



This is a repository copy of *Identification of Nonlinear Processes in the Magnetospheric Dynamics and Forecasting of Dst Index*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/84379/>

Monograph:

Boaghe, O.M., Balikhin , M.A., Billings, S.A. et al. (1 more author) (1999) Identification of Nonlinear Processes in the Magnetospheric Dynamics and Forecasting of Dst Index. Research Report. ACSE Research Report 761 . Department of Automatic Control and Systems Engineering

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

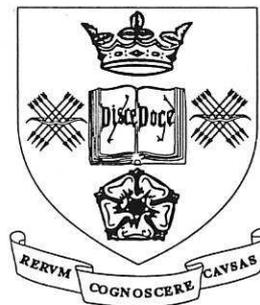
X

Identification of Nonlinear Processes in the Magnetospheric Dynamics and Forecasting of D_{st} Index

O.M. Boaghe, M.A. Balikhin, S.A. Billings, H. Alleyne
Department of Automatic Control and Systems Engineering
University of Sheffield
Mappin Street, Sheffield S1 3JD
United Kingdom

Research Report No. 761

October 1999



University of Sheffield

200453370



Identification of Nonlinear Processes in the Magnetospheric Dynamics and Forecasting of D_{st} Index

O.M. Boaghe, M.A. Balikhin, S.A. Billings, H. Alleyne

Department of Automatic Control and Systems Engineering, University of Sheffield

Abstract. The NARMAX approach is used to analyse simultaneous measurements of the geomagnetic D_{st} index and VB_s , the merging rate of the IMF and the geomagnetic field. The nonlinear discrete relation which describes the dynamics of the D_{st} index driven by VB_s is identified directly from experimental measurements. This identified model is then used to forecast the evolution of the D_{st} index. Analysis of the identified model in the frequency domain provides information about the types of nonlinearities involved in the energy storage process in the magnetosphere. It is also shown how the nonlinear frequency response functions derived from the data can be used to validate theoretical models proposed to describe the evolution of the magnetosphere.

Introduction

Two complementary approaches to the investigation of the dynamics of the magnetosphere under the influence of the solar wind exist in the current literature (see the review [Klimas *et al.*, 1996]). One approach is based on deriving a mathematical description of the dynamics of the magnetosphere from basic physical principles. Once derived the mathematical model can then be used to forecast disturbances of the ground magnetic field caused by the solar wind. Thus it would be possible to forecast magnetic storms and substorms which affect our everyday life causing currents in pipelines, problems in communications and even blackouts in the electrical power supply. However, magnetospheric events such as storms or substorms

represent chains of complex physical processes (e.g. [Fairfield *et al*, 1992]). If all the links in these chains could be completely understood they could be conjugated into a theory of magnetospheric dynamics. However at the present time even a complete quantitative theory of some individual processes, links in this chain, are far from completion. Difficulties in finding a solution to these important problems from basic physical principles stimulated the second approach based on the identification of dynamical analogue models of the magnetosphere, which can in turn be used both for physical interpretation and for forecasting. In this approach the magnetosphere is considered as a gray box nonlinear system with solar wind as the input and indices of geomagnetic activity as the output. Continuous time dynamical models were constructed using this approach, both for magnetospheric storms and substorms by considering simplified physical analogies, for example the dripping faucet model [Baker *et al*, 1990], directly driven model [Goertz *et al*, 1993], Faraday loop model [Klimas *et al*, 1996] for auroral electrojet indices or the Burton model [Burton *et al*, 1975] for the evolution of the D_{st} index. Once the physical structure of the analogue model has been formulated, data from the measured solar wind parameters and geomagnetic indices is used to estimate the unknown parameters in the model. Finally the relevance of the models derived is usually assessed by the performance over a data not used in the estimation.

However consideration of physical analogies is not the only way to construct dynamical models. The dynamics of the system can be identified directly from the input-output data. One approach which has become very popular recently is to use a neural network. The application of neural networks leads to significant progress in forecasting geomagnetic activity in general and the evolution of D_{st} in particular, see for example Wu and Lundsted [1997]. Neural networks are known to be good approximators for predictions, and the multi-layer perceptron network can approximate arbitrarily well any continuous function. However, in application to magnetosphere data the goal of system identification is not only to provide accurate forecasts but also to yield a mathematical model which can be related back to

the underlying physics. This involves determining the mathematical representation of the system from the data and often leads to models which forecast well, which provide good generalisation characteristics, and which can be used to provide insight into the underlying system. Such mathematical models are more difficult to estimate than training a neural network for example, but the reward is a parsimonious mathematical model which can readily be analysed to reveal the characteristic dynamical features of the system under investigation.

If some assumptions about the dynamical structure of nonlinear system are made local linear predictions can be used to identify a model of the system [Klimas *et al*, 1997]. However in the present paper a more general method is used to identify nonlinear discrete difference equations which describe the dynamics of the D_{st} index under the influence of the solar wind. The analysis of the data sets which led to the identification of these equations, was not based on any apriori information about physical processes in the magnetospheric system. The structure, form and dynamical features are learnt from the data. It is shown that these models can then be readily mapped into the frequency domain to reveal energy transfer mechanisms that characterise the magnetosphere dynamics. Finally these results are used to derive a continuous time nonlinear differential equation model of the D_{st} evolution driven by the solar wind as we will do in our next publication.

Brief introduction to the identification of nonlinear dynamical systems with applications to the dynamics of the magnetosphere

Assume that the system can be represented by a nonlinear black box system which is characterised by a state vector \vec{x} of dimension n . In many cases the values of the state vector and procedures to measure the states are not available. Sometimes the dimension n is unknown. What is usually known is that the evolution of $\vec{x}(t)$ is determined by previous values, by external factors $\vec{u}(t)$ and by noise $\vec{e}_1(t)$ so that

$$\frac{d\vec{x}}{dt} = F(\vec{x}(t); \vec{u}(t); \vec{e}_1(t)) \quad (1)$$

If there are some unknown inputs to the system these can be considered as a part of the noise. The nonlinear system can usually be observed by making some measurements $\vec{y}(t)$ which are assumed to be a function of the state of the system, and can be corrupted by errors $\vec{e}_2(t)$

$$\vec{y}(t) = G(\vec{x}(t)) + \vec{e}_2(t) \quad (2)$$

It is customary to refer to y as the output of the system. In the vast majority of cases the only available information are discrete data sets of previous values of the input

$$\dots, \vec{u}_{k-2}, \vec{u}_{k-1}, \vec{u}_k, \dots$$

and the output

$$\dots, \vec{y}_{k-3}, \vec{y}_{k-2}, \vec{y}_{k-1}, \dots$$

The identification of a nonlinear system, as described above, involves problems such as:

1. *Forecasting*: What will be the future evolution of the output $y_k, y_{k+1} \dots$?
2. *Model validation*: If some theoretical models for functions F and G are derived how is it possible to validate them?
3. *Identification of the system dynamics*: If theoretical models of the nonlinear system are not clear, how to derive an analytical relation between \vec{u} and \vec{y} which can be used to uncover the physical processes underlying the dynamics of the system under investigation.

In the case of the terrestrial magnetosphere the state of the system $\vec{x}(t)$ is unknown. There is not even a clear understanding of what minimal set of parameters completely determines the state of the magnetosphere. However it is known that the evolution of the magnetosphere is determined by its history and by the solar wind. Thus parameters of the solar wind can be considered as an independent variable or an input that affects the magnetosphere. Geomagnetic indices are the output of such a system.

It is well known that the evolution of a linear system is determined by *the impulse response*

function $h_1(t)$. The output is related to the input through the convolution integral

$$y(t) = \int_0^{\infty} h_1(\tau)u(t - \tau)d\tau$$

The Fourier transform of $h_1(t)$ is the linear frequency response function $H_1(f)$ and if $U(f)$ and $Y(f)$ are the Fourier transforms of input and output then

$$Y(f) = H_1(f)U(f)$$

Thus the absolute value and the phase of $H_1(f)$ determines the amplification of a spectral component at the frequency f and its phase respectively. Both descriptions, the impulse response function in the time domain and the linear frequency response in the frequency domain, are equivalent. Thus the output of linear systems is fully determined by the input and by either the impulse response function $h_1(t)$ or the linear frequency response $H_1(f)$.

The Volterra series representation is a generalisation of this approach to nonlinear systems

$$\begin{aligned} y(t) &= \int_0^{\infty} h_1(\tau)u(t - \tau)d\tau + \\ &+ \int_0^{\infty} \int_0^{\infty} h_2(\tau_1, \tau_2)u(t - \tau_1)u(t - \tau_2)d\tau_1 d\tau_2 + \dots \\ &+ \int_0^{\infty} \dots \int_0^{\infty} h_i(\tau_1, \dots, \tau_i)u(t - \tau_1) \dots u(t - \tau_i)d\tau_1 \dots \tau_i + \dots \end{aligned} \quad (3)$$

where $h_i(\tau_1, \dots, \tau_i)$ is referred to as the i th order Volterra kernel.

In the case of discrete measurements (3) takes the form

$$\begin{aligned} y(k) &= \sum_{r_1 \geq 0}^K h_1(r_1)u(k - r_1) + \\ &+ \sum_{r_1, r_2 \geq 0}^K h_2(r_1, r_2)u(k - r_1)u(k - r_2) + \\ &+ \sum_{r_1, r_2, \dots, r_i \geq 0}^K h_i(r_1, r_2, \dots, r_i)u(k - r_1)u(k - r_2) \dots u(k - r_i) + \dots \end{aligned} \quad (4)$$

In the frequency domain the system is characterised by taking the Fourier transform of equation (3) to yield

$$\begin{aligned} Y(f) &= H_1(f)U(f) + \sum_{f_1, f_2; f_1+f_2=f} H_2(f_1, f_2)U(f_1)U(f_2) + \\ &+ \sum_{f_1, f_2, f_3; f_1+f_2+f_3=f} H_3(f_1, f_2, f_3)U(f_1)U(f_2)U(f_3) + \dots \\ &+ \sum H_i(f_1, f_2, \dots, f_i)U(f_1)U(f_2) \dots U(f_i) + \dots \end{aligned} \quad (5)$$

where $H_l(f_1, f_2, \dots, f_l)$ are referred to as higher or nonlinear frequency response functions or more commonly Generalised Frequency Response Functions (GFRF). $H_1(f)$ is equivalent to the well known linear frequency response function. The absolute value and phase of $H_1(f)$ determine the amplification and phase delay for each spectral component of the output. The magnitude of $H_2(f_1, f_2)$ determines the strength of coupling between two spectral components f_1 and f_2 of the input that transfer the energy to the spectral component on the summation frequency $f_1 + f_2$ in the output. Similarly H_l describes the energy transfer from l spectral components in the input to the summation frequency $f_1 + f_2 + \dots + f_l$ in the output etc. Similar to the linear systems, the infinite sets of $h_1; h_2; h_3; \dots; h_l; \dots$ in the time domain or $H_1; H_2; \dots; H_l; \dots$ in the frequency domain provide a complete description of the nonlinear system. They can be used to determine the future output of the system if the input is known. Interpretation of $H_2(f_1, f_2)$, $H_3(f_1, f_2, f_3)$ and higher order GFRF's exposes energy transfer effects and often provides considerable insight into the physical mechanisms associated with the underlying system.

Time domain analysis is more convenient for forecasting the output of the system. But the frequency domain analysis is designed to reveal invariant system features which should exist irrespective of the model or the estimation method. The GFRF's can also be used to reconstruct continuous time nonlinear differential equation models from discrete data sets. Such models can be built term by term, again with the objective of finding the simplest model which represents the system.

In practice the Volterra decomposition must be truncated to a finite number of terms. The class of systems which can be approximated by Volterra series was studied by *Boyd and Chua* [1985]. They concluded that systems with *fading memory* may be approximated arbitrarily well by truncated Volterra series. Mathematically the condition of *fading memory* is a stronger version of *continuity* [Pearson and Ogunnaike, 1997]. The precise definition can be found in [Boyd and Chua, 1985]. Generally speaking a fading memory system is one

where the dependence on the input decreases rapidly enough with time.

The NARMAX (Nonlinear AutoRegressive Moving Average models with eXogenous inputs) model is more general than the Volterra series. NARMAX models can represent a wide class of nonlinear systems and there is no restriction to systems with properties of fading memory. Once the NARMAX model of the system has been determined this can be used to analytically calculate any component H_i of the GFRF's. The NARMAX representation, proposed by *Leontaritis and Billings* [1985], is given as

$$y(k) = F[y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u), \xi(k-1), \dots, \xi(k-n_\xi)] + \xi(k) \quad (6)$$

where $F[\cdot]$ denotes a nonlinear function, u and y are the discrete-time input and output signals with corresponding maximum lagged values of n_u and n_y . The quantity $\xi(k)$ accounts for possible noise, and uncertainties, n_ξ represents the maximum noise lag. The nonlinear function F can be a polynomial or rational function, radial basis functions, wavelet decomposition or any other function.

It has been shown that a large class of nonlinear systems accept a NARMAX representation. In particular Hammerstein, Wiener, bilinear and Volterra systems can be regarded as special cases or subsets of the class of NARMAX systems. The NARMAX representation of nonlinear systems has sound theoretical foundations and many algorithms for structure detection, parameter estimation and model validation of the NARMAX model have been developed.

The structure and parameters in the NARMAX mathematical model can be identified using an orthogonal least-squares algorithm, which searches through all the potential model terms and selects the final model terms according to the contribution that they make to the variance of the system output. This allows the user to build the simplest possible model using the most significant model terms. Model validation techniques are then applied to confirm that the model is adequate. These involve high order correlations which insure that the residuals are unpredictable from all past values of the input and output.

It is the NARMAX approach which will be used in the present work to identify features of the nonlinear dynamics of the terrestrial magnetosphere.

Magnetic storms and substorms are the result of the transfer of directed kinematic energy of the solar wind flow initially into magnetic energy stored in the magnetotail and then into thermal energy in the plasma sheet, ring current, auroral particles and Joule heating of the ionosphere [Gonzalez *et al*, 1994]. It is believed that the largest portion of this energy during geomagnetic disturbances goes into the buildup of the storm ring current. The D_{st} index was proposed to quantify the intensity of the ring current. However magnetopause currents and maybe induced currents in the solid Earth, also affect the D_{st} value.

The reconnection of the geomagnetic field and the southwardly directed interplanetary magnetic field is widely accepted as the mechanism for energisation of the magnetosphere and for magnetic storms [Gonzalez *et al*, 1994]. Previous studies have led to an understanding about the rate of energy transfer from the solar wind into the magnetosphere via the reconnection processes. A list of the main coupling functions between the solar wind and the terrestrial magnetosphere can be found in the review [Gonzalez *et al*, 1994]. Numerous studies of correlations between solar wind parameters and magnetospheric disturbances have resulted in the conclusion that the product of solar wind velocity V and the southward component of the magnetic field represented by B_s ($B_s = 0$ if $B_z > 0$ otherwise $B_s = B_z$) represents the input to the magnetosphere reasonably well.

