Incorporating an Optical Clock into a Time Scale at NIST: Simulations and Preliminary Real-Data Analysis

Jian Yao, Jeffrey Sherman, Tara Fortier, Thomas Parker, Judah Levine, Joshua Savory, Stefania Romisch, William McGrew, Xiaogang Zhang, Daniele Nicolodi, Robert Fasano, Stephan Schaeffer, Kyle Beloy, and Andrew Ludlow

Time and Frequency Division, National Institute of Standards and Technology (NIST), Boulder, Colorado, USA

Abstract

This paper describes the recent NIST work on incorporating an optical clock into a time scale. We simulate a time scale composed of continuously-operating commercial hydrogen masers and an optical frequency standard that does not operate continuously as a clock. The simulations indicate that to achieve the same performance of a continuously-operating Cs-fountain time scale, it is necessary to run an optical frequency standard 12 min per half a day, or 1 hour per day, or 4 hours per 2.33 day, or 12 hours per week. Following the simulations, a Yb optical clock at NIST was frequently operated during the periods of 2017 March – April and 2017 late October – late December. During this operation, comb-mediated measurements between the Yb clock and a hydrogen maser had durations ranging from a few minutes to a few hours, depending on the experimental arrangements. This paper analyzes these real data preliminarily, and discusses the results. More data are needed to make a more complete assessment.

Key Words

Time scale, optical clock, Cs fountain, hydrogen maser, Kalman filter.

I. Introduction

The Al+ [1] and the Sr/Yb [2-3] frequency standards show a stability of better than 1×10^{-16} at an averaging time of several minutes, and even more stable devices are being developed. However, it is challenging to run an optical frequency standard continuously as a clock, making it difficult to incorporate such a device into a conventional time scale, which is based on time differences. In this paper, we present the results of a simulation of a time scale that can accept a device that provides only occasional frequency data in addition to the more traditional time-differences. We also show the preliminary result of incorporating this type of data into the time scale that is used to realize UTC(NIST).

Section II presents the details of the simulation. Also, we describe a time scale using the Kalman filter that can realize the combination of devices we have simulated. Section III presents the simulation results of a time scale that incorporates an optical clock. In particular, we show how long and how often an optical frequency standard must be run in order to make the time scale comparable to the Cs-fountain time scale. The simulated scale is based on free-running continuously-operating hydrogen masers whose characteristics are derived from the real masers in the NIST clock ensemble. Section IV shows the preliminary result of incorporating a real optical clock, i.e., Yb clock, into the NIST time scale. The preliminary real-data result matches the simulation. Section V concludes this paper.

II. Simulation Procedures

There are three steps in the simulation. First, we simulate a free-running time scale, a Cs fountain, and an optical frequency standard. The time scale algorithm is based on a measurement of the time differences every 720 s, which is the measurement interval that is used in the NIST ensemble. A conventional Kalman filter algorithm is used to estimate the frequency and frequency drift of the free-running time scale by using data from either the optical device or the cesium fountain. The frequency stability of the steered time scale is computed.

II(a). Clock Simulation

We model the characteristics of a "typical" hydrogen maser by using the data from our clock ensemble. The blue curve in Fig. 1 shows the Modified Allan Deviation of the measured time differences between two such masers. We assume that both masers have the same characteristics, so that the red curve in the figure, which is scaled by $1/\sqrt{2}$ relative to the blue curve, is our estimate of the stability of a single maser. The dotted lines in the figure show our model of the observed stability.
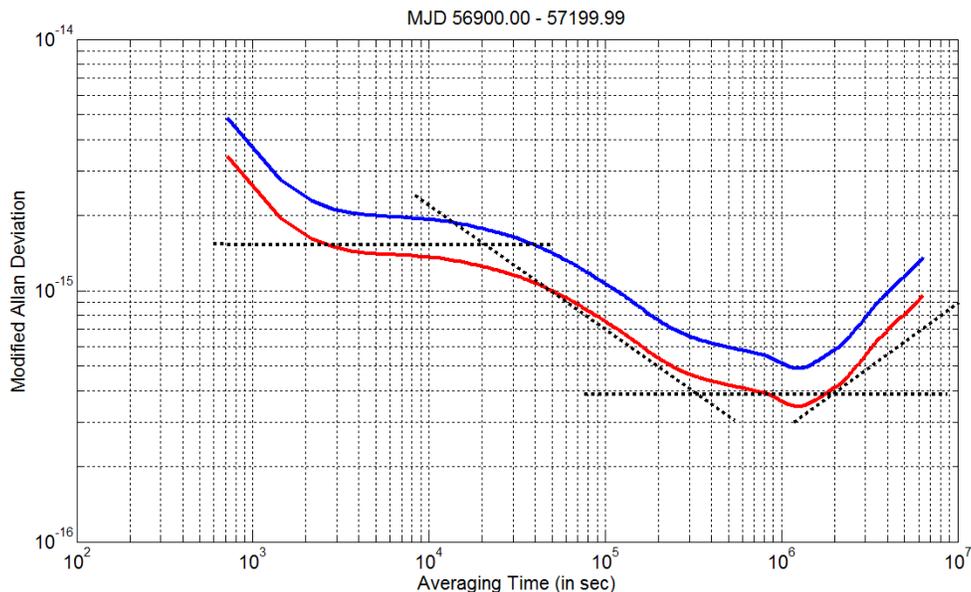

Figure 1. The blue curve shows the frequency stability of the time difference between two NIST H-masers. The red curve shows the estimated frequency stability of a single H-maser, assuming that the two masers have the same statistical characteristics. We model the stability as shown by the black dotted lines.

