Translation correlations in anisotropically scattering media

Benjamin Judkewitz[1,2], Roarke Horstmeyer[1], Ivo M. Vellekoop[3] and Changhuei Yang[1]
1: California Institute of Technology, Departments of Electrical and Bioengineering
1200 E California Blvd, Pasadena, CA 91125
2: Exzellenzcluster NeuroCure, Charité Berlin, Humboldt University
Charitéstr. 1, 10117 Berlin, Germany
3: MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente,
P.O. Box 217, 7500 AE Enschede, Netherlands
Correspondence to: benjamin.judkewitz@charite.de

Controlling light propagation across scattering media by wavefront shaping holds great promise for a wide range of applications in biomedical imaging. However finding the right wavefront to shape is a challenge when the scattering transmission matrix is not known. Correlations in transmission matrices, especially the so-called memory-effect, have been exploited to address this limitation. However, the traditional memory-effect applies to thin scattering layers at a distance from the target, which precludes its use within thick scattering media. Here, we report on analogous transmission matrix correlations within thick anisotropically scattering media, with wide-ranging implications for biomedical imaging. We use a simple conceptual framework to explain these findings and relate them to the traditional memory effect.

1. INTRODUCTION

Focusing light through strongly scattering media is an important goal of biomedical optics. Long considered impossible, recent advances in the field of wavefront-shaping [1,2] changed this view in demonstrating that diffuse light can be focused through inhomogeneous media – as long as the correct input wavefront is used. With direct optical access to the target plane, the correct wavefront can be obtained by iterative optimization [2], phase-conjugation [3], or by measuring the transmission matrix [4,5]. In many imaging scenarios, however, there is no direct access to the target plane. In those cases, nonlinear [6], fluorescent [7], acousto-optic [8-10] and photo-acoustic [11-13] guide-stars can be used as reference beacons. However, these techniques only provide wavefront information for one target location at a time. While transmission matrices can be sampled quickly with a photo-acoustic approach [14], this method requires absorbing samples. As a result, many samples' transmission matrices can only be sampled sparsely. Correlations within a transmission matrix can compensate for sparse sampling and could enable high-speed imaging. One of the most widely known transmission matrix correlations is the so-called "memory effect" [15,16], which describes the following phenomenon: when an input wavefront reaching a diffusing sample is tilted within a certain angular range, the output wavefront is equally tilted, resulting in the translation of the far-field speckle pattern at a distance behind the sample (see Fig. 1).

The translation or field-of-view (FOV) within which this effect holds, is inversely proportional to diffuser thickness t and directly proportional to the distance s of the diffuser from the screen. It can be approximated by the equation $FOV \approx s\lambda/\pi t$ [17-20].

The memory effect has found numerous applications for point-scanning [21,22], direct image transfer [18] as well as for computational image recovery [19,20]. Yet, in all of these applications, the target plane was at a distance from the diffuser with free space in between – which has limited use for imaging inside thick scattering media. Because such samples are neither thin nor at a distance from the target area of interest, the correlations predicted by the "traditional" memory effect should be minimal [23]. Here, we set out to ask whether there are other correlations that apply to such samples. We show that significant transmission correlations can exist in thick scattering media at zero distance, as long as scattering is directional.

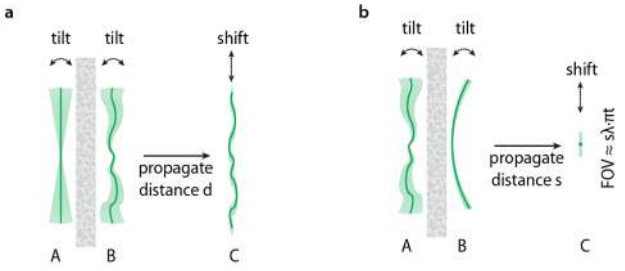

FIG. 1 The traditional memory effect. (a) the traditional memory effect as described for light propagating through thin diffusing slabs. Tilting the input wavefront (plane A) reaching the slab tilts the scattered wavefront at the output (plane B), which shifts the far-field intensity speckle pattern projected on a screen (plane C). (b) when the input wavefront is shaped to converge at a spot on the screen, tilting the input wavefront scans the spot laterally, which can be used for imaging by point scanning. The field-of-view of this approach is approximated by the equation $FOV \approx s\,\lambda / \pi\, t$.

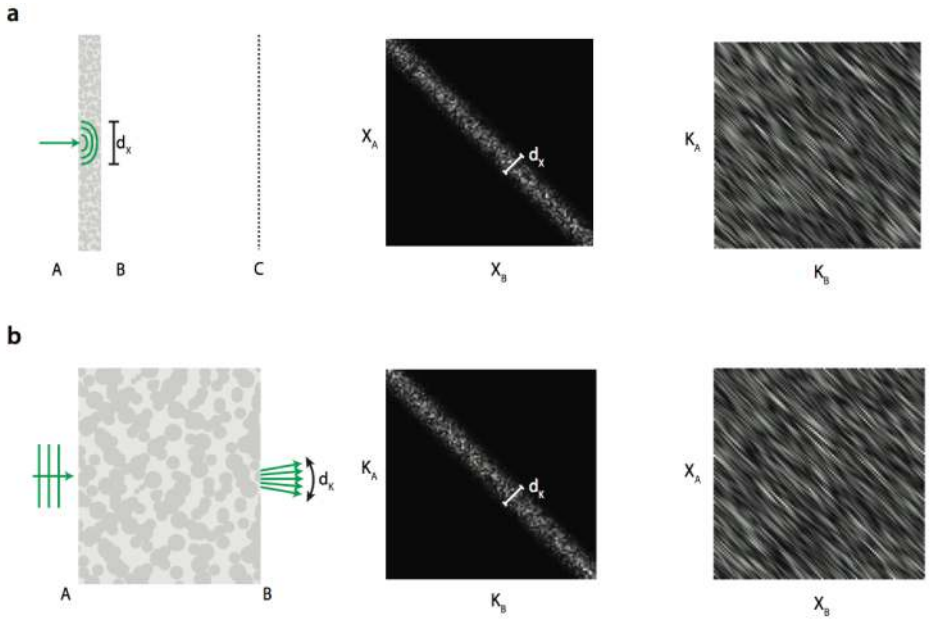

FIG. 2 Correlations within transmission matrices (simulated). (a) the traditional (tilt/tilt) memory effect explained in terms of transmission matrix correlations. A pencil beam illuminating a thin slab, will cause a diffuse spot at the output surface, whose diameter d_X is on the order of the slab thickness. The profile of the spot will be apparent in the (ordered) X/X transmission matrix, and results in strong near-diagonal components. The corresponding K/K transmission matrix is diagonally smeared (since X/X and K/K transmission matrices are related by the 2D Fourier transform). Hence, a tilt (k-shift) at the input plane, results in a corresponding tilt (k-shift) at the output plane. (b) in anisotropically scattering media of finite thickness the directionality of the input light may be preserved. A result, a plane wavefront illuminating the sample results in a limited angular range d_K of output wavefronts. This suggests that the K/K transmission matrix of such samples has elements of higher amplitude near the diagonal. This would results in a diagonally smeared X/X transmission matrix, indicating shift/shift correlations.

