

PAPER • OPEN ACCESS

Robust sparse linear regression for tokamak plasma boundary estimation using variational Bayes

To cite this article: Vít Škvára *et al* 2018 *J. Phys.: Conf. Ser.* **1047** 012015

View the [article online](#) for updates and enhancements.

Related content

- [Plasma boundary identification in HL-2A by means of the finite current element method](#)
You Tian-Xue, Yuan Bao-Shan, Liu Li *et al.*
- [Spatial hierarchical Bayes estimation of mean years of schooling](#)
Dwi A S Wahyuni, Sutarman Wage and Open Darnius
- [Plasma Density Control in a Tokamak](#)
Atsushi Ogata and Teruhiko Tazima



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

Robust sparse linear regression for tokamak plasma boundary estimation using variational Bayes

Vít Škvára^{1,2,3}, Václav Šmídl^{1,4}, Jakub Urban²

¹Institute of Information Theory and Automation, The Czech Academy of Sciences, Prague, Czech Republic

²Institute of Plasma Physics, The Czech Academy of Sciences, Prague, Czech Republic

³Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

⁴Regional Innovation Center for Electrical Engineering, Pilsen, Czech Republic

Abstract. Precise control of the shape of plasma in a tokamak requires reliable reconstruction of the plasma boundary. The problem of boundary estimation can be reduced to a simple linear regression with a potentially infinite amount of regressors. This regression problem poses some difficulties for classical methods. The selection of regressors significantly influences the reconstructed boundary. Also, the underlying model may not be valid during certain phases of the plasma discharge. Formal model structure estimation technique based on the automatic relevance principle yields a version of sparse least squares estimator. In this contribution, we extend the previous method by relaxing the assumption of Gaussian noise and using Student's *t*-distribution instead. Such a model is less sensitive to potential outliers in the measurement. We show on simulations and real data that the proposed modification improves estimation of the plasma boundary in some stages of a plasma discharge. Performance of the resulting algorithm is evaluated with respect to a more detailed and computationally costly model which is considered to be the "ground truth". The results are also compared to those of Lasso and Tikhonov regularization techniques.

1. Introduction

The tokamak is one of the most promising concepts for future thermonuclear fusion reactors. In a tokamak, hot plasma is confined by a combination of strong toroidal and poloidal magnetic fields. The almost negligible weight of the plasma dictates that magnetohydrodynamic (MHD) equilibrium, i.e. a force-free state, describes its macroscopic evolution. The MHD description also includes the plasma boundary (the so-called last closed flux surface), which must be controlled during a tokamak discharge because of performance and safety reasons. Real-time feedback control is necessary since the plasma behaviour cannot be reliably predicted, particularly because the equilibrium is unstable. Such a feedback control requires the plasma boundary as an input in real-time. See e.g. [1] or [2] for more details.

One of the possibilities for real-time plasma shape reconstruction is offered by magnetic sensors near the plasma boundary. However, full reconstruction of the equilibrium magnetic field (which includes the plasma shape) is a complex task due to non-linear behavior of the plasma. In [3], it was shown that the plasma boundary can be modeled by a linear combination



of a specific kind of basis functions. Since the basis functions are known, the estimation problem can be formulated as a simple linear regression with a potentially infinite number of regressors. However, many of these regressors will be uninformative or extremely sensitive to noise. Thus, our task is the selection of a relevant subset of the basis functions of the model.

This task is a well studied problem in statistics and optimization. Examples of classical methods are the Tikhonov [4] or the Lasso [5] regularizations. Crucial problem of these methods is selection of the tuning parameter. This problem is elegantly solved by the Bayesian approach which treats all tuning parameters as unknowns and estimates them from the data. Therefore, we focus on the Bayesian model known as automatic relevance determination (ARD) [6]. It is closely related to scaled l_1 optimization [7]. In effect, the penalization forces weights of spurious regressors to zero, effectively removing them from the regression. This approach was shown to be a suitable model for estimation of the plasma boundary in [8].

In this paper, we extend the previous approach to recursive evaluation of the model in time during the full duration of a discharge. Moreover, we will assume that the model noise (or model residue) is not Gaussian distributed but that it may contain outliers or systematic errors. We will use the model of [9] to estimate the parameters. The resulting algorithm is essentially an iterated weighted least squares algorithm with adaptive estimation of the weight and regularization matrices.

The performance of the presented algorithm will be compared to traditional approaches using simulated and real measurements from a tokamak. We will attempt to provide physical interpretation of the introduced latent variables.

2. Plasma boundary model

Feedback control of plasma in a tokamak requires on-line reconstruction of the plasma boundary. The quality of the reconstruction is highly significant for tokamak operation safety and performance. The reconstruction of the boundary is done in two steps: i) the spatial distribution of scalar magnetic flux $\psi(r, z)$ is determined from the measurements, ii) the boundary is identified as the largest closed flux isoline. In this paper, we will be primarily concerned with the first step, however the quality of the reconstruction will be judged by output of the second step.

