

Scientific Data Processing of a Fiber Network for Optical Frequency Transfer: Methods and Studies

Mads Tønnes, Etienne Cantin, Dan Xu, Olivier Lopez, Anne Amy-Klein, and Paul-Éric Pottie

Abstract – In recent years, the use of optical frequency dissemination through telecommunication fibers has steadily grown both in the context of state-of-the-art metrology like optical clocks and in the use of ultrastable reference signals for precision measurements in other fields of physics or earth science. Evaluating the performance of the fiber link can be difficult, as outliers and dysfunctioning regions lead to missing data. In this article, we discuss these challenges and investigate three different approaches to dealing with the missing data of a 1410-km-long optical link. We find a benefit in replacing the missing data with simulated noise with a statistical behavior similar to that of the original signal.

1. Introduction

Coherent optical fiber links is a rapidly emerging technology, enabling ultrastable optical frequency transfer over thousands of kilometers. Compared to the traditional methods of frequency transfer through the Global Navigation and Satellite Services (GNSS), it operates with orders-of-magnitude lower noise. This makes it suitable for the comparison of optical clocks and cavities [1], very long baseline interferometry (VLBI) [2], spectroscopy [3], along with many other applications. In France in particular, the research infrastructure Réseau fibré métrologique à vocation européenne aims at providing an ultrastable frequency reference to more than 32 partner laboratories throughout France [4]. It is evident that such a network will

Manuscript received 6 December 2021.

Mads Tønnes is with LNE-SYRTE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, 61 Avenue de l'Observatoire, 75014 Paris, France; e-mail: mads.tonnes@obspm.fr.

Etienne Cantin is with Laboratoire de Physique des Lasers, Université Sorbonne Paris Nord, CNRS, 99 Avenue Jean-Baptiste Clément, 93430 Villetaneuse, France; e-mail: etienne.cantin@univ-paris13.fr.

Dan Xu is with LNE-SYRTE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, 61 Avenue de l'Observatoire, 75014 Paris, France; e-mail: dan.xu@obspm.fr.

Olivier Lopez is with Laboratoire de Physique des Lasers, Université Sorbonne Paris Nord, CNRS, 99 Avenue Jean-Baptiste Clément, 93430 Villetaneuse, France; e-mail: olivier.lopez@univ-paris13.fr.

Anne Amy-Klein is with Laboratoire de Physique des Lasers, Université Sorbonne Paris Nord, CNRS, 99 Avenue Jean-Baptiste Clément, 93430 Villetaneuse, France; e-mail: amy@univ-paris13.fr.

Paul-Éric Pottie is with LNE-SYRTE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, 61 Avenue de l'Observatoire, 75014 Paris, France; e-mail: paul-eric.pottie@obspm.fr.

consist of many branches and have several simultaneous users. Supervision and data processing of the full network are therefore of the utmost importance. In the following, we will consider problems occurring when evaluating the stability of the frequency dissemination in the case of missing data. Similar concerns have been considered for GNSS [5–7], where the main difference to this work is that the fiber links tend to operate in the phase-coherent regime, leading to complexities in the evaluation of the phase of the signal in the case of missing data. In [8], a method of applying a correction to the Allan variance of a data set with missing data is presented.

2. The Challenge of Missing Data

In this article, we show a case study of a 2×705 -km-long optical fiber link between the two French cities of Paris and Strasbourg. The link is actively compensated, and the optical signal is regenerated three times along the way [9]. We measure the end-to-end signal of the 1410-km link, which is the frequency difference between the original signal and the disseminated light. This is the only data we are discussing in this article.

The end-to-end frequency is measured using a dead-time free frequency counter. This measures the optical beat $f(t)$ continuously. However, due to technical challenges, not all the recorded data will be valid in the end, leading to gaps in the frequency data [4]. When calculating the phase evolution of the signal in postprocessing, we usually express the phase in terms of the *time error*, $x(t)$, which has the unit of seconds. The time error is calculated from the relative frequency fluctuations $y(t) = \frac{f(t) - \nu_0}{\nu_0}$, with ν_0 being the nominal frequency, by the integral

$$x(t) = \int_{t_1}^{t_2} y(t) dt, \quad (1)$$

with $t_2 - t_1$ being the time interval of the integration. As this integral assumes a continuous frequency measurement, the periods of missing data will lead to unknown phase evolution, impairing the phase coherence of the link. This is problematic for the postprocessing of the data. Therefore, we need to have an evaluation of the effects of missing data and the bias it introduces in the statistical evaluation of the fiber link. If we want to be able to evaluate the long-term behavior of the link, we will need to deal with the missing data in a way that biases the evaluation of the link as little as possible.

3. Comparison of Methods Dealing With Missing Data

We investigated three different methods of dealing with missing data in the link signal. To analyze the effects of each method, we use a reference data set that is perfectly continuous. In a controlled way, we remove data points at random, deal with the missing data each of three different ways, analyze the statistical consequences, and compare the results. In the following, we discuss three different approaches: concatenation of the resulting data set, keeping the phase constant across the gap left by the missing data, and the more novel approach of replacing the missing data with simulated data, which follows a simple noise model of the link.

