The generalized optical memory effect

Gerwin Osnabrugge*,1, Roarke Horstmeyer†,2,3 Ioannis N. Papadopoulos,2 Benjamin Judkewitz,2 and Ivo M. Vellekoop1

1Biomedical Photonic Imaging Group, MIRA Institute for Biomedical Technology & Technical Medicine, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands
2Bioimaging and Neurophotonics Lab, NeuroCure Cluster of Excellence, Charité Berlin and Humboldt University, Berlin, Germany
3Future address: Biomedical Engineering Department, Duke University, Durham, NC 27708, USA

(Dated: September 23, 2018)

The optical memory effect is a well-known type of wave correlation that is observed in coherent fields that scatter through thin and diffusive materials, like biological tissue. It is a fundamental physical property of scattering media that can be harnessed for deep-tissue microscopy or 'through-the-wall' imaging applications. Here we show that the optical memory effect is a special case of a far more general class of wave correlation. Our new theoretical framework explains how waves remain correlated over both space and angle when they are jointly shifted and tilted inside scattering media of arbitrary geometry. We experimentally demonstrate the existence of such coupled correlations and describe how they can be used to optimize the scanning range in adaptive optics microscopes.

I. INTRODUCTION

It is challenging to record clear images from deep within biological tissue. As an optical field passes through tissue its spatial profile becomes randomly perturbed, which subsequently results in a blurry image of the features that lie underneath. Luckily, even highly scattered optical fields still maintain a certain degree of correlation. Such scattering-based correlations have recently enabled new ‘hidden imaging’ approaches [1–4], which reconstruct clear images from behind diffusive materials. These prior investigations have primarily exploited what is traditionally referred to as the optical memory effect [5, 6]. This effect predicts that a scattered wavefront will tilt, but otherwise not change, when the beam incident upon a scattering material is also tilted by the same amount (see Fig. 1a). These correlations have been observed through both thin isotropically scattering screens [7] as well as thick forward scattering tissue [8].

Recently, we reported a new type of ‘shift’ memory effect, illustrated in Fig. 1b, that occurs primarily in anisotropically scattering media [9]. This form of correlation is especially important in biomedical imaging, as it offers the ability to physically shift (as opposed to tilt) a focal spot formed deep within scattering tissue by translating an incident optical beam.

Here, we show how the optical ‘tilt’ and ‘shift’ memory effects are manifestations of a more general source of correlation within a scattering process, which depends upon how an incident wavefront is both tilted and shifted (see Fig. 1c). Our new ‘generalized memory effect’ model offers a complete description of the first-order spatial correlations present within scattering media. While it applies to coherent waves in general, we limit our attention to the optical regime, where its applications range from maximizing the isoplanatic patch of adaptive optics microscopes [10] to ‘hidden imaging’ approaches [1–4]. Here we develop and verify a general model that predicts an optimal imaging/scanning strategy for a given sample.

II. MODEL

In our model, we consider propagation of monochromatic and coherent light from an input plane a to a target plane b. We limit ourselves to scalar waves. The forward-propagating field at the input surface, $E_a$, and

FIG. 1. Three different types of spatial correlations in disordered media. (a) The optical ‘tilt’ memory effect [6], where an input wavefront tilt leads to a tilt at the target plane. (b) The anisotropic ‘shift’ memory effect [9], where an input wavefront shift also shifts the target plane wavefront. (c) Our new generalized memory effect, relying on both tilts and shifts, can maximize correlations along the target plane for a maximum imaging/focus scanning area.
at the target plane, $E_b$, are coupled through

$$E_b(r_b) = \int T(r_b, r_a)E_a(r_a)d^2r_a. \quad (1)$$

It should be noted here that our model is completely general and does not make any assumptions about the location or orientation of the two planes, nor about the symmetry, shape, or other properties of the scattering sample. For example, for transmission through a slab of scattering material, the target plane is located at the back surface of the sample and $T(r_b, r_a)$ is the sample’s transmission matrix. Conversely, in a biomedical microscope setup, the target plane is typically located inside a scattering sample like tissue, and $T$ is a field propagator connecting position $r_a$ to a position $r_b$ located inside the tissue.

At this point let’s define exactly what type of correlations we are interested in. Suppose that an input field $E_a$ results in a target field $E_b$. We hypothesize that when we shift/tilt the incident field with respect to the medium, the new target field will also experience a shift/tilt and also remain somewhat similar to the original target field. This similarity can be expressed as a correlation function involving the original matrix $T$ and a new matrix $\tilde{T}$ associated with the shift/tilted field. We describe ‘tilting’ the incident wave as a multiplication with a phase ramp, $\exp(i\Delta k_a \cdot r_a)$, resulting in a wavefront that is tilted by $\Delta k_a$ with respect to plane $a$. Likewise, we describe the corresponding tilt at the target plane via a multiplication with $\exp(-i\Delta k_b \cdot r_b)$.