We generate a time series containing the sum of white frequency noise, flicker frequency noise and random-walk frequency noise with noise parameters chosen so that the time scale, which consists of N identical masers, is more stable than a single H-maser by a factor of $1/\sqrt{N}$. (This improvement is rigorously true for white noise processes where there is no correlation between the noise contributions of each of the masers. Our experience is that it is also true for any noise process in an ensemble with no correlations.) Second, we apply the phenomenologically-determined moving average algorithm in Eq. (1). This algorithm is chosen so as to make the frequency-stability curve flat for averaging times less than or equal to 20 000 s so that the simulation is close to the statistics of our masers. (The white phase noise for averaging times less than 2000 s in Fig. 1. is included in the model for the time difference measurements.) In the following equation, $x(n)$ denotes the generated data value at epoch n and $y(n)$ denotes the output of the averaging process at the same epoch. The Modified Allan Deviation of the simulated maser is shown by Figure 2. An ensemble of N masers would have a Modified Allan Deviation that was smaller by a factor of $1/\sqrt{N}$.

$$y(n) = \frac{1}{3}x(n) + \frac{2}{3} \cdot \frac{1}{3}x(n-1) + \left(\frac{2}{3}\right)^2 \cdot \frac{1}{3}x(n-2) + \left(\frac{2}{3}\right)^3 \cdot \frac{1}{3}x(n-3) + \cdots \tag{1}$$
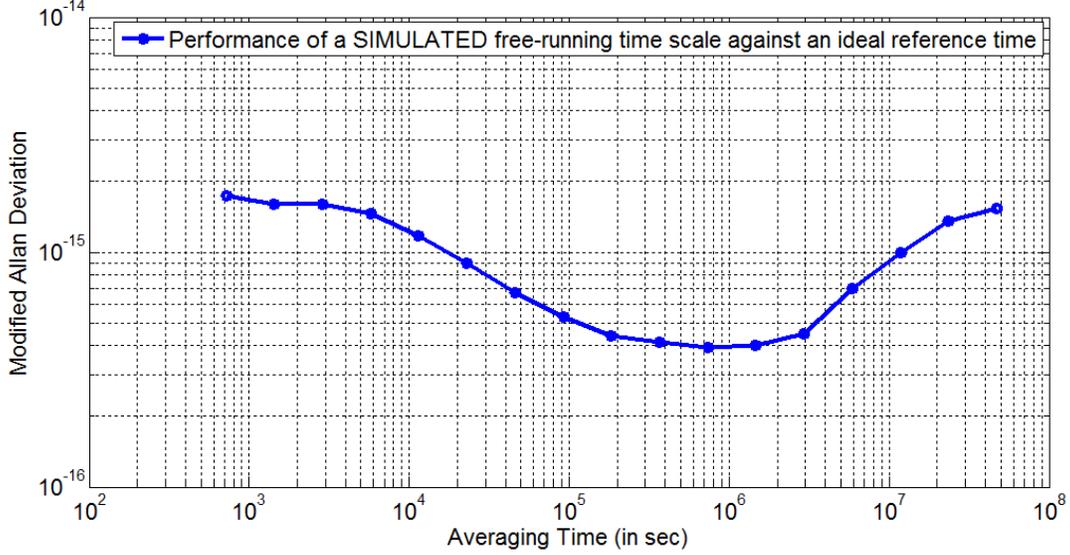
Figure 2. Frequency stability of a simulated free-running time scale with one maser against an ideal reference time.

The continuously-operating cesium fountain is simulated by white frequency noise with a fractional frequency stability of 1.4×10^{-15} at 1 day, which is typical Cs fountain performance at low atom density [4-6]. We simulate an optical clock as having white frequency noise with fractional frequency stability of 3.4×10^{-16} at 1 s, which is typical for the Sr/Yb optical-lattice clock [2-3]. The simulated optical-clock data are truncated to match various operational schedules as discussed below.

II(b). Steering Algorithm

The Kalman filter is used to estimate the frequency and frequency drift of the free-running time scale with respect to the Cs fountain or the optical clock, and the frequency and frequency drift of the time scale are steered based on this estimate.

There are two basic equations in the Kalman filter [7]. Eq. (2) is the system model, which predicts the state of the system at epoch $k+1$ based on its state at epoch k. Here, $X(k)$ is the estimate state vector of the system at epoch k, and $X(k+1|k)$ is the predicted state vector of the system at epoch $k+1$. Φ is the transition matrix, which links $X(k)$ and $X(k+1|k)$; it is determined by the physical properties of the system. u is the process noise, which is characterized by the Q matrix. Eq. (3) is the measurement model. The H matrix gives the relation between the state vector X and the measurement vector Z. v is the measurement noise, which is characterized by the R matrix. The R matrix is determined by the details of the measurement process, and its detailed specification is outside of the scope of this simulation. We assume that the white frequency noise of the maser will make a significant contribution to the R matrix, and that other contributions will be much smaller.

$$X(k + 1|k) = \Phi \cdot X(k) + u, \quad u \sim N(0, Q). \tag{2}$$

$$Z(k + 1) = H \cdot X(k + 1|k) + v, \quad v \sim N(0, R). \tag{3}$$

The estimated state vector at epoch $k+1$ is

$$X(k + 1) = X(k + 1|k) + K \cdot (Z(k + 1) - H \cdot X(k + 1|k)), \tag{4}$$

where K is the Kalman gain matrix, which is given by [7, Chapter 4],

$$K = P(k + 1|k) \cdot H^T \cdot [H \cdot P(k + 1|k) \cdot H^T + R]^{-1}, \tag{5}$$

where $P(k + 1|k) = \Phi \cdot P(k) \cdot \Phi^T + Q$, and $P(k + 1) = [I - K \cdot H] \cdot P(k + 1|k)$.

