SAMPLING WITHOUT TIME:
RECOVERING ECHOES OF LIGHT VIA TEMPORAL PHASE RETRIEVAL

Ayush Bhandari[1], Aurélien Bourquard[2], and Ramesh Raskar[1]

[1]Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA.

[2]Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA.

Email: ayush@MIT.edu   •   aurelien@MIT.edu   •   raskar@MIT.edu

ABSTRACT

This paper considers the problem of sampling and reconstruction of a continuous-time sparse signal without assuming the knowledge of the sampling instants or the sampling rate. This topic has its roots in the problem of recovering multiple echoes of light from its low-pass filtered and auto-correlated, time-domain measurements. Our work is closely related to the topic of sparse phase retrieval and in this context, we discuss the advantage of phase-free measurements. While this problem is ill-posed, cues based on physical constraints allow for its appropriate regularization. We validate our theory with experiments based on customized, optical time-of-flight imaging sensors. What singles out our approach is that our sensing method allows for temporal phase retrieval as opposed to the usual case of spatial phase retrieval. Preliminary experiments and results demonstrate a compelling capability of our phase-retrieval based imaging device.

Index Terms— Computational imaging, phase-retrieval, sampling theory, sparsity and time–of–flight imaging.

1. INTRODUCTION

During a recent visit to meet a collaborator, the author (who happens to be an avid photographer) saw a stark reflection of a local monument on the window panes of the Harvard science center. This is shown in Fig. 1. This is a common place phenomenon at a macro scale as well as a micro scale (for example microscopy). At the heart of this problem is a fundamental limitation, that is, all of the conventional imaging sensors are agnostic to the time information. Alternatively stated, the image–formation process is insensitive to the potentially distinct times that the photons can spend traveling between their sources and the detector.

To elaborate, consider the *Gedankenexperiment* version of Fig. 1 described in Fig. 2 (a). Let us assume that a light source is co-located with the imaging sensor (such as a digital camera). The reflection from the semi-reflective sheet (such as a window pane) arrives at the sensor at time $t_1 = 2d_1/c$ while the mannequin is only observable on or after $t_2 = 2d_2/c$. Here, c is the speed of light. While we typically interpret images as a two dimensional, spatial representation of the scene, let us for now, consider the photograph in Fig. 1 along the time-dimension. For the pixel (x_0, y_0), this time-aware image is shown in Fig. 2 (b). Mathematically, our time-aware image can be written as a 2–sparse signal (and as a K–sparse signal in general),

$$\boxed{\text{Time Aware Image} \atop m\left(\boldsymbol{r}, t\right) = \left(\lambda_1 \Gamma_1\right)\left(\boldsymbol{r}\right) \delta\left(t - t_1(\boldsymbol{r})\right) + \left(\lambda_2 \Gamma_2\right)\left(\boldsymbol{r}\right) \delta\left(t - t_2(\boldsymbol{r})\right)} \tag{1}$$

where $\{\Gamma_k\left(\boldsymbol{r}\right)\}$ are the constituent image intensities, $\{\lambda_k\left(\boldsymbol{r}\right)\}$ are the

Fig. 1: Reflection from semi-reflective surfaces. The Memorial Church can be seen imprinted on the glass facade of the Harvard University Science Center.

reflection coefficients, $\boldsymbol{r} = (x, y)^\top$ is the 2D spatial coordinate, and $\delta(\cdot)$ is the Dirac distribution.

A conventional imaging sensor produces images by marginalizing time information, resulting in the 2D photograph,

$$\int_0^{\Delta \gg t_2} m\left(\boldsymbol{r}, t\right) dt = \left(\lambda_1 \Gamma_1\right)\left(\boldsymbol{r}\right) + \left(\lambda_2 \Gamma_2\right)\left(\boldsymbol{r}\right) \equiv I\left(\boldsymbol{r}\right). \tag{2}$$

Recovering $\{\Gamma_k\}, k = 1, 2$ given I, more generally, K echoes of light,

$$m\left(\boldsymbol{r}, t\right) = \sum_{k=0}^{K-1} \left(\lambda_k \Gamma_k\right)\left(\boldsymbol{r}\right) \delta\left(t - t_k(\boldsymbol{r})\right) \tag{3}$$

is an ill-posed problem. Each year, a number of papers [1–6] attempt to address this issue by using regularization and/or measurement-acquisition diversity based on image statistics, polarization, shift, motion, color, or scene features. Unlike previous works, here, we ask the question: Can we directly estimate $\{\Gamma_k\}_{k=0}^{K-1}$ in (1)? In practice, sampling (3) would require exorbitant sampling rates and this is certainly not an option. Also, we are interested in $\{\Gamma_k\}_{k=0}^{K-1}$ only, as opposed to $\{\Gamma_k, t_k\}_{k=0}^{K-1}$ [7] where t_k is a non-linear argument in (3). As a result, our goal is to recover the intensities of a sparse signal. For this purpose, we explore the idea of *sampling without time*—a method for sampling a sparse signal which does not assume the knowledge of sampling instants or the sampling rate. In this context, our contributions are twofold:

1) For the general case of K-echoes of light, our work relies on estimating the constituent images $\{\Gamma_k\}_{k=0}^{K-1}$ from the filtered, auto-correlated, time-resolved measurements. This is the distinguishing feature of our approach and is fundamentally different from methods proposed in literature which are solely based on spatio-angular
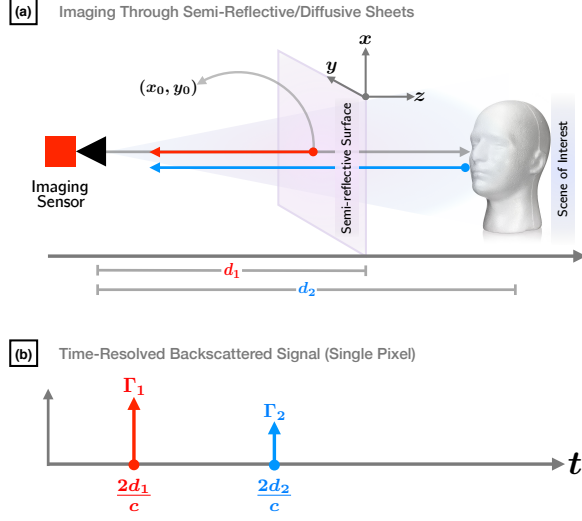
(a) Imaging Through Semi-Reflective/Diffusive Sheets

(b) Time-Resolved Backscattered Signal (Single Pixel)

Fig. 2: Time-aware imaging. (a) Exemplary setting for imaging through semi-reflective media. (b) Corresponding time-aware image at pixel $\boldsymbol{r}_0 = (x_0, y_0)^\top$.

information (cf. [1–6] and references therein).

2) As will be apparent shortly, our work is intimately linked with the problem of (sparse) phase retrieval (cf. [8–12] and references therein). Our "sampling without time" architecture leads to an interesting measurement system which is based on time-of-flight imaging sensors [13] and suggests that $K^2 - K + 1$ measurements suffice for estimation of K echoes of light. To the best of our knowledge, neither such a measurement device nor bounds have been studied in the context of image source separation [1–5].

For the sake of simplicity, we will drop the dependence of m and Γ on spatial coordinates $\boldsymbol{r}$. This is particularly well suited for our case because we do not rely on any cross-spatial information or priors. Also, scaling factors λ_k are assumed to be absorbed in Γ_k.

Significance of Phase-free or Amplitude-only measurements: It is worth emphasizing that resolving spikes from a superposition of low-pass filtered echoes, is a problem that frequently occurs in many other disciplines. This is a prototype model used in the study of multi-layered or multi-path models. Some examples include seismic imaging [14], time-delay estimation [15], channel estimation [16], optical tomography [17], ultrasound imaging [18], computational imaging [13,19–22] and source localization [23]. Almost all of these variations rely on amplitude and phase information or amplitude and time-delay information. However, recording amplitude-only data can be advantageous due to several reasons. Consider a common-place example based on pulse-echo ranging. Let $p(t) = \sin(\omega t)$ be the emitted pulse. Then, the backscattered signal reads $r(t) = \Gamma_0 \sin(\omega t - \theta_0)$. In this setting, on-chip estimation of phase (θ_0) or time delay (t_0) [21],

- can either be computationally expensive or challenging since t_k is a non-linear parameter in (3), and hence, in $r(t)$.

- requires more measurements. For instance, 2 measurements suffice for amplitude-only estimation (Γ_0) while amplitude-phase pair requires 4 measurements [24]. This aspect of phase estimation is an important bottleneck as the the frame rate of an imaging device is inversely proportional to number time-domain measurements.

- is prone to errors. In many applications, multiple-frequency measurements, $\omega = k\omega_0$, are acquired [21,24,25] assuming that phase and frequency, $\theta_0 = \omega t_0$ are linearly proportional. However, this is not the case in practice. The usual practice is to oversample [26,27].

In all such cases, one may adopt our methodology of working with intensity-only measurements.

