A APPENDIX

A.1 Illumination and viewing angles

We provide a justification for Eq. (11), stating that field correlation is a function of only the angle between the illumination and viewing points, rather than their actual distance.

We start by more precisely defining ME field correlation. For illumination and viewing displacements $\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}$ we want to measure correlation as

$$C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}) = \sum_{\mathbf{i}_{x,y}} \sum_{\mathbf{v}_{x,y}} C_f\left(u^{\mathbf{i}_{x,y}-\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} - \frac{\Delta_{\boldsymbol{v}}}{2}\right), u^{\mathbf{i}_{x,y}+\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} + \frac{\Delta_{\boldsymbol{v}}}{2}\right)\right). \tag{27}$$

However as we deal with complex numbers, even if the correlation $|C_f\left(u^{\mathbf{i}_{x,y}-1/2\Delta_{\boldsymbol{i}}}\left(\mathbf{v}_{x,y} - 1/2\Delta_{\boldsymbol{v}}\right) u^{\mathbf{i}_{x,y}+1/2\Delta_{\boldsymbol{i}}}\left(\mathbf{v}_{x,y} + 1/2\Delta_{\boldsymbol{v}}\right)\right)|$ is large, it can be a complex number and the phase of these correlations can vary for different pairs. Thus we should sum correlations subject to proper phase correction:

$$C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}) = \sum_{\mathbf{i}_{x,y}} \sum_{\mathbf{v}_{x,y}} e^{ik\phi(\mathbf{i}_{x,y}, \mathbf{v}_{x,y})} C_f\left(u^{\mathbf{i}_{x,y}-\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} - \frac{\Delta_{\boldsymbol{v}}}{2}\right), u^{\mathbf{i}_{x,y}+\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} + \frac{\Delta_{\boldsymbol{v}}}{2}\right)\right). \tag{28}$$

The ideal phase correction is the conjugate of the correlation phase

$$\phi(\mathbf{i}_{x,y}, \mathbf{v}_{x,y}) = -\angle C_f\left(u^{\mathbf{i}_{x,y}-\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} - \frac{\Delta_{\boldsymbol{v}}}{2}\right), u^{\mathbf{i}_{x,y}+\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} + \frac{\Delta_{\boldsymbol{v}}}{2}\right)\right), \tag{29}$$

leading to

$$C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}) = \sum_{\mathbf{i}_{x,y}} \sum_{\mathbf{v}_{x,y}} \left| C_f\left(u^{\mathbf{i}_{x,y}-\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} - \frac{\Delta_{\boldsymbol{v}}}{2}\right), u^{\mathbf{i}_{x,y}+\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} + \frac{\Delta_{\boldsymbol{v}}}{2}\right)\right)\right|. \tag{30}$$

In practice, we follow Osnabrugge et al. [2017] and restrict the discussion to sinusoidal phase correction terms. This restriction is convenient, first because it simplifies the analysis, allowing for analytic results; and second, because it is easy to implement in experimental setups, being equivalent to a tilt of an incoming or outgoing beam. This leads to the definition of the tilt-shift correlation:

$$C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}, \boldsymbol{\theta}) = \sum_{\mathbf{i}_{x,y}} \sum_{\mathbf{v}_{x,y}} e^{ik(\boldsymbol{\theta}\cdot(\mathbf{v}_{x,y}-\mathbf{i}_{x,y}))} C_f\left(u^{\mathbf{i}_{x,y}-\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} - \frac{\Delta_{\boldsymbol{v}}}{2}\right), u^{\mathbf{i}_{x,y}+\frac{\Delta_{\boldsymbol{i}}}{2}}\left(\mathbf{v}_{x,y} + \frac{\Delta_{\boldsymbol{v}}}{2}\right)\right). \tag{31}$$

Note that $\boldsymbol{\theta}$ and the displacement $\mathbf{v}_{x,y} - \mathbf{i}_{x,y}$ are 2D vectors, and the phase in Eq. (31) is their inner product.

To emphasize that Eq. (31) depends on the distance of the illumination and viewing planes we sometimes denote $C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}, \boldsymbol{\theta}|z_i, z_v)$

Tilt-shift correlation conversions. We show that we can easily convert tilt-shift correlation at one plane to a tilt-shift correlation at a different plane, through a change in the shift.

Claim 3.

$$C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}, \boldsymbol{\theta}|z_i^1, z_v^1) = C_f(\Delta_{\boldsymbol{i}} + (z_i^1 - z_i^2)\boldsymbol{\theta}, \Delta_{\boldsymbol{v}} - (z_v^1 - z_v^2)\boldsymbol{\theta}, \boldsymbol{\theta}|z_i^2, z_v^2). \tag{32}$$

Proof. We prove the case of changing viewing planes z_v^1, z_v^2, the case of changing illumination planes is similar.

We denote by $u_{\Delta,\boldsymbol{\theta},z_v}$ a field focused at z_v after some tilt and shift:

$$u_{\Delta,\boldsymbol{\theta},z_v}(\boldsymbol{\tau}) = u_{z_v}(\boldsymbol{\tau} - \Delta)e^{ik(\boldsymbol{\theta}\cdot(\boldsymbol{\tau}-\Delta))}. \tag{33}$$

We consider the correlation $C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}, \boldsymbol{\theta})$ and define the component of it involving a fixed illumination point $\mathbf{i}$

$$C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}, \boldsymbol{\theta}|\mathbf{i}, z_v) = E\left[\sum_{\boldsymbol{\tau}} u^{\mathbf{i}_{x,y}-\frac{\Delta_{\boldsymbol{i}}}{2}}_{-1/2\Delta_{\boldsymbol{v}},-1/2\boldsymbol{\theta},z_v}(\boldsymbol{\tau}) \cdot u^{\mathbf{i}_{x,y}+\frac{\Delta_{\boldsymbol{i}}}{2}}_{1/2\Delta_{\boldsymbol{v}},1/2\boldsymbol{\theta},z_v}(\boldsymbol{\tau})^*\right]. \tag{34}$$

Clearly $C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}, \boldsymbol{\theta}|z_v)$ as defined in Eq. (31) is

$$C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}, \boldsymbol{\theta}|z_v) = \sum_{\mathbf{i}} C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}, \boldsymbol{\theta}|\mathbf{i}, z_v), \tag{35}$$

thus if we prove the relation outlined in the claim for each $\mathbf{i}$, it will also hold for the full correlation.

