Fluorescent wavefront shaping using incoherent iterative phase conjugation Supplementary material

DROR AIZIK[1,*], IOANNIS GKIOULEKAS[2], AND ANAT LEVIN[1]

[1] Department of Electrical and Computer Engineering, Technion, Haifa, Israel
[2] Robotics Institute, Carnegie Mellon University, PA, USA
[*] Corresponding author: droraizik@campus.technion.ac.il

Compiled November 18, 2022

This article provides supplement information to "Fluorescent wavefront shaping using incoherent iterative phase conjugation".

http://dx.doi.org/10.1364/optica.XX.XXXXXX

1. CONVERGENCE ANALYSIS FOR INCOHERENT TRANSMISSION

We analyze the convergence of our iterative algorithm, while modeling the incoherent summation of different fluorescent emitters. We show that incoherence leads to asymptotically faster convergence when compared to the coherent case.

Model. We introduce some notation that we will use to rewrite the transmission image formation model in an equivalent form that is more convenient for our analysis. For this, we recall from the main paper our assumption that the transmission matrices for excitation $\mathcal{T}^i$ and emission $\mathcal{T}^o$ are transposes of one another, $\mathcal{T}^o = {\mathcal{T}^i}^\top$. Therefore, the rows of $\mathcal{T}^i$ equal the columns of $\mathcal{T}^o$. We write $\mathcal{T}^i_{k,:}, \mathcal{T}^o_{:,k}$ for the k-th row and column of these $\mathcal{T}^i$ and $\mathcal{T}^o$, respectively. We also denote by n_k the norm of the k-th row of $\mathcal{T}^i$ and k-th column of $\mathcal{T}^o$,

$$n_k \equiv \sum_x |\mathcal{T}^i_{k,x}|^2 = \sum_x |\mathcal{T}^o_{x,k}|^2, \tag{1}$$

where x is a position on the input or sensor plane, for excitation and emission respectively. n_k can be lower than one, because some light emitted by the fluorescent particles scatters at angles higher than the numerical aperture of the objective and does not reach the sensor. We use $\boldsymbol{T}^i, \boldsymbol{T}^o$ to denote the matrices $\mathcal{T}^i, \mathcal{T}^o$ after normalizing their rows and columns, respectively, to have unit-norm:

$$\boldsymbol{T}^i_{k,:} \equiv \frac{1}{\sqrt{n_k}} \mathcal{T}^i_{k,:}, \quad \boldsymbol{T}^o_{:,k} \equiv \frac{1}{\sqrt{n_k}} \mathcal{T}^o_{:,k}. \tag{2}$$

We will account for the norm explicitly in the image formation model. Then, linear fluorescence emission from the k-th location inside tissue is proportional to

$$|\boldsymbol{\nu}_k|^2 = e_k n_k |\boldsymbol{T}^i_{k,:}\mathbf{u}|^2, \tag{3}$$

where we use e_k to denote the emission power of the k-th particle (for simplicity of exposition, in the main text we have absorbed e_k into the unnormalized transmission matrix). Similarly, the measured emission intensity is

$$\sum_k |\mathcal{T}^o_{:,k}|^2 |\boldsymbol{\nu}_k|^2 = \sum_k |\boldsymbol{T}^o_{:,k}|^2 n_k |\boldsymbol{\nu}_k|^2. \tag{4}$$

With this notation we can express the combined transmission operator as

$$\mathcal{T}^a \equiv \boldsymbol{T}^o \boldsymbol{W} \boldsymbol{T}^i, \tag{5}$$

where $\boldsymbol{W}$ is a diagonal matrix with non-negative diagonal entries

$$w_k \equiv n_k \sqrt{e_k}. \tag{6}$$

These entries encode the power of the fluorescent emitter at the k-th focus location, as well as the amount of energy transferred on the k-th row and column of $\mathcal{T}^i$ and $\mathcal{T}^o$, respectively.

To further simplify our analysis, we also assume that the wavefonts emitted by different fluorescent particles are sufficiently random, and their correlation

$$\varepsilon_{k,\ell} \equiv \sum_x \boldsymbol{T}^o_{x,k} {\boldsymbol{T}^o_{x,\ell}}^* \tag{7}$$

is sufficiently small. For the rest of this derivation we will assume $\varepsilon_{k,\ell} \approx 0$ and can be neglected. We note that the memory effect implies that the rows of the transmission matrix are correlated *shifted* versions of each other; that is $\boldsymbol{T}^o_{x,k} \approx \boldsymbol{T}^o_{x+\boldsymbol{\Delta},\ell}$ where $\boldsymbol{\Delta}$ is the displacment between the k, ℓ particles. However, even

in the presence of ME correlation, at the zero shift we consider in Eq. (7), such rows are uncorrelated, as effectively, the row entries are pseudo-random patterns.

With this decorrelation assumption, $\boldsymbol{T^i}, \boldsymbol{T^o}$ are orthogonal matrices, and the diagonal entries w_k of the matrix $\boldsymbol{W}$ in Eq. (5) will correspond to the eigenvalues of the transmission operator $\boldsymbol{\mathcal{T}^a}$.

