IOP Publishing Journal of Physics D: Applied Physics
J. Phys. D: Appl. Phys. **52** (2019) 144001 (15pp) https://doi.org/10.1088/1361-6463/aafe88

Compressed Hadamard microscopy for high-speed optically sectioned neuronal activity recordings

Vicente J Parot[1,2,9], Carlos Sing-Long[3,4,8,9], Yoav Adam[5], Urs L Böhm[5], Linlin Z Fan[5], Samouil L Farhi[5] and Adam E Cohen[5,6,7]

[1] Biophysics Program, Harvard University, Cambridge, MA, United States of America
[2] Division of Health Science and Technology, Massachusetts Institute of Technology, Cambridge, MA, United States of America
[3] Institute for Mathematical and Computational Engineering, School of Engineering and Faculty of Mathematics, Pontificia Universidad Católica de Chile, Santiago, Chile
[4] Institute for Biological and Medical Engineering, Schools of Engineering, Medicine and Biological Sciences, Pontificia Universidad Católica de Chile, Santiago, Chile
[5] Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, United States of America
[6] Department of Physics, Harvard University, Cambridge, MA, United States of America
[7] Howard Hughes Medical Institute, Harvard University, Cambridge, MA, United States of America

E-mail: vparot@fas.harvard.edu and cohen@chemistry.harvard.edu

Received 26 October 2018, revised 4 January 2019
Accepted for publication 15 January 2019
Published 4 February 2019

Abstract

Structured illumination microscopies achieve optical sectioning via differential modulation of in-focus and out-of-focus contributions to an image. Multiple wide-field camera images are analyzed to recreate an optical section. The requirement for multiple camera frames per image entails a loss of temporal resolution compared to conventional wide-field imaging. Here we describe a computational structured illumination imaging scheme, compressed Hadamard imaging, which achieves simultaneously high spatial and temporal resolution for optical sectioning of 3D samples with low-rank dynamics (e.g. neurons labeled with fluorescent activity reporters). We validate the technique with numerical simulations, and then illustrate with wide-area optically sectioned recordings of membrane voltage dynamics in mouse neurons in an acute brain slice and of calcium dynamics in zebrafish brain *in vivo*.

Keywords: optical sectioning, neuronal activity recording, microscopy

(Some figures may appear in colour only in the online journal)

1. Introduction

Fluorescence microscopy of dynamic biological samples is a powerful tool for studying physiology in its native context. An ever growing repertoire of fluorescent indicators reports on physiological variables, such as pH, calcium, glutamate, and membrane voltage [1, 2]. One would like to measure these signals with simultaneously high spatial resolution, high temporal resolution, optical sectioning and wide field of view. However, this task is challenging in turbid tissues such as the brain, because background fluorescence and light scattering mix in-focus and out-of-focus signals.

Optical sectioning techniques often entail tradeoffs between imaging parameters. For instance, one- and two-photon point scanning techniques provide high spatial resolution and good background rejection, but are limited in imaging speed by their point-scanning nature, and ultimately by the electronic excited state lifetime of the fluorophores which sets the

[8] Note: Carlos contributed to this work while affiliated to the Chemistry and Chemical Biology Department, Harvard University, Cambridge, MA, United States of America, as a Research Scholar.
[9] These authors contributed equally to this work.

minimum time between pixels to ∼10 ns for high-brightness fluorophores whose lifetimes are typically 2–4 ns. Camera-based systems offer intrinsically high data rates, good sensitivity and are readily coupled with wide-area optics, but do not natively provide optical sectioning. Spinning disk confocal microscopy provides high frame rates and optical sectioning, but the microlens array used in this technique does not provide adequate etendue for high light collection efficiency over a wide field of view.

Camera based optical sectioning using structured illumination has been used for decades to reject background signals [3]. In structured illumination microscopy (SIM) one projects a series of spatially patterned beams onto the image plane and uses a camera to acquire wide-field fluorescence images [4]. In-focus and out-of-focus contributions to the images are modulated differently between illumination patterns, enabling computation of an optical section. SIM was first implemented with multiple phases of sinusoidal illumination patterns, together with a simple demodulation algorithm ('stripe SIM') [5]. Interest in physiological signals such as neuronal activity has motivated approaches to increase time resolution by acquiring fewer frames to calculate an optical section. A fast SIM technique (HiLo) used only two illumination patterns: one (pseudo)-random speckle pattern and one homogeneous pattern [6–8].

We recently introduced a SIM technique called Hadamard microscopy, depicted in figure 1, where adjacent regions of the focal plane were illuminated with a set of temporally orthogonal patterns, and optically sectioned images were computed via matched filtering of the corresponding images [9]. Hadamard microscopy was slower than other SIM techniques because it required many camera frames (typically 12) to yield a single optical section. The maximum frame rate was 5.6 Hz in our recent implementation. This speed was fast enough to image a nuclear-localized Ca^{2+} indicator, but led to artifacts for signal sources whose intensity changed substantially during the time to acquire an optical section.

All three SIM techniques (stripe, HiLo, and Hadamard) yielded point-spread functions (PSFs) with identical full-width at half maximum (FWHM) along the axial and lateral line-sections, but Hadamard microscopy outperformed the other SIM techniques by other measures. The stripe SIM and HiLo PSFs had substantial out-of-focus conical lobes (lying neither along the axial nor lateral directions), while Hadamard did not. These lobes are an unavoidable consequence of the use of a small number of structured illumination patterns: a fluorescent source that is simultaneously axially and laterally displaced from a given in-focus pixel will emit some light along the same rays as the source at the targeted pixel. There is no way, even in principle, to distinguish these rays. By using many illumination patterns, Hadamard microscopy resolves this ambiguity in signal assignment.

In images of acute brain slices, the conical lobes in stripe SIM and HiLo caused out-of-focus cells to have a halo structure, which prevented clear resolution of single cells. Further, the (pseudo)-random illumination in HiLo imprinted static random noise, beyond the unavoidable photon shot-noise. This technical noise has been analyzed previously [10].

Ideally, one would like to combine the compact and low-noise PSF of Hadamard microscopy and the high time resolution of HiLo. Here we introduce a variant of Hadamard microscopy, termed compressed Hadamard imaging (CHI), which produces an optically sectioned movie at half the frame rate of the camera, matching the time resolution of HiLo. To achieve this dramatic speedup, we use our foreknowledge of the statistical structure of neural recordings. Namely, movies of neuronal dynamics tend to have a low-rank decomposition. The complete set of pixel intensities as a function of space and time can be approximately decomposed into a sum of images of individual neurons, each modulated with its own time-trace. Each neuron covers multiple pixels, so while a complete image of the neuron is not acquired in each camera frame, enough information is recorded to obtain an estimate of its fluorescence. Recent advances in microscopy signal processing have exploited the spatiotemporal structure of neuronal signals for unmixing, parallel detection, and computational volume reconstruction [11–13].

We validate the technique of CHI by applying it to simulated movies of 3D samples where the ground-truths of the source shapes and source dynamics are known. Then we apply this method to neuronal recordings using a genetically-encoded membrane voltage indicator in acute mouse brain slices, and to a genetically encoded cytosolic Ca^{2+} indicator in live zebrafish brain. These demonstrations achieve a ∼100-fold improvement in the time resolution of Hadamard microscopy compared to the previous demonstration [9], and match the speed of HiLo imaging.

2. Methods

2.1. Computational modeling

We validated CHI using numerical simulations of the optical system and of a dynamic brain-like 3D sample.

2.1.1. Simulating the sample. We simulated a 25 × 25 × 24 μm volume comprised of dynamically fluctuating cells in a foreground volume, plus static intensity cells in a background plane. The foreground volume was discretized into four parallel planes at depths $z = 0$, 4, 8, and 12 μm. 24 cells were arranged among these planes, 14 of which changed intensity over time. Each cell was represented by a 2D ellipse with a major axis that was selected uniformly at random between 1.35 μm and 1.65 μm, and a minor axis selected similarly between 0.9 μm and 1.1 μm, and random in-plane orientation. Cells in the same foreground plane were not allowed to overlap. The background plane was located at $z = 24$ μm, and included 100 static cells, determined as in the foreground planes except that overlaps were allowed. For all cells, the position, intensity, and orientation were randomly determined within predefined ranges. Each plane was sampled on a 64 × 64 equispaced grid ($\Delta x = \Delta y = 0.40$ μm). Camera images were simulated with the same dimensions and spacing.

Each active cell was assigned a firing pattern, drawn from a shifted exponential distribution of inter-spike intervals with a decay parameter of 2 s and minimum interval of 50 ms.

The mean firing rate per cell was thus once per 2.05 s. Each spike drove a uniform increase in cell brightness. The fluorescence timecourse of cell i was computed by convolving the firing times with an exponential response function, of the form $a_i t\, e^{-t/\tau}$ for $t > 0$ with $\tau = 500$ ms. The response amplitude a_i of cell i was selected from a uniform distribution on [0.75,1]. These parameters resulted in dynamic cells with an expected time-averaged intensity of approximately 1/2 of their maximum amplitude.

2.1.2. Simulating the optical system. The microscope was modeled as a space-invariant intensity-linear system comprising an illumination component, which projected patterned illumination onto the sample, and a collection component, which received light and formed an image in the camera. In experiments, the patterns of illumination were created by a digital micromirror device (DMD) which modulated a laser beam. The plane of the DMD was re-imaged onto the focal plane of the microscope.

Illumination and coding patterns. The illumination patterns comprised binary arrays on a rectangular grid with spacing corresponding to the pixel size of the DMD when projected into the focal plane at $z = 0$. In our simulations, which were matched to an experimental apparatus, each DMD pixel projected onto an area of $\approx 0.8\ \mu\text{m} \times 0.8\ \mu\text{m}$. Thus each DMD pixel mapped to 2×2 pixels on the camera plane.

The illumination pattern on each source plane was calculated by performing a discrete convolution between the binary code at $z = 0$ and a 3D illumination PSF, p_I, corresponding to a Gaussian beam focused at $z = 0$ given by

$$p_I(\vec{r}, z) = I_0 \left(\frac{w_0}{w(z)} \right)^2 \exp\left(\frac{-2|\vec{r}|^2}{w(z)^2} \right), \qquad (2.1)$$

where $\vec{r}$ represents a position on the sampled xy-plane, I_0 is the intensity at the center of the beam, set to $2/\pi w_0^2$ to normalize the transversal integral, and

$$w(z) = w_0 \sqrt{1 + \left(\frac{z}{z_R} \right)^2}$$

with $\lambda = 488$ nm and $w_0 = 0.75\ \mu$m. The Rayleigh range is defined as $z_R = \pi\, w_0^2/\lambda$. For the present example, $z_R = 3.62\ \mu$m.

The fluorescence emitted in each source plane was obtained by a point-wise multiplication between the illumination and the brightness of the cells.