We note that $C_f(\Delta_{\boldsymbol{i}}, \Delta_{\boldsymbol{v}}, \boldsymbol{\theta}|\mathbf{i}, z_v)$ in Eq. (34) is the expected inner product of two fields shifted and tilted at the opposite directions. We seek a relation between the inner product of the fields at planes z_v^1, z_v^2. For that we recall, from Parseval's theorem, that the inner products in the prime and Fourier domains are equivalent:

$$\sum_{\boldsymbol{\tau}} u^1_{-1/2\Delta_{\boldsymbol{v}},-1/2\boldsymbol{\theta},z_v^1}(\boldsymbol{\tau}) \cdot u^2_{1/2\Delta_{\boldsymbol{v}},1/2\boldsymbol{\theta},z_v^1}(\boldsymbol{\tau})^* = \tag{36}$$

$$\sum_{\boldsymbol{\omega}} \mathcal{F}\left(u^1_{-1/2\Delta_{\boldsymbol{v}},-1/2\boldsymbol{\theta},z_v^1}\right)(\boldsymbol{\omega}) \cdot \mathcal{F}\left(u^2_{-1/2\Delta_{\boldsymbol{v}},1/2\boldsymbol{\theta},z_v^1}\right)(\boldsymbol{\omega})^*, \tag{37}$$

where we use u^1, u^2 as shorthand for $u^{\mathbf{i}_{x,y}-1/2\Delta_{\boldsymbol{i}}}, u^{\mathbf{i}_{x,y}+1/2\Delta_{\boldsymbol{i}}}$, and $\mathcal{F}$ denotes the Fourier transform.

We show below that the transformation between fields at plane z_v^1 to fields at plane z_v^2 is equivalent to multiplying the Fourier transform with a filter that has varying phase but fixed amplitude. Thus, such a filter does not change the inner product, allowing us to derive a direct relation between inner products at different planes.

To transform a field at plane z_v^1 to a field at plane z_v^2, we need to convolve $u_{\Delta,\boldsymbol{\theta},z_v^1}$ with a kernel we denote by $P(z_v^1 - z_v^2)$. The Fourier transform of the defocus blur is simply a quadratic phase,

$$\mathcal{F}(P(z_v^1 - z_v^2)) = e^{ik\frac{z_v^1 - z_v^2}{2}|\boldsymbol{\omega}|^2}. \tag{38}$$

As this kernel has a uniform amplitude and only varies in phase, it cancels out in the inner product defined above. However, to exactly exactly the transformation between the two planes, we need to look more carefully at the coordinate change introduced by the tilt and shift.

We want to express $u_{\Delta,\boldsymbol{\theta},z_v^1} * P(z_v^1 - z_v^2)$ as the field at plane z_v^2, with some tilt and shift. We will show that

$$u_{\Delta,\boldsymbol{\theta},z_v^1} * P(z_v^1, z_v^2) = u_{\Delta-\boldsymbol{\theta}(z_v^1-z_v^2),\boldsymbol{\theta},z_v^2} \cdot e^{ikc}, \tag{39}$$

with $c = (z_v^1 - z_v^2)/2 \cdot \boldsymbol{\theta}^2$. Namely, when refocusing at a different plane, the same tilt $\boldsymbol{\theta}$ is maintained, but the shift changes according to the tilt direction to $\Delta - \boldsymbol{\theta}(z_v^1 - z_v^2)$.

For that, let us denote the Fourier transform of the field by $U_{z_v^1} = \mathcal{F}(u_{z_v^1})$. We recall that a shift in prime space becomes a tilt in Fourier space, and conversely a tilt in prime space becomes a shift in Fourier space. Thus the Fourier transform of the tilted-shifted field u_{Δ,θ,z_v} is related to U_{z_v} as

$$\mathcal{F}(u_{\Delta,\theta,z_v}) = U_{z_v}(\omega - \theta)e^{ik(\Delta\cdot(\omega-\theta))}. \tag{40}$$

As mentioned in Eq. (38), the Fourier transform of the defocus blur is simply a quadratic phase. A short calculation shows that

$$\mathcal{F}(u_{\Delta,\theta,z_v^1} * P(z_v^1 - z_v^2)) = \tag{41}$$

$$U_{z_v^1}(\omega - \theta)e^{ik(\Delta\cdot(\omega-\theta))} \cdot e^{ik\frac{z_v^1-z_v^2}{2}\omega^2} = \tag{42}$$

$$U_{z_v^1}(\omega - \theta)e^{ik((\Delta-(z_v^1-z_v^2)\theta)\cdot(\omega-\theta))} \cdot e^{ik\frac{z_v^1-z_v^2}{2}(\omega-\theta)^2}e^{ik\frac{z_v^1-z_v^2}{2}\theta^2}, \tag{43}$$

where we get from Eq. (42) to Eq. (43) by rearranging the quadratic terms in the exponent. We note that Eq. (43) is essentially the Fourier transform of a field $u_{z_v^2,\Delta-(z_v^1-z_v^2)\theta,\theta} \cdot e^{ikc}$, focused at z_v^2, with the same tilt θ, but with a different shift $\Delta - (z_v^1 - z_v^2)\theta$.

