High precision radio pulsar timing

Janssen, G.H.

CHAPTER 6

Using and optimising multi-telescope radio pulsar timing

G. H. Janssen and B. W. Stappers

For the European Pulsar Timing Array
To be submitted to Astronomy & Astrophysics

Abstract There are three projects around the world that are aiming to try and detect gravitational waves using a pulsar timing array. This is an ambitious project needing the highest timing precision over a large number of pulsars and requires that precision be maintained over at least 5 years. It is recognised that this is unlikely to be able to be achieved using a single telescope, due to the large number of observations of a large number of pulsars required. The EPTA is the first of the projects designed to combine data from multiple telescopes. Ultimately, sharing data from all collaborations is likely to be necessary in order to achieve gravitational wave detection on the timescale of the next few years. Combining pulsar timing data sets that originate from multiple telescopes has a number of very clear advantages; as each observatory covers a different time span, although usually overlapping, the total time is generally extended. It also results in better orbital phase sampling for binaries, and when observations at multiple frequencies are available, provides a tool to probe (changes in) the interstellar weather. However, every observatory uses different procedures and methods to acquire and process the data, and also the process of calculating the times of arrival (TOAs) for each observatory may vary. Calculating TOAs is by itself already a complex procedure that can be influenced by several effects. Comparing TOAs from multiple telescopes therefore requires consideration of a multitude of differences to optimise the timing solution. We combined data sets of several pulsars generated by multiple EPTA telescopes. To measure and monitor differences in arrival times, and to exclude intrinsic pulsar variations that may give

rise to extra differences, we carried out a few sessions of simultaneous observations using two or more telescopes. For non-millisecond, closeby pulsars, the offsets between telescopes are dominated by template offsets. These can be eliminated by aligning the templates before calculating TOAs, or reprocessing the observations using the same standard ephemerides for each pulsar at each observatory. In other cases, for example when a pulsar signal is affected by DM variations along the line of sight, or pulsars that show profile frequency evolution, aligning is more complicated and requires more modelling and measurements to be resolved. We provide basic guidelines to optimally combine timing observations from multiple telescopes. The detection of gravitational waves in the pre-SKA era is likely to require the combination of data from many different telescopes around the world.

6.1 Introduction

The European Pulsar Timing Array (EPTA) network is now sharing pulsar timing data sets on a regular basis. For most pulsars, the best timing solution is found by using as many data points as possible. Generally, this allows for extending the time span of observations to a longer period. For overlapping periods, the observations are usually independently scheduled between observatories, resulting in better distribution of the observations over time. As all telescopes have different capabilities at different observing frequencies, there are usually TOAs from more observing frequencies available than when using just one telescope. When the pulsars are regularly observed at the separate frequencies, this allows for monitoring, and correcting for, changes in the interstellar weather along the line-of-sight towards the pulsar (Foster & Cordes 1990; You et al. 2007a).

So far, when combining data sets for pulsar studies, it has been standard procedure to use one constant offset for each added data set from an additional telescope. However, over the last couple of years it has become evident that this approach is not sufficient for all pulsars, and to improve timing solutions to higher precision requires a more sophisticated analysis of the TOA calculation procedure at the different observatories.

The overall process, starting from observing a pulsar using a radio telescope to the calculation of a TOA for that observation, involves a large number of separate processes and steps in analysing the data. Although observing procudures at different telescopes are quite comparable, high precision timing depends on extremely accurate time measurements. First of all, each observatory uses its own reference clock to put a time stamp on observations. These clocks are then referred to an international time standard, however errors in the local clock may occur that need to be detected and accounted for.

As pulsar timing uses broad-band observations, and usually does not work on single pulses, the observation of a pulsar needs to be dedispersed and folded to determine one profile for each observation. This is done using an ephemeris for the pulsar, and between observatories these ephemerides are usually slightly different. To calculate the TOA, the measured profile of the observation is compared to a high signal-to-noise (S/N) template profile, again each observatory uses their own template generated from integrating a large number of high S/N observations. As the zero-phase reference point on this profile is an arbitrary choice, differences in TOAs result from this step as well.

In this Chapter we will describe all effects that may influence the calculation of TOAs. We will present the results of combining data sets from different telescopes and describe the specific problems that need to be taken care of to optimally align the data sets.

