Audio Engineering Society

Convention Paper

Presented at the 129th Convention
2010 November 4–7 San Francisco, CA, USA

Impulse Response Measurements in the Presence of Clock Drift

Nicholas J. Bryan[1], Miriam A. Kolar[1], Jonathan S. Abel[1]

[1] *Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA, 94305, USA*

Correspondence should be addressed to Nicholas J. Bryan (njb@ccrma.stanford.edu)

ABSTRACT

There are many impulse response measurement scenarios in which the playback and recording devices maintain separate unsynchronized digital clocks resulting in clock drift. Clock drift is problematic for impulse response measurement techniques involving convolution, including sinusoidal sweeps and pseudo-random noise sequences. We present analysis of both a drifting record clock and playback clock, with a focus on swept sinusoids. When using a sinusoidal sweep without accounting for clock drift, the resulting impulse response is seen to be convolved with an allpass filter having the same frequency trajectory form as the input swept sinusoid with a duration proportional to the input sweep length. Two methods are proposed for estimating the clock drift and compensating for its effects in producing an impulse response measurement. Both methods are shown to effectively eliminate any clock effects in producing room impulse response measurements.

1. INTRODUCTION

When measuring room impulse responses (IRs), it is occasionally difficult to playback and record the test signal and response using a single device or to synchronize the digital clocks between separate devices. Such situations arise in room acoustics and archaeoacoustic IR measurements where large or constrained architectural dimensions force the use of two audio interface devices—one for playback and record [1]. In such situations, the recording and playback signals will run off separate internal digital clocks, creating a small difference between the two sampling rates or clock drift. While the difference in clock speeds is small, such errors can accumulate over time and become problematic for longer duration recordings used for robust convolution-based IR measurement methods such as sinusoidal-sweeps (linear and exponential) and pseudo-random noise sequences.

Over time, the clock drift causes a misalignment between the output test signals and the recorded response, corrupting the deconvolution process. Fig. 1 illustrates the effect of drifting clocks between professional quality playback and recording devices imposed on a two-minute long periodic impulse train. Sampled impulses are displayed at 10 second increments to illustrate the gradual, nearly linear drift of one clock relative to the other. Fig. 2 demonstrates the same effect of clock drift applied to an actual room impulse response measured via a linearly-convolved sine sweep. In this scenario, clock drift significantly corrupts the impulse response measurement.

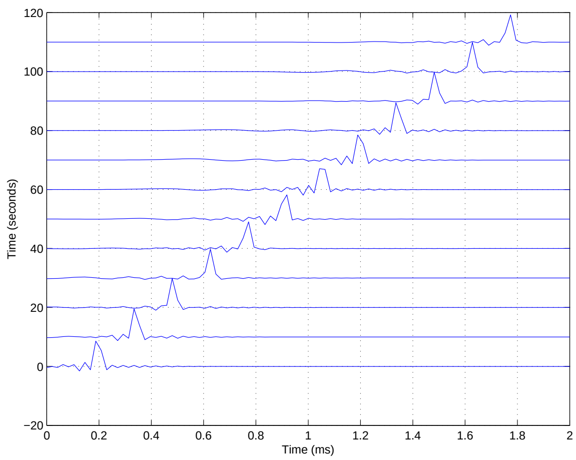

Fig. 1: Recorded Impulse Train with Clock Drift. A two-minute long recorded impulse train in the presence of clock drift.

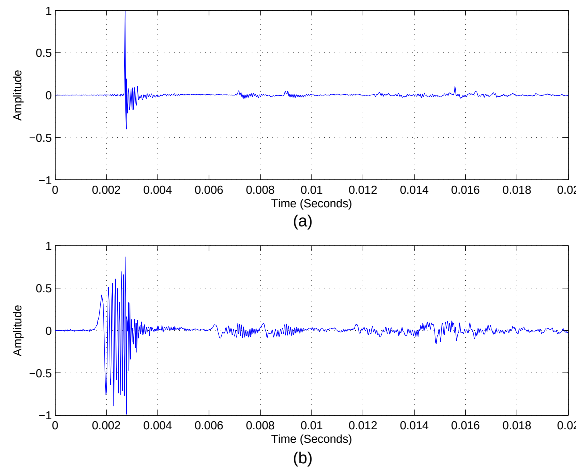

Fig. 2: (a) Room Impulse Response without Clock Drift. A reference room impulse response measured without clock drift using a single audio interface for playback and recording. (b) Room Impulse Response with Clock Drift. A room impulse response measured in the presence of clock drift using two unsynchronized audio interfaces for playback and recording.