Having derived a simple expression for the refocused field, we can return to the correlation. As stated in Eqs. (36) and (37), the correlation is just the expected inner product between the fields from different illuminators. Using Parseval's theorem, the inner products in the prime and Fourier domains are equivalent:

$$\sum_{\tau} u^1_{-1/2\Delta_v,-1/2\theta,z_v^1}(\tau) \cdot u^2_{1/2\Delta_v,1/2\theta,z_v^1}(\tau)^* = \tag{44}$$

$$\sum_{\omega} \mathcal{F}\left(u^1_{-1/2\Delta_v,-1/2\theta,z_v^1}\right)(\omega) \cdot \mathcal{F}\left(u^2_{-1/2\Delta_v,1/2\theta,z_v^1}\right)(\omega)^*. \tag{45}$$

Using Eq. (42), in the Fourier domain the only effect of changing the focus is a multiplication with a quadratic phase $e^{ik(z_v^1-z_v^2)/2|\omega|^2}$, which is not changing the inner product, thus

$$\sum_{\omega}\mathcal{F}\left(u^1_{-\frac{\Delta_v}{2},-\frac{\theta}{2},z_v^1}\right)(\omega)\cdot\mathcal{F}\left(u^2_{\frac{\Delta_v}{2},\frac{\theta}{2},z_v^1}\right)(\omega)^* =$$

$$\sum_{\omega}\mathcal{F}\left(u^1_{-\frac{\Delta_v}{2},-\frac{\theta}{2},z_v^1}*P(z_v^1-z_v^2)\right)(\omega)\cdot\mathcal{F}\left(u^2_{\frac{\Delta_v}{2},\frac{\theta}{2},z_v^1}*P(z_v^1-z_v^2)\right)(\omega)^*. \tag{46}$$

As a result we can express:

$$E\left[\sum_{\tau} u^1_{-\frac{\Delta_v}{2},-\frac{\theta}{2},z_v^1}(\tau)\cdot u^2_{\frac{\Delta_v}{2},\frac{\theta}{2},z_v^1}(\tau)^*\right] =$$

$$E\left[\sum_{\tau}\left(u^1_{-\frac{\Delta_v}{2},-\frac{\theta}{2},z_v^1}(\tau)*P(z_v^1-z_v^2)\right)\cdot\left(u^2_{\frac{\Delta_v}{2},\frac{\theta}{2},z_v^1}(\tau)*P(z_v^1-z_v^2)\right)^*\right] =$$

$$E\left[\sum_{\tau} u^1_{-\frac{1}{2}(\Delta_v-(z_v^1-z_v^2)\theta)),-\frac{\theta}{2},z_v^2}(\tau)\cdot u^2_{\frac{1}{2}(\Delta_v-(z_v^1-z_v^2)\theta)),\frac{\theta}{2},z_v^2}(\tau)^*\right]. \tag{47}$$

□

Optimal tilt. We have seen that tilt-shift correlation can be easily converted between different illumination and viewing planes. The next question is what values of tilt and shift maximize correlation. For that we argue that these optimal values are obtained when

$$\theta^{\text{opt}}(\Delta_i) = \frac{\Delta_i}{z_i} = \frac{\Delta_v}{z_v}, \tag{48}$$

that is, when the tilt angle θ equals the angle formed between the target, and illumination and viewing points at displacements Δ_i, Δ_v, as illustrated in Fig. 21(a). This also implies that the optimal view-plane displacement is

$$\Delta_v^{\text{opt}}(\Delta_i) = \frac{z_v}{z_i}\Delta_i. \tag{49}$$

This is harder to justify with an exact analytic argument as above, but we can validate it using the simulator of Bar et al. [2019; 2020], as evaluated in Fig. 21.

Alternatively, rather than rely on a numerical simulator, we can derive an approximate closed-form expression for the *single scattering* component of the correlation [Bar et al. 2021], including only photon paths of length one, that is, paths that scatter at a single particle. The contribution from such paths can be computed in closed-form and can help understand the shape of the correlation. For this derivation, we assume $z_i = z_v$, that is, that the camera is focused at the illuminator plane. In this case, we can show that correlation happens only for $\Delta_i = \Delta_v = \Delta$. In particular, Bar et al. [2021] express the correlation as a function of the z plane on which the scattering event happens, deriving the following analytical result.

Claim 4. *We denote $\tau = \mathbf{v}^1 - \mathbf{i}^1 = \mathbf{v}^2 - \mathbf{i}^2$, and let $\phi(z) = \arcsin(|\tau|/(z_i - z))$ be the angle formed between $\mathbf{i}^1, \mathbf{v}^1$ and a scattering point at depth z. The single scattering covariance resulting from scattering points at depth z is*

$$C_f^{\mathbf{i}^1,\mathbf{i}^2}(\mathbf{v}^1,\mathbf{v}^2)(z) = \rho(\phi(z))\sigma_s e^{-\sigma_t L}e^{-ik\left(\tau\cdot\frac{\Delta}{z_i-z}\right)}, \tag{50}$$

where $k = 2\pi/\lambda$ is the wavenumber, σ_t is the extinction coefficient of the medium, and $\rho(\phi)$ its phase function. The total covariance corresponding to an integration over all z positions is

$$C_f^{\mathbf{i}^1,\mathbf{i}^2}(\mathbf{v}^1,\mathbf{v}^2) = \int_{z=z_{min}}^{z_{max}} C_f^{\mathbf{i}^1,\mathbf{i}^2}(\mathbf{v}^1,\mathbf{v}^2)(z). \tag{51}$$

From Eq. (50), we can deduce that, for the part of the correlation resulting from scatterers on one z plane, the tilt angle θ is a linear function of the displacement Δ, leading to a sinusoid of frequency $\theta = \Delta/(z_i - z)$. This θ is actually the angle formed between two illuminators located at depth z_i and displaced by Δ, and a scattering point at plane z of the sample. That is, the ideal tilt angle is just the angle formed by the two illuminators and the target, as shown in Fig. 21(a). The challenge is that, when the sample thickness is not negligible, every z plane inside the sample corresponds to a slightly different θ angle. The full correlation in Eq. (51) is the mean of correlation values from different planes. Although the integral in Eq. (51) cannot be computed in closed-form, we expect it will have the phase of one of the mean angles in the volume. We can show that the resulting phase is approximately:

$$\theta \approx \frac{\Delta}{z_i - z^*}, \tag{52}$$

where z^* is the depth of a plane in the middle of the volume $z^* = 0.5(z_{\min} + z_{\max})$. In Fig. 2, we chose to place $z_{\min} = -L/2, z_{\max} = L/2$, simplifying the formula to $\theta = \Delta/z_i$.