Power method under coherent illumination. We start by reviewing the principles of the power method [1] considering a coherent illumination model. For simplicity, we assume that we can measure the phase of wavefronts rather than only their intensities. We will later extend this analysis to the incoherent case.

To apply the power method for the coherent illumination case, we start with a random illumination pattern $\mathbf{u}^0$. At each iteration, we illuminate the tissue sample with a wavefront $\mathbf{u}^t$. The propagation through the sample and back to the sensor produces a wavefront $\boldsymbol{\mathcal{T}^a}\mathbf{u}^t$. Assuming we can measure both amplitude and phase of the resulting wavefront, we update the illumination wavefront as

$$\mathbf{u}^{(t+1)} \equiv \frac{(\boldsymbol{\mathcal{T}^a}\mathbf{u}^{(t)})^*}{\|\boldsymbol{\mathcal{T}^a}\mathbf{u}^{(t)}\|}, \tag{8}$$

where we normalize $\mathbf{u}^{(t+1)}$ to fix the total energy of the excitation pattern $\|\mathbf{u}^{(t+1)}\|$ at each iteration, as determined by the power of the excitation laser. We note that, in practice, we use a phase-only SLM, and thus we only display the phase of $\mathbf{u}^{(t+1)}$, dropping its amplitude.

To understand the convergence of this algorithm, we denote by β_k^t the energy scattered from the k-th particle inside the tissue sample at the t-th iteration,

$$\beta_k^{(t)} \equiv w_k \boldsymbol{T^i_{k,:}}\mathbf{u}^{(t)}. \tag{9}$$

With this notation, the operation of the transmission matrix $\boldsymbol{\mathcal{T}^a}$ on $\mathbf{u}^{(t)}$ equals the sum of the columns of $\boldsymbol{T^o}$ weighted by $\beta_k^{(t)}$,

$$\boldsymbol{\mathcal{T}^a}\mathbf{u}^{(t)} = \sum_k \beta_k^{(t)} \boldsymbol{T^o_{:,k}} \tag{10}$$

Therefore, the illumination pattern used at the next iteration will equal a weighted combination of the conjugate columns,

$$\mathbf{u}^{(t+1)} = \left(\boldsymbol{\mathcal{T}^a}\mathbf{u}^{(t)}\right)^* = \sum_k \beta_k^{(t)} \boldsymbol{T^o_{:,k}}^*, \tag{11}$$

$\mathbf{u}^{(t+1)}$ is then normalized as in Eq. (8).

Applying $\boldsymbol{T^i}$ on $\mathbf{u}^{(t+1)}$ can be expressed as a summation over all entries x. Using the decorrelation assumption of Eq. (7), and Eq. (11) this reduces to

$$\boldsymbol{T^i_{k,:}}\mathbf{u}^{(t+1)} = \sum_x \boldsymbol{T^i_{k,x}}\mathbf{u}_x^{(t+1)} = \sum_\ell \beta_\ell^{(t)} \sum_x \boldsymbol{T^i_{k,x}}\boldsymbol{T^o_{x,l}}^* = \beta_k^{(t)}. \tag{12}$$

Therefore,

$$\beta_k^{(t+1)} = w_k \boldsymbol{T^i_{k,:}}\mathbf{u}^{(t+1)} = w_k \beta_k^{(t)}. \tag{13}$$

A simple recursion implies that $\beta_k^{(t)}$ follows an exponential series of the form

$$\beta_k^{(t)} = (\lambda_k)^t c_k \tag{14}$$

where $\lambda_k \equiv w_k, c_k \equiv \beta_k^0$. Eq. (14) states that the entries of $\beta^{(t)}$ scale exponentially as a function of the iteration number. This implies that the gap between the largest entry of β and the next one increases with each iteration. It is easy to show that the sequence converges quickly into a one-hot vector, which is non-zero at a single entry.

Power method under incoherent illumination. The coherent case is similar to the classical application of the power method. We now make the necessary adaptations for the incoherent case. The convergence rate we achieve is asymptotically faster than the exponential convergence we derived in Eq. (14). Throughout this derivation we assume the power of the fluorescent sources is constant during optimization and ignore effects such as blinking or bleaching.

To study the incoherent emission of fluorescent sources, we start by deriving the corresponding image formation model. At the t-th iteration, we excite the tissue sample with a wavefront $\mathbf{u}$, and measure

$$I^{(t)} = \sum_k |\boldsymbol{T^o_{:,k}}|^2 \alpha_k^{(t)}, \tag{15}$$

where $\alpha_k^{(t)}$ denotes the incoherent equivalent of $\beta_k^{(t)}$, the energy emerging from the k-th emitter, times the norm of the k-th column. Following the definitions in Eqs. (3), (4) and (6), we use:

$$\alpha_k^{(t)} = n_k |\boldsymbol{\nu}_k|^2 = w_k^2 |\boldsymbol{T^i_{k,:}}\mathbf{u}^{(t)}|^2. \tag{16}$$

Our goal is to show that, as in the coherent case, within a small number of iterations, $\alpha_k^{(t)}$ converges to a one-hot vector.