The most important criterion on which we are evaluating each method is how well they are able to keep the statistical behavior as close as possible to the reference. To make this assessment, we use two different tools to analyze the signals. The first is the phase power spectral density, which is calculated from the Fourier transform of the autocorrelation function of the time error, $S_{\phi}(f)$. This is used to analyze the amount of phase noise at a given Fourier frequency and is a good method of determining periodic perturbations to the signal. The second tool we use is the modified Allan deviation, $\text{Mod } \sigma_y(\tau)$, calculating the stability of the signal as a function of integration time τ [10]. This is a tool commonly used in time and frequency metrology, and in this context, we will use it to evaluate the phase coherence of the link both with and without missing data.

3.1 Concatenating the Remaining Data

One can choose to ignore the problem altogether. This effectively concatenates the data set, creating a length mismatch as compared to the reference. This creates a bias in the Fourier domain, where it becomes difficult to identify periodic perturbations to the signal. This is illustrated in Figure 1, where we show the phase noise of the reference signal (blue) with the phase noise of the three different methods. It is clear that many periodic peaks have either vanished or are blue-shifted (red curve) after removing 10% of the data.

Finally, this method does not address the problem of the deterioration of the phase coherence, which can be seen in the deterioration of the stability for an increasing amount of removed data, as shown in Figure 2.

3.2 Keeping the Phase Constant

A second method is to keep the phase constant across the gap left by the missing data. From a conceptual point of view, this is more reasonable, as it keeps the length of the data set intact, which makes for a more reasonable Fourier analysis, as seen by the peaks in the yellow trace in Figure 1. However, it still has the problem of not addressing the deterioration of the coherence, and having periods of a constant phase

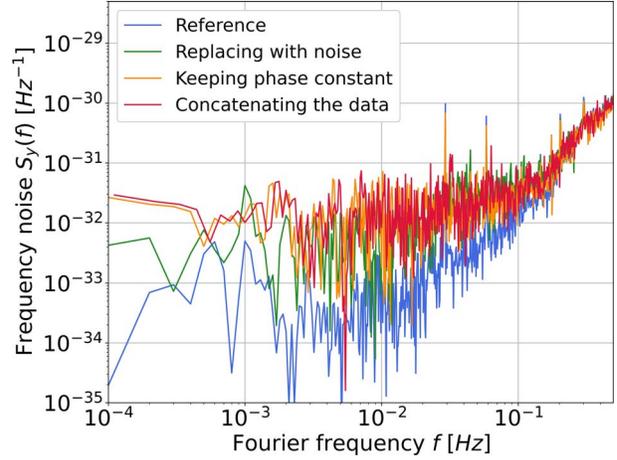


Figure 1. Phase noise of the reference signal (blue) and the three different methods of dealing with missing data. Ten percent of the data has been removed, and the data set has been concatenated (red), the phase has been kept constant across the missing data (orange), and the missing data have been replaced with statistically similar simulated data (green).

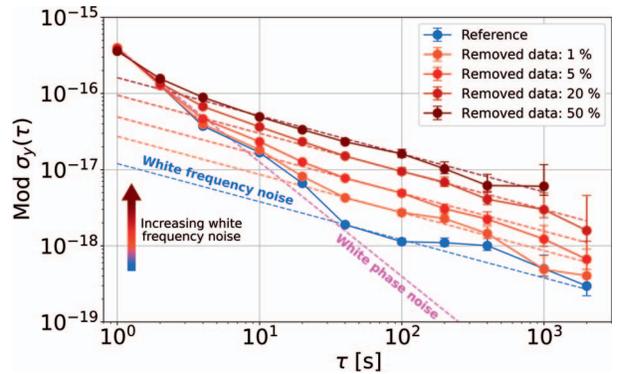


Figure 2. Modified Allan deviation of the reference signal without missing data (blue) and the concatenated signal of an increasing amount of removed data (increasingly dark red).

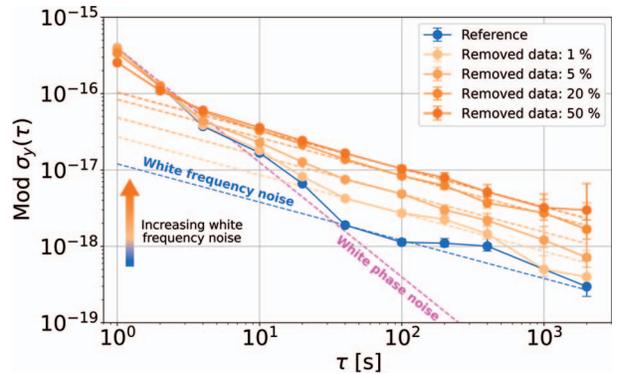


Figure 3. Modified Allan deviation of the reference signal without missing data (blue) and with an increasing amount of removed data, where the phase is kept constant across the gap of the missing data (increasingly dark orange).