2. IMAGE–FORMATION MODEL FOR TOF SENSORS

Optical ToF sensors are active devices that capture 3D scene information. We adopt the generic image–formation model used in [13] which is common to all ToF modalities such as lidar/radar/sonar, optical coherence tomography, terahertz, ultrasound, and seismic imaging. In its full generality, and dropping dependency on the spatial coordinate for convenience, one can first formalize this ToF acquisition model as:

$$\boxed{\text{Image–Formation Model} \qquad p \to \boxed{h} \to r \to \boxed{\phi} \to y \xrightarrow{\text{Sampling}} \boldsymbol{y} \mapsto \underbrace{m = \mathcal{C}\left[\boldsymbol{y}\right]}_{\text{Intensity}}} \tag{4}$$

where,

▷ $p(t) > 0$ is a T-periodic probing function which is used to illuminate the scene. This may be a sinusoidal waveform, or even a periodized spline, Gaussian, or Gabor pulse, for instance.

▷ $h(t, \tau)$ is the scene response function (SRF). This may be a filter, a shift-invariant function $h_{\text{SI}}(t, \tau) = h(t - \tau)$, or a partial differential equation modeling some physical phenomenon [25].

▷ $r(t) = \int p(\tau) h(t, \tau)\, d\tau$ is the reflected signal resulting from the interaction between the probing function and the SRF.

▷ $y(t) = \int r(\tau) \phi(t, \tau)\, d\tau$ is the continuous-time signal resulting from the interaction between the reflected function and the instrument response function (IRF), or ϕ, which characterizes the electro-optical transfer function of the sensor.

▷ $\boldsymbol{y}$ is a set of discrete measurements of the form $y(t)|_{t = n\Delta}$.

▷ $\mathcal{C}[\boldsymbol{y}] = \boldsymbol{y} * \overline{\boldsymbol{y}}$ is the cyclic auto-correlation of $\boldsymbol{y}$, where $*$ and $\overline{(\cdot)}$ denote the convolution and time reversal operators, respectively.

The interplay between p, h, and ϕ results in variations on the theme of ToF imaging [13]. In this work, we will focus on an optical ToF setting. Accordingly:

1) The probing function corresponds to a time-localized pulse.

2) The SRF, accounting for the echoes of light, is a K-sparse filter,

$$h_{\text{SI}}(t, \tau) \equiv h_K(t - \tau) = \sum_{k=0}^{K-1} \Gamma_k \delta(t - \tau - t_k). \tag{5}$$

3) The IRF is fixed by design as $\phi_{\text{SI}}(t, \tau) = p(t + \tau)$. This implements the so-called homodyne, lock-in sensor [13].

Due to this specific shift-invariant structure of the SRF and IRF, we have, $\boldsymbol{y} = p * \overline{p} * \overline{h}_K \equiv \mathcal{C}[p] * \overline{h}_K$. Finally, the measurements read,

$$m = \mathcal{C}[\boldsymbol{y}] = (\mathcal{C}[\varphi] * \mathcal{C}[h_K]), \quad \varphi = \mathcal{C}[p]. \tag{6}$$

3. SAMPLING ECHOES OF LIGHT

3.1. Bandlimited Approximation of Probing Function

Due to physical constraints inherent to all electro-optical systems, it is reasonable to approximate the probing signal as a bandlimited function [7,25]. We use L–term Fourier series with basis functions $u_n(\omega_0 t) \stackrel{\text{def}}{=} e^{j\omega_0 n t}$, $\omega_0 T = 2\pi$, to approximate p with,

$$p(t) \approx \widetilde{p}(t) = \sum_{|\ell| \leqslant L} \widehat{p}_\ell u_\ell(\omega_0 t), \tag{7}$$

where the $\widehat{p}_\ell$ are the Fourier coefficients. Note that there is no need to compute $p(t)$ explicitly as we only require the knowledge of $\mathcal{C}[p]$ (cf. (6)). In Fig. 3, we plot the calibrated $\mathcal{C}[p]$ and its Fourier coefficients. In this case, $T = 232.40$ ns. We also plot its approximation, $\widetilde{\varphi} = \mathcal{C}[\widetilde{p}]$ with $L = 25$ together with the measured $\mathcal{C}[p]$ and φ.