Klimas *et al* [1997] proposed the use of local-linear filters to capture the underlying dynamics of the terrestrial magnetosphere which lead to the evolution of the D_{st} index. This approach was based on the fact that any nonlinear system can be linearised in a small locality of any point in the phase space. Thus the input-output relation can be represented by a linear polynomial (4) in the vicinity of any point of phase space. This linear polynomial can be used to estimate the linear frequency response function $H_1(f)$. Klimas *et al* [1997] mapped $H_1(f)$ into a continuous linear differential relation between VB_s and D_{st} of the

form

$$\frac{d^m D_{st}}{dt^m} + \alpha_{m-1} \frac{d^{m-1} D_{st}}{dt^{m-1}} + \alpha_0 D_{st} = \beta V B_s$$

But local linear models cannot always capture the correct underlying nonlinear characteristics, and they do not reveal any information about nonlinear energy transfer mechanisms. In the present study therefore nonlinear mathematical representations will be identified. These will then be used to compare the corresponding generalised frequency response functions, and to provide an interpretation of the dynamics of the terrestrial magnetosphere.

Data Analysis in the Time Domain

The input-output data analysed in this paper will consist of 4344 hours of data for D_{st} and the solar wind VB_s over the time period from January to June 1979. The data is illustrated in Figure 1. The data with an initial resolution of 5 minutes were provided by A. Klimas and Y. Kamide. The D_{st} data has not been pressure corrected, and the VB_s data has been propagated ballistically to the magnetopause. Inspection of Figure 1 shows that a few strong magnetospheric storms $D_{st} < -100$ took place during the time period under investigation.

Initial preprocessing of the data led to the conclusion that the data were oversampled. The data were therefore decimated by a factor of 12, to yield a data set of 4344 points sampled with a 1 hour resolution. This data set will be used as the basis for all the results in the present study. The set of 1000 points between samples [501;1500] of the data set was chosen as the estimation data. Applying the NARMAX procedure to the data, using the Orthogonal Least Squares estimation and a polynomial expansion of $F[\cdot]$, the model of

process terms was identified in the form

$$\begin{aligned}
D_{st}^m(k) = & 1.18D_{st}^m(k-1) - 4.64VB_s(k-1) - 0.12D_{st}^m(k-2)VB_s(k-1)+ \\
& + 0.07D_{st}^m(k-3)VB_s(k-1) - 0.73D_{st}^m(k-2) + 0.67D_{st}^m(k-3)- \\
& - 0.41D_{st}^m(k-4) + 0.19D_{st}^m(k-5) - 1.03VB_s(k-1)VB_s(k-6)+ \\
& + 0.93VB_s(k-2) + 0.96VB_s(k-1)VB_s(k-7) + 0.81VB_s(k-1)VB_s(k-13) \\
& - 0.67VB_s(k-3)VB_s(k-15) - 0.85VB_s(k-1)VB_s(k-8)+ \\
& + 0.31VB_s(k-3)VB_s(k-8)
\end{aligned} \tag{7}$$

The model (7) consists of 15 process terms, and these contributed with 98.95% to the total variance in the output. The notation D_{st}^m is used for the identified geomagnetic index.

The validity of the model equation (7) can be assessed in two ways, based on the performance of the model as predictor for the rest of the data set and by the statistical properties of the one step ahead prediction errors

$$e(k) = D_{st}(k) - D_{st}^m(k)$$

where D_{st}^m denotes the identified geomagnetic index, while the notation D_{st} is used for the measured geomagnetic index. If the identified model in equation (7) has captured the dynamics of the system which is present in the data set, $e(k)$ should be unpredictable from all past inputs, outputs and predictions errors. This can be tested using the model validity tests developed by [Billings and Voon, 1983] and these are illustrated in Figure 2 for the model in equation (7). It can be seen that all correlations in Figure 2 are inside the dashed 95% confidence bands, which suggests that the polynomial model in equation (7) is a good representation of the input-output relation between D_{st} and VB_s . Validation based on more recent tests which are based on the output D_{st} and the residuals are illustrated in Figure 3. These tests are more sensitive and it can be seen that the first test is just outside the 95% confidence bands at lag 3, while the second test is outside at approximately lag 20. Since the true confidence bands bell out, the latter test is probably acceptable. Fine tuning of the

model did not improve the validation tests and the model was accepted as the best fit.

The reliability of the identified model equation (7) can also be judged by the predictive performance of the model. The simplest type of prediction is *one-step ahead* forecasting, in which the predicted values of the storm index at the moment k $D_{st}^{1s}(k)$ are calculated from previous *measured* values of D_{st} and VB_s using the identified NARMAX model (7)

$$D_{st}^{1s}(k) = F^d [D_{st}(k-1), \dots, D_{st}(k-n_o); VB_s(k-1), \dots, VB_s(k-n_i)] \quad (8)$$

The result of such a prediction is plotted in Figure 4a (a) where the one step ahead predictions, in blue, are compared to the real data, in red, over the whole data set. Figure 4a (a) clearly shows that the predictions are in excellent agreement with the measurements. The agreement is particularly impressive for points outside the data used for estimation, interval [501; 1500].

However, while this is often the only type of pendulum used by many authors to validate identified models, previous experience shows that sometimes even poor models give good one step ahead predictions. This is because the prediction is essentially reset, using the data set, at each step. A much better test is to compute the many steps ahead predictions D_{st}^{ms} , often called the model predicted output. This is computed without using any measured output values, except a few initial values, and is defined as

$$D_{st}^{ms}(k) = F^d [D_{st}^{ms}(k-1), \dots, D_{st}^{ms}(k-n_o); VB_s(k-1), \dots, VB_s(k-n_i)] \quad (9)$$

The identified discrete relation (7) depends upon 5 previous output measurements, therefore only the initial 5 measured output values are required to compute the many steps ahead predictions. The results of the many steps ahead predictions over the whole data set are illustrated in Figure 4a (c). As expected the agreement is not as good as with one step predictions, but these are good long term predictions and they give confidence in the identified model.

The disagreement between these predictions and the measured values of D_{st} should be

attributed to the existence of other inputs to the system. In practice one step ahead (or 1-hour ahead) predictions are not very useful, as it is difficult during a few minutes to collect all data from both satellite measurements and ground based magnetometers and to feed them into relation (7). On the other hand forecasting D_{st} half a year ahead of the real measurements is not required. To be practically useful, the prediction should be made on some time scale which is intermediate between these two extreme cases. A ten hours ahead prediction based on (7) is plotted in Figure 4a (b). These predictions are almost in the same good agreement with the real measurements over the whole data set as the one step ahead predictions.

Summarising the results represented in Figures 2, 3 and 4a it can be concluded that the discrete relation (7) appears to provide a very good representation of the dynamics between the input VB_s and the output D_{st} index. It is important to emphasise that the model equation (7) was identified based only on the data in the interval [501;1500] but the model prediction tests show good agreement over the whole data set [1;4344], and not only on the estimation set.

Frequency domain analysis

The time domain relation (7) is convenient for forecasting and validation purposes. But frequency domain analysis is more suitable for analysis of physical aspects of the dynamics. Once the NARMAX model (7) has been identified and validated a simple analytical procedure can be applied to calculate any order of the GFRF's. This avoids all the limitations and assumptions required in FFT based methods. The first three GFRF's are plotted in Figures 5 a,b ($H_1(f)$), Figures 4 b,c ($H_2(f_1, f_2)$) and Figure 7 ($H_3(f_1, f_2, f_3)$).

The magnitude of $H_1(f)$ plotted in Figure 5a represents the linear amplification of the spectral components. The maximum amplification occurs at very low frequency $1/f > 100\text{hours}$. The second local maximum can be seen at $f \approx 0.3\text{hour}^{-1}$. The transfer function

resembles that of the linear system with two resonance frequencies one of which is very low. For example a double pendulum with two characteristic lengths will have such a resonance curve if one of the lengths is infinitely long. The phase of $H_1(f_1)$ is plotted in Figure 5b.

The second order frequency response function is plotted in the Figure 4b as an isometric surface plot from which it is easy to see the relative magnitudes of the various peaks, ridges and valleys in H_2 . However a plan view of this surface, shown in Figure 4c, is more suitable to study the positions of these maxima. The physical interpretation of the maxima in $H_2(f_1, f_2)$ observed at the coordinates $f_1 = f_k$ and $f_2 = f_i$ is that a nonlinear coupling between two spectral components of the input observed at frequencies f_k and f_i is present with a transfer of energy from these spectral components to the spectral component of the output at the summation frequency $f_j = f_i + f_k$.

Analysis of any sets of measurements taken with a sampling frequency of f_{samp} can not provide information about frequencies higher than the Nyquist frequency $f_{Nyq} = 1/2f_{samp}$. Data sets which were used to identify the NARMAX model in equation (7) have a sampling frequency of $1/hour$ and so $H_2(f_1, f_2)$ can only be determined for f_1 and f_2 such that the sum does not exceed the Nyquist frequency. Thus H_2 in Figure 4b-4c is plotted only for points for which $f_1 + f_2 \leq 0.5hour^{-1}$. Maxima which are located along the line $f_2 = -f_1$ prevail on the H_2 surface. These positions of the main maxima indicate that the dominant three-wave process in the dynamics of the system under consideration is the coupling of the spectral components observed at frequencies of virtually the same magnitude but opposite signs. Interaction between these components should result in the transfer of energy to the summation frequency $F_{sum} = F_1 + (-F_1) = 0$ that is to zero frequency, which corresponds to the main maximum in H_1 .

This process has a very simple interpretation in terms of energy storage. If energy in the signal is stored at the spectral component f the characteristic time of release τ of this energy is proportional to $\tau \propto 1/f$. Assume that energy is supplied to a linear system by

spectral components f_1 and f_2 . As the linear system is not able to transfer energy between spectral components the energy will be released on the time scales $\tau_{1l} \propto 1/f_1$ and $\tau_{2l} \propto 1/f_2$ that is on the same time scales at which energy is consumed from the input. However if the system is nonlinear and frequencies f_1 and f_2 have almost the same magnitude but opposite sign, then the coupling between these two spectral components will transfer energy to the summation frequency $f_1 + f_2 = \epsilon \ll |f_1|, |f_2|$. In this case the characteristic scale on which energy will be released is $\tau_{nonlin} \propto 1/\epsilon$ which is much higher than the temporal scale on which energy is supplied. Thus the storage of energy by a nonlinear resonator will take place.

Another much less prominent feature of the $H_2(f_1, f_2)$ surface is a ridge-like local maximum around frequency $f_1 = 0$ for $|f_1| < \delta f \approx 0.03 \text{hour}^{-1}$ and $0 \leq f_2 \leq 0.15$. This indicates nonlinear coupling which involves spectral components at very low frequencies which transfers energy to higher frequencies. This processes is opposite to the storage described above and will be referred to as the release of energy.

The third order GFRF $H_3(f_1, f_2, f_3)$ is defined on the three dimensional space $R^1 \times R^1 \times R^1$. Thus the only way to plot this is to plot slices. The plots of H_3 on 16 slices which correspond to planes $f_3 = \text{const}$ as plan views are shown in Figure 7a. Values of f_3 are shown above each sub-plot. $H_3(f_1, f_2, f_3)$ is plotted only for points for which $f_1 + f_2 + f_3 \leq f_{Nyq}$. All plots are dominated by a maximum ridge defined by the line $f_1 + f_2 = -f_3$. On the upper left plot ($f_3 = -0.5$) the dominate maximum is along the line $f_1 + f_2 = 0.5$. With an increase in f_3 the position of the maximum shifts to the left, until it almost disappears in the bottom left hand corner of the $f_3 = 0.4375$ plot. The intensity of this maximum increases as f_3 approaches 0. Similar to the case for $H_2(f_1, f_2)$ this dominant feature of $H_3(f_1, f_2, f_3)$ also corresponds to the processes

$$f_1 + f_2 + f_3 \Rightarrow 0,$$

that is interaction of three waves in the input with the transfer of energy to the spectral

component at a very low frequency. This can again be interpreted as energy storage.

Another feature of $H_3(f_1, f_2, f_3)$ which can be seen in Figure 7a are the maxima located around the lines $f_1 = f_2 = 0$, $f_1 = f_3 = 0$, $f_2 = f_3 = 0$. The last two of these are most evident in the plot which corresponds to the $f_3 = 0$ as maxima around the lines $f_1 = 0$ and $f_2 = 0$, particularly in Figure 7b where H_3 is plotted on a larger scale. The third maximum, $f_1 = f_2 = 0$, is represented by local maxima around points $f_1 = 0; f_2 = 0$ on the plots for $f_3 \in \{-0.0625, 0.0, 0.0625, 0.125\}$ in Figure 7b. These maxima represent processes similar to the release of energy discussed above in relation to $H_2(f_1, f_2)$, that is the transfer of energy from very low frequencies to higher frequencies.

In numerical methods various discrete difference equations can correspond to the same continuous one. Similarly different polynomial relations of the type (1) can correspond to the same continuous dynamical process. However if all these relations are derived by proper methods their representations in the frequency domain should be very close to each other. Thus to study if the described features of H_1 , H_2 and H_3 are real properties of the dynamical system or are related to the concrete discrete representation (1), we have derived other discrete representations. In order to derive these other discrete representations we have changed the decimation factor and the training interval. For example the representation

$$\begin{aligned}
D_{st}^m(k) = & 1.16D_{st}^m(k-1) - 1.97VB_s(k-2) - 0.45D_{st}^m(k-2) + 0.25D_{st}^m(k-3) \\
& + 0.75VB_s(k-4) - 0.058VB_s(k-2)VB_s(k-6)VB_s(k-6) \\
& + 0.29VB_s(k-2)VB_s(k-8) - 0.0093VB_s(k-9)VB_s(k-9)VB_s(k-9) \\
& - 0.03VB_s(k-3)VB_s(k-3)VB_s(k-3) + 0.0035VB_s(k-5)VB_s(k-5)VB_s(k-10) \\
& - 0.0047D_{st}^m(k-1)VB_s(k-2)VB_s(k-2) + 0.33VB_s(k-6) \\
& + 0.0018D_{st}^m(k-3)VB_s(k-1)VB_s(k-4) - 0.20VB_s(k-2)VB_s(k-2) \\
& + 0.012D_{st}^m(k-3)VB_s(k-2)
\end{aligned} \tag{10}$$

was derived by using 4000 training points with a 40 minutes sample time (decimation factor

8). One step ahead, 10 hours ahead and model predicted output for the model (10) are shown in Figure 4a (d)-(e). Visual comparison of the measured values of D_{st} with the predictions again shows good agreement. The coherency

$$\gamma(f) = \frac{\langle Y_{f \text{ pred}} Y_{f \text{ real}}^* \rangle}{\langle |Y_{f \text{ pred}}|^2 \rangle \langle |Y_{f \text{ real}}|^2 \rangle}$$

provides insight into the relationship between $Y_{f \text{ pred}}$ and $Y_{f \text{ real}}$, the spectral components at the frequency f of the predicted and real data sets respectively, and $\langle \cdot \rangle$ denotes averaging. The value of the coherency is bounded between 0 and 1. A value of coherency close to one means high correlation between the spectral components of the two data sets. A value close to zero means complete independence between two signals at that particular frequency. The magnitude and phase of the coherency function calculated for the real measurements of D_{st} and for D_{st}^* are plotted in Figure 8 a,b. The coherency estimated from a finite point data set is a statistical value, the distribution of which is related to the real value of the coherency between the two signals. In particular the variance of the phase is determined by the real value of the coherency. The coherency between the real measurements and the predictions computed using the NARMAX model (10) exceeds 0.8 for the low frequencies ($f < 0.1 \text{ hour}^{-1}$). For higher frequencies it is slightly lower, but still mainly has values between 0.5 and 0.75. The phase of the coherency function varies in the vicinity of 0, indicating the absence of time shift between the model (10) and the actual data. The small variance of the phase is in accordance with high values of the magnitude of coherency.