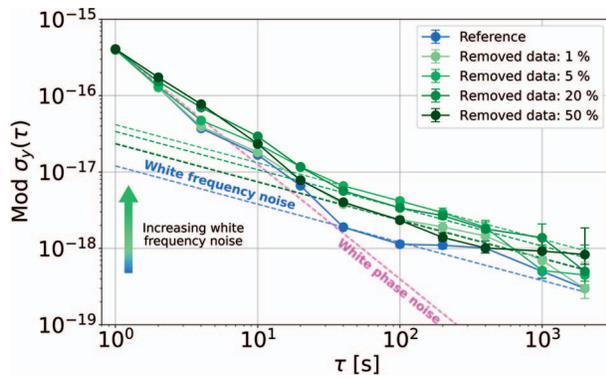


Figure 4. Modified Allan deviation of the reference signal without missing data (blue) and with an increasing amount of removed data, where removed data are replaced with simulated noise (increasingly dark green).

will bias the short-term noise level, which is seen in the stability curves in Figure 3.

3.3 Replacing With a Noise Model

The last method we investigated is a more novel approach: we made a simple method to produce noise with statistical behavior largely comparable to that of the link, consisting of white phase noise and white frequency noise. A more detailed model will be described in a future publication. We used this model to simulate noise that we could place instead of the missing data, effectively “filling” up the gaps in the data set. In this way, we kept the original length of the data set, and we avoided the bias in the short-term noise level. We furthermore observed a decrease in the deterioration of the coherence of the link, which is seen in the modified Allan deviation in Figure 4: the increase in white frequency noise with an increased amount of missing data is lower. This directly corresponds to a lower effective decoherence of the data as compared to the two previous cases shown in Figures 2 and 3.

4. Conclusion

We have discussed the challenges with missing data in the data analysis of coherent optical fiber links. We have compared different approaches of handling the missing data and found that it is possible to reduce the aliasing effects by replacing the missing data with a simple noise model with similar statistical behavior as that of the fiber link itself. This is an important step for

exploring long-term effects on fiber links, where missing data are inevitable, and finding line identifications in their Fourier analysis at very low frequency. We believe that this work may find its application in earth science and fundamental science, like the exploration of the Sagnac effect and other tests of relativity [11] or the search for dark matter.

5. References

1. C. Lisdat, G. Grosche, N. Quintin, C. Shi, S. M. F. Raupach et al., “A Clock Network for Geodesy and Fundamental Science,” *Nature Communications*, **7**, August 2016, pp. 12443.
2. C. Clivati, R. Aiello, G. Bianco, C. Bortolotti, P. De Natale, et al., “Common-Clock Very Long Baseline Interferometry Using a Coherent Optical Fiber Link,” *Optica*, **7**, 2020, pp. 1031-1037.
3. R. Santagata, D. B. A. Tran, B. Argence, O. Lopez, S. K. Tokunaga, et al., “High-Precision Methanol Spectroscopy With a Widely Tunable SI-Traceable Frequency-Comb-Based Mid-Infrared QCL,” *Optica*, **6**, 2019, pp. 411-423.
4. E. Cantin, M. Tønnes, R. Le Targat, A. Amy-Klein, O. Lopez, et al., “An Accurate and Robust Metrological Network for Coherent Optical Frequency Dissemination,” *New Journal of Physics*, **23**, March 2021, pp. 053027.
5. I. Sesia and P. Tavella, “Estimating the Allan Variance in the Presence of Long Periods of Missing Data and Outliers,” *Metrologia*, **45**, December 2008, pp. S134-S142.
6. I. Sesia, E. Cantoni, A. Cernigliaro, G. Signorile, G. Fantino, et al., “An Efficient and Configurable Preprocessing Algorithm to Improve Stability Analysis,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **63**, October 2015, pp. 575-581.
7. C. Xu, “Reconstruction of Gappy GPS Coordinate Time Series Using Empirical Orthogonal Functions,” *Journal of Geophysical Research: Solid Earth*, **121**, December 2016, pp. 9020-9033.
8. L. Galleani and I. Sesia, “The Corrected Allan Variance: Stability Analysis of Frequency Measurements With Missing Data,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **66**, July 2019, pp. 1667-1683.
9. N. Chiodo, N. Quintin, F. Stefani, F. Wiotte, E. Camisard, et al., “Cascaded Optical Fiber Link Using the Internet Network for Remote Clocks Comparison,” *Optics Express*, **23**, 2015, pp. 33927-33937.
10. D. W. Allan and J. A. Barnes, “A Modified ‘Allan Variance’ with Increased Oscillator Characterization Ability,” *Thirty Fifth Annual Frequency Control Symposium*, May 1981, pp. 470-475.
11. P. Delva, J. Lodewyck, S. Bilicki, E. Bookjans, G. Vallet et al., “Test of Special Relativity Using a Fiber Network of Optical Clocks,” *Physical Review Letters*, **118**, June 2017, 221102.