2. TRADITIONAL MEMORY EFFECT

While the memory effect has been extensively derived from first principles [16], these derivations relied upon assuming perfectly diffuse scattering – which does not apply to many biological samples. Here we will take a more heuristic approach using amplitude transmission matrices. Specifically, we are interested in the matrix $T_x = T(x_a, x_b)$ which maps input spatial modes to output spatial modes. For simplicity of graphical representation, we assume propagation of 1D wavefronts in a 2D geometry, but all conclusions will be generalizable to 2D wavefronts in a 3D geometry. Due to its discrete nature, the transmission matrix is especially amenable to experimental observation. We will first use our framework to analyze the traditional memory effect, and then calculate the speckle correlations in thick anisotropic media.

When assuming highly randomizing transmission, but incomplete measurement of input and output channels, the transmission matrix is often modeled as a random matrix with complex Gaussian elements. However, in the case of a thin scattering slab, the transmission matrix will have an additional macroscopic structure: a point-source on the input plane of the slab would spread to a bell-shaped diffuse spot at the output plane (whose diameter would be on the order of t for slabs whose thickness t is larger than the transport mean free path – as predicted by the diffusion approximation). As a result, even though individual transmission matrix elements may not be known, a bell-shaped profile will be represented in the envelope of transmission matrix columns (see Figure 2a).

To recognize tilt correlations, we are interested in the effect of this macroscopic structure on the spatial frequency domain (k-space) representation. Every spatial-domain transmission matrix T_x can the transformed into the corresponding frequency-domain transmission matrix $T_k = T(k_a, k_b)$ by the following operation: $T_k = F^{-1}T_xF$, where F is a discrete Fourier transform (DFT) matrix. Due to the Fourier Inversion Theorem, we can express the inverse Fourier transform of T_x as the Fourier transform of T_x, horizontally flipped: $F^{-1}T_XF = (FT_X)^{\leftrightarrow}F = (FT_XF)^{\leftrightarrow}$, where $\leftrightarrow$ denotes the horizontal flip operator. This is analogous to performing the 2D Fourier transform of T_x and flipping it horizontally:

$$T_K = (\mathcal{F}_{2D}T_X)^{\leftrightarrow} \tag{1}$$

Expressing T_k in terms of the 2D Fourier transform of T_x provides a straightforward explanation of how the macroscopic structure of T_x influences correlations in T_k: a macroscopic envelope with strong near-diagonal components in T_x acts as low pass filter along the diagonal of T_k. This convolution of matrix elements along the diagonal of T_k corresponds to the traditional memory effect, in which a tilt of the input wavefront (shift in k-space) causes a tilt of the output wavefront, resulting in a shifted speckle pattern at a distant screen. For this reason, and to distinguish the traditional memory effect from further correlations described below, we also refer to the traditional memory effect as "tilt/tilt correlation".

More exactly, we can define the intensity propagator $P_x(x_a, x_b) \equiv \langle |T_x(x_a, x_b)|^2 \rangle$ and use the cross-correlation theorem to write

$$\mathcal{F}_{2D}^{x_b\to\Delta k_b, x_a\to\Delta k_a} P_x = \sum_{k_a,k_b}\langle T_k(k_a,k_b)T_k^*(k_a-\Delta k_a, k_b-\Delta k_b)\rangle \propto C_k(\Delta k_a,\Delta k_b) \quad (2)$$

with $C(\Delta k_a, \Delta k_b)$ the tilt/tilt correlation function. Since, generally, P_x only depends on the difference between the coordinates, this equation reduces to

$$C_k(\Delta k_a,\Delta k_b) \propto \delta_{\Delta k_a,\Delta k_b}\mathcal{F}^{\Delta x\to\Delta k_b} P_x(\Delta x), \quad (3)$$

which is the well known memory effect [24].

These considerations reconfirm our expectation that the traditional memory effect may be minimal in thick biological media: First, sample thickness would lead to an increased spread of the input light, reducing tilt/tilt correlations between input and output plane. Second, the plane of interest would not be at a distance from the sample, which means that the tilt at the output plane would not translate into a shift at the plane of interest.

3. CORRELATIONS IN ANISOTROPIC MEDIA

We therefore asked whether there may be other types of transmission matrix correlations in thick samples, especially biological media. We started by recognizing that in many samples scattering is anisotropic and occurs primarily in the forward direction. Scattering is particularly anisotropic in biological media, where the anisotropy parameter g (the average cosine of the scattering angle) typically ranges from 0.9 to 0.98 [25,26]. This means that after a limited number of scattering events, the directionality of an input beam will be preserved to some extent as it reaches the output plane. In other words, one input plane wave (one mode in k-space) will result in a limited angular span of output waves.

As a result of such preserved directionality, anisotropically scattering media will have a macroscopic structure in T_k (rather than T_x), with high near-diagonal amplitudes (see Fig. 2b). By analogy to our prior reasoning for the traditional memory effect, a macroscopic envelope in T_k will lead to a diagonally convolved T_x, which means that a shift at the input plane will cause a shift at the output plane (this is in contrast to the traditional memory effect, in which a tilt at the input plane causes a tilt at the output plane).

The range of correlations in T_x will depend on the envelope of the columns in T_k. Equivalent to Eq. 2 we define $P_k(k_a,k_b) \equiv \langle |T_k(k_a,k_b)|^2\rangle$ and find

$$\mathcal{F}_{2D}^{k_b\to\Delta x_b, k_a\to\Delta x_a} P_k = \sum_{x_a,x_b}\langle T_x(x_a,x_b)T_x^*(x_a-\Delta x_a, x_b-\Delta x_b)\rangle \propto C_x(\Delta x_a,\Delta x_b) \quad (4)$$

Here, the k-space propagator P_k includes anisotropy's effects, and reduces to

$$C_x(\Delta x_a,\Delta x_b) \propto \delta_{\Delta x_a,\Delta x_b}\mathcal{F}^{\Delta k\to\Delta x_b} P_k(\Delta k), \quad (5)$$

when only dependent upon difference coordinates. This result is the exact Fourier conjugate of the traditional (tilt/tilt) memory effect. This result also suggests that the correlation function can be predicted by a simple experiment, namely by illuminating the sample with a plane wave $E_p(x_a) = 1$ and measuring the output wavefront $E_p\ (x_b) = T_x E_p(x_a)$. The electric field correlation between the transmitted outputs of two spatially shifted inputs can then be approximated by

$$I_k(k) \equiv \left|\mathcal{F}^{\Delta x_b\to k} E_p(x_b)\right|^2 \quad (6)$$

which, in effect, means that the shape of the shift/shift correlation function equals the autocorrelation of $E_p(x_b)$.