We may now describe the new shifted/tilted situation by writing the corresponding shift/tilt matrix as $\tilde{T}(r_b, r_a) = \exp(-i\Delta k_b \cdot r_b)T(r_b + \Delta r_b, r_a + \Delta r_a)\exp(i\Delta k_a \cdot r_a)$, where $\tilde{T}$ is shifted by $\Delta r_a$ and $\Delta r_b$ at the input and target plane, respectively. By calculating the ensemble averaged value $\langle T^{+}\tilde{T}^{*}\rangle$, we can find the corresponding correlation coefficient between the original fields (from $T$) and the tilted/shifted fields (from $\tilde{T}$). However, we found that the problem is expressed more naturally if we first introduce the center-distance coordinates $r_b^+ \equiv r_a + \Delta r_a/2, r_b^- \equiv r_a - \Delta r_a/2$ (and the same for $r_b$). Using the same reasoning as above, we now find a symmetric expression for the field correlation function:

$$C_W(\Delta r_b, \Delta k_b; \Delta r_a, \Delta k_a) \equiv \int \int \langle T(r_b^+, r_a^+)^*T^*(r_b^-, r_a^-) \rangle e^{i(\Delta k_a \cdot r_a - \Delta k_b \cdot r_b)}d^2r_a d^2r_b. \quad (2)$$

In the special case that $\Delta k_a = \Delta k_b$ and $\Delta r_a = \Delta r_b = 0$, equation $2$ reduces to the optical ‘tilt’ memory effect [6], whereas for $\Delta r_a = \Delta r_b$ and $\Delta k_a = \Delta k_b = 0$, the correlation function corresponds to the anisotropic ‘shift’ memory effect described in Ref. [9].

Before proceeding to calculate $C_W$ in terms of the sample properties, we introduce the Wigner distribution function (WDF) of the optical field. The WDF describes an optical field as a joint ‘phase space’ function of two variables: one spatial variable and one wavevector variable. We choose to work with the WDF because it will allow us to convert equation $1$ into a function of both space and wavevector, which we may easily connect to our new correlation function $C_W$ of similar variables. We define the WDF of field $E$ as

$$W(r, k) \equiv \int E(r^+)E^*(r^-)e^{-i(k \cdot r)}d^2r d^2k. \quad (3)$$

While not always exact, it is convenient to connect the value $W(r, k)$ to the amount of optical power at point $r$ that is propagating in direction $k$ [11].

To describe the scattering of incident light over space and wavevector, we introduce the ‘average light field transmission function’, $P_W$. This function maps the incident light field $W_a(r_a, k_a)$ of the input field to the ensemble averaged transmitted light field $\langle W_b(r_b, k_b)\rangle$ at the target plane:

$$\langle W_b(r_b, k_b)\rangle = \frac{1}{(2\pi)^2} \int \int P_W(r_b, k_b; r_a, k_a)W_a(r_a, k_a)d^2r_a d^2k_a. \quad (4)$$

Equation $4$ is the phase space equivalent of equation $1$, now considered with an ensemble average. It is well-known that as a transformation that connects two WDFs, $P_W$ takes the form of a double-Wigner distribution [11]:

$$P_W(r_b, k_b; r_a, k_a) \equiv \int \int \langle T(r_b^+, r_a^+)^*T^*(r_b^-, r_a^-) \rangle e^{i(\Delta r_a \cdot k_a - \Delta r_b \cdot k_b)}d^2r_a d^2k_b. \quad (5)$$

Although $P_W$ is a function of 4 variables, it only depends upon the scattering system’s 2-variable transmission matrix, $T$, and obeys the same properties as a WDF (e.g., rephasal).

In equation $5$, we recognize a Fourier transform from $\Delta r_b$ to $k_b$ and an inverse Fourier transform from $\Delta r_a$ to $k_a$. Performing the reversed transforms on both sides of equation $5$ yields,

$$\langle T(r_b^+, r_a^+)^*T^*(r_b^-, r_a^-) \rangle = \frac{1}{(2\pi)^2} \int \int P_W(r_b, k_b; r_a, k_a)e^{i(-\Delta r_a \cdot k_a + \Delta r_b \cdot k_b)}d^2k_a d^2k_b, \quad (6)$$

which can be inserted into equation $2$ to arrive at our central result:
Here, \( \Delta r_b, \Delta k_b; \Delta r_a, \Delta k_a \) = \( \frac{1}{(2\pi)^3} \int \int \int P_W(\mathbf{r}_b, \mathbf{k}_b; \mathbf{r}_a, \mathbf{k}_a) e^{i(\Delta r_a \cdot \mathbf{k}_a + \Delta r_b \cdot \mathbf{k}_b + \Delta k_a \cdot \mathbf{r}_a - \Delta k_b \cdot \mathbf{r}_b)} d^2\mathbf{k}_a d^2\mathbf{r}_a d^2\mathbf{k}_b d^2\mathbf{r}_b. \) (7)

In short, \( P_W \) and \( C_W \) are connected through two forward and two inverse 2D Fourier transforms. The correlation function in equation (7) is the formulation of our new generalized memory effect, generalizing the well-known ‘tilt’ memory effect into a full class of interrelated shift and tilt correlations. At this point we want to emphasize that equation (7) is still valid for any scattering medium or geometry. Also, these spatio-angular correlations are an intrinsic property of the scattering medium and will thus be present regardless of the form of the input field.

III. APPROXIMATE SOLUTION FOR FORWARD SCATTERING

We are now left with the challenge of calculating \( P_W \). Although we could use the radiative transfer equation in the most general case, we prefer an approximate solution that gives a simple analytical form for \( P_W \) under the condition that light is mainly scattered in the forward direction. Furthermore, we assume paraxial light propagation, as the WDF’s joint description of light across \( \mathbf{r} \) and \( \mathbf{k} \) is very closely related to the light field description of rays [12]. The idea is to approximate the specimen as a series of thin phase plates. Each phase plate slightly changes the propagation direction of the rays while maintaining their position. Between the phase plates, the light propagates along straight rays and maintains its directionality. In the limit of infinitely small steps, this picture translates to a Fokker-Planck equation with the following solution [13] (full derivation in the Supplementary Information A):

\[
P_{FP}^W(\hat{\mathbf{r}}, \hat{\mathbf{k}}) = \frac{12\ell_T^2}{k_0^2 L^4} \exp \left( \frac{6\ell_T}{L} \left[ \frac{\hat{\mathbf{r}}^2}{\ell_T^2} - \frac{\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}}{k_0 L} + \frac{\hat{\mathbf{k}}^2}{3k_0^2} \right] \right), \tag{8}
\]

Here, \( \hat{\mathbf{k}} \equiv \mathbf{k}_b - \mathbf{k}_a, \hat{\mathbf{r}} \equiv \mathbf{r}_b - \mathbf{r}_a - L\mathbf{k}_a/k_0, \) \( k_0 \) is the wavenumber, \( \ell_T \) is the transport mean free path, and \( L \) is the separation between the input and target plane (that is, the target plane depth).