In this work, we present analysis of how clock drift affects room IR measurement techniques, with an emphasis on sinusoidal sweep test signals. It will be shown that sine sweeps (SS) are particularly well suited for measuring impulse responses using unsynchronized playback and record devices in addition to the previously known benefits of a straightforward implementation, robustness against noise, and insensitivity to weak speaker nonlinearities [2, 3, 4, 5, 6]. These benefits motivate the focus on SS methods and provide a suitable scope for this work.

The scenario of a drifting clock applied to sine sweep test signals was considered in [4] and described as having the effect of causing a "skewed" impulse response. A brief discussion on compensation was also presented. Here, closed-form expressions describing the effect of clock drift on the resulting impulse responses are presented, focusing on swept sinusoidal methods. By assuming that the clocks are running at a constant drift rate, and establishing the most trusted clock as a reference (with the opposing clock as drifting), both scenarios of a drifting playback clock and drifting record clock are explored using linear convolution methods. Circular convolution IR techniques are less useful and not explicitly considered, as the recorded and inverse signals differ in length and therefore in periodicity. Two methods of compensation are suggested and provide a post-processing solution to correct SS IR measurements.

An overview of sinusoidal sweeps is presented in §2, general clock drift analysis applied to convolution-based IR techniques in §3, and the specific consequences of clock drift on sine sweeps in §4. Techniques for clock drift compensation are found in §5 with a discussion on clock drift estimation in §6, and final remarks in §7.

2. SINE SWEEP OVERVIEW

Over the past decade, sinusoidal sweeps have become a prominent method for measuring room impulse responses [3, 4, 5, 6]. The sinusoidal sweep $s(t)$ is generated according to a frequency trajectory $\omega(t)$

$$s(t) = \sin \phi(t), \tag{1}$$

$$\phi(t) = \int_0^t \omega(\nu) d\nu, \tag{2}$$

where $\phi(t)$ is the instantaneous phase, the integral of the frequency trajectory $\omega(t)$ over time $t \in [0, \tau]$. Notable frequency trajectories include linear and exponential, defined as

$$\omega_{lin}(t) = \Big(\frac{\omega_1 - \omega_0}{\tau}\Big) t + \omega_0 \tag{3}$$

$$\omega_{exp}(t) = \omega_0 \exp\Big\{ \frac{\mathrm{t}}{\tau} \ln \omega_1/\omega_0 \Big\}, \tag{4}$$

where the trajectory is designed to sweep from ω_0 at time $t = 0$ to ω_1 at time $t = \tau$. Linear chirps spend equal lengths of time transversing any fixed bandwidth interval and provide a constant signal-to-noise ratio (SNR) across all frequencies. Exponential chirps spend equal lengths of time transversing any given octave and provide an SNR gain that is inversely proportional to frequency. Exponential chirps are typically more appropriate for measuring room responses, which have similar noise floor frequency characteristics.

2.1. Sine Sweep Transforms

The frequency response $S(\omega)$ of the sweep can be formulated via a magnitude and phase decomposition $|S(\omega)|e^{j\phi(w)}$. Equations (5) and (6) show an approximation of the magnitude and exact formulation of the phase,

$$|S(\omega)| \approx 1 \Big/ \sqrt{\frac{1}{2}\Big|\frac{d\gamma(\omega)}{d\omega}\Big|}, \tag{5}$$

$$\phi_s(\omega) = -\int_0^\omega \gamma(\nu) d\nu, \tag{6}$$

where the group delay $\gamma(\omega)$ is the functional inverse of $\omega(t)$. The magnitude and phase functions of the group delay only.

The compact representation of the magnitude (5), follows from noting that the signal energy in the frequency domain between two nearby frequencies $[\omega_-, \omega_+]$ is approximated by the signal energy in the time domain over the corresponding time interval $[\gamma(\omega_-), \gamma(\omega_+)]$. In essence, the sinusoid sweep power at any given frequency is proportional to the "time spent" on that frequency—the inverse frequency trajectory derivative. The original derivation and more complete explanation is found in [2], with a similar formulation applied to all-pass chirps in [7].

