Bootstrapping cascaded random matrix models: Correlations in permutations of matrix products

Niall Byrnes [1], Gary R. W. Greaves [2], and Matthew R. Foreman [1,3,*]

[1]*School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore*
[2]*School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371, Singapore*
[3]*Institute for Digital Molecular Analytics and Science, 59 Nanyang Drive, Singapore 636921, Singapore*

Random matrix theory is a useful tool in the study of the physics of multiple scattering systems, often striking a balance between computation speed and physical rigour. Propagation of waves through thick disordered media, as arises, for example, in optical scattering or electron transport, typically necessitates cascading of multiple random matrices drawn from an underlying ensemble for thin media, greatly increasing the computational burden. Here we propose a dual pool based bootstrapping approach to speed up statistical studies of scattering in thick random media. We examine how potential matrix reuse in a pool based approach can impact statistical estimates of population averages. Specifically, we discuss how both bias and additional variance in the sample mean estimator are introduced through bootstrapping. In the diffusive scattering regime, the extra estimator variance is shown to originate from samples in which cascaded transfer matrices are permuted matrix products. Through analysis of the combinatorics and cycle structure of permutations we quantify the resulting correlations. Proofs of several analytic formulas enumerating the frequency with which correlations of different strengths occur are derived. Extension to the ballistic regime is briefly considered.

I. INTRODUCTION

Computational modeling of wave scattering in random disordered media is a difficult problem that has been researched in earnest in recent decades. A full finite element based solution of the relevant wave equation offers the greatest rigour but consequently suffers from limited simulation volumes and large computational cost [1]. Alternatively, more approximate methods, such as diffusion-, Monte Carlo-, or Green's function-based approaches, balance computation time with physical rigour to varying degrees [2–5]. Random matrix approaches in particular typically sacrifice system-specific details in favor of significantly reduced computation times, allowing statistical properties, such as mean transmission, spectral distributions, and phase transitions, to be numerically investigated [6,7]. Knowledge of such general features of random scattering media has afforded great physical insights [8–12] and enabled development of a number of useful techniques, for example, for imaging through scattering media and wave front control [13–15].

Early random matrix models, such as the circular and Gaussian ensembles proposed by Wigner and Dyson [16–18], were based on the assumption of isotropic scattering which often poorly approximates reality. More sophisticated random matrix models have however been developed in an attempt to capture additional symmetries, constraints or system specific properties while preserving numerical speed [7,19,20]. For example, filtered matrix ensembles can more accurately simulate transmission in systems where the input

and output measurement channels are a finite subset of all potentially available channels [21], whereas Euclidean matrix ensembles can describe random Green's matrices relevant to propagation of waves in collections of point-like scattering centers [22–24]. Wishart and Jacobi ensembles meanwhile are also classic ensembles useful, for instance, in the description of chaotic cavities [25,26]. Cascaded, or coupled, random matrix models have also seen significant study in the literature, finding applications in, e.g., electronic transport in wires, quantum chaos, wireless communications, the study of spin glasses [6,14,27,28], and more. Such models describe the linear scattering properties of a system, as described by its (random) transfer matrix $\mathbb{T}$, through the correctly ordered product of individual transfer matrices $\mathbb{T}^\delta$ (drawn from an appropriate underlying matrix ensemble) of constituent scattering sites or system components. Cascaded models can, for example, describe universal conductance fluctuations [29] and naturally introduce a notion of length to the system such that variations in total transmission and reflection with system size can be modeled (in contrast with more traditional matrix ensembles). Of particular significance is the Dorokhov-Mello-Pereyra-Kumar (DMPK) cascaded matrix model [30,31] and its higher-dimensional and electromagnetic generalizations [32–34], which recognize that a thick scattering system can be considered as a chain of thinner, weakly scattering, media. Each successive matrix in the associated product acts as a perturbation, enabling description of various wave transport regimes in a scattering medium (e.g., ballistic, diffusive or localized) through cascades of different lengths. In the so-called ballistic scattering regime, the average number of scattering events of a wave propagating in the medium is $\ll 1$, whereas in the diffusive

*Contact author: matthew.foreman@ntu.edu.sg

regime, corresponding to thicker media, transmitted waves undergo many scattering events.

Although allowing a more physically accurate description of a large variety of scattering systems, cascaded random matrix techniques can suffer from reduced speed since calculation of the transfer matrix for a single realization of a thick medium can necessitate generation of a large number of thin slab matrices. Moreover, statistical convergence is slower for large thicknesses due to the greater variances typically involved, thereby requiring more realizations to be simulated and longer computation times. In this work, our aim is to investigate computational gains that can be made in cascaded random matrix models through a bootstrapping approach employing pools of precalculated matrices. Moreover, we seek to assess the extent to which use of a pool based bootstrapping approach can degrade statistical estimates of mean scattering properties of thick scattering media. In Sec. II, we first introduce and discuss a dual pool bootstrapping method, which can reduce both the number of matrix generations and matrix products required for statistical studies of system properties. In Sec. III, we investigate estimator bias and variance when mean properties are derived from transfer matrices formed from products of matrices sampled from a matrix pool $\mathcal{P}$. We show that bootstrapping can introduce additional estimator fluctuations, the prevalence of which are enumerated in Sec. IV (supporting proofs based on combinatorics of permutations are presented in the Appendix). Finally, in Sec. V, we briefly discuss extending our findings to more complex scenarios.

II. DUAL POOL BOOTSTRAPPING OF CASCADED RANDOM MATRICES

Computational gains in cascaded random matrix models can be made if a bootstrapped approach is adopted in which individual matrices $\mathbb{T}^{\delta}$ are resampled from an existing set instead of on-the-fly generation. This sample could, for example, represent a pool of pregenerated transfer matrices or experimental data. To illustrate this principle, consider simulating N_r realizations of the transfer matrix of a medium of thickness L through cascading N_{δ} slabs of thickness $\delta L = L/N_{\delta}$ using on-the-fly calculations. As indicated in Table I this would require $N_r N_{\delta}$ thin-slab matrices to be generated and calculation of $N_r(N_{\delta} - 1)$ matrix products. Alternatively, pregeneration of a single pool $\mathcal{P}_1$ of N_1 such matrices, from which $\mathbb{T}^{\delta}$ are resampled as needed, can greatly reduce the number of matrix generations required if $N_1 \ll N_r N_{\delta}$. With a single matrix pool, however, the required number of matrix products is unaffected (see Table I). In scenarios where the number of relevant scattering modes can be very large,

TABLE I. Number of calculations required for on-the-fly, single pool and dual pool approaches.

