X-RAY TOMOGRAPHY ON TCV

M. Anton, H. Weisen, M.J. Dutch,
F. Bühlmann, R. Chavan, B. Marlétaz,
P. Marmillod, P.J. Paris

X-Ray Tomography on TCV

M. Anton, H. Weisen, M. J. Dutch,

F. Buhlmann, R. Chavan, B. Marletaz, P. Marmillod, P. Paris

Centre de Recherches en Physique des Plasmas

Association EURATOM – Confédération Suisse

Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

short title: X-ray Tomography on TCV

CC: <u>5270</u>, 5225P, 5235

1

Abstract

The TCV Tokamak offers an outstanding variability of the plasma shape. Using X-ray tomography, the shape of the inner flux surfaces of a poloidal cross section of the plasma can be reconstructed, including fast variations due to MHD activity. The hardware as well as the software of the 200 channel system developped for TCV is described.

A new, 'dynamical' calibration is used. The actual plasma temperature and some global profile parameters serve to determine the spectrum-dependent efficiency of the photodiodes. Compared to a 'static' calibration with constant calibration factors, an enhanced quality of the reconstructed images is observed.

The tomographic inversion is performed using a variety of methods such as Maximum Entropy, linear Regularisation and a new method making use of the Fisher information of the emissivity distribution. The merits of the different algorithms which have been implemented as MATLAB functions are compared.

The tomographic inversion results are analysed with the help of the biorthogonal decomposition, allowing *e.g.* identification of MHD modes without using any *a priori* information on the poloidal mode structure.

2

Recent results on the dependence of sawtooth activity on the plasma triangularity are presented to demonstrate the performance of the soft X-ray tomography system.

1 Introduction

The flexibility of the tokamak TCV, the *Tokamak à Configuration Variable* is unique. Design parameters of TCV are a $1MA$ maximum plasma current, a toroidal field of $1.5T$ and a maximum elongation of $\kappa = 3$. The major radius is $R = 0.88m$. Width and height of the vessel are $0.55m$ and $1.55m$, respectively, with an almost rectangular cross section. About 70 % of the interior walls of the vacuum vessel are covered with carbon tiles. TCV is regularly boronised.

One of the main objectives of TCV is the study of the importance of plasma shaping for MHD stability and confinement. Elongations of up to $\kappa = 2$ and triangularities in the range $-0.7 \leq \delta \leq 1$ have already been achieved [1]. The shape is determined by magnetic diagnostics in connection with the inverse equilibrium reconstruction code LIUQE [2].

A second diagnostic for the plasma shape is soft X-ray tomography, which in addition allows access to transient phenomena due to MHD activity, which play an important role in confinement. The problem consists of determining the distribution of the emissivity of soft X-rays in a poloidal cross section of the plasma from a number of line integrated measurements. A set of pinhole

"cameras" holding several X-ray detectors each is placed around the vacuum vessel. A cone is defined by the relative position of each detector and the corresponding aperture, usually a slit. The field of view should be sufficiently narrow to justify the approximation of the cone by a line of sight. X-ray emission along these lines of sight is integrated to yield a set of chordal measurements. The inversion of the line integrated data to obtain the two–dimensional emissivity distribution is an ill-posed problem. A great variety of algorithms to solve this underdetermined problem exists, some of which will be discussed below. A second problem is the precise relative calibration of the detectors which is required to obtain a reasonable inversion, which will also be adressed in the following.

TCV is equipped with a system of 9 pinhole cameras for soft X-ray tomography with 200 lines of sight altogether, which will be described in section 2. In section 3, we will formulate the mathematical problem of tomography and present a survey of the different algorithms which we use to solve it, one of which has been newly developped for TCV. A paragraph is dedicated to the description of a new 'dynamical' calibration method which has proven very useful. The biorthogonal decomposition finally is used to analyse the inverted data.

Section 4 describes some numerical experiments which have been carried out to access the performance of our inversion algorithms. Some experimental data recently obtained on TCV have been analysed to investigate the influence of the triangularity of the plasma on the sawtooth activity. Results are presented in section 5. We conclude with a brief summary and an outlook on the future development of TCV's soft X-ray tomography system.

2 Experimental Setup

The ability of reconstructing complex plasma shapes is strongly dependent on the number of cameras arranged around the plasma cross section. The system on TCV makes use of 10 cameras with 20 detector elements each (see figure 1). The design was also constrained by the available ports and the necessity of vessel bakeout to up to $400^{\circ}C$.

Due to different sizes of the portholes available on the TCV vessel, three different types of cameras had to be built, which only differ in the size of the housing and the positioning of the detector head, the principle of construction being essentially identical.

The cylindrical camera housings are made of stainless steel. The walls are

formed by three layers of 1mm thick stainless steel with $1-2mm$ spacing each. The space between the inner two layers is connected to a water-cooling circuit while a second one can be evacuated to improve the thermal insulation (see figure 2). The walls of the concentric cylinders are welded to the first front plate of the cameras via a thin, bended structure to reduce conductive heat transfer from the ports during bakeout. The first front plate holds a small aperture which serves to collimate the field of view in poloidal direction. A second front plate, which is placed several centimeters behind the carbon tiles of the vessel's inner wall serves as a heat shield and has a slit which limits the view toroidally, but not poloidally. An Aluminium plate which closes the camera rear holds a vacuum feedthrough for the electrical connections as well as the connections to provide the primary vacuum required inside the camera to avoid absorption of the soft X-rays on the way between the Be-absorber and the detectors. The plate can be removed, allowing access to the detectors without breaking the torus vacuum.

The poloidally collimating aperture is made from Tungsten to prevent transparency of the trapezoidal edges (see detail of figure 2). A precise shaping of the edges with a minimum thickness of $20\mu m$ was obtained by electroerosive processing. The optical throughput or *étendue géométrique*

7

of the aperture/detector arrangement varies between $5.7 \cdot 10^{-9} m^2 sr$ and $7.7 \cdot 10^{-8} m^2 sr$.

Separation of torus vacuum and camera vacuum is provided by Beryllium foils with a thickness of $47 \mu m$ which at the same time block off photons with an energy of less than about $1 keV$. The Beryllium foil is curved to provide equal absorber thickness for all detectors, which is essential for a precise relative calibration.

The detectors are CENTRONIC LD20-5T linear arrays of 20 pn-photodiodes which were designed for visible and ultraviolet radiation. To use them as soft X-ray detectors, the glass window was removed. The active surface of a single diode is $4 \times 0.95 mm^2$, the length of an array is $20 mm$. The arrays are mounted on disk-shaped printed circuits with a diameter of $55 mm$. To minimise the distance between detector and the preamplifiers, 20 SMD operational amplifiers with a transimpedance gain of $2.2 \cdot 10^5 V/A$ and a bandwidth of $200 kHz$ are mounted on the back side of each printed circuit. Detectors and preamplifiers are held by aluminium housings which can be fixed to the inner front plate of the camera which also holds the aperture and the Beryllium foil.

The output of the preamplifiers is transmitted by twisted pairs of cables,

with a ground and a signal for each of the detectors. The 22 pairs of cables (20 signals and a $\pm 15V$ power supply) are grouped inside every camera, shielded and connected to the exterior via a vacuum feedthrough. The signals are then fed to remotely controllable amplifiers with gains of 1, 10, 100 or 1000. A two–stage low–pass filter is usually set to the Nyquist frequency to avoid aliasing. The maximum acquisition frequency of the $12bit$ analog–to–digital converters is $10kHz$.

3 Data Analysis

The following paragraphs provide the theoretical background of the tomography software package developped for TCV. After an outline on tomographic inversion algorithms, we briefly describe the way we simulate emissivity data which are used as well for the calibration as for performance tests. Inverted data are analysed with the help of the biorthogonal decomposition, which will be presented in the following paragraph. A few notes on the implementation of the software conclude this section.

3.1 Introductory Remarks

The goal of soft X-ray tomography is to reveal the spatial distribution of the so-called emissivity g. The emissivity is determined by the plasma parameters such as temperature, density, and the distribution of impurities and by the absorbers which define the spectral range which can be observed, usually thin Beryllium or Aluminium foils. To define the tomography problem precisely and to clarify to what extent approximations and simplifications are used, some definitions are at hand.

The dimension of the *spectral* emissivity $G = G(\vec{r}, \nu)$ is *power per volume and frequency interval*, where ν is the frequency of the radiation. Let $\Omega_\ell(\vec{r})$ be the solid angle subtended by one of n_ℓ detectors with efficiency $\eta_\ell(\nu)$. If we assume that the power is radiated isotropically, the total power P_ℓ detected by detector $\#\ell$ equals

$$P_\ell = \int\limits_{S_\ell} d\vec{r} \int d\nu \; \frac{\Omega_\ell(\vec{r})}{4\pi} \cdot G(\vec{r}, \nu) \cdot \eta_\ell(\nu) \quad \ell = 1 \ldots n_\ell \tag{1}$$

where the integrals run over the entire frequency range and the cone defined by the detector–aperture geometry. If the field of view of the detector is sufficiently narrow, we can assume that the emissivity does not vary on a surface perpendicular to the central line of the cone S_ℓ, which is in the following

10

referred to as "line of sight". The volume integral is thus transformed to a line integral $d\vec{r} \to A(s) \times ds$, where ds is a line element along the line of sight. This leads to

$$P_\ell \approx \frac{(A\Omega)_\ell}{4\pi} \cdot \int_{S_\ell} ds \int d\nu \ G(\vec{r}, \nu) \ \eta_\ell(\nu). \tag{2}$$

where the factor $(A\Omega)_\ell$, the *étendue géometrique* or optical throughput, is taken outside the integral, which is justified by a theorem of geometrical optics which states that the throughput is conserved (see *e.g.* [3]). The chord brightness f_ℓ is defined as

$$f_\ell = \frac{P_\ell}{(A\Omega)_\ell / 4\pi} \tag{3}$$

and is measured in $[W \cdot m^{-2}]$. A further, generally applied approximation is to assume that possible differences in detector response can be allowed for using a calibration factor c_ℓ for every diode (see section 3.4), which leads to

$$f_\ell = c_\ell \cdot \int_{S_\ell} ds \ g(\vec{r}) \tag{4}$$

where the emissivity

$$g(\vec{r}) = \int d\nu \ G(\vec{r}, \nu) \tag{5}$$

has been introduced.

11

3.2 Mathematical Definition of the Tomography Problem

The task of X-ray tomography is to reconstruct the two–dimensional distribution of the local emissivity g from a limited number of line integrated measurements f_ℓ. Mathematically, the problem consist of solving the system of integral equations

$$f_\ell = \int\limits_{S_\ell} ds \ g, \quad \ell = 1 \dots n_\ell \tag{6}$$

where the integral is along the line of sight and n_ℓ is the number of available measurements. This system of inhomogeneous *Fredholm* equations of the first kind [4] is always underdetermined, since an infinite number of measurements f_ℓ would be required to determine g exactly.

Due to restricted space, the number of line integrated data available in fusion research is usually limited to the order of some 10^2, whereas in medical tomography some 10^5 are available.

3.3 Possible Solutions to the Tomography Problem

There are essentially two ways to adress the tomography problem:

- It is possible to reduce the number of degrees of freedom by expanding the emissivity distribution in a set of orthogonal functions. If we choose *e.g.* polar coordinates, a Fourier decomposition for the angular part and a polynomial expansion for the radial part of g can be used. The underdetermined system of integral equations is thus transformed to an overdetermined system of algebraic equations which can be solved for the coefficients of the base functions using a least-squares fit. This is done by the *Cormack-Granetz* algorithm [5], which is widely used in fusion research.

- The area of the poloidal cross section where we want to reconstruct the emissivity distribution is subdivided into pixels. The size of the pixels has to be sufficiently small to justify the assumption of approximately constant emissivity within one pixel. At the same time, the number of pixels has to be sufficiently small to provide a tractable system of equations.

We shall concentrate here on the second possibility.

Many pixel shapes can be thought of: square pixels, radial concentric pixels with and without angular separations [6], concentric pixels determined

13

by the flux surface geometry [7] and even hybrids of pixel and orthogonal function expansions have been used successfully [5, 8]. In the following, we shall restrict ourselves to a rectangular grid of square pixels with a horizontal direction x and a vertical direction y.

3.3.1 Pixel Methods: Some Generalities

An advantage of the pixel Ansatz is that the system (6) is transformed to a system of algebraic equations in a very natural way. We can store the $n_{pixel} = n_x \cdot n_y$ values g_i of our two–dimensional emissivity distribution as lines of the column vector $\mathbf{g}$ (n_x and n_y are the number of pixels in horizontal and vertical direction, respectively). The n_ℓ line integrated data are put into a column vector $\mathbf{f}$. We get

$$f_\ell = \sum_{i=1}^{n_{pixel}} T_{\ell i} g_i \quad (\ell = 1 \dots n_\ell) \tag{7}$$

or simply

$$\mathbf{f} = \mathbf{T} * \mathbf{g} \tag{8}$$

where $*$ denotes usual matrix multiplication. In the simplest approximation, the matrix element $T_{\ell i}$ is equal to the length of the chord $\# \ell$ in pixel $\# i$.

The size of **T** is $n_\ell \times n_{pixel}$, the number of lines of sight times the number of pixels.

The most obvious idea to solve (8) is to invert **T**. In most cases this is impossible, either because there are less equations than unknowns (*i.e.* the inverse of **T** does not exist) or, even if we have $n_\ell = n_{pixel}$, the matrix **T** may be badly conditioned.

If we had *more* line integrated measurements than pixels, *i.e.* $n_\ell > n_{pixel}$, we would try to

$$minimise \quad \chi^2 \tag{9}$$

with

$$\chi^2 = \sum_\ell (\frac{\sum_i T_{\ell i} g_i - f_\ell}{\sigma_\ell})^2 \tag{10}$$

which is the same as

$$\chi^2 = (\tilde{\mathbf{T}} * \mathbf{g} - \tilde{\mathbf{f}})^T * (\tilde{\mathbf{T}} * \mathbf{g} - \tilde{\mathbf{f}}) \tag{11}$$

The exponent ^{T} denotes transposition. For convenience, we have used the abbreviations $\tilde{T}_{\ell i} = T_{\ell i}/\sigma_\ell$ and $\tilde{f}_\ell = f_\ell/\sigma_\ell$, where σ_ℓ is the standard deviation of f_ℓ. The solution of the set of normal equations

$$\tilde{\mathbf{T}}^T * \tilde{\mathbf{T}} * \mathbf{g} = \tilde{\mathbf{T}}^T * \tilde{\mathbf{f}} \tag{12}$$

15

then yields a least-squares-fit solution to the tomography problem. This can work very well, as has been demonstrated *e.g.* by *Decoste* [9].

In the limit of much fewer equations, hence f_l's, than unknown g_i's, we can always achieve $\chi^2 = 0$, because then there is an infinite number of solutions ("overfitting"). The general idea to obtain a unique and sensible solution in case of an underdetermined problem is to look for a minimum of a functional ϕ which may be written as

$$minimise \quad \phi = \frac{1}{2}\chi^2 + \alpha \, \mathcal{R} \qquad (13)$$

instead of searching a minimum of χ^2 alone. $\mathcal{R}$ is a regularising functional, the regularisation parameter α is a positive number which determines the weighting between the goodness-of-fit, represented by χ^2, and the requirements imposed on the solution **g** by the functional $\mathcal{R}$, *e.g.* the smoothness of the solution. In the limit $\alpha \to 0$, the solution is determined by χ^2 alone as above, in the limit of very large α it is only the smoothing (or whatever we may have required) which determines the solution. The most likely solution must be somewhere in between, so one part of the problem is to find a way to choose the "correct" value of the regularisation parameter, a second part is to find a solution g for a given value of α.

16

Three different choices of ϕ will be discussed in the following: Linear regularisation, a new method relying on the Fisher information and Maximum Entropy.