The spatial distribution of the magnetic flux $\psi(r, z)$ can be decomposed into an infinite series of basis functions f_i (see [3] for details). Using the assumption of symmetry of the three-dimensional plasma column in the toroidal direction, for each point in the vicinity of the plasma with coordinates (r, z) in the poloidal plane the flux distribution can be decomposed into an infinite series:

$$\psi(r, z) = \sum_{i=0}^{\infty} \theta_i f_i(r, z), \quad (1)$$

where f_i are functions based on Legendre functions with half integer degree [10], closely related to toroidal harmonics, and $\theta_i \in \mathbb{R}$, $i \in \mathbb{N}$ are multipliers associated with each basis function. For increasing i , the basis functions f_i represent higher frequencies of magnetic flux isolines. The expected shape of the boundary is relatively smooth, hence it is expected that only M lower order harmonics will be sufficient for its reconstruction. The magnetic field effects are measured by N magnetic probes around the tokamak chamber. These measurements are transformed by inverse operation into vector $y \in \mathbb{R}^N$ such that

$$y = X\theta + e, \quad (2)$$

where $X \in \mathbb{R}^{N \times M}$ is a design matrix whose rows are composed of evaluations of basis functions $f_i(r_j, z_j)$, $i = 1, \dots, M$, where (r_j, z_j) are coordinates of the j -th probe, $j = 1, \dots, N$, $\theta = (\theta_1, \dots, \theta_M)$ is a vector of unknown coefficients and $e \in \mathbb{R}^N$ is the vector of noise/residue. When the θ vector is identified, we are able to evaluate $\psi(r, z)$ in an arbitrary point, enabling

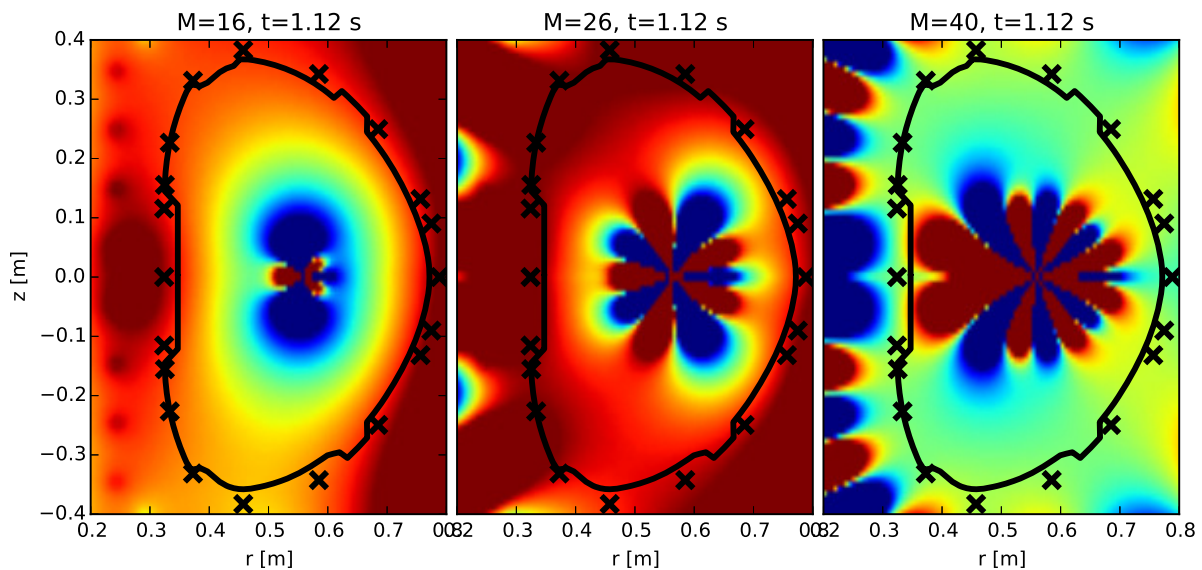


Figure 1. Illustration of the influence of the order M on the spatial distribution of magnetic flux $\psi(r, z)$ during a tokamak discharge when coefficients θ of each plot are obtained using ordinary least squares without regularization. The black line is the tokamak chamber wall, black crosses denote positions of the measurement probes.

us to identify the plasma boundary as the largest closed isoline. Illustration of the spatial distribution of the magnetic flux reconstructed from recorded data using least squares solution of (2) for different values of M are displayed in Fig. 1.

Note that the order of decomposition M significantly influences the spatial distribution of the magnetic flux making the task of plasma boundary identification sensitive to it. While the largest closed isoline in the right plot of Fig. 1 can be recognized around the yellow area inside the tokamak chamber, finding such isolines in the remaining two pictures is impossible. Therefore, we seek an optimal selection of the relevant elements of the model (2). For this purpose, we use a statistical model known as the automatic relevance determination (ARD) that identifies the parameters that are relevant and the remaining are set to zero. Using such an approach, we can then construct an ill-posed problem (2) even with $M > N$ and the algorithm should automatically select the necessary regressors. The selected number M is thus interpreted as an upper bound on the necessary coefficients and when selected sufficiently high, the results should not be sensitive to it.

The proposed algorithm is essentially an iterative solution of the least squares with Tikhonov regularization

$$J(\theta) = \|y - X\theta\|_2^2 + \sum_i \alpha_i \theta_i^2, \quad (3)$$

where the Tikhonov matrix $\text{diag}(\alpha)$ is adaptively updated in each iteration. This is in contrast with the classical Tikhonov regularization where the matrix $\text{diag}(\alpha)$ is assumed to be known. Although the proposed adaptation of α implies higher computation cost than fixed value of α , the increase is still acceptable and can be computed in real time in a closed control loop.