	Matrix generation	Matrix product
On-the-fly	$N_r N_{\delta}$	$N_r(N_{\delta} - 1)$
Single pool	N_1	$N_r(N_{\delta} - 1)$
Dual pool	N_1	$N_2 n_{\Delta} + N_r(N_{\Delta} - 1)$

e.g., optical scattering, products of large transfer matrices can hence still limit total computation time and it is thus preferable to also reduce the number of matrix products. To this end, a dual pool approach can be adopted whereby the pool $\mathcal{P}_1$ of N_1 thin-slab matrices is again pregenerated, however, additionally a second pool $\mathcal{P}_2$ of N_2 transfer matrices $\mathbb{T}^{\Delta}$ for sections of intermediate thickness ΔL is calculated. Each matrix in $\mathcal{P}_2$ is found by cascading $n_{\Delta} = \Delta L/\delta L$ thin-slab matrices which are themselves drawn from $\mathcal{P}_1$. The dual pool approach then allows the number of matrix products to be reduced to $N_2 n_{\Delta} + N_r(N_{\Delta} - 1)$, where $N_{\Delta} = L/\Delta L$. To assess whether a dual pool approach is advantageous, a determination of the computational complexity, time and the required memory for each operation must be made. It is also worthwhile to note that matrix pools can be stored and reused. By way of an example of potential gains, we note that in our earlier work [34] use of the dual pool approach allowed the number of matrix generations and products to be reduced by approximately three and two orders of magnitude, respectively. Consequently, prohibitive computation times were overcome enabling statistical characterization of a variety of polarization scattering phenomena, such as light transmission and optical depolarization, in multiply scattering media up to thicknesses of 30 mean-free paths, using 10^4 realizations for each length. Physical parameters extracted from our simulations using the bootstrapped approach were found to be in agreement with previously reported results.

Resampling, or bootstrapping [35–37], approaches, such as the single or dual pool technique described, although enabling computational gains, do come at the potential cost of matrix reuse since, typically, matrices are sampled with replacement. In a single pool approach, a given transfer matrix $\mathbb{T}^{\delta}$ sampled from $\mathcal{P}_1$ could in principle be used multiple times in calculation of $\mathbb{T}$. Similarly, for a dual pool approach a given $\mathbb{T}^{\delta}$ from $\mathcal{P}_1$ could be used several times when computing single, or different, $\mathbb{T}^{\Delta}$ matrices. Moreover, a $\mathbb{T}^{\Delta}$ drawn from $\mathcal{P}_2$ could be reused during generation of one or more instance of $\mathbb{T}$. Such matrix reuse is undesirable since it can introduce residual and unphysical biases and correlations between individual random matrices. It is therefore important to establish the nature and quantify the magnitude of these detrimental effects so as to assess the viability of use of the bootstrapped approach. This task forms the focus of the remainder of this work.

III. BOOTSTRAPPED ESTIMATION OF POPULATION AVERAGES

We begin our analysis by first adopting a more general and convenient notation to more easily accommodate the different cases discussed above. Specifically, we consider an ensemble of random $M \times M$ transfer matrices $\mathbb{X}$ governed by the probability density function $p_{\mathbb{X}}(\mathbb{X})$. As an attempt to avoid notational clutter in what follows, we assume that transfer matrices are real, although extension of our results to complex ensembles is straightforward, as discussed below. We also construct a pool $\mathcal{P}$ of N_X transfer matrices $\mathbb{X}^{(j)}$, where we label each matrix in the pool with the superscript $j \in [1, N_X]$. For on-the-fly calculations we can form a single realization of a cascaded transfer matrix $\mathbb{Z}$ by generating N individual independent $\mathbb{X}$ matrices according to the true underlying

ensemble probability distribution $p_{\mathbb{X}}(\mathbb{X})$. The $\mathbb{Z}$ matrices produced in this manner define a second ensemble governed by the probability density function $p_{\mathbb{Z}}(\mathbb{Z})$. Alternatively, when using a pool-based approach we instead construct realizations of cascaded transfer matrices $\mathbb{Z}^{\boldsymbol{\alpha}^{(l)}}$ by drawing matrices $\mathbb{X}^{(j)}$ from the pool $\mathcal{P}$ and evaluating the ordered matrix products

$$\mathbb{Z}^{\boldsymbol{\alpha}^{(l)}} = \prod_{p=1}^{N} \mathbb{X}^{(\alpha_p^{(l)})} = \mathbb{X}^{(\alpha_1^{(l)})} \mathbb{X}^{(\alpha_2^{(l)})} \cdots \mathbb{X}^{(\alpha_N^{(l)})}, \quad (1)$$

where $\alpha_p^{(l)}$ is the pth member of a sequence $\boldsymbol{\alpha}^{(l)}$. The members of this sequence correspond to the labels of N matrices drawn from the pool, such that $\boldsymbol{\alpha}^{(l)}$ defines the lth realization of $\mathbb{Z}$. Note that the order of matrix products defined in Eq. (1) will be assumed throughout this work. As an illustration to clarify the connection with our earlier discussion, we can make the associations $\mathbb{X} = \mathbb{T}^{\delta}$ and $\mathbb{Z} = \mathbb{T}^{\Delta}$, as corresponds to a dual pool approach. One possible sampling sequence $\boldsymbol{\alpha}^{(1)} = (2, 2, 1, 2, 1)$, then means that the first realization of $\mathbb{T}^{\Delta} = \mathbb{Z}^{\boldsymbol{\alpha}^{(1)}}$ is formed from sampling the first matrix in $\mathcal{P} = \mathcal{P}_1$ ($\mathbb{X}^{(1)}$) twice and the second matrix ($\mathbb{X}^{(2)}$) thrice, and taking the appropriately ordered matrix product. Generation of multiple realizations of $\mathbb{Z}^{\boldsymbol{\alpha}^{(l)}}$ can then be collected to form $\mathcal{P}_2$. Alternative associations include $(\mathbb{X}, \mathbb{Z}) = (\mathbb{T}^{\delta}, \mathbb{T})$ for the single pool approach or $(\mathbb{T}^{\Delta}, \mathbb{T})$ for the dual pool approach.

