nature photonics

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41566-020-0630-0

In the format provided by the authors and unedited.

Fluorescence imaging through dynamic scattering media with speckle-encoded ultrasound-modulated light correlation

Haowen Ruan[1,2]✉, Yan Liu[1,2]✉, Jian Xu[1], Yujia Huang[1] and Changhuei Yang[1]✉

[1]Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA. [2]These authors contributed equally: Haowen Ruan, Yan Liu.
✉e-mail: hruan@caltech.edu; yanliu@caltech.edu; chyang@caltech.edu

Fluorescence imaging through dynamic scattering media with speckle encoded ultrasound modulated light correlation

Haowen Ruan[1,†,*], Yan Liu[1,†,*], Jian Xu[1], Yujia Huang[1], and Changhuei Yang[1,*]

[1]Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, 91125

[†]These authors contributed equally to this work

[*]Contact emails: hruan@caltech.edu, yanliu@caltech.edu or chyang@caltech.edu

Supplementary Figures

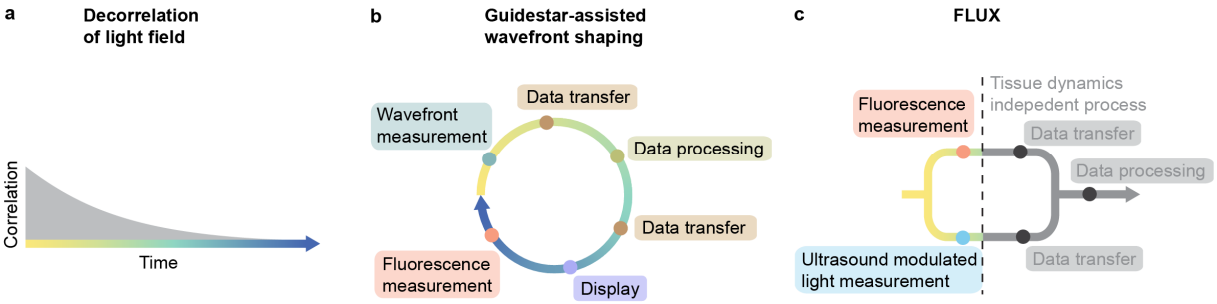

Figure S1 | Comparison of FLUX with the guidestar-assisted wavefront shaping technique in terms of the vulnerability to speckle decorrelation. a. Dynamic samples such as *in vivo* tissue lead to exponential light field decorrelation over time. **b**. In digital optical phase conjugation, the fastest guidestar-assisted wavefront shaping technique, a sequence of steps needs to be completed before the shaped light returns to tissue and excites a fluorescent target. Therefore, the field decorrelation during the runtime of the entire cycle determines the system performance. As such, wavefront shaping is not effective because of the significant field decorrelation shown in the figure. **c**. Signal flow of FLUX. Fluorescence measurement and ultrasound modulated light measurement take place simultaneously, after which the signals will not interact with the sample again, and the speckle decorrelation during data transfer and processing decouple from the system performance. Therefore, this method is much more immune to speckle decorrelation.