At each iteration, our algorithm needs to estimate some phase from the speckle intensity image $I^{(t)}$. As we mention in the main paper, we use a phase diversity acquisition scheme. As this scheme is based on optimizing a non-linear score, analyzing its convergence is not straightforward. To this end, we start by considering a simpler acquisition scheme based on point diffraction interferometry [2, 3]. This scheme is highly-sensitive to noise, and implementing it using weak fluorescent sources is impractical. However, the advantage of this scheme is in providing a closed-form expression for the contribution of different sources to the estimated phase, allowing for simple analysis. We will later use numerical simulations to compare the convergence of phase diversity against the analytical expressions we derive from point diffraction interferometry.

Point diffraction interferometry. Consider the field $\boldsymbol{T^o_{x,k}}$ generated by the k-th fluorescent source at image point x. We decompose it as

$$\boldsymbol{T^o_{x,k}} = \hat{\boldsymbol{T}}^{\boldsymbol{o}}{}_{x,k} + {\tau^o}_k \tag{17}$$

where ${\tau^o}_k$ is its complex mean

$$ {\tau^o}_k = \sum_x \boldsymbol{T^o_{x,k}} \tag{18}$$

and $\hat{\boldsymbol{T}}^{\boldsymbol{o}}{}_{x,k} = \boldsymbol{T^o_{x,k}} - {\tau^o}_k$.

Point diffraction interferometry captures $J \geq 3$ images using the SLM in the Fourier plane of the imaging arm. It changes phases at a single spot corresponding to the 0-th (central) frequency. Placing phase ϕ_j at the zero frequency only changes the mean of the signal, and the intensity to be measured at pixel x of the image plane from the k'th source corresponds to

$$\begin{aligned} I^{(t,j)}_{k,x} &= \alpha_k^{(t)} \left|\hat{\boldsymbol{T}}^{\boldsymbol{o}}{}_{x,k} + e^{i\phi_j}{\tau^o}_k\right|^2 \\ &= \alpha_k^{(t)} \left(\left|\hat{\boldsymbol{T}}^{\boldsymbol{o}}{}_{x,k}\right|^2 + |{\tau^o}_k|^2 + 2\Re\left(e^{i\phi_j}{\tau^o}_k \hat{\boldsymbol{T}}^{\boldsymbol{o}*}_{x,k}\right)\right), \end{aligned} \tag{19}$$

where $\Re$ denotes the real component, and $\alpha_k^{(t)}$ defined in Eq. (16) corresponds to the energy emitted by the k-th fluorescent particle given the current excitation wavefront. In the presence of

multiple incoherent sources, we measure the incoherent summation of the intensity speckle patterns produced by each of them,

$$I_x^{(t,j)} = \sum_k \alpha_k^{(t)} I_{k,x}^{(t,j)} \tag{20}$$

In point diffraction interferometry, we capture $J \geq 3$ images with equally-spaced phase shifts $\phi_j = [1 \dots J]\frac{2\pi}{J}$. Standard phase shifting interferometry techniques [4] imply that, by summing measurements with different phase shifts, we can extract

$$\mathbf{u}_x^{(t+1)} = \sum_j e^{-i\phi_j} I_x^{(t,j)} = \sum_k \alpha_k^{(t)} {\tau^o}_k \hat{\boldsymbol{T}}^{\boldsymbol{o}*}_{x,k}. \tag{21}$$

Thus, point diffraction interferometry extracts a weighted combination of the wavefronts emerging from all incoherent sources. The weights correspond to the intensity they receive from the previous excitation pattern, weighted by a complex scalar corresponding to the mean ${\tau^o}_k$.

In the next iteration of the power method, we excite the tissue with the extracted wavefront normalized to have unit energy:

$$\frac{\mathbf{u}^{t+1}}{\|\mathbf{u}^{t+1}\|} \tag{22}$$

Convergence. We now show that the sequence α_k^t converges to a one-hot vector, which implies that the iterative approach we described above converges. To this end, we note that when we excite the tissue sample with illumination $\mathbf{u}^{t+1}(x)$, we effectively multiply $\mathbf{u}^{t+1}(x)$ by $\boldsymbol{T^i}$. Using the decorrelation assumption in Eq. (7), we can write the energy at the k-th fluorescent particle as

$$\boldsymbol{T^i_{k,:}}\mathbf{u}_x^{(t+1)} = \sum_\ell \alpha_\ell^{(t)} {\tau^o}_\ell \sum_x \boldsymbol{T^i}_{k,x} \hat{\boldsymbol{T}}^{\boldsymbol{o}*}_{x,\ell} = \alpha_k^{(t)} {\tau^o}_k. \tag{23}$$

Thus, using the definition of α_k in Eq. (16) and ignoring the normalization in Eq. (22), we have

$$\alpha_k^{(t+1)} = (w_k \alpha_k^{(t)} |{\tau^o}_k|)^2. \tag{24}$$

Using recursion, this leads to

$$\begin{aligned} \alpha_k^{(t)} &= (w_k)^{2(2^{t+1}-1)} |{\tau^o}_k|^{2(2^t-1)} |\boldsymbol{T^i_{k,:}}\mathbf{u}^{(0)}|^{2^{t+1}} \\ &= (\lambda_k)^{2^t} \cdot c_k. \end{aligned} \tag{25}$$

where

$$\lambda_k = w_k^4 |{\tau^o}_k|^2 |\boldsymbol{T^i_{k,:}}\mathbf{u}^{(0)}|^2, \quad c_k = w_k^{-2} |{\tau^o}_k|^{-2}. \tag{26}$$

To understand the difference between this result and the coherent case in Eq. (14), we note that in the coherent case the leading term converges as λ_k^t, which is an exponential sequence. In the incoherent case we have another exponent, and the leading term in Eq. (25) is of the form $(\lambda_k)^{2^t}$. This is known as a *doubly exponential series*, which will converge into a one-hot vector much faster than the exponential series of Eq. (14).