3.3.2 Linear Regularisation

The first method we want to discuss is the so-called linear regularisation method. All linear regularisation algorithms have in common that the solution, apart from providing a reasonable fit to the data, is expected to be smooth (see *Press et al.* [4] and references therein). Combinations with other algorithms like the *Gilbert–Backus* method [4, 10, 11] have also been used successfully. A comparison to a Neural Networks approach has been published by *Steriti* and *Fiddy* [12]. We present a two–dimensional extension of the one dimensional case described in [4].

If we expect a smooth solution, the functional $\mathcal{R}$ has to measure the roughness of the solution somehow. The simplest approach is to require the solution vector g to have minimum norm,

$$\mathcal{R} = \|\mathbf{g}\|^2 = \mathbf{g}^T * \mathbf{g} \tag{14}$$

where $\| \cdot \|$ denotes the usual euclidean vector norm. This is called zeroeth

order regularisation.

First order linear regularisation is equivalent to require the solution to have small gradients. We can choose

$$\mathcal{R} = \|\mathbf{g_x}\|^2 + \|\mathbf{g_y}\|^2 \tag{15}$$

as the regularising functional where $\mathbf{g_x}$ and $\mathbf{g_y}$ denote the partial derivatives with respect to x and y, respectively. If we assume that ∇_x and ∇_y are finite-difference matrix representations of the corresponding differential operators, we can write

$$\mathcal{R} = (\nabla_x \ \mathbf{g})^{\mathbf{T}} * (\nabla_{\mathbf{x}} \ \mathbf{g}) + (\nabla_{\mathbf{y}} \ \mathbf{g})^{\mathbf{T}} * (\nabla_{\mathbf{y}} \ \mathbf{g}) \tag{16}$$

or

$$\mathcal{R} = \mathbf{g}^T * (\nabla_{\mathbf{x}}^T * \nabla_{\mathbf{x}} + \nabla_{\mathbf{y}}^T * \nabla_{\mathbf{y}}) * \mathbf{g}. \tag{17}$$

With the help of the definition

$$\mathbf{H} = \nabla_{\mathbf{x}}^T * \nabla_{\mathbf{x}} + \nabla_{\mathbf{y}}^T * \nabla_{\mathbf{y}} \tag{18}$$

we get from (11) and (13)

$$minimise \quad \phi = \frac{1}{2}(\tilde{\mathbf{T}} * \mathbf{g} - \tilde{\mathbf{f}})^T * (\tilde{\mathbf{T}} * \mathbf{g} - \tilde{\mathbf{f}}) + \mathbf{g}^T * \mathbf{H} * \mathbf{g} \tag{19}$$

Setting all partial derivatives $\partial \phi / \partial g_i$ to zero, we obtain the normal equations

$$(\tilde{\mathbf{T}}^T * \tilde{\mathbf{T}} + \alpha \mathbf{H}) * \mathbf{g} = \tilde{\mathbf{T}}^T * \tilde{\mathbf{f}} \qquad (20)$$

which have to be solved for **g**, which can be done by standard methods like _e.g._ LU-decomposition[1].

Second order linear regularisation is equivalent to looking for a solution fitting the line integrated data and having minimum curvature. The regularisation functional can be expressed as

$$\mathcal{R} = \|\triangle \mathbf{g}\|^2 = (\triangle \mathbf{g})^T * (\triangle \mathbf{g}) \qquad (21)$$

If $\triangle$ denotes a finite difference matrix representation of the Laplacian, we can write

$$\mathcal{R} = (\triangle * \mathbf{g})^T * (\triangle * \mathbf{g}) = \mathbf{g}^T * \triangle^T * \triangle * \mathbf{g} \qquad (22)$$

which means that the matrix **H** of eqn (20) is given by

$$\mathbf{H} = \triangle^T * \triangle \qquad (23)$$

.

It has to be stressed that the equation (20) is the _same_ for _all_ orders of linear regularisation. It is just the matrix **H** which has to be changed, for an

---

[1]If you are a lucky user of MATLAB, you just employ the '\'- operator.

n'th order regularisation, H contains the (finite difference) matrices of the n'th derivative of the solution, analogous to (21). For zero order regularisation (14), the matrix H is simply the unit matrix. Higher order algorithms are possible, one can even think of a mixture of different orders ("solution close to a differential equation", see [4]).

The "correct" α has to be determined iteratively, *i.e.* for a fixed α eq (20) has to be solved. It has to be verified if a certain criterion is fulfilled, if not, α has to be changed and a new solution has to be calculated. A possible criterion is

$$\chi^2 \approx n_\ell \tag{24}$$

provided the measurement errors σ_ℓ are known sufficiently well (see the discussion of Maximum Entropy below).

3.3.3 Minimum Fisher Information

Reinmuth has shown [13] that the so–called Fisher information can successfully be used for the assessment of inverse problems. The Fisher Information of a probability distribution is defined as

$$I_F = \int \frac{g'(x)^2}{g(x)} \, dx \tag{25}$$

where the prime denotes the derivative with respect to x and we assume that the integral $\int g(x)dx$ equals unity.

If we identify the distribution g with the soft X-ray emissivity and x with the pixel index i, the motivation to use the Fisher information as a regularising functional is evident. We see immediately from the definition (25), that minimising the Fisher information of g implies a minimisation of the absolute value of the first derivative of g, as is the case for first order linear regularisation (see above). But the denominator of the integrand in eqn (25) *weighs* the smoothing in the sense that for a fixed contribution to the integral, the absolute value of the derivative is allowed to be larger if the value of g itself is also large than in a case where g itself is small. This means that the smoothing is strongest where the values of g are small. For soft-X-ray tomography, this seems reasonable: small values of g correspond to low emissivity, hence border regions of the plasma with low temperature. The contribution of these areas to the chord brightness f_ℓ are small due to the transmission function of the Be-absorbers, so not much information about the low–emissivity regions is contained in the line integrated data f_ℓ. On the other hand we do not want to smooth structure in the center of the plasma where the emissivity is high, so the weighted smooth provided by a

minimisation of the Fisher information of g seems to suit our problem very well.

Instead of treating the fully non-linear problem of minimising the Fisher information under the constraints (6) as sketched in *Reinmuth's* diploma thesis [13], and by *von der Linden* [14], we rather exploited the central idea to conceive a weighted linear regularisation method.

Starting from first order linear regularisation as described above, we are free to add a diagonal weight matrix $\mathbf{W}$, as long as all elements are greater than zero ($W_{ii} > 0$):

$$\mathbf{H} \rightarrow \nabla_{\mathbf{x}}^T * \mathbf{W} * \nabla_{\mathbf{x}} + \nabla_{\mathbf{y}}^T * \mathbf{W} * \nabla_{\mathbf{y}} \tag{26}$$

To minimise the Fisher information of the distribution $\mathbf{g}$, we can not directly insert $1/g_i$ as a weight, since this would make the method nonlinear. Instead, we start with $\mathbf{W}^{(0)} \equiv 1$, the unit matrix, which is equivalent to first order linear regularisation. We solve the normal equations (20) with $\mathbf{H}$ defined by equation (26) and use the solution $\mathbf{g}^{(1)}$ thus obtained to determine a new weight matrix $\mathbf{W}^{(1)}$ such that

$$
\begin{aligned}
W_{ij}^{(0)} &= \delta_{ij} \\
W_{ij}^{(n)} &= \frac{1}{g_i^{(n)}} \cdot \delta_{ij}, \quad g_i^{(n)} > 0 \quad i = 1 \ldots n_{pixel}
\end{aligned}
\tag{27}
$$

and

$$W_{ij}^{(n)} = W_{max} \cdot \delta_{ij} > 0, \quad g_i^{(n)} \leq 0$$

where the superscript (n) denotes the solution of the n-th iteration and W_{max} is an upper limit for the weights. This can be written as

$$(\tilde{\mathbf{T}}^T * \tilde{\mathbf{T}} + \alpha \mathbf{H}^{(n)}) * \mathbf{g}^{(n+1)} = \tilde{\mathbf{T}}^T * \tilde{\mathbf{f}} \qquad (28)$$

where $\mathbf{H}^{(n)}$ is defined by (26) and (27) and $\mathbf{g}^{(n+1)}$ is the new solution. This procedure can be continued until the weight matrix elements remain constant within a certain limit. It turns out that, although 4-5 iterations may be needed to have stable $\mathbf{W}$-matrix elements, the solution $\mathbf{g}$ remains almost unchanged after the first iteration, $i.e.$ according to our notation, $\mathbf{g}^{(2)}$ represents a reasonable approximation to the sought–for solution. We have thus introduced a regularisation which remains linear, but the solution of which should have minimum Fisher information.

3.3.4 Maximum Entropy

Several algorithms exploiting in one way ore another the information entropy of a probability distribution exist. The common feature of all Maximum Entropy algorithms is that the values of the emissivity assigned to the pixels

g_i are considered as statistically independent, random variables. In addition to a reasonable fit to the data, the information entropy of the emissivity distribution is assumed to attain its maximum value.

Different degrees of sophistication and depth of philosophical background are possible. A brief introduction is provided by [4], although the *Bayesian*[2] aspect of Maximum Entropy is somewhat played down by the authors. A palatable survey on the theory and different applications of *Bayesian* Maximum Entropy is given in a fairly recent book editet by *Buck* and *Mackaulay* [15]. A review of a variety of MaxEnt algorithms has been given by *Djafari* [16]. We will discuss a Bayesian MaxEnt algorithm described by *von der Linden* [17, 18].

Bayes' theorem relates conditional probabilities P from two events A and

---

[2]It was *Sir Harold Jeffreys* who re-discoverd *Bayes* theorem to approach statistics in a philosophically different way than the usual "frequentist" school of thought, dominated by *Fisher*. The general idea of Bayesian statistics is that we assign probabilities in case of *incomplete knowledge* of the system under consideration, *i.e.* probabilities are rather degrees of plausibility than relative frequencies of occurence as in the usual frequentist point of view. Probability theory thus becomes an extension of logic rather than just a toolbox to handle (seemingly) random data.

B in the following way:

$$P(A|B) = P(A) \cdot \frac{P(B|A)}{P(B)} \qquad (29)$$

$P(A|B)$ is the conditional probability of A given that B has occured, $P(A)$ and $P(B)$ are unconditional probabilities (see [4, 19] or any textbook on probability theory).

For so-called inverse problems like tomographic inversion, the theorem can be rewritten as [17, 18, 19]

$$P(\mathbf{g}|\mathbf{f}_{exp}, I) = P(\mathbf{g}|I) \cdot \frac{P(\mathbf{f}_{exp}|\mathbf{g}, I)}{P(\mathbf{f}_{exp}|I)}, \qquad (30)$$

where names and meanings of the symbols are listed below:

$\mathbf{f}_{exp}$ experimental, line integrated data (chord brightness)

I any other available information *a priori*

$\mathbf{g}$ the sought-for emissivity distribution

$P(\mathbf{g}|I)$ the probability of the solution prior to any experiment

$P(\mathbf{f}_{exp}|I)$ the probability for the experimental data, given only the

                  prior information I; $P(\mathbf{f}_{exp}|I)$ usually remains unknown

$P(\mathbf{f}_{exp}|\mathbf{g}, I)$ the probability for the measured data $\mathbf{f}_{exp}$ if $\mathbf{g}$ and I

                  were known, known as the likelihood function

$P(\mathbf{g}|\mathbf{f}_{exp}, I)$ the so-called posterior probability for the solution $\mathbf{g}$,

                  given the experimental data $\mathbf{f}_{exp}$ and the *a priori* in-

                  formation I

The general idea is to search for a maximum of the posterior probability $P(\mathbf{g}|\mathbf{f}_{exp}, I)$, *i.e.* to look for the most probable solution, which is sometimes referred to as the MAP (Maximum A Posteriori) estimator.

The denominator on the rhs of equation (30) usually remains unknown and serves only as a normalisation constant. The likelihood function is well-known and can be expressed as

$$P(\mathbf{f}_{exp}|\mathbf{g}, I) \propto exp \left(-\frac{1}{2}\chi^2 \right) \tag{31}$$

with χ^2 given by eqn (11), as usual. *Skilling* [20] showed that the most uninformative, unbiased prior is the entropic one, given by

$$P(\mathbf{g}|I) = (\frac{\alpha}{2\pi})^{n_{pixel}/2} \cdot \frac{1}{\sqrt{\Pi_i\, g_i}} \cdot exp(\alpha\, S). \tag{32}$$

The entropy S is defined with respect to a default model $\mathbf{m}$ of the emissivity $\mathbf{g}$ as

$$S = \sum_i g_i - m_i - g_i\, ln(\frac{g_i}{m_i}) \tag{33}$$

The entropy attains its maximum for $\mathbf{g}$ given by the default model. If we combine equations (31) and (32), we see that a maximum of the posterior probability occurs at a maximum of

$$\phi(\alpha,\mathbf{g}) = -\frac{1}{2}\chi^2 + \alpha\, S, \tag{34}$$

the positive number α still being the regularisation parameter (compare eqn (13)).

For α fixed, we have to find the solution $\mathbf{g}^*$. Following *von der Linden* [17, 18], this can be done introducing n_ℓ Lagrangian parameters λ_ℓ, thus reducing the number of unknowns from n_{pixel} to n_ℓ. Maximising (34) under the exact "constraints"

$$0 \quad = \quad F_\ell - f_\ell(\mathbf{g})$$

with

$$f_\ell(\mathbf{g}) \quad = \quad \sum_i T_{\ell i}\, g_i \tag{35}$$

$$\tag{36}$$

is equivalent to maximising

$$\phi_\lambda = \alpha\ S - \frac{1}{2}\sum_\ell (\tilde{f}_\ell^{\,exp} - \tilde{F}_\ell)^2 + \alpha\ \sum_\ell \lambda_\ell (F_\ell - f_\ell(\mathbf{g})) \tag{37}$$

with the ~ designing division by the error σ_ℓ as above and F_ℓ designing

an auxiliary variable which will immediately disappear. Setting all partial

derivatives to zero yields

$$g_i = m_i\ exp(-\sum_\ell \lambda_\ell T_{\ell i}), \tag{38}$$

which automatically assures that the emissivity is nonnegative. The resulting

system of n_ℓ implicit equations for the Lagrangian parameters

$$f_\ell^{\,exp} - \sum_i T_{\ell i}\ m_i\ exp\left(-\sum_{\ell'} \lambda_{\ell'}\ T_{\ell' i}\right) + \alpha\ \sigma_\ell^2\ \lambda_\ell = 0 \tag{39}$$

has a unique solution [17, 18] which can be determined using a Newton-

Raphson scheme [4].

Determining the maximum posterior probability (30) as a function of

the regularisation parameter is a rather formidable task, because it involves

28

the evaluation of n_{pixel}-dimensional integrals. As indicated in [17, 18], the integrals become tractable if we expand the integrand into a Taylor series around the solution $\mathbf{g}^*$. In that case the posterior probability can be expressed as

$$P(\mathbf{g}|\mathbf{f}_{exp}, I) = \left(\frac{1}{2\pi}\right)^{n_{pixel}/2} \cdot \left[det(\ \delta_{ll'} + \frac{1}{\alpha}\sum_i \tilde{T}_{li}\ g_i^*\ \tilde{T}_{l'i})\ \right]^{-\frac{1}{2}} \cdot exp(\ \phi(\alpha, \mathbf{g}^*)\ )$$

(40)

Although quite costly in terms of computer time, this approach to determine α is much more satisfactory than the "historic" criterion (24), which is used in the so-called Classical Maximum Entropy Methods as well as in linear regularisation. It turns out that the χ^2 of a solution $\mathbf{g}$ with maximum posterior probability can be much smaller than n_l. As has been discussed by *Skilling* [20] and *von der Linden* [17, 18], the "historic" criterion leads to an underestimate of the information provided by the data. The application of (24) would be justified if we had n_l measurements of the *same* quantity, which is clearly not the case in tomography or other fields where inverse methods are used.