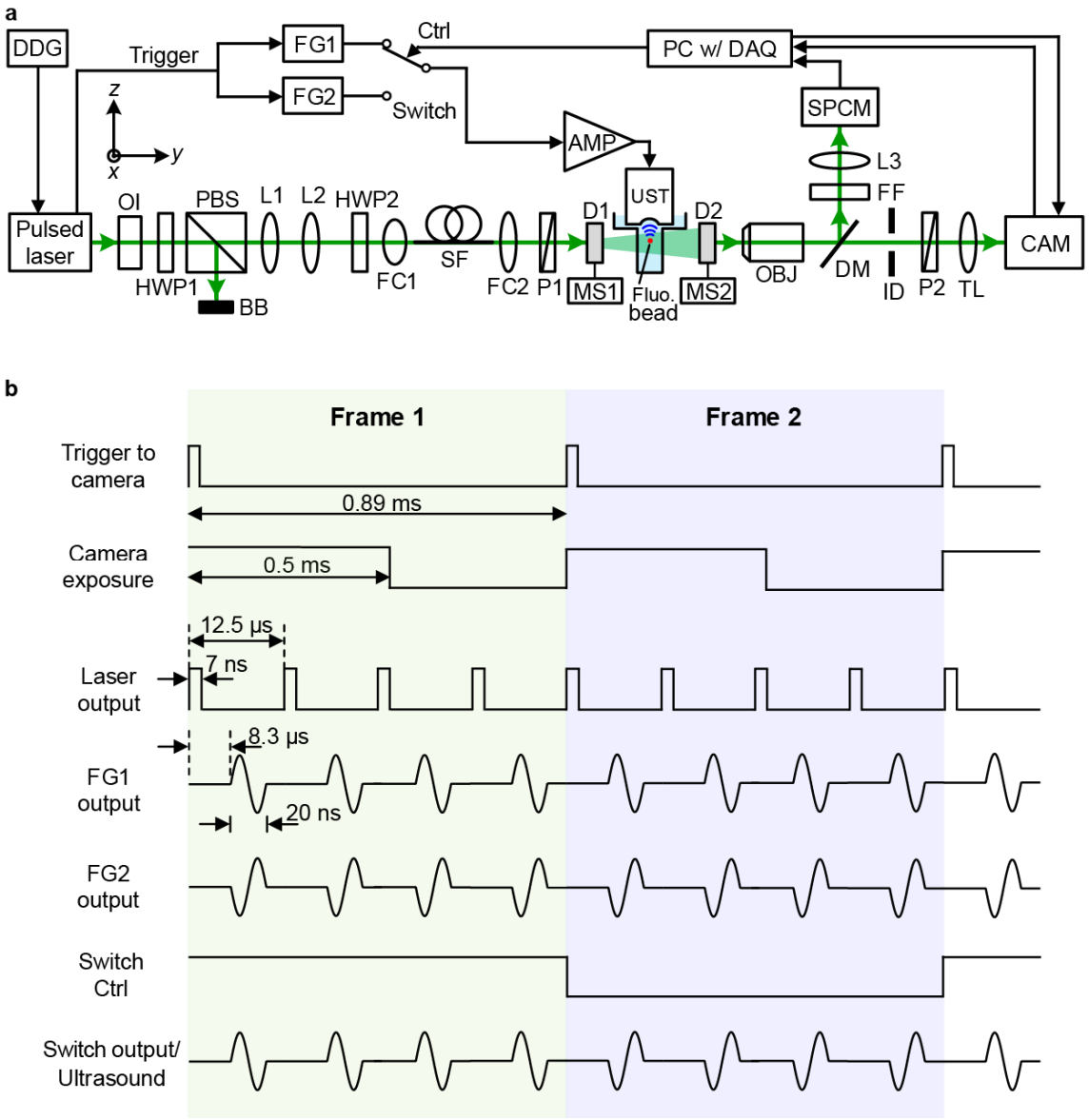

Figure S2 | Schematic of the experimental setup and electrical signal flow diagram. a. Setup. Abbreviations: AMP, amplifier; BB, beam block; CAM, camera; Ctrl, control port of the switch; D, light-scattering diffuser; DAQ, multifunctional data acquisition card; DDG, digital delay generator; DM, dichroic mirror; FC, fibre coupler; FF, fluorescent filter; FG, function generator; Fluo, fluorescent; HWP, half-wave plate; ID, iris diaphragm; L, lens; MS, motorized stage; OBJ, objective; OI, optical isolator; P, polarizer; PBS, polarizing beamsplitter; PC, personal computer; SF, single-mode fibre; SPCM, single photon counting module; TL, tube lens; UST, ultrasound transducer. **b**. Electrical signal flow diagram. Due to multiple time scales involved, the drawing is not to scale, but the timing is labelled by the texts.

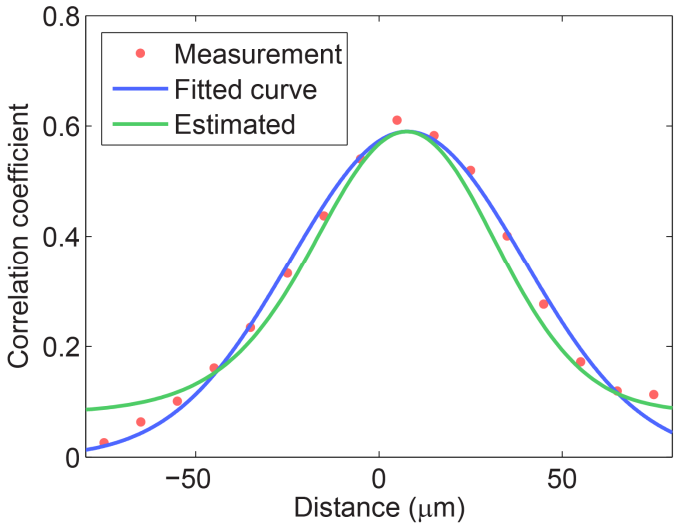

Figure S3 | Comparison of the experimental and theoretical imaging resolution. The experimental data (red dots) and the fitted curve (blue) show the one-dimensional image of a fluorescent particle along the ultrasound transverse direction. This image has a full width at half maximum (FWHM) of 74.8 ± 3.5 μm, the same as the resolution characterization shown in Fig. 2c. We compute the theoretical image of the object (green curve) by cross-correlating the object function with the speckle pattern autocorrelation function and with the ultrasound intensity profile, and find the FWHM of the image (green curve) to be 61.2 μm. Here, the fluorescent particle has a diameter of 10 μm; the FWHM of the speckle pattern autocorrelation function (i.e. speckle size) is 39.0 μm; and the FWHM of the ultrasound lateral profile is measured to be 38.8 μm.

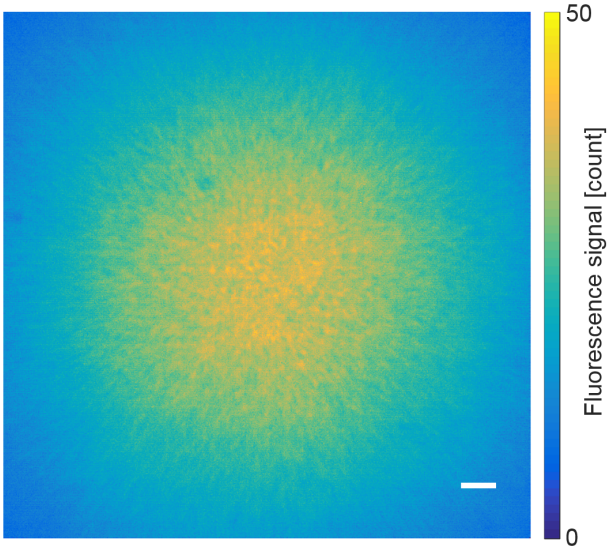

Figure S4 | Image of the fluorescent target acquired by a conventional wide-field fluorescence imaging system in the presence of the scattering diffusers. The FWHM spot size is 1.3 mm. Scale bar, 100 μm.