Collection. Images were formed by convolving the emitted fluorescence across all source planes with a 3D PSF focused at $z = 0$, assumed to be the same as the illumination PSF (equation (2.1)).

Confocal simulation. To characterize performance of different dynamic estimation methods, simulated results were compared with a simulated confocal movie. This movie consisted of an instantaneously acquired optical section of the simulated sample at each timepoint, resulting from z-integrating the sample convolved with a 3D PSF focused at $z = 0$, assumed to be the product of illumination and collection PSFs. The PSFs were assumed to be Gaussian (equation (2.1)) with widths set by the combination of diffraction and DMD or camera pixel size, respectively. The calculation in appendix A gives $w_0 = 1.13\ \mu$m for excitation PSF width and 0.77 μm for collection PSF width. This collection PSF is equivalent to setting the confocal pinhole diameter to 1 Airy Unit $= 1.21\pi w_0$, corresponding to 2.92 μm at the sample. Confocal simulations were scaled to match the amplitude of reconstructed CHI sections at long code lengths. With this scaling, the mean intensity of raw camera frames in CHI was 11.36 fold larger than the mean intensity of confocal signals, reflecting the fact that CHI removes background computationally while confocal microscopy does so physically.

2.1.3. Sampling. A sequence of N_t camera images was simulated, each with a distinct illumination pattern, from $t = 0$ s to $t = 30$ s. Cell firing patterns were simulated from $t = -8$ s to achieve a steady activity by $t = 0$ s. The simulations assumed ideal behavior for the DMD (instantaneous switching of patterns) and for the camera (instantaneous acquisition of images). In some simulations we added Poisson shot noise to each pixel, with variance equal to the counts.

2.2. Microscope

Imaging experiments used a custom epi-fluorescence microscope equipped with a 488 nm laser (Coherent OBIS) under intensity control by an acousto-optic tunable filter (Gooch and Housego 48058-2.5-.55) and a 640 nm laser (DILAS MB-638.3-8C-T25-SS4.3). Both colors were patterned by a digital micromirror device (VIALUX V-7000), where half of the chip was used to pattern each wavelength, permitting simultaneous independent patterning of both colors. The two patterned beams were focused independently to correct for chromatic aberrations, and then combined via a dichroic beam splitter. The beams were then reflected by a dichroic mirror (Semrock Di01-R405/488/561/635). An astigmatic aberration due to reflection off a warped dichroic was compensated by a cylindrical pair tube lens. Excitation light was focused on the sample from the top by a 20× water-immersion objective (Olympus XLUMPLFLN, NA 1.0), then imaged onto a sCMOS camera (Hamamatsu Orca Flash 4.0). The imaging magnification was 20×, and projection magnification was 28×, resulting in camera and DMD pixel linear dimensions of 325 nm and 493 nm in the sample plane, respectively.

To synchronize illumination and imaging, the DMD was triggered to expose the sample only during the full camera exposure of a limited field of view (FOV), thus avoiding rolling shutter artifacts. For voltage imaging, the camera was set to record a FOV comprising 48 rows at 1 kHz frame rate, allowing 760 μs of full exposure in each frame, over a FOV 15.6 μm along the camera rows. This configuration constrained the illumination time to 76% of the total time. As patterns turned 50% of the pixels on, there was overall 38% time-averaged use of the available laser power. For Ca^{2+} imaging, a FOV with 1264

camera rows was acquired at 100 Hz, with 3.7 ms of illumination exposure time, and illumination efficiency 18%. In both cases, the FOV along the camera columns was not limited by frame rate. The camera was configured to run in synchronous acquisition mode and to output a 100 kHz master clock signal which synchronized the rest of the apparatus.

2.3. Animal experiments

All animal procedures were conducted following the National Institutes of Health (NIH) guide for the care and use of laboratory animals and approved by the Institutional Animal Care and Use Committee (IACUC) at Harvard University.

2.3.1. Brain slice preparation. A P1 mouse (C57BL/6, Charles River Labs #027) was anesthetized in ice for 5 min. Constructs coding for (1) hSynapsin-Cre and (2) Cre-dependent co-expression of soma-localized transmembrane Archaerhodopsin-derived voltage indicator and a soma-localized channelrhodopsin were packaged in AAV2/9 viral capsids (Massachusetts Eye and Ear Infirmary Virus Core and Janelia Virus Service Facility, respectively) and transcranially co-injected in unequal molar ratios -8.4×10^9 GC ml^{-1} for (1) and 1.1×10^{12} GC ml^{-1} for (2). This allowed a small portion of Cre-infected neurons to express voltage indicator and channelrhodopsin at high levels.

At age P40 the animal was anesthetized by isoflurane inhalation, and transcardially perfused with ice-cold slicing solution containing (in mM): 110 choline chloride, 25 $NaHCO_3$, 2.5 KCl, 7 $MgCl_2$, 0.5 $CaCl_2$, 1.25 NaH_2PO_4, 25 glucose, 11.6 ascorbic acid, and 3.1 pyruvic acid (310 mOsm kg^{-1}). The brain was cut into 300 μm thick slices using a vibratome (Leica VT1200S) and transferred to a 34 °C artificial cerebrospinal fluid (aCSF) bath containing (in mM): 125 NaCl, 2.5 KCl, 25 $NaHCO_3$, 2 $CaCl_2$, 1 $MgCl_2$, 1.25 NaH_2PO_4, 25 glucose (295 mOsm kg^{-1}). Slices were allowed to recover for 45 min at 34 °C and then transferred to room temperature. During imaging, room temperature aCSF was perfused at 1 ml min^{-1}. Solutions were continuously bubbled with carbogen (95% O_2 and 5% CO_2).

2.3.2. Zebrafish preparation. Zebrafish expressing cytosolic GCaMP6f under a neuronal promoter were used (Tg(elavl3:GCaMP6f)^{jf1} [14], a gift from Florian Engert). At day four post fertilization a fish was immersed for 10 min in a 1 mg ml^{-1} solution of α-bungarotoxin (Life Technologies B1601) to paralyze it, then embedded in 1.5% low melting point agarose in the 10 mm well of a glass-bottom 35 mm cell culture dish. The animal was incubated for 45 min in 75 mM pentylenetetrazole (PTZ, Sigma P6500) to induce seizure-like brain activity immediately prior to imaging.

2.4. Illumination patterns

Series of Hadamard-coded patterns were designed to illuminate neighboring sample locations with orthogonal time-sequences of intensity. Let $\boldsymbol{H}$ denote a Hadamard matrix of size m with elements in $\{-1, 1\}$ satisfying $\boldsymbol{H}^T\boldsymbol{H} = m\boldsymbol{I}_m$, where $\boldsymbol{I}_m$ is the identity of size m, and having ones in the first column and first row. The illumination P_{ij} of sequence i at time-step j was defined as $P_{ij} = (H'_{ij} + 1)/2$, where $\boldsymbol{H}'$ is the matrix formed by the last $m-1$ columns of $\boldsymbol{H}$. Orthogonality between illumination codes is verified by $\boldsymbol{P}^T\boldsymbol{H}' = (m/2)\boldsymbol{I}_{m-1}$. These codes were tiled spatially, assigning code $k_{ij} \in \{1, ..., m-1\}$ to the projector pixel (i, j) with $k_{ij} = \mathrm{mod}(i \cdot q + j, m-1) + 1$, where q was an offset parameter to maximize the spatial separation of repeating codes. To further reduce spurious crosstalk between repeating codes, a random binary mask B inverted the polarity of 50% of projected pixels. This process is shown in figure 1(b). Ca^{2+} activity was measured setting $m = 36$ and $q = 10$, while voltage recordings used $m = 20$ and $q = 5$.

2.5. Hadamard demodulation

Hadamard microscopy typically entails a tradeoff between the quality of optical sectioning (requiring more illumination patterns) and temporal resolution (requiring fewer illumination patterns). First, we review briefly the image reconstruction algorithm for Hadamard microscopy, then we show how to relax this tradeoff by taking advantage of the low-rank statistical structure of the underlying signals.

In Hadamard microscopy, one first acquires a set of calibration images $\boldsymbol{C}$ by projecting the illumination patterns onto a flat and homogeneous fluorescent film (often a coverslip with a stripe of fluorescent highlighter ink). Throughout this document, a video sequence of L frames will be represented as a matrix in which each column is a vectorized video frame, as $\boldsymbol{C} = [\boldsymbol{c}^1 \ldots \boldsymbol{c}^L]$. In this case, $\boldsymbol{c}^p$ denotes the video frame corresponding to the pth Hadamard calibration pattern. One then repeats the acquisition procedure on the real sample. Let $\boldsymbol{X} = [\boldsymbol{x}^1 \ldots \boldsymbol{x}^L]$ denote the resulting raw camera video. For the moment, we assume that both $\boldsymbol{C}$ and $\boldsymbol{X}$ have length L, equal to one repetition of the Hadamard code. The algorithm to compute a Hadamard optical section is:

$$\boldsymbol{x}^S := \sum_{p=1}^{L} \delta\boldsymbol{c}^p \odot \boldsymbol{x}^p$$

where $\delta\boldsymbol{c}^p = \boldsymbol{c}^p - \langle\boldsymbol{C}\rangle_L$ represents the deviation of each pixel from its average over all frames, and $\odot$ represents the entry-wise product between two arrays. A diagram of this calculation is shown in figure 1(c). Multiple sets of L images can be concatenated to construct an optically sectioned movie, albeit with a frame rate L-fold slower than the raw video.

2.6. Theory of dynamic optical sectioning

Our objective is to devise a practical optical sectioning procedure that matches the performance of an idealized procedure that we now describe. Suppose we had access through an oracle to a set of coded video sequences $\boldsymbol{X}^p$ for all p, i.e. at each time t we had access to images of the sample illuminated by all L different Hadamard patterns. In this case we could perform optical sectioning at the full frame rate by using the Hadamard image reconstruction algorithm described above for each time t.

This idealized method suggests an approach to attaining similar performance with physically attainable data. From a given recording $\boldsymbol{X}$, we seek to estimate all video sequences $\boldsymbol{X}^p$ where the sample has been illuminated with a fixed pattern across all frames. If we can perform this estimate for each p then we can reconstruct the high time-resolution optically sectioned movie.

Let $\boldsymbol{X}^w$ be the movie that would have been acquired with homogeneous wide-field illumination, i.e. conventional epifluorescence. The key assumption is that if $\boldsymbol{X}^w$ is approximately low-rank, then the temporal components of the factorized $\boldsymbol{X}^w$ are good approximations of the temporal components of each $\boldsymbol{X}^p$. In the following two sections we discuss: (i) how to estimate the constant-pattern data $\boldsymbol{X}^p$ from the wide-field data $\boldsymbol{X}^w$; and (ii) how to estimate the wide-field video $\boldsymbol{X}^w$ from data acquired with patterned illumination.