6.2 Origin of time of arrival differences

The time of arrival (TOA) that is calculated for a pulsar timing observation can be summarized as follows:

$$T_{\mathrm{obs}} = T_{\mathrm{PSR}} + \Delta T_{\mathrm{cable}} + \Delta T_{\mathrm{freq}} + \Delta T_{\mathrm{prof}} + \Delta T_{\mathrm{clock}} + \Delta T_{\mathrm{bary}} \tag{6.1}$$

where T_{obs} is the arrival time as measured at the telescope. This is the arrival time of the signal coming from the pulsar, with several additions resulting from external or intrinsic effects. When comparing TOAs from two telescopes, one or more of these effects can be different.
T_{PSR} is the effective arrival time at the telescope of the signal as transmitted by the pulsar, including interstellar medium (ISM) effects. The ISM effect is dependent on frequency and therefore the error on the measurement of this parameter may be correlated with ΔT_{freq} if the observations at the telescopes that are being compared were not taken at the same frequency, or with the same bandwidths.
$\Delta T_{\mathrm{cable}}$ is any cable-induced delay in arrival time between the telescope and its pulsar backend.
ΔT_{freq} is the difference induced by small observing frequency differences. Although all EPTA telescopes have similar observing frequencies around 1400 MHz, the mid-frequencies of their respective observing bandwidths are not exactly the same. When the dispersion measure of the pulsar is not known exactly, or when it is changing with time, this leads to differences in arrival times calculated from different frequencies. Moreover, some pulsars show profile evolution dependent on observing frequency (e.g. Kramer et al. 1999b). This will in some cases lead to alignment errors when using multiple observing bands. The use of frequency-dependent, synthetic templates built from Gaussian profiles at all observatories will ultimately remove the errors generated by ΔT_{freq}. An extensive study on synthetic profiles will be presented in a forthcoming paper (Purver et al. 2009, in prep.).
ΔT_{prof} is the result induced by using different ephemerides for folding at the different telescopes. The high S/N templates that are used for TOA calculations are independently made at each telescope, and as S/N ranges differ for each telescope, those may be slightly different. Furthermore, choosing a different reference point within the profile will give a phase offset. Although this difference will be the largest among all effects, it is constant as long as observatories do not change their reference template, and can be easily found by comparing the templates and inserting the corresponding offset in the timing ephemeris.
$\Delta T_{\mathrm{clock}}$ represents local clock errors. They may be of any size, although large clock errors are easily identified and accounted for. However, if clock "glitches" occur with sizes as small as the TOA measurement error, they will stay unnoticed.
ΔT_{bary} is the difference when referring each TOA to the Solar system barycentre (SSB). This transfer is dependent on the observing frequency, and the accuracy of the positions of the pulsars as used in the ephemeris.

Table 6.1: Properties of the data sets of the pulsars that are presented in this study.

Pulsar	Telescope	Time range	Frequency (MHz)
J1518+4904	WSRT	1999-2007	840,1380,2300
	EFF	2002-2007	1400
	NCY	2004-2007	1368
	JB	1995-2006	400,600,1400
	GBT	1994-2003	350,600,800
J1012+5307	WSRT	1999-2008	1380
	EFF	1997-2008	1400
	NCY	2005-2008	1368
	JB	1993-2003	1400
J1713+0747	WSRT	1999-2008	1380
	EFF	1996-2008	1410
B1937+21	WSRT	1999-2008	840,1380,2300
	EFF	1997-2008	1400,2600
	NCY	1997-2008	1400,2018

If all terms are added correctly, the barycentric arrival times should be equal for simultaneous measurements at multiple telescopes. Usually, the combined errors will result in differences between the individual TOAs measured at different telescopes. If one of the effects dominates over the others, the result on the combined timing residuals will look like a constant offset or a trend between two data sets. Only if all errors can be disentangled from each other, and corrected for, will it be possible to use pulsar timing for measurements that require large time spans and the highest precision TOAs available. We discuss below some of the details of what can be achieved with the existing EPTA data sets.

6.3 Results

Over the last few years a couple of papers have been published that were based on timing solutions for which TOAs from more than one, or even all, EPTA telescopes have been combined. For PSR J1518+4904 (Janssen et al. 2008a, Chapter 3), we had five separate data sets available. We used different smaller combinations of the data sets to try to determine specific parameters to better accuracy. For example, when refining binary parameters for a short-orbit system, better results may be gained by using a shorter data set that has the highest precision TOAs. On the other hand, determination of long-term parameters is only possible by using many years of observations. For this pulsar, in all cases the best result was gained by using the most extended data set. We note that for PSR J1518+4904 there were no large differences in TOA accuracy between the separate data sets. However, if the differences are larger, the most probable outcome is that using a combination with the best TOAs and longest time span will provide the most constrained parameters.