Supplementary Notes

Supplementary Note 1 | FLUX analytical model and signal-to-noise ratio expression

In this section, we derive a formula for *FLUX* and produce an analytical form for its expectation and variance.

To start, we assume that the excitation light field illuminates a given volume of a fluorescent sample (V_0) with an average intensity of I_0. The ultrasound focus occupies a smaller volume (V) within this illuminated volume and has an average intensity of I_U. The number of independent optical speckle grains within V is given by M_{in} (see Fig. S5). The number of independent optical speckle grain within V_0 but outside of V is given by M_{out}. The number of independent speckle grain in volume V_0 is given by $M_{all} = M_{in} + M_{out}$. We can split V_0 into a grid of voxels, with the voxel size equal to the speckle grain size. The number of fluorophores in the i-th voxel is given by N_i. We further use $i \in M_{in}$ to denote the voxels that are within the ultrasound focus and $i \in M_{out}$ to denote the voxels that are outside of the ultrasound focus.

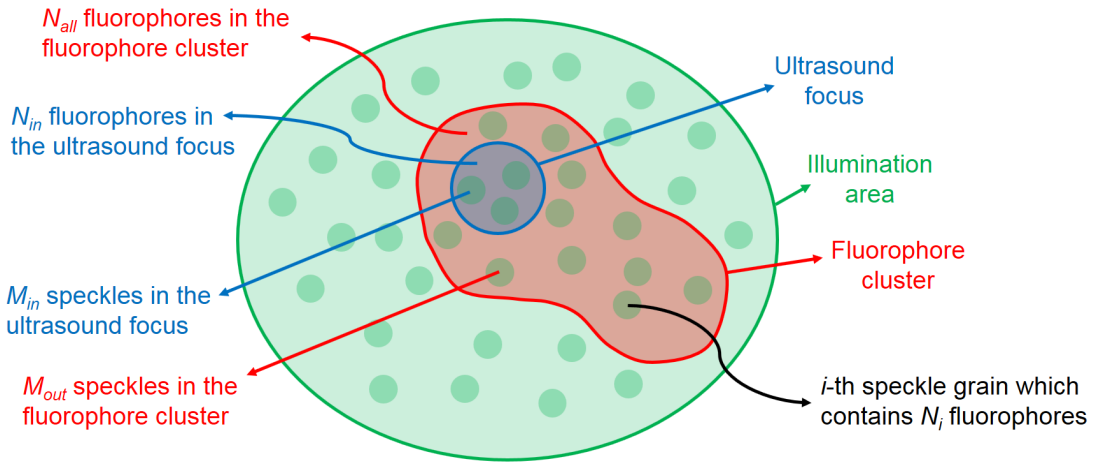

Figure S5 | Illustration of different variables. A fluorophore cluster (shown in red) is located in the illumination area (shown in green), and an ultrasound focus (shown in blue) is located on the fluorophore cluster.

In this formulation, the total fluorescence photon count rate detected from the sample can be expressed as:

$$
\begin{aligned}
F(t) &= F_{dc} + F_{ac}(t) + n_f(t), \\
F_{dc} &= \sum_{i \in M_{in}} I_0 q_f N_i + \sum_{i \in M_{out}} I_0 q_f N_i, \\
F_{ac}(t) &= \sum_{i \in M_{in}} I_0 q_f N_i (S_i(t) - 1) + \sum_{i \in M_{out}} I_0 q_f N_i (S_i(t) - 1), \\
n_f(t) &= \sum_{i \in M_{in} + M_{out}} n_{f,i}(t),
\end{aligned}
\tag{S1}
$$

where $I_0S_i(t)$ denotes the illumination intensity time trace at the i-th voxel. $S_i(t)$ is a dimensionless quantity with a mean of 1 and a variance of 1, representing full developed speckles. $S_i(t)$ decorrelates with a time constant τ_d. q_f is a constant that encapsulates various experimental parameters related to the fluorescence generation and detection process. It is defined as the probability of an illumination photon being absorbed by a fluorophore and resulting in a fluorescence photon being detected by the system, multiplied by the inverse of the photon quantum energy. In Eq. S1, we separate the summation of $F_{ac}(t)$ into two subsets associated with M_{in} and M_{out} for subsequent calculation, because the former contributes to the signal while the latter contributes to noise. $n_{f,i}(t)$ is the fluorescence shot noise associated with the i-th voxel. It is white in spectrum, zero mean and has the following property:

$$\left\langle \left(\int_0^{\delta t} n_{f,i}(t)dt \right)^2 \right\rangle = I_0 q_f N_i \delta t, \tag{S2}$$

where δt is a small time increment.

For completeness, the factor q_f defined in Eq. S1 can be written as:

$$q_f = \frac{\eta_f \sigma \eta_{det}}{h\nu}, \tag{S3}$$

where η_f is the fluorophore quantum efficiency, σ is the fluorophore absorption cross section, η_{dec} is the emitted photon collection efficiency, h is the Planck constant, v is the frequency of the photon.