A. Estimator bias

Consider now the process of estimating some mean property, $F = \langle f(\mathbb{Z}) \rangle$, of the true matrix ensemble, where $\langle \ldots \rangle$ denotes the expectation with respect to $p_{\mathbb{Z}}(\mathbb{Z})$ and where f is some arbitrary function. We may, for instance, wish to estimate mean transmission, channel capacity, or medium depolarization [11,38,39]. In practice, to estimate F we would consider the sample mean $\hat{F} = \bar{f} = N_Z^{-1} \sum_{l=1}^{N_Z} f(\mathbb{Z}^{\boldsymbol{\alpha}^{(l)}})$, where N_Z is the number of realizations of $\mathbb{Z}$ that we choose to generate (or sample from a pool) and where $\hat{F}$ denotes the estimator of F. The quality of $\hat{F}$ can be assessed through evaluation of the estimator bias

$$\text{bias}[\hat{F}] = \langle \hat{F} \rangle - \langle f \rangle. \quad (2)$$

As an example, consider estimating the mean transfer matrix $F = \langle \mathbb{Z} \rangle$, i.e., $f(\mathbb{Z}) = \mathbb{Z}$, using the sample mean

$$\bar{\mathbb{Z}} = \frac{1}{N_Z} \sum_{l=1}^{N_Z} \mathbb{Z}^{\boldsymbol{\alpha}^{(l)}} = \frac{1}{N_Z} \sum_{l=1}^{N_Z} \prod_{p=1}^{N} \mathbb{X}^{(\alpha_p^{(l)})}. \quad (3)$$

Notably, if all indices in each sequence $\boldsymbol{\alpha}^{(l)}$ are distinct, then taking the ensemble average of Eq. (3) reduces to taking the product of the ensemble means of $\mathbb{X}$, i.e.,

$$\langle \bar{\mathbb{Z}} \rangle = \frac{1}{N_Z} \sum_{l=1}^{N_Z} \prod_{p=1}^{N} \langle \mathbb{X}^{(\alpha_p^{(l)})} \rangle = \frac{1}{N_Z} \sum_{l=1}^{N_Z} \langle \mathbb{X} \rangle^N = \langle \mathbb{Z} \rangle, \quad (4)$$

since each sample of $\mathbb{X}$ is drawn independently from the underlying ensemble and is identically distributed. In this scenario it is seen that the sample mean of the matrices $\mathbb{Z}^{\boldsymbol{\alpha}^{(l)}}$ formed from sampling the underlying pool $\mathcal{P}$ is an unbiased estimator of the true ensemble mean. It is simple to show this is true for more general functions $f(\mathbb{Z})$. When sampling

matrices from the (finite) pool $\mathcal{P}$, however, there is a nonzero probability that the same matrix is drawn multiple times when sampling is performed with replacement, i.e., there may be repeated indices in $\boldsymbol{\alpha}^{(l)}$. Such matrix repetitions, in general, will destroy the unbiased nature of our estimator. Continuing our illustrative example, we can express the sample mean of the (j, k)th matrix element, Z_{jk}, in terms of explicit sums over the matrix elements of the constituent $\mathbb{X}^{(\alpha_p^{(l)})}$ matrices such that

$$\langle \bar{Z}_{jk} \rangle = \frac{1}{N_Z} \sum_{l=1}^{N_Z} \sum_{u_1, \ldots, u_{N-1}} \left\langle \prod_{p=1}^{N} X_{u_{p-1}u_p}^{(\alpha_p^{(l)})} \right\rangle, \quad (5)$$

where the sums over $u_1, \ldots, u_{N-1}$ run from 1 to M and for notational simplicity we let $u_0 = j$, $u_N = k$. The average of the products appearing in Eq. (5) can be rewritten as the product of averages as before; however, factors for which $\alpha_p^{(l)}$ are equal must be grouped together. For instance, again considering the example whereby $\boldsymbol{\alpha}^{(1)} = (2, 2, 1, 2, 1)$, we can write

$$\left\langle \prod_{p=1}^{N} X_{u_{p-1}u_p}^{(\alpha_p^{(1)})} \right\rangle = \langle X_{u_0u_1}^{(2)} X_{u_1u_2}^{(2)} X_{u_3u_4}^{(2)} \rangle \langle X_{u_2u_3}^{(1)} X_{u_4u_5}^{(1)} \rangle. \quad (6)$$

In general, the average of a product, is not commensurate with the product of averages (e.g., $\langle X_{u_2u_3}^{(1)} X_{u_4u_5}^{(1)} \rangle \neq \langle X_{u_2u_3}^{(1)} \rangle \langle X_{u_4u_5}^{(1)} \rangle$) and can moreover be nonzero, particularly when the indices are the same (e.g., $u_2 = u_4$, $u_3 = u_5$). Consequently, it follows that when matrices are sampled from $\mathcal{P}$ with replacement the sample mean typically constitutes a biased estimator of the ensemble mean. Critically, the bias originates from the finite size of the pool from which we sample $\mathbb{X}^{(j)}$ and would not arise for on-the-fly calculations and is thus an artifact of the bootstrapped approach. A notable exception is when the $\mathbb{X}$ matrices belong to a zero-mean Gaussian ensemble and each matrix product $\mathbb{Z}^{\boldsymbol{\alpha}}$ is formed from the product of an odd number of matrices (N is odd). Since all odd moments of a zero-mean Gaussian process are identically zero and at least one of the factors in Eq. (5) will necessarily be the average of an odd number of terms, it follows that Eq. (5) reduces to zero. More generally, bias can be avoided when $\mathbb{X}^{(j)}$ are sampled from $\mathcal{P}$ without replacement when generating individual realizations of $\mathbb{Z}^{\boldsymbol{\alpha}}$, as assumed in the remainder of this work (although we note different $\mathbb{Z}^{\boldsymbol{\alpha}}$ matrices may crucially still have common $\mathbb{X}^{(j)}$ factors). Sampling without replacement also more closely matches physical reality in the sense that the probability of any two (or more) sections of a thick random media being identical is infinitesimally small, such that if this possibility arises it is, again, purely an artifact of the sampling procedure that should be avoided.