In this contribution, we extend the previous approach by considering unknown precision of the measurements. While the accuracy of the measuring probes is approximately similar, the error of the model may vary across the probes. Specifically, the inverse transformations required to compute vector y from the measurements is nonlinear and the variance of the model residue

e may be time- and space-dependent. Since the exact model of the residue is unknown, we consider the variance of each residue to be unknown and we will estimate it from the data using a variational Bayes approach. We chose a prior model that implies Student's-t distribution of the residue. In effect, the resulting variational Bayes algorithm will solve the weighted least squares with regularization

$$J(\theta) = \|T(y - X\theta)\|_2^2 + \sum_i \alpha_i \theta_i^2, \quad (4)$$

with adaptively estimated weight matrix T in tandem with the Tikhonov matrix $\text{diag}(\alpha)$.

3. Bayesian model selection

Hierarchical priors are a commonly used tool to achieve model structure selection. The conventional least squares formulation of (2) corresponds to maximum likelihood estimation of likelihood function given by multivariate Gaussian probability distribution

$$p(y|X, \theta, \beta) = \mathcal{N}(y|X\theta, \beta^{-1}I). \quad (5)$$

where $\beta \in \mathbb{R}$ is a precision parameter of the model error. The likelihood model alone cannot decide which structure of the model is relevant. This can be achieved by Bayesian treatment with carefully chosen prior distribution on unknown parameters [11]. An extension of the model by the following prior

$$p(\theta|\alpha) = \mathcal{N}(\theta|0, A^{-1}), \quad (6)$$

$$A = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_M),$$

leads to a maximum likelihood solution that is equivalent to Tikhonov regularization (3) where A is the Tikhonov matrix. The choice of diagonal matrix corresponds to the assumption that the elements of θ are a priori independent, each with its own prior precision α_i . In our approach, the precision parameters are treated as unknown and are estimated jointly with the parameters θ . Prior distribution of the precisions α_i and β is chosen in the form of gamma distribution:

$$p(\alpha_i) = \mathcal{G}(\alpha_i|a_0, b_0), \quad (7)$$

$$p(\beta) = \mathcal{G}(\beta|c_0, d_0). \quad (8)$$

Here, values of parameters a_0, b_0, c_0, d_0 are assumed to be known. This leads to a more complicated model but removes the burden of expert selection of the precision parameters.

Now, the task is to obtain the posterior distributions of parameters θ, A, β . This is done using the variational Bayes method. It consists of factorizing the joint posterior and then marginalizing it over individual parameters. For details see [12]. The approximate posterior densities are identified to be of forms conjugate to the priors:

$$\tilde{p}(\alpha_i|y) = \mathcal{G}(\alpha_i|a_{\alpha_i}, b_{\alpha_i}), \quad (9)$$

$$\tilde{p}(\beta|y) = \mathcal{G}(\beta|c, d), \quad (10)$$

$$\tilde{p}(\theta|y) = \mathcal{N}(\theta|\hat{\theta}, S), \quad (11)$$

with equations that govern the shaping parameters given in the appendix. The resulting estimation algorithm is essentially a least squares estimation with ARD adaptation of the regularization (LS-ARD) [8]. The LS-ARD algorithm is guaranteed to converge to a local minimum of the Kullback-Leibler divergence from the estimated to the true posterior. We assume that prior parameters a_0, b_0, c_0 and d_0 are chosen to be 10^{-10} to yield non-informative prior.

4. Robust Least Squares using Student's t noise model

To improve the results of plasma reconstruction, we alter the original algorithm from Section 3 by allowing a richer error distribution which can model outliers and systematic errors, as explained in Section 2. Specifically, we replace the Gaussian noise distribution by a Student's t-distribution with unknown degrees of freedom d . For low values of d , this distribution has heavier tails than the Gaussian. The estimation algorithm can thus identify outliers and provide more robust estimate of the parameter of interest.

While prior distributions for θ and A are the same as in the previous model, i.e. (6), (7), the likelihood model (5) is changed to

$$p(y|X, \theta, \omega, z) = \mathcal{N}(y|X\theta, (\omega \text{diag}(u))^{-1}) \quad (12)$$

$$p(\omega) = \mathcal{G}(\omega|\rho_0, \kappa_0), \quad (13)$$

$$p(u_i|d) = \mathcal{G}(u_i|d/2, d/2), \forall i \quad (14)$$

$$p(d) = \mathcal{G}(d|\phi_0, \psi_0). \quad (15)$$

The values of u_i (more exactly its square root) represent the weights in a diagonal matrix T of the weighted least square problem (4). The Student's t-distribution of the noise arises from the fact that

$$S(e|0, \omega, d) = \int_0^\infty \mathcal{N}(e|0, \omega^{-1}u^{-1}) \mathcal{G}(u|d/2, d/2) du$$

Again, we use the VB framework to obtain equations for shaping parameters of the posteriors. Their forms are

$$\tilde{p}(\theta|y) = \mathcal{N}(y|\hat{\theta}, S), \quad (16)$$

$$\tilde{p}(\alpha_i|y) = \mathcal{G}(\alpha_i|a_{\alpha i}, b_{\alpha i}), \quad (17)$$

$$\tilde{p}(\omega|y) = \mathcal{G}(\omega|\rho, \kappa), \quad (18)$$

$$\tilde{p}(u_i|y) = \mathcal{G}(u_i|a_{u i}, b_{u i}), \quad (19)$$

$$\tilde{p}(d|y) = \mathcal{G}(d|\phi, \psi), \quad (20)$$

with shaping parameters in the appendix. The final estimation algorithm – robust least squares with ARD adaptation (RLS-ARD) algorithm – is given in Algorithm 1. This more elaborate structure of the model may provide us with further information on the physical system we are describing as will be discussed in the next section.

Algorithm 1 RLS-ARD - iterative algorithm for structure selection of robust least squares problem.