3.2. Sampling Theory Context

The shift-invariant characterization of the ToF image–formation model allows to re-interpret the sampled version of (6) as the filtering of a sparse signal $\mathcal{C}[h_K]$ with a low-pass kernel $\psi = \mathcal{C}[\varphi]$ (cf. [7, 28]).

$$m(t) \overset{(6)}{=} \mathcal{C}[h_K] * \mathcal{C}[\varphi] \equiv \mathcal{C}[h_K] * \psi. \tag{8}$$

Note that the properties of the auto-correlation operation imply that the sparsity of $\mathcal{C}[h_K]$ is $K^2 - K + 1$, unlike h_K that is K-sparse and is completely specified by $(K^2 - K)/2 + 1$ due to symmetry. Based on the approximation (7) and the properties of convolution and complex exponentials, and defining $\widehat{\psi}_\ell = |\widehat{p}_\ell|^4$, we rewrite $m(t)$ as,

$$m(t) = \sum_{|\ell| \leq L} \widehat{\psi}_\ell u_\ell(\omega_0 t) \underbrace{\int \mathcal{C}[h_K](\tau) u_\ell^*(\omega_0 \tau) d\tau}_{\text{Fourier Integral}}, \tag{9}$$

Finally, the properties of the Fourier transform imply that sampling the above signal $m(t)$ at time instants $t = n\Delta$, results in discrete measurements of the form $\boldsymbol{m} = \mathbf{U}\mathbf{D}_{\widehat{\psi}}\widehat{\boldsymbol{s}}$, which corresponds to the available acquired data in our acquisition setting. Combining all the above definitions, it follows that:

- $\boldsymbol{m} \in \mathbb{R}^N$ is a vector of filtered measurements (cf. (8)).

- $\mathbf{U} \in \mathbb{C}^{N \times (2L+1)}$ is a DFT matrix with elements $[u_n(\omega_0 \ell \Delta)]_{n,\ell}$.

- $\mathbf{D}_{\widehat{\psi}} \in \mathbb{C}^{(2L+1) \times (2L+1)}$ is a diagonal matrix with diagonal elements $\widehat{\psi}_\ell$. These are the Fourier-series coefficients of $\mathcal{C}[\varphi]$.

- $\widehat{\boldsymbol{s}} \in \mathbb{R}^{2L+1}$ is a phase-less vector containing the Fourier coefficients of $\mathcal{C}[h_K]$, which is obtained by sampling the Fourier transform $\widehat{s}(\omega)$ of $\mathcal{C}[h_K]$ at instants $\omega = \ell\omega_0$. The signal $\widehat{s}(\omega)$ directly depends on the quantities $|\Gamma_k|$ of interest to be retrieved and is expressed as

$$\widehat{s}(\omega) = \left| \sum_{k=0}^{K-1} \Gamma_k u_{t_k}(\omega) \right|^2 \equiv |\widehat{h}_K(\omega)|^2 \tag{10}$$

$$= \underbrace{\sum_{k=0}^{K-1} |\Gamma_k|^2}_{a_0} + 2 \sum_{k=0}^{K-1} \sum_{m=k+1}^{K-1} \underbrace{|\Gamma_k||\Gamma_m|}_{a_{k,m}} \cos(\omega t_{k,m} + \angle\Gamma_{k,m})$$

where $t_{k,m} = t_k - t_m$ and $\angle\Gamma_{k,m} = \angle\Gamma_k - \angle\Gamma_m$.

It is instructive to note that the relevant unknowns $\{\Gamma_k\}_{k=0}^{K-1}$ can be estimated from $\widehat{\boldsymbol{s}} \in \mathbb{R}^{2L+1}$ in (10) which in turn depend only on L as opposed to sampling rate and sampling instants. This aptly justifies our philosophy of *sampling without time*.

4. RECONSTRUCTION VIA PHASE RETRIEVAL

Given data model $\boldsymbol{m} = \mathbf{U}\mathbf{D}_{\widehat{\psi}}\widehat{\boldsymbol{s}}$ (9), we aim to recover the image parameters $|\Gamma_k|$ at each sensor pixel. For this purpose, we first estimate samples of $\widehat{s}(\omega)$ as $\widehat{\boldsymbol{s}} = \mathbf{D}_{\widehat{\psi}}^{-1}\mathbf{U}^+\boldsymbol{m}$, where $\mathbf{U}^+$ is the matrix pseudo-inverse of $\mathbf{U}$. This is akin to performing a weighted deconvolution, knowing that $\mathbf{U}$ is a DFT matrix. Next, we solve the problem of estimating $|\Gamma_k|$ in two steps (1) First we estimate $a_{k,m}$, and then, (2) based on the estimated values, we resolve ambiguities due to $|\cdot|$.
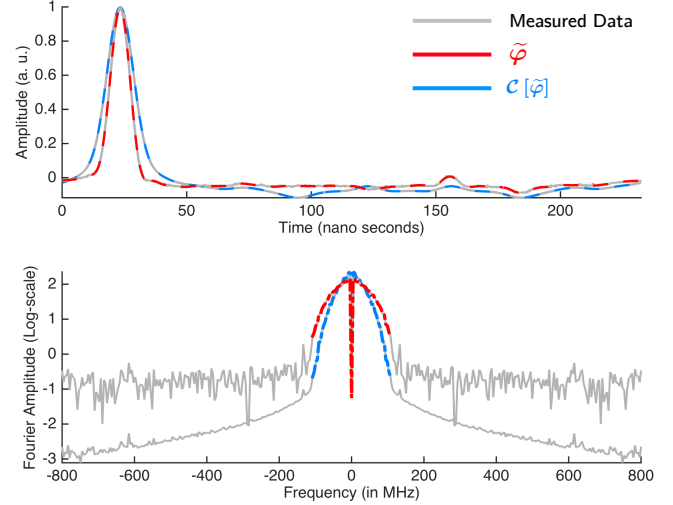