B. Estimator variance

Having evaluated the bias of $\hat{F}$, we now consider the estimator variance $\text{var}[\hat{F}] = \langle \hat{F}^2 \rangle - \langle \hat{F} \rangle^2$. Using the definition of the sample mean it is simple to show

$$\text{var}[\hat{F}] = \frac{\text{var}[f(\mathbb{Z})]}{N_Z} + \sum_{l_1=1}^{N_Z} \sum_{\substack{l_2=1 \\ l_1 \neq l_2}}^{N_Z} \frac{\text{cov}[f(\mathbb{Z}^{\boldsymbol{\alpha}^{(l_1)}}), f(\mathbb{Z}^{\boldsymbol{\alpha}^{(l_2)}})]}{N_Z^2},$$

$$(7)$$

where $\mathrm{cov}[x, y] = \langle xy \rangle - \langle x \rangle \langle y \rangle$ denotes the covariance between x and y. The first term in Eq. (7) represents the usual convergence in the mean as governed by the law of large numbers [40] and hence corresponds to the variance in the sample mean $\hat{F}$ for on-the-fly calculations. Additional fluctuations in $\hat{F}$, however, can arise when different samples of $\mathbb{Z}^{\boldsymbol{\alpha}}$ are correlated to some degree, as embodied in the second term of Eq. (7). When the size of the pool $\mathcal{P}$ is finite, specific samples of $\mathbb{X}^{(\alpha_p^{(l)})}$ can be present in different realizations of $\mathbb{Z}^{\boldsymbol{\alpha}}$, implying that these correlations are indeed nonzero when the bootstrapping approach is used.

To assess the strength of these additional fluctuations, we again consider the simple case for which $f(\mathbb{Z}) = \mathbb{Z}$. We therefore seek to assess the covariance matrix $\mathbb{C}$ for two ordered matrix products $\mathbb{Z}^{\boldsymbol{\alpha}^{(1)}}$ and $\mathbb{Z}^{\boldsymbol{\alpha}^{(2)}}$. Element-wise, we therefore wish to evaluate

$$C_{mn}^{jk} = \langle Z_{jk}^{\boldsymbol{\alpha}^{(1)}} Z_{mn}^{\boldsymbol{\alpha}^{(2)}} \rangle - \langle Z_{jk}^{\boldsymbol{\alpha}^{(1)}} \rangle \langle Z_{mn}^{\boldsymbol{\alpha}^{(2)}} \rangle. \quad (8)$$

For simplicity and without loss of generality, we assume that the first sequence of sampled matrices is $\boldsymbol{\alpha}^{(1)} = (1, 2, \ldots, N)$, whereas the second sampling sequence $\boldsymbol{\alpha}^{(2)}$ is arbitrary, albeit each element is unique since we assumed $\mathcal{P}$ is sampled without replacement to avoid estimator bias (relaxation of this assumption is discussed in Sec. V). Accordingly we drop the $\boldsymbol{\alpha}^{(1)}$ dependence of Eq. (8) and let $\boldsymbol{\alpha}^{(2)} = \boldsymbol{\alpha}$. Expressing the matrix products of Eq. (1) as sums over components, we can explicitly write the element-wise covariance in the form

$$C_{mn}^{jk} = \sum_{\substack{u_1,\ldots,u_{N-1} \\ v_1,\ldots,v_{N-1}}} \left[\left\langle \prod_{q=1}^{N} X_{u_{q-1}u_q}^{(q)} X_{v_{q-1}v_q}^{(\alpha_q)} \right\rangle \right.$$
$$\left. - \left\langle \prod_{p=1}^{N} X_{u_{p-1}u_p}^{(p)} \right\rangle \left\langle \prod_{q=1}^{N} X_{v_{q-1}v_q}^{(\alpha_q)} \right\rangle \right], \quad (9)$$

where again for notational convenience we let $v_0 = m$, $v_N = n$ and the sums over $v_1, \ldots, v_{N-1}$ run from 1 to M. With the sampling assumptions given above and further assuming that each $\boldsymbol{\alpha}$ is equally likely, the probability that the two samples of $\mathcal{P}$ contain at least one common matrix is

$$P\left(\{\alpha_j^{(1)}\} \cap \{\alpha_j^{(2)}\} \neq \emptyset\right) = 1 - \binom{N_X - N}{N} \Big/ \binom{N_X}{N},$$

where $\binom{n}{k}$ denotes the binomial coefficient. Noting, as before, that samples of $\mathbb{X}$ used to generate $\mathcal{P}$ are taken independently and are by definition identically distributed, we can match any pairs of $X^{(q)}$ and $X^{(\alpha_k)}$ in Eq. (9) for which $q = \alpha_k$. A matching pair will generate a factor of the form $\langle X_{u_{q-1}u_q}^{(\alpha_q)} X_{v_{q-1}v_q}^{(\alpha_q)} \rangle$ in the first term of Eq. (9) and implies that the matrix $\mathbb{X}^{(\alpha_k)}$ is common to both samples of the matrix pool $\mathcal{P}$. Remaining unpaired matrices, however, decouple and reduce to a product of their means. To facilitate notation, we can momentarily neglect the ordering of the elements of $\boldsymbol{\alpha}$ and think of it as an integer set. We then introduce the set function $\xi_{\boldsymbol{\alpha}}$ associated with $\boldsymbol{\alpha}$ defined such that $\xi_{\boldsymbol{\alpha}}(\{\alpha_{q_1} \ldots \alpha_{q_k}\}) = \{q_1, \ldots, q_k\}$. In words, given some collection of elements of $\boldsymbol{\alpha}$, $\xi_{\boldsymbol{\alpha}}$ picks out the indices corresponding to those elements. Defining now the sets $\mathcal{M} = \xi_{\boldsymbol{\alpha}}(\{1, \ldots, N\} \cap \boldsymbol{\alpha})$, $\mathcal{N} = \{1, \ldots, N\} \setminus \boldsymbol{\alpha}$, and $\mathcal{N}' = \xi_{\boldsymbol{\alpha}}(\boldsymbol{\alpha} \setminus \{1, \ldots, N\})$, where

$\setminus$ denotes the set difference, we find that Eq. (9) can be written as

$$C_{mn}^{jk} = \sum_{\substack{u_1,\ldots,u_{N-1} \\ v_1,\ldots,v_{N-1}}} \left[\prod_{q \in \mathcal{M}} \left\langle X_{u_{(\alpha_q-1)}u_{\alpha_q}}^{(\alpha_q)} X_{v_{q-1}v_q}^{(\alpha_q)} \right\rangle \right.$$
$$\left. - \prod_{q \in \mathcal{M}} \left\langle X_{u_{q-1}u_q}^{(q)} \right\rangle \left\langle X_{v_{q-1}v_q}^{(\alpha_q)} \right\rangle \right]$$
$$\times \prod_{q \in \mathcal{N}} \left\langle X_{u_{p-1}u_p}^{(p)} \right\rangle \prod_{q \in \mathcal{N}'} \left\langle X_{v_{p-1}v_p}^{(\alpha_p)} \right\rangle. \quad (10)$$