Figures 4a (d)-(e) and Figure 8 suggest that the identified model in equation (10) can be used for the reliable prediction of the D_{st} index. $H_1(f)$, $H_2(f_1, f_2)$, $H_3(f_1, f_2, f_3)$ for the model equation (10) are illustrated in Figures 4d-4e and Figures 9-11. The main maximum of $H_1(f)$ in Figure 9a is also located at very low frequencies. The second distinct local maximum observed in Figure 5a at $f \approx 0.25 - 0.3 \text{ hour}^{-1}$ is absent in Figure 9a. However the change of the slope of H_1 at frequencies $0.2 - 0.3 \text{ hour}^{-1}$ may be an indication of a weak resonance at these frequencies. Similar to the previous model the storage maxima

$f_1 + f_2 \Rightarrow 0$ and $f_1 + f_2 + f_3 \Rightarrow 0$ dominate both $H_2(f_1, f_2)$ (Figures 4 d-e and Figure 10) and $H_3(f_1, f_2, f_3)$. However the release maxima $f + 0 + 0 \Rightarrow f$ are not visible in $H_3(f_1, f_2, f_3)$. The relative amplitude of the release maximum in H_2 ($f + 0 \Rightarrow f$) is higher than in the previous model. Note that the limits of the gray scale in the plan view of $H_2(f_1, f_2)$ (Figure 10) were suppressed to make these effects more evident.

In addition to model equations (7) and (10) other discrete relations were fitted to the data using various training intervals and decimation factors. In spite of the fact that the discrete relations differ, as expected because of the sample rate, the representation in the frequency domain through $H_1(f_1)$, $H_2(f_1, f_2)$ and $H_3(f_1, f_2, f_3)$ have a close resemblance to the two models discussed above. The linear part of the transfer function always had the main maximum at the very low frequencies, with secondary maxima or changes in the slope at frequencies $0.2 - 0.3 \text{ hours}^{-1}$.

The dominant feature of $H_2(f_1, f_2)$ and $H_3(f_1, f_2, f_3)$ for all the derived GFRF's were storage processes of the type $\sum f_i \Rightarrow 0$. However the release processes were visible only in some plots.

Discussion

The most prominent features of GFRF's for the identified NARMAX models were storage maxima observed in both $H_2(f_1, f_2)$ and $H_3(f_1, f_2, f_3)$ corresponding to the process

$$\sum f_i \Rightarrow \delta f \quad (11)$$

where δf is almost zero. The fact that all the derived GFRF's had this feature strongly suggests that the process of equation (10) is important in energy storage by the magnetosphere. Such processes of energy storage have been discussed often in relation to various nonlinear systems. For example in the process of the Langmuir collapse, energy from a collapsing cavern is transferred to a low frequency condensate by the so called inverse cascade which is the same type of process as equation (10) [Zaharov, 1972]. The simplest differential equa-

tion which has similar features in the GFRF's is a linear oscillator with a nonlinear external input $u(t)$ for which the output $y(t)$ satisfies

$$\frac{d^2y(t)}{dt^2} + \alpha \frac{dy(t)}{dt} + \beta y(t) = P[u(t)]$$

where $P[u(t)]$ denotes a polynomial in the input variable $u(t)$, of an order higher than 2. The imaginary parts of both eigen frequencies of such an oscillator should be much higher than the real parts.

The process of equation (11) is not the only possibility which can result in energy storage. The simplest storage is a time delay $t_d(f)$ for each spectral component with no energy transfer between them. In that case the storage will be a linear process which will be reflected only in the phase of the linear part of the GFRF, $\Phi_{H_1}(f)$. For example, if the time delay does not depend upon the frequency component then $\Phi_{H_1}(f)$ will increase with the frequency as $\Phi_{H_1}(f) = ft_d$.

The release of energy feature of the derived GFRF's is not as certain as the process in equation (11). The processes which correspond to the release of energy from very low frequencies ($\delta f_1; \delta f_2$) ≈ 0 to the finite frequency f are described by

$$f + \delta f_1 + \delta f_2 \Rightarrow f \quad (12)$$

This process was observed only in some of the GFRF's, and the intensity was always much lower than for the process in equation (11). The above analysis does not provide confidence that energy release via process (12) is a significant part of D_{st} . It is also possible that the time resolution of the data sets analysed was too coarse to provide comprehensive information about this type of energy release. As mentioned above a second paper will be devoted to the derivation of a continuous time nonlinear model relating D_{st} to VB_s and these interpretations will be revisited then.

The features in the generalised frequency response functions can be considered as fingerprints of the nonlinear processes which characterise the system. To illustrate this consider

the model of D_{st} dynamics proposed by *Burton et al* [1997]. In this model the evolution of the D_{st} index under the influence of solar wind is governed by the following equation

$$\frac{dD_{st}(t)}{dt} = -c_1 D_{st}(t) + c_1 c_2 \sqrt{P(t)} + \frac{d}{dt} c_2 \sqrt{P(t)} + c_4 E_{eff}(t - t_d) + c_5, \quad (13)$$

where $P(t)$ is the solar wind dynamic pressure, t_d is some time delay due to the magnetospheric processes, the effective E_{eff} component of the interplanetary electric field, and c_i represent various constants. The interplanetary electric field E_y is proportional to VB_s . In *Burton et al* [1975] $E_{eff} = 0$ during periods of northward IMF. When the IMF has a southward component *Burton* considered $E_{eff} = E_y$ except for time intervals in which strong high frequency oscillations in E_y were observed. For later periods E_{eff} was the result of filtering of E_y with a low-pass filter. In the case of the assumption made regarding the statistical independence between E_y (VB_s) and P , equation (13) represents a linear relation between VB_s and D_{st} . Thus for this mathematical model all of the GFRF's except the linear $H_1(f)$ case must vanish. Hence the nonlinear storage process (11) cannot be incorporated into (13). According to the GFRF's identified in the present study, the correct model of D_{st} must include nonlinear relationships between D_{st} and VB_s . In fact the existence of a nonlinear relation between these was proved by *Klimas et al* [1996] by demonstrating temporal changes in the coefficients of the local linear continuous time models.

In the model of *Burton et al* [1975] the energy storage from VB_s in (13) is provided via a time delay t_d which is the same for all spectral components. As discussed above such a constant delay should lead to a linear increase of the phase of H_1 with frequency which was not observed in the GFRF's. On the contrary in the derived H_1 the phase decreases with frequency, see Figure 5b. The rate of this decrease also depends upon frequency. That corresponds to a more complex process than a simple time delay.

Conclusions

The identification of a mathematical model relating the geomagnetic index D_{st} to the geomagnetic field $V B_s$ has been performed using methods from nonlinear systems identification. A concise representation of the system was identified using the NARMAX approach and model validation procedures produced good short and long term model predictions of D_{st} index. The dominant nonlinear characteristics of the geomagnetic index were studied using the Generalised Frequency Response Functions computed directly by mapping the identified NARMAX model into the frequency domain. This has revealed the existence of processes of energy storage which involve multiwave coupling.

This is just the first part of this study which will be continued in a companion paper where the identification and interpretation of a nonlinear differential equation model will be investigated.

Acknowledgements. M. Balikhin was supported by a grant from PPARC UK. The authors are grateful to the Prof. A. Klimas and Prof. Y. Kamide for providing data. The authors acknowledge very useful discussions with A. Klimas and R.L. McPherron. M. Balikhin is also grateful to V. Krasnoselskikh and O. Pokhotelov for their comments. O.M. Boaghe gratefully acknowledges financial support from the University of Sheffield. S.A. Billings gratefully acknowledges that part of this work was supported by EPSRC.

References

- Baker D. N., A. J. Klimas, R.L. McPherron, and J Buchner, The evolution from weak to strong geomagnetic activity: An interpretation in terms of deterministic chaos, *Geophys. Res. Lett.*, **17**, 41, 1990.
- Billings, S.A., Voon, W.S.F., Structure detection and model validity tests in the identification of nonlinear systems, *IEE Proceedings*, **130**, 4, 193-199, 1983.
- Boyd S., and L.O. Chua, Fading Memory and the problem of approximating Nonlinear Operators with Volterra series, *IEEE trans, Circuits and Systems*, **32**, 1150-1161, 1985.
- Burton, R.K., R.L. McPherron, C.T. Russell, An empirical relationship between interplanetary conditions and D_{st} , *J. Geophys. Res.*, **80**(31), 4204, 1975.
- Fairfield, D.H., Advances in Magnetospheric storm and Substorm Research:1989-1991, *J. Geophys. Res.*, **97**, 10865, 1992.
- Goertz, C. K.,L. H. Shan, R. A. Smith, Prediction of geomagnetic activity, *J. Geophys. Res.*, **98**, 7673, 1993.
- Gonzalez, W.D., J. A. Joselyn, Y. Kamide, H.W. Kroehl, G. Rostoker, B.T. Tsurutani, and V.M. Vasiliunas, What is a geomagnetic storm?, *J. Geophys. Res.*, **99**, 5771, 1994.
- Klimas, A. J., D. Vassiliadis, D. N. Baker and D. A. Roberts, The organised nonlinear dynamics of the magnetosphere, *J. Geophys. Res.*, **101**, 13089, 1996.
- Klimas, A. J., D. Vassiliadis, D. N. Baker, Data-derived analogues of the magnetospheric dynamics, *J. Geophys. Res.*, **102**, 26993, 1997.
- Leontaritis, I.J., Billings, S.A., Input-output parametric models for nonlinear systems, Part II: Stochastic nonlinear systems, *International Journal of Control*, **41**, 329-344, 1985.
- Pearson, R. K., and B. A. Ogunnaike, Nonlinear Process Identification, in *Nonlinear Process Control*, ed. by Seborg M. A. and D.Henson, Prentice-Hall, International (UK) Limited, London, p.11, 1997.
- Wu J.-G. and H. Lundsted, Neural Network modelling of Solar wind magnetosphere interaction, *J. Geophys. Res.*, **102**, 14457, 1997.
- Zaharov, V. E., Collapse of Langmuir waves, *Soviet Physics JETP*, **35**, 908, 1972.

(Received _____)

Paper number 6.

0148-0227/93/92JB-04601\$05.00

Figure Captions.

- Fig. 1. Input to the system VB_s and output D_{st} plotted during the studied time interval
- Fig. 2. Correlation between input VB_s and residuals e for the model (7)
- Fig. 3. Correlation between input D_{st} and residuals e for the model (7)
- Fig. 4a. Predicted (blue) and real values (red) of D_{st}
- Fig. 4b Magnitude of H_2 for the model (7) - isometric projection
- Fig. 4c Magnitude of H_2 for the model (7) - plan view
- Fig. 4d Magnitude of H_2 for the model (10) - isometric projection
- Fig. 4e Magnitude of H_2 for the model (10) - plan view
- Fig. 5 Magnitude of H_2 for the model (7) - plan view in expanded scale
- Fig. 6a Magnitude of H_1 for the model (7)
- Fig. 6b Phase of H_1 for the model (7)
- Fig. 7a Magnitude of H_3 for the model (7) - plan view
- Fig. 7b Magnitude of H_3 for the model (7) - plan view in expanded scale
- Fig. 8a Magnitude of coherency between real and predicted values of D_{st} for the model (10)
- Fig. 8b Phase of coherency between real and predicted values of D_{st} for the model (10)
- Fig. 9a Magnitude of H_1 for the model (10)
- Fig. 9b Phase of H_1 for the model (10)
- Fig. 10 Magnitude of H_2 for the model (10) - plan view in expanded scale
- Fig. 11a Magnitude of H_3 for the model (10) - plan view
- Fig. 11b Magnitude of H_3 for the model (10) - plan view in expanded scale

BOAGHE ET AL: MAGNETOSPHERE DYNAMICS

BOAGHE ET AL: MAGNETOSPHERE DYNAMICS

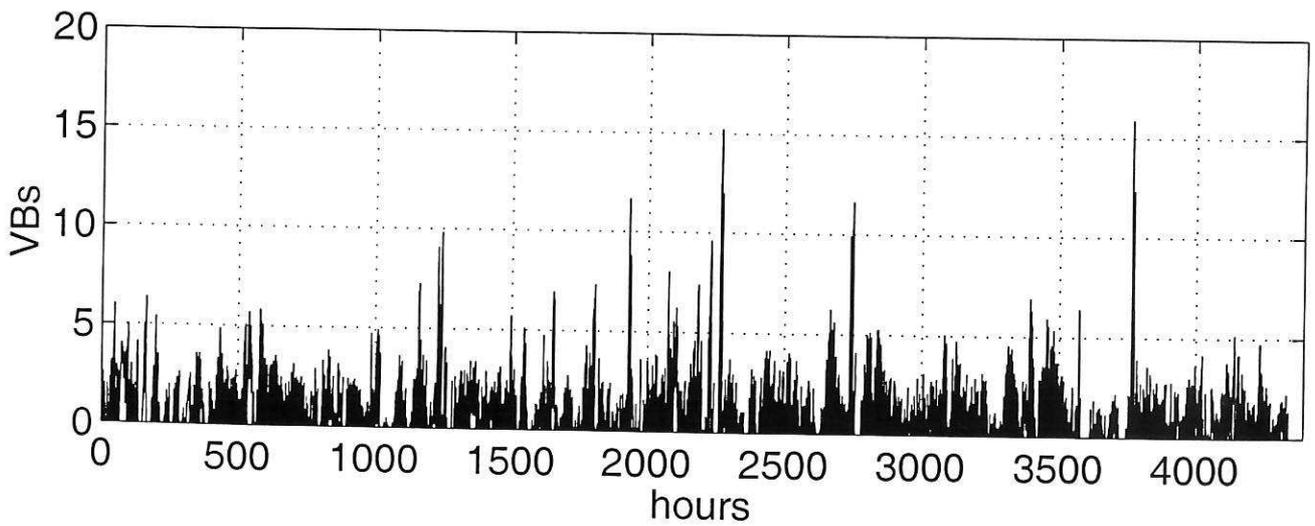
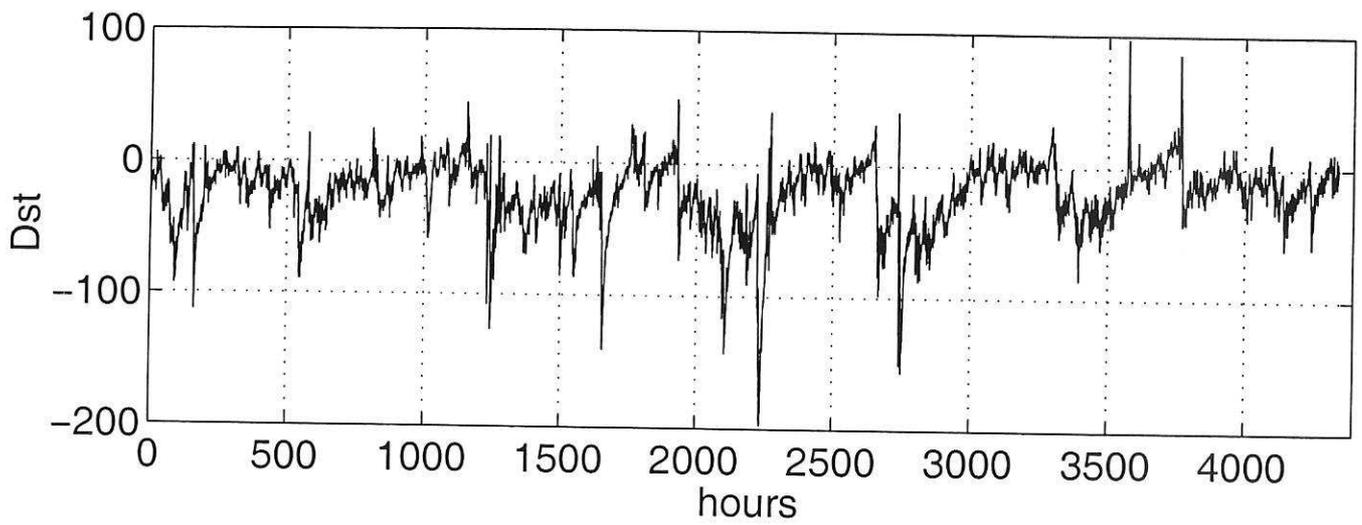