Off-line:

- Select prior parameters $a_{0i}, b_{0i}, \rho_0, \kappa_0, \phi_0$ and ψ_0 .

On-line:

- Initialize $\hat{\omega}$ and \hat{u}_i with chosen values.
 - Repeat until convergence:
 - update shaping parameters of θ , using (B.1),
 - update shaping parameters of A , using (B.2),
 - update shaping parameters of ω , using (B.3),
 - update shaping parameters of u , using (B.4),
 - update shaping parameters of d , using (B.5).
 - Report final estimate $\hat{\theta}$.
-

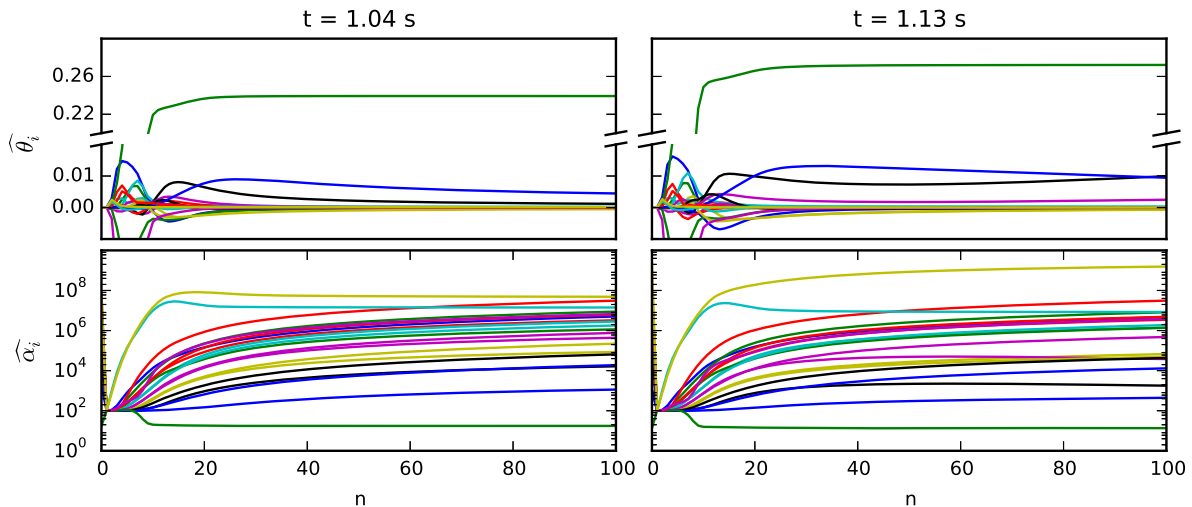


Figure 2. Convergence paths of different elements of the estimated $\hat{\theta}_i$ during a run of the RLS-ARD algorithm for two fixed time samples during a discharge (Upper row), and convergence paths of the corresponding precisions $\hat{\alpha}_i$ (Lower row), n denotes the iteration number.

5. Experimental Validation

5.1. Tokamak COMPASS and data from a validated numerical model

Both simulated data and measurements from operation of the tokamak COMPASS will be used in this section. Simulated data were obtained as a fit to real measurements with the EFIT++ [13] plasma reconstruction code. This code evaluates a complex numerical model of magnetohydrodynamic equilibrium which includes the full toroidal flow and anisotropy changes to the Grad-Shafranov equation. The result of the EFIT++ reconstruction is also the plasma boundary, however obtained at high computational cost. Therefore, when using the simulated measurements, we can use the boundary from the EFIT++ as “ground truth” to compare the performance of the proposed algorithm. Because we are mainly interested in the quality of the plasma boundary reconstruction, we use the Hausdorff distance [14] to compare the estimated 2D shape of the plasma boundary to the ground truth. This metric represents the maximum difference between the two boundaries evaluated on discretized angular grid.

For the purpose of this paper, all simulated data were generated with the configuration which agrees with the real geometry of the tokamak and measuring devices. The EFIT++ reconstructs the plasma field from measurements at standard sampling period 1ms, which is sufficient for real time shape control. Thus the plasma boundary ground truth is available in a time series. The proposed algorithm is designed to reconstruct the plasma boundary at each sampling time independently. Therefore, we will show results of reconstruction at selected time samples. In each time sample, an iterative algorithm (such as Algorithm 4 for RLS-ARD) will be run until convergence. The relatively slow temporal evolution of the shape can be used to warm start the iterative algorithm by a solution obtained in the previous time instant. This will be important to achieve computational cost permitting real time evaluation.

5.2. Model structure selection

First, we demonstrate how the ARD property selects the optimal regularization coefficients for reconstruction of the magnetic flux at a single time instance. We have selected data from two time samples of a discharge and run the RLS-ARD algorithm (Algorithm 4) for $n = 100$

iterations with initialization $\alpha_{0i} = 100$. Convergence of the estimated values of $\hat{\theta}_i$ and their corresponding precision parameters (penalization coefficients) $\hat{\alpha}_i$ are displayed in Fig. 2. With increasing precision $\hat{\alpha}_i$ grows prior probability that $\hat{\theta}_i$ is zero. This has a direct link to Tikhonov regularization as was hinted in section 2 where α_i is the penalization parameter that governs how much the squared value of a regression coefficient is penalized.

For these particular reconstructions, one coefficient out of 20 (i.e. $M = 20$) in the decomposition (2) seems to have a major contribution. Clearly, its precision/penalization converges to a smaller number than the common starting value of 10^2 while the other coefficients have larger precisions and thus are more penalized.