Phase diversity acquisition. As we mentioned above, the point diffraction interferometry scheme is useful for analysis, as it leads to closed-form expressions. In practice, this approach is very sensitive to noise, and implementing it with weak fluorescent sources is unrealistic. Instead, our implementation uses a phase diversity acquisition scheme [5]. We place $J = 5$ known modulation patterns H_j on the SLM of the imaging arm, and measure speckle intensity images of the form

$$I^{(t,j)} = \sum_k |h^j \star \boldsymbol{T^o_{:,k}}|^2 \alpha_k^{(t)}, \tag{27}$$

where $\star$ denotes convolution, and h^j is the Fourier transform of the pattern we placed on the SLM. We use gradient descent optimization to find a complex wavefront $\mathbf{u}^{(t+1)}$ minimizing

$$\sum_j \left| I^{(t,j)} - |h^j \star \mathbf{u}^{(t+1)}|^2 \right|^2. \tag{28}$$

This optimization is subject to local minima, and it is hard to give any analytic guarantees about its convergence. Below we conduct numerical simulations comparing its empirical convergence to the analytical predictions from point diffraction interferometry.

Numerical evaluation. We sample transmission matrices $\boldsymbol{T^o}, \boldsymbol{T^i}$ such that each row is a random i.i.d. complex Gaussian random vector. For simplicity all rows have the same mean ${\tau^o}_k$. We transform the noise vectors to the Fourier domain and set to zero any frequency above $NA = 0.5$. In the primal domain, we limit the speckles in a Gaussian window of STD $20\mu m$, as the speckles imaged in our setup have a limited support and do not spread over the full sensor. We assume $K = 20$ fluorescent sources. We initialize with a uniform excitation and apply the power method as we described above. We simulate phase acquisition with ideal noise-free point diffraction interferometry, and also by solving the phase diversity optimization of Eq. (28), which may converge to local optima. In Fig. 1(a-b) we plot the vectors $\alpha^{(t)}$ we obtain in the first three iterations of the algorithm. For ease of visualization, we sort the entries of this vector by decreasing order of w_k, so that the maximal eigenvalue is always at $k = 1$. We also normalize the plotted vectors to sum to 1. With both acquisition schemes, within a small number of iterations $\alpha^{(t)}$ is a one-hot vector. In each case we plot a dashed line with the expected values following the model of Eq. (25). Even the point diffraction interferometry values do not match this model precisely, because the transmission matrices we sample have random rows which have low correlation, yet their correlation is not precisely zero as assumed in Eq. (7). Despite this difference, the convergence rate of both schemes qualitatively agrees with the model predictions.

To statistically assess the differences between the model of Eq. (25) and the empirical phase retrieval results, we sample 50 random transmission matrices and apply on each the first iteration of the power method. For this, we use an initial excitation such that $|\boldsymbol{T^i_{k,:}}\mathbf{u}^{(0)}|$ is uniform.We measure intensities using the point diffraction interferometry or phase diversity schemes, recover the phase of $\mathbf{u}^{(1)}$, and compute the vector

$$\alpha_k^{(1)} = w_k^2 |\boldsymbol{T^i_{k,:}}\mathbf{u}^{(1)}|^2. \tag{29}$$

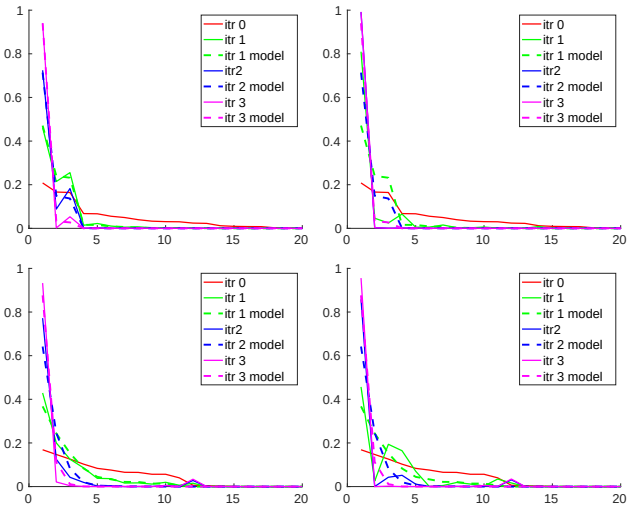
To assess the sparsity of this vector, we measure