Interestingly, \( P_{FP}^W \) is only a function of two variables. The reduction of \( \mathbf{r}_a \) and \( \mathbf{r}_b \) to a single difference coordinate \( \hat{\mathbf{r}} \) is possible because the Fokker-Planck equation is shift invariant. A similar simplification was used in the original derivation of the ‘tilt’ optical memory effect [6]. By assuming the average scattered intensity envelope only depended upon relative position, the resulting memory effect correlation reduced to a function of only one tilt variable, which is now a commonly applied simplification in many experiments [13]. Note that in this paraxial model the target intensity distribution is additionally offset by \( L\mathbf{k}_a/k_0 \), which is exactly what is expected from pure ballistic propagation through a trans-

IV. EXPERIMENTAL VALIDATION

We now measure \( C_W \) and \( P_W \) for a forward scattering medium in two separate experiments. These experiments will verify the \( P_W-C_W \) relation predicted in equation (7) and the Fokker-Planck model in equation (8). For simplicity, we consider \( C_W \) and \( P_W \) only along the horizontal dimension, with \( x_b \) and \( k_b \) as the horizontal components of the position and the wavevector at the sample back
surface, respectively. We created two samples of 3μm-
diameter silica microspheres immersed in agarose gel with
thicknesses of $L = 258 \pm 3\mu m$ and $520 \pm 5 \mu m$. Using Mie
theory, we find an anisotropy factor of $g = 0.980 \pm 0.007$
and a scattering mean free path of $\ell_s = 0.296 \pm 0.016$
mm at a wavelength of 632.8 nm, yielding a transport
mean free path of $\ell_P = 14.8 \pm 5.2 \text{ mm}$.

In the experimental setup, light from a 632.8 nm HeNe
laser is expanded and split into a reference path and the
sample path. The light is focused onto the scattering
sample using an objective (100x/0.8), and we image the
sample backsurface (target plane) with a CCD camera
through a second identical objective and tube lens ($f = 150\text{ mm}$). The phase of light transmitted through the
sample is determined by means of off-axis holography.
The sample is placed on a translation stage for lateral
movement. Additionally in the experiments, either a pin-
hole or diffuser is placed in front of the first microscope
objective. The experimental setup is illustrated in the
Supplementary Figure.

A. Measurements of the average light field
transmission function

From equation (4), we know that a finite input beam
$E_{in}$ will result in an average WDF at the target plane,
$W_{a}$, that is a convolution between the desired average
light field transmission function $P_W$ and the WDF of the
input field, $W_{a}$. We are thus able to determine $P_W$ by
using a pencil beam for sample illumination (size = 2.0
μm) at $x_a = 0$ and $k_a = 0$, which approximately forms
the input WDF $W_a(x_a,k_a) \propto \delta(x_a)\delta(k_a)$. $P_W$ is then
found by measuring the average scattered light at the
target plane across both space and angle. The pencil
beam is formed by placing a 500 μm-wide pinhole close
to the back-aperture of the first objective. After mea-
suring the scattered field, we numerically calculate the
WDF $W_b(x_b,k_b)$, using equation (3), to find the scat-
ered intensity as a function of both $x_b$ and $k_b$. We av-
erage the WDF over 300 measurements, translating the
sample over a distance of 10 μm in between each measure-
ment for statistical variation. To facilitate comparison of
our measurements with the Fokker-Planck model, which
does not include ballistic light, we chose to remove the
contribution of ballistic light by subtracting the average
transmitted field from every measured field before calculat-
ing its WDF. Following prior work, it is possible to
modify our model to also include ballistic light at the
expense of some added complexity [14, 15].

In Fig. 2 we compare our measured average light field
transmission functions, $P_{av}^W$, to those computed with
the Fokker-Planck model ($P_{FP}^W$) in equation (5). The effect of
the cross-term between $x_b$ and $k_b$, as predicted in $P_{FP}^W$,
is clearly visible in our measurements from the fact that
the distributions are sheared. This shear implies that
the light at the edges of the diffuse spot (large $x_b$) con-
tinues to propagate, on average, in a radially outward
manner (large $k_b$) after scattering. The $P_{av}^W$ measure-
ments are less spread out than the Fokker-Planck model,
which might be a result of the limited optical sensitivity
at the edges of the objective lenses. Both a large field
of view and a large numerical aperture of the microscope
objective are required to measure the full extent of the
average light field transmission function.