Equation (10) is valid quite generally, i.e., for arbitrary $p_{\mathbb{X}}(\mathbb{X})$, however to proceed further we must make some restrictions on the statistical properties of $\mathbb{X}$. As a simple case, we assume that individual transfer matrices, $\mathbb{X}$, are defined by a maximum entropy (i.e., Gaussian) ensemble [31], as is, for example, appropriate for scattering in the diffusive regime. For this model, $\langle \mathbb{X} \rangle = \mathbb{O}$ and $\langle X_{ab}^{(q)} X_{cd}^{(q)} \rangle = r^2 \delta_{ac} \delta_{bd}$, where $\mathbb{O}$ is the null matrix and δ_{jk} is the Kronecker δ function, i.e., each individual element of $\mathbb{X}$ has equal variance, whereas distinct elements are uncorrelated. While this model neglects more complex correlations that may exist in transfer matrices, such as the memory effect, physics imposed symmetry constraints or reflection-transmission correlations [41–44], it serves to sufficiently capture boot-strapping induced correlations that may arise in cascaded models. Moreover, in practice, constituent transfer matrices used in matrix chained models are frequently drawn from simpler matrix ensembles such as the circular or Gaussian ensembles [28,31,45]. The ballistic scattering regime is discussed further below in Sec. V. With these assumptions C_{mn}^{jk} is only nonzero when $\boldsymbol{\alpha}$ is an N permutation of $(1, 2, \ldots, N)$, that is, when the sets of sampled matrices $\mathcal{N}$ and $\mathcal{N}'$ are empty and $\mathcal{M} = \{1, \ldots, N\}$, such that we can again enumerate over q, whereby

$$C_{mn}^{jk} = \sum_{\substack{u_1,\ldots,u_{N-1} \\ v_1,\ldots,v_{N-1}}} \prod_{q=1}^{N} \left\langle X_{u_{\alpha_q-1}u_{\alpha_q}}^{(\alpha_q)} X_{v_{q-1}v_q}^{(\alpha_q)} \right\rangle. \quad (11)$$

The summations over $v_1, \ldots, v_{N-1}$ appearing in Eq. (11) can be performed analytically yielding

$$C_{mn}^{jk} = r^{2N} \sum_{u_1,\ldots,u_{N-1}} \delta_{u_{\alpha_1-1}v_0} \delta_{v_N u_{\alpha_N}} \prod_{q=1}^{N-1} \delta_{u_{\alpha_{(q+1)}-1}u_{\alpha_q}}. \quad (12)$$

To help evaluate C_{mn}^{jk} we define the multiset β of all indices appearing in Eq. (12) (allowing for repeated indices) such that $\beta = \{u_{\alpha_1-1}, v_0, u_{\alpha_2-1}, u_{\alpha_1}, u_{\alpha_3-1}, u_{\alpha_2}, \ldots, u_{\alpha_N}, v_N\}$, which has cardinality $|\beta| = 2(N + 1)$, corresponding to the $N + 1$ different Kronecker δ functions appearing in Eq. (12). The elements $\{u_0, u_N, v_0, v_N\}$ each have a multiplicity of one in β, whereas $\{u_1, u_2, \ldots, u_{N-1}\}$ have multiplicities of two. Mirroring the structure of Eq. (12) we further define the multisets $\gamma = \{u_{\alpha_1-1}, v_0, u_{\alpha_N}, v_N\}$ and $\zeta = \{u_{\alpha_2-1}, u_{\alpha_1}, u_{\alpha_3-1}, u_{\alpha_2}, \ldots\}$ such that $\beta = \gamma \cup \zeta$. Note that $u_0 \notin \zeta$ if and only if $\alpha_1 - 1 = 0$. Similarly, $u_N \notin \zeta$ if and only if $\alpha_N = N$. With these definitions, consider then the form of Eq. (12) when $\{u_0, u_N\} \notin \zeta$ which reduces to (recalling that

we previously set $u_0 = j$, etc. for convenience)

$$C_{mn}^{jk} = r^{2N} \delta_{jm} \delta_{kn} \sum_{u_1,\ldots,u_{N-1}} \prod_{q=1}^{N-1} \delta_{u_{\alpha_{(q+1)}-1} u_{\alpha_q}}. \tag{13}$$

The indices appearing in the summations are those contained in $\{u_1, u_2, \ldots, u_{N-1}\}$ each with multiplicities of two. We can thus group together indices that form disjoint closed cycles. For example, consider the permutation $(1)(2,3,5)(4)$, which we have written in standard cycle notation [46]. The summation in this case would take the form

$$\left[\sum_{u_1} \delta_{u_1 u_1} \right] \left[\sum_{u_2} \sum_{u_3} \sum_{u_5} \delta_{u_2 u_3} \delta_{u_3 u_5} \delta_{u_5 u_2} \right] \left[\sum_{u_4} \delta_{u_4 u_4} \right],$$

where the indices are grouped according to the cycle structure. Given that each index is summed from 1 to M it is quickly seen that each disjoint set of summations totals M such that

$$C_{mn}^{jk} = r^{2N} \delta_{jm} \delta_{kn} M^K, \tag{14}$$

where K is the number of disjoint cycles of indices in the sum (which we note do not contain u_0 or u_N).