An advantage of MaxEnt becomes clear from eqn (38): non–physical negative emissivity values are automatically excluded. A disadvantage of MaxEnt

is the amount of calculations involved as well as the fact that pixel emissivities are treated as completely independent, which results in reconstructions which tend to be less "smooth" than expected. A smoothing can be introduced applying a so-called "preblur" [19].

3.4   Simulation and Calibration

Artificial data for the emissivity distribution and the chord brightness are generated using the poloidal flux surfaces obtained from the equilibrium reconstruction code LIUQE for actual TCV discharges. We assume that the temperature and the density are constant on isocontours of the normalised flux $\Psi = \psi/\psi_{axis}$. Temperature and density profiles are assumed to depend on Ψ like

$$n_e \quad \propto \quad n_o \cdot \Psi^{a_{n_e}} \tag{41}$$

$$T_e \quad \propto \quad T_o \cdot \Psi^{a_{T_e}} \tag{42}$$

where the exponents a_{n_e} and a_{T_e} are adjusted to realistically model actual profiles measured using the Thomson scattering diagnostic [21]. The central electron temperature T_o is usually taken from Thomson scattering or measurements of the central T_e by a soft X-ray filter method.

Assuming a pure Hydrogen or Deuterium plasma, the spectral emissivity depends on n_e and T_e like (see *e.g.* [22])

$$G_o(\vec{r}, \nu) \propto n_e^2 \cdot \frac{1}{\sqrt{k_B T_e}} \cdot \exp\left(-\frac{h\nu}{k_B T_e}\right) \qquad (43)$$

where k_B is Boltzmann's constant, h is Planck's constant and the dependence of G_o on $\vec{r}$ is implicit by $\Psi = \Psi(\vec{r})$. The distribution (43) describes continuum radiation, *i.e.* free-free and free-bound transitions. The distribution is cut off at low energies by the transmission τ of our Beryllium absorber, which leads to

$$G(\vec{r}, \nu) \quad = \quad \tau(\nu) \cdot G_o(\vec{r}, \nu) \qquad (44)$$

with

$$\tau(\nu) \quad = \quad \exp\left(-\alpha(\nu)\right), \qquad (45)$$

where data for the absorption coefficients $\alpha(\nu)$ were taken from a compilation by *Veigele* [23]. If we are only interested in the spectral distribution and not in absolute values, (43) is also valid for homogeneous distribution of light impurities up to Oxygen since the $47\mu m$ thick *Be* absorbers mounted on TCV do not transmit line radiation with $h\nu < 1keV$.

Two different sets of line integrated data f_ℓ are calculated from the simulated spectral emissivity distribution: one assuming ideal detectors, *i.e.* de-

tectors with a spectral efficiency $\eta(\nu)$ identical to unity for all frequencies, which is also used to obtain the emissivity distribution g. The second set is obtained using a model of the spectral efficiency of the photodiodes and experimentally obtained parameters of our actual detectors, which has been described in [24]. This second set serves to check our calibration method.

A dead layer on top of the photodiodes as well as the finite depth of the sensitive region lead to an angular dependence of the spectrum averaged efficiency $< \eta_l >_\nu$ as defined in [24] and in eqn (47) below. The spectrum average efficiency can be identified with the calibration factor c_l from eqn(4). Essentially, the integral on the rhs of eqn (2) is approximated by

$$\int_S ds \int d\nu \; G(\vec{r},\nu) \; \eta(\nu) \quad \approx \quad \frac{\int_S \int d\nu \; \tilde{G}(\vec{r},\nu) \; \eta(\nu)}{\int_S \int d\nu \; \tilde{G}(\vec{r},\nu)} \cdot \int_S ds \; g \qquad (46)$$

or

$$c_l = < \eta_l >_\nu \quad \equiv \quad \frac{\int_{S_l} \int d\nu \; \tilde{G}(\vec{r},\nu) \; \eta_l(\nu)}{\int_{S_l} \int d\nu \; \tilde{G}(\vec{r},\nu)} \qquad (47)$$

where $\tilde{G}$ is a reasonable approximation to the unknown actual spectral emissivity G, for example (44), and g is given by (5). Instead of calculating the calibration factors once and for all as usual, the calibration factors used on TCV are recalculated for every discharge or even separately for different phases of one discharge using the actual central electron temperature and the

profile parameters a_{n_e} and a_{T_e}. The effects of *not* proceeding like this have for the first time been described in [25], we will show an example in section 4.2.1.

In addition to the equilibria, we have also simulated MHD mode structures. We start from a simulation of an equilibrium as described above and add a perturbation to the spectral emissivity distribution which can be described by

$$\Delta g(\Psi, t) = \Delta g(\Psi_m) \cdot \left[1 + \left(\frac{\Psi - \Psi_m}{\Psi_{HWHM}} \right)^2 \right]^{-1} \cdot \cos\left(m \cdot \theta + \phi(t) \right) \qquad (48)$$

The amplitude $\Delta g(\Psi_m)$, the flux coordinate of the mode Ψ_m with the poloidal harmonic m as well as the half-width Ψ_{HWHM} and the time–dependent phase $\phi(t)$ are considered as free parameters, θ is the poloidal angle. Simulations of such time series of emissivity distributions have been used to check the precision of the different reconstruction methods as shown below (section 4.1.2).

3.5 The Biorthogonal Decomposition

One of the strengths of X-ray tomography is the possibility to investigate MHD activity, *i.e.* the main interest is rather the investigation of a time

series of reconstructions than just the reconstruction of a few "stills". A mathematical tool which has proven very useful for the investigation of such time series is the so–called biorthogonal decomposition or BD, see for example *Dudok De Wit et al.* [26] and references therein. Technically speaking, the biorthogonal decomposition can easily be performed using the singular value decomposition, in the following referred to as SVD.

Suppose we have a matrix X with M lines and N columns. The result of an SVD of X are three matrices U, S and V so that X can be recovered by forming their product:

$$X = U * S * V^T \qquad (49)$$

where ^{T} denotes transposition. If X were a square matrix, U and V would contain the eigenvectors of X, whereas S would be a diagonal matrix of the corresponding eigenvalues. The SVD provides a generalisation to rectangular matrices. The matrix X is decomposed into a finite number K of "modes" where K is the minimum of its number of rows and columns. (For a discussion of the orthogonality properties and the mathematical definitions see [4] or [26]).

Contrary to previous publications of a successful use of SVD for the

analysis of soft X-ray data, we rather use SVD *after* the tomographic re-construction: If the lines of **X** contain the time series of the different pixel emissivities[3], the columns of the matrix **U** represent the spatial "eigen"-modes, in the following called *topos*, whereas the corresponding temporal information is contained in the columns of **V** or *chronos*[4]. The pairs of *topos* and *chronos* are coupled by the elements of the diagonal matrix **S**, their corresponding singular values. The elements of **S** are stored in descending order, so that the first topos/chronos pair contains the most important features, the second contains the most important changes to the "average" described by pair one and so on.

If we define the global signal energy [26] of **X** as

$$E_X = \sum_m^M \sum_n^N (X_{m,n})^2, \tag{50}$$

it can be shown that E_X is equal to the sum of the squares of the diagonal elements of **S**:

$$E_X = \sum_k^K S_k^2 \qquad K = min(M, N) \tag{51}$$

---

[3]This means that the columns of **X** are formed by the emissivity vectors **g** as defined above.

[4]We adopt here the names used by *Dudok De Wit et al.* which stem from the greek words for 'place' and 'time'.

The relative signal energy

$$p_k = S_k^2 / E_X \qquad (52)$$

is a useful quantity to measure the contributions of different features of the reconstruction, as will be shown below (sections 4.1.2 and 5).

3.6 Implementation of the Software

A software package which includes all the features described in the preceding sections has been written in the high–level language MATLAB. A library of processor–optimized numerical routines as described in the LAPACK manual [27] is provided by DEC. An interface from MATLAB to these so–called DXML routines has been written by *Moret* [28]. Matrix multiplication and LU-decomposition via DXML speeds up the inversion algorithms by a factor of five. A graphic user interface facilitates data retrieval, calibration and inversion procedures. Many graphical display options are available. Inversion results are stored in a common database.

4 Performance of the Tomography System

4.1 Comparison of different algorithms using artificial data

We have determined the optimum pixel size in a series of tests, as well with simulated as with experimental data. As a result, we use a number of 15 pixels radially, which corresponds to a size of 3.8×3.8 cm^2. The number of vertical pixels depends on the elongation of the plasma under consideration. The grid is usually placed with the help of the last closed flux surface (LCFS) as given by the equilibrium reconstruction from the magnetic diagnostics. The outermost rows and columns of the pixelgrid are always chosen such that their entire pixels are outside the LCFS. The emissivity of those pixels is forced to zero during inversion by introducing artificial lines of sight. To assess the performance of the different reconstruction algorithms, we have used simulated data.

4.1.1 Simulated Equilibrium

To test the capability of our reconstruction algorithms to reconstitute the emissivity of an equilibrium, we have used the normalised flux from the equilibrium reconstruction of TCV discharge #8100 at 0.4 seconds (see figure 4). This was a limiter configuration in the equatorial plane of the vessel with an elongation of $\kappa_{95} = 1.6$ and a triangularity δ_{95} of +0.12. The central electron temperature used to generate the simulated emissivity distribution with the help of the equations given in section 3.4 was $700eV$, the profile parameters a_{T_e} and a_{n_e} being 0.8 and 0, respectively, assuming ideal detectors ($\eta \equiv 1$). The simulated chord brightness is displayed in figure 3, the results of the reconstructions are shown in figure 4 and in table 1.

Figure 4 displays the simulated emissivity distribution (left) and reconstructions from artificial data without any noise, using four different methods, namely Bayesian Maximum Entropy, Minimum Fisher regularisation and first and second order linear regularisation (left to right), as described in section 3.3. The second row shows reconstructions obtained from *noisy* data, *i.e.* we added normally distributed random noise with zero mean and a standard deviation of 2.5% of the maximum value of f_ℓ on all channels. The dashed lines

show the poloidal flux isocontours to facilitate a comparison of the different results.

The Maximum Entropy algorithm yields the "noisiest" reconstructions, while second order regularisation is smoothest. First and second order regularisation tend to reconstruct an emissivity profile which is slightly broader than the simulated one. This is not the case for the Minimum Fisher method, which yields reconstructions slightly less smooth. With respect to noise, the second order linear regularisation seems to be the least sensitive.

Table 1 provides some figures describing the performance of the algorithms: The number of floating point operations, the elapsed time on a DECα station and MATLAB V4.1 are listed as well as an average reconstruction error σ which we have defined as

$$\sigma = \sqrt{< \left(g_i - g_i^{sim}\right)^2 >} \tag{53}$$

where $< \cdot >$ stands for an average over all pixels and g_i^{sim} is the simulated emissivity distribution.

The results do not seem to justify the 7×10^8 floating point operations required for the Maximum Entropy inversion. Concerning speed and quality of the reconstruction, this example seems to favor the second order regular-

isation method, although the reconstruction error σ of the Minimum Fisher Regularisation is less or equal to that of second order regularisation. We will see in the following section that this first conclusion will have to be corrected.

4.1.2 Simulated MHD Mode activity

Since one of the main objectives of the soft X-ray diagnostic is the investigation of MHD activity in the plasma center, we generated a series of emissivity distributions and artificial line integrated data with a small temporal variation of the emissivity, simulating a rotating $m = 2$ island structure.

We chose TCV discharge # 9243 at $0.746s$, with negative triangularity $\delta_{95} = -0.23$ as a base for our simulation. The choice was motivated by the fact that similar TCV discharges tend to exhibit $m = 2$ modes which occasionally lock.

The parameters of the simulation are a central electron temperature of $600eV$ and profile parameters a_{T_e} and a_{n_e} of 0.8 and 0, respectively. We assumed that an $m = 2$ mode is situated relatively near the center at a normalised flux of $\Psi_m = 0.85$ with a half-width Ψ_{HWHM} of 0.15 and a relative amplitude of 0.1, which means that the variation of the emissivity due to mode activity is 10% of the equilibrium value at $\Psi = 0.85$ (see section

3.4). We calculated 41 successive emissivity distributions and allowed the phase angle to vary with time so as to simulate a slowing down rotation of the mode structure. The eight first distributions are displayed in figure 5.

To analyse the simulation and the performance of the different reconstruction algorithms, we use the biorthogonal decomposition or SVD as described in section 3.5 or by [26].

The emissivity is stored in a matrix X, the lines of which correspond to different pixels and the columns of which represent different times. Figure 6 displays the result of an SVD of the *simulated* emissivity matrix. The logarithm of the singular values S_k is plotted in the top left of figure 5 vs. the index k. The top right shows a plot of the maximum emissivity as a function of time. The second row of figure 5 represents the first three spatial modes or *topos* of the reconstruction, in descending order of the corresponding singular values (left to right). The bottom row displays the most important three temporal modes or *chronos* as a function of time.

From figure 5, we conclude that all the structure of the simulated emissivity distribution is contained in the first three *topos/chronos* pairs, since the higher singular values are smaller by 14 orders of magnitude. The first *topos/chronos* pair represents a kind of average, in our case the equilibrium,

while the higher terms represent changes with respect to the first pair.

Equal singular values as in this example usually indicate a correlation of two *topos/chronos* pairs. In our case, the second and third pair together represent the rotating island structure. For the topos, broken contours in figure 5 indicate negative values while positive values are plotted using continuous contours. The relative phase shift of the temporal functions together with the positive and negative sign of the *topos* allows to recover the rotating mode structure.

Reconstructions were produced from artificial line integrated data using three different levels of noise: zero, 1% and 2.5% of noise, respectively, the latter representing a rather pessimistic case compared to available experimental data. We have treated the matrices containing the reconstructed emissivity distributions in the same manner as the simulated data. One example is shown in figure 7, the analysis of the Minimum Fisher reconstruction from noisy data (1%). The noise is clearly visible on the maximum emissivity as well as on the temporal functions. But in spite of the noise, the first three *topos* are almost unchanged as well as the relative signal energies (see table 2). The noise has mainly increased the singular values of the higher order components ($k > 3$), which means that it is contained in the *topos/chronos*

pairs with with $k > 3$. (We could thus obtain a filtering of the reconstruction by leaving those components aside.)

We have summarized the results of the comparison in table 2. The first column indicates the data set (simulated or reconstructed), the second column lists the noise level of the artificial line integrated data used for the inversion. Column three displays the average reconstruction error σ as defined in equation (53). The following three columns contain the relative signal energy of the first three *topos/chronos* pairs a defined in section 3.5. The last column finally indicates if the mode can be easily recovered from the reconstruction (+), if it can be recognised, even though hardly (0) or if the information is completely lost (-).

One important information which is not contained in table 2 is the precision concerning location and spread of the island structure. To judge the precision of the reconstruction, please refer to figure 8, where the second spatial "eigenmode", *topos* # 2, is displayed for all different algorithms and noise levels investigated: The top left frame displays *topos* # 2 from the simulated distribution. The other frames of the first row show the same *topos* from an SVD of reconstructions using artificial data without any noise, using Maximum Entropy, Minimum Fisher Regularisation and 1st and 2nd order

43

linear regularisation (from left to right). The second and third row show the topos obtained from corresponding reconstructions using noisy data with noise levels of 1% and 2.5%, respectively.

If noise is negligible, the ability of Bayesian MaxEnt and Minimum Fisher Regularisation to reveal the mode structure are about the same, although MaxEnt takes six times more CPU time (compare table 1). In spite of the very promising reconstructions of equilibria using second order linear regularisation, both first and second order regularisation are unsatisfactory because we want to determine the location, spread and the signal energy of island structures precisely. What seemed only a slight tendency to "broaden" the reconstructions with respect to the simulated distribution in case of the equilibrium now turns out to completely smear out features of interest.