Figure 1

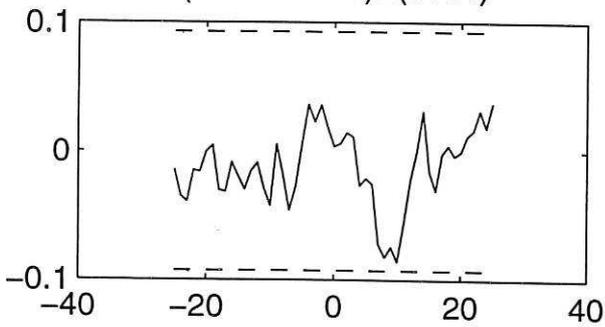
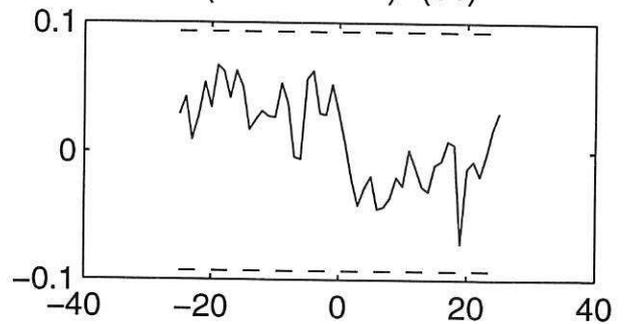
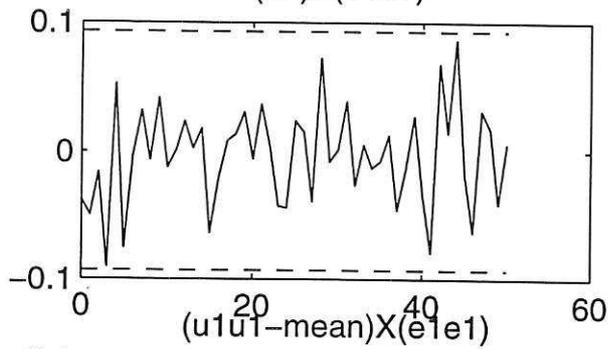
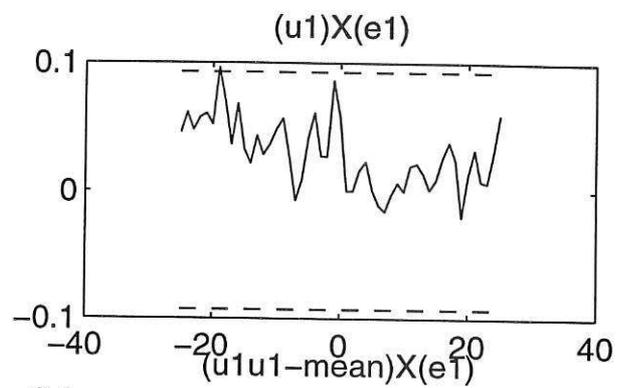
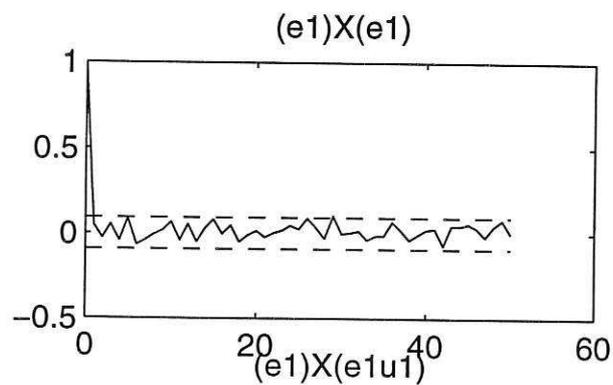


Figure 2

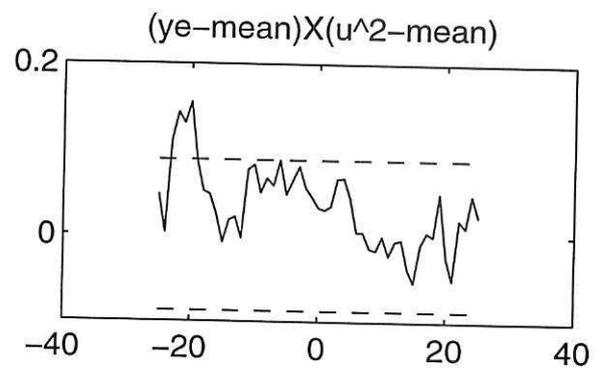
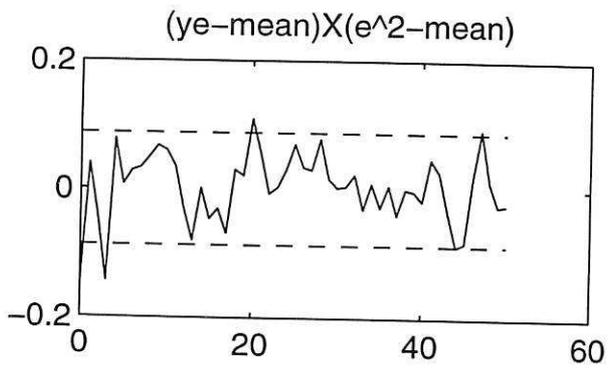