2.7. The relation between widefield and constant-pattern data.

We can always decompose the wide-field data $\boldsymbol{X}^w$ as a superposition of rank-1 video sequences by SVD:

$$\boldsymbol{X}^w = \boldsymbol{U}^w \boldsymbol{V}^{wT} = \sum\nolimits_i \boldsymbol{u}_i^w \otimes \boldsymbol{v}_i^w. \tag{2.2}$$

Here $\otimes$ represents the outer product, and ^{T} denotes transpose. For our simulated data $\boldsymbol{u}_i^w$ is a vector representation of a 64×64 image and $\boldsymbol{v}_i^w$ is a $T \times 1$ time modulation vector, where T is the number of frames. We assume the set $\{\boldsymbol{u}_i^w\}$ is orthogonal. We can consider the same decomposition for each hypothetical movie $\boldsymbol{X}^p$ where all frames are illuminated with the same single pattern p:

$$\boldsymbol{X}^p = \boldsymbol{U}^p \boldsymbol{V}^{pT} = \sum\nolimits_i \boldsymbol{u}_i^p \otimes \boldsymbol{v}_i^p.$$

We propose the following model for the relation between the spatial and temporal components of each decomposition. If there is only one term in (2.2) then it must be the case that the decomposition for each $\boldsymbol{X}^p$ has only one term: the effect of projecting a constant pattern onto a sample can certainly reduce the rank, but not increase it. In this case, it is reasonable that the temporal dynamics will be the same under either homogeneous or patterned illumination, i.e. that $\boldsymbol{v}_1^p = \boldsymbol{v}_1^w$, provided that a representative portion of the dynamic object is still illuminated.

Our key assumption is that this relation holds even when the video sequences have more than one component, and thus $\boldsymbol{v}_i^p = \boldsymbol{v}_i^w$ for every component i of the movie and illumination pattern p. Under this assumption,

$$\boldsymbol{X}^p = \boldsymbol{U}^p \boldsymbol{V}^{wT} = \sum\nolimits_i \boldsymbol{u}_i^p \otimes \boldsymbol{v}_i^w, \tag{2.3}$$

and if $\boldsymbol{V}^w$ is known, estimating $\boldsymbol{X}^p$ reduces to estimating $\boldsymbol{U}^p$.

2.8. Dynamic optical sectioning from widefield data and coded illumination patterns

Let $\boldsymbol{Y}$ be a video sequence where L coded illumination patterns have been used and repeated R times, and suppose first we also have access through an oracle to the wide-field data $\boldsymbol{X}^w$. Since each pth coding pattern is repeated R times, we have observed exactly R frames of the sequence $\boldsymbol{X}^p$. Therefore, the coded illumination data provides us with *incomplete* information about each $\boldsymbol{X}^p$. In figure 2, $\boldsymbol{Y}$ is represented as the structured illumination measured data, $\boldsymbol{X}^w$ is represented as paired frame widefield, and $\boldsymbol{X}^p$ is represented as estimated constant pattern data.

By our key assumption, if we have access to the wide-field sequence $\boldsymbol{X}^w$ we also know the dynamic components of each $\boldsymbol{X}^p$. This leads us to an estimation problem. Let T_p be the set of sampled times at which the pth pattern has been projected; each one of such sets has exactly R elements. To simplify the notation, let $\boldsymbol{m}_t^w$ be the tth *row* of the matrix $\boldsymbol{V}^w$. This allows us to write the consistency constraint

$$t \in T_p \quad \Rightarrow \quad \boldsymbol{y}_t = \boldsymbol{x}_t^p = \boldsymbol{U}^p \boldsymbol{m}_t^w. \tag{2.4}$$

We propose to estimate the matrix $\boldsymbol{U}^p$ by solving

$$\text{minimize} \quad \sum\nolimits_{t \in T_p} \|\boldsymbol{y}_t - \boldsymbol{U}^p \boldsymbol{m}_t^w\|_F^2 \tag{2.5}$$

where $\|\cdot\|_F^2$ denotes the Frobenius norm squared, i.e. the sum of the squared entries of its argument. In the above, we assume we are trying to find as many components as there are present in the widefield data. In this case, as long as the number of repetitions of the patterns is larger than the number of components we are solving for, the solution is unique. The first-order optimality condition characterizes the optimal solution $\widehat{\boldsymbol{U}}^p$ as the solution to

$$\widehat{\boldsymbol{U}}^p \sum\nolimits_{t \in T_p} \boldsymbol{m}_t^w \otimes \boldsymbol{m}_t^w = \sum\nolimits_{t \in T_p} \boldsymbol{y}_t \otimes \boldsymbol{m}_t^w, \tag{2.6}$$

from where it is clear the solution is unique as long as the second factor in the left-hand side is non-singular; this is generically true if the number R of repetitions is at least the number of components of the widefield sequence. If one can cleanly divide the movie into non-overlapping spatial domains with independent dynamics, then each domain can be analyzed separately. In this case, R must only exceed the rank of the domain with the largest number of components. We define the estimate for $\boldsymbol{X}^p$ as

$$\widehat{\boldsymbol{X}}^p = \widehat{\boldsymbol{U}}^p \boldsymbol{V}^{wT}.$$

This allows us to calculate a Hadamard optical section for time t as

$$\begin{aligned} \widehat{\boldsymbol{x}}_t^S &= \sum\nolimits_{p=1}^{L} \delta \boldsymbol{c}^p \odot \widehat{\boldsymbol{x}}_t^p \\ &= \sum\nolimits_{p=1}^{L} \delta \boldsymbol{c}^p \odot (\widehat{\boldsymbol{U}}^p \boldsymbol{m}_t^w) \\ &= \sum\nolimits_i \left(\sum\nolimits_{p=1}^{L} \delta \boldsymbol{c}^p \odot \widehat{\boldsymbol{u}}_i^p \right) \otimes \boldsymbol{v}_i^w. \end{aligned}$$

Consequently, our estimate of the optically sectioned movie $\widehat{\boldsymbol{X}}^p$ corresponds to the static Hadamard demodulation of the estimated spatial components, time-modulated by the temporal components of the widefield data.

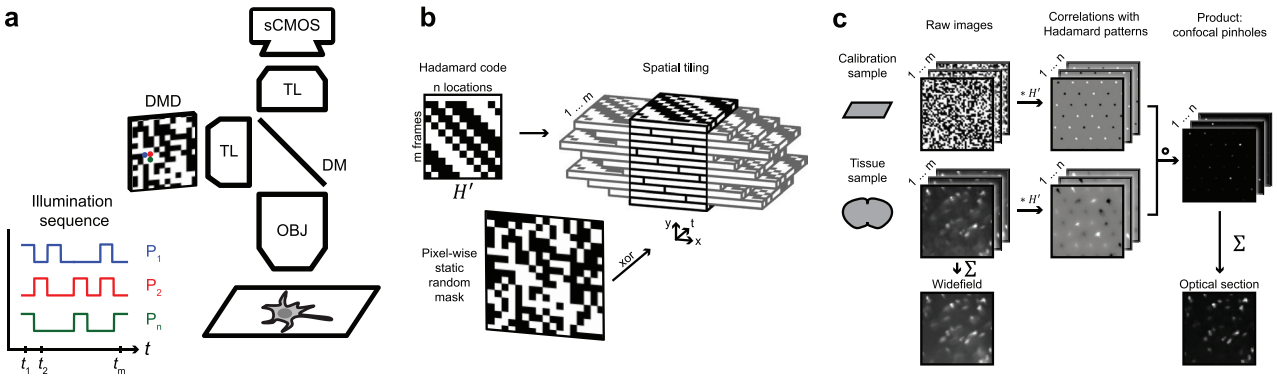

Figure 1. Hadamard optical sectioning microscopy. (a) A microscope projects orthogonal functions of illumination versus time in neighboring regions of the sample. Light is modulated by a DMD and measured with a camera. (b) Hadamard codes are tiled to fill the image FOV, and a random binary mask inverts 50% of pixels to reduce crosstalk between repeated code locations. (c) Raw images from a tissue sample are demodulated by matched filtering with raw data from a calibration sample. Filtered correlation maps are summed to form an optical section. Raw tissue data are summed to obtain a widefield image. Adapted from [9].

2.9. Estimating widefield data from coded illumination patterns

The construction of our estimated optical section relies on the computation of the right-singular vectors of $\boldsymbol{X}^w$. As we cannot *both* acquire a structured illumination sequence and a uniform illumination sequence simultaneously, we designed illumination patterns to overcome this problem. At the cost of reducing the dynamic reconstruction frame-rate by a factor of two, a complement-interleaved sequence of patterns enabled estimation of the widefield sequence *and* the dynamic Hadamard optical section, as follows.

Consider the video sequence where L coded illumination patterns have been used and repeated R times. We interleave the acquisition of each pattern with the acquisition of its complement, i.e. the binary pattern obtained by inverting each entry. Adding both patterns yields a uniform illumination pattern, i.e.

$$\widehat{\boldsymbol{x}}_t^w = \boldsymbol{y}_t + \bar{\boldsymbol{y}}_{t+\delta t} \tag{2.7}$$

where $\boldsymbol{y}_t$ is the data image acquired at time t and $\bar{\boldsymbol{y}}_{t+\delta t}$ is the data image acquired one camera frame later, with the complementary illumination pattern.

We can now estimate the temporal components of the widefield sequence, which we denote as $\widehat{\boldsymbol{V}}^w$; similarly, we let $\widehat{\boldsymbol{m}}_t^w$ be the estimate for $\boldsymbol{m}_t^w$. The dynamic Hadamard optical section becomes

$$\widehat{\boldsymbol{x}}_t^S = \sum\nolimits_i \left(\sum\nolimits_{p=1}^{L} \delta \boldsymbol{c}_p \odot \widehat{\boldsymbol{u}}_i^p \right) \otimes \widehat{\boldsymbol{v}}_i^w$$

where $\widehat{\boldsymbol{U}}^p$ is now computed from (2.6) by replacing $\boldsymbol{m}_t^w$ by its estimate $\widehat{\boldsymbol{m}}_t^w$.

In our implementation, we found small fluctuations in illumination that varied with the same period as the sequence of illumination patterns, a consequence of an un-identified instrumentation artifact. To correct this artifact, the widefield estimated movies were digitally divided by an estimate of the intensity variations.

2.10. Pseudocode

A pseudocode of the data analysis method is shown in algorithm 1. The computational complexity not including the computation of the SVD is $O\left(N \times R \times L \times P\right)$ where N is the number of SVD components, and P the number of camera pixels. MATLAB source code that generates figure 4(b) from raw data is available at https://github.com/adamcohenlab/Compressed-Hadamard-Code.