5.3. Reconstruction of a full discharge

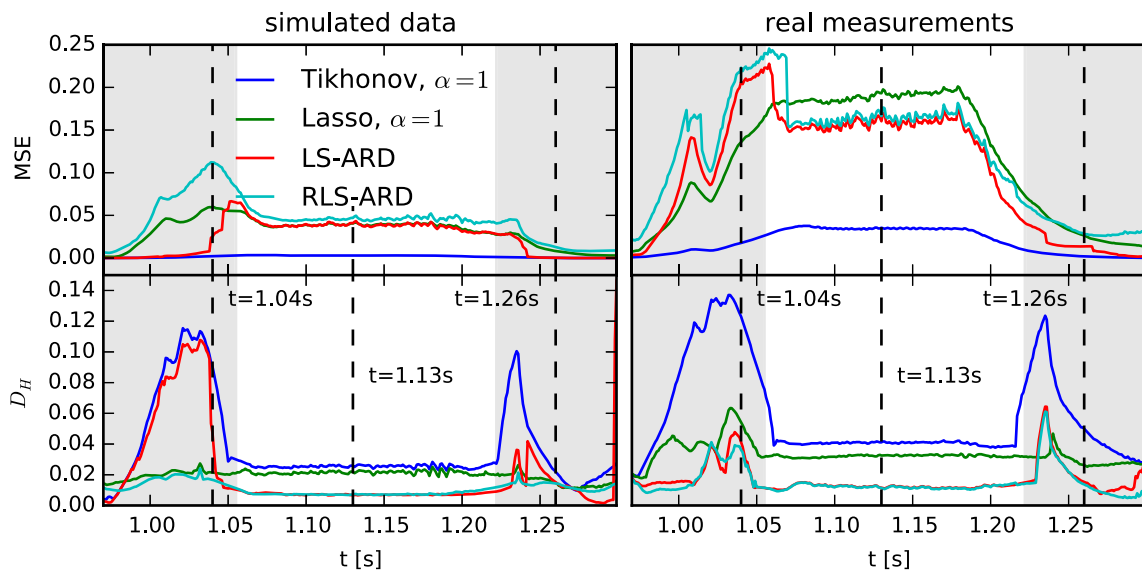


Figure 3. Comparison of model residues and Hausdorff distance for simulated data (left column) and real measurements (right column). **Top row:** mean square error of the model residue of the linear problem (2) during a plasma discharge as fitted by different methods. **Bottom row:** Hausdorff distance D_H between the reconstructed plasma boundary and ground truth (or the output of a validated numerical model in case of real measurements).

The shaded areas on the left and right distinguish the initial and terminal phase of the discharge as identified by presence of x-point computed by EFIT++. Dashed vertical lines denote time samples at which the reconstruction results are presented in Fig. 5.

Next, we compare the performance of tested algorithms on full duration of a selected plasma discharge. To understand the results better, we will describe the individual phases of a tokamak plasma discharge, which are also distinguished by shading in Fig. 3. When the discharge starts, the 2D cross section of the plasma column is elliptic. Afterwards, it undergoes a transition which, in an ideal case, ends with the plasma being elongated with a pointed tip (x-point). This part is marked by gray background color at the left side of the discharge figures. When the x-point is established, the plasma rapidly reaches the desired equilibrium state (white background). After that, the plasma collapses back to an elliptic shape before completely vanishing (gray background color on the right side of each plot).

Since the x-point is a saddle point in the $\psi(r, z)$ contours and it is usually close to the tokamak chamber wall, the identification of the plasma shape is very sensitive to its predicted position.

If a saddle point is predicted to be outside of the chamber, even just a few centimeters off its true position, the reconstructed plasma shape is elliptic and very different from the true shape.

In Fig. 3 the MSE of model residue of the linear equation (2) is plotted for the whole plasma discharge for different approaches to regularization and for both simulated and real measurements. For the Tikhonov regularization, the regularization constant α was set to be equal for all regressors. The MSE of the model residue is accompanied by values of the Hausdorff distance D_H to the ground truth for each algorithm.

Although the fit using Tikhonov has the lowest value of the residue at the measurement points during the whole discharge, the final boundary is not very close to the ground truth. The remaining methods clearly gain advantage by using stricter regularization and thus achieving a better generalization beyond the measurement points, which is crucial for identifying the boundary. Also, the plasma computed with tested methods may be predicted to have an x-point when the ground truth does not have one or vice versa, which leads to high peaks in the D_H metric. Lasso algorithm for l_1 optimization yields results very similar to those of LS-ARD or RLS-ARD.