In this section we will describe a few procedures of combining data sets that were used to find the solutions as described in earlier EPTA papers. We investigate how important the effects as described in Sect. 6.2 are, and if there are dominant contributions of the various effects for pulsars with specific properties.

6.3.1 Offsets from template differences

The procedure that was followed to align the separate data sets can be considered as representative of other previously published timing solutions from EPTA combinations. For all timing analysis presented in this Chapter, we have used the timing software package TEMPO2 (Hobbs et al. 2006). Firstly, each individual data set was fitted separately to scale the errors on the TOAs of each telescope to return a reduced $\chi^2 \sim 1$. The same standard ephemeris was used for each individual data set. Because some of the data sets were limited in observing time span, we held values for long-term parameters like proper motion fixed, to prevent covariances affecting other parameters. Using similar reasoning, for observatories that observed the particular pulsar at one frequency only, the DM parameter was fixed. Apart from scaling the errors, this procedure provides a check for consistency between the data sets, and to exclude possible systematic errors like unmodelled cable delays or large clock errors. The parameter files should give similar results if the offsets are only caused by the use of different hardware, or by processing and analysis of the data using different software and templates. Finally, all individuall data sets were put together by inserting one constant offset for each telescope.

As described above, the largest expected difference between TOAs from separate telescopes comes from the fact that the reference point on the high S/N templates is an arbitrary choice. However, as the resulting offset will be constant as long as the templates are not changed, this can be easily confirmed by plotting the templates on top of each other. An example is shown in Fig. 6.1. Using a common standard which is scaled appropriately for the properties of each individual observing hardware and exact centre frequency will eliminate the profile offset term, ΔT_{prof}. As mentioned before, a study on the advantages of using synthetic profiles will be presented in a future paper (Purver et al. 2009, in prep.).

6.3.2 PSRs J1012+5307 and J1713+0747: monitoring offsets

Both PSRs J1012+5307 and J1713+0747 can be timed to high precision, and therefore the continuity of the offsets between telescopes may be monitored with better accuracy than for PSR J1518+4904. For both pulsars, we only used the observations that were taken around 1400 MHz from each telescope, as they generally yield the best timing accuracy. These pulsars have low DM values that are previously determined to high accuracy and have not shown significant DM variations before (Lange et al. 2001; Splaver et al. 2005).

We used a slightly different procedure in aligning the data sets for these pulsars. We first selected the occasions where the pulsars were observed coincidentally by two different telescopes within one day. To exclude biases in the timing solution, we did not include those TOAs when determining the initial timing solution. The resulting best-fit timing solution was

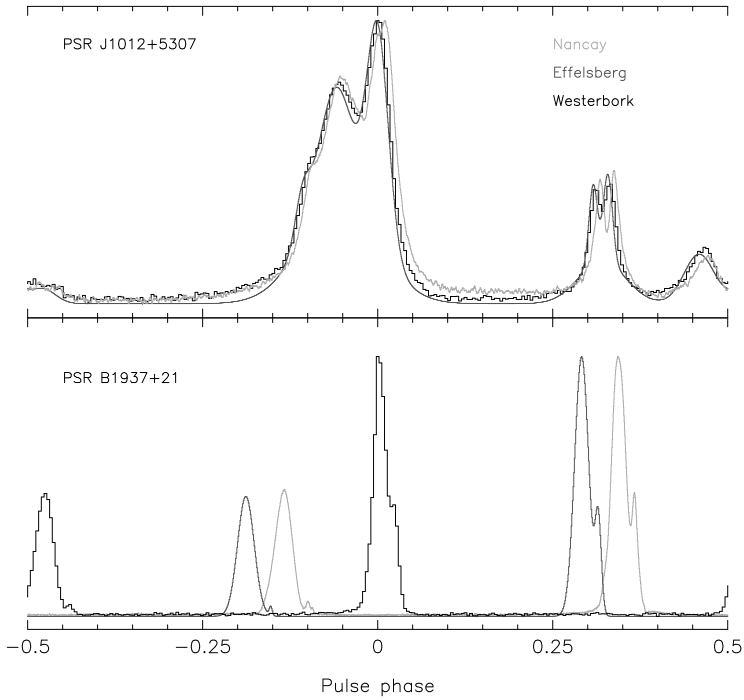

Figure 6.1: Template profiles as used for PSR J1012+5307 and PSR B1937+21 at Westerbork, Effelsberg and Nançay. For PSR J1012+5307 this figure shows that the WSRT and Effelsberg profiles are aligned and should not give an error. The small phase shift of the NRT profile is the cause of the jump required to align the TOA sets. For PSR B1937+21 the template offsets are larger, which is also visible in Fig. 6.4. Not only the phase shift, but also, for example, the subtle differences in the pulse width may cause small changes in the templates. If the shapes shown here represent the true instrumental response of the pulsar backends and telescopes, then any synthetic template system should take that into account.