Fig. 3: Bandlimited approximation $\widetilde{\varphi} \approx \mathcal{C}[p]$ with $L = 25$ (red curves) and its cyclic auto-correlation $\mathcal{C}[\widetilde{\varphi}]$ (blue curve). Shown are the spatial-domain signals (top) and the corresponding Fourier-domain magnitudes in log scale (bottom).

Parameter Identification via Spectral Estimation: Based on the coefficients $\widehat{\boldsymbol{s}}$, one can then retrieve the amplitude and frequency parameters that are associated with the oscillatory terms as well as the constant value in (10). The oscillatory-term and constant-term parameters correspond to $\{a_{k,m}, t_{k,m}\}$ and a_0, respectively. All parameter values are retrievable from $\widehat{\boldsymbol{s}}$ through spectral estimation [29]; details are provided in [24] for the interested reader.

Note that, given the form of (10) and our acquisition model, the sparsity level of the sequence $\widehat{\boldsymbol{s}}$—corresponding to the total amount of oscillatory and constant terms—is $(K^2 - K)/2 + 1$. The magnitude values $a_{k,m}$ and a_0 can thus be retrieved if at least $L \geq (K^2 - K) + 1$ autocorrelation measurements are performed [24, 30], which is the case in the experiments described in Section 5.

Resolving Ambiguities: Based on the aforementioned retrieved parameters, one wishes to then deduce $\{|\Gamma_k|\}_{k=0}^{K-1}$. The estimated cross terms $\{a_{k,m}, t_{k,m}\}$ allow to retrieve the values of $|\Gamma_k|$ through simple pointwise operations. Here we will focus on the case, $K = 2$. Due to space limitations, we refer the interested readers to our companion paper [24] for details on the general case ($K > 2$). The case when $K = 2$,

$$\widehat{s}(\omega) \overset{(10)}{=} |\Gamma_0 u_{t_0}^*(\omega) + \Gamma_1 u_{t_1}^*(\omega)|^2 + \underbrace{\varepsilon_\omega(\Gamma_k, t_k)_{k>2}}_{\approx 0}$$

$$= \underbrace{|\Gamma_0|^2 + |\Gamma_1|^2}_{a_0} + \underbrace{2|\Gamma_0||\Gamma_1|}_{a_{01}} \cos(\omega t_{0,1} + \angle\Gamma_{0,1}),$$

may be a result of two distinct echoes (Fig. 2) or due to approximation of higher order echoes by two dominant echoes (due to inverse-square law). In this case, we effectively estimate two terms: a_0 and a_{01} (10). The set of retrieved magnitude values then amount to solving [24]:

$$||\Gamma_0| \pm |\Gamma_1|| = \sqrt{a_0 \pm a_{01}}, \qquad a_0, a_{01} > 0. \tag{11}$$

Thanks to the *isoperimetric property* for rectangles: $a_0 \geq a_{01}$, the r.h.s above is always positive unless there is an estimation error in which case, an exchange of variables leads to the unique estimates,