If noise is present, Bayesian MaxEnt does not even recover the mode, since the reconstruction itself is too "noisy". The newly developped Minimum Fisher algorithm on the other hand exhibits a remarkable precision and good stability with respect to noisy data at reasonable expenses of CPU time. In fact, it is the only algorithm capable of recovering the mode at 2.5% noise level without loosing precision on the location and without loss of relative signal energy to higher components, as is clearly revealed from table 2.

As a consequence, our new method, Minimum Fisher Regularisation, has been adopted as the "default" reconstruction method for the soft X-ray tomography system of TCV.

4.2 Limitations due to systematic errors

Due to restricted space and the resulting compact camera design, several sources of systematic errors connected to geometry can be identified. Another, independent source of error is magnetic pickup noise.

4.2.1 Geometrical Errors: Angular Factors and Calibration

The use of planar arrays of photodiodes led to angles of incidence of the different detectors which vary between zero and up to 50°, measured with respect to the surface normal of the diodes. In connection with the rather small detector–to–aperture distances involved, this leads to large variations of the angular factors $A\Omega$, which may differ in some cases by almost an order of magnitude. Of course, there is also an error caused by the fact that the position of the line of sight is incorrect.

The first reconstructions were tried using positions and distances as given by the technical drawings. The results indicated that there was a slight prob-

lem. This was solved using a consistency check between the tomographic reconstruction and the equilibrium reconstruction from the magnetics for a series of quiescent discharges with low elongation, small plasma current and therefore comparatively localised soft X-ray emissivity. We varied the relative position of the detector array and the aperture, being the most critical parameter, and compared simulated emissivity distributions which were based on the equilibrium reconstruction as described above (section 3.4), to our experimental results. The positions used to calculate the geometry of the lines of sight and the angular factors were adapted until an agreement between theory and experiment was achieved consistently for similar discharges at different vertical positions of the vessel. Figure 9 shows an example of a reconstruction before and after the consistency checks. As a result, only the relative poloidal position of the detector and the apertures had to be modified by about or less than $1/10mm$. The mechanical precision of the camera / detector head assembly is better than 1/10 of a mm. The difficulties faced were probably due to the tolerances of the detector assembly itself, $i.e.$ the Silicon chips were fixed to the housings with limited precision.

Another problem caused by the compact design is the influence of the spectral and angular dependence of the detector efficiency, which has been

discussed in detail in [24]. In short, the highly doped zone on the surface of the diode which is needed to obtain a depletion zone near the surface forms an absorptive layer for low energy photons. The effective thickness of this dead layer depends on the angle of incidence, which decreases the efficiency with increasing angle of incidence for the low energy part of the spectrum. The increase of the effective thicknesses of the depletion zone and the absorber with the angle of incidence on the other hand increases the efficiency for high energy photons. This causes a complicated angular and spectral dependence of the detector efficiency.

To overcome this problem, we use the dynamical calibration as sketched in section 3.4. Figure 10 shows an example of reconstructions with and without calibration using the spectrum averaged efficiency. Simulations as well as experimental results indicate that for temperatures characteristic of ohmic plasmas $(T_e < 1keV)$ it is sufficient to use identical spectral distributions for all lines of sight. The use of individual spectral distributions led to no significant differences with respect to this simpler approximation. This may change if the plasma temperature exceeds the $1keV$–range.

Reconstructions of soft X-ray emissivity distributions from discharges with high emissivity such as during H-mode exhibited distortions with re-

spect to the magnetic reconstruction which were very similar to those observed for low emissivity discharges without calibration. This may hint to the fact that our estimates for the dead layers of the photodiodes are not sufficiently pessimistic: *Krumrey* and *Tegeler* [29, 30] showed that Silicon diodes with n—substrate and a highly doped p^+ region on the front side like the ones we are actually using, may in fact not be adequate for the use with soft X-rays. They observed an increase of the dead layer thickness with increasing irradiation, due to a charging of the diode surface and a resulting increase of the surface recombination rate.

To avoid any problem related to calibration in the near future, we are going to use a new type of diodes especially designed for the detection of soft X-rays with near-theoretical (ideal) efficiency for photon energies up to $6keV$ and angles of incidence up to $60°$ [31, 32].

4.2.2 Magnetic Pickup

Another source of error which is not directly linked to geometry is magnetic pickup from the poloidal field coils. There are essentially two different types: one with a constant amplitude and a frequency of about $1350Hz$ which slightly decreases during a discharge. This is due to the rectifiers of the power

supplies for the magnetic field coils. (The observed frequency equals five times the frequency of our flywheel generator), a second one is due to rapid changes of the coil currents. We have calculated coefficients for every detector which allow to subtract the pickup from the diode signals given the derivatives of the poloidal magnetic field dB_r/dt and dB_z/dt at all detector positions. The coefficients were obtained using so called stray shots — discharges without plasma, where all coils are switched on and off successively. Usually, a subtraction of this kind of pickup is not necessary during a flat top phase of the plasma discharge and is thus only applied if needed. The amplitude of the $1350Hz$ pickup is usually negligible, it is only significant at very low signal levels.

4.2.3   Treatment

The errors induced by incorrect geometry and calibration have largely been eliminated by the procedures described above. To take residual systematic errors into account, we choose sufficiently large values for the statistical errors. This is done in a systematic manner, *i.e.* the standard deviations σ_ℓ of the chord brightness f_ℓ used in all inversion algorithms are weighted with the inverse of the angular factor $A\Omega$, since in our case small angular factors

imply large angles of incidence and thus larger uncertainties on the angular factors, misestimated efficiency and amplified pickup noise. Due to the weighted errors, the effective number of independent data as defined in [17] is reduced by less than 20%.

5 A Case Study: Plasma Triangularity and Sawtooth Oscillations

This section is an illustration of the possibilities offered by the TCV tomography system and not meant to be an extensive physics study.

In TCV a strong correlation is observed between sawtooth oscillations and plasma shape. A series of discharges has been obtained with a scan of the triangularity during the flat top, other plasma parameters such as elongation or safety factor q_{95} were fixed. One of these discharges is shot # 9650. The plasma current during a triangularity rampdown from $\delta_{95} = +0.4$ to $\delta_{95} = -0.3$ decreased from $380kA$ to $305kA$. The elongation varied only slightly between $\kappa_{95} = 1.51$ and $\kappa_{95} = 1.45$. The safety factor q_{95} varied at the same time from 2.8 at the beginning to 2.6 at the end of the δ-rampdown.

The central plasma electron temperature was $\approx 650eV$ at a line density of $5.6 \cdot 10^{-19} m^{-2}$.

We investigated several time intervals during this rampdown phase with 100 samples each which had been acquired at a frequency of $10kHz$, such that two complete sawteeth were contained in each interval. The line integrated data were calibrated and then inverted using the Minimum Fisher Regularisation method with a spatial resolution of 15 radial times 23 vertical pixels. The results of the tomographic inversion were analysed with the help of the SVD as described in sections 3.5 and 4.1.2. As an example, figure 11 displays the SVD analysis of the time interval between $t = 0.455 - 0.4669s$ at an average triangularity of $\delta_{95} \approx 0$, the meaning of the different parts of the figure being the same as described in section 4.1.2.

The first *topos / chronos* pair represents the equilibrium or a kind of average profile while pairs #3 and #4 represent a mode. The second *topos* characterises a profile change. A look on the corresponding *chronos* reveals that this part represents sawtoothing. (broken lines indicate negative values of the displayed distribution, continuous lines represent positive values). The shape of *topos* #2 suggests an interpretation of its zeroes as the sawtooth inversion radius. A comparison with the location of the minimum and the

maximum of *topos* # 3 and # 4 shows that this interpretation is consistent with the usual identification of the sawtooth inversion radius with the $q = 1$-surface, since the mode observed is of the type $m = 1$, m being the poloidal harmonic of the mode. As has been shown in section 4.1.2, our new method yields reliable results concerning the location of the mode structures, so we conclude that a determination of the inversion radius should be possible within the limits of the spatial resolution, which is in this case $\approx \pm 2cm$.

From the series of reconstructions we can thus deduce two figures of interest for the investigation of sawteeth: The first is the relative signal energy connected to sawtoothing (compare section 3.5), the second is the inversion radius or, more precisely, two numbers characterising the inversion radius, a horizontal radius a and a vertical radius b, thus taking the elongation of the plasma into account. In figure 12, the relative signal energy p_k of the sawtooth is displayed as a function of the triangularity δ_{95}. We observe that the sawteeth have almost vanished for negative values of δ. The inversion radius as deduced from the relevant topos is shown in figure 13 for the different time intervals investigated. In figure 14 the geometrical mean of the horizontal and the vertical radius, *i.e.* $r_{inv} = \sqrt{(a \cdot b)}$, is displayed as a function of the triangularity. Within the limits of error, the inversion radius seems

52

to remain constant for $\delta_{95} > -0.1$. Only for large negative values a slight reduction may be inferred.

To be able to interpret these results, more analyses will have to be carried out where all important plasma parameters should be considered. Here, we showed that soft X-ray tomography in combination with Singular Value Decomposition can add useful figures to a sawtooth database.

6 Summary

We have shown that the 200–channel soft X-ray tomography diagnostic of the TCV tokamak works satisfactory. The spatial resolution we can achieve using square pixels with constant emissivity as a base for our different reconstruction algorithms is approximately $3.8 \times 3.8 cm^2$.

The quality of the inversions is improved by the application of a spectrum and angular dependent calibration, a subtraction of magnetic pickup and an appropriate quantification of the experimental errors.

We have compared the performance of four different algorithms using square pixels, notably a Bayesian Maximum Entropy Method, first and second order Linear Regularisation and a new method we called Minimum Fisher

Regularisation. A comparison of the reconstruction of simulated MHD activity showed that our new method is for the moment the best choice: First and second order linear regularisation smooth the reconstructed emissivity distribution too strongly which leads to loss of information concerning the localisation of island structures. Maximum Entropy on the other hand yields too noisy reconstructions, thereby consuming six times more CPU time than our new method. The reason for the comparatively poor performance of Max-Ent for soft X-ray tomography is probably caused by the major assumption that the emissivity values of the different pixels are statistically independent random variables, which is simply not true[5]. The good performance of our Minimum Fisher Regularisation algorithm is probably also due to the fact that the assumptions, the *prior* information which is put into the algorithm, is well adapted to the problem: Smoothing is less strong where the soft X-rays emissivity is high, *i.e.* at the center of the plasma where we expect to find the dynamical structures due to MHD activity. Thus less information on

---

[5]We have tested the *same* algorithm for spectrum deconvolution. For this kind of application where photon counting is involved, the assumptions of MaxEnt are adequate. As a consequence, the Bayesian MaxEnt algorithm as described above and by *von der Linden* [17] yielded by far the best results!

54

the high emissivity regions is lost compared to linear regularisation.

A very useful tool for the analysis of tomographic inversion is the biorthogonal decomposition or SVD, as we have demonstrated with simulated data as well as with a case study where the relative importance of sawtoothing and the inversion radius could be determined from inverted experimental data where the triangularity of the plasma had been varied.

Further improvements will probably be achieved by the installation of new soft X-ray detectors.

Acknowledgements

First of all, we wish to thank the entire TCV technical team for their continuous effort during construction, installation and operation of the soft X–ray tomography system.

We would also like to thank *Dr. W. von der Linden* for his help with the Bayesian MaxEnt algorithm and for pointing out the possibility of using the Minimum Fisher information.

It is a pleasure to acknowledge the contributions of *Dr. Yves Peysson*, Cadarache, and *S. Sagbo*, EPFL, as well as a hint by *Dr. B. Möbus*, PTB Berlin.

References

[1] *MORET, J M et al.:*

Plasma Phys. Control. Fusion **37** 11A pA215 (1995)

[2] *HOFMANN, F and TONETTI, G:*

Nuclear Fusion **28** p1871 (1988)

[3] *THORNE A P:*

"Spectrophysics", second edition, Chapman and Hall, London (1988)

[4] *PRESS, W H, TEUKOLSKY, S A, VETTERLING, W T, and FLAN-*
NERY, B P:

"Numerical Recipes in FORTRAN: The Art of Scientific Computing",

Second Edition, Cambridge University Press (1992)

[5] *GRANETZ, R S and SMEULDERS, P:*

Nuclear Fusion **28** 3 (1989)

[6] *MLYNÁŘ, J:*

Czechoslovak Journal of Physics, 45 10 p799 (1995)

[7] *WILLIAMSON, J H and EVANS, D E:*

IEEE Transactions on Plasma Science **PS-10** 2 p82 (1982)

[8] *FUCHS, G, MIURA, Y and MORI, M:*

Plasma Phys. Control. Fusion **36** p307 (1994)

[9] *DECOSTE, R:*

Rev. Sci. Instrum. **56** 5 p806 (1985)

[10] *PARKER, R L in:*

Annual Review of Earth and Planetary Science **5** p35 (1977)

[11] *PENG, C, RODI, W and TOKSÖZ, M N:*

Inverse Problems **9** p339 (1993)

[12] *STERITI, R and FIDDY, M A in*

SPIE Vol. **1767** Inverse Problems in Scattering and Imaging, p112 (1992)

[13] *REINMUTH, J:*

Diplomarbeit, TU München / IPP Garching (1994)

[14] *VON DER LINDEN, W:*

private communication (1995)

[15] *BUCK, B and MACKAULAY, V A (eds.):*

"Maximum Entropy in Action" Oxford University Press (1991)

[16] *DJAFARI, A M:*

Traitement du Signal 11 2 p87 (1994)

[17] *VON DER LINDEN, W:*

Appl. Phys. A **60** p155 (1994)

[18] *VON DER LINDEN, W:*

private communication (1994)

[19] *GULL, S:* in

"Maximum Entropy and Bayesian Methods", Ed. J.Skilling, Kluwer

Academ. Publ. p53 (1989)

[20] *SKILLING, J:* in

"Maximum Entropy and Bayesian Methods", Ed. J. Skilling, Kluwer

Academ. Publ. p45 (1989)

[21] *BEHN R, FRANKE S, PIETRZYK Z A, ANTON M, NIESWAND C,*

WEISEN H and MARLETAZ B:

Proceedings of the 7th International Symposium on Laser–Aided Plasma

Diagnostics, Fukuoka, Japan 5-8 December 1995 p392 (1995)

[22] *MIYAMOTO, K:*

"Plasma Physics for Nuclear Fusion",MIT Press (1976)

[23] *VEIGELE, W M J:*

Atomic Data **5**, No. 1 (1973)

[24] *ANTON, M, DUTCH, M J and WEISEN, H:*

Rev. Sci. Instrum. **66** 7 p3762 (July 1995)

[25] *ANTON, M, DUTCH, M J and WEISEN, H:*

22nd EPS conference on CONTROLLED FUSION AND PLASMA

PHYSICS, Bournemouth UK, 3rd-7th July 1995, Contributed Papers,

Volume19C Part II p389 (1995)

[26] *DUDOK DE WIT, T, PECQUET, A-L, VALLET, J C*

and LIMA, R:

Phys. Plasmas **1** (10) p3288 (1994)

[27] *ANDERSON E et al.:*

LAPACK user's guide, Society for Industrial and Applied Mathematics,

Philadelphia (1992)

[28] *MORET J M:*

private communication

[29] *KRUMREY, M and TEGELER, E:*

Rev. Sci. Instrum. **63** (1) (1992)

[30] *TEGELER, E and KRUMREY, M:*

Nucl. Inst. Meth. **A 282** p701 (1989)

[31] *KORDE, R and CANFIELD, L R:*

SPIE **1140** p126 (1989)

[32] *WENZEL, K W, LI, C K, PAPPAS, D A and KORDE, R:*

IEEE **5** p274

Reconstruction of simulated data TCV discharge # 8100 at 0.4s				
Method	flops	time [s]	$\sigma_{0\%}$	$\sigma_{2.5\%}$
Bayesian MaxEnt	700M	124	0.936	1.50
Minimum Fisher	22M	17	0.175	0.382
1st order Regul.	1.4M	9	0.503	0.518
2nd order Regul.	1.8M	6.5	0.303	0.349

Table 1: Performance of different algorithms using square pixels. The number of floating point operations, the CPU time on a DECα station with MATLAB 4.1 and the reconstruction error (see text) for two different noise levels are compared.

