum, and the ASDEX Upgrade Team. Estimation of edge electron temperature profiles via forward modelling of the electron cyclotron radiation transport at ASDEX Upgrade. *Plasma Physics and Controlled Fusion*, 55(2):025004, February 2013.

- [39] I. T. Chapman, V. Igochine, M. Maraschek, P.J. McCarthy, G. Tardini, the AUG ECRH Group, and the ASDEX Upgrade Team. Sawtooth control using electron cyclotron current drive in the presence of energetic particles in high performance ASDEX Upgrade plasmas. *Plasma Physics and Controlled Fusion*, 55(6):065009, June 2013.
- [40] B. Zehnwirth. A Generalization of the Kalman Filter for Models With State-Dependent Observation Variance. *Journal of the American Statistical Association*, 83(401):164–167, March 1988.
- [41] R. K. Mehra. On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control*, 15(2):175–184, April 1970.
- [42] J. Svensson, M. von Hellermann, and R. W. T. König. Analysis of JET charge exchange spectra using neural networks. *Plasma Physics and Controlled Fusion*, 41(2):315–338, February 1999.
- [43] S. Haykin. *Neural Networks: a comprehensive foundation*. prentice-Hall International, London, Great-Brittain, second edition, 1999.
- [44] F. Felici, C. J. Rapson, W. Treutterer, L. Giannone, E. Maljaars, H. van den Brand, M. Reich, O. Sauter, A. Teplukhina, D. Kim, P. Piosevan, C. Piron, L. Barrera, M. Willensdorfer, A. Bock, E. Fable, B. Geiger, G. Tardini, the ASDEX Upgrade Team, and the EUROfusion MST1 Team. Real-time plasma profile state reconstruction on ASDEX-Upgrade. In *42nd EPS Conference on Plasma Physics*, page O4.127, 2015.
- [45] C. Piron, F. Felici, D. Kim, C. J. Rapson, M. Reich, O. Sauter, T. Goodman, W. Treutterer, H. van den Brand, I. Chapman, the ASDEX Upgrade Team, P. Piosevan, C. Taliercio, the RFX-mod Team, and the EUROfusion MST1 Team. Real-time simulation of internal profiles in the presence of sawteeth using the RAPTOR code and applications to ASDEX Upgrade and RFX-mod. In *42nd EPS Conference on Plasma Physics*, page P1.145, 2015.
- [46] R. Fischer, C. J. Fuchs, B. Kurzan, W. Suttrop, E. Wolfrum, and the ASDEX Upgrade Team. Integrated Data Analysis of Profile Diagnostics at ASDEX Upgrade. *Fusion Science and Technology*, 58(2):675–684, October 2010.