Figure 3

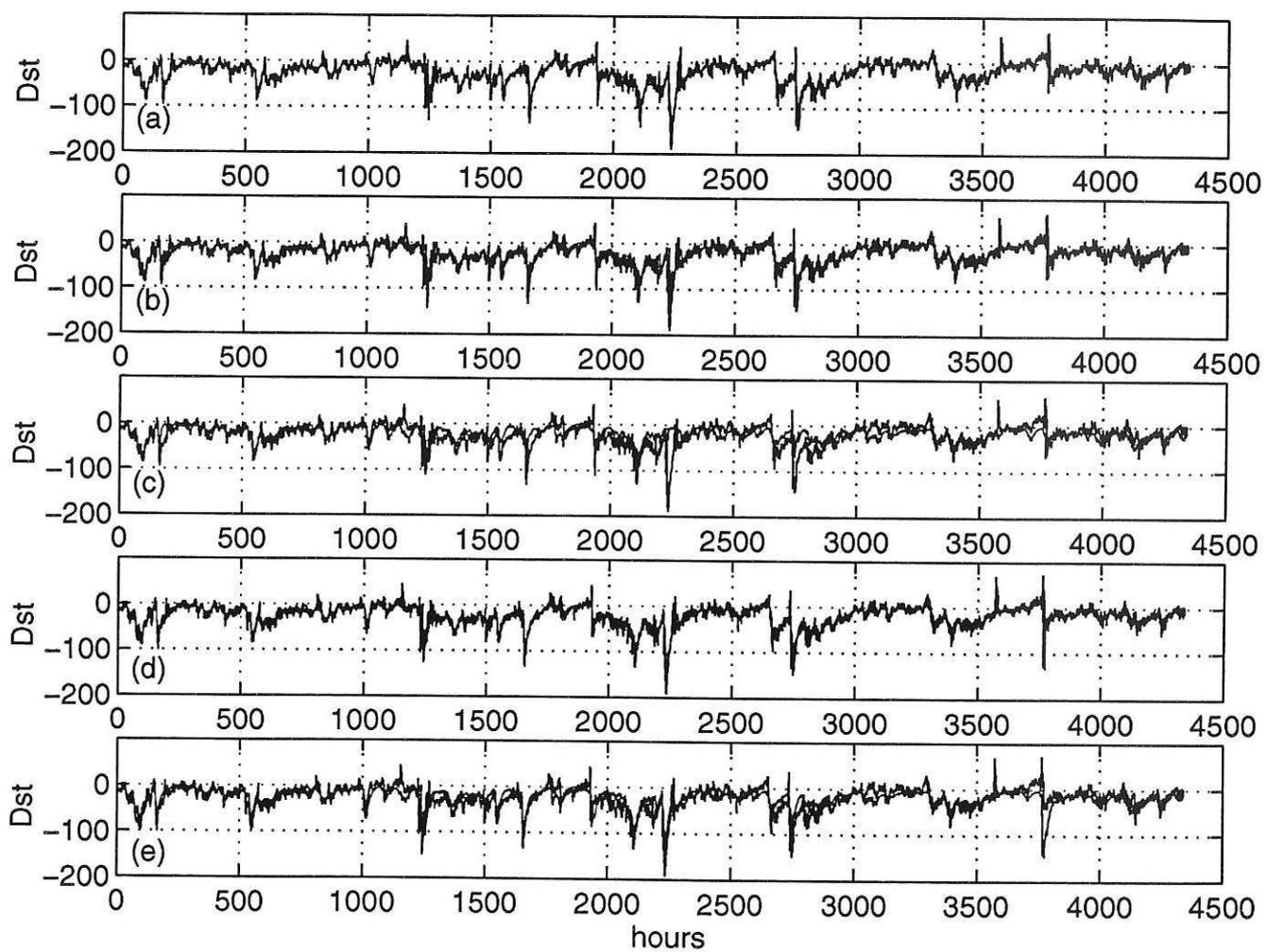


Figure 4a

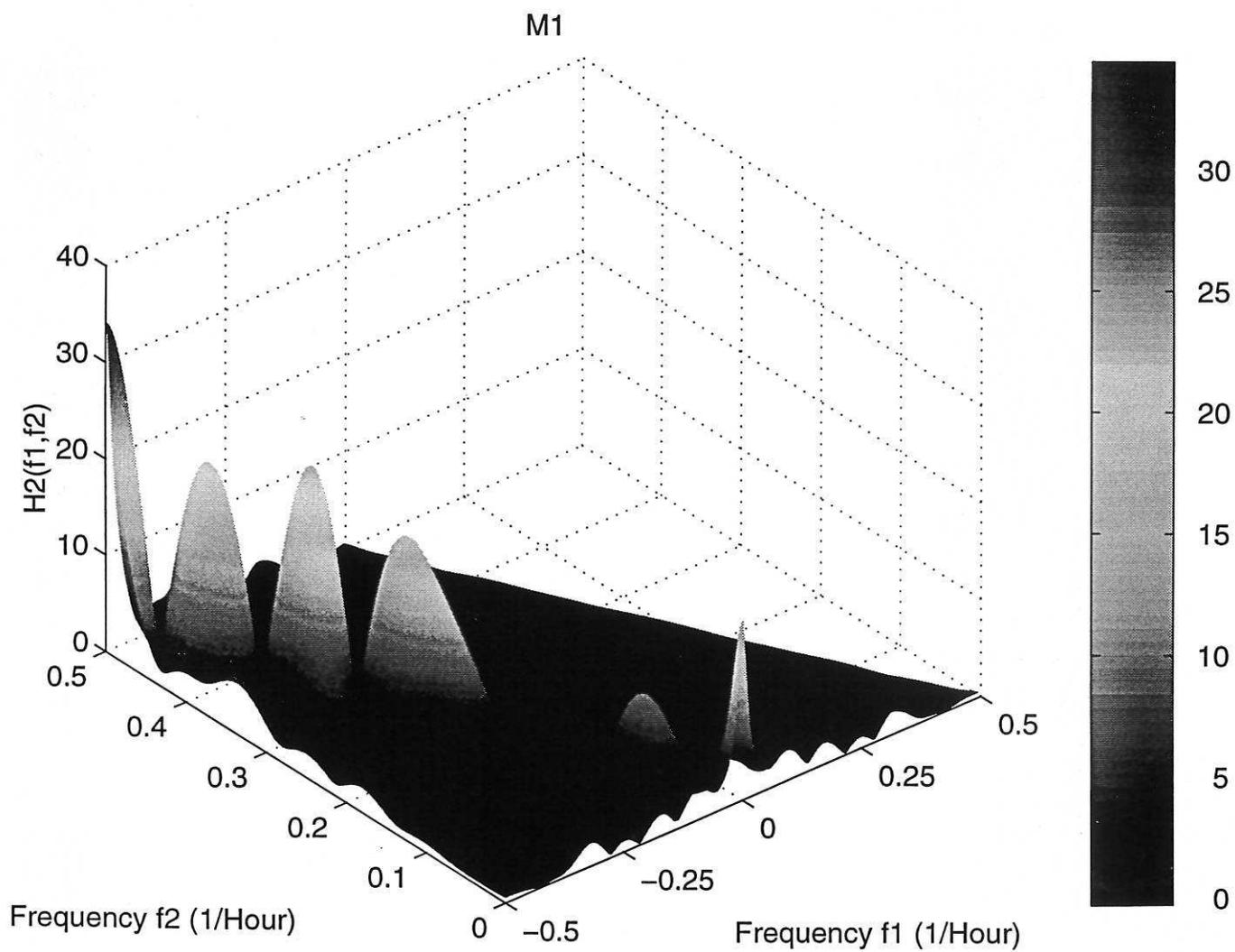


Fig. 4b

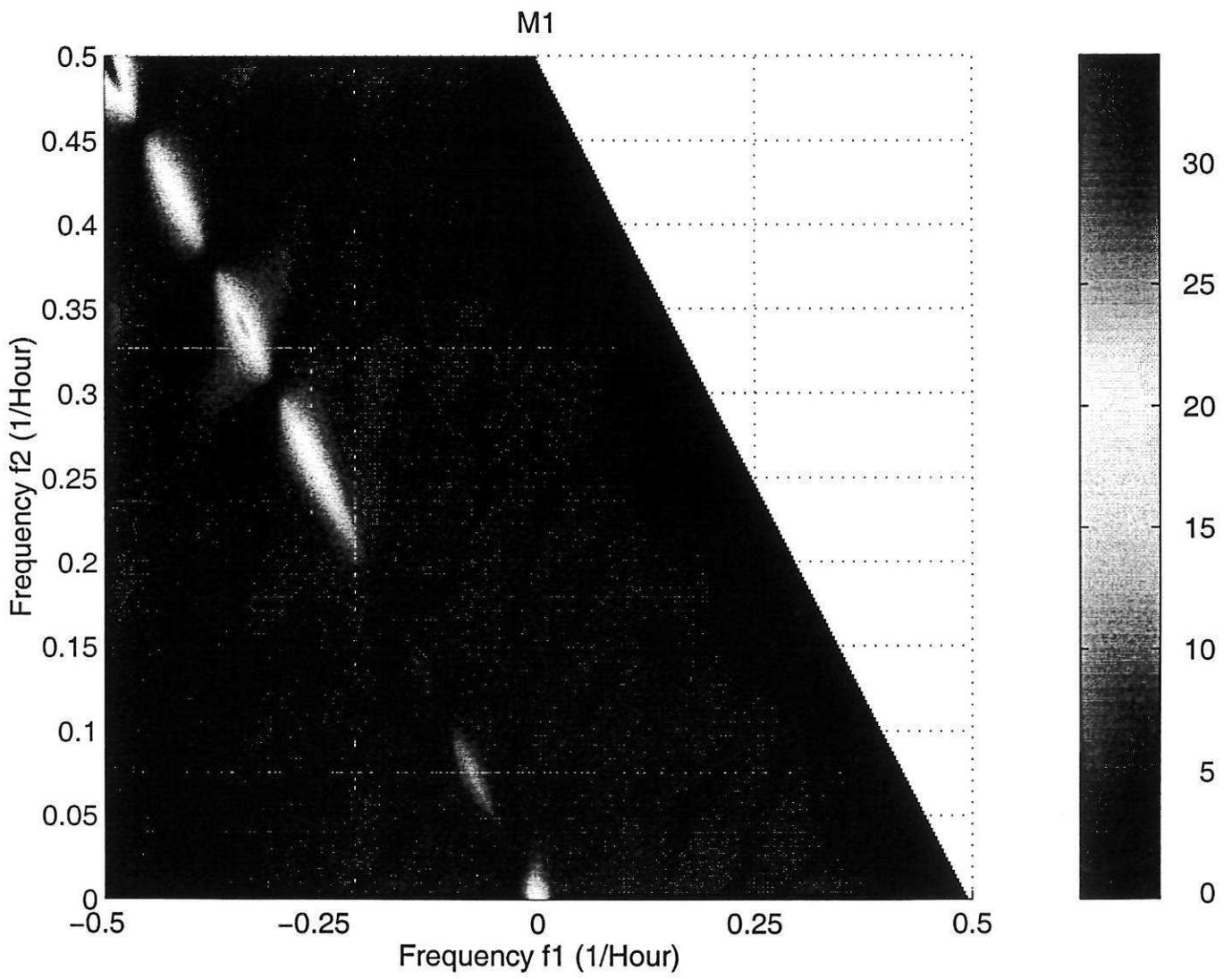


Fig. 4c

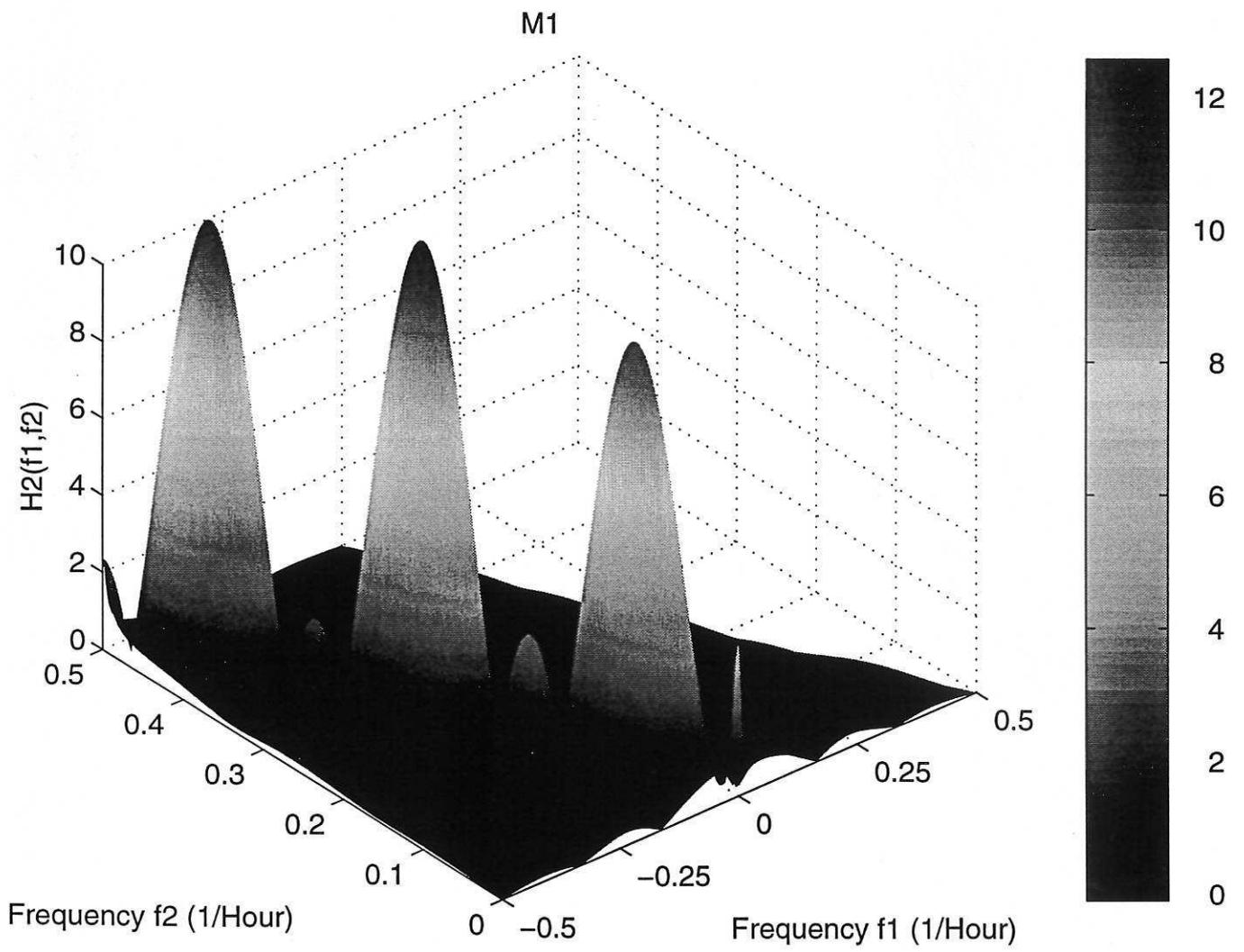


Fig-4 d

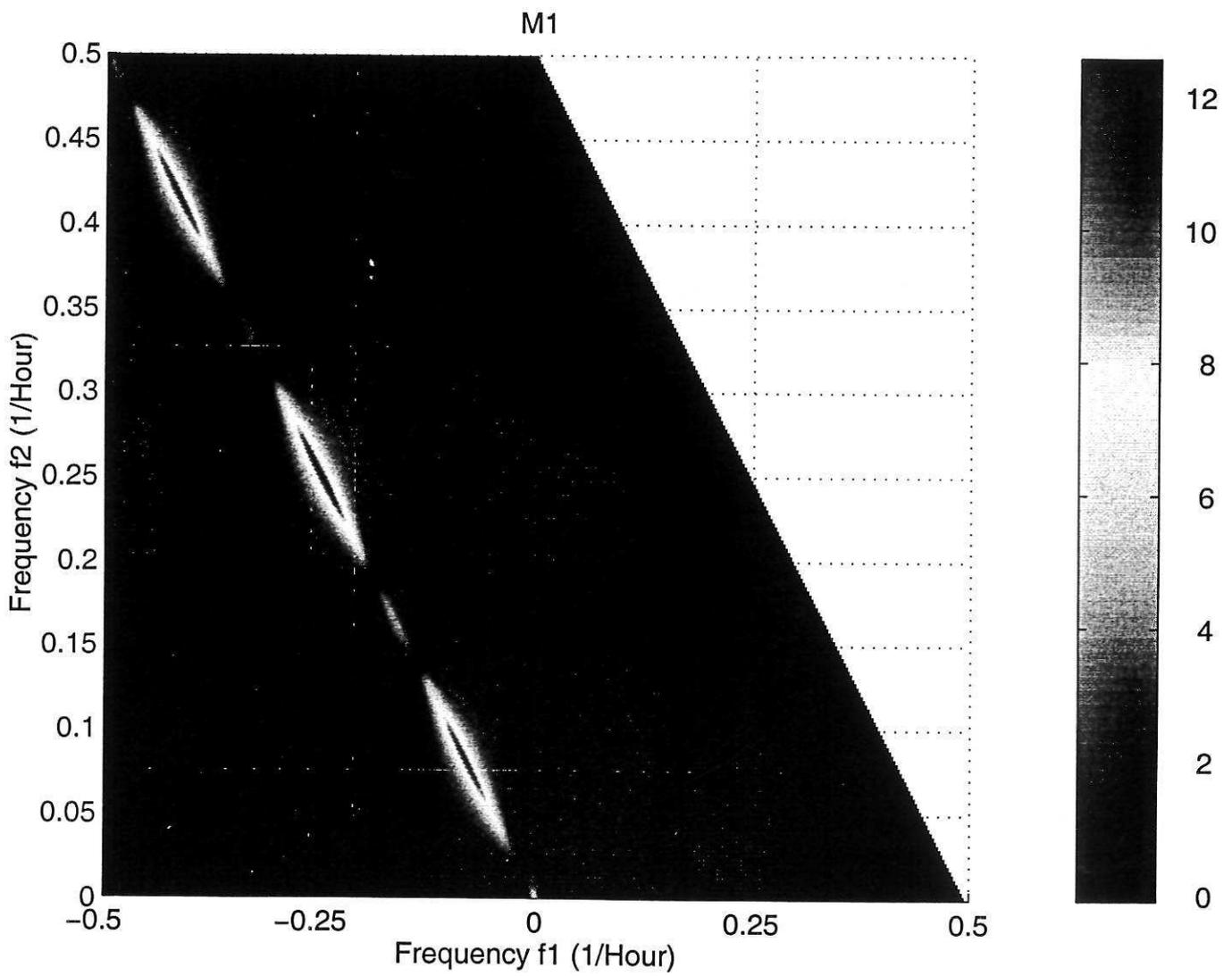


Fig-4e

