

Using Empirical Mode Decomposition to Process Marine Magnetotelluric Data

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Jin Chen

Kiel, Dezember 2013

Referent:	Dr. M. Jegen-Kulcsar
Koreferent:	Prof. Dr. W. Rabbel
Tag der mündlichen Prüfung:	21. February 2014
Zum Druck genehmigt:	21. February 2014

Der Dekan

Abstract

One of the most important steps in magnetotellurics (MT) studies is the derivation of frequency-based impedances from the measured time-varying electromagnetic fields. Accurate estimations of impedances, which depends on the quality of the collected data sets, can greatly improve the reliability of the subsurface resistivity structure model. Most current MT data processing methods use Fourier transform based procedures to estimate power spectral densities and, therefore, assume that the signals are stationary over the analyzed time segments. However, stationarity in geomagnetic data is not always ensured given the variety of source mechanisms causing the geomagnetic variations at different time and spatial scales. Additional complication and bias may arise from the presence of noise in the recorded electric and magnetic fields, which is particularly evident in marine MT measurements.

I explore a new heuristic method for dealing with the non-stationarity of MT time series based on Empirical Mode Decomposition (EMD). It is a dynamic time series analysis method, in which time series containing stationary and / or non-stationary components can be decomposed into a finite and small number of Intrinsic Mode Functions (IMF). Physical meaningful instantaneous information (e.g. frequencies and amplitudes) can be derived from each IMFs. In this thesis it is the first time that the EMD method is used in combination with instantaneous spectra to determine impedance tensor estimates. The estimation is performed in the time-frequency domain using instantaneous spectra rather than in pure frequency domain as in traditional processing procedures. The procedure is tested on synthetic and field MT data.

In synthetic tests the reliability of impedance estimates from EMD is compared to the synthetic responses of a 1D layered model. While the derived apparent resistivities and phases appear to be well recovered, a small systematic downward bias relative to the synthetic response values is observed. Through an analysis of the complete EMD procedure, I conclude that the bias is associated with numerical noise that is introduced by the decomposition used in EMD. Furthermore I find that this bias can be strongly reduced by either applying a second EMD algorithm to create a numerical reference time series of the magnetic field or by using a reference magnetic field time series measurements. To examine how impedance estimates are affected by noise, stochastic stationary and non-stationary noise are added on the time series. Comparisons reveal that estimates by EMD are generally more stable than those by simple Fourier analysis. At present stage, no robust estimate procedures are implemented in the EMD approach. However, I also compare the results derived from EMD to those derived by a commonly used data processing software (BIRRP), which incorporates additional sophisticated robust estimations to deal with noise issues. It is revealed that the results from both methods are already comparable.

The processing scheme is then also applied to marine MT field data, which were measured at

the Costa Rican subduction zone under the framework of collaborative research center SFB 574. Firstly, the non-stationary characteristics of motion induced noise is analyzed. Through analyzing the instantaneous spectra of the noisy section it can be shown that the motion induced noise is present over the whole frequency range. However, the effect of this noise decreases with the decreasing frequency. Further testing is performed on short, relatively quiet segments of several data sets, which exhibit different noise levels. Compared to a regular processing algorithm (BIRRP), the new method gives similar impedance estimates, but the impedance estimates are extended to lower frequencies and less noise biased estimates with smaller error bars are obtained at high frequencies. Finally, the EMD method is applied to long segments of data with many non-stationary noise packages. The estimates results derived from EMD are generally stable and less biased at all frequencies. In the low frequency range, results are comparable to those derived from BIRRP, and at high frequencies the results have smaller error bars.

As a conclusion, non-stationary noise can heavily affect Fourier-based MT data processing. I show that by deriving impedance estimates based on the EMD method, a new methodology especially designed to handle non-stationary data, comparable or better impedance estimates over a larger frequency range can be derived without employing any sophisticated estimation procedures, which usually need to be invoked in FT based processing. Especially for marine applications, where compared to land data non-stationary noise is more pronounced and where the signal to noise ratio is reduced at high frequencies due to absorption by the overlying ocean layer, the new processing methodology I developed represents an important step towards deriving a better resolved Earth model to greater depth underneath the seafloor.

Zusammenfassung

Einer der wichtigsten Schritte in Magnetotellurik (MT)-Studien besteht aus der Herleitung der frequenzabhängigen Impedanzen aus den gemessenen Zeitreihen. Die Genauigkeit der Impedanzabschätzung, welche von der Qualität der gemessenen Zeitreihen abhängt, bestimmt die Genauigkeit und Auflösung des Widerstandmodells des Untergrundes. Gängige Methoden der MT-Datenverarbeitung verwenden die Fouriertransformation um spektrale Energiespektren zu berechnen. Hierbei wird für die analysierten Zeitsegmente angenommen, dass die Signale stationär sind. Diese Stationarität ist aber nicht immer gewährleistet, da diverse Quellmechanismen des Erdmagnetfeldes zeitlich und räumlich variierende Quellfelder erzeugen. Rauschen in den elektrischen und magnetischen Messungen, das besonders bei marinen Messungen signifikant sein kann, stellt ein zusätzliches Problem dar.

Mit der Empirical Mode Decomposition (EMD) wird eine neue, heuristische Methode untersucht, um mit der fehlenden Stationarität von MT-Zeitreihen umzugehen. Die EMD ist eine dynamische Zeitreihenanalyse, bei der Zeitreihen mit stationären und nicht-stationären Anteilen in eine finite und kleine Anzahl von Intrinsic Mode Functions (IMF) zerlegt werden. Von diesen IMFs können physikalisch aussagekräftige, instantane Informationen (z.B. Frequenzen, Amplituden) abgeleitet werden. In dieser Arbeit wird die EMD-Methode in Kombination mit den instantanen Spektren verwendet, um die Impedanztensoren zu berechnen. Im Gegensatz zur traditionellen Prozessierung im Frequenzbereich, werden die Berechnungen hier mittels instantaner Spektren im Zeit-Frequenz Bereich durchgeführt. Die generelle Eignung der Methode wird sowohl anhand synthetischen Daten als auch an Feldmessungen untersucht.

Bei den synthetischen Tests wird zunächst die Plausibilität der mittels EMD berechneten Impedanzen durch Vergleiche mit analytischen Lösungen getestet. Obwohl die ermittelten scheinbaren Widerstände und Phasen plausibel sind, zeigt sich im Vergleich zu den analytischen Lösungen ein Bias. Dieser wird durch numerisches Rauschen der Zerlegung erklärt. Zur Reduzierung des Bias können Referenzzeitreihen verwendet werden. Diese können entweder durch Verwendung eines zweiten EMD Algorithmuses numerisch rekonstruiert werden oder aus einer zweiten gemessenen Zeitreihe des Magnetfeldes stammen. Des weiteren wird untersucht, inwieweit die berechneten Impedanzen durch stochastisch stationäres und nicht-stationäres Rauschen beeinflusst werden. Vergleiche der mit der EMD Methode berechneten Ergebnisse mit der Fouriertransformation zeigen, dass die Ergebnisse der EMD Methode stabiler sind als die der Fouriertransformation. Obwohl robuste Algorithmen zur Datenprozessierung im EMD Ansatz noch nicht implementiert wurden, ist die Qualität der Ergebnisse mit denen einer spezialisierten Prozessierungssoftware (BIRRP) vergleichbar, die solche robuste Algorithmen zur Bewältigung von Rauschen verwendet.

Als Anwendungsbeispiel wird die neuartige Prozessierung von marinen MT Daten gezeigt, welche an der Costa Rica Subduktionszone im Rahmen des Sonderforschungsbereichs SFB 574 gesammelt wurden. Als erstes wird das durch die Bewegung des Meßsystems verursachte, nicht-stationäre Rauschen der Daten analysiert. Die verrauschten Sektionen des instantanen Spektrums zeigen, dass das Bewegungsrauschen im kompletten Frequenzspektrum beobachtet werden kann, jedoch lässt dieser Effekt mit abnehmender Frequenz nach. Für verschiedenen Datensätze wird die neuartige Prozessierung dann mit einer spezialisierten Prozessierungssoftware (BIRRP) für kurze, ungestörte Zeitabschnitte verglichen. Beide Methoden liefern ähnliche Impedanzen, die neuartige Prozessierung erweitert aber den nutzbaren Frequenzbereich zu tieferen Frequenzen und liefert bei hohen Frequenzen kleinere Fehlerbalken. Zuletzt wird die EMD Methode auf lange Datensegmente mit vielen nicht-stationären, verrauschten Abschnitten angewendet. Auch hier sind die mit der EMD berechneten Ergebnisse bei allen Frequenzen generell stabiler und weniger von Rauschen beeinflusst.

Schlussfolgernd zeigt sich, dass nicht-stationäres Rauschen die klassische Fourierbasierte MT Datenprozessierung stark beeinflussen kann. Impedanzwerte, die durch EMD, eine Methode die speziell auf nicht stationäre Daten zugeschnitten ist, hergeleitet werden, liefern gleichgute oder bessere Impedanzwerte über einen größeren Frequenzbereich und das ohne statistisch anspruchsvolle Abschätzungen zu benutzen, wie sie bei der Fourierbasierten Auswertung herangezogen werden müssen. Die von mir entwickelte EMD Datenprozessierungsmethode eignet sich damit besonders gut für die Verarbeitung von marinen Daten. Denn im Vergleich zu Landdaten sind marine Daten durch größeres, nicht-stationäres Rauschen und kleinere Signalrauschverhältnisse, besonders bei hohen Frequenzen durch die Absorption des aufliegenden Ozeans, charakterisiert. Durch die hier vorgeschlagene verbesserte Datenverarbeitung kann somit in Zukunft ein besser aufgelöstes und tieferes Erdmodell unterhalb des Meeresbodens hergeleitet werden.

摘要

在大地电磁(magnetotelluric (MT))研究方法中, 其中一个非常重要的步骤是从所测量的电磁时间序列得到随频率变化的阻抗(impedance)。准确估算的阻抗能大大提高地下电阻率结构模型的可靠度, 而能否准确估算取决于所采集的数据质量和所使用的方法。目前大多数数据处理的方法是使用傅立叶变换来估算功率谱, 因而假定被选取的数据段都是平稳的。然而, 因为大地电磁信号源的机制原理的多样性导致地磁数据也随时间、地点而变化, 所以平稳性在地磁数据中不是总能满足的。另外, 当所测的电场、磁场数据受到噪声污染时会进一步增加数据处理的难度和导致估算有所偏差, 这在处理海洋 MT 数据时尤为突出。

针对大地电磁数据的非平稳性, 本论文采用了一种基于经验模态分解(EMD)的新的探索式的方法来处理 MT 数据。经验模态分解是一种动态的时间序列处理方法, 能将一个无论是平稳的还是非平稳的时间序列, 分解成有限的固有模态函数(IMF)。而从每个固有模态函数能够得到时间序列具有物理意义的瞬时信息(如频率、振幅等)。在此论文中, 第一次尝试使用经验模态分解方法结合所得的瞬时频谱, 来估算阻抗张量。与传统的在单纯的频率域来估算阻抗的过程不同, 新方法是在时间-频率域进行的。新的估算方法在模拟数据和真实数据中进行了测试。

在使用模拟数据的测试中, 为了评估新方法的可靠性, 将估算的阻抗结果和一维层状模型的解析结果做了比较。由新方法所得的视电阻率和相位基本上是令人满意的, 但相对模型的解析结果, 视电阻率整体上有小小的向下的偏差。经过分析测试, 偏差的原因是来自经验模态分解方法本身。此偏差可以通过使用参照的磁场数据来大大减小。如没有参照数据可用另一种经验模态分解算法来分解磁场数据得到“参照数据”, 也能达到同样的效果。更进一步, 在模拟数据被随机平稳噪声和非平稳噪声污染的情况下, 对新的阻抗估算方法进行了测试。测试结果显示, 用经验模态分解方法得到的结果比用简单的傅立叶变换得到的结果好并且更稳定。虽然到目前为止在新方法中还没用采用任何鲁棒(robust)的估算方法, 但我们还是把所得的结果和用通常使用的数据处理软件(BIRRP)所得的结果进行了比较, 而在 BIRRP 软件里有针对处理噪声的一系列的复杂的鲁棒算法。比较的结果显示用两种方法处理所得的估算已经非常的接近。

在真实数据的测试中, 使用的是采集于哥斯达黎加俯冲带的海洋 MT 数据。该数据是在合作项目 SFB 574 的框架下采集的。首先, 对数据中由于仪器摆动而引起的噪声的非平稳特征进行了分析。通过对含噪声的信号片断的瞬时频谱的分析, 可以看到因仪器摆动而引起的噪声分布在所有的频率段, 然而噪声在各个频段的影响随着频率的变小而减少。随后, 针对受不同程度噪声污染的几个台站的数据, 节取了较短、相对平静的数据片断, 用新的方法进行了处理。和使用常规的处理软件(BIRRP)的结果相比, 用新方法也得到了相近的阻抗估算, 并且用新方法在低频范围得到了更多的阻抗估算, 而在高频范围得到的估算受噪声影响的偏差也相对较小且有较小的误差棒。最后, 用 EMD 方法对含有很多非平稳噪声片断的长段信号进行了处理。所得的估算结果在所有频率段受噪声的影响都较小。在低频范围, 所得结果比较稳定, 和使用 BIRRP 得到的结果接近。在高频范围, 新方法所得的结果有较小的误差棒。

本论文得到的结论是, 即便是使用了复杂的鲁棒算法, 非平稳的噪声污染还是会严重影响基于

傅立叶变换的 MT 数据处理方法的结果。而 EMD 方法是一种新的特别适合处理非平稳数据的方法，用基于 EMD 的处理方法，即使没有使用鲁棒算法，所得到的阻抗估算不仅拓宽了估算的频率范围而且和用鲁棒方法所得结果相近甚至更好。新方法在海洋大地电磁的研究应用中尤为重要，因为和陆地研究不同，非平稳噪声在海洋研究中更为显著，并且由于海水的吸收，在高频段海洋电磁信号的信噪比非常低。新的数据处理方法在构建具有更好解析度、更深的海底地球模型上迈出了重要的一步。

Contents

1	Introduction	1
2	Introduction of Magnetotelluric Method	5
2.1	Electrical resistivity of Earth materials	5
2.2	Magnetotellurics: Natural source EM method	6
2.3	Applications of MT in geoscience	8
3	Maxwell's Equations and Transfer Function	11
3.1	Maxwell's equations	11
3.2	1D homogeneous half space	14
3.3	1D layered Earth response	17
3.4	Simple 2-D Earth and E- and B-polarization	20
3.5	3D case	22
4	From Measured Time Series to Impedance Tensor	23
4.1	Short review of MT data processing	23
4.2	Spectral estimation	26
4.3	Least square approach	27
4.4	Remote reference method	29
4.5	Robust estimation I: M-estimator	30
4.6	Robust estimation II: Bounded influence estimation	32
4.7	Error estimate	34
4.7.1	Parametric error estimate	34
4.7.2	Non-parametric error estimate	35
4.8	Bounded influence remote reference processing	38
4.9	Estimator in time-frequency domain	39
4.10	Short summary	41
5	EMD Method and Instantaneous Spectrum	43
5.1	Motivation	43
5.2	Empirical mode decomposition	46
5.3	Instantaneous frequency and spectrum	58
5.3.1	Concept of instantaneous frequency	58
5.3.2	IF based on Hilbert transform	60
5.3.3	IF based on direct quadrature	70
5.4	Is an intrinsic mode function a mono-component signal?	73
5.5	Instantaneous spectrum and Fourier spectrum	76
5.6	Some limitations of the EMD method	87

5.6.1	Effects from end point extension	90
5.6.2	Effects from interpolation method	95
5.6.3	Effects from stop criterion	104
5.7	Short summary	105
6	Using EMD to Process Synthetic Data	107
6.1	1-D Earth model and synthetic time series	107
6.2	Impedance and the error estimates	114
6.3	The bias problem and its solution	120
6.4	Noise test	124
6.4.1	Adding stationary colored noise	124
6.4.2	Adding non-stationary noise	131
7	Using EMD to Process Field Data	137
7.1	Fluids in subduction zones and SFB 574	137
7.2	MT experiment in the Costa Rican subduction Zone	138
7.2.1	MT profile	138
7.2.2	Data quality	141
7.2.3	Preprocessing of MT time series	144
7.3	Processing the quiet segments of data	146
7.4	Noise effect analysis and processing the noisy data	155
7.5	Short summary and outlook	161
8	Summary - Conclusion - Outlook	167

List of Figures

2.1	Magnetic field of the Earth distorted by the solar wind. (Figure comes from: http://phyworld.ed.hkedcity.net/iq/aurora/magnetopause_e.gif)	7
3.1	The homogeneous half space.	15
3.2	The 1D n-layered half space.	18
3.3	Left panel: A horizontally layered model with three layers. Right panel: The corresponding apparent resistivity & phase curves of the model.	20
3.4	Simple 2D model composed of two vertical contacted zones with different conductivities σ_1 and σ_2 . The striking is oriented in x-direction.	21
4.1	The procedure to construct a complex transient signal $x_f(t)$ around a desired estimation frequency ω_0 from a real signal $x(t)$. The real and imaginary parts of $x_f(t)$ are the narrow band filtered signal of $x(t)$ and its Hilbert transform, respectively.	40
5.1	Example of three day long period of noisy marine MT measurements collected at Costa Rican subduction zone in 2007. All E- and B-field components and tilt movements are shown. It is obvious that both E - and B -field data are contaminated by noise associated with the tilt movements.	44
5.2	Illustration of the "sifting" process. One signal can always be decomposed into a fast oscillate signal plus a sum of slow oscillate signals. In this example only two oscillations are present.	47
5.3	Synthetic harmonic time series of example 1 (first panel) and its three constituent components (panel 2 to 4).	49
5.4	This figure illustrates how the first sifting iteration of the EMD method is performed on the times series shown in Fig. 5.3. (Here only half of the total time interval is shown). In the first panel the original signal and its local maxima and minima (black and green dots) are plotted. From the extrema the upper and lower envelope are determined (black and green line) that are used to calculate the local mean values (red line) required for the first sifting process. The second panel shows the signal after the first iteration (namely h_1 in Eq. 5.1) which will be used as the input signal for next sifting iteration. The third panel shows the residue (namely m_1 in Eq. 5.1) that corresponds to the local mean value in the first panel.	51
5.5	The six IMFs (blue curves) obtained for the signal of example 1 (see first panel in Fig. 5.3). The first three IMFs in principal represent the three constituent components of the original signal (red curves, also see panel 2 - 4 in Fig. 5.3), and the remaining three IMFs, which have relatively lower amplitudes, are related to some numerical inaccuracies like e.g. boundary problems.	52

5.6	Two sinusoidal frequency-modulated (FM) waves (panel 1 - 2) and a Gaussian wavepacket (panel 3). These signals are combined to construct the synthetic time series of example 2 (see upper panel of Fig. 5.7).	53
5.7	Synthetic time series of the example 2 (upper panel) and the frequency & amplitude contents of its three components (lower panel) that are plotted in Fig. 5.6. The color coding is associated with the (maximum) amplitudes.	54
5.8	Fourier power spectra of the three components of the synthetic time series shown in Fig. 5.6. (To obtain these spectra a discrete prolate spheroidal sequence with a time-bandwidth of 1 as tapering window is applied on overlapping sections.)	55
5.9	The first six IMFs (blue curves) of the synthetic time series of example 2. For comparison, the three original components are also plotted (red curves) in the panels with IMF 1 - IMF 3.	56
5.10	IFs calculated by using Eq. 5.14 for three harmonic sinusoidal signals (Eq. 5.24) with different constant amplitude shift.	62
5.11	Upper panel: The synthetic time series from example 1 (blue curve) and its Hilbert transform (red curve). Lower panel: The IFs computed directly from the analytic signal (without decomposing the signal with EMD beforehand).	64
5.12	Upper panel: The synthetic time series from example 2 (blue curve) and its Hilbert transform (red curve). Lower panel: The IFs computed directly from the analytic signal (without decomposing the signal with EMD beforehand).	65
5.13	The first three IMFs (blue curves) of the synthetic time series of example 1 and their Hilbert transform (red curves).	67
5.14	The instantaneous spectra of the first three IMFs of the synthetic time series of example 1 (see Fig. 5.13). Color coding indicates the amplitude level.	68
5.15	The instantaneous spectra of the first three IMFs of the synthetic time series of example 2 (the true instantaneous frequency & amplitude variations with time are shown in lower panel of Fig. 5.7). Color coding indicates the amplitude level. . . .	69
5.16	Illustration of the demodulation process. Upper panel: The IMF 1 (blue curve) of the synthetic time series of example 2, the absolute value of this IMF 1 (red dashed curve) and two envelopes calculated from this IMF. One envelope is obtained from HT-based analytic signal $ u(t) $ (black curve), the other is the first envelope $e_1(t)$ (see Eq. 5.35) derived from the demodulation procedure (magenta curve). Lower panel: The IMF 1 before (blue curve) and after (red dashed curve) normalization with $e_1(t)$. Note that both figures are zoomed in on a small time interval from 50 to 150 s.	74
5.17	Upper and lower panels show the instantaneous spectrum of the first three IMFs of the synthetic time series of example 1 and 2, respectively. In both figures the color coding indicates the amplitude level. As a comparison IFs which are calculated by using HT-based analytic signal are plotted as thin grey curves. For these instantaneous spectra no amplitude information are given (see Figs. 5.14 and 5.15 for amplitude information).	75
5.18	Upper panel: The marginal spectrum of the synthetic time series of example 1. Lower panel: The Fourier amplitude spectrum of the synthetic time series of example 1.	77
5.19	Upper panel: The marginal spectrum of the synthetic time series of example 2. Lower panel: The Fourier amplitude spectrum of the synthetic time series of example 2.	78

5.20	Comparison of the Fourier and marginal power density spectra of the sinusoid signal $x(t) = \sin(\pi t/256)$ with $t = 0, 1, \dots, 1023$. The numbers given for the different marginal power densities correspond to the number of bins in frequency direction and, hence, indicate the frequency resolutions.	81
5.21	Comparison of the Fourier and marginal power density spectra of synthetic time series of example 1. The numbers given for the marginal power densities correspond to the numbers of equally sized bins used in frequency direction.	82
5.22	Comparison of the Fourier and marginal power density spectra of synthetic time series of example 2. The numbers given for marginal power densities correspond to the number of bins in frequency direction.	83
5.23	Upper panel: Two time series of a sinusoidal signal with lengths of $t = 1023$ s and $t = 800$ s. Lower panel: Comparison of the Fourier and marginal power density spectra of the sinusoidal signal for both time lengths. The same binning in frequency direction (512 bins) is used for the calculation of the spectra.	85
5.24	Comparison of the Fourier and marginal power density spectra of synthetic time series of example 1 for time intervals of 10 s and 8.8 s (sampling interval $\Delta t = 0.01$ s). The bin number in frequency direction (512 bins) is the same for both spectra.	86
5.25	Upper panel: The real part of the complex spectra of a sinusoidal signal for time intervals of 1023 s and 800 s. Lower panel: The corresponding imaginary parts of the same complex spectra.	88
5.26	Upper panel: The real part of the complex spectra of the synthetic time series of example 1 for time intervals of 1023 s and 800 s (sampling interval is $\Delta t = 0.01$ s). Lower panel: The corresponding imaginary parts of the same complex spectra.	89
5.27	First panel: The first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto the left boundary) using free boundaries (the first and last data points are not considered as extrema). Lower three panels: The first three IMFs of the same time series extracted by using free boundary interpolations.	91
5.28	First panel: The first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto the left boundary of the signal) using free boundaries (the first and last data points are considered as extrema). Lower three panels: The first three IMFs of the same time series extracted by using free boundary interpolations.	93
5.29	First panel: The first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto the left boundary of the time series) using the mirror strategy from Rilling et al. [2003] (the first and last data points are not considered as extrema). Lower three panels: The extracted first three IMFs of the same time series.	94
5.30	First panel: The first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto left boundary of the time series) using the strategy from Rato et al. [2008] (the first and last data points are not considered as extrema). Lower three panels: The extracted first three IMFs of the same time series.	96
5.31	Cubic spline interpolation of the upper & lower envelopes at the first iteration to extract the IMF 2 from synthetic time series of example 1 (zoomed onto a small time interval of the time series). Some of the intervals where the upper envelope (lower envelope) have smaller (larger) values as the actual signal are numbered A, B and C (D, E and F).	98

5.32	The upper envelopes are obtained by using three different interpolation methods (linear, piecewise cubic Hermite and cubic spline) at the first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto a small time interval of the time series). Note that all three methods have overshooting problems (see e.g. points A, B and C).	99
5.33	The second IMF (upper panel) and the third IMFs (middle panel) extracted by using three different interpolation methods (see Fig. 5.32 of the synthetic time series example 1) and their total residuals $x(t) - \sum_{i=1}^3 c_i(t)$ (last panel).	100
5.34	Sketch illustrating the parabolic strategy proposed by Rato et al. [2008] to modify the position and value of an extremum for sampled data.	101
5.35	Comparison of two upper envelopes in the first sifting iteration for the synthetic example 1 using different sampling frequencies of 100 Hz (upper panel) and 50 Hz (lower panel). Blue curves are interpolated by using the maxima at the sample locations, green curves are interpolated by using the new maxima determined by Rato et al. [2008]'s parabolic strategy.	102
5.36	The total residuals, $x(t) - \sum_{i=1}^3 c_i(t)$, of the synthetic time series of examples 1 obtained by using Rato et al.'s envelopes (red dashed lines) and Huang et al.'s envelopes (blue lines) in the shifting procedure. The used sampling frequencies in the upper and middle panel are 100 Hz and 50 Hz, respectively. In the lower panel an magnified interval of the middle panel is shown.	103
6.1	Nine day long section of B-field data with sampling frequency 1 Hz measured onshore Costa Rica in 2007. These data are used as the B-field time series in the synthetic tests.	109
6.2	The Fourier power spectra of the nine day long section of B-field data. The spectra are calculated from simple Fourier transform of the detrended and windowed (hanning window) time series.	110
6.3	Synthetic layered Earth model to test the new EMD based impedance estimation method.	110
6.4	First panel: The real and imaginary parts of the impedance Z determined analytically from the synthetic layered Earth model shown in Fig. 6.3. Middle and lower panels: The corresponding apparent resistivities and phases. . . .	111
6.5	Synthetic E-field time series are constructed from the magnetic field data using the analytically determined impedance $Z(\omega)$ from the layered model.	112
6.6	The real and imaginary parts of the $E_x(\omega)$ and $E_y(\omega)$ which are calculated by multiplication of $B_y(\omega)$ with $Z(\omega)$ and $B_x(\omega)$ with $Z(\omega)$, respectively.	113
6.7	Estimated apparent resistivity and phase by using simple Fourier transform (red dashed curves and green curves). The estimated values are calculated from the average auto- and cross-power densities of small sections (4096 and 32768 data points long, respectively) by using Eq. 4.23.	114
6.8	One short section x-component magnetic time series (left upper panel) and its 13 IMFs obtained by using EMD method.	116

6.9	The instantaneous spectra of one short section x-component magnetic time series. The color bar indicates the logarithm of the instantaneous amplitudes. The instantaneous spectrum of the last IMF, which is the trend of the time series and has much high amplitude than other IMFs, is not included. For convenience reason, the figure only shows 10000 seconds (≈ 2.78 hours) length time interval. One frequency window centered at 5.6×10^{-3} Hz (blue box) is highlighted and will be further explained in Fig. 6.10.	117
6.10	The real and imaginary parts of the complex signal $B_{xf}(t)$ of the short section of the magnetic time series corresponds the time interval and the highlighted frequency band shown in Fig. 6.9. It can be seen that certain time segments contain no energy in this particular frequency band.	118
6.11	Apparent resistivity (upper panel) and phase (lower panel) for the synthetic time series. Blue curves are the analytic values and red circles are the results from EMD. It should be noted that the apparent resistivities are slightly downward biased.	119
6.12	Estimated apparent resistivity (upper panel) and phase (lower panel) for the synthetic time series by using formula Eq. 6.1. Blue curves are the analytic values and red circles are the results from EMD. Notice that the apparent resistivities are slightly upward biased.	121
6.13	Apparent resistivity (upper panel) and phase (lower panel) estimates using a "numeric" reference time series for the B-field data. For the reference the same B-field data are used as for the original, however, the EMD algorithms applied on the time series are different. Blue curves are the analytic values and red circles are the estimated results from EMD. It should be noted that the downward bias of apparent resistivities (see Fig. 6.11) is not present any more.	123
6.14	Fourier amplitude spectra of synthetic time series and their corresponding smoothed spectra of noise: B_y (upper panel) and E_x (lower panel).	125
6.15	Three levels of noise added on B_y (upper panel) and E_x (lower panel).	126
6.16	The time series after three levels of noise are added: B_y (upper panel) and E_x (lower panel).	127
6.17	Apparent resistivity (upper panel) and phase (lower panel) estimates from EMD using a "numeric" reference time series for the B-field data. In contrast to Fig. 6.13, uncorrelated colored noise, which is $\sim 2\%$ of the amplitude spectra of the associated time series, are added to all time series. In addition to the estimates from EMD (red circles), estimates from Fourier transform without any robust processing steps (blue diamonds) and estimates from BIRRP (black square) are presented.	128
6.18	Analogue description as in Fig.6.17, except that the noise level here is $\sim 5\%$	129
6.19	Analogue description as in Fig.6.17, except that the noise level here is $\sim 10\%$	130
6.20	The tilt measurements T_x (upper panel) and T_y (lower panel) at marine MT station m05 in Costa Rica.	132
6.21	Noisy (blue curves) and original (red curves) synthetic electric (upper panel) and magnetic (lower panel) field data. Non-stationary noise added on data are obtained from the detrending tilt measurements (Fig. 6.20) multiplied by constants. T_x multiplied by 1 is added on E_x and T_y multiplied by 200 is added on B_y	133

6.22	Normalized power spectral densities of noisy and original synthetic magnetic field data, which are calculated by using the multitaper method.	134
6.23	Apparent resistivity (upper panel) and phase (lower panel) estimates from EMD using a "numeric" reference time series for the B-field data. In contrast to Fig. 6.13, uncorrelated non-stationary noise are added to all time series. In addition to the estimates from EMD (red circles), estimates from Fourier transform without any robust processing steps (blue diamonds) and estimates from BIRRP (black square) are presented.	135
7.1	Tectonic setting of Central America. The countries are labeled as follows: Mexico (M), Belize (B), Guatemala (G), El Salvador (S), Honduras (H), Nicaragua (N), Costa Rica (CR) and Panama (P). Map is modified from Rogers et al. [2002].	139
7.2	Topographic map and MT profile across the Costa Rican subduction zone. Offshore MT stations were deployed by GEOMAR, Kiel, in 2007 and land stations were installed by the Free University of Berlin in 2008.	140
7.3	Noisy time series example measured at station m05. Different time sections are displayed: A) 5 days, B) 1 day and C) 7 hours. B_x, B_y, B_z are three orthogonal components of the magnetic field; E_x, E_y are two orthogonal components of electric field; t_x, t_y are two components of tilt measurements (pitch and roll angles). The time series exhibit significant noise over the entire measuring period, due to the affection of motion induced by tilt variations.	142
7.4	Coherency between 5-day-long tilt components and horizontal magnetic components time series measured at station m05. The time series are shown in Fig. 7.3 A).	143
7.5	Overview of all identified quiet time sections of all marine stations which are used for data processing. Red bars denote periods of quiet time sections. As a comparison, land data has good quality during all measure time. Figure comes from Worzewski [2011].	145
7.6	Upper panel: station map including bathymetry. Middle panel: two-dimensional inversion model of electrical resistivity below the Costa Rican subduction zone. The logarithmical color scale indicates the electrical resistivity ranging from conductive (least resistive) in red to highly resistive in blue. Lower panel: seismic information superimposed on inversion result-seismic boundaries (dashed brown) and velocities; blue circles, interplate earthquakes; pink circles, seismogenic-zone earthquakes from H.R.DeShon et al. [2006]. Figure comes from Worzewski et al. [2011].	147
7.7	One long quiet section of E- and B-field time series at MT station 11 in Costa Rica (sampling frequency is 1 Hz).	148
7.8	B-field time series from a land station that are used as the remote reference site. The time window shown here corresponds to the quiet section of station m11 shown in Fig. 7.7.	149
7.9	Station m11: comparison of results derived from BIRRP processing and EMD processing. Left panels: off-diagonal (upper) and diagonal (lower) apparent resistivities. Right panels: off-diagonal (upper) and diagonal (lower) phases.	151
7.10	Station m10: comparison of results derived from BIRRP processing and EMD processing. Left panels: off-diagonal (upper) and diagonal (lower) apparent resistivities. Right panels: off-diagonal (upper) and diagonal (lower) phases.	152

7.11	Station m09: comparison of results derived from BIRRP processing and EMD processing. Left panels: off-diagonal (upper) and diagonal (lower) apparent resistivities. Right panels: off-diagonal (upper) and diagonal (lower) phases.	153
7.12	Station m05: comparison of results derived from BIRRP processing and EMD processing. Left panels: off-diagonal (upper) and diagonal (lower) apparent resistivities. Right panels: off-diagonal (upper) and diagonal (lower) phases.	154
7.13	A short segment of x-component magnetic field and tilt data at station m05.	156
7.14	The 1st, 4th, 7th and 10th IMFs of the short segment of x-component magnetic field and tilt measurements at station m05 (Fig. 7.13).	157
7.15	Three dimension shaded surface plot of instantaneous spectra of the tilt data (see lower panel in Fig. 7.13). The color code indicates the logarithmic scale of the instantaneous amplitudes.	158
7.16	Three dimension shaded surface plot of instantaneous spectra of the magnetic data (see upper panel in Fig. 7.13). The color code indicates the logarithmic scale of the instantaneous amplitudes.	159
7.17	One of the long segment of data measured at station m05.	160
7.18	Station m05: estimated off-diagonal elements of the apparent resistivities (upper panel) and phases (lower panel) from three long noisy sections by using BIRRP. The results from corresponding short quiet sections estimated by using BIRRP are also plotted as a comparison.	162
7.19	Station m05: estimated diagonal elements of the apparent resistivities (upper panel) and phases (lower panel) from three long noisy sections by using BIRRP. The results from corresponding short quiet sections estimated by using BIRRP are also plotted as a comparison.	163
7.20	Estimated off-diagonal elements of the apparent resistivities (upper panel: ρ_{xy} , lower panel: ρ_{yx}) from three long noisy sections at station m05 by using EMD. The results estimated by using BIRRP are also plotted as a comparison.	164

Chapter 1

Introduction

Magnetotelluric (MT) studies use natural electromagnetic source fields generated in the Ionosphere and Magnetosphere as a source, which is assumed to be quasi-uniform. The horizontal components of the time-varying natural electric and magnetic fields are measured at the Earth's surface and their spectral ratio gives the impedance as a function of frequency. This impedance can be used to determine the resistivity structure in the subsurface. In order to obtain a reliable subsurface structure model, processing measured time-varying data to derive frequency based impedance is the first very important and fundamental step in MT studies. In standard data processing routines, the Fourier transform is used to determine the spectra [e.g. Vozoff, 1972, Sims et al., 1971] and impedance estimates are determined through a combination of methods like remote reference [e.g. Gamble et al., 1979a,b], multitaper spectral analysis [Thomson, 1982], robust estimates [e.g. Egbert and Booker, 1986, Chave and Thomson, 1989] and bounded-influence remote reference processing, BIRRP, [Chave and Thomson, 2004]. Especially the lattermost modern MT processing codes have greatly improved the reliability of impedance calculations. However, sometimes even sophisticated methods can provide poorly determined impedance estimates [Szarka, 1988, Qian and Pedersen, 1991].

The main reason for these problems is that in practice, field data contains signal components which violate the source assumption made in MT processing. Electromagnetic signals measured on land may be contaminated by artificial noise originating from e.g. power lines, electrical fences and other industrial sources [Szarka, 1988, Junge, 1996]. In addition, high power transients with a distinct non-stationary behavior are often present. Marine measurements are often noisy because of the surface waves, passing ships and the noise associated with tilt movements of the MT instruments. Because these signal components are usually present in both the electric and magnetic fields, they cause high coherence in the associated time series and may be accordingly difficult to remove by standard processing techniques. Therefore, noisy sections of the time series are typically removed before processing. For low quality data this means that only a small portion of data may remain for impedance calculation.

Furthermore, one underlying assumption for methods using power spectral densities calculated through Fourier analysis is that the time series are in a wide-sense stationary. However, many geomagnetic time series recorded for MT investigations are characterized by a changing frequency content with time mainly due to contamination by transient local natural or artificial noise signals. Chant and Hastie [1992, 1993] demonstrate that MT signals can

be considered as non-stationary within typical recording lengths and the impedance tensor is affected significantly by such non-stationary source field fluctuations when using Fourier transform based processing.

To process non-stationary MT signals, several techniques are used, such as time-frequency distribution (TFD) analysis techniques [Chant and Hastie, 1992, 1993], wavelet analysis [e.g. Alexandrescu et al., 1995, Suto et al., 2006, Garcia and Jones, 2008] and robust processing techniques [Chave and Thomson, 2004]. Different techniques have some advantages and shortcomings (see section 5.1). Another way to address the problem of non-stationarity is to use Instantaneous Frequency (IF) and Instantaneous Amplitude (IA), which allow for imaging the frequency content as a function of time. However, calculation of IF is only physically meaningful when the signal is mono-component, which means that the signal is zero-mean and, amplitude and / or frequency modulated (AM-FM) [Boashash, 1992a,b, Loughlin and Tracer, 1996, Huang et al., 2009]. Since geomagnetic time series, like most other time series, typically are not mono-component, IFs directly calculated from these time series are generally not physically meaningful.

With the development of the Empirical Mode Decomposition (EMD) introduced by Huang et al. [1998], a processing method exists which decomposes non-stationary time series into a sequence of mono-component signals, termed intrinsic mode functions (IMFs). Accordingly, physically meaningful instantaneous frequencies can be calculated for all IMFs. Compared to traditional Fourier analysis, EMD works directly in the time domain rather than in frequency domain and results are typically presented as time-frequency-energy representations. However, due to its definition, EMD is an empirical method in contrast to the mathematically well-defined Fourier transform.

Over the last ten years, the heuristic EMD approach has demonstrated its applicability in a wide range of geoscience studies, as well as in many other different research fields. For instance, Vasudevan and Cook [2000] use EMD to extract the scaling behavior of reflected waves. Huang et al. [2001] apply EMD to seismic time series from a large earthquake and conclude that Fourier analysis seriously under-represents low frequency energy due to artificial high frequency harmonics. Zhang et al. [2003a,b] and [Zhang, 2006] use EMD in tracing seismic wave propagation. Applications of EMD in meteorological studies have been summarized by Duffy [2004], and Huang and Wu [2008] give a review of applications in atmosphere, climate and oceanographic studies.

In this thesis, I investigate how EMD, combined with instantaneous spectra calculation, is suited for the processing of marine MT data, which is typically characterized by non-stationary noise. The procedure is applied to both synthetic and field data. The recorded electric and magnetic time series are first decomposed into intrinsic mode functions using EMD. Then the instantaneous frequencies of each intrinsic mode function are calculated through an empirical amplitude and frequency demodulation method proposed by Huang et al. [2009]. Finally, the obtained time-frequency distributions are used for impedance estimation. In contrast to traditional processing, the impedance estimation is performed in the time-frequency-domain using instantaneous spectra [Berdichevsky et al., 1973] rather than in the frequency domain. To evaluate the results the impedance estimates are compared to analytical results for the case of synthetic data, and with estimates from a standard MT data processing technique [Chave and Thomson, 2004] for the case of real data examples.

The thesis is arranged as follows. Firstly, chapter 2 gives a brief introduction of magnetotelluric method and its applications in geoscience. The relevant theoretical aspects of the MT method and the concepts of impedance, apparent resistivity and phase are shortly summarized in chapter 3. Traditional MT data processing methods and their technique problems and limitations are discussed in chapter 4. Chapter 5 focusses on the EMD method. How the EMD method works, how instantaneous parameters such as instantaneous frequencies and instantaneous amplitudes as functions of time can be derived, and further some limitations of the EMD method are discussed in detail. In chapter 6, by means of a layered Earth model, synthetic time series are constructed and the EMD impedance estimation procedure is tested. In order to investigate how far the impedance estimates are affected by noise, stochastic stationary and non-stationary noise are added to the synthetic time series and the resulting estimations are analyzed as well. Finally, in chapter 7 the new estimation procedure is applied to several marine MT field data sets measured at Costa Rica subduction zone. Both short quiet segments and long noisy segments of the data are tested, and the results are compared with the ones obtained from processing the data with BIRRP [Chave and Thomson, 2004]. The last chapter concludes with a summary of the study and provides an outlook for future investigations.

The main results of this thesis are presented in the publication of Chen et al. [2012], including parts of chapters 5, 6 and 7.

Chapter 2

Introduction of Magnetotelluric Method

In geophysics many methods are applied to obtain information about the subsurface of the Earth which is not available from surface geological observations. To investigate the electrical conductivity structure of the Earth, numerous electromagnetic (EM) methods have been used. Among them, Magnetotellurics (MT) is one of these techniques that uses the natural source and has the ability to image Earth's structure from the near surface down to depths of several hundreds kilometers. The MT method has been introduced in 1950s and has been increasingly used in geological applications and in the petroleum industry over the past decades.

In this chapter I give a short overview of the electrical resistivity of Earth materials, the basic concepts of magnetotellurics and some applications of the MT method in Earth's investigation.

2.1 Electrical resistivity of Earth materials

The electrical resistivity/conductivity of rocks can vary by many orders of magnitudes. Therefore, it is a suitable parameter to distinguish different subsurface between structures. Electrical resistivity is the reciprocal of electrical conductivity, both terms are often used alternatively in the community, depending on the focus of the target.

Generally, the electrical resistivity of a rock depends on 1) rock matrix, 2) porosity and temperature, 3) fluid content in pore space and 4) fluid type. In metallic ores resistivity can be very low, e.g., for a high-grade pyrrhotite ore it is of the order of $10^{-5} \Omega m$, while in dry rocks it can be a very high, e.g., in dry marble it is around $10^8 \Omega m$. The transmission of electrical currents occurring in rocks and minerals can be classified mainly in three types of conductive mechanisms: electronic conduction, semi-conduction and electrolytic conduction [Simpson and Bahr, 2005].

Electronic conduction occurs in metallic ore minerals, in which the electrons can freely move within the material and transport charges. In the presence of an interconnected network of highly conducting matter, such as graphite or ores, the conductivity of the rock increases by orders of magnitude.

Semi-conduction occurs in dry unaltered rocks in the crust and upper mantle and some crystals, such as the silicate mineral. Such minerals have resistivities higher than a conductor since they contain fewer conduction electrons, but lower than an insulator due to the fact that electrons are not rigidly bound to atoms as they are in an insulator. The energy needed to release additional electrons from the atoms of such minerals is relatively low, such that thermal excitation is sufficient. For example, the conductivity of the olivine increases by several orders of magnitude (from 10^{-6} to 10^{-3} S/m) with rising temperature (from about $700^{\circ}C$ to $1400^{\circ}C$) [Constable et al., 1992].

Electrolytic conduction occurs in aqueous solutions in which the electric current is accompanied by the movement of matter in form of ions. Although, the effect of electrolytic conduction is generally smaller than electronic conduction, the resistivity of rocks can be strongly influenced by the presence of water in the pore space, which acts as an electrolyte. This is especially important in porous sediments and sedimentary rocks, since the rock materials are generally much poorer conductors than water and overall conductivities of the rocks increase with the amount of water they contain. An empirical formula to describe the resistivity of sedimentary rocks containing water is given by Archie's law (Archie [1942]):

$$\rho = a \cdot \phi^{-m} \cdot S^{-n} \cdot \rho_w \quad (2.1)$$

where a , m and n are empirically determined parameters which vary depending on the cementation of the sediments, ϕ is the porosity, S is the water saturation and ρ_w is the resistivity of the water.

2.2 Magnetotellurics: Natural source EM method

Magnetotelluric (MT) technique, which has been pioneered by Tikhonov and Cagniard (Tikhonov [1950], Cagniard [1953]), determines the electrical properties of the subsurface through the analysis of natural magnetic and telluric (electric) field variations at the Earth's surface.

The external magnetic field of the Earth is encompassed by a region called magnetosphere, which has a tear-drop shape due to the solar wind plasma interacting with the intrinsic magnetic field of the Earth (see Fig. 2.1). The Earth's atmosphere is contained in the magnetosphere. When electrical properties are concerned, the atmosphere can be classified into neutrosphere and ionosphere. The neutrosphere is the lower part of atmosphere and electrically neutral. The atmosphere above the neutrosphere is the ionosphere, which is ionized by solar radiation and is, therefore, electrically conductive. The transition region between neutrosphere and ionosphere is around 50 - 100 km above the Earth's surface.

The Earth's natural magnetic field variations used in MT are caused by two types of sources. One is large scale currents in ionosphere which are arising from the interaction between the ionized particles in solar wind and the Earth's magnetic field. This interaction distorts the Earth's magnetic field and results in magnetic field variations that can be measured at the surface of the Earth. The typical frequencies of these variations range from about 10^{-5} to 1

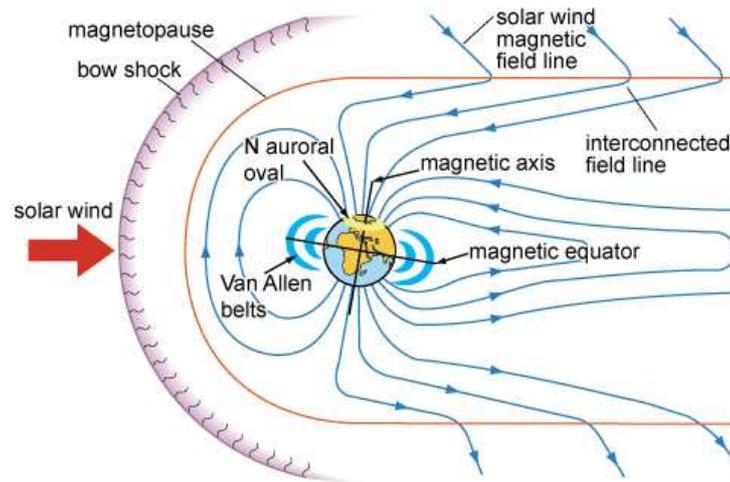


Figure 2.1: Magnetic field of the Earth distorted by the solar wind. (Figure comes from: http://phyworld.ed.hkedcity.net/iq/aurora/magnetopause_e.gif)

Hz.

Another source is the worldwide lightning activities, which generate magnetic waves above 1 Hz. The electromagnetic waves emitted by individual lightning get partially trapped in a waveguide located between the conductive ionosphere and the Earth. These waves can travel long distances. Lightning around the world is enough to provide a continuous source at any location of the Earth's surface.

Both source mechanisms are of extremely large scale, such that by first approximation, the sources which induce magnetic field can be assumed to be quasi-uniform in space [e.g. Schmucker, 1995].

The fluctuating magnetic fields caused by these two sources induce secondary electric currents within the Earth which are referred to as "telluric" currents. The strength of telluric current depends on electrical conductivity of the Earth. As will be discussed in next chapter, from simultaneously measured time variations of magnetic and electric fields at Earth's surface, a frequency-dependent transfer function can be determined, which allows the geophysicists to map the Earth's electrical conductivity as a function of depth and position.

Generally, two horizontal orthogonal components of electric signals and three orthogonal components of magnetic signals are measured at the surface of the Earth. The measured magnetic field is often a superposition of the small variations of solar and atmospheric signals (intensity in the order of $\approx 10^{-1} - 10^2$ nT, depending on the solar activity) and the large static primary magnetic field of Earth (intensity in the order of $\approx 10^4$ nT). However, since the primary field can be regarded static during MT experiment, the fluctuation can be clearly distinguished from the primary field. In addition, the measured signals always contain not only the natural source fields but also some artificial noise. This increases the difficulties in the next important step of a MT experiment, the derivation of a transfer function from the field variations, as will be discussed in Chapter 4.

In MT studies, the frequency range of the measured fields is from $\sim 10^{-5}$ Hz to $\sim 10^3$ Hz, and the depths range that can be imaged is from ~ 160 m to > 500 km. Compared to other EM methods in which the maximum penetration depth is always limited by the size and power of the available source and the source-receiver configurations, the large depths that can be imaged by using MT technique is one advantage of the method. Another advantage is that no man-made source is needed, which makes MT a logistically easy exploration method. Therefore, MT investigations become more and more important in Earth exploration and is accompanied by advancement of instruments and the development of better data processing and modelling algorithms.

2.3 Applications of MT in geoscience

MT method has been used on land to image the geologic structure since 1950s [Vozoff, 1972] and in ocean since the 1980s [e.g. Law, 1983, Heinson and Constable, 1992, Palshin et al., 1995, Baba, 2005]. Due to the huge frequency range and depth of penetration, MT can be used for a large number of applications. Some typical applications are summarized in the following:

1. *Studies of deep crust and mantle:* The aim of crustal and mantle studies is usually to understand the structure and tectonic history of a certain region on a large scale. For example, Wannamaker et al. [1989] observed very low resistivity in the sedimentary wedge of the abyssal Cascadia Basin which is interpreted as that the sediments are off-scraped or lose most of their interstitial water before being carried down the trench. They also observed moderately low resistivities from about 35 - 215 km depth in the oceanic upper mantle of the Juan de Fuca plate, which can be interpreted as up to several percent partial melt attending regional upwelling in the vicinity of the ridge. Lezaeta et al. [2000] used magnetotellurics to image the crust and upper mantle in the backarc of the northwest Argentinean Andes. Brasse and Soyer [2001] revealed a zone of moderately high conductivities in the deep crust below the active volcanic arc that extends eastwards to the Argentinean border. Jones et al. [2001] presented result from three MT experiments on the Slave craton and found diamondiferous kimberlite pipes. Ritter et al. [2003] discovered that under the very resistive upper crust, two subvertical conductors in upper to mid-crustal levels correlate with major tectonic zone boundaries and they found out that there exists a highly conductive middle to lower crust zone in the southern part in Damara Belt of Namibia. Wannamaker et al. [2004] studied the structure and thermal regime beneath the South Pole region from magnetotelluric measurements. Jiracek et al. [2007] established the evidence for the occurrence of trapped fluidized zones in the ductile crust through the study of three continental plate boundaries in New Zealand, Himalaya, and California-USA. Brasse et al. [2009] revealed the deep electrical resistivity structure of northwestern Costa Rica. Worzewski et al. [2011] presented the MT image of the fluid cycle in the Costa Rican subduction zone. Some other examples are e.g. studies of fluids in the Tibetan crust [Wei et al., 2001] and fluids at intraplate earthquake zones in Japan [Ogawa et al., 2001].
2. *Studies in oil and gas exploration:* MT is an important supplementary method for

petroleum exploration where seismic investigations are hampered. Petroleum and gas are often trapped in pockets between rock and salt. These pockets were formed by the upward movement of salt bodies towards the surface, penetrating and bending the overlying and neighboring rocks. To identify the location of reservoirs, mapping the top and bottom of salt structures is therefore important. Because salt structures contain entrained sediments that produce significant scattering and have strong reflecting boundaries, this produces ambiguities in the interpretation of seismic reflection data, especially in the steeply dipping sides and the base of the salt. However the resistivity of salt is often more than ten times greater than the surrounding sediments, such that MT can help to resolve these problems. Hoversten et al. [1998] showed on the synthetic examples that MT responses at frequencies from 10^{-3} to 1 Hz are sensitive to the salt structures at depth substantial for petroleum exploration. Key et al. [2006] used data, which were collected over a salt body, to investigate the effectiveness of using 2D marine MT methods to map 3D salt structures. Through synthetic model studies, Avdeeva et al. [2012] demonstrated that 3D MT inversion can be applied to decide whether the overhang is present in the shallow salt structure even in the case when only profile data are available. Jegen et al. [2009], Heincke et al. [2010], Moorkamp et al. [2011] discovered that the sub-basalt structure can be better imaged by using a joint inversion of combining Marine MT, gravity and velocity data sets, which is useful in oil and gas exploration.

3. *Studies of geothermal:* Geothermal studies examine the structures of geothermal areas and help to understand basic properties of these systems and the connection to large scale tectonic processes. Heise et al. [2007] collected MT data along a profile across the Taupo volcanic zone and constructed an electrical conductivity model which showed that the resistive (500 - 1500 Ωm) crust is thinned underneath the volcanic zone from 20 to 10 km. Beneath the thinned crust they observed an area of high conductivity (3 - 30 Ωm), which they interpret as a zone of melt accumulation. Ingham and Reeves [1993] made MT soundings at 15 locations in and around the Tokaanu geothermal field in the North Island of New Zealand. Their models along two transects clearly identify the locations of the eastern and southeastern boundaries of the field. The existence of low resistivity zone at depth is interpreted to correspond to the upwelling of geothermal fluid and associated hydrothermal alteration.
4. *Studies in mining exploration and ground water monitoring:* Electromagnetic methods are sensitive to the conductive structures in the underground and are therefore widely used in finding minerals and ground water. Tuncer et al. [2006] used EM methods to map basement conductors at depth in the Athabasca Basin. Their study has shown that audiomagnetotelluric exploration, which is a higher-frequency magnetotelluric technique for shallower investigations, is an effective tool for mapping basement conductors to a depth of 2-3 km. Pedersen et al. [2005] used another higher-frequency magnetotelluric technique, radiomagnetotelluric (14-250 kHz), combined with controlled-source MT (1-12 kHz) measurements to the exploration of groundwater located in sandy formations at depths as great as 20 m below thick clay lenses. Compared to the results from bore-holes and shallow-reflection seismic sections, their electromagnetic results show that their interpretation largely agrees with the former and provides useful information for interpreting the latter by identifying lithological boundaries between the clay and sand and between the sand and crystalline basement.

5. *Studies of the environment:* Buried waste often has a much higher electrical conductivity than the surrounding host material, therefore, economical EM methods seem suitable for investigating the extent of the waste deposits. MT has been shown to be very useful for buried waste characterization when used at radio frequencies (from 10 kHz up to 1 MHz). E.g. Newman et al. [2003] collected 320 radio magnetotelluric measurements (frequencies varying from 18.3 and 234 kHz) over a buried waste site near Cologne, Germany. They successfully recovered the shape and the extent of the contamination, which agrees with the borehole data.

Chapter 3

Maxwell's Equations and Transfer Function

In this chapter, the most relevant theoretical aspects of the MT method are summarized. Since the behavior of any electromagnetic field is described by Maxwell's equations, they provide a starting point to understand the MT method. Therefore, Maxwell's equations and their simplified forms used in MT studies are described in section 1. In section 2, based on a simple 1D homogeneous half space model, important concepts like the impedance tensor, apparent resistivity & phase and skin depth are introduced, as well as their relation to the resistivity properties of the Earth. The calculation of the impedance for a layered 1D Earth model is shown in section 3. A short introduction of how Maxwell's equations can be split into two independent signals, the TE and TM modes, for 2D resistivity models is given in section 4. In the last section some characteristics of the most general 3D case of the impedance tensor are described.

3.1 Maxwell's equations

In the time domain Maxwell's equations relating the changing electric and magnetic fields in an arbitrary medium are given by:

$$\nabla \times \mathbf{e}(t) = -\frac{\partial \mathbf{b}(t)}{\partial t}, \quad (3.1)$$

$$\nabla \times \mathbf{h}(t) = \mathbf{j} + \frac{\partial \mathbf{d}(t)}{\partial t}, \quad (3.2)$$

$$\nabla \cdot \mathbf{d}(t) = q, \quad (3.3)$$

$$\nabla \cdot \mathbf{b}(t) = 0. \quad (3.4)$$

where $\mathbf{e}(t)$ is the electric field in V/m , $\mathbf{b}(t)$ is the magnetic induction in $V \cdot s/m^2$ or *Tesla*, $\mathbf{h}(t)$ is the magnetic field strength in A/m , \mathbf{j} is current density in A/m^2 , $\mathbf{d}(t)$ is the electric displacement in C/m^2 and q is the charge density in C/m^3 . Eq. 3.1 states that circulating electric fields are induced by time-varying magnetic fields (Faraday's law), Eq. 3.2 states that circulating magnetic fields are induced by the vector sum of electric currents and time-varying electric fields (Ampere's law), and Eqs. 3.3 and 3.4 state that the electric field

diverges from electric charges (Gauss's law for electricity) and that there are no magnetic monopoles (Gauss's law for magnetism), respectively.

For magnetotellurics studies, Maxwell's equations can be simplified by a number of assumptions.

- (1). All media are assumed to be isotropic, locally homogeneous, and possess electrical properties which are independent of time.
- (2). The Earth is assumed to be a conductor which obeys Ohm's law, such that current density \mathbf{j} and electric field \mathbf{E} are linearly related by the scalar electrical conductivity σ in *siemens/m* ($\sigma = 1/\rho$, ρ is resistivity in Ωm),

$$\mathbf{j} = \sigma \mathbf{e}. \quad (3.5)$$

- (3). Changes in electrical permittivities ε and magnetic permeabilities μ in different Earth materials can be neglected in most cases such that

$$\begin{aligned} \mathbf{d} &= \varepsilon_0 \mathbf{e}, \\ \mathbf{b} &= \mu_0 \mathbf{h}, \end{aligned} \quad (3.6)$$

where $\mu_0 = 4\pi \cdot 10^{-7} H/m$ is the permeability and $\varepsilon_0 = 8.85 \cdot 10^{-12} farads/m$ is the capacity of free space.

- (4). There exists no current sources within the Earth,

$$q = 0. \quad (3.7)$$

Maxwell's equations are then represented by:

$$\nabla \times \mathbf{e} = -\frac{\partial \mathbf{b}}{\partial t}, \quad (3.8)$$

$$\nabla \times \mathbf{b} = \mu_0 \sigma \mathbf{e} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{e}}{\partial t}, \quad (3.9)$$

$$\nabla \cdot \mathbf{e} = 0, \quad (3.10)$$

$$\nabla \cdot \mathbf{b} = 0. \quad (3.11)$$

To derive the wave equations for electric and magnetic fields, the vector identity

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \Delta \mathbf{A}, \quad (3.12)$$

will be used.

Applying the curl operator to Eq. 3.8 and substitution into Eqs. 3.9 and 3.10 yields:

$$\begin{aligned}\nabla(\nabla \cdot \mathbf{e}) - \Delta \mathbf{e} &= \nabla \times (\nabla \times \mathbf{e}) \\ &= -\frac{\partial}{\partial t} \nabla \times \mathbf{b} = -\mu_0 \sigma \frac{\partial \mathbf{e}}{\partial t} - \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{e}}{\partial t^2},\end{aligned}\tag{3.13}$$

\implies

$$\Delta \mathbf{e} = \mu_0 \sigma \frac{\partial \mathbf{e}}{\partial t} + \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{e}}{\partial t^2}.\tag{3.14}$$

Since the media is assumed to be locally homogeneous ($\nabla \sigma = 0$), similarly by applying the curl operator to Eq. 3.9 and substitution into Eq. 3.11, one obtains:

$$\begin{aligned}\nabla(\nabla \cdot \mathbf{b}) - \Delta \mathbf{b} &= \nabla \times (\nabla \times \mathbf{b}) = \nabla \times (\mu_0 \sigma \mathbf{e}) + \nabla \times (\mu_0 \varepsilon_0 \frac{\partial \mathbf{e}}{\partial t}) \\ &= \mu_0 \sigma \nabla \times \mathbf{e} - \mu_0 \mathbf{e} \times \nabla \sigma + \mu_0 \varepsilon_0 \frac{\partial}{\partial t} \nabla \times \mathbf{e} \\ &= -\mu_0 \sigma \frac{\partial \mathbf{b}}{\partial t} - \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{b}}{\partial t^2},\end{aligned}\tag{3.15}$$

\implies

$$\Delta \mathbf{b} = \mu_0 \sigma \frac{\partial \mathbf{b}}{\partial t} + \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{b}}{\partial t^2}.\tag{3.16}$$

Combining Eqs. 3.14 and 3.16 together yields two damped wave equations:

$$\Delta \begin{bmatrix} \mathbf{e} \\ \mathbf{b} \end{bmatrix} = \mu_0 \sigma \frac{\partial}{\partial t} \begin{bmatrix} \mathbf{e} \\ \mathbf{b} \end{bmatrix} + \mu_0 \varepsilon_0 \frac{\partial^2}{\partial t^2} \begin{bmatrix} \mathbf{e} \\ \mathbf{b} \end{bmatrix}.\tag{3.17}$$

These equations can be easily transferred from time domain into frequency domain:

$$\begin{aligned}\mathbf{e}(t) &\xrightarrow{\mathcal{F}} \underline{\mathbf{E}}(\omega), \\ \mathbf{b}(t) &\xrightarrow{\mathcal{F}} \underline{\mathbf{B}}(\omega),\end{aligned}\tag{3.18}$$

by using the Fourier transform, $\mathcal{F}[f(t)] = F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt$ and $\mathcal{F}^{-1}[F(\omega)] = f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} d\omega$, to obtain Helmholtz equations:

$$\Delta \begin{bmatrix} \underline{\mathbf{E}} \\ \underline{\mathbf{B}} \end{bmatrix} = i\omega \mu_0 \sigma \begin{bmatrix} \underline{\mathbf{E}} \\ \underline{\mathbf{B}} \end{bmatrix} - \omega^2 \mu_0 \varepsilon_0 \begin{bmatrix} \underline{\mathbf{E}} \\ \underline{\mathbf{B}} \end{bmatrix}.\tag{3.19}$$

As mentioned in chapter 2, the frequency range of naturally induced electromagnetic fields in the Earth is $\sim 10^{-5}$ Hz to $\sim 10^3$ Hz, whereas the scalar electrical resistivity $\rho (= 1/\sigma)$ in the Earth ranges from $\sim 10^{-5}$ Ωm to $\sim 10^8$ Ωm . Comparing the two terms in the right hand side of Eq. 3.19,

$$\frac{\omega^2 \mu_0 \varepsilon_0 \mathbf{A}}{i \omega \mu_0 \sigma \mathbf{A}} = \frac{\omega \varepsilon_0}{i \sigma} \approx 10^{-14} \text{ to } 10^{-9}, \quad \text{where } \mathbf{A} = \underline{\mathbf{E}}, \underline{\mathbf{B}}, \quad (3.20)$$

one can observe that the second terms on the right hand side can be neglected and, therefore, Eq. 3.19 is further simplified as two diffusion equations:

$$\Delta \begin{bmatrix} \underline{\mathbf{E}} \\ \underline{\mathbf{B}} \end{bmatrix} = i \omega \mu_0 \sigma \begin{bmatrix} \underline{\mathbf{E}} \\ \underline{\mathbf{B}} \end{bmatrix}. \quad (3.21)$$

These two equations describe the most general form of the electric and magnetic fields in MT studies. In order to explain more basic concepts of MT, firstly, the situation is simplified to a homogeneous half space.

3.2 1D homogeneous half space

The simplest conceivable geoelectric model is a homogeneous half space, composed of an insulated air layer overlying a homogeneous subsurface of constant conductivity σ , as shown in Fig. 3.1. In such a 1D model, the electromagnetic waves, $\underline{\mathbf{E}}$ and $\underline{\mathbf{B}}$, only vary with time and depth and, have no variations in x- and y- directions, which means that, $\frac{\partial \underline{\mathbf{E}}}{\partial x} = 0$, $\frac{\partial \underline{\mathbf{E}}}{\partial y} = 0$, $\frac{\partial \underline{\mathbf{B}}}{\partial x} = 0$ and $\frac{\partial \underline{\mathbf{B}}}{\partial y} = 0$.

Then, the two diffusion equations in Eq. 3.21 can be simplified to:

$$\frac{\partial^2 \underline{\mathbf{E}}}{\partial z^2} = i \omega \mu_0 \sigma \underline{\mathbf{E}}, \quad (3.22)$$

$$\frac{\partial^2 \underline{\mathbf{B}}}{\partial z^2} = i \omega \mu_0 \sigma \underline{\mathbf{B}}. \quad (3.23)$$

These two second-order differential equations have solutions of the form:

$$\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{-\sqrt{i \omega \mu_0 \sigma} z} + \underline{\mathbf{E}}_1 e^{\sqrt{i \omega \mu_0 \sigma} z}, \quad (3.24)$$

$$\underline{\mathbf{B}} = \underline{\mathbf{B}}_0 e^{-\sqrt{i \omega \mu_0 \sigma} z} + \underline{\mathbf{B}}_1 e^{\sqrt{i \omega \mu_0 \sigma} z}. \quad (3.25)$$

Because all sources of the electromagnetic fields in MT studies are assumed to be outside the Earth (see section 3.1), the electromagnetic energy can only be absorbed within the Earth. Therefore the electromagnetic field will be decreasing with increasing depth and arbitrary

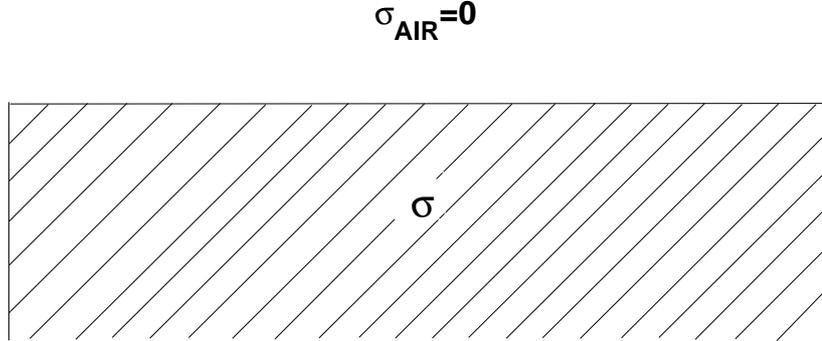


Figure 3.1: The homogeneous half space.

large electromagnetic field amplitudes cannot be present in Earth. This implies $\underline{\mathbf{E}}_1 = 0$ and $\underline{\mathbf{B}}_1 = 0$, therefore one obtains:

$$\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{-\sqrt{i\omega\mu_0\sigma} z}, \quad (3.26)$$

$$\underline{\mathbf{B}} = \underline{\mathbf{B}}_0 e^{-\sqrt{i\omega\mu_0\sigma} z}, \quad (3.27)$$

where, $\underline{\mathbf{E}}_0$ and $\underline{\mathbf{B}}_0$ are the electromagnetic fields at the Earth's surface.

Impedance, Apparent Resistivity and Phase

In frequency domain, Maxwell equation 3.8 is represented as:

$$\nabla \times \underline{\mathbf{E}}(\omega) = -i\omega \underline{\mathbf{B}}(\omega). \quad (3.28)$$

Let us assume that $\underline{\mathbf{E}} = (E_x, E_y, E_z)$ and $\underline{\mathbf{B}} = (B_x, B_y, B_z)$, note that $\underline{\mathbf{E}}$ and $\underline{\mathbf{B}}$ have no variations in x- and y- directions. Inserting Eqs. 3.26 and 3.27 into Eq. 3.28 yields:

$$-\frac{\partial E_y}{\partial z} = \sqrt{i\omega\mu_0\sigma} E_y = -i\omega B_x \quad \text{and} \quad \frac{\partial E_x}{\partial z} = -\sqrt{i\omega\mu_0\sigma} E_x = -i\omega B_y. \quad (3.29)$$

Then for a homogeneous half space the basic equations used in MT are derived:

$$Z_{xy} = \mu_0 \frac{E_x}{B_y} = \frac{i\omega\mu_0}{\sqrt{i\omega\mu_0\sigma}}, \quad (3.30)$$

$$Z_{yx} = \mu_0 \frac{E_y}{B_x} = \frac{-i\omega\mu_0}{\sqrt{i\omega\mu_0\sigma}}. \quad (3.31)$$

These two equations reveal that the conductivity of the half space, σ , can be calculated from the ratio of the orthogonal components of the electric and magnetic fields at a given frequency. This fact was the motivation for early magnetotelluric surveys.

Eqs. 3.30 and 3.31 can be written in a matrix form:

$$\left(\begin{array}{c} E_x \\ E_y \end{array} \right) = \frac{1}{\mu_0} \left(\begin{array}{cc} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{array} \right) \left(\begin{array}{c} B_x \\ B_y \end{array} \right) \quad \text{and} \quad \left. \begin{array}{l} Z_{xy} = -Z_{yx} \\ Z_{xx} = Z_{yy} = 0 \end{array} \right\} \quad \text{1-D.} \quad (3.32)$$

The quantity $Z = \left(\begin{array}{cc} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{array} \right)$ (in Ω) is known as the *magnetotelluric impedance tensor*. It does not depend on the source field and is only a function of Earth's resistivity (conductivity). Therefore it is an useful quantity in the theory of magnetotellurics.

Since Z is complex, it can be expressed in amplitude and phase. The *apparent resistivity* ρ (in Ωm) is a scaled amplitude of Z :

$$\rho_{ij} = \frac{1}{\mu_0 \omega} |Z_{ij}(\omega)|^2 \quad i, j = x \text{ or } y, \quad (3.33)$$

and the *phase* ϕ (in degree) is given as

$$\phi_{ij} = \tan^{-1} \left(\frac{\text{imag}(Z_{ij})}{\text{real}(Z_{ij})} \right) \quad i, j = x \text{ or } y. \quad (3.34)$$

Apparent resistivity and phase are equivalent representation of the impedance and are commonly used to represent and interpret MT data.

For a homogeneous half space, it is easy to derive that:

$$\rho_{xy} = \frac{1}{\mu_0 \omega} |Z_{xy}(\omega)|^2 = \frac{1}{\mu_0 \omega} \left| \frac{i^2 \omega^2 \mu_0^2}{i \omega \mu_0 \sigma} \right| = \frac{1}{\sigma} = \rho, \quad (3.35)$$

$$Z_{xy} = \sqrt{i} \sqrt{\frac{\omega \mu_0}{\sigma}} = \sqrt{\frac{\omega \mu_0}{2\sigma}} + i \sqrt{\frac{\omega \mu_0}{2\sigma}} \implies \phi_{xy} = \text{atan } 1 = 45^\circ. \quad (3.36)$$

This means that the apparent resistivity ρ_{xy} is exactly the resistivity of the half space and the phase ϕ_{xy} is constant with frequency. Similarly, one can derive that $\rho_{yx} = \rho$ and $\phi_{yx} = 135^\circ$.

When the electromagnetic waves propagate into the subsurface, the wave energies are absorbed by the conductive medium. Therefore in a homogeneous half-space the waves decay exponentially with depth in the Earth, as shown in Eqs. 3.26 and 3.27. The penetration depths of the EM waves depend on the frequencies and the Earth's conductivity structure. The *skin depth* is defined as the depth in the Earth where the amplitudes of EM waves are

reduced to $1/e$ ($\approx 37\%$) of their amplitudes measured at the Earth's surface.

In a homogeneous half-space, taking the magnetic field $\underline{\mathbf{B}}$ as an example, one can obtain:

$$|\underline{\mathbf{B}}_0|e^{-1} = |\underline{\mathbf{B}}_0| |e^{-\sqrt{i\omega\mu_0\sigma} p}| = |\underline{\mathbf{B}}_0| |e^{-\frac{1+i}{\sqrt{2}}\sqrt{\omega\mu_0\sigma} p}| = |\underline{\mathbf{B}}_0| e^{-\sqrt{\omega\mu_0\sigma/2} p}.$$

This means the skin depth can be estimated by

$$p = \sqrt{\frac{2}{\omega\mu_0\sigma}} \approx \sqrt{\frac{\rho}{f}} \cdot 500, \quad (3.37)$$

where $f = \frac{\omega}{2\pi}$ is the frequency.

Although Eq. 3.37 is only exact for a homogeneous half space, it is also often used in more complex environments to get a first idea of the penetration depth at a given frequency. The general range of the penetration depths in MT studies is ~ 160 m to more than 500 km. Higher frequency waves give information on the shallow Earth, whereas deeper information is provided by low-frequency waves.

3.3 1D layered Earth response

For a further step, now a layered Earth model is considered. Fig. 3.2 shows a 1D n -layered Earth model with different thicknesses h_i and conductivities σ_i , $i = 1, \dots, n$.

In each layer, the partial differential equations 3.22 and 3.23 are still satisfied by electric field and magnetic field, respectively. The electric field can then be expressed as:

$$E_x^i = E_0^i e^{-\alpha_i z} + E_1^i e^{\alpha_i z}, \quad i = 1, \dots, n \quad (3.38)$$

where $\alpha_i = \sqrt{i\omega\mu_0\sigma_i}$. In this case, $E_1^i \neq 0$, for $i = 1, \dots, n-1$, since each layer has limited thickness such that z cannot be arbitrary large, while $E_1^n = 0$ because the lowermost layer is a half space.

