h Reconstruction

draft 2

May 7, 1999

Contents

1 Introduction

1.1 Definition

Reconstruction is intended as the first step to all data analyses. Its primary goal is to disentangle - as much as possible - the data from any bias introduced by the detector. This includes unfolding the frequency dependent response of the detector in the output time series recorded at the antisymmetric port of the interferometer, as well as subtracting to the measured signal those contributions which are known to be noise.

More generally, reconstruction can include all the pre-processing which is common to subsequent analysis algorithms, and which has to be done in a consistent way.

At first order, reconstruction is simply unfolding the frequency dependent response of the detector in the noisy output time series, i.e. computing the relative differential length variation of the long ITF cavities which would have produced the same output, regardless of whether this differential length variation is due to a true GW, or even whether the detector output has been caused by a differential length variation of the long ITF cavities.

As a second step, though, the process of removing some noise can be introduced in the reconstruction, i.e. removing from the unfolded detector output some contributions that we know are not coming from GW signal. This assumes that additional information can be used in order to remove such noise contributions. Some contributions such as harmonics from the 50 Hz line for instance could be evaluated directly from the interferometer main channel, using the fact that such noise is periodic. More generally, noise contributions induced by the environment (electro-magnetic noise, seismic noise...) can be reduced using the information recorded in the environment channels. Other channels such as the signals collected by the photo-diodes monitoring the secondary beams could also be used - if well under control - to clean up the signal induced by a differential movement of the long cavities from contributions induced by other interferometer vibration modes.

The contents of the reconstruction procedure is therefore very likely to evolve with time, as we learn about the detector and the noise sources, starting from a basic unfolding operation and evolving towards more and more complex noise removal.

The reconstruction procedure could also include the pre-whitening of the interferometer output. Strictly speaking, whitening is beyond the scope of reconstruction, but on-line whitening can be thought of as pre-processing common to several algorithms and be done at the reconstruction stage, if necessary.

In the same way, decimation is likely to be needed at this stage, to be able to provide reduced data streams to those algorithms which do not need the full bandwidth. Reconstruction should therefore also provide some reconstructed data sampled at a reduced frequency with respect to the original acquisition frequency.

1.2 Requirements

The needs on the precision of the reconstruction process arise from two distinct requirements.

1. One set of requirements arises from the necessity to be able to compare in a useful way the GW signals measured in different detectors, in order to extract valuable information on the signals and their sources. Such comparisons are based essentially on two parameters of the measured signals: amplitude and arrival time. This leads to requirements on the absolute timing and absolute gain correction performed in the reconstruction.

   Uncertainties in this process should be less than those arising from the moderate signal to noise ratio of the measurement. Assuming a signal to noise ratio of at most 20, the uncertainty on signal amplitude from an overall gain uncertainty - over the whole bandwidth - of the detector response should stay within 5%. A signal to noise ratio of 20 should result into a resolution for time of flight measurements between two detectors around 100 μs, for the shortest signals (1 ms long). The intrinsic timing ability of the GPS system being 1 μs, uncertainties in the delays of the instrument could reasonably lead to a timing uncertainty within 20 μs, consistent with the overall goal.

2. Another set of requirements is related to the issue of signal detection in VIRGO itself, especially detection of signals from coalescing binaries. The search for such signals is based on optimal filtering which relies on a close match between the expected signal and the signal measured by the instrument. Errors in the transfer function can result into a reconstructed signal which is slightly mismatched to the expected signal, thus reducing the detection ability.

   Point (1.) was concerned with an overall gain uncertainty of the detector response, which is irrelevant here since it affects both noise and signal in a uniform way and thus does not affect the signal to noise ratio performance of a matched filter. On the other hand, errors in the detector response which are not uniform over the whole bandwidth can result in a loss of signal to noise ratio. Such effects can arise from random noise affecting the measurement of the detector response. Also, uncertainties on the values of the characteristic time constants involved in the detector response (e.g.

cutoff frequency of the Fabry-Perot cavities) can lead to the same kind of effects, especially undesirable phase mismatches.

The actual loss in SNR is related to the mismatch between the imperfectly reconstructed signal and the expected signal, and exact requirements depend on the shapes of the signal and detector noise. However, a conservative specification is to require that uncertainties in the detector response be such that the - local - loss in SNR is less than 5% in every point of the frequency bandwidth, to make sure that the overall loss in the integrated SNR be also within 5%. This conservative specification is fulfilled for instance if the real and imaginary parts of the transfer function are known within ± 5% at any frequency.