Osnabrugge et al. [2017] also attempt to derive an analytic expression for the tilt shift correlation based on an approximate differential equation. They arrive at a slightly different result, stating that the optimal tilt depth is 2/3 of the way to the end, rather than exactly the mean. Namely, they suggest that the tilt angle should be computed according to the plane $z^* = (1/3)z_{\min} + (2/3)z_{\max}$, rather than according to the plane $z^* = 0.5(z_{\min} + z_{\max})$. The numerical evaluation below suggests that the difference between the $1/2$ and the $2/3$ rules is minor.

In Fig. 21, we numerically evaluate the effect of the tilt angle using the Monte Carlo simulator of Bar et al. [2019; 2020], without relying on the single scattering approximation. We test the correlation produced by tilt angles of the form $\boldsymbol{\theta} = \Delta/(z_i - z^*)$, for different selections of the plane z^*. For both near-field and far-field configurations, the largest correlation is obtained when $\boldsymbol{\theta}$ is selected at $z^* = 0.5(z_{\min} + z_{\max})$. However, the selection $z^* = (1/3)z_{\min} + (2/3)z_{\max}$ provides results of very similar quality.

Optimal $\boldsymbol{\Delta_v}$ at other planes. The above discussion shows that, for $z_i = z_v$, the optimal angle should be selected as $\boldsymbol{\theta} = \boldsymbol{\Delta_i}/z_i = \boldsymbol{\Delta_v}/z_v$. We are still left with the task of showing that this expression holds for other view plane selections. From Claim 3, we know that if we vary the view plane from z_i to any other z_v plane, the displacement should be adjusted as

$$\boldsymbol{\Delta_v} = \boldsymbol{\Delta_i} + \boldsymbol{\theta}(z_v - z_i). \tag{53}$$

If we replace $\boldsymbol{\theta} = \boldsymbol{\Delta_i}/z_i$ and rearrange terms, we get

$$\boldsymbol{\Delta_v} = \frac{z_v}{z_i}\boldsymbol{\Delta_i}, \tag{54}$$

as required.

General phase corrections. To conclude this discussion, we state that the above derivation approximated the field covariance of Eq. (30) using a sinusoidal phase. We used this restricted setting for tractability, as it allowed us to obtain analytical results. In practice, the optimal phase correction is just the phase of the covariance, which is not exactly sinusoidal. If we use this optimal phase correction, the field correlation will be somewhat higher than that produced using the sinusoidal phase (Figs. 4 and 21). In turn, the invariance of ME to illumination depth only holds approximately.

The difference between field and intensity correlations. We note that, though we have shown that the decay of field correlation is invariant to the distance of the illuminator plane, this is not the case for intensity correlation. To see this, we repeat for convenience the definitions of the two types of correlations in Eqs. (5) and (6):

$$\boldsymbol{C_f}(\boldsymbol{\Delta_i}) = \sum_{\mathbf{v}_{x,y}} \left| \boldsymbol{C_f}(u^{\mathrm{i}^1}(\mathbf{v}_{x,y}), u^{\mathrm{i}^2}(\mathbf{v}_{x,y} + \boldsymbol{\Delta_v})) \right|, \tag{55}$$

$$\boldsymbol{C_I}(\boldsymbol{\Delta_i}) = \sum_{\mathbf{v}_{x,y}} \boldsymbol{C_I}(I^{\mathrm{i}^1}(\mathbf{v}_{x,y}), I^{\mathrm{i}^2}(\mathbf{v}_{x,y} + \boldsymbol{\Delta_v})). \tag{56}$$

For every sensor position in the summation, we have

$$\boldsymbol{C_I}(I^{\mathrm{i}^1}(\mathbf{v}_{x,y}), I^{\mathrm{i}^2}(\mathbf{v}_{x,y} + \boldsymbol{\Delta_v})) = \left| \boldsymbol{C_f}(u^{\mathrm{i}^1}(\mathbf{v}_{x,y}), u^{\mathrm{i}^2}(\mathbf{v}_{x,y} + \boldsymbol{\Delta_v})) \right|^2, \tag{57}$$

but $\boldsymbol{C_I}(\hat{\boldsymbol{\Delta}}_{\boldsymbol{i}})$ is not the squared amplitude of $\boldsymbol{C_f}(\hat{\boldsymbol{\Delta}}_{\boldsymbol{i}})$. As a results, results derived for $\boldsymbol{C_f}(\hat{\boldsymbol{\Delta}}_{\boldsymbol{i}})$ cannot be directly applied to $\boldsymbol{C_I}(\hat{\boldsymbol{\Delta}}_{\boldsymbol{i}})$.

A.2 SNR Derivation

Proof of Claim 2. We define SNR as:

$$\mathrm{SNR} = \frac{E\left[\boldsymbol{c}_{\mathrm{emp}}(\mathbf{i}^1_{x,y}, \Delta)\right]^2}{\mathrm{Var}\left[\boldsymbol{c}_{\mathrm{emp}}(\mathbf{i}^1_{x,y}, \Delta)\right]}. \tag{58}$$

As $\boldsymbol{C_I}(\Delta, \boldsymbol{\tau}) = E\left[I(\mathbf{i}^1_{x,y} + \boldsymbol{\tau})I(\mathbf{i}^1_{x,y} + \boldsymbol{\tau} + \Delta)\right]$ (we have already subtracted the mean of I in pre-processing), from the definition of $\boldsymbol{c}_{\mathrm{emp}}$:

$$E\left[\boldsymbol{c}_{\mathrm{emp}}(\mathbf{i}^1_{x,y}, \Delta)\right]^2 = \left| \sum_{\boldsymbol{\tau}} w(\Delta, \boldsymbol{\tau})\boldsymbol{C_I}(\Delta, \boldsymbol{\tau}) \right|^2. \tag{59}$$