2.2. Sine Sweep Inverse Filter

In addition to the magnitude and phase decomposition, it is also useful to formulate an inverse filter $f(t)$, such that $s(t) * f(t) \approx \delta(t)$. The filter $f(t)$ can be computed using various methods including:

- time-reversal mirror plus whitening filter approach [3]
- closed form approximation [2]
- numerical inversion via $\mathrm{IFFT}(\frac{1}{\mathrm{FFT}(s)})$ with sufficient zero-padding

Fig. 3 depicts the resulting impulse response of a sine sweep convolved with its inverse filter using each of the three methods. As seen, the time-reversal mirror and closed-form approaches are roughly bandlimited to the frequency range of the sine sweep. The numerical inversion approach improves upon both alternative methods and is suggested for implementation.

For analysis, however, the time-reversal mirror and closed-form approaches yield considerable insight. The time-reversal mirror approach defines the inverse as a windowed time-flipped version of the original sweep. The closed-form approach extends this interpretation by providing a general expression for the "whitening" window. The closed-form representation of the inverse filter $f(t)$ is

$$f(t) \approx \Big|\frac{d\omega(-t)}{dt}\Big| s(-t). \tag{7}$$

The filter $f(t)$ is the time-flipped sinusoid, windowed in proportion to the length of time the sinusoid

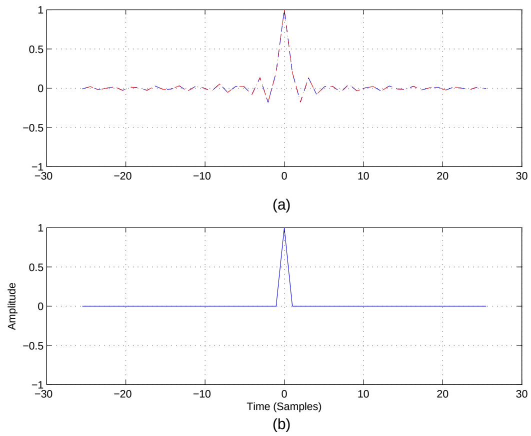

Fig. 3: (a) Impulse Responses Using Time-Reversal Mirror and Closed-Form Inverse Approaches. Impulse responses computed using inverse filters via time-reversal mirror (dashed dot) and closed-form (dashed) approaches showing roughly bandlimited inversion. (b) Impulse Response Using the Numerical Inversion Inverse Approach. An impulse response computed using an inverse filter created via numerical inversion showing an improved result.

spends on each frequency. The magnitude and phase response are

$$|F(\omega)| \approx \sqrt{\frac{1}{2}\left|\frac{d\gamma(\omega)}{d\omega}\right|} \quad (8)$$

$$\phi_f(\omega) = \int_0^{\omega} \gamma(\nu) d\nu, \quad (9)$$

which are the corresponding inverses of (5) and (6) respectively.

3. CLOCK DRIFT ANALYSIS

Consider the linear time-invariant system with input test signal $s(t)$, recorded output $r(t)$, impulse response $h(t)$, and additive Gaussian noise $n(t) \sim \mathcal{N}(0, \sigma^2)$,

$$r(t) = s(t) * h(t) + n(t). \quad (10)$$

The recorded system response to the test signal $s(t)$ is corrupted by additive noise, limiting the ability to identify the true linear system response. The system impulse response $h(t)$ can be recovered by convolving the recorded test signal $r(t)$ with an inverse filter $f(t)$. Assuming for the moment that $n(t) = 0$,

$$\begin{aligned} h(t) &= r(t) * f(t) \\ &= h(t) * s(t) * f(t). \end{aligned} \quad (11)$$

where $s(t) * f(t) \approx \delta(t)$.

With drifting playback and record time scales ($\alpha_p t$ and $\alpha_r t$), the playback and record clock are not synchronized. For convenience of having only one drifting clock, a single clock can be chosen as a reference operating on a time scale t with $\alpha = 1$, with the opposing clock encompassing the drift. The more trusted or convenient clock should be considered as reference and assumed to be of perfect accuracy. This simplification results in two useful scenarios when applied to convolution-based methods: a drifting playback clock and a drifting record clock, considered in §3.1 and §3.2 respectively.