B. Measurements of the generalized correlation
function

Next, we experimentally measure the generalized cor-
relation function $C_{W}^{ex}$. We will directly compare these
measurements to our correlation model for $C_{W}^{FP}$ in
equation (10). We will also compare these measurements
to the 2D Fourier transform of the average light field
transmission functions in Fig. 2, which will verify our
main result in equation (7). For the correlation measure-
ments, we replace the pinhole in front of the first objec-
tive with a diffuser that forms a randomized input field
(average speckle size = 400 nm) at the sample surface.
We use a diffuser to minimize correlations within the
input field, which manifest themselves as a convolution
in our measurement of $C_{W}^{ex}$ (see Supplementary Information
B). We tilt the random input field by translating the dif-
funser at the objective back aperture. We record a total
of 625 scattered fields by illuminating the sample at 25
unique spatial locations and under 25 unique angles of
incidence. From this 252 data cube, we compute $C_{W}^{ex}$ by
taking the ensemble average of all the absolute values of

![Fig. 2. Results of the average light field transmission function ($P_W$) experiment. We compare our measurements, $P_{av}^W$, to the Fokker-Planck model prediction ($P_{FP}^W$, from equation (5)) for samples with (a-b) $L = 258 \mu m$ and (c-d) $L = 520 \mu m$. Colorbar indicates the normalized transmitted intensity as function of $x_b$ and $k_b$.](image-url)
FIG. 3. Results of the generalized memory effect correlation ($C_W$) experiments. Measurements of $C_{Wex}$ are compared to $C_{WPW}$ and the Fokker-Planck correlation model $C_{WFP}$ from equation (10) for samples with $L = 258 \, \mu m$ (a-c) and $L = 520 \, \mu m$ (d-f). Dashed lines indicate the optimal scanning condition in equation (11). (g-i), Cross sections of the 2D correlation functions in (a-c), evaluated (g) along the horizontal axis at $\Delta k_b = 0$, (h) along vertical axis at $\Delta x_b = 0$, and (i) along the optimal scan line for the 258 $\mu m$-thick sample. (j-l), Same cross sections for the 520 $\mu m$-thick sample. The black dashed line denotes the $C_{WFP}$ model, while red and blue stars denote the measured $C_{Wex}$ and $C_{WPW}$, respectively.

the correlation coefficients between two fields separated by the same amount of shift $\Delta x_b$ and tilt $\Delta k_b$. Finally, we normalize $C_{W}$ after subtracting the correlation value at the maximum shift, corresponding with the correlations in the ballistic light.

Fig. 3 presents the results of our $C_W$ experiments. For the two different sample thicknesses of $L = 258 \, \mu m$ and $L = 520 \, \mu m$, the measured 2D correlation functions in (a,d) are compared to $C_{WPW}$, the Fourier transform of $P_{Wex}^{\tau}$ in (b,e), and the Fokker-Planck model in (c,f). The dashed line indicates the optimal tilt/scan condition as predicted in equation (11). We also show cross sections through $C_{Wex}$ (red stars), $C_{WPW}$ (blue stars) and $C_{WFP}$ (dashed black) on the right. These are given along the $\Delta k_b = 0$ line (g,j), the $\Delta x_b = 0$ line (h,k), and the optimal target plane scan condition (i,l). These last two plots demonstrate how jointly considered tilts and shifts can increase correlations along $\Delta x_b$ to increase the scan range (i.e., isoplanatic patch) at the target plane. Our correlation measurements in Fig. 3 closely match the Fokker-Planck model. The cross sections for $C_{Wex}$ and $C_{WPW}$ are also in good agreement, except for Fig. 3k. This data corresponds to the thickest sample, for which the limited field of view of the objective lens prevented us from fully measuring $P_{Wex}^{\tau}$ along $x_b$. The truncation of $P_{Wex}^{\tau}$, already observed in Fig. 2b, results in a broadening of its Fourier transform $C_{WPF}^{FP}$.

V. MAXIMIZING THE ISOPLANATIC PATCH

Having experimentally confirmed the validity of the generalized memory effect in forward scattering media, we now examine its application to adaptive optics (AO) systems. AO allows light to be focused inside scattering media by correcting for the induced wavefront distortions by means of a deformable mirror or spatial light modulator. Generally, AO systems are limited to a single plane of wavefront correction which can be tilted to scan the focus. However, a single correction plane cannot correct for a full scattering volume and therefore the focus scan range is limited to a small area termed the isoplanatic patch. A central question in AO is where to conjugate the correction plane to maximize the isoplanatic patch [10]. In the case of conjugate AO, this correction plane is located at the sample’s top (input) surface, whereas in pupil AO it is effectively located at an infinite distance from the sample. In Fig. 4, we diagram how conjugate AO and pupil AO are analogous to experiments that measure the ‘tilt’ (Fig. 4a) and ‘shift’ memory effects (Fig. 4b), respectively. Put in terms of our model, maximizing the isoplanatic patch equates to finding an optimal combination of $\Delta r_a$ and $\Delta k_a$ to maximize $C_{WFP}$ at a shifted target position $\Delta r_b$.

In equation (10), we considered $C_{WFP}$ only in terms of translation and rotation of the target field $\Delta x_b$ and $\Delta k_b$. However, $C_{WFP}$ can also be expressed as a func-
tion of the translation and rotation of the incident field $\Delta r_a$ and $\Delta k_a$. Using the delta function relations from equation (9), we can substitute $\Delta r_b = \Delta r_a + L\Delta k_a/k_0$ and $\Delta k_b = \Delta k_a$ in equation (11) to find the optimal tilt/shift combination at the input plane for a given amount of target shift $\Delta r_b$:

$$\Delta k_a^{opt} = \frac{3k_0\Delta r_b}{2L} \quad \text{and} \quad \Delta r_a^{opt} = -\frac{\Delta r_b}{2}. \quad (12)$$

Hence, to scan a focus by a distance of $\Delta r_b$, the optimal strategy is to shift the incident field by $\Delta r_b/2$ in the opposite direction and then tilt until the desired $\Delta r_b$ is reached. In other words, the optimal scanning mechanism is achieved by conjugating the AO correction plane inside the scattering sample at a depth of $L/3$, which geometrically corresponds to the ideal tilt/shift rotation plane. We illustrate this mechanism for optimized focus scanning in Fig. 4.