Algorithm 1. Compressed Hadamard estimation.

parameters:
  N ▷ number of components for truncated SVD
inputs:
  $\boldsymbol{M}$ ▷ tissue activity movie with $R \times 2 \times L$ patterned frames
  $\boldsymbol{C}$ ▷ calibration movie with $2 \times L$ patterned frames
initialization:
  $map \leftarrow [1, 2, 3, ..., 2 \times L, 1, 2, 3, ...]$ ▷ pattern order in $\boldsymbol{M}$
procedure CHI($\boldsymbol{M}, \boldsymbol{C}$)
  $\boldsymbol{M}^w \leftarrow \boldsymbol{M}_{even} + \boldsymbol{M}_{odd}$
  $\boldsymbol{C}^\delta \leftarrow \boldsymbol{C}_{even} - \boldsymbol{C}_{odd}$
  $\boldsymbol{U}, \boldsymbol{S}, \boldsymbol{V} \leftarrow \mathrm{SVD}(\boldsymbol{M}^w, N)$ ▷ $\boldsymbol{USV}^T$ approximates $\boldsymbol{M}^w$ with rank N
  $\boldsymbol{M}^r \leftarrow \boldsymbol{M}^w$
  for $i \leftarrow 1$ to N **do**
    $\boldsymbol{v} \leftarrow \boldsymbol{S}_{i,i}\boldsymbol{V}_i$

for $j \leftarrow 1$ to $2 \times L$ **do**
$\widehat{\boldsymbol{U}}_j^p \leftarrow \sum_{k\in map=j} \boldsymbol{M}_k^r \boldsymbol{v}_k / \sum_{k\in map=j} \boldsymbol{v}_k^2$ ▷ estimation
end for
$\widehat{\boldsymbol{U}}_i^w \leftarrow \sum_{k=1}^{2\times L} \widehat{\boldsymbol{U}}_k^p$
$\widehat{\boldsymbol{U}}^{p\delta} \leftarrow \widehat{\boldsymbol{U}}_{even}^p - \widehat{\boldsymbol{U}}_{odd}^p$
$\widehat{\boldsymbol{U}}_i^H \leftarrow \sum_{k=1}^{L} \widehat{\boldsymbol{U}}_k^{p\delta} \odot \boldsymbol{C}_k^\delta$ ▷ Hadamard demodulation
$\boldsymbol{M}^r \leftarrow \boldsymbol{M}^r - \widehat{\boldsymbol{U}}_{map}^p \boldsymbol{v}^T$ ▷ remove estimated component from residual
end for
end procedure
outputs:
$\widehat{\boldsymbol{M}}^w \leftarrow \widehat{\boldsymbol{U}}^w \boldsymbol{S}\boldsymbol{V}^T$ ▷ widefield movie
$\widehat{\boldsymbol{M}}^H \leftarrow \widehat{\boldsymbol{U}}^H \boldsymbol{S}\boldsymbol{V}^T$ ▷ optical section movie

2.11. Robustness

Accurate estimation of the spatial components relies on being able to: (i) estimate accurately the widefield sequence; and (ii) compute the time-modulation factors according to the rank of the sequence. Several factors may impact either step. For instance, the changes in sample intensity might be too fast relative to the rate at which the interleaved complementary patterns are illuminated, leading to an inaccurate estimate of the widefield sequence. Even if the widefield sequence is estimated accurately, its rank could be too high relative to the number of repetitions, and therefore the number of time-modulation factors might be underestimated.

This section addresses the variability of the proposed estimate for $\boldsymbol{U}^p$ with respect to these two factors. Suppose the widefield sequence $\boldsymbol{X}^w$ has a rank N_w, which we do not assume to be small, and consider a rank $\widehat{N}_w$ estimate $\widehat{\boldsymbol{X}}^w$ of $\boldsymbol{X}^w$ with $\widehat{N}_w \leqslant N_w$. Furthermore, assume the number of repetitions to be greater than the number of components to be estimated ($\widehat{N}_w \leqslant R$). As indicated before, this makes the estimate uniquely defined. In movies with e.g. a modest number of neurons, this is often the case. If the sources are compact and one can divide the movie into sub-regions each with a modest number of sources, one can then decrease the requirement on R in each sub-region.

In this setting $\boldsymbol{U}^p$ has N_w columns and each $\boldsymbol{m}_t^w$ has N_w entries. We can decompose their product as

$$\boldsymbol{U}^p \boldsymbol{m}_t = \begin{bmatrix} \boldsymbol{U}_0^p & \boldsymbol{U}_R^p \end{bmatrix} \begin{bmatrix} \boldsymbol{m}_{0,t}^w \\ \boldsymbol{m}_{R,t}^w \end{bmatrix} = \boldsymbol{U}_0^p \boldsymbol{m}_{0,t}^w + \boldsymbol{U}_R^p \boldsymbol{m}_{R,t}^w,$$

where $\boldsymbol{U}_0^p$ has now $\widehat{N}_w$ columns and $\boldsymbol{m}_{0,t}^w$ has $\widehat{N}_w$ entries. In this decomposition, $\boldsymbol{U}_0^p$ represents the $\widehat{N}_w$ components that we ought to estimate accurately, whereas $\boldsymbol{U}_R^p$ represents the remaining $N_w - \widehat{N}_w \geqslant 0$ components that we are neglecting. From (2.6), we see that $\widehat{\boldsymbol{U}}^p$ is determined by

$$\begin{aligned} \widehat{\boldsymbol{U}}^p \sum_{t\in T_p} \widehat{\boldsymbol{m}}_t^w \otimes \widehat{\boldsymbol{m}}_t^w &= \sum_{t\in T_p} \boldsymbol{y}_t^w \otimes \widehat{\boldsymbol{m}}_t^w \\ &= \sum_{t\in T_p} (\boldsymbol{U}^p \boldsymbol{m}_t^w) \otimes \widehat{\boldsymbol{m}}_t^w \\ &= \boldsymbol{U}_0^p \sum_{t\in T_p} \boldsymbol{m}_{0,t}^w \otimes \widehat{\boldsymbol{m}}_t^w + \boldsymbol{U}_R^p \sum_{t\in T_p} \boldsymbol{m}_{R,t}^w \otimes \widehat{\boldsymbol{m}}_t^w. \end{aligned}$$

A straightforward algebraic manipulation shows that

$$\begin{aligned} \left(\widehat{\boldsymbol{U}}^p - \boldsymbol{U}_0^p\right) \sum_{t\in T_p} \widehat{\boldsymbol{m}}_t^w \otimes \widehat{\boldsymbol{m}}_t^w &= \boldsymbol{U}_R^p \sum_{t\in T_p} \boldsymbol{m}_{R,t}^w \otimes \widehat{\boldsymbol{m}}_t^w \\ &+ \boldsymbol{U}_0^p \sum_{t\in T_p} (\boldsymbol{m}_{0,t}^w - \widehat{\boldsymbol{m}}_t^w) \otimes \widehat{\boldsymbol{m}}_t^w. \end{aligned}$$

Consequently, when the second factor in the left-hand side is well-conditioned, the error $\widehat{\boldsymbol{U}}^p$ incurs in estimating $\boldsymbol{U}_0^p$ is essentially controlled by two dependent factors: (i) the decay in the singular values of the widefield sequence $\boldsymbol{X}^w$ as quantified by the size of the modulation factors $\boldsymbol{m}_{R,t}^w$ for the components we are *not* estimating; and (ii) the accuracy in the estimate of the modulation factors $\boldsymbol{m}_{0,t}^w$ for the components we are estimating.

This analysis shows that our approach to estimating $\boldsymbol{X}^w$ should be evaluated according to these criteria. The first criterion depends on the properties of the sample being imaged, namely the activity of the cells during the observation and the number of cells active in the field of view at any given time. It thus depends on the specifics of the experimental setup. The second criterion can be analyzed as follows. If we let $\overline{\boldsymbol{U}}^p$ be the components corresponding to the inverted pth pattern and the pth pattern is illuminated at time t, the estimate (2.7) implies

$$\begin{aligned} \widehat{\boldsymbol{x}}_t^w &= (\boldsymbol{U}^p + \overline{\boldsymbol{U}}^p)\boldsymbol{m}_t^w + \overline{\boldsymbol{U}}^p(\boldsymbol{m}_{t+\delta t}^w - \boldsymbol{m}_t^w) \\ &\approx \boldsymbol{x}_t^w + \overline{\boldsymbol{U}}^p(\boldsymbol{m}_{t+\delta t}^w - \boldsymbol{m}_t^w) \end{aligned}$$

where we consider that $\boldsymbol{U}^p + \overline{\boldsymbol{U}}^p \approx \boldsymbol{U}^w$. Consequently, the error when estimating the components of the widefield sequence will be mostly due to the rate of change of the temporal components, which depends on the rate at which each cell in the sample activates. Dynamics faster than the temporal resolution (half the camera frame rate) will not be accurately captured by the proposed method. Large variations of fluorescence between the instants when the sample is illuminated by a given pattern and by its complement will give rise to artifacts. In the estimated widefield movie, a spatial artifact corresponding to the momentary illumination pattern will arise. This effect is formalized in appendix B.

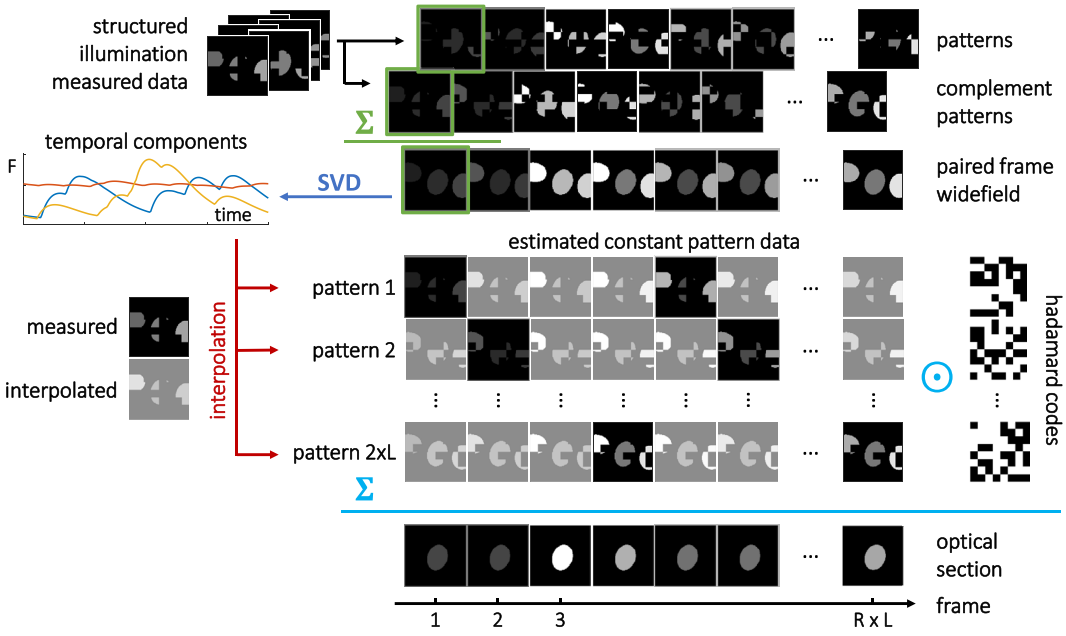

Figure 2. Estimation of optical section dynamics from compressed Hadamard measurements. Green: pairs of images acquired with complement illumination patterns are summed to form a half-frame rate widefield movie. Blue: low rank temporal components are calculated using singular value decomposition (SVD). Red: missing frames in constant-pattern movies are interpolated from measured patterns and temporal evolution of components. Cyan: Hadamard demodulation of estimated patterned frames results in a dynamic optical section.