SVD analysis of simulated data and reconstructions TCV discharge # 9243 at 0.746-0.75s						
			p_1	p_2	p_3	Mode
Simulation			9.98E-1	1.08E-3	7.63E-4	+
Reconstruction	noise	σ				
Bayesian MaxEnt	0%	0.080	9.98E-1	1.03E-3	8.72E-4	+
	1.0%	0.120	9.88E-1	1.39E-3	1.24E-3	0
	2.5%	0.230	9.40E-1	6.80E-3	5.00E-3	-
Minimum Fisher	0%	0.045	9.99E-1	7.47E-4	6.03E-4	+
	1.0%	0.050	9.98E-1	8.29E-4	6.11E-4	+
	2.5%	0.078	9.94E-1	9.62E-4	8.52E-4	+
1st order Regul.	0%	0.084	9.99E-1	3.19E-4	2.67E-4	+
	1.0%	0.080	9.99E-1	3.59E-4	2.84E-4	+
	2.5%	0.079	9.94E-1	7.17E-4	5.73E-4	0
2nd order Regul.	0%	0.054	9.99E-1	2.98E-4	2.39E-4	+
	1.0%	0.052	9.99E-1	3.42E-4	2.56E-4	+
	2.5%	0.056	9.96E-1	5.93E-4	5.59E-4	-

Table 2: Performance of different tomography algorithms concerning the reconstruction of MHD island structures from simulated data is compared using the SVD analysis for different levels of noise. For the definitions of the reconstruction error σ and the relative signal energy p_k see text. The last column indicates wether the mode was well recovered (+), hardly recognisable (0) or lost (-). Compare figure 8

Figure 1: A view of the experimental setup of TCV's X-ray tomography system. All 9 cameras are placed around one poloidal cross section of the torus. The vacuum vessel with the Carbon tiles as well as the poloidal field coils and the central solenoid are shown.

Figure 2: A cross section of one of the pin-hole cameras: 1: CENTRONIC LD20-5T photodiode array with premplifier circuit; 2: $47\mu m$ thick curved Be foil; 3: poloidally limiting W-aperture, the detail shows the trapezoidal shape of its edges; 4: front plate / toroidally limiting aperture; 5: outer spacing can be evacuated to improve thermal insulation; 6: water–cooled spacing between inner walls; 7: vacuum feedthrough for the electrical connections; 8: to primary vacuum pumps; Further explanations see text.

Figure 3: Artificial data f_l for TCV discharge #8100 at 0.4s. open circles: data with 2.5% noise added (see text). Continuous curve: fit provided by the Minimum Fisher Regularisation algorithm

Figure 4: Reconstruction of the simulated emissivity of TCV discharge # 8100 at 0.4s. The simulated emissivity distribution is shown as well (top left) as the different reconstructions. The first row shows reconstructions from artificial data without noise, the second row was obtained from artificial data with 2.5% of noise added (see text)

Figure 5: The eight first simulated emissivity distribution for TCV discharge #9243 at 0.7460-0.7467s exhibiting $m = 2$ mode activity.

Figure 6: SVD analysis of the time series of simulated emissivity distributions. Top left: log plot of the singular values S_k vs. the index $k = 1 \ldots 9$. Top right: plot of the maximum emissivity [$a.u.$] as a function of time. Middle, from left to right: spatial "eigen"modes or *topos*, $k = 1 \ldots 3$. Bottom: corresponding temporal modes or *chronos*

Figure 7: Same plot as 6, but this time for the emissivity reconstructed from noisy data (1%) using the Minimum Fisher Regularisation algorithm

Figure 8: Comparison of *topos* # 2 for different algorithms and different noise levels. Top left: simulated distribution. First row: reconstruction from artificial data without noise, from left to right: MaxEnt, Minimum Fisher, 1st and 2nd order Regularisation. Middle row: same as above, but reconstructed from noisy data (1%). Bottom: Reconstructions from noisy data (2.5%). Note that the rotation of the reconstruction with respect to the original is unimportant, the correct phase is recovered from *two* topos/chronos pairs.

Figure 9: A comparison of a reconstruction from experimental data before (left) and after (right) the consistency checks of the geometry. The data were taken from TCV discharge # 8117, a single null divertor configuration, during an H-mode phase at 0.5s.

Figure 10: A comparison of reconstructions from calibrated and uncalibrated data. Left, top: simulated without calibration Left, bottom: simulated data calibrated using the spectrum average efficiency, same spectral distribution for all channels. Right, top: reconstruction from uncalibrated experimental data. Right, bottom: same data, but calibrated using the spectrum average efficiency, identical spectral distributions. Data were taken from TCV discharge # 8100 at 0.4s (compare figure 4)

Figure 11: SVD analysis of a time series of emissivity distributions reconstructed from experimental data, TCV discharge # 9650 at $0.455 - 0.4669s$, $\delta_{95} \approx 0$. Top left: log plot of the singular values S_k vs. the index $k = 1 \ldots 8$. Top right: plot of the maximum emissivity $[a.u.]$ as a function of time. Middle, from left to right: spatial "eigen"modes or *topos*, $k = 1 \ldots 4$. Bottom: corresponding temporal modes or *chronos*

Figure 12: Relative signal energy p_k of the *topos / chronos* pair representing sawtoothing as a function of the plasma triangularity δ_{95}. The results from TCV discharge # 9650 are indicated by the open circles. For comparison, data from a similar discharge (# 9553), indicated by the crosses, are also shown.

Figure 13: Inversion radius deduced from the *topos* are shown together with the last closed flux surface for the different time intervals, hence different triangularities of TCV discharge #9650

Figure 14: Sawtooth inversion radius as a function of the plasma triangularity. The radius is defined as the geometrical mean of the vertical and the horizontal radius: $r_{inv} = \sqrt{(a \cdot b)}$. Errorbars are given by the pixelsize of $\approx 4cm$

68
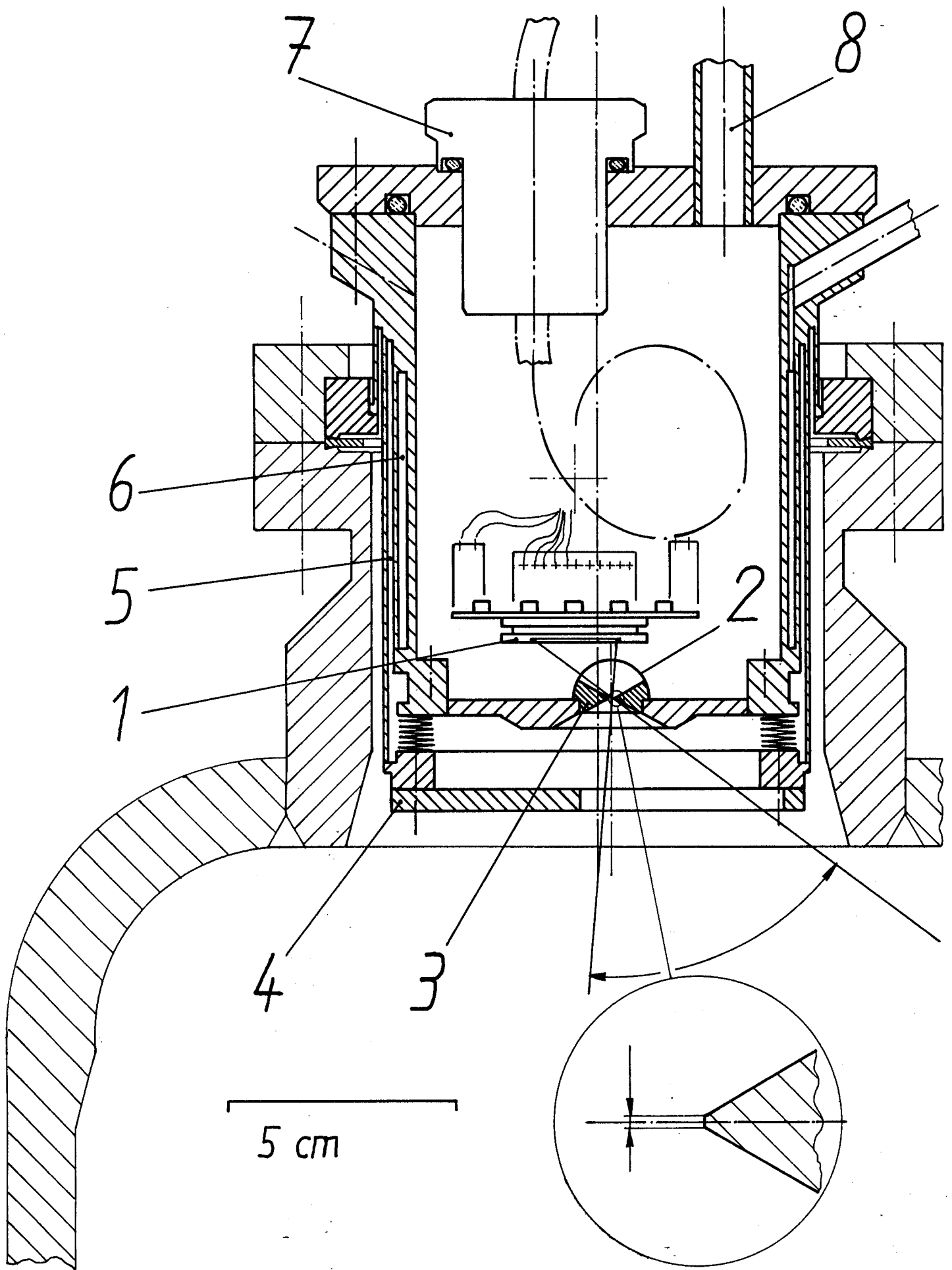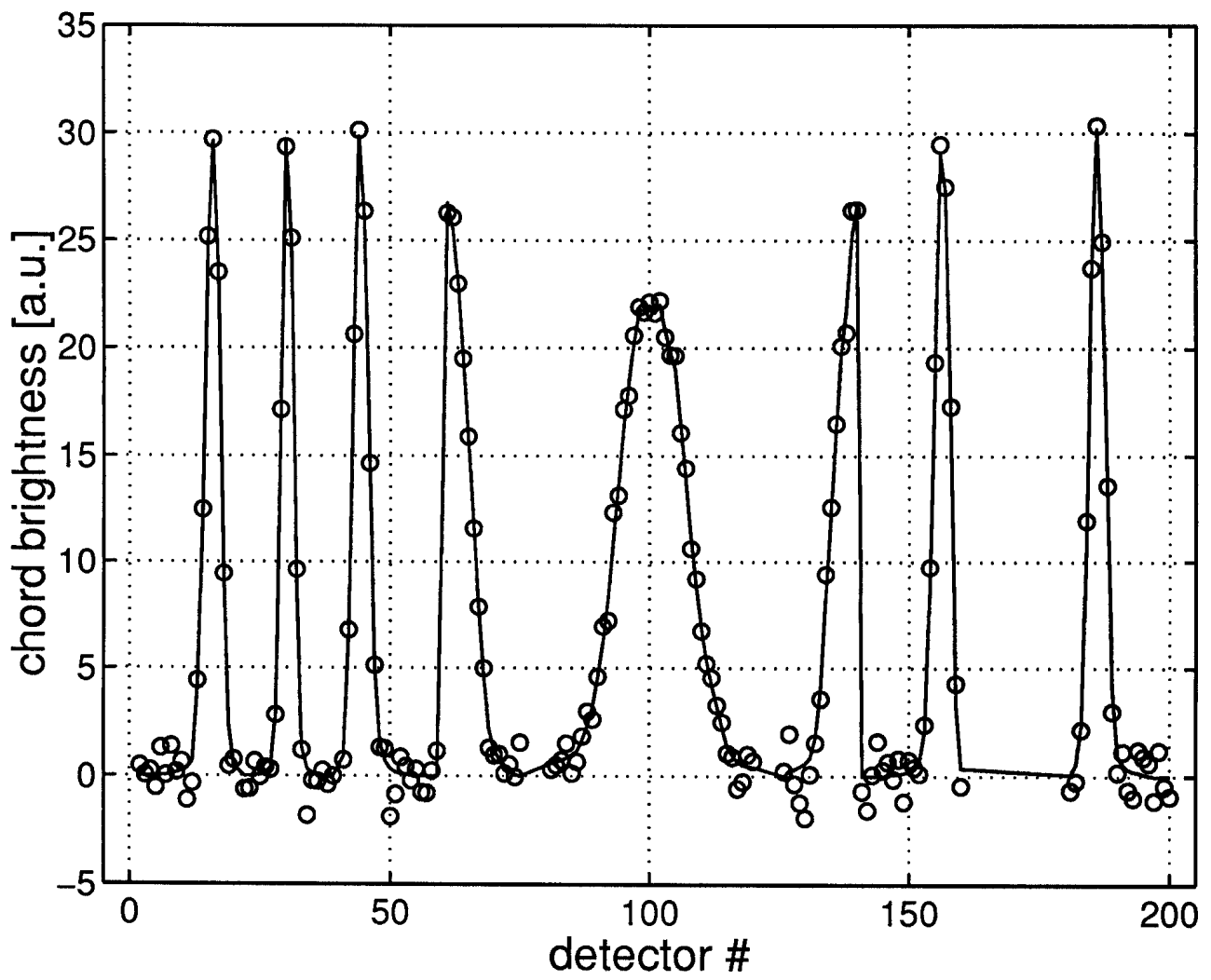
FIGURE 2