For our clock system, Eq. (2) becomes

$$\begin{pmatrix} f_{diff}(k+1|k) \\ d_{diff}(k+1|k) \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} f_{diff}(k) \\ d_{diff}(k) \end{pmatrix} + \begin{pmatrix} \eta \\ \zeta \end{pmatrix}, \tag{6}$$

where f_{diff} is the fractional frequency difference between the time scale and the external frequency standard, d_{diff} is the fractional frequency-drift difference, and Δt is the time interval between epoch k and epoch $k+1$. The constant time interval is 720 s. The parameters η and ζ are the noise terms in frequency and frequency drift, respectively [8-9]. Since the frequency noise of the external frequency standard is assumed to be negligible, η mainly comes from the frequency noise of the free-running time scale. ζ is typically too small to be observed for a measurement time interval of 720 s. Thus, we set the variance of ζ to a small value. In other words, the model assumes that the frequency drift is essentially constant over the measurement interval of 720 s.

If the Cs fountain is used as the reference, the measurement vector Z is the scalar frequency difference between the time scale and the Cs fountain measured every 720 s. If the optical clock is used as the reference, we do the pre-processing to get the average frequency difference between the time scale and the optical clock during the operation period of the optical clock. (As discussed in the previous paragraph, this assumes that the frequency drift of the maser is constant during this interval.) Then we use the average frequency difference as the measurement vector Z.

The measurement process provides an estimate of frequency, so that the H matrix of the measurement system is

$$H = (1 \quad 0). \tag{7}$$

The frequency and frequency drift at epoch $k+1$ are given by Eq. 4

$$\begin{pmatrix} f_{diff}(k+1) \\ d_{diff}(k+1) \end{pmatrix} = \begin{pmatrix} f_{diff}(k+1|k) \\ d_{diff}(k+1|k) \end{pmatrix} + K \cdot \left(Z(k+1) - H \cdot \begin{pmatrix} f_{diff}(k+1|k) \\ d_{diff}(k+1|k) \end{pmatrix} \right), \tag{8}$$

where K is the Kalman gain matrix, which can be calculated using Eq. (5).

Finally, we use the updated estimates to steer the time scale by adjusting its frequency and frequency drift.

III.   Simulation Results

III(a). Simulated Performance of a Cs-Fountain Time Scale

In this section, we first show the simulated performance of a time scale steered to a continuously-operating cesium fountain. The results of the simulation are shown in fig. 3. The green dashed line is the estimate of the stability of a typical cesium fountain operating at low atom density. The solid curves show the stability of a steered time scale containing one or four masers. The stability of an ensemble of N masers would be scaled by a factor of $1/\sqrt{N}$ relative to the one-maser values.
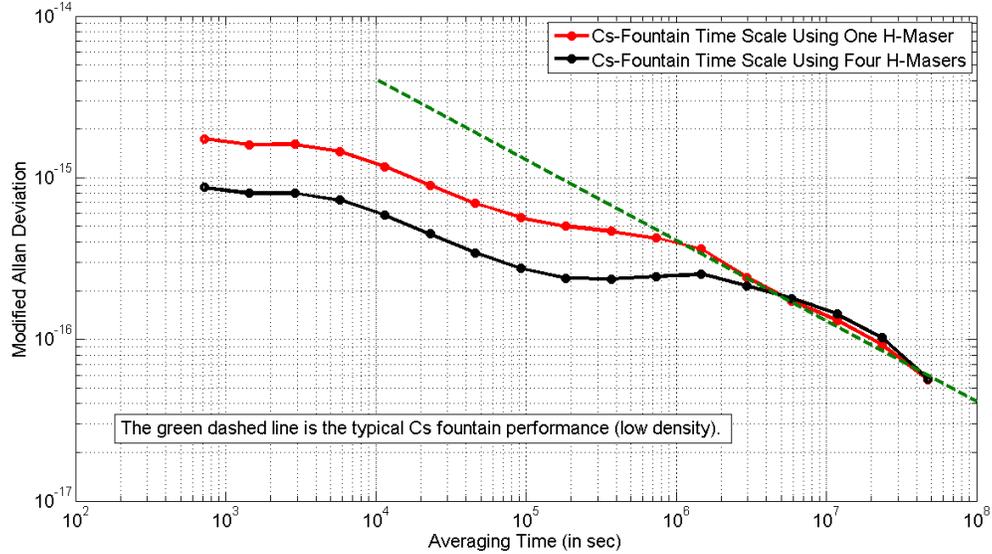
Figure 3. Performance of a time scale steered to a continuously-operating cesium fountain. The green dashed line shows the performance of the cesium fountain. The solid curves show the stability of a steered time scale that contains either one or four masers.

All of the ensembles are more stable than the cesium fountain standard for averaging times less than about 10 days. The figure illustrates the statistical advantage of steering an ensemble of multiple masers rather than a single device. This point is discussed in more detail in section V. This advantage is in addition to the ability of an ensemble algorithm to minimize the impact of the failure of any one of the masers and to provide a flywheel should the link to the cesium fountain fail for any reason.

A real-world Cs-fountain time scale, such as UTC(OP) [6] and UTC(PTB) [5], has a quite similar performance to the red curve in Figure 3. For example, analyzing the UTC(OP) data during Modified Julian Date 56649 – 57514 indicates that the frequency stability of UTC(OP) is about 4.8×10^{-16} at 10 days, and 2.4×10^{-16} at 60 days. The small difference between UTC(OP) and our simulation may come from the time-transfer noise and the fact that the Cs fountain does not run 100 % of the time.