3. Results

3.1. Numerical experiments

We simulated movies of a dynamic 3D sample, approximately matched to brain tissue expressing a nuclear-localized Ca^{2+} indicator. This process is illustrated in figure 3(a), left panel. For each set of sample dynamics, we simulated several different imaging modalities. (1) Spatially uniform illumination and wide-field imaging onto a camera ($\boldsymbol{X}^w$); (2) patterned illumination and widefield imaging ($\boldsymbol{Y}$); (3) instantaneous confocal acquisition of the whole FOV; (4) a hypothetical set of movies where each illumination pattern was acquired for all frames ($\boldsymbol{X}^p$); and (5) a calibration pattern dataset ($\boldsymbol{C}$) that simulated projection on a thin constant fluorescent layer in the focal plane. Figure 3(a) shows calculations of input datasets in the center panel, and resulting optical sections in the right panel. These experiments investigate how the proposed method is affected by measurement noise, and by parameter choices in the code length, the number of repetitions of periodic codes, and the number of SVD components.

To first characterize the static Hadamard demodulation method, we simulated the ideal (noiseless) acquisition of a static sample using different lengths of Hadamard code. Ideal Hadamard demodulation results in a confocal-like optical section [9]. However our simulation represented two non-ideal aspects of practical Hadamard microscopy: (1) codes were repeated to tile the FOV, introducing crosstalk between spatial repetitions of the same code; (2) the calibration method measured the illumination pattern convolved with the collection PSF, whereas one would ideally measure the illumination pattern at the sample itself. Errors were calculated as the root mean squared deviations for all pixels and all timepoints, normalized by the mean of the reference movie. As anticipated, for a time-invariant sample, increasing Hadamard code length led to improved correspondence of Hadamard demodulation and simulated confocal (figure 3(b)).

We then tested the dynamic estimation method with simulated movies of time-varying samples. For a desired length of experiment, camera frame rate, and total number of pictures, one could choose more repetitions of a short code, or fewer repetitions of a long code. To characterize this tradeoff, we simulated an ideal (noiseless) movie for a range of Hadamard code lengths (8, 12, 20, 24, 32, 40, 48, and 60), using periodic repetitions of the complement-interleaved pattern sequence (60, 40, 24, 20, 15, 12, 10, and 8 repetitions, respectively) for a total of 960 frames in all cases. Estimated CHI optically sectioned movies ($\widehat{\boldsymbol{X}}^S$) were compared to confocal images (accounting for the confocal point-spread function, but assuming instantaneous image acquisition across the whole field of view) and to 'full demodulation' Hadamard optical section movies ($\boldsymbol{X}^S$, calculated by demodulating the full set $\boldsymbol{X}^p$ for each t, i.e. assuming that images with all Hadamard illumination patterns were available at each time-point).

Figure 3(c) shows the discrepancy between CHI versus the confocal simulations (blue circles), indicating an intermediate value of code length as optimal. The comparison of CHI with 'full demodulation' Hadamard showed the reconstruction error due to the dynamic estimation process, while avoiding confusion with systematic errors (relative to confocal) that were present in both Hadamard demodulation processes, i.e. from tile-to-tile optical crosstalk. Figure 3(c) (red squares) shows that the dynamic estimation error increased with code length, as the number of code repetitions decreased. We ascribe this effect to rank under-representation at longer code

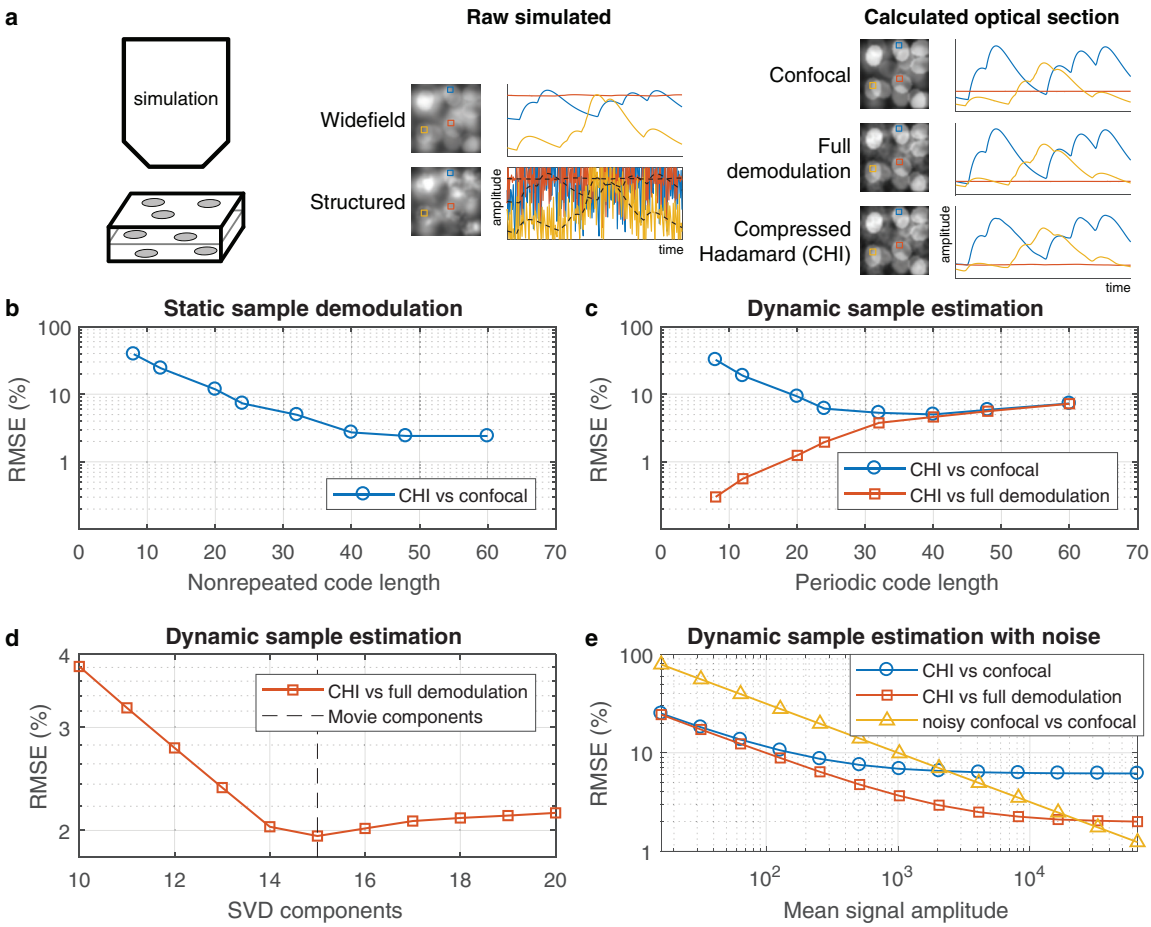

Figure 3. Compressed optical sectioning in simulated data. (a) Left: simulations modeled a 3D sample of neuronal activation under patterned illumination. Center: simulated datasets used in calculation. The fluctuations in the 'Structured' traces represent the effect of modulated illumination. Right: optical section datasets computed by different methods. 'Confocal' is an idealized instantaneously acquired confocal image. 'Full demodulation' is an idealized Hadamard section in which images with all illumination patterns are known for each frame. 'Compressed Hadamard' represents a physically realizable approach in which one illumination pattern is used per frame, and the missing data are inferred (CHI). (b) In a static sample, increasing the Hadamard code length increased the size of the spatial tiles, and reduced systematic error due to tile-to-tile crosstalk. (c) In a dynamic sample with a fixed number of frames (enforcing a tradeoff between code length L and number of repeats R), short codes introduced errors into Hadamard reconstructions due to tile-to-tile crosstalk. Long codes had less tile-to-tile crosstalk, but fewer repetitions, R, leading CHI to suffer relative to 'full demodulation'. (d) For a fixed code length, the CHI error relative to full demodulation was minimized when the number of SVD components used (e.g. in equation (2.2)) was equal to the rank of the data. (e) CHI reconstructions encountered a shot noise limited regime at low counts under Poisson noise (where error is inversely proportional to square root of intensity), but at higher counts errors approached the level of systematic errors present in the noiseless case. The shot noise limited error in CHI was ∼3-fold smaller than in confocal when signal levels were matched. The two techniques were comparable when raw photon counts were matched.

lengths. The dynamic estimation error was less than 2% for code length 24, the longest code length for which the simulated sample rank (15) was over-represented by the number of code repetitions (20).

CHI uses the singular value decomposition (SVD) to represent movie dynamics. To characterize the effect of the number of SVD components, we applied CHI with code length 24 (corresponding to $R = 20$ replicates) to a simulated movie with 15 components. We calculated reconstruction error as a function of number of SVD components. The error, shown in figure 3(d), was minimized when using the same number of components as in the data, but did not suffer very much from using more than the ideal number of SVD components.

To test the effect of shot noise on CHI reconstructions, we simulated camera movies with mean counts per pixel ranging from ∼1 to 5×10^4 and imposed Poisson-distributed noise. We used CHI to estimate optically sectioned movies, and compared with noiseless confocal and noiseless 'full demodulation' Hadamard movies (figure 3(e)). Two noise regimes were apparent. At low counts, shot noise dominated and error was inversely proportional to the square root of the signal amplitude. At high counts, the noise asymptotically approached the level of the noiseless case. No instabilities or numerical artifacts were observed in estimates from noisy datasets.