A more detailed derivation of Eq. (1)-(6) may be found in the Supplementary Material

4. EXPERIMENTAL VALIDATION

The shift/shift correlations apply to any input field, including fields that are shaped to converge to a sharp focus. To demonstrate the use of these correlations for scanning a point focus, we first used digital optical phase conjugation [3] to focus light (532 nm DPSS laser) through 500 µm to 1 mm thick slices of chicken muscle tissue (reduced scattering coefficient l: 1-2 mm, anisotropy factor g: ~0.97), employing off-axis holography for wavefront measurement and a spatial-light-modulator (SLM) for wavefront shaping (Fig. 3a) [10,27]. We projected a point-

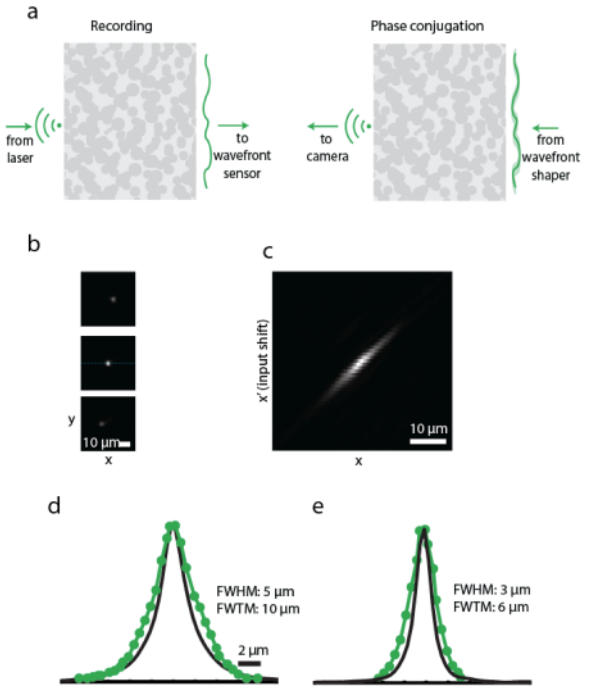

FIG. 3 Measured shift/shift correlations (a) Experimental setup (b) the time-reversed spot (middle) and shifted foci resulting from laterally shifting the phase conjugated wavefront at the sample. (c) line scan (intensity profile) along the blue dotted line in b while shifting the input wavefront, d and e: focus peak intensity as a function of shifted location for 1000/500 µm (d/e) slices. Black curve: prediction based on the speckle autocorrelation measured during plane-wave illumination (Eq. 2).

source at one surface of the tissue slice (surface A) and detected the scattered wavefront propagating from this point through the tissue (exiting at surface B) to the SLM-plane. In the next step, we displayed the phase-conjugate of the detected wavefront, which travelled back through the tissue and formed a focus on surface A.

To validate the predicted shift/shift correlations, we then shifted the phase-conjugated wavefront laterally at surface B, testing whether the focus at surface A would be preserved and whether it moved.

As expected, we noticed that motion of the shaped wavefront resulted in concurrent movement of the focus (Figure 3b-c), while the focus intensity decreased with distance from the original position, following a bell-shaped curve (Figure 3d-e). For the 500 µm slice the full width at half of the maximum (FWHM) was 5 µm, while the full-width at tenth of the maximum (FWTM) was 10 µm. In the case of the 1000 µm slice, the FWHM was 3 and the FWTM was 6 µm.

Next, we illuminated the sample with a plane wavefront and asked whether the shape of the spatial autocorrelation of the resulting amplitude speckle pattern (E_p) followed the profile of the shift/shift correlations (C_x), as derived in Eq 6. Indeed, Figure 3d-e shows that both profiles are in good experimental agreement.

5. DISCUSSION

The traditional (tilt/tilt) memory effect has recently enabled the development of several modalities to image through scattering 'walls' [18-22]. Intriguing as these methods are, they suffer from two limitations: the sample should be thin, and the object should be placed at a distance behind the sample.

Here, we demonstrated a complementary type of memory effect that suffers from neither limitation: the correlations are present even inside thick scattering media, as long as scattering is anisotropic and the mapping between input and output wavefronts preserves any level of directionality. This is the case up to a depth of one transport mean free path.

We showed that the shift/shift memory effect is the Fourier complement analog of the traditional (tilt/tilt) memory effect, and that the extent of correlations can be directly determined from the spatial speckle autocorrelation function during plane-wave illumination.

Our results pave the way for extending memory-effect-based imaging methods [18-22] to also work inside biological tissue. Based on our measurements, we expect such methods to achieve diffraction-limited resolution at a depth of 1 mm inside muscle tissue, albeit at a limited field-of-view of <10 µm initially.

We foresee several possibilities to further increase the field-of-view of our method, including tiling neighboring fields-of-view using multiple corrections. Also, our results suggest that the extent of correlations will be largest for those photons which have undergone few scattering events and little angular deviation – also called snake-photons.

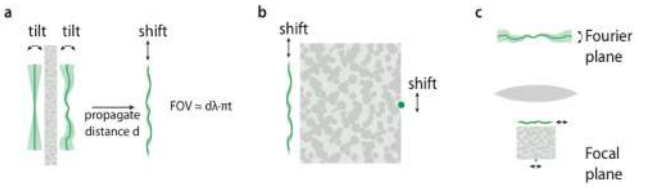

FIG. 4 **Comparison between correlations and relation to adaptive optics** (a) the traditional memory effect, in which a tilt at the input leads to a tilt at the output face of the sample. (b) shift/shift correlations, in which a point can be scanned by shifting the corrected input wavefront. (c) typical setup in adaptive optics microscopy, where corrected wavefronts are tilted in the Fourier-plane. This leads to a shift at the sample surface, analogous to the correlations described here.

Hence, selective measurement and correction of snake-photons (e.g. by temporal gating, coherence gating or spatial filtering) may considerably increase the extent of correlations and the imaging field-of-view.

Finally, we note that tilt/tilt and shift/shift correlations are not mutually exclusive. For instance, we assumed that the average T_x matrix is translationally invariant, which is a pessimistic assumption even in thick biological tissue. Furthermore, we anticipate that there may be additional correlations present in biological media. Future work measuring complete transmission matrices in the adaptive optics and the complex wavefront shaping regime will shed light on spectral, temporal and spatial correlations. They may ultimately be utilized in combination with the shift/shift correlations reported here.