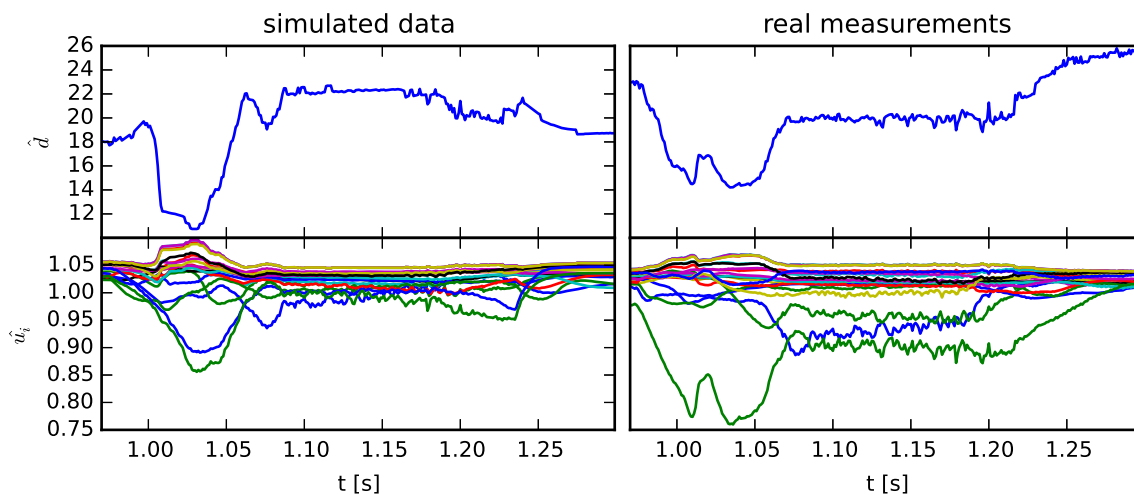


Figure 4. Measures of non-Gaussianity of the robust linear regression. **Top row:** Estimate of the degrees of freedom of the Student's noise parameter d during the discharge. **Bottom row:** estimates of the weights u_i of individual measurements.

Interestingly, the LS-ARD algorithm does not perform well on simulated data in the transition phases. The value of D_H is higher in the transition phases than for the other approaches. Higher deviation from the EFIT boundary in the transition phases means that assumptions for one of the models (EFIT or toroidal harmonics) may be violated. Also, the assumption of non-Gaussian residue of the RLS-ARD model may be more appropriate in the transition phases because of the nonlinear propagation of errors as is demonstrated by the results produced using synthetic data. Apart from the transition phases, the difference of results between LS and RLS-ARD is marginal. This similarity is also apparent from the estimated degrees of freedom of the Student's t -distribution (parameter d) from the RLS-ARD algorithm is displayed in Fig. 4. Note that it decreases during the first transition phases. Smaller values of d can be interpreted as a deviation of the measurement errors from the Gaussian distribution. This behaviour indicates the presence of outliers or a systematic error in the model. However, because we observe the deviation even on simulated data, the most likely explanation is insufficiency in the used model as discussed above. Also, we can track which measurements are suppressed from the fit since

these measurements have the value of u_i smaller than 1.

The effect of the magnetic flux reconstruction on the shape of the reconstructed plasma boundary is displayed in Fig. 5, where the estimated boundary shapes are compared to the ground truth results at different moments of the discharge. These moments were selected from different stages of the discharge, see dashed lines in Fig. 3. Again, we see that better results are obtained in the stationary phase of a discharge. The value of u_i estimates at a given time is illustrated at the corresponding measurement points. Those measurement points with $u_i < 1$ are highlighted with a larger circle. This can provide experiment designers with additional information about the precision of measurements.

While the RLS-ARD algorithm appeared to improve results on the simulation data, its improvement on real data is not so significant. Moreover, due to its lower computational cost, the LS-ARD was selected for implementation. The computational cost of this algorithm can be lowered by initialization of the parameters from those obtained in previous time sample (warm starts). The effect of this strategy is displayed in Fig. 6, where the number of iterations needed for convergence of the LS-ARD algorithm are shown for two variants: offline, where the initial guesses of parameters are the same for each time of the discharge, and online, where the parameters estimated in the previous time are used as initialization values for the algorithm. This results in a significant speedup, where only a few iterations are needed.

5.4. Evaluation of multiple discharges

A quantitative study of quality of reconstruction was tested on a data set of ten plasma discharges. The relative performance of the tested algorithms remains the same for all discharges. At two discharges, the RLS-ARD algorithm was worse than the LS-ARD in the initial phase of the discharge. We conjecture that this is caused by incorrect suppression of the relevant measurements (i.e. incorrectly classifying them as outliers). Statistical evaluation of the distribution of the mean and median of the Hausdorff distance for all times of the ten discharges is displayed in Fig. 7. From these results, the LS-ARD method seems to provide the most consistent and reliable results.

6. Conclusion

In this paper, we have described an inverse problem that arises in the field of tokamak plasma reconstruction and the difficulties it presents for classical methods. We have proposed a Bayesian approach to the problem resulting in an algorithm that selects a sparse and robust solution. The proposed method has been shown to provide consistently better results compared to the conventional approach. It has been shown that a model with fewer automatically selected parameters avoids the overfitting problem and thus achieves better reconstruction of the overall shape of the plasma boundary. This has been achieved on both simulated and measured data. The complexity of the computation is intentionally kept low so that the algorithm can be run in real time on existing hardware. The robust algorithm was able to improve quality of reconstruction only at the non-stationary discharge phases in simulation. However, its improvement on real data was marginal and not very reliable. Since the distribution of the error was estimated to be almost Gaussian in the most important stationary regimes, we selected the sparse LS-ARD algorithm for final implementation.

Acknowledgments

This work was partially funded by the Czech Technical University through the grant SGS15/214/OHK4/3T/14 and by the Ministry of Education, Youth and Sports of the Czech Republic under the RICE – New Technologies and Concepts for Smart Industrial Systems, project No. LO1607. COMPASS tokamak experiments are supported by MEYS projects LM2015045 and 8D15001.