III(b). Simulated Performance of a Time Scale with an Intermittently-Operating Optical Clock

In Figure 4, the orange dotted line is the modeled stability of a Sr/Yb optical clock, with a stability dominated by white frequency noise with an amplitude of 3.4×10^{-16} at 1 s. The green dotted line is the estimate of the stability of a continuously-operating cesium fountain as in the previous figures. The solid curves show the stability of a time scale ensemble of one maser when its frequency and frequency drift are estimated and corrected from the optical clock data.

In the upper left plot of Figure 4, the optical clock runs every half a day for either 12 or 24 minutes. In both of these cases, we can achieve the same performance as that of a continuously-operating Cs fountain time scale. The operation time of the optical clock is only 1.66 %, for the blue curve. We can also see that there is little improvement if we double the running time to 24 min every 12 hours. This is due to the maser's flat noise characteristics at times on the order of 1000 s (see Figure 2). The upper right plot shows the simulated performance of the same time scale with an optical clock running for various time intervals once a day. If we run an optical clock 12 hours a day (grey curve), the frequency stability of the time scale will be improved to three times better than the Cs-fountain time scale for an averaging time of greater than 10 days. Even with a 50 % duty cycle, the stability of the steered time scale is significantly worse than the stability of the optical clock itself. The bottom left plot shows the simulated performance of the same time scale with an optical clock running three times every week. To achieve performance similar to that of a Cs-fountain time scale, we need to run an optical clock for at least four hours, three times a week. The bottom right plot shows the simulated performance of the steered time scale with an optical clock running once a week. The

optical clock must operate for 12 hours each week to make the optical-clock time scale comparable to a Cs-fountain time scale. Practically, a free-running time scale can be significantly pulled by an abnormal hydrogen maser, and the error becomes larger over time. Thus, a frequent evaluation of the free-running time scale using the optical clock is necessary to avoid a large timing error. From this perspective, the actual performance of a time scale steered once a week is unlikely to be as good as the simulation indicated in the bottom right plot.

Figure 5 shows the improvement that can be realized by steering an ensemble of N masers rather than just a single device as in the previous figures. The optical clock runs for one hour per day in these simulations. The stability has an improvement of $\sqrt{N}$, for all averaging times. Comparing Figure 5 with Figure 3, for an averaging time of greater than 10 days, using an ensemble of masers, instead of a single maser, does not reduce the instability of a Cs-fountain time scale, while there is an improvement of $\sqrt{N}$ for an optical-clock time scale by using an ensemble of masers. Comparing the black curve in Figure 5 with the orange curve of the upper right plot of Figure 4, the time scale composed of four masers and one optical clock running an hour a day is better than the time scale composed of one maser and one optical clock running two hours a day.

In summary, to achieve the same performance as a Cs-fountain time scale, we could run an optical clock according to one of the following options: 12 min per half a day, 1 hour per day, 4 hours per 2.33 day, or 12 hours per week. The results are not sensitive to the assumed stability of the optical clock, provided only that it is much more stable than the free-running time scale.
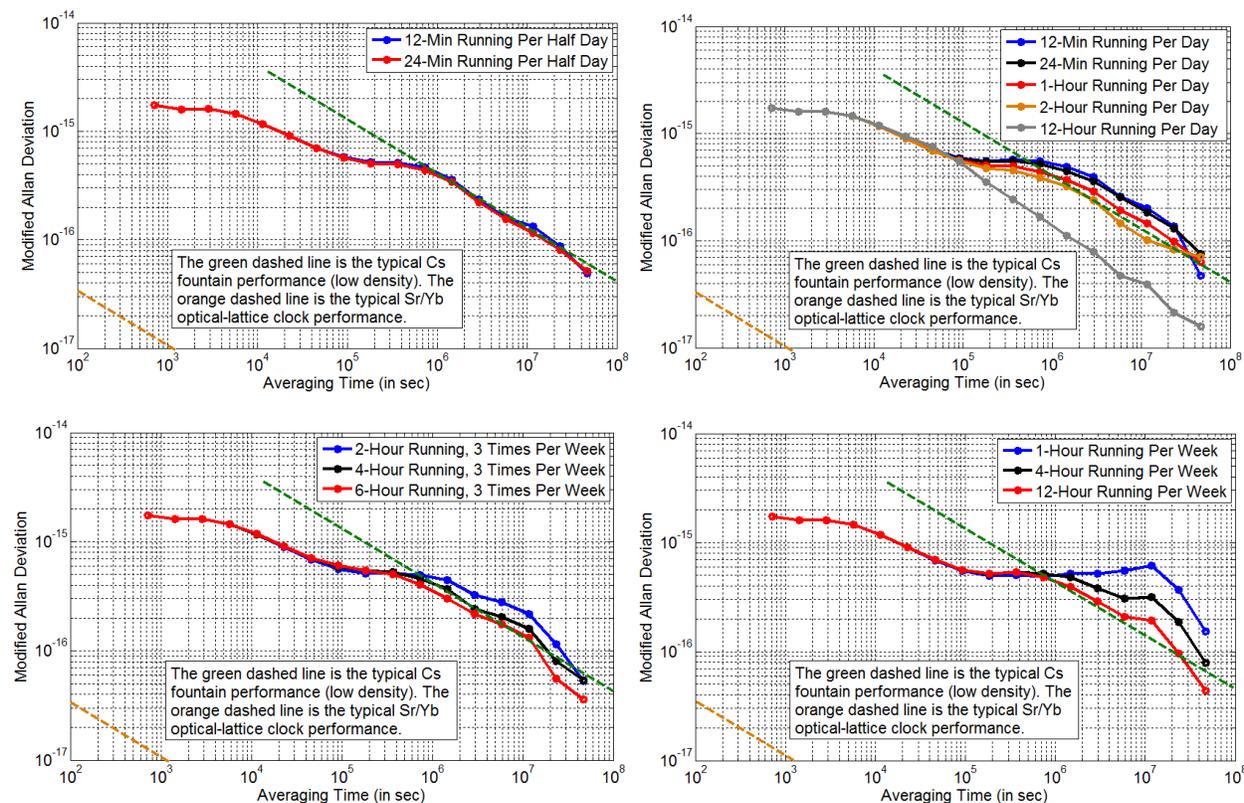