The total ultrasound modulated photon count rate detected from the sample can be expressed as:

$$\begin{aligned} U(t) &= U_{dc} + U_{ac}(t) + n_u(t), \\ U_{dc} &= \sum_{i \in M_{in}} I_0 q_u \frac{V}{M_{in}} I_U, \\ U_{ac}(t) &= \sum_{i \in M_{in}} I_0 q_u \frac{V}{M_{in}} I_U (S_i(t) - 1) + \sum_{i \in M_{in}} n_{u,i}(t), \end{aligned} \tag{S4}$$

where q_u is a constant that encapsulates various experimental parameters related to the ultrasound modulated photon generation and detection process. $n_{u,i}(t)$ is the ultrasound modulated photon shot noise associated with the i-th voxel. It is white in spectrum, zero mean and has the following property:

$$\left\langle \left(\int_0^{\delta t} n_{u,i}(t)dt \right)^2 \right\rangle = I_0 q_u \frac{V}{M_{in}} I_U \delta t, \tag{S5}$$

where δt is a small time increment.

The FLUX signal is a time integration of the product between $F_{ac}(t)$ and $U_{ac}(t)$:

$$FLUX = T \int_0^T F_{ac}(t) U_{ac}(t) dt, \tag{S6}$$

where T is the total measurement time duration.

The expectation of *FLUX* is given by:

$$\langle FLUX \rangle = T\int_0^T \langle F_{ac}(t)U_{ac}(t)\rangle dt = (I_0 q_f N_{in} T)\times(I_0 q_u V I_U T)\times\frac{1}{M_{in}}, \quad \text{S7}$$

where N_{in} equals $\sum_{i\in M_{in}} N_i$ and is simply the total number of fluorophores within the ultrasound focus. The reduction of Eq. S7 to its simplified form on the right hand side makes use of the fact that for fully developed speckles

$$\langle S_i(t)S_j(t)\rangle = \begin{cases} 2, & \text{for } i = j \\ 1, & \text{for } i \neq j \end{cases}. \quad \text{S8}$$

The individual factors in the product of Eq. S7 have simple interpretations. The first factor is equal to the total fluorescence photons detected from within the ultrasound focal volume. The second factor is the total ultrasound modulated photons detected from that same volume.

The variance of *FLUX* can be derived and expressed as:

$$\begin{aligned} &\mathrm{var}(FLUX) = \langle FLUX^2\rangle - \langle FLUX\rangle^2 \\ &= \gamma\tau_d T^3 (I_0 q_f)^2 \left(I_0 q_u \frac{V}{M_{in}} I_U\right)^2 M_{in} \sum_{i\in M_{in}+M_{out}} N_i^2 \\ &+ T^3 (I_0 q_f)\left(I_0 q_u \frac{V}{M_{in}} I_U\right)^2 M_{in} N_{all} \\ &+ T^3 (I_0 q_f)^2 \left(I_0 q_u \frac{V}{M_{in}} I_U\right) M_{in} \sum_{i\in M_{in}+M_{out}} N_i^2 . \end{aligned} \quad \text{S9}$$

Here, N_{all} equals $\sum_{i\in M_{in}+M_{out}} N_i$ and is simply the total number of fluorophores within the illumination light field; $N_i = c_i V/M_{in}$ where c_i is the fluorophore concentration in the i-th voxel; γ is a dimensionless quantity and is equal to the expected integral $\frac{1}{T\tau_d}\left\langle \int_0^T \int_0^T dt_1 dt_2 (S(t_1)-1)(S(t_2)-1)\right\rangle$ when T is much greater than the speckle decorrelation time τ_d, where $S(t)$ has the same definition as $S_i(t)$ mentioned above. We can also express γ as $\frac{2}{\tau_d}\int_0^T g_{s-1}(t)(1-\frac{t}{T})dt$ where $g_{S-1}(t) = \langle (S(\tau)-1)(S(\tau-t)-1)\rangle_\tau$ is the decorrelation function of $S(t) - 1$. As a reference point, if the fluctuation is flat in frequency content and truncated at $2\pi/\tau_d$, γ equals $\pi/4$. We will use $\gamma = 1$ for simplicity in later calculation.

The first term in the summation of Eq. S9 is the variance term related to the stochastic fluctuations arising from a finite measurement time (T) of the FLUX signal (i.e. speckle statistical

noise). An intuitive interpretation of this is that, the average of a sinusoidal signal is zero, but simply taking a finite integration of a sinusoidal signal will leave a non-zero residual that is phase dependent. It also depends on the spectral frequency content of the speckle fluctuations. The second term in the summation of Eq. S9 is the variance term related to the fluorescence shot noise. The third term in the summation of Eq. S9 is the variance term related to the shot noise of the ultrasound modulated light field. In practice, it is likely that one of these three terms will dominate over the others, and the overall FLUX signal-to-noise ratio (SNR) calculations will be impacted by that term. In the case where the first term dominates, var(*FLUX*) is inversely proportional to M_{in}^2 when the change of M_{in} is caused by a change of speckle size while the ultrasound focal size is fixed.