By using $\cosh(x) = (e^x + e^{-x})/2$ and $\sinh(x) = (e^x - e^{-x})/2$, Eq. 3.38 can be rewritten as:

$$E_x^i = u_i \cosh(\alpha_i z) + v_i \sinh(\alpha_i z). \quad (3.39)$$

From Maxwell equation 3.28, the magnetic field can be expressed as:

$$B_y^i = -\frac{1}{i\omega} \frac{\partial E_x}{\partial z} = -\frac{\alpha_i}{i\omega} (u_i \sinh(\alpha_i z) + v_i \cosh(\alpha_i z)). \quad (3.40)$$

Let us assume:

$$Q_i = -\mu_0 \frac{E_x^i}{i\omega B_y^i} = -\frac{Z_i}{i\omega}, \quad (3.41)$$

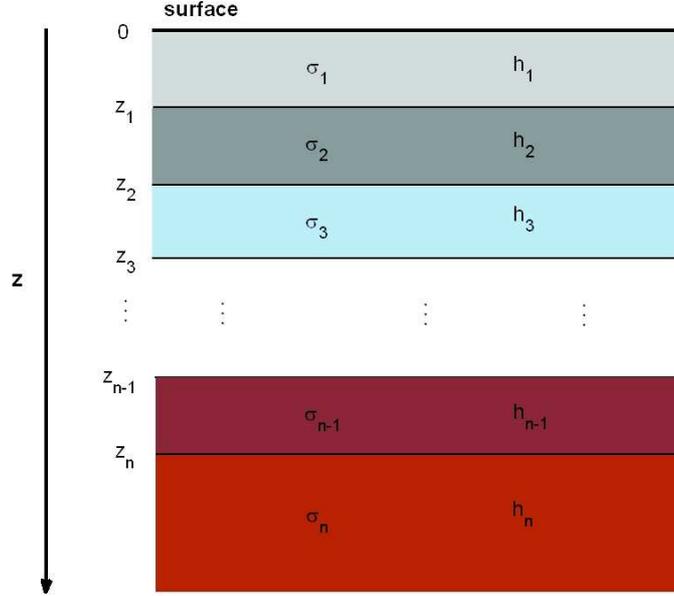


Figure 3.2: The 1D n-layered half space.

then one obtains:

$$Q_i = \frac{\mu_0(u_i \cosh(\alpha_i z) + v_i \sinh(\alpha_i z))}{\alpha_i(u_i \sinh(\alpha_i z) + v_i \cosh(\alpha_i z))}. \quad (3.42)$$

At the top of the i th layer $z = z_{i-1}$, one has:

$$Q_i(z_{i-1}) = \frac{\mu_0(u_i \cosh(\alpha_i z_{i-1}) + v_i \sinh(\alpha_i z_{i-1}))}{\alpha_i(u_i \sinh(\alpha_i z_{i-1}) + v_i \cosh(\alpha_i z_{i-1}))}. \quad (3.43)$$

Similarly, at the bottom of the i th layer $z = z_i$, one has:

$$Q_i(z_i) = \frac{\mu_0(u_i \cosh(\alpha_i z_i) + v_i \sinh(\alpha_i z_i))}{\alpha_i(u_i \sinh(\alpha_i z_i) + v_i \cosh(\alpha_i z_i))}. \quad (3.44)$$

By re-arranging Eq. 3.44, one can derive:

$$\frac{u_i}{v_i} = \frac{-Q_i(z_i) \cosh(\alpha_i z_i) + \frac{\mu_0}{\alpha_i} \sinh(\alpha_i z_i)}{Q_i(z_i) \sinh(\alpha_i z_i) - \frac{\mu_0}{\alpha_i} \cosh(\alpha_i z_i)}. \quad (3.45)$$

Substituting Eq. 3.45 into Eq.3.43 to eliminate u_i and v_i yields:

$$\begin{aligned}
Q_i(z_{i-1}) &= \frac{\mu_0 \frac{u_i}{v_i} \cosh(\alpha_i z_{i-1}) + \sinh(\alpha_i z_{i-1})}{\alpha_i \frac{u_i}{v_i} \sinh(\alpha_i z_{i-1}) + \cosh(\alpha_i z_{i-1})} \\
&= \frac{\mu_0 - Q_i(z_i) \cosh(\alpha_i(z_i - z_{i-1})) + \frac{\mu_0}{\alpha_i} \sinh(\alpha_i(z_i - z_{i-1}))}{\alpha_i Q_i(z_i) \sinh(\alpha_i(z_i - z_{i-1})) - \frac{\mu_0}{\alpha_i} \cosh(\alpha_i(z_i - z_{i-1}))} \quad (3.46) \\
&= \frac{\mu_0 - Q_i(z_i) + \frac{\mu_0}{\alpha_i} \tanh(\alpha_i h_i)}{\alpha_i Q_i(z_i) \tanh(\alpha_i h_i) - \frac{\mu_0}{\alpha_i}}
\end{aligned}$$

Both the horizontal electric and magnetic field are continuous across the layer boundary, hence their ratio Q has to be continuous as well such that:

$$Q_i(z_i) = Q_{i+1}(z_i). \quad (3.47)$$

Inserting Eq. 3.47 into Eq. 3.46 results in a recursive formula:

$$Q_i(z_{i-1}) = \frac{\mu_0 - Q_{i+1}(z_i) + \frac{\mu_0}{\alpha_i} \tanh(\alpha_i h_i)}{\alpha_i Q_{i+1}(z_i) \tanh(\alpha_i h_i) - \frac{\mu_0}{\alpha_i}}. \quad (3.48)$$

The lowermost layer is a homogeneous half space and according to Eq. 3.30, one obtains:

$$Q_n = -\mu_0 \frac{E_x^n}{i\omega B_y^n} = -\frac{\mu_0}{\alpha_n}. \quad (3.49)$$

The impedance Z on top of a n -layered model can then be calculated recursively (starting from the lowest layer and ending at the first layer) by using formula Eq. 3.48.

While in a homogeneous half space, the apparent resistivity is exactly the resistivity of the half space (see Eq. 3.35), the apparent resistivity of a layered Earth model is approximately an average of the Earth's resistivity within the penetration depth of different frequency. The phase is furthermore not constant, but changes with frequency. As an example, a 1D layered Earth model consisting of two horizontal layers and a homogeneous half-space with resistivities of $10 \Omega m$, $1 \Omega m$ and $1000 \Omega m$, and thickness of 1 km, 2 km, respectively, is shown on left panel of Fig. 3.3. The apparent resistivities and phases (ϕ_{xy}) derived from impedances, which are calculated from above recursive formulae for a frequency range of 10^{-6} Hz to 0.5 Hz, are shown on right panel of Fig. 3.3. From high frequency to low frequency (from shallow to deep), the apparent resistivity curve changes continuously from $10 \Omega m$ (first layer) to $1 \Omega m$ (second layer) then to $1000 \Omega m$ (third layer), while the phases are larger than 45° in higher frequency range which is consistent with the decrease in resistivity

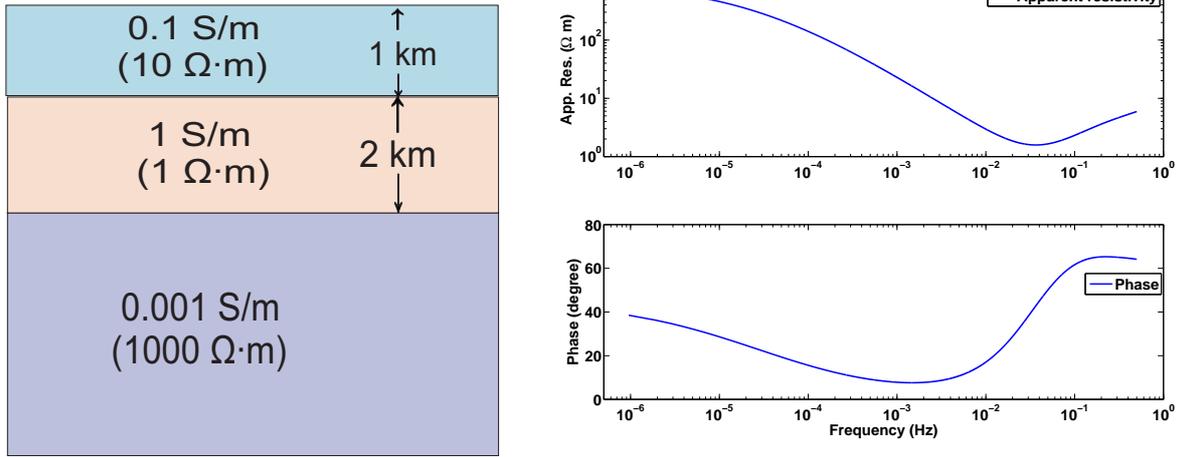


Figure 3.3: Left panel: A horizontally layered model with three layers. Right panel: The corresponding apparent resistivity & phase curves of the model.

from the first layer to the second layer, then decrease below 45° in lower frequency range which is consistent with the increase in resistivity from the second layer to the third layer [e.g. Parker, 1983, Simpson and Bahr, 2005].

3.4 Simple 2-D Earth and E- and B-polarization

Now a more realistic 2-D Earth model is considered, where resistivity changes with depth and one lateral direction. Fig. 3.4 shows a very simple 2D Earth model consisting of a vertical contact between two zones with different conductivity, σ_1 and σ_2 , and extending to infinity along x-direction.

For such 2-D models, for which the electric and magnetic fields vary in the z- and y-directions, independent of the x coordinate (i.e., $\partial/\partial x = 0$), Maxwell equations 3.27 and 3.28 can be simplified by separating them into two sets of self-consistent, differential equations. One set describes currents flowing along strike direction and relates E_x , B_y and B_z with each other. This mode is referred to as *E-polarization*, *transverse electric* or *TE mode* (see Fig. 3.4).

$$\left. \begin{aligned} \frac{\partial E_x}{\partial y} &= i\omega B_z \\ \frac{\partial E_x}{\partial z} &= -i\omega B_y \\ \frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z} &= \mu_0 \sigma E_x \end{aligned} \right\} \text{TE mode} \quad (3.50)$$

The another set describes currents flowing along the $z - y$ plane and relates E_y , E_z and B_x with each other. This mode is referred to as *B-polarization*, *transverse magnetic* or *TM mode*

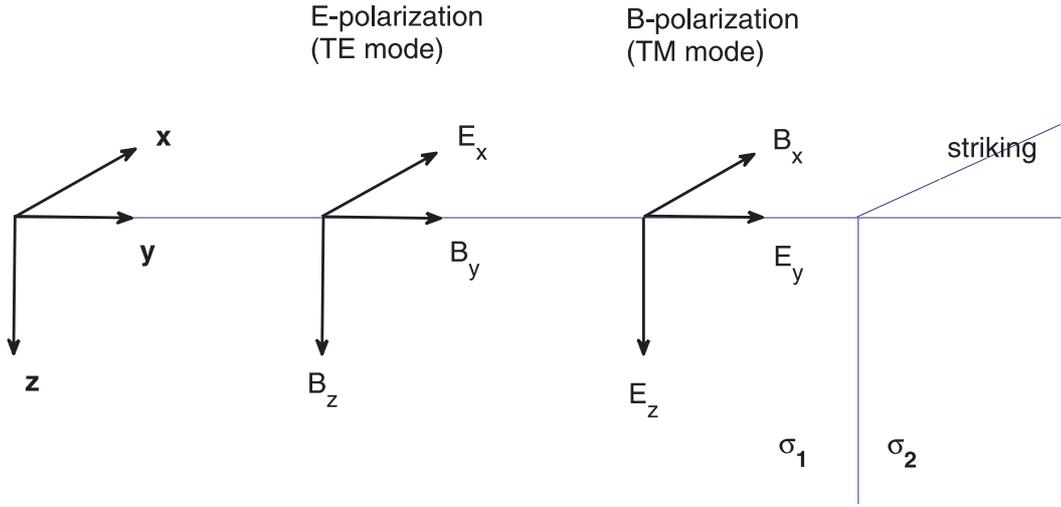


Figure 3.4: Simple 2D model composed of two vertical contacted zones with different conductivities σ_1 and σ_2 . The striking is oriented in x-direction.

(see Fig. 3.4).

$$\left. \begin{aligned} \frac{\partial B_x}{\partial y} &= -\mu_0 \sigma E_z \\ \frac{\partial B_x}{\partial z} &= \mu_0 \sigma E_y \\ \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} &= -i\omega B_x \end{aligned} \right\} \text{TM mode} \quad (3.51)$$

The Maxwell equations for both modes can be combined such that two independent diffusion equations are obtained:

$$\begin{aligned} \frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} &= i\omega \mu_0 \sigma E_x & \text{TE mode,} \\ \frac{\partial^2 B_x}{\partial y^2} + \frac{\partial^2 B_x}{\partial z^2} &= i\omega \mu_0 \sigma B_x & \text{TM mode.} \end{aligned} \quad (3.52)$$

For 2D Earth model in which currents are parallel (TE mode) or perpendicular (TM mode) to the strike direction, the diagonal elements of impedance tensor remain zero ($Z_{xx} = Z_{yy} = 0$), and the off-diagonal elements now can be considered as two independent measurements for the underlying resistivity structure, $Z_{xy} \neq -Z_{yx}$ [e.g., Simpson and Bahr, 2005]. Eq. 3.32 can then be written as :

$$\left(\begin{array}{c} E_x \\ E_y \end{array} \right) = \frac{1}{\mu_0} \left(\begin{array}{cc} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{array} \right) \left(\begin{array}{c} B_x \\ B_y \end{array} \right) \quad \text{and} \quad \left. \begin{array}{l} Z_{xx} = Z_{yy} = 0 \\ Z_{xy} \neq Z_{yx} \end{array} \right\} \text{2-D.} \quad (3.53)$$

In a more general case, when electric and magnetic fields are not oriented parallel or perpendicular to the strike direction, the impedances Z_{xx} and Z_{yy} are not zero and always different from each other, but have opposite sign, whereas Z_{xy} and Z_{yx} are different. Under ideal conditions, by rotating the impedance tensor, one can find the strike direction in which the condition that $Z_{xx} = Z_{yy} = 0$ is satisfied. However, in practice with measured data, this is often not possible. Instead, the strike direction can be determined by rotating the impedance tensor to an angle such that the diagonal elements are minimized.

3.5 3D case

Generally, the conductivity distribution in the Earth is three dimensional, that is resistivity changes with depth and in both lateral directions. For such 3D cases, the corresponding electric and magnetic fields are no longer mutually orthogonal and do not decouple in two modes for fields aligned in or orthogonal to strike direction. All impedance tensor elements are now non-zero and differ from each other:

$$\left. \begin{aligned} \begin{pmatrix} E_x \\ E_y \end{pmatrix} = \frac{1}{\mu_0} \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} \begin{pmatrix} B_x \\ B_y \end{pmatrix} \quad \text{and} \quad \begin{matrix} Z_{xx} \neq Z_{yy} \\ Z_{xy} \neq Z_{yx} \end{matrix} \right\} \text{3-D.} \quad (3.54)$$

In MT surveys, sometimes it is difficult to determine in which dimension the collected data should be interpreted. When the MT data measured at sites with 2D or 3D structures are interpreted as 1D models, or data measured at sites with 3D structures are interpreted as 2D models, that may yield unreliable models and interpretations. A way to solve such problems is the dimensionality analysis, which has been investigated for several decades (see [e.g. Swift, 1967, Word et al., 1970, Bahr, 1988] and the review paper by Ledo et al. [2002]). Through dimensionality analysis of the impedance tensor one can approximately determine the dimensionality of the subsurface electrical conductivity anomalies/structures, and provides a more reliable imaging of the subsurface.

Chapter 4

From Measured Time Series to Impedance Tensor

MT exploration surveys generally include the design of MT survey appropriate for targets, collecting MT data for a period of time, processing raw time-series data to obtain frequency based impedance tensor and interpreting impedance tensor to create a subsurface resistivity model. In this chapter, the discussion will concentrate on one aspect of the MT method - the data processing. The manner in which data are acquired and the problems of designing suitable MT experiments will be shortly considered in chapter 7.

In order to obtain a reliable subsurface resistivity structure model, data processing is a very important interpretation step in MT studies. In traditional MT data processing methods, Fourier analysis is used to transfer measured time series into frequency domain to estimate impedance tensor. One underlying assumption for Fourier analysis is that the analyzed segments of time series are stationary. However time series used in MT always contain noise due to the configuration and accuracy of the measurement equipment, some artificial noise, and additionally short time scale episodes of non-stationary source signal. This causes a bias problem in impedance estimates and makes the estimation procedure fundamentally a statistical process. To avoid the bias, more sophisticated methods such as using remote reference time series and various robust estimate methods are involved. However, sometimes even these more sophisticated methods fail to give physically interpretable responses ([e.g. Szarka, 1988, Qian and Pedersen, 1991, Schultz et al., 1993, Junge, 1996, Garcia et al., 1997, Egbert et al., 2000]). In this chapter, traditional MT data processing methods and their technical problems and limitations will be discussed in more detail. To address these limitations, a new time series processing method which does not require stationary - Empirical Mode Decomposition will be used, which will be discussed in more detail in the next chapter.

4.1 Short review of MT data processing

As discussed in the previous chapter, the transfer function between the simultaneously measured horizontal electric and magnetic fields theoretically only depends on the resistivity distribution in the Earth. In the absence of noise and with precise data, the frequency-dependent tensor relationship between the electric and magnetic fields can be written as (see Eq. 3.53):

$$\mu_0 \underline{\mathbf{E}} = \underline{\mathbf{Z}} \underline{\mathbf{B}}. \quad (4.1)$$

In an ideal situation the calculation of the transfer function is simple and straightforward. Transform electromagnetic fields time series from time domain into frequency domain and then calculate the power spectra densities to obtain impedance tensor $\underline{\mathbf{Z}}$. However in practice such a simple data precessing procedure will fail, since the measured data generally contain noise which in a general sense can be considered as any electromagnetic disturbances that violate the source assumption made in MT and hence do not satisfy Eq. 4.1.

Therefore, in reality it becomes necessary to estimate both impedance and its uncertainty in a statistical manner. The classical approach to determine $\underline{\mathbf{Z}}$ is to solve a linear regression problem by minimize the residual \mathbf{r} in a least square sense:

$$\min \quad |\mathbf{r}|^2 = |\mu_0 \underline{\mathbf{E}} - \underline{\mathbf{Z}} \underline{\mathbf{B}}|^2, \quad \implies \quad \underline{\mathbf{Z}} = \mu_0 (\underline{\mathbf{E}} \underline{\mathbf{B}}^*) (\underline{\mathbf{B}} \underline{\mathbf{B}}^*)^{-1}, \quad (4.2)$$

where the superscript $*$ denotes the complex conjugate. Terms in parentheses are the cross-power and auto-power spectral densities of the field data. When the errors are independent and identical Gaussian distributed, a least square estimate of $\underline{\mathbf{Z}}$ in Eq. 4.2 gives a maximum likelihood, minimum variance, unbiased estimate.

The commonly used spectral analysis methods is Fourier transform. However, in order to obtain reliable spectra estimates, prior to a Fourier transform, first the time series are visually inspected to identify unusual segments with severe noise parts that need to be removed. For example spikes will distort all frequency estimates since their frequency contents cover the entire spectrum. In addition, the time series need to be pre-conditioned to reduce the effect of the finite observation length and trends, e.g., pre-whitening, detrending and windowing the time series. Pre-whitening is generally achieved by time domain filtering with a short autoregressive (AR) sequence fit to the time series. Detrending the time series is used to prevent that Fourier transform injects a significant amount of power at the frequency of the long term variations. Multiplication of the time series with a window function can effectively avoid spectral leakage. Conventionally used Hamming, Hanning, Parzen etc windows can perform a reasonable job. Thomson et al. [1976] has shown that for the geomagnetic data the best windows are "prolate spheroidal" windows, which offer the greatest side-lobe protection.

The general Fourier-based estimation procedure can be shortly summarized as following:

1. Depending on the longest period needed, cut the time-series in segments of equal length.
2. Multiply each segment with a suitable window function.
3. Transform each segment into frequency domain using Fourier transform.
4. Calculate the auto- and cross-power spectral densities of E - and B - fields for each segment.
5. Estimate the resulting transfer function and the associated error.

The classical Fourier-based spectral analysis procedures and least square regression used in Eq. 4.2 are derived from simple, wide-sense stationary and Gaussian models. However,

However, magnetotelluric data seldom fulfill this condition (e.g., Chave et al. [2012, p. 168 - 169]). Non-stationary phenomena in MT measurements due to the possibility of sudden spatial and temporal variations in the naturally occurring source field and some measurement errors can destroy conventional spectral estimates. This is the motivation to develop methods which are robust and relatively insensitive to non-stationarity and extreme points (outliers in electric field data and leverage points in magnetic field data, will be discussed in more detail in section 4.5 and 4.6). In the last thirty years, three kind of developments have dramatically improved the Fourier based impedance estimations.

First, Gamble et al. [1979a,b] developed the remote reference method, in which auxiliary MT measurements from a second location, consisting usually of horizontal magnetic field data, are used to minimize bias in impedance estimates. It uses the fact that in most instances noise in the magnetic field and the reference field are uncorrelated, such that remote referencing may be used to significantly improve the transfer function estimates over conventional approaches.

Second, to eliminate the influence of the outliers in electric field data, a variety of data-adaptive robust M-estimators have been developed [e.g. Egbert and Booker, 1986, Chave et al., 1987, Chave and Thomson, 1989, Larsen, 1989, Sutarno and Vozoff, 1989, 1991, Larsen et al., 1996, Egbert and Livelybrooks, 1996, Egbert, 1997, Oettinger et al., 2001, Smirnov, 2003]. The performance of these robust methods was compared by Jones et al. [1989]. Moreover, the bounded influence function has been implemented in impedance estimation by Chave and Thomson [2003, 2004] to reduce the influence of leverage points in magnetic field data. The robust M-estimators and bounded influence estimators have now become standard in MT data processing.

Third, non-parametric jackknife or bootstrap method have been introduced ([Chave and Thomson, 1989]) to improve conventional error estimates in the MT response tensor that are often biased due to the inaccuracies of parametric estimates.

In contrast to traditional processing methods mentioned above which estimate the impedance tensor directly in frequency domain, Kunetz [1972], Spagnolini [1994] used estimators in the time domain based on the convolution theorem and, Berdichevsky et al. [1973] proposed an impedance tensor estimate in the time-frequency-domain, in which the impedance can be estimated using transient or instantaneous spectra. Furthermore, instead of Fourier transform, Chant and Hastie [1992, 1993] used time-frequency distribution (TFD) analysis techniques to obtain the impedance tensor estimates and Garcia and Jones [2008] demonstrated a robust MT data processing in audiomagnetotelluric dead band (1 kHz - 5 kHz) by using the continuous wavelet transform.

In subsequent sections these methods are described in more detail. The fundamental problem of spectral estimation is shortly introduced in section 2. In section 3, simple least square estimates of the impedance and the bias problem are described. Using remote reference method to avoid bias is discussed in section 4. Robust M-estimators and bounded influence estimators are summarized in section 5 and 6 respectively. Section 7 describes the error estimate in impedance tensor using parametric and non-parametric methods. The bounded influence remote reference processing (BIRRP) code is shortly summarized in section 8. Impedance estimate in time-frequency domain [Berdichevsky et al., 1973] is introduced in section 9. A

short summary of MT data processing is given in last section.

4.2 Spectral estimation

The first step and an important aspect in MT data processing is to calculate the spectral densities of the raw E - and B -fields data. Traditionally, the Fourier transform is used and the most basic method is the nonparametric method, *periodogram*, which consists of a mean spectral densities of different segments of the time series.

$$S(\omega) = \frac{1}{L} \sum_{j=1}^L \left(\frac{1}{M} \left| \sum_{n=1}^M x_j(t_n) e^{-i\omega t_n} \right|^2 \right), \quad (4.3)$$

where M is the length of segments, L is the number of segments which can be either completely separated or partially overlap. A periodogram is a natural estimator and easy to implement. However, it has an inconsistent variance because of the sharp truncation of the sequence, and suffers from spectra bias problems due to numerous side-lobes which lead to spectral leakage. Multiplying the segments with a window function w before spectra calculation (a form of modified periodogram [e.g. Bartlett, 1967, Welch, 1967]) can improve the estimation results:

$$S(\omega) = \frac{1}{L} \sum_{j=1}^L \left(\frac{1}{M} \left| \sum_{n=1}^M x_j(t_n) w_n e^{-i\omega t_n} \right|^2 \right). \quad (4.4)$$

Furthermore, to obtain a smoothed estimate, the spectral densities may be convoluted with a second window G in the frequency domain,

$$\tilde{S}(\omega) = S(\omega) * G(\omega). \quad (4.5)$$

The two operations are generally considered to be unrelated. While the data window w mainly controls the bias, the smoothing window G controls the variance.

Thomson [1982] introduced a *multitaper spectral estimator* which improves the spectra bias problems of conventional Fourier analysis and can produce more accurate spectral estimates. In the multitaper spectral estimator, the variance is controlled by averaging many approximately independent direct spectral estimates. These estimates are constructed using the mutually orthogonal discrete prolate spheroidal sequences. The zeroth-order sequence is used to obtain direct spectral estimates with minimum leakage, and higher-order sequences are used to compute the spectral estimates which are included in the average to obtain the multitaper spectral estimates:

$$\hat{S}^{(mt)}(\omega) = \frac{1}{L} \sum_{j=1}^L \left[\frac{1}{K} \sum_{k=1}^K \left(\frac{1}{M} \left| \sum_{n=1}^M x_j(t_n) w_n^{(k)} e^{-i\omega t_n} \right|^2 \right) \right], \quad (4.6)$$

where $w_n^{(k)}$ is the n th element of k th order discrete prolate spheroidal sequence of the window function.

By using the specially designed orthogonal tapers, this estimator reduces optimally spectral leakage and variance in most instances, and is therefore used in some standard MT processing packages (e.g. BIRRP by Chave and Thomson [2004]).

4.3 Least square approach

The frequency domain MT tensor is a 2×2 complex tensor as shown in Eq. 4.1, which consists of a linear system with two inputs and two outputs. As mentioned before, in practice these equations do not hold exactly due to noise. Assuming that one has n segments for both, E - and B -field, time series, which are transformed into frequency domain and denoted by $E_x^{(i)}, E_y^{(i)}, B_x^{(i)}, B_y^{(i)}$ with $i = 1, \dots, n$, Eq. 4.1 is replaced by an equivalent matrix form as (notice that compare to Eq. 4.1 the order of \mathbf{B} and \mathbf{Z} is changed):

$$\mu_0 \begin{pmatrix} E_x^{(1)} & E_y^{(1)} \\ E_x^{(2)} & E_y^{(2)} \\ \vdots & \vdots \\ E_x^{(n)} & E_y^{(n)} \end{pmatrix} = \begin{pmatrix} B_x^{(1)} & B_y^{(1)} \\ B_x^{(2)} & B_y^{(2)} \\ \vdots & \vdots \\ B_x^{(n)} & B_y^{(n)} \end{pmatrix} \cdot \begin{pmatrix} Z_{xx} & Z_{yx} \\ Z_{xy} & Z_{yy} \end{pmatrix} + \mathbf{r} \quad \text{or} \quad \mu_0 \mathbf{E} = \mathbf{BZ} + \mathbf{r}, \quad (4.7)$$

where \mathbf{r} is $n \times 2$ residual vector.

Impedance tensor \mathbf{Z} can be estimated in the least square sense by minimizing $|\mathbf{r}|^2 = \mathbf{r}^* \mathbf{r}$. Let's consider the equations involving $E_x^{(i)}$ as an example (the derivation of equations involving $E_y^{(i)}$ is analog). Z_{xx} and Z_{xy} can be estimated by

$$\begin{aligned} \min \quad & \sum_{i=1}^n |r^{(i)}|^2 = \sum_{i=1}^n |\mu_0 E_x^{(i)} - Z_{xx} B_x^{(i)} - Z_{xy} B_y^{(i)}|^2 \\ & = \sum_{i=1}^n (\mu_0 E_x^{(i)*} - Z_{xx}^* B_x^{(i)*} - Z_{xy}^* B_y^{(i)*}) \cdot (\mu_0 E_x^{(i)} - Z_{xx} B_x^{(i)} - Z_{xy} B_y^{(i)}). \end{aligned} \quad (4.8)$$

Setting the derivatives of the equation with respect to the real and imaginary parts of Z_{xx} and Z_{xy} to zero will minimize the error $|\mathbf{r}|^2$ and yield:

$$\mu_0 \sum_{i=1}^n E_x^{(i)} B_x^{(i)*} = Z_{xx} \sum_{i=1}^n B_x^{(i)} B_x^{(i)*} + Z_{xy} \sum_{i=1}^n B_y^{(i)} B_x^{(i)*}, \quad (4.9)$$

and

$$\mu_0 \sum_{i=1}^n E_x^{(i)} B_y^{(i)*} = Z_{xx} \sum_{i=1}^n B_x^{(i)} B_y^{(i)*} + Z_{xy} \sum_{i=1}^n B_y^{(i)} B_y^{(i)*}. \quad (4.10)$$

Notice that the summations in Eqs. 4.9 and 4.10 represent the auto- and cross-power spectral densities. Strictly speaking, these two equations are valid only at a given frequency. However, in practice, the components of Z are slowly varying functions of frequency, therefore the summations may be taken as the averages of auto- and cross-power spectral densities over some finite frequency bandwidth. Let " $\overline{\quad}$ " denote the average, Eqs. 4.9 and 4.10 can then be rewritten as:

$$\mu_0 \overline{E_x B_x^*} = Z_{xx} \overline{B_x B_x^*} + Z_{xy} \overline{B_y B_x^*}, \quad (4.11)$$

and

$$\mu_0 \overline{E_x B_y^*} = Z_{xx} \overline{B_x B_y^*} + Z_{xy} \overline{B_y B_y^*}. \quad (4.12)$$

Z_{xx} and Z_{xy} derived from Eqs. 4.11 and 4.12 will minimize the error caused by noise on E_x . It is also possible to define other least square estimates as in Eq. 4.8 which minimize other types of noise. For example, if one takes

$$\min \sum_{i=1}^n \left| \mu_0 \frac{E_x^{(i)}}{Z_{xx}} - B_x^{(i)} - \frac{Z_{xy}}{Z_{xx}} B_y^{(i)} \right|^2, \quad (4.13)$$

the error introduced by noise on B_x will be minimized.

Therefore, taking into consideration noise on E_x , E_y , B_x and B_y respectively, one can obtain four distinct equations. Besides Eqs. 4.11 and 4.12, they are:

$$\mu_0 \overline{E_x E_x^*} = Z_{xx} \overline{B_x E_x^*} + Z_{xy} \overline{B_y E_x^*}, \quad (4.14)$$

and

$$\mu_0 \overline{E_x E_y^*} = Z_{xx} \overline{B_x E_y^*} + Z_{xy} \overline{B_y E_y^*}. \quad (4.15)$$

Any two of the four Eqs. 4.11, 4.12, 4.14 and 4.15, can be used to estimate the impedance, which leads to six possible estimates for each of the four tensor elements [Sims et al., 1971]. For example, the six estimates for Z_{xy} are:

$$Z_{xy} = \mu_0 \frac{(\overline{B_x B_x^*})(\overline{E_x B_y^*}) - (\overline{B_x B_y^*})(\overline{E_x B_x^*})}{(\overline{B_x B_x^*})(\overline{B_y B_y^*}) - (\overline{B_x B_y^*})(\overline{B_y B_x^*})}, \quad (4.16)$$

$$Z_{xy} = \mu_0 \frac{(\overline{B_x E_x^*})(\overline{E_x E_y^*}) - (\overline{B_x E_y^*})(\overline{E_x E_x^*})}{(\overline{B_x E_x^*})(\overline{B_y E_y^*}) - (\overline{B_x E_y^*})(\overline{B_y E_x^*})}, \quad (4.17)$$

$$Z_{xy} = \mu_0 \frac{(\overline{B_x E_x^*})(\overline{E_x B_x^*}) - (\overline{B_x B_x^*})(\overline{E_x E_x^*})}{(\overline{B_x E_x^*})(\overline{B_y B_x^*}) - (\overline{B_x B_x^*})(\overline{B_y E_x^*})}, \quad (4.18)$$

$$Z_{xy} = \mu_0 \frac{(\overline{B_x E_x^*})(\overline{E_x B_y^*}) - (\overline{B_x B_y^*})(\overline{E_x E_x^*})}{(\overline{B_x E_x^*})(\overline{B_y B_y^*}) - (\overline{B_x B_y^*})(\overline{B_y E_x^*})}, \quad (4.19)$$

$$Z_{xy} = \mu_0 \frac{(\overline{B_x E_y^*})(\overline{E_x B_x^*}) - (\overline{B_x B_x^*})(\overline{E_x E_y^*})}{(\overline{B_x E_y^*})(\overline{B_y B_x^*}) - (\overline{B_x B_x^*})(\overline{B_y E_y^*})}, \quad (4.20)$$

and

$$Z_{xy} = \mu_0 \frac{(\overline{B_x E_y^*})(\overline{E_x B_y^*}) - (\overline{B_x B_y^*})(\overline{E_x E_y^*})}{(\overline{B_x E_y^*})(\overline{B_y B_y^*}) - (\overline{B_x B_y^*})(\overline{B_y E_y^*})}. \quad (4.21)$$

In the case of absence of noise, all six impedance estimates are equivalent. However, in the presence of noise, each estimate is different. The difference arises from the combination

of random errors in the cross- and auto-power densities. For example, assuming that the measured \tilde{B}_x is the real B_x plus noise ε (uncorrelated with B_x):

$$\tilde{B}_x = B_x + \varepsilon \quad \text{yields} \quad \tilde{B}_x \tilde{B}_x^* = B_x B_x^* + \varepsilon \varepsilon^*, \quad (4.22)$$

The auto-power density of the measured data $\tilde{B}_x \tilde{B}_x^*$ is systematically larger than that of the real data $B_x B_x^*$ as the noise auto-power density increases, which will lead to a downward bias of the estimator in Eq. 4.16. On the other hand noise cross-power density is cancelled out if the noise in different channels is uncorrelated.

Among the six estimators, the estimation based on Eq. 4.16 is downward biased by random noise on the magnetic field signals but is unaffected by random noise on the electric field signals. Estimation based on Eq. 4.18 is upward biased by random noise on the electric field signals but is unaffected by random noise on the magnetic field signals. Estimates derived from the remaining four estimators exhibit a mixed bias. It is obvious that downward and upward biased estimates give an envelope within which the true impedance should lie.

Typically in MT application, the electric fields data are much noisier than the magnetic fields data, therefore, Eq. 4.16 generally gives both the least biased and the statistically best estimate. Accordingly, the corresponding estimate in the matrix form of Eq. 4.7 is usually given by:

$$\mathbf{Z} = \mu_0 (\mathbf{B}^* \mathbf{B})^{-1} (\mathbf{B}^* \mathbf{E}). \quad (4.23)$$

4.4 Remote reference method

To address the bias problems associated with impedance estimation in the presence of noise, Gamble et al. [1979a,b] introduced a method that uses additional correlated magnetic field time series to yield estimates unbiased of noise, called *remote reference method*.

Under the assumptions of homogeneous source and planar wave, in a purely layered Earth model the magnetic field should be equal at any recording site. In the more general Earth model, magnetic field components B_{xr} and B_{yr} measured at a reference station should be coherent with B_x and B_y which are simultaneously measured at a local MT station. Furthermore, the noise in local and reference stations should be uncorrelated. For example, assuming that the measured \tilde{B}_x and \tilde{B}_{xr} are the real B_x and B_{xr} plus noise: $\tilde{B}_x = B_x + \varepsilon$ and $\tilde{B}_{xr} = B_{xr} + \varepsilon_r$, and ε and ε_r are uncorrelated, one can calculate the cross-power density of B_x and B_{xr} instead of the auto-power density calculated in Eq. 4.22:

$$\tilde{B}_x \tilde{B}_{xr}^* = B_x B_{xr}^* + B_x \varepsilon_r^* + \varepsilon B_{xr}^* + \varepsilon \varepsilon_r^* = B_x B_{xr}^*. \quad (4.24)$$

Now the cross-power density of measured data $\tilde{B}_x \tilde{B}_{xr}^*$ is equal the cross-power density of real data $B_x B_{xr}^*$.

Thus if B_x^* and B_y^* are replaced with B_{xr}^* and B_{yr}^* respectively in Eq. 4.16, the noise induced bias vanishes. The remote reference estimator for Z_{xy} is then given by:

$$Z_{xy} = \mu_0 \frac{(\overline{B_x B_{xr}^*})(\overline{E_x B_{yr}^*}) - (\overline{B_x B_{yr}^*})(\overline{E_x B_{xr}^*})}{(\overline{B_x B_{xr}^*})(\overline{B_y B_{yr}^*}) - (\overline{B_x B_{yr}^*})(\overline{B_y B_{xr}^*})}. \quad (4.25)$$

The estimators for Z_{xx} , Z_{yx} and Z_{yy} may be derived in analogue to Eq. 4.25.

The corresponding matrix form in Eq. 4.23 now becomes:

$$\mathbf{Z} = \mu_0(\mathbf{B}_r^*\mathbf{B})^{-1}(\mathbf{B}_r^*\mathbf{E}). \quad (4.26)$$

The remote reference estimator is unbiased if the noise in the local and remote field data is uncorrelated and the coherence between the real B and B_r is high.

In practice, the choice of a reference site, ensuring lack of correlation of the noise at the two sites, depends on the local geology, the source field behavior, instrument noise and cultural noise. The horizontal wavelength of the MT source is much larger than the lateral induction scale length, therefore, for long-period data, stations up to several hundred kilometers apart are usually suitable as remote reference sites. However it should be noticed that remote reference results are not minimum variance and are always associated with larger errors than standard least-square ones [Jones et al., 1989]. Also, if the noise in the local and the reference sites are correlated, remote reference processing will be characterized by a bias.

While the ordinary remote reference method uses one reference site time series, Egbert [1997], Chave and Thomson [2004] introduced a generalization that may consist of many reference channels. This generalization can improve MT impedance estimates if different channels become noisy at different times.

4.5 Robust estimation I: M-estimator

The estimators obtained by the least square method as described in the previous sections are based on Gaussian maximum likelihood approach. The least square method yields the best unbiased linear estimate if errors are uncorrelated and share a common variance, independent of any distributional assumptions. The advantages of this method is its simplicity and the optimal properties established by the Gauss-Markov theorem. However most data contain a small fraction of unusual observations that do not fit the model distribution, such as point defects caused by transient instruments failures, spike noise from lightening or other local non-stationarity. In linear regression, these unusual observations are classified into two types of point: outliers and leverage points. An outlier is an observation which has normal independent variable value (generally in B -field in a MT context) but unusual dependent variable value (generally in E -field in a MT context), while a leverage point is an observation that has unusual independent variable value. Least square methods are very sensitive to these anomalous points such that sometimes the obtained estimates are seriously biased, which led to the development of the robust MT data processing procedures. In this section the robust M-estimator which minimizes the influence of outliers will be considered. Bounded influence estimator which minimizes the influence of leverage points will be discussed in the next section.

Generally, there are two major types of robust procedures which are less sensitive to outliers, the L-estimators and M-estimators [e.g. Huber, 1981]. The L-estimator is based on combinations of the order statistics, but suffers from the fact that the result is not data adaptive because the truncation point must be fixed in advance. The suitable robust estimator for MT data is the M-estimator which is a variant of maximum likelihood method.

The M-estimator is similar to the least square method. Both minimize a norm of the random errors as in Eq. 4.2. However the M-estimator is based on a measure which avoids that a few extreme values (outliers) in the electric field time series dominate the result. It is obtained by solving

$$\min \sum_{i=1}^n \rho\left(\frac{r_i}{d}\right) \quad \text{with} \quad (r_1 \ r_2 \ \cdots \ r_n)^T = \mathbf{r} = \mu_0 \mathbf{E} - \mathbf{BZ}, \quad (4.27)$$

where the superscript T denotes the transpose of the matrix, d is a scalar parameter which normalizes the residuals r_i and ρ is a suitable loss function which is a measure of the distance between the true and estimated values of the impedance. If the loss function $\rho(x)$ is chosen as $x^2/2$, Eq. 4.27 reduces to standard least square. If $\rho(x)$ is chosen to be $-\log[f(x)]$, where $f(x)$ is the probability density function of the residuals, the M-estimator is a maximum likelihood estimator.

Analog to normal least square solution (by setting the derivative of $\sum_{i=1}^n \rho(r_i/d)$ with respect to the real and imaginary parts of \mathbf{Z} to zero), the solution of the minimization problem in Eq. 4.27 can be obtained:

$$\sum_{i=1}^n \Psi(r_i/d) \frac{\partial r_i}{\partial Z} = \sum_{i=1}^n \Psi(r_i/d) B_i = 0, \quad \text{or in matrix form} \quad \mathbf{B}^* \mathbf{\Psi} = 0, \quad (4.28)$$

where $\Psi(x) = \partial \rho(x)/\partial x$ and $\mathbf{\Psi} = (\Psi(r_1/d) \ \Psi(r_2/d) \ \cdots \ \Psi(r_n/d))^T$. Generally, while Eq. 4.28 are nonlinear equations, it is easiest to rewrite them as weighted least square problems and to get a linear approximation.

Let us define the weights $w_i = \Psi(r_i/d)/r_i$, $i = 1, 2, \dots, n$ and $n \times n$ diagonal weight matrix $\mathbf{w} = \text{diag}(w_1 \ w_2 \ \cdots \ w_n)$, then Eq. 4.28 becomes:

$$\sum_{i=1}^n w_i r_i \frac{\partial r_i}{\partial Z} = \sum_{i=1}^n w_i r_i B_i = 0 \quad \text{or} \quad \mathbf{B}^* \mathbf{w} \mathbf{r} = \mathbf{B}^* \mathbf{w} (\mu_0 \mathbf{E} - \mathbf{BZ}) = 0. \quad (4.29)$$

Eq. 4.29 is exactly the equation which can be obtained if one solves the weighted minimization problem: $\min \sum_{i=1}^n w_i r_i^2$. The weights are chosen to minimize the influence of the data with large residuals, and can be computed iteratively based on the residuals and scale estimate from the previous iteration. The initial value can be chosen as the obtained values from the ordinary least square estimate.

Therefore, the robust impedance estimator is given by:

$$\mathbf{Z} = \mu_0 (\mathbf{B}^* \mathbf{w} \mathbf{B})^{-1} (\mathbf{B}^* \mathbf{w} \mathbf{E}). \quad (4.30)$$

There are many possible choices for the loss function ρ . The most widely used one is Huber loss function defined as [Huber, 1981]:

$$\rho(x) = \begin{cases} x^2/2 & |x| \leq \alpha \\ \alpha|x| - \alpha^2/2 & |x| > \alpha \end{cases}. \quad (4.31)$$

The choice $\alpha = 1.5$ gives better than 95% efficiency with outlier-free Gaussian data.

Its derivative, function Ψ , is given by:

$$\Psi(x) = \begin{cases} x & |x| \leq \alpha \\ \alpha \operatorname{sgn}(x) & |x| > \alpha \end{cases}, \quad (4.32)$$

and the corresponding weight matrix have diagonal elements

$$w_i = \begin{cases} 1 & |x| \leq \alpha \\ \alpha/|x| & |x| > \alpha \end{cases}. \quad (4.33)$$

In each iteration, when $|r_i/d| > \alpha$, the residuals are regarded as large and are down-weighted. The scale parameter d must be chosen carefully, Egbert and Booker [1986] chose d as a robust estimate of the residual standard deviation, Chave et al. [1987], Chave and Thomson [2004] give two choices of d using the median absolute deviation and interquartile distance (the difference between the third (75%) and first (25%) quartile points of the residuals).

It should be noted that the robust M-estimator cannot eliminate random noise in the magnetic field time series. Therefore, in practice, M-estimator is adapted to use one or more remote references to avoid the bias. Such that Eq. 4.23 becomes

$$\mathbf{Z} = \mu_0(\mathbf{B}_r^* \mathbf{W} \mathbf{B})^{-1}(\mathbf{B}_r^* \mathbf{W} \mathbf{E}). \quad (4.34)$$

Furthermore, the robust M-estimator can effectively reduce the outliers influence of electric field in impedance estimation, but is typically not sensitive to leverage points in magnetic field data which have undue influence on the outcome. The estimates obtained from any conventional robust M-estimator can be severely biased when leverage points are present, which leads to the bounded influence estimator discussed in next section.

4.6 Robust estimation II: Bounded influence estimation

The basic measures of the robustness of an estimator are its breakdown point and the influence function [Hampel, 1974, Hampel et al., 1986]. The breakdown point, roughly speaking, is the smallest percentage of gross errors that may cause an estimator to take arbitrary large values. For the standard least square solution the breakdown point is zero, which means

that even one single point might have a strong influence on the final estimate. The influence function measures the effect of adding one observation to a very large sample on the estimate. It may be bounded or unbounded when the estimator is not or is sensitive to a bad observation. Hence, a robust estimator should have a bounded influence function. However, the M-estimator described in previous section also possesses a very low breakdown point if the leverage points are present. A single leverage point might completely dominate the estimate, since it might just about be different enough and has strong influence to skew the estimate.

In regression theory, the hat matrix is a widely used quantity to detect high-leverage data [Hoaglin and Welsch, 1978]. In standard regression problem given by Eq. 4.7, the impedance tensor \mathbf{Z} is estimated by $\mu_0(\mathbf{B}^*\mathbf{B})^{-1}(\mathbf{B}^*\mathbf{E})$ (see Eq. 4.23). Let $\hat{\mathbf{E}}$ denotes the electric field vector predicted by \mathbf{Z} , one obtains:

$$\hat{\mathbf{E}} = \frac{1}{\mu_0}\mathbf{BZ} = \mathbf{B}(\mathbf{B}^*\mathbf{B})^{-1}(\mathbf{B}^*\mathbf{E}) = \mathbf{HE}, \quad (4.35)$$

where

$$\mathbf{H} = \mathbf{B}(\mathbf{B}^*\mathbf{B})^{-1}\mathbf{B}^*. \quad (4.36)$$

\mathbf{H} is a $n \times n$ matrix, and is often called "hat matrix". It is a projection matrix that maps \mathbf{E} onto $\hat{\mathbf{E}}$, therefore, \mathbf{H} is a Hermitian and idempotent ($\mathbf{H}^2 = \mathbf{H}$) matrix. It is easy to show that \mathbf{H} has eigenvalues of either 0 or 1, and its diagonal elements satisfy $0 \leq h_{ii} \leq 1$ since

$$\mathbf{H} = \mathbf{H}^2 \quad \implies \quad h_{ii} = \sum_{j=1}^n h_{ij}^2 = h_{ii}^2 + \sum_{j \neq i} h_{ij}^2. \quad (4.37)$$

From Eq. 4.35, the i -th residual then can be written as

$$r_k^{(i)} = E_k^{(i)} - \hat{E}_k^{(i)} = (1 - h_{ii})E_k^{(i)} - \sum_{j \neq i} h_{ij}E_k^{(j)}, \quad k = x \text{ or } y. \quad (4.38)$$

Two extreme cases $h_{ii} = 0$ and $h_{ii} = 1$, can be interpreted as follows. From Eq. 4.37, it is easy to see that whenever $h_{ii} = 0$ or $h_{ii} = 1$, one has $h_{ij} = 0$ for all $j \neq i$. If $h_{ii} = 0$, then the corresponding entry $\hat{E}_k^{(i)} = 0$, hence $\hat{E}_k^{(i)}$ is not affected by $E_k^{(i)}$ or by any other $E_k^{(j)}$. If $h_{ii} = 1$, then $\hat{E}_k^{(i)} = E_k^{(i)}$, which means that the predicted value fits the observed value exactly. In this case only the i -th row of \mathbf{B} has an influence on the regression problem, which indicates that the i -th row of \mathbf{B} will excessively influence the estimate and, therefore, it is a high leverage point analogue to an outlier in \mathbf{E} . Thus, the hat matrix diagonal elements h_{ii} are a measure of the amount of leverage in the magnetic field data.

To determine whether a point is leverage point or not, a threshold value related to h_{ii} can be given. Theoretically, the value is not well defined, but empirically a threshold value larger than $4/n$ or $6/n$ are suitable [Hoaglin and Welsch, 1978].

The hat matrix and its properties for standard least square estimation discussed above can be easily generalized to the robust estimation in Eqs. 4.27 - 4.30 if one defines:

$$\mathbf{H} = \sqrt{\mathbf{w}} \mathbf{B}(\mathbf{B}^* \mathbf{w} \mathbf{B})^{-1} \mathbf{B}^* \sqrt{\mathbf{w}}. \quad (4.39)$$

When the leverage points are determined, it is possible to down weight them in regression procedure to reduce their effect on the estimator. The M-estimator in Eqs. 4.28 - 4.30 then can be extended to a robust estimator which can detect both outliers based on the regression residuals and leverage points based on the hat matrix diagonal elements [Mallows, 1975]. Analogous to Eqs. 4.28 and 4.30, the robust estimator is given by solving:

$$\mathbf{B}^* \Psi \mathbf{v} = 0, \quad (4.40)$$

where $\mathbf{v} = \text{diag}(v_{11} \ v_{22} \ \cdots \ v_{nn})$ is leverage weight matrix, and Ψ is given as in Eq. 4.28. This yields:

$$\mathbf{Z} = \mu_0 (\mathbf{B}^* \mathbf{w} \mathbf{v} \mathbf{B})^{-1} (\mathbf{B}^* \mathbf{w} \mathbf{v} \mathbf{E}). \quad (4.41)$$

The derived estimator has a higher breakdown point and a bounded influence function, therefore, is often termed *bounded influence estimator* (or generalized M-estimator).

The weights v_{ii} in Eq. 4.40 are chosen to minimize the influence of the data with high leverage on the estimate and can be computed iteratively. Chave and Thomson [2003] suggested that the leverage weights are initialized with unity on the main diagonal, and at the k -th iteration the i -th diagonal element is given by

$$v_{ii}^{(k)} = v_{ii}^{(k-1)} \exp(e^{-\chi^2}) \exp(-e^{\chi(u_i - \chi)}), \quad (4.42)$$

where $u_i = \text{trace}(\mathbf{w}^{(k-1)} \mathbf{v}^{(k-1)}) \cdot h_{ii}^{(k)} / 2$ and

$$h_{ii}^{(k)} = \sqrt{w_{ii}^{(k-1)} v_{ii}^{(k-1)}} B_i (\mathbf{B}^* \mathbf{w}^{(k-1)} \mathbf{v}^{(k-1)} \mathbf{B})^{-1} B_i^* \sqrt{w_{ii}^{(k-1)} v_{ii}^{(k-1)}}.$$

χ is a free parameter to determine at which leverage point down-weighting begins, and can be empirically chosen between 2 and 4.

In practice one or more remote references can be incorporated in bounded influence estimator to eliminate the bias caused by noise on magnetic time series.

4.7 Error estimate

Since the impedance estimate is actually a statistical problem, the estimation of the uncertainty, including covariance and confidence interval, is as important as impedance estimation. Uncertainty estimation can be obtained through two approaches: one is a parametric approach which depends on explicit statistical models based on a Gaussian distribution for the data, and another is a more computationally expensive non-parametric approach in which the statistical assumptions can be relaxed.

4.7.1 Parametric error estimate

Traditionally, in standard least square estimator given in Eqs. 4.7 and 4.23, a basic assumption is that the residuals \mathbf{r} are n -variate, complex, normally distributed with zero mean and common variance σ^2 , that is $\mathbf{r} \sim \mathbf{C}_n(\mathbf{0}, \sigma^2)$ (implies that \mathbf{r} are identically distributed). The least square estimator \mathbf{Z} given in Eq. 4.23 is then also the maximum likelihood estimate. The regression quantities are distributed as:

$$\begin{aligned}\mathbf{Z} &\sim \mathbf{C}_2[\mathbf{Z}_{\text{true}}, \sigma^2(\mathbf{B}^*\mathbf{B})^{-1}], \\ \mathbf{E} &\sim \mathbf{C}_n[\mathbf{B}\mathbf{Z}_{\text{true}}, \sigma^2\mathbf{B}(\mathbf{B}^*\mathbf{B})^{-1}\mathbf{B}^*], \\ \frac{(n-2)\sigma^2}{\sigma^2} &\sim \chi_{2(n-2)}^2\end{aligned}\tag{4.43}$$

where $\sigma^2 = \frac{(\mathbf{E} - \mathbf{B}\mathbf{Z})^*(\mathbf{E} - \mathbf{B}\mathbf{Z})}{n-2}$ and $\chi_{2(n-2)}^2$ is the chi-square distribution with $2(n-2)$ degrees of freedom.

Eq. 4.43 can be used to estimate the covariance and confidence interval of \mathbf{Z} . In robust M-estimators and bounded influence estimators, Chave and Thomson [2004] verified that the individual response function elements are also approximately Gaussian distributed.

In practical MT applications, the assumption on \mathbf{r} is rarely valid. For example, the finite duration of many geomagnetic or cultural events result in data anomalies that violate independent residual requirements. Furthermore due to local non-stationarity, the residual distribution is typically very long-tailed around a Gaussian center since extreme residuals are much more common with MT data than is expected for a Gaussian model. Therefore, the information about the data distribution is not clear, which leads to an unknown level of accuracy for parametric estimates. In addition, even if the parameters are correctly estimated, computing the correct number of degrees of freedom in the presence of correlated spectral estimates also make the estimate procedure more complicated. To get an acceptable result, often numerous approximations are needed, however the approximations are sometimes even more sensitive to outliers or other data anomalies than the least square estimates themselves.

As a consequence, instead of assuming a particular form for the distribution and estimating its parameters, non-parametric confidence interval estimators, in which only fewer assumptions on distribution are required, are increasingly used. Among the non-parametric methods the simplest method is the *jackknife* method, which is a linear approximation to the more general *bootstrap* estimator.

4.7.2 Non-parametric error estimate

Non-parametric error estimation relies only on observations and does not require prior knowledge on the population density function.

Bootstrap method

Originally, bootstrap method was introduced to compute standard error of an arbitrary estimator by Efron [1979]. Its basic idea is to use the data samples themselves to compute a

statistic and to estimate the sampling distribution without any prior model assumption.

Assuming that n samples $x = (x_1, x_2, \dots, x_n)$ are available from an unknown population density function, one wants to assess the accuracy (mean, standard error, etc.) of an arbitrary estimator $\hat{\theta} = s(x)$. Choosing one sample randomly from x and setting it as x_1^* , yields, when repeated n times $x^{*(1)} = (x_1^*, x_2^*, \dots, x_n^*)$. $x^{*(1)}$ is an independent identically distributed (i.i.d.) sample from x , since each x_i^* is chosen independently with equal probability from n samples. Note that some of x_i^* may be identical. Repeating this procedure B times, yielding B samples $x^{*(1)}, x^{*(2)}, \dots, x^{*(B)}$, the *bootstrap samples*.

The statistic $\hat{\theta}$, which one wants to estimate, can be computed with each bootstrap sample, is called a bootstrap replication:

$$\hat{\theta}^{*(b)} = s(x^{*(b)}), \quad \forall b \in \{1, 2, \dots, B\}. \quad (4.44)$$

$\hat{\theta}^{*(1)}, \hat{\theta}^{*(2)}, \dots, \hat{\theta}^{*(B)}$, constitutes a random sample from the distribution of $\hat{\theta}$, which contains all the information related to $\hat{\theta}$. Hence it can be used to approximate every quantity of interest of $\hat{\theta}$ (e.g. mean, standard deviation, percentiles and etc). Moreover, a histogram constitutes an estimate of the unknown density of the $\hat{\theta}$, such that other quantities like skewness may be derived as well.

The bootstrap estimate of the standard error (SE) of $\hat{\theta}$ is then given by

$$SE_B(\hat{\theta}) = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{\theta}^{*(b)} - \hat{\theta}^*)^2}, \quad (4.45)$$

where

$$\hat{\theta}^* = \frac{1}{B} \sum_{b=1}^B \hat{\theta}^{*(b)}. \quad (4.46)$$

Similarly an estimate of the bias of $\hat{\theta}$ is obtained as

$$Bias(\hat{\theta}) = \hat{\theta}^* - \hat{\theta}. \quad (4.47)$$

The bootstrap method has the advantage that the quantity of interest can be computed based on all samples without consideration of error propagation, since the errors are automatically propagated to the final result. It can furthermore handle data with skewed distributions because the samples are just as skewed as the original data. Its disadvantage is that it is computationally intensive and furthermore requires reasonably large amount of samples of the original data.

Jackknife method

The jackknife is, somewhat similar to bootstrap, another resampling method discovered a few decades earlier. Tukey [1958] firstly use the name of "jackknife" to include variance

estimation. It is a "leave-one-out" type of cross validation.

Assuming again n samples $x = (x_1, x_2, \dots, x_n)$ available from an unknown population density function, one wants to assess the accuracy of an arbitrary estimator $\hat{\theta} = s(x) = s(x_1, x_2, \dots, x_n)$. The data is subdivided into n groups of size $n - 1$ by deleting one sample. Estimations obtained without the i -th observation are called the i -th jackknife replication, and are denoted by $\hat{\theta}_{-i} = s(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$.

The difference between the whole sample estimate and the jackknife replication is called pseudo-value $\hat{p}_i^* = n\hat{\theta} - (n - 1)\hat{\theta}_{-i}$, $i = 1, 2, \dots, n$. These pseudo-values reduce the linear bias of the jackknife replication because the bias is eliminated by the subtraction between the two estimates. The jackknife estimate of $\hat{\theta}$ is obtained as the mean of the pseudo-values

$$\hat{\theta}^* = \frac{1}{n} \sum_{i=1}^n \hat{p}_i^*. \quad (4.48)$$

The variance of the pseudo-values is obtained with the usual formula:

$$\hat{\sigma}_{\hat{p}_i^*}^2 = \frac{1}{n-1} \sum_{i=1}^n (\hat{p}_i^* - \hat{\theta}^*)^2. \quad (4.49)$$

The standard error of the parameter estimates can be obtained from the variance of the pseudo-values from the usual formula for the standard error of the mean as:

$$\hat{\sigma}_{\hat{\theta}^*}^2 = \frac{\hat{\sigma}_{\hat{p}_i^*}^2}{n} = \frac{\sum_{i=1}^n (\hat{p}_i^* - \hat{\theta}^*)^2}{n(n-1)}. \quad (4.50)$$

The jackknife variance always slightly exceeds the true variance even if the data is not identically distributed. This standard error can then be used to compute confidence intervals for the estimation of the parameter. Under the independence assumption, this estimation is distributed as a Student's t distribution with $n - 1$ degrees of freedom. Specifically, a $1 - \alpha$ confidence interval can be computed as

$$\hat{\theta}^* \pm t_{\alpha, \nu} \hat{\sigma}_{\hat{\theta}^*}, \quad (4.51)$$

with $t_{\alpha, \nu}$ being the α -level critical value of a Student's t distribution with $\nu = n - 1$ degrees of freedom.

The bias of the jackknife is given by

$$\text{Bias}_j = \hat{\theta} - \hat{\theta}^*. \quad (4.52)$$

A major advantage of the jackknife is its computational simplicity. Compared to a parametric approach in which the distribution function needs to be known, all information needed for jackknife is an estimator for the statistic of interest. However, the jackknife method can fail if the statistic parameter of interest is not smooth.

Error estimate in MT

Thomson and Chave [1991] gave a detailed description of implementation and performance of the jackknife method in spectral analysis, and Chave and Thomson [1989] described how to use it to estimate confidence intervals on the MT response tensor. The delete-one estimates of Z can be derived for either the standard least square solution, the robust estimator, or robust remote reference estimators by deleting a row from E and B in turn. The jackknife is directly used to yield an estimate for the variance of response estimates of the real and imaginary parts of the impedance. The variance is used to derive the standard error on the apparent resistivity and phase by using

$$\Delta\rho_{ij} = 2\mu_0|z_{ij}|\Delta z_{ij}/\omega, \quad (4.53)$$

$$\Delta\phi_{ij} = \sin^{-1}(\Delta z_{ij}/|z_{ij}|), \quad (4.54)$$

where $i, j = x$ or y and Δz_{ij} is the jackknife estimate of the standard error. Confidence intervals of these quantities are calculated in the standard way. Compared to parametric estimates for real data, Chave and Lezaeta [2007] concluded that the estimated results using jackknife are accurate enough and simple to use and to implement.

4.8 Bounded influence remote reference processing

The bounded influence remote reference processing (BIRRP) code has been introduced by Chave and Thomson [2004], and is now frequently used as a MT data processing code. BIRRP program uses a bounded influence, remote reference method, along with an implementation of jackknife error estimates on the impedance estimations. It also incorporates a method for controlling leverage points and removal of outlier in both local magnetic and electric field time series.

In BIRRP, after the data are read in, they are pre-whitened using an autoregressive filter and the Fourier transforms of the sections are computed after tapering with a prolate data window as described in section 4.2. The algorithm can be summarized as follows:

1. Given n estimates of the Fourier transforms of the horizontal electric and magnetic field time series, calculate the least square estimate $\mathbf{Z} = \mu_0(\mathbf{B}_r^*\mathbf{B})^{-1}(\mathbf{B}_r^*\mathbf{E})$.
2. Compute the residual $\mathbf{r} = \mu_0\mathbf{E} - \mathbf{BZ}$, the residual sum of squares $\mathbf{r}^*\mathbf{r}$, the scale parameter $d = s_{MAD}/\sigma_{MAD}$ and the diagonal elements of hat matrix $\mathbf{H} = \mathbf{B}_r(\mathbf{B}_r^*\mathbf{B}_r)^{-1}\mathbf{B}_r^*$.
3. Set χ in Eq. 4.42 just below the largest normalized diagonal element of \mathbf{H} .
4. Compute the Huber weights and hat matrix weights, and compute the bounded influence remote reference estimate $\mathbf{Z} = \mu_0(\mathbf{B}_r^*\mathbf{wvB})^{-1}(\mathbf{B}_r^*\mathbf{wvE})$.
5. Compute the residual $\mathbf{r} = \mu_0\mathbf{E} - \mathbf{BZ}$, the scale d , and the weighted residual sum of squares $\mathbf{r}^*\mathbf{wvr}$.
6. Repeat steps 4 and 5 until the residual sum of squares converge.
7. Decrement χ , repeat steps 3,4 and 5 until χ reaches the given minimum value.
8. Compute the covariance matrix and confidence intervals.

4.9 Estimator in time-frequency domain

The estimators discussed above are all given in the frequency domain, because of the simplicity of the fast Fourier transform. In fact, the impedance tensor can also be estimated in time domain [Kunetz, 1972, Spagnolini, 1994] according to the convolution theorem or in time-frequency domain [Berdichevsky et al., 1973] using transient complex spectra analysis. In the following, the latter method will be introduced in more detail.

The basic idea of the time-frequency domain method is that a real time series $x(t)$ can be written as a complex time series since its spectrum is conjugate symmetric, that is

$$\hat{x}(\omega) = (\hat{x}(-\omega))^*, \quad (4.55)$$

where $*$ denotes the complex conjugate. Therefore $x(t)$ can be expressed as

$$\begin{aligned} x(t) &= \int_{-\infty}^{\infty} \hat{x}(\omega) e^{2\pi\omega t} d\omega \\ &= \int_0^{\infty} \hat{x}(\omega) e^{2\pi\omega t} d\omega + \int_{-\infty}^0 \hat{x}(\omega) e^{2\pi\omega t} d\omega \\ &= \int_0^{\infty} \hat{x}(\omega) e^{2\pi\omega t} d\omega + \left(\int_0^{\infty} \hat{x}(\omega) e^{2\pi\omega t} d\omega \right)^* \\ &= \operatorname{Re} \left\{ \int_0^{\infty} 2\hat{x}(\omega) e^{2\pi\omega t} d\omega \right\} = \operatorname{Re}(x_c(t)), \end{aligned} \quad (4.56)$$

where

$$x_c(t) = \int_0^{\infty} 2\hat{x}(\omega) e^{2\pi\omega t} d\omega, \quad (4.57)$$

$x_c(t)$ is the *complex signal* or *analytic signal* of $x(t)$, with the Fourier spectrum of

$$\hat{x}_c(\omega) = \begin{cases} 2\hat{x}(\omega) & \omega \leq 0 \\ 0 & \omega < 0 \end{cases}. \quad (4.58)$$

The imaginary part of $x_c(t)$ is exact the Hilbert transform of $x(t)$: $\mathcal{H}\{x(t)\} = \frac{1}{\pi} \wp \int_{-\infty}^{\infty} \frac{x(t')}{t-t'} dt'$, and the real and imaginary parts of $x_c(t)$ have a $\pi/2$ phase shift.

In the method suggested by Berdichevsky et al. [1973], the magnetotelluric time series are first filtered using a narrow band filter around the desired estimation frequency $[\omega_0 - \Delta\omega, \omega_0 + \Delta\omega]$. Then, for each filtered time series, a complex time series $x_f(t)$, which is called a complex transient time series, is constructed as shown in Fig. 4.1.

Assuming that $E_{xf}(t), E_{yf}(t), B_{xf}(t), B_{yf}(t)$ are complex transient time series corresponding to the frequency ω_0 (the subscript f indicates the filtered time series), the following equations can be derived:

$$\begin{pmatrix} \sum_{n=0}^N E_{xf}(t_n) B_{yf}^*(t_n) \\ \sum_{n=0}^N E_{yf}(t_n) B_{xf}^*(t_n) \end{pmatrix} = \frac{1}{\mu_0} \begin{pmatrix} Z_{xx}(\omega_0) & Z_{xy}(\omega_0) \\ Z_{yx}(\omega_0) & Z_{yy}(\omega_0) \end{pmatrix} \begin{pmatrix} \sum_{n=0}^N B_{xf}(t_n) B_{yf}^*(t_n) \\ \sum_{n=0}^N B_{yf}(t_n) B_{xf}^*(t_n) \end{pmatrix}, \quad (4.59)$$

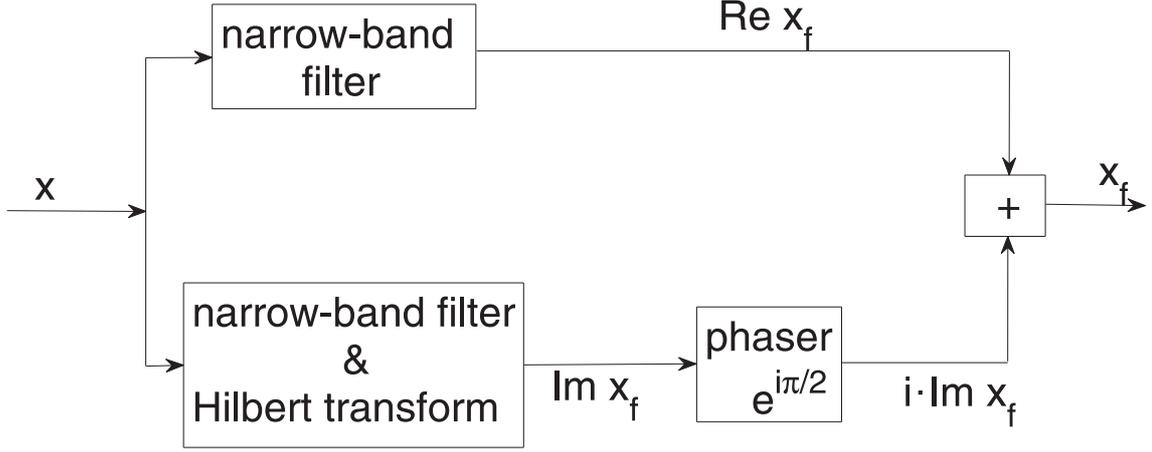


Figure 4.1: The procedure to construct a complex transient signal $x_f(t)$ around a desired estimation frequency ω_0 from a real signal $x(t)$. The real and imaginary parts of $x_f(t)$ are the narrow band filtered signal of $x(t)$ and its Hilbert transform, respectively.

and

$$\begin{pmatrix} \sum_{n=0}^N E_{x_f}(t_n) B_{x_f}^*(t_n) \\ \sum_{n=0}^N E_{y_f}(t_n) B_{x_f}^*(t_n) \end{pmatrix} = \frac{1}{\mu_0} \begin{pmatrix} Z_{xx}(\omega_0) & Z_{xy}(\omega_0) \\ Z_{yx}(\omega_0) & Z_{yy}(\omega_0) \end{pmatrix} \begin{pmatrix} \sum_{n=0}^N B_{x_f}(t_n) B_{x_f}^*(t_n) \\ \sum_{n=0}^N B_{y_f}(t_n) B_{x_f}^*(t_n) \end{pmatrix}, \quad (4.60)$$

where N is the number of sampling points in time domain.

From Eqs. 4.59 and 4.60, the impedance component Z_{xy} can be derived as

$$Z_{xy} = \frac{\sum_{n=0}^N E_{x_f}(t_n) B_{y_f}^*(t_n) \sum_{n=0}^N B_{x_f}(t_n) B_{x_f}^*(t_n) - \sum_{n=0}^N E_{x_f}(t_n) B_{x_f}^*(t_n) \sum_{n=0}^N B_{x_f}(t_n) B_{y_f}^*(t_n)}{\sum_{n=0}^N B_{x_f}(t_n) B_{x_f}^*(t_n) \sum_{n=0}^N B_{y_f}(t_n) B_{y_f}^*(t_n) - \sum_{n=0}^N B_{x_f}(t_n) B_{y_f}^*(t_n) \sum_{n=0}^N B_{y_f}(t_n) B_{x_f}^*(t_n)} \mu_0. \quad (4.61)$$

Other components Z_{xx}, Z_{yx}, Z_{yy} can be calculated similarly. It is easy to show that the standard least square solution in Eq. 4.23 is equivalent to Eq. 4.61.

From Eq. 4.61, the impedances can be calculated in time-frequency domain through complex transient time series. As will be shown in next chapter, the instantaneous information of a time series, which can also be represented in time-frequency domain, can be derived from a new method - empirical mode decomposition. Consequently, Eq. 4.61 can be used to calculate the impedance through instantaneous information in chapter 6 and 7.

4.10 Short summary

In this chapter, traditional impedance estimators in MT data processing and the estimation of the uncertainties are discussed. Most of data processing techniques are based on Fourier analysis due to the fact that impedance is defined in frequency domain, and furthermore use approximately a least-squares method. For basic least square estimator which was originally with assumption of a Gaussian statistical model, the estimated impedance is always biased when noise is present in data. To address this bias, a remote reference estimator may be used under the assumption that the noise at the site and at the reference stations are uncorrelated. Furthermore, while least-squares based methods are sensitive to small amounts of anomalous data in E- or B-field, therefore, in practice more sophisticated robust M-estimators and bounded influence estimators are needed. Such sophisticated estimators have greatly improved the reliability of impedance calculations and become nowadays standard in MT data processing.

However, many geomagnetic time series recorded for MT investigations are characterized by a changing frequency content with time mainly due to contamination by transient local natural or artificial noise signals, and can be considered as non-stationary within typical recording lengths [Chant and Hastie, 1992, 1993]. Fourier-based sophisticated MT processing techniques in some instances can handle non-stationarities through strategies such as tapering of the time series, but sometimes can also provide strongly biased and unusable impedance estimates. Until now it is not completely clear how extreme non-stationarity may affect the resulting impedances.

Historically, to process non-stationary MT time series, Chant and Hastie [1992, 1993] provided stable and reliable impedance estimates through an evolutive estimate using time-frequency distribution (TFD) analysis techniques. Wavelet analysis, as a different approach to handle non-stationary data, has also been discussed by e.g. Alexandrescu et al. [1995], Suto et al. [2006], Garcia and Jones [2008]. As a possible alternative, I will examine a heuristic method called empirical mode decomposition, which is developed to handle arbitrary non-stationary time series, for MT data processing. The EMD method will be introduced in the next chapter.

Chapter 5

EMD Method and Instantaneous Spectrum

In this chapter, a recently developed data-driving non-stationary time series processing method - empirical mode decomposition (EMD), is introduced. By using EMD, a time series can be adaptively decomposed into several components called intrinsic mode functions (IMF). These IMFs have the characteristic that physically meaningful instantaneous frequencies can be derived for all time samples. Knowledge of time-varying instantaneous frequencies is important for many aspects in data analysis and particularly relevant to understand the mechanisms involved in nonlinear and non-stationary processes.