Figure 4. Performance of a time scale steered to an intermittently-operating optical clock. A variety of running schedules are studied in this figure.
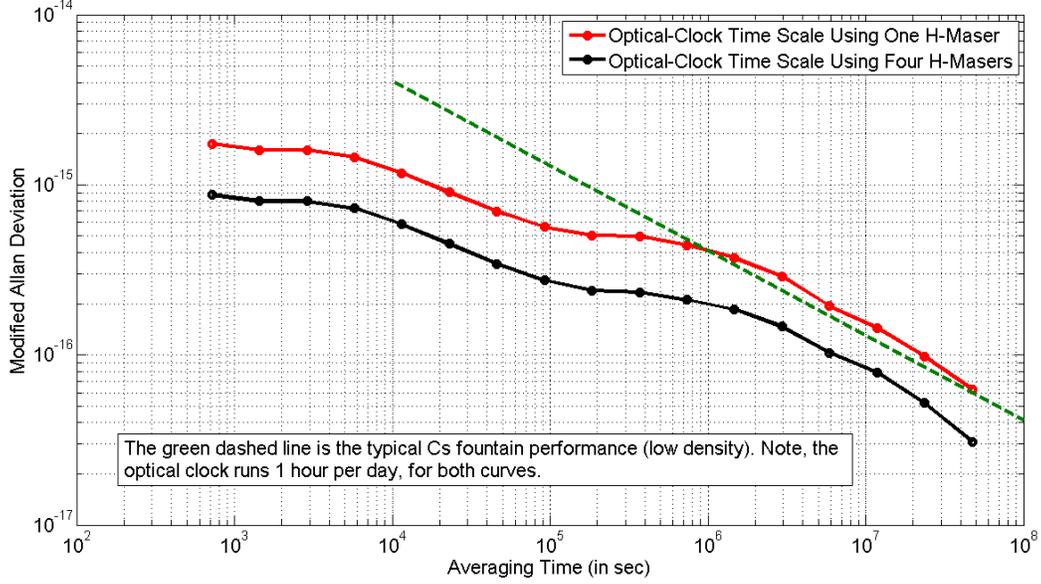
Figure 5. The simulated results of steering an ensemble of masers to an optical frequency reference that runs one hour per day. An ensemble of four masers steered to an optical clock running one hour per day is more stable than a cesium fountain for all averaging times.

IV. Preliminary Real-Data Analysis

IV(a). Initial Campaign in 2017 March and April

Following the simulations, the ytterbium optical clock at NIST [2] has been run regularly in 2017 March and April. Individual run times range from 48 min to 6.9 hours, depending on the experimental arrangement. The Yb frequency standard was down-converted to around 1 GHz via a Ti:sapphire frequency comb. Then this microwave signal was compared with the up-converted hydrogen maser signal, via a mixer (Note, the 5 MHz signal from a hydrogen maser is up-converted to 1 GHz). The frequency difference between the two microwave signals was measured by a commercial frequency counter. The fractional frequency difference between the Yb clock and the hydrogen maser can be derived based on this measurement. In addition, the gravitational correction due to the height from the geoid was applied to the Yb clock so that we know the fractional frequency difference between the "geoid" Yb clock and the hydrogen maser. For simplicity, the Yb clock mentioned below means the "geoid" Yb clock.

Figure 6 shows the fractional frequency difference between Yb and a NIST hydrogen maser ST15. Similar to what we have done in the simulations (see Section II(b)), we average each individual run and compute the mean fractional frequency difference between Yb and ST15. According to the simulation result in Figure 5, an ensemble steered to an optical clock is more stable than a single clock steered to an optical clock, for all averaging times. Thus, to completely exploit the advantages of having an optical clock, we steer the free-running AT1 time scale to the optical clock, instead of steering a single H-maser to the optical clock. Admittedly, in practice, the measurement noise of the link between the H-maser and the time scale could make this architecture noisier. However, this noise is mainly white phase noise and is averaged down to become negligible after tens of minutes. Thus, the architecture of steering the time scale to an optical clock is still more favorable in practice. Figure 7 shows the average fractional frequency difference between AT1 and the Yb clock, for each individual run. The variation between AT1 and Yb is about 5e-15 s/s. The frequency offset shown in Figure 7 mainly comes from the frequency offset of AT1. We can also observe a tiny frequency drift of AT1 from Figure 7. The information of the frequency drifts of AT1 and hydrogen masers based on an optical clock is one of a few benefits of having an optical clock running. This information is helpful for the timekeeping at NIST. The steer of AT1 to Yb is achieved by the Kalman filter as mentioned in Section II(b). An intuitive understanding of this filter is: the longer operation time of the Yb clock, the larger weight the run gets; the longer time interval between

the previous run and the current run, the larger weights the current run gets. At the current stage, the steer is done on paper, instead of tuning an AOG (auxiliary output generator) physically.