Moving now to a more complex case whereby we allow $u_0 \in \zeta$, but still restrict $u_N \notin \zeta$, we have

$$C_{mn}^{jk} = r^{2N} \delta_{kn} \sum_{u_1,\ldots,u_{N-1}} \delta_{u_0 u_{\alpha_p}} \delta_{v_0 u_{\alpha_1-1}} \prod_{\substack{q=1 \\ q \neq p}}^{N-1} \delta_{u_{\alpha_{(q+1)}-1} u_{\alpha_q}}, \tag{15}$$

where, since $u_0 \in \zeta$, we note p is the index such that $\alpha_{p+1} = 1$ where $p \neq 0$. We also have that $\alpha_1 - 1 > 0$. We can once again group indices into summations over closed disjoint cycles of indices, which requires that the cycles do not contain u_0 or v_0. Each set of indices that can be so factored contributes a multiple of M to C_{mn}^{jk}. We denote the remaining multiset of indices $\psi = \{u_0, u_{\alpha_p}, u_{\alpha_1-1}, v_0, u_{\alpha_p}, u_{\alpha_1-1}, u_f, u_g, \ldots\}$, where u_f and u_g denote general indices which will have multiplicity of two. From the structure of the indices appearing in Eq. (12), it however follows that indices with multiplicity of two appear as different arguments to the Kronecker δ functions, i.e., the remaining summation can be written as

$$\sum_{u_{\alpha_p}} \sum_{u_{\alpha_1-1}} \cdots \sum_{u_f} \delta_{u_0 u_{\alpha_p}} \delta_{u_{\alpha_p} u_f} \cdots \delta_{u_f u_{\alpha_1-1}} \delta_{u_{\alpha_1-1} v_0} = \delta_{jm}.$$

We thus find that C_{mn}^{jk} is again given by Eq. (14). Analogous arguments show the same to be true for the case that $u_0 \notin \zeta$, but $u_N \in \zeta$. The remaining case ($\{u_0, u_N\} \in \zeta$) can be considered in a similar manner, factoring out disjoint cycles of indices not containing u_0, u_N, v_0, and v_N, each of which contributes a factor of M. The remaining indices reduce upon summation to factors of the form $\delta_{jk} \delta_{mn}$ or $\delta_{jm} \delta_{kn}$ depending on the initial permutation, such that C_{mn}^{jk} is given either by Eq. (14) or by

$$C_{mn}^{jk} = r^{2N} \delta_{jk} \delta_{mn} M^K. \tag{16}$$

Given these results we observe that the elements of the correlation matrix formed from C_{mn}^{jk} are either zero or a fixed

value ($= r^{2N} M^K$) for any given permutation $\boldsymbol{\alpha}$ (note $\mathbb{C}$ is not diagonal). Moreover, the number of nonzero elements is

$$\sum_{j,k,m,n=1}^{M} \delta_{jk} \delta_{mn} = \sum_{j,k,m,n=1}^{M} \delta_{jm} \delta_{kn} = M^2 \tag{17}$$

for all possible permutations. Returning to Eq. (7) for $f(\mathbb{Z}) = \mathbb{Z}$ and considering the total variance in our estimate, as shall be quantified by taking the 1-norm (denoted $\| \ldots \|_1$), we can write

$$\|\text{var}[\hat{\mathbb{Z}}]\|_1 = \frac{\|\text{var}[\mathbb{Z}]\|_1}{N_{\mathbb{Z}}} + \sum_{l_1=1}^{N_{\mathbb{Z}}} \sum_{\substack{l_2=1 \\ l_1 \neq l_2}}^{N_{\mathbb{Z}}} \frac{C(K, M, N)}{N_{\mathbb{Z}}^2}, \tag{18}$$

where K is a function of the specific sampled sequences $\boldsymbol{\alpha}^{(l_1)}$ and $\boldsymbol{\alpha}^{(l_2)}$ (or equivalently the sequences $(1, 2, \ldots, N)$ and $\boldsymbol{\alpha}^{(l)}$ under a suitable transformation) and

$$C(K, M, N) = \|\mathbb{C}\|_1 = \sum_{j,k,m,n=1}^{M} C_{mn}^{jk} = r^{2N} M^{K+2}. \tag{19}$$

The total covariance $C(K, M, N)$ hence characterizes the degree to which $\mathbb{Z}^{(1,\ldots,N)}$ and $\mathbb{Z}^{\boldsymbol{\alpha}}$ are correlated for a given permutation $\boldsymbol{\alpha}$ (note that with a slight abuse of notation we use α to denote the permutation described by the sequence $\boldsymbol{\alpha}$). Naturally $C(K, M, N)$ is largest when α is the identity (whereby $K = N$), i.e., $\mathbb{Z}^{\boldsymbol{\alpha}} = \mathbb{Z}^{(1,\ldots,N)}$. Observing further that $\|\text{var}[\mathbb{Z}]\|_1 = C(N, M, N)$, the relative magnitude of each contribution in the second term of Eq. (18) with respect to the first term is $\rho/N_{\mathbb{Z}}$, where

$$\rho = \frac{C(K, M, N)}{C(N, M, N)} = \frac{1}{M^{N-K}} \tag{20}$$

is the Pearson correlation coefficient, which decreases with matrix size M since $K \leqslant N$. For cascaded transfer matrix models, the correlations between different matrix products hence become less significant for systems with large numbers of scattering modes, therefore promoting the use of bootstrapping techniques in such scenarios. Likewise as the number of matrices N within each cascade increases, so the potential correlation between samples decreases in strength.

From a practical standpoint, it may also be of interest to consider the additional number of realizations of $\mathbb{Z}$ that must be generated using the bootstrapping method as compared with the on-the-fly approach so as to produce the same total estimator variance $\|\text{var}[\hat{\mathbb{Z}}]\|_1$. To answer this we recall Eq. (18) and note that if only the first term is considered the result corresponds to the total variance of $\hat{\mathbb{Z}}$ for on-the-fly calculations. Therefore, denoting the total number of realizations required to produce a given and fixed $\|\text{var}[\hat{\mathbb{Z}}]\|_1$ using a bootstrapped or on-the-fly approach explicitly as $N_{\mathbb{Z}}^{\text{bs}}$ and $N_{\mathbb{Z}}^{\text{otf}}$, it follows [by equating Eq. (18) for the two cases and solving the resulting quadratic equation in $N_{\mathbb{Z}}^{\text{bs}}$] that

$$\frac{N_{\mathbb{Z}}^{\text{bs}}}{N_{\mathbb{Z}}^{\text{otf}}} \approx 1 + \sum_{l_1=1}^{N_{\mathbb{Z}}^{\text{bs}}} \sum_{\substack{l_2=1 \\ l_1 \neq l_2}}^{N_{\mathbb{Z}}^{\text{bs}}} \frac{1}{M^{N-K}}, \tag{21}$$

where we have also used Eq. (20) and assumed that $\sum_{l_1} \sum_{l_2 \neq l_1} \rho \ll N_Z^{\text{otf}}$. It therefore follows that the additional number of realizations required to produce a desired variance also decreases with increasing matrix size M and number of matrices N.