To benchmark the photon economy of CHI reconstructions, we compared CHI and confocal when both were subject to Poisson-distributed shot noise. The two sets of images were scaled to have the same mean counts, and the error in each was calculated relative to a noiseless confocal dataset. As expected, the confocal shot noise error was inversely proportional to the square root of mean counts. At low intensity, the confocal error was larger than the CHI error by a factor

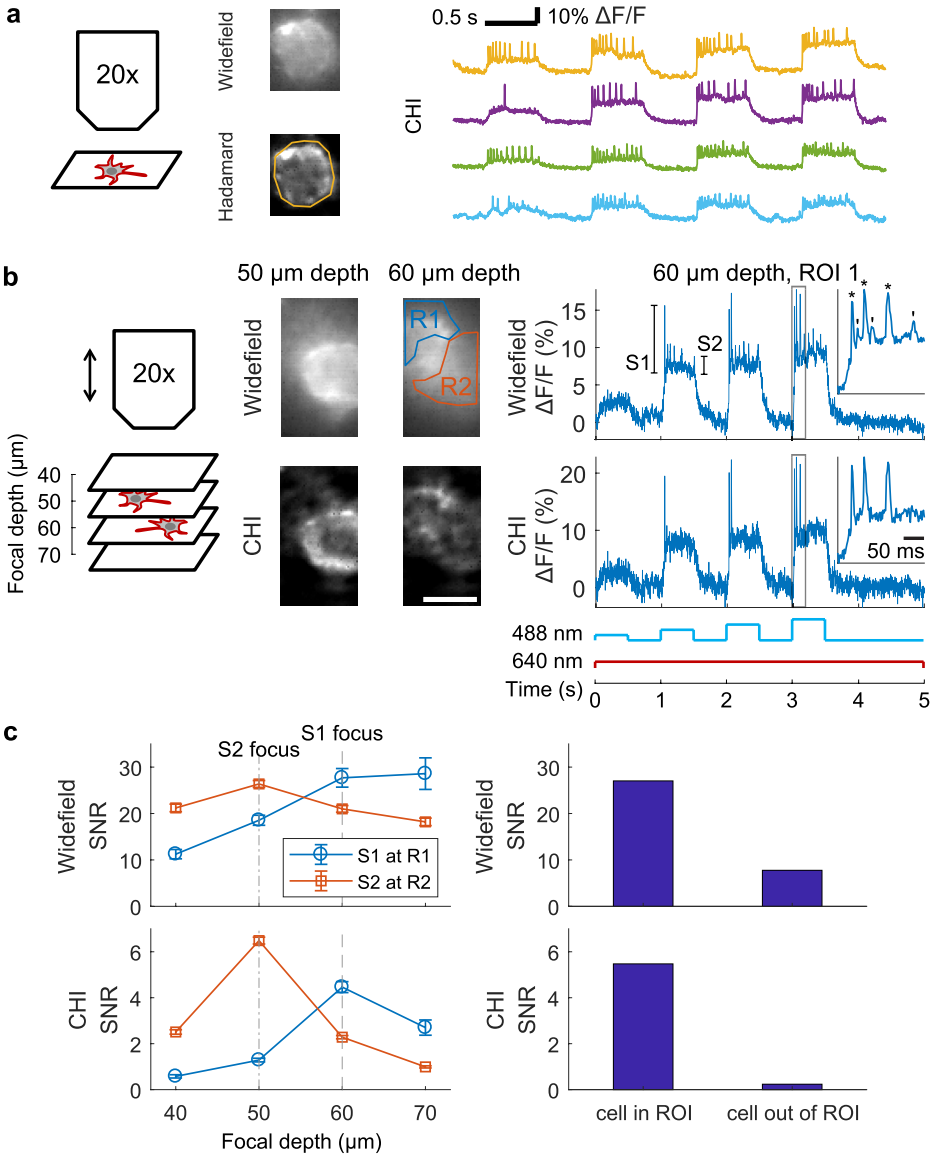

Figure 4. Compressed Hadamard imaging of high-speed neuronal voltage dynamics in brain tissue. CHI optical sections were computed at a 500 Hz reconstructed frame rate in acute mouse brain slices expressing an Optopatch construct. (a) Example single-neuron recordings showing optically evoked and optically recorded action potentials. The CHI cell images showed increased membrane contrast and background suppression relative to wide-field images. The first two traces come from the cell and ROI in the image. The last two come from another cell. (b) Two partially overlapping neurons (S1 and S2) were recorded in the same FOV, and the measurement repeated at multiple depths. Scale bar 10 μm. Two ROIs (R1 and R2) reported mixed signals from both neurons in the wide-field images. The same ROIs reported single-cell spike patterns in the CHI images. (c) Quantification of action potential SNR as a function of focal plane depth. Left, top: in wide-field recordings, both cells contributed to the spikes in both ROIs at all focal depts. In CHI, the signal from each cell was localized around the corresponding focal plane. Error bars represent s.e.m. Right: Hadamard removed spurious signals from out-of-focus cells, increasing the signal contrast. Maximum SNR was reduced by a factor of ≈5 in CHI sections compared to the widefield reference.

of 3.28. This discrepancy reflects the 11.36-fold greater mean counts in the raw CHI frames ($\sqrt{11.36} = 3.37$), i.e. the two techniques have comparable shot noise properties when raw photon counts are matched. We attribute the favorable noise properties of CHI to the denoising properties of the SVD truncation step, which has been characterized previously [15, 16]. In general, the noise properties of CHI will depend on the amount of background autofluorescence (the image reconstruction process removes this background on average, but not its shot noise), whereas the noise properties of confocal are insensitive to the background because background is rejected physically before photon detection. Practical implementations of camera-based spinning disk and point-scanning confocal systems can produce optical sections with minimal noise from background fluorescence. Spinning disk confocal can readily image at a 1 kHz camera frame-rate, while point-scanning is typically slower due to limitations on galvo mirror acceleration. However neither spinning disk nor point-scanning confocal techniques readily extend to high NA large FOV optics, as further elaborated upon in the Discussion section.

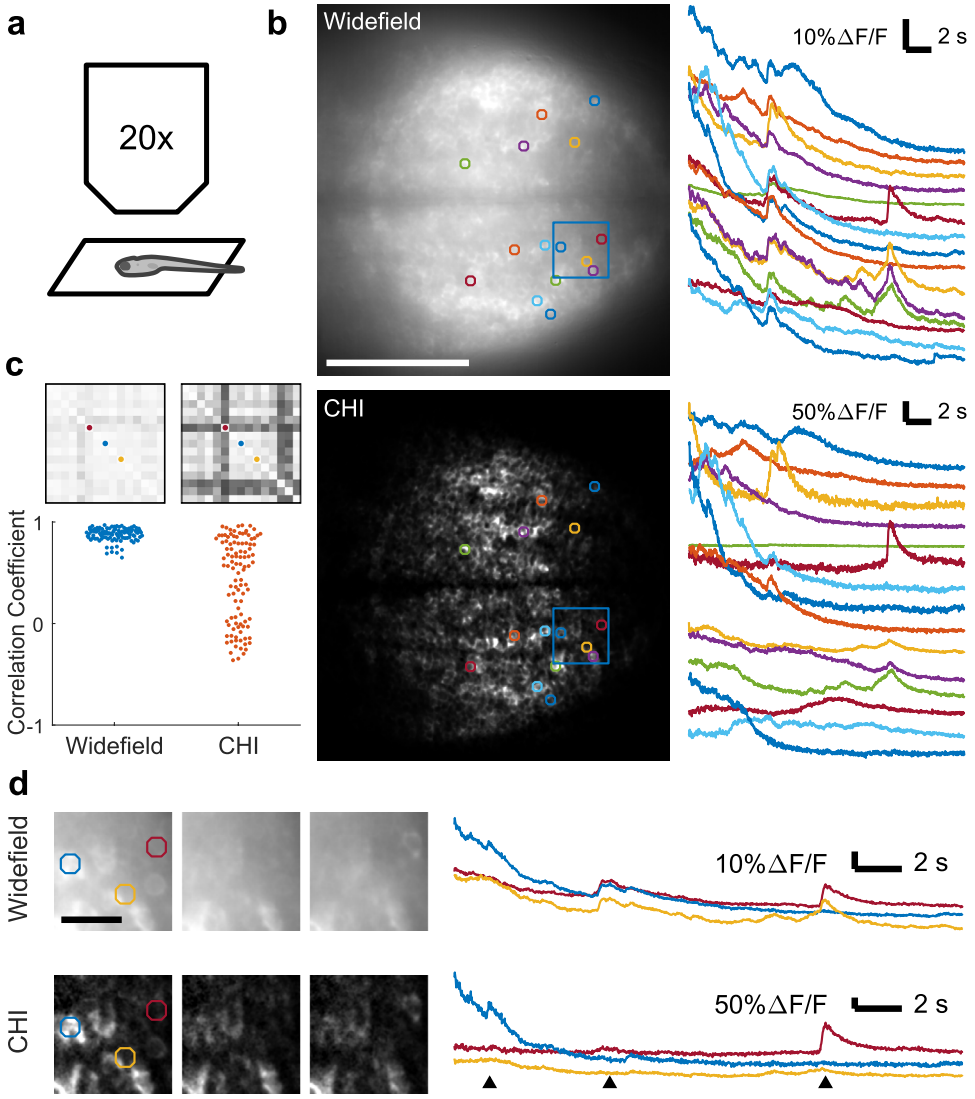

Figure 5. CHI high speed optical sectioning of large FOV, high dynamic rank data. (a) Experimental setup for *in vivo* zebrafish imaging. (b) Brain-wide FOV containing hundreds of neurons reconstructed at 50 Hz time resolution. Top: widefield; bottom: CHI. Images show the time average of movies. Scale bar 100 μm. Traces correspond to ROIs sorted from top to bottom. (c) CHI recordings showed increased contrast relative to widefield in the cell-to-cell cross correlation matrix, reflecting rejection of out-of-focus common mode signals. Top: cross-correlation matrices for widefield (left) and CHI (right) signals. Bottom: distribution of the off-diagonal correlation coefficients showing the spuriously high widefield cross-correlations due to out-of-focus signals. (d) Small ROI corresponding to one reconstructed block (from blue frame in (b)), showing selected Widefield and Hadamard movie frames and ROIs. Triangles indicate timepoints of selected movie frames. Scale bar 20 μm.

3.2. Compressed Hadamard imaging in live tissue

All-optical electrophysiology in acute brain slice. We tested CHI in acute mouse brain slice. A sparse subset of neurons co-expressed a blue-excited optogenetic actuator and a red-excited genetically encoded voltage indicator. Targeted blue light stimulation evoked action potentials (APs), which were detected in changes of the fluorescence of the voltage indicator (figure 4(a), left). Similar preparations have previously been used for high-throughput all-optical mapping of neuronal excitability in intact brain tissue [17]. Data acquisition at a 1 kHz framerate allowed activity reconstructions at 500 Hz. This experiment used a Hadamard code length of 20 patterns in a complement-interleaved sequence, repeated 240 times over 9.6 s.