To evaluate the performance of the steered AT1, we compare the steered AT1 with the UTC via the TWSTFT (two-way satellite time and frequency transfer) link. The result is shown by the left plot of Figure 8. We can see that there is a constant frequency offset of -9.14e-15 s/s between the steered AT1 and the UTC. This offset could come from a variety of sources, e.g. the Yb clock, the frequency comb, the measurement system, or an analysis error. Indeed during these measurements, the Yb clock frequency was being intentionally varied as part of an independent systematic-uncertainty-evaluation, and the frequency counter used to count the comb repetition rate was being operated in a mode where known biases exist. In a more recent campaign (see Section IV(b)), where systems were operated normally and the frequency counting was done in a high accuracy configuration, we observe no such offset. The right plot of Figure 8 shows the residual of the steered AT1 against UTC (red curve), and the residual of the AT1 against UTC (blue curve) as a comparison. The standard deviation of the residual of the steered AT1 is only 0.12 ns, and the corresponding Allan deviation is about 4.4e-16 at 5 days with an upper bound of 7.9e-16 and a lower bound of 3.4e-16. Also, we can see that the frequency drift in AT1 is corrected by steering AT1 to the Yb clock.
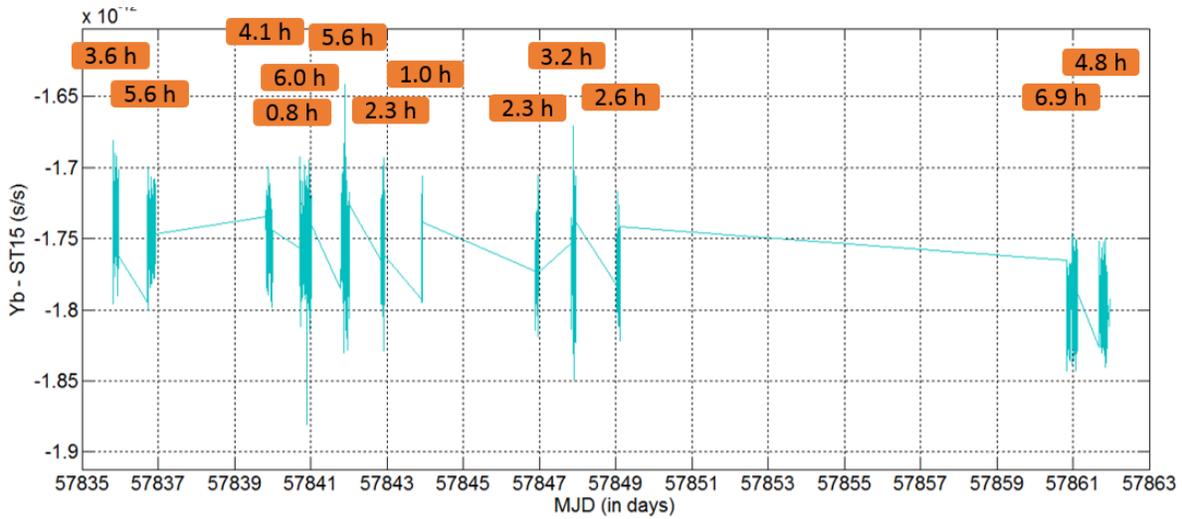

Figure 6. Comparison of the Yb optical clock and a NIST hydrogen maser (i.e., ST15), for the initial campaign.
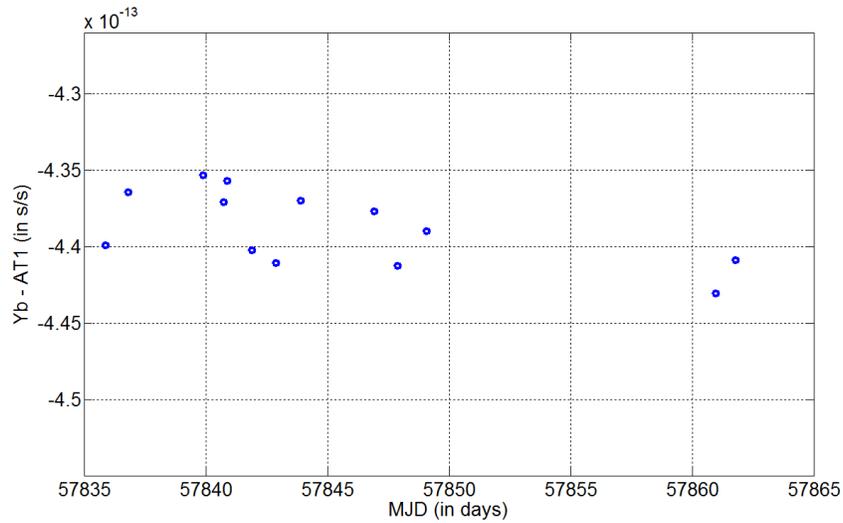

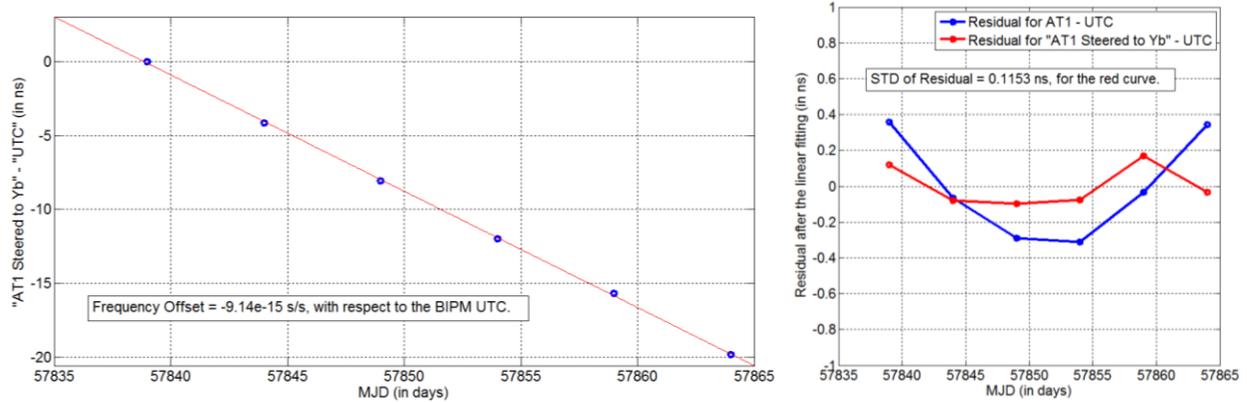Figure 7. Average of "Yb – AT1," for the initial campaign.