The general SNR equation can be written as:

$$
\begin{aligned}
SNR_{FLUX} &= \frac{FLUX}{\sqrt{\mathrm{var}\left(FLUX\right)}} \\
&= \frac{(I_0 q_f N_{in} T) \times (I_0 q_u \frac{V}{M_{in}} I_U T)}{\sqrt{\begin{array}{l} \gamma T^3 \tau_d \left(I_0 q_f\right)^2 \left(I_0 q_u \frac{V}{M_{in}} I_U\right)^2 M_{in} \sum_{i \in M_{in}+M_{out}} N_i^2 \\ +T^3 \left(I_0 q_f\right) \left(I_0 q_u \frac{V}{M_{in}} I_U\right)^2 M_{in} N_{all} \\ +T^3 \left(I_0 q_f\right)^2 \left(I_0 q_u \frac{V}{M_{in}} I_U\right) M_{in} \sum_{i \in M_{in}+M_{out}} N_i^2 \end{array}}} \\
&= \sqrt{\frac{T}{\tau_d}} \frac{N_{in}}{\sqrt{\begin{array}{l} \gamma M_{in} \sum_{i \in M_{in}+M_{out}} N_i^2 \\ +\left(I_0 q_f \tau_d\right)^{-1} M_{in} N_{all} \\ +\left(I_0 q_u \frac{V}{M_{in}} I_U \tau_d\right)^{-1} M_{in} \sum_{i \in M_{in}+M_{out}} N_i^2 \end{array}}}.
\end{aligned}
\qquad \text{S10}
$$

Eq. S10 provides a general prescription for calculating the SNR of FLUX.

Supplementary Note 2 | The impact of detection etendue of ultrasound modulated light on the SNR of FLUX

Because the ultrasound modulated light field emerged from the scattering sample is itself a speckle field, the detection of the ultrasound modulated photon count is confounded by the statistical nature of that speckle field as well. We can determine the impact of a finite detection etendue by introducing an additional term in Eq. S4 given by:

$$U(t) = U_{dc} + U_{ac}(t) + n_u(t) + n_{etendue}(t)$$
$$n_{etendue}(t) = \sum_{i \in M_{in}} n_{entendue,i}(t) = \sum_{i \in M_{in}} [\frac{1}{\beta} I_0 q_u \frac{V}{M} I_U (S_i(t) - 1) \sum_{j=1}^{\beta} (\beta_j - 1)], \quad \text{S11}$$

where β is the total number of emerging speckles through which we are detecting the ultrasound modulated light signal. β_j is the normalized exponentially distributed random variable with mean of 1 and variance of 1. In experiments where the ultrasound modulated light signal is collected through an optical system with limited numerical aperture and cameras with finite pixel counts, β can be interpreted as the number of independent optical modes through which we detect the ultrasound modulated light signal.

This term eventually introduces an additional variance term in Eq. S9 as shown below:

$$\begin{aligned}
&\mathrm{var}\left(FLUX\right) = \left\langle FLUX^2 \right\rangle - \left\langle FLUX \right\rangle^2 \\
&= \gamma \tau_d T^3 \left(I_0 q_f\right)^2 \left(I_0 q_u \frac{V}{M_{in}} I_U \right)^2 M_{in} \sum_{i \in M_{in} + M_{out}} N_i^2 \\
&+ T^3 \left(I_0 q_f\right) \left(I_0 q_u \frac{V}{M_{in}} I_U \right)^2 M_{in} N_{all} \\
&+ T^3 \left(I_0 q_f\right)^2 \left(I_0 q_u \frac{V}{M_{in}} I_U \right) M_{in} \sum_{i \in M_{in} + M_{out}} N_i^2 \\
&+ \gamma T^3 \min(\tau_d, \tau_{US\text{-}detector}) \left(I_0 q_f\right)^2 \left(I_0 q_u \frac{V}{M_{in}} I_U \right)^2 \frac{M_{in}}{\beta} M_{in} \sum_{i \in M_{in} + M_{out}} N_i^2,
\end{aligned} \quad \text{S12}$$

where $\tau_{US\text{-}detector}$ is the speckle decorrelation time associated with the dynamic scattering medium between the ultrasound focus and the detector. In the reflection detection geometry, $\tau_{US\text{-}detector} = \tau_d$. In this case, Eq. S10 becomes

$$SNR_{FLUX} = \frac{FLUX}{\sqrt{\mathrm{var}\left(FLUX\right)}} = \sqrt{\frac{T}{\tau_d}} \frac{N_{in}}{\sqrt{\begin{aligned} &\gamma M_{in} (1 + \frac{M_{in}}{\beta}) \sum_{i \in M_{in} + M_{out}} N_i^2 \\ &+ \left(I_0 q_f \tau_d\right)^{-1} M_{in} N_{all} \\ &+ \left(I_0 q_u \frac{V}{M_{in}} I_U \tau_d \right)^{-1} M_{in} \sum_{i \in M_{in} + M_{out}} N_i^2 \end{aligned}}}. \quad \text{S13}$$

The *FLUX* variance term introduced by the detection etendue is simply a M_{in}/β scaled version of the first variance term in Eq. S9. As long as $M_{in}/\beta \ll 1$, this term can be neglected in the variance and SNR calculation. In our experiments, $M_{in}/\beta < 0.01$, and thus the variance term introduced by the detection etendue can be ignored without significant effects. In this case, Eq. S13 reduces to Eq. S10.

Supplementary Note 3 | Analysing the feasibility of FLUX for imaging in tissue

In this case study, we apply the FLUX SNR equation to an imaging scenario where a fluorescent object is located at 3 mm deep inside biological tissue. The primary aim of this exercise is to explore the theoretical feasibility of FLUX and examine the possible technical challenges associated with adapting FLUX for a practically useful application.