As speckle patterns in different pixels are independent random variables, we can compute the variance as:

$$\begin{aligned} &\mathrm{Var}\left[\boldsymbol{c}_{\mathrm{emp}}(\mathbf{i}^1_{x,y}, \Delta)\right] = \\ &= \sum_{\boldsymbol{\tau}} |w(\Delta, \boldsymbol{\tau})|^2 \mathrm{Var}\left[I(\mathbf{i}^1_{x,y} + \boldsymbol{\tau})I(\mathbf{i}^1_{x,y} + \boldsymbol{\tau} + \Delta)\right] &(60)\\ &= \sum_{\boldsymbol{\tau}} |w(\Delta, \boldsymbol{\tau})|^2 \mathrm{Var}\left[I(\mathbf{i}^1_{x,y} + \boldsymbol{\tau})\right] \mathrm{Var}\left[I(\mathbf{i}^1_{x,y} + \boldsymbol{\tau} + \Delta)\right]. &(61) \end{aligned}$$

We assume that, after adding contributions from all illuminators, the speckle spread on the sensor and thus the variance are both roughly uniform, i.e., $\mathrm{Var}\left[I(\mathbf{i}^1_{x,y} + \boldsymbol{\tau})\right]$ is constant for all pixels. Then,

$$\mathrm{Var}\left[I(\hat{\mathbf{v}})\right] = \frac{1}{P}\sum_{\boldsymbol{\tau}=1}^{P} \mathrm{Var}\left[\sum_{k=1}^{K} I^{\hat{\mathbf{i}}^k}(\boldsymbol{\tau})\right] = \frac{1}{P}\sum_{\boldsymbol{\tau}=1}^{P}\sum_{k=1}^{K} \mathrm{Var}\left[I^{\hat{\mathbf{i}}^k}(\boldsymbol{\tau})\right], \tag{62}$$

where we used the fact that the intensity image is a sum of K independent sources. Noting that the variance from a single source is equivalent to the covariance at displacement $\Delta = 0$, we have

$$\sum_{\boldsymbol{\tau}} \mathrm{Var}\left[I^{\hat{\mathbf{i}}^k}(\boldsymbol{\tau})\right] = \sum_{\boldsymbol{\tau}} \boldsymbol{C_I}(0, \boldsymbol{\tau}). \tag{63}$$

Substituting Eq. (63) into Eq. (62) gives us

$$\mathrm{Var}\left[I(\hat{\mathbf{v}})\right] = \frac{K}{P}\sum_{\boldsymbol{\tau}} \boldsymbol{C_I}(0, \boldsymbol{\tau}) = \alpha \sum_{\boldsymbol{\tau}} \boldsymbol{C_I}(0, \boldsymbol{\tau}). \tag{64}$$

Substituting Eq. (64) into Eq. (61) yields

$$\mathrm{Var}\left[\boldsymbol{c}_{\mathrm{emp}}(\mathbf{i}^1_{x,y}, \Delta)\right] = \alpha^2 \sum_{\varphi} |w(\Delta, \varphi)|^2 \left| \sum_{\boldsymbol{\tau}} \boldsymbol{C_I}(0, \boldsymbol{\tau}) \right|^2. \tag{65}$$

Combining Eqs. (59) and (65) provides the desired Eq. (20).

$$\mathrm{SNR} = \frac{|\sum_{\boldsymbol{\tau}} w(\Delta, \boldsymbol{\tau})\boldsymbol{C_I}(\Delta, \boldsymbol{\tau})|^2}{\alpha^2 \cdot \sum_{\varphi} |w(\Delta, \varphi)|^2 \cdot |\sum_{\boldsymbol{\tau}} \boldsymbol{C_I}(0, \boldsymbol{\tau})|^2}. \tag{66}$$

We now turn to determining the optimal weights $w(\Delta, \boldsymbol{\tau})$. This is equivalent to maximizing

$$\frac{|\sum_{\boldsymbol{\tau}} w(\Delta, \boldsymbol{\tau})\boldsymbol{C_I}(\Delta, \boldsymbol{\tau})|^2}{\sum_{\boldsymbol{\tau}} |w(\Delta, \boldsymbol{\tau})|^2}, \tag{67}$$

or equivalently, maximizing $\sum_{\boldsymbol{\tau}} w(\Delta, \boldsymbol{\tau})\boldsymbol{C_I}(\Delta, \boldsymbol{\tau})$ under the constraint $\sum_{\boldsymbol{\tau}} |w(\Delta, \boldsymbol{\tau})|^2 = 1$. It is easy to see (e.g., using Lagrange multipliers) that the maximum is achieved when $w(\Delta, \boldsymbol{\tau}) \propto \boldsymbol{C_I}(\Delta, \boldsymbol{\tau})$. With this choice, Eq. (66) reduces to:

$$\mathrm{SNR}_{\mathrm{matched}} = \frac{\sum_{\boldsymbol{\tau}} |\boldsymbol{C_I}(\Delta, \boldsymbol{\tau})|^2}{\alpha^2 \cdot |\sum_{\boldsymbol{\tau}} \boldsymbol{C_I}(0, \boldsymbol{\tau})|^2}. \tag{68}$$