Figure 8. Steered AT1 versus UTC (left) and the residual after the linear fitting (red curve in the right plot), for the initial campaign.

IV(b). Ongoing Campaign since 2017 late October

Encouraged by the above resulting statistical performance, since 2017 late October, we have conducted a focused campaign with the ytterbium clock and optical frequency comb. In particular, the Yb clock is run in normal operation including a formal accounting of systematic clock shifts. Furthermore, frequency offsets in the frequency-comb-based optical-to-microwave synthesis were characterized and confirmed to be small, and the frequency counter was operated in a mode with essentially negligible bias. As mentioned above, during this campaign, no sizeable frequency offset was measured versus UTC. This campaign is ongoing, and the expected duration of more than a few months can provide better statistics than the initial campaign.

Figure 9 shows the data from Oct. 26, 2017 to Dec. 28, 2017. The error bar of each data point mainly depends on the instability of AT1, if we assume that the noise from the Yb clock and the measurement system is negligible. Depending on the time duration of the comparison between Yb and AT1, the error bar could vary from high 10^{-16} to low 10^{-15}. From the data shown, there were two measurements with the optical-clock that lasted for only 7 min, which are labeled by the red ellipses. These two data points have big error bars and should not be considered as outliers. They are given the weights that they deserve, in the Kalman filter. From Figure 9, we observe a fractional frequency change in "Yb – AT1" of about $+1\times10^{-15}$ from the first group of data sets (i.e., MJD 58053 – 58075) to the second group of data sets (i.e., MJD 58083 – 58110). Comparing AT1 with the BIPM UTC cannot resolve any obvious frequency drift in AT1. When comparing the maser-based measurement of Yb clock with the BIPM UTC shown by Figure 10, we see a difference in the two data sets of about $+1\times10^{-15}$.

Over the same time period, another very preliminary independent study has compared the measured Yb frequency with the Cs frequency (i.e., PSFS – Primary and Secondary Frequency Standards) published in the BIPM Circular T. For the first group of data sets, the average frequency difference between Yb and PSFS is -1.2×10^{-15}, with an estimated uncertainty of 5×10^{-16}. This uncertainty includes a dead time uncertainty of 3.7×10^{-16}, a frequency transfer uncertainty of 2.6×10^{-16}, and a PSFS uncertainty of 2.0×10^{-16}. It does not include the u_A and u_B uncertainties for Yb. The inherent uncertainties for Yb should be very small, but there may be an additional u_A uncertainty due to a possible frequency error that may occur in reducing the optical frequency down to the 1 GHz range where it can be measured relative to a hydrogen maser with a counter. This additional uncertainty is currently being evaluated. For the second group of data sets, the average frequency difference between Yb and PSFS is -1×10^{-17}, with an uncertainty of 6×10^{-16} (again not including u_A and u_B). Thus, the two groups of data sets agree with each other at the 1.5 sigma level. Since the campaign is ongoing, more data will be collected to make a more complete assessment.
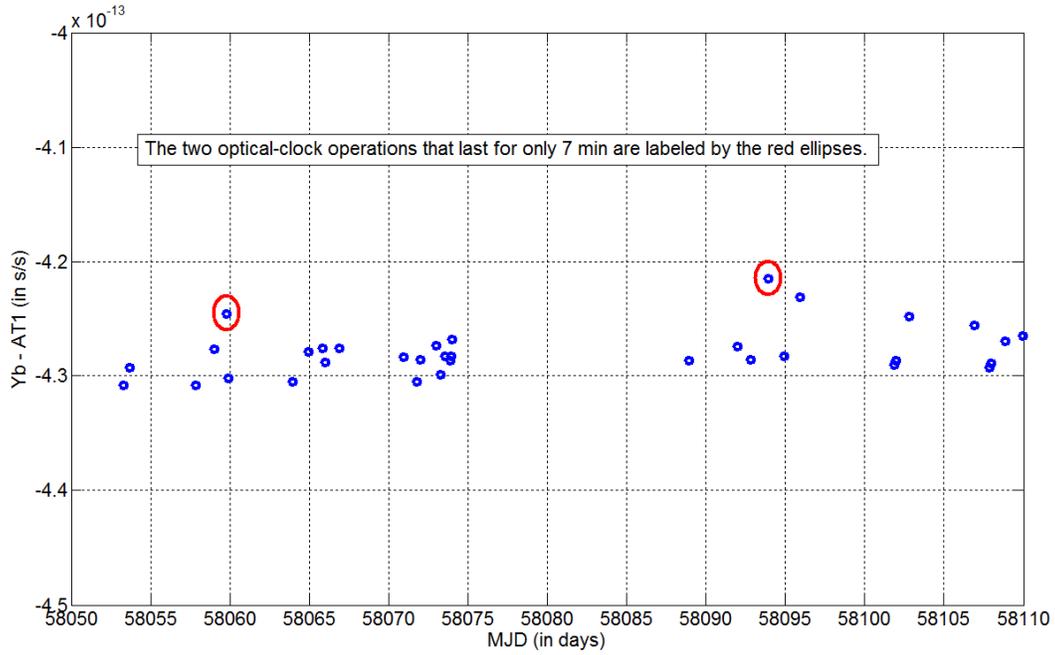
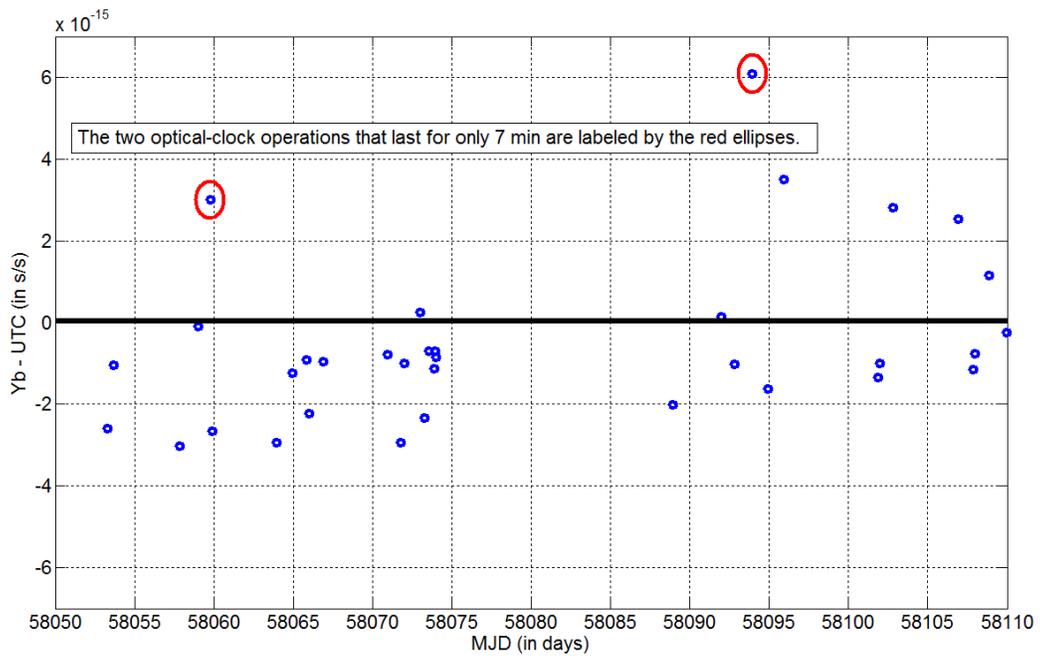Figure 9. Average of "Yb – AT1," for the ongoing campaign.