fixed and then the coincident TOAs were reintroduced. The now complete TOA set was fitted using the aforementioned fixed timing solution, and for each occurence of a coincicent TOA pair, the offset for that day was measured. This offset was calculated by taking the differences of the residuals of both TOAs to the timing model, where the error on the offset was based on the uncertainties of the involved TOAs. If the separation between the data sets would be constant over time, and the timing model correct (optimal) for both data sets, the difference in postfit residuals would also be constant within the combined error.

Fig. 6.2 shows the offsets as measured for PSR J1012+5307 using the four long data sets that were available in the EPTA database. For all combinations, the offset was constant within the measurement errors. We note that for PSR J1012+5307, the reference point taken in the high S/N template was the same for the WSRT and Effelsberg 1400 MHz data, although the templates were generated from different data sets. This shows nicely that when the profiles are aligned, and no other large effects are present, the offset reduces to zero. The TOAs from Nançay and Jodrell Bank were not calculated with respect to the same reference point in their template profiles, and therefore showed a non-zero offset. Before measuring changes in the offsets, we inserted the template differences with respect to the WSRT template as fixed offsets into the timing solution. Then the remaining offsets between data sets as described above were calculated. The result is that, in Fig. 6.2, all deviations of offsets between observatories are grouped around zero. For all observatories, apart from the constant offset between templates as mentioned above, there was no indication of any deviation within the error. As can be seen in Fig. 6.2, the combined error from each of the TOAs in a pair is still quite large and thus only shows there is agreement to about the microsecond level. Higher precision TOAs would be needed to check for smaller effects on these timescales.

PSR J1713+0747 is among the best timed pulsars. Its very sharp profile, combined with excellent timing stability allows for individual TOA errors of less than 1 μs. Splaver et al. (2005) showed that using time spans of more than 12 years, an overall rms of the timing solution of less than 2 μs can be reached. Although the measurements of WSRT, Effelsberg and NRT give very good precision on individual TOA measurements, even with our combined data sets, the overall timing solution needed substantial scaling of the individual errors to result in a reduced $\chi^2 \sim 1$. Unfortunately, it was therefore not possible to monitor the telescope offsets to the sub-microsecond level precision that we expected for this pulsar. Again, we used a best-fit model with fixed parameters for a single data set, and then inserted additional data sets without fitting for a jump between the data sets. For this pulsar we did detect a change in the offset of about 3 μs between the WSRT and Effelsberg data around July 2005. Although we have no conclusive way of deciding, the distribution of residuals close to this epoch suggested that the effect is included in the WSRT TOAs. Possible explanations are changing templates, changing ephemerides or changes in sampling time or observing frequency of the observations. We did not find any evidence of such changes. Also, when compared with the offsets monitoring from other pulsars like J1012+5307 or B1937+21, which should be precise enough to confirm clock errors with that size, showed no WSRT clock error at that specific epoch. As this offset has no clear cause, and may be due to a combination of effects as described in Sect. 6.2 and a non-optimal timing solution, more work is needed to explain the differences in the residuals.

6.3.3 PSR B1937+21: DM variations and profile evolution

PSR B1937+21 was the first millisecond pulsar (MSP) discovered and therefore has been timed for a very long time at most observatories. Its profile is very suitable for high precision timing, as it is bright, has a narrow peak and an interpulse, see Fig. 6.1. However, the pulsar is known to show a lot of timing-noise-like behaviour, and possibly related DM variations on several timescales (Kaspi et al. 1994; Ramachandran et al. 2006).