3.1. Drifting Playback Clock

With a drifting playback clock, the system output $\tilde{r}(t) = h(t) * s(\alpha_p t)$ results in the desired impulse response convolved with a residual filter $d_p(t)$,

$$\begin{aligned} \tilde{h}(t) &= \tilde{r}(t) * f(t) \\ &= h(t) * (s(\alpha_p t) * f(t)) \quad (12) \\ &= h(t) * d_p(t) \quad (13) \end{aligned}$$

$$d_p(t) = s(\alpha_p t) * f(t). \quad (14)$$

The deconvolution residual or applied drift filter is the convolution between a time-scaled input test signal and its true inverse.

3.2. Drifting Record Clock

With a drifting record clock, the system output $\tilde{r}(t) = h(\alpha_r t) * s(\alpha_r t)$ produces an impulse response described as

$$\begin{aligned} \tilde{h}(t) &= \tilde{r}(t) * f(t) \\ &= h(\alpha_r t) * (s(\alpha_r t) * f(t)). \end{aligned} \quad (15)$$

A drifting record clock results in an equivalently filtered impulse response as above with an additional step of resampling. The applied drift filter $d_r(t)$ is

$$d_r(t) = s(\alpha_r t) * f(t). \tag{16}$$

For typical situations, the applied drift filter has the most significant drift effect when compared to the resampling of $h(t)$. The resampling of $h(t)$ has an imperceivable psychoacoustic effect and can be neglected unless otherwise deemed significant. Now, a more specific look at clock drift applied to linearly convolved sinusoidal sweeps is presented in §4.

4. SINE SWEEP CLOCK DRIFT ANALYSIS

As shown in (12) and (15), clock drift results in an undesirably modified measured impulse response. The inverse filtering process does not adequately invert the input signal $s(t)$ and effectively applies a drift filter $d(t)$ on the system. Extending this analysis to sine sweeps, we first notice that for slowly varying sweeps, a stretch α in the time scale of the sweep will result in a stretch the frequency trajectory $w(t)$. Knowing that the group delay $\gamma(\omega)$ is the functional inverse of $w(t)$, the clock drift error can then be conveniently expressed as a simple modification to the group delay

$$\begin{aligned}\tilde{\gamma}(\omega) &\approx \alpha\gamma(\omega) \\ &\approx (1+\epsilon)\,\gamma(\omega),\end{aligned} \tag{17}$$

where ϵ and is the fraction of drift away from the desired sample rate.

Given that the magnitude and phase of a sweep can be solely expressed as a function of group delay as outlined in §2.1, the drifting sweep magnitude $|\tilde{S}(\omega)|$ and phase $\tilde{\phi}_s(\omega)$ are

$$|\tilde{S}(\omega)| \approx \sqrt{1+\epsilon}\,|S(\omega)| \tag{18}$$

$$\tilde{\phi}_s(\omega) = -(1+\epsilon)\,\phi(\omega). \tag{19}$$

Applying this in the frequency domain to (11), the drift filter is expressed as

$$\begin{aligned}D(\omega) &\approx \sqrt{1+\epsilon}\,\exp\{-\mathrm{j}\,\epsilon\,\phi_s(\omega)\} \\ &\approx \exp\{-\mathrm{j}\,\epsilon\,\phi_{\mathrm{s}}(\omega)\}\end{aligned} \tag{20}$$

In the time domain, this applied drift filter is

$$d(t) \approx \sqrt{\frac{d\omega(t)}{dt}}\sin\left(\epsilon\int_0^t \omega(\nu)d\nu\right) \tag{21}$$

As shown, the applied drift filter approximates an allpass filter of the same frequency trajectory form as the input sweep. This dependence causes the drift filter group delay to be proportional to the signal length and dependent on the sweep frequency trajectory.

To illustrate the dependence on input signal length, Fig. 4 shows six impulse responses measured in the presence of clock drift with varying length input sine sweeps. The signal lengths are 2^N samples, where $N \in \{16, 17, 18, 19, 20, 21\}$ and of linear trajectory. To illustrate the dependence of the trajectory type, Fig. 5 depicts three impulse responses. The first IR is provided as reference to the second and third responses, which were measured with exponential and linear frequency trajectories of the same length. The amplitude window of Fig. 5 (c) compared to Fig. 5 (b) is explained via the differing trajectories. For an equivalent amount of time, the exponential trajectory only covers a smaller range of high frequencies, resulting in the small visual footprint.