FIGURE 3

TCV Soft X-ray Tomography Setup
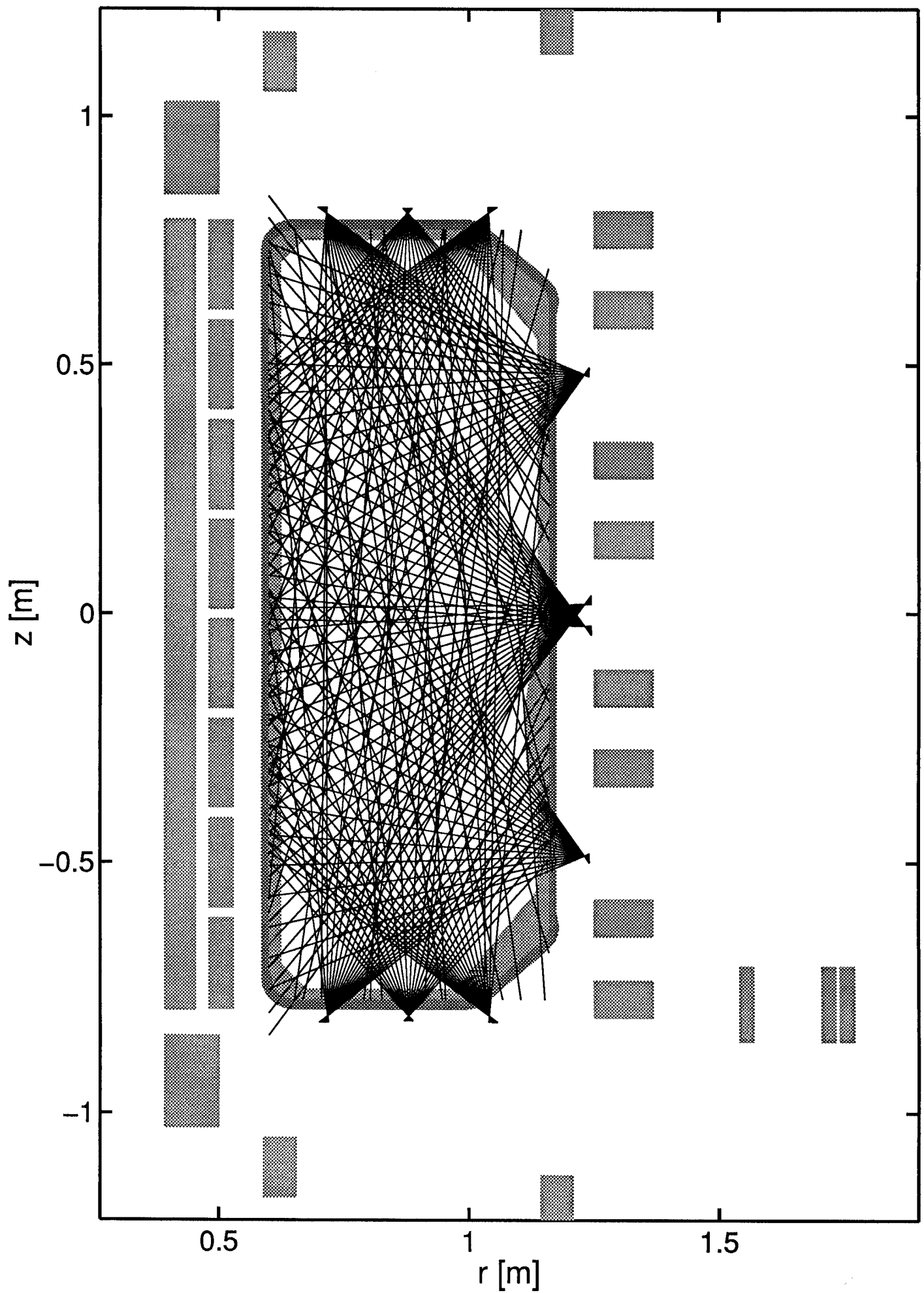
FIGURE 1
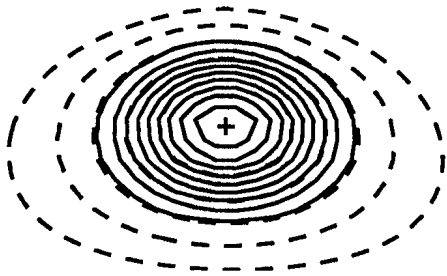
Simulation Bayesian MaxEnt Mininum Fisher 1st order Regul. 2nd order Regul.
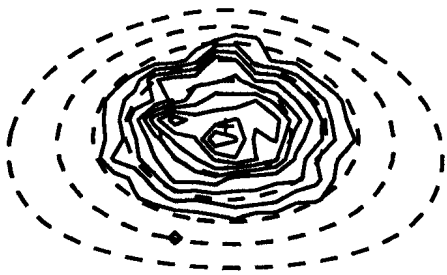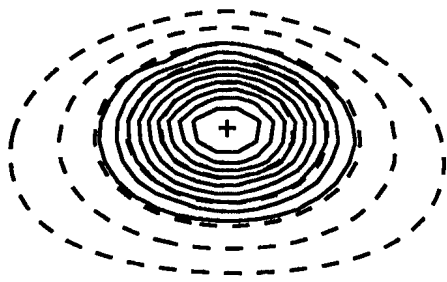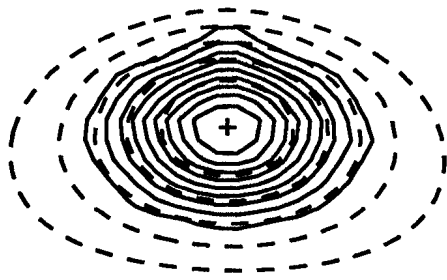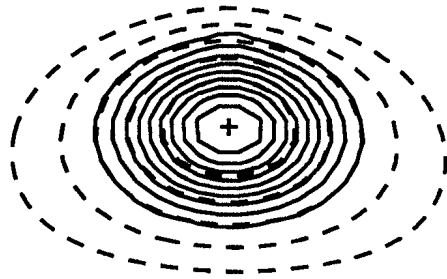
#8100@0.4s
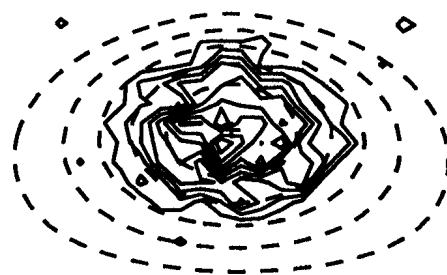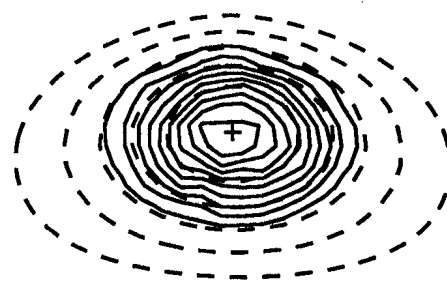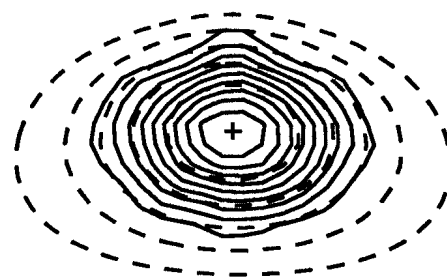
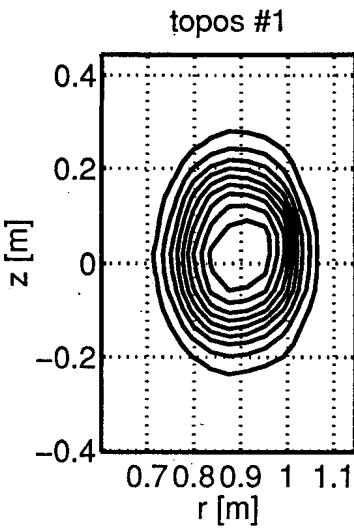Bayesian MaxEnt Mininum Fisher 1st order Regul. 2nd order Regul.
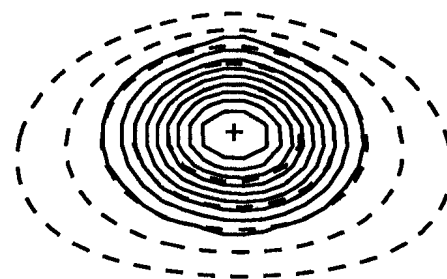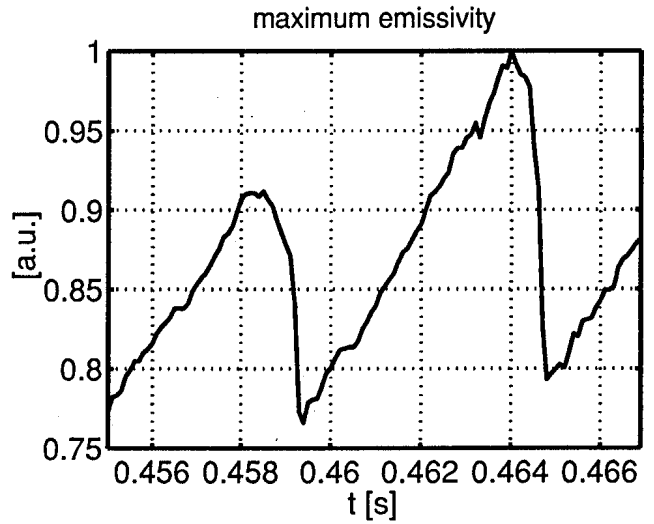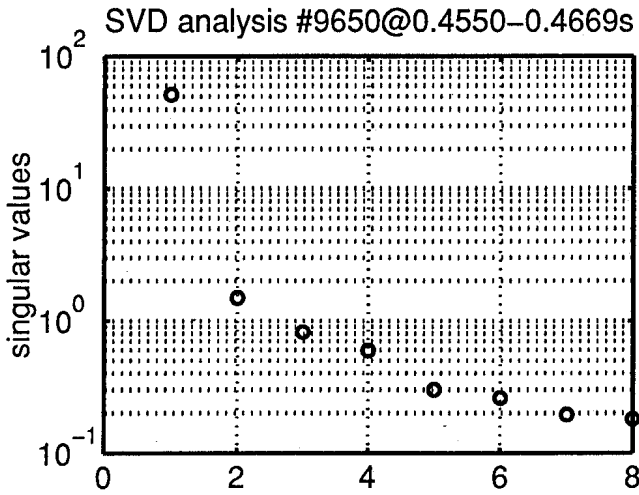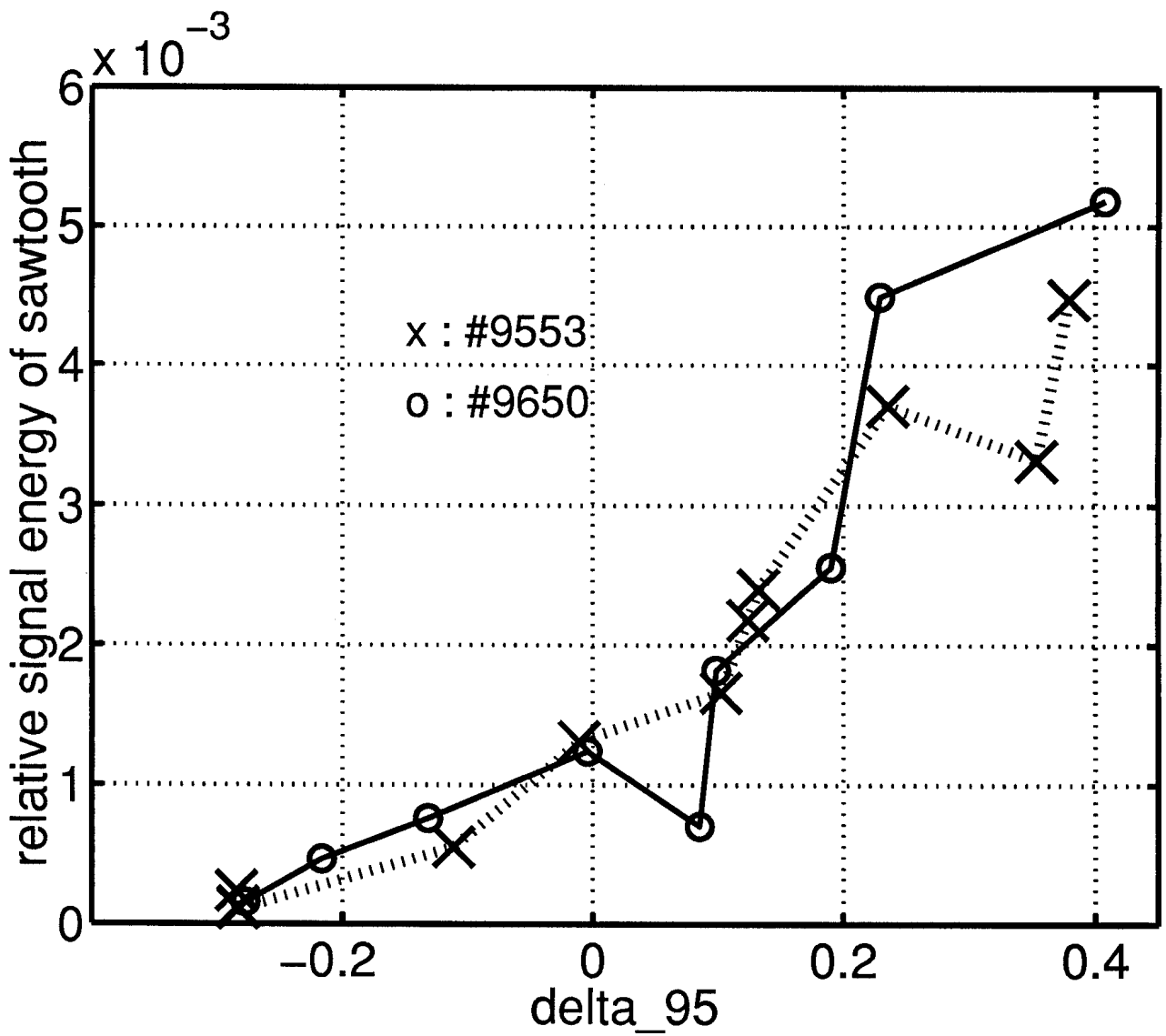
FIGURE 4
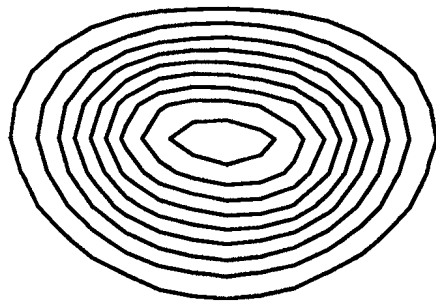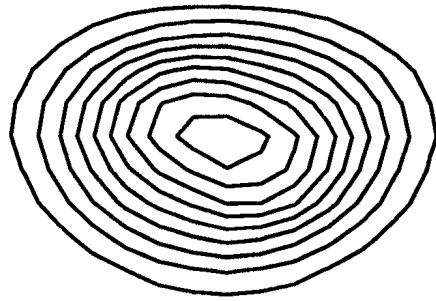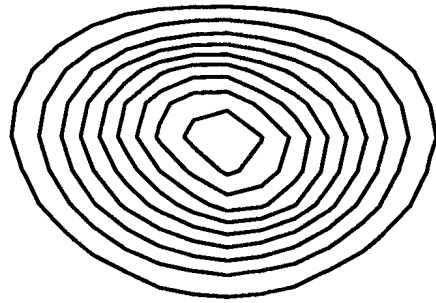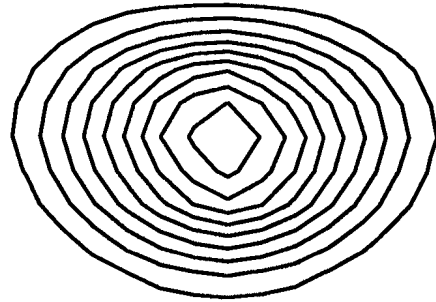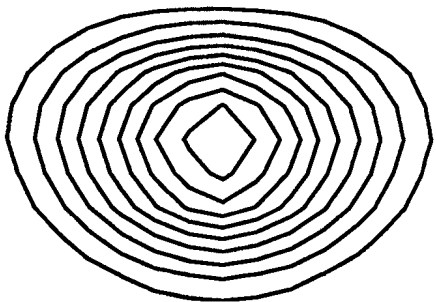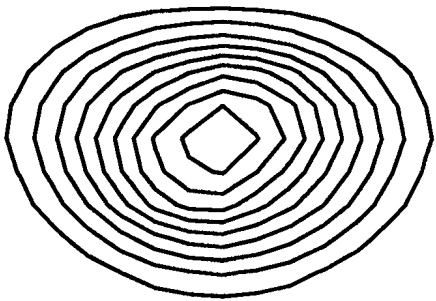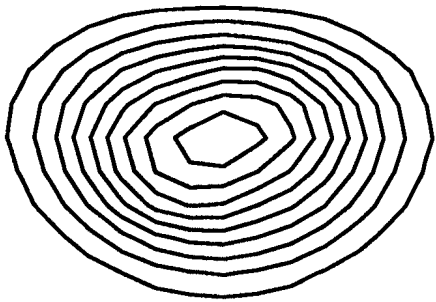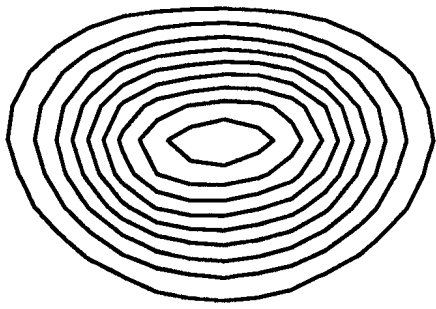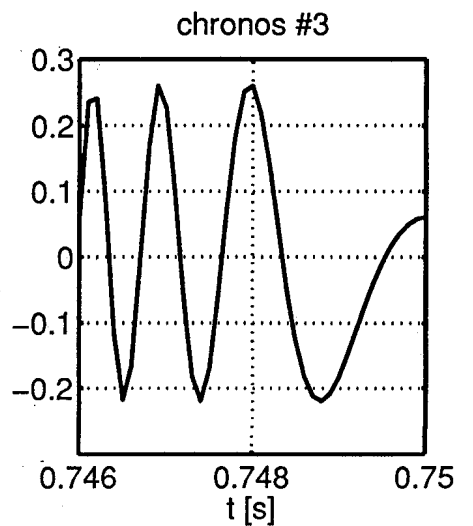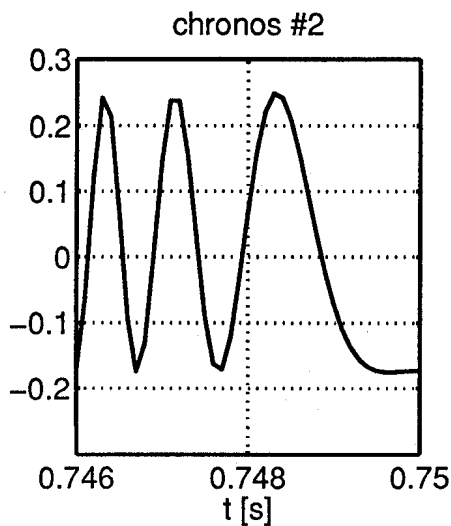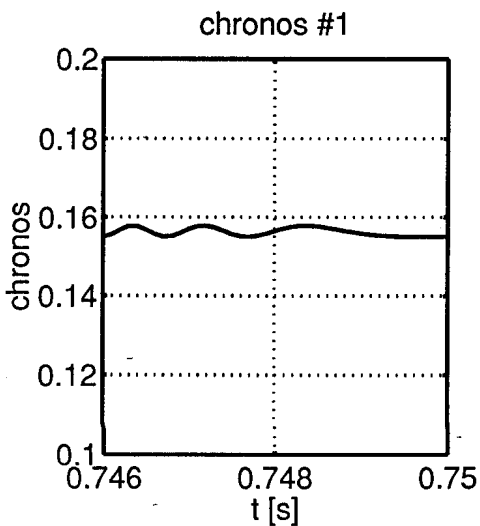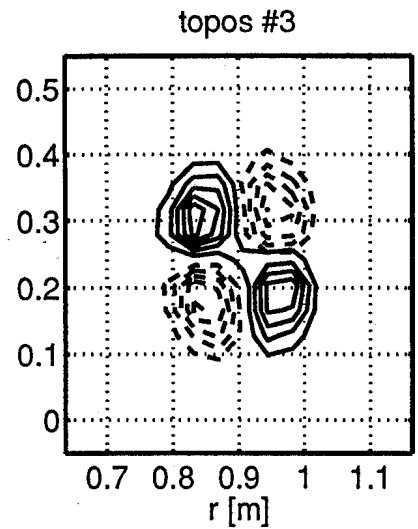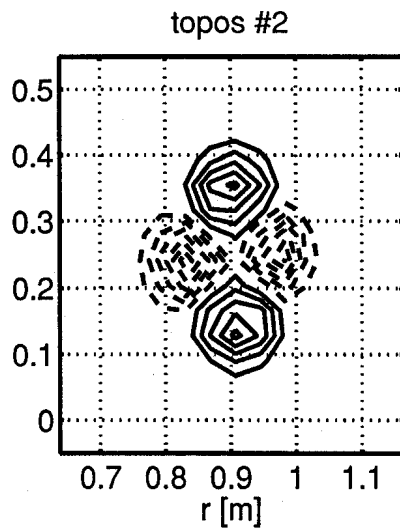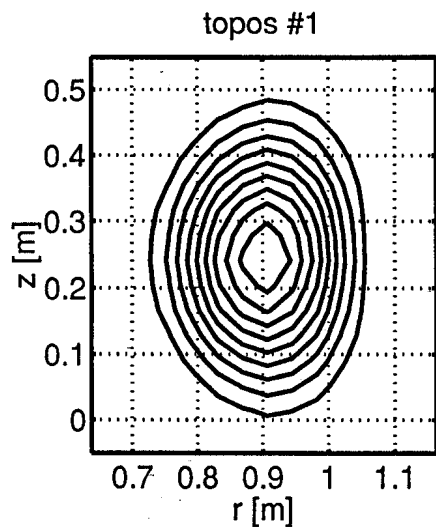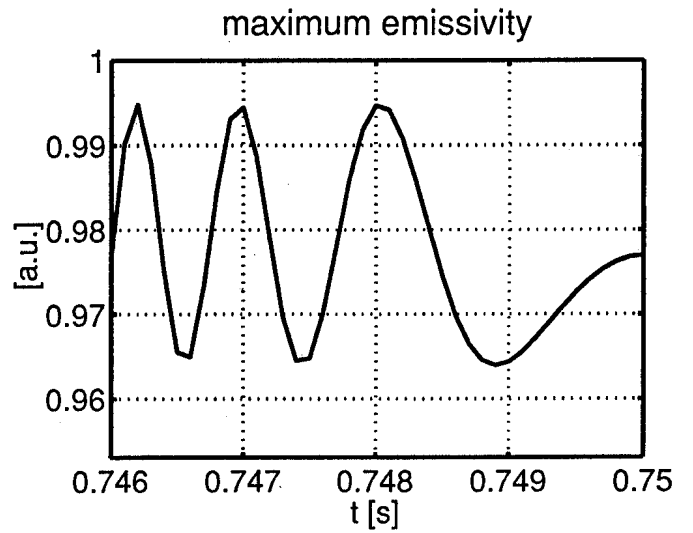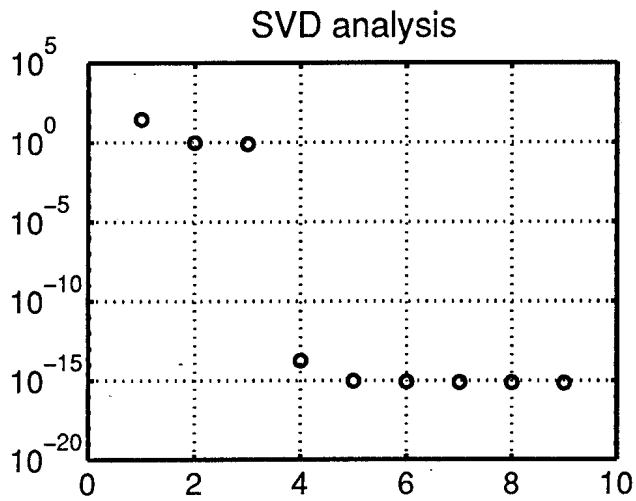
FIGURE 5