$$s = \max_k \frac{\alpha_k^{(1)}}{\sum_k \alpha_k^{(1)}}. \tag{30}$$

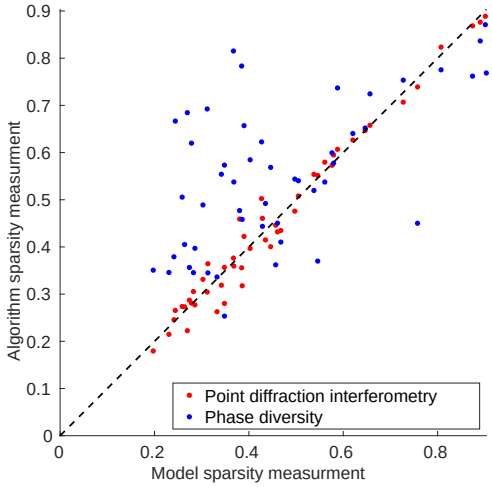
Ideally we want s to be as close to 1 as possible. According to the model in Eq. (25) the sparsity of the eigenvalues after one iteration should be equivalent to

$$s_o = max_k \frac{w_k^6}{\sum_k w_k^6}. \tag{31}$$

In Fig. 1(c), we evaluate s_{pdi} and s_{pd} using the point diffraction interferometry and phase diversity schemes for 50 different transmission matrices. For the k-th random transmission



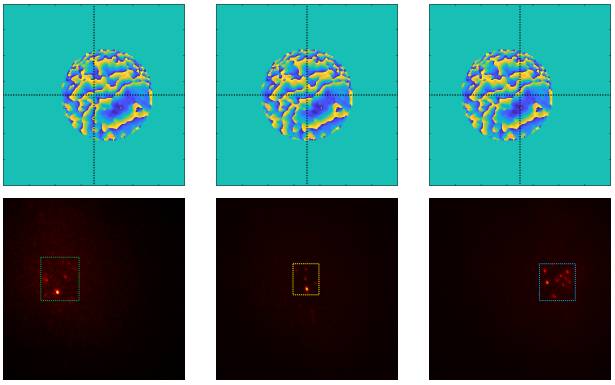
(a) Point diffraction interf. (b) Phase diversity



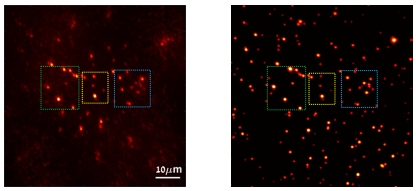
(c) Sparsity statistics

Fig. 1. Numerical convergence evaluation. (a-b) two typical examples for the $\alpha^{(t)}$ values corresponding to the energy in different fluorescent sources for the first three iterations of the algorithm. We compare the convergence with an idealized noise free point diffraction interferometry scheme and the phase diversity optimization we use in practice. In each case we plot in dashed lines the prediction of the model in Eq. (25), with qualitative match to what we measure in practice. The graphs are plotted as a function of the bead index k, where for ease of visualization we sort the eigenvalues in decreasing order of the bead strength w_k so that the strongest eigenvalue appears in the first place. (c) Comparing the sparsity of $\alpha^{(1)}$ to the model prediction for 50 random transmission matrices. Examples that match the model prediction should lie on the dashed diagonal line and indeed point diffraction interferometry results are concentrated around the diagonal line. Phase diversity optimization is often sparser than the model prediction (above the diagonal line) but can also be of lower quality.

matrix, we plot the 2D points (s_o^k, s_{pdi}^k) and (s_o^k, s_{pd}^k). The plot demonstrates that, in practice, the sparsity of phase diversity is equivalent or even better than point diffraction interferometry. If s_{pdi} and s_{pd} matched the s_o prediction, all points should lie on the diagonal dashed line marked in the figure. We see that, for point diffraction interferometry, s_{pdi} is proportional to s_o, but is not exactly equivalent to it, as the decorrelation assumption of Eq. (7) does not hold exactly. We obtained the result of phase diversity acquisition using gradient descent optimization,



Three shifts of correction pattern and the resulting images.



Combined reconstruction Reference

Fig. 2. Using the tilt-shift memory effect to see a wide area behind the tissue. The top row demonstrates three different shifts of the recovered correction pattern. The second row demonstrates the image we capture by placing this shifted mask on the SLM of the imaging arm. Each shift allows us to see a different sub-region of fluorescent sources. By merging 21×21 such shifts we get the wider image in the lowest row. Compare this reconstruction against the reference from the validation camera.

which does not always converge to a global optimum. The plot in Fig. 1(c) illustrates that, in most cases, this solution is actually better than the s_o prediction (points above the dashed line), though for some transmission matrices the solution is worse, and the points lie below the dashed lines.