$$\left\{|\widetilde{\Gamma}_k|\right\}_{k=1,2} = \frac{\sqrt{\widetilde{a}_0 + \widetilde{a}_{01}} \pm \sqrt{\widetilde{a}_0 - \widetilde{a}_{01}}}{2}.$$
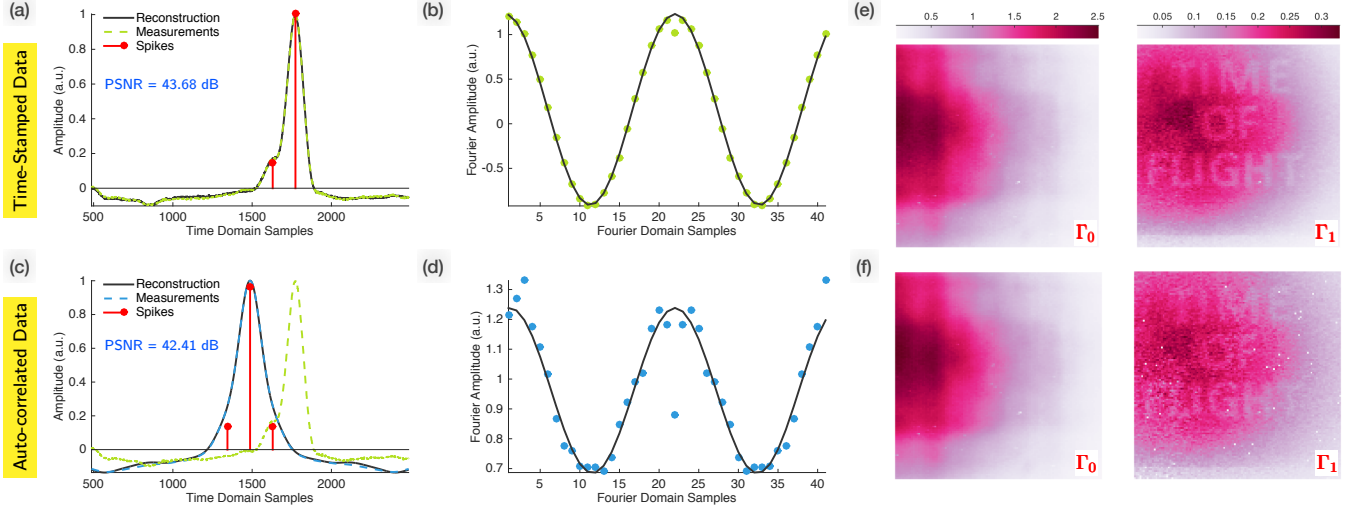
Fig. 4: Experimental results on acquisition and reconstruction of echoes of light in a scene. (a) Single pixel, time-stamped samples $y = \mathcal{C}[p] * \overline{h}_2$ serve as our ground truth because the phase information is known. We also plot the estimated, 2-sparse SRF (5) in red ink. (b) Fourier domain frequency samples of the SRF, $\widehat{h}_2(\ell\omega_0)$ with $L = [-20, 20]$ and $\omega_0 = 2\pi/\Delta$, $\Delta = 70$ ps. We plot the measured data in green and the fitted data in black. The frequencies are estimated using Cadzow's method. (c) Same pixel, time-stampless, low-pass filtered, auto-correlated samples (8) together with the estimated, auto-correlated SRF $\mathcal{C}[h_2]$ in red ink. (d) Fourier domain samples $\widehat{s}(\ell\omega_0)$ (10) and its fitted version (black ink). (e) Ground truth images for the experiment. (f) Estimated images using temporal phase retrieval.

5. EXPERIMENTAL VALIDATION

5.1. Simulation

Noting that measurements $\boldsymbol{m} = \mathbf{U}\mathbf{D}_{\hat{\psi}}\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{s}}$ are linked by an invertible, linear system of equations, knowing $\boldsymbol{m}$ amounts to knowing $\hat{\boldsymbol{s}}$. We have presented detailed comparison of simulation results in [24] for the case where one directly measures $\hat{\boldsymbol{s}}$. In this paper, we will focus on practical setting where the starting point of our algorithm is (9).

5.2. Practical Validation

Our experimental setup mimics the setting of Fig. 2. A placard that reads "Time–of–Flight" is covered by a semi-transparent sheet, hence $K = 2$. The sampling rate is $\Delta \approx 70 \times 10^{-12}$ seconds using a our custom designed ToF imaging sensor [13]. Overall, the goal of this experiment is to recover the magnitudes $\{|\Gamma_k|\}_{k=0}^{k=1}$ given auto-correlated measurements $m(n\Delta) \stackrel{(6)}{=} \mathcal{C}[y](n\Delta), \mathbf{m} \in \mathbb{R}^{2795}$. To be able to compare with a "ground truth", we acquire time-domain measurements $y(n\Delta)$ before autocorrelation, whose Fourier-domain phases are intact. In Fig. 4(a), we plot the non-autocorrelated measurements $\mathbf{y}$ while phase-less measurements $\mathbf{m}$ are shown in Fig. 4(c) from which we note that the samples are symmetric in time domain due to $m = \mathcal{C}[y]$ (cf. 6). In the first case (cf. Fig. 4(a)) where $y = \mathcal{C}[p] * \overline{h}_K$, we have [13],

$$\widehat{y}(\omega) = |\widehat{p}(\omega)|^2 h_K^*(\omega), \quad h_K(\omega) = \sum_{k=0}^{K-1} \Gamma_k u_{t_k}(\omega).$$