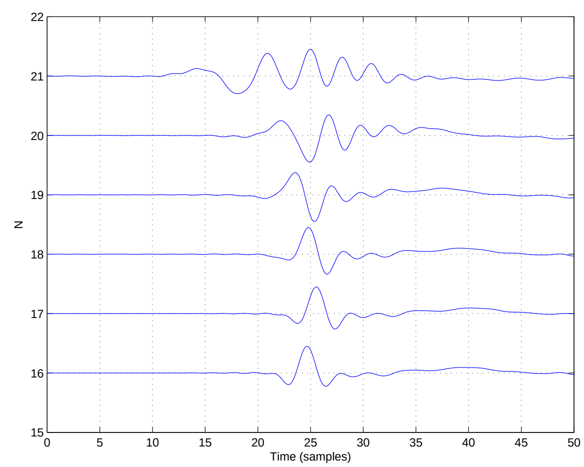

Fig. 4: Various Length Impulse Responses with Clock Drift. Impulse responses in the presence of clock drift computed using various length input sine sweeps.

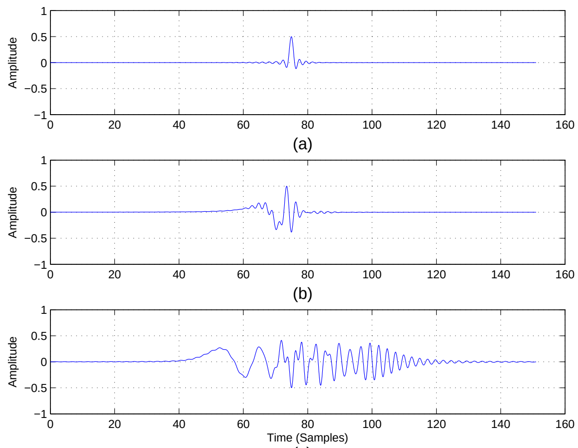

Fig. 5: (a) Reference Impulse Response. A reference impulse response without clock drift. (b) Impulse Response via Exponential Sine Sweep. An impulse response in the presence of clock drift computed using an exponential frequency trajectory sine sweep input. (c) Impulse Response via Linear Sine Sweep. An impulse response in the presence of clock drift computed using a linear frequency trajectory sine sweep input.

5. COMPENSATION

To compensate for the clock drift, two methods are proposed including a pre-convolution resampling method as well as a post-convolution compensation filtering method discussed in §5.1 and §5.2.

5.1. Resampling

The first method resamples either the recorded test signal or inverse filter to equalize the different clock rates prior to convolution. Once the recorded test signal and inverse filter are at equal clock rates, the unwarped impulse response can be computed as before via (11). As will be discussed in §6, the resampling can be done explicitly by first estimating the clock drift rate and interpolating accordingly or implicitly estimating the clock drift rate and resampling in one process by using a direct electronic loopback recording of either the input test signal or inverse filter.

5.2. Compensation Filtering

The second method of compensation involves first constructing a stretched impulse response via (12) or (15). Once the stretched IR is constructed, a compensation filter can be used to remove the applied drift filter. As discussed in §4, the applied drift filter is approximately an allpass filter, allowing the compensation filter to be constructed via a time-flipped version of (14) or (16).

The subtle difference between the two compensation methods can be seen as a simple reordering of operations for linear convolution. When using the same clock drift estimation method, resampling or applying a compensation filter should achieve nearly identical results for sinusoidal sweep test signals. Depending on what is more convenient or accessible at the time, one method or the other maybe preferable.

6. DRIFT RATE ESTIMATION

Before we can compensate for the unwanted effects of clock drift, the clock drift rate must be estimated. Two methods of estimation are proposed: explicit and implicit estimation. Explicit estimation involves a direct estimation of the clock drift rate using a direct electronic loopback recording of a periodic pulse train as discussed in §6.1. Implicit estimation is done by making a direct electronic loopback recording of the actual sine sweep or inverse filter, allowing the system to internally estimate the clock drift and resample the signal at the same time. As stated in §6.2, this method is suggested for actual implementation because it is considerably easier to implement and requires significantly less computation.