We note that the described shift/shift correlations are consistent with the setup geometry of adaptive optics microscopy (Fig 4), where wavefronts are corrected in the conjugate Fourier plane of the microscope objective. Tilting the incoming wavefront in the Fourier plane (e.g. in a laser scanning microscope) leads to a shift of the wavefront reaching the sample with a resulting shift of the focus [28]. In other words, adaptive optics microscopy implicitly already takes advantage of shift/shift correlations, albeit in the ballistic regime – as such it can be interpreted as a special case of the general shift/shift correlations derived here.

With further study of spatial spectral and temporal transmission matrix correlations, these advances may lead to a unified understanding of adaptive optics and complex wavefront shaping and extend their use in thick biological tissues, enabling versatile imaging and photostimulation in a wide range of biologically relevant media.

We thank Robert Chen for providing very helpful feedback on this manuscript. This work was supported by NIH 1DP2OD007307-01 and the Wellcome Trust (B.J.).

B.J. and R.H. contributed equally to this work.

[1] A. P. Mosk, A. Lagendijk, G. Lerosey, and M. Fink, Nature Photonics **6**, 283 (2012).
[2] I. M. Vellekoop and A. P. Mosk, Opt Lett **32**, 2309 (2007).
[3] Z. Yaqoob, D. Psaltis, M. S. Feld, and C. Yang, Nature Photonics **2**, 110 (2008).
[4] S. Popoff, G. Lerosey, M. Fink, A. C. Boccara, and S. Gigan,

Nat Commun **1**, 81 (2010).
[5] W. Choi, A. P. Mosk, Q.-H. Park, and W. Choi, Phys. Rev. B **83**, (2011).
[6] C.-L. Hsieh, Y. Pu, R. Grange, and D. Psaltis, Opt Express **18**, 12283 (2010).
[7] X. Tao, J. Crest, S. Kotadia, O. Azucena, D. C. Chen, W. Sullivan, and J. Kubby, Opt Express **20**, 15969 (2012).
[8] X. Xu, H. Liu, and L. V. Wang, Nature Photonics **5**, 154 (2011).
[9] Y. M. Wang, B. Judkewitz, C. A. DiMarzio, and C. Yang, Nat Commun **3**, (2012).
[10] B. Judkewitz, Y. M. Wang, R. Horstmeyer, A. Mathy, and C. Yang, Nature Photonics **7**, 300 (2013).
[11] F. Kong, R. H. Silverman, L. Liu, P. V. Chitnis, K. K. Lee, and Y. C. Chen, Opt Lett **36**, 2053 (2011).
[12] D. B. Conkey, A. M. Caravaca-Aguirre, J. D. Dove, H. Ju, T. W. Murray, and R. Piestun, arXiv Preprint arXiv:1310.5736 (2013).
[13] P. Lai, L. Wang, J. W. Tay, and L. V. Wang, arXiv (2014).
[14] T. Chaigne, O. Katz, A. C. Boccara, M. Fink, E. Bossy, and S. Gigan, Nature Photonics **8**, 58 (2014).
[15] I. Freund, M. Rosenbluh, and S. Feng, Phys Rev Lett **61**, 2328 (1988).
[16] S. Feng, C. Kane, P. A. Lee, and A. D. Stone, Phys Rev Lett **61**, 834 (1988).
[17] I. Freund, Physica a: Statistical Mechanics and Its Applications **168**, 49 (1990).
[18] O. Katz, E. Small, and Y. Silberberg, Nature Photonics **6**, 549 (2012).
[19] J. Bertolotti, E. G. van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, Nature **491**, 232 (2012).
[20] O. Katz, P. Heidmann, M. Fink, and S. Gigan, Nature Photonics (2014).
[21] C.-L. Hsieh, Y. Pu, R. Grange, G. Laporte, and D. Psaltis, Opt Express **18**, 20723 (2010).
[22] I. M. Vellekoop and C. M. Aegerter, Opt Lett **35**, 1245 (2010).
[23] D. Psaltis and I. N. Papadopoulos, Nature **491**, 197 (2012).
[24] J. H. Li and A. Z. Genack, Phys. Rev. E **49**, 4530 (1994).
[25] S. L. Jacques, Phys Med Biol **58**, R37 (2013).
[26] W.-F. Cheong, S. A. Prahl, and A. J. Welch, Quantum Electronics, IEEE Journal of DOI - 10.1109/JQE.1986.1073146 **26**, 2166 (1990).
[27] Y. M. Wang, 1 (2013).
[28] N. Ji, T. R. Sato, and E. Betzig, Proc Natl Acad Sci USA **109**, 22 (2012).

Supplementary material for: Translation correlations in anisotropically scattering media

Appendix A: The traditional memory effect as a diagonal envelope

Here, we connect our experimentally driven explanation of the "traditional" memory effect to previous theoretical derivations. By "traditional" memory effect, we refer to the commonly observed [Feng et al. 1988, Li and Genack 1994, Bertolotti et al. 2012] distance over which a speckle pattern can shift on a screen placed a distance behind a scatterer, when light incident upon the scatterer is tilted in angle. Our goal is to show that under appropriate conditions, this correlative effect is linked to a diagonally enveloped real space transmission matrix.

To begin, we will first define an intensity transmission matrix in k-space suitable for our scenario of a widely-illuminated scattering slab (K/K matrix, Fig. 2(a) middle). We will then transform this matrix into a matrix containing the complex field transmission coefficients in real space (X/X matrix, Fig. 2(a) right). This transformation will reveal a diagonal multiplicative envelope whose width is defined by the average intensity distribution across the scatterer's output (back) surface caused by a point source placed at the input (front) surface.

We assume a scattering slab containing disordered material of thickness L, width W (on both input and output surface, and scattering mean free path l. Following prior work [Feng et al. 1988], we assume a linear optical medium $kl \gg 1$ (weak disorder) and $L \gg l$ (multiple scattering). These assumptions are not necessary to show the memory effect (either traditional or anisotropic) exists in general, but help connect our approach to this prior work. A more rigorous derivation performed without such assumptions is in Appendix C. We start by describing this slab with a transmission matrix $\boldsymbol{T_I}(k_a, k_b)$, containing the intensity transmission coefficients of wavevector k_a on its input surface scattering into wavevector k_b on its output surface, under illumination of coherent light of wavelength λ. The "resolution" of each matrix element is set at $\delta k_a = \delta k_b = 2NA/\lambda$, where the numerical aperture NA is a function of the maximum acceptance angle of light on the input surface, typically set by the illumination/detection optics. For example, in our experiment we have the approximate values $\lambda = 532$ nm, $l \approx 1 - 2$ mm, $L = 0.5$ mm, NA = 0.75.