In Table I we compare the isoplanatism provided by the three different memory effects. Here, we define isoplanatism as the maximum allowed scan distance $r_n$ at the target plane that maintains 1/e correlation, which we solve for with appropriately defined input variables in equation (10). Interestingly, all memory effects decrease at the same rate of $\sqrt{L}\ell_r/k_0^2 L$ as the target plane is placed deeper inside the medium. However, the scan range of the tilt/tilt memory effect is always a factor of $\sqrt{3}$ larger than the shift/shift memory effect. By exploiting the optimal tilt/shift combination, as given in equation (12), the scanning improvement is increased to a factor of 2.

| Memory effect  | Adaptive Optics | Tilt plane | Scan range |
|----------------|-----------------|------------|------------|
| Shift          | Pupil           | $-\infty$  | $\sqrt{2\ell_r/k_0^2 L}$ |
| Tilt           | Surface Conjugate | 0         | $\sqrt{6\ell_r/k_0^2 L}$ |
| Generalized    | Sample Conjugate | $L/3$     | $\sqrt{8\ell_r/k_0^2 L}$ |

TABLE I. Comparison of the performance of the three different memory effects in terms of adaptive optics scan range.

VI. DISCUSSION

The optical memory effect as reported in Ref. [6] has paved the way for several new imaging techniques that can ‘see through’ thin scattering layers [1–4]. These techniques require the object of interest to be positioned at a distance behind the thin layer and are thus not immediately applicable to situations in biomedical imaging where the object of interest is embedded in scattering tissue. The anisotropic memory effect [9] showed that translational spatial correlations also exists within the scattering sample, which have been demonstrated to be suitable for imaging or focus scanning inside biological tissue [16,17]. Our new generalized memory effect model offers a new theoretical framework, combining the two known memory effects into a single complete model and offering a means to increase the imaging/scanning area.

Using the Fokker-Planck light propagation model, we have shown that the ‘tilt’ memory effect is not only present behind, but also inside scattering layers, proving to be a factor $\sqrt{3}$ more effective as a scanning technique than the ‘shift’ memory effect alone. This finding supports the field-of-view (FOV) advantage of conjugate AO over pupil AO discussed in Ref. [10], as the ‘tilt’ and ‘shift’ memory effects are utilized by conjugate AO and pupil AO, respectively. Our optimal joint tilt/shift scheme, which corresponds to an optimal AO conjugation tilt plane, maximizes the corrected FOV beyond what is predicted independently by only the ‘tilt’ memory effect and only the ‘shift’ memory effect. The scan ranges given in Table I are strictly the correlations that result from the transmission matrix. Correlations in the input beam further extend the scan range of the different memory effects (see Supplementary Information B). The Fokker-Planck model assumes continuous scattering throughout the sample and neglects back-scattered light, which might not hold in strong isotropic scattering media. However, in this case the Fourier relation between $P_V$ and $C_W$ is still valid. The generalized correlation function is the phase-space equivalent of the $C_1$ intensity correlations introduced by Feng et al. [6]. We envisage that there also exist phase-space equivalents of $C_2$ and $C_3$ correlations. These higher-order correlations have recently been exploited to focus light through scattering media with an unexpectedly high efficiency, and investigating their phase-space equivalents may prove equally useful [18].

Concluding, our new generalized memory effect model generalizes the known optical memory effects without making any assumptions about the geometry or scattering properties of the scattering system. The predicted Fourier relation between the average light field transmission function and generalized correlation function has been experimentally verified in forward scattering media. Furthermore, we found that the simple Fokker-Planck model for light propagation is surprisingly accurate in describing the full set of first-order spatial correlations inside forward scattering media. Our new
The generalized memory effect model predicts the maximum distance that a scattered field can be scanned while remaining correlated to its unshifted form. In other words, the generalized memory effect provides the optimal scan mechanism for deep-tissue focusing techniques.

[1] Katz, O., Small, E. & Silberberg, Y. Looking around corners and through thin turbid layers in real time with scattered incoherent light. *Nat. Photon.* **6**, 549–553 (2012).
[2] Bertolotti, J. et al. Non-invasive imaging through opaque scattering layers. *Nature* **491**, 232–234 (2012).
[3] Katz, O., Heidmann, P., Fink, M. & Gigan, S. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nat. Photon.* **8**, 784–790 (2014).
[4] Yang, X., Pu, Y. & Psaltis, D. Imaging blood cells through scattering biological tissue using speckle scanning microscopy. *Opt. Express* **22**, 3405–3413 (2014).
[5] Freund, I., Rosenbluh, M. & Feng, S. Memory effects in propagation of optical waves through disordered media. *Phys. Rev. Lett.* **61**, 2328 (1988).
[6] Feng, S., Kane, C., Lee, P. A. & Stone, A. D. Correlations and fluctuations of coherent wave transmission through disordered media. *Phys. Rev. Lett.* **61**, 834 (1988).
[7] Li, J.H., Genack, A. Z. Correlation in laser speckle. *Phys. Rev. E* **49**, 4530–4533 (1994).
[8] Schott, S., Bertolotti, J., Léger, J.-F., Bourdieu, L. & Gigan, S. Characterization of the angular memory effect of scattered light in biological tissues. *Opt. Express* **23**, 13505–13516 (2015).
[9] Judkewitz, B., Horstmeyer, R., Vellekoop, I. M., Papadopoulos, I. N. & Yang, C. Translation correlations in anisotropically scattering media. *Nature Phys.* **11**, 684–689 (2015).
[10] Mertz, J., Paudel, H. & Bifano, T. G. Field of view advantage of conjugate adaptive optics in microscopy applications. *Appl. Opt.* **54**, 3498–3506 (2015).
[11] Bastiaans, M. J. Wigner distribution function and its application to first-order optics. *J. Opt. Soc. Am.* **69**, 1710–1716 (1979).
[12] Zhang, Z. & Levoy, M. Wigner distributions and how they relate to the light field. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, 1–10 (IEEE, 2009).
[13] Liu, H.-Y. et al. 3d imaging in volumetric scattering media using phase-space measurements. *Opt. Express* **23**, 14461–14471 (2015).
[14] Wax, A. & Thomas, E. Measurement of smoothed Wigner phase-space distributions for small-angle scattering in a turbid medium. *J. Opt. Soc. Am. A* **15**, 1896–1908 (1998).
[15] Cheng, C.-C. & Rayner, M. Propagation of transverse optical coherence in random multiple-scattering media. *Phys. Rev. A* **62**, 023811 (2000).
[16] Papadopoulos, I., Jouhanneau, J., Poulet, J. & Judkewitz, B. Scattering compensation by focus scanning holographic aberration probing (f-sharp). *Nat. Photon.* **11**, 116–123 (2017).
[17] Park, J.-H., Sun, W. & Cui, M. High-resolution in vivo imaging of mouse brain through the intact skull. *Proc. Nat. Acad. Sci.* **12**, 9236–9241 (2015).
[18] Hsu, C. W., Liew, S. F., Goetschy, A., Cao, H. & Stone, A. D. Correlation-enhanced control of wave focusing in disordered media. *Nature Phys. Advance online publication* (2017).
[19] Tanski, I. A. Fundamental solution of Fokker-Planck equation. *arXiv preprint nlin/0407007* (2004).
SUPPLEMENTARY INFORMATION A: FOKKER-PLANCK MODEL