FIGURE6

FIGURE 7
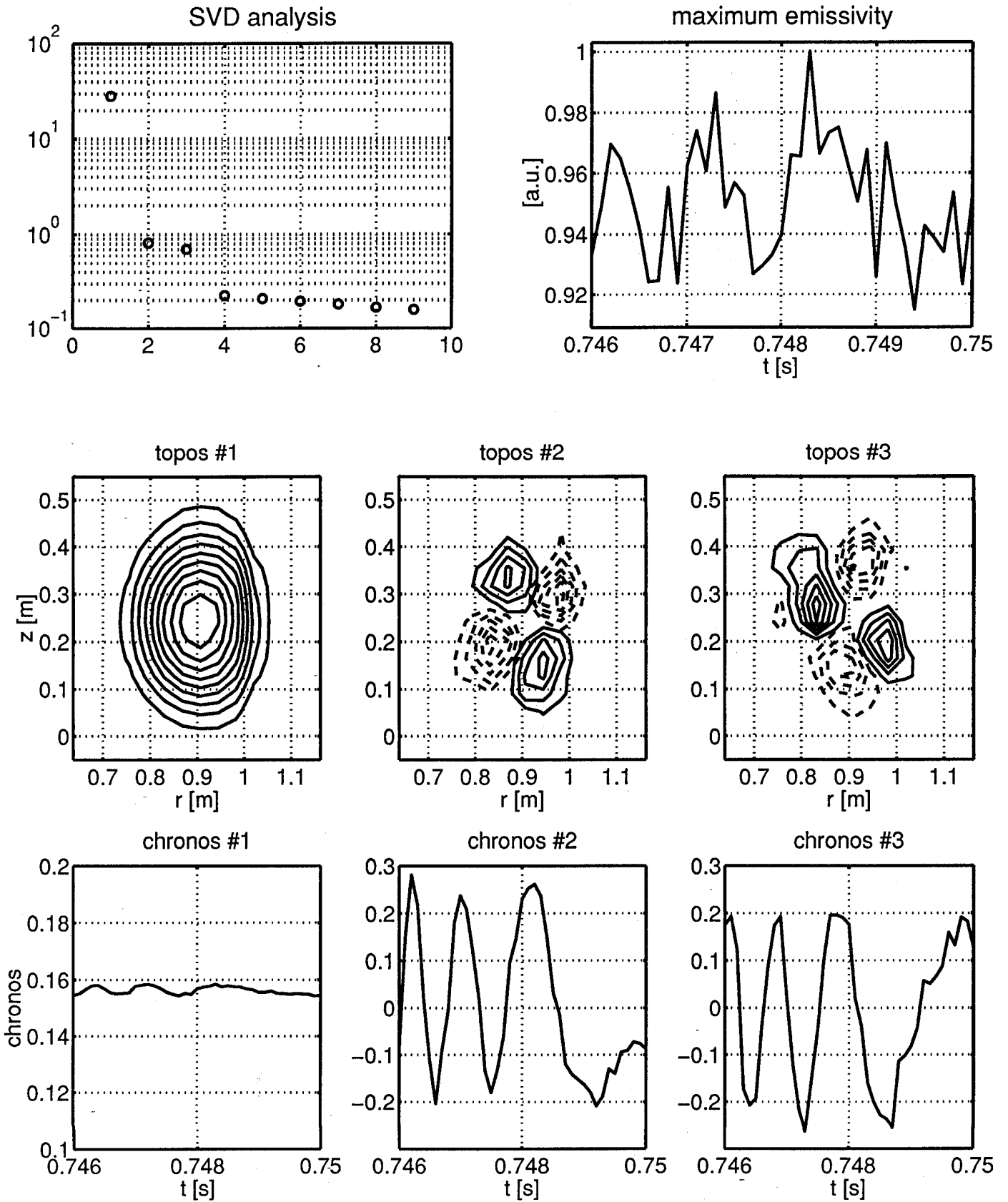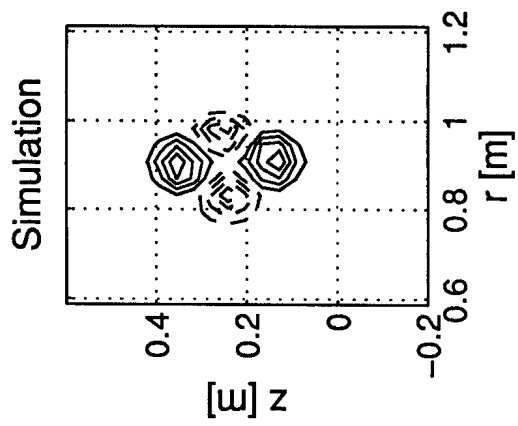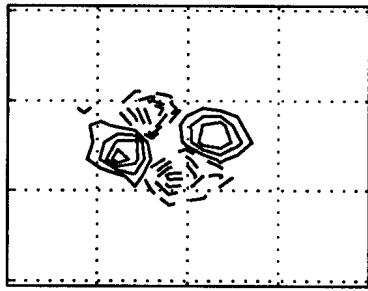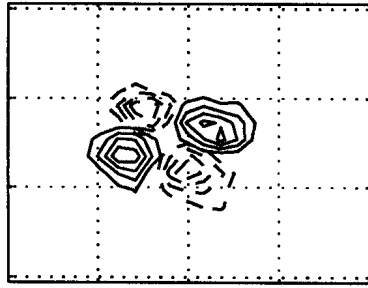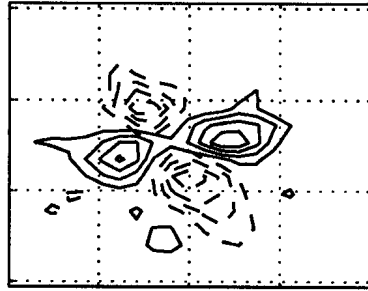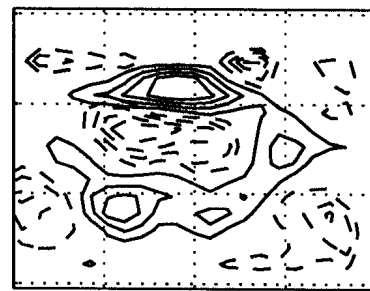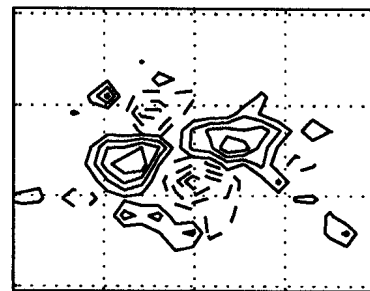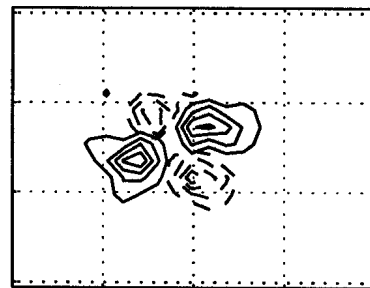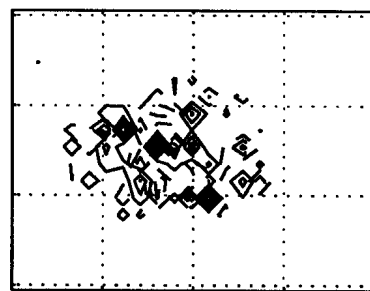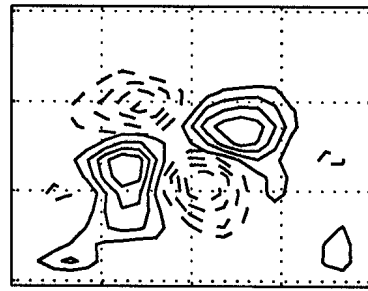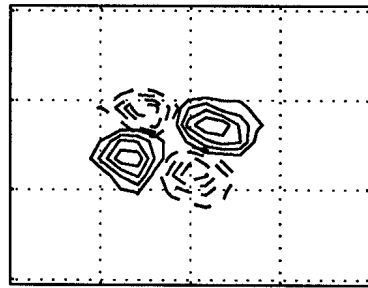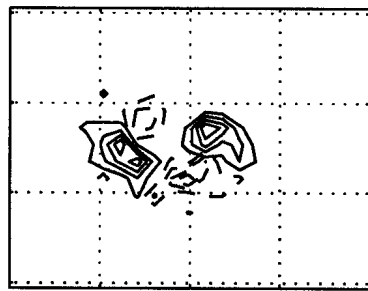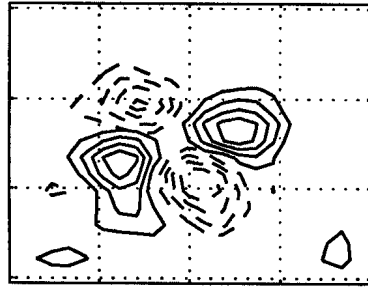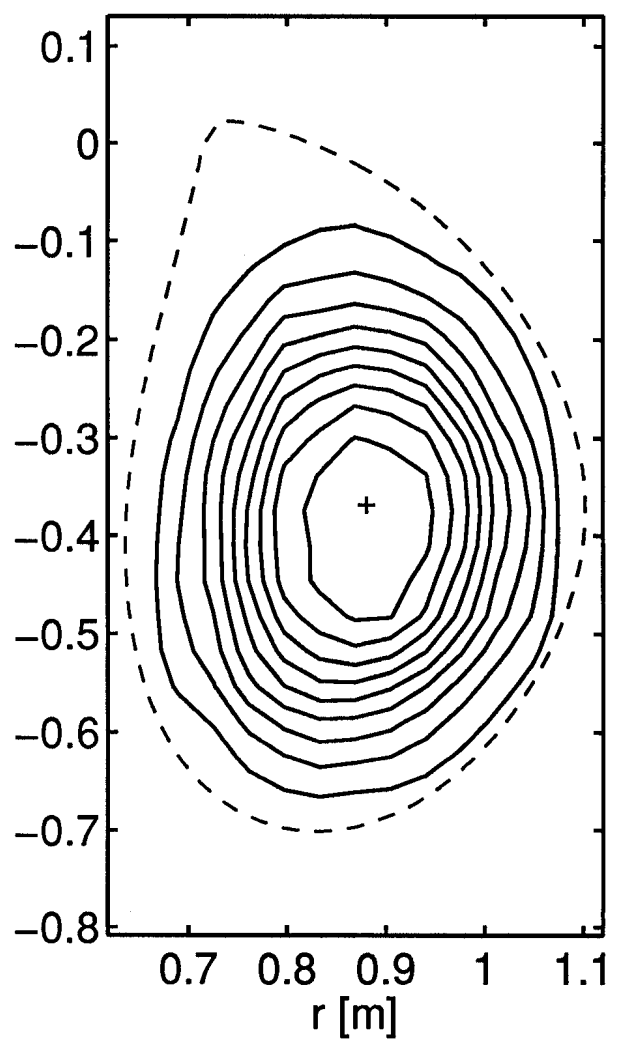