The motivation of why this new method is applied to process MT data is presented in the first section. The EMD method and the determination of the associated IMFs are introduced in section 2. Concept and calculation of physically meaningful instantaneous frequencies are summarized in section 3. A number of details of the EMD method and their implementations are summarized in the remaining sections. In this context some critical aspects and limitations of EMD method are furthermore discussed.

5.1 Motivation

As mentioned in chapter 4, the noise issues cause a bias problem in impedance estimates in traditional Fourier-based MT data procedures. This is particularly true when MT time series are contaminated with local non-stationary noise.

The local non-stationarities are often caused by signal components which violate the source assumptions made in MT analysis, such as high power transients caused by magnetic storms. On land MT measurements can in addition be contaminated by artificial noise due to ground currents from electric fences and railways, mining equipment and electric trains, as reviewed by Szarka [1988] and Junge [1996]. Since these events are usually present in both the electric and magnetic fields and cause high coherence in the associated time series, robust processing procedures both with and without remote reference will produce biased response function estimates. To avoid such a bias, these sections are typically removed before processing in conventional MT procedures.

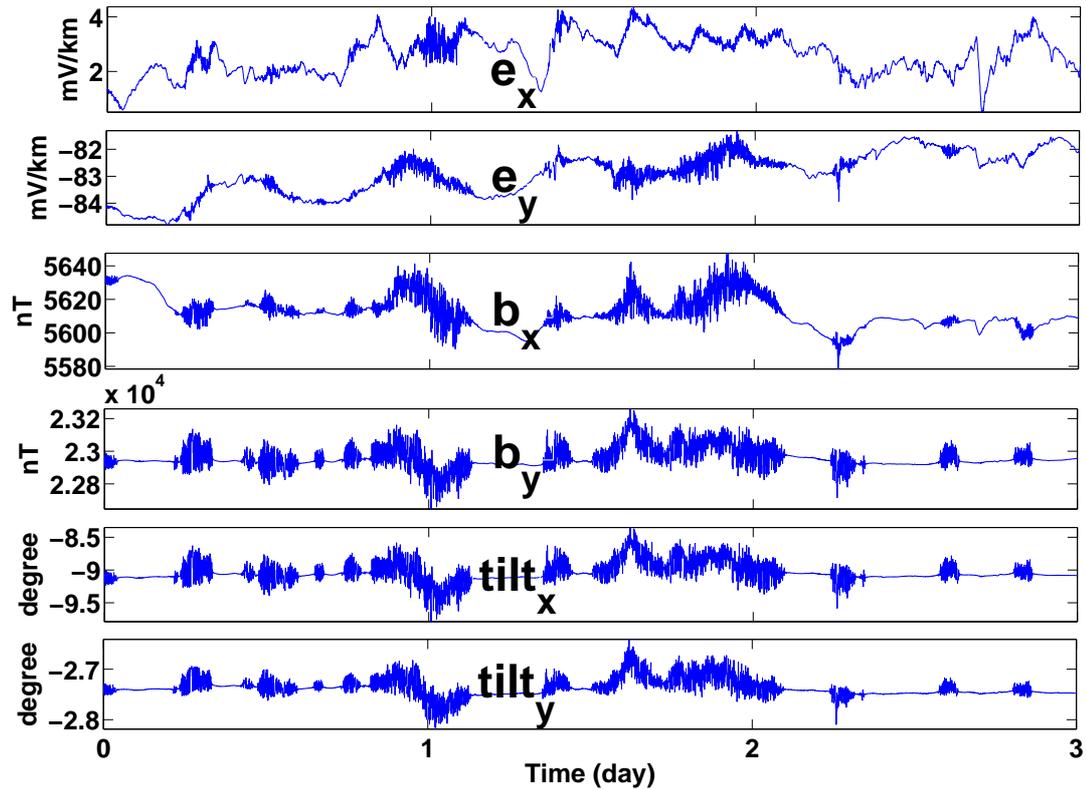


Figure 5.1: Example of three day long period of noisy marine MT measurements collected at Costa Rican subduction zone in 2007. All E - and B -field components and tilt movements are shown. It is obvious that both E - and B -field data are contaminated by noise associated with the tilt movements.

In marine environment such non-stationary effects are even more pronounced. Measured MT time series are often contaminated with non-stationary noise from sources like water waves, passing ships and tilt movements of the instruments. As an example, Fig. 5.1 shows marine MT data from the Costa Rican subduction zone (see also Chapter 7), where tilt movements generate noise in the time series of the E - and B -field components. It can be easily observed that time-periods with high-frequency tilt movements (see panels 5-6) strongly correlates with sections of non-stationary noise in all E - and B -field components (panels 1-4). If these sections are removed before processing as done in the traditional procedures, then only a small portion of data remains for impedance calculation and accordingly no information about lower frequencies remains.

The reason, why Fourier-based MT data processing methods have difficulties with such noise contaminations arise from how non-stationarities affect the power spectral density and the error estimate. If a non-stationary event is present in a time section many additional harmonic components are needed to describe this event. This is particularly true if the non-stationary

event has an impulsive character, which results in a spreading of the power density over a wide frequency range. Constrained by the energy conservation, these spurious harmonics and the wide frequency spectrum cannot faithfully represent the true power density in the frequency space [Huang et al., 1998]. Furthermore, for least-square based impedance estimators discussed in chapter 4, errors in each segments are assumed to be uncorrelated and have a Gaussian distribution with common variance for each frequency. The Eq. 4.43 shows that also the impedance estimate itself is then Gaussian distributed. However, when non-stationarities exist, the estimated impedances are not Gaussian distributed any more because more extreme residuals exist than would be expected. This impedance distribution is then typically very long-tailed with a peak value that can strongly differ from the mean value. Accordingly, impedance estimate and its error estimates can be seriously biased.

Robust processing techniques introduced in chapter 4 try to handle non-stationarity by adapting the segment lengths in time to the frequencies of interest. Such techniques have the advantage that their bias and statistical properties are well-developed. Chant and Hastie [1992, 1993] used time-frequency distribution (TFD) (see [e.g. Cohen, 1989]) analysis techniques to process non-stationary MT signals, and obtained stable and reliable impedance estimates through an evolutive estimate. However estimates from TFD techniques are still considered problematic since dependent on the complexity of MT signals, different time-lag kernel functions are required. Another popular method used to handle non-stationary data is the wavelet analysis [e.g. Alexandrescu et al., 1995, Suto et al., 2006, Garcia and Jones, 2008]. However, since wavelet analysis gives time-frequency information over time-variable windows, a wavelet basis has to be selected before processing and it can be difficult to evaluate whether the chosen wavelet basis is suited for all signals in the time series or not.

The limitations of these traditional methods motivate one to search for other methods for MT data processing. To find such a method, some questions are arisen at the very beginning that need to be answered:

- Are there some methods beyond Fourier analysis that can handle the non-stationarity?
- If the answer is yes. Can one use them to detect and visualize non-stationary noise effects (e.g. the tilt movements) effects in MT field time series?
- Finally, can data sections contaminated by non-stationary noise still be utilized in impedance tensor estimation rather than simply removing them before processing?

A method for which all three questions can be answered with "yes" surely need different properties as Fourier analysis which globally defines uniform harmonic components with constant amplitude during a time interval. The new methods should have local and adaptive properties to represent a non-stationary time series. The requirement of locality is obvious, because non-stationary events do not repeat with time but are located at specific times. Accordingly the time of their occurrences have to be identified and also the time-varying amplitude & frequency should be determined. The adaptivity is required because the locations of non-stationarities are not known in beforehand. Therefore, the method has to be data-driven (see for more details in Huang et al. [1998]). One cannot expect to find a predetermined basis that explains all of non-stationary time series.

In this context, *Instantaneous Frequency* (IF) seems to be a very important parameter. The concept of IF was developed for applications in signal processing [e.g. Carson and Fry, 1937, Gabor, 1946], and is considered as especially helpful for non-stationary signals, whose spectral characteristics are varying with time. One problem with IFs is that they have a real physical meaning only for a very specific type of signals that are called *mono-component* [Boashash, 1992a,b, Loughlin and Tracer, 1996, Huang et al., 2009]. Unfortunately, there is no rigorous definition for a mono-component signal and it will be illuminated in more detail what different people consider as "mono-component" in section 5.4. However, independent of the actual considered definition, geomagnetic time series, like most other time series, are typically not mono-component such that IFs calculated directly from these time series are generally not physically meaningful, as will be shown in section 5.3.

With the development of the *Empirical Mode Decomposition* (EMD) introduced by Huang et al. [1998], a new processing method exists which decomposes non-stationary time series into a sequence of *intrinsic mode functions* (IMFs). The big advantage of this composition is that for each IMF physically meaningful instantaneous frequencies can be calculated for all times. This makes it possible to handle non-stationarities for arbitrary time series - including of course also geomagnetic time series.

5.2 Empirical mode decomposition

The basic concept of the empirical mode decomposition (EMD) is to identify empirically the intrinsic oscillatory modes by their characteristic time variations in the data, and to decompose the data into a set of components that are called intrinsic mode functions (IMFs). As will be discussed later, the IMFs have the characteristic that physically meaningful IFs can be calculated at every time sample and EMD allows, accordingly, localization of any event on the time as well as the frequency axis. Very important in this context is that the IMFs represent the local variations of the data and, therefore, are adaptive. EMD works directly in the time domain rather than in frequency domain and results are typically presented as time-frequency-energy representations of the IMFs. Due to its definition, EMD is an empirical method based on a relatively simple numerical algorithm and is therefore quite different to the mathematically and analytically well-defined Fourier transform. Over the last ten years, more than hundred papers about EMD were published dealing with different theoretical and technical aspects and applications. Applications in various scientific fields and in engineering are collected in Huang and Shen [2005] and Huang and Attoh-Okine [2005]. A review of applications in geosciences are summarized in Huang and Wu [2008].

Referring to Huang et al. IMFs have to satisfy two conditions [Huang et al., 1998, Rilling et al., 2003, Flandrin, 2004], in order to achieve a decomposition of the time series from which physically meaningful IFs can be derived:

- (i). In each IMF, the number of extrema and the number of zero crossings must either equal or differ at most by one; and
- (ii). at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

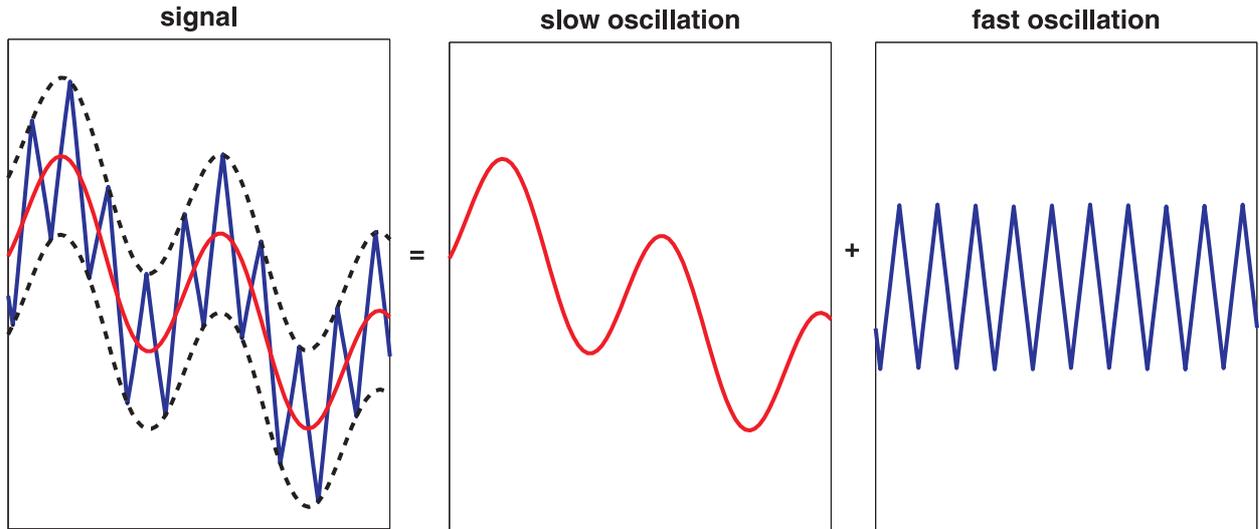


Figure 5.2: Illustration of the "sifting" process. One signal can always be decomposed into a fast oscillate signal plus a sum of slow oscillate signals. In this example only two oscillations are present.

These requirements for the intrinsic mode functions have to be reached by adaption because they represent oscillation modes that are imbedded in the time series. Each IMF represents a simple oscillatory mode, whose frequency and amplitude contents vary with time as will be shown later, has a totally other meaning as a simple harmonic function that is associated with a single Fourier frequency. The decomposition of an arbitrary signal into IMFs that satisfy the above mentioned definitions, is an iterative process called sifting and has been proposed by Huang et al. [1998]. The basic idea of "sifting" is that any signal $x(t)$ can always be decomposed into a faster (high frequency) oscillation and a sum of the remaining slower (low frequency) oscillations. For example, the signal, that is shown as a blue curve in the left panel of Fig. 5.2, can be decomposed into a slower oscillation, which is the local mean of the signal (red curve in middle panel), and a faster oscillation (blue curve in the right panel).

The whole decomposition proceeds as follows [Huang et al., 1998, Rilling et al., 2003, Flandr in, 2004]:

1. Identify all local extrema of $x(t)$.
2. Interpolate all local maxima and local minima by cubic splines to obtain upper and lower envelopes, respectively. Most parts of the time series are usually located between the upper and lower envelopes.
3. The mean of upper and low envelope value designated as m_1 (i.e. red curve in Fig. 5.2) is subtracted from the signal $x(t)$, i.e.

$$x(t) - m_1 = h_1. \quad (5.1)$$

If the difference h_1 (i.e. the blue curve in right panel of Fig. 5.2) fulfills conditions (i) and (ii) for IMFs, h_1 is the first IMF component of $x(t)$.

4. If h_1 does not fulfill both conditions, h_1 is treated as the original signal and steps 1 to 3 are repeated,

$$h_1 - m_{11} = h_{11}. \quad (5.2)$$

The procedure of step 4 (also called "sifting") is repeated as long as the two conditions are fulfilled. At the final iteration k the component h_{1k} that is determined from

$$h_{1(k-1)} - m_{1k} = h_{1k}, \quad (5.3)$$

is assigned to

$$c_1 = h_{1k}. \quad (5.4)$$

c_1 is now considered as the first IMF component of the original time series. It typically contains the finest scale or the shortest period component of the signal.

5. To obtain now the next higher IMF, c_1 is separated from $x(t)$ by

$$r_1 = x(t) - c_1. \quad (5.5)$$

The remaining signal r_1 is treated now as the original time series and the above mentioned processes (steps 1 - 4) are repeated until the second IMF component c_2 of $x(t)$ has been derived.

6. The above process is repeated n times until n IMFs of the signal $x(t)$ have been determined,

$$\begin{aligned} r_1 - c_2 &= r_2, \\ &\vdots \\ r_{n-1} - c_n &= r_n. \end{aligned} \quad (5.6)$$

The decomposition process is stopped when r_n becomes a monotonic function from which no more IMFs can be extracted. Finally one obtains

$$x(t) = \sum_{j=1}^n c_j + r_n. \quad (5.7)$$

The signal has now been decomposed into n empirical modes and a residue r_n , which is the mean low-frequency trend of $x(t)$. The sum of all IMFs and the residue completely recovers the signal. Furthermore, the IMFs c_1, c_2, \dots, c_n include different frequency bands, where highest frequencies are usually contained in the first IMF and lower frequencies in subsequent IMFs. Because frequency content in each IMF typically changes with time, it is, however, common that different IMFs occupy the same frequency band at different sections in time. The number n of IMFs needed to represent the signal depends on the complexity of the signal. For white noise Wu and Huang [2004] showed that EMD works as a dyadic filter resulting in a number of IMFs close to $\log_2(N)$, where N is the number of samples.

I will now demonstrate the capability of EMD to decompose a multi-component time series using two examples.

Example 1

Consider a synthetic harmonic time series $x(t)$, which is the sum of a 15 Hz and a 1 Hz sine wave and a 5 Hz sine wave modulated by a 0.1 Hz sine wave,

$$x(t) = 2\sin(2\pi \times 15t) + \sin(2\pi \times 5t)\sin(2\pi \times 0.1t) + 4\sin(2\pi \times t). \quad (5.8)$$

The time series and its three components are plotted in Fig. 5.3. In this example the time series is discretized with a sampling frequency of 100 Hz.

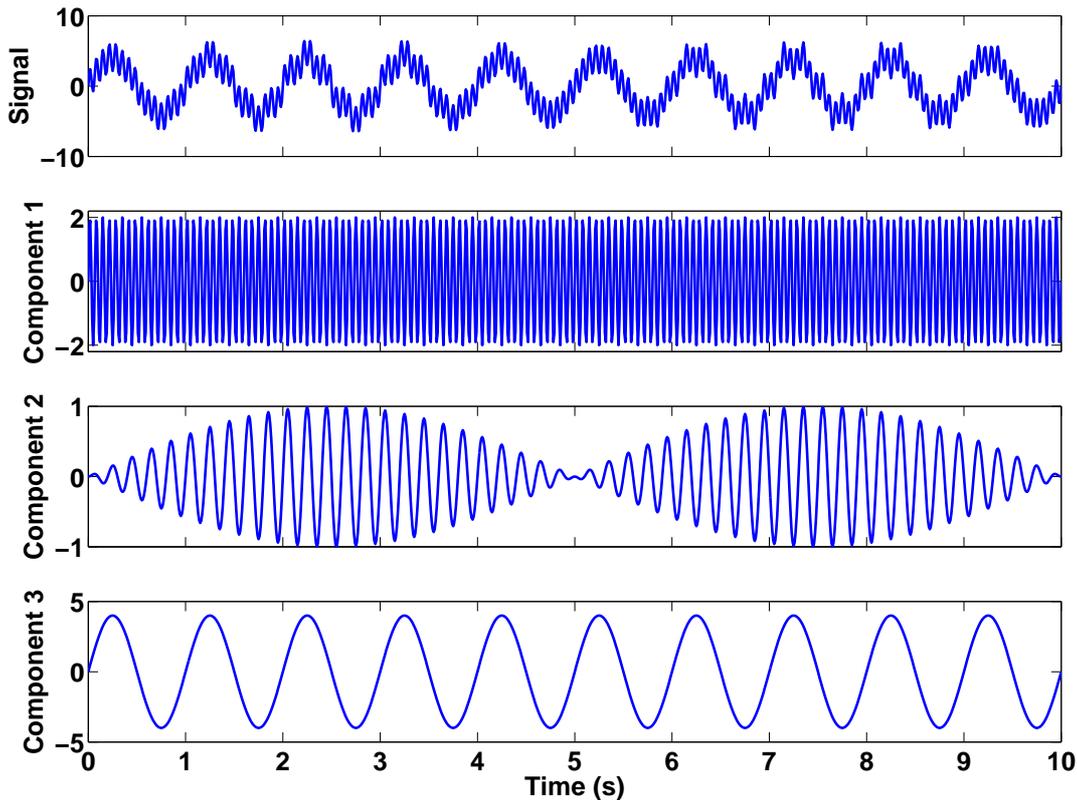


Figure 5.3: Synthetic harmonic time series of example 1 (first panel) and its three constituent components (panel 2 to 4).

Fig. 5.4 shows the first sifting iteration of the EMD method applied to this time series. The original signal and its local extrema, the upper & lower envelopes and the mean of envelopes are shown in the first panel (step 1 to 2). The signal shown in the second panel is the new signal obtained after the first sifting iteration that is used then again as the input time series for the next sifting iteration (step 3). The residue which is equal to the local mean of the signal (see red curve in first panel) is shown in the third panel. The sifting process needs to be repeated several times to obtain the first IMF; in this example, two iterations are needed. The final decomposition of the time series consists of six IMF's which are shown in Fig. 5.5 as blue curves. For comparison, the three original components of the signal are added as red curves to the same panels. It is obvious that the first IMF mainly recovers the 15 Hz sine

wave component, the second IMF the 5 Hz amplitude-modulated signal and the third IMF the 1 Hz sine wave component of the time series. The remaining three IMFs have relatively small amplitudes and are induced by numerical inaccuracies like boundary problems which will be discussed later. Since in this example the frequencies of components 1 - 3 are separated from higher to lower in the whole time interval, they are successfully extracted from the time series by using EMD.

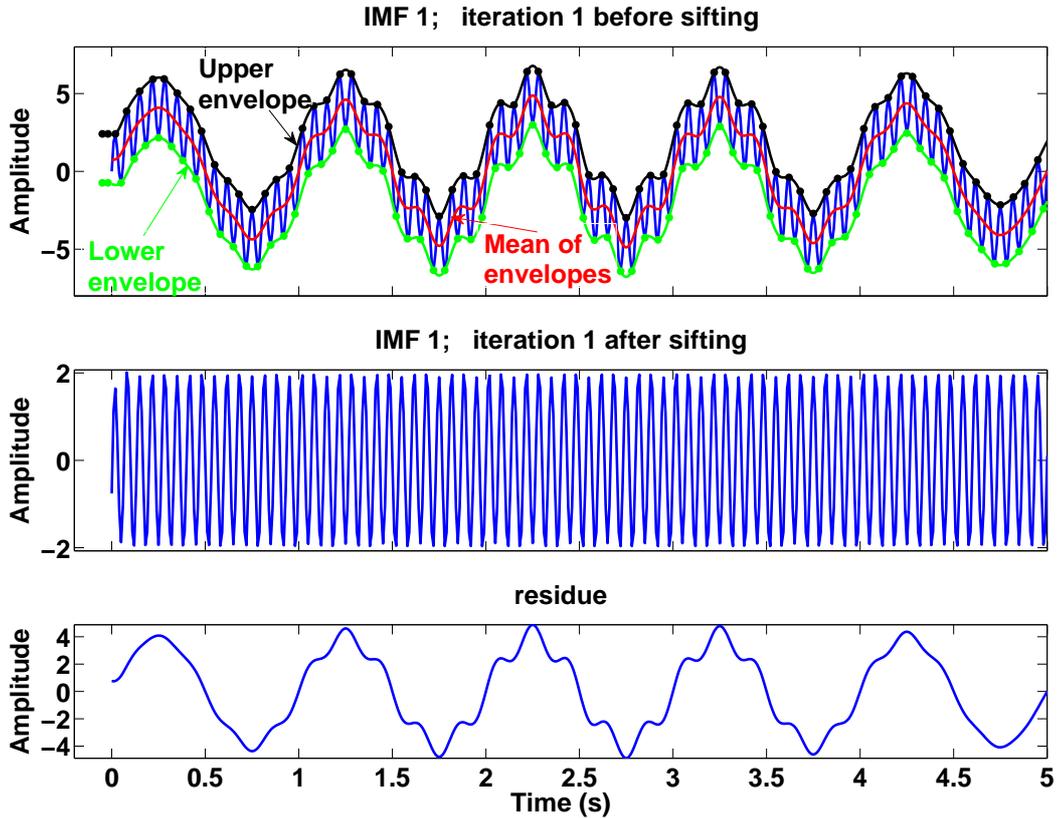


Figure 5.4: This figure illustrates how the first sifting iteration of the EMD method is performed on the times series shown in Fig. 5.3. (Here only half of the total time interval is shown). In the first panel the original signal and its local maxima and minima (black and green dots) are plotted. From the extrema the upper and lower envelope are determined (black and green line) that are used to calculate the local mean values (red line) required for the first sifting process. The second panel shows the signal after the first iteration (namely h_1 in Eq. 5.1) which will be used as the input signal for next sifting iteration. The third panel shows the residue (namely m_1 in Eq. 5.1) that corresponds to the local mean value in the first panel.

Example 2

The second example is based on a times series that is presented in Rilling et al. [2003]. It contains two sinusoidal frequency-modulated (FM) waves with maximum amplitudes of 1 and 2, respectively, and a Gaussian wavepacket. The three components are plotted in Fig. 5.6. The composed signal and time dependent frequency & amplitude information of the three components are shown in Fig. 5.7. The time series is discretized with a sampling

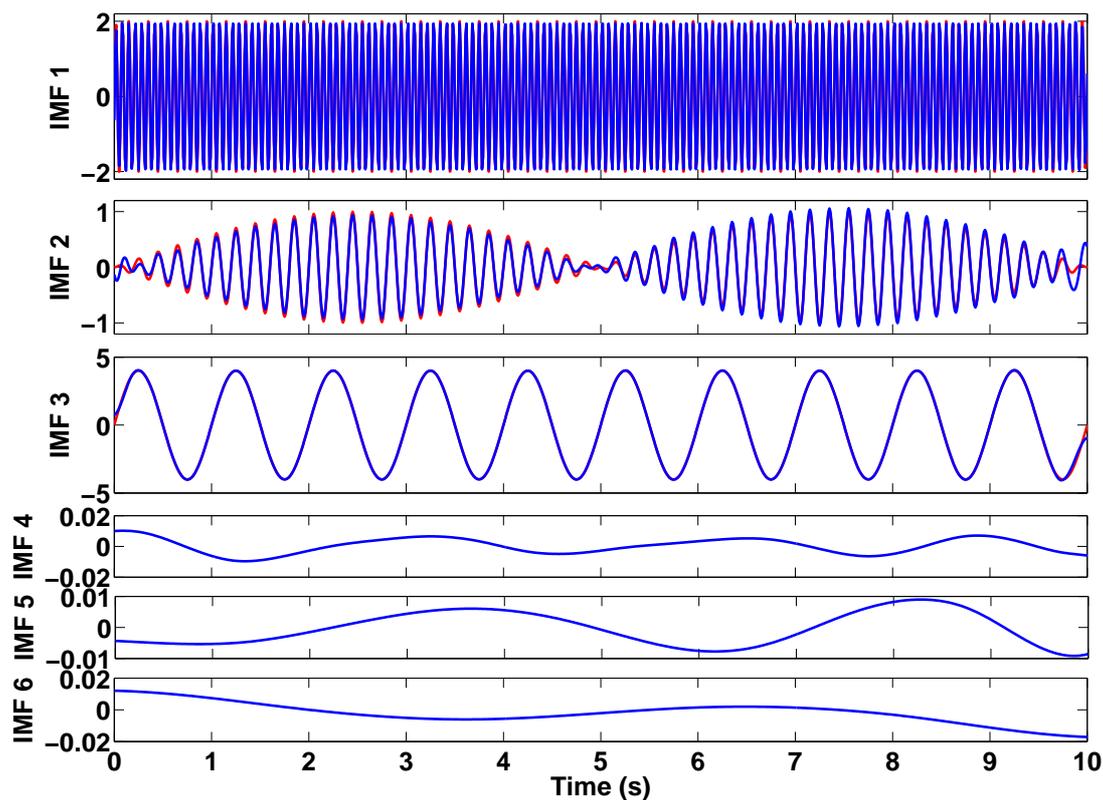


Figure 5.5: The six IMFs (blue curves) obtained for the signal of example 1 (see first panel in Fig. 5.3). The first three IMFs in principal represent the three constituent components of the original signal (red curves, also see panel 2 - 4 in Fig. 5.3), and the remaining three IMFs, which have relatively lower amplitudes, are related to some numerical inaccuracies like e.g. boundary problems.

frequency of 10 Hz in this example. It is obvious that the statistical properties of this time series change over time and, therefore, time series is non-stationary.

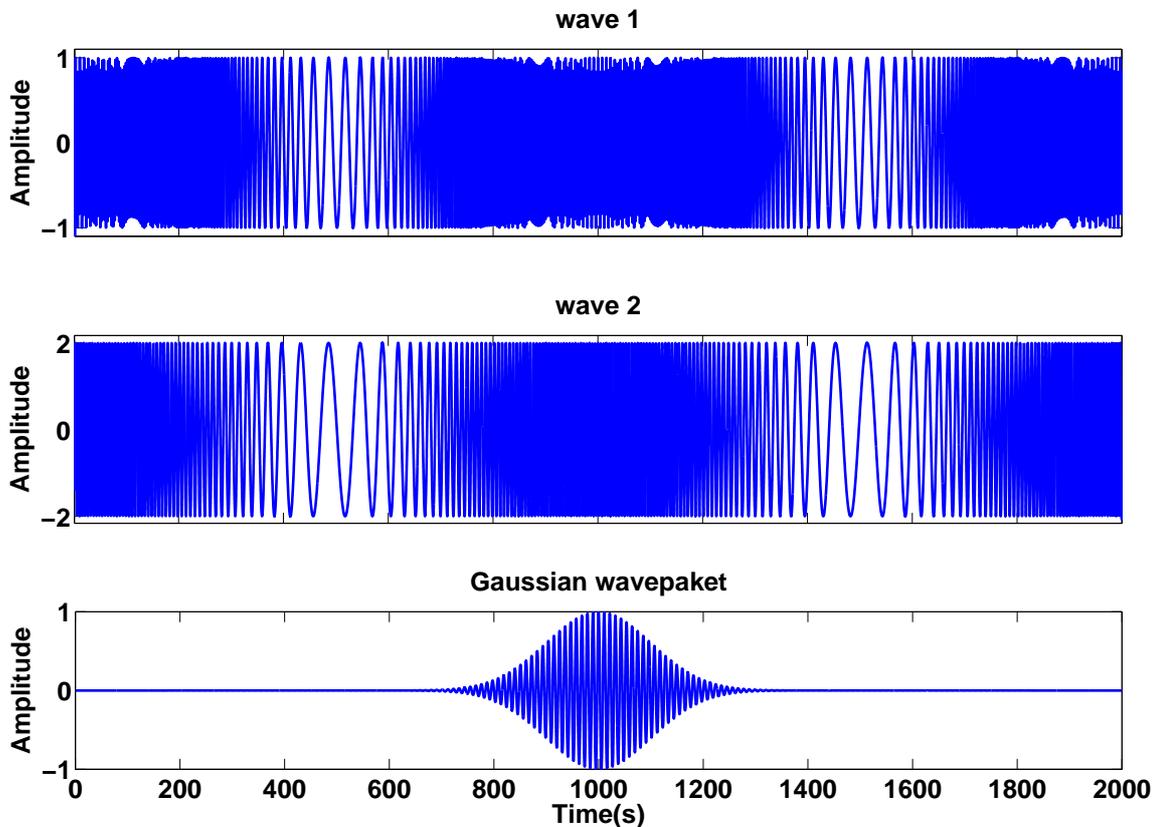


Figure 5.6: Two sinusoidal frequency-modulated (FM) waves (panel 1 - 2) and a Gaussian wavepacket (panel 3). These signals are combined to construct the synthetic time series of example 2 (see upper panel of Fig. 5.7).

Because the Fourier spectra of the three time-series are overlapping in frequency domain, it is not possible to separate the signal characteristics of the individual components by Fourier transform. Fig. 5.8 shows the power spectra for the three components of the time series. The spectra are robustly estimated by using a discrete prolate spheroidal sequence with a time-bandwidth of 1 as tapering window and overlapping sections [Chave et al., 1987].

Using EMD to decompose the time series, eleven IMFs are obtained which contain components of the time series from higher to lower frequencies. The first six IMFs (IMF 1 - IMF 6) are plotted in Fig. 5.9 as blue curves. As for example 1, the original three components are added as red curves to the panels showing the IMF 1 - IMF 3. It can be seen that the first three IMFs mainly recover three original components signals; IMF 1 - 2 the two FM components and IMF 3 the Gaussian wavepacket. Again, the remaining eight IMFs with relatively small amplitudes are induced by the numerical procedure of the EMD method.

These two synthetic examples show that the EMD has the capability to decompose an arbi-

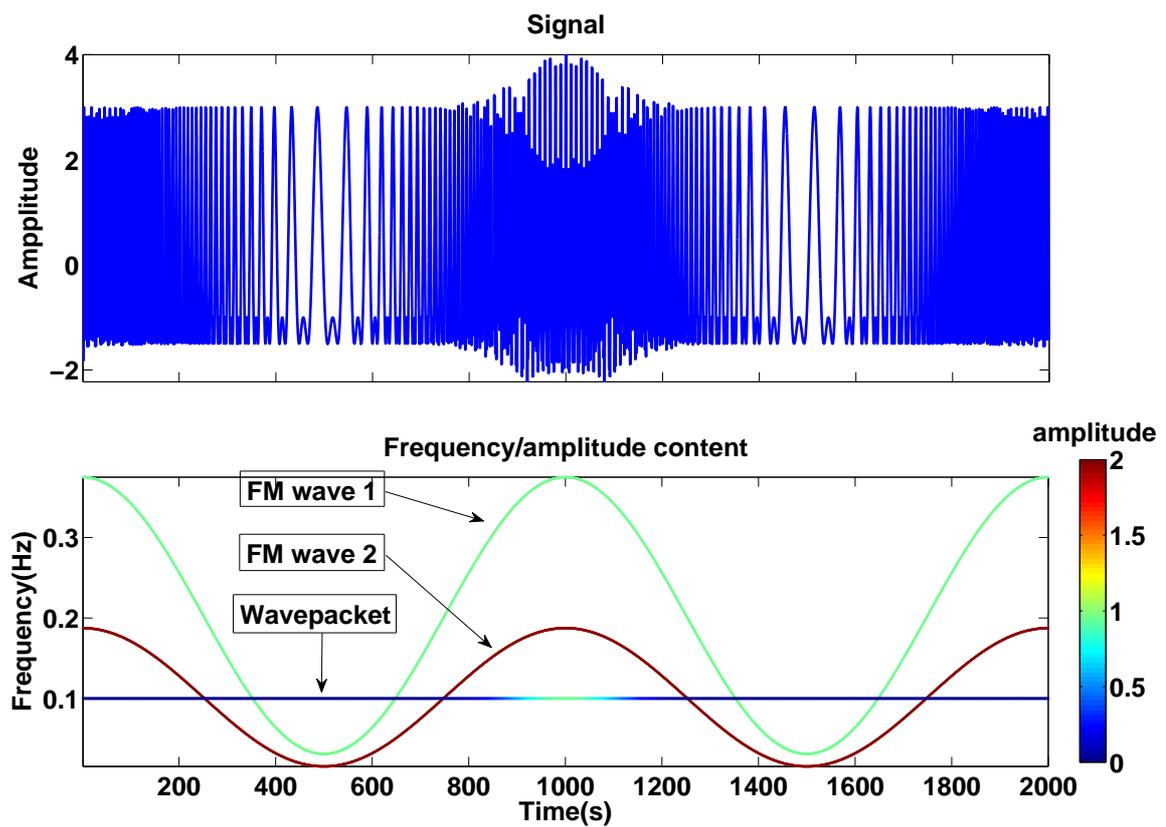


Figure 5.7: Synthetic time series of the example 2 (upper panel) and the frequency & amplitude contents of its three components (lower panel) that are plotted in Fig. 5.6. The color coding is associated with the (maximum) amplitudes.

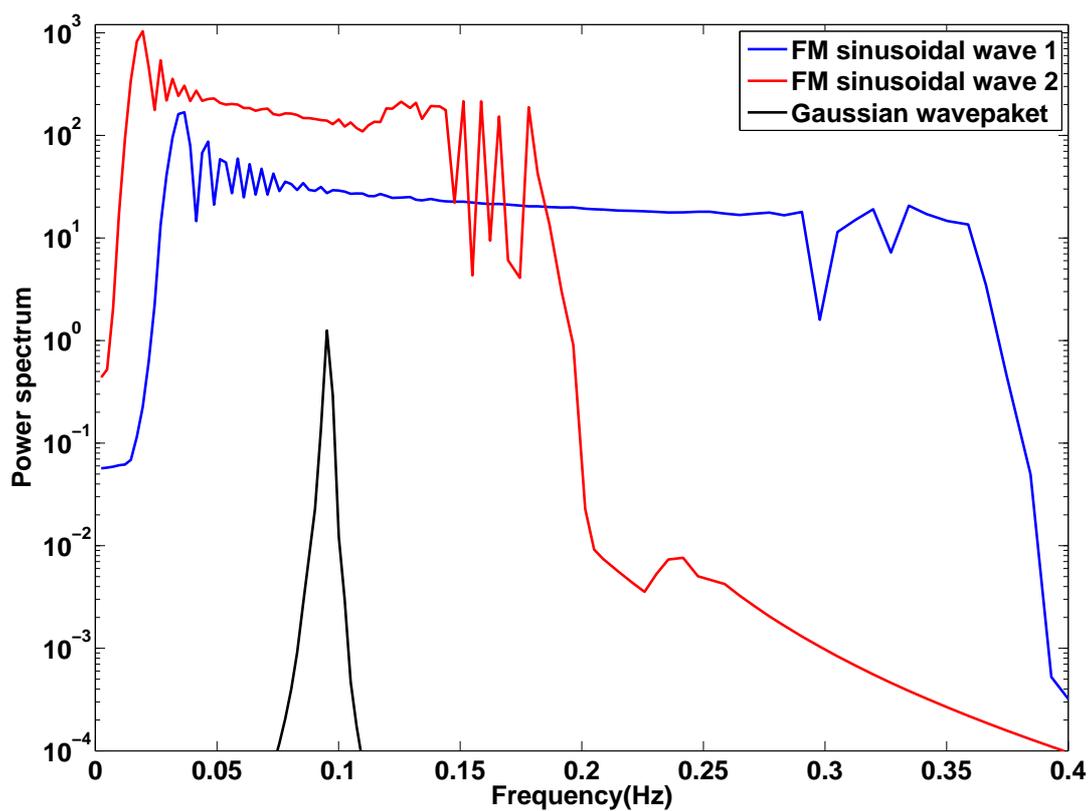


Figure 5.8: Fourier power spectra of the three components of the synthetic time series shown in Fig. 5.6. (To obtain these spectra a discrete prolate spheroidal sequence with a time-bandwidth of 1 as tapering window is applied on overlapping sections.)

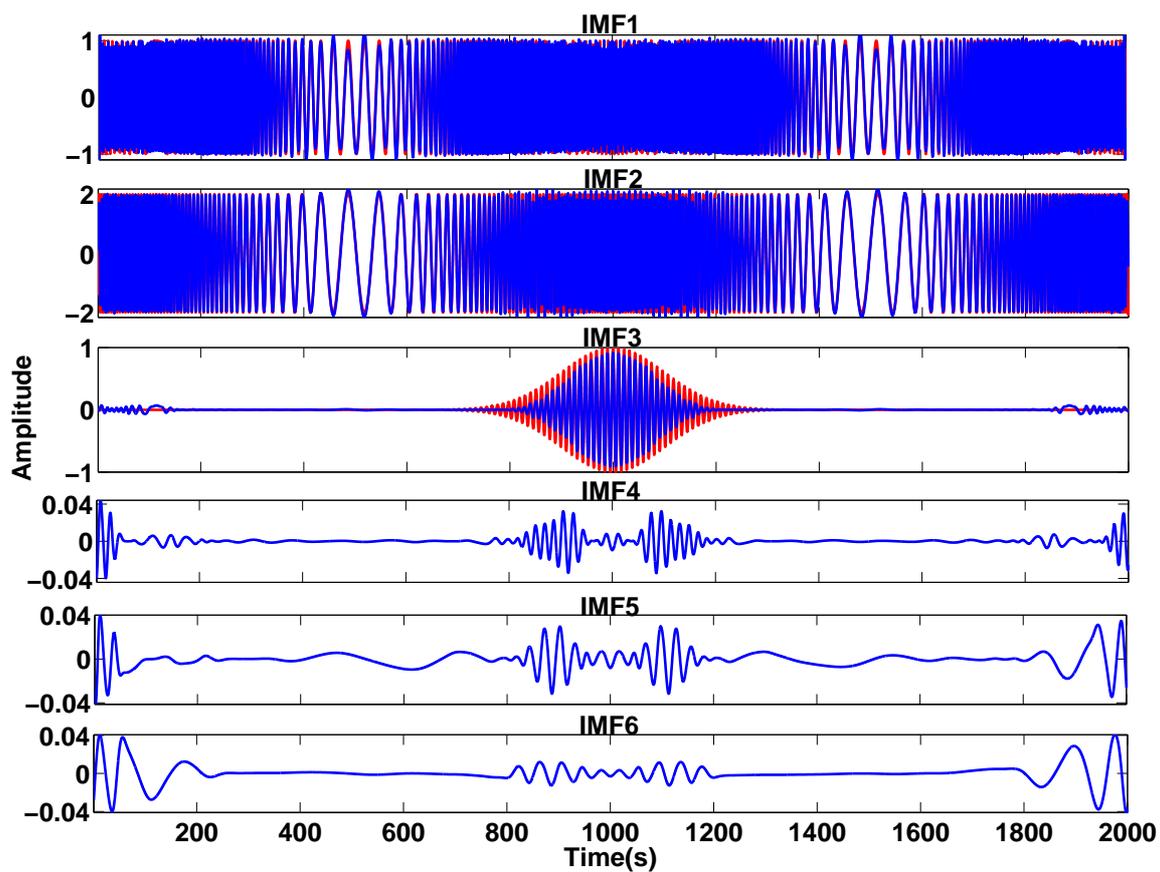


Figure 5.9: The first six IMFs (blue curves) of the synthetic time series of example 2. For comparison, the three original components are also plotted (red curves) in the panels with IMF 1 - IMF 3.

rary (both stationary and non-stationary) time series into its original components by using its local and adaptive properties. While a time series is expanded into a combination of harmonic sinusoidal waves in Fourier analysis, it is expanded into fewer non-harmonic intrinsic mode functions by using EMD. Although the decomposition procedures are very different for Fourier analysis and EMD, the EMD method is considered by Huang et al. [1998] as a generalization of the Fourier transform in which IMFs are regarded as basis functions.

Completeness and orthogonality of IMFs

In harmonic/non-harmonic signal analysis, it is very important that the basis, which is used to expand the time series, is complete and orthogonal. Due to the fact that the sum of all IMFs and the residue completely recovers the time series (see Eq. 5.7), the completeness of the IMFs is ensured. However, the orthogonality of IMFs is theoretically not guaranteed. Huang et al. [1998] concluded that the orthogonality is typically satisfied in a practical sense though, since locally each IMF contains lower frequency oscillations than all the others that are extracted before. Another important aspect in EMD method is that for different time series the obtained IMFs can also be totally different. This means that the basis functions (IMFs) changes with the signal characteristics and it is generally not possible to reconstruct any arbitrary time series by a linear combination of the fixed IMFs like it is the case for the harmonics in the Fourier analysis. In this context, Huang et al. [1998] proposed two criteria to quantify the orthogonality numerically:

1. The overall index of orthogonality, IO , defined as

$$IO = \sum_t \left(\sum_{j=1, j \neq k}^{n+1} \sum_{k=1}^{n+1} |c_j(t)c_k(t)|/x^2(t) \right) \quad (5.9)$$

and

2. orthogonality for any two components

$$IO_{jk} = \sum_t \frac{|c_j(t)c_k(t)|}{c_j^2(t) + c_k^2(t)}. \quad (5.10)$$

The first criteria can vary from 0.0 to 1.0 and zero values indicate that the IMFs span a perfect orthogonal system. The second criteria varies from 0.0 to 0.5 and zero values indicate that the associated two IMFs are orthogonal.

Now for the two synthetic examples described in the previous section, I check that to which extent their IMFs describe an orthogonal system. For the first example, the overall index of orthogonality IO is 0.0010 and the orthogonality matrix of the six IMFs is given in table 5.1. For the second example, the overall index of orthogonality is 0.0190 and the orthogonality matrix for the first 4 IMFs is shown in table 5.2. The low values of the IO and all IO_{ij} for $i \neq j$ for both examples indicate that the IMFs represent, to a first approximation, an orthogonal system.

Although a number of publications [e.g. Chen et al., 2006, Xu and Zhang, 2009, Qian et al., 2010] deal with mathematical aspects of the orthogonality of IMFs, a complete answer to

IO_{ij}	c_1	c_2	c_3	c_4	c_5	c_6
c_1	0.5					
c_2	0.0154	0.5				
c_3	0.0044	0.0073	0.5			
c_4	0.0038	0.0003	0.0018	0.5		
c_5	0.0004	0.0024	0.0010	0.0110	0.5	
c_6	0.0010	0.0017	0.0010	0.0149	0.2436	0.5

Table 5.1: Components of the orthogonality matrix for the six IMFs of synthetic example 1.

IO_{ij}	c_1	c_2	c_3	c_4
c_1	0.5			
c_2	0.0156	0.5		
c_3	0.0027	0.0039	0.5	
c_4	0.0000	0.0000	0.0005	0.5

Table 5.2: Components of the orthogonality matrix for the first four IMFs of synthetic example 2.

this issue is still not available. Some other authors also proposed modifications of the EMD method to improve the orthogonality of IMFs. Junsheng et al. [2006] suggested an energy difference tracking method in the sifting procedure that forces IMFs to meet the orthogonality condition and thereby reduce the index of orthogonal (IO) to even smaller values. However, this method suffers from increased computation time due to the fact that many more sifting iterations are needed. Huang et al. [2008] presented a modification of the EMD method called OEMD (orthogonal empirical mode decomposition) that is based on the Gram-Schmidt orthogonalization method which guarantees complete orthogonality of all IMFs. Nevertheless, the approximation of the orthogonality obtained from the original procedure from Huang et al. [1998] can be considered as adequate for most practical purposes.

5.3 Instantaneous frequency and spectrum

In a second step, I will now present that the obtained IMFs can be used as a basis to calculate the instantaneous frequencies and amplitudes.

5.3.1 Concept of instantaneous frequency

The instantaneous frequency (IF) was originally defined in the context of frequency modulation (FM) in communication. In contrast to the instantaneous energy or the instantaneous envelope, whose definitions are clear and well accepted, the concept of the IF is, somehow, controversial. The main reasons for this arise from the frequency definition and the deep influence of the Fourier spectral analysis which has dominated signal processing in the past decades and accordingly biased the understanding of what one considers as a frequency nowadays. Traditionally, the frequency is defined as the number of oscillations per unit time

in a periodic motion. Thus, one needs at least one full oscillation cycle to define the local frequency value. Using such a definition there is apparently a paradox to assign a specific "frequency" to a specific "instantaneous" time. Furthermore, calculation of frequencies in the classic Fourier analysis is based on sine or cosine functions that have constant amplitudes for the whole data length. If a signal can be fully represented by a superposition of such functions, the frequencies are well defined. However, in practice, signals are rarely truly sinusoidal or superpositions of sinusoidal components. This is especially true for non-stationary signals for which the frequency content changes with time. In this case, such a frequency definition and calculation would not make sense. Instead a parameter is required that can characterize the time-varying nature of the process. This has led to the concept of *Instantaneous Frequency*.

Many authors have contributed to the study of IF [e.g. Carson and Fry, 1937, Van der Pol, 1946, Gabor, 1946, Ville, 1948, Cohen, 1966, Rihaczek, 1968, Ackroyd, 1970, Griffiths, 1975, Gupta, 1975]. In the following, some important ideas about how to define the IF are introduced.

Carson and Fry [1937] considered a complex purely frequency modulated wave in the form of, $x(t) = \exp(i\phi(t))$, where the trend of $\phi(t)$ is mainly linear and accordingly is dominated by a constant frequency. They defined the IF as the rate of change of phase angle $\phi(t)$ with respect to time t , which is a generalization for the definition of constant frequencies. Actually, when higher order terms in a Taylor expansion of phase function $\phi(t)$ around time t are neglected, the first order derivative of $\phi(t)$ takes over the role of the constant frequency, that is:

$$\begin{aligned} \phi(t) \approx \phi(t_0) + \phi'(t_0)(t - t_0) &\implies & x(t) \approx \exp(i\phi(t_0) + i\phi'(t_0)(t - t_0)) \\ & & = e^{i(\phi(t_0) - \phi'(t_0)t_0)} [\cos(\phi'(t_0)t) + i \sin(\phi'(t_0)t)]. \end{aligned} \quad (5.11)$$

In contrast to the complex representation of the signal used by Carson and Fry [1937], Van der Pol [1946] considered a simple real amplitude modulated and frequency modulated (AM-FM) signal of the form, $x(t) = a(t) \cos(\phi(t))$, where $a(t)$ changes slowly with time and the trend of $\phi(t)$ is mainly linear. Using these assumptions they obtained a similar definition for IFs: $f_{ins}(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt}$.

The next important step in understanding the concept of IF was made by Gabor [1946], who proposed to define the IF by generating an unique complex analytic signal $u(t)$ from a real signal, $x(t)$:

$$u(t) = x(t) + i\mathcal{H}\{x(t)\} = a(t)e^{i\phi(t)} \quad (5.12)$$

where

$$\begin{aligned} a(t) &= \sqrt{x^2(t) + (\mathcal{H}\{x(t)\})^2} \\ \phi(t) &= \text{atan}\left(\frac{\mathcal{H}\{x(t)\}}{x(t)}\right) \end{aligned} \quad \text{and} \quad \mathcal{H}\{x(t)\} = \frac{1}{\pi} \wp \int_{-\infty}^{\infty} \frac{x(t')}{t - t'} dt' \quad (5.13)$$

He called $a(t)$ the *instantaneous amplitude* and $\phi(t)$ the *instantaneous phase*. $\mathcal{H}\{x(t)\}$ is the Hilbert transform (HT) of $x(t)$ and \wp indicates the Cauchy principle value. The instantaneous frequency was then defined by him as:

$$\omega(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt}. \quad (5.14)$$

Another way to determine the IFs is based on the use of Time-Frequency Distribution (TFD). Ville [1948] unified the work done by Carson and Fry [1937] and Gabor [1946], and defined IFs for signals of the form $x(t) = a(t)\cos\phi(t)$ through a signal representation both in time and frequency, $W(t, f)$. The representation he used is commonly referred as the *Wigner-Ville Distribution* (WVD):

$$W(t, f) = \int_{-\infty}^{+\infty} u(t + \tau/2)u^*(t - \tau/2)e^{-2\pi i\omega\tau} d\tau. \quad (5.15)$$

where $u(t)$ is complex analytic signal as defined in Eq. 5.12 and the definition he obtained for the IF is:

$$f_{ins}(t) = \frac{\int_{-\infty}^{+\infty} fW(t, f)df}{\int_{-\infty}^{+\infty} W(t, f)df} \quad (5.16)$$

Some other energy or intensity distributions, which represent a signal both in time and frequency domain, were also used to define IFs [e.g. Rihaczek, 1968, Page, 1952]. Cohen [1966] and Cohen and Lee [1988] developed a generalized formulation for the distribution of energy in time and frequency and defined the IF as the average of the frequencies that exists in the time-frequency plane at a given time. More details about the time-frequency distributions are presented in Cohen [1989], and more information about the relationship between the IF and TFD is given by Boashash [1992a,b].

The non-uniqueness of the definition makes the usefulness of IF to a subject of controversy. In the following sections, it will focus on Gabor's definition (see Eq. 5.14), which is widely used in areas such as communication theory. Under which conditions physically meaningful IF can be calculated will be discussed in more detail as well.

5.3.2 IF based on Hilbert transform

In Eq. 5.14 Gabor used the Hilbert transform of the signal $x(t)$ as the imaginary part to construct a complex analytic signal. Using the Fourier transform \mathcal{F} , the Hilbert transform has the frequency response:

$$\mathcal{F}\{\mathcal{H}\{x(t)\}\} = \mathcal{H}(\omega) \cdot \mathcal{F}\{x(t)\}(\omega), \quad (5.17)$$

where

$$\mathcal{H}(\omega) = -i \cdot \text{sgn}(\omega) = \begin{cases} -i & \text{for } \omega > 0 \\ 0 & \text{for } \omega = 0 \\ i & \text{for } \omega < 0 \end{cases}. \quad (5.18)$$

This means that the Hilbert transform has the effect of shifting the negative and positive frequency components of $x(t)$ by 90° and -90° , respectively. In other words, for any real-valued signal $x(t)$, the Fourier transform of the corresponding analytic signal $u(t)$ given in Eq. 5.12 has only non-zero coefficients for positive frequencies and all coefficients for negative frequencies are zeros. In theory, the signal pair, $x(t)$ and $\mathcal{H}\{x(t)\}$, are out of phase by $\pi/2$

and they are often said to be in *quadrature*.

To understand the definition given in Eq. 5.14, a simple cosine signal with constant frequency f is considered as an example:

$$x(t) = a \cos(2\pi ft). \quad (5.19)$$

First, the Hilbert transform is calculated, the analytic signal and its phase (see Eqs. 5.12 and 5.13) are then expressed as:

$$\mathcal{H}\{x(t)\} = a \sin(2\pi ft) \implies \begin{array}{l} \text{analytic signal: } u(t) = ae^{i2\pi ft}, \\ \text{phase function: } \phi(t) = 2\pi ft. \end{array} \quad (5.20)$$

Then, Eq. 5.14 is used to determine the IF. The IF is exactly f and, accordingly, it is physically meaningful:

$$f_{ins} = \frac{1}{2\pi} \frac{d\phi(t)}{dt} = f. \quad (5.21)$$

The real and imaginary parts of the analytic signal $u(t)$ have exactly a phase shift of $\pi/2$.

In Eq. 5.14 IF is given as the first order derivative of the phase function with respect to time. If one sets $y(t) = \mathcal{H}\{x(t)\}$ and considers the phase function $\phi(t) = \text{atan}(y(t)/x(t))$, Eq. 5.14 can be expressed in another form:

$$\omega(t) = \frac{1}{2\pi} \frac{x(t)y'(t) - x'(t)y(t)}{x(t)^2 + y(t)^2}, \quad (5.22)$$

in which two differentiations have to be calculated to determine the IF. However, in practise, measured time series generally exist in discrete form and one needs the discrete versions of the formula to calculate the IF. Numerically, there are two implementation problems that need to be solved; one is the calculation of Hilbert transform and the other is the determination of the derivatives. The most convenient way to implement the Hilbert transform is through the fast Fourier transform (FFT). However, because of the finite data length, then the spectral leakage caused by numerous side-lobes must be considered, which in practice is generally handled through strategies such as tapering of the time series. To solve the derivatives numerically, several neighboring samples are generally used to approximate the derivatives at one given sample point. Barnes [1992] summarized that three discrete formulae based on Eq. 5.22 exist to calculate the IFs effectively and quickly. They are a) a two-points formula [Scheuer and Oldenburg, 1988] which introduces a half-sample time shift and generates no instantaneous frequencies greater than the Nyquist frequency, b) a three-points formula [Boashash, 1992b] which has no time shift and also generates no IFs greater than the Nyquist frequency, and c) the two approximations of the derivation formula proposed by Claerbout [1976, p. 20] and Yilmaz [1987, p. 521] which introduces a half-sample time shift and obtained IFs could be greater than the Nyquist frequency.

All three formulae generally produce results that are acceptable close to the true values and for most applications any of the three would suffice. However, since the derivatives at one sample are only approximated by two or three neighbor points, the obtained time-varying IFs are often not smooth and small oscillations and some peaks exist close to the extrema of the signal. Typically the time series of IFs produced by the three-point operator is slightly

smoother because it involves twice as much samples for averaging. The expression for the three-point formula is [Boashash, 1992b]:

$$f_{ins}(t) = \frac{1}{4\pi T} \operatorname{atan} \left[\frac{x(t-T)y(t+T) - x(t+T)y(t-T)}{x(t-T)x(t+T) + y(t-T)y(t+T)} \right], \quad (5.23)$$

where T is the sampling interval. In the remaining part of this section, this formula is chosen to calculate the IFs.

Theoretically, for any signal one can construct a complex analytic signal and calculate its IFs from Eq. 5.23. However, one has not considered yet, if the calculated IFs are then also necessarily physically meaningful. To address this issue a sine wave with constant frequency of 0.00195 Hz and a constant amplitude shift a is considered:

$$x(t) = a + \sin(\pi t/256), \quad \text{with} \quad t = 0, 1, \dots, 1023, \quad (5.24)$$

The time interval is set as $t = 0, 1, \dots, 1023$ such that the first and the last data point are both zero to avoid the spectral leakage. Three different values are used for a ($a = 0$, $a = 0.5$ and $a = 2$) and the IFs are calculated by using Eq. 5.23 (see Fig. 5.10). For $a = 0$, all obtained IFs have the same positive value and are accordingly meaningful as one expected. For $0 < a < 1$ the obtained IFs differ with time, but are still always positive and hence meaningful. However, for $a > 1$ for some time intervals with negative IFs exist that are physical meaningless.

Already this simple example shows clearly that positive meaningful IFs are not obtained for all signals. Therefore, it is important to identify for which type of signals physically meaningful IFs are guaranteed. Boashash [1992a] and Cohen [1995] stated that for a "mono-component" signal, the calculated IF is physically meaningful. Unfortunately, until now there is no clear definition for a "mono-component" signal, which makes it difficult to judge whether a signal is mono-component or not. Intuitively, a mono-component signal should be a zero-mean signal and possibly modulated in amplitude and/or frequency. Which type of signal can be considered as a "mono-component" signal will be discussed in section 5.4 in more detail. Furthermore, independent of the actual definition most of the measured signals are too complex to be considered as simple mono-component signals. Instead they are considered as multi-component and as long as such signals are not decomposed into components with a "mono-component" characteristic, the commonly accepted fact is that the calculated IFs are always meaningless.

This discussion illustrates that it is surely not an easy task to identify and select functions for which positive meaningful IFs are guaranteed. Fortunately, Huang et al. [1998] solved this problem to a large extent. They proposed that meaningful instantaneous frequencies can only be guaranteed for functions that have locally zero mean values and are locally symmetric with respect to the time axis (For the simple sine function example these assumptions are fulfilled for $a = 0$.) Based on these local restrictions, Huang et al. [1998] developed the EMD method to decompose any kind of signal into IMFs which have zero mean values and symmetric properties and hence, allow to calculate physically meaningful IFs for all times. Since they originally calculated the IFs for each IMF by using Gabor's definition that is based on the Hilbert transform, they called their procedure *Hilbert-Huang transform* (HHT).

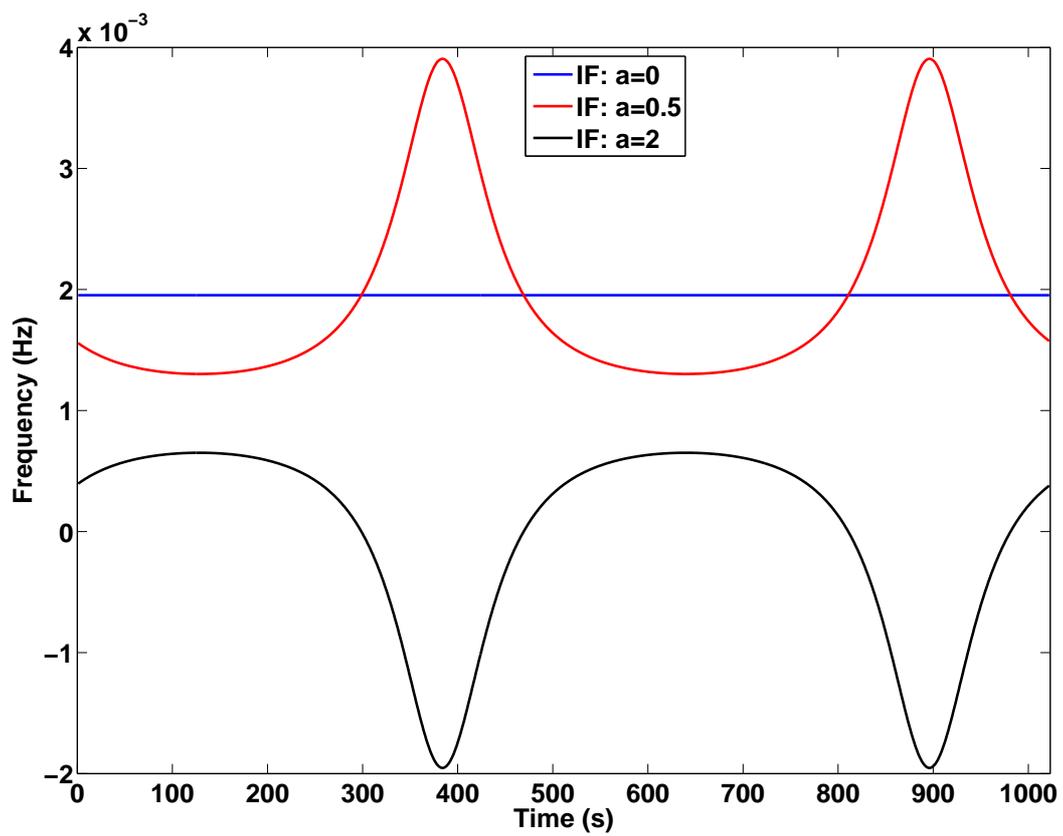


Figure 5.10: IFs calculated by using Eq. 5.14 for three harmonic sinusoidal signals (Eq. 5.24) with different constant amplitude shift.

To investigate the calculation of IFs on real time series, I consider in the following the examples 1 and 2 in section 5.2 which are both 3-component signals (see first panels in Figs. 5.3 and 5.7, respectively). Firstly, in order to verify the statement that the IFs are not necessarily physically meaningful for the multi-component signals, the IFs are calculated directly from the signals of example 1 and 2 without applying a decomposition beforehand. Both signals consist of three components with a mono-component characteristic (see Figs. 5.3 and 5.7) and can accordingly be considered as multi-component signals. The signals and their Hilbert transforms are shown in the upper panels of Fig. 5.11 (example 1) and Fig. 5.12 (example 2), and the corresponding IFs obtained by using the three-points formula Eq. 5.23 are shown in the lower panels.

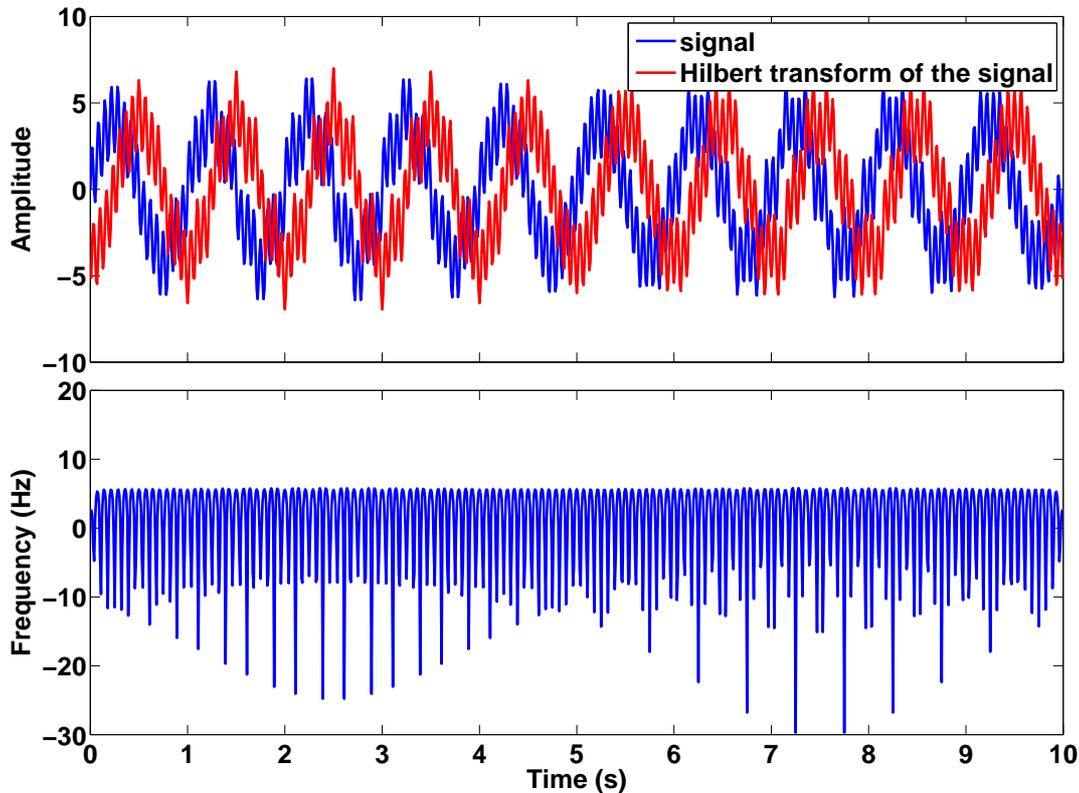


Figure 5.11: Upper panel: The synthetic time series from example 1 (blue curve) and its Hilbert transform (red curve). Lower panel: The IFs computed directly from the analytic signal (without decomposing the signal with EMD beforehand).

The obtained instantaneous frequencies in both examples vary very rapidly with time accommodating for the fact that at least two frequencies are present at each point in time. Moreover, negative instantaneous frequencies are present for many time samples. Obviously such a procedure does not recover the frequency characteristics of the signal and the derived IFs are not physically meaningful. Therefore a decomposition of the signal into "mono-component" functions is needed before IFs can be calculated. Like mentioned before, the EMD method exactly does this and accordingly resolves one crucial problem for computing

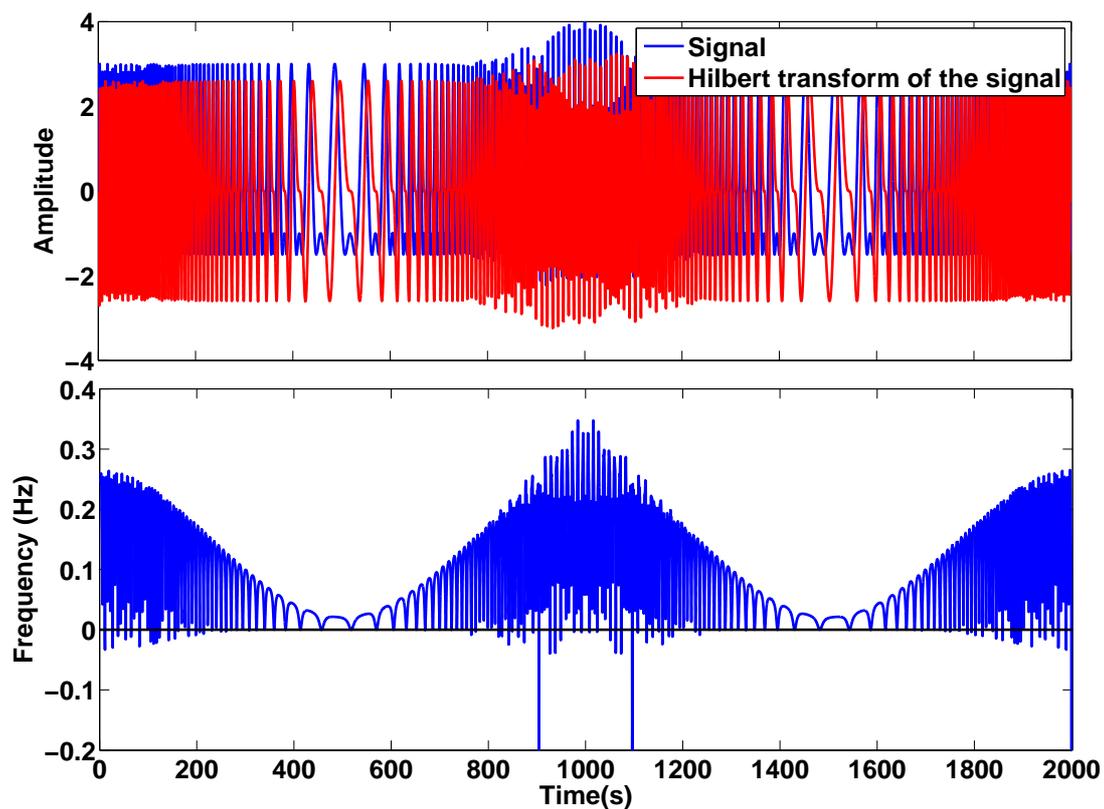


Figure 5.12: Upper panel: The synthetic time series from example 2 (blue curve) and its Hilbert transform (red curve). Lower panel: The IFs computed directly from the analytic signal (without decomposing the signal with EMD beforehand).

physically meaningful IFs from a multi-component signal.

The whole procedure of the Hilbert-Huang transform to determine the IFs from multi-component signals can be summarized as follows:

- 1). the EMD processing is performed and the signal is expressed as:

$$x(t) = \sum_{i=1}^n c_i(t). \quad (5.25)$$

In this case the monotonic residual r_n in Eq. 5.7 is omitted

- 2). For each IMF $c_i(t)$, an analytic signal is constructed with the Hilbert transform of $c_i(t)$ as the imaginary part:

$$u_i(t) = c_i(t) + i\mathcal{H}\{c_i(t)\} = a_i(t)e^{i\phi_i(t)} \quad (5.26)$$

$$\text{with } a_i(t) = \sqrt{c_i^2(t) + (\mathcal{H}\{c_i(t)\})^2}, \quad \phi_i(t) = \text{atan}(\mathcal{H}\{c_i(t)\}/c_i(t)).$$

- 3). The IFs are calculated for each IMF: $\omega_i(t) = d\phi_i(t)/dt$, $i = 1, 2, \dots, n$. By using the instantaneous amplitudes and frequencies the analytic signal can be written in the alternative form: $u_i(t) = a_i(t)e^{i\int\omega_i(t)dt}$, where each IMF $c_i(t)$ and the signal $x(t)$ can be expressed as:

$$c_i(t) = \text{Re}\{u_i(t)\} = a_i(t) \cos[i \int \omega_i(t)dt],$$

$$x(t) = \text{Re} \left\{ \sum_{i=1}^n a_i(t) e^{i \int \omega_i(t)dt} \right\} = \sum_{i=1}^n a_i(t) \cos \left[\int \omega_i(t)dt \right]. \quad (5.27)$$

Eq. 5.27 enables one to represent each IMF as a zero-mean amplitude-modulated & frequency-modulated (AM-FM) signal. In addition, Eq. 5.27 assigns amplitude and frequency content of the signal to each point in time and accordingly provides frequency-time distributions of the signal amplitudes. Huang et al. [1998] designated this distribution as *Hilbert spectrum* $S(\omega, t) = \sum_{i=1}^n S_i(\omega, t)$ (where $S_i(\omega, t)$ is frequency-time distribution of i -th IMF), since the Hilbert transform of all IMFs are involved to calculate the instantaneous frequencies. However, since it will be shown in the next subsection that IFs can be calculated by other methods as the HT, it is better to call $S(\omega, t)$ *instantaneous spectrum* instead.

Huang et al. [1998] moreover defined the *marginal spectrum* $s(\omega)$ as:

$$s(\omega) = \int_0^T S(\omega, t)dt, \quad (5.28)$$

where T is the total data length. The marginal spectrum offers a measure of the total amplitude contribution for each frequency value and represents the cumulated amplitude over

the entire data span in a probabilistic sense.

Eq. 5.27 also shows that, in contrast to generally infinite number of sine and cosine terms with constant amplitude and frequency in Fourier expansion, EMD expands signal $x(t)$ into a finite number of zero-mean AM-FM signals components. Although the expansion of the Fourier analysis and the EMD (see Eq. 5.27) are similar, the basis functions (IMFs) of the EMD are time variant which is not the case for the ones of the Fourier transform. Therefore the EMD can be considered as a generalization of the Fourier expansion. The variable amplitude and the instantaneous frequency distribution demonstrates that the EMD is highly flexible and can even handle non-stationary data.

Now the procedure of the Hilbert-Huang transform is illustrated by means of two examples. After decomposing the signal of the first example (see Fig. 5.5) using EMD, the first three IMFs of the synthetic time series and their Hilbert transform are plotted in Fig. 5.13. Their instantaneous spectra (IFs calculated by Eq. 5.23 and IAs calculated by Eq. 5.26) are shown in Fig. 5.14. In this plot, the color coding indicates the amplitude values of the analytic signal. The first three IMFs resemble the three components of the time series and their instantaneous frequencies and amplitudes are very similar to the constant frequencies of 15 Hz, 5 Hz and 1 Hz and the maximum amplitudes of 2, 1 and 4 of these components. The peaks of IFs for the second IMF can be explained with numeric instabilities due to very small values (see the second panel in Fig. 5.13) of the denominator in Eq. 5.23.

For example 2 the instantaneous spectra of the first three IMFs are plotted in Fig. 5.15. As for the first example the instantaneous frequencies and amplitudes of the first three IMFs largely reflect the frequency and amplitude contents of the three original components (see lower panel in Fig. 5.7). Again, some numerical induced peaks of IFs exist for the third IMF (Gaussian wavepacket).

In both synthetic time series examples, the instantaneous spectra of the IMFs show that the main characteristics of the original time series are recovered. This means that the obtained instantaneous frequencies from IMFs can be considered as physically meaningful and the EMD method indeed has the ability to decompose a multi-component time series into "mono-component" signals.