1.3 Outputs

The output of the reconstruction step will not be unique but will consist of several channels (*list to be completed according to needs*):

- one or more channels (corresponding to different reconstruction methods) with reconstructed data sampled at the original acquisition frequency (20 kHz)
- one or more channels with reconstructed data cleared of some identified noise contributions
- one or more channels with reconstructed data sampled at reduced frequencies (e.g. 2 kHz)
- one or more channels with whitened data.

A naming convention will be used for the different output channels.

1.4 Software Issues

1.4.1 Transfer Function Storage

The information about the detector transfer function will be stored both in a database and in the frames themselves, using the static data structure (`FrStatData`).

1.4.2 Reconstructed Data Storage

The reconstructed data will be added in the original frames using the processed data structure (`FrProcData`). The frames can contain as many such structures as necessary, thus allowing for the data processed with different reconstruction methods to be stored together in a non-exclusive way. In the same way, the result of off-line re-processing can be added in the frames while keeping the original on-line reconstructed data.

1.4.3 Reconstruction Algorithm

The idea is to have a first reconstruction performed on-line, in order to be able to send reconstructed data to the trigger algorithms. Technically the reconstruction algorithm will be considered in the same way as the others and will follow the same standards as trigger algorithms [1].

2 Transfer Function Determination

2.1 Indirect Method

The overall detector transfer function can be determined from the knowledge of the different contributions and the relevant parameters. Depending on the stability of the transfer function, this a priori guess of the transfer function can either be used directly to reconstruct the data, or rather be used as a starting point to the direct determination of the transfer function from calibration runs (see next section). A tentative list of the different TF contributions and of the relevant parameters follows:

- optics

  In the considered bandwidth, the ITF optical transfer function is basically that of a first order low-pass filter whose gain and cutoff frequency depend on a number of parameters of the ITF. Parameters liable to vary in time include laser power and modulation depth. The optical transfer function is also sensitive to offsets in the locking system that may shift the ITF working point and change the power stored in the cavities.

  Changes in the optical transfer function might be tracked on line through monitoring of the relevant parameters (TBC), a global measurement being provided by the calibration runs.

- control

  The transfer function of the ITF control system may have a complex shape and may evolve in time following changes in the control system. However this should be well under control because the system is digital, the parameters of the locking transfer function being known and stored in an appropriate database.

- electronics

  The electronics brings several contributions to the detector transfer function. One contribution is due to the anti-aliasing filters used in front of the

ADCs. They are analog filters and therefore their parameters may slightly vary in time.

Another contribution is due to the (analog) pre-whitening filters used in front of the ADCs and the corresponding (digital) restoring filters. Those filters are designed such that the product of both transfer functions be unity, so that the overall contribution to the detector transfer function should vanish. However,like the anti-aliasing filter, the analog part may suffer from time variations, which can be monitored only through calibration runs.

- DAQ

  The readout and acquisition of the signals introduce extra delays in the detector response. Those delays will be measured a priori. However, during operation, possible changes can me monitored only through calibration runs.

2.2 Direct Method

2.2.1 Calibration Principle

The configuration of the calibration system [2] is summarized in figure 2.2.1. The idea is to perform a direct measurement of the detector response by forcing a known differential displacement on the mirrors of the ITF and recording the resulting output signal. The forces needed to do so will be applied on the mirrors through two different ways: either by using the coil-magnet pairs used as locking actuators or by using the radiation pressure of a secondary laser beam.

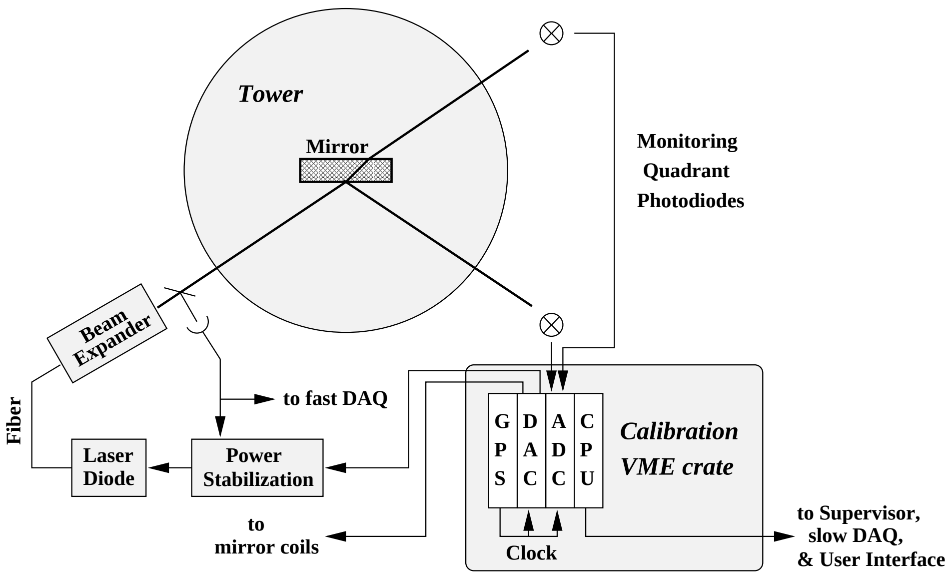

Figure 1: Configuration of the calibration system.