We have three long multi-frequency data sets available for this pulsar, see Table 6.1. Aligning data sets is complicated for pulsars like PSR B1937+21, as the expected DM variations could influence our offset measurements. When fitting an offset to two sets of TOAs, TEMPO2 tries to minimise the difference between the total data sets. The result can be that

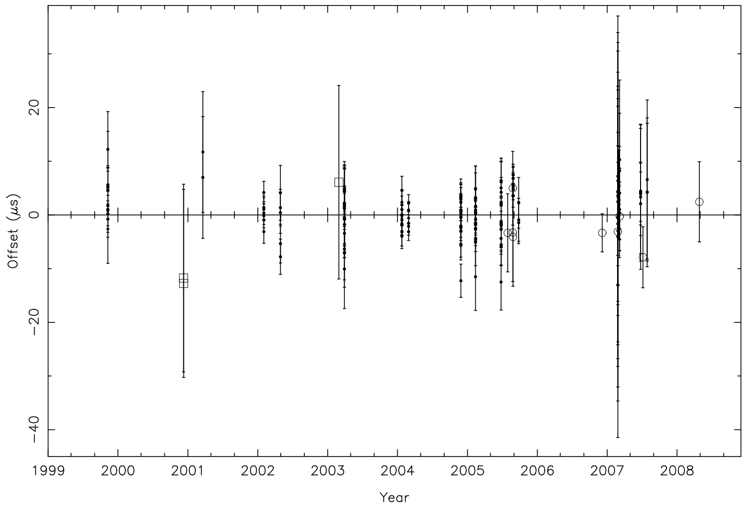

Figure 6.2: Offsets between TOAs from WSRT with respect to TOAs from other observatories (dots: Effelsberg; open circles: NRT; open squares: Jodrell Bank) for PSR J1012+5307. Each point represents the difference between TOAs from the two observatories, taken with a maximum of 1 day separation. Within the error, the offsets are constant over the whole observing time span. The error bars represent the uncertainties in the differences, calculated from the TOA errors corresponding to the observations that are used.

the offset absorbs part of other variations and the measured value is incorrect. To avoid including any DM variations in the offset measurements, we started the aligning process for PSR B1937+21 by using only the TOAs that were calculated from observations around 1400 MHz.

A Nançay clock error?

The three data sets were again first individually fitted to scale the errors to result in $\chi^2 \sim 1$ for each data set. In the top panel of Fig 6.3, the separate 1400 MHz data set of Nançay is shown. Due to unmodelled DM variations, the residuals still show timing noise on timescales of a few months to years. Although apparently a good timing solution, as shown in the middle panel of Fig. 6.3 when compared to the overlapping residuals of WSRT or Effelsberg, a small change in the offset of only 7.6 μs was detected around July/August 2003. Apart from this change the residuals follow the WSRT and Effelsberg data very nicely, indicating that the jump in residuals must have occured in an instantaneous event, that would not cause the offset to increase with time. At Nançay, all data before 2005 were processed with the same pulsar backend, and the same ephemeris and template were used to calculate the TOAs. This indicates that the Nançay offset is probably caused by an unmodelled clock error. This could be confirmed at a later stage by comparing TOAs for another pulsar timed with high precision at both Nançay and another telescope.