Finally, we consider how the results given generalize when the constituent matrices $\mathbb{X}$ are complex. Notably, all results pertaining to estimator bias presented in Sec. III A are unchanged. When considering estimator variance, for proper complex random matrices $\mathbb{X}$ with zero mean and equal variance, i.e., $\langle \mathbb{X} \rangle = \mathbb{O}$ and $\langle X_{ab}^{(q)} X_{cd}^{(q)*} \rangle = r^2 \delta_{ac} \delta_{bd}$, formally identical results to those given in Sec. III B can also be derived, albeit using complex generalizations to the correlation functions [such as Eq. (8)], in which the second matrix factor is conjugated. Note that pseudocorrelation functions are identically zero for proper random variables, i.e., $\langle X_{ab}^{(q)} X_{cd}^{(q)} \rangle = 0$, and $\text{var}[\text{Re}(\mathbb{X})] = \text{var}[\text{Im}(\mathbb{X})] = \text{cov}[\mathbb{X}, \mathbb{X}^*]/2$ [47]. For improper $\mathbb{X}$ the pseudocorrelation would, however, be nonzero requiring further analysis, for which specification of both $\langle X_{ab}^{(q)} X_{cd}^{(q)} \rangle$ and $\langle X_{ab}^{(q)} X_{cd}^{(q)*} \rangle$ would be needed.

IV. ENUMERATION OF PERMUTED MATRIX CORRELATIONS

In addition to quantifying the total correlation between permuted matrix products for a given permutation, it is also relevant to enumerate the frequency with which each value of correlation occurs across the symmetric group S_N and thus how often each would appear in the summation of Eqs. (18) and (21). To begin to answer this we seek a more convenient way to determine K. To do so we revisit the form of Eq. (12) and define ordered $(N+1)$ tuples comprising of the indices found in the first and second position of each Kronecker delta, respectively, i.e., $\mathbf{p}_1 = (u_{\alpha_1-1}, u_{\alpha_2-1}, \ldots, u_{\alpha_N-1}, v_N)$ and $\mathbf{p}_2 = (v_0, u_{\alpha_1}, u_{\alpha_2}, \ldots, u_{\alpha_{N-1}}, u_{\alpha_N})$. Noting that $\alpha_q - 1$ spans $[0, N-1]$ and α_q spans $[1, N]$ for $q \in [1, N]$, it follows that the subscripts are unique within each tuple and we can equivalently consider the subscripts themselves whereby $\mathbf{p}_1 = (\alpha_1 - 1, \alpha_2 - 1, \ldots, \alpha_N - 1, N)$ and $\mathbf{p}_2 = (0, \alpha_1, \alpha_2, \ldots, \alpha_{N-1}, \alpha_N)$. The cycles of the $(N+1)$ permutation σ that maps $\mathbf{p}_2 \to \mathbf{p}_1$ (i.e., considering σ as a permutation matrix $\mathbf{p}_1^T = \sigma \mathbf{p}_2^T$) therefore correspond to the cycles of indices in which we are interested. Specifically, K is the number of disjoint cycles of σ which do not contain 0 or N, whereby it follows that $K \leqslant N - 1$.

To express σ in terms of α we let $\mathbf{p}_0 = (0, 1, 2, \ldots, N)$ and define the $(N+1) \times (N+1)$ permutation matrices σ_1 and σ_2 whereby $\mathbf{p}_k^T = \sigma_k \mathbf{p}_0^T$ for $k = 1, 2$. We thus immediately see that σ_2 has the structure

$$\sigma_2 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \sigma_\alpha \end{bmatrix}, \tag{22}$$

where σ_α is the $N \times N$ permutation matrix associated with α and $\mathbf{0}$ is an N-element column vector of zeros. We also note $\mathbf{p}_1^T = \sigma_1 \sigma_2^{-1} \mathbf{p}_2^T$ such that $\sigma = \sigma_1 \sigma_2^{-1}$. Defining the

cyclic-shifting operator

$$\sigma_+ = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 & 1 \\ 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 & 0 \end{bmatrix}, \tag{23}$$

which shifts elements in a tuple p to the right in a cyclic manner, and noting that shifting to the left (as performed by $\sigma_- = \sigma_+^{-1} = \sigma_+^T$) is equivalent to subtraction of unity modulo $N + 1$, we can write

$$\sigma_1 = \sigma_- \sigma_2 \sigma_+ = \begin{bmatrix} \sigma_\alpha & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{24}$$

yielding

$$\sigma = \sigma_- \sigma_2 \sigma_-^{-1} \sigma_2^{-1}. \tag{25}$$

From Eq. (25) it follows simply that $\det[\sigma] = 1$ implying that σ is always an even permutation [46]. Furthermore, since σ cannot map the first element of $\mathbf{p}_2$ (which is zero by construction) to the last element of $\mathbf{p}_1$ (which is N), it follows that element $(N + 1, 1)$ of σ is zero, i.e., $\sigma_{N+1,1} = 0$.

With this formalism we can thus more simply numerically enumerate the frequency of $C(K, M, N)$ across all possible permutations $\alpha_N \in S_N$, which we denote $\nu(K, N)$. Note that the argument M has been omitted since the frequency derives purely from combinatorial aspects of matrix products and is hence independent of matrix size. Results are presented in Table II for small values of N. We note the following results (which have been numerically verified via exhaustion up to $N = 13$):

$$\nu(N - 1, N) = 1, \tag{26}$$

$$\nu(N - 2, N) = N - 1, \tag{27}$$

$$\nu(N - 3, N) = \binom{N-1}{N-3} + \binom{N}{N-3} + \binom{N+1}{N-3}, \tag{28}$$

$$\nu(0, N) = N! / \lfloor (N+2)/2 \rfloor, \tag{29}$$

for all N. We observe Eqs. (28) and (29) correspond to integer sequences A005718 and A107991 respectively [48,49]. Mathematical proofs of Eqs. (26)–(29) are presented in the Appendix. Specifically, we prove Eqs. (26)–(28) through exhaustion and further conjecture that the graphical approach employed to prove Eq. (28) can be extended to other cases. For Eq. (29), however, we present an alternative approach based on establishing a bijection to known combinatorial results.