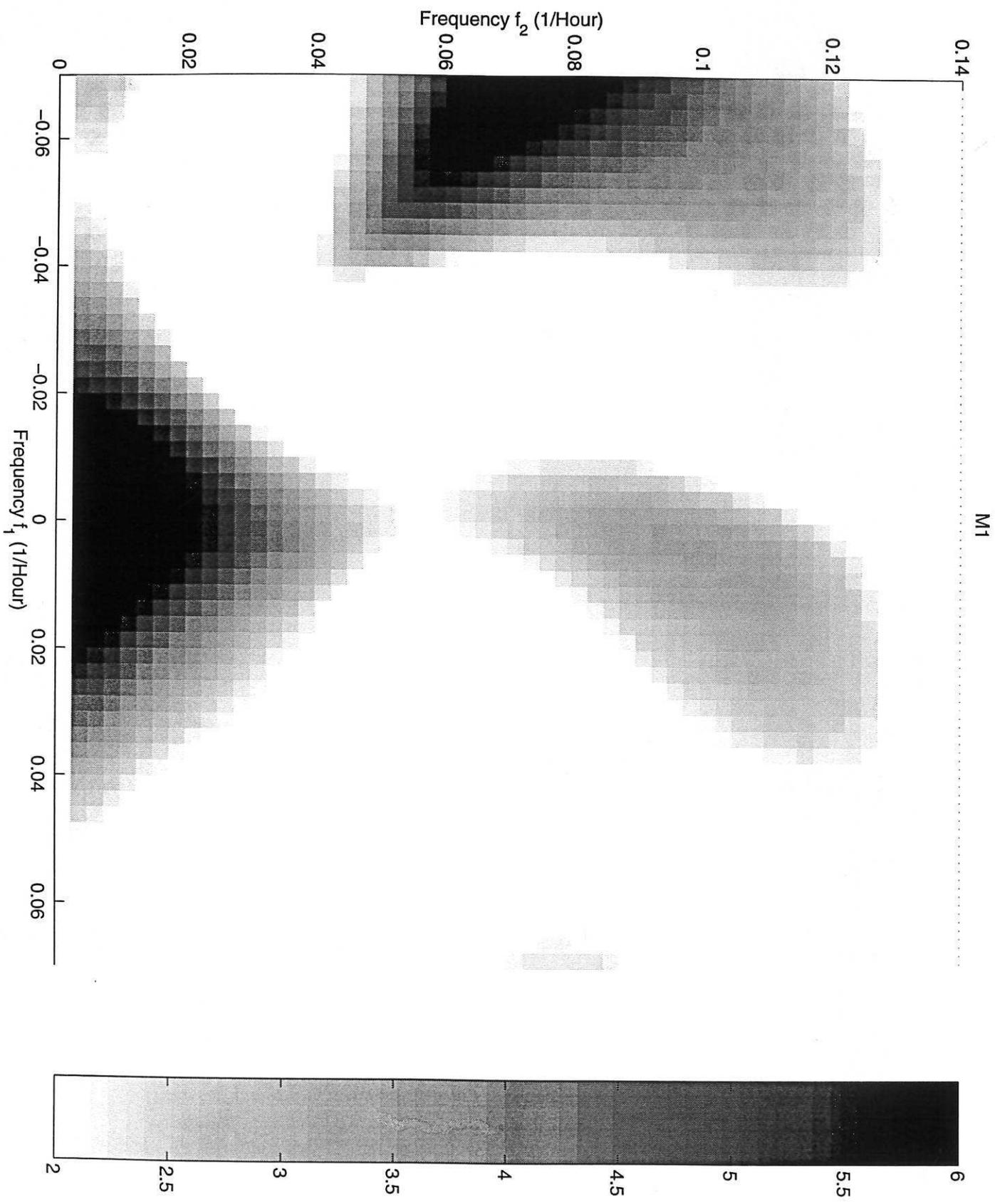
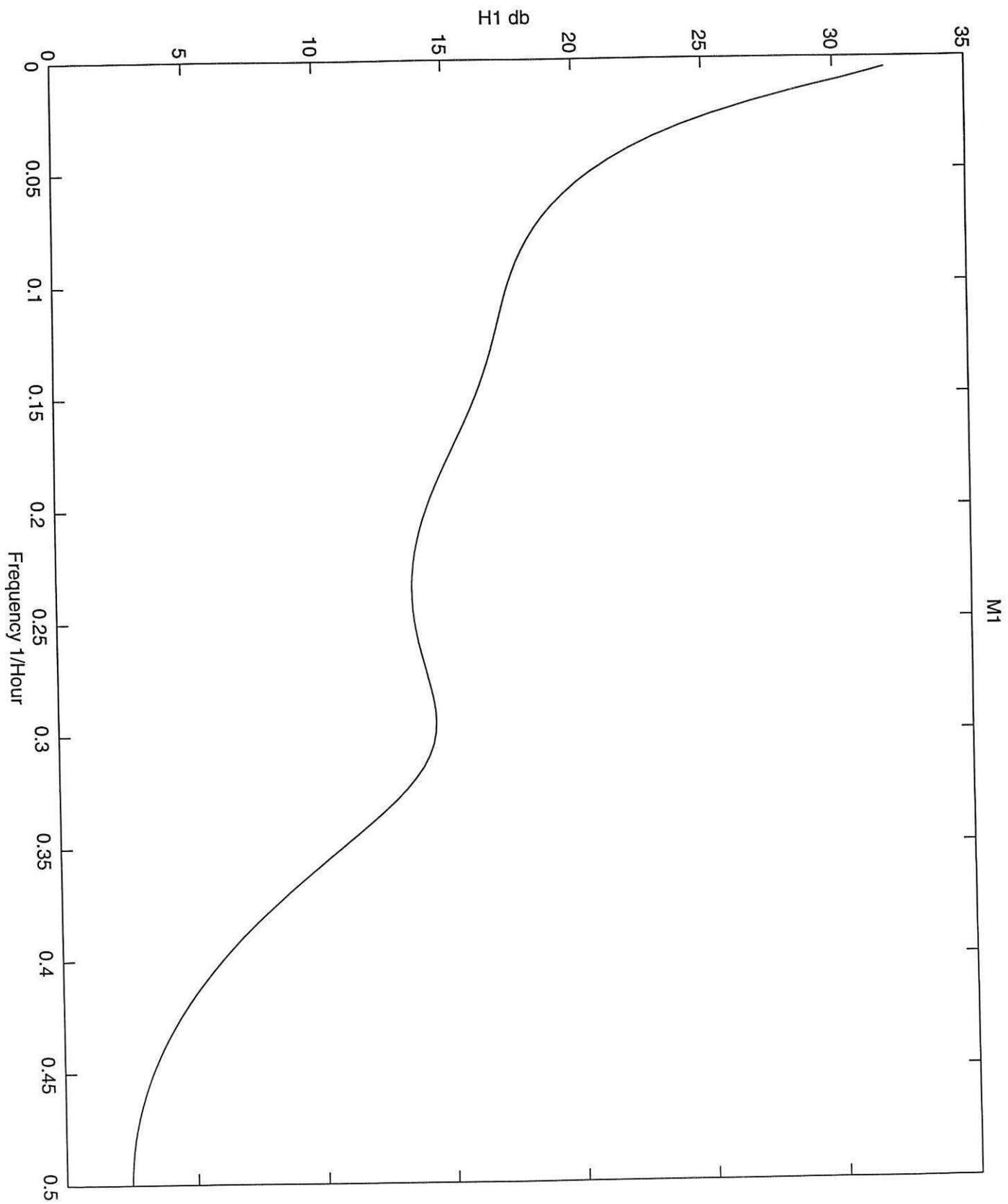
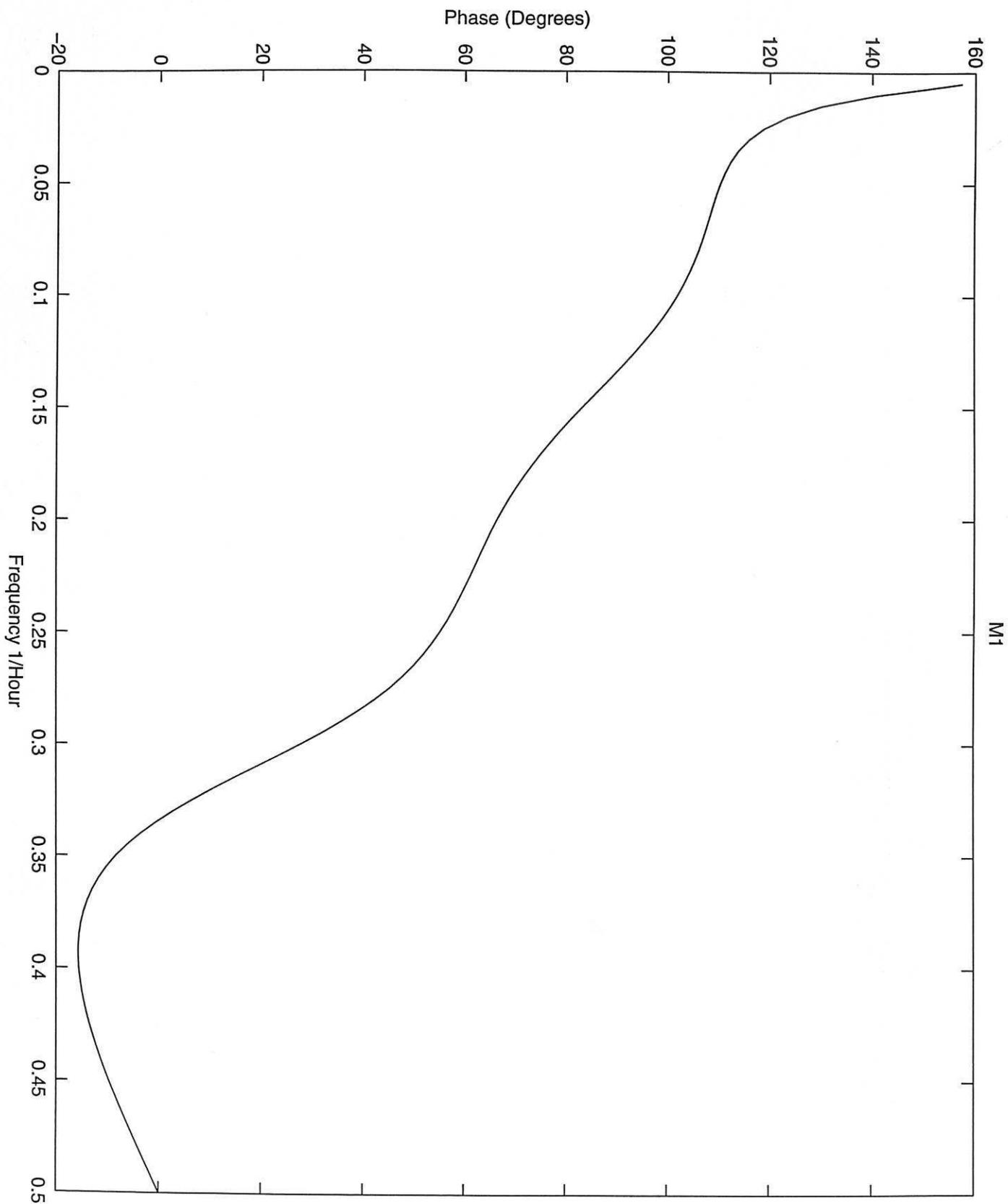


Fig. 5



M1

Fig. 6a



M1

Fig. 6d

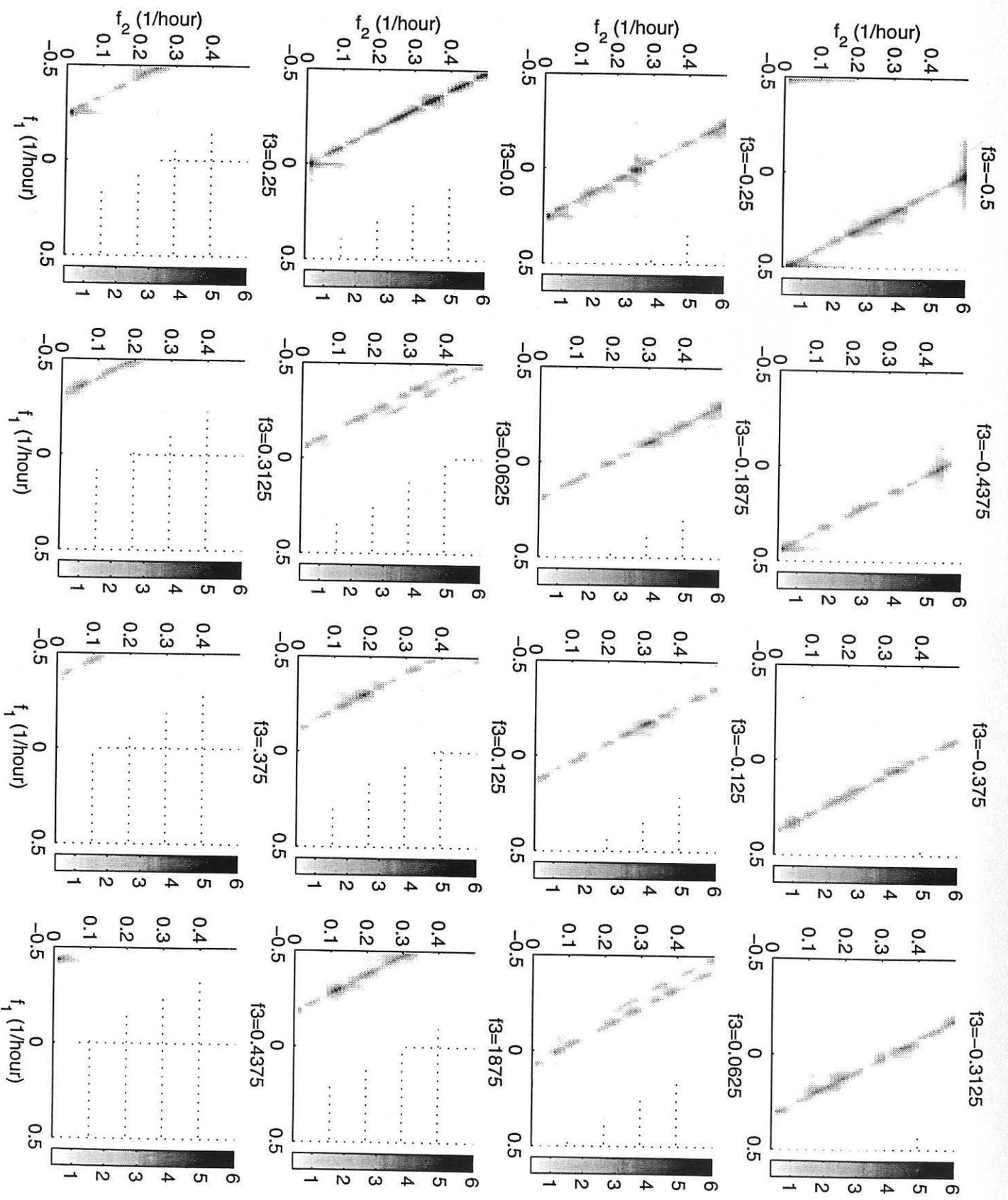
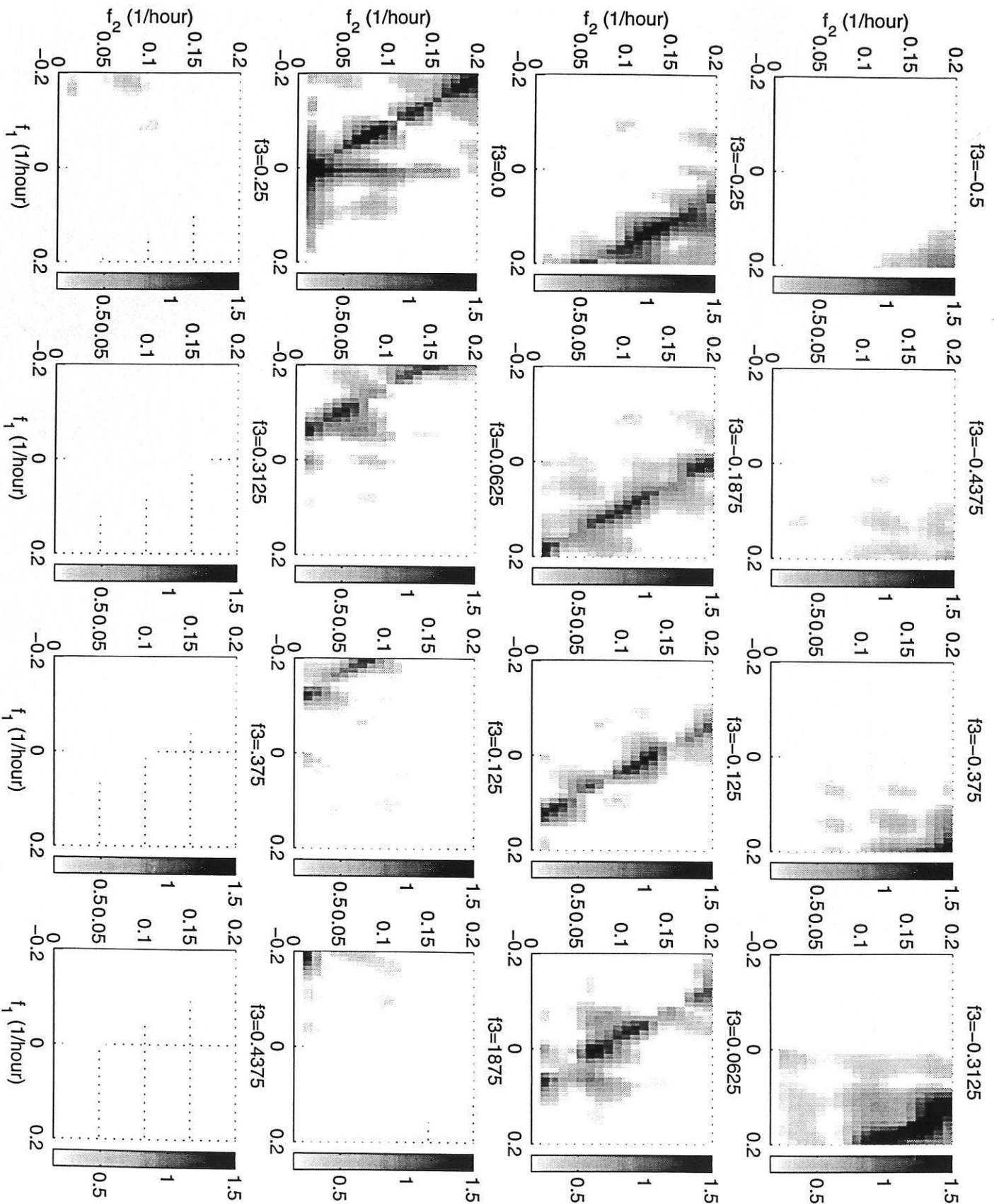
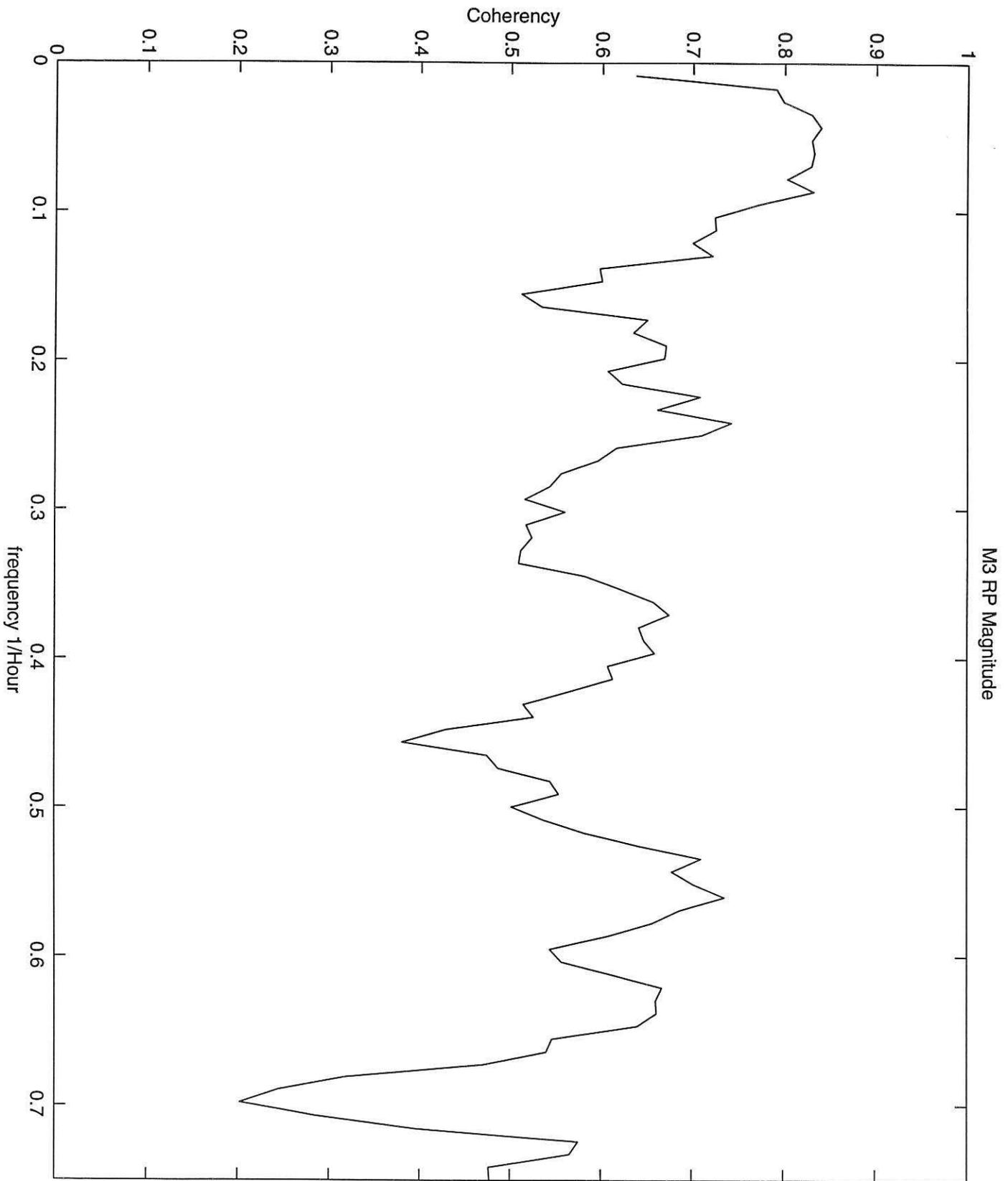