The correlation of this matrix $\boldsymbol{T_I}$ with itself, considering many different random scatterer configurations, takes the form,

$$C_{k_a,k_b,k_a',k_b'} = \langle (\boldsymbol{T_I}(k_a, k_b) - \mu_I)(\boldsymbol{T_I}(k_a', k_b') - \mu_I') \rangle, \qquad \text{(A1)}$$

where $\mu_I = \langle \boldsymbol{T_I}(k_a, k_b) \rangle$, the average coefficient value. Prior work expresses Eq. (A1) as three unique correlations [Feng et al. 1988, Li and Genack 1994]:

$$C_{k_a,k_b,k_a',k_b'} = C^{(1)}{}_{k_a,k_b,k_a',k_b'} + C^{(2)}{}_{k_a,k_b,k_a',k_b'} + C^{(3)}{}_{k_a,k_b,k_a',k_b'} \quad \text{(A2)}$$

The first "memory effect" term, $C^{(1)}$, is the dominant term and traditionally takes the form,

$$C^{(1)}{}_{k_a,k_b,k_a',k_b'} = C^{(I)}{}_{\Delta k_a,\Delta k_b} = A^2 \delta_{\Delta k_a,\Delta k_b} \left(\frac{\Delta k_b L}{\sinh \Delta k_b L}\right)^2, \quad \text{(A3)}$$

where $\Delta k_a = k_a - k_a'$ and $\Delta k_b = k_b - k_b'$ are shifts in input and output wavevector, respectively, the δ-function indicates that the correlation is not present unless both shifts are examined simultaneously, and A^2 is a constant. Eq. (A3)'s correlations define $\boldsymbol{T_I}$'s diagonal "smearing" as depicted in Fig. 2(a), middle. Its δ-function holds valid so long as the beam exiting the scatterer remains wide ($I'(x_b) \gg L$), which is satisfied in our current consideration of plane-wave input illumination. Given a narrow illumination beam, we may replace Eq. (A3)'s delta function with a finite-width function $|g(\Delta k_a - \Delta k_b)|^2$, which may be interpreted as finite-width speckle (in k-space) at the input and output surface [Li and Genack 1994, Bertolotti et al. 2013]. Sampling conditions may easily be redefined to reduce $g(\cdot)$ to a discrete δ-function, thus removing off-diagonal matrix correlations from the following analysis.

Eq. (A3)'s $\sinh(\cdot)$ fraction originates from the influence of an intensity propagation function $P(x_b - x_a)$ connecting the spatial coordinates of the input surface x_a to the output surface x_b [Berkovitz et al. 1989]. Prior work [Li and Genack 1994] suggests a more general form of Eq. (A3) based upon the intensity propagator P:

$$C^{(I)}{}_{\Delta k_a,\Delta k_b} = A^2 \delta_{\Delta k_a,\Delta k_b} \left|\mathcal{F}^{x_b \to \Delta k_b}[P(x_b)]\right|^2, \quad \text{(A4)}$$

where $\mathcal{F}$ is a one-dimensional Fourier transform and $P(x)$ connects the average intensity distribution on the output surface $I(x_b)$ to the input surface $I(x_a)$: $I(x_b) = \sum_{x_a} I(x_a)\, P(x_b - x_a)$. We may find $P(x)$ by measuring or deriving the average output surface intensity envelope $I'(x_b)$ due to input illumination from a discrete spatial δ -function: $I'(x_b) = P(x_b)$ given $I(x_a) \approx \delta(0)$. Following our matrix sampling conditions, an input illumination spot of approximate size $\lambda/2NA$ is sufficiently narrow to satisfy this discrete delta criterion, as Fig. 2(b) illustrates. Likewise, its relative position is not critical – our assumption of a spatially homogenous slab requires a shifted input spot $I(x_a) \approx \delta(d)$ to create a shifted yet otherwise unchanged output $I'(x_b - d)$. $P(x)$ and this measured or derived $I'(x_b)$ are thus easily interchanged.

Next, we transform Eq. (A4) from an intensity correlation into a similar function for a complex optical field. Following [Goodman 2007, Pnini and Shapiro 1989], we recognize that the correlation function of a circularly complex Gaussian random process $C^{(E)}$ is connected to the correlation function of the same random process's intensity $C^{(I)}$ through a simple square root relationship:

$$C^{(E)}{}_{k_a,k_b,k_a',k_b'} = \langle \boldsymbol{T_E}(k_a, k_b) \boldsymbol{T_E^*}(k_a', k_b') \rangle = \sqrt{C^{(I)}{}_{k_a,k_b,k_a',k_b'}} \quad \text{(A5)}$$

Here, the ensemble average is over many random complex transmission matrices $\boldsymbol{T_E}(k_a, k_b)$, we've used $\mu_E = \langle \boldsymbol{T_E}(k_a, k_b) \rangle = 0$ for a zero-mean complex speckle process, and we require $C^{(E)}$ to be real and positive. The symbol E indicates $\boldsymbol{T_E}$ and $C^{(E)}$ apply to the complex field. The main text replaces E with subscript x or k for both matrices for notational clarity. Inserting Eq. (A4) into Eq. (A5) and replacing $P(x_b)$ with $I'(x_b)$ gives an expression for the field transmission matrix $\boldsymbol{T_E}$'s correlation:

$$C^{(E)}{}_{\Delta k_a, \Delta k_b} = A\delta_{\Delta k_a, \Delta k_b} \mathcal{F}^{x_b \to \Delta k_b}[I'(x_b)], \tag{A6}$$

which mirrors Eq. (3) in the main text. Our final step is to show that Eq. (A6)'s memory effect correlation, when expressed in the spatial domain, appears as a diagonally enveloped transmission matrix $\boldsymbol{T_E}(x_a, x_b)$. Since the speckle field is a spatially stationary random process (i.e., whose statistics do not vary with position), we may rewrite Eq. (A5) as a discrete autocorrelation:

$$C^{(E)}{}_{\Delta k_a, \Delta k_b} = \sum_{k_a} \sum_{k_b} \boldsymbol{T_E}(k_a, k_b) \boldsymbol{T_E^*}(k_a - \Delta k_a, k_b - \Delta k_b)\,. \tag{A7}$$