We tested the ability of CHI to register the short fluorescence transients associated with APs. The blue stimulation pattern consisted of 500 ms every 1 s, increasing in intensity from 10 to 40 mW cm^{-2}. To reconstruct images, each acquisition was analyzed with the same algorithm, selecting five components for the SVD. This procedure yielded estimates of the widefield epifluorescence movie and Hadamard optical section. Figure 4(a) (center panel) shows time-averaged images of an example movie reconstructed in Widefield (top) and CHI (bottom), showing an increased membrane contrast and background rejection in the CHI optical section; the right panel shows time traces integrated from a cell mask ROI in multiple CHI movies. Slow transients uncorrelated with optogenetic stimulation were attributed to sub-threshold potential fluctuations.

To test the ability to reject out-of-focus signals, we studied two nearby, partially overlapping neurons, located at 50 and 60 μm below the tissue surface. We repeated the stimulation and imaging protocol while translating the objective z position to focus at different depths ($z = 40$, 50, 60 and 70 μm below the tissue surface, left panel of figure 4(b). Widefield and Hadamard movies were reconstructed at 500 Hz from each recording. Figure 4(b) shows in the center panel time averaged images at the depths focused on each neuron soma. CHI images showed reduced background fluorescence and higher cell membrane contrast, compared to the widefield images.

Two ROIs were manually defined to cover part of each neuron and to avoid their region of overlap. ROI-integrated time traces clearly showed APs in each recording. In widefield movies, the APs were mixed in each ROI, with distinct amplitudes from the two cells (figure 4(b), right panel insets, asterisks and apostrophes). The SNR was defined as the ratio of mean AP spike height to baseline standard deviation during a spike-free interval. In CHI movies, the signal from each cell was localized to the corresponding focal plane, and there was little crosstalk between ROIs (figure 4(c), left). These metrics indicate that CHI provided fast optical sectioning and reduced crosstalk relative to widefield Imaging. However, the CHI signals had reduced SNR relative to widefield by a factor of four for neuron 1, and by a factor of seven for neuron 2 (figure 4(c), left), reflecting the fact that CHI optical sections sampled data from only a single focal plane in the cell, whereas widefield images included fluorescence from the whole cell.

Brain-wide neuronal recording in live zebrafish. We implemented CHI in a larger FOV format to make neuronal recoding across a zebrafish brain densely expressing the Ca^{2+} indicator GCaMP6 (figure 5(a)). The animal was paralyzed with α-bungarotoxin to avoid motion artifacts, and the drug PTZ was administered to induce brain-wide epileptiform activity. No optogenetic stimulation was used. We recorded at a framerate of 100 Hz, with a 410 μm FOV. These parameters yielded brain-wide recordings with CHI reconstructions at 50 Hz, sufficient to capture Ca^{2+} dynamics. Movie data were analyzed block-wise (40 μm side square blocks, 13 μ m sliding step, with posterior weighted average reconstruction) under the assumption of locally low rank dynamics (SVD with 50 components per block). This experiment used a Hadamard code length of 36 patterns in a complement-interleaved sequence, repeated 30 times over 21.6 s.

From one acquisition, widefield and Hadamard movies were estimated with CHI. Figure 5(b) shows the dramatic difference in contrast between widefield and CHI in the time-averaged images. To favor an unbiased ROI-based analysis, 15 cell-sized ROIs were automatically selected by Gram–Schmidt orthogonalization of the widefield movie, iteratively selecting ROIs of largest variance over time in the residual movie after projecting out the signals from previous ROIs. This process selected ROIs with the most different and significative signals in the widefield movie. ROIs in widefield data revealed a diversity of transients, with a large common mode decay beginning at $t = 0$, attributed to visually-evoked neural activity (figure 5(b), right panel). The CHI movie showed large initial transients in only a subset of ROIs. To quantify how different were the neural signals among each set, we calculated their normalized cross-correlations, shown in figure 5(c). Figure 5(d) shows the detail of one reconstruction block, with selected ROIs and corresponding time traces, illustrating the suppression of out-of-focus fluorescence and mixed signal transients. This experiment demonstrated the implementation of high spatiotemporal bandwidth computational optical sectioning, allowing wide-brain neural recording of Ca^{2+} transients.

4. Discussion

CHI microscopy is a novel form of computational optical sectioning using structured illumination and ideas from compressed sensing to achieve simultaneously high spatial and temporal resolution. We performed detailed numerical simulations to characterize the properties of the technique. The CHI technique relies on the low-rank statistical structure of neural recordings: while the number of pixels in each image is large, the number of independently varying fluorescence components is much smaller. The spatial and corresponding temporal components of the signals are estimated separately, and then combined to yield a high spatio-temporal resolution reconstructed movie.

We demonstrated these advances by recording wide-area optically sectioned movies in live tissue, including voltage imaging of neurons in acute mouse brain slice with 500 Hz time resolution, and brain-wide Ca^{2+} imaging in live zebrafish with 50 Hz time resolution. These recordings represent a $\sim$100-fold increase in time resolution compared to the previous report of Hadamard microscopy (without compressed sensing dynamic reconstruction) in which Ca^{2+} dynamics were recorded at 5.6 Hz [9]. CHI can be readily adapted to measure other fast physiological signals with fluorescence microscopy, such as pH, glutamate, acetylcholine, and membrane tension, enabling wide-area high speed optically sectioned recordings while avoiding off-focal contaminating signals.

How does CHI compare to other optical sectioning microscopy techniques? CHI allows optical sectioning at half the maximum pixel rate of the camera, far surpassing point-scanning methods in temporal resolution. The time resolution of CHI is equivalent to that of HiLo microscopy [18, 19], but we previously showed that Hadamard image reconstruction offers better background rejection than HiLo and avoids systematic speckle noise that occurs in HiLo due to the use of a single pseudo-random illumination mask in that technique [9, 10]. Both CHI and HiLo microscopy are susceptible to shot noise from out-of-focus background. The extent to which this noise degrades the signal depends on the sample structure. CHI is anticipated to have lower shot noise than HiLo in the computed images due to the filtering effect of the low-rank representation of the data; though we note that similar post-processing could be applied to HiLo data. Commercially available spinning disk confocal systems offer high-speed imaging and physical optical sectioning, but are incompatible with large-area, high numerical aperture objectives. The micromirror-based optical patterning

system of CHI is compatible with any objective. The micromirror system also facilitates combination of the technique with patterned illumination for optogenetic stimulation or photochemistry. To retrofit a fluorescence microscope with a CHI system is approximately 1/5 of the cost of a spinning disk system and 1/10 of the cost of a two-photon imaging system.

We note that the dynamic estimation algorithm could be adapted to work with any structured illumination technique, including stripe SIM [4, 5] or HiLo [7]. For instance, the PSF of stripe SIM could be improved by successive illumination with stripes of different orientation and spatial frequency. With conventional image reconstruction algorithms this approach would decrease time resolution; but with compressed sensing image reconstruction, the full time resolution would be maintained. Hadamard encoding achieves the shortest possible packing of orthogonal codes in binary matrices, and is optimal to minimize the code length while having a flexible choice of spatial code separation.

For practical CHI implementations, acquisition parameters have to be chosen to match experimental needs. The biological process and sample of interest will determine the necessary spatial and temporal resolution. Here follows a step-by-step guide to choose these parameters:

(i) Imaging resolution, including optical system and camera spatial resolution, should be determined by the size of static sample features that need to be resolved.
(ii) DMD projection elements should be resolvable by the imaging system, but smaller than the dynamic objects to be measured.
(iii) Camera framerate should be twice the maximum temporal resolution of interest.
(iv) Recording length should be maximized within the limits set by sample stability and technical limitations such as data storage.
(v) Finally, the code length L should be maximized, but constrained such that code repetitions R surpass the rank of dynamic data in the sample FOV to be analyzed. For example, if 1000 frames are to be recorded, and the rank is expected to be 10, the code length should not exceed 50 (100 frames are required to cycle through the code and its interleaved complement).

The reconstruction algorithm for optical section movies has only one adjustable parameter: the number of SVD components should be set to exceed the rank of the underlying sample, and it does not need to be higher than the number of code repetitions. An upper bound to the rank of the sample can be estimated by calculating how many neurons would contribute distinguishable activity to a widefield movie.

Compressed sensing theory has been used in other microscopy applications [11–13, 20–22]. These other techniques require explicit regularization to constrain solutions with sparsity properties, together with computationally intensive, parameter-dependent calculations to minimize an objective function. In this work, measurements are compressed but the requirement for low-rank sample representation is implicit, with optical section estimations that result from an analytic and sample-independent direct calculation. There are no adjustable parameters to enforce spatial or temporal sparsity, iterative optimization, or convergence criteria.

5. Conclusion

This work presents an extension to existing computational optical sectioning methods, to record neuronal activity at high temporal resolution by leveraging the natural redundant structure of these biological signals. Future work could extend the algorithm to accommodate moving or shape-changing samples, that is, samples with low dynamic rank under a more complex signal generation model.

Acknowledgments

This work was supported by the Harvard Data Science Initiative, the Harvard Brain Initiative, and the Howard Hughes Medical Institute. V J P was supported by a Becas Chile scholarship. C S L was supported by a HBI Young Scientist grant awarded to Vicente Parot, and was partially supported by Fondecyt grant #11160728.

Appendix A. The relation between confocal and Hadamard data

This section explores the optical sectioning property of Hadamard demodulation, as applied to static samples, and considering a practical calibration method for illumination patterns. The system is modeled as a continuous intensity-linear space-invariant optical system with patterned illumination. The illumination is further assumed to consist of spatially-uncorrelated illumination functions of time. Demodulation using matched filtering results in a confocal optical section.