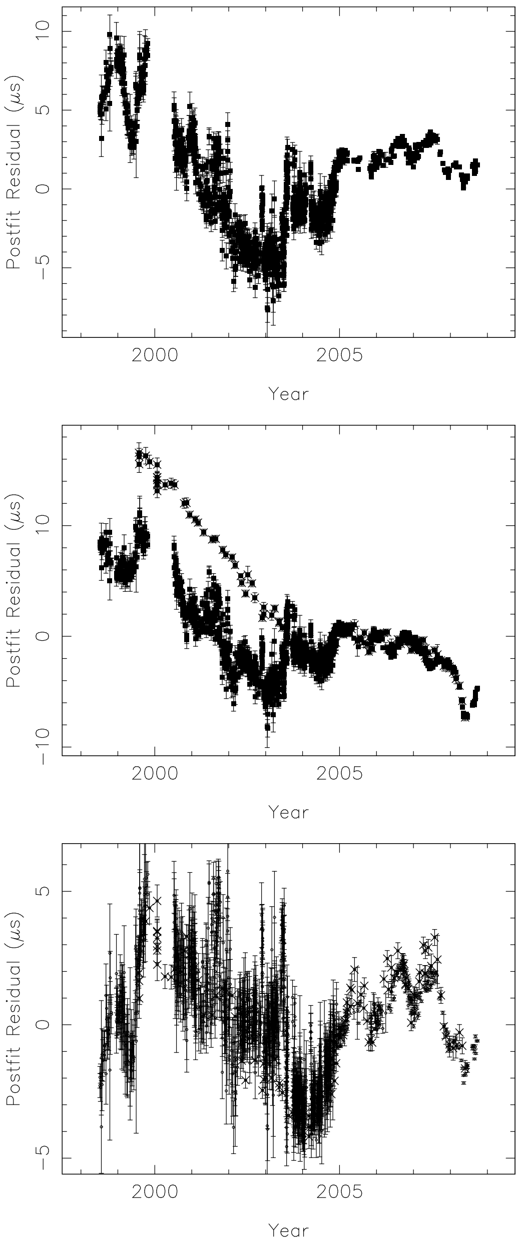

Figure 6.3: Top: Best-fit residuals for Nançay TOAs taken at 1400 MHz for PSR B1937+21. Middle: When combined with WSRT TOAs at similar observing frequency, it becomes visible that the constant offset that has been used to align the data sets is not representing the differences between the two data sets correctly, see Sect. 6.3.3 for a discussion. Bottom: Best-fit residuals for optimal alignment of all 1400 MHz TOAs of WSRT, Effelsberg and Nançay, with the extra offset corrected.

Although the explanation for this change in offset is not yet confirmed, it has to be stressed that only by using multi-telescope timing do we have an independent tool to find delays like this for pulsars that are timed to microsecond precision levels. If this is truly a clock error of this magnitude, it should manifest itself in more than one pulsar and thus should be able to be confirmed in an independent way by the individual observatory. However, if it were smaller or more long term, as mentioned above, it may only be seen when comparing between observatories, where it will show up as a trend. If the jump is actually specific to this pulsar it could be explained by changes in the hardware which affect the pulse shape, intrinsic time resolution, or a change in the template used for TOA determination.

Combining data with DM variations

After determining the constant offset for data sets with comparable centre frequencies, accounting for clock errors, and having found a representative timing solution for the combined set it may be possible to insert data points from additional frequencies and (while keeping offsets and other parameters fixed) fit for DM. If the DM is constant over time this should give no problems. We note that it is not always clear at what stage of the alignment procedure DM variations are best accounted for. In our example of PSR B1937+21, part of the DM variations may already have been absorbed in the telescope offsets, as all data sets, even though similar frequencies are used, have different centre frequencies. In contrast, DM variations are better measurable when a wider frequency range is available.

PSR B1937+21 is known to show DM variations. Previously, the variations could be fitted over a long timescale with a steady DM derivative (e.g. Kaspi et al. 1994; You et al. 2007a) however from Figs. 6.4 it is clear that there appears to be a cutoff in the slope around mid-2004.

The top plot in Fig. 6.4 shows the usual starting point when combining data sets, now showing TOAs from all frequencies. To show the separate data sets from the three observatories more clearly, we have inserted the offsets here manually and they do not represent the actual template offset sizes as discussed in Sect. 6.3.3. The middle plot shows the result after inserting a constant offset for each extra telescope and the correction for the Nançay event of 2003. The DM variations that are not yet accounted for are visible in both panels as deviations of the TOAs from additional frequencies (represented by diamonds and crosses).

To probe the changes in DM, we used the fitting routine "stridefit" in TEMPO2 that enables fitting small segments of data for a specified parameter. Effectively the DM is measured for consecutive overlapping segments of the data, and it is possible to adapt the time range of the fit and shifting timestep to optimise measuring the actual changes. If the fitted timerange is too small, the measured DM at that epoch may be dominated by TOAs from one observatory or be limited in the total frequency range that is probed. On the contrary, we need to avoid averaging out smaller variations by covering a range that is too large.

For 1937 we used fitting timespans of 300 days and shifted that range by 100 days at a time, the result is shown in Fig. 6.5. We corrected all TOAs to have the corresponding DM value and refitted the complete data set. The result is shown in the bottom panel of Fig. 6.4. It is clear that the procedure is not optimal: around 2004 and 2005, the TOAs appear to be overcorrected by a few μs on multiple occasions. This may be an effect of the DM