We may then use a Fourier transform to examine Eq. (A7) in the spatial domain. From the Weiner-Khinchen theorem, we know that a function's autocorrelation and its power spectral density (PSD) are Fourier pairs. We may thus write the inverse Fourier transform $\hat{\mathcal{F}}$ of the two-dimensional autocorrelation $C^{(E)}{}_{\Delta k_a, \Delta k_b}$ as,

$$\hat{\mathcal{F}}_{2D}^{\Delta k_a \to x_a, \Delta k_b \to x_b}\left[C^{(E)}{}_{\Delta k_a, \Delta k_b}\right] = \langle |\boldsymbol{T_E}(x_a, x_b)|^2 \rangle, \tag{A8}$$

where $\boldsymbol{T_E}(x_a, x_b)$ is the spatial (XX) optical field transmission matrix described as a 2D Fourier transform of the KK transmission matrix in Eq. (1): $\boldsymbol{T_E}(x_a, x_b) = \mathcal{F}_{2D}[\boldsymbol{T_E}(k_a, k_b)]$. The PSD's ensemble average is also taken over many possible scatterer configurations. Applying the same inverse two-dimensional Fourier transform along $(\Delta k_a, \Delta k_b)$ to the right side of Eq. (A6) leads to,

$$\hat{\mathcal{F}}_{2D}^{\Delta k_a \to x_a, \Delta k_b \to x_b}\left[A\delta_{\Delta k_a, \Delta k_b} \mathcal{F}^{x_b \to \Delta k_b}[I'(x_b)].\right] = AI'(x_b - x_a). \tag{A9}$$

Here we again use the modified 2D Fourier transform relationship from Eq. (1) and change coordinates from x_b to $x_b - x_a$ via the δ-function. After constructing a Toeplitz mask matrix $\boldsymbol{M}(x_a, x_b) = I'(x_b - x_a)$ for all (x_a, x_b), we may equate Eq. (A8) and Eq. (A9) to find,

$$\langle |\boldsymbol{T_E}(x_a, x_b)|^2 \rangle = \sqrt{A}\, \boldsymbol{M}(x_a, x_b). \tag{A10}$$

In summary, a K/K transmission matrix $\boldsymbol{T_E}(k_a, k_b)$ exhibiting the $C^{(1)}$ memory effect reduces to a Toeplitz diagonal masking matrix when examined in the spatial domain. This masking effect on a single XX matrix is depicted in Fig. 2(a) left, and is valid given a relatively wide input illumination beam ($I'(x_b) \gg L$). The diagonal mask's shape is set by the average intensity distribution on the output surface of the scatterer, $I'(x_b)$, from a point source placed on the input surface. Again, we note that all above assumptions (except the requirement of a linear scattering process) are not necessary to show a finite memory effect correlation exists for any optical scattering material. The same holds for the anisotropic memory effect, which we explicitly show in Appendix C.

Appendix B: The anisotropic memory effect

Building upon our model in Appendix A, we now derive our new anisotropic scattering –based correlation function. Here, we will again connect our derivation to as much prior work as possible, which will necessitate several assumptions. Appendix C offers a stand-alone derivation of the anisotropic memory effect that does not require any of the assumptions noted below, save a linear scattering process.

This "anisotropic" memory effect manifests itself as correlations within the spatial (X/X) transmission matrix, much the same way as the "traditional" memory effect influences the k-space (K/K) transmission matrix. Specifically, certain properties of a scattering slab may cause a spatially invariant, finite angular spread of wavevectors to emerge from its output surface when we illuminate it with a plane wave. We first connect this finite wavevector spread to a function modulating the anti-diagonal of this scattering slab's average K/K transmission matrix (see sketch in Fig. 2(b), middle). Then, we transform this modulated K/K transmission matrix into an X/X transmission matrix to reveal our new anisotropic correlation along its diagonal (Fig. 2(b), right).

We begin by assuming a scattering slab with similar properties as in Appendix A under near plane-wave illumination, but with two important modifications. First, we now assume the average intensity spread at the output surface due to a point source extends nearly the full surface width ($I'(x_b)$~W), leading to a negligible traditional memory effect:

$$C_{k_a,k_b,k_a',k_b'} \approx C^{(I)}{}_{\Delta k_a,\Delta k_b} = A^2 \delta_{\Delta k_a,\Delta k_b} \left|\mathcal{F}^{x_b \to \Delta k_b}[I'(x_b)]\right|^2 \approx A^2 \delta_{\Delta k_a} \delta_{\Delta k_b} \qquad \text{(B1)}$$

As in Eq. (A3), we again neglect the higher-order $C^{(2)}$ and $C^{(3)}$ correlations and assume a sufficiently wide $I'(x_b)$ to ensure the above Fourier transform reduces to a discrete delta function (within our $2NA/\lambda$ sampling rate). $C^{(1)}$ correlations may be added back into $\boldsymbol{T_I}(k_a, k_b)$ without significantly impacting the following results, but lead to a more complex analysis.

Second, we assume the scattering process within the slab is anisotropic with parameter $g > 0$ (i.e., primarily forward-scattering) and the directionality of multiply scattered photons is sufficiently preserved, such that when the input surface is illuminated with a plane wave $E'(x_a) = e^{ikx_a\theta}$ across its spatial extent W, only a finite spread of wavevectors appears at the output surface. Defining the field at the output surface as $E'(x_B)$ from this plane wave input, we may find its spatial frequency power spectrum via a Fourier transform: $|\mathcal{F}[E'(x_B)]|^2 = \left|\hat{E}'(k_B)\right|^2 = I'(k_B)$. This spectrum's ensemble average over many random possible anisotropic scatterer configurations will follow a band-limited curve:

$$\langle I'(k_B)\rangle = P_k(k_B). \qquad \text{(B2)}$$

$P_k(k_B)$ describes an average radiance distribution, here as a function of wavevector, emerging from the scatterer output surface when illuminated with a plane wave. In

practice, $P_k(k_B)$ will be defined by the scattering slab's anisotropy g, mean-free path l and thickness L (ignoring boundary effects). $P_k(k_B)$'s envelope width will decrease as g varies from 0 to 1, but may be determined for any anisotropy level and in general any linear scattering medium either experimentally (via direct intensity measurement), through simulation (e.g., a Monte Carlo model) or theoretically (e.g. by the diffusion approximation).

When the input plane wave $E'(x_a)$ is traveling at $\theta = 0$, $P_k(k_B)$ will describe a single column of the ensemble-averaged K/K intensity transmission matrix:

$$\langle \boldsymbol{T_I}(0, k_b) \rangle = P_k(k_b) \tag{B3}$$

As the illumination angle is tilted, the average distribution of radiance emerging from the output surface of a homogeneous scattering slab will rotate unchanged (ignoring refractive effects at extreme angles). In other words, if we tilt the input plane wave $E'(x_a)$ by angle θ, we expect the average radiance spectrum on the output surface will tilt to form $\langle I'(k_B - \theta) \rangle = P_k(k_B - \theta)$. Thus, we may fill each column of the average K/K intensity transmission matrix with a shifted version of the average radiance distribution P_k:

$$\langle \boldsymbol{T_I}(k_a, k_b) \rangle = \langle |\boldsymbol{T_E}(k_a, k_b)|^2 \rangle = P_k(k_b - k_a). \tag{B4}$$

Here, we've also re-expressed the intensity transmission matrix as the square of the field matrix, $\boldsymbol{T_E}(k_a, k_b)$. Eq. (B4) defines the average anisotropic KK transmission matrix with an envelope P_k along its anti-diagonal, as illustrated by Fig. 2(b), middle. (Note P_k's uniform angular shifting is in no way related to the traditional memory effect, but is instead directly analogous to Appendix A's assumption that intensity envelope $I'(x_B)$ is independent of spatial location).