In this section, we derive the Fokker-Planck equation used to model the generalized memory effect, and show how it leads to the simple two-dimensional correlation function over space and angle in equation [10]. We use a paraxial model to describe light propagation inside a forward scattering medium. We assume that light rays propagate through the medium at an angle \( \arcsin(k/k_0) \approx k/k_0 \), with \( k \equiv (k_x, k_y) \) the perpendicular component of the wave vector. The scattering sample is modeled as a stack of scattering phase plates [13], where each plate slightly changes the direction of the beam (i.e. \( k \)), while preserving the position. Between the phase plates, the rays propagate along straight trajectories. In the limit where the number of phase plates per unit length becomes infinite, this step-scatter model reduces to the following Fokker-Planck type partial differential equation for the radiance:

\[
\frac{\partial I(r, k, z)}{\partial z} = -\frac{k_x}{k_0} \frac{\partial I(r, k, z)}{\partial x} - \frac{k_y}{k_0} \frac{\partial I(r, k, z)}{\partial y} + \beta k^2 \left( \frac{\partial^2 I(r, k, z)}{\partial k^2_x} + \frac{\partial^2 I(r, k, z)}{\partial k^2_y} \right),
\]

(A1)

where \( r \equiv (x, y) \) and \( z \) is the optical axis. Here, the first two terms correspond to propagation of light along straight trajectories, while the last two terms describe scattering. We will find an expression for the scattering strength \( \beta \) later.

The model in equation [A1] neglects all backscattered light. Furthermore, the use of a Fokker-Planck equation (also known as a Kolmogorov forward equation) implies a continuous diffusion process, i.e., even after a very small propagation distance all light has been scattered, albeit at a very small angle. Physically, this model corresponds having the limit of the scattering length approach zero, while keeping fixed the transport length. The important implication of this assumption is that there is no ballistic light in the Fokker-Planck model.

Under the initial condition \( I(r, k, z = 0) = (2\pi)^2 \delta(k - k_a) \delta(r - r_a) \), we find the following solution at \( z = L \) using the method outlined in Ref. [19]

\[
I(\hat{r}, \hat{k}, z = L) = \frac{3}{k_0^2 \beta^2 L^4} \exp \left( -\frac{3}{\beta L} \left[ \frac{|\hat{r}|^2}{L^2} \right] - \frac{\hat{r} \cdot \hat{k}}{Lk_0} + \frac{|k|^2}{3k_0^2} \right).
\]

(A2)

Here, we introduced the variables \( \hat{k} \equiv k - k_a \) and \( \hat{r} \equiv r - r_a - L k_a / k_0 \). It can be easily verified through substitution that equation [A2] is a (normalized) solution to equation [A1]. The difference coordinate variables \( \hat{k} \) and \( \hat{r} \) now define the two dimensional light field transmission function \( I_W(\hat{r}, \hat{k}, L) \equiv I(\hat{r}, \hat{k}, L) \) that we introduce in equation [8].

It is worth noting that \( I(\hat{r}, \hat{k}, L) \) describes a joint Gaussian distribution along \( \hat{r} \) and \( \hat{k} \) with covariance matrix,

\[
\Sigma = \beta L \begin{bmatrix}
\frac{2L^2}{3} & 0 & k_0 L & 0 \\
0 & \frac{2L^2}{3} & 0 & k_0 L \\
k_0 L & 0 & 2k_0^2 & 0 \\
0 & k_0 L & 0 & 2k_0^2
\end{bmatrix}
\]

(A3)

At this point, we still need to determine the scattering strength \( \beta \). In order to do so, we compare our results to a discrete scattering model, which assumes that light propagates an average distance of \( l_{sc} \) between each scattering event (i.e., \( l_{sc} \) is the medium’s scattering coefficient).

For this comparison, we first use equation [A2] to calculate the average angular spread of light, \( \langle |\hat{k}|^2 \rangle \), as a function of depth:

\[
\langle |\hat{k}|^2 \rangle \equiv \frac{1}{(2\pi)^2} \int \int |\hat{k}|^2 I(\hat{r}, \hat{k}, L) d\hat{r} d\hat{k} = 4k_0^2 L \beta.
\]

(A4)

Here, we see that the variance of the angular distribution grows linearly with propagation distance \( L \), consistent with a diffusion process in the \( k_0 \) coordinate system at the target plane.