2. TILT-SHIFT CORRECTION

Below we explain the acquisition of the last row of Fig. 5 in the main paper. Given a wavefront shaping modulation that applies to one fluorescent particle inside the tissue sample, we correct nearby ones using the tilt-shift memory effect. For that, we denote by $u_x^{\mathbf{o}_1}, u_x^{\mathbf{o}_2}$ two speckle fields obtained on the sensor plane of our main camera (where x denotes spatial position on this plane), generated by fluorescent particles at $\mathbf{o}_1, \mathbf{o}_2$. We focus the objective such that the sensor plane is conjugate to the plane containing the fluorescent sources. The tilt-shift memory effect [6, 7] implies that, for small displacements, $u^{\mathbf{o}_1}$ is correlated with a tilted and shifted version of $u^{\mathbf{o}_2}$:

$$u_x^{\mathbf{o}_1} \sim u_{x+\mathbf{\Delta}}^{\mathbf{o}_2} e^{ik\alpha<\mathbf{\Delta},x>} \tag{32}$$

with $\mathbf{\Delta} = \mathbf{o}_2 - \mathbf{o}_1$ the displacement between the sources. If there was no tilt, and the speckle at the image plane could be explained by pure shift, placing in the Fourier plane the Fourier transform of $u_x^{\mathbf{o}_1}$ would correct the emission from $\mathbf{o}_1$ and the emission from nearby points $\mathbf{o}_2$. Given the tilt, the Fourier correction for $\mathbf{o}_2$ should be a shifted version of the Fourier correction of $\mathbf{o}_1$. To account for this, we place in the Fourier plane of our imaging arm shifted versions of our recovered mask. Fig. 2 illustrates

that each such shift allows us to see the fluorescent particles in a different local region. By scanning multiple shifts of the modulation mask, we construct a wider image of the fluorescent particles inside the tissue sample, as shown in the last row of Fig. 2 and in Fig. 5 of the main paper.

3. ADDITIONAL RESULTS

Most experiments in this paper used chicken breast tissue, whose optical properties have been characterized by [8], reporting an anisotropy parameter $g = 0.965$ and a mean free path (MFP) around $43.7\mu m$. In practice, these numbers may vary significantly between different tissue slices. We also demonstrate results on two other materials whose optical properties are better characterized. First, as in [6] we used $10\mu m$ polystyrene micro-spheres dispersed in agarose. Using Mie theory we compute the anisotropy parameter of this dispersion to $g = 0.98$. Sample thickness was $500\mu m$ and we measured its optical depth as $OD = 5.9$. Results on this sample are demonstrated in Fig. 3. In addition, we used parafilm, whose optical properties were characterized by [9]. This has an anisotropy $g = 0.77$ and a MFP around $170\mu m$, where each layer is $120\mu m$ thick. In Fig. 4 we demonstrate focusing through one and two parafilm layers. While the OD here is not high, the parafilm has a much wider scattering angle and the speckle spread on the sensor is very wide. As the fluorescent emission is weak in power, for the two layer example the speckle images we measure involved a lot of shot noise and the algorithm convergence was not very stable.

4. CALIBRATION AND ALIGNMENT

Below we elaborate on various calibration and alignment details.

First, to correctly modulate the Fourier transform of the wave, the illumination SLM needs to be at the focal plane of the lens right after it ($L2$ in the system figure), and the imaging SLM at the focal plane of the lens before it ($L5$). We do this alignment using another camera focused at infinity. We use this camera to view the SLM through the relevant lens, forming a relay system. We adjust the distance between the SLM to the lens until the calibration camera can see a sharp image of the SLM plane. We also ensure that the distance between the sensor of the main/validation cameras and the lenses $L6/L7$ attached to them is set such that the cameras focuse at infinity.

A second step of the alignment is to focus the excitation laser and the system camera on the same target plane. In our setup the sample and the objective of the validation camera are mounted on two motorized z-axis (axial) translation stages. We use fluorescent beads with no tissue and adjust the axial distance between the sample and the objective of the main camera (Obj1 in the setup figure) such that the main camera sees a sharply focused image of the bead. Then we adjust the distance of the validation objective (Obj2 in the setup figure) from the beads so that we see a sharp image of the same beads in the validation camera. We then want the laser to generate its sharpest spot on the same plane. Assuming the validation and main camera are focused at the same plane, we adjust the position of the lens $L3$ until the validation camera sees a sharp laser spot.

After the system has been aligned we need to determine two mappings. The first one is between frequencies to pixels on the SLM. A second, more challenging one is the registration between the two SLMs, so that we can map a pixel on the imaging SLM to a pixel on the illumination SLM controlling the same frequency. We start with a mapping between frequencies to the SLM on the imaging arm. We first put a calibration camera that can image the camera SLM plane directly when it receives light from fluorescent beads. This allows us to see an illuminated circle on the SLM plane, corresponding to the numerical aperture of the imaging system. The center of this circle gives us a first estimate of the zero (central) frequency of the Fourier transform. Assuming we know the focal length of $L5$, the SLM pitch and the wavelength of the emitted light, we can map frequencies to SLM pixels using simple geometry. Alternatively we can display on the SLM sinusoidals of various frequencies. This shifts the image on the sensor plane. By measuring the shift resulting from each sinusoidal we can calibrate the mapping between frequencies to SLM pixels. To align between the two SLMs we find a region behind the tissue where a single isolated bead is excited so that optimizing the phase diversity cost provides the correct modulation pattern with a single power iteration. We need to determine how to position this modulation on the SLMs keeping in mind that tilt and shift on these planes may impact the results. For the imaging SLM this is less of an issue because we have already marked the zero frequency and because a tilt of the imaging SLM only shifts the position of the spot on the sensor. However, if the illumination SLM is not registered correctly we may see a sharp spot behind the tissue but it will be shifted from the bead of interest and will not excite it. Thus, we tilt and shift the modulation on the illumination SLM until we see the bead is excited in the validation camera, or alternatively, until the intensity we measure on the main camera (when the modulation correction is on) is maximized. After this is achieved we can fine tune the shift on the imaging SLM, which is equivalent to the position of the zero (central) frequency that we have previously marked by looking at the illuminated circle.