Similar to $\mathbf{m}$ in (9), in this case, the measurements read $\mathbf{y} = \mathbf{U}\mathbf{D}_{\widehat{\varphi}}\widehat{\mathbf{h}}$ and we can estimate the complex-valued vector $\widehat{\mathbf{h}}$. The phase information in $\widehat{\mathbf{h}}$ allows for the precise computation of $\{\Gamma_k, t_k\}_{k=0}^{K-1}$ [7,13]. These "intermediate" measurements serve as our ground truth. The spikes corresponding to the SRF (5) are also marked in Fig. 4(a). Fourier-domain measurements $\widehat{h}_2^*$ linked with $y(n\Delta)$ are shown in Fig. 4(b). Accordingly, $\widehat{h}_2^*(\omega)$, $\omega = \ell\omega_0, l = \{-L, \ldots, L\}$ where $\omega_0 = 2\pi/\Delta$ and $L = 20$. We estimate the frequency parameters $\{\Gamma_k, t_k\}_{k=0}^{K-1}$ using Cadzow's method [31]. The resulting fit is plotted in Fig. 4(b).

In parallel to Figs. 4(a),(b), normalized, autocorrelated data $\mathbf{m} = \mathcal{C}[\mathbf{y}]$ are marked in Fig. 4(c). The signal $\mathbf{y}$ is also shown as reference in green dashed line. We also plot the estimated $\mathcal{C}[h_2]$ (autocorrelated SRF). In Fig. 4(d), we plot measured and deconvolved vector $\widehat{\boldsymbol{s}} = \mathbf{D}_\psi^{-1}\mathbf{U}^+\boldsymbol{m}$ from which we estimate $a_0, a_{k,m}$ (10). The result of fitting using Cadzow's method with 41 samples is shown in Fig. 4(d).

The reconstructed images, $|\Gamma_0|$ and $|\Gamma_1|$, due to amplitude-phase measurements (our ground truth, Fig. 4(a)) are shown in Fig. 4(e). Alternatively, the reconstructed images with auto-correlated/intensity-only information are shown in 4(f). One can observe the great similarity between the images obtained in both cases, where only a few outliers appear in the phase-less setting. The PSNR values between the maps reconstructed with and without the phase information are of 30.25 dB for $|\Gamma_1|$ and 48.88 dB for $|\Gamma_0|$. These numerical results indicate that, overall, the phase loss that occurs in our autocorrelated measurements still allows for accurate reconstruction of the field of view of the scene.

Finally, in order to determine the consistency of our reconstruction approach, the PSNR between the original available measurements and their re-synthesized versions—as obtained when reintroducing our reconstructed profiles $|\Gamma_k|$ into our forward model—are also provided. As shown in Figs. 4(a) and (c), the PSNR values in the oracle and phaseless settings correspond to 43.68 dB and 42.41 dB, respectively, which confirms that our reconstruction approach accurately takes the parameters and structure of the acquisition model into account.

6. CONCLUSIONS

In this paper, we have introduced a method to satisfactorily recover the intensities of superimposed echoes of light, using a customized ToF imaging sensor for acquisition and temporal phase retrieval for reconstruction. Up to our knowledge, this is the first method that performs time-stampless sampling of a sparse ToF signal, in the sense that we only measure the amplitudes sampled at uniform instants without caring about the particular sampling times or sampling rates. This innovation can potentially lead to alternative hardware designs and mathematical simplicity as phases need not be estimated in hardware.

7. REFERENCES

[1] N. Ohnishi, K. Kumaki, T. Yamamura, and T. Tanaka, "Separating real and virtual objects from their overlapping images," in *Lecture Notes in Computer Science*. Springer Science + Business Media, 1996, pp. 636–646.

[2] A. Levin, A. Zomet, and Y. Weiss, "Separating reflections from a single image using local features," in *Proc. of IEEE Comp. Vis. and Patt. Recognition*, vol. 1, 2004, pp. I–306.

[3] A. M. Bronstein, M. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi, "Sparse ICA for blind separation of transmitted and reflected images," *Int. J. Imaging Syst. Technol.*, vol. 15, no. 1, pp. 84–91, 2005.

[4] N. Kong, Y.-W. Tai, and S. Y. Shin, "High-quality reflection separation using polarized images," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3393–3405, 2011.

[5] Y. Li and M. S. Brown, "Single image layer separation using relative smoothness," in *Proc. of IEEE Comp. Vis. and Patt. Recognition*, 2014, pp. 2752–2759.

[6] P. Chandramouli, M. Noroozi, and P. Favaro, "ConvNet-based depth estimation, reflection separation and deblurring of plenoptic images," University of Bern, submitted, 2016.

[7] A. Bhandari, A. M. Wallace, and R. Raskar, "Super-resolved time-of-flight sensing via FRI sampling theory," in *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Sig. Proc. (ICASSP)*, Mar. 2016, pp. 4009–4013.