6.1. Explicit Estimation

A direct electronic loopback recording of a periodic pulse train can be used to estimate the relative speed between the differing clocks rates. As one clock drifts from another, the time indices of the recorded pulses will gradually drift, compressing or expanding the period of the pulse train by the drift rate α. To obtain an accurate estimate of α, the local maxima and corresponding time values of the recorded signal can be found using quadratic peak interpolation [8]. A least-squares estimate is then made for α by minimizing the norm error between the input time indices ($\mathbf{t}$) and the recorded time indices ($\mathbf{t}_{meas}$) via

$$\min_\alpha \parallel \mathbf{t} - \mathbf{t}_{meas}\alpha \parallel . \quad (22)$$

For a more robust measurement, a short duration allpass chirp signal can be convolved with the impulse train prior to recording. The recorded allpass impulse train can then be correlated with a single allpass chirp to obtain the desired impulse train as discussed above. Using a cascade of 20 first-order allpass filters with coefficient of .4, a confident estimate of the drift rate α was made with an average norm error of $\approx$ .002 samples at a 48kHz sampling rate. Observed clock speed differences were seen to be approximately 1-2ms/minute between professional audio interfaces, matching the typical hardware specifications of crystal oscillators used in digital clock circuitry.

Equivalent results were found independent of which device was used as playback and record. Additionally, measurements taken before and after a five hour recording session showed a relatively constant drift rate estimate of ($\approx \pm$.001) over time. From day-to-day recording sessions with identical setup, however, drift rate measurements did show significant variation suggesting a single estimate of the drift rate for a given recording session was needed.

To demonstrate the ability to remove any unwanted clock drift effects, Fig. 6 first shows a drifted impulse response followed by Fig. 7 showing a compensated version of the drifted IR plotted against a reference IR. The reference IR was taken with an identical setup, but at a different point in time and using the same clock for recording and playback. Using explicit estimation for the clock drift rate and windowed sinc interpolation resampling, near perfect compensation results were achieved.

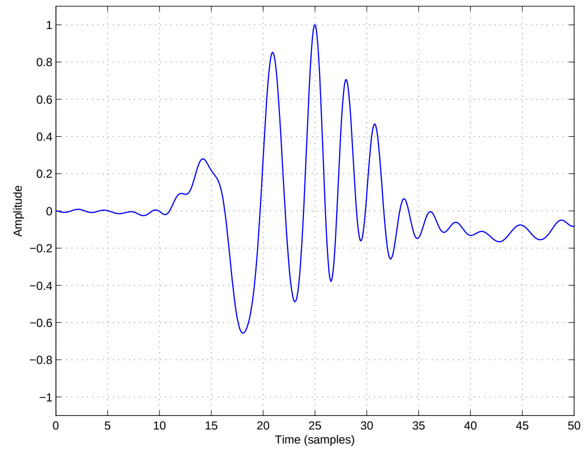

Fig. 6: Impulse Response Direct Path With Clock Drift. The direct path of an impulse response in the presence of clock drift used for comparison with Fig. 7 and Fig. 8.

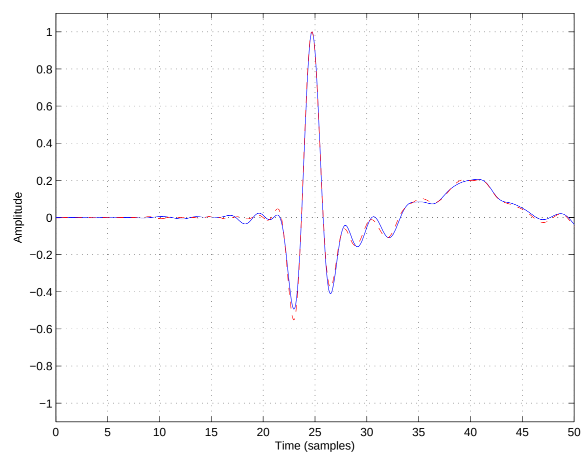

Fig. 7: Impulse Response with Compensation. Impulse response with drift compensation (dashed) using explicit clock drift estimation and windowed sinc interpolation plotted against a reference impulse response with no drift (solid).