Second, we calculate the expected angular spread as a function of depth for the discrete scattering model. After propagating a short distance \( L \ll l_{sc} \), a fraction of \( 1 - \exp(L/l_{sc}) \) of the light is scattered. The angular spread of the scattered light is usually quantified with the anisotropy factor \( g = \langle \cos \theta \rangle \), with \( \theta \) the scattering angle. Of course, the non-scattered light will still maintain its original direction. When \( L \) is small, we can write \( g = \langle \cos \theta \rangle \approx 1 - \langle |k_0|^2 \rangle / (2k_0^2) \) and we find that

\[
\langle |\hat{k}|^2 \rangle = 2k_0^2 (1 - e^{-L/l_{sc}}) (1 - g) \approx 2k_0^2 L \frac{1 - g}{l_{sc}}.
\]

(A5)
In the limit of $L \to 0$, we can combine equation (A4) and equation (A5) to find
\[
\beta = \frac{1 - g}{2l_{sc}} = \frac{1}{2l_r},
\]
with $l_r \equiv l_{sc}/(1 - g)$ denoting the transport mean free path. Substituting this result into equation (A2) finally gives us equation (8) in the main text.

To find $C_{WP}^{FP}$, we need to evaluate equation (7) using the solution for $P_W = P_W^{FP}$ in equation (8). Since equation (8) only depends on the difference coordinates $\hat{r}$ and $\hat{k}$, we first rewrite the integrals over $r_a$ and $k_a$ to integrals over $\hat{r}$ and $\hat{k}$ by applying the following changes of variables: $r_b \to \hat{r} + r_a + Lk_a/k_0$ and $k_b \to \hat{k} + k_a$. Doing so leads to,
\[
C_W(\Delta r_b, \Delta k_b; \Delta r_a, \Delta k_a) =
\frac{1}{(2\pi)^4} \iint P_W^{FP}(\hat{r}, \hat{k}) e^{i(\Delta r_a \cdot k_a + \Delta r_b \cdot (k + k_a) + \Delta k_a \cdot r_a - \Delta k_b \cdot (\hat{r} + r_a + Lk_a/k_0))} d^2k d^2\hat{r} d^2\hat{k} a d^2 r_a =
\frac{1}{(2\pi)^4} \iint P_W^{FP}(\hat{r}, \hat{k}) e^{i(\Delta r_b \cdot \hat{k} - \Delta k_a \cdot \hat{r})} d^2k d^2\hat{r}.
\]

The integrals over $k_a$ and $r_a$ reduce to the delta function product $(2\pi)^4 \delta(\Delta r_b - \Delta r_a - \Delta k_b L/k_0) \delta(\Delta k_a - \Delta k_b)$, or, equivalently $(2\pi)^4 \delta(\Delta r_b - \Delta r_a - \Delta k_a L/k_0) \delta(\Delta k_b - \Delta k_a)$. We can now write the original four dimensional correlation function as,
\[
C_W(\Delta r_b, \Delta k_b; \Delta r_a, \Delta k_a) =
(2\pi)^2 \delta(\Delta r_b - \Delta r_a - \Delta k_a L/k_0) \delta(\Delta k_b - \Delta k_a) C_{WP}^{FP}(\Delta r_b, \Delta k_b),
\]
with
\[
C_{WP}^{FP}(\Delta r_b, \Delta k_b) \equiv \frac{1}{(2\pi)^2} \iint P_W^{FP}(\hat{r}, \hat{k}) e^{i(\Delta r_b \cdot \hat{k} - \Delta k_a \cdot \hat{r})} d^2k d^2\hat{r}.
\]

Here, $C_{WP}^{FP}$ is a generalized correlation function that is the forward and inverse 2D Fourier transforms along $\hat{r}$ and $\hat{k}$ of the associated average light field transmission function, $P_W^{FP}(\hat{r}, \hat{k})$, for the shift and tilt invariant medium. We use $P_W = P_W^{FP}$ for the particular case of comparing our results to the Fokker-Planck model, described in equation (10). However, equation (A9) holds for any shift/tilt invariant light field transmission function. In the case of the Fokker-Planck model, the total transmission is unity, and the correlation coefficient in equation (A9) is normalized. Otherwise, we may choose to normalize the correlation coefficient by dividing by $\frac{1}{(2\pi)^2} \iint P_W(\hat{r}, \hat{k}) d^2k d^2\hat{r}$, which is the total transmissivity of the sample.

**SUPPLEMENTARY INFORMATION B: CORRELATION MEASUREMENT**

In this section, we present additional mathematical details about our second experiment, where we directly measured the generalized memory effect correlation function. We performed our correlation measurements by first measuring many scattered field responses at the target plane for an input speckle field that we both shift and tilt via translation stages. Here, the sample is physically shifted by a distance $\Delta r$ and the incident wavefront is tilted by $\Delta k$ by translating the diffuser in front of the microscope objective.

Hence, a single shifted/tilted measurement will take the form,
\[
E_{\Delta r, \Delta k}^{ex}(r_b) = \int T(r_b - \Delta r, r_a - \Delta r) E(r_a) e^{i\Delta k \cdot r_a} d^2r_a.
\]

We then perform a pairwise analysis of the measured fields. For each pair of fields, we have $\Delta r_a = \Delta r_1 - \Delta r_2$ and $\Delta k_a = \Delta k_1 - \Delta k_2$. For simplicity, and without loss of generality, we now only consider the symmetric case $\Delta r_1 = -\Delta r_2 = 2\Delta r_a$ and $\Delta k_1 = -\Delta k_2 = 2\Delta k_a$.