5.3.3 IF based on direct quadrature

The definition of IF based on the analytic signal is an important step forward in understanding the concept. However, Shekel [1953] pointed out that the polar coordinates expression $a(t)e^{i\phi(t)}$ in Eq. 5.12 is not unique which means that the combination of frequency and amplitude modulation is not unique, and Boashash [1992a] also demonstrated that one may define other methods for constructing a complex signal from a real signal as the Hilbert transform. Furthermore, even for mono-component or zero-mean AM-FM signals, there still are remaining difficulties if IFs and IAs are calculated by analytic signals that are based on the Hilbert transform [Boashash, 1992a, Huang and Wu, 2008, Huang et al., 2009].

In the most simple case, the Hilbert transform of a harmonic cosine function $\mathcal{H}\{a \cos(2\pi ft)\}$ is a sine function $a \sin(2\pi ft)$ and the real and imaginary parts of the analytic signal have

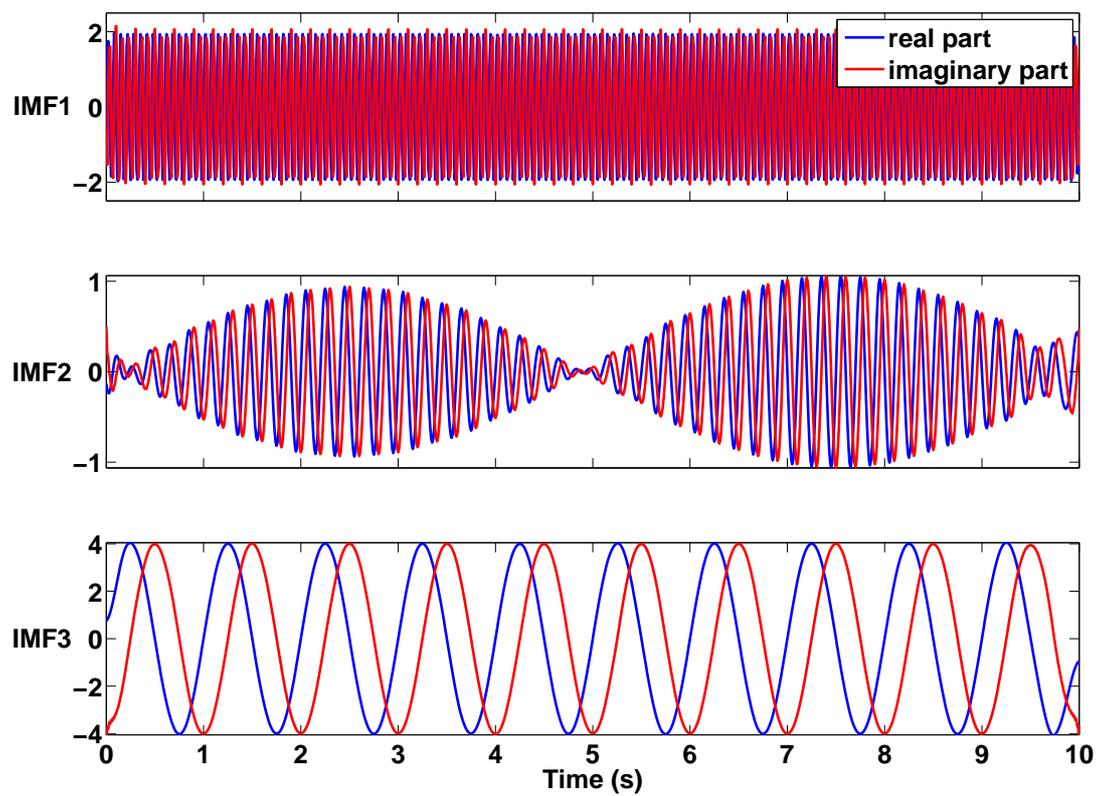


Figure 5.13: The first three IMFs (blue curves) of the synthetic time series of example 1 and their Hilbert transform (red curves).

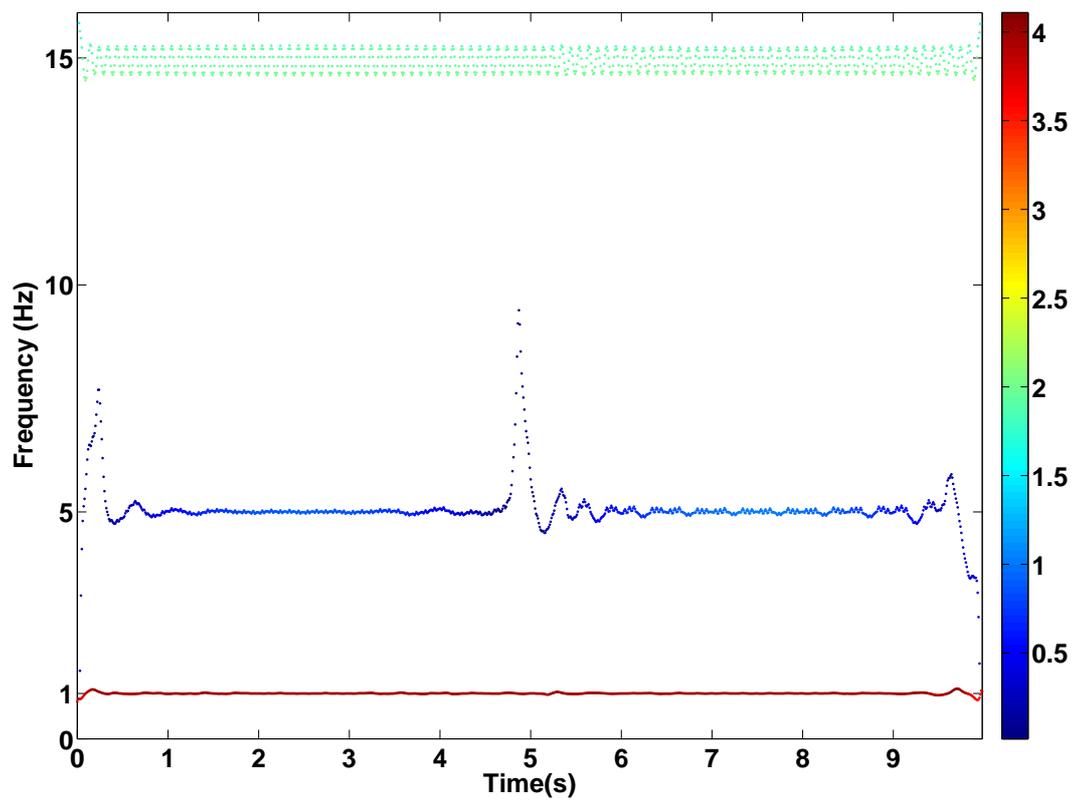


Figure 5.14: The instantaneous spectra of the first three IMFs of the synthetic time series of example 1 (see Fig. 5.13). Color coding indicates the amplitude level.

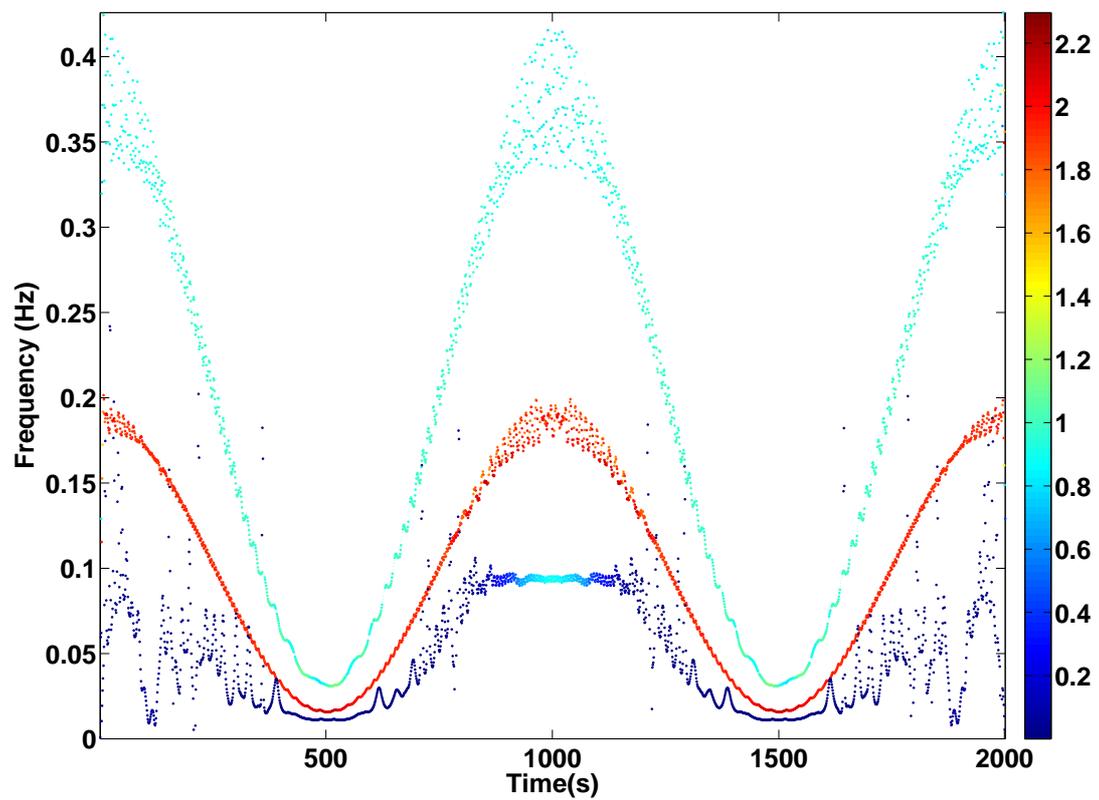


Figure 5.15: The instantaneous spectra of the first three IMFs of the synthetic time series of example 2 (the true instantaneous frequency & amplitude variations with time are shown in lower panel of Fig. 5.7). Color coding indicates the amplitude level.

everywhere a phase shift of $\pi/2$. However, this is not always true for a more general AM-FM signal of the form $x(t) = a(t) \cos \phi(t)$. There are two limitations on applying the HT-based analytic signal as discussed by Bedrosian [1963] and Nuttall and Bedrosian [1966].

Firstly, Bedrosian [1963] found out under which condition the Hilbert transform has the wanted property:

$$\mathcal{H}\{a(t) \cos \phi(t)\} = a(t)\mathcal{H}\{\cos \phi(t)\}, \quad (5.29)$$

such that the Hilbert transform produces a simple $\pi/2$ phase shift. He concluded that Eq. 5.29 is valid only if $a(t)$ varies slower with time than $\cos \phi(t)$ and the Fourier spectra $\mathcal{F}\{a(t)\}$ and $\mathcal{F}\{\cos \phi(t)\}$ are separated in frequency. If the two spectra are not separated in frequency, then the HT will give a result that represents the overlapping and phase-distorted functions. He showed that the Hilbert transform works as a type of high frequency selector and that the predominantly high frequency portion of the signal becomes the complex phase term.

For example, for a signal of the simple harmonic form of:

$$x(t) = \cos(2\pi f_1 t) \cos(2\pi f_2 t) \quad \text{with} \quad f_2 > f_1, \quad (5.30)$$

according to Bedrosian's theorem, its Hilbert transform will be:

$$\mathcal{H}\{x(t)\} = \cos(2\pi f_1 t)\mathcal{H}\{\cos(2\pi f_2 t)\} = \cos(2\pi f_1 t) \sin(2\pi f_2 t), \quad (5.31)$$

such that the analytic signal $u(t)$ is given by:

$$u(t) = x(t) + i\mathcal{H}\{x(t)\} = \cos(2\pi f_1 t)e^{i2\pi f_2 t}. \quad (5.32)$$

Secondly, Nuttall and Bedrosian [1966] furthermore questioned whether

$$\mathcal{H}\{\cos \phi(t)\} = \sin \phi(t), \quad (5.33)$$

is always true. They provided a mathematical proof that when the signal has a dominate frequency f_0 , i.e., $\phi(t) = 2\pi f_0 t + \varphi(t)$, Eq. 5.33 is valid if and only if the Fourier spectrum of $\cos \phi(t) + i \sin \phi(t)$ is zero for $f < -f_0$:

$$F(f) = \int_{-\infty}^{+\infty} [\cos \phi(t) + i \sin \phi(t)] e^{-i2\pi f t} dt = 0 \quad \text{when} \quad f < -f_0. \quad (5.34)$$

The critical issue here is that the HT-based analytic signal may not preserve the true phase function.

Both limitations discussed by Bedrosian [1963] and Nuttall and Bedrosian [1966] are theoretically verified and must be satisfied when HT-based analytic signal is used to calculate the IF. Otherwise, the calculated IF may lose its physical meaning.

When Huang et al. proposed in 1998 to calculate the IFs by using a HT-based analytic signal for each IMF, these limitations were not mentioned. To address these limitations, they proposed ten years later in Huang et al. [2009] an iterative amplitude normalization

scheme to demodulate AM and FM parts of the signal. The normalized FM has then the characteristic that the quadrature can be directly calculated and accordingly IFs can be determined through the direct quadrature function. Very important is that this procedure separates the AM and FM parts in a way that the difficulties pointed out by Bedrosian and Nuttall are resolved as will be seen later.

The demodulation procedure can be summarized as follow for a given IMF $c(t)$:

1. All local maxima of the absolute value of $c(t)$ are identified.
2. All maxima are connected with a cubic spline curve. This spline curve $e_1(t)$ is designated as the empirical envelope of $c(t)$. Next $c(t)$ is normalized by

$$y_1(t) = c(t)/e_1(t). \quad (5.35)$$

Ideally, all the values of $y_1(t)$ should vary between -1 and 1, in this case $y_1(t)$ is the empirical FM part of the IMF $c(t)$. However, this behavior is usually not derived after the first iteration and in this case $y_1(t)$ is taken as a new input signal. Step 1 and 2 are then repeated n times until all values of $y_n(t)$ vary between -1 and 1:

$$\begin{aligned} y_2(t) &= y_1(t)/e_2(t), \\ \dots & \end{aligned} \quad (5.36)$$

$$y_n(t) = y_{n-1}(t)/e_n(t).$$

3. The obtained $y_n(t)$ is designated as the empirical FM part of the IMF $c(t)$ that can be expressed as:

$$y_n(t) = \cos \phi(t) = F(t). \quad (5.37)$$

Since $y_n(t)$ is characterized by a change in frequency but not in amplitude over time, $F(t)$ is surely a purely frequency modulated function with unity amplitude.

4. After the FM part is determined the AM part $A(t)$ is simply determined by,

$$A(t) = e_1(t)e_2(t) \cdots e_n(t). \quad (5.38)$$

It is obvious that

$$c(t) = A(t)F(t) = A(t) \cos \phi(t). \quad (5.39)$$

Thus, one accomplished the empirical AM-FM decomposition through repeated normalization.

The *direct quadrature* (DQ) of the normalized IMF $F(t)$ is easily derived as

$$\sin \phi(t) = \sqrt{1 - F^2(t)}. \quad (5.40)$$

The instantaneous frequency can then be obtained from the phase function $\phi(t)$:

$$f_{ins} = \frac{d\phi(t)}{dt} \quad \text{where} \quad \phi(t) = \text{atan} \frac{F(t)}{\sqrt{1 - F^2(t)}}. \quad (5.41)$$

From the DQ of the frequency modulated signal $F(t)$ (see Eq. 5.40), the direct quadrature for the IMF $c(t)$ can then be simply determined as:

$$c_{DQ}(t) = A(t) \sin \phi(t). \quad (5.42)$$

This is the reason why the method is called direct quadrature method. Bedrosian theorem is then obviously satisfied due to the fact that only the pure FM part is transformed and the AM part remains the same.

A complex signal can then be constructed by taking $c(t)$ as real part and its direct quadrature as the imaginary part,

$$u_{DQ}(t) = c(t) + ic_{DQ}(t) = A(t) \cos \phi(t) + iA(t) \sin \phi(t). \quad (5.43)$$

The complex pair is not necessarily analytic, since the imaginary part typically differ from the one of the HT (see Eq. 5.26). However, in contrast to the HT a phase shift of $\pi/2$ between the real and imaginary part is always guaranteed and accordingly the correct phase function is preserved without any distortions. This means Nuttall theorem is satisfied since the only difference between the real and imaginary parts is that the FM signal is changed from cosine to sine.

It should be noted that due to numerical inaccuracies, the normalization process may not perfectly separate the FM and AM parts of the original IMF from each other. However, since both the positions of the extrema and the positions of the zero-crossing points remain totally unchanged by the normalization process, the amount of the associated deformation is negligible.

Example illustrating the demodulation procedure

Now the demodulation procedure is illustrated by means of the first IMF of the synthetic time series of example 2 (see section 5.2). The upper panel of Fig. 5.16 shows the IMF 1 (blue curve), the absolute value of this IMF 1 (red dashed curve), the envelope derived from HT-based analytic signal $|u(t)|$ (black curve) and the first envelope $e_1(t)$ (see Eq. 5.35) derived from the demodulation procedure (magenta curve). It can be seen that the empirical envelope $e_1(t)$ is smoother and devoid of higher frequency fluctuations and overshoots in comparison to the envelope derived from the HT-based analytic signal. The IMF 1 (blue curve) and the IMF 1 normalized by $e_1(t)$ (red dashed curve) are shown in the lower panel of Fig. 5.16. It is evident that the amplitude variation of the normalized IMF 1 are reduced. After a second iteration applied on the normalization (not shown here), all extrema have amplitude values of exactly 1 or -1 and the normalized IMF 1 now represents the pure FM part of the IMF 1. From the FM part, instantaneous frequencies of the IMF 1 can then be calculated by Eq. 5.41 and the AM part can also be obtained by Eq. 5.38.

The instantaneous spectra derived by using the DQ method for the first three IMFs of the two synthetic examples are shown in Fig. 5.17. The color coding represents again the amplitude level. As a comparison, the IFs calculated from the HT-based analytic signal are also plotted as a thin grey curves without amplitude information. It can be observed in Fig. 5.17 that for

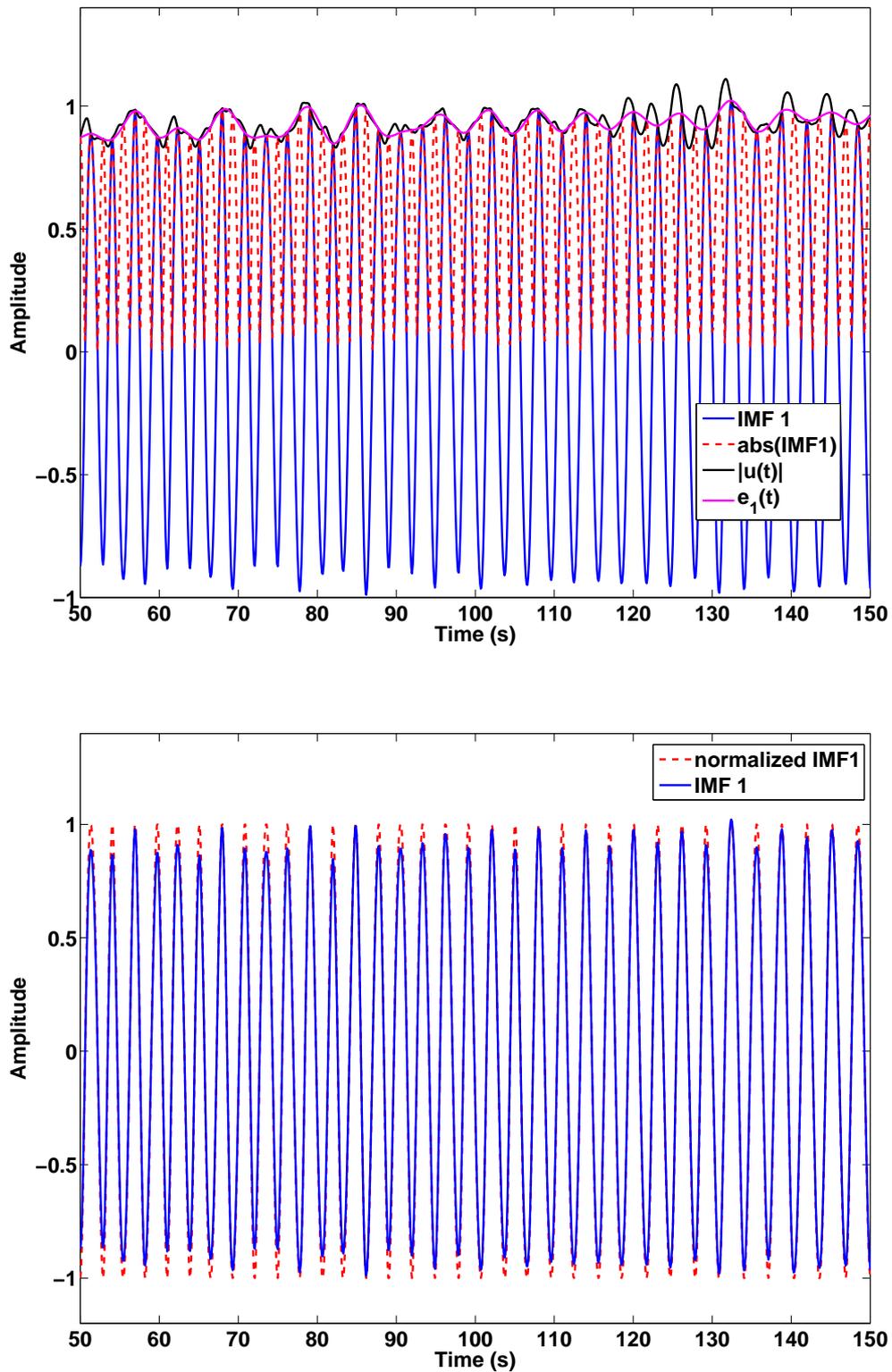


Figure 5.16: Illustration of the demodulation process. Upper panel: The IMF 1 (blue curve) of the synthetic time series of example 2, the absolute value of this IMF 1 (red dashed curve) and two envelopes calculated from this IMF. One envelope is obtained from HT-based analytic signal $|u(t)|$ (black curve), the other is the first envelope $e_1(t)$ (see Eq. 5.35) derived from the demodulation procedure (magenta curve). Lower panel: The IMF 1 before (blue curve) and after (red dashed curve) normalization with $e_1(t)$. Note that both figures are zoomed in on a small time interval from 50 to 150 s.

both examples frequency and amplitude contents of the three component of the time series are correctly resolved by the first three IMFs. The IF calculation is, moreover, significantly more robust and accurate when the DQ method is used. This observation is in agreement with the conclusions by Huang et al. [2009].

The direct quadrature method offers an easy and direct way to compute the IF from any IMF. In contrast to the HT-based analytic signal, the quadrature is derived without any integration or transform, which indicates that the local characteristics of the time series are extracted and no global information is needed. More accurate IF can be calculated by using the DQ method since limitations that are associated with Bedrosian's and Nuttall's theorems for Hilbert transform are avoided. Therefore I will always consider in the following the direct quadrature as the imaginary part of a time series in both the construction of a complex signal and the calculation of IFs.

5.4 Is an intrinsic mode function a mono-component signal?

The results from the two synthetic examples demonstrate that a multi-component time series can be decomposed into intrinsic mode functions using the EMD method. The instantaneous frequencies and amplitudes calculated from each intrinsic mode function largely coincide with the true frequency and amplitude contents of the original components of the time series. This suggests that the intrinsic mode functions can be generally considered as mono-component time series. Although, many authors believe this is true and state that EMD can decompose a multi-component signal into mono-component signals [e.g. Rilling and Flandrin, 2008, Kopsinis and McLaughlin, 2008b, Stevenson et al., 2010], this conclusion is unfortunately not proven.

The concept of a "mono-component" signal has been used widely in engineering as a loose concept. However, a number of people [e.g. Boashash, 1992a, Huang et al., 1998] state that the precise criteria for what actually constitutes a "mono-component" signal has not been developed yet, probably for the simple reason that no decomposition methodology had existed prior to EMD to split up a signal into something that could be viewed as "mono-component" sequences. Accordingly the discussion of "mono-component" signals always tend to be rather vague in literature.

Cohen [1992] considered a signal as mono-component if its time-frequency distribution is a single "ridge", i.e. a single delineated region of energy concentration. Boashash [1992a,b] defined the mono-component signal as one whose time-varying IFs $f_{ins}(t)$ are invertible ($f_{ins}^{-1}(t)$ exists) and accurately represent the frequency modulation law [e.g. Delprat, 1997].

Intuitively, a mono-component signal should be a complex sinusoid possibly modulated in amplitude and/or frequency, i.e. a signal of the form $a(t)e^{i\phi(t)}$, where $a(t)$ is real and positive and $\phi(t)$ is real. Picinbono [1997] mentioned that physical meaningful instantaneous frequencies based on Hilbert transform are only guaranteed for complex signals of the form

$$z(t) = a(t)e^{i\phi(t)} \quad (5.44)$$

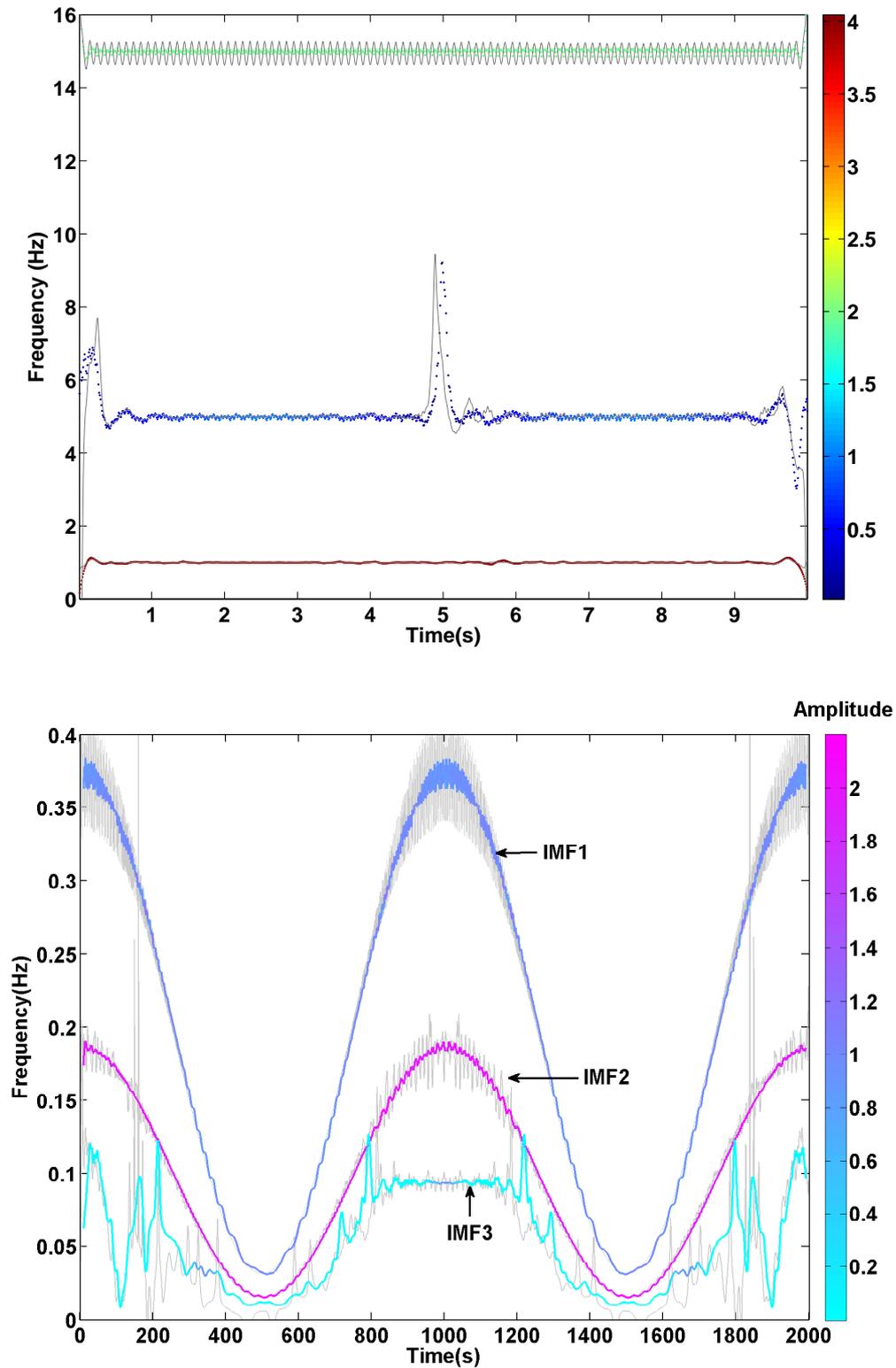


Figure 5.17: Upper and lower panels show the instantaneous spectrum of the first three IMFs of the synthetic time series of example 1 and 2, respectively. In both figures the color coding indicates the amplitude level. As a comparison IFs which are calculated by using HT-based analytic signal are plotted as thin grey curves. For these instantaneous spectra no amplitude information are given (see Figs. 5.14 and 5.15 for amplitude information).

where $a(t)$ is nonnegative and the time derivative of the phase function $\phi(t)$ is the instantaneous frequency. Since Eq. 5.27 can be associated with the real part of eq. 5.44, this comment is particularly important to understand why IMFs can also be considered as mono-component (as long as numeric inaccuracies during the sifting process are ignored). Considering the sifting process described in Eqs. 5.1 - 5.4, it is an iterative way of removing the dissymmetry between the upper and lower envelopes in order to transform the original signal into a signal with symmetric envelopes, i.e. an AM signal. Therefore, the obtained IMFs are amplitude modulated. When numeric inaccuracies are ignored, the criteria for IMFs are responsible for continuously undulating zero mean signals that are of the form $a(t) \cos[\int_0^t \omega(\tau) d\tau]$ or more simplified of the form $a(t) \cos[\theta(t)]$ (see Eq. 5.27 and Huang et al. [1998, p. 928]), which means the instantaneous frequency $\omega(t)$ of an IMF changes with time and an IMF is also frequency modulated. Hence, IMFs have the form of AM-FM signals. Thus one can say that the EMD is nothing else but a decomposition which separates a signal $x(t)$ into a set of AM-FM signals.

On the other hand, it is obvious that sinusoidal signals $\sin(\omega t)$ or $\cos(\omega t)$ for any real ω and t fulfill the two conditions posed on the IMFs. Rato et al. [2008] pointed out that this is also true for the pure FM signals $\sin[\phi(t)]$ and $\cos[\phi(t)]$ with continuous function $\phi(t)$. However, an AM-FM signal in the form $A(t) \cos[\phi(t)]$ does not necessarily fulfill the two conditions even if $A(t)$ is a sinusoid unless that $A(t)$ changes slowly compared to changes in $\cos[\phi(t)]$. It is easy to understand that in the EMD procedure the envelopes can not vary as fast as the signal $x(t)$. The reason for this is that when the sifting is performed, low frequency components are removed and a high frequency component remains and IMFs appear in a high to low frequency order, such that the EMD behaves like a bank of filters [Flandrin et al., 2004]. Therefore, one can draw a conclusion that IMFs can accordingly considered as mono-component signals, since IMFs are AM-FM signals whose AM parts change slowly compared to changes in FM parts.

Of course, the above discussion is surely dependent on the definitions of physically meaningful instantaneous frequencies and mono-component signals. Without accurate definitions, it is difficult to obtain more general understanding of instantaneous attributes and mono-component signals. In this way one can think that EMD is quite remarkable, because it really helps to understand these concepts better.

5.5 Instantaneous spectrum and Fourier spectrum

When Huang et al. [1998] introduced the instantaneous spectrum (Hilbert spectrum) $S(\omega, t)$ (Eq. 5.27) and the marginal spectrum $s(\omega)$ (Eq. 5.28), they mentioned that both of them have a totally different meaning from the Fourier spectrum. The existence of energy at the frequency ω in the instantaneous spectrum or marginal spectrum means that an oscillation with this frequency occurred during a point in time or a time interval within the time span of the data, respectively. In the Fourier representation energy at a frequency ω means that a component of a sine or a cosine signal persists through the entire time span of the data. Such discrepancies had made it difficult to compare the Hilbert spectrum with the traditional Fourier spectrum.

Before Huang et al.'s publication from 2011, the true meaning and the properties of the

Hilbert spectral representation had not been well understood. Except for the first publication by Huang et al. [1998] probably only Wen and Gu [2009] pointed out that an original marginal spectrum which consists of the integration of the Hilbert energy spectrum with time gives the total energy rather than the energy density as in the Fourier spectral representation. Figs. 5.18 and 5.19 show the marginal spectra and Fourier amplitude spectrum of the synthetic time series of example 1 and 2, respectively. The marginal spectra are calculated from Eq. 5.28, and the Fourier amplitude spectrum are calculated by using a multi-taper method which uses orthogonal discrete prolate spheroidal sequences tapers to compute the combinations of modified periodograms. In both examples (except for the y - scales which are totally different), the spectra are very similar. Especially in example 1, one can clearly see that the amplitude peaks in both spectra are at frequencies 1 Hz, 5 Hz and 15 Hz, which correspond to the frequencies of the three components of the original signal.

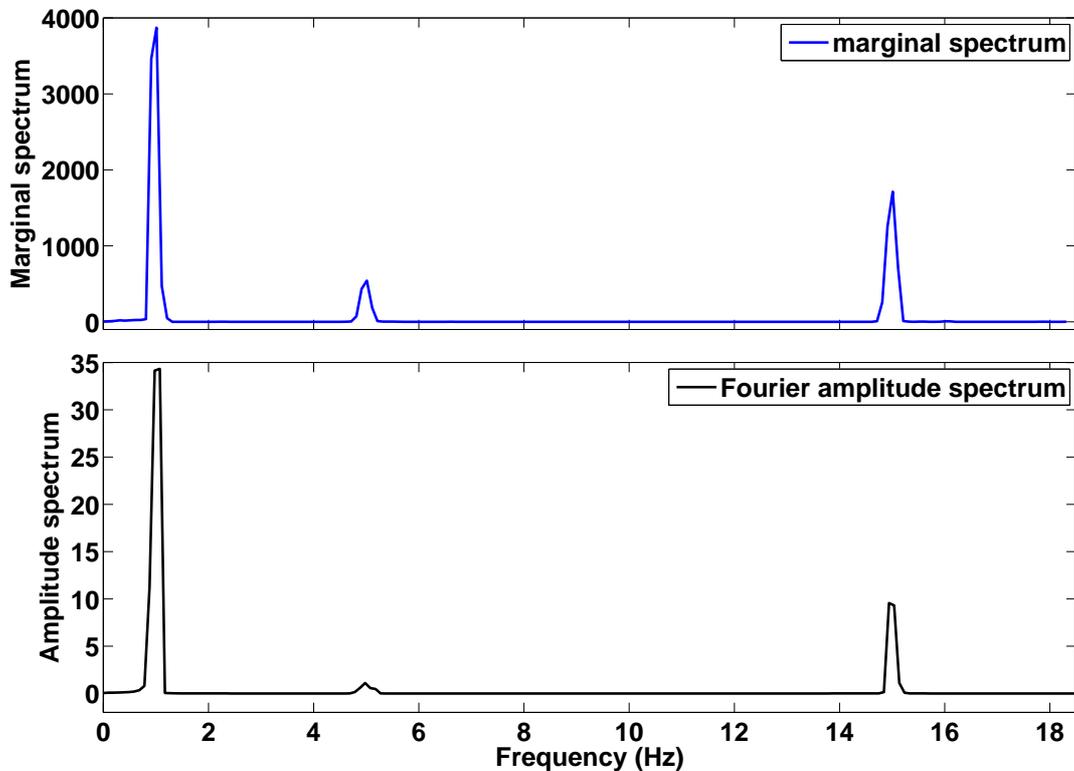


Figure 5.18: Upper panel: The marginal spectrum of the synthetic time series of example 1. Lower panel: The Fourier amplitude spectrum of the synthetic time series of example 1.

Huang et al. [2011] investigated the difference between the Hilbert spectrum and Fourier spectrum in more detail. In order to make them comparable, they rigorously defined the Hilbert energy density and amplitude density in the time-frequency space.

According to Eq. 5.28, the time-frequency space is divided into $m_1 \times m_2$ bins with equal size

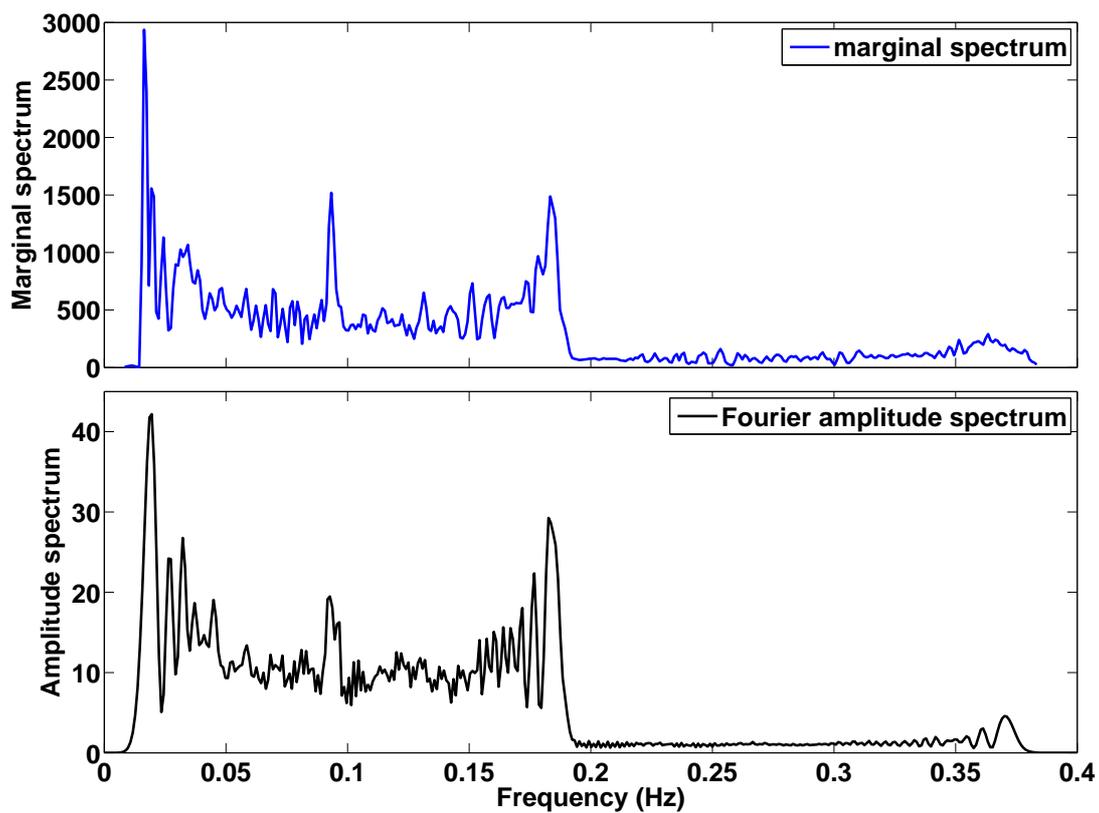


Figure 5.19: Upper panel: The marginal spectrum of the synthetic time series of example 2. Lower panel: The Fourier amplitude spectrum of the synthetic time series of example 2.

$\Delta t \times \Delta \omega$:

$$\begin{aligned} t_0, t_0 + \Delta t, \dots, t_0 + i\Delta t, \dots, t_1, & \quad i = 1, \dots, m_1, \\ \omega_0, \omega_0 + \Delta \omega, \dots, \omega_0 + j\Delta \omega, \dots, \omega_1, & \quad j = 1, \dots, m_2, \end{aligned} \quad (5.45)$$

thus

$$t_i = t_0 + i\Delta t \quad \text{and} \quad \omega_j = \omega_0 + j\Delta \omega, \quad (5.46)$$

where the numbers and sizes of the time and frequency bins could be selected independently and arbitrarily, as long as they obey the following constrains:

1. t_0 and t_1 have to be within the data interval $[0, T]$.
2. Δt cannot be smaller than the sampling rate.
3. ω_0 should not be smaller than the Nyquist frequency.

Assume a real signal $x(t)$ after EMD decomposition is expressed as: $x(t) = \sum_{i=1}^n a_i(t) \cos[\int \omega_i(t) dt]$ (see Eq. 5.27). Then in each bin, the Hilbert energy density spectrum, $S_{i,j}$, and the amplitude density spectrum, $A_{i,j}$, are respectively defined as:

$$\begin{aligned} S_{i,j} &= \frac{1}{\Delta t \times \Delta \omega} \left\{ \sum a_k^2(t) : t \in \left(t_i - \frac{\Delta t}{2}, t_i + \frac{\Delta t}{2} \right], \omega \in \left(\omega_j - \frac{\Delta \omega}{2}, \omega_j + \frac{\Delta \omega}{2} \right] \right\}, \\ A_{i,j} &= \frac{1}{\Delta t \times \Delta \omega} \left\{ \sum a_k(t) : t \in \left(t_i - \frac{\Delta t}{2}, t_i + \frac{\Delta t}{2} \right], \omega \in \left(\omega_j - \frac{\Delta \omega}{2}, \omega_j + \frac{\Delta \omega}{2} \right] \right\}, \end{aligned} \quad (5.47)$$

where k indicates k -th IMFs, $k = 1, 2, \dots, n$.

In the discrete version of the traditional Fourier spectral representation, the time scaling does not appear in the formulation and the frequency scaling is determined by the sampling rate Δt and the data span T , through the frequency resolution $\Delta \omega = \frac{1}{T}$ and the Nyquist frequency $\frac{1}{2\Delta t}$. However, in Hilbert spectrum both the time and frequency scaling Δt and $\Delta \omega$ appear explicitly.

From Eq. 5.47 the marginal energy density, $h(\omega_j)$, and the marginal amplitude spectrum, $h_a(\omega_j)$, respect to frequency ω_j are then defined as:

$$\begin{aligned} h(\omega_j) &= \sum_{i=1}^{m_1} S_{i,j} \Delta t = \frac{1}{\Delta \omega} \sum_{i=1}^{m_1} a_i^2(t), \\ h_a(\omega_j) &= \sum_{i=1}^{m_1} A_{i,j} \Delta t = \frac{1}{\Delta \omega} \sum_{i=1}^{m_1} a_i(t), \end{aligned} \quad (5.48)$$

where m_1 is number of bins with respect to time axis.

Huang et al. [2011] concluded that the conversion factor between marginal and Fourier spectra is T/N (where N is the number of time samples), which corresponds exactly to the sampling rate Δt . This means that Fourier spectral values need to be multiplied by N/T to convert Fourier spectral density to marginal energy density. Conversely, marginal energy density need to be multiplied by T/N to make it comparable to Fourier spectral density.

The actual determination of the marginal power density is, however, not such straightforward, because one can select an arbitrary frequency resolution bin size and an arbitrary frequency range. This freedom leads to a modification of the conversion factor. For different frequency resolutions $\Delta\omega$ and $\Delta\omega_1$ the corresponding marginal power density spectra $h(\omega)$ and $h_1(\omega)$ have the relation:

$$h_1(\omega) = \frac{\Delta\omega}{\Delta\omega_1} h(\omega). \quad (5.49)$$

For example, for a time series given by

$$x(t) = \sin(\pi t/256) \quad \text{with} \quad t = 0, 1, \dots, 1023, \quad (5.50)$$

the energy density is $\sum x_i^2(t)/1024 = 0.5$. When the time series is decomposed by EMD, only one IMF (and a residue that has everywhere zero values) is obtained. The IFs of this IMF are 0.002 everywhere. The marginal power densities are then calculated using 256, 512, 1024 and 2048 equally sized bins in frequency direction ranging from 0 Hz to the Nyquist frequency corresponding to frequency resolutions of 0.5/256 Hz, 0.5/512 Hz, 0.5/1024 Hz and 0.5/2048 Hz, respectively. A comparable Fourier spectrum is determined without applying any de-trending or other type of tapering to the original time series. The Fourier and the marginal power densities of the data are shown in Fig. 5.20. Note that the two densities are equal when the number of bins in frequency direction is same for the marginal power density and Fourier spectrum (512 frequency bins).

In Figs. 5.21 and 5.22 the Fourier power density spectra and marginal power density spectra with different frequency resolutions are compared for the synthetic time series of example 1 and 2. Both examples show that the two types of spectra are quantitatively comparable when the simple conversion formula described above is used. However, there are also some distinct differences between the two spectra; in particular for the synthetic example 2. Because of the non-stationarity of this time series, many additional harmonic components are needed when Fourier transform is applied on the data without de-trending and tapering the time series before. As a result, the energy is spread over a wider frequency range up to ~ 0.4 Hz. Constrained by the energy conservation, these spurious harmonics and the wide frequency spectrum cannot faithfully represent the true energy density in the frequency domain. One can observe in Fig. 5.22 that the Fourier power density spectrum exhibits some ringing which is associated with spectral leakage. This situation can be improved when a multi-taper method is involved in calculation of the Fourier power density spectrum as shown in Fig. 5.19.

As already mentioned above that in Fourier spectral analysis the frequency resolution is determined entirely by the length of the data set and its sampling rate, while for the marginal spectrum both time and frequency resolutions can be arbitrarily selected. This means that the frequency resolution of the EMD method is independent of the length of the data set

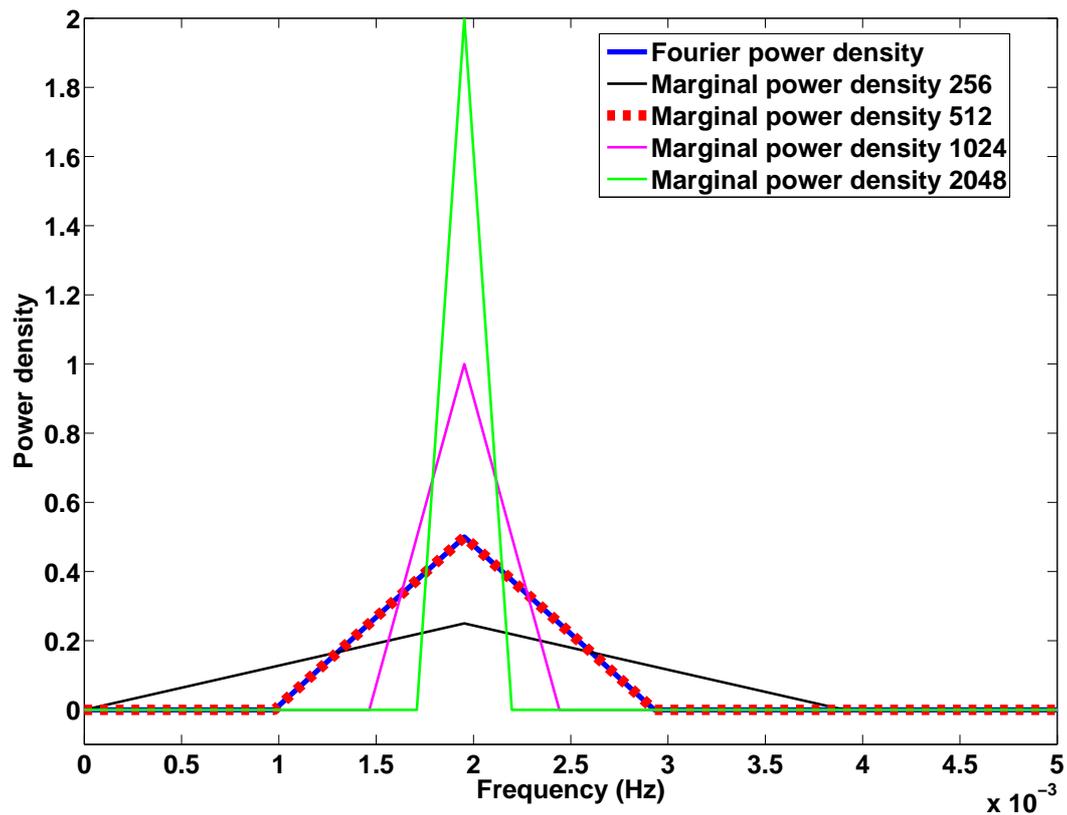


Figure 5.20: Comparison of the Fourier and marginal power density spectra of the sinusoid signal $x(t) = \sin(\pi t/256)$ with $t = 0, 1, \dots, 1023$. The numbers given for the different marginal power densities correspond to the number of bins in frequency direction and, hence, indicate the frequency resolutions.

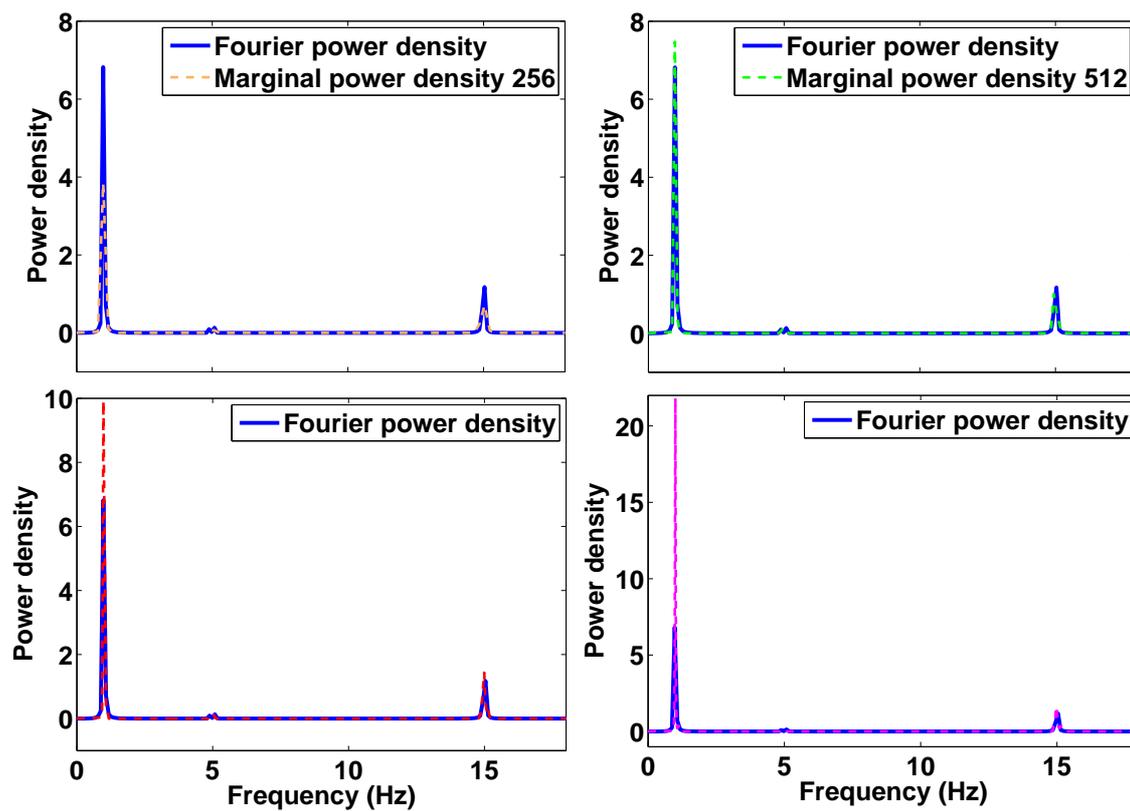


Figure 5.21: Comparison of the Fourier and marginal power density spectra of synthetic time series of example 1. The numbers given for the marginal power densities correspond to the numbers of equally sized bins used in frequency direction.

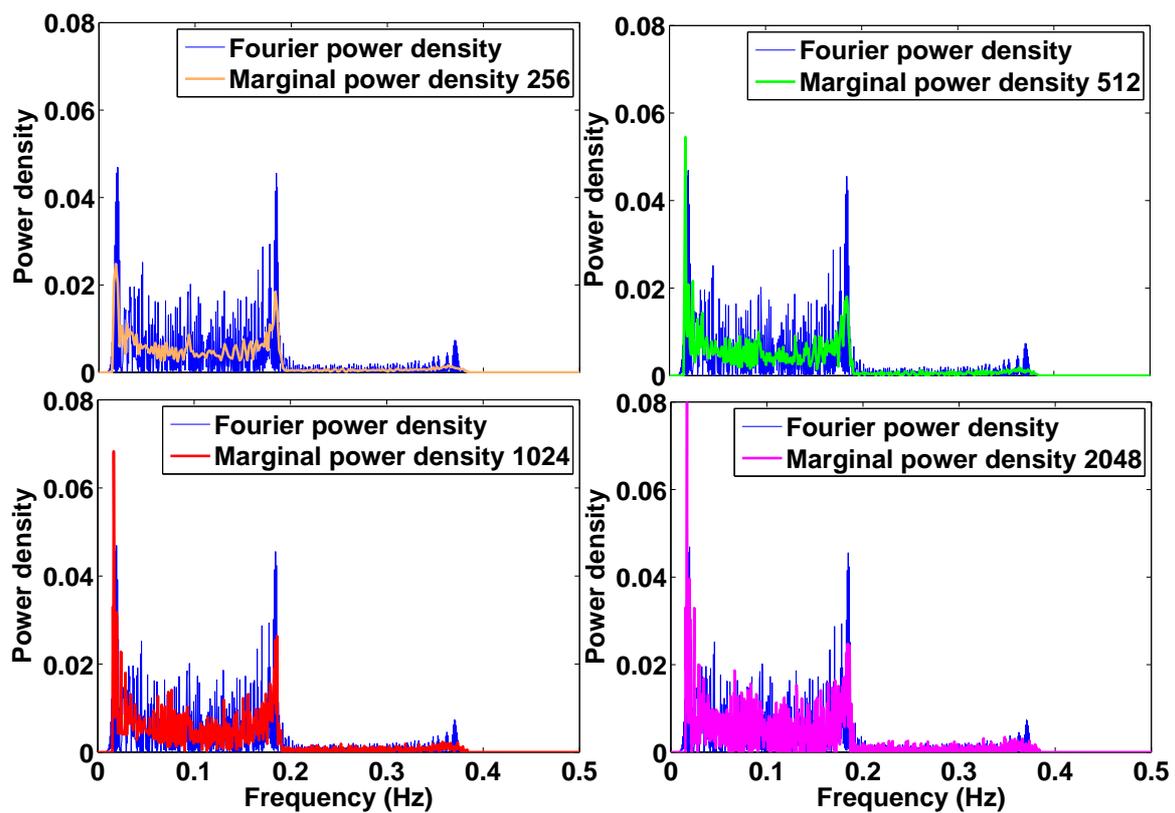


Figure 5.22: Comparison of the Fourier and marginal power density spectra of synthetic time series of example 2. The numbers given for marginal power densities correspond to the number of bins in frequency direction.

and the sampling rate (as long as its inverse does not exceed the Nyquist frequency). If one considers Figs. 5.20 - 5.22 it becomes clear that the width of the energy peaks is narrowed for smaller sampling intervals, however, otherwise the marginal power density spectra are rather similar for different frequency resolutions. Another advantage of the marginal spectrum is that bins sizes in time direction can also be chosen flexibly. This allows to zoom in onto some special time intervals of the data to better understand the local physical properties. Another limitation of the Fourier transform is related to the assumption that the time series is considered as an oscillating signal that is repeated after the last sample of the actual time series. Therefore the best way to choose the length of the time series for Fourier analysis is that the time series contains whole cycles of the signal or that it has zero amplitudes and slopes at the beginning and end points which is typically reached by using tapers. If this is not done artificial harmonics are introduced that distort the whole Fourier spectrum. In contrast, the length of the time series can be flexibly chosen for EMD without distorting the overall frequency content because this method works directly in the time domain.

Let us consider these effects on a simple sinusoidal signal with a phase shift of $\pi/3$,

$$x(t) = \sin(\pi t/256 + \pi/3), \quad (5.51)$$

and two different lengths of the time series of 1023 s and 800 s are considered (see Fig. 5.23). The signal energy densities for the two intervals are 0.5 and 0.516, respectively. When EMD is used to decompose the data, for both time lengths one IMF is obtained that contains the data itself with $IF = 0.002$ Hz at every data point and a residue that has everywhere zero values. The corresponding Fourier and marginal power densities are shown in the lower panel of Fig. 5.23, where the numbers of frequency bins in both Fourier and marginal power densities are chosen to be equal (512 bins). For the time interval of 1023 s, there is no energy leakage in the Fourier spectrum, since the time series contains exact two circles of sinusoidal wave. However, for the shorter time interval of 800 s, a leakage occurs in Fourier spectrum due to the Gibbs phenomenon. Technically, the Gibbs phenomenon can be taken care of by applying tapering on the data before Fourier transform. However, in the marginal power density, this situation does not occur since there is no periodic restriction on the data in the EMD method, which makes tapering unnecessary.

The same effect can be illustrated using the previous synthetic time series of example 1. Two different time intervals of 10 s and 8.8 s with a constant sampling interval of $\Delta t = 0.01$ s are used. The first time interval contains a number of complete cycles for all three components that build up the total signal of example 1. While for the second time interval the last cycle of all three components is incomplete. The Fourier and marginal spectra of the two time intervals are shown in Fig. 5.24. If one compares the two Fourier power densities (blue curve for longer interval, black curve for shorter interval) and two marginal power densities (red dashed curve for longer interval, green curve for shorter interval), it can be observed that there are distinct differences in the two Fourier spectra but not in the two marginal spectra.

In addition to the power density spectra, now the complex spectra of Fourier analysis and the EMD method will be also compared with each other. This point is considered interesting because it gives an idea how the phase spectra of the different methods are affected by the the length of the time series, a point that has not been discussed yet in literature. Compared to Fourier analysis, which transfers the signal into frequency domain, the EMD

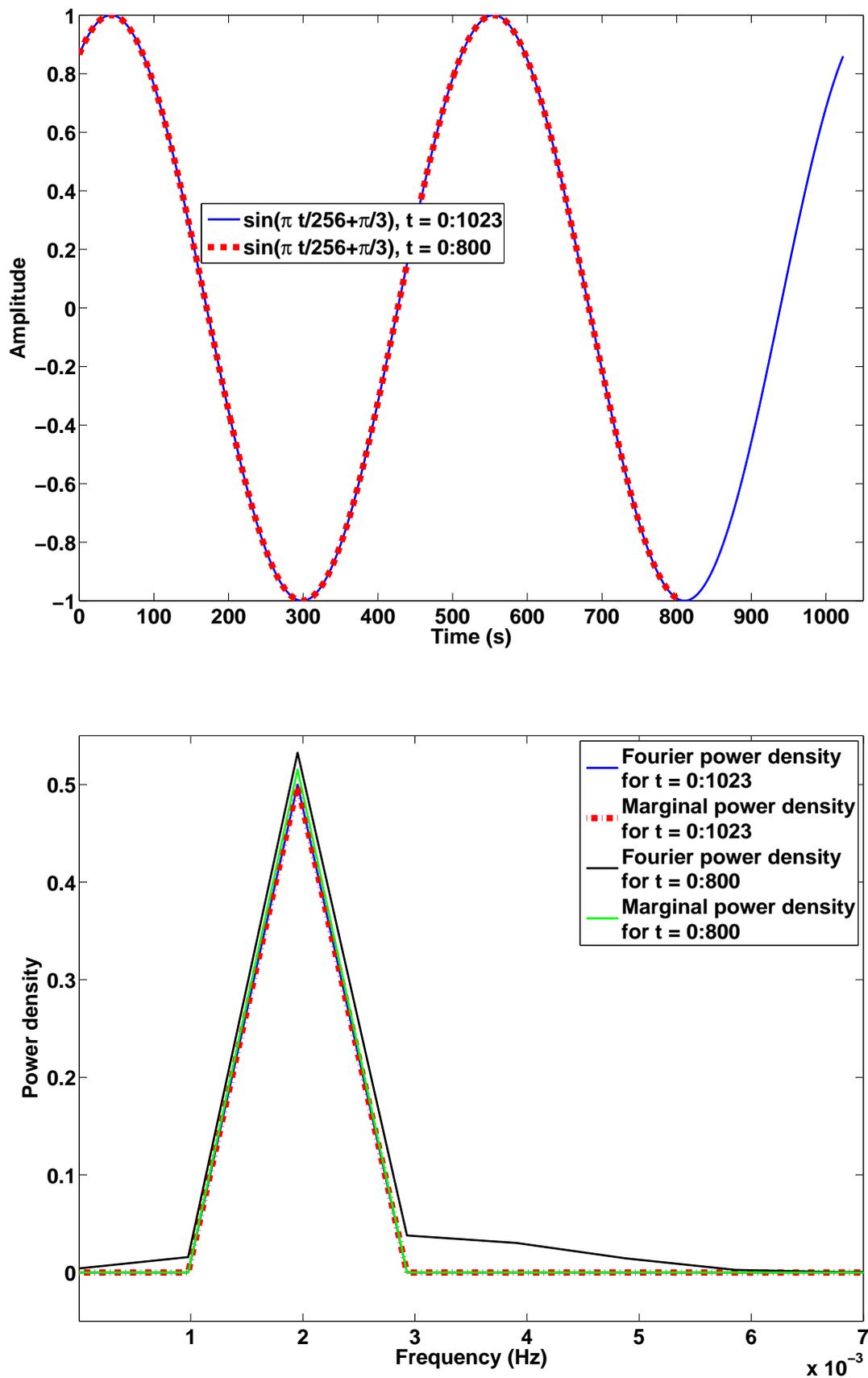


Figure 5.23: Upper panel: Two time series of a sinusoidal signal with lengths of $t = 1023$ s and $t = 800$ s. Lower panel: Comparison of the Fourier and marginal power density spectra of the sinusoidal signal for both time lengths. The same binning in frequency direction (512 bins) is used for the calculation of the spectra.

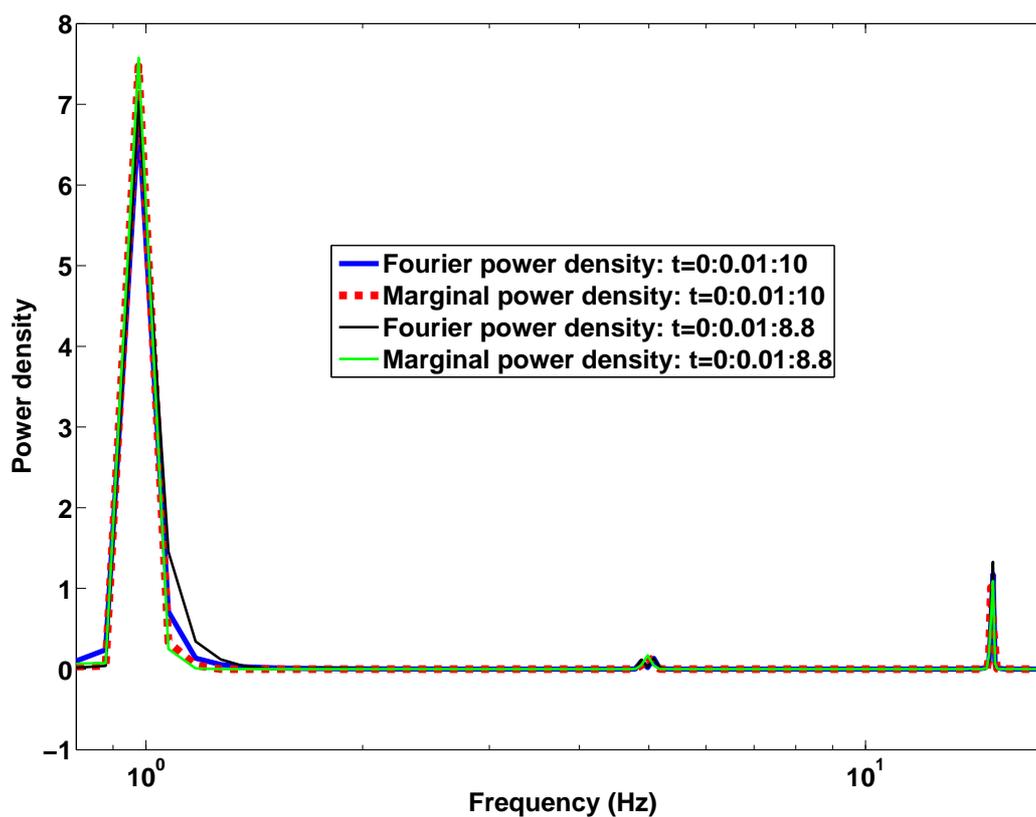


Figure 5.24: Comparison of the Fourier and marginal power density spectra of synthetic time series of example 1 for time intervals of 10 s and 8.8 s (sampling interval $\Delta t = 0.01$ s). The bin number in frequency direction (512 bins) is the same for both spectra.

represents the signal in time-frequency domain. To make their complex spectra comparable, one can transfer the time-frequency representation of EMD into frequency domain. This can be done by summing up all contributions of the IMFs in each predefined frequency bin, $(\omega_j - \Delta\omega/2, \omega_j + \Delta\omega/2]$, in time domain,

$$x_{\omega_j}(t) = \sum_{i=1}^n \{c_i(t) : \text{instantaneous frequency } \omega \text{ satisfies } \omega_j - \Delta\omega/2 < \omega \leq \omega_j + \Delta\omega/2\}. \quad (5.52)$$

It is trivial to show that $\sum_{j=1}^{m_2} x_{\omega_j}(t) = x(t)$, where m_2 is the number of the frequency bins. Then, similar to discrete Fourier transform, in each frequency bin the obtained time series is multiplied by exponential function $e^{-i\omega t}$ to yield a spectrum, which is called the marginal complex spectrum:

$$x_{EMD}(\omega_j) = \sum_{k=1}^N x_{\omega_j}(t_k) e^{-i2\pi\omega_j t_k}, \quad (5.53)$$

where N is the number of data sample.

For the simple sinusoidal signal given by Eq. 5.51 with a time interval of 1023 s, the time series $x_{\omega=0.002}(t)$ in Eq. 5.52 is exactly the signal itself, since the IMF 1 with IFs of 0.002 for all data points is identical to the signal. Therefore, the real and imaginary parts of complex spectrum are identical with the Fourier complex spectrum, as can be seen by the red dashed curves and the blue curves shown in Fig. 5.25. The black (Fourier) and green (marginal) curves plotted in Fig. 5.25 are the complex spectra for the shorter time interval of 800 s. In contrast, the Fourier complex spectrum is obviously distorted for this shorter time interval due to the discontinuity at the boundary points (Gibbs phenomenon), and much ringing appears in both real and imaginary parts.

When the synthetic time series of example 1 with the time intervals of 10 s and 8.8 s is tested, one observes a similar effect as described above. The real and imaginary parts of both spectra are plotted in Fig. 5.26. The spectrum corresponding to the shorter time interval exhibits energy leakage in the Fourier complex spectrum but not in the marginal complex spectrum.

The results show that both the Fourier power density and the complex Fourier spectrum are distorted due to changes in the length of the data set, unless some type of windowing or tapering is applied prior to the transform. In contrast the marginal power density and complex spectrum are not distorted. With the freedom of choosing arbitrary time and frequency bin sizes, the EMD method can give results with high frequency resolution and can reveal more details in local time intervals. This demonstrates that the local properties of the EMD method can be useful in many ways.

5.6 Some limitations of the EMD method

EMD is an empirical method for which explanations and proofs are still missing for a number of aspects. Furthermore, different numerical implementations exist for the EMD method to reach the proposed criteria for IMFs and a number of parameters has to be set such that the

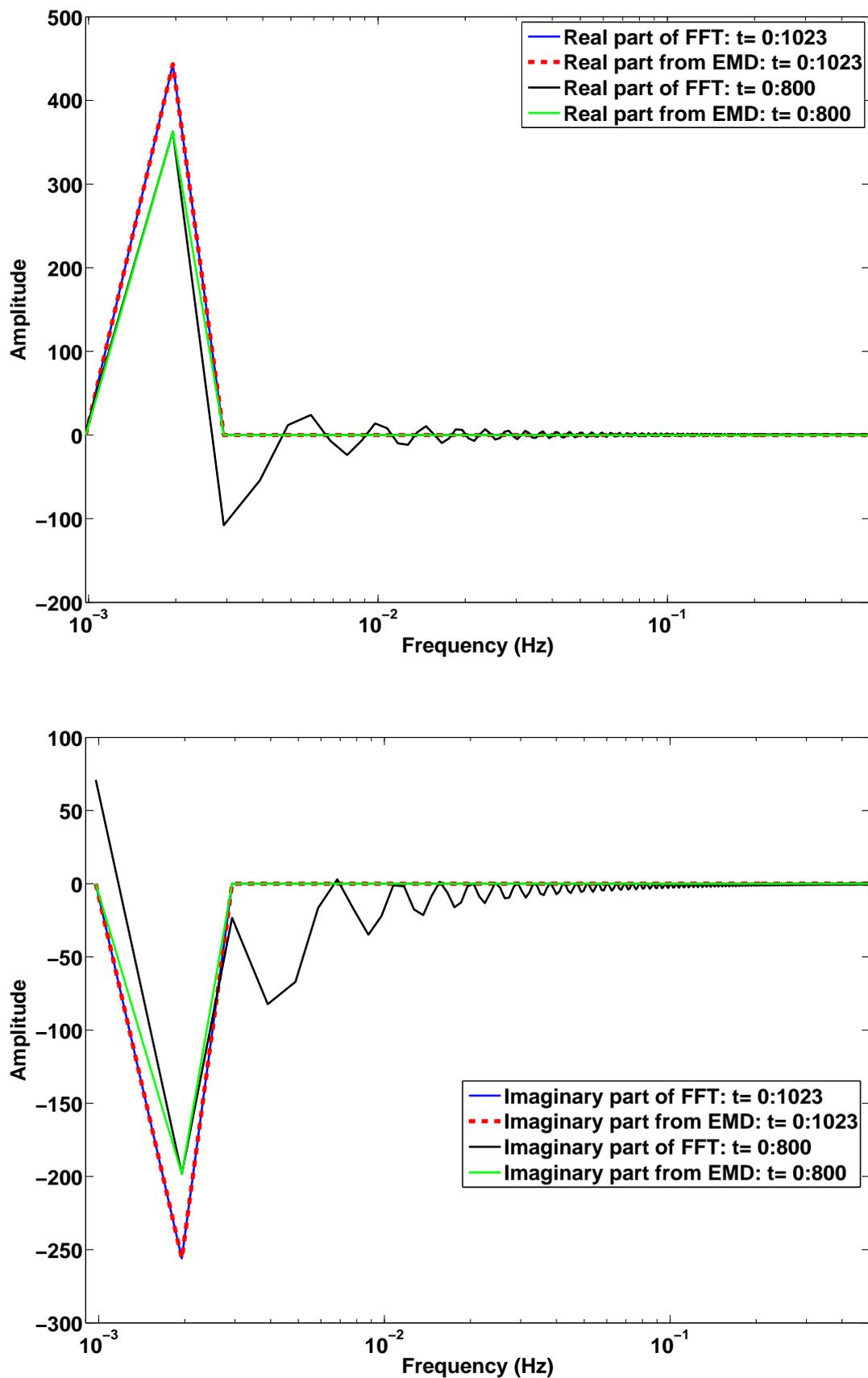


Figure 5.25: Upper panel: The real part of the complex spectra of a sinusoidal signal for time intervals of 1023 s and 800 s. Lower panel: The corresponding imaginary parts of the same complex spectra.