6.2. Implicit Estimation

As an alternative to explicitly measuring the clock drift rate and applying resampling accordingly, the clock drift can be implicitly measured and compensated. A direct electronic (or loopback) recording can be made through the drifting system using the sine sweep or inverse filter. The direct loopback recording automatically resamples the input signal or inverse, which is then be used to recover the unwarped room impulse response. Using the same drifted impulse response of Fig. 6, implicit estimation and compensation results are presented in Fig. 8. As with explicit estimation, near perfect

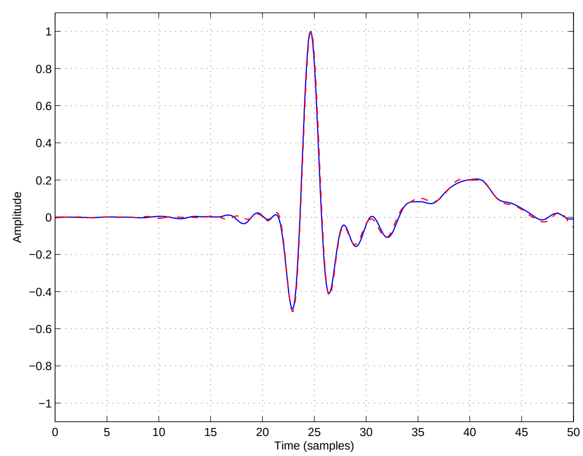

Fig. 8: Impulse Response with Compensation. Impulse response with drift compensation (dashed) using implicit clock drift estimation and compensation plotted against a reference impulse response with no drift (solid).

compensation is achieved.

The benefits of implicit estimation are simplicity, minimal additional computation, and essentially perfect estimation (although explicitly unknown). The downside of implicitly estimating the drift is the additional filtering applied at DC and the band edge due to dc blocking analog-to-digital (A/D) and digital-to-analog (D/A) conversion. By using a loopback recording to resample, the converter filtering is applied twice rather than once. Psychoacoustically, however, this additional filtering is imperceivable, suggesting the implicit estimation and resampling method to be preferable over explicit methods.

7. CONCLUSION

To summarize, the problem of measuring impulse responses in the presence of clock drift is explained. By assuming a linear drift between the two devices, a closed-form approximation shows that the unwanted drift imposes allpass filtering on the resulting room impulse response. Two methods of compensation are proposed including resampling and compensation filtering. For proper compensation, accurate clock drift rate estimation is made via implicit or explicit estimation. Regardless of which methods are used, example results yield near perfect clock drift compensation, demonstrating an ability to eliminate any noticeable clock drift effects when measuring room impulse responses.

8. ACKNOWLEDGEMENTS

This work was supported in part by a grant from SiCa, the Stanford Institute for Creativity and the Arts, for the Chavín de Huántar Archaeological Acoustics Project (see `https://ccrma.stanford.edu/groups/chavin/`).

9. REFERENCES

[1] J. S. Abel, J. W. Rick, P. Huang, M. A. Kolar, J. O. Smith, J. M. Chowning, "On the Acoustics of the Underground Galleries of Ancient Chavín de Huántar, Peru," Invited paper Acoustics '08, Paris, France, July 2008.

[2] J. S. Abel, D. P. Berners, "MUS424/EE367D: Signal Processing Techniques for Digital Audio Effects," Unpublished Course Notes, CCRMA, Stanford University, Stanford, CA, April 2004.

[3] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept- sine technique," Convention Paper 5093. Presented at the 108th Audio Engineering Society Convention, Paris, February 19-22, 2000.

[4] A. Farina, "Advancements in impulse response measurements by sine sweeps," Convention Paper 7121. Presented at the 122nd Audio Engineering Society Convention, Vienna, Austria, May 5-8, 2007.

[5] S. Müller, P. Massarani. "Transfer-function measurement techniques with sweeps," Journal of the Audio Engineering Society, vol. 49, no. 6, p. 443, 2001.

[6] G. Stan, J. J. Embrechts, D. Archambeau. "Comparison of different impulse response measurement techniques," Journal of the Audio Engineering Society, Vol. 50, No. 4, p. 249, 2002.

[7] V. Välimäki, J. S. Abel, J. O. Smith III, "Spectral delay filters," Journal of the Audio Engineering Society, vol. 57, no. 7/8 pp. 521-531, 2009.

[8] J. O. Smith. Spectral Audio Signal Processing, October 2008 Draft, `http://ccrma.stanford.edu/~jos/sasp/`, online book.