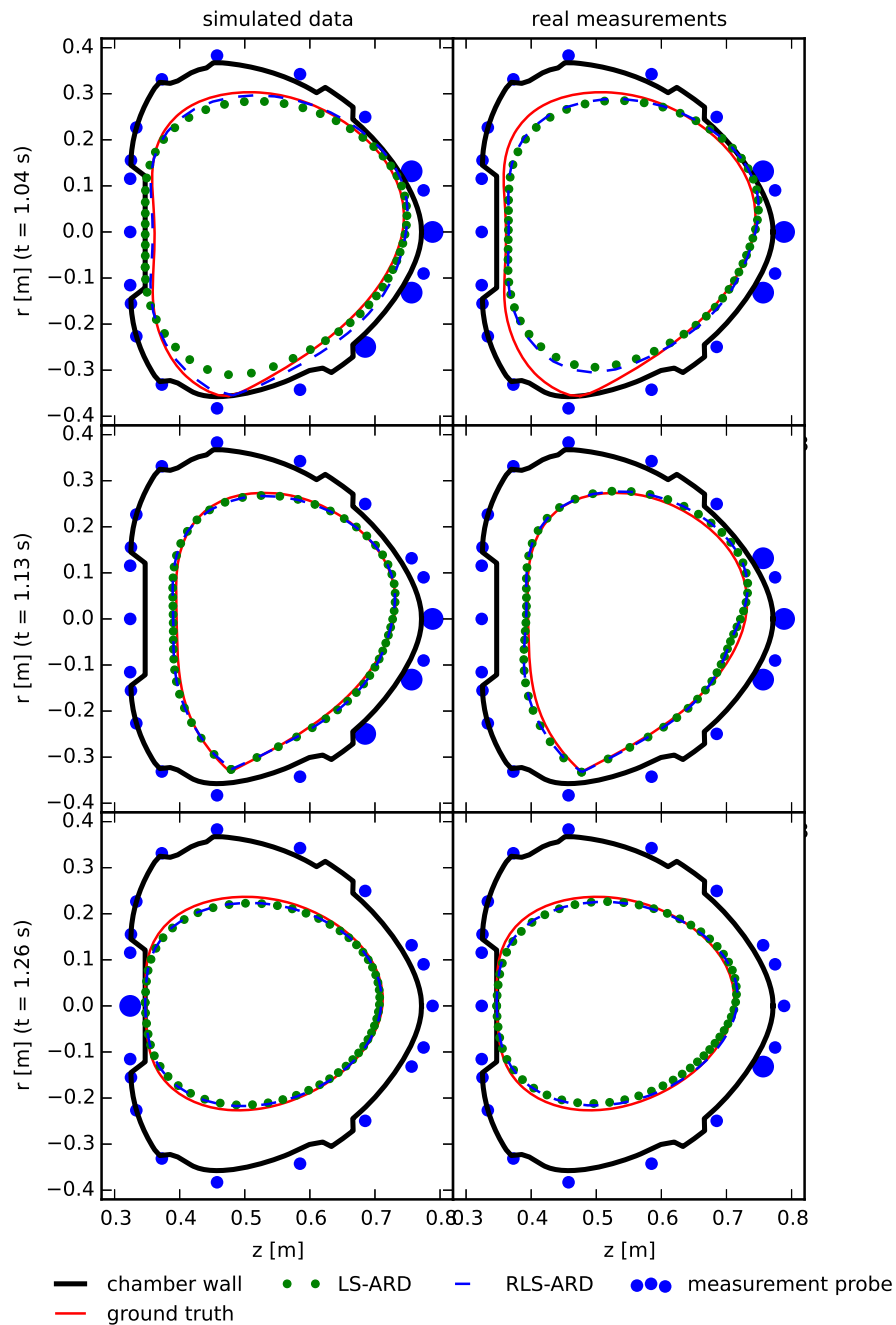


Figure 5. Estimates of the plasma boundary: ground truth from the EFIT model (red solid line), LS-ARD (green dotted line) and RLS-ARD (blue dashed line) algorithms for different data sets and different time samples of the discharge. Inner boundary of the tokamak chamber is in black. Also, the locations of the sensors are displayed (blue dots), those that have the value of estimate of u_i smaller than 1 are enlarged.

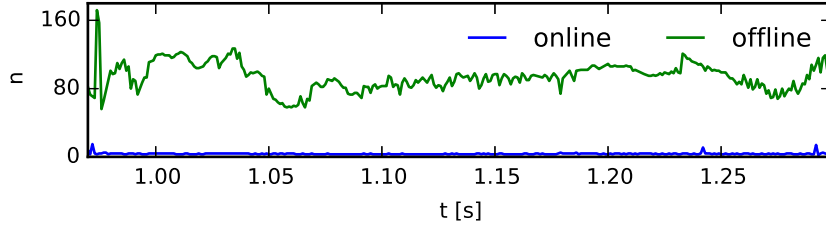


Figure 6. Number of iterations n needed for convergence of the LS-ARD algorithm during a tokamak discharge. In online mode, parameters are initialize with previous results, when in offline mode, the same initial conditions are used at every sampling time.

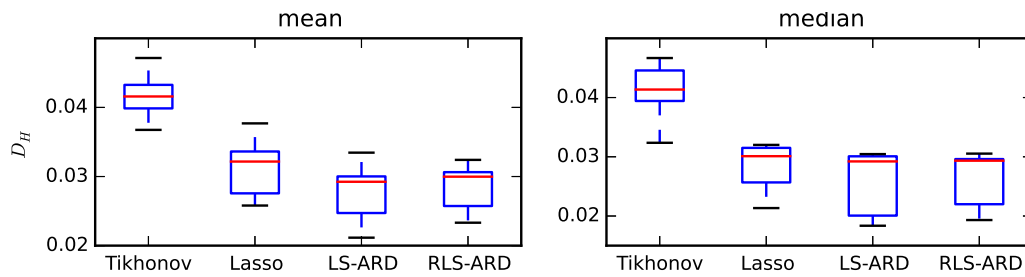


Figure 7. Mean and median Hausdorff distance D_H for 10 plasma discharges and different algorithms.