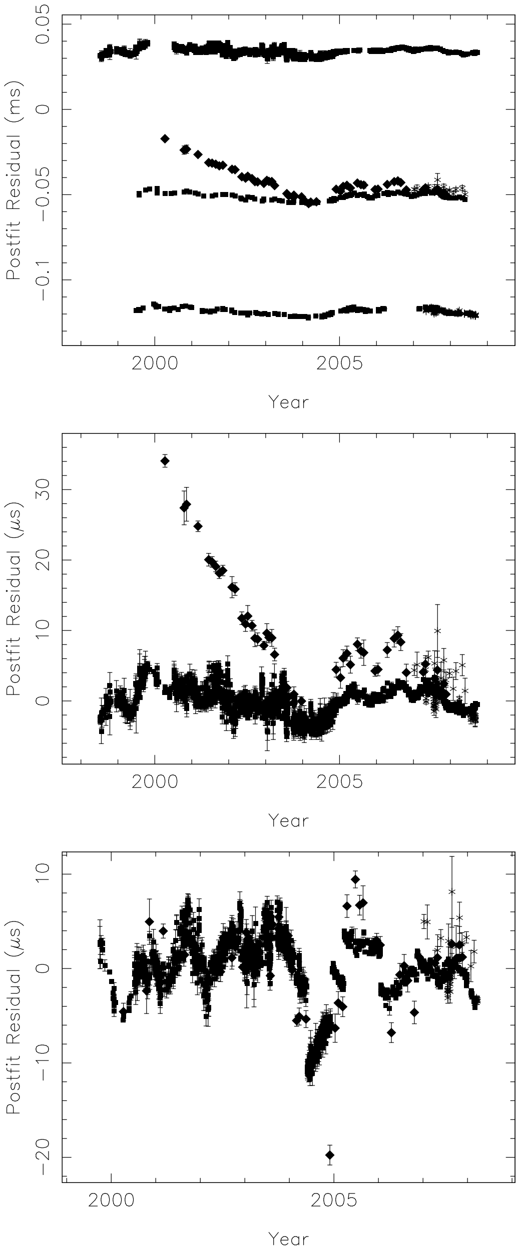

Figure 6.4: The procedure of fitting for DM variations for PSR B1937+21. The top panel shows the data sets grouped by telescope, additional frequencies are represented by different symbols. The offsets between the data sets are inserted manually to show more clearly which data sets have additional frequencies. From top to bottom: Nançay, Westerbork, Effelsberg. The middle plot shows all TOAs together, using the fixed constant offsets as determined by aligning the 1400 MHz TOAs (Fig. 6.3). The bottom plot shows the result of correcting for DM variations as determined from the stridefit which is described in Sect. 6.3.3.

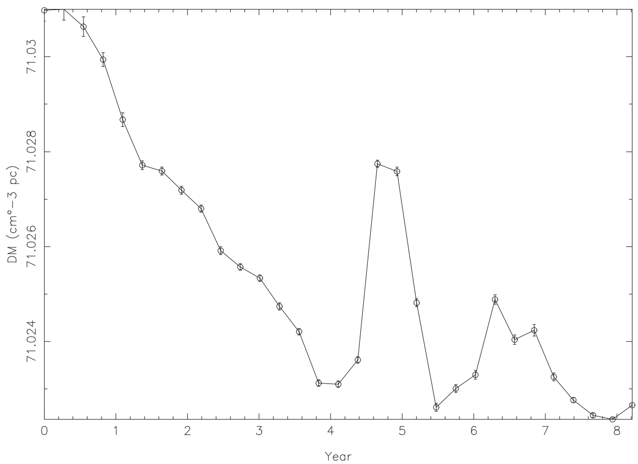

Figure 6.5: The result of the DM "stridefit" corresponding to Fig. 6.4. Each point represents a fit for DM for all TOAs in a range of 300 days around that epoch.

correction being measured by averaging too many TOAs, where any large DM variations on short timescales may affect later TOAs in the averaging process in an unsatisfactory way. This could be resolved by allowing the DM to change by small amounts on shorter timescales. However, we need to be careful not to solve every unmodelled feature with an extra DM step.

Frequency evolution of the pulse profile

Even though the 1400 MHz data were aligned carefully, and the DM variations appear to be monitorable to high precision, when using TOAs from three widely separated observing frequencies they could not all be aligned together for PSR B1937+21. In this example, most of the DM measurements were largely based on comparing TOAs from all telescopes taken at frequencies around 1400 MHz with 840 MHz TOAs from WSRT. Two things are to be noted from the bottom plot of Fig. 6.4: firstly, where there are not many 840 MHz observations available, the errors in DM are larger. Furthermore, the high-frequency data of the other telescopes does not align well with the other TOAs, even after DM correction. Explanations for these effects could be that either the DM variations and the template offsets are influencing each other too much, or the profile is actually evolving with frequency.

In the previous section, we have implicitly assumed that the profiles that were used to calculate TOAs at different observing frequencies were aligned correctly. However, profiles are not constant with frequency and to optimally align them, frequency evolution has to be taken into account. Furthermore, the quality of the template is affected by the way it it produced, and more investigation is needed on the optimal way of comparing templates.