Once the system has been calibrated and aligned the algorithm can proceed as described above. For the phase diversity we use $j = 5$ random, phase-only, modulation masks H^j. We start by sampling values for each SLM pixel independently, but we then low pass the masks H^j to limit the spread of the convolution kernels h^j in the image domain, so that the limited fluorescent energy is not split between too many sensor pixels. We chose the support of h^j to be about the same as the spread of the speckle pattern we observe in an unmodulated image.

To use a recovered modulation pattern to image a larger region of beads we use the tilt shift memory effect. To apply the scan we need to recover the parameter α of Eq. (32), determining the ratio between the tilt and shift. For that, after we recover the modulation pattern we place it on the illumination SLM and use the validation camera to view the focused spot. We then adjust the ratio between tilt and shift of the modulation pattern so that we can move the focused spot in the validation camera, while preserving maximal intensity.

REFERENCES

1. L. N. Trefethen and D. Bau, *Numerical Linear Algebra* (SIAM, 1997).
2. R. Smartt and W. Steel, "Theory and application of point-diffraction interferometers," Jpn. J. Appl. Phys. **14**, 351 (1975).
3. V. Akondi, A. Jewel, and B. Vohnsen, "Digital phase-shifting point diffraction interferometer," Opt. Lett. **39**, 1641–1644 (2014).
4. P. Hariharan, B. Oreb, and T. Eiju, "Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm," Appl. Opt. **26**, 2504–2506 (1987).

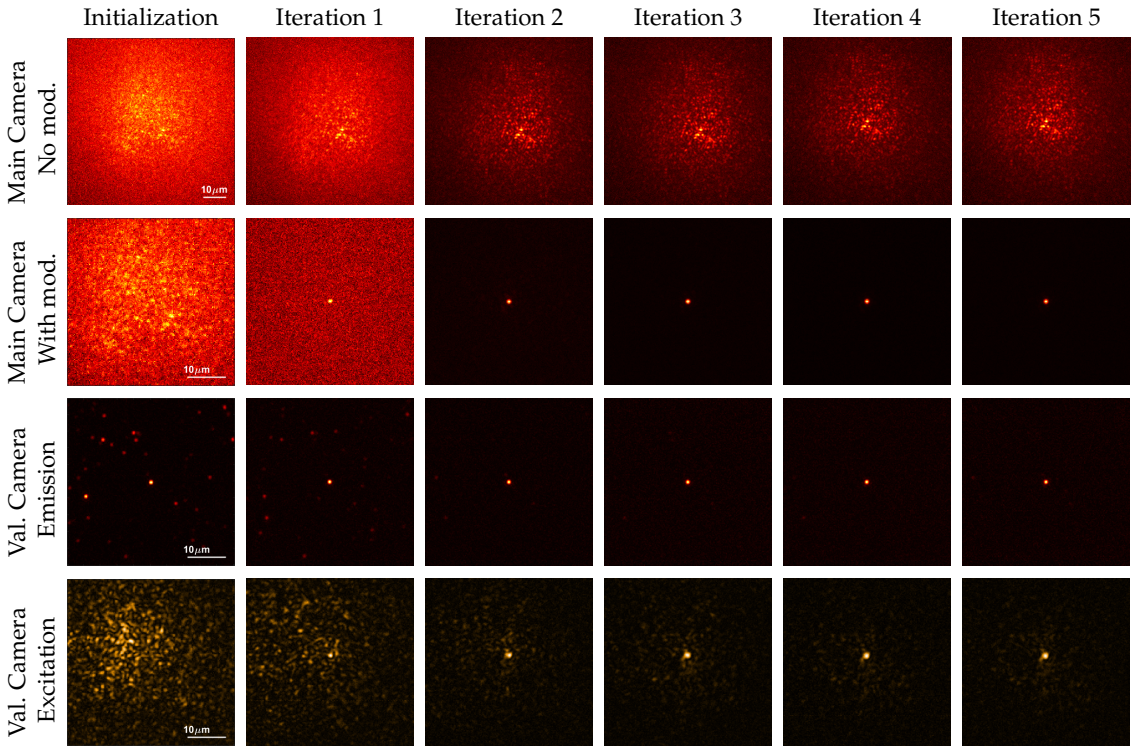


Fig. 3. Algorithm convergence on an agarose scattering phantom of $OD = 5.9$. We demonstrate views via the main camera seeing the front of the tissue with and without the modulation correction, and the validation camera observing fluorescent beads directly. To better appreciate the focusing we used the validation camera to capture both the excitation and emission wavelengths. In the first iteration we see a speckle image, but as power iterations proceed the illumination wavefront converges and focuses on a single bead. When the same modulation pattern is placed at the imaging arm, imaging aberrations are corrected and one can see a sharp image of the excited bead. Note that images in different iterations have very different ranges, and for better visualization each image was normalized to its own maximum. Note also the different scale bar in different rows, some rows zoom only on the center of the speckle pattern.