For ease of calculation, we make two assumptions. First, we assume that the fluorophores are uniformly distributed within the fluorescent object; this allows us to simplify N_i for all speckles within the object to a constant number. Second, we assume the feature size of the fluorescent object is larger than the spatial resolution of the FLUX imaging system. In other words, we assume all the speckle grains within the ultrasound focus are fluorophore bearing, when the ultrasound is focused within the object.

The FLUX SNR equation (Eq. S10) can then be simplified as:

$$
\begin{aligned}
SNR_{FLUX} &= \frac{FLUX}{\sqrt{\mathrm{var}\left(FLUX\right)}} \\
&= \sqrt{\frac{T}{\tau_d}} \frac{N_{in}}{\sqrt{\gamma M_{in} N_{all} (\frac{N_{in}}{M_{in}}) + \left(I_0 q_f \tau_d\right)^{-1} M_{in} N_{all} + \left(I_0 q_u \frac{V}{M_{in}} I_U \tau_d\right)^{-1} M_{in} N_{all} (\frac{N_{in}}{M_{in}})}} \\
&== \sqrt{\frac{T}{\tau_d}} \sqrt{\frac{N_{in}}{N_{all}}} \frac{1}{\sqrt{\gamma + \left(I_0 q_f \frac{N_{in}}{M_{in}} \tau_d\right)^{-1} + \left(I_0 q_u I_U \frac{V}{M_{in}} \tau_d\right)^{-1}}} \\
&= \sqrt{\frac{T}{\tau_d}} \sqrt{\frac{N_{in}}{N_{all}}} \frac{1}{\sqrt{\gamma + \frac{M_{in}}{I_0 q_f N_{in} \tau_d} + \frac{M_{in}}{I_0 q_u I_U V \tau_d}}}.
\end{aligned}
\qquad \text{S10.1}
$$

The three terms under the big square root of Eq. S10.1 describes the three noise factors for FLUX. The first term is associated with the stochastic fluctuation arising from the finite measurement time (T) of the FLUX signal. The second term is associated with the fluorescence shot noise of the fluorescence measurement process, and it has a simple interpretation – it is the inverse of the average number of fluorescence photons originating from a single independent illumination speckle grain that is detected per speckle decorrelation time τ_d. The third term is associated with the shot noise from the ultrasound modulated light measurement process. It has the corresponding interpretation – it is the inverse of the average number of ultrasound modulated photons from a single speckle grain that is detected per speckle decorrelation time τ_d.

The fraction N_{in}/N_{all} can be replaced by the ratio of the speckle count in the ultrasound focus to the speckle count of all fluorophore bearing speckles M_{in}/M_{all}. This translation is only valid because we assume the fluorophore concentration is uniform within the fluorescent object. Nevertheless, this translation maps the FLUX SNR equation into a form that is likely more familiar to the wavefront shaping community where mode counting is important.

Overall this equation lends itself well to physical interpretations. However, we note that it is an approximation that does not account for the impact that a heterogeneous fluorophore concentration distribution within the fluorescent object would have on the SNR. In fact, the mode counting framework used by the wavefront shaping community fails in FLUX when heterogeneous fluorescent targets are involved. For a more accurate SNR analysis, Eq. S10 should be used.

For our case study, we assume that the ultrasound used has a frequency of 50 MHz and the ultrasound lateral full width at half maximum focal spot size d_{US} = 40 μm. The ultrasound is pulsed and each pulse is comprised of a single cycle – giving us an axial focal spot length of 30 μm. This implies that the volume targeted by the ultrasound focus is roughly 4.8×10^{-14} m^3.

We assume that the tissue is semi-infinite and it has typical optical properties[1, 2]: scattering coefficient μ_s = 100 cm^{-1}, absorption coefficient μ_a = 0.1 cm^{-1}, anisotropy g = 0.9, reduced scattering coefficient $\mu_s^{'} = \mu_s(1-g)$ = 10 cm^{-1}. The fluorescent object is located at a depth of 3 mm in the tissue. It is a cubic object with a side length of 170 μm (total volume = 5×10^6 μm^3). The fluorophore concentration is assumed to be uniform (concentration = 10 μM) within the fluorescent object for simplicity, which allows us to directly apply Eq. S10.1 for our calculations. We illuminate the tissue from the surface at an average intensity I_{th} = 200 mW/cm^2 (within the safety limit[3]) and a wavelength λ = 800 nm. The light is pulsed and synchronized to the ultrasound pulse train, but the time-averaged intensity conforms to the quantity stated above. We assume that the emission wavelength is 800 nm as well for simplicity.

The fluence rate in tissue can be calculated by using the diffusion theory for a planar source[1,2]: $\phi_{sz}(z) = \mu_{eff} / (2\mu_a) \times \exp(-\mu_{eff} z) \times I_{th}$, where z is the depth, $\mu_{eff} = \sqrt{3\mu_a(\mu_a + \mu_s^{'})}$ is the effective attenuation coefficient. At a depth of 3 mm, the light field would become a full developed speckle field. If we assume that each speckle grain (lateral dimension is $\lambda/2$ in all three dimensions) within the fully developed speckle field at 3 mm depth is independent, this implies that $M_{in} = 7.5 \times 10^5$. Suppose we collect ultrasound modulated photons through 10^7 optical modes using a commercially available high-speed camera (e.g. Phantom from Vision Research Inc.); this gives us $M_{in}/\beta = 0.075$ and thus makes the impact of ultrasound modulated photon detection etendue on the SNR of FLUX negligible.