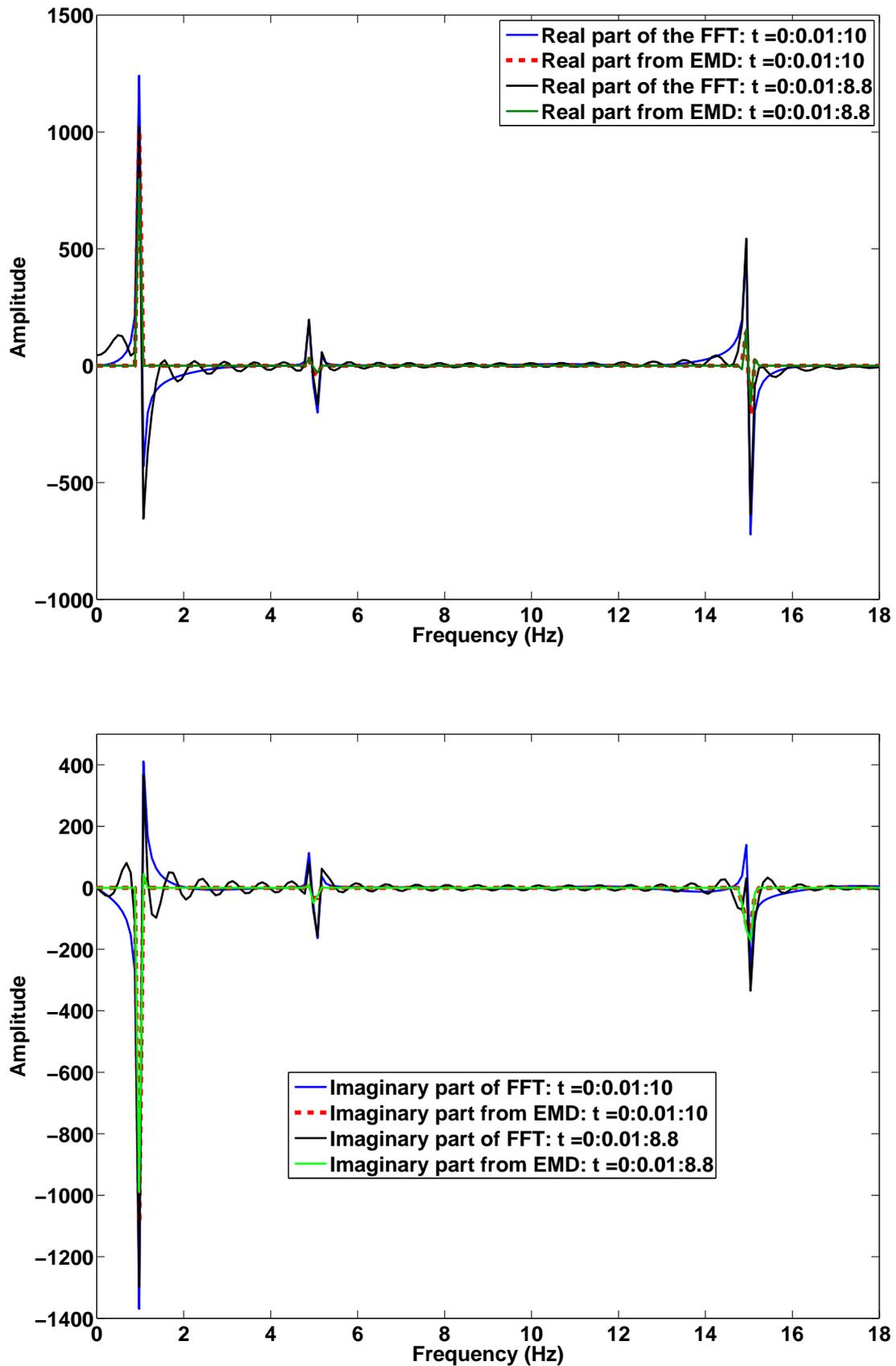


Figure 5.26: Upper panel: The real part of the complex spectra of the synthetic time series of example 1 for time intervals of 1023 s and 800 s (sampling interval is $\Delta t = 0.01$ s). Lower panel: The corresponding imaginary parts of the same complex spectra.

decomposition of time series into IMFs is non-unique. Finally, inaccuracies are introduced and some difficulties arise from the assumptions made in the sifting procedure such as the choices made for end point extension, extrema interpolation and sifting stop criterion.

In the following, synthetic time series 1 will be taken as the example to illustrate these limitations and numeric inaccuracies.

5.6.1 Effects from end point extension

Undesired effects from the beginning and the end points of time series (End effects) are unavoidable for all known data analysis methods. In Fourier analysis, the traditional way to treat end effects is to use some kind of tapering windows, which forces the amplitude values to converge towards zero at the ends. Such procedures inevitably sacrifice some precious data points near the ends, which may cause some very serious distortions when the whole time series is short.

In the EMD procedure, end effects are introduced first of all when the extrema are interpolated to determine the upper and lower envelopes of the time series. Because of the finite length of the data, serious problems due to large swings of the spline fitting can occur near the ends. These swings can eventually propagate into the data range during the sifting procedure and can especially distort the low-frequency components. The Hilbert Transform introduces, moreover, another end effect, when the HT-based analytic signal is used to calculate the IFs, since the numerical method to determine the HT is based on the Fourier transform. However, in the latter case, proper window tapering can be used to reduce these end effects. Another option to determine the IFs would be the use of the direct quadrature. In this case an iterative procedure is used (see section 5.3.3) to separate the IMFs into the AM and FM components, which is similar to the sifting procedure. Accordingly also at this stage inaccuracies are introduced at the end points by the spline fitting.

Firstly, end effects due to interpolation is considered. For example, if one uses the free boundary interpolation, where only the local maxima/minima (the first and last data points are not considered as extrema) are used to interpolate the envelopes, the extracted IMF may be seriously distorted as shown in Fig. 5.27. The first panel of Fig. 5.27 shows the first iteration to extract the IMF 2 of synthetic time series of example 1. The blue curve is the residual signal after IMF 1 is subtracted, magenta and black dashed curves are the interpolated upper and lower envelopes of residual signal, respectively, and the red curve is the mean of the two envelopes. (Here the sifting process to extract the IMF 1 is not shown, because the end effect is typically not so obvious when the extracted component is a sine or cosine wave with constant amplitude.) The lower three panels of Fig. 5.27 show the first three IMFs (blue curves) extracted by using free boundaries and the original three components of time series (red curves), respectively. One can see in the first panel that the two envelopes are not well controlled at the boundary because no more points beyond the data range are used in the interpolation. This causes very large swings of envelopes at the boundary, which propagate into the IMFs as shown in IMF 2 - IMF 3 in the third and fourth panels of the figure.

The situation can be slightly improved when the first and last points are considered as ex-

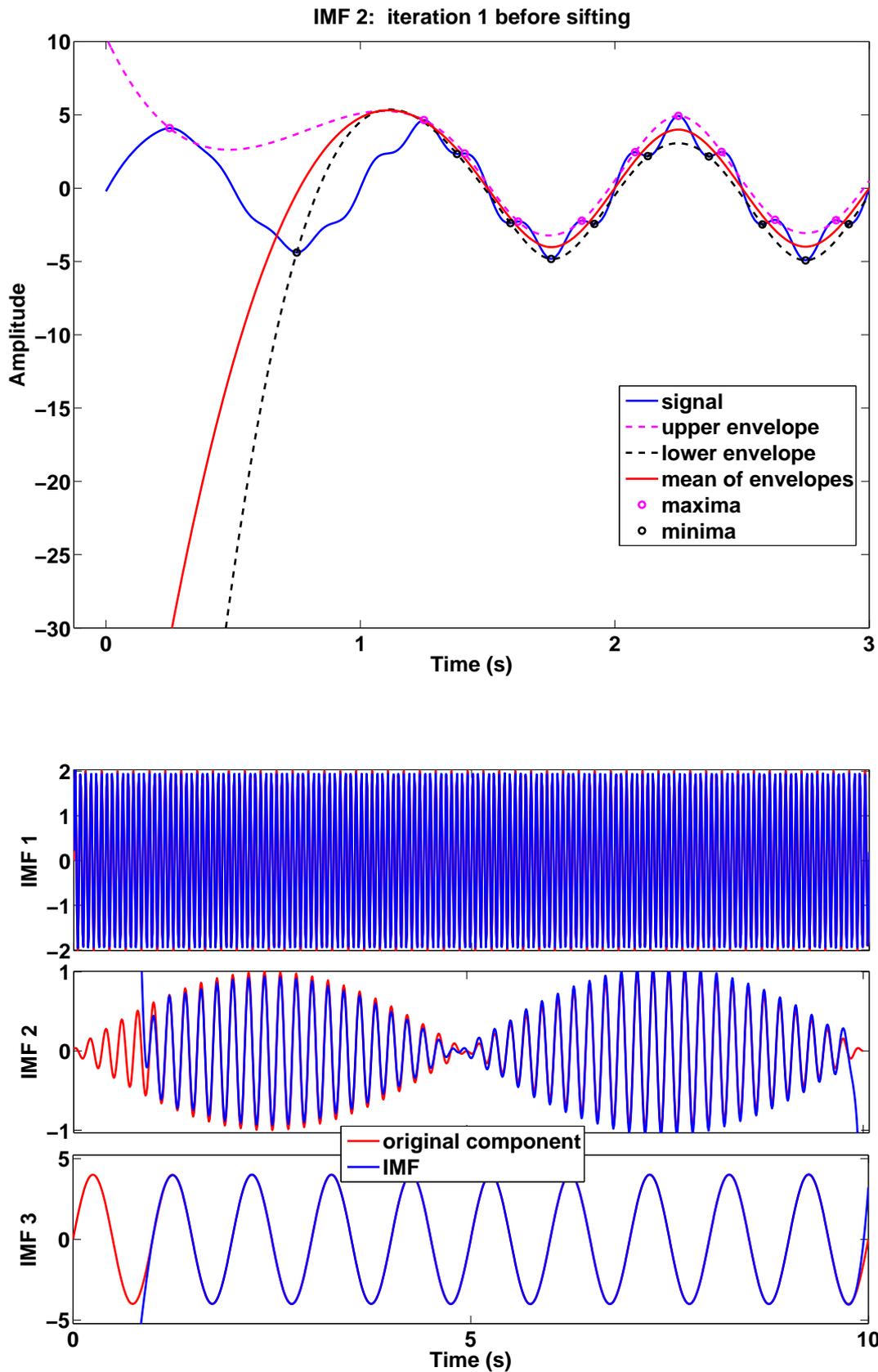


Figure 5.27: First panel: The first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto the left boundary) using free boundaries (the first and last data points are not considered as extrema). Lower three panels: The first three IMFs of the same time series extracted by using free boundary interpolations.

rema, as shown in Fig. 5.28. By comparing Figs. 5.27 and 5.28, one can see that the large swings of the envelopes at the boundary is better controlled in Fig. 5.28. The extracted first three IMFs shown in the lower three panels of Fig. 5.28 are also less effected by the end points than the ones shown in Fig. 5.27, although their artifacts are still too large that the results can be considered as satisfying.

This example indicates that interpolation of the existing extrema is not sufficient to stabilize the behavior of the two envelopes at the boundaries. Therefore another strategy is required. One option could be to add some extrema beyond the boundaries before applying the interpolation. Generally, two ways are proposed to extend data at their ends. One is to predict the data beyond their endpoints [e.g., Deng et al., 2001, Kopsinis and McLaughlin, 2008a]. However, data prediction is a difficult procedure even for linear and stationary processes and becomes hardly possible in nonlinear and non-stationary stochastic processes. Since EMD is developed for the processing of nonlinear and non-stationary data sets such a prediction is not considered as particularly suitable for this method. Another way is to add extrema through mirror symmetry with respect to the end points or with respect to the extrema which are closest to the end points [e.g. Huang et al., 1998, Rilling et al., 2003, Rato et al., 2008, Wu and Riemenschneider, 2010]. In the strategy presented by Rilling et al. [2003], the mirror axis is located at the first or last extrema and one or more adjacent extrema are mirrored. Using this strategy the end effects introduced in the sifting process are significantly weaker as for the free boundary interpolations (see Fig. 5.29). In the first panel of the figure, two adjacent extrema (maxima depicted as magenta circles and minima as black circles) are symmetrically mirrored with respect the axis (black line). Compared to the first panels in Figs. 5.27 and 5.28, the end effects in both envelopes are more efficiently controlled by using Rilling et al.'s strategy. Similar improvements can be observed in the first three IMFs (see the lower three panels of Fig. 5.29).

However, Rato et al. [2008] showed that the extrapolation of the extrema (not data) might be an even better strategy to avoid end effects than mirroring end points. They proposed a simple procedure to extrapolate one maximum point and one minimum point at the boundaries. Assuming that the first extreme point of the data is a minimum, the extrapolation procedure is then performed as follows:

- Find the locations of the first minima, t_1 and the first maxima, T_1 , having amplitudes values of m_1 and M_1 , respectively.
- The data should have a maximum adjacent to the first minimum, therefore a new maximum, $M_0 = M_1$, is inserted at the location $T_0 = -t_1$. In the same way a new minimum, $m_0 = m_1$, is added at the location $t_0 = -T_1$.
- The extrapolation at the end of the data is similar.

Based on the results from some simple time series and a complex Electrocardiography (ECG) signal, Rato et al. [2008] concluded that their strategy works quite well and that the end effects on the IMFs are almost invisible. I modify Rato et al.'s strategy by simultaneously inserting one minimum and one maximum at the locations $T_0 = -t_1$ and $t_0 = -T_1$ in order to better control the slope of the envelopes. The result for the synthetic time series of example 1 is shown in Fig. 5.30. In comparison to the other strategies (see Figs. 5.27, 5.28 and 5.29), the two envelopes (see first panel of Fig. 5.30) are much better controlled at the

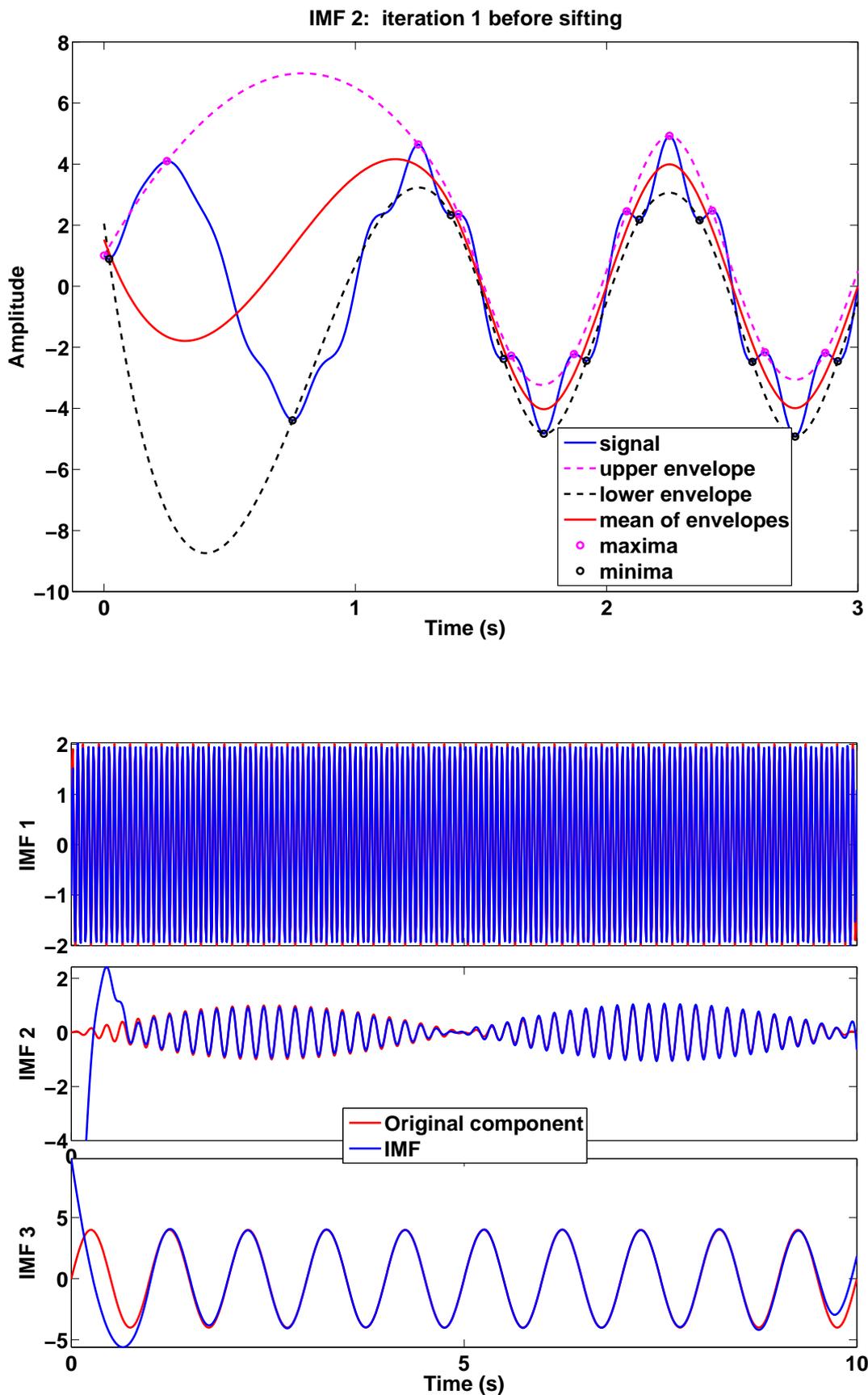


Figure 5.28: First panel: The first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto the left boundary of the signal) using free boundaries (the first and last data points are considered as extrema). Lower three panels: The first three IMF's of the same time series extracted by using free boundary interpolations.

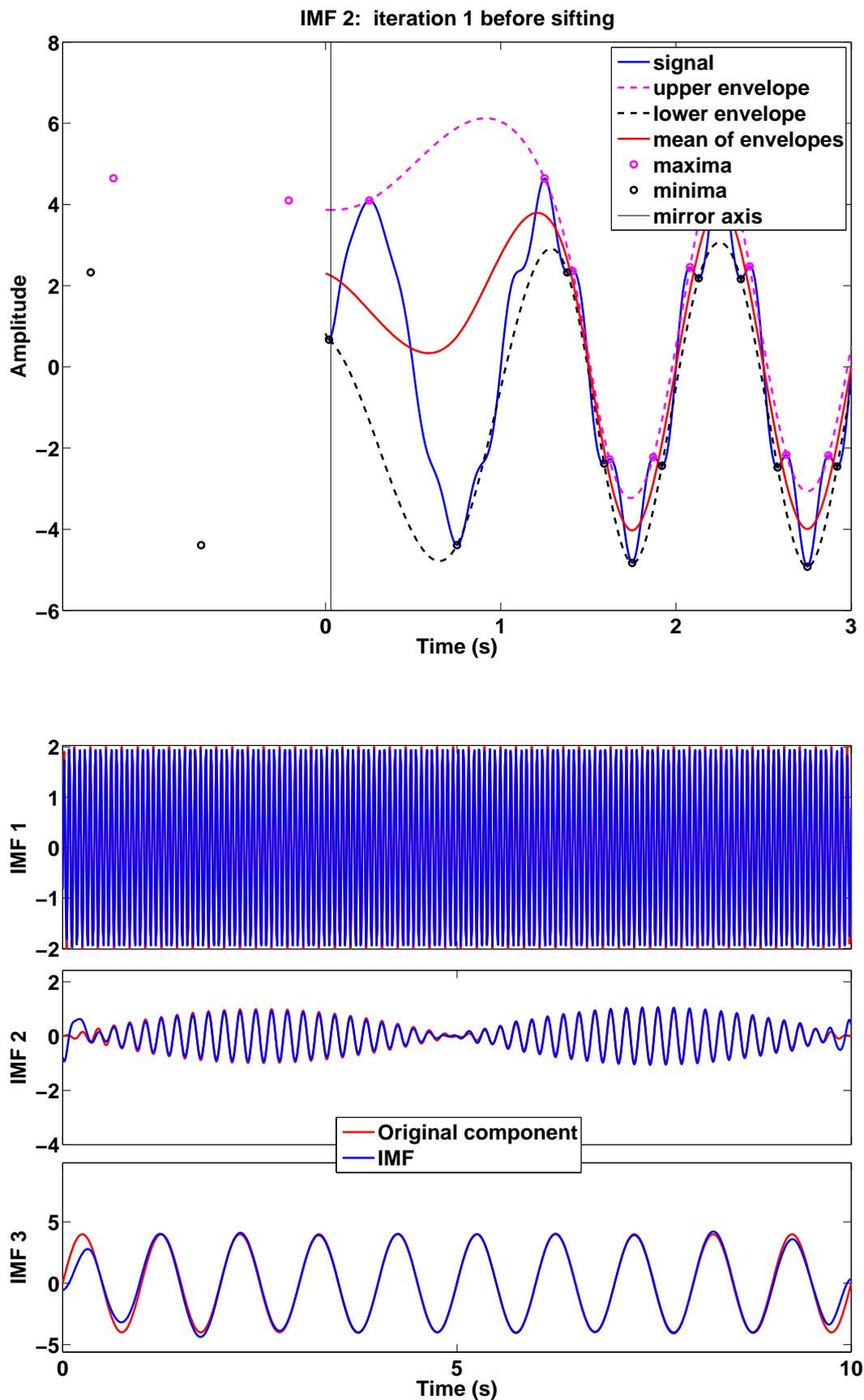


Figure 5.29: First panel: The first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto the left boundary of the time series) using the mirror strategy from Rilling et al. [2003] (the first and last data points are not considered as extrema). Lower three panels: The extracted first three IMFs of the same time series.

boundary and hence, the end effects are efficiently reduced by introducing new extrema with extrapolation (magenta and black circles). The more stable behavior at the boundaries can also be observed in the first three IMFs (see lower three panels of Fig. 5.30).

However, removal of artificial end effects in EMD method is still a not totally solved problem and further research is needed in future. Since Rato et al.'s strategy shows the best performance to reduce the end effects, this strategy is chosen in all EMD examples (see e.g. results in Figs. 5.5 and 5.9 in section 5.2 and all decompositions presented in the following chapters).

5.6.2 Effects from interpolation method

The principle of EMD is to extract variations from the data by repeatedly subtracting the local mean obtained by performing a spline interpolations of the extrema. Therefore, the spline interpolation is an essential step in the sifting process and in generation of intrinsic mode functions. However, up to now research is still conducted in assessing the effects of different interpolation methods on the final EMD results.

It is easy to show that all IMFs (except for the first one) can be considered as a summation of interpolated functions. So, one obtains from Eqs. 5.1 - 5.4 for the first IMF,

$$c_1 = x(t) - (m_{1k} + m_{1(k-1)} + \dots + m_{11} + m_1), \quad (5.54)$$

in which all the intermittent mean m_1 and $m_{1i}, i = 1, \dots, k$ are generated by interpolated functions. Therefore, after the first IMF is extracted one has the residual of the form:

$$r_1 = x(t) - c_1 = m_{1k} + m_{1(k-1)} + \dots + m_{11} + m_1, \quad (5.55)$$

which is totally determined by the combination of interpolated functions. Consequently, all the subsequent IMFs can also be expressed by combining interpolated functions.

Now the question arises, which interpolation method is best suited for EMD analysis and how one can quantify the suitability of the selected spline function. Many authors discussed the advantages and limitations of different interpolation methods [e.g. Huang et al., 1998, Rilling et al., 2003, Chen et al., 2006, Washizawa et al., 2006, Qin and Zhong, 2006, Kopsinis and McLaughlin, 2008a,b, Huang and Wu, 2008]. The choice of the interpolation method has not only an influence on the form of the obtained IMFs, but also on the convergence behavior of the EMD method. "Convergence" means in this context that the EMD will produce a finite set of IMF components. Intuition and experiences suggest that the EMD procedure should be convergent, since the signal always has a finite number of extrema and the number of extrema should decrease after each extraction of an IMF. Therefore, an adequate interpolation method does not introduce additional extrema during the sifting procedure.

Intuitively, to avoid introducing extra extrema, in one sifting procedure between two adjacent maximum and minimum (or minimum and maximum) the subtracting mean which is determined by the sum of the two interpolations should have maximal two extrema and one inflection point. However, a general and complete mathematical proof for this issue is still missing although some preliminary efforts have been made [e.g., Chen et al., 2006, Xu and Zhang, 2009]. Based on the experiments, Huang et al. [1998] and Rilling et al. [2003] stated

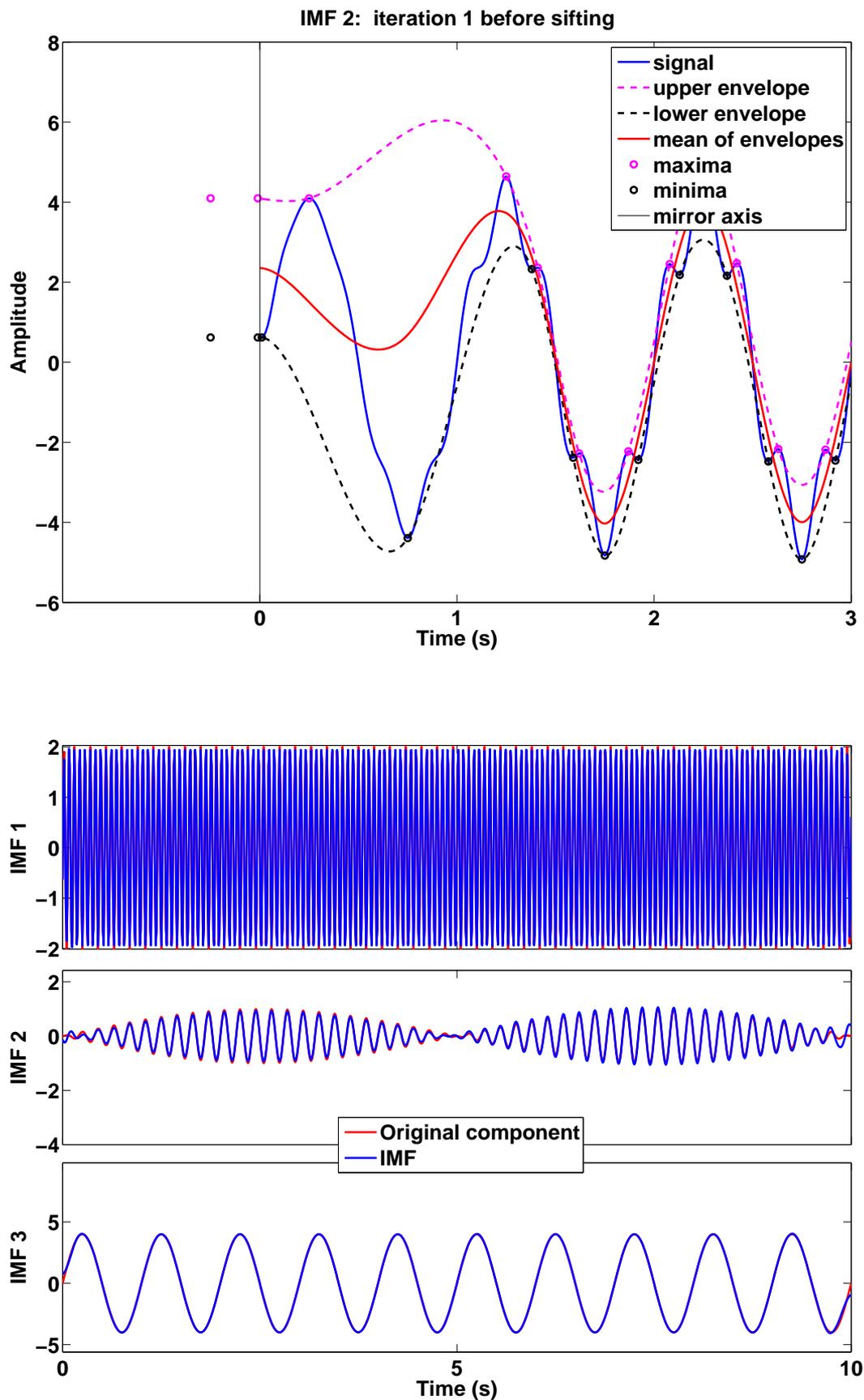


Figure 5.30: First panel: The first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto left boundary of the time series) using the strategy from Rato et al. [2008] (the first and last data points are not considered as extrema). Lower three panels: The extracted first three IMFs of the same time series.

that the "cubic spline" is an adequate choice and should be preferred.

The basic idea of cubic spline is that for given number of data points (x_i, y_i) , $i = 1, 2, \dots, n$, a piecewise function can be constructed as:

$$s(x) = \begin{cases} s_1(x) & x_1 \leq x < x_2 \\ s_2(x) & x_2 \leq x < x_3 \\ \vdots & \\ s_{n-1}(x) & x_{n-1} \leq x < x_n \end{cases}, \quad (5.56)$$

where s_i are $n - 1$ third-order polynomial functions, given by:

$$s_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i, \quad i = 1, 2, \dots, n - 1. \quad (5.57)$$

The coefficients, a_i , b_i , c_i and d_i , are determined such that the function, $s(x)$, intersects each data point, (x_i, y_i) , and the second-order derivative of $s(x)$ is continuous.

Although the cubic spline interpolation has both overshoot and undershoot problems (i.e. the envelopes typically intersect the signal at different locations), it generally works well in most cases. Fig. 5.31 depicts the first iteration of sifting process when IMF 2 is extracted from synthetic time series of example 1. The figure is zoomed onto a small time interval in order to observe regions with overshooting (A, B and C) and with undershooting (D, E and F). An alternative is to use more sophisticated high-order spline methods [Kopsinis and McLaughlin, 2008a,b] or B-spline method [Chen et al., 2006], however, increasing complexity of the interpolation results an increasing computing time.

Fig. 5.32 shows three different types of upper envelopes used for the first sifting iteration to extract IMF 2 from the data of the synthetic example 1. (The reason why the sifting process of extracting IMF 1 is not shown here is that the choice of the interpolation methods has less effect on extracting IMF 1, since the component of the time series with the highest frequency is a harmonic sine wave.) The interpolations used here are 1) a linear interpolation, 2) a piecewise cubic Hermite interpolation and 3) a cubic spline interpolation. The extracted second and third IMFs from different interpolation methods and their total residuals, $x(t) - \sum_{i=1}^3 c_i(t)$, are shown in Fig. 5.33. Because the first IMFs derived from three method are only slightly different, they are not plotted in the figure.

From Fig. 5.32 it is evident that all three interpolation methods have overshooting problems. The IMFs extracted by using piecewise cubic Hermite and cubic spline interpolations are very similar except at the boundaries. Linear interpolations yields the poorest results and the largest total residual. Moreover in this example, linear interpolation shows no convergence in the EMD analysis. Therefore in this thesis the cubic spline is always used as the interpolation method in EMD.

Independent of the chosen interpolation method the quality of the envelope interpolation is also influenced by the sampling rate [e.g. Rilling et al., 2003, Rilling and Flandrin, 2006, Rato et al., 2008]. In practice, due to the discrete time sampling an error in the location of an extremum generally occurs that can reach up to half the sampling interval. Therefore, some special attention has to be paid to correctly identify the extrema. Rilling et al. [2003]

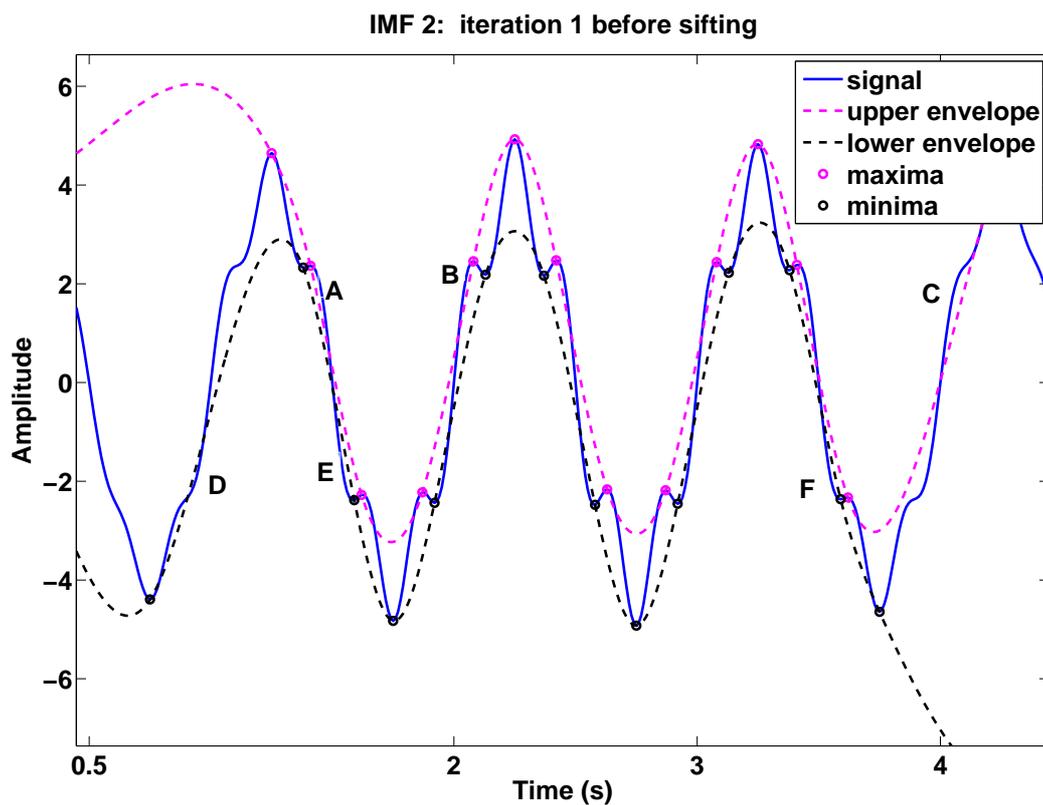


Figure 5.31: Cubic spline interpolation of the upper & lower envelopes at the first iteration to extract the IMF 2 from synthetic time series of example 1 (zoomed onto a small time interval of the time series). Some of the intervals where the upper envelope (lower envelope) have smaller (larger) values as the actual signal are numbered A, B and C (D, E and F).

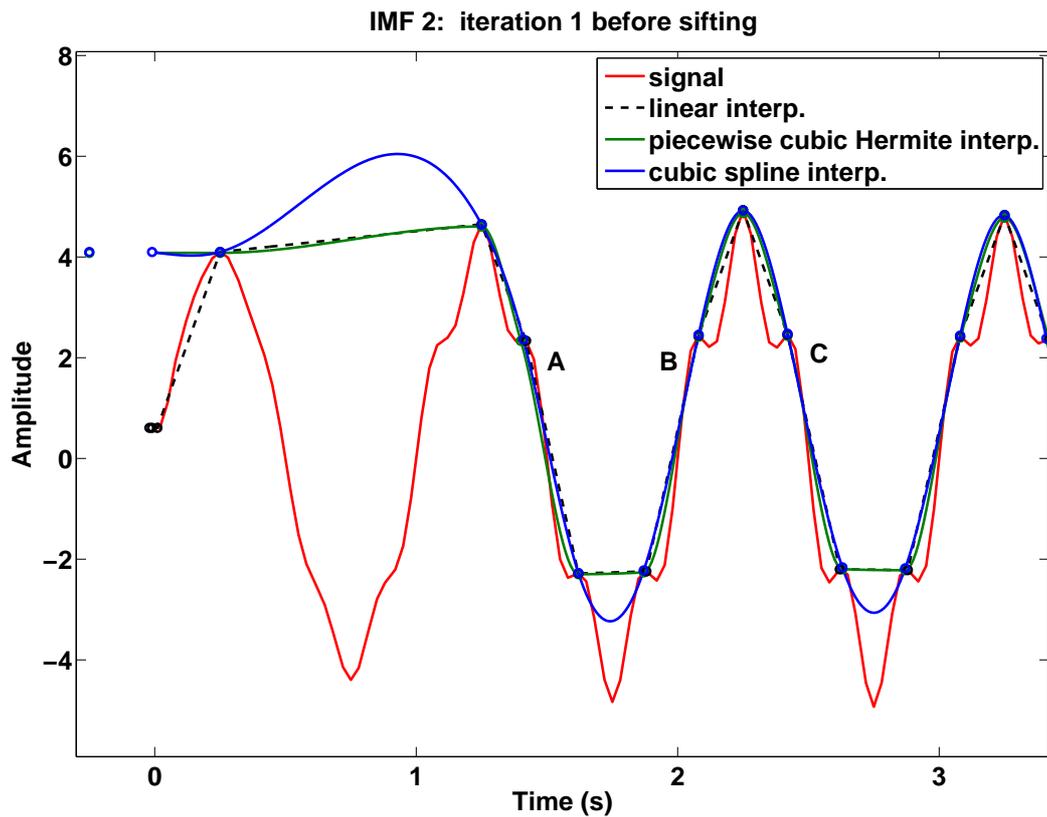


Figure 5.32: The upper envelopes are obtained by using three different interpolation methods (linear, piecewise cubic Hermite and cubic spline) at the first iteration to extract the IMF 2 of synthetic time series of example 1 (zoomed onto a small time interval of the time series). Note that all three methods have overshooting problems (see e.g. points A, B and C).

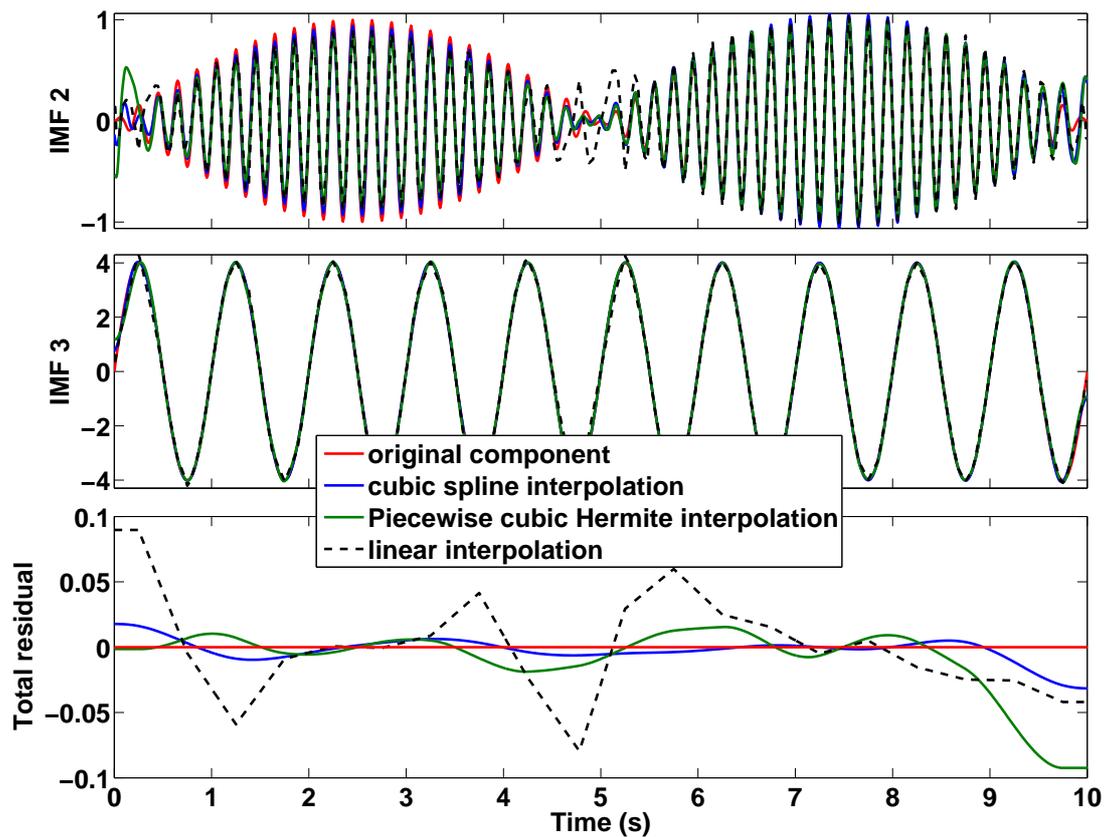


Figure 5.33: The second IMF (upper panel) and the third IMFs (middle panel) extracted by using three different interpolation methods (see Fig. 5.32 of the synthetic time series example 1) and their total residuals $x(t) - \sum_{i=1}^3 c_i(t)$ (last panel).

proposed the use of a fair amount of over-sampling, however, this approach suffers from an increase in computing time particularly for large data sets. Rato et al. [2008] proposed a parabolic strategy to calculate more accurate positions and values of the extrema. In this scheme the sample with the local extremum and its left and right neighboring data points are used to estimate a parabolic function. The extremum of the parabola is then taken as the new local extremum of the signal. The cubic spline functions that interpolate new extrema are re-sampled onto the original sampling points to obtain the upper and lower envelopes. The procedure of the scheme is illustrated in Fig. 5.34. The larger blue dot indicates the local maximum from the sampled data, the red curve indicates the parabolic curve estimated by three adjacent data points and blue star indicates the new local maximum. Note that a slight time shift of new maximum can be observed in Fig. 5.34.

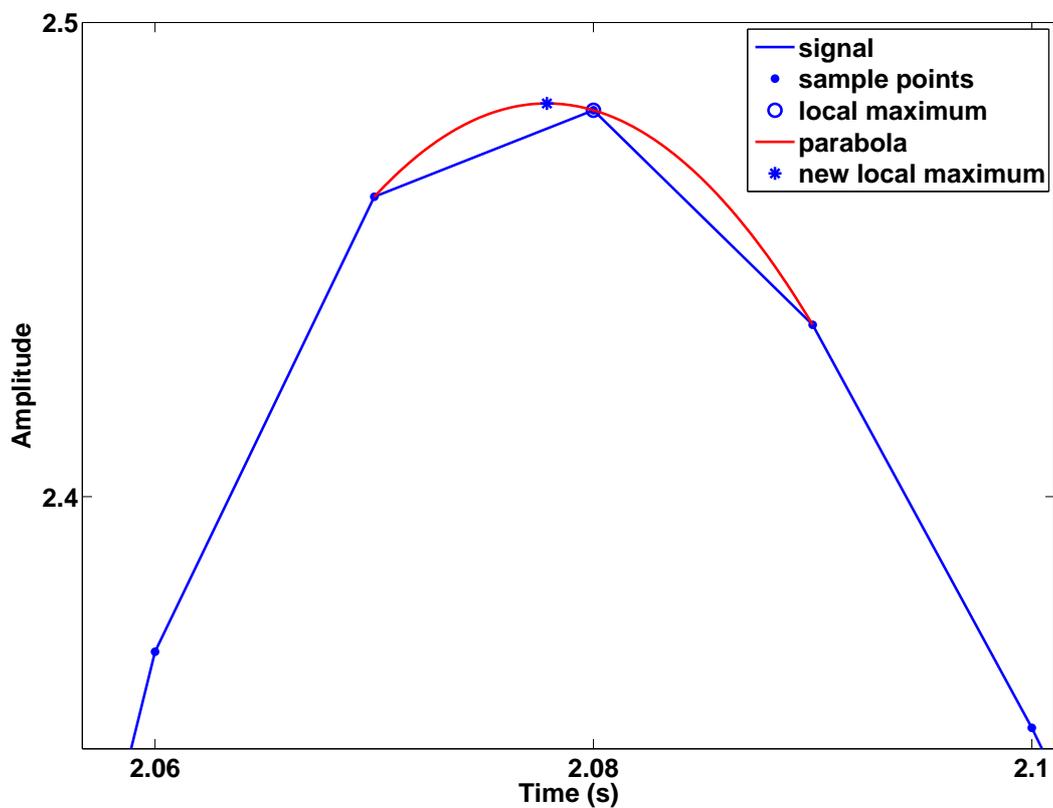


Figure 5.34: Sketch illustrating the parabolic strategy proposed by Rato et al. [2008] to modify the position and value of an extremum for sampled data.

The difference between these new envelopes (designated as Rato et al.'s envelope) and the envelopes interpolated by using the extrema of sampled data points (designated as Huang et al.'s envelope) is shown in Fig. 5.35 for different sampling rates. Fig. 5.35 shows the first sifting iteration to extract IMF 1 from the synthetic time series of example 1. The time series is originally sampled with a frequency of 100 Hz. For this relative dense sampling the two kinds of envelopes are only slightly different (see the upper panel of Fig. 5.35). However, when the sampling frequency is changed to a coarser sampling of 50 Hz, the differences

become significantly larger (see the lower panel of the Fig. 5.35).

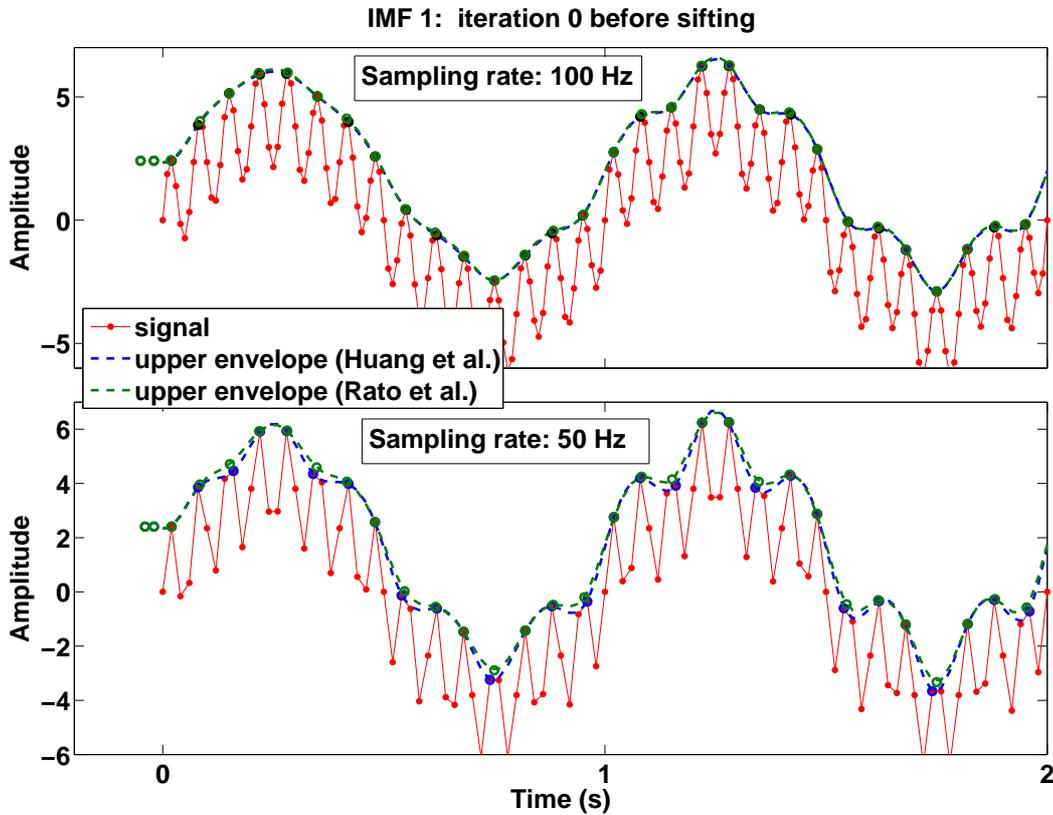


Figure 5.35: Comparison of two upper envelopes in the first sifting iteration for the synthetic example 1 using different sampling frequencies of 100 Hz (upper panel) and 50 Hz (lower panel). Blue curves are interpolated by using the maxima at the sample locations, green curves are interpolated by using the new maxima determined by Rato et al. [2008]’s parabolic strategy.

Due to the slight difference of the determined extrema by the two algorithms, after several sifting iterations for each IMF, the differences between the two types of envelopes become quite significant for this example no matter which sampling frequency is chosen. In addition, the total residuals, $x(t) - \sum_{i=1}^3 c_i(t)$, describing the difference between the time series and the sum of the three IMF 1 - IMF 3 extracted by EMD, are quite different (see Fig. 5.36). Because all three components of the synthetic example 1 are zero mean signals, one would expect that they are completely recovered by the IMFs and that this residual is everywhere equal to zero when the EMD method would work perfectly. However, in practise, numerical problems and limitations of the method cause that the residual generally alternates around zero and that the amplitudes of this residual correspond to the numerical inaccuracies of the EMD method. The total residuals obtained for the two different algorithms are shown for 100 Hz and 50 Hz sampling rates in the upper and middle panel of Fig. 5.36, respectively. It can be observed that the positions of the extrema of the total residuals are partly different for the two algorithms. This indicates that both algorithms introduce different numerical inaccuracies in the decomposition procedure.

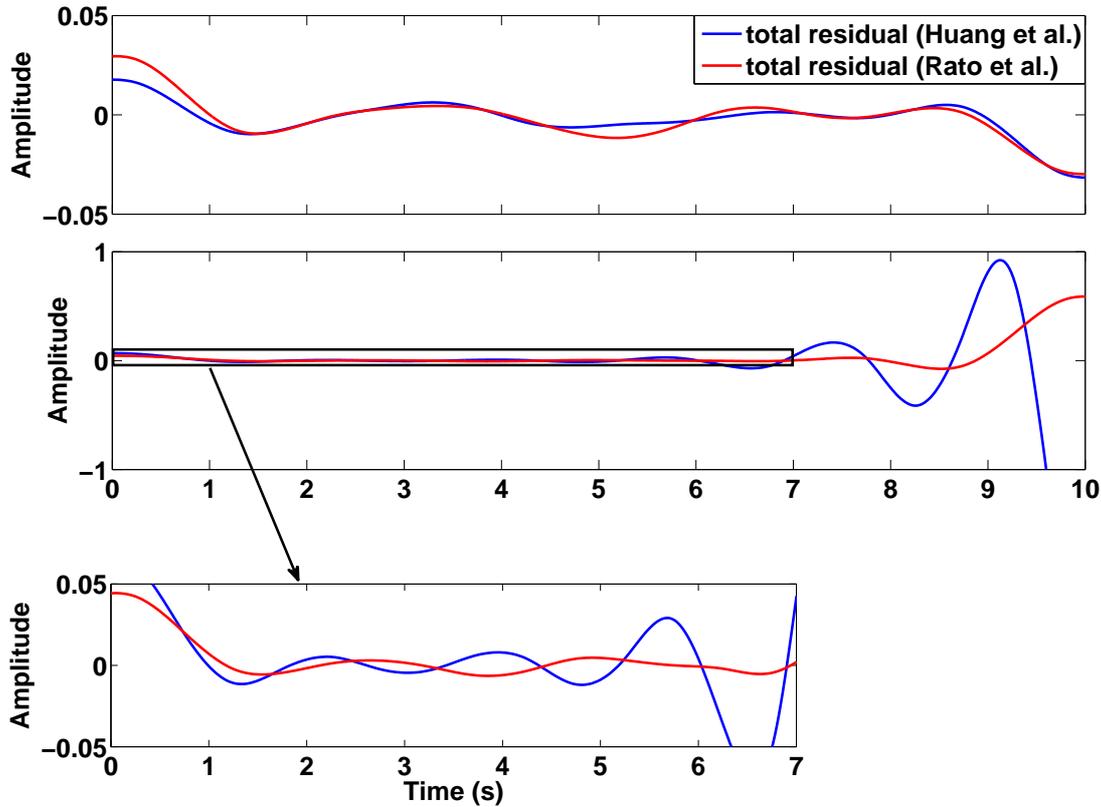


Figure 5.36: The total residuals, $x(t) - \sum_{i=1}^3 c_i(t)$, of the synthetic time series of examples 1 obtained by using Rato et al.'s envelopes (red dashed lines) and Huang et al.'s envelopes (blue lines) in the shifting procedure. The used sampling frequencies in the upper and middle panel are 100 Hz and 50 Hz, respectively. In the lower panel an magnified interval of the middle panel is shown.

However, for example 1 only very minor differences between the corresponding first three IMFs are derived from these two algorithms and the associated total residuals have the same amplitude range (see Fig. 5.36). Therefore, it is hard to conclude which algorithm gives better results. However, it will be seen later in section 6.3 that the numerical inaccuracies introduced by the two algorithms are indeed different and uncorrelated and that therefore a combination of both schemes can be used to minimize systematic errors in impedance estimates caused by numerical noise in EMD.

5.6.3 Effects from stop criterion

The purpose of the sifting procedure in EMD is to obtain signals that satisfy the conditions of IMFs by eliminating riding waves and smoothing uneven amplitudes. In section 5.2 the two conditions that define an IMF have been introduced: a) the number of extrema and the number of zero-crossings must differ at most by 1, and b) the mean between the upper and lower envelopes must be zero. Numerically, the first condition is easy to implement, but for

the second condition some kind of criterion is required that defines when the sifting process should be stopped.

A very weak criterion will generate components that deviate very much from ideal IMFs, but a very rigorous criterion will generate more or less IMFs with uniform amplitudes [Huang et al., 1998, Junsheng et al., 2006, Wang et al., 2010], which would destroy the intrinsic amplitude variations and make the results physically less meaningful. To preserve the natural amplitude variations of the oscillations, sifting should be limited to as few steps as possible. Therefore, the criteria in the sifting process should be chosen cautiously. Originally, Huang et al. [1998] proposed a Cauchy type criterion by limiting the size of the standard deviation which is computed from the two consecutive sifting results as

$$SD = \sum_{t=0}^T \frac{|h_{1(k-1)}(t) - h_{1k}(t)|^2}{h_{1(k-1)}^2(t)}. \quad (5.58)$$

A typical value for SD varies between 0.2 and 0.3 [Huang et al., 1998]. This criterion, however, suffers from the fact that it does not ensure that the results are indeed IMFs, since satisfying the criterion has nothing to do with neither the numbers of extrema and zero-crossings nor the symmetry of the envelopes.

As an improvement Rilling et al. [2003] introduced a new mean value criterion, which is now commonly used. It is based on two threshold values, θ_1 and θ_2 , which guarantee global small fluctuations in the mean while allowing locally large excursions. Let us denote n_{zero} as the number of zero-crossings, n_{ext} as the number of extrema, and e_{up} , e_{low} as the upper and lower envelope values, respectively. The amplitude of the mean m_{amp} and the amplitude of the envelope e_{amp} are given as:

$$m_{amp} = (e_{up} + e_{low})/2, \quad e_{amp} = (e_{up} - e_{low})/2. \quad (5.59)$$

This stop criterion compares now the amplitude of the mean with the amplitude of the envelope. When a) the numbers of zero-crossing and extrema differ with less than one, and b) the ratio, $|m_{amp}/e_{amp}| < \theta_1$, for some pre-defined fraction $(1 - \alpha)$ of the total data duration, c) and $|m_{amp}/e_{amp}| < \theta_2$ for the remaining fraction, the sifting process is stopped. The stop criterion can mathematically be written as:

$$\begin{aligned} |m_{amp}/e_{amp}| < \theta_2 & \quad \text{for all data} \\ \text{and} \\ \frac{\text{the number of data with } |m_{amp}/e_{amp}| > \theta_1}{\text{the data length}} < \alpha & \quad (5.60) \end{aligned}$$

and

$$|n_{zero} - n_{ext}| \leq 1$$

Referring to [Rilling et al., 2003] a typically settings are $\theta_1 \approx 0.05$, $\theta_2 \approx 10 \theta_1$ and $\alpha \approx 0.05$, which means that for 5% data points only have to fulfill the weaker threshold criterion associated with θ_2 , but not the strict criterion associated with θ_1 .

Rilling et al.'s criterion is certainly better than the Cauchy type one, because it is related to the definition of IMF by forcing that the mean of envelopes to be small and that the envelopes are symmetric. It is therefore used as the stopping criterion in this thesis.

The number of sifting iterations needed to extract one IMF is generally higher for more stringent stop criteria. My experiences is that only a few shifting iterations are needed to extract an IMF that represents a harmonic components of a time series; even if the stop criterion is stringent. However, to extract IMFs that represents non-harmonic components of the time series, more sifting iterations are needed and the number of iterations increases significantly for more stringent stop criterion.

5.7 Short summary

In this chapter, a relatively new EMD method is introduced. By means of some examples it was shown that EMD has the ability to adaptively decompose a time series into several intrinsic mode functions (IMFs) which are zero-mean AM-FM mono-component signals. From IMFs physically meaningful instantaneous frequencies and amplitudes can be calculated. The instantaneous spectra of the IMFs faithfully characterize the local properties of the time series. Compared to Fourier analysis, the marginal power spectrum and complex spectrum have more freedom in choosing time and frequency resolution and are less effected by the periodicity and length of the data. Although the EMD method is an empirical method without strong mathematical background and has some limitations in its implementation, it has been found useful to deal with non-stationary time series in many engineering and research areas. Since geomagnetic time series recorded for MT investigations are usually characterized by non-stationarity, EMD method appears to be a very promising technique to handle the non-stationarity and analyze the noise effects in these types of time series.

In the next chapter, I will develop an approach to use this new method instead of the traditional Fourier analysis to process the MT data and test the approach on synthetic MT time series.

Chapter 6

Using EMD to Process Synthetic Data

As a first test, the EMD method is applied to process synthetic MT data with a known impedance Z derived from a constructed 1-D Earth model. In the first section of this chapter, the construction of a synthetic model and time series are described. Impedance and error estimates derived by using EMD method and Berdichevsky's formulae Eqs. 4.59 - 4.61 are presented in section 2. However, a systematic downward bias in the apparent resistivity estimation is observed. In section 3 the bias problem is solved by analyzing the bias and by introducing "numeric" reference time series. In the last section, the estimation procedure is tested by adding several levels and types of noise and the results are compared with those derived by using simple Fourier analysis and robust processing method BIRRP [Chave and Thomson, 2004].

6.1 1-D Earth model and synthetic time series

As discussed in chapter 4, the estimation of the Earth transfer functions in MT prospecting method is an elaborate task. Therefore, many advanced data processing techniques have been developed. In order to assess the performance of each technique, synthetic data with different known noise types simulating the observed signals have frequently been used for testing different estimation techniques of the MT impedance tensors. Many authors contributed to the generation of synthetic MT time series for simulation purposes [e.g. Goubau et al., 1978, McMechan and Barrodale, 1985, Yee and Paulson, 1988, Egbert, 1992, Egbert and Booker, 1992, Varentsov and Sokolova, 1995, Larsen et al., 1996, Loddo et al., 2002].

To construct the synthetic MT time series, a relatively simple way according to the method of Larsen et al. [1996] is used. Measured real magnetic time series are used as synthetic B-field data. The reason for such a choice is to recover realistic spectra of the Earth's magnetic field variations. The construction of the synthetic E-field data is as follows.

The real B-field time series are firstly transferred to the frequency domain by using the Fourier transform to obtain $B_y(\omega)$ and $B_x(\omega)$.

Then the impedance $Z(\omega)$ of a simple 1-D Earth model is multiplied with $B_y(\omega)$ and $B_x(\omega)$ to determine the Fourier spectra $E_x(\omega)$ and $E_y(\omega)$ of E-field time series,

respectively.

Subsequent, the electric field spectra are transformed back into time domain to obtain synthetic E-field data $E_x(t)$ and $E_y(t)$.

The magnetic field data measured in the real world are non-synthetic and therefore may also contain both correlated and uncorrelated noise. However, the construction of electric field data guarantees that for each data sample the impedance relationship $E(\omega) = Z(\omega) \cdot B(\omega)$ holds. In this sense the data sets are noise free, because the magnetotelluric noise consist of those parts of the electric-/magnetic- field data which do not follow the impedance relationship.

It should be noted that in such a construction all time series are assumed to be stationary, although it is actually proposed to demonstrate that EMD can also handle non-stationary time series. There are two reasons for this, one is that it is hard to construct non-stationary time series for all field components using the Fourier transform, another is that it would be better to test how well EMD works through comparing the results of this new approach with analytic results. In this thesis, it is the first time that the EMD method is used for impedances estimates. At this stage it is not clear what kind of inaccuracies are introduced by the numeric algorithms and therefore I consider it as most adequate to start with totally noise-free conditions. Afterwards I will also consider non-stationary noise and apply EMD to process non-stationary data.

The synthetic B-field data which is used is a nine-day-long magnetic time series measured onshore Costa Rica in 2007 by GEOMAR (see chapter 7). The time series is measured at a sampling frequency of 1 Hz, its x- and y-components are shown in Fig. 6.1. The frequency range of the data is from about 10^{-6} Hz to 0.5 Hz, as shown by their Fourier power spectra presented in Fig. 6.2. The power spectra are simply calculated from Fourier transform of windowed (Hanning window) and detrended time series.

The synthetic 1-D Earth model that is used consists of a three layered half space model of the Earth as shown in Fig. 6.3. The three layers have the thicknesses $h_1 = 1$ km, $h_2 = 2$ km and $h_3 = \infty$ with resistivities of $\sigma_1 = 10 \Omega m$, $\sigma_2 = 1 \Omega m$ and $\sigma_3 = 1000 \Omega m$. The impedance of the model can be analytically determined through the formulae Eqs. 3.41, 3.48 and 3.49 in chapter 3. Fig. 6.4 shows the impedance (first panel) of the model, and the apparent resistivity and phase which are calculated from the impedance (middle and lower panels). Such an Earth model with large resistivity contrasts was chosen in order to produce an apparent resistivity curve with strong slopes that give clear indication whether the estimated impedance values are correct.

The synthetic E-field time series which have been constructed through the procedure are shown in Fig. 6.5, and their real and imaginary parts of the complex spectra $E_x(\omega)$ and $E_y(\omega)$ are shown in Fig. 6.6.

Since the above construction procedure ensures that the impedance relationship holds for each data sample, classical MT processing methods based on Fourier analysis should recover the impedance Z as well as the apparent resistivity and phase. For example, one can choose

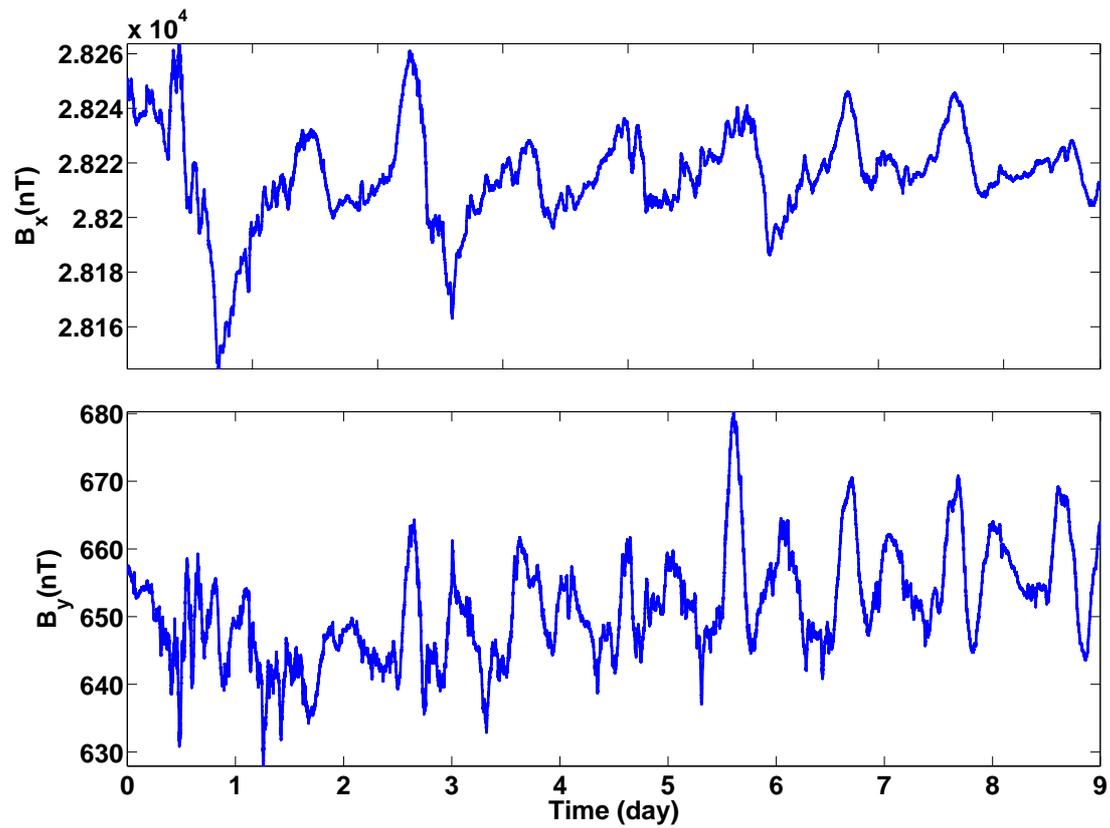


Figure 6.1: Nine day long section of B-field data with sampling frequency 1 Hz measured onshore Costa Rica in 2007. These data are used as the B-field time series in the synthetic tests.

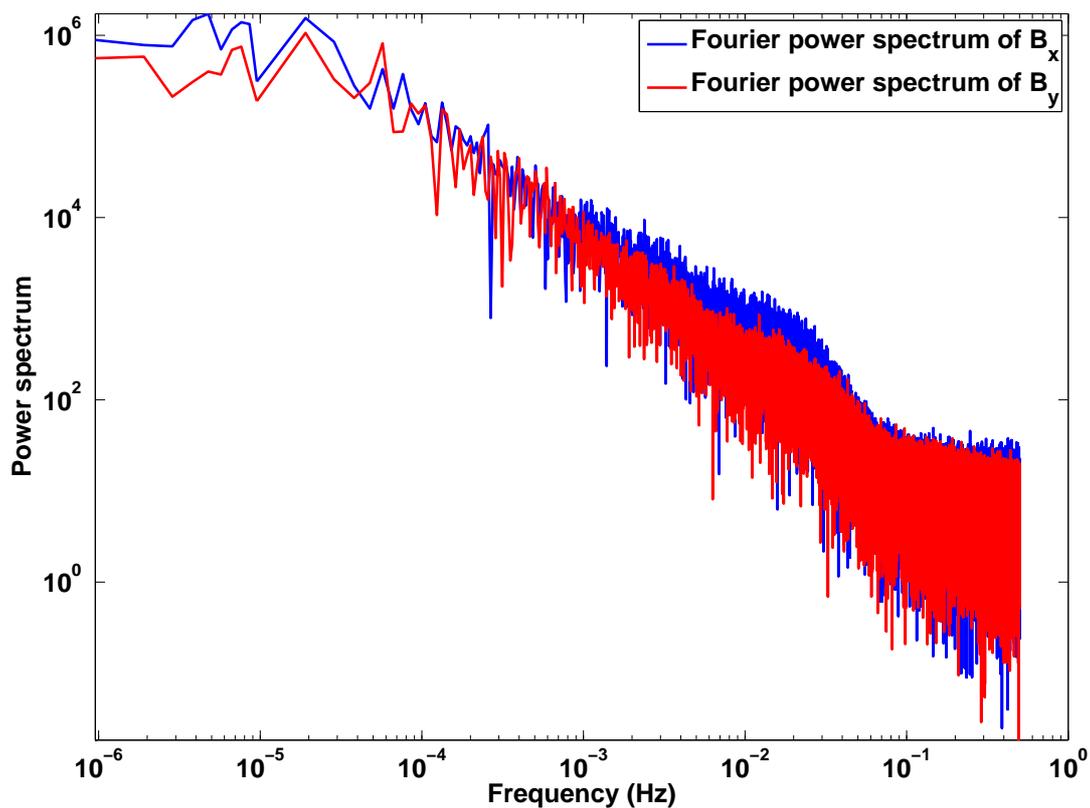


Figure 6.2: The Fourier power spectra of the nine day long section of B-field data. The spectra are calculated from simple Fourier transform of the detrended and windowed (hanning window) time series.

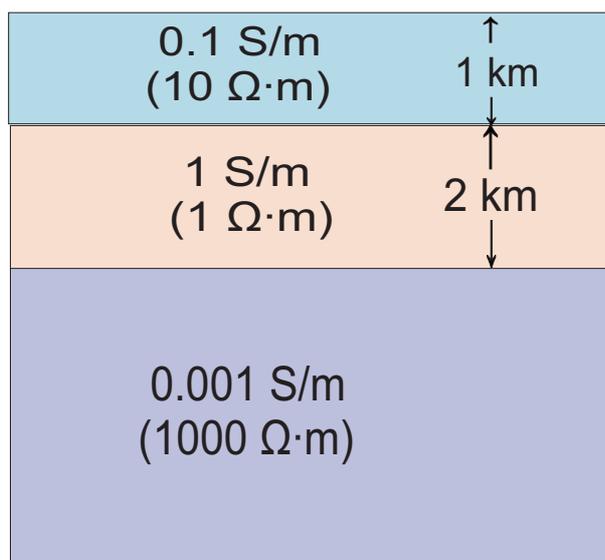


Figure 6.3: Synthetic layered Earth model to test the new EMD based impedance estimation method.

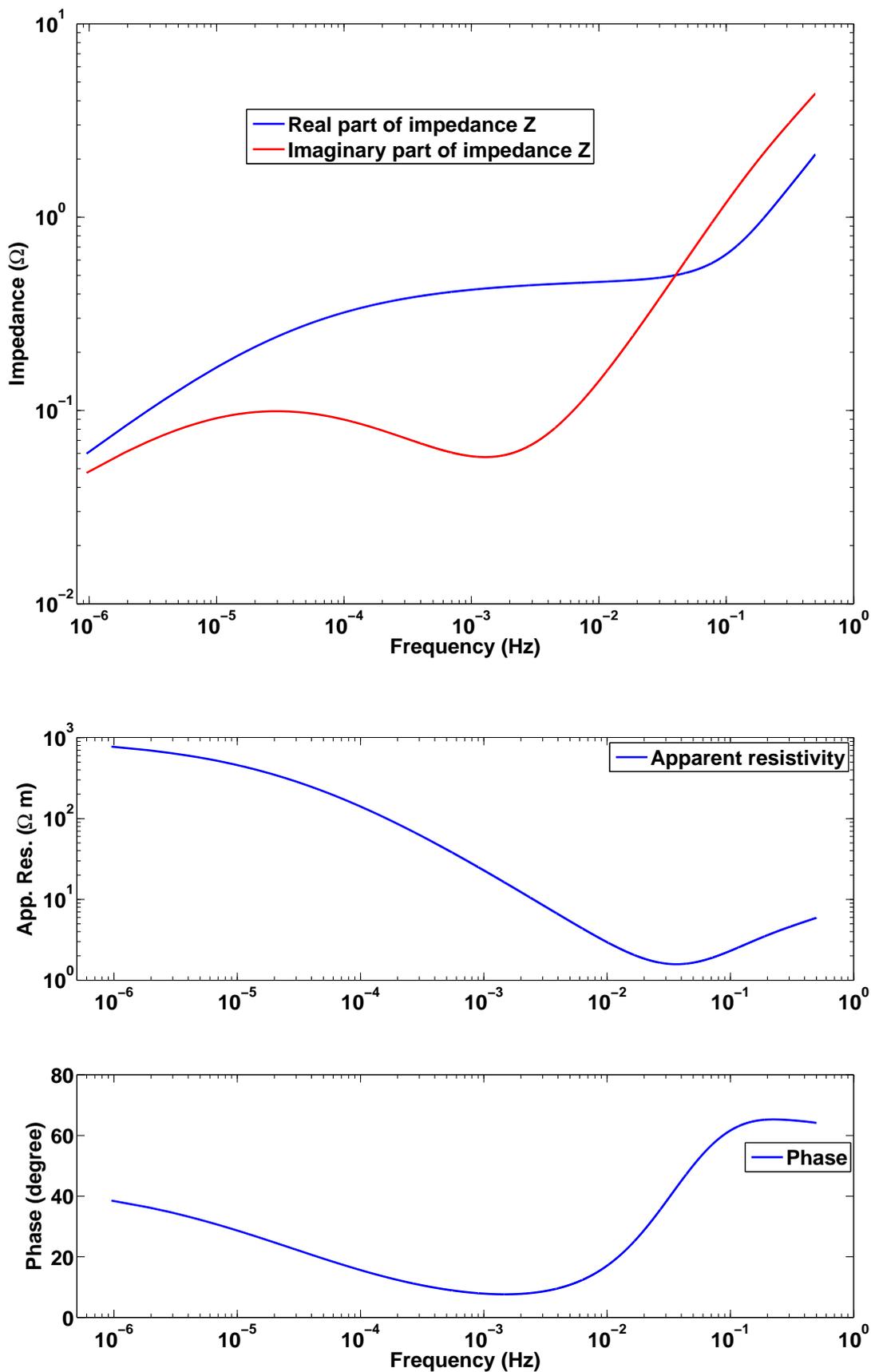


Figure 6.4: First panel: The real and imaginary parts of the impedance Z determined analytically from the synthetic layered Earth model shown in Fig. 6.3. Middle and lower panels: The corresponding apparent resistivities and phases.