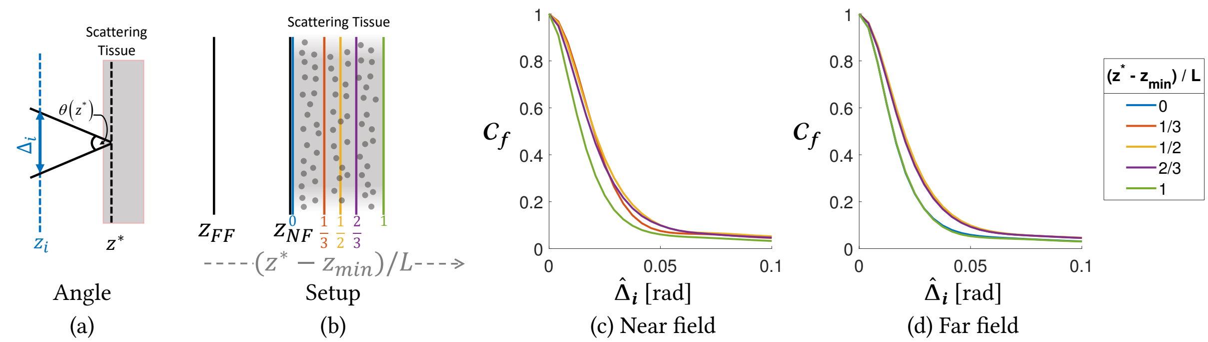

Fig. 21. **Tilt-shift correlation.** We evaluate the tilt-shift correlation $C_f(\Delta_i, \Delta_v, \theta)$ (Eq. (31)), where θ is taken to be the angle illuminators at displacement Δ_i form with a point at depth z^*, as shown in (a). (b) shows a few selections of the plane z^*, and (c,d) show evaluattions of near-field and far-field correlations for the corresponding θ values. In both near-field and far-field configurations, the viewing plane was set as $z_v = z_i$. The best result is obtained when θ is selected using a z^* plane in the middle of the volume.

A.3 Gradient evaluation

The derivative of the loss of Eq. (24) with respect to O is:

$$\frac{\partial E}{\partial O} = 2\sum_j w_j^{2\tau} \cdot \left(e \star \mathbf{F}(w_j^{\Delta} \cdot \bar{I})\right) + w_j^{\Delta} \cdot \left((w_j^{2\tau} \cdot \bar{I}) \star e\right), \qquad (69)$$

where $\mathbf{F}$ denotes the flipping of the image around its center, $\cdot$ is element-wise multiplication, and e is the error of the current guess, $e = \bar{I}_{w_j^{\tau}} \star \bar{I}_{w_j^{\Delta}} - O_{\bar{w}_j^{2\tau}} \star O_{w_j^{\Delta}}$. The gradient can be computed in the frequency domain as

$$2\sum_j w_j^{2\tau} \mathcal{F}^{-1}\left(\mathcal{F}(e)\mathcal{F}(w_j^{\Delta} \cdot \bar{I})\right) + w_j^{\Delta} \mathcal{F}^{-1}\left(\mathcal{F}(w_j^{2\tau} \cdot \bar{I})\mathcal{F}(e)^*\right), \qquad (70)$$

where $\mathcal{F}, \mathcal{F}^{-1}$ denote Fourier transform and its inverse, respectively.

A.4 Additional results

In Fig. 22, we show an example demonstrating that, as illuminator density increases, eventually our local approach fails as well.

In Fig. 23 we show additional reconstructions from our single-capture setup, where a mask target is simultaneously illuminated by a spatially-incoherent LED source, as in Fig. 14.

In Fig. 24, we show a result that demonstrates the increased range our algorithm provides. We select a few local subwindows from the first pattern in Fig. 11, and display the maximal window for which the full-frame auto-correlation was successful. Each subwindow pair demonstrates a small window where reconstruction succeeds, and a slightly bigger one where reconstruction fails.

A.5 Selecting local window parameters.

Our local algorithm has two hyperparameters, T_τ and T_Δ. For our experiments, we set these parameters manually to maximize image quality. Fig. 25 evaluates the sensitivity of our algorithm to their exact values. Very large values of T_τ bring our algorithm closer to the full-frame auto-correlation algorithm, and often result in convergence failure. Increasing T_Δ improves the results, so long as some correlation remains (in the example of Fig. 25, correlation between the far illuminators did not completely decay to 0). This is because larger values of T_Δ provide more constraints on the latent image. By contrast, for very small T_Δ values, reconstruction becomes poor because there are not enough constraints. We chose to use a binary approximation to the matched filter, as the exact matched filter is unavailable in a general experimental setting where material parameters are unknown. However, in our implementation, as we image scattering through individual positions of the laser source behind the tissue, we can also compute the exact speckle covariance empirically as in Fig. 6, giving us access to the exact matched filter. In Fig. 25, we compare the results produced using the exact matched filter (limited to a maximal displacement of $T_\Delta = 500$ pixels) and the binary approximation. We observe that they lead to similar results.

An additional consideration for parameter selection is that decreasing T_τ or increasing T_Δ increases computational complexity.

It is important to emphasize that our algorithm provides improved performance only under the conditions it was designed for: namely when the speckle pattern due to one illuminator has local support. As the support size increases, our algorithm becomes equivalent to the full-frame autocorrelation approach. We have empirically observed that our algorithm provides improved performance when the speckle support size is smaller than the range of memory effect correlation—that is, when the true values of T_τ, T_Δ satisfy $T_\tau < T_\Delta$. In particular, as the true value of T_τ decreases or the true value of T_Δ increases, the robustness of our algorithm and the number of independent sources it can reliably detect both improve. By contrast, if the true value of T_τ is greater than that of T_Δ, our algorithm becomes equivalent to the full-frame auto-correlation, and reconstruction will be limited by the range of the memory effect.