It is necessary to implement frequency evolution in pulse profile templates, which can be achieved by using synthetic templates. An ongoing project will construct frequency-dependent profiles to be used for TOA calculations over a wide range of frequencies (Purver et al. 2009, in prep.). This will ultimately dissolve the factors ΔT_{freq} and ΔT_{prof} as shown in Eq. 6.1.

6.3.4 Simultaneous multi-telescope observations

Based on our experiences with the existing data sets we planned a number of observations with all EPTA telescopes where a set of pulsars were observed simultaneously, or quasi-simultaneously. As shown in Sect. 6.3.2, so far all offsets that were calculated from (quasi-) simultaneous observations were dominated by errors in the TOA calculation itself. However, when TOA accuracy improves, for example when coherent dedispersion becomes standard for all observatories, regular monitoring of offsets will give better insight in the different delays as presented in Eq. 6.1.

While not sensitive to small deviations, the simultaneous observations showed that they can be carried out and should be, and that there were no larger effects that had been somehow missed previously. When better levels of accuracy are within reach, we can think of several projects to monitor: what are the variations in offsets per pulsar when using multiple days of simultaneous observations? Is there any correlation in offsets for different pulsars? How accurate can TOAs be determined when all delays are measured separately and accounted for?

The majority of the telescopes in the EPTA are, or will soon be, also used for VLBI observations. This offers an interesting opportunity to correct for the cable delay terms in Eq. 6.1. Using the delays required for finding fringes in the VLBI data, we can model the majority of the time delay up to the point where the data streams are split for VLBI and the pulsar backends. For example, early experiments with the WSRT showed that there were cable delays of around 12 μs contributing to both VLBI and pulsar observations. Determining such delays is also going to be essential for the LEAP project which will use the delays between the different observatories to combine the telescopes of the EPTA coherently. We note also that using and determining such delays requires at least one pair of observatories to observe at the same frequency. As at this moment the WSRT is the only telescope that is capable of observing at frequencies below 400 MHz, for the lower frequency observations this may be more problematic.

6.4 Conclusions

One of the goals of the EPTA is to share data to improve science on individual systems. Using the maximum number of TOAs that are available for finding a timing solution gives in most cases the best results. However, when combining pulsar timing data from multiple telescopes, great caution has to be taken to the procedure of aligning the data sets. Depending on the precision of the measurements, and the properties of the pulsar and its sensitivity to interstellar medium effects, the procedure is as follows:

- Use the data set that covers the longest time span to find a general timing solution which can be used to scale the errors for each individual data set.

- Fit all data sets separately to obtain appropriate scaling factors for TOAs from different data sets. For data sets covering a limited time span, it may be necessary to fix some parameters that need long-term coverage to be fitted (e.g. proper motion).

- Use a representative timing model with all parameters fixed to determine offsets between data sets. When there is a hint of changes in the offsets, determine the (quasi-) simultaneous observations and use the differences in their post-fit residuals (with respect to the fixed timing model) to probe the size of the offset over the total time range.

- When multiple frequencies (per observatory) are available, DM variations can be traced and resolved.

As shown in Fig. 6.4, to account for DM variations for pulsars like PSR B1937+21 it is necessary to take great caution in aligning data sets and measuring the DM at different epochs. To distinguish the DM variations from all other effects as presented in Eq.6.1, multi-frequency multi-telescope observations are needed on a regular basis. We have shown that by carefully aligning data sets, accounting for clock errors and using DM-corrected TOAs, we are able to reach sub-microsecond rms levels over a long term period for multi-telescope data. Simultaneous observations at similar centre frequencies, at regular time intervals will allow for finding local clock errors and cable delays. Using the same folding ephemerides for pulsars across observatories should then result in equal TOAs at the SSB.

We have shown that using data sets from multiple telescopes can greatly improve the timing precision achievable. However, it is clear that it also introduces a number of extra fit parameters that have to be carefully considered and modelled. Ideally, we want to be able to know all the terms in Eq. 6.1 before we combine the data sets, and we have given some recipes for achieving that. We also note that with 3 of the 5 telescopes in the EPTA going over to using the same hardware, and the other two both already using coherent dedispersion systems, that a number of the problems highlighted here will be resolved. With care and due attention the combination of data sets from multiple telescopes can be optimally achieved, and therefore brings the goal of detecting gravitational waves much closer.