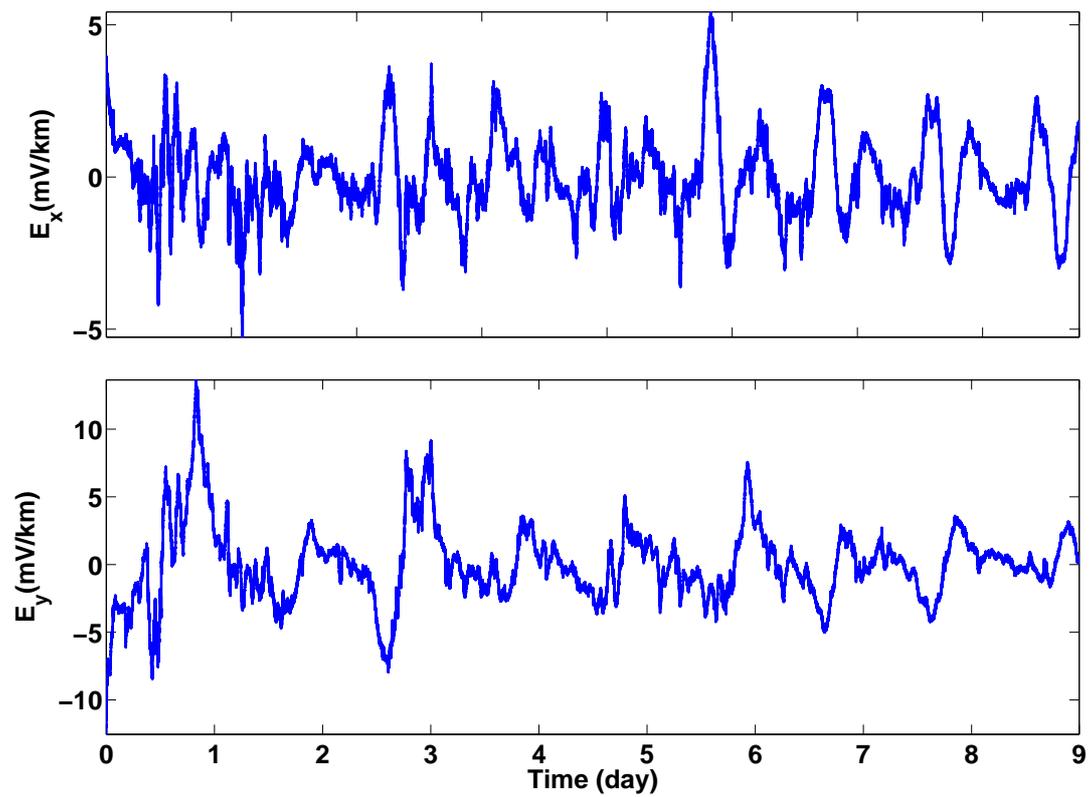


Figure 6.5: Synthetic E-field time series are constructed from the magnetic field data using the analytically determined impedance $Z(\omega)$ from the layered model.

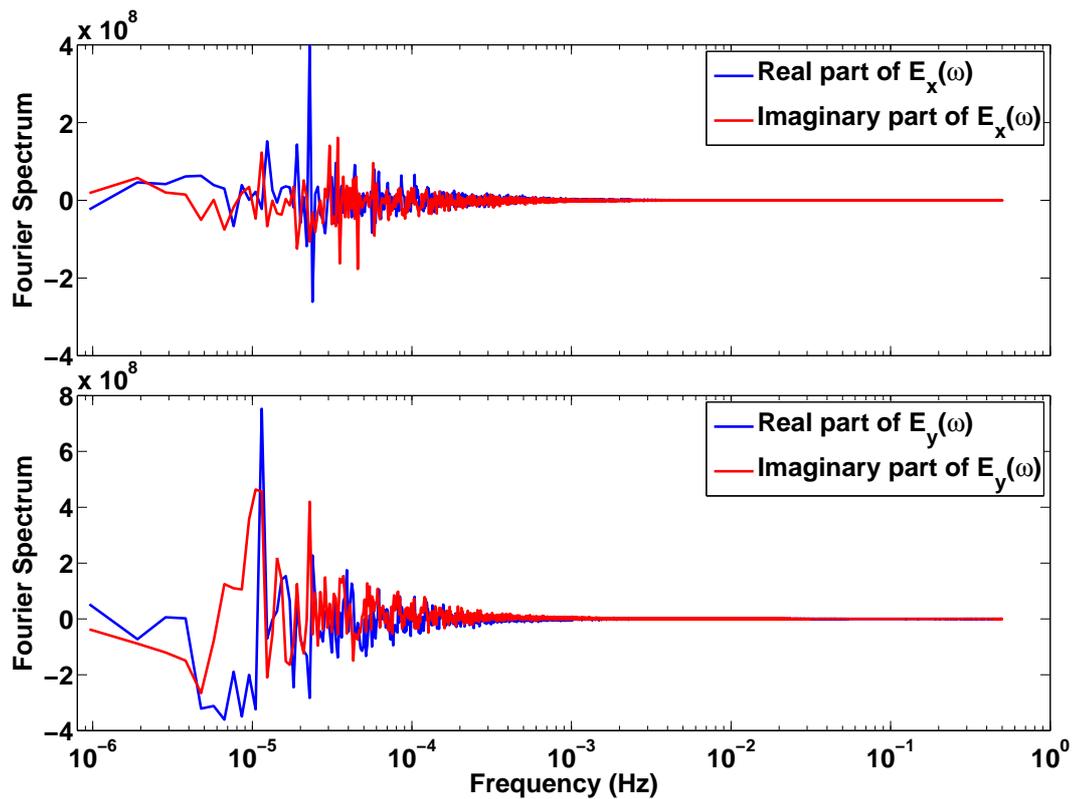


Figure 6.6: The real and imaginary parts of the $E_x(\omega)$ and $E_y(\omega)$ which are calculated by multiplication of $B_y(\omega)$ with $Z(\omega)$ and $B_x(\omega)$ with $Z(\omega)$, respectively.

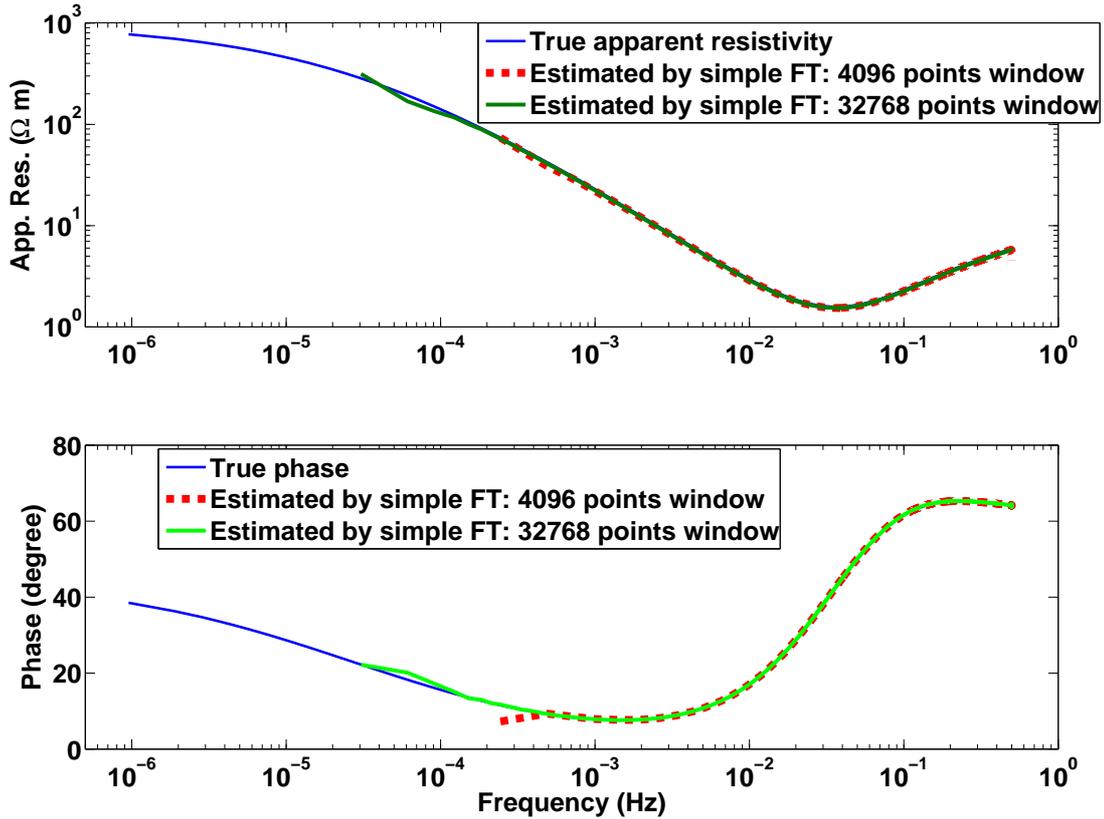


Figure 6.7: Estimated apparent resistivity and phase by using simple Fourier transform (red dashed curves and green curves). The estimated values are calculated from the average auto- and cross-power densities of small sections (4096 and 32768 data points long, respectively) by using Eq. 4.23.

two kind of window lengths as $2^{12} = 4096$ and $2^{15} = 32768$ data points, move the window without overlapping, and use simple Fourier transform to calculate the auto- and cross-power densities of E- and B-field data for each window. The obtained average power densities are then used to calculate the impedance from Eq. 4.23. Fig. 6.7 shows that the apparent resistivity and phase derived from impedance curves are identical to the true values. The lowest frequency range is determined by the window length ($1/4096 = 2.44 \cdot 10^{-4}$ Hz and $1/32768 = 3.05 \cdot 10^{-5}$ Hz); the longer the window length, the lower the frequency range.

6.2 Impedance and the error estimates

Most of the traditional MT data processing methods discussed in chapter 4 use Fourier analysis to transfer the time series into frequency domain. The reason for that is that the basic impedance estimate formula Eq. 4.23 and its other forms (Eq. 4.26 in remote reference estimation and Eqs. 4.34 and 4.31 in robust estimation) are given in frequency domain. However, the EMD method gives a time varying instantaneous frequency and amplitude representation of the time series. The purely frequency based formulae are therefore not ap-

plicable for the estimation of impedances based on EMD. Hence I choose the time-frequency domain estimator (Eqs. 4.59 - 4.61) proposed by Berdichevsky et al. [1973] to estimate the impedance tensor.

As mentioned in chapter 4, the time-frequency impedance estimator in Eqs. 4.59 - 4.61 is actually equivalent to the standard least square estimator of Eq. 4.23. Through the time-frequency estimator, the impedance at a certain frequency can be estimated by using a transient complex signal corresponding to a frequency band which is centered at this particular target frequency. These transient complex signals of electromagnetic time series can be obtained by using different methods. Berdichevsky et al. [1973] proposed Fourier-based narrow-band filters. The transient complex signals consist of the real parts of the band-limited amplitude modulated signals (which are filtered signals of electromagnetic time series around the target frequency) and their Hilbert transform constitutes the imaginary parts.

EMD gives another way to obtain such transient complex signals and avoids using Fourier-based filters. After instantaneous frequencies of the IMFs of the electromagnetic time series are calculated, the real parts of the transient complex signals for different frequency bands can be obtained by summing up all the IMFs which have IFs in the particular frequency-band at each point in time (see Eq. 5.52). The imaginary parts can come from either the Hilbert transform or direct quadrature of these IMFs (see section 5.3.3). However, since the IMFs are amplitude modulated and frequency modulated time series, which, as has been discussed in section 5.3.3, their Hilbert transform can become inaccurate [Bedrosian, 1963, Nuttall and Bedrosian, 1966]. Therefore, the direct quadrature (DQ) is chosen as the imaginary part of the complex signal. The direct quadrature guarantees the $\pi/2$ phase shift from the original signal.

The procedure of using EMD method to process the MT data is as follows.

Prior to EMD decomposition, both E and B field time series are cut into 18 short sections with a length of about 12 hours each ($N \approx 43,000$ samples). The choice of such a window length is based on two reasons: first the length is comparable to the length of relatively quiet sections measured at a station located at the sea-floor in Costa Rica, which will be examined in the next chapter. Second it saves computation time of the decomposition procedure.

The next step is to decompose all the short time series sections into IMFs by using EMD. As an example, one short section of x-component B-field time series and its 13 IMFs are shown in Fig. 6.8.

After decomposing each section into IMFs, instantaneous amplitudes & frequencies are calculated and the corresponding complex time series for each IMF are constructed based on the direct quadrature method. The instantaneous spectra of one section of x-component magnetic field data are shown in Fig. 6.9.

Finally, the time-frequency domain estimator (Eqs. 4.59 - 4.61) is used to estimate the impedance tensor.

One practical problem in using the Berdichevsky's formulae is that there are rarely any frequencies ω_0 for which frequency information exists for all time samples and in all different

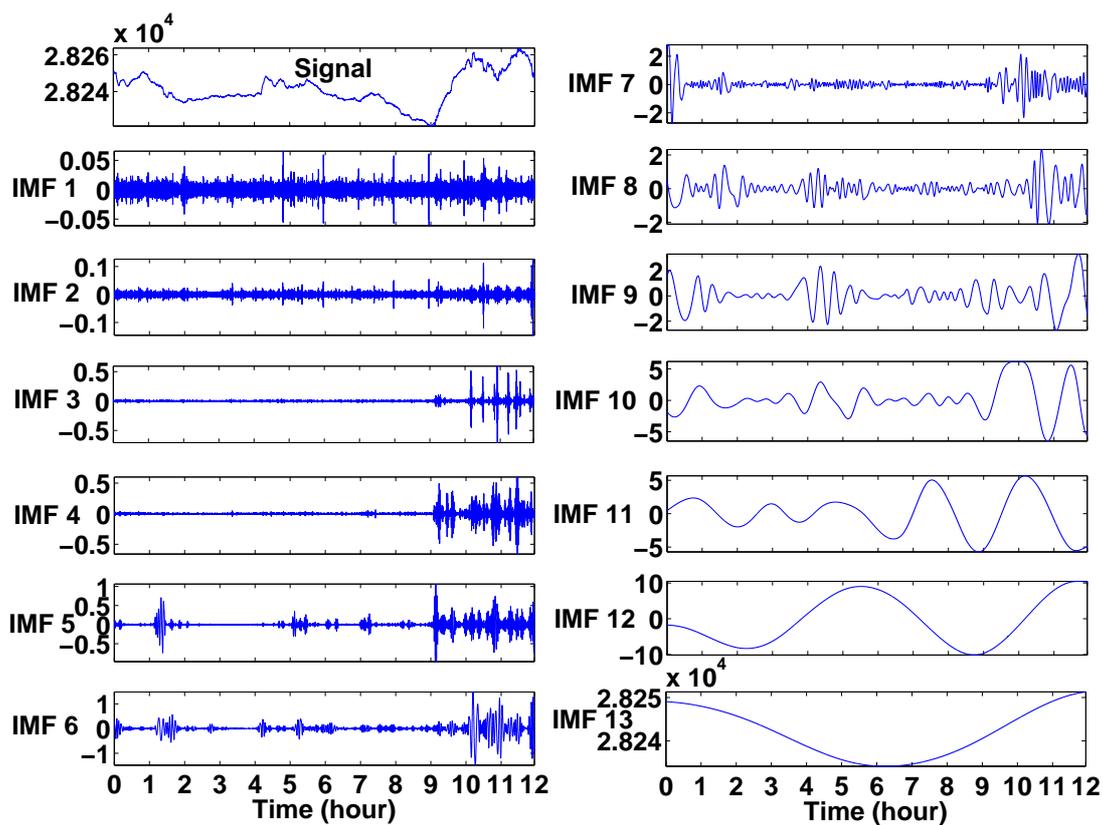


Figure 6.8: One short section x-component magnetic time series (left upper panel) and its 13 IMFs obtained by using EMD method.

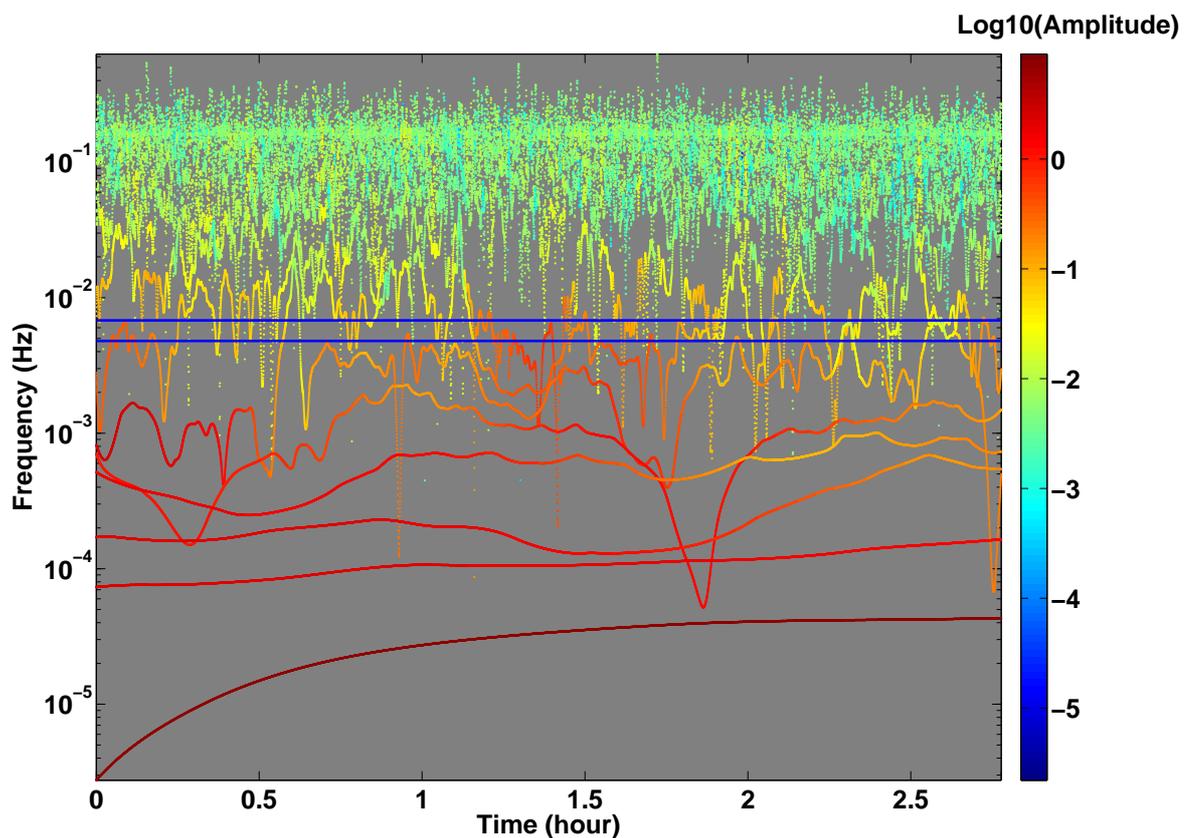


Figure 6.9: The instantaneous spectra of one short section x-component magnetic time series. The color bar indicates the logarithm of the instantaneous amplitudes. The instantaneous spectrum of the last IMF, which is the trend of the time series and has much high amplitude than other IMFs, is not included. For convenience reason, the figure only shows 10000 seconds (≈ 2.78 hours) length time interval. One frequency window centered at 5.6×10^{-3} Hz (blue box) is highlighted and will be further explained in Fig. 6.10.

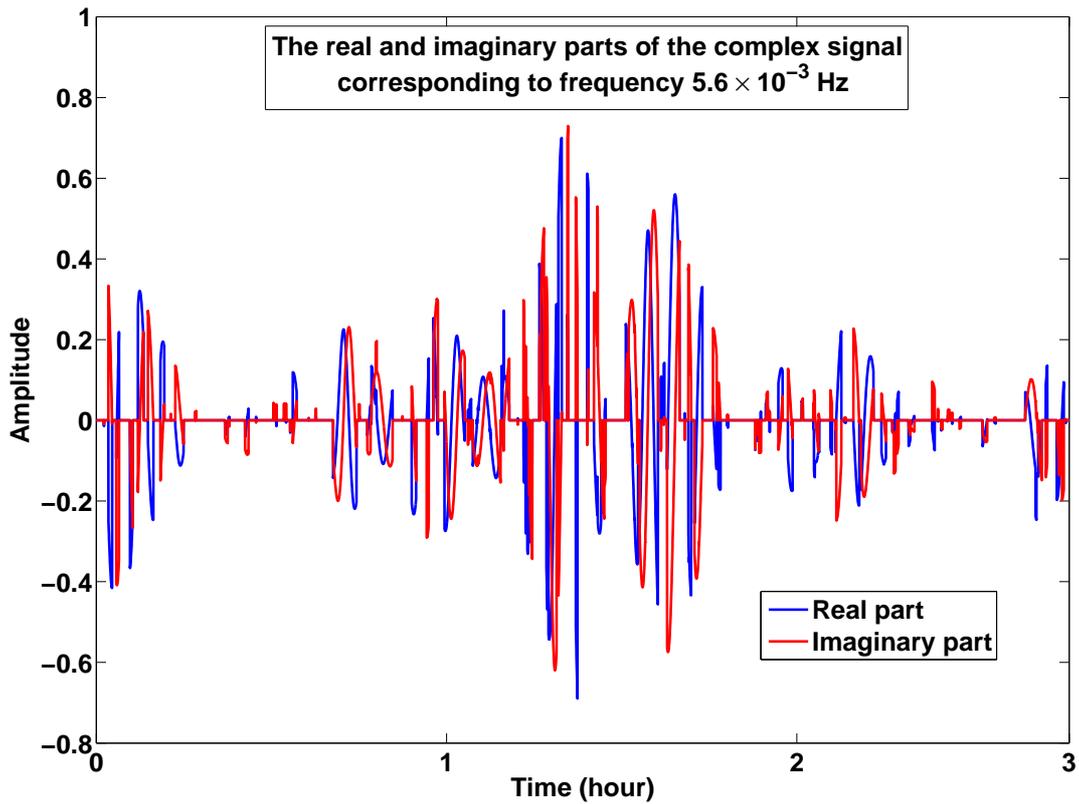


Figure 6.10: The real and imaginary parts of the complex signal $B_{xf}(t)$ of the short section of the magnetic time series corresponds the time interval and the highlighted frequency band shown in Fig. 6.9. It can be seen that certain time segments contain no energy in this particular frequency band.

E- and B-field components simultaneously. This is because individual IMFs have only one specific frequency value at each time sample and these IMFs typically vary with time. To make Eqs. 4.59 - 4.61 applicable, the frequency axis is therefore divided into a number of frequency windows. In this synthetic test, the spectral range is divided into 25 frequency windows varying from $5.6 \cdot 10^{-5}$ Hz to 0.25 Hz, each with identical band-width on a logarithmic scale. In my analysis, I use the center frequencies of each window as ω_0 , for which tensor estimates should be calculated. One frequency band centered at 5.6×10^{-3} Hz is highlighted as an example in Fig. 6.9. Corresponding to this frequency band and time interval shown in Fig. 6.9, the real and imaginary parts of the complex signal $B_{xf}(t)$ which has been constructed based on DQ, are shown in Fig. 6.10.

For each B- and E-field component I determine for each time sample and for each IMF the frequency window into which the corresponding IMFs fall. The values are used in further processing only if a value in the frequency window is present for all components simultaneously. For each section and each frequency-band, the cross- and auto-correlation terms are calculated based on the sum defined in Eq. 4.61. This yields 8 terms for each section and each frequency-band. In the following one will see that such an approximation is adequate

for impedance calculation as long as the size of the frequency windows are not too large, since impedances only vary gradually with frequencies.

To obtain estimates of the impedance and their error in each frequency band the bootstrap method [Efron, 1979] (see section 4.7.2) is used. For each section, the auto- and cross-correlation terms of electromagnetic time series in Eq. 4.61 are saved as a matrix. Each of the 18 matrices is considered as a data point, that is one has 18 samples $x = (x_1, x_2, \dots, x_{18})$. One section (or matrix) is chosen randomly from x with equal probability $1/18$, and the process is repeated 18 times. This yields 18 sections (or matrices) $x^{*(1)} = (x_1^*, x_2^*, \dots, x_{18}^*)$. It should be noted that among them, some sections x_i^* may be the same section, since each section is chosen independently with equal probability from the entire sampling space. The corresponding auto- and cross-correlation terms of the chosen 18 matrices, $x^{*(1)}$, are then used in Eq. 4.61 to obtain one estimate of the impedance \mathbf{Z}_1 .

The whole procedure is repeated 200 times such that 200 estimates of impedances $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_{200}$ are obtained. The mean values obtained by the bootstrapping are considered as the impedance estimates and confidence intervals of 95% are derived from standard deviations.

The apparent resistivity and phase derived from the impedance together with their confidence intervals are shown as red circles and error bars in Fig. 6.11. As a comparison, the analytic apparent resistivity and phase values are shown as blue curves.

The derived apparent resistivities and phases appear to be reasonable. However, one can observe a small systematic downward bias relative to the analytically determined apparent resistivities. In the next section the downward bias is analyzed, and it is shown that the bias is actually caused by numerical noise introduced in the decomposition procedure of EMD.

6.3 The bias problem and its solution

For field data which contains noise, a downward bias in impedance or apparent resistivities derived from Eq. (4.2) (see also explicit form in Eq. 4.16) is usually observed [Sims et al., 1971]. This is due to the reason that, as has been discussed in section 4.2, the auto-power density BB^* of measured data is systematically larger due to correlated noise. However, the cross-power densities are usually unaffected. The same situation will occur when the impedance is estimated by Berdichevsky's formula Eq. 4.61, since the denominator in Eq. (4.61) contains an autocorrelation term $\sum_{n=0}^N B_{xf}(t_n)B_{xf}^*(t_n) \sum_{n=0}^N B_{yf}(t_n)B_{yf}^*(t_n)$, such that noise in the magnetic fields increases the autocorrelation values, while the numerator is unaffected by noise.

Therefore, the downward bias shown in Fig. 6.11 indicates that the data sets are biased due to noise. To verify this, the impedance estimation using the alternative formula, i.e. Eq. 4.18, is tested. In section 4.3, it has been discussed that when Eq. 4.18 is used to estimate the impedance, the estimate result is upward biased by random noise on the electric field signals [Sims et al., 1971] because the auto-power density EE^* now systematically increases the

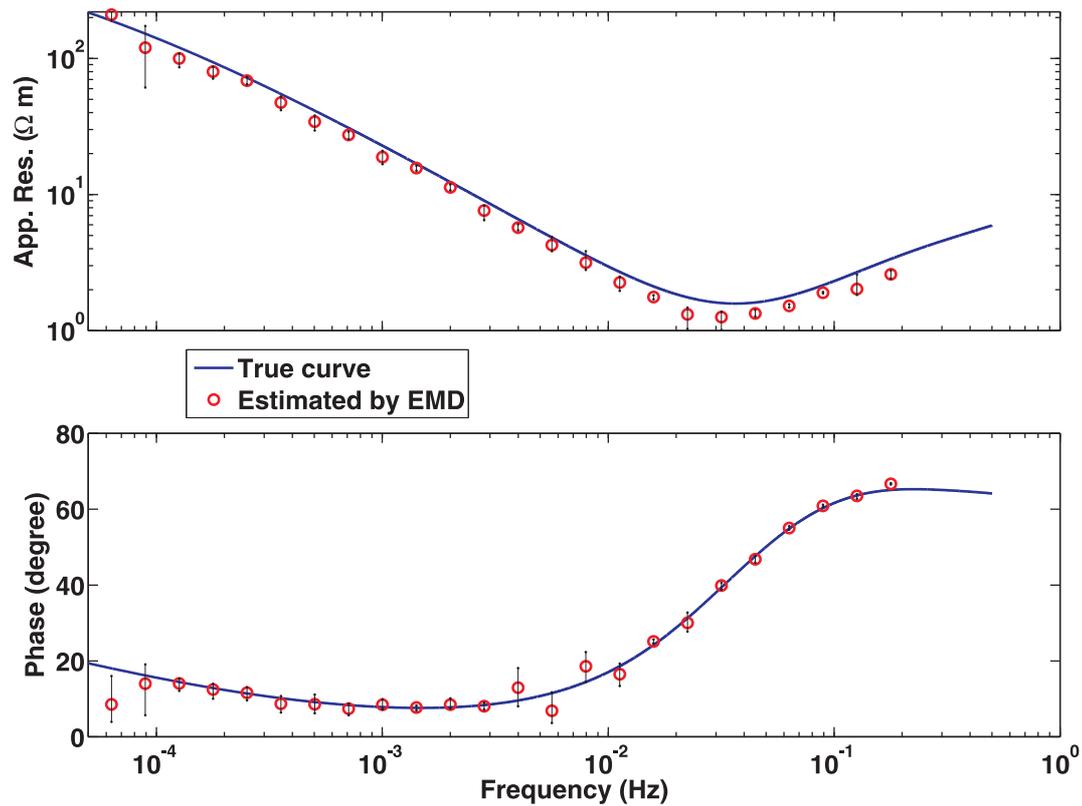


Figure 6.11: Apparent resistivity (upper panel) and phase (lower panel) for the synthetic time series. Blue curves are the analytic values and red circles are the results from EMD. It should be noted that the apparent resistivities are slightly downward biased.

numerator, while the cross spectra in the denominator are unaffected.

Corresponding to Eq. 4.18, a time-frequency domain impedance estimate formula, similar to Eq. 4.61, can be derived by taking $E_{yf}^*(t_n)$ and $E_{xf}^*(t_n)$ instead of $B_{yf}^*(t_n)$ and $B_{yf}^*(t_n)$ in Eqs. 4.59 and 4.60, respectively, with the auto-power term $E_{xf}(t_n)E_{xf}^*(t_n)$ in the numerator:

$$Z_{xy} = \mu_0 \frac{\sum_{n=0}^N E_{xf}(t_n)E_{xf}^*(t_n) \sum_{n=0}^N B_{xf}(t_n)E_{yf}^*(t_n) - \sum_{n=0}^N B_{xf}(t_n)E_{xf}^*(t_n) \sum_{n=0}^N E_{xf}(t_n)E_{yf}^*(t_n)}{\sum_{n=0}^N B_{xf}(t_n)E_{yf}^*(t_n) \sum_{n=0}^N B_{yf}(t_n)E_{xf}^*(t_n) - \sum_{n=0}^N B_{xf}(t_n)E_{xf}^*(t_n) \sum_{n=0}^N B_{yf}(t_n)E_{yf}^*(t_n)} \quad (6.1)$$

When this formula is used, the resulting apparent resistivity and phase, as shown in Fig. 6.12, now exhibit a systematical upward bias, which indicates that the observed systematical bias can be explained by the noise.

However, according to the construction of the synthetic data sets, no measurement noise is present in the synthetic time series. Therefore, the noise which causes downward bias in Fig. 6.11 (or upward bias in Fig. 6.12) must have been introduced by numerical noise within the method itself, i.e. through the EMD and IF/IA calculation procedures.

An analysis of the complete EMD procedure discussed in section 5.2 suggests that the introduction of noise occurs through the iterative sifting procedure in the decomposition. Due to the use of a stop criterion and the limited accuracy of the interpolation of the upper and lower envelopes, the subtracted mean is not really the "true" mean value of the signal. Hence, if this mean is subtracted from the signal, some inaccuracies from the interpolation are added as noise into the signal. After each sifting process, these inaccuracies are summed up and create numerical noise in the estimated IMFs.

A downward bias in the impedances due to the noise on B-Field measurements is commonly suppressed by using the remote referencing method [Gamble et al., 1979a,b], as described in section 4.4. Remote referencing entails the substitution of the autocorrelation term with a cross correlation term between the magnetic field and a remote reference magnetic field measured in the vicinity of MT site. Assuming that the noise in the magnetic field and the reference magnetic field are uncorrelated, the denominator is then unaffected by noise and the bias vanishes.

Since there is no "remote reference" time series for the synthetic example, a "reference" signal is created for the impedance calculation by using another algorithm for the decomposition of the same B-field time series. To highlight that the reference signal is not taken from another station, in the following the term "numeric reference" is used instead of "remote reference". So, for the "numeric reference", the algorithm from Rato et al. [2008] and not the one from Huang et al. [1998] is used to determine the upper and lower envelopes in the EMD procedure. The algorithm from Rato et al. [2008] does not use the actual minima and maxima at the discrete sample positions, but re-positions the extrema by a parabolic

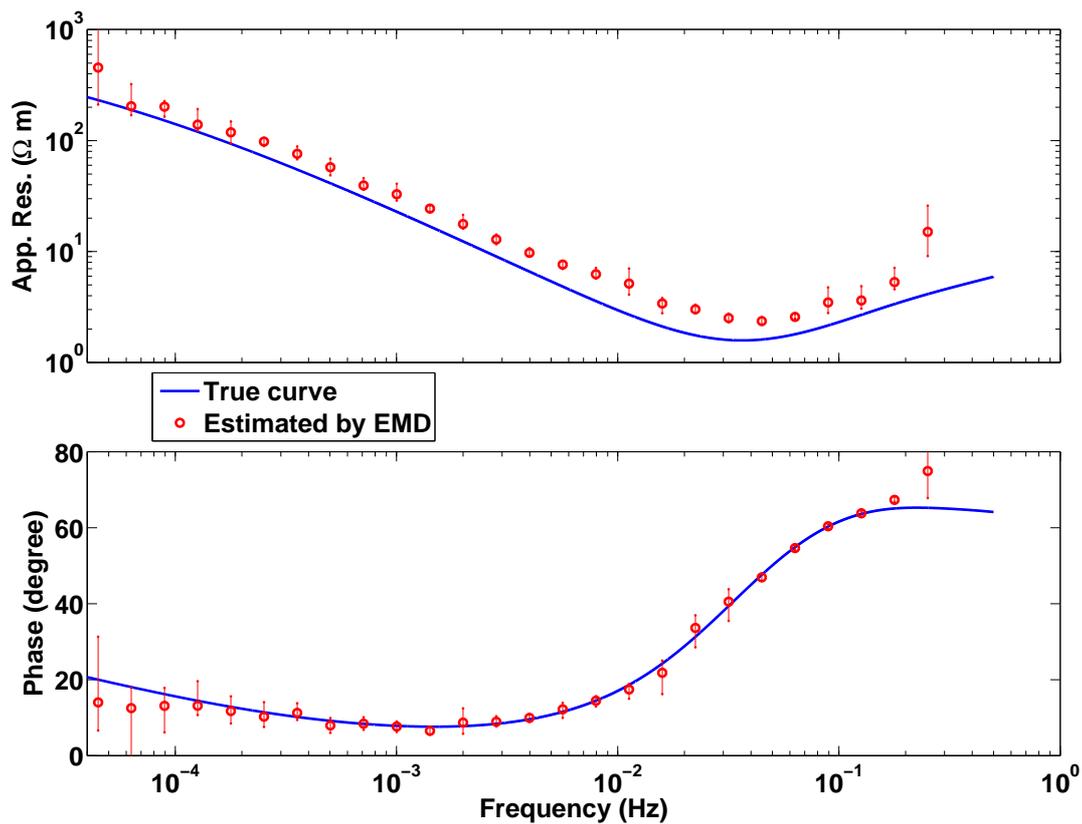


Figure 6.12: Estimated apparent resistivity (upper panel) and phase (lower panel) for the synthetic time series by using formula Eq. 6.1. Blue curves are the analytic values and red circles are the results from EMD. Notice that the apparent resistivities are slightly upward biased.

interpolation considering amplitudes from one extremum and its two neighboring samples (see more details in section 5.6.2). Because shifting processes are repeated several times for each IMF, introduced numeric inaccuracies are significantly differently distributed over the IMFs for both algorithms. This means that the error contribution in the two final decompositions are mostly uncorrelated for the different frequency bands and the impact of these inaccuracies onto the impedance calculation should be reduced.

To estimate the impedance using the "numeric" referencing technique, the general impedance estimator given by Eq. 4.61 is still used, but it has to be modified to

$$Z_{xy} = \mu_0 \frac{\sum_{n=0}^N E_{xf}(t_n) B_{yf}^{r*}(t_n) \sum_{n=0}^N B_{xf}(t_n) B_{xf}^{r*}(t_n) - \sum_{n=0}^N E_{xf}(t_n) B_{xf}^{r*}(t_n) \sum_{n=0}^N B_{xf}(t_n) B_{yf}^{r*}(t_n)}{\sum_{n=0}^N B_{xf}(t_n) B_{xf}^{r*}(t_n) \sum_{n=0}^N B_{yf}(t_n) B_{yf}^{r*}(t_n) - \sum_{n=0}^N B_{xf}(t_n) B_{yf}^{r*}(t_n) \sum_{n=0}^N B_{yf}(t_n) B_{xf}^{r*}(t_n)}, \quad (6.2)$$

where $B_{xf}^r(t_n)$ and $B_{yf}^r(t_n)$ denote the numeric reference time series. Note that the auto-power term of B-field data in the denominator of Eq. 4.61 is substituted by cross-power term of B-field and "numeric" reference time series in Eq. 6.2.

In the impedance estimation procedure, the same 18 sections of B-field data are decomposed by using EMD with Rato et al.'s envelope strategy to obtain IMFs as "numeric" reference data. The calculation of IFs and construction of the complex signals for each IMF are exactly the same as described in previous section. The only difference is that the values are used in further processing not only if a value in the frequency window is present for all B- and E-components simultaneously, but also for all "numeric" reference components. With this procedure the obtained impedance estimates are shown in Fig. 6.13.

The estimated impedances are now satisfactory, the downward bias of apparent resistivities has vanished. However, there are still some slight deviations from the analytically determined apparent resistivity curve, which are most likely caused by remaining numerical noise in the empirical mode decomposition and could be improved by using a more sophisticated robust estimation technique.

The synthetic time series test results provide the convincing evidence that the EMD and instantaneous spectra can be used as an alternative method in impedance estimation. Although numerical noise is introduced by the method and causes systematic downward shifts in the apparent resistivity estimates, the shifts can be strongly reduced by applying a second EMD algorithm to create a numerical reference time series.

However, as mentioned in section 6.1, up to now the time series I used are noise-free. The next logic step is to test the method under various noise conditions, since in reality, the measured data sets always contain different type of noise.

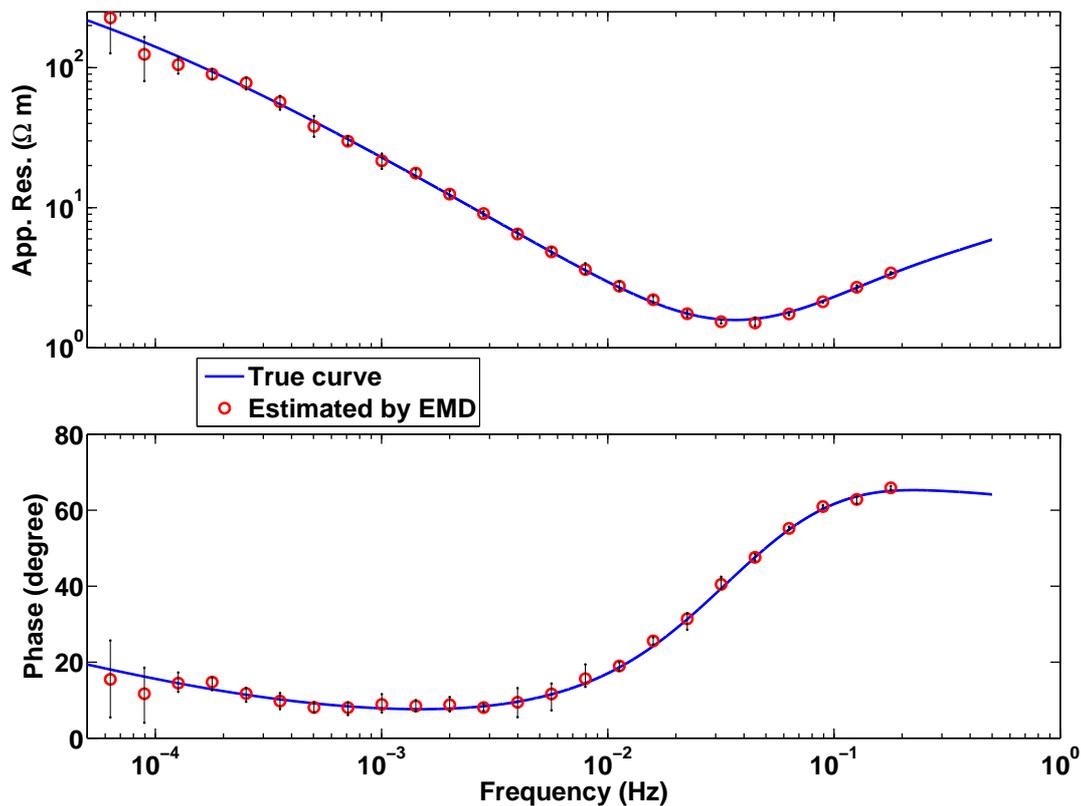


Figure 6.13: Apparent resistivity (upper panel) and phase (lower panel) estimates using a "numeric" reference time series for the B-field data. For the reference the same B-field data are used as for the original, however, the EMD algorithms applied on the time series are different. Blue curves are the analytic values and red circles are the estimated results from EMD. It should be noted that the downward bias of apparent resistivities (see Fig. 6.11) is not present any more.

6.4 Noise test

6.4.1 Adding stationary colored noise

To investigate how far impedance estimates from the EMD based procedure are affected by noise, different levels of uncorrelated noise are added to all time series and the resulting impedance estimations are analyzed. In order to make it easier to evaluate how the impedance estimation procedure treats noise in the different frequency ranges, colored noise is added in such a way that the signal-to-noise ratio is almost the same for all frequencies (remember that the MT field spectra are non-white).

To determine the noise contributions, at first the Fourier amplitude spectra of the synthetic time series from all field components are smoothed, and then $\sim 2\%$, $\sim 5\%$ and $\sim 10\%$ of all amplitude spectra are taken as noise levels. Fig. 6.14 shows examples of the Fourier amplitude spectra of time series B_y (upper panel) and E_x (lower panel), and their corresponding smoothed amplitude spectra of noise with $\sim 2\%$, $\sim 5\%$ and $\sim 10\%$ levels. Here, in order to correspond the real amplitude spectra of the time series, the Fourier amplitude spectra are directly calculated without detrending and windowing on purpose.

Then, for the each noise level, the phase spectra is randomly chosen from an uniform distribution defined on the interval $[-180, 180]$, which ensures that the phase spectra of the different noise channels are uncorrelated. The phase spectra are then added with the amplitude spectra to obtain complex Fourier spectra of the noise. Finally, the entire noise sequences are transferred back to the time domain. Examples of the resulting "noisy" time sections for three levels of noise are shown in Fig. 6.15.

It is ensured that all noise sequences are uncorrelated before they are finally added to the corresponding time series of the different field components. Examples of the resulting "noisy" time series with different levels of noise are plotted in Fig. 6.16.

The impedance and error estimate procedures are the same as described in section 6.2 and 6.3. Apparent resistivity and phase estimates derived from impedance estimates for noise levels of 2%, 5% and 10% are presented in Figs. 6.17 - 6.19, respectively. In addition to this EMD based procedure, estimates are also determined from the simple Fourier derived impedance formula (Eq. 4.16) and the robust processing method BIRRP [Chave and Thomson, 2004]. For the simple Fourier based estimates the same window length as for the EMD procedure is used. To make the results more stable the linear trends in the time series are removed and Hanning windows are applied before the Fourier transforms are performed. To get estimates for the error variances, the bootstrapping strategy as described in section 6.2 is performed.

From Figs. 6.17 - 6.19, one can observe that the mean values from the EMD and simple Fourier derived impedance formula fit similarly well for low frequencies in the apparent resistivity and phase estimates. However, for higher frequencies the mean values from the EMD often fit better than the ones from simple Fourier based formula and the error variances are typically smaller. The results from the EMD are typically in a similar range or slightly worse than those from the robust processing method BIRRP. In particular, for the phase one obtains less precise estimates from the EMD than from BIRRP for all three noise

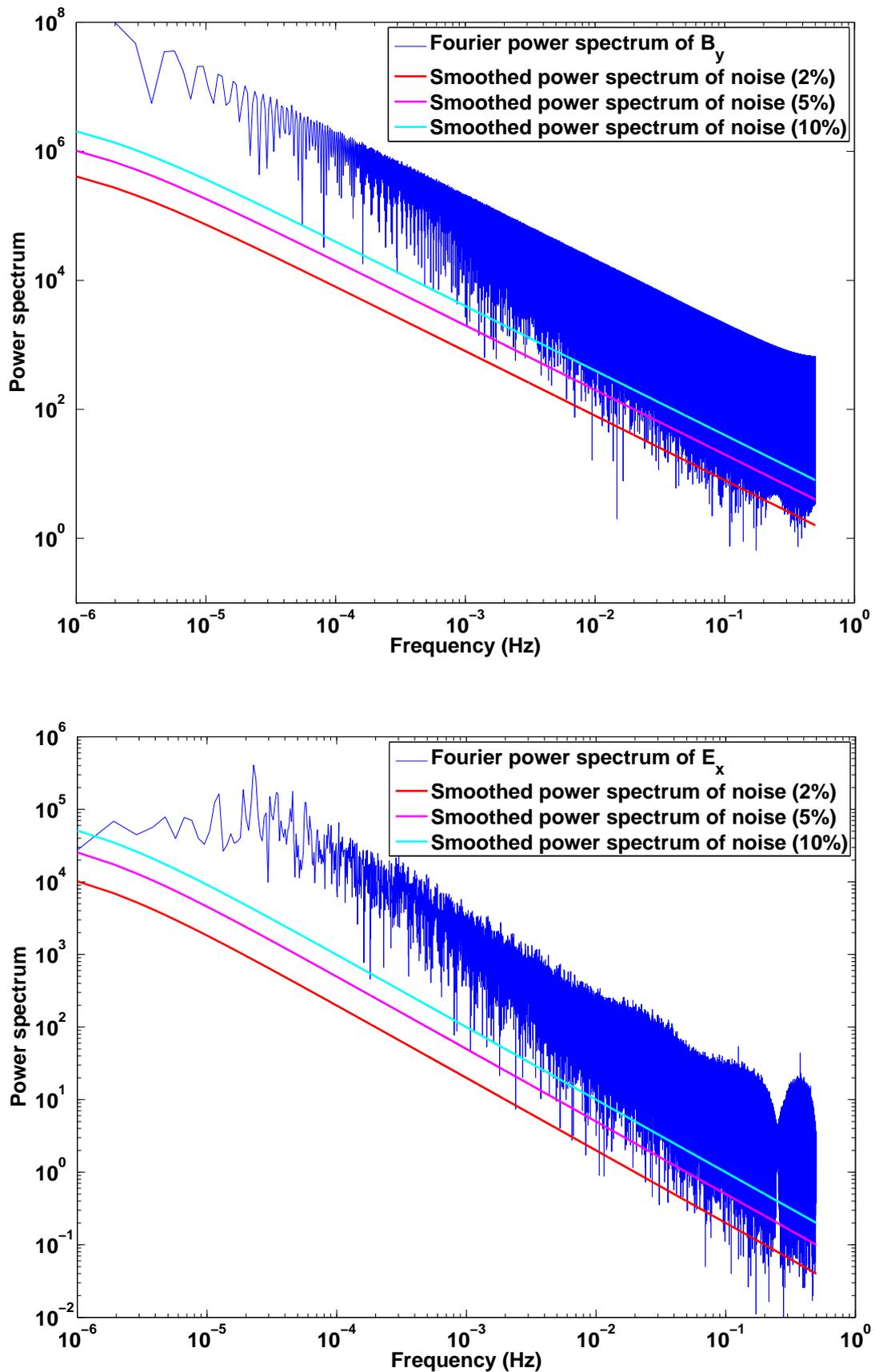


Figure 6.14: Fourier amplitude spectra of synthetic time series and their corresponding smoothed spectra of noise: B_y (upper panel) and E_x (lower panel).

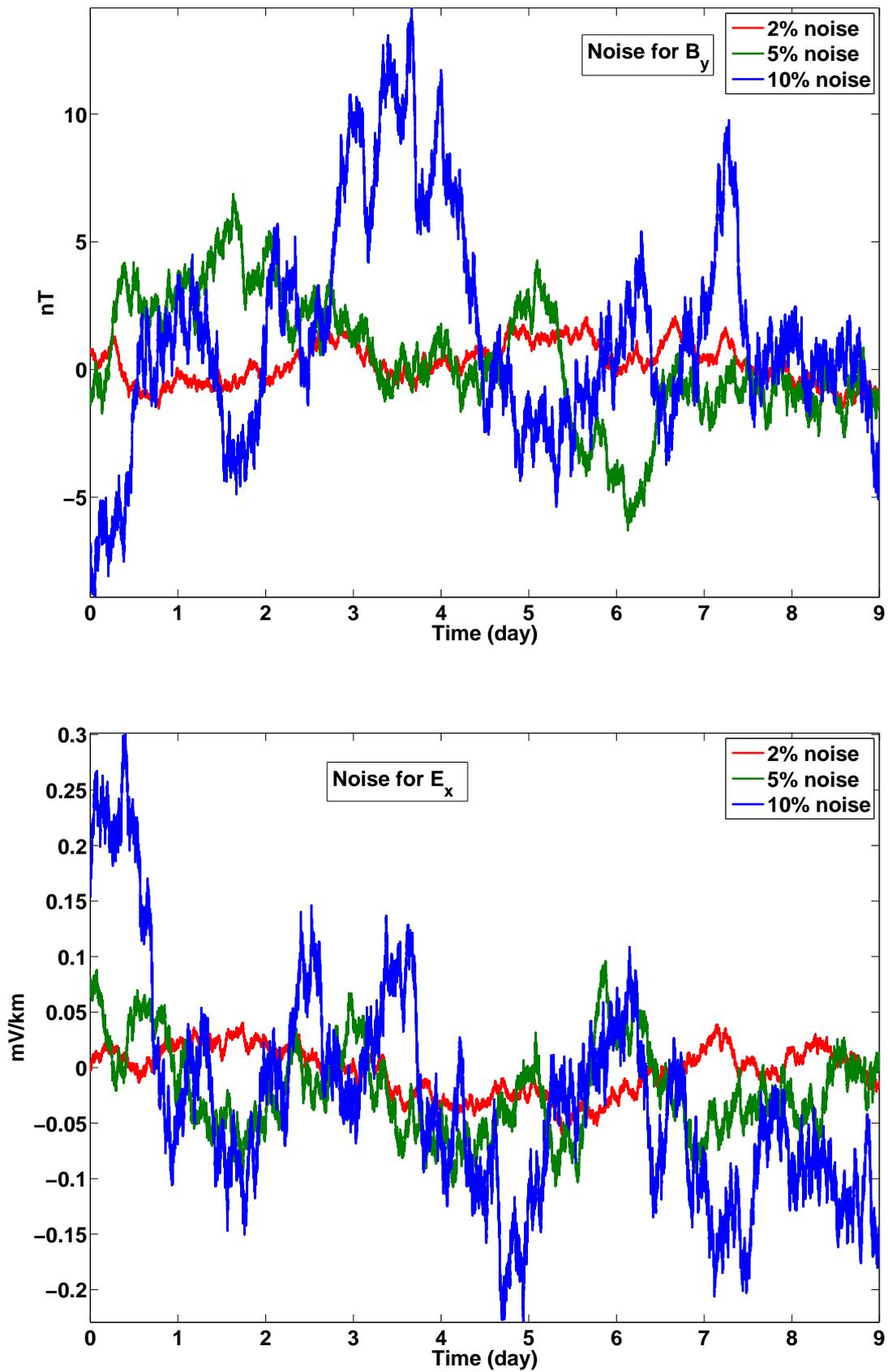


Figure 6.15: Three levels of noise added on B_y (upper panel) and E_x (lower panel).

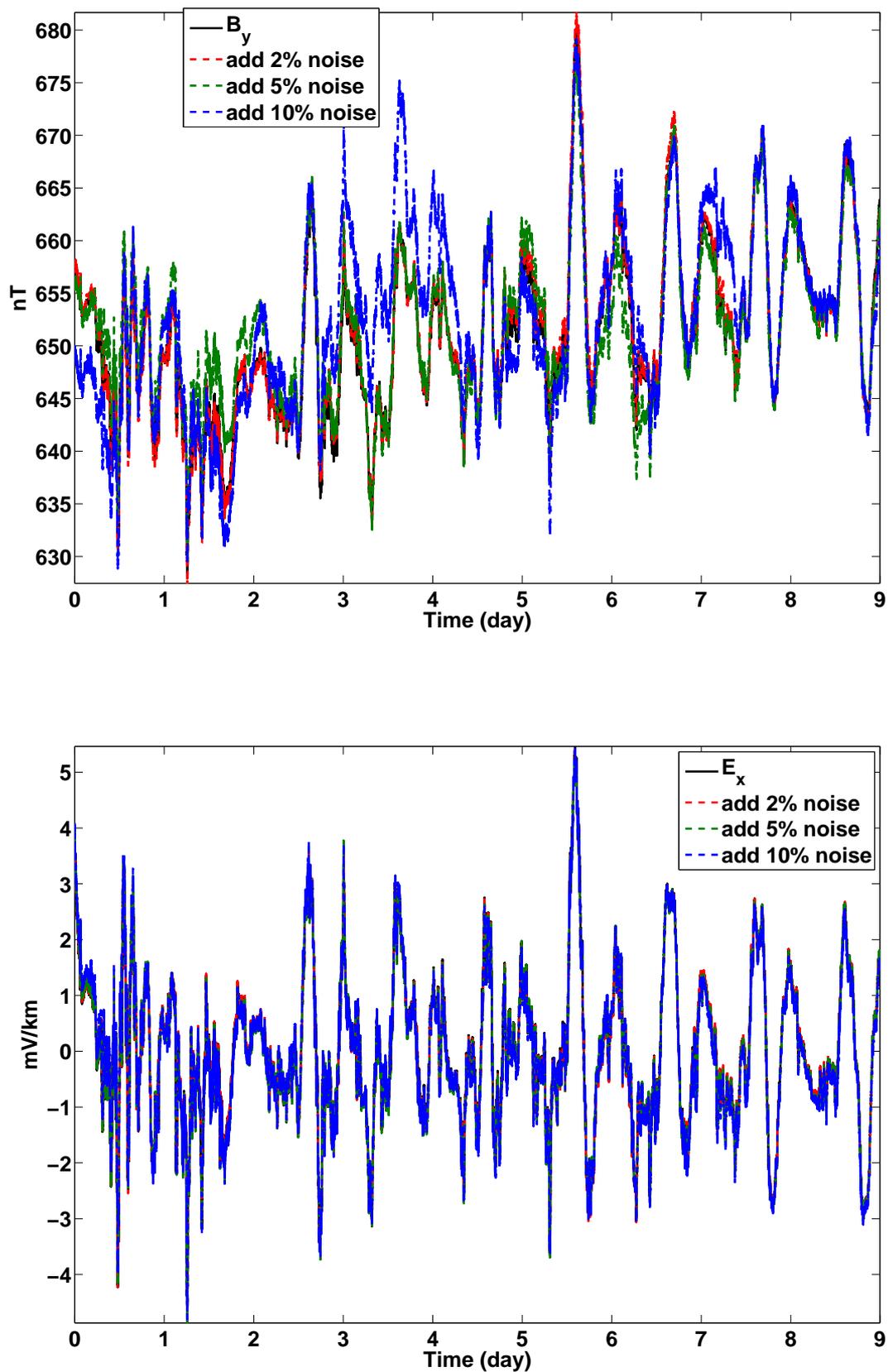


Figure 6.16: The time series after three levels of noise are added: B_y (upper panel) and E_x (lower panel).

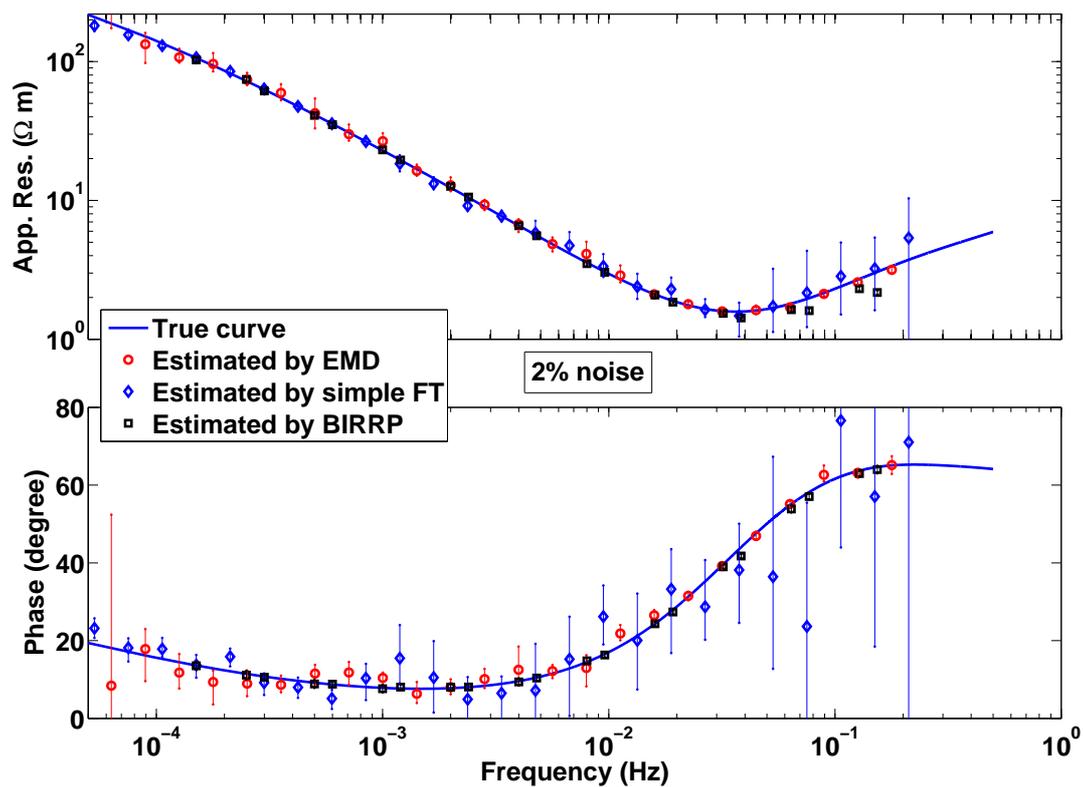


Figure 6.17: Apparent resistivity (upper panel) and phase (lower panel) estimates from EMD using a "numeric" reference time series for the B-field data. In contrast to Fig. 6.13, uncorrelated colored noise, which is $\sim 2\%$ of the amplitude spectra of the associated time series, are added to all time series. In addition to the estimates from EMD (red circles), estimates from Fourier transform without any robust processing steps (blue diamonds) and estimates from BIRRP (black square) are presented.

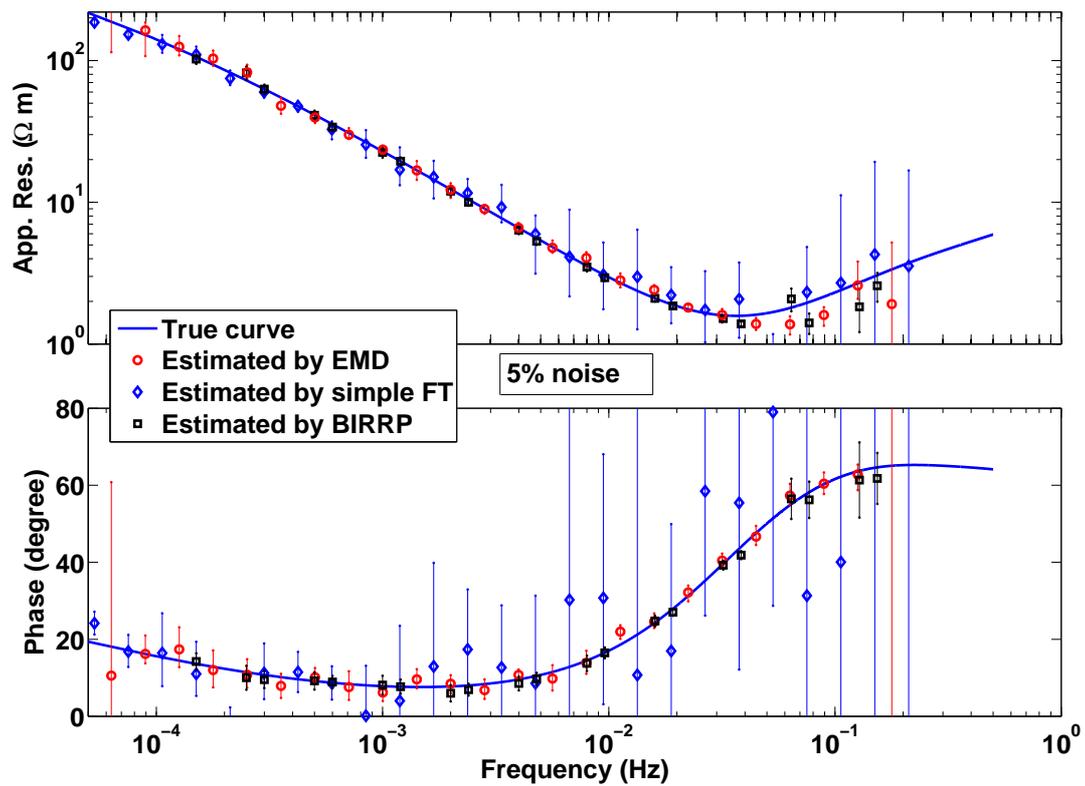


Figure 6.18: Analogue description as in Fig.6.17, except that the noise level here is $\sim 5\%$.

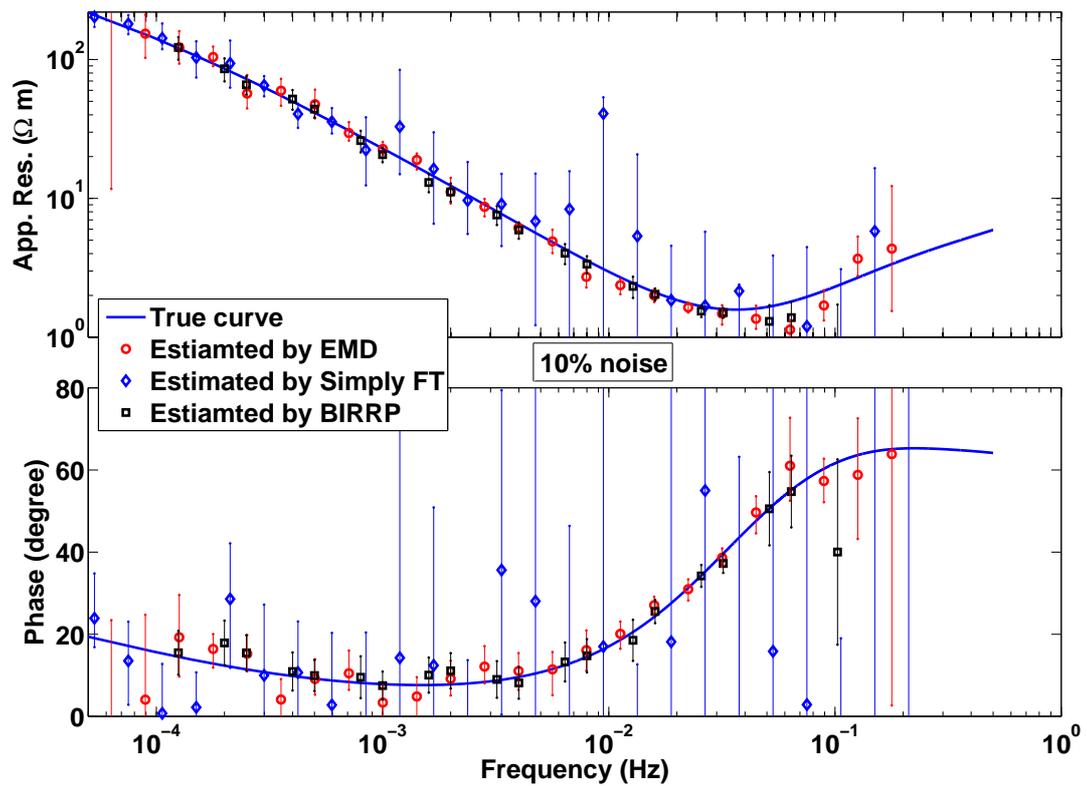


Figure 6.19: Analogue description as in Fig.6.17, except that the noise level here is $\sim 10\%$.

levels, but still better than those from simple Fourier based formula.

Comparison of the results from the different methods demonstrates that EMD provides reliable impedance estimates and it can be concluded that the EMD procedure has the potential to handle noisy data sets. It should be noted that in the EMD procedure no robust processing steps are implemented yet and accordingly one would expect that EMD results should be comparable to those of the simple Fourier derived impedance formula. However, results from the EMD procedure are better! Furthermore they are as good as those from BIRRP method even without using sophisticated robust estimation procedures. This is most likely due to the fact, that

- a) noise added to the data acts locally for EMD processing and is not, as in the Fourier approach, smeared over different frequencies, and
- b) while the noise that has been added to the data is stationary over the entire time series, cutting the time series into segments may introduce non-stationary effects for the shorter time segments, which are better handled by EMD.

6.4.2 Adding non-stationary noise

In addition to stationary noise tests, the effect of non-stationary noise on impedance estimates is investigated. The reason for doing this is that, as discussed in the beginning of chapter 5, in some instances, MT time series (especially marine MT time series) are superimposed with short time scale episodes of non-stationary noise which cause major problems in the MT data processing. In marine MT experiment, the ocean bottom MT instruments are moving with water current and wave, hence the recorded tilt movements exhibit highly non-stationary character. Tilt movements influence both, electric and magnetic magnetic field measurements, however, magnetic data are more sensitive to the tilt movement, because the Earth's static total magnetic field vector has such a large magnitude (from 25,000 nT to 65,000 nT) that even a very small displacement of the magnetometer would cause tens of nT variation in the measured data (see next chapter for a discussion).

Since the relationship between the tilt and the electric and magnetic field variation is quite complex and not fully understood yet, I simply use the tilt measurements, which are recorded at station m05 in Costa Rica, multiply it by some constants and add it as non-stationary noise to the synthetic time series. The measured tilt variations recorded in Costa Rica (magnetic inclination (2007 - 2008): 34°) are shown in Fig. 6.20. To simulate noise in E_x of approximate 10% magnitude, the detrended T_x is simply added to the electric field values. To simulate noise in B_y , the detrended tilt variations T_y is multiplied with a factor of 200, which corresponds to a maximum noise amplitude of ~ 20 nT. This noise level corresponds to the expected noise level at 34° magnetic inclination in Costa Rica for a 0.06° coil angle variation. The resulting synthetic time series with noise (blue) and the original ones (red) are shown in Fig. 6.21.

From Fig. 6.20, it is easy to observe that the non-stationary noise added on both, electric and magnetic time series, are dominated by high frequency variations. This can also be seen in the comparison of Fourier power spectral densities of noisy and original magnetic time series B_y as shown in Fig. 6.22. The power spectral densities are calculated by using the multitaper method. The frequency range of the noise is from about 0.001 Hz to 0.5 Hz.

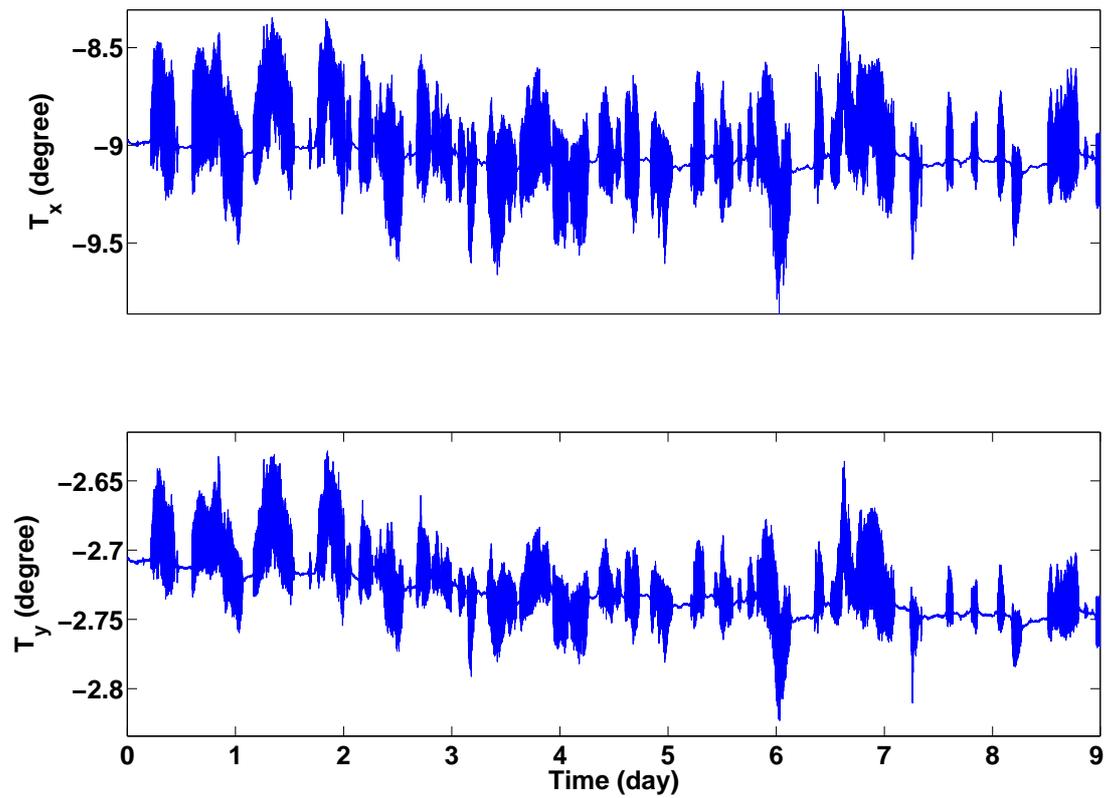


Figure 6.20: The tilt measurements T_x (upper panel) and T_y (lower panel) at marine MT station m05 in Costa Rica.

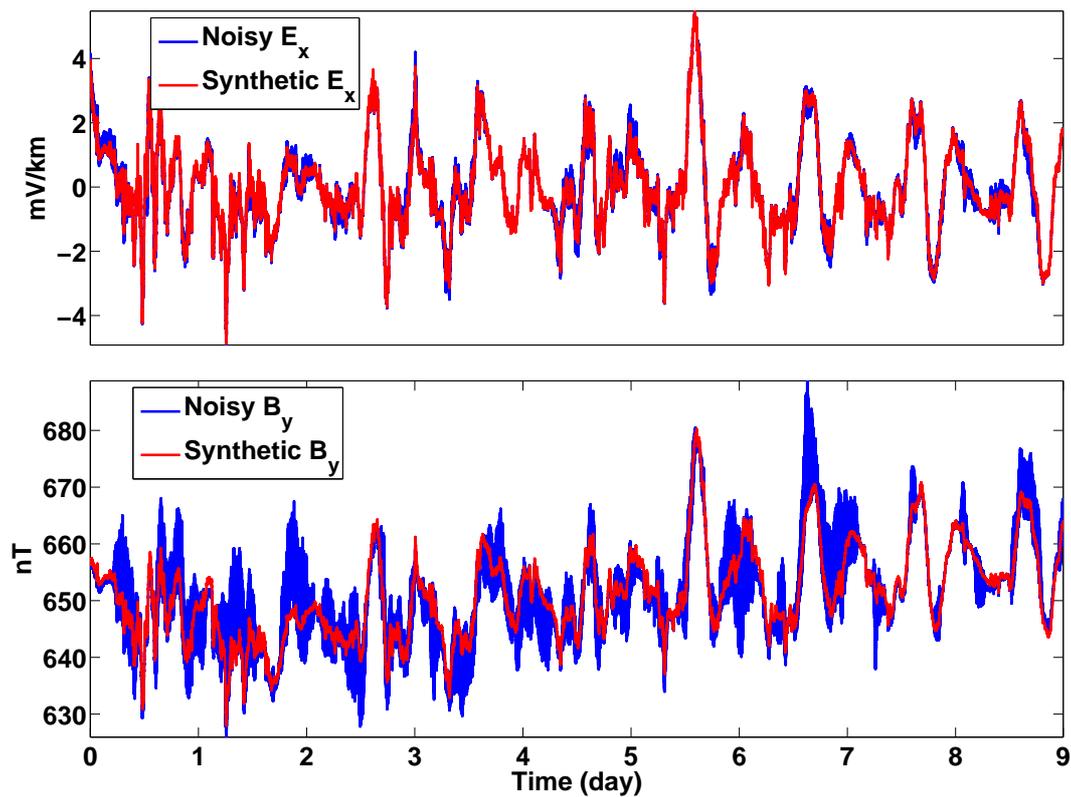


Figure 6.21: Noisy (blue curves) and original (red curves) synthetic electric (upper panel) and magnetic (lower panel) field data. Non-stationary noise added on data are obtained from the detrending tilt measurements (Fig. 6.20) multiplied by constants. T_x multiplied by 1 is added on E_x and T_y multiplied by 200 is added on B_y .