5. L. M. Mugnier, A. Blanc, and J. Idier, "Phase diversity: A technique for wave-front sensing and for diffraction-limited imaging," (Elsevier, 2006), pp. 1–76.
6. G. Osnabrugge, R. Horstmeyer, I. N. Papadopoulos, B. Judkewitz, and I. M. Vellekoop, "Generalized optical memory effect," Optica. (2017).
7. C. Bar, M. Alterman, I. Gkioulekas, and A. Levin, "Single scattering modeling of speckle correlation," in *ICCP,* (2021).
8. S. Schott, J. Bertolotti, J.-F. Léger, L. Bourdieu, and S. Gigan, "Characterization of the angular memory effect of scattered light in biological tissues," Opt. Express (2015).
9. A. Boniface, B. Blochet, J. Dong, and S. Gigan, "Noninvasive light focusing in scattering media using speckle variance optimization," Optica (2019).

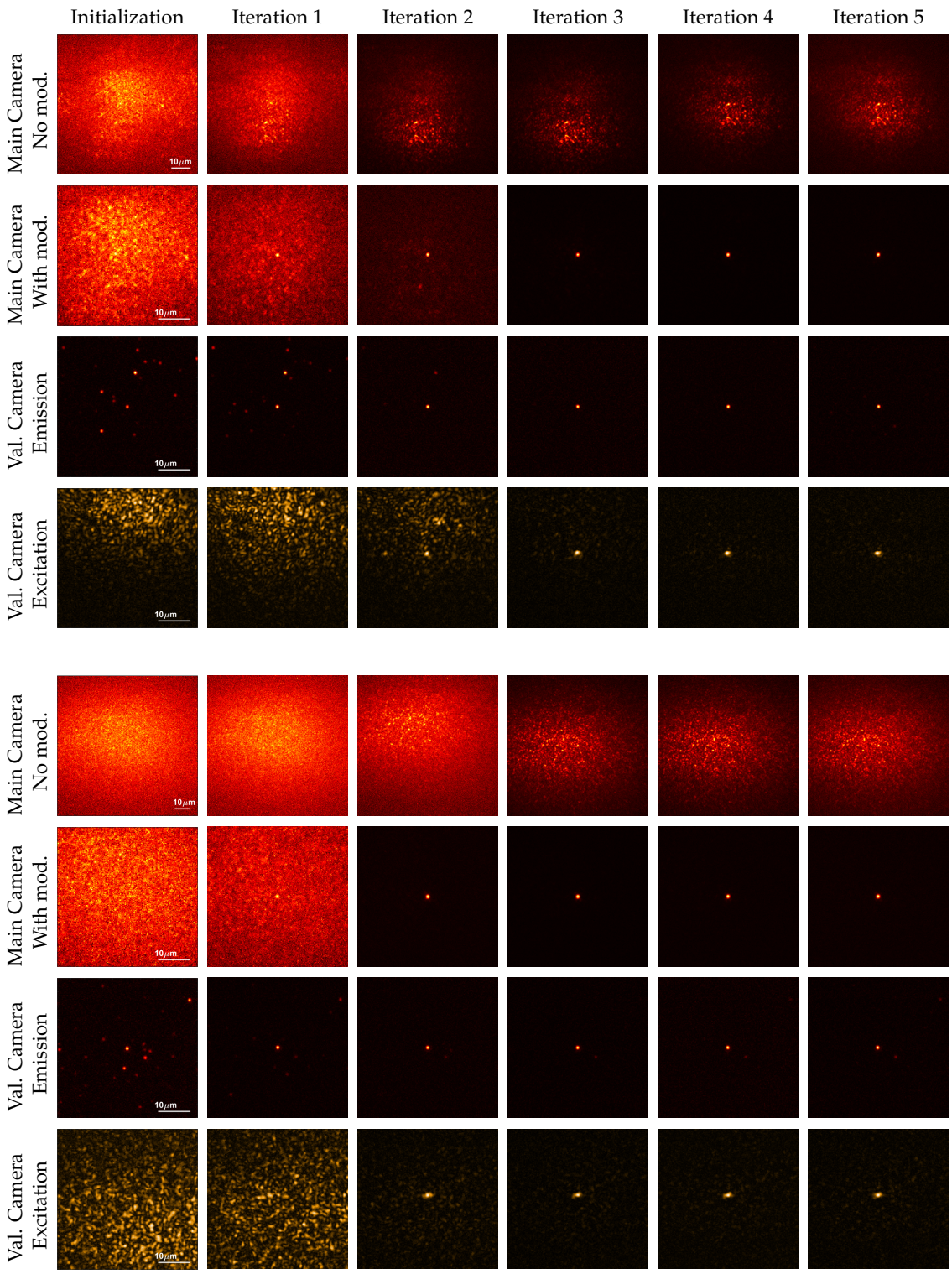


Fig. 4. Algorithm convergence on a parafilm scattering phantom. The top example uses a single parafilm layer and the second one images through two such layers. We demonstrate views via the main camera seeing the front of the tissue with and without the modulation correction. We also demonstrate the view from the validation camera observing fluorescent beads directly, with and without a bandpass filter. Note the different scale bar in different rows, some rows zoom only on the center of the speckle pattern.