A.6 Fluorescent beads imaging

We provide details for the fluorescent bead imaging experiment in Fig. 20. We place *FluoSpheres* beads of diameter 2 μm on a 170 μm cover glass attached at the back of a 100 μm chicken breast tissue slice. We use the setup of Fig. 26 to image the samples. The beads

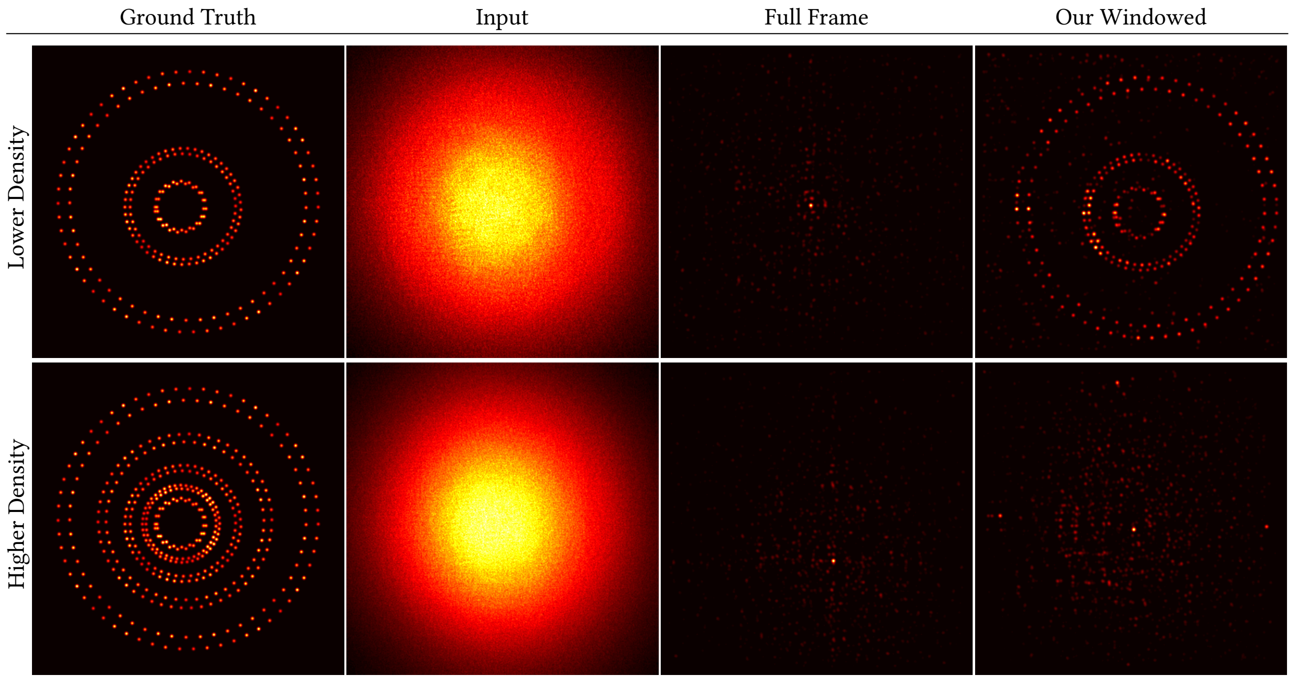

Fig. 22. **Recoverable density.** Reconstructions at two densities of a far-field pattern. As illuminator density increases, our local-cost fails as well.

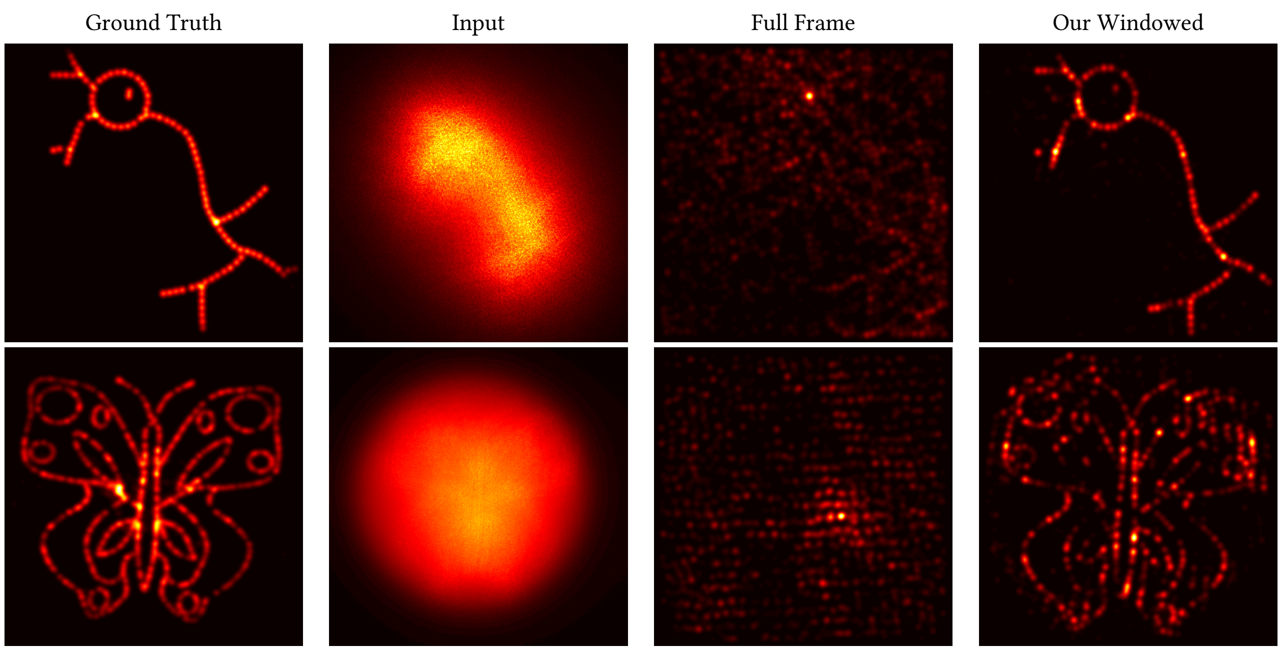

Fig. 23. **Far-field reconstructions using an LED source.** Our local correlation approach outperforms the classical full-frame auto-correlation approach.

are excited by a wide-area 637 nm laser beam, and emit light at a spectral range of about $20 - 30$ nm centered at 680 nm. The emitted light scatters through the tissue and we use a sensor to image it from the other side of the sample. We use a dichroic filter and a 10 nm band-pass filter centered at 680 nm, to filter out the laser illumination used for excitement, and to limit the bandwidth of the fluorescent emission. The sensor captures the input image in Fig. 20.