Following Appendix A's derivation in reverse, we can inverse Fourier transform both sides of Eq. (B4) to obtain our desired "anisotropic memory effect" correlations in the spatial domain (i.e., within the XX transmission matrix $\boldsymbol{T_E}(x_a, x_b)$). We first apply a two-dimensional Fourier transform to $P_k(k_b - k_a)$ on Eq. (B4)'s right side to obtain,

$$\begin{aligned} \mathcal{F}_{2D}^{k_a \to \Delta x_a, k_b \to \Delta x_b}[P_k(k_b - k_a)] &= \mathcal{F}^{k_a \to \Delta x_a}\left[P_k(\Delta x_b) e^{i k_a \Delta x_b}\right] \\ &= \hat{P}_k(\Delta x_b)\delta(\Delta x_b - \Delta x_a) \end{aligned} \tag{B5}$$

where $\hat{P}_k$ is the one-dimensional Fourier transform of P_k. In the first step we've used the Fourier shift theorem and in the second we've converted the multiplication into a convolution, as required, using the same directional flip from Eq. (1).

Finally, we perform a two-dimensional Fourier transform on Eq. (B4)'s left side and apply the Wiener-Khinchin theorem to our stationary speckle process to find,

$$\mathcal{F}_{2D}^{\Delta k_a \to x_a, \Delta k_b \to x_b}\left[\langle |\boldsymbol{T_E}(k_a, k_b)|^2 \rangle\right] = \langle \boldsymbol{T_E}(x_a, x_b) \boldsymbol{T_E^*}(x_a', x_b') \rangle, \tag{B6}$$

where the ensemble average is over many random configurations of the same anisotropic scatterers. Combining Eq. (B5) with Eq. (B6) and renaming $\langle \boldsymbol{T_E}(x_a, x_b) \boldsymbol{T_E^*}(x_a', x_b') \rangle$ as the complex field correlation function $C^{(E)}$ gives,

$$C^{(E)}{}_{\Delta x_a,\Delta x_b} = \delta(\Delta x_b - \Delta x_a)\hat{P}_k(\Delta x_b). \tag{B7}$$

Eq. (B7) is analogous to Eq. (A4) but with a correlation that is now a function in the spatial instead of angular domain and now depends upon a finite output wavevector spread instead of a finite spatial intensity spread. Correspondingly, P_k plays the Fourier dual role of the propagator function P from Appendix A, which becomes significant enough to induce correlations when anisotropy is high. We note the same field-intensity squaring relationship for correlation from Appendix A will yield,

$$C^{(I)}{}_{\Delta x_a,\Delta x_b} = \delta_{\Delta x_a,\Delta x_b}\left|\hat{P}_k(\Delta x_b)\right|^2, \tag{B8}$$

which shares direct parallels with Eq. (A3)'s "traditional" memory effect, $C^{(I)}{}_{\Delta k_a,\Delta k_b}$. By following Appendix A's steps, we have tried to explicitly show our new "anisotropic" memory effect is indeed the "traditional" memory effect's Fourier dual. We next use an alternative approach, requiring much fewer assumptions, to construct a more general anisotropy-based correlation function.

Appendix C: Rigorous derivation of the anisotropic memory effect

Here we present a rigorous derivation of the anisotropic memory effect. We prove that the anisotropic memory effect is present for all scattering objects, including ordered or absorbing ones, as long as 1) light propagation is linear, and 2) the directionality of the incident beam is maintained to any extent (the angular intensity propagator is not constant across angles).

Wave propagation through any linear medium can be described by a field transmission matrix $\boldsymbol{T_E}(x_a, x_b)$ such that,

$$E(x_b) = \int \boldsymbol{T_E}(x_a, x_b)\, E(x_a)\mathrm{d}^2x_a. \tag{C1}$$

Here, $E(x_a)$ is the incident field, $E(x_b)$ is the transmitted field, and x_a and x_b are two-dimensional coordinates in arbitrary planes in front of and behind the sample, respectively.

We now proceed to define the angular intensity propagator, $P_k(k_b, k_a)$. The angular intensity propagator gives the transmitted intensity $I(k_b)$ when the scatterer is illuminated by an incident plane wave with unit power. By convention, we define the intensity as a function of propagation direction k_b as, $I(k_b) \equiv |E(k_b)|^2$.

First, we construct a truncated plane wave with unit power:

$$E'(x_a) = \frac{H_A(x_a)}{\sqrt{A}}\, e^{ik_a \cdot x_a}, \quad H_A(x_a) \equiv \begin{cases} 1 & \text{for } x_a \text{ inside area } A \\ 0 & \text{otherwise} \end{cases} \tag{C2}$$

Here, we have $\int_A |E'(x_a)|^2 \mathrm{d}^2x_a = 1$. Note that while Eq. (C2)'s wave is truncated to a square area A, we will take the limit of $A \to \infty$ below. Following Eq. (C1), the incident field $E'(x_a)$ will generate the following transmitted field:

$$E'(x_b) = \frac{1}{\sqrt{A}} \int \boldsymbol{T_E}(x_a, x_b) H_A(x_a) e^{ik_a \cdot x_a} \, \mathrm{d}^2 x_a. \tag{C3}$$

The resulting angular intensity propagator is thus defined as,

$$\begin{aligned} P_k(k_b, k_a) &\equiv \left| \int e^{-ik_b \cdot x_b} \, E'(x_b) \, \mathrm{d}^2 x_b \right|^2 \\ &= \frac{1}{A} \iint e^{ik_b \cdot x_b} \, \boldsymbol{T_E}(x_a, x_b) \, H_A(x_a) e^{-ik_a \cdot x_a} \mathrm{d}^2 x_a \mathrm{d}^2 x_b \\ &\times \iint e^{-ik_b \cdot x_{b'}} \, \boldsymbol{T_E}^*(x_{a'}, x_{b'}) \, H_A(x_{a'}) e^{ik_a \cdot x_{a'}} \, \mathrm{d}^2 x_{a'} \mathrm{d}^2 x_{b'} \end{aligned} \tag{C4}$$