Fig. 7a



Handwritten signature or initials.



M3 RP Magnitude

Fig. 8a

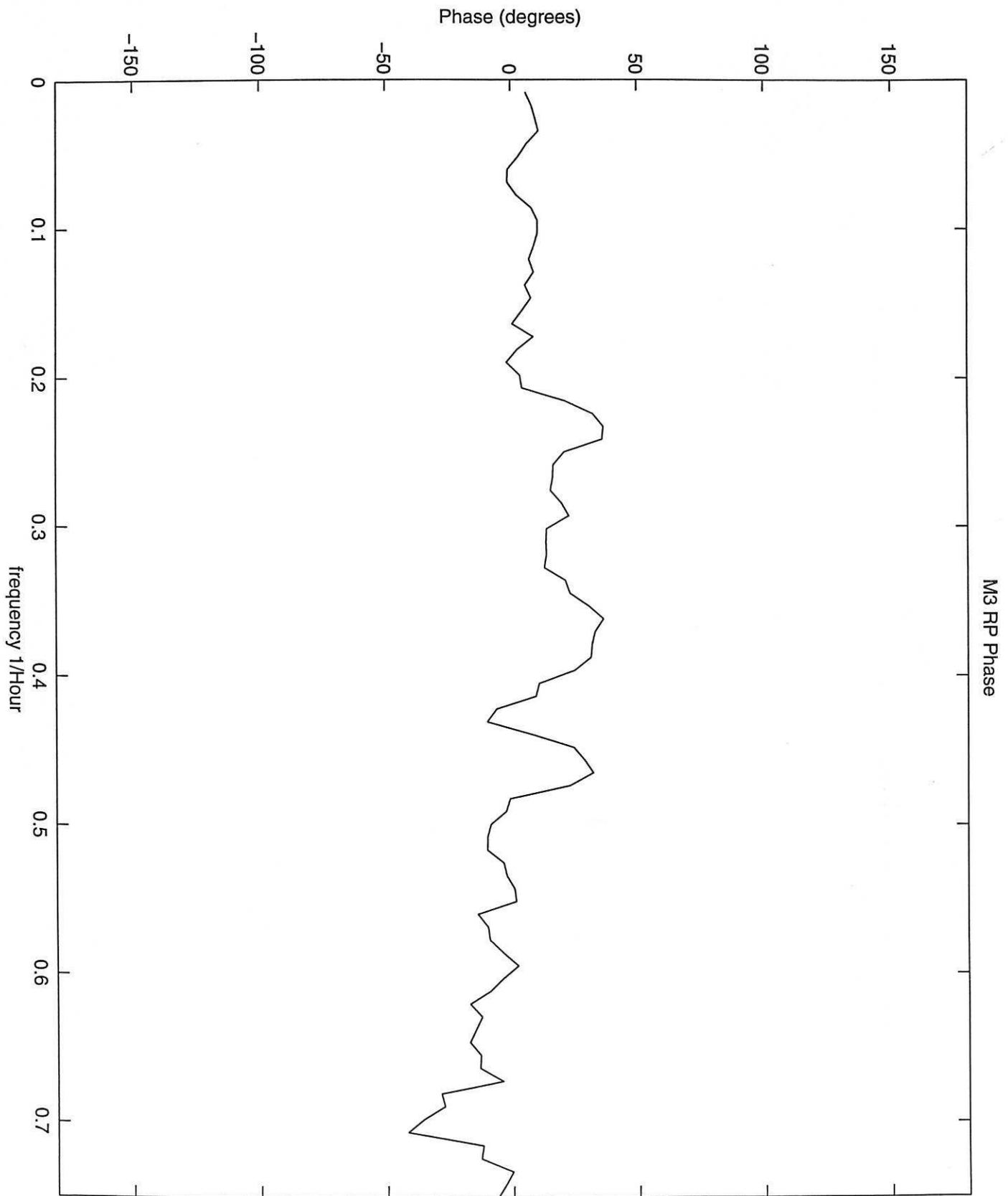
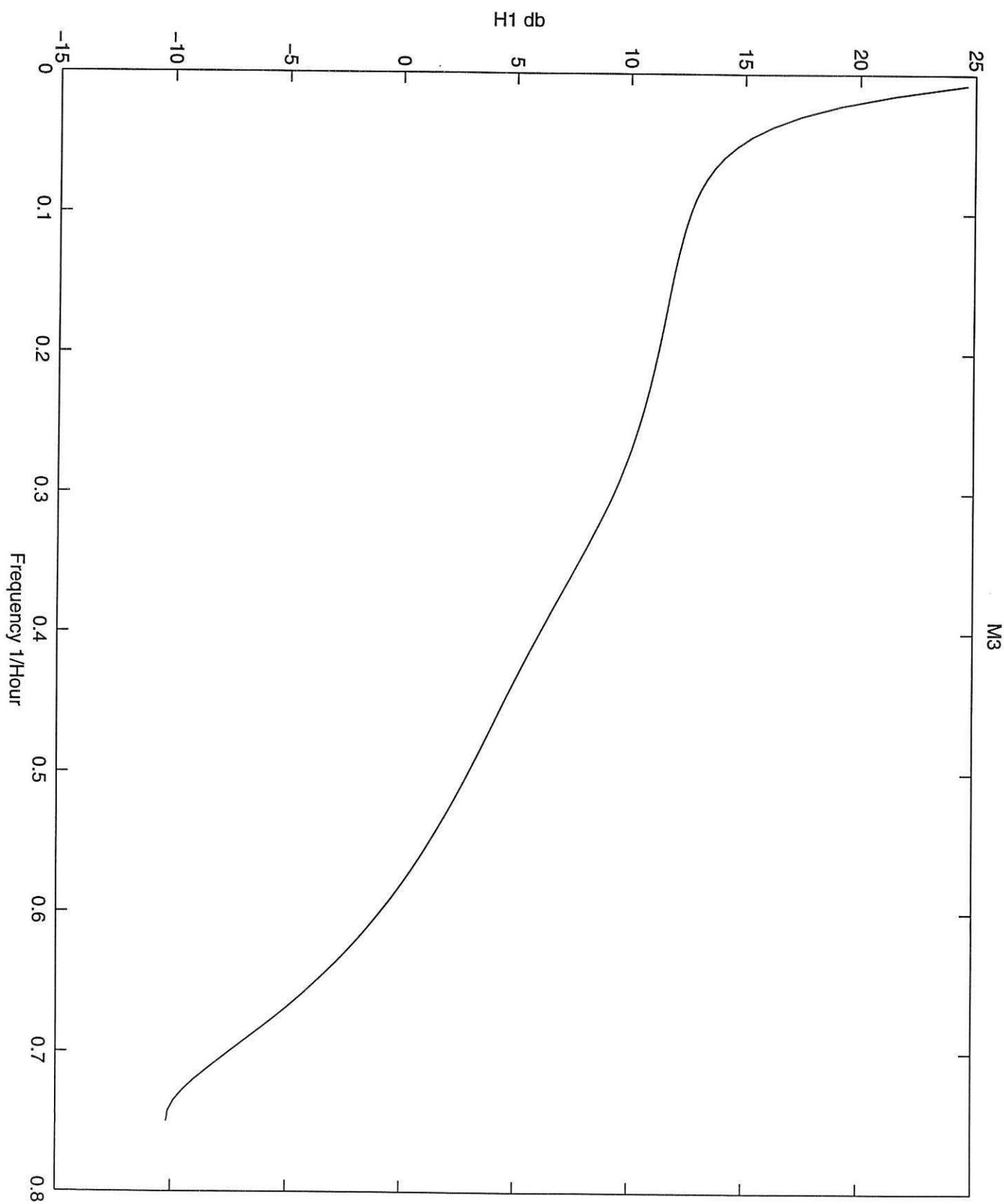
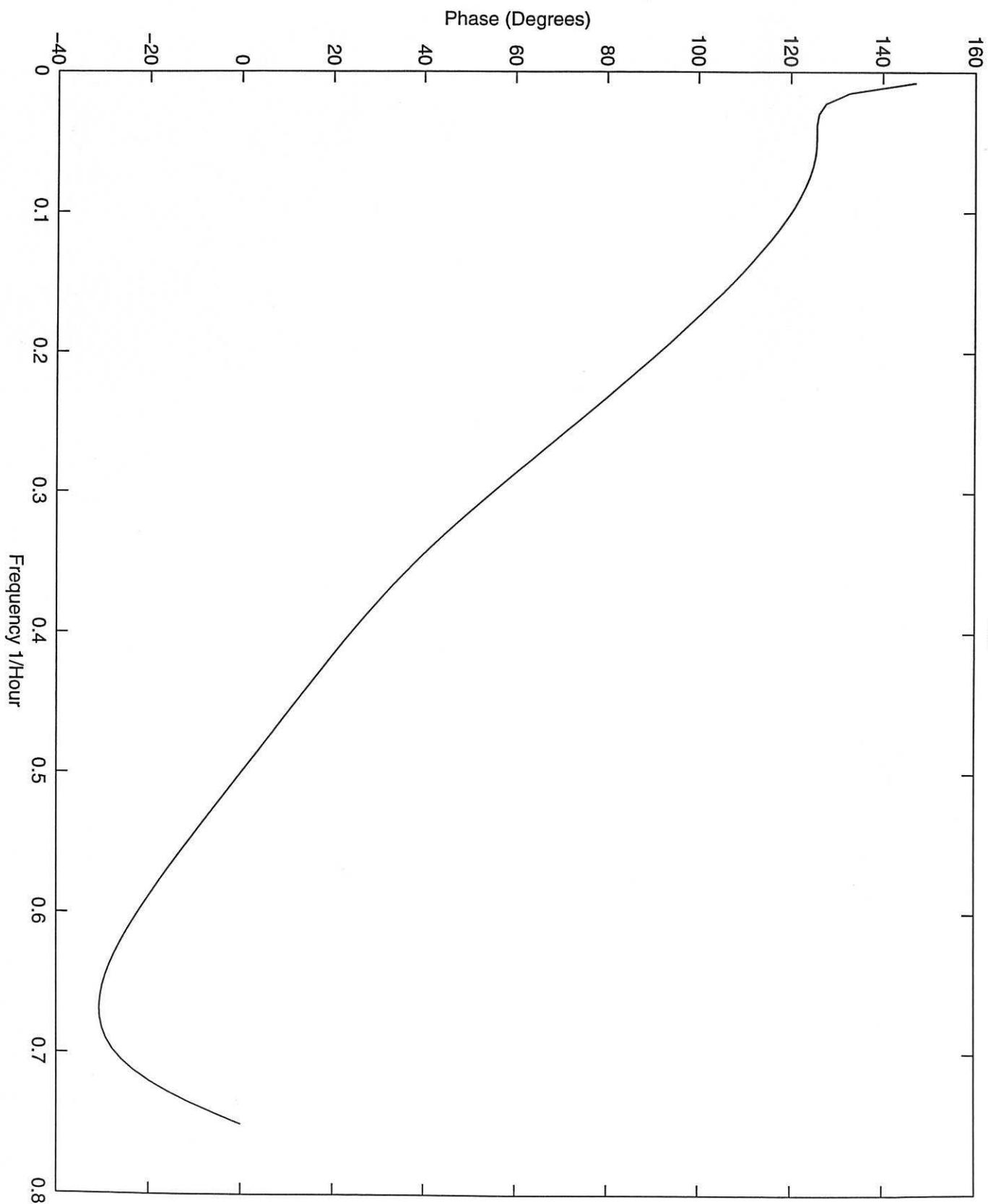