The model for image formation is

$$u(\boldsymbol{p},t) = \int_{z\geqslant 0} p_C(\boldsymbol{p}-\boldsymbol{q},z) f(\boldsymbol{q},z) \mu(\boldsymbol{q},z,t)\, d\boldsymbol{q}dz$$

where p_C is the collection PSF, f is the sample fluorescence and μ is the illumination onto the sample. Here we have assumed the camera is imaged to the xy-plane at the $z=0$ coordinate, and that the sample is on the positive z-axis. The illumination process is modeled as

$$\mu(\boldsymbol{q},z,t) = \int p_I(\boldsymbol{q}-\boldsymbol{r},z)\pi(\boldsymbol{r},t)\, d\boldsymbol{r},$$

where p_I is the illumination PSF, and π is the illumination pattern. Using this model, the practical calibration data can be ideally represented as

$$c(\boldsymbol{p},t) = \int p_C(\boldsymbol{p}-\boldsymbol{q},z_0)\mu(\boldsymbol{q},z_0,t)\, d\boldsymbol{q},$$

i.e. the calibration screen corresponds to a uniform f fully concentrated on the plane $z=z_0$. The location z_0 represents

deviation between the calibration plane and the focal plane, normally set to $z_0 = 0$. The positive-only illumination pattern is compensated by

$$c^{\delta}(\boldsymbol{p},t) = c(\boldsymbol{p},t) - \int_{t\geqslant 0} c(\boldsymbol{p},t)\,dt,$$

and matched filtering demodulation using the calibration data becomes

$$\int_{t\geqslant 0} c^{\delta}(\boldsymbol{p},t)u(\boldsymbol{p},t)\,dt = \int_{z\geqslant 0} p_C(\boldsymbol{p}-\boldsymbol{q}',z_0)\,p_C(\boldsymbol{p}-\boldsymbol{q},z) f(\boldsymbol{q},z) \times \left(\int_{t\geqslant 0} \mu(\boldsymbol{q}',z_0,t)\mu(\boldsymbol{q},z,t)\,dt\right) d\boldsymbol{q}d\boldsymbol{q}'dz$$

whence

$$\int_{t\geqslant 0} \mu(\boldsymbol{q}',z_0,t)\mu(\boldsymbol{q},z,t)\,dt = \int_{z\geqslant 0} p_I(\boldsymbol{q}'-\boldsymbol{r}',z_0)\,p_I(\boldsymbol{q}-\boldsymbol{r},z)\left(\int_{t\geqslant 0}\pi(\boldsymbol{r}',t)\pi(\boldsymbol{r},t)\,dt\right) d\boldsymbol{r}'d\boldsymbol{r}.$$

We assume a full orthonormal set of illumination patterns such that

$$\int_{t\geqslant 0}\pi(\boldsymbol{r}',t)\pi(\boldsymbol{r},t)\,dt = \delta(\boldsymbol{r}-\boldsymbol{r}')$$

and it follows that

$$\int_{t\geqslant 0}\mu(\boldsymbol{q}',z_0,t)\mu(\boldsymbol{q},z,t)\,dt = \int p_I(\boldsymbol{q}'-\boldsymbol{r}',z_0)\,p_I(\boldsymbol{q}-\boldsymbol{r}',z)d\boldsymbol{r}' = \int p_I(\boldsymbol{r}',z_0)\,p_I(\boldsymbol{q}-\boldsymbol{q}'-\boldsymbol{r}',z)d\boldsymbol{r}'$$

where we used the change of variables $\boldsymbol{r}' \to \boldsymbol{r}'+\boldsymbol{q}'$ to obtain the last equality. Since the change of variables $\boldsymbol{q}' \to \boldsymbol{q}'+\boldsymbol{q}$ yields

$$\int p_C(\boldsymbol{p}-\boldsymbol{q}',z_0)\,p_C(\boldsymbol{p}-\boldsymbol{q},z)\,p_I(\boldsymbol{r}',z_0)\,p_I(\boldsymbol{q}-\boldsymbol{q}'-\boldsymbol{r}',z)\,d\boldsymbol{q}'d\boldsymbol{r}' = \int p_C(\boldsymbol{p}-\boldsymbol{q}-\boldsymbol{q}',z_0)\,p_C(\boldsymbol{p}-\boldsymbol{q},z)\,p_I(\boldsymbol{r}',z_0)\,p_I(\boldsymbol{q}'-\boldsymbol{r}',z)\,d\boldsymbol{q}'d\boldsymbol{r}'$$

we conclude

$$\int_{t\geqslant 0} c^{\delta}(\boldsymbol{p},t)u(\boldsymbol{p},t)\,dt = \int_{z\geqslant 0} p_H(\boldsymbol{p}-\boldsymbol{q},z)f(\boldsymbol{q},z)d\boldsymbol{q}dz$$

with

$$p_H(\boldsymbol{p},z) = p_C(\boldsymbol{p},z)\int p_C(\boldsymbol{p}-\boldsymbol{q}',z_0)\left(\int p_I(\boldsymbol{r}',z_0)\,p_I(\boldsymbol{q}'-\boldsymbol{r}',z)\,d\boldsymbol{r}'\right)d\boldsymbol{q}'.$$

The optical section PSF given by p_H is the product of the collection PSF and the illumination PSF, modified by the in-plane imaging quality of the calibration pattern used for demodulation.

Appendix B. Estimating the widefield sequence from complementary illumination patterns

This appendix considers the effect of fast fluorescence dynamics, e.g. when there is a transient increase in intensity that persists during one illumination pattern but not its complement. The estimated widefield movie then includes an artifact comprising the change in intensity, multiplied by the illumination pattern used at that time.

Appendix A shows that the estimate of the widefield sequence for a time-dependent fluorescence is given by

$$u(\boldsymbol{p},t+\Delta t)+u(\boldsymbol{p},t) = \int_{z\geqslant 0} p_C(\boldsymbol{p}-\boldsymbol{q},z)(f(\boldsymbol{q},z,t+\Delta t) \times \mu(\boldsymbol{q},z,t+\Delta t)+f(\boldsymbol{q},z,t)\mu(\boldsymbol{q},z,t))\,d\boldsymbol{q}dz.$$

If the illumination at $t+\Delta t$ is complementary to that used at t we have both $\pi(\boldsymbol{r},t+\Delta t)+\pi(\boldsymbol{r},t)=1$ and $\mu(\boldsymbol{q},z,t+\Delta t)+\mu(\boldsymbol{q},z,t)=1$. It follows that

$$f(\boldsymbol{q},z,t+\Delta t)\mu(\boldsymbol{q},z,t+\Delta t)+f(\boldsymbol{q},z,t)\mu(\boldsymbol{q},z,t) = f(\boldsymbol{q},z,t)+\mu(\boldsymbol{q},z,t+\Delta t)(f(\boldsymbol{q},z,t+\Delta t)-f(\boldsymbol{q},z,t)).$$

A straightforward manipulation shows that

$$u(\boldsymbol{p},t+\Delta t)+u(\boldsymbol{p},t) = u_w(\boldsymbol{p},t)+\int_{z\geqslant 0} p_C(\boldsymbol{p}-\boldsymbol{q},z)(f(\boldsymbol{q},z,t+\Delta t) \times -f(\boldsymbol{q},z,t))\mu(\boldsymbol{q},z,t+\Delta t)\,d\boldsymbol{q}dz,$$

where u_w denotes the widefield sequence. Observe the additive component depends *only* on the variations of the fluorescence at times t and $t+\Delta t$.

ORCID iDs

Vicente J Parot https://orcid.org/0000-0002-9261-6667
Urs L Böhm https://orcid.org/0000-0002-7958-4509
Linlin Z Fan https://orcid.org/0000-0001-6921-2659

References

[1] Rodriguez E A, Campbell R E, Lin J Y, Lin M Z, Miyawaki A, Palmer A E, Shu X, Zhang J and Tsien R Y 2017 The growing and glowing toolbox of fluorescent and photoactive proteins *Trends Biochem. Sci.* **42** 111–29
[2] Sample V, Mehta S and Zhang J 2014 Genetically encoded molecular probes to visualize and perturb signaling dynamics in living biological systems *J. Cell Sci.* **127** 1151–60
[3] Neil M A A, Juškaitis R and Wilson T 1997 Method of obtaining optical sectioning by using structured light in a conventional microscope *Opt. Lett.* **22** 1905–7
[4] Saxena M, Eluru G and Gorthi S S 2015 Structured illumination microscopy *Adv. Opt. Photonics* **7** 241–75

[5] Gustafsson M G L 2000 Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy *J. Microsc.* **198** 82–7
[6] Mertz J 2011 Optical sectioning microscopy with planar or structured illumination *Nat. Methods* **8** 811
[7] Lim D, Chu K K and Mertz J 2008 Wide-field fluorescence sectioning with hybrid speckle and uniform-illumination microscopy *Opt. Lett.* **33** 1819–21
[8] Kim D H, Kim J, Marques J C, Grama A, Hildebrand D G C, Gu W, Li J M and Robson D N 2017 Pan-neuronal calcium imaging with cellular resolution in freely swimming zebrafish *Nat. Methods* **14** 1107
[9] Farhi S L *et al* 2018 Wide-area all-optical neurophysiology in acute brain slices (https://doi.org/10.1101/433953)
[10] Mazzaferri J, Kunik D, Belisle J M, Singh K, Lefrançois S and Costantino S 2011 Analyzing speckle contrast for hilo microscopy optimization *Opt. Express* **19** 14508–17
[11] Pnevmatikakis E A *et al* 2016 Simultaneous denoising, deconvolution, and demixing of calcium imaging data *Neuron* **89** 285–99
[12] Nöbauer T, Skocek O, Pernía-Andrade A J, Weilguny L, Traub F M, Molodtsov M I and Vaziri A 2017 Video rate volumetric Ca 2 + imaging across cortex using seeded iterative demixing (SID) microscopy *Nat. Methods* **14** 811
[13] Pégard N C, Liu H-Y, Antipa N, Gerlock M, Adesnik H and Waller L 2016 Compressive light-field microscopy for 3D neural activity recording *Optica* **3** 517–24
[14] Dunn T W, Mu Y, Narayan S, Randlett O, Naumann E A, Yang C-T, Schier A F, Freeman J, Engert F and Ahrens M B 2016 Brain-wide mapping of neural activity controlling zebrafish exploratory locomotion *Elife* **5** e12741
[15] Tufts D W, Kumaresan R and Kirsteins I 1982 Data adaptive signal estimation by singular value decomposition of a data matrix *Proc. IEEE* **70** 684–5
[16] Furnival T, Leary R K and Midgley P A 2017 Denoising time-resolved microscopy image sequences with singular value thresholding *Ultramicroscopy* **178** 112–24
[17] Lou S *et al* 2016 Genetically targeted all-optical electrophysiology with a transgenic Cre-dependent optopatch mouse *J. Neurosci.* **36** 11059–73
[18] Lim D, Ford T N, Chu K K and Metz J 2011 Optically sectioned *in vivo* imaging with speckle illumination hilo microscopy *J. Biomed. Opt.* **16** 016014
[19] Lauterbach M A, Ronzitti E, Sternberg J R, Wyart C and Emiliani V 2015 Fast calcium imaging with optical sectioning via HiLo microscopy *PLoS One* **10** e0143681
[20] Studer V, Bobin J, Chahid M, Mousavi H S, Candès E and Dahan M 2012 Compressive fluorescence microscopy for biological and hyperspectral imaging *Proc. Natl Acad. Sci.* **109** E1679–87
[21] Zhu L, Zhang W, Elnatan D and Huang B 2012 Faster storm using compressed sensing *Nat. Methods* **9** 721
[22] Meiniel W, Spinicelli P, Angelini E D, Fragola A, Loriette V, Orieux F, Sepulveda E and Olivo-Marin J 2017 Reducing data acquisition for fast structured illumination microscopy using compressed sensing *IEEE 14th Int. Symp. on Biomedical Imaging* pp 32–5