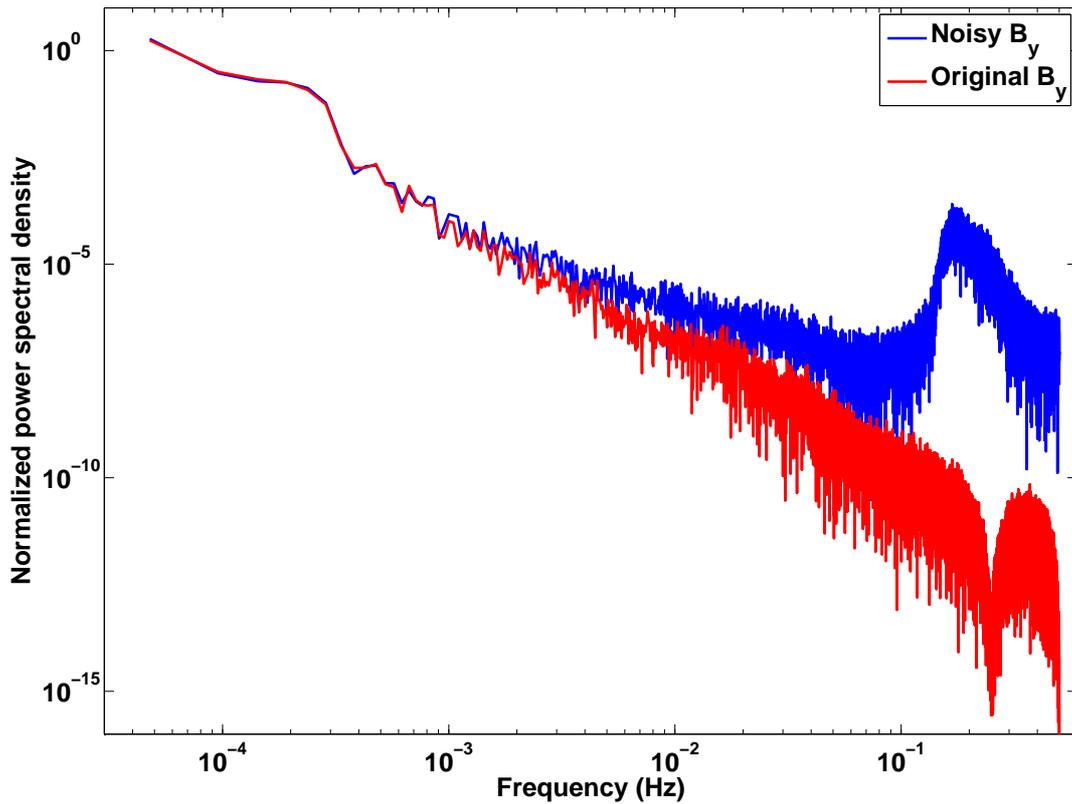


Figure 6.22: Normalized power spectral densities of noisy and original synthetic magnetic field data, which are calculated by using the multitaper method.

Now the EMD, the simple FT and the BIRRP method are used to estimate the impedance as described in the previous section. However, no reference time series is used except the "numerical reference" in the EMD procedure to avoid a bias caused by the numerical inaccuracies. The results of apparent resistivities and phases are shown in Fig. 6.23.

From Fig. 6.23, one can observe that in the frequency range lower than 0.001 Hz the apparent resistivity and phase estimates derived from all three methods fit similarly well. In the frequency range from 0.001 Hz to 0.01 Hz, the results from EMD and BIRRP also fit similarly well, however, those from simple FT are strongly downward biased. At high frequencies (>0.01 Hz) the results from the EMD are slightly worse than those from the robust processing method BIRRP, but better than those from simple FT method. The comparison indicates that the EMD method can also provides impedance estimates as reliable as BIRRP, when non-stationary noise are present in data, even without relying on robust estimation procedures as employed by BIRRP.

In next chapter, the new method is tested to process the real measured MT data.

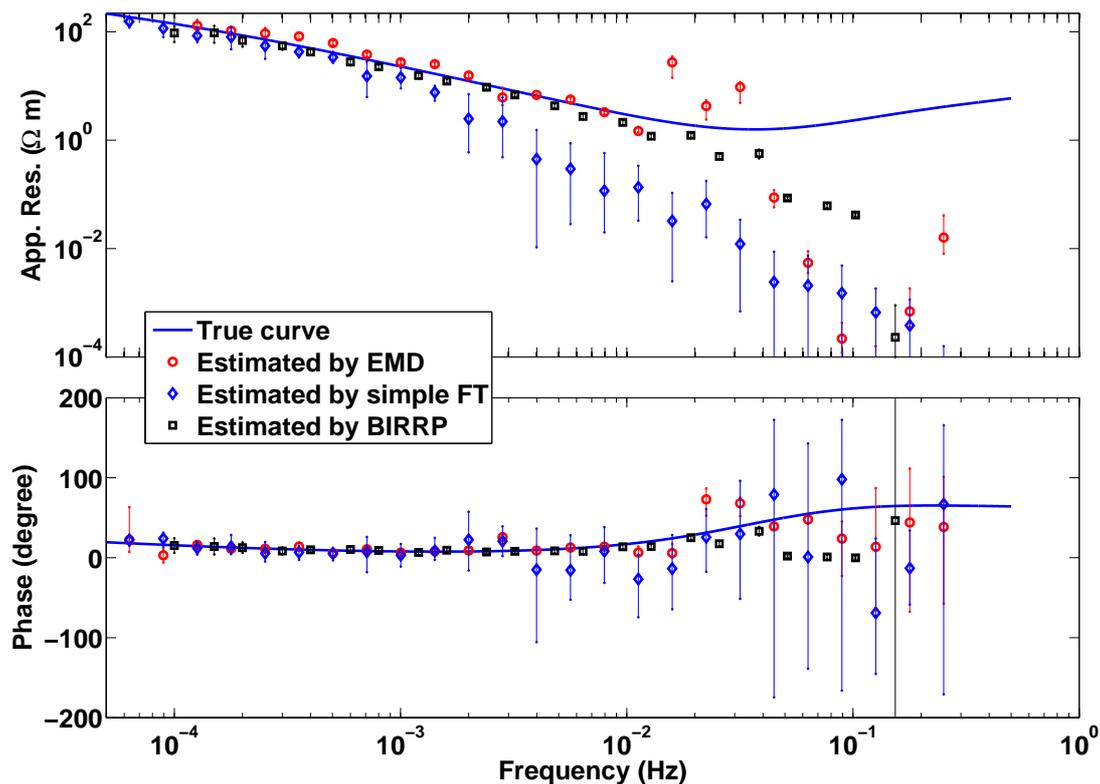


Figure 6.23: Apparent resistivity (upper panel) and phase (lower panel) estimates from EMD using a "numeric" reference time series for the B-field data. In contrast to Fig. 6.13, uncorrelated non-stationary noise are added to all time series. In addition to the estimates from EMD (red circles), estimates from Fourier transform without any robust processing steps (blue diamonds) and estimates from BIRRP (black square) are presented.

Chapter 7

Using EMD to Process Field Data

In this chapter, the EMD method is applied to real MT data sets collected in Costa Rica within the framework of SFB 574, Volatiles and Fluids in Subduction Zones. The purpose of the MT experiment was to image the hydration and dehydration of a subduction zone. First some background on the experiment itself is given, then the quality of the acquired data sets and noise characteristics are discussed. In the next step, the EMD method is exemplarily applied to process the relative quiet segments of the data on some sites which contain different level of motion-induced noise, and the results are compared with those estimated by the robust processing BIRRP. Then the instantaneous spectrum of the noisy marine MT data is analyzed and some ideas on possible future development in using EMD to process noisy MT data are presented. Last the EMD method is applied to process a high-noise-level data set and again the results are compared with those from BIRRP.

7.1 Fluids in subduction zones and SFB 574

The largest earthquakes and volcanic eruptions occur at subduction zones, where two tectonic plates collide and where one plate (always oceanic plate) slides under an overriding plate. The subducting plate transfers material which contains water to greater depths into the Earth's interior. These fluids play an important role in understanding the subduction processes and in the emergence of natural hazards.

The collaborative research project SFB 574 (SonderForschungsBereich 574: Volatiles and Fluids in Subduction Zones, 2001 - 2012) at University of Kiel and GEOMAR undertook a large multi-disciplinary investigation of fluids and volatiles in subduction zones. The aims of the project were to reveal the flow of volatiles and fluids into and out of subduction zones and to understand trigger mechanisms for natural disasters and their complex effects on the Earth's climate. One of the main research area of the project was the Central American subduction zone which is the home to several active volcanoes and is frequently struck by strong earthquakes.

Based on the results of SFB 574 and other related research, much has been revealed about the structure and dynamics of the Central American subduction zone, its hydration and dehydration processes and the estimations of fluid budgets. The incoming oceanic crust and lithosphere is already hydrated during its creation at mid-ocean ridges [Mevel, 2003].

Fluids are incorporated into the extrusives of the upper oceanic crust near the spreading center and pervades the sedimentary sequence within clays. The incoming plate bends when it approaches the trench and the bending is thought to cause extension and normal faulting. Fluids are proposed to penetrate through these faults into the crust and upper mantle [Ranero et al., 2003, DeShon and Schwartz, 2004, Grevemeyer et al., 2005, H.R.DeShon et al., 2006, Grevemeyer et al., 2007, Worzewski et al., 2011]. Water is carried into the subduction zone as free pore water or chemically bound in minerals of sediments, crust and mantle. In the subduction channel, that is the region between the subducting and overriding plates, dehydration process occurs instead of hydration process. Free fluids are squeezed out from gradually closing fractures and pore spaces along the upper few kilometers of the subduction channel and mineral-bound water is released along the deeper part of the channel. Portion of fluids are also released into the mantle wedge, where they contribute to melting which then rises upwards by magma. Fluids released at greater depth may find their way to the surface through volcanic gas emissions and magma extrusions. However, research within the SFB 574 indicated that there is a discrepancy between the input flux and the flux leaving the subduction system through volcanic emissions, fluid venting and transportation to the deep mantle [Rüpke et al., 2004, Rüpke, 2004, Worzewski et al., 2011, Worzewski, 2011]. Furthermore, it is revealed that fluids can change the dynamics and strength of rocks and thus influence earthquake activities [Peacock, 1990, Byerlee, 1993, Hacker et al., 2003]. The largest earthquakes happen in the so-called seismogenic zone, which is the shearing area between the subducting and the overriding slabs. These transition from aseismic to seismogenic behavior is association with a decrease in fluid content.

In order to further understanding the fluid distribution on the subduction zone from the outer arc high to the volcanic front, the magnetotelluric (MT) technique was applied on the Costa Rican subduction zone under the framework of SFB 574. The MT technique is particularly suitable in this context, since electrical resistivity is very sensitive to the temperature and the presence of highly conductive phases such as fluids.

7.2 MT experiment in the Costa Rican subduction Zone

7.2.1 MT profile

Fig. 7.1 shows the tectonic setting of the Costa Rican subduction zone. The Cocos plate and the Nazca plate both originated 22.7 Ma ago from the breakup of the Farallon plate [Barckhausen et al., 2001]. The Cocos Plate is outlined by the East Pacific Rise in the West, the Cocos-Nazca Spreading Center in the South, the Panama Transform in the East, and the Middle America Trench in the North-East. The East Pacific Rise has a fast spreading rate of 10-12 cm/yr [Carbotte and Macdonald, 1992, Madsen et al., 1992], which leads to the production of a thin, smooth and uniform oceanic crust. In contrast, the Cocos-Nazca Spreading Center is a relatively slow spreading ridge with an increasing rate from 2.6 cm/yr in the West to 7.5 cm/yr in the East [Allerton et al., 1996, Lonsdale, 1998], which causes an anti-clockwise rotation of the Cocos Plate and leads to the production of a rougher crust. The Cocos plate converges at different rates to the Middle America Trench. The rate is approximately 9 cm/yr in the South and 7 cm/yr in the North. In Costa Rica, the Cocos plate is thrust beneath the Caribbean plate at an orthogonal convergence rate of approx-

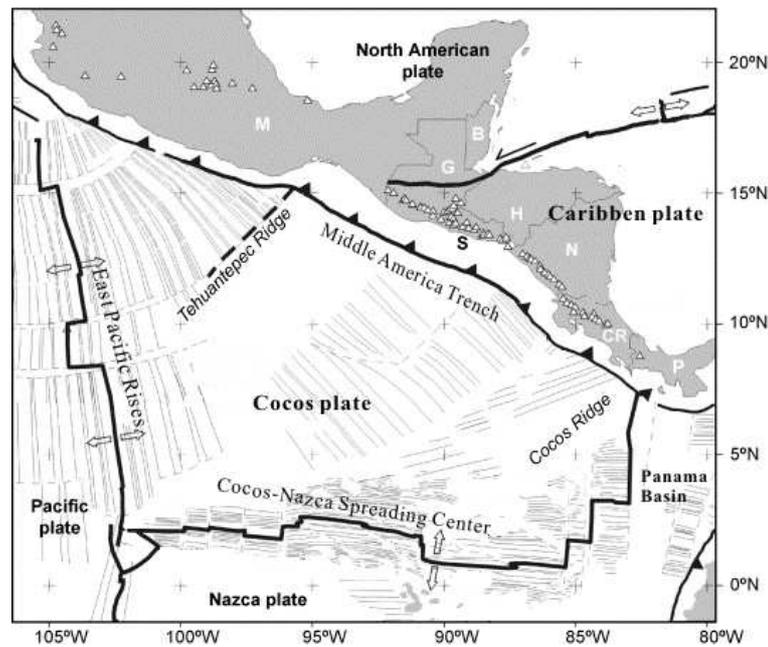


Figure 7.1: Tectonic setting of Central America. The countries are labeled as follows: Mexico (M), Belize (B), Guatemala (G), El Salvador (S), Honduras (H), Nicaragua (N), Costa Rica (CR) and Panama (P). Map is modified from Rogers et al. [2002].

imately 8 cm/yr [DeMets, 2001]. Around Nicoya peninsula in northern Costa Rica, which is characterized by active volcanism, the crust is 22 Ma old with 6 km thickness [Walther and Flueh, 2002].

In 2007-2008, an offshore magnetotelluric experiment was performed along a profile crossing the trench [Worzewski et al., 2011, Worzewski, 2011]. Fig. 7.2 shows the topographic map and MT profile across the Costa Rican subduction zone. Offshore stations were deployed by GEOMAR in 2007 and the land stations were installed by the Free University of Berlin in 2008. The total length of profile length is about 370 km. The onshore profile acquired by the Free University of Berlin was analyzed by Brasse et al. [2009].

The instruments used offshore were the first generation of Ocean Bottom MagnetoTelluric (OBMT) instruments developed at GEOMAR and University of Kiel. The frame of the OBMT instrument was constructed based on the ocean bottom seismometers designed from GEOMAR [Bialas and Flueh, 1999]. It consists of a recording instrument in a titanium cylinder, flotation, anchor, an acoustic releaser and four 5 m plastic tubes with electrodes fixed at the far end spanning two orthogonal electric dipoles of 10 m length. The recording instrument includes a three component fluxgate magnetometer, a two component electric field amplifier, dual axes tilt meters for measuring pitch and roll, a temperature sensor, an internal data logger, a synchronized high precision clock and a compact flash card for data storage. With the equipment, the natural temporal geomagnetic and geoelectric field variations in a period range from $T = 0.2$ s to DC (possible sampling rates of 10Hz, 5Hz, 1Hz and 1/60Hz.) can be measured.

The offshore instruments were deployed free falling from the ship and recorded MT signals from 5. Sept. 2007 to 22. Jan. 2008. After the measurement the anchor was released

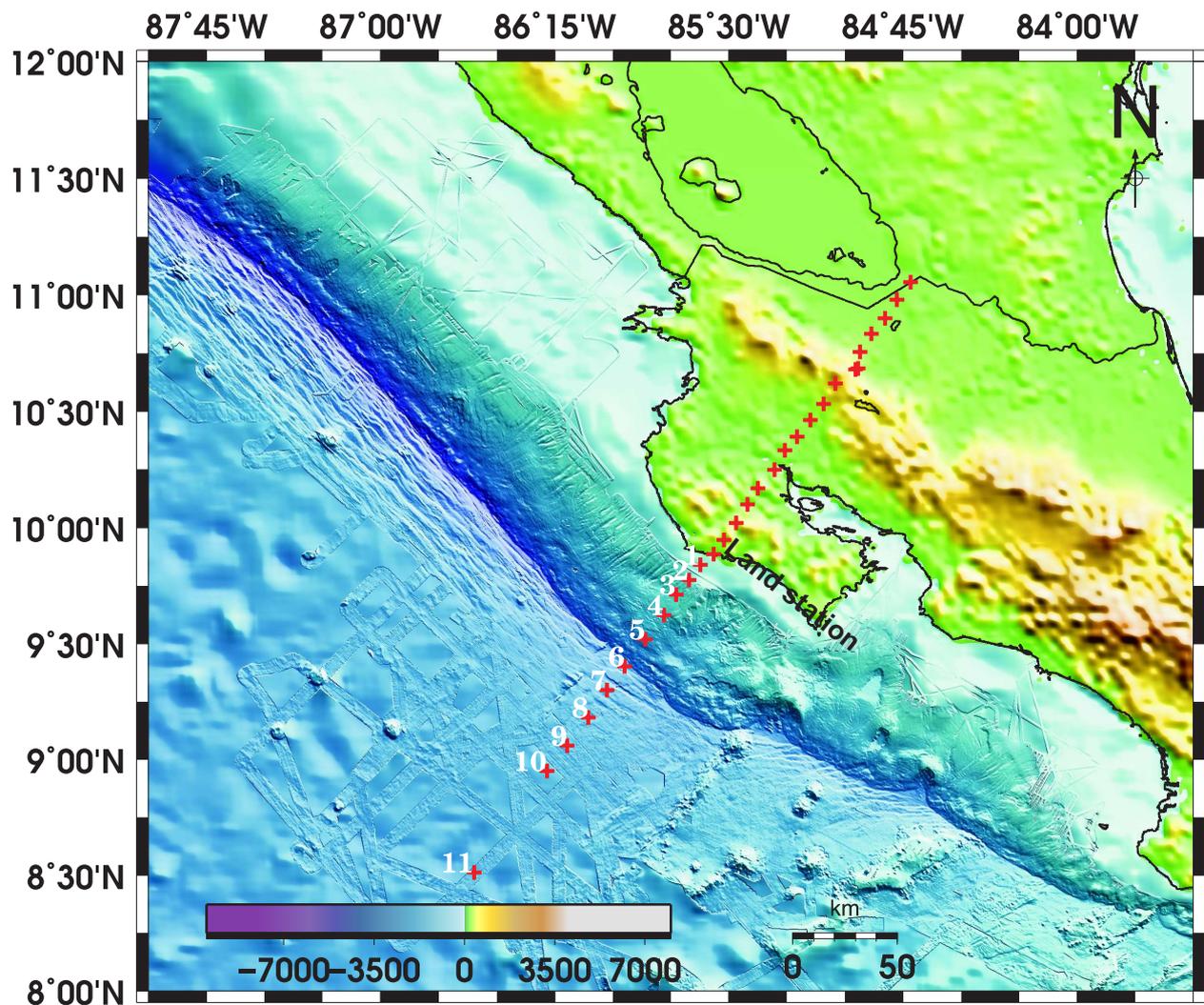


Figure 7.2: Topographic map and MT profile across the Costa Rican subduction zone. Offshore MT stations were deployed by GEOMAR, Kiel, in 2007 and land stations were installed by the Free University of Berlin in 2008.

through an acoustic signal and the instrument rose to the surface to be recovered. The data were recorded at sampling rate of 1 Hz. Simultaneously, reference magnetic field variation were recorded by one onshore station installed near the coast line (see Fig. 7.2).

On land, the MT instrument can be installed horizontally and orientated to a fixed direction, whereas the position and orientation direction of the OBMT instrument can not be controlled due to free falling deployment. Furthermore due to seafloor topography the instruments therefore exhibit a tilt. For a static tilt, measured data can be corrected using measured tilt angles. However, due to water currents and waves the tilt of the OBMT instrument varies. These variations in tilt cause strong artificial magnetic field variations. This is because, when the instrument is stable, the measured projectional components in the horizontal and vertical directions of the Earth's static field (see section 2.2) are also static and can be accounted for. However, this is not the case anymore when the angles of pitch and roll of the instrument are changing. Compared with the Earth's static field (intensity in order of $\sim 10^4$ nT) the Earth's natural magnetic field variations (intensity in order of $\sim 10^{-1} - 10^2$ nT) used by MT are very small. A slight change of the tilt angle of the instrument changes the magnetic flux through the coil, and results in large variations (relative to the natural variations) of the magnetic field unrelated to the induction process. For example, if the magnetometer position is displaced by a very small angle, say 0.1° , from the horizontal, the measured fields will differ from tens of nT in both horizontal and vertical directions, which is already in the order of natural magnetic fluctuation. The marine MT data collected offshore Costa Rica exhibited extraordinarily large distortions due to instrument movement and would be regarded as highly noisy in comparison to land measurements.

7.2.2 Data quality

The data quality ranges from good to extremely noisy depending on the instrument's position and tilt variations due to ocean currents. At the stations closer to shore, stations m01, m03, m04 and m05 (unfortunately, station m02 was lost), the measured data show very strong noise. For station m01 the noise was particularly high such that no impedance estimates could be obtained with any processing technique. At station m06 - m10, the data exhibit intermediate noise. At station farthest from the coast, m11, data exhibits only small noise. Fig. 7.3 shows the example of a high-noise-level time series measured at station m05 for different time sections: 5 days, 1 day and 7 hours. In each panel, three orthogonal components of magnetic field (B_x, B_y, B_z), two components of electric field (E_x, E_y) and two components of tilt movements (t_x, t_y) are displayed. Fig. 7.3 clearly illustrates that most of the noise can be tracked back to water induced movement of the instrument as shown by tiltmeter recording. Both magnetic and electric field measurements are influenced by the movement of the instrument, however, magnetic field recording are more sensitive to the tilt movement. Throughout the full four month long time series measured at station m05, noise appears every few hours and lasts for several hours and only several small sections with several hours length can be identified in which the instrument was relative stable. These sections are referred to as quiet sections, and are shown in Fig. 7.3 C.

A more quantitative assessment through the coherencies between 5-day-long tilt and horizontal magnetic components time series measured at station m05 (see Fig. 7.3 A)) is shown in Fig. 7.4. The power spectral densities in the calculation of coherency are estimated by

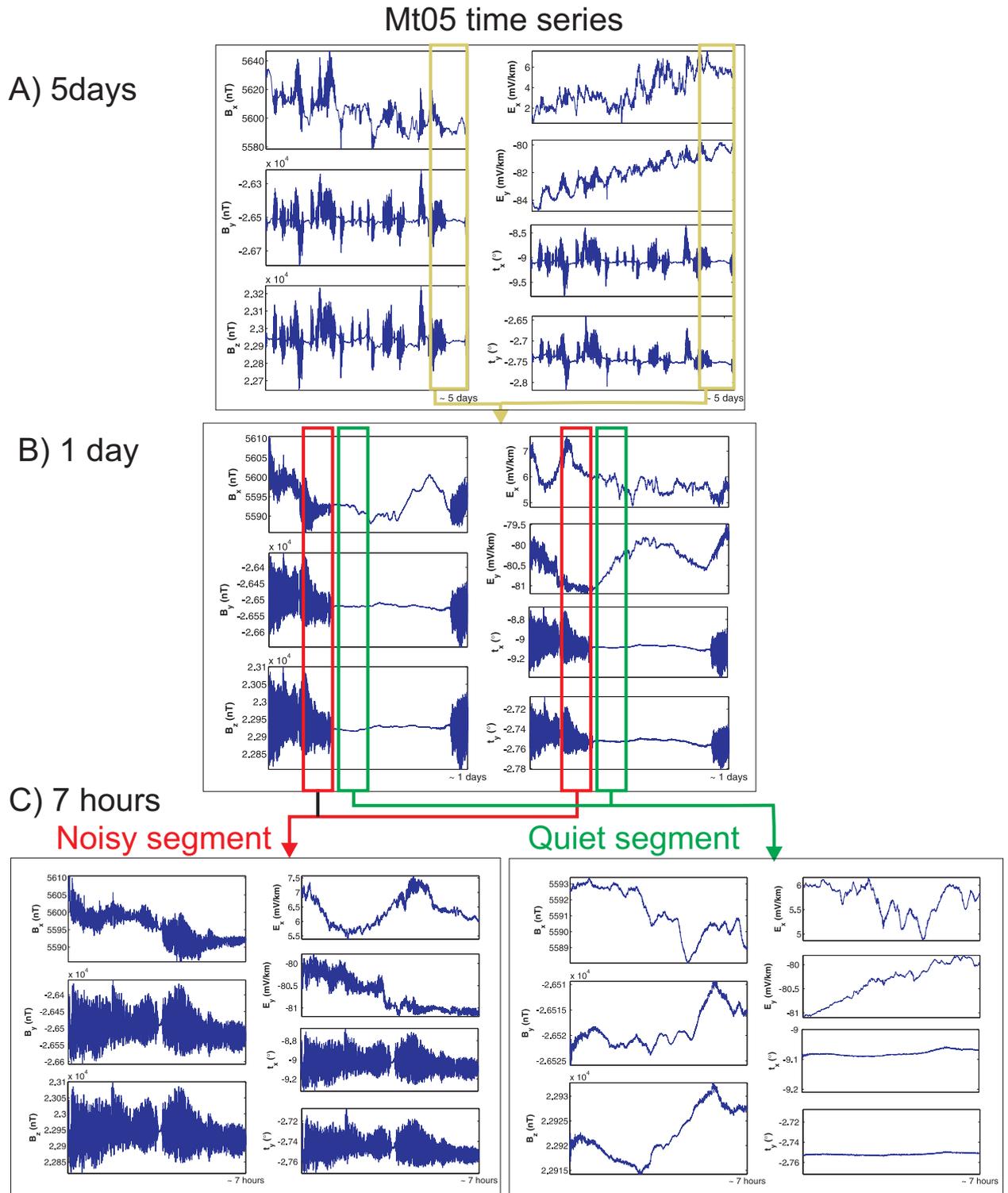


Figure 7.3: Noisy time series example measured at station m05. Different time sections are displayed: A) 5 days, B) 1 day and C) 7 hours. B_x, B_y, B_z are three orthogonal components of the magnetic field; E_x, E_y are two orthogonal components of electric field; t_x, t_y are two components of tilt measurements (pitch and roll angles). The time series exhibit significant noise over the entire measuring period, due to the affection of motion induced by tilt variations.

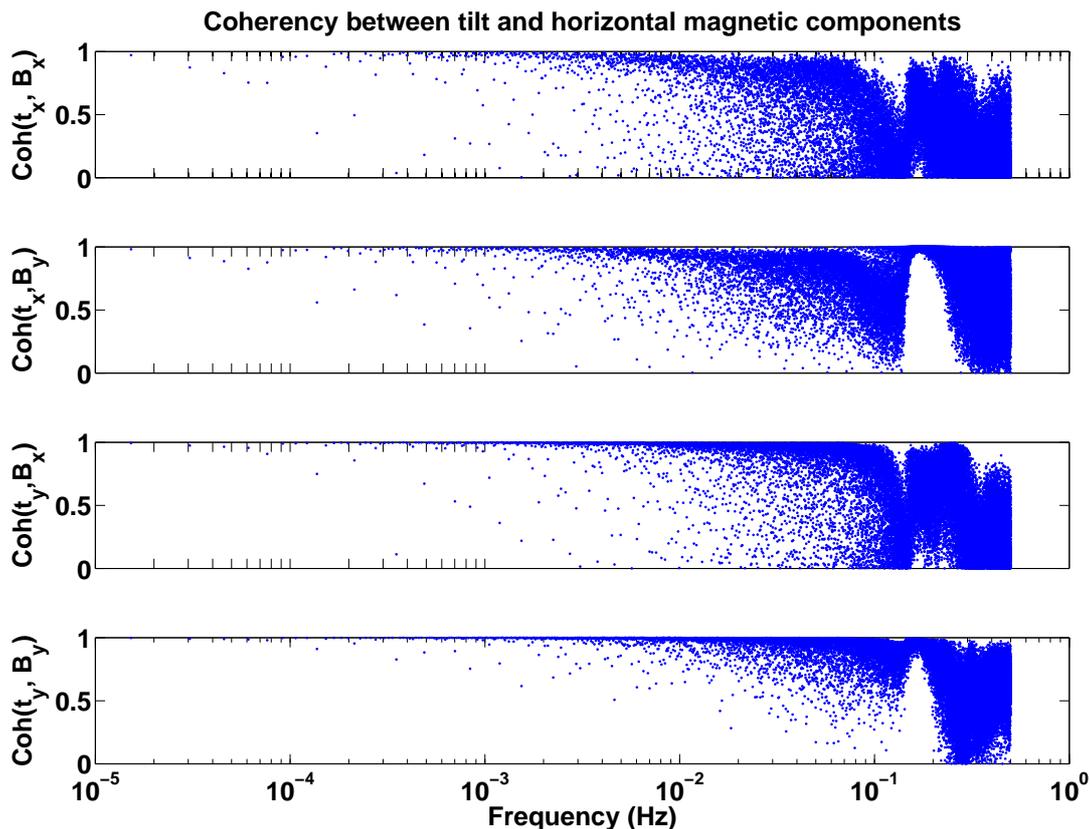


Figure 7.4: Coherency between 5-day-long tilt components and horizontal magnetic components time series measured at station m05. The time series are shown in Fig. 7.3 A).

using a modified periodogram method with a Hanning window [Welch, 1967]. Fig. 7.4 shows that a very high coherency at all frequencies between one tiltmeter component and one horizontal magnetic field component exists, which implies that any movement of the instrument directly influences the magnetic field registration.

7.2.3 Preprocessing of MT time series

In the following, the preprocessing is summarized. Further details can be found in Worzewski [2011]. All the data sets which are used from this section to the end of this chapter are rotated to the horizontal plane and to North.

In order to derive an electromagnetic transfer function, the extraction of signal without motion induced noise is the first important step in the data processing sequence. Different techniques were utilized in an attempt to reduce the noise caused by the instrument's shaking. One attempt was made to rotate back the magnetic field time series at each sample using the displacement angles measured by the tiltmeter with the rotational matrices described in Fitterman and Yin [2004]. Additionally, several filter options were tested, including different type of filters designed in frequency domain and a time domain adaptive filter described by Lezaeta et al. [2005] which is designed to clean the time series using a variable filter that

permanently adapts its entries according to changes in the tiltmeter time series. Furthermore, a frequency dependent transfer function between noise vector (tiltmeter components) and horizontal magnetic field vectors was searched for in order to transform back the time series vector.

However, all of these attempts failed, especially for sections with high noise level. Neither of the attempts reduced the noise sufficiently to a level lower than the natural electromagnetic signals. In some sections the noise are generated due to tilt variations with amplitude higher than 0.1° , at some stations even reaching several degrees. These movements infer a response in the magnetic field sensor that ranges over tens to hundreds of nT, which is much higher than natural signals measured at the ocean bottom. Moreover, as shown in Fig. 7.4, the periodical motions dominate over natural magnetic field variations at all frequencies, therefore, the Fourier-based frequency-domain filters also fail. When such a filter is used, not only the noise induced by instrument's shaking but also the natural signals are filtered out. At such high noise levels, time domain adaptive filtering technique also fail as has been observed by Lezaeta et al. [2005]. Furthermore, noise that arises from movement is not unnecessary uncorrelated for E- and B-fields, thus, the standard processing techniques, which are described in chapter 4, cannot provide clean transfer functions.

After all attempts using the complete information of the full time series to calculate reasonable transfer functions failed, a more straight forward strategy was adopted by Worzewski [2011], in which all noisy sections which are defined by the tiltmeter measurements are discarded and from the remaining data sets only quiet time sections were used for processing. In parallel some other different strategies to deal with the non-stationary noise (including the method described in this thesis) and new correction methods are still ongoing. Recently, Neska et al. [2013] reported that a new correction approach called tilt-response correction can better reduce the motion noise and provide an alternative source of information about the Earth's total magnetic field and obtain good-quality transfer function estimates.

In the original approach by Worzewski [2011], both tilt variation and the coherency between quiet section and section measured simultaneously at a land station have been taken into consideration in the selection of quiet sections. The quiet sections are identified by requiring that the tilt variations are smaller than 0.01 (for extremely noisy stations a higher threshold value had to be accepted) and that marine and land magnetic field variations are highly coherent.

Fig. 7.5 shows an overview of all identified quiet time sections of all marine stations which are used for data processing. As a comparison, data recorded at land station has good quality throughout the entire measure time.

It can be clearly observed that at station m01 the data is so noisy that no section fulfills the criterion for a quiet section. Stations close to the coast line (m03 - m05) are so disturbed that the data only exhibit very short quiet sections. At station m03, any section of comparably low noise level has to be picked out to ensure enough data for processing, resulting in about 60 single quiet sections of only a few hours length each. More than 90% of the time series had to be discarded. Stations further offshore (m06 - m11) display good quality and noisy sections are cut out liberally. Fortunately, the long recording time of all stations assure that sufficient low noise data could be identified to determine meaningful transfer

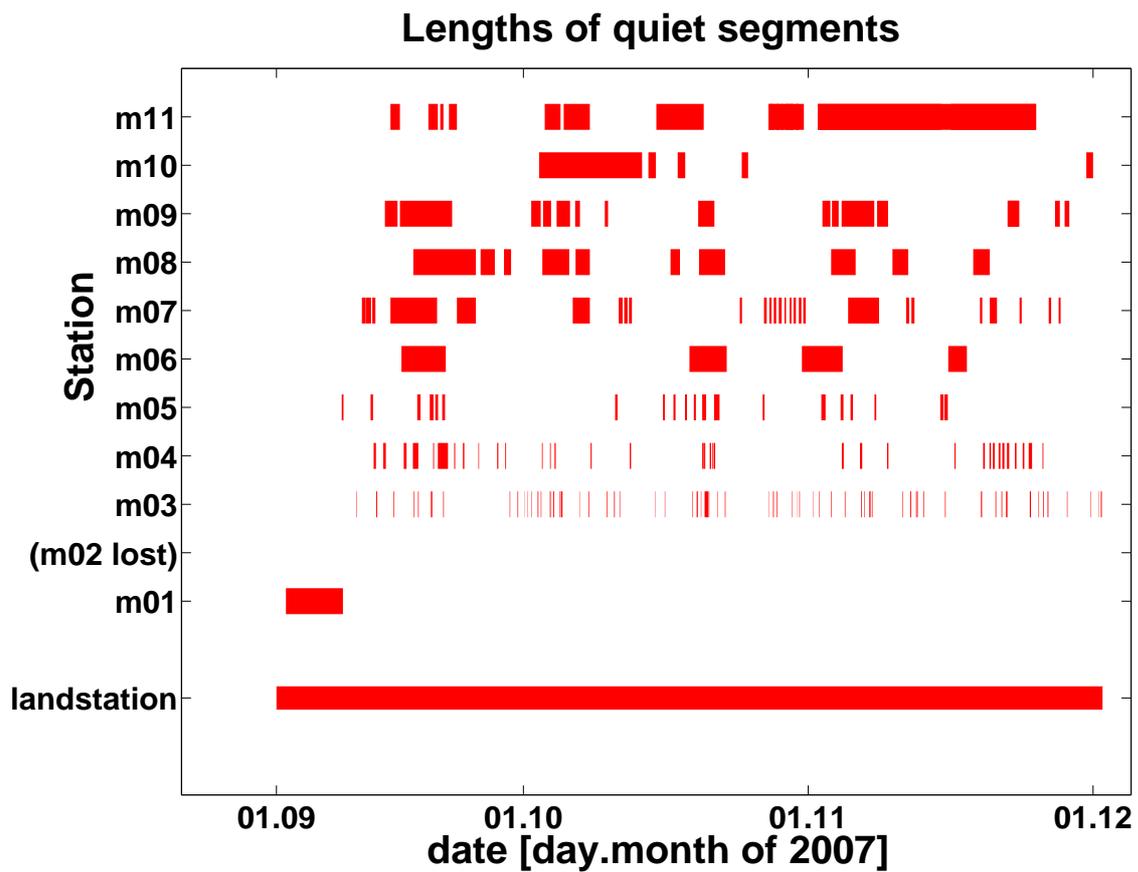


Figure 7.5: Overview of all identified quiet time sections of all marine stations which are used for data processing. Red bars denote periods of quiet time sections. As a comparison, land data has good quality during all measure time. Figure comes from Worzewski [2011].

functions. However, on the other hand, as a payment, transfer function at lower frequencies could not be determined due to the limited length of each quiet section, which means that the resistivity information of deeper subsurface can not be resolved.

The selected quiet segments of time series in each marine stations have been processed with BIRRP code [Chave and Thomson, 2004] to estimate the transfer function. Based on the results, a two-dimensional inversion model of electrical resistivity below Costa Rica subduction zone was obtained [Worzewski et al., 2011, Worzewski, 2011], as shown in Fig. 7.6. In the lower panel of the figure, six main resistivity anomalies are marked as A, B, C, D, E and G. Anomaly A is interpreted as the normal oceanic lithosphere in terms of electrical resistivity. The resistivity increases from the low resistive ($< 10 \Omega m$) water-saturated oceanic sediments to high resistive ($> 1000 \Omega m$) upper mantle. Then at depths beneath 40 km it decreases due to increasing temperatures. Low-resistive anomaly B located beneath ~ 80 km depth, could possibly be associated with an asthenosphere containing minor amounts of partial melts. Where the incoming plate bends, anomaly C was observed, which shows that the highly resistive oceanic lithosphere exhibiting a moderate reduction in electrical resistivity down to deep crustal and possibly upper-mantle regions ($> 1,000 \Omega m$ to $50 \Omega m$). This was interpreted as hydration of the oceanic crust by sea water penetrating bending-related faults, possibly down to upper mantle depth, promoting serpentinization of the upper mantle. Farther landward in the subduction process, hydration is superseded by dehydration. Low resistive anomaly D was detected at 12 - 15 km depth due to the dehydration of sediments. Underneath anomaly D a cluster of earthquakes defining the seismogenic zone is observed, showing that the transition from aseismic to seismogenic behavior correlates with a declining amount of fluid. Furthermore, at a distance of 30 km seaward from the volcanic arc and at depths about 20 - 30 km an accumulation of fluids (anomaly G) was revealed by the land data, and is most probably associated with fluid release from the oceanic crust or heating and dewatering of the upper part of the serpentinized mantle wedge. An effective dewatering through de-serpentinization at depths greater than ~ 100 km might be associated with a weakly resolved anomaly E. More detail about the model can be found in Worzewski et al. [2011], Worzewski [2011], Worzewski et al. [2012] and Brasse et al. [2009].

7.3 Processing the quiet segments of data

Since tilt-induced noise in marine MT time series is highly variable in time and is therefore non-stationary, it makes it difficult to process the data by using traditional Fourier based procedures which implicitly depend on the stationarity assumption on the data. Consequently, it is worth to try the newly developed method to process the noisy data, which has the ability to handle the non-stationarity as discussed before. In a first step, the method is tested with quiet segments and the results are compared to those obtained by BIRRP code.

According to the noise level, one can divided measured data into three groups: mostly quiet (m11), medium noise level (m06 - m10) and high noise level (m03 - m05). Several data sets at different noise levels are tested. Compared to other stations, station m11 data has the longest quiet segments, therefore, all the quiet segments at this station are firstly chosen to demonstrate the processing method below. The longest quiet electric and magnetic time series (23 days, from ~ 01 . Nov. 2012 to ~ 23 . Nov. 2012) for station 11 are shown in

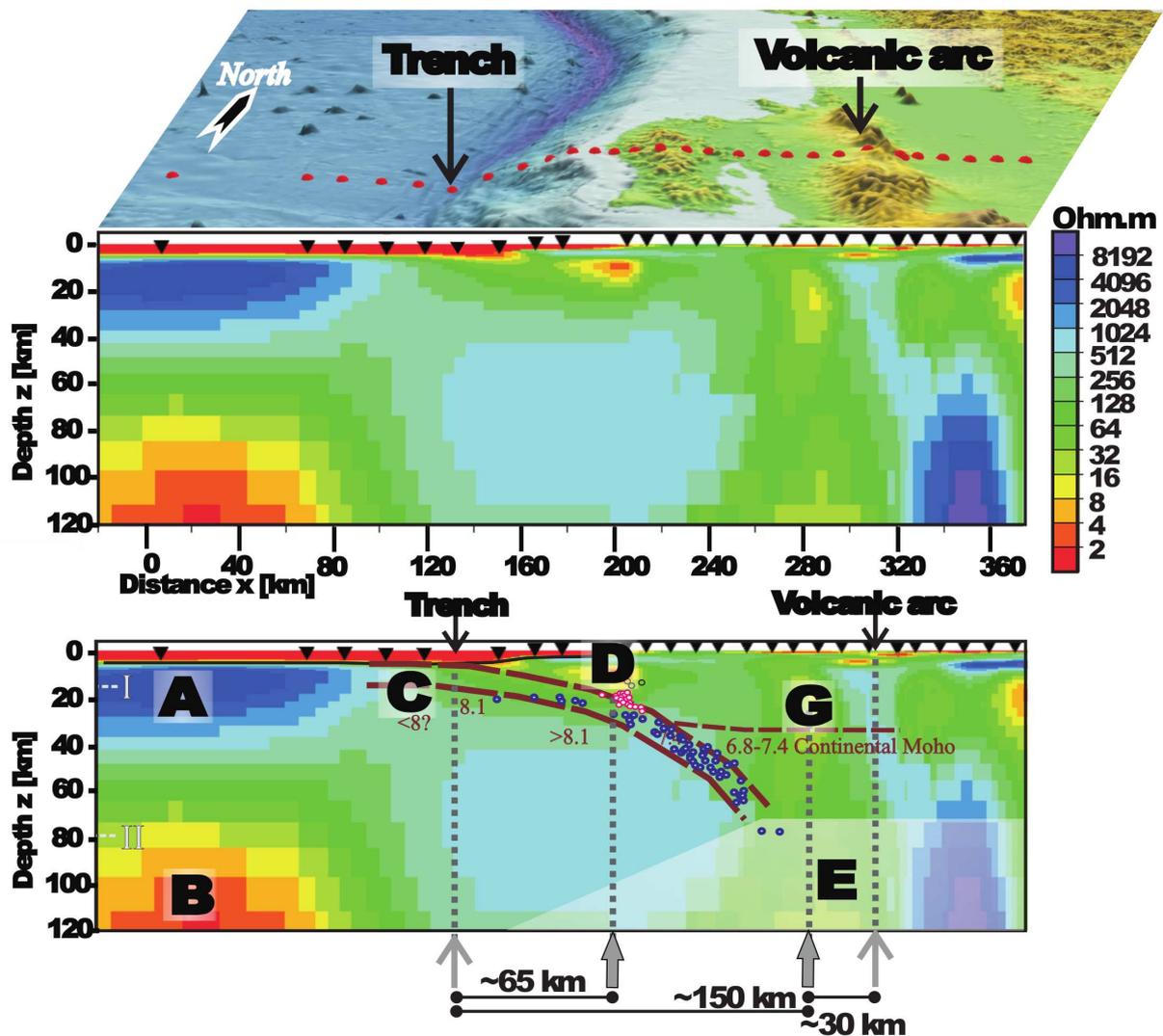


Figure 7.6: Upper panel: station map including bathymetry. Middle panel: two-dimensional inversion model of electrical resistivity below the Costa Rican subduction zone. The logarithmical color scale indicates the electrical resistivity ranging from conductive (least resistive) in red to highly resistive in blue. Lower panel: seismic information superimposed on inversion result—seismic boundaries (dashed brown) and velocities; blue circles, interplate earthquakes; pink circles, seismogenic-zone earthquakes from H.R.DeShon et al. [2006]. Figure comes from Worzewski et al. [2011].

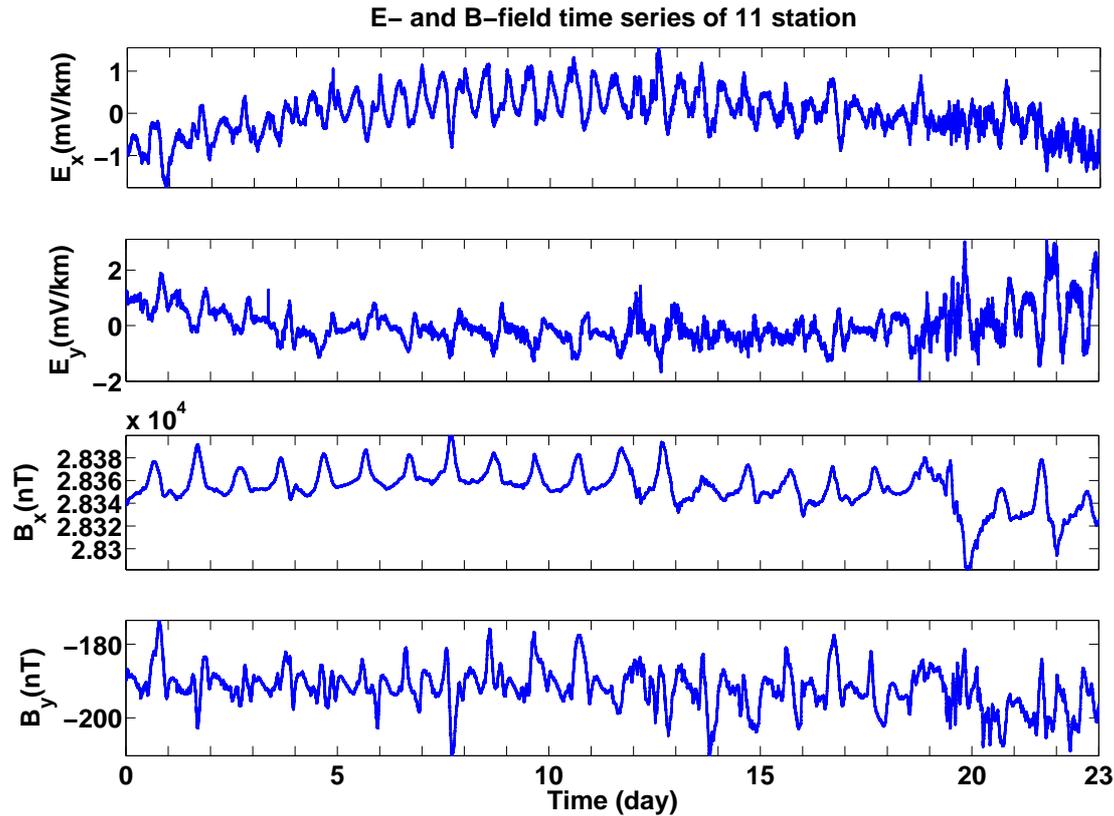


Figure 7.7: One long quiet section of E- and B-field time series at MT station 11 in Costa Rica (sampling frequency is 1 Hz).

Fig. 7.7, and the simultaneously measured magnetic time series for reference land station in Fig. 7.8.

From Fig. 7.5 one can see that the length of the quiet sections at station m11 varies between 8 hours and 23 days. In order to save computation time, longer quiet sections are cut into shorter sections of up to 24 hours length for the empirical mode decomposition. This results in a total of 35 sections. The frequency range between $5.6 \cdot 10^{-5}$ Hz and 0.2 Hz is divided into 24 bands with equal width on a logarithmic scale. Similar to the synthetic example discussed before, EMD method is applied to all electric and magnetic field components of the marine station and the magnetic field components of the reference land stations. From the calculated IMFs, instantaneous spectra are computed and their contributions are summed up using Eq. 4.61 for times at which the instantaneous spectra of all components fall in the same frequency band as outlined in previous chapter. The error bars are estimated using the bootstrap method and constitute a 95% confidence interval in the same manner as used for the synthetic data. It should be noticed that in real data test, a "numerical" reference was not used to reduce the bias caused by inaccuracies of the method as it was done in the synthetic data test. Instead, the data from the land station are used as remote reference signals. The numerical inaccuracies induced by the EMD method in both data from OBM T station and land station are mostly uncorrelated for different frequency bands such that the impact of these inaccuracies onto the impedance calculation are reduced. The whole

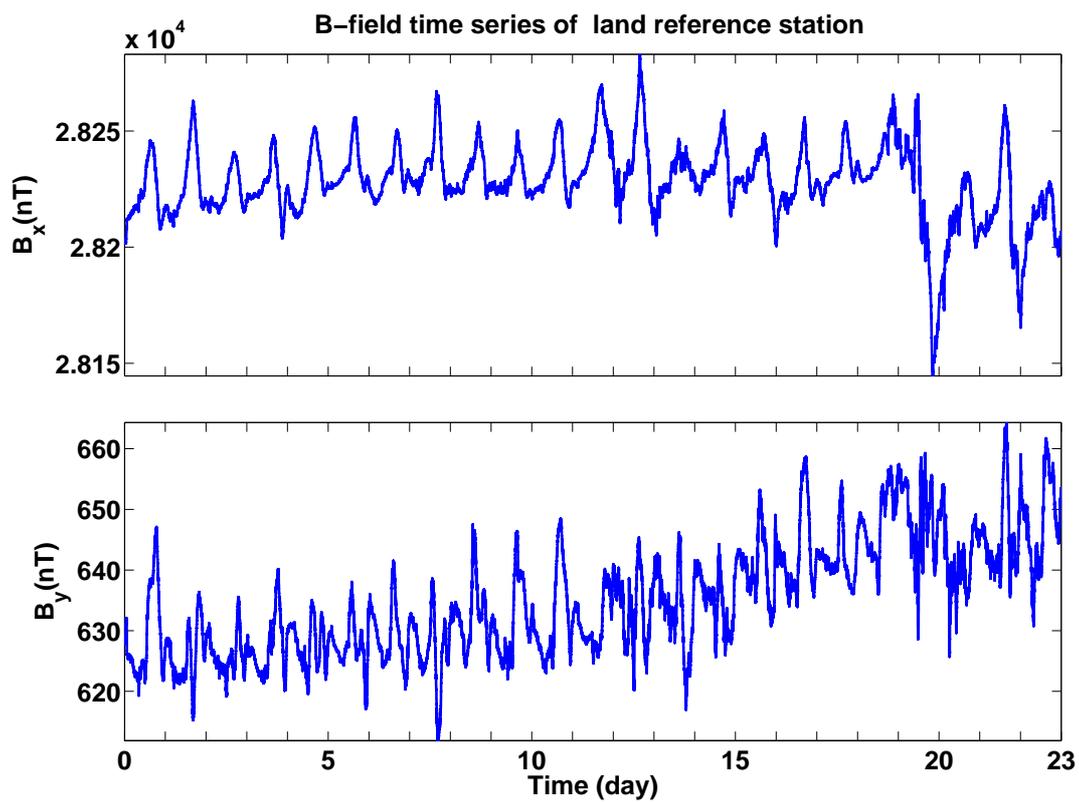


Figure 7.8: B-field time series from a land station that are used as the remote reference site. The time window shown here corresponds to the quiet section of station m11 shown in Fig. 7.7.

computation time is approximately 6 hours using a 3.2GHz computer with 2 CPUs. The obtained apparent resistivity and phase curves for the off-diagonal and diagonal components are shown in Fig. 7.9 together with the results obtained by BIRRP.

For medium noise level data sets, stations m10 and m09 are chosen, and for high noise level data station m05 was chosen. The lengths of the quiet sections that can be selected are decreasing from station m10 to m03 (towards direction to coast). At station m10, five quiet sections are selected. Their lengths vary from about 10 hours to 11 days. At station m09, 16 quiet sections are selected with a shortest length of 8 hours and a longest length of 5.5 days. At station m05, 20 quiet sections are selected with lengths only varying from 4.5 hours up to 14 hours. The results estimated from the quiet segments of stations m10, m09 and m05 are shown in Fig. 7.10, Fig. 7.11 and Fig. 7.12, respectively.

As all stations are located in deep water, due to the strong attenuation of the ocean the measured MT signals have very low signal to noise ratio in the high-frequency range and, therefore, it is difficult to estimate the impedance in the high-frequency range. However, compared to the results obtained with BIRRP at high frequencies, the results obtained from EMD are less downward biased and have relative small error bar. This coincides with the observations of non-stationary noise test in section 6.4.2, that EMD method has the advantage that it is less effected by the low signal to noise ratio. Furthermore the impedance can be estimated by EMD method for more than half decade deeper in low-frequency range compared to BIRRP. This is because that at least several cycles of signal are needed to estimate the spectrum by the BIRRP method, while in principal half a cycle in signal is necessary for EMD method, as discussed in chapter 5.

Furthermore, it can be observed that at station m05 (Fig. 7.11) the phase wanders through all quadrants in both the EMD and BIRRP results for certain frequencies. This kind of observations have been reported and explained as the marine coast effect. The effect can be regarded as an extreme bathymetric feature distorting the data, and it depends on several physical parameters such as distance from the coast, frequency, ocean depth and bulk resistivities. Worzewski et al. [2012] gave a more detailed discussion and definition about the coast effect based on numerical modelling of several ocean-coast scenarios.

Although neither any robust estimation nor other sophisticated fitting procedure as implemented in BIRRP are used in EMD method, the results obtained with EMD method from different noise level data sets are satisfactory and comparable to the results obtained with BIRRP. Apart from cutting out the noisy sections, the data sets have not been preprocessed or filtered in any way. High frequency range, results are even less downward biased by noise and have smaller error estimates. In the low frequency range, impedance values for more than half decade of frequencies can be obtained by EMD method. This is useful to obtain more information about the deeper structure of the investigation areas. These first results demonstrate that EMD technique is a suitable alternative processing method for magnetotelluric data.

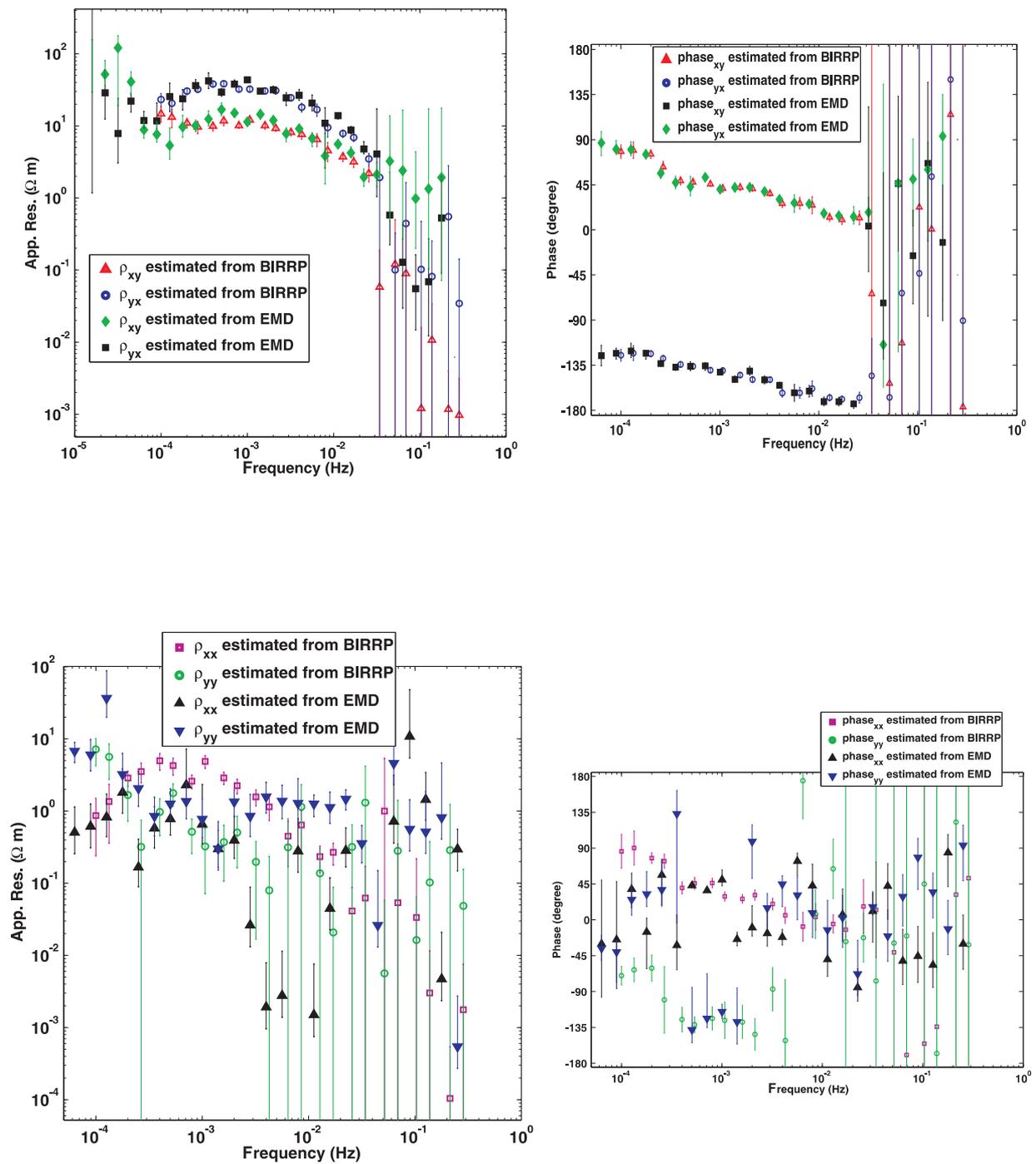


Figure 7.9: Station m11: comparison of results derived from BIRRP processing and EMD processing. Left panels: off-diagonal (upper) and diagonal (lower) apparent resistivities. Right panels: off-diagonal (upper) and diagonal (lower) phases.

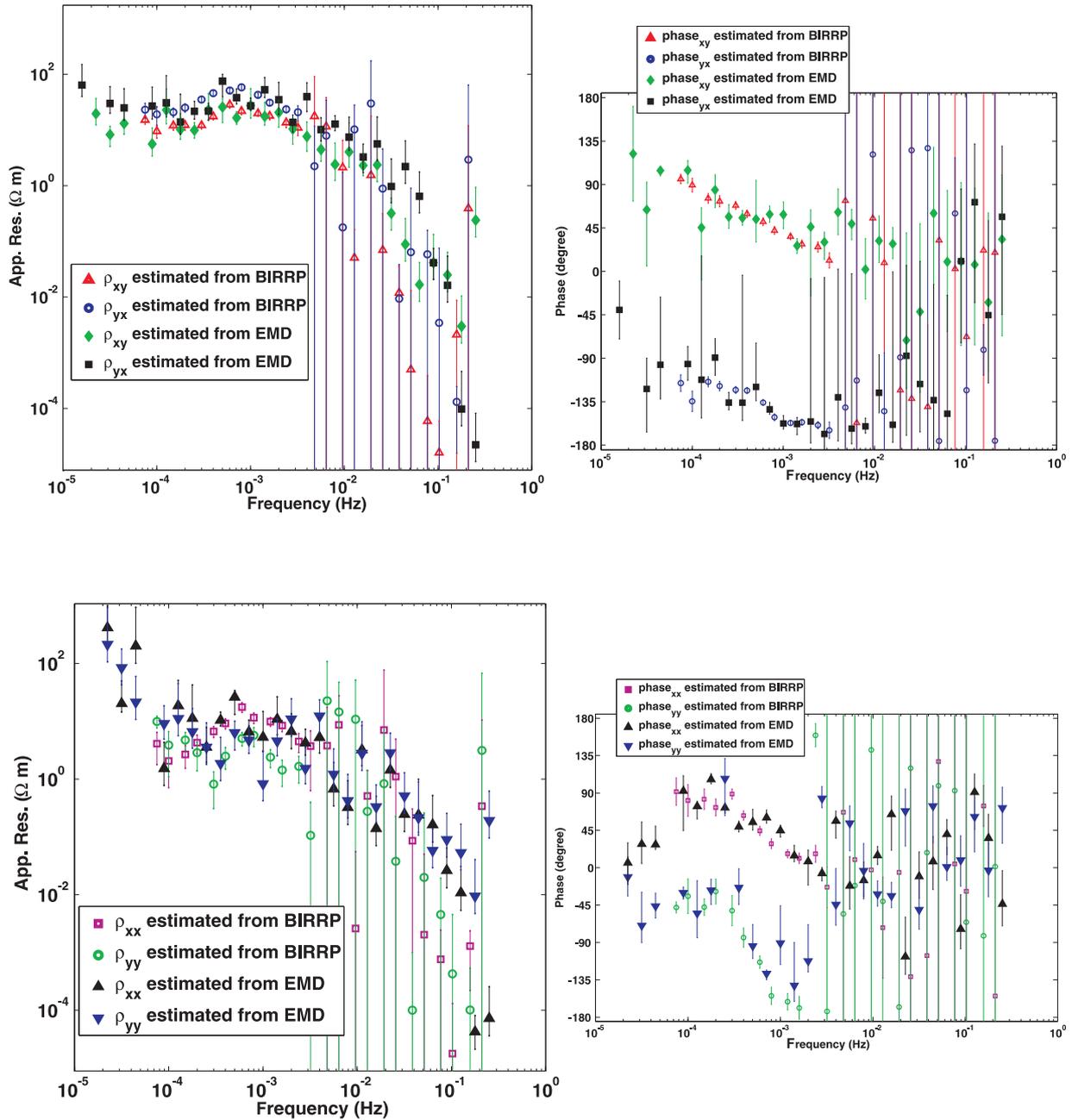


Figure 7.10: Station m10: comparison of results derived from BIRRP processing and EMD processing. Left panels: off-diagonal (upper) and diagonal (lower) apparent resistivities. Right panels: off-diagonal (upper) and diagonal (lower) phases.

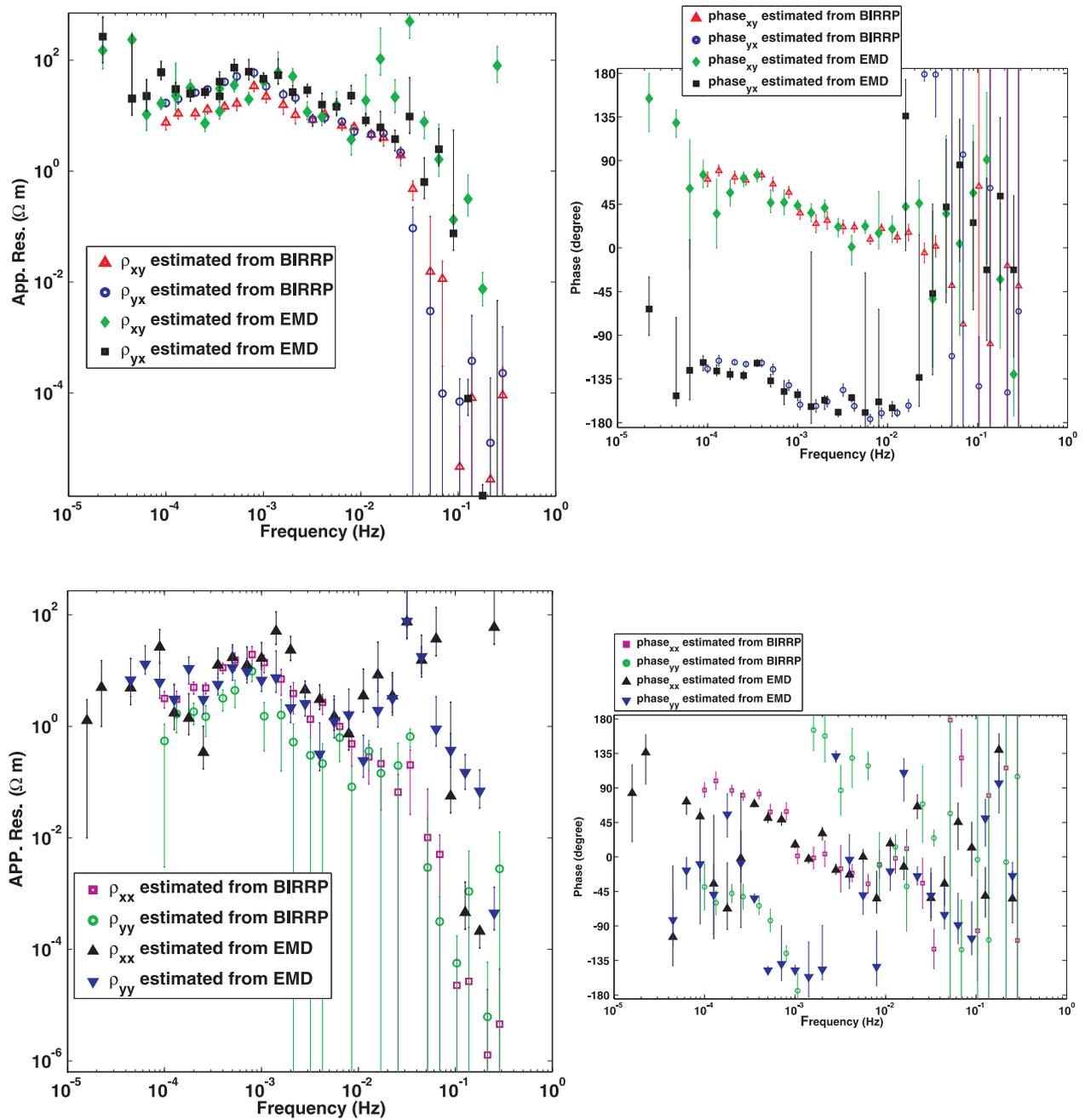


Figure 7.11: Station m09: comparison of results derived from BIRRP processing and EMD processing. Left panels: off-diagonal (upper) and diagonal (lower) apparent resistivities. Right panels: off-diagonal (upper) and diagonal (lower) phases.

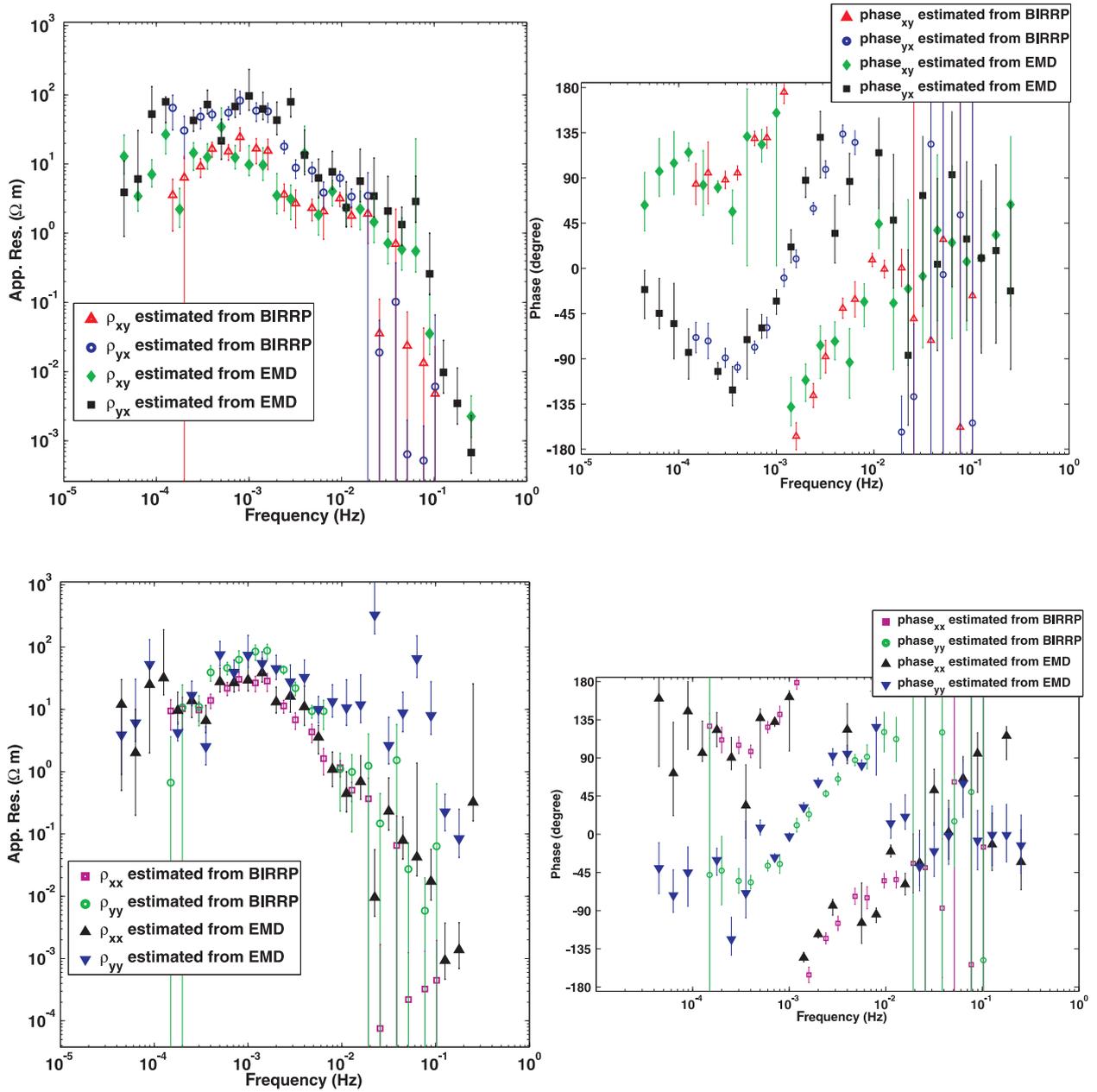


Figure 7.12: Station m05: comparison of results derived from BIRRP processing and EMD processing. Left panels: off-diagonal (upper) and diagonal (lower) apparent resistivities. Right panels: off-diagonal (upper) and diagonal (lower) phases.

7.4 Noise effect analysis and processing the noisy data

Until now, the data used to estimate the impedance tensor for both BIRRP and EMD method consist of relatively quiet segments. When noisy segments are used to process the data, the results estimated by sophisticated Fourier-based method are very poor [Worzewski, 2011]. This is mainly due to the reasons that a) the spectra of motion-induced noise spread in all frequency range, b) the Fourier spectra estimates, due to its formulation as an integral, seriously smear the motion-induced noise over frequency and c) there is no simple removal techniques for motion-induced noise exist to date since the relationship between tilt and both magnetic and electric field variations is actually quite complex. However, EMD method, as a suitable tool to deal with the non-stationary time series, has the advantage to reveal the local instantaneous information of the signal in both time and frequency domain. The revealed instantaneous information at the quiet segments will be rarely effected by the existence of the noisy segments, which coincides with the observations in the non-stationary noise test in section 6.4.2. There it was shown that in the lower frequency range EMD can still provide reliable impedance estimates. Moreover, EMD decomposes signal into different band-limited intrinsic oscillations (IMFs), which makes it possible to analyze the noise effect in each individual IMFs in both time and frequency range and thereby better assess its impact on impedance estimation.

Firstly, the noise effect in each individual IMFs is investigated by analyzing a noisy short section of signal. Fig. 7.13 shows one short noisy segments (about 15.3 hours) of magnetic field B_x and tilt measurements T_x at station m05 (see also Fig. 7.3). There are two periods with strong instrument motion ($> 0.1^\circ$) in this time section, which cause strong variations in the measured magnetic field data. To investigate their instantaneous spectra, both time series are decomposed into IMFs. Their 1st, 4th, 7th and 10th IMFs are shown in Fig. 7.14. One can easily observe that the tilt movement is present in all frequency bands (in all IMFs). However, comparing the amplitudes of the magnetic time series of the quiet sections with the noisy sections shows the influences of the tilt movement decreases from high frequencies (e.g., IMF 1) to low frequencies (e.g., IMF 10). This can be more easily observed in the three dimension shaded surface of the tilt and magnetic segments in Figs. 7.15 and 7.16.

By comparing Fig. 7.14 with Fig. 7.15, one can observe that in the two periods with strong instrument motion, magnetic field data are mainly effected by tilt movements in very high frequency range and that the effect decreases with descending frequency. This indicates that at low frequency range the data could still be used to estimate the impedance. Therefore, it is not necessary to discard the whole section of noisy data in preprocessing procedure like is usually done in BIRRP. With EMD one has now the chance to identify and discard noisy time-frequency sections (i.e., high frequency in noisy sections) of the data and use the remaining time-frequency sections to estimate the impedances.

Secondly, the noise impact on the impedance estimation is tested by using long sections of signals with many noise packages. At station m05, after discarding extremely noisy segments of data, three long sections varying from ~ 10 days to ~ 16 days are chosen to estimate the impedance tensor. One of them is shown in Fig. 7.17, and shows that the data, especially the B-field, exhibits highly non-stationary character.