After recording, we digitally shift and tilt the first field by $-\Delta r_b/2$ and $-\Delta k_b/2$, respectively, and we shift and tilt the second field by $\Delta r_b/2$ and $\Delta k_b/2$. Finally, we determine the overlap of the two fields, effectively calculating
\[
C_W^{ex}(\Delta r_b, \Delta k_b; \Delta r_a, \Delta k_a) = \left[ \int T(r_b + r_a + \Delta r_a/2) E(r_a') e^{i\Delta k_a \cdot r_a'/2} d^2r_a' \right] \times \left[ \int T(r_b - r_a/2) E(r_a'') e^{-i\Delta k_a \cdot r_a''/2} d^2r_a'' \right] \times e^{-i\Delta k_b \cdot r_b}.
\]
with $\Delta r_b = \Delta r_b \pm \Delta r_b / 2$. At this point, we wish to distinguish the effects of correlations in the transmission matrix and correlations in the incoming field. For this purpose, we introduce the ambiguity function $\mathcal{A}(\kappa, \rho)$:

$$\mathcal{A}(\kappa, \rho) \equiv \int E(r_a + \rho/2)E^*(r_a - \rho/2)e^{-i\kappa r_a}d^2 r_a.$$  

(B3)

By applying an inverse Fourier transform, we find

$$E(r_a + \rho/2)E^*(r_a - \rho/2) = \frac{1}{(2\pi)^2} \int \mathcal{A}(\kappa, \rho)e^{i\kappa r_a}d^2 \kappa.$$  

(B4)

Next, we identify $r_a' = r_a + \rho/2$ and $r_a'' = r_a - \rho/2$ so that we can substitute this result into equation (B2)

$$C_W^F(\Delta r_b, \Delta k_b; \Delta r_a, \Delta k_a) = \frac{1}{(2\pi)^2} \int \int T(r_b^+, r_a + \rho/2 + \Delta r_a/2) \times$$

$$T^*(r_b^-, r_a - \rho/2 - \Delta r_a/2)e^{i(\Delta k_a r_a - \Delta k_b r_b + \kappa r_a)}A(\kappa, \rho)d^2 r_b d^2 r_a d^2 \kappa d^2 \rho.$$  

(B5)

We can now substitute the definition of $C_W$ in equation (2) to arrive at

$$C_W^F(\Delta r_b, \Delta k_b; \Delta r_a, \Delta k_a) = \frac{1}{(2\pi)^2} \int \int C_W(\Delta r_b, \Delta k_b; \Delta r_a, \Delta k_a + \kappa)A(\kappa, \rho)d^2 \kappa d^2 \rho.$$  

(B6)

In words, the measured correlation function is the actual correlation function of the sample, convolved with the ambiguity function of the incident field mirrored along both $\kappa$ and $\rho$.

Let us evaluate equation (B6) for a correlation function for a shift/tilt invariant medium, as we found with the Fokker-Planck model. Substituting equation (A9) into equation (B6) gives

$$C_W^F(\Delta r_b, \Delta k_b; \Delta r_a, \Delta k_a) = \int \int C_W^F(\Delta r_b, \Delta k_b)\delta(\Delta r_b - \Delta r_a - \rho - (\Delta k_a + \kappa)L/k_0)\delta(\Delta k_b - \Delta k_a - \kappa)A(\kappa, \rho)d^2 \kappa d^2 \rho =$$

$$C_W^F(\Delta r_b, \Delta k_b)A(\Delta k_b - \Delta k_a, \Delta r_b - \Delta r_a - \Delta k_b L/k_0).$$  

(B7)

Let us now consider some typical experimental scenarios. First of all, for any incident beam with unit power, the ambiguity function will have a maximum at $A(0, 0) = 1$. If the incident beam approximates a plane wave, then $A$ will be sharp in the direction of $\kappa$, and almost constant in the direction of $\rho$. As a result, equation (B7) predicts that we can vary $\Delta r_b$ without observing an effect on the correlation. This result is to be expected since the incident beam is basically translational invariant, shifting it has no effect. This symmetry does not hold for the scattered wave, and thus $C_W^F$ decreases with $\Delta r_b$ in the same way as $C_W^{FP}$ decreases with $\Delta r_b$. Finally, the fact that $A$ is sharp in the direction of $\kappa$ makes that tilt/tilt correlations are only observed when $\Delta k_b \approx \Delta k_a$, as is well-known from experimental observations of the tilt/tilt optical memory effect.

Second, if we focus the incident light onto the sample with a high numerical aperture, $A$ will be sharp in the direction of $\rho$, and almost constant in the direction of $\kappa$. In this case, the incident beam is tilt invariant, so we expect $C_W^F$ not to remain constant when the incident beam is tilted (within the numerical aperture).

To summarize, we found that the angular correlations for an incident plane wave are only determined by the sample, whereas for a finite-size beam the correlations in the incident beam give rise to observed correlations over a larger angular range. This result is in line with observations by Li and Genack [7].

By creating a highly randomized input field for our correlation measurements, we were able to create a field with an ambiguity function that is sharp both along $\kappa$ and $\rho$, thus minimizing the impact of $A$ in this convolution and obtaining an accurate measurement of $C_W$. Additionally, in our experiment, we normalize the measured correlation coefficient and take the average over all pairs of measurement with the same spacing $\Delta r_a$ and $\Delta k_a$ to average over disorder.
Schematic of the experimental setup used for measuring both the average light field transmission ($P_W$) and generalized correlation functions ($C_W$). The pinhole and diffuser are used in the $P_W$ and $C_W$ experiments, respectively. Both the diffuser and the sample holder are placed on a translation stage.