The fluorophore concentration is 10 μM, therefore the number of fluorophores per speckle grain n is 3.9×10^2. The fluorophore is assumed to have a quantum efficiency η_f of 50% and an absorption cross section[4] σ of 1×10^{-19} m^2. We assume that the fluorescence is detected through an ideal detector of area A_f = 1 mm^2 placed on the surface of the tissue, where the average fluence rate of light emerged from each speckle at 3 mm depth can be calculated by employing the diffusion theory for a point source[1, 2] $\phi_{zs}(r) = 1/(4\pi D r) \times \exp(-\mu_{eff} r) \times P_0$. Here, $D = 1/[3(\mu_a + \mu_s^{'})]$ is the diffusion coefficient; r is the distance between the source point and the field point; P_0 is the power of the point source. The fluorescence photon detection efficiency $\eta_{det} = \int_{A_f} \phi_{zs}(r) dA \Big/ P_0$. From Eq. S3, $q_f = \eta_f \sigma \eta_{det} / h\nu$ = 9.5 $\times10^{-3}$ m^2·W^{-1}·s^{-1}. We further assume that this tissue has a decorrelation time τ_d of 1 ms at 3 mm depth. Then, in Eq. 10.1, the average number of fluorescence

photons originating from a single independent speckle grain at 3 mm depth that is detected per speckle decorrelation time $N_f = I_0 q_f \left(N_{in}/M_{in}\right)\tau_d = \phi_{sz}(z) q_f n \tau_d = 37$. Similarly, the average number of ultrasound modulated photons originating from a single independent speckle grain at 3 mm depth that is detected per speckle decorrelation time $N_u = \frac{\phi_{sz}(z)(\lambda/2)^2 \tau_d \eta_{modu} \eta_{u,det}}{h\nu} \approx 2.6 \times 10^3$, where η_{modu} = 1% is the ultrasound modulation efficiency, $\eta_{u,det} = \int_{A_u} \phi_{zs}(r) dA \Big/ P_0$ is the detection efficiency of ultrasound modulated light, A_u = 1 cm^2 is the area of the camera sensor used to detect the ultrasound modulated light.

Putting these parameters into Eq. S10.1, because $\gamma \gg \left(1/N_f\right)$ and $\gamma \gg \left(1/N_u\right)$, the *FLUX* variance term associated with the finite measurement time dominates over the other two shot noise related variance terms. This reduces Eq. S10.1 to

$$SNR_{FLUX} = \sqrt{\frac{T}{\tau_d}} \sqrt{\frac{N_{in}}{N_{all}}}. \qquad \text{S14}$$

From Eq. S14, we note that in this regime, the SNR of FLUX is independent of the number of speckle grains within the ultrasound focus (we assume the ultrasound focal size does not change and the speckle size change causes the number of speckle grains within the ultrasound focus to change). The fraction of fluorophore inside the ultrasound focus compared to the whole fluorescent object N_{in}/N_{all} = 0.01. If we desire a FLUX SNR of 10, we would need a total measurement time of 10 s. This is a long measurement time, but not impractically so. FLUX is mainly developed to overcome the fast speckle decorrelation caused by light scattered by moving red blood cells in blood vessels, which poses a grand challenge for wavefront shaping techniques to form a focus in living tissue. In practice, the red blood cells in blood vessels are highly dynamic but the fluorophores can be quasi-static. FLUX has the advantage of allowing fluorescence imaging at ultrasound resolution (~40 μm) at a depth where photons are mostly diffusive; at 3 mm depth, the diffuse optical tomography method generally provides a resolution of ~1 mm at best. According to Eq. S14, the SNR of FLUX is proportional to the square root of the ratio between the number of fluorophores within the ultrasound focus and the total number of fluorophores. As such, FLUX may be suited for applications where the fluorophores are clustered but the number of clusters is low (i.e. sparse objects). In this case study, the number of background fluorophores outside the ultrasound focus is roughly 1/0.01 = 100 times greater than the number of fluorophores inside the ultrasound focus. Alternatively, FLUX can potentially be used in continuous functional monitoring of fluorophore-labelled targets within scattering samples. In this case, because long time measurement is available, FLUX may be used in samples with a larger fluorophore-tagged volume compared to the imaging applications. We further note that in the case where autofluorescence or non-specific labelling are present, the SNR of FLUX will reduce. These effects can be modelled by Eq. S10, which is capable of analysing samples with arbitrary fluorescence concentration distribution. When autofluorescence or imperfect labelling are present, N_i in Eq. S10 should include the counts of those fluorophores as well. Some of the methods that mitigate these issues in

conventional fluorescence imaging, such as using near-infrared (NIR) wavelengths for illumination to reduce autofluorescence[5, 6], can potentially be applied.