Appendix A. Parameters of least squares ARD algorithm

Hyperparameters of normal and gamma posteriors (9)-(11) are given by

$$a_{\alpha i} = a_{0i} + \frac{1}{2}, \quad b_{\alpha i} = b_{0i} + \frac{1}{2} \widehat{\theta}_i^2. \quad (\text{A.1})$$

$$c = c_0 + \frac{N}{2}, \quad d = d_0 + \frac{1}{2} y^T y - y^T X \widehat{\theta} + \frac{1}{2} \text{tr} (X^T X \widehat{\theta \theta}^T) \quad (\text{A.2})$$

$$\widehat{\theta} = \widehat{\beta} S X^T y, \quad S = (\widehat{A} + \widehat{\beta} X^T X)^{-1} \quad (\text{A.3})$$

Notation $\widehat{\cdot}$ stands for expected value of the variable. Other required moments are

$$\begin{aligned} \widehat{\alpha}_i &= \frac{a_{\alpha i}}{b_{\alpha i}}, & \widehat{\beta} &= \frac{c}{d}, \\ \widehat{\theta}_i^2 &= S_{ii} + \widehat{\theta}_i^2, & \widehat{\theta \theta}^T &= \widehat{\theta} \widehat{\theta}^T + S. \end{aligned}$$

Appendix B. Parameters of robust least squares ARD algorithm

Hyperparameters of distributions (16)-(20) have the following form

$$\hat{\theta} = \hat{\omega} S X^T \widehat{\text{diag}}(u) y, \quad S = (\hat{A} + \hat{\omega} X^T \widehat{\text{diag}}(u) X)^{-1}, \quad (\text{B.1})$$

$$a_{\alpha i} = a_{0i} + \frac{1}{2}, \quad b_{\alpha i} = b_{0i} + \frac{1}{2} \hat{\theta}_i^2, \quad (\text{B.2})$$

$$\rho = \rho_0 + \frac{N}{2}, \quad \kappa = \kappa_0 + \frac{1}{2} y^T \widehat{\text{diag}}(u) y - y^T \widehat{\text{diag}}(u) X \hat{\theta} + \frac{1}{2} \text{tr}(\hat{\theta} \hat{\theta}^T X^T \widehat{\text{diag}}(u) X), \quad (\text{B.3})$$

$$a_{ui} = \frac{1}{2} (\hat{d} + 1), \quad b_{ui} = \frac{1}{2} \hat{d} + \frac{1}{2} y_i^2 \hat{\omega} - y_i \hat{\omega} (X \hat{\theta})_i + \frac{1}{2} \hat{\omega} (X \hat{\theta} \hat{\theta}^T X^T)_i, \quad (\text{B.4})$$

$$\phi = \phi_0 + \frac{1}{2} N, \quad \psi = \psi_0 - \frac{1}{2} N + \frac{1}{2} \sum_{i=1}^N \hat{u}_i - \frac{1}{2} \sum_{i=1}^N \ln \hat{u}_i. \quad (\text{B.5})$$

Note that for derivation of ψ we used Stirling's formula for the term $\ln(\Gamma(d/2))$ as in [9]. The required moments of these distributions are:

$$\begin{aligned} \hat{\theta}_i^2 &= S_{ii} + \hat{\theta}_i^2, & \hat{\theta} \hat{\theta}^T &= \hat{\theta} \hat{\theta}^T + S, & \hat{u}_i &= \frac{a_{ui}}{b_{ui}}, \\ \hat{\alpha}_i &= \frac{a_{\alpha i}}{b_{\alpha i}}, & \hat{\omega} &= \frac{\rho}{\kappa}, & \hat{d} &= \frac{\phi}{\psi}. \end{aligned}$$

References

- [1] Ambrosino G and Albanese R 2005 *IEEE Control Systems Magazine* **25** 76–92
- [2] Ariola M and Pironti A 2008 *Magnetic control of tokamak plasmas* (Springer Science & Business Media)
- [3] Faugeras B, Blum J, Boulbe C, Moreau P and Nardon E 2014 *Plasma Physics and Controlled Fusion* **56** 114010
- [4] Hoerl A E and Kennard R W 1970 *Technometrics* **12** 55–67
- [5] Tibshirani R 1996 *Journal of the Royal Statistical Society. Series B (Methodological)* 267–288
- [6] Tipping M 2001 *The journal of machine learning research* **1** 211–244
- [7] Wipf D P and Nagarajan S S 2007 A new view of automatic relevance determination *Advances in Neural Information Processing Systems* pp 1625–1632
- [8] Škvára V, Šmídl V and Urban J 2015 *Journal of Physics: Conference Series* **659** 012010
- [9] Christmas J and Everson R 2011 *IEEE Transactions on Signal Processing* **59** 48–57
- [10] Lebedev N N and Silverman R A 1972 *Special functions and their applications* (Courier Corporation)
- [11] Bishop C 2006 *Pattern recognition and machine learning* vol 1 (Springer New York)
- [12] Šmídl V and Quinn A 2005 *The Variational Bayes Method in Signal Processing* (Springer)
- [13] Fitzgerald M, Appel L C and Hole M J 2013 *Nuclear Fusion* **53** 113040
- [14] Huttenlocher D P, Klanderman G, Rucklidge W J *et al.* 1993 *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **15** 850–863