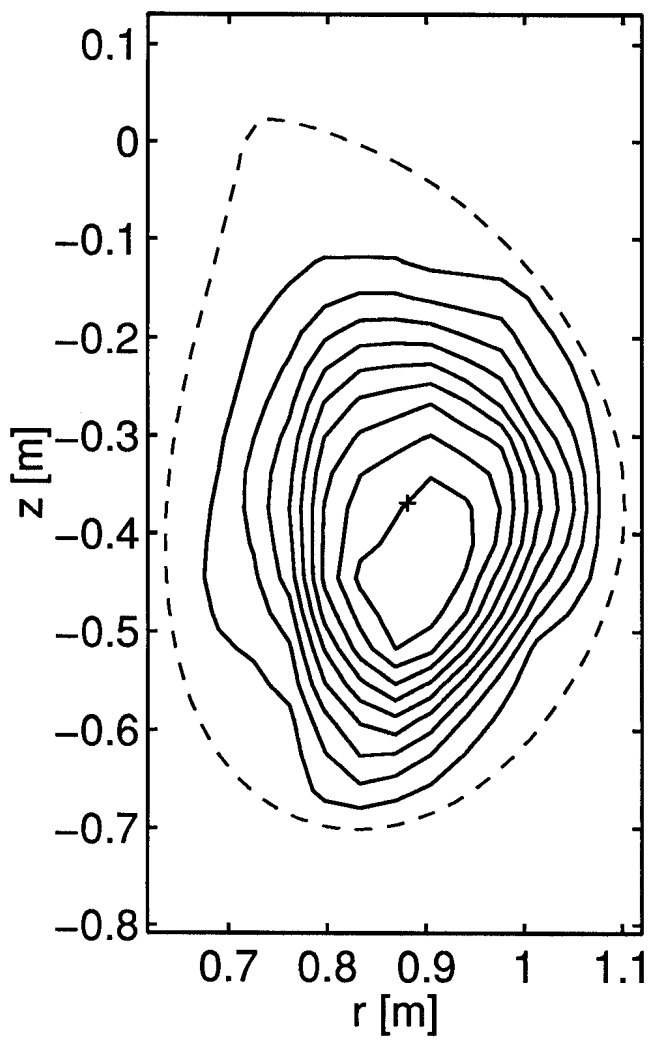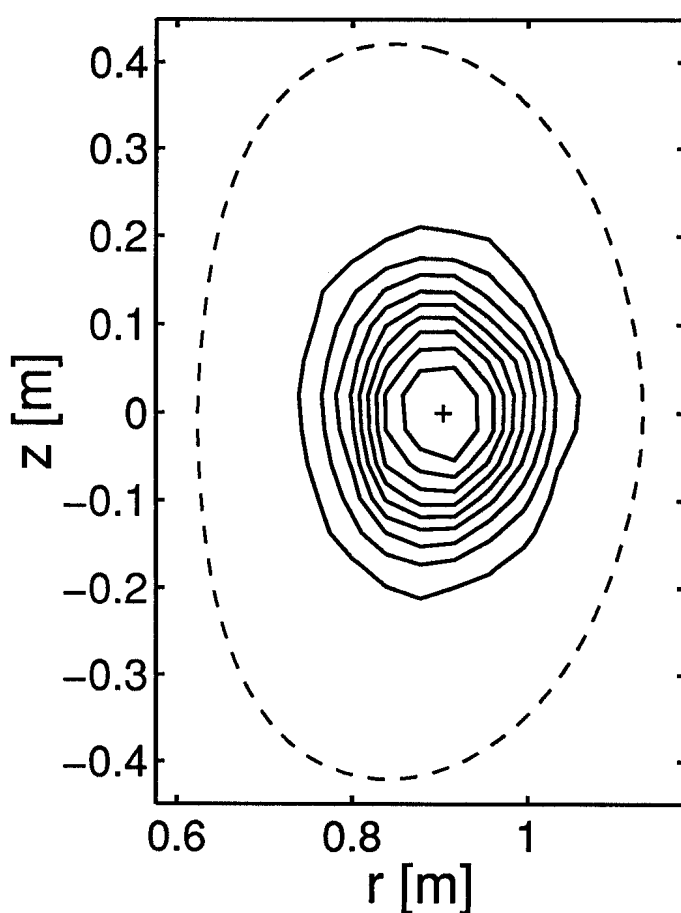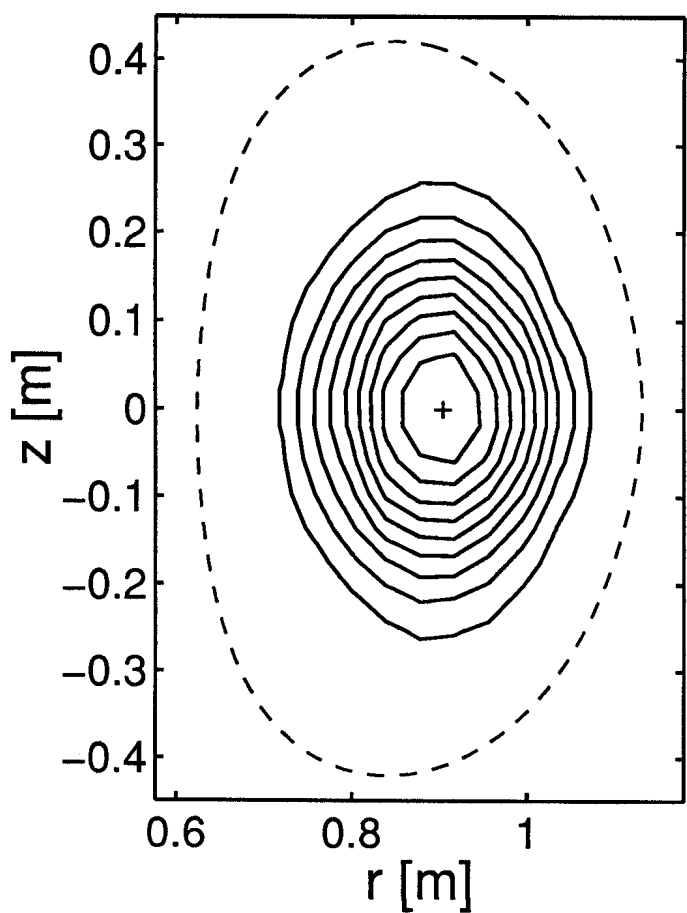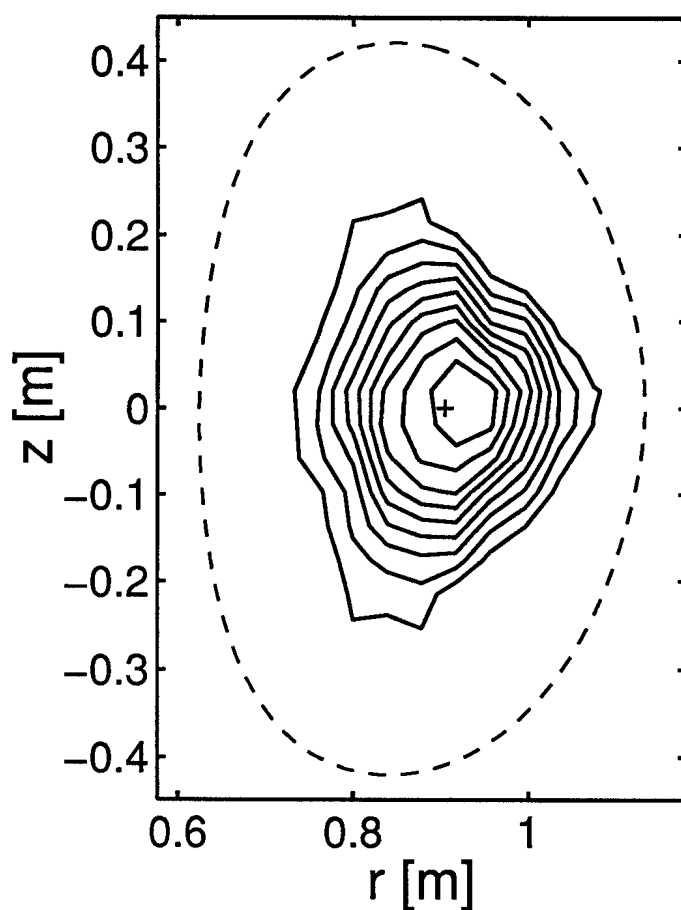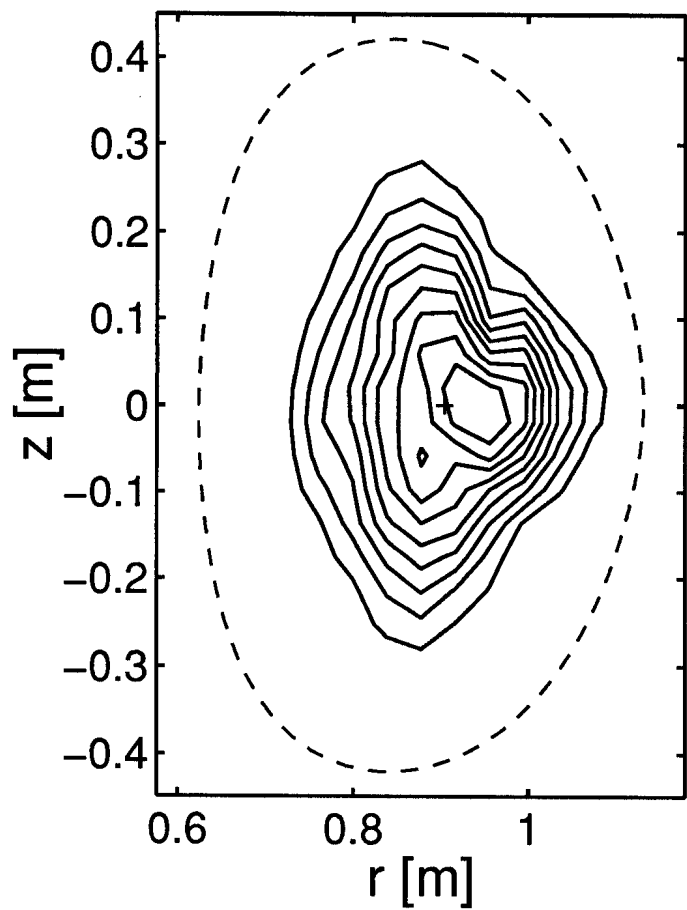
FIGURE 8

FIGURE 9

FIGURE 10

FIGURE 11
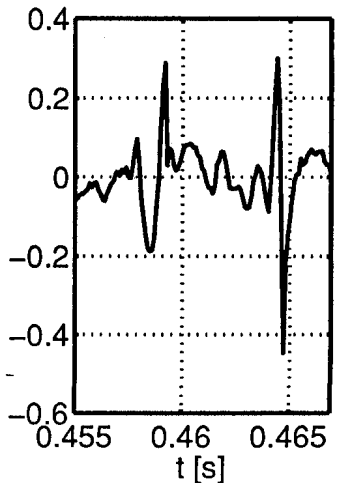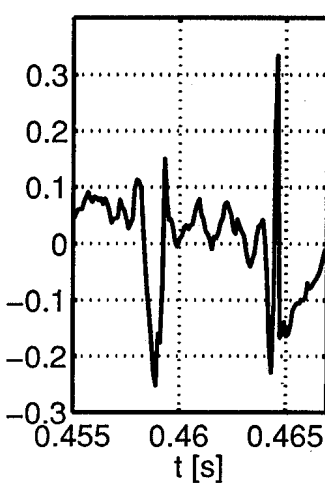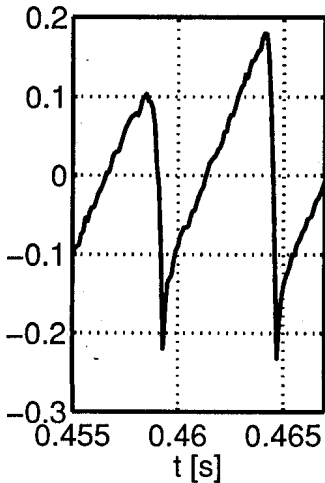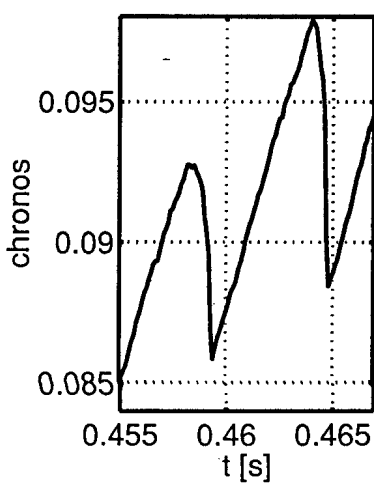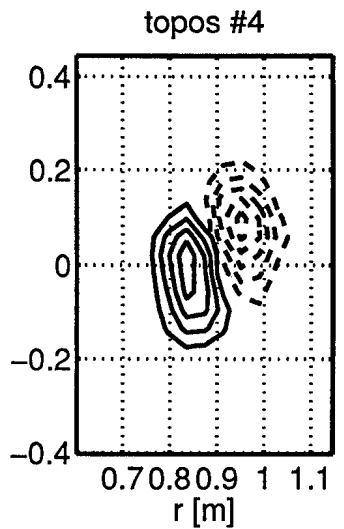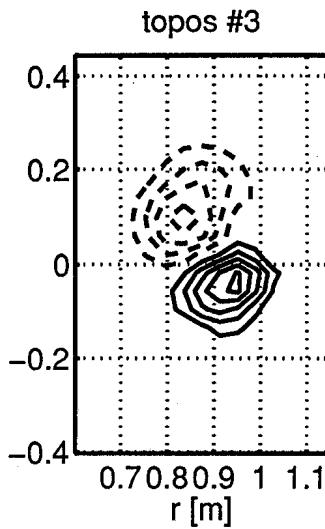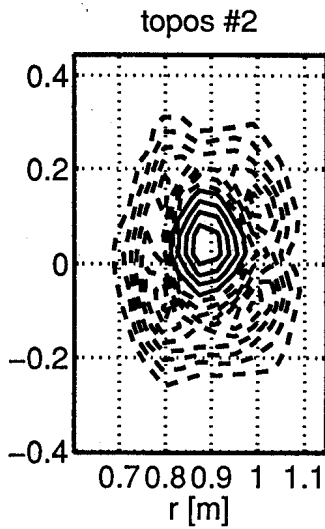
FIGURE 12

FIGURE 13

FIGURE 14