[8] Y. C. Eldar, N. Hammen, and D. G. Mixon, "Recent advances in phase retrieval [Lecture Notes]," *IEEE Signal Process. Mag.*, vol. 33, no. 5, pp. 158–162, Sep. 2016.

[9] H. Qiao and P. Pal, "Sparse phase retrieval with near minimal measurements: A structured sampling based approach," in *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Sig. Proc. (ICASSP)*, Mar. 2016, pp. 4722–4726.

[10] B. Rajaei, E. W. Tramel, S. Gigan, F. Krzakala, and L. Daudet, "Intensity-only optical compressive imaging using a multiply scattering material and a double phase retrieval approach," in *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Sig. Proc. (ICASSP)*, Mar. 2016.

[11] K. Huang, Y. C. Eldar, and N. D. Sidiropoulos, "Phase retrieval from 1D Fourier measurements: Convexity, uniqueness, and algorithms," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6105–6117, Dec. 2016.

[12] Y. M. Lu and M. Vetterli, "Sparse spectral factorization: Unicity and reconstruction algorithms," in *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Sig. Proc. (ICASSP)*, May 2011.

[13] A. Bhandari and R. Raskar, "Signal processing for time-of-flight imaging sensors," *IEEE Signal Processing Magazine*, vol. 33, no. 4, pp. 45–58, Sep. 2016.

[14] S. Levy and P. K. Fullagar, "Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution," *Geophysics*, vol. 46, no. 9, pp. 1235–1243, Sep. 1981.

[15] J. J. Fuchs, "Multipath time-delay detection and estimation," *IEEE Trans. Signal Process.*, vol. 47, no. 1, pp. 237–243, Jan. 1999.

[16] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli, "Estimation of sparse MIMO channels with common support," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3705–3716, Dec. 2012.

[17] C. S. Seelamantula and S. Mulleti, "Super-resolution reconstruction in frequency-domain optical-coherence tomography using the finite-rate-of-innovation principle," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 5020–5029, Oct. 2014.

[18] P. Boufounos, "Compressive sensing for over-the-air ultrasound," in *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Sig. Proc. (ICASSP)*, May 2011.

[19] A. Velten, R. Raskar, D. Wu, B. Masia, A. Jarabo, C. Barsi, C. Joshi, E. Lawson, M. Bawendi, and D. Gutierrez, "Imaging the propagation of light through scenes at picosecond resolution," *Commun. ACM*, vol. 59, no. 9, pp. 79–86, Aug. 2016.

[20] D. Shin, F. Xu, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, "Computational multi-depth single-photon imaging," *Optics Express*, vol. 24, no. 3, p. 1873, Jan. 2016.

[21] A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar, "Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization," *Optics Letters*, vol. 39, no. 6, pp. 1705–1708, Mar. 2014.

[22] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar, "Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles," *ACM Trans. Graphics*, vol. 32, no. 6, pp. 1–10, Nov. 2013.

[23] J. J. Fuchs and H. Chuberre, "A deconvolution approach to source localization," *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1462–1470, Jun. 1994.

[24] A. Bhandari, A. Bourquard, S. Izadi, and R. Raskar, "Time-resolved image demixing," in *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Sig. Proc. (ICASSP)*, Mar. 2016, pp. 4483–4487.

[25] A. Bhandari, C. Barsi, and R. Raskar, "Blind and reference-free fluorescence lifetime estimation via consumer time-of-flight sensors," *Optica*, vol. 2, no. 11, pp. 965–973, Nov. 2015.

[26] K. Hibino, B. F. Oreb, D. I. Farrant, and K. G. Larkin, "Phase-shifting algorithms for nonlinear and spatially nonuniform phase shifts," *Jour.l of the Opt. Soc. of America A*, vol. 14, no. 4, p. 918, Apr. 1997.

[27] P. Hariharan, B. F. Oreb, and T. Eiju, "Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm," *Appl. Opt.*, vol. 26, no. 13, p. 2504, Jul. 1987.

[28] H. Pan, T. Blu, and M. Vetterli, "Towards generalized FRI sampling with an application to source resolution in radioastronomy," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 821–835, Feb. 2017.

[29] P. Stoica and R. L. Moses, *Introduction to spectral analysis*. Prentice Hall, Upper Saddle River, 1997, vol. 1.

[30] D. Potts and M. Tasche, "Parameter estimation for exponential sums by approximate prony method," *Signal Processing*, vol. 90, no. 5, pp. 1631–1642, May 2010.

[31] J. A. Cadzow, "Signal enhancement–A composite property mapping algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 1, pp. 49–62, 1988.