From three long noisy sections, the estimated results of apparent resistivities and phases by

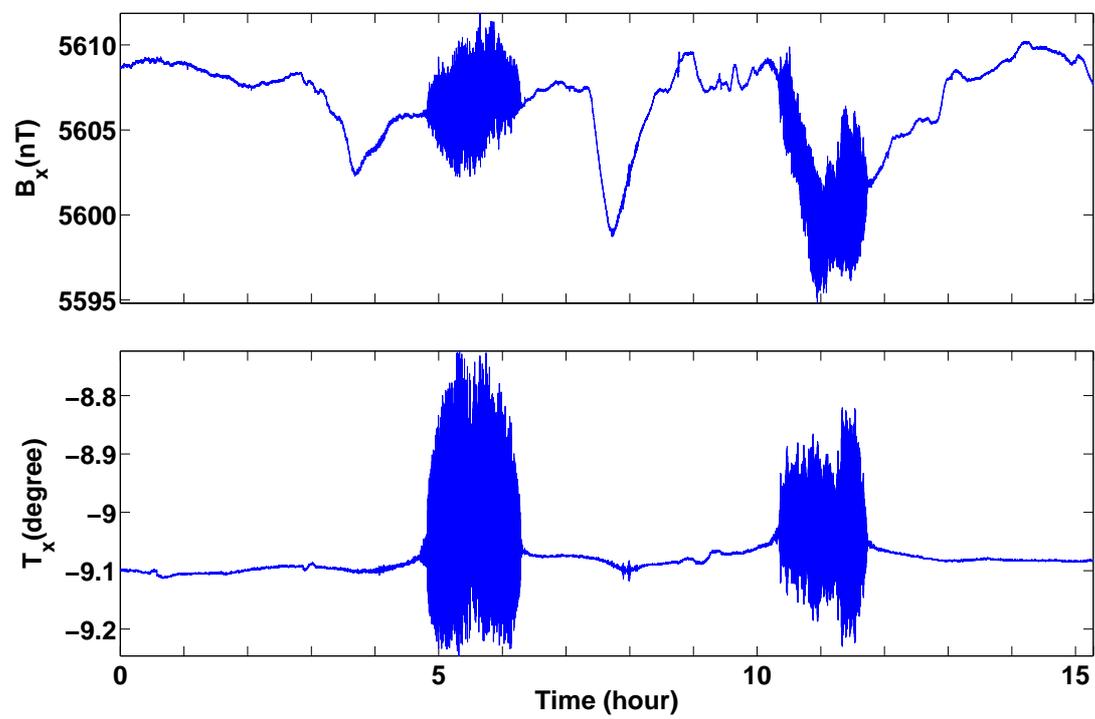


Figure 7.13: A short segment of x-component magnetic field and tilt data at station m05.

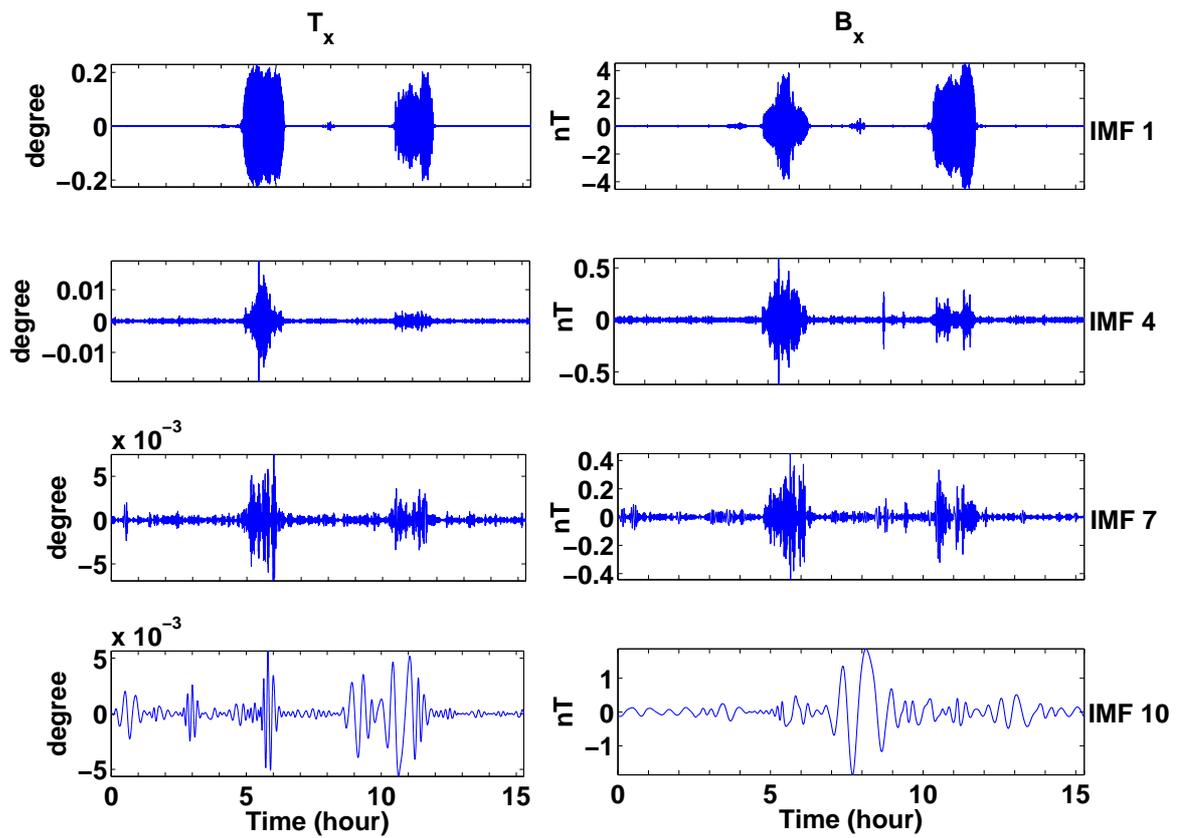


Figure 7.14: The 1st, 4th, 7th and 10th IMFs of the short segment of x-component magnetic field and tilt measurements at station m05 (Fig. 7.13).

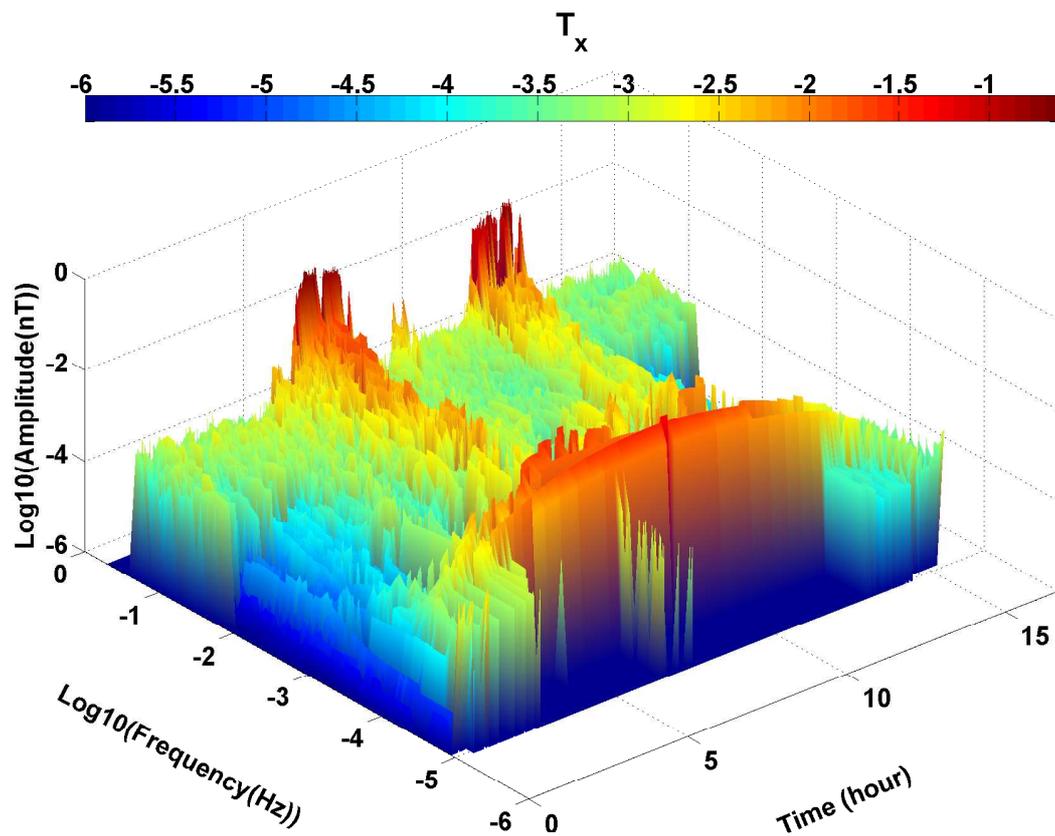


Figure 7.15: Three dimension shaded surface plot of instantaneous spectra of the tilt data (see lower panel in Fig. 7.13). The color code indicates the logarithmic scale of the instantaneous amplitudes.

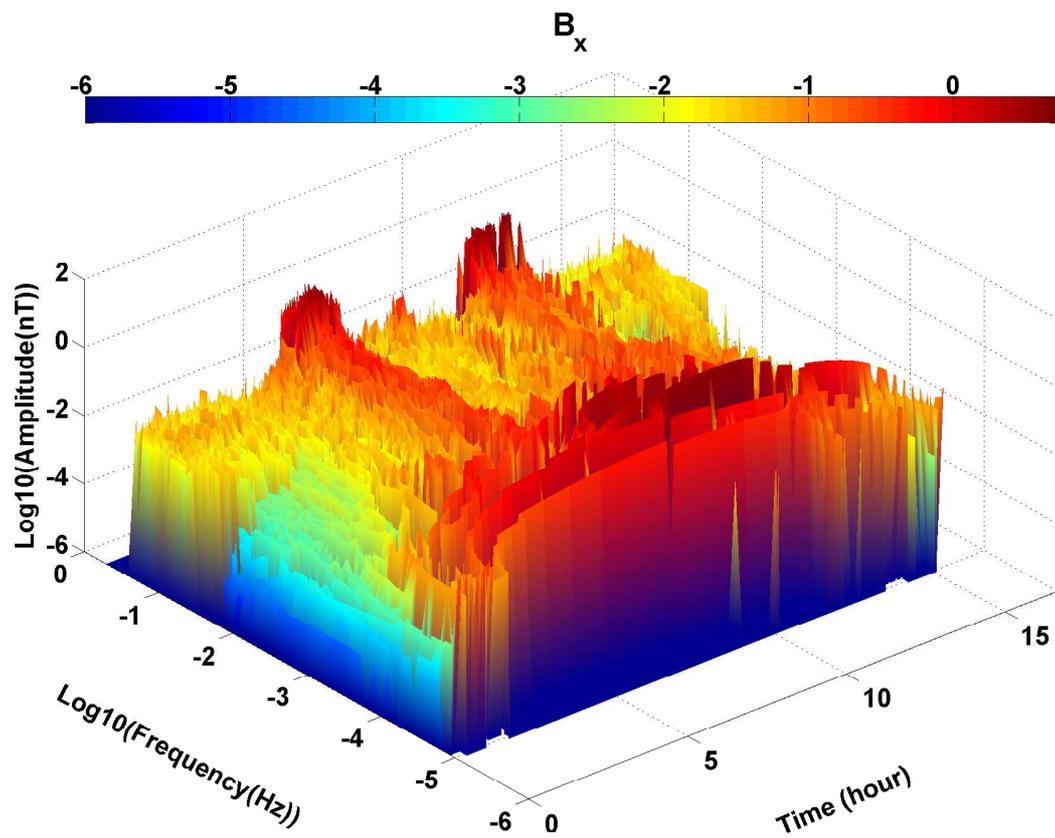


Figure 7.16: Three dimension shaded surface plot of instantaneous spectra of the magnetic data (see upper panel in Fig. 7.13). The color code indicates the logarithmic scale of the instantaneous amplitudes.

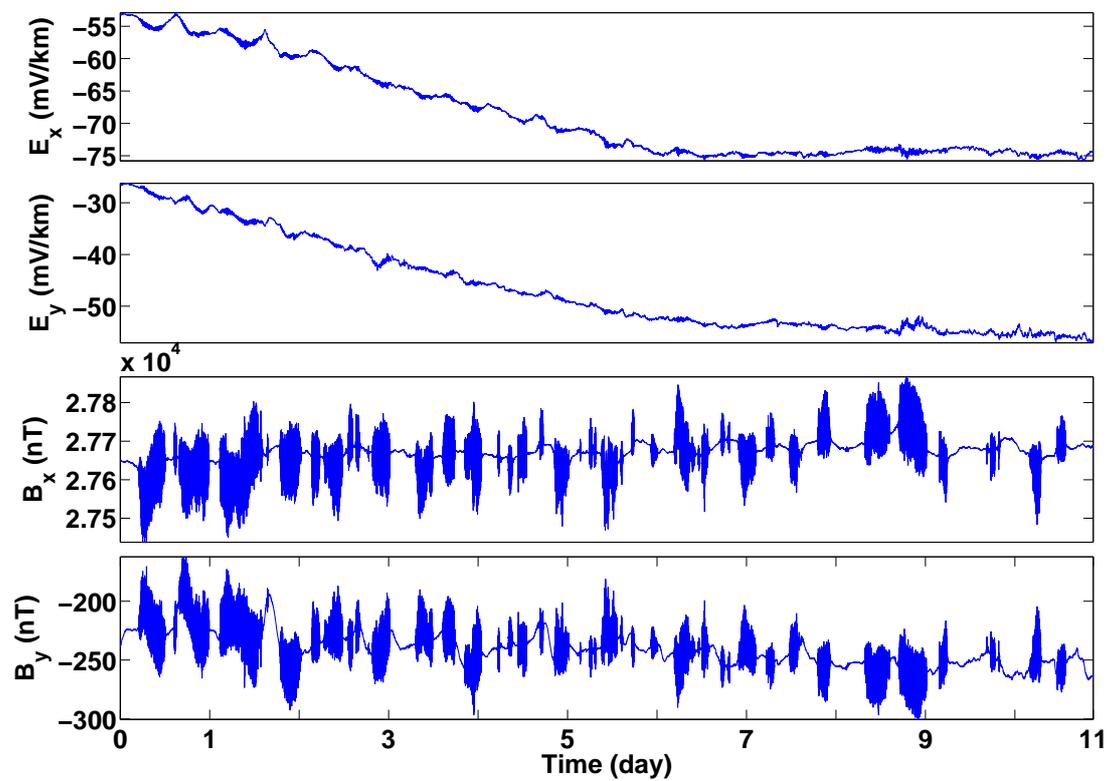


Figure 7.17: One of the long segment of data measured at station m05.

using BIRRP are shown in Fig. 7.18 (off-diagonal, red triangles for ρ_{xy} and green squares for ρ_{yx}) and Fig. 7.19 (diagonal, magenta squares for ρ_{xx} and green diamonds for ρ_{yy}). The corresponding results derived through BIRRP from short quiet sections (see Fig. 7.12) are also plotted in the two figures. Comparing the results from long noisy and short quiet sections, it can be observed that those from long noisy sections are biased at all frequencies, but especially for high frequencies. This is due to the fact that the spectra of motion induced noise which are estimated by Fourier transform are smeared. Furthermore, since the noisy segments are much longer than the quiet segments, one would expect that more impedance values at low frequencies could be estimated. However, since the coherency values between E- and B-field data at lower frequencies become very small due to noise bias in the Fourier spectra, no impedance values can be estimated by BIRRP.

The results estimated by using EMD for the noisy data are shown in Fig. 7.20 (just the off-diagonal elements of apparent resistivities are shown). As a comparison, the figure also includes the results from BIRRP for quiet segments. Both ρ_{xy} and ρ_{yx} values estimated by the EMD method are relative stable in the lower frequency range and are extended to more frequencies. This coincides the conclusion that the noise effects are less at lower frequencies. In high frequency range, the results from EMD are also less downward biased by noise than those from BIRRP and have relative narrow error bar. This indicates that, without any robust process as in BIRRP, EMD method can provide reliable estimate results even for noisy data. Although there are still some slight deviations in apparent resistivity curve, I believe that it could be improved by using a more sophisticated robust estimation technique.

7.5 Short summary and outlook

In this chapter, first the EMD method is utilized to estimate the impedance tensor from quiet segments of real marine MT data, and as from the synthetic data the results obtained are satisfactory. Compared to the results estimated using BIRRP, low frequency range results from EMD are reliable and high frequency range estimates are less downward biased and exhibit smaller error estimates. Then the instantaneous spectra of the noisy segments is analyzed by using EMD method and it is observed that the motion induced noise differs at different frequencies. The influence is higher in high frequency range and decreases with the descending of the frequency. This makes it possible in the future to use the instantaneous information in lower frequency part of noisy segment in the processing procedure rather than discarding the whole noisy segment. Furthermore, the impedance tensor is estimated by using EMD from the long segment of data containing many non-stationary noisy packages. The results are also stable at lower frequencies and less downward biased at higher frequencies and coincide with the observations in non-stationary noise test in synthetic time series.

All examples demonstrate that EMD method has the potential to handle noisy data and can be a new alternative method to process MT data. Compared to the robust Fourier-based processing methods, BIRRP, EMD method has some advantages by utilizing the instantaneous information of the data. BIRRP method works well for the short quiet segments of data but fails for the long segments of data with many non-stationary noise packages. The results derived from BIRRP for the short quiet segments are limited in bandwidth at low

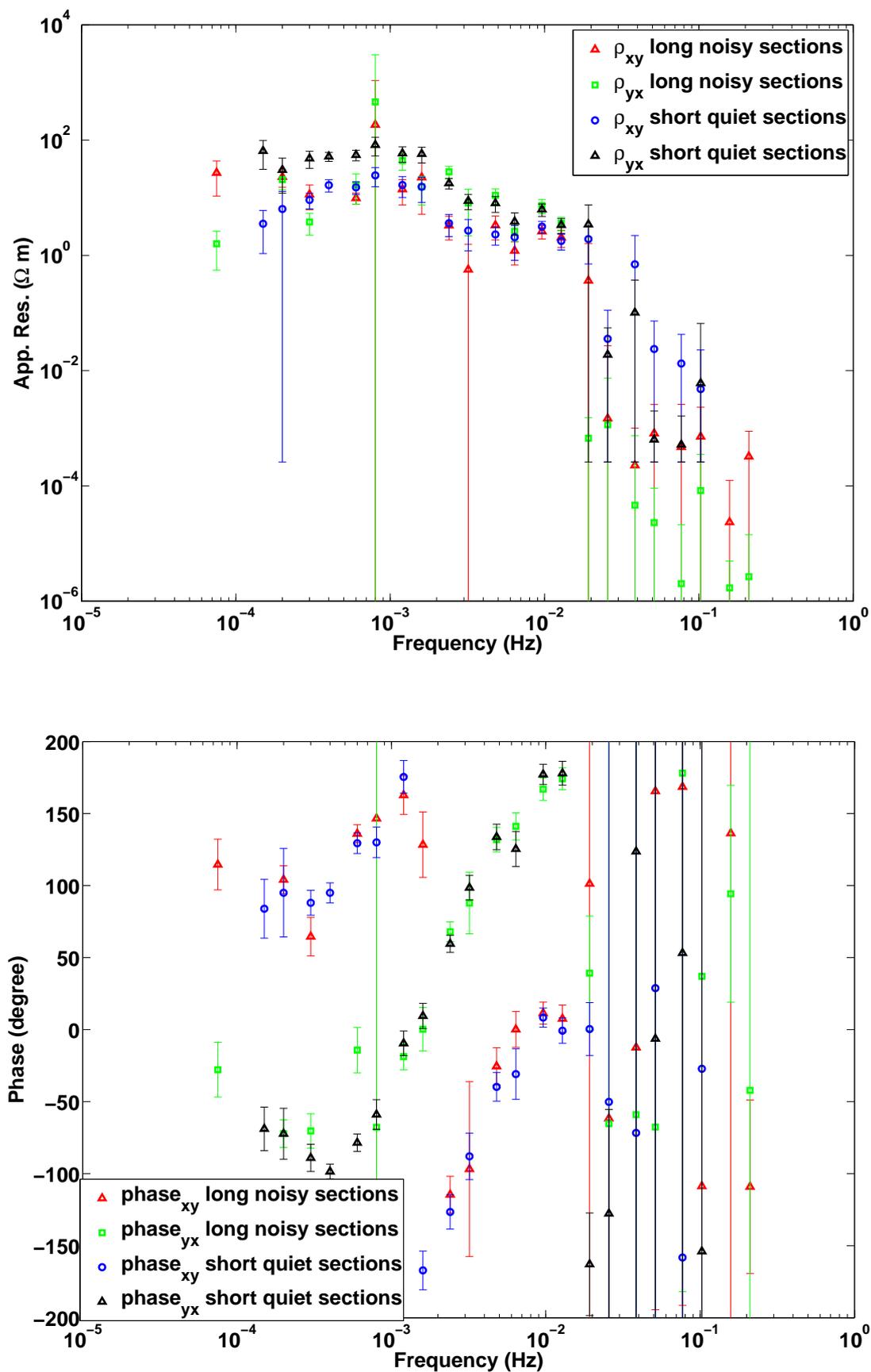


Figure 7.18: Station m05: estimated off-diagonal elements of the apparent resistivities (upper panel) and phases (lower panel) from three long noisy sections by using BIRRP. The results from corresponding short quiet sections estimated by using BIRRP are also plotted as a comparison.

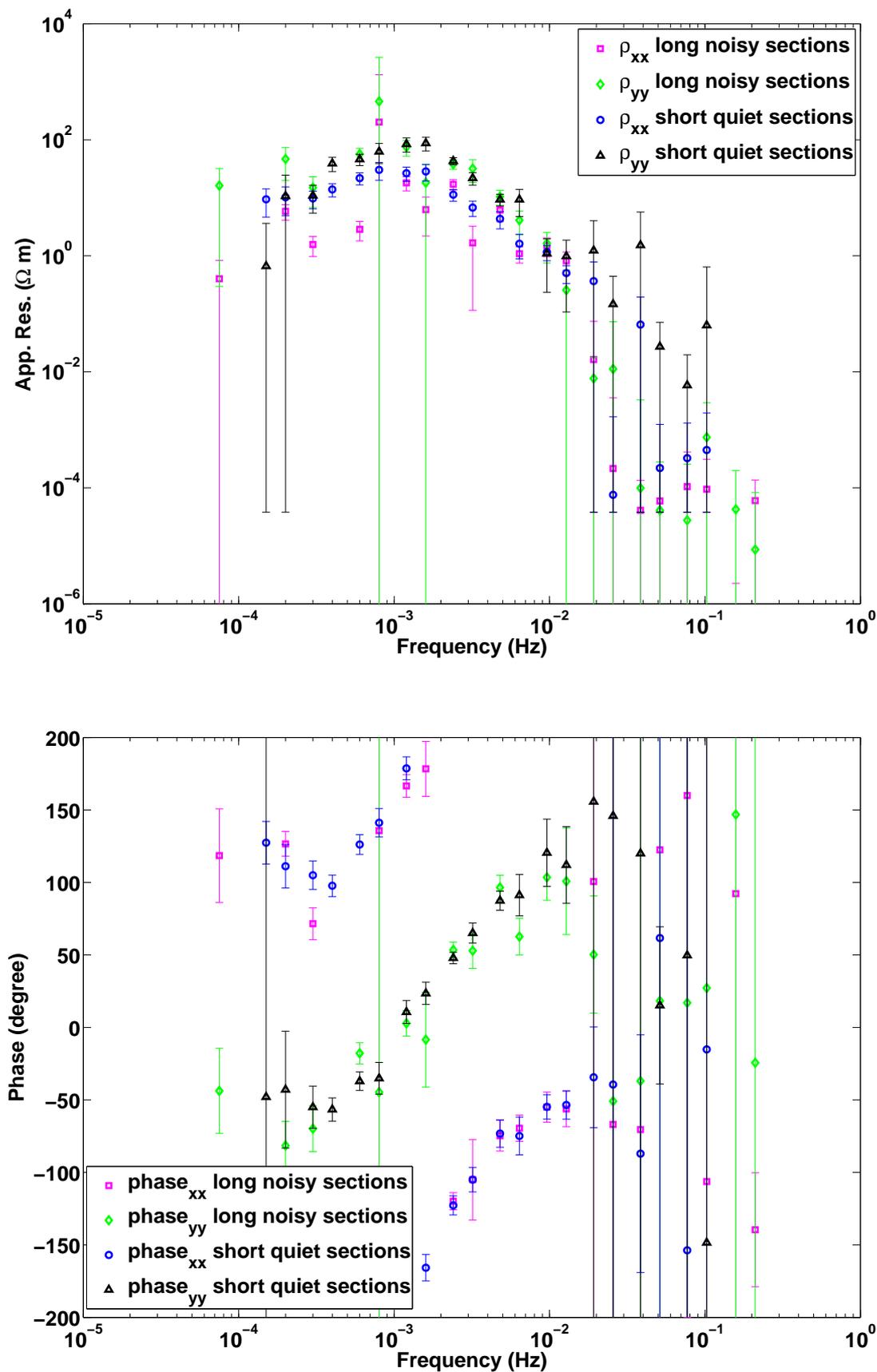


Figure 7.19: Station m05: estimated diagonal elements of the apparent resistivities (upper panel) and phases (lower panel) from three long noisy sections by using BIRRP. The results from corresponding short quiet sections estimated by using BIRRP are also plotted as a comparison.

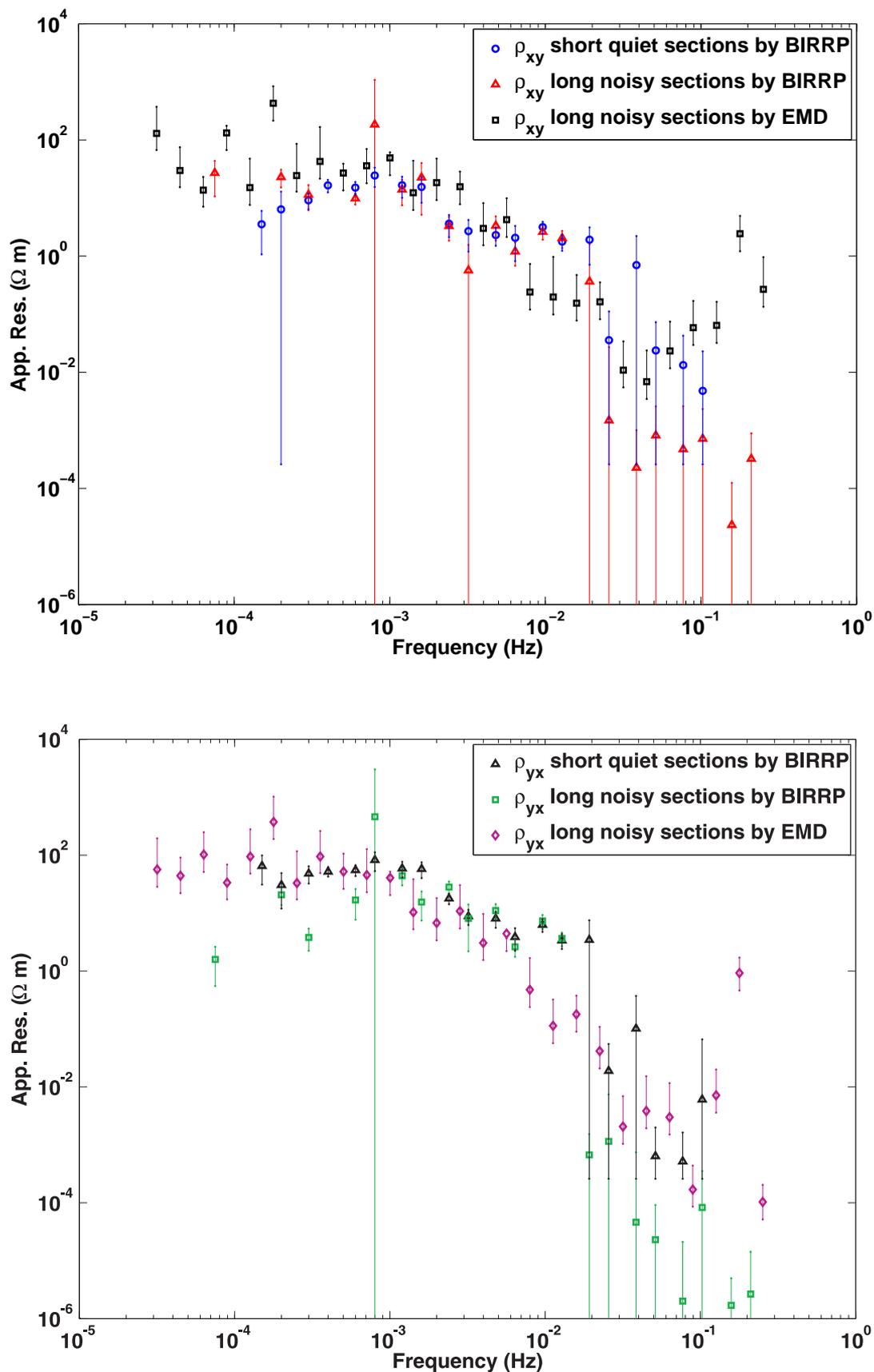


Figure 7.20: Estimated off-diagonal elements of the apparent resistivities (upper panel: ρ_{xy} , lower panel: ρ_{yx}) from three long noisy sections at station m05 by using EMD. The results estimated by using BIRRP are also plotted as a comparison.

frequencies due to the lengths of the segments and, are seriously downward biased and with large error estimates at high frequencies due to the low signal-noise ratios. Furthermore, even for long-high-noisy segments, BIRRP is also limited in bandwidth at low frequencies due to noise bias in the Fourier spectra. However, for short quiet segments EMD method gives reliable and stable impedance estimates even without any robust procedures as implemented in BIRRP. For long noisy segments, the results derived from EMD are also less biased at all frequencies and, exhibit smaller error estimates at high frequencies. EMD method gives more impedance estimates at low frequencies due to the fact that EMD does not need several cycles to estimate the spectrum at one frequency as in Fourier-based method. Furthermore, since the EMD method operates in a time-frequency domain, it potentially enables to control the data quality through analyzing the change of instantaneous parameters at different time and frequency intervals, which can be further investigated in a future work.

Chapter 8

Summary - Conclusion - Outlook

In this thesis I introduce a new method, Empirical Mode Decomposition (EMD), to process magnetotelluric (MT) data. EMD adaptively decomposes a time series into a number of mono-component basis functions (Intrinsic Mode Functions (IMFs)), which are nearly orthogonal from each other. From IMFs, physical meaningful instantaneous frequencies and spectra at every time sample of the data can be derived. Compared to the fixed harmonic basic functions of the Fourier transform, IFMs have the advantages of being data-driving and extracting local instantaneous information of the data, which makes EMD suitable to deal with non-stationary data and noisy data as encountered in MT data, especially in marine MT data. It is the first time to utilize EMD method and the instantaneous parameters to calculate estimates of the impedance tensor. The obtain results with both synthetic and real MT data are satisfactory. Since the method is not based on the assumption that the recorded electromagnetic fields are stationary as is case for a Fourier based technique, it is applicable on data with any degree of non-stationarity and has the potential to be an alternative MT processing technique when other methods fail.

By means of synthetic tests on a 1-D model, I demonstrated how this decomposition can be practically implemented. Results show that the apparent resistivities and phase estimates derived from this impedance tensor are generally stable. However, systematic downward shifts are present in the final apparent resistivity estimates. These downward shifts are associated with numerical noise that is induced by the decomposition and it is found that these shifts can be strongly reduced by applying a second EMD algorithm to create a numerical reference time series of the B-field or using a reference B-field time series. From additional synthetic tests it can be concluded that the EMD based procedure provides better results than simple Fourier derived impedance formulas in the presence of stochastic stationary and non-stationary noise. However, when compared to impedances determined using the robust processing method BIRRP, the results are slightly worse. This is to be expected as the implementation of EMD does not include robust processing techniques at the present stage.

For tests on real data sets, I chose several marine MT data sets with different levels of instrument motion-reduced noise which were measured on Costa Rica subduction zone. First, through analyzing the instantaneous spectra of a noisy data segment by the EMD method, it is revealed that the motion-induced noise is present over the whole frequency range, however, it differs for different frequency ranges. The noise is largest at high frequencies and decreases with the descending frequency. Then I test impedance estimates from the EMD method and BIRRP by using relative quiet segments of the data sets. The results derived

by two methods are very similar, and in high frequency range the results derived by EMD are even less biased by noise and exhibit smaller error bars, and furthermore the impedance estimates are extended to lower frequencies. This indicates that EMD method is suitable for practical application. At last, the impedance estimates from EMD method and BIRRP are tested to one data set with high noise level, the results derived by both methods are similar in low frequencies and, in high frequencies the results from EMD are less biased by noise and with smaller error bars.

All tests on synthetic and real MT data indicate that EMD method, as an alternative method in MT data processing beyond Fourier-based methods, does have many advantages, especially in processing the non-stationary and noisy marine MT data. Although EMD does not include at the present stage robust processing techniques, it can already provide comparable estimations to robust processing method BIRRP. Furthermore, impedances estimated by EMD contain more information at lower frequencies, with which deeper Earth resistivity structures can be imaged. They are also less biased by noise at higher frequencies and exhibit smaller error bars, with which the resolution of the shallow resistivity structures can be improved.

Based on the results and analysis, one can see potential to further improve EMD analysis, such as to get rid of some aspects of numerical noise (e.g., the boundary extension, interpolation method, determination of stop criterion and calculation of instantaneous frequency.) and to combine the EMD method with robust processing procedures such as those employed in BIRRP to make impedance estimates less sensitive to noise and unusual observations (outliers and leverage points). These will take this research from a proof of principle to a stage, at which it can be used in practise. To make the new method robust, the determination of the outliers will be challenging since the new methodology operates in a time-frequency domain rather than in a pure frequency domain as in Fourier-based methods. At last, the coherence between the electric and magnetic field data should also be taken into consider, some threshold values need to be determined to discard data points with lower coherence values.

Another attractive feature of the proposed method is that one can potentially remove noise before processing by excluding some IMFs that were identified as noisy. This would still allow one to use the remaining data without having to discard the entire noisy section as did in conventional approaches or filter the data prior to impedance estimation. Obviously, this can only work as long as the noise is limited to a small amount of IMFs and can be reliably identified. Future experiments will investigate if such a noise removal procedure can be performed effectively and in an automated manner. One of the practical implications for such a noise removal procedure could be in marine MT processing, where the recorded magnetic field is often contaminated with high frequency noise caused by instrument motion due to deep water or tidal currents.

Bibliography

- M. H. Ackroyd. Instantaneous and time varying spectra – an introduction. *J. Radio and Electron. Eng.*, 39:no. 3, 1970.
- M. Alexandrescu, D. Gilbert, G. Hulot, J.-L. Le Mouel, and G. Saracco. Detection of geomagnetic jerks using wavelet analysis. *Journal of Geophysical Research*, 100:12557–12572, 1995.
- S. Allerton, H. U. Worm, and L. B. Stokking. Paleomagnetic and rock magnetic properties of hole 896A. *Proc. Ocean Drill. Program Sci. Results*, 148:217226, 1996.
- G. Archie. The electrical resistivity log as an aid in determining some reservoir characteristics. *Petroleum Transactions of AIME*, 146:5456, 1942.
- A. Avdeeva, D. Avdeev, and M. Jegen. Detecting a salt dome overhang with magnetotellurics: 3D inversion methodology and synthetic model studies. *GEOPHYSICS*, 77(4):E251–E263., 2012.
- K. Baba. Electrical structure in marine tectonic settings. *Surveys in Geophysics*, 26:701–731, 2005.
- K. Bahr. Interpretation of the magnetotelluric impedance tensor: regional induction and local telluric distortion. *Journal of Geophysics*, 62:119–127, 1988.
- U. Barckhausen, C. R. Ranero, R. von Huene, S. C. Cande, and H. A. Roeser. Revised tectonic boundaries in the Cocos Plate off Costa Rica: Implications for the segmentation of the convergent margin and for plate tectonic models. *Jornnal of geophysical research*, 106 (NO. B9):19,207–19,220, 2001.
- A. E. Barnes. Short note: The calculation of instantaneous frequency and instantaneous bandwidth. *Geophysics*, 57, No. 12:1520–1524, 1992.
- M. S. Bartlett. Some remarks on the analysis of time series. *Biometrika*, 54:25–38, 1967.
- E. Bedrosian. A product theorem for Hilbert Transforms. *Proceedings of the IEEE*, 51, Issue:5:868–869, 1963.
- M. N. Berdichevsky, I. A. Bezruk, and O. M. Chinavera. Magnetotelluric sounding with the use of mathematical filters (in Russian). *Izv. Akad. Nauk SSSR Fiz. Zeml.*, 3:76–92, 1973.
- J. Bialas and E. Flueh. Ocean bottom seismometers. *Sea Technology*, 40 (4):41 – 46, 1999.
- B. Boashash. Estimating and Interpreting The Instantaneous Frequency of a Signal–Part 1: Fundamentals. *Proceedings of the IEEE*, 80, No. 4:520–538, 1992a.

- B. Boashash. Estimating and Interpreting The Instantaneous Frequency of a Signal—Part 2: Algorithms and Applications. *Proceedings of the IEEE*, 80, No. 4:540–568, 1992b.
- H. Brasse and W. Soyer. A magnetotelluric study in the Southern Chilean Andes. *Geophysical Research Letters*, 28 (19):3757–3760, 2001.
- H. Brasse, G. Kapinos, L. Mtschard, G. Alvarado, T. Worzewski, and M. Jegen. Deep electrical resistivity structure of northwestern Costa Rica. *Geophysical Research Letters*, 36:doi:10.1029/2008GL036397, 2009.
- J. Byerlee. Model for episodic flow of high-pressure water in fault zones before earthquakes. *Geology*, 21:303–306, 1993.
- L. Cagniard. Basic theory of the magneto-telluric method of geophysical prospecting. *Geophysics*, 18:605–635, 1953.
- Carbotte and K. Macdonald. East Pacific Rise 8-10 31°N: Evolution of ridge segments and discontinuities from SeaMARC II and three-dimensional magnetic studies. *Journal of Geophysical Research*, 97:69596982, 1992.
- J. Carson and T. Fry. Variable frequency electric circuit theory with application to the theory of frequency modulation. *Bell System Tech. J.*, 16:513 – 540, 1937.
- I. J. Chant and L. M. Hastie. Time-frequency analysis of magnetotelluric data. *Geophysical Journal International*, 111:399–413, 1992.
- I. J. Chant and L. M. Hastie. A comparison of the stability of stationary and non-stationary magnetotelluric analysis methods. *Geophysical Journal International*, 115:1143–1147, 1993.
- A. Chave, A. Jones, P. Weidelt, R. Evans, A. Viljanen, C. Weiss, W. Rodi, R. Mackie, and I. F. and. *The Magnetotelluric Method: Theory and Practice*. Cambridge University Press, 2012.
- A. D. Chave and P. Lezaeta. The statistical distribution of magnetotelluric apparent resistivity and phase estimates. *Geophys. J. Int.*, 171:127–132, 2007.
- A. D. Chave and D. J. Thomson. Some Comments on Magnetotelluric Response Function Estimation. *Journal of Geophysical Research*, 94 (No. B10):14, 215 – 14, 225, 1989.
- A. D. Chave and D. J. Thomson. A bounded influence regression estimator based on the statistics of the hat matrix. *Royal Statistical Society*, 52:307–322, 2003.
- A. D. Chave and D. J. Thomson. Bounded influence magnetotelluric response function estimation. *Geophys. J. Int.*, 157:988–1006, 2004.
- A. D. Chave, D. J. Thomson, and M. E. Ander. On the robust estimation of power spectra, coherences, and transfer functions. *Journal of Geophysical Research*, 92 (No. B1):633–648, 1987.
- J. Chen, B. Heincke, M. Jegen, and M. Moorkamp. Using empirical mode decomposition to process marine magnetotelluric data. *Geophys. J. Int.*, 190:293–309, 2012.

- Q. Chen, N. E. Huang, S. Riemenschneider, and Y. Xu. A B-spline approach for empirical mode decompositions. *Advances in computational mathematics*, 24:171–195, 2006.
- J. F. Claerbout. *Fundamentals of geophysical data processing: with applications to petroleum prospecting*. McGraw-Hill Book Co., 1976.
- L. Cohen. Generalised phase-space distributions. *J. Math. Phys.*, 7:781–786, 1966.
- L. Cohen. Time-Frequency Distributions-A Review. *proceedings of the IEEE*, 77, NO. 7:941–981, 1989.
- L. Cohen. What is a multicomponent signal? In *in Proc. ICASSP, San Francisco, CA, 1992, vol. 5, pp. 113116.*, 1992.
- L. Cohen. *Time-frequency analysis*. Prentice Hall, Englewood Cliffs, NJ, 1995.
- L. Cohen and C. Lee. Instantaneous frequency, its standard deviation and multicomponent signals. *Advanced Algorithms and Architectures for Signal Processing III, Franklin T. Luk, Ed., Proc. SPIE*, 975:186–208, 1988.
- S. Constable, T. J. Shankland, and A. Duba. The electrical conductivity of an isotropic olivine mantle. *Journal of Geophysical Research*, 97:3397, 1992.
- N. Delprat. Global Frequency Modulation Laws Extraction from the Gabor Transform of a Signal: A First Study of the Interacting Components Case. *IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING*, 5, No.1:64–71, January 1997.
- C. DeMets. A new estimate for present-day Cocos-Caribbean plate motion: Implications for slip along the Central American volcanic arc. *Geophysical research letters*, 28:4043–4046, 2001.
- Y.-J. Deng, W. Wang, C.-C. Qian, Z. Wang, and D.-J. Dai. Boundary-processing-technique in EMD method and Hilbert transform. *Chinese Science Bulletin*, 46:954–960, 2001.
- H. DeShon and S. Y. Schwartz. Evidence for serpentinization of the forearc mantle wedge along the Nicoya Peninsula, Costa Rica. *Geophysical research letters*, 31:L21611, 1–4, 2004.
- D. G. Duffy. The application of Hilbert-Huang transforms to meteorological data sets. *J. Atmos. Oceanic Technol.*, 21:599–611, 2004.
- B. Efron. Bootstrap methods: another look at the jackknife. *The Annals of statistics*, Vol. 7, No. 1:1–26, 1979.
- G. D. Egbert. Noncausality of the discrete-time magnetotelluric impulse-response. *Geophysics*, 57:1354–1358, 1992.
- G. D. Egbert. Robust multiple-station magnetotelluric data processing. *Geophysical Journal International*, 130:475–496, 1997.
- G. D. Egbert and J. R. Booker. Robust estimation of geomagnetic transfer functions. *Geophysical Journal of the Royal Astronomical Society*, 87:173–194, 1986.

- G. D. Egbert and J. R. Booker. Very long period magnetotellurics at Tucson observatory: Implications for mantle conductivity. *Journal of Geophysical Research*, 97:15099, 1992.
- G. D. Egbert and D. Livelybrooks. Single station magnetotelluric impedance estimation: coherence weighting and regression M-estimate. *Geophysics*, 61:964–970, 1996.
- G. D. Egbert, M. Eisel, O. S. Boyd, and H. F. Morrison. DC trains and PC3s: Source effects in mid-latitude geomagnetic transfer functions. *Geophysical research letters*, 124:25–28, 2000.
- D. V. Fitterman and C. Yin. Effect of bird maneuver on frequency-domain helicopter EM response. *Geophysics*, 69:1203–1215, 2004.
- P. Flandrin. Empirical mode decompositions as data-driven wavelet-like expansions. *International Journal of Wavelets, Multiresolution and Information Processing*, 2, No.4:1–20, 2004.
- P. Flandrin, G. Rilling, and P. Goncalves. On emperical mode decomposition as a filter bank. *IEEE Signal Processing Letters*, 11, No. 2:112 – 114, 2004.
- D. Gabor. Theory of communication. *Journal of the Institution of Electrical Engineers*, 93 (26):429 – 457, 1946.
- T. Gamble, W. Goubau, and J. Clarke. Magnetotellurics with a remote magnetic reference. *Geophysics*, 44:53–68, 1979a.
- T. Gamble, W. Goubau, and J. Clarke. Error analysis for remote reference magnetotellurics. *Geophysics*, 44:959–968, 1979b.
- X. Garcia and A. Jones. Robust processing of magentotelluric data in the AMT dead band using the continuous wavelet tranform. *Geophysics*, 73 No.6:F223–F234, 2008.
- X. Garcia, A. D. Chave, and A. D. Jones. Robust processing of magnetotelluric data from the auroral zone. *J. Geomag. Geoelectr.*, 49:1451–1468, 1997.
- W. Goubau, T. Gamble, and J. Clarke. Magnetotelluric data analysis: removal of bias. *Geophysics*, 43:1157–1166, 1978.
- I. Grevemeyer, N. Kaul, J. L. Diaz-Naveas, H. W. Villinger, C. R. Raneroa, and C. Reichert. Heat flow and bending-related faulting at subduction trenches: Case studies offshore of Nicaragua and Central Chile. *Earth and Planetary Science Letters*, 236:238–248, 2005.
- I. Grevemeyer, C. R. Ranero, E. R. Flueh, D. Kläschen, and J. Bialas. Passive and active seismological study of bending-related faulting and mantle serpentinitization at the Middle America trench. *Earth and Planetary Science Letters*, 258:528–542, 2007.
- L. Griffiths. Rapid measurement of digital instantaneous frequency. *IEEE Trans. on Acoust. Speech, Signal Processing*, 23, 1975.
- M. S. Gupta. Definition of instantaneous frequency and frequency measurability. *Amer. J. Phys.*, 43:no. 12, 1975.

- B. Hacker, S. M. Peacock, G. A. Abers, and S. D. Holloway. Subduction factory 2. Are intermediate-depth earthquakes in subducting slabs linked to metamorphic dehydration reactions? *Journal of Geophysical Research*, 108:1–11, 2003.
- F. R. Hampel. The influence curve and its role in robust estimation. *J. Am. Stat. Assoc.*, 69:383 – 393, 1974.
- R. F. Hampel, M. E. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. *Robust Statistics. The Approach Based on Influence Functions*,. Wiley, New York, 1986.
- B. Heincke, M. Jegen, M. Moorkamp, J. Chen, and R. W. Hobbs. Adaptive coupling strategy for simultaneous joint inversions that use petrophysical information as constraints. In *SEG Technical Program Expanded Abstracts, Denver, USA*, volume 29, pages 2805–2809, 2010.
- G. Heinson and S. Constable. The electrical conductivity of the oceanic upper mantle. *Geophysical Journal International*, 110:159–179, 1992.
- W. Heise, H. M. Bibby, T. G. Caldwell, S. C. Bannister, Y. Ogawa, S. Takakura, and T. Uchida. Melt distribution beneath a young continental rift: the Taupo Volcanic Zone, New Zealand. *Geophysical Research Letters*, 34:L14313, 2007.
- D. C. Hoaglin and R. E. Welsch. The hat matrix in regression and ANOVA. *Am. Stat.*, 32: 17 – 22, 1978.
- G. M. Hoversten, H. F. Morrison, and S. C. Constable. Marine magnetotellurics for petroleum exploration, Part II: Numerical analysis of subsalt resolution. *Geophysics*, 63:826840, 1998.
- H.R.DeShon, S. Y. Schwartz, A. V. Newman, V. Gonzalez, M. Protti, L. M. Dorman, T. H. Dixon, D. E. Sampson, and E. R. Flueh. Seismogenic zone structure beneath the Nicoya Peninsula, Costa Rica, from three-dimensional local earthquake P- and S-wave tomography. *Geophys. J. Int.*, 164:109–124, 2006.
- N. E. Huang and N. O. Attoh-Okine. *Hilbert-Huang Transforms in Engineering*. CRC Press, Boca Raton, Fla., 2005.
- N. E. Huang and S. S. P. Shen. *Hilbert-Huang Transform and Its Applications*. World Sci., Singapore, 2005.
- N. E. Huang and Z. Wu. A review on Hilbert-Huang transform: method and its applications to geophysical studies. *Reviews of Geophysics*, 46, RG2006:doi:10.1029/2007RG000228, 2008.
- N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A*, 454:903–995, 1998.
- N. E. Huang, C. C. Chern, K. Huang, L. W. S. S. R. Long, and K. L. Fan. A new spectral representation of earth-quake data: Hilbert spectral analysis of station TCU129, Chi-Chi, Taiwan, 21 September 1999. *Bull. Seismol. Soc. Am.*, 91:1310–1338, 2001.
- N. E. Huang, Z. Wu, and S. R. Long. On instantaneous frequency. *Advance in Adaptive Data Analysis*, 1, No.2:177–229, 2009.

- N. E. Huang, X. Y. Chen, M. T. Lo, and Z. Wu. On Hilbert Spectral representation: A true time-frequency representation for nonlinear and nonstationary data. *Advances in Adaptive Data Analysis*, Vol. 3, Nos. 1 & 2:63–93, 2011.
- T. Huang, W. Ren, and M. Lou. The orthogonal Hilbert-Huang transform and its application in earthquake motion recordings analysis. In *Proceedings of the 14th World Conference on Earthquake Engineering, October 12-17, 2008, Beijing, China*, 2008.
- P. Huber. *Robust Statistics*. John Wiley, New York, 1981.
- M. R. Ingham and R. Reeves. Magnetotelluric soundings and structure of the Tokaanu geothermal-field, New-Zealand. *J. Geomagn. Geoelectr.*, 45:729–740, 1993.
- M. D. Jegen, R. W. Hobbs, P. Tarits, and A. Chave. Joint inversion of marine magnetotelluric and gravity data incorporating seismic constraints: Preliminary results of sub-basalt imaging off the Faroe Shelf. *Earth and Planetary Science Letters*, 282:47–55, 2009.
- G. R. Jiracek, V. M. Gonzalez, T. G. Caldwell, P. E. Wannamaker, and D. Kilb. Seismogenic, Electrically Conductive, and Fluid Zones at Continental Plate Boundaries in New Zealand, Himalaya, and California-USA. pages 1–21, 2007.
- A. Jones, A. Chave, G. Egbert, D. Auld, and K. Bahr. A comparison of techniques for magnetotelluric response function estimation. *Journal of Geophysical Research*, 94:14201–14213, 1989.
- A. G. Jones, I. J. Ferguson, A. D. Chave, R. L. Evans, and G. W. McNeice. Electric lithosphere of the slave craton. *Geology*, 29:423–426, 2001.
- A. Junge. Characterization of and correction for cultural noise. *Surveys in Geophysics*, 17: 361–391, 1996.
- C. Junsheng, Y. Dejie, and Y. Yu. Research on the intrinsic mode function (IMF) criterion in EMD method. *Mechanical systems and signal processing*, 20:817–824, 2006.
- K. W. Key, S. C. Constable, and C. J. Weiss. Mapping 3D salt using the 2D marine magnetotelluric method: Case study from Gemini Prospect, Gulf of Mexico. *Geophysics*, 71 (NO. 1):B17–B27, 2006.
- Y. Kopsinis and S. McLaughlin. Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach. *IEEE transactions on signal processing*, 56, No. 1:1–13, 2008a.
- Y. Kopsinis and S. McLaughlin. Improved EMD Using Doubly-Iterative Sifting and High Order Spline Interpolation. *EURASIP Journal on Advances in Signal Processing*, Volume 2008:8 pages, 2008b.
- G. Kunetz. Processing and Interpretation of Magnetotelluric Soundings. *Geophysics*, 37: 1005–1021, 1972.
- J. Larsen. Transfer functions: smooth robust estimates by least-squares and remote reference methods. *Geophysical Journal International*, 89:645–663, 1989.

- J. C. Larsen, R. L. Mackie, A. Manzella, A. Fiordelisi, and S. Rieven. Robust smooth magnetotelluric transfer functions. *Geophysical Journal International*, 124:801–819, 1996.
- L. K. Law. Marine electromagnetic research. *Geophysical surveys*, 6:1 – 213, 1983.
- J. Ledo, P. Queralt, A. Marti, and A. G. Jones. Two-dimensional interpretation of three-dimensional magnetotelluric data: an example of limitations and resolution. *Geophys. J. Int.*, 150 (1):127 – 139, 2002.
- P. Lezaeta, M. Muñoz, and H. Brasse. Magnetotelluric image of the crust and upper mantle in the backarc of the NW Argentinean Andes. *Geophysical Journal International*, 142: 841–854, 2000.
- P. F. Lezaeta, A. D. Chave, and R. L. Evans. Correction of shallow-water electromagnetic data for noise induced by instrument motion. *Geophysics*, 70 (No.5):G127–G133, 2005.
- M. Loddo, D. Schiavona, and A. Siniscalchi. Generation of synthetic wide-band electromagnetic time series. *Annals of Geophysics*, 45, N. 2, April:289 – 301, 2002.
- P. Lonsdale. Structural pattern of the Galapagos microplate and evolution of the Galapagos triple junction. *Journal of Geophysical Research*, 93:1355113574, 1998.
- P. J. Loughlin and B. Tracer. On the amplitude- and frequency-modulation decomposition of signals. *J. Acoust. Soc. Am.*, 100:1594–1601, 1996.
- J. A. Madsen, D. J. Fornari, M. H. Edwards, D. G. Gallo, and M. R. Perfit. Kinematic framework of the Cocos-Pacific plate boundary from 13N to the Orozco transform fault: Results from an extensive magnetic and SeaMARC II survey. *Journal of Geophysical Research*, 97:7011 – 7024, 1992.
- C. L. Mallows. On some topics in robustness. Technical report, Bell Telephone Laboratories Technical Memorandum, Murray Hill, New Jersey, 1975.
- G. A. McMechan and I. Barrsdale. Processing electromagnetic data in the time domain. *Geophys. J. R. astr. Soc.*, 81:277–293, 1985.
- C. Mevel. Serpentinization of abyssal peridotites at mid-ocean ridges. *Comptes Rendus Geosciences*, 335:825 – 852, 2003.
- M. Moorkamp, B. Heincke, M. Jegen, A. W. Roberts, and R. W. Hobbs. A framework for 3-D joint inversion of MT, gravity and seismic refraction data. *Geophysical Journal International*, 184:477–493, 2011.
- A. Neska, K. Nowozynski, J. Reda, and M. Jegen. Reducing motion noise in marine magnetotelluric measurements by means of tilt records. *Geophysical Journal International*, 194, Issue 1:304 – 315, 2013.
- G. A. Newman, S. Recher, B. Tezkan, and F. M. Neubauer. Case history - 3D inversion of a scalar radio magnetotelluric field data set. *Geophysics*, 68:791–802, 2003.
- A. H. Nuttall and E. Bedrosian. On the Quadrature Approximation to the Hilbert Transform of Modulated Signals. *Proceedings of the IEEE*, 54, Issue:10:1458–1459, 1966.

- G. Oettinger, V. Haak, and J. C. Larsen. Noise reduction in magnetotelluric time-series with a new signal-noise separation method and its application to a field experiment in the saxonian granulite massif. *Geophys. J. Int.*, 146:659–669, 2001.
- Y. Ogawa, M. Mishina, T. Goto, H. Satoh, N. Oshiman, T. Kasaya, Y. Takahashi, T. Nishitani, S. Sakanaka, M. Uyeshima, Y. Takahashi, Y. Honkura, and M. Matsushima. Magnetotelluric imaging of fluids in intraplate earthquake zones, NE Japan back arc. *Geophysical research letters*, 28:3741–3744, 2001.
- C. H. Page. Instantaneous power spectra. *J. Appl. Phys.*, 23:103 – 106, 1952.
- N. A. Palshin, Y. M. Abramov, A. De Santis, A. Meloni, A. M. Poray-Koshits, V. S. Shneyer, and L. M. Abramova. Magnetovariational gradient sounding in the Tyrrhenian basin. *Physics of the Solid Earth*, 31:353–355, 1995. Russian Edition April 1995.
- R. Parker. The magnetotelluric inverse problem. *Geophysical Surveys*, 46:5763–5783, 1983.
- S. M. Peacock. Fluid processes in subduction zones. *Science*, 248:329 – 337, 1990.
- L. B. Pedersen, M. Bastani, and L. Dynesius. Groundwater exploration using combined controlled-source and radiomagnetotelluric techniques. *Geophysics*, 70:G8–G15, 2005.
- B. Picinbono. On instantaneous amplitude and phase signals. *IEEE Trans. Signal Processing*, 45:552–560, 1997.
- T. Qian, R. Wang, Y.-S. Xu, and H.-Z. Zhang. Orthonormal bases with nonlinear phases. *Adv. Comput. Math.*, 33:75 – 95, 2010.
- W. Qian and L. B. Pedersen. Industrial interference magnetotellurics: an example from the Tangshan area, China. *Geophysics*, 56:265–273, 1991.
- S. R. Qin and Y. M. Zhong. A new envelope algorithm of Hilbert-Huang transform. *Mechanical systems and signal processing*, 20:1941–1952, 2006.
- C. R. Ranero, J. P. M., K. McIntosh, and C. Reichert. Bending-related faulting and mantle serpentinization at the Middle America trench. *Nature*, 425:367–373, 2003.
- R. T. Rato, M. D. Ortigueira, and A. G. Batista. On the HHT, its problems, and some solutions. *Mechanical systems and signal processing*, 22:1374–1394, 2008.
- A. Rihaczek. Signal energy distribution in time and frequency. *IEEE Trans. Inform. Theory*, IT-14:369 – 374, 1968.
- G. Rilling and P. Flandrin. On the influence of sampling on the empirical mode decomposition. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing, ICASSP*, pages III-444 – III-447, 2006.
- G. Rilling and P. Flandrin. One or Two Frequencies? The Empirical Mode Decomposition Answers. *IEEE Trans. on Signal Proc.*, 56, No. 1:85 – 95, 2008.
- G. Rilling, P. Flandrin, and P. Goncalves. On empirical mode decomposition and its algorithms. In *In Proceedings of the 6th IEEE/EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP '03), Grado, Italy*, 2003.

- O. Ritter, U. Weckmann, T. Vietor, and V. Haaka. A magnetotelluric study of the Damara Belt in Namibia 1. Regional scale conductivity anomalies. *Physics of the Earth and Planetary Interiors*, 138:71–90, 2003.
- R. Rogers, H. Karason, and R. van der Hilst. Epeirogenic uplift above a detached slab in northern Central America,. *Geology*, 30:1031–1034, 2002.
- J. P. M. Rüpke, L.H., M. Hort, and J. A. Connolly. Serpentine and the subduction zone water cycle. *Earth and Planetary Science Letters*, 223:17–34, 2004.
- L. H. Rüpke. *Effects of Plate Subduction on the Earths Deep Water Cycles*. PhD thesis, Christian-Albrechts-Universität zu Kiel, 2004.
- T. Scheuer and D. W. Oldenburg. Aspects of time-variant filtering. *Geophysics*, 53:1399–1409, 1988.
- U. Schmucker. Electromagnetic induction in thin sheets: integral equations and model studies in two dimensions. *Geophysical Journal International*, 121:173–190, 1995.
- A. Schultz, R. D. Kurtz, A. D. Chave, and A. G. Jones. Conductivity discontinuities in the upper mantle beneath a stable craton. *Geophys. Res. Lett.*, 20:2941–2944, 1993.
- J. Shekel. Instantaneous frequency. *Proc. IRE*, 41:548, 1953.
- F. Simpson and K. Bahr. *Practical Magnetotellurics*. Cambridge University Press, 2005.
- W. Sims, F. Bostick, and H. Smith. The estimation of magnetotelluric impedance tensor elements from measured data. *Geophysics*, 36:938–942, 1971.
- M. Y. Smirnov. Magnetotelluric data processing with a robust statistical procedure having a high breakdown point. *Geophysical Journal International*, 152:1–7, 2003.
- U. Spagnolini. Time domain estimation of MT impedance tensor. *Geophysics*, 59:712–721, 1994.
- N. J. Stevenson, M. Mesbah, and B. Boashash. Multiple-view time-frequency distribution based on the empirical mode decomposition. *IET Signal Process.*, 4, Iss. 4:447–456, 2010.
- D. Sutarno and K. Vozoff. Robust M-estimation of the magnetotelluric impedance tensor. *Expl. Geophys.*, 20:383–398, 1989.
- D. Sutarno and R. Vozoff. Phase-smoothed robust M-estimation of magnetotelluric impedance functions. *Geophysics*, 56:1999–2007, 1991.
- N. Suto, M. Harada, J. Izutsu, and T. Nagao. Time variation of the electromagnetic transfer function of the earth estimated by using wavelet transform. *Proc. Jpn. Acad., Ser. B*, 82:175–180, 2006.
- C. M. Swift. *A Magnetotelluric Investigation of an Electrical Conductivity Anomaly in the Southwestern United States*. PhD thesis, Dept. of Geology and Geophysics, MIT, Cambridge, Mass., 1967.
- L. Szarka. Geophysical Aspects of Man-made Electromagnetic Noise in the Earth - A Review. *Surveys in Geophysics*, 9:287–318, 1988.

- D. J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70: 1055–1096, 1982.
- D. J. Thomson and A. D. Chave. Jackknifed error estimates for spectra, coherences, and transfer functions. *Advances in spectrum analysis and array processing*, 1:58–113, 1991.
- D. J. Thomson, M. F. Robbins, C. G. Maclellan, and U. Lanzerotti. Spectral and windowing techniques in power spectral analyses of geomagnetic data. *Physics of the Earth and Planetary interiors*, 12:217–231, 1976.
- A. N. Tikhonov. On the determination of electric characteristics of deep layers of the Earth's crust. *Dokl. Akad. Nauk SSSR*, 73:295–297, 1950.
- J. W. Tukey. Bias and confidence in not quite large samples (abstract). *Annals of Mathematical Statistics*, 29:614, 1958.
- V. Tuncer, M. J. Unsworth, W. Siripunvaraporn, and J. A. Craven. Exploration for unconformity-type uranium deposits with audiomagnetotelluric data: A case study from the McArthur River mine, Saskatchewan, Canada. *Geophysics*, 71:B201–B209, 2006.
- B. Van der Pol. The fundamental principles of frequency modulation. *Proc. IEE*, 93 (III): 153 – 158, 1946.
- I. M. Varentsov and E. Y. Sokolova. Generation of synthetic magnetotelluric data. *Izvestiya, Physics of the Solid Earth*, 30:554 – 562, 1995.
- K. Vasudevan and F. A. Cook. Empirical mode skeletonization of deep crustal seismic data: Theory and applications. *J. Geophys. Res.*, 105:7845–7856, 2000.
- J. Ville. Theorie et application de la notion de signal analytique. *Cables et Transmissions*, 2A(1):61 – 74, 1948.
- K. Vozoff. The magnetotelluric method in the exploration of sedimentary basins. *Geophysics*, 37:98–141, 1972.
- C. Walther and E. Flueh. Remnant of the ancient Farallon Plate breakup: A low-velocity body in the lower oceanic crust off Nicoya Peninsula, Costa Rica - evidence from wide-angle seismics. *Geophysical research letters*, 29:45–1 – 45–4, 2002.
- G. Wang, X.-Y. Chen, F.-L. Qiao, Z. Wu, and N. E. Huang. On intrinsic mode function. *Advance in Adaptive Data Analysis*, 2, No. 3:277 – 293, 2010.
- P. E. Wannamaker, J. R. Booker, A. G. Jones, A. D. Chave, J. H. Filloux, H. S. Waff, and L. K. Law. Resistivity cross-section through the Juan de Fuca subduction system and its tectonic implications. *Journal of Geophysical Research*, 94:14,127–14,144, 1989.
- P. E. Wannamaker, J. A. Stodt, L. Pellerin, S. L. Olsen, and D. B. Hall. Structure and thermal regime beneath the South Pole region, East Antarctica, from magnetotelluric measurements. *Geophys. J. Int.*, 157:36–54, 2004.
- Y. Washizawa, T. Tanaka, D. P. Mandic, and A. Cichocki. A flexible method for envelope estimation in empirical mode decomposition. In *Proc. of 10th International Conference on Knowledge-Based & Intelligent Information & Engineering Systems (KES2006)*, pages 1248 – 1255, 2006.

- W. Wei, M. Unsworth, A. Jones, J. Booker, H. Tan, D. Nelson, L. Chen, S. Li, K. Solon, P. Bedrosian, S. Jin, M. Deng, J. Ledo, D. Kay, and B. Roberts. Detection of Widespread Fluids in the Tibetan Crust by Magnetotelluric Studies. *Science*, 292(DOI: 10.1126/science.1010580):716–719, 2001.
- P. D. Welch. The use of the fast Fourier transform for estimation of spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.*, AU-15:70–74, 1967.
- Y. K. Wen and P. Gu. HHT-based simulation of uniform hazard ground motions. *Advance in Adaptive Data Analysis*, 1:71 – 87, 2009.
- D. R. Word, H. W. Smith, and F. X. Bostick. An investigation of the magnetotelluric tensor impedance method. Technical report, Tech. Rep. 82, Electr. Geophys. Res. Lab., University of Texas at Austin, 1970.
- T. Worzewski. *Marine Magnetotellurics on a Continental Margin: Imaging the Hydration and Dehydration Cycle of the Costa Rican Subduction Zone*. PhD thesis, Christian-Albrechts-Universitaet zu Kiel, 2011.
- T. Worzewski, M. Jegen, H. Kopp, H. Brasse, and E. T. Castillo. Magnetotelluric image of the fluid cycle in the Costa Rican subduction zone. *Nature Geoscience*, 4:108 – 111, 2011.
- T. Worzewski, M. Jegen, and A. Swidinsky. Approximations for the 2-D coast effect on marine magnetotelluric data. *Geophys. J. Int.*, 189:357 – 368, 2012.
- Q. Wu and S. D. Riemenschneider. Boundary extension and stop criteria for empirical mode decomposition. *Advances in Adaptive Data Analysis*, 2, No. 2:157 – 169, 2010.
- Z. Wu and N. E. Huang. A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. R. Soc. Lond. A*, 460:1597–1611, 2004.
- Y.-S. Xu and H.-Z. Zhang. Recent mathematical developments on Empirical Mode Decomposition. *Advances in Adaptive Data Analysis*, 1, No. 4:681 – 702, 2009.
- E. Yee and K. V. Paulson. Concerning dispersion relations for the magnetotelluric impedance tensor. *Geophysical Journal*, 95:549 – 559, 1988.
- Ö. Yilmaz. *Seismic data processing*. Soc. Expl. Geophys., 1987.
- R. R. Zhang. Characterizing and quantifying earthquake-induced site nonlinearity. *Soil Dyn. Earthquake Eng.*, 26:799–812, 2006.
- R. R. Zhang, S. Ma, and S. Hartzell. Signatures of the seismic source in end-based characterization of the 1994 north-ridge, california, earthquake recordings. *Bull. Seismol. Soc. Am.*, 93:501–518, 2003a.
- R. R. Zhang, S. Ma, E. Safak, and S. Hartzell. Hilbert-Huang transform analysis of dynamic and earthquake motion recordings. *J. Eng. Mech.*, 129:861–875, 2003b.

Acknowledgements

I own my deepest gratitude to my PhD-supervisor Dr. Marion Jegen-Kulcsar. *Dear Marion, Many many thanks! Without your guidance and constant support in the last seven years, this work could not be ongoing to its completion. Thanks for your incessant encouragements, and for your scientific and personal counsel to make me scientifically growing up, and for the homey environment you have created for a foreign student - I was so enjoyed working in your group and could never forget all you have done for me. Thanks for your endless patience to correct my paper and thesis, to endure my "Chinglish". Thanks also for all the fruitful discussion we had about science, about the education of children and the differences between the cultures, and for your always caring of my family and very thoughtful helps to the birth and growing-up of my daughter Leonie. I have learnt so many things from you. Also thanks for your family always caring about me. I owe you!*

I am sincerely thankful to Prof. Dr. Wolfgang Rabbel for agreeing to be my co-referee and taking time to read my thesis.

I am also sincerely thankful to Prof. Dr. László Szarka and Prof. Dr. Lars Rüpke for taking time to read my thesis and agreeing to be members of the examination committee.

I am so deeply grateful to Dr. Björn Heincke. *Dear Björn, thank you for your support and for always taking time to help me. Thanks for your providing me so many intelligent ideas and suggestions, especially for reasonable scientific thinking in the past years. Thanks for your endless patience to discuss with me, to explain your joint inversion code, to share your opinions and to correct my paper and thesis. I am also so grateful for your helps at the beginning of my arrival in Kiel.*

Many, many thanks to Dr. Max Moorkamp. *Dear Max, you are a super expert of data processing. You provided me so many invaluable ideas and suggestions, without which my work could not go to this end. Thanks for your patience and taking time to make all the corrections.*

I am also deeply thankful for Dr. Tamara Worzewski. *Dear Tami, I am so lucky to have had a co-PhD fellow like you. You are so nice, kindly, friendly, accommodating, merry and Thanks for providing me the rotated MT data and sharing the ideas about the data. You gave me so many helps that I could not make a list. I was very happy to work with you. All things you have done for me will remain in my mind forever.*

I wish to warmly and sincerely thank all other members in Marion's EM group for being so supportive, for making environment most pleasant: Dr. Sebastian Hölz, Dr. Anna Avdeeva, Dr. Andrei Swidinsky, Dr. Reza Mir, Romina Gehrmann, Malte Sommer.

Dear Anna, thank you for always caring about the progress of my work and taking time to listen my problems.

Dear Sebi, thank you for your advice, helpfulness and patience. You are a super Matlab expert, have helped me so much with my codes. Also thanks for providing me so many nice barbecues.

Dear Andrei and Reza, thanks for reading through my paper and providing me so many fruitful comments and corrections of the language.

I wish to thank Dr. Alan Chave, Dr. Xavier Garcia and Maik Neukirch for thoughtful peer-reviewing of the manuscript and providing important comments, also thank for valuable suggestions from editor Dr. Oliver Ritter. All of these improved the manuscript until it was published.

I would like to thank for very helpful discussions with Prof. Ulrich Schmucker, Prof. Peter Weidelt, Prof. Andreas Junge, Prof. Gary Egbert, Maik Neukirch, Pierre Wawrzyniak and all other persons.

I sincerely thank all members of the Geodynamics Section of GEOMAR and SFB 574 for helpful discussions. Specially thank Erna Lange, Silke Schenck and Martina Noeske for their pleasant company and constant support in making everything easier every day.

I am very thankful to my beloved wife, Bolian Fu, for her encouragements and taking care of me, my son Weiahao and my daughter Leonie Tingyan. *My dear Bolian, without your love, your support and selfless dedications, I could not finish my work. I know it is a very hard time for you in these years. Most heartfelt gratitude to you as the person being always at my side, always supporting and encouraging me to hold on.*

My dear Weihao and Leonie, I am so proud of you. You give me so many amusements and you are the motives of my work.

I wish to give my special thanks to my beloved parents, Xianzhen Chen and Zhengbi Li, and my sisters, Xiaorong, Ningchuan and Mei and their families, for their unconditional love and consistent encouragements and supporting me to come in Germany to study. Also special thanks to my cousin, Qinghui He and other friends and relatives.

Also I want to thank all friends and colleagues of my previous worked university (CQUPT) in China. I give my sincerely thanks to all friends of Chinesischer Bibelkreis Kiel for their always helps and encouragements.

Curriculum Vitae

Personal data:

Name: Jin Chen

Date and Place of birth: 25 September 1967, Qinghai, P. R. China

Nationality: P. R. China

Education:

09. 1974 - 07. 1981 Bayi School of Geermu, Qinghai, P. R. China.

09. 1981 - 07. 1983 The Third high-school of Geermu, Qinghai, P. R. China.

09. 1983 - 07. 1987 Fudan University, Shanghai, P.R.China.

Bachelor of Science in Mathematics.

09. 1996 - 12. 1999 Chongqing University, Chongqing, P.R.China.

Master of Science in Applied Mathematics.

10. 2001 - 08. 2003 University of Kaiserslautern, Germany.

Master of Science in Industrial Mathematics (Specialization: Mathematical Modelling and Scientific Computing)

03. 2006 - now PhD student of GEOMAR and University of Kiel (CAU), Germany.

Working experiences:

07. 1987 - 10. 2001 Lecturer of mathematics at Chongqing University of Posts and Telecommunications, Chongqing, P. R. China.

10. 2003 - 02. 2006 Doing Research work at Technical University of Kaiserslautern and Fraunhofer Institut of Techno- und Wirtschaftsmathematik (ITWM), Germany.

Declaration

I hereby declare that I am the only author of this work and that no sources other than those listed have been used in this work.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass die vorliegende Abhandlung, abgesehen von der Beratung der Betreuer, nach Inhalt und Form meine eigene Arbeit ist. Bisher ist die Arbeit noch nicht an anderer Stelle im Rahmen eines Prüfungsverfahrens vorgelegt worden. Die Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden.

Kiel, Dezember 2013

Jin Chen