2.2.2 Calibration Signals

The generator used to provide the driving signals of the calibrator is based on a CPU, and can therefore generate a variety of signal shapes very easily. Some possibilities are listed below, together with a short discussion of their respective advantages and drawbacks:

- swept sines

  This is the type of signal used in a standard calibration method. The response is measured at a given frequency at a time, and the frequency is swept to cover the full bandwidth. The points where the measurement is performed can be chosen so as to lie at frequencies where the detector sensitivity is good (avoiding resonances) and the measurement can be done with good signal to noise ratio. The drawback of this method is that it can take some time, since the measurement is repeated at one frequency after the other. Since the signal generator we have at our disposal is based on a CPU and is not limitated to deliver signals at a single frequency at a time, the swept sines can be replaced advantageously by a sine comb.

- sine comb

  This type of signal has the same advantages as the swept sines without the drawback. Namely the measurement can be done in all the frequency points at the same time, and the sine frequencies can be chosen to lie at places where sensitivity is good. Such a method seems most appropriate in the case we wish to record the transfer function in the frequency domain (see section 2.4), with the possibility to interpolate between the measured points to obtain the response function in as many points as necessary for the reconstruction process.

- white noise

  On the other hand, white noise may be more appropriate if the goal is to obtain a parametric identification of the detector response (see section 2.4), which can be achieved through analysis of the correlation between the driving signal and the output signal. In the case a white excitation is used, it might be necessary to filter out the low frequency part to avoid exciting the SA chain holding the mirror.

2.2.3 Duration of a Calibration Run

The calibrator has been designed to be able to deliver signal amplitudes such that the spectral density of the resulting mirror displacements is comparable to or better than the VIRGO sensitivity, so that good signal to noise ratio can be obtained in a reasonable time. The exact duration of calibration runs is TBD, according to the precision we want to reach on the transfer function determination.

2.3 Transfer Function Monitoring

During interferometer operation, a permanent calibration signal of small amplitude will be injected in order to monitor changes in the transfer function and decide when a new calibration run is needed. Typically this permanent calibration signal will be a comb of sines located at critical frequencies (e.g. cutoff frequency of Fabry-Perot cavities, intermediate – coalescing binary – frequency range).

Other information than permanent calibration signals could be used to monitor changes in the transfer function, like the amplitude of the violin modes in the output spectrum (TBC).

2.4 Transfer Function Representation

There are essentially two different ways of representing the interferometer transfer function, which are closely related to the method used to reconstruct the data in the time domain (see section 3).

2.4.1 Non Parametric TF Representation

In this approach of a frequency domain representation, the transfer function is described by a set of numbers representing its complex amplitude at a number of frequencies covering the measurement bandwidth. These values are extracted from the calibration runs, typically by comparing the spectra of the interferometer output and of the driving calibration signal.

2.4.2 Parametric TF Representation

Another approach is to use a parametric representation of the transfer function. Such a representation could presumably be more compact, and is natural in the perspective of reconstructing the data in the time domain. Moreover, the a priori knowledge of the transfer function that we may have (from previous determination or from the indirect method) could be used in the identification process.

In this approach, the transfer function is described by an ARMA process whose parameters are assumed to be estimated from the calibration runs:

$$itf(n) \;=\; -\sum_{k=1}^{P} a_k \; itf(n-k) \;+\; \sum_{k=0}^{Q} b_k \; h(n-k)$$

where itf is the time sequence of the interferometer output, h is the calibration h-like differential signal, and the a_k and b_k are the parameters of the ARMA process describing the response of the system.

3 Data Reconstruction in Time Domain

There are two general ways to compute time sequences of reconstructed data. The first one is to switch to the frequency domain to unfold the interferometer response, and the other one is to use a digital filter to perform the unfolding directly in the time domain. These two choices are closely related to the choice of the representation of the transfer function.