Equation S14 highlights the interplay between the ratio of the in-focus fluorophores to the background fluorophores and the ratio of the total measurement time to the speckle decorrelation time. Interestingly, if the speckle decorrelation time can be further reduced, SNR_{FLUX} can be boosted, or, in turn, FLUX can work at lower in-focus versus background fluorophore count ratios. Reducing speckle decorrelation time is not a difficult task. One approach would be to introduce a separate vibration to the tissue to quickly agitate the inherent scatterers within the tissue. This reduction of decorrelation time to boost SNR has a limit, as too short a decorrelation time will cause the variance term associated with fluorescence shot noise to begin dominating in Eq. S10.1. In this case study, this transition point is reached for a decorrelation time of ~0.03 ms. In addition to the strategy of directly perturbing the tissue, it is also possible to consider using a well-conditioned fluctuating illumination light field to change the decorrelation time. On a separate note, the *FLUX* variance term associated with fluorescence shot noise can also dominate if the fluorescence concentration is low, the q_f term is low and/or M_{in} is high (see Eq. S10.1). In this study, the parameter set is realistically chosen, and we considered a worst case (for M_{in}) where we assumed the speckle field is fully developed and the power fluctuation of each speckle grain is independent (see the next paragraph for a discussion on this assumption). Even in that case, the variance term associated with fluorescence shot noise is still below the dominance level. The variance term associated with ultrasound modulated light shot noise is generally much smaller than the variance term associated with fluorescence shot noise, and is generally unlikely to dominate the variance of *FLUX*.

The assumption that each speckle grain within the fully developed speckle field at the depth of 3 mm is independent may be overly stringent. A speckle grain in the interior of a volume receives power solely through the speckle grains on the surface of the volume; therefore, to the first order approximation, its power fluctuation should be a scaled summation of the power fluctuations from the surface speckle grains. Thus, the number of truly independent speckle grains can be lower than we calculated. In this case, FLUX can be expected to work better, and the transition point (as quantified by the decorrelation time) where the *FLUX* variance term associated with fluorescence shot noise begins dominating over the variance term associated with a finite measurement time in Eq. S10.1 can be pushed further down.

Supplementary Note 4 | Potential technical improvement

The theoretical analysis shows that it is feasible and promising for this new technique to be applied to fluorescence imaging through highly dynamic scattering media. We have demonstrated FLUX through a prove-of-concept experimental setting. We anticipate that the performance of the experimental setup can be significantly improved and approach the theoretical limit.

On the ultrasound modulated light detection side, we currently take the difference of two frames to obtain the ultrasound modulated light signal. While this method is simple, it is also sensitive to the speckle decorrelation between the two frames, which results in unwanted

difference. Therefore, the SNR of the ultrasound modulated light signal can be improved by shortening the time interval between the two frames, using a single frame approach such as off-axis holography[7], lock-in camera detection[8], and speckle contrast measurement[9], or using optical filter based methods such as spectral hole burning[10]. In addition, the number of ultrasound-modulated speckles can be reduced by using a higher ultrasound frequency, a higher harmonic ultrasound modulation[11], or a longer illumination wavelength. We also realized that the laser pulse timing jitter has an impact on the SNR and the axial resolution of the ultrasound modulated light signal, therefore a pulsed laser with lower jitter and higher stability is desired.

On the fluorescence detection front, the dead time of our single photon detector limited the maximum photon counting rate to the laser pulse repetition rate. The SNR of the fluorescence detection can be significantly improved by using avalanche photodiodes in a linear mode in the future. In addition, a large detection etendue for the fluorescence detection will help to improve the signal collection efficiency and thus improve the SNR.

Finally, we note that the fluorescence targets used in our experiments are made of polystyrene microspheres, whose optical and acoustic refractive indices do not well match with those of tissue phantoms such as water-based gels. This results in unwanted diffraction of optical and acoustic waves. We anticipate that better refractive index matching would help to increase the SNR as well as the correlation between the fluorescence and ultrasound modulated light signals.

References

1. L. V. Wang and H. Wu, Biomedical optics : principles and imaging, Wiley-Interscience (2007).

2. S. L. Jacques and B. W. Pogue, "Tutorial on diffuse light transport," *J. Biomed. Opt.* **13**, 041302 (2008).

3. American National Standards Institute Inc. (ANSI), ANSI Z136.1-2007: American National Standard for Safe Use of Lasers (2007).

4. J. R. Lakowicz, Principles of fluorescence spectroscopy, Springer (2013).

5. Frangioni, J.V. In vivo near-infrared fluorescence imaging. *Curr. Opin. Chem. Biol.* **7**, 626-634 (2003).

6. Hong, G., Antaris, A.L. & Dai, H. Near-infrared fluorophores for biomedical imaging. *Nature Biomedical Engineering* **1**, 0010 (2017).

7. Cuche, E., Marquet, P. & Depeursinge, C. Spatial filtering for zero-order and twin-image elimination in digital off-axis holography. *Appl. Opt.* **39**, 4070-4075 (2000).

8. Liu, Y., Shen, Y., Ma, C., Shi, J. & Wang, L.V. Lock-in camera based heterodyne holography for ultrasound-modulated optical tomography inside dynamic scattering media. *Appl. Phys. Lett.* **108**, 231106 (2016).

9. Li, J., Ku, G. & Wang, L.V. Ultrasound-modulated optical tomography of biological tissue by use of contrast of laser speckles. *Appl. Opt.* **41**, 6030-6035 (2002).

10. Li, Y. et al. Pulsed ultrasound-modulated optical tomography using spectral-hole burning as a narrowband spectral filter. *Appl. Phys. Lett.* **93**, 011111 (2008).

11. Selb, J., Pottier, L. & Boccara, A.C. Nonlinear effects in acousto-optic imaging. *Opt. Lett.* **27**, 918-920 (2002)