Fig. 8b



M3

Fig. 9a



M3

Fig. 9b

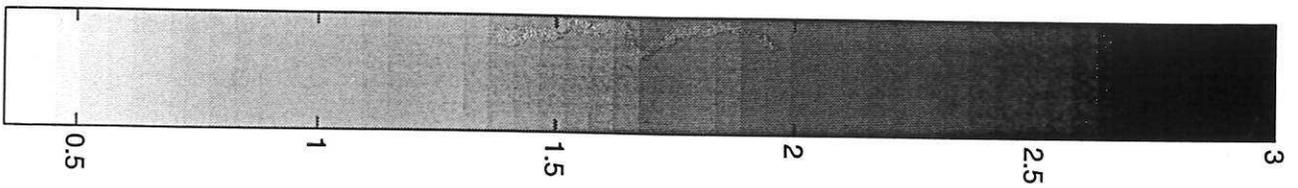
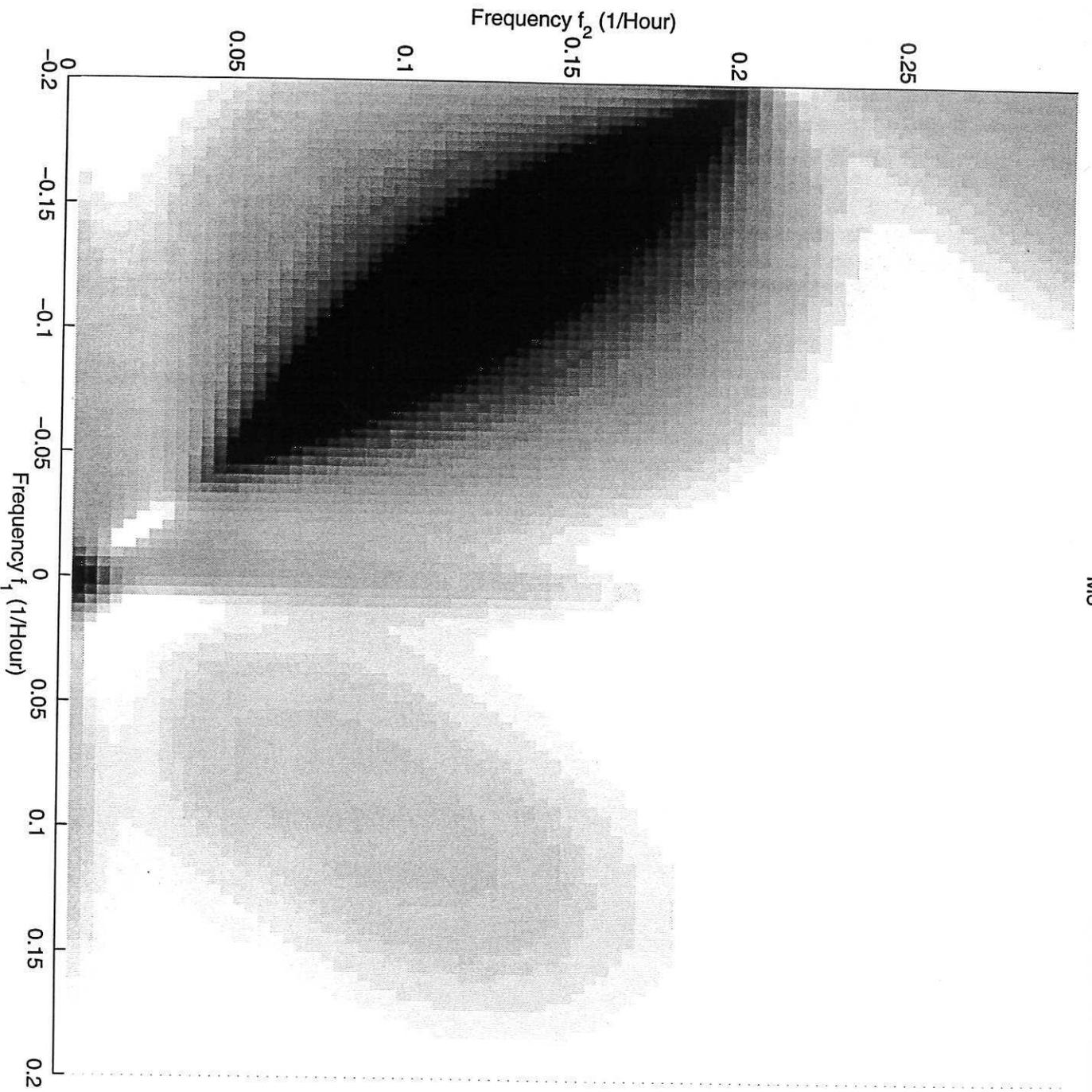


Fig. 10

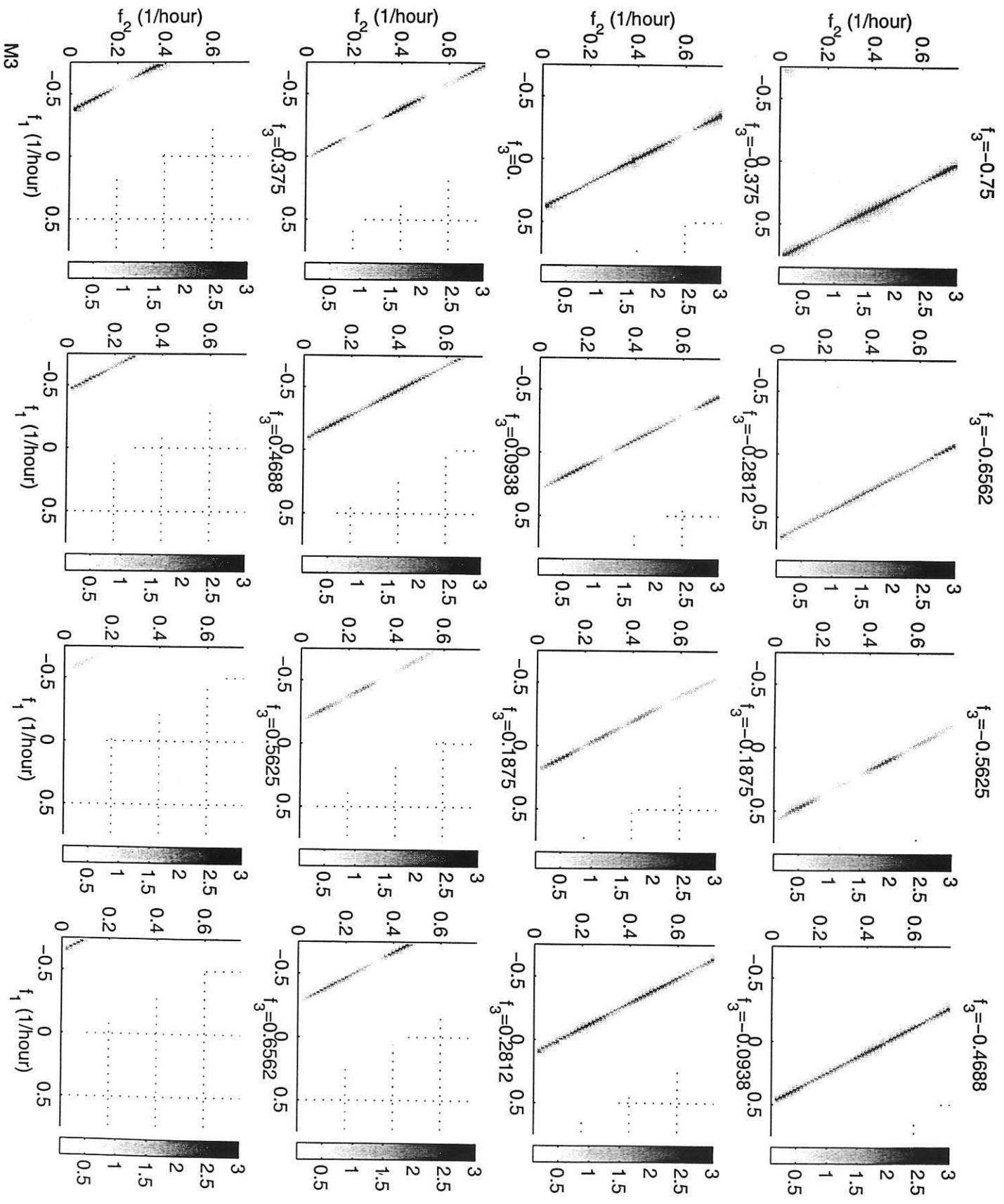


Fig. 11a

$f_3 = -0.75$ $f_3 = -0.6562$ $f_3 = -0.5625$ $f_3 = -0.4688$

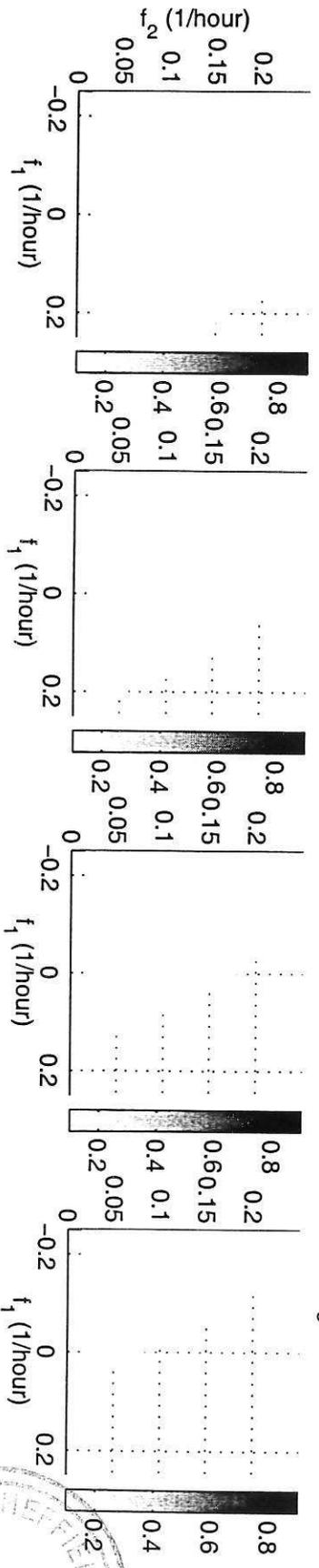
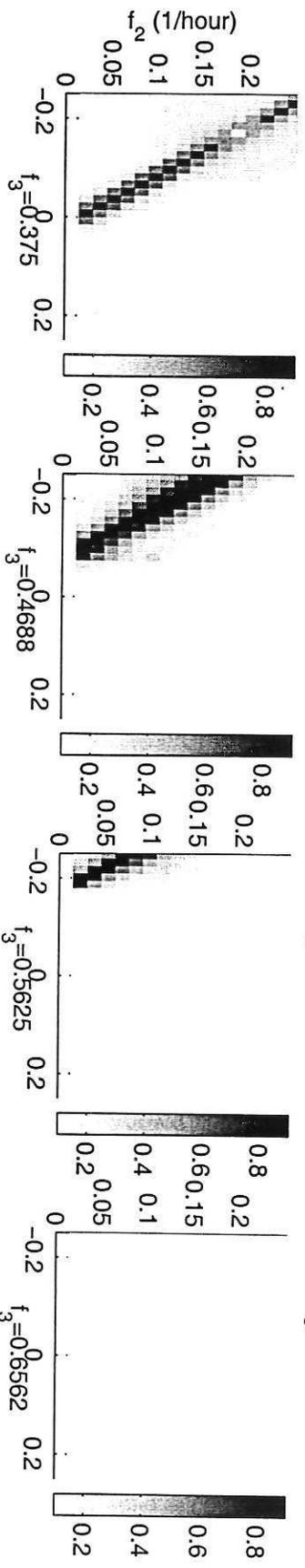
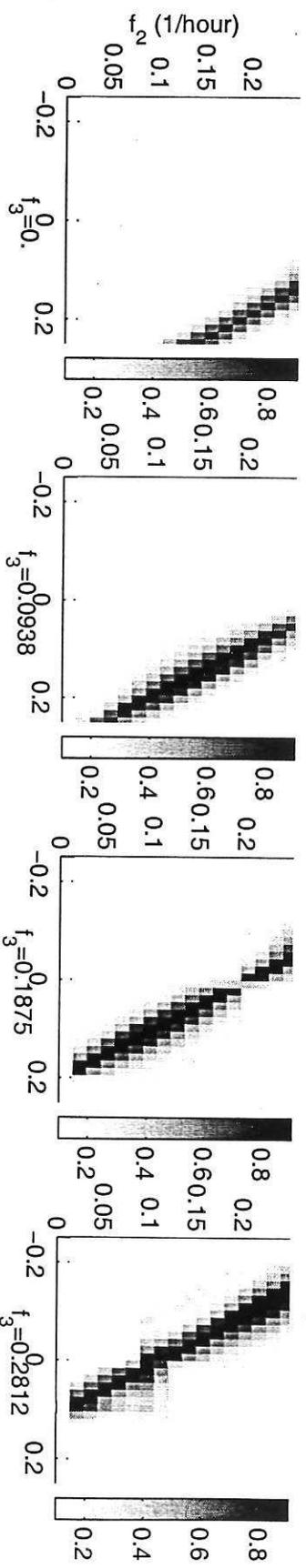
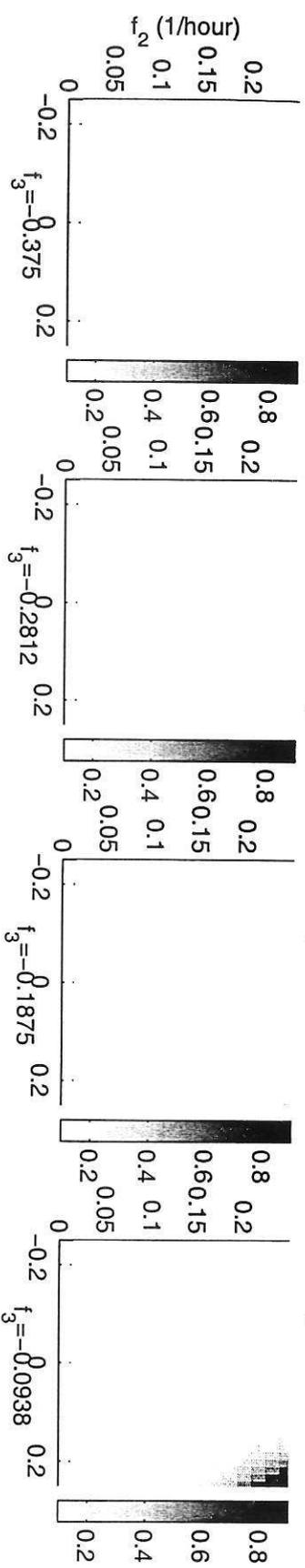


Fig. 11.2