3.1 Frequency Domain Filtering

In the case the ITF transfer function is described in the frequency domain by a set of complex numbers, the basic operation to compute the time sequence of reconstructed data is of the type:

$$h = FFT^{-1}\left\{ \frac{FFT\{itf\}}{TF} \right\}$$

This is a simple operation which needs however some special care to work:

- Overlapping data sets should be used in order to avoid end effects.
- The dynamic range of the itf spectrum should be small enough to avoid numerical problems. This is the most critical point, since the ITF output signal has typically a large low-frequency component and large noise peaks dominating over the standard noise level. If this issue turns out to be severe, it could be necessary to filter out the low frequency component and to remove the noise contribution from resonances (see section 4) prior to the reconstruction operation.

3.2 Time Domain Filtering

Reconstructing data directly through time domain filtering is possible if a parametric description of the transfer function has been obtained. The method relies on some implementation of the basic inverse filtering operation:

$$h(n) \;=\; \frac{1}{b_0}\left\{ -\sum_{k=1}^{Q} b_k\, h(n-k) \;+\; \sum_{k=0}^{P} a_k\, itf(n-k) \right\}$$

4 Noise Removal

One important point in the reconstruction process is to remove from the measured signal contributions that can be identified as noise. This can be done provided some additional information is available.

Notes:

- It is yet TBD whether the noise removal should take place after the time domain data reconstruction or before. The latter solution could be especially interesting in the perspective of using frequency domain filtering to reconstruct the data, if numerical problems arise due to the large dynamic range of the signal spectrum, since the range could be reduced through removing noise.
- The noise removal process is related to the issue of assessing the quality of the data, and the results obtained in the clean-up operation could be used for diagnostics and by the Data Quality on-line task.

4.1 Noise removal based on main channel

Some noise contributions can be identified and removed using the ITF output signal alone, under the assumption that the noise is periodic. Such line removal algorithms [3] can be used to remove noise contributions consistent with a sine wave, typically noise from the 50 Hz line harmonics or from violin modes. Some tests have shown that they could be applied valuably to data from a gravitational wave interferometer [4].

4.2 Noise removal based on additional channels

More generally, most noise contributions to the ITF signal come from environmental sources like seismic noise, electro-magnetic noise, or even Newtonian noise, which can be monitored independently. Assuming the relevant parameters are suitably monitored, the information thus recorded can be used to clean-up the signal of interest.

Work in [5] show promising results about automatic cross-talk removal from multi-channel data applied to data from the Caltech 40 meter prototype. The technique used is based on assuming that the signal of interest is a linear combination of some true signal plus noise uncorrelated with the true signal, and that the noise can be related to the signals measured by environment channels through unknown transfer functions. The transfer functions are estimated through analysis of the observed correlation between the different channels, which leads to the possibility of subtracting the environmental contamination in the main channel.

The technique just mentioned relies on the assumption that the auxiliary channels are not correlated with the true signal we want to measure. In practice we will also have some auxiliary signals that *are* correlated with the true signal, namely the signals collected by the auxiliary photo-diodes, which could nevertheless be used to disentangle the contributions from the various ITF vibration modes to the signal measured at the antisymmetric port. This should be done with care, and requires that the relationships between the ITF modes and the photo-diode signals are well under control.

5 Other Common Processing

This section deals with the processing which has to be done prior to search analysis and could therefore be included in the reconstruction process, though its purpose lies somewhat outside the scope of reconstruction.

5.1 Decimation

In many cases, search algorithms do not need data sampled at 20 kHz but can be applied on data sampled at a reduced frequency. A typical example is the search for coalescing binary signals, which does not benefit from a bandwidth extending above say, 1kHz.

The reconstruction stage can therefore provide several output channels with data resampled at a lower frequency after proper low-pass filtering in order to avoid aliasing. This allows to reduce both the data flow and the computing effort for subsequent analyses.

5.2 Whitening

If possible, on-line spectral estimation and whitening of the measured signal [6] would be an efficient way to provide suitable data to the analysis algorithms based on matched filtering. As such, it could be implemented at the reconstruction stage.

References

[1] D.Buskulic, F.Marion, B.Mours and D.Verkindt, *A Proposal for Data Analysis Algorithm Software Specifications*, VIRGO note VIR-SPE-LAP-5400-106 (1999)

[2] VIRGO Coll., Final Design Report (1997)

[3] D.B.Percival and A.T.Walden, *Spectral Analysis for Physical Applications*, Cambridge University Press (1993)

[4] B.Allen et al, *GRASP: a data analysis package for gravitational wave detection*, version 1.8.3. Manual and package at `www.lsc-group.phys.uwm.edu/`.

[5] B.Allen, W.Hua, A.C.Ottewill, *Automatic cross-talk removal from multichannel data*, preprint WISC-MILW-99-TH-04 and LIGO-P9990002-00-E (1999)

[6] E.Cuoco, PhD Thesis, Università degli Studi di Pisa (1997)