Figure 10. Average of "Yb – UTC," for the ongoing campaign.

V.   Future Work

The measurement campaign described in Section IV is ongoing, and with additional data in the coming months, we will know more about the statistics of Yb versus UTC/PSFS. Furthermore, NIST has two other optical clocks which

are sometimes operated: the Al+ clock at NIST [1] and a Sr clock at JILA [3]. We plan to explore the scheme of having multiple optical clocks in a future time scale. This can distribute the work to different groups of people and thus ease the work intensity.

References

[1] C. W. Chou, D. B. Hume, J. C. J. Koelemeij, D. J. Wineland, and T. Rosenband, "Frequency comparison of two high-accuracy Al$^+$ optical clocks," Phys. Rev. Lett. 104, 070802, 2010.
[2] N. Hinkley, J. A. Sherman, N. B. Phillips, M. Schioppo, N. D. Lemke, K. Beloy, M. Pizzocaro, C. W. Oates, and A. D. Ludlow, "An atomic clock with 10^{-18} instability," Science, vol. 341, pp. 1215-1218, 2013.
[3] B. J. Bloom, T. L. Nicholson, J. R. Williams, S. L. Campbell, M. Bishof, X. Zhang, W. Zhang, S. L. Bromley, and J. Ye, "An optical lattice clock with accuracy and stability at the 10^{-18} level," Nature, vol. 506, pp 71-75, 2014.
[4] T. P. Heavner, S. R. Jefferts, E. A. Donley, J. H. Shirley, and T. E. Parker, "NIST-F1: recent improvements and accuracy evaluations," Metrologia, vol. 42, pp. 411-422, 2005.
[5] A. Bauch, S. Weyers, D. Piester, E. Staliuniene, and W. Yang, "Generation of UTC(PTB) as a fountain-clock based time scale," Metrologia, vol. 49, pp. 180-188, 2012.
[6] G. D. Rovera, S. Bize, B. Chupin, J. Guena, Ph. Laurent, P. Rosenbusch, P. Uhrich, and M. Abgrall, "UTC(OP) based on LNE-SYRTE atomic fountain primary frequency standards," Metrologia, vol. 53, pp. S81-S88, 2016.
[7] A. Gelb, "Applied optimal estimation," MIT Press, 1974.
[8] L. Galleani, and P. Tavella, "On the use of the Kalman filter in timescales," Metrologia, vol. 40, pp. S326-S334, 2003.
[9] J. Levine, "Invited review article: The statistical modeling of atomic clocks and the design of time scales," Review of Scientific Instruments, vol. 83, 021101, 2012.
[10] N. Ashby, "Probability distributions and confidence intervals for simulated power law noise," IEEE Transactions on UFFC, vol. 62, pp. 116-128, 2015.
[11] C. Grebing, A. Al-Masoudi, S. Dorscher, S. Hafner, V. Gerginov, S. Weyers, B. Lipphardt, F. Riehle, U. Sterr, and C. Lisdat, "Realization of a timescale with an accurate optical lattice clock," Optica, vol. 3, pp. 563-569, 2016.
[12] T. Ido, H. Hachisu, F. Nakagawa, and Y. Hanado, "Time scale steered by an optical clock," Proceedings of the 48th Precise Time and Time Inverval Meeting, pp. 15-17, 2017.