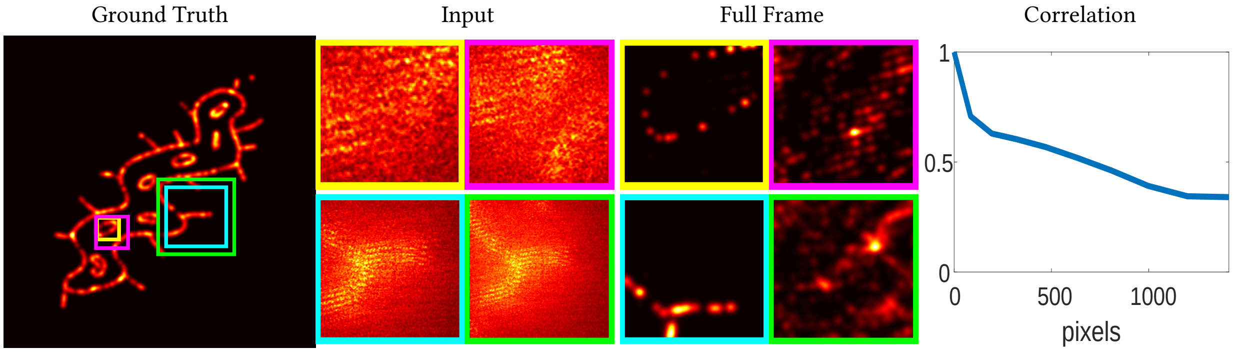

Fig. 24. **Full-frame auto-correlation algorithm applied on small crops of the patterns in Fig. 11.** The yellow and cyan sub-windows demonstrate areas where reconstruction roughly succeeds, and the magenta and green ones show slightly larger areas where reconstruction fails. To the right, we plot correlation as a function of displacement length $|\Delta|$, as measured for the corresponding tissue slice.

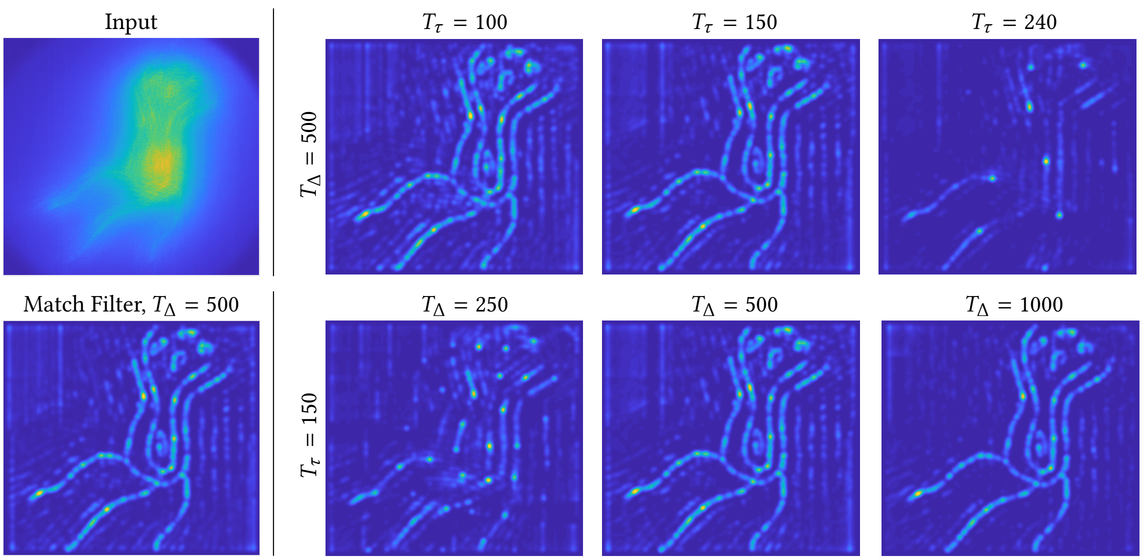

Fig. 25. **Sensitivity to window parameters.** We visualize results from our algorithm when varying the support of the binary matched filter, T_τ, T_Δ. We also compare with an exact matched filter computed from empirical speckle correlations, utilizing a maximal displacement of $T_\Delta = 500$ pixels.

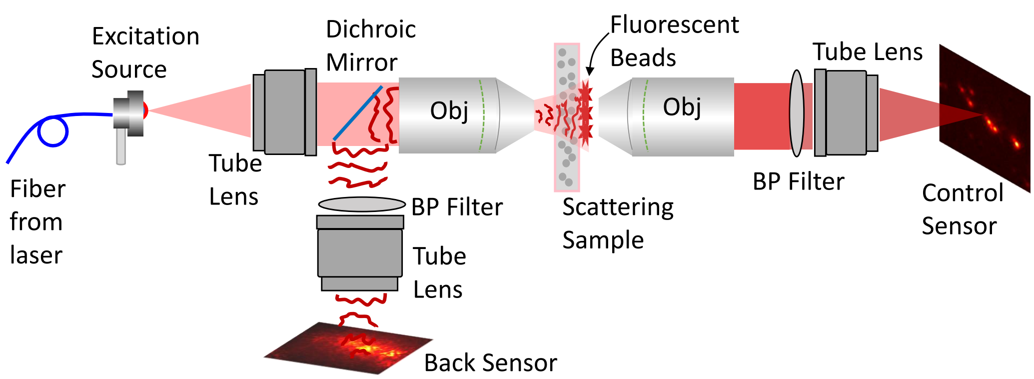

Fig. 26. **Schematic of near-field fluorescent setup.** Fluorescent beads are excited by a laser and emit light in a higher spectral band. Part of the emitted light passes though the scattering sample, a dichroic mirror, and a band-pass (BP) filter, and is imaged by the back sensor. Additionally, another part of the emitted light is imaged by the control sensor, for the purposes of ground truth acquisition.