Next, we apply the coordinate transform $x_{a'} \to x_a - \Delta x_a$, $x_{b'} \to x_b - \Delta x_b$ and reorder Eq. (C4) to find,

$$\begin{aligned} P_k(k_b, k_a) = \frac{1}{A} \iiiint & \boldsymbol{T_E}(x_a, x_b) \boldsymbol{T_E}^*(x_a - \Delta x_a, x_b - \Delta x_b) H_A(x_a) \, H_A(x_a - \Delta x_a) \\ & \times e^{-ik_a \cdot \Delta x_a} e^{ik_b \cdot \Delta x_b} \, \mathrm{d}^2 x_a \, \mathrm{d}^2 x_b \, \mathrm{d}^2 \Delta x_a \, \mathrm{d}^2 \Delta x_b. \end{aligned} \tag{C5}$$

We now define our shift-shift correlation function C_x in the limit $A \to \infty$:

$$\begin{aligned} C_x(\Delta x_a, \Delta x_b) &\equiv \lim_{A \to \infty} \frac{1}{A} \iint H_A(x_a) H_A(x_a - \Delta x_a) \\ &\qquad \times \boldsymbol{T_E}(x_a, x_b) \, \boldsymbol{T_E}^*(x_a - \Delta x_a, x_b - \Delta x_b) \, \mathrm{d}^2 x_a \, \mathrm{d}^2 x_b, \\ &= \lim_{A \to \infty} \frac{1}{A} \int_A \int_A \boldsymbol{T_E}(x_a, x_b) \, \boldsymbol{T_E}^*(x_a - \Delta x_a, x_b - \Delta x_b) \, \mathrm{d}^2 x_a \, \mathrm{d}^2 x_b. \end{aligned} \tag{C6}$$

Note that Eq. (C6)'s correlation function can always be defined. However, defining it only makes sense when the medium is statistically invariant to translations over the area of interest. When the medium is ergodic, spatial averaging can be replaced by ensemble averaging, although this is not a requirement for the current derivation. We may now insert Eq. (C6) into Eq. (C5) to find,

$$P_k(k_b, k_a) = \iint C_x(\Delta x_a, \Delta x_b) \, e^{ik_b \cdot \Delta x_b} e^{-ik_a \cdot \Delta x_a} \, \mathrm{d}^2 \Delta x_a \, \mathrm{d}^2 \Delta x_b. \tag{C7}$$

Inverting this Fourier transform gives,

$$C_x(\Delta x_a, \Delta x_b) = \iint P_k(k_b, k_a) \, e^{-ik_b \cdot \Delta x_b} e^{ik_a \cdot \Delta x_a} \, \mathrm{d}^2 k_a \, \mathrm{d}^2 k_b, \tag{C8}$$

which is the continuous equivalent to Eq. (4) in the main manuscript and connects to Eq. (B5)-(B7) above. For convenience, we can change the coordinates of P_k to $k \equiv k_a$ and $\Delta k \equiv k_b - k_a$ and write,

$$C_x(\Delta x_a, \Delta x_b) = \iint P'_k(k; \Delta k) \, e^{-i\Delta k \cdot \Delta x_b} e^{ik \cdot (\Delta x_a - \Delta x_b)} \mathrm{d}^2 k \, \mathrm{d}^2 \Delta k, \tag{C9}$$

which is still a general expression. Note that in most circumstances $C_x(\Delta x_a, \Delta x_b) \neq 0$, even for $\Delta x_a \neq \Delta x_b$.

Interpretation

In its most general form, Eq. (C9)'s shift-shift correlation function depends on both Δx_a and Δx_b. If we want to calculate the magnitude of the anisotropic memory effect, we

evaluate the correlation function at $\Delta x_a = \Delta x_b = \Delta x$ to find that it is simply the Fourier transform of an angle-averaged angular intensity propagator:

$$C_x(\Delta x) = \int \overline{P_k}(\Delta k)\, e^{-i\Delta k \cdot \Delta x}\, \mathrm{d}^2\Delta k. \tag{C10}$$

Here, $\overline{P_k}$ is an angle-averaged angular intensity propagator that we define as, $\overline{P_k}(\Delta k) \equiv \int P_k'(k;\Delta k)\, \mathrm{d}^2 k$. The width of $\overline{P_k}(\Delta k)$ depends on the scattering object's mean free path, anisotropy coefficient, thickness, reflections at the object surface, and possibly other parameters. If a disordered scattering object's width L is much thicker than the transport mean free path for light l, all directionality is lost and $\overline{P_k}(\Delta k)$ will be a constant function across Δk. In this case, the width of $\overline{P_k}(\Delta k)$ will still be limited to $2k_0$, with $k_0 \equiv 2\pi/\lambda$. This finite support of $\overline{P_k}$ gives rise to trivial correlations on the scale of half a wavelength.

If a disordered scattering object's width L is thinner than one transport mean free path l, any level of directionality is preserved and $\overline{P_k}(\Delta k)$ will be narrower than for case when $L \gg l$. The narrower the angular intensity propagator's support, the larger the distance over which the sample can be translated while preserving correlations in the speckle at the scatterer's back surface.

Special case: P_k is separable

If Eq. (C9)'s intensity propagator $P_k(k;\Delta k)$ is separable into $P_k^0(k)$ and $P_k^\Delta(\Delta k)$, we may rewrite it in the form,

$$C_x(\Delta x_a, \Delta x_b) = \int P_k^\Delta(\Delta k)\, e^{-i\Delta k \cdot \Delta x_b} \mathrm{d}^2\Delta k \int P_k^0(k) e^{ik\cdot(\Delta x_b - \Delta x_a)}\, \mathrm{d}^2 k, \tag{C11}$$

which is the product of the Fourier transform of P_k^Δ with the inverse Fourier transform of P_k^0. In the special case where P_k only depends on the difference between the incident and transmitted wave angle Δk, we find that P_k^0 is constant. Under this condition, Eq. (C11) simplifies to,

$$C_x(\Delta x_a, \Delta x_b) = \delta(\Delta x_b - \Delta x_a) \int P_k(\Delta k) e^{-i\Delta k \Delta x_b}\, \mathrm{d}^2\Delta k. \tag{C12}$$

In Eq. (C12)'s final form, it is immediately clear that the anisotropic memory effect is the real-space analogy to the traditional memory effect.

Conclusion

The anisotropic memory effect exists *for any linear medium*, as long as any level of directionality is maintained. The effect of shifting the medium is always robustly defined as the Fourier transform of the angular spread function $\overline{P_k}(\Delta k)$. In the C_1 approximation, the intensity-intensity correlation function can simply be obtained by taking the absolute square of the field-field correlation function C_x defined above.