The results of Table II and the analytic formulas for $\nu(K, N)$ highlight a number of important trends with respect to correlations between cascaded transfer matrices. Most importantly, it is evident that the relative frequency

$$\nu_{\text{rel}} = \nu(K, N)/N! \tag{30}$$

of the higher-correlation cases (i.e., smaller $S = N - K$) across the symmetric group becomes smaller for larger N. Consequently, the majority of thick media transfer matrices found from cascading large numbers of thin-section transfer matrices are only weakly correlated with each other, if

TABLE II. Table of frequency $\nu(N, K)$ of each value of $C(N, K)$ for all permutations $\alpha_N \in S_N$.

N \ K	0	1	2	3	4	5	6	7	8	9
1	1									
2	1	1								
3	3	2	1							
4	8	12	3	1						
5	40	44	31	4	1					
6	180	324	145	65	5	1				
7	1260	1784	1499	370	120	6	1			
8	8064	16 288	9772	5180	805	203	7	1		
9	72 576	120 672	113 868	39 032	14 833	1568	322	8	1	
10	604 800	1 327 680	958 956	570 044	126 861	37 149	2814	486	9	1

at all. Correlations thus decrease in both relative magnitude and relative frequency as the number of matrices in a given product increases and consequently the additional fluctuations introduced in an estimator $\hat{F}$ will also typically be lower. The frequency function $\nu(K, N)$ is nevertheless peaked at a nonzero correlation, showing a nonmonotonic dependence on K for a fixed N.

V. ROUTES TO GENERALIZATION

In the above analysis we made a number of restrictive assumptions. We now briefly discuss and outline how generalization of our results could be sought by the interested reader.

First, we note that the statistical model assumed thus far for the thin section transfer matrix $\mathbb{X}$ was appropriate for the diffusive scattering regime. Alternatively, we can con-sider the case for which a thin slab only weakly perturbs the field incident upon it as is more relevant for calculations in the ballistic regime and DMPK-type models. For weakly scattering slabs the transfer matrix can be written in the form $\mathbb{X} = \mathbb{I} + \Delta\mathbb{X}$, where $\Delta\mathbb{X}$ describes the scattering based perturbation from the identity matrix $\mathbb{I}$. As a simple model we can assume now $\langle \Delta\mathbb{X} \rangle = \mathbb{O}$, where $\mathbb{O}$ is the null matrix and $\langle \Delta\mathbb{X}_{ab} \Delta\mathbb{X}_{cd} \rangle = r^2 \delta_{ac}\delta_{bd}$. It immediately follows that $\langle X_{ab}X_{cd} \rangle = \delta_{ab}\delta_{cd} + r^2 \delta_{ac}\delta_{bd}$ such that each individual element of $\mathbb{X}$ has equal variance, whereas distinct off-diagonal elements are uncorrelated. Note, diagonal elements of the transfer matrix $\mathbb{X}$ describing direct transmission possess a nonzero mean in contrast with the diffusive result.

For the case that $\boldsymbol{\alpha}$ is an N permutation of $(1, 2, \ldots, N)$ we note that

$$C_{mn}^{jk} = \sum_{\substack{u_1, \ldots, u_{N-1} \\ v_1, \ldots, v_{N-1}}} \left[\prod_{q=1}^{N} \left\langle X_{u_{\alpha_{q-1}} u_{\alpha_q}}^{(\alpha_q)} X_{v_{q-1} v_q}^{(\alpha_q)} \right\rangle - \prod_{q=1}^{N} \left\langle X_{u_{q-1} u_q}^{(q)} \right\rangle \left\langle X_{v_{q-1} v_q}^{(\alpha_q)} \right\rangle \right] \tag{31}$$

$$= \sum_{\substack{u_1, \ldots, u_{N-1} \\ v_1, \ldots, v_{N-1}}} \left[\prod_{q=1}^{N} \left(\delta_{u_{\alpha_{q-1}} u_{\alpha_q}} \delta_{v_{q-1} v_q} + r^2 \delta_{u_{\alpha_{q-1}} v_{q-1}} \delta_{u_{\alpha_q} v_q} \right) - \prod_{q=1}^{N} \delta_{u_{q-1} u_q} \delta_{v_{q-1} v_q} \right]. \tag{32}$$

Expanding the first product term gives a power series in r^2, for which a general term can be written in the form

$$r^{2|\gamma|} \sum_{\substack{u_1, \ldots, u_{N-1} \\ v_1, \ldots, v_{N-1}}} \prod_{q \in \beta} \delta_{u_{\alpha_{q-1}} u_{\alpha_q}} \delta_{v_{q-1} v_q} \prod_{p \in \gamma} \delta_{u_{\alpha_{p-1}} v_{p-1}} \delta_{u_{\alpha_p} v_p}, \tag{33}$$

where β and γ are sets of indices dependent on which term in the series we consider and $\beta \cup \gamma = \{1, 2, \ldots, N\}$. Following a similar logic to above, we can sum over $v_1, \ldots, v_{N-1}$ such that a general term in the expansion is given by

$$r^{2|\gamma|} \sum_{u_1, \ldots, u_{N-1}} \delta_{u_{\tau_1-1} v_0} \delta_{u_{\tau_{|\gamma|}} v_N} \prod_{j=1}^{|\beta|} \delta_{u_{\rho_{j-1}} u_{\rho_j}} \prod_{k=1}^{|\gamma|-1} \delta_{u_{\tau_k} u_{\tau_{k+1}-1}}, \tag{34}$$

where $\rho_j = \alpha_{\beta_j}$ and $\tau_k = \alpha_{\gamma_k}$. Equation (34) can then be evaluated by counting the cycles of the permutation which transforms the tuple $\mathbf{p}_1$ to $\mathbf{p}_2$ for

$$\mathbf{p}_1 = (\rho_1 - 1, \ldots, \rho_{|\gamma|} - 1, \tau_1 - 1, \tau_2 - 1, \ldots, \tau_{|\tau|} - 1, N)$$

and

$$\mathbf{p}_2 = (\rho_1, \ldots, \rho_{|\gamma|}, 0, \tau_1, \ldots, \tau_{|\tau|-1}, N).$$

The analysis for a single term in Eq. (32) can thus proceed in an analogous manner to that given above for the diffusive regime. Given the indices involved in each term, as defined by the sets β and γ [cf. Eq. (34)], differ for each term, it follows that the relevant permutation and the resulting con-tribution to the total covariance also varies term by term. Nevertheless, the route to the ballistic regime is apparent, if not somewhat tedious. Ultimately, similar trends in the bootstrapped induced fluctuations in estimates of population averages, in terms of matrix size M and number of factors in the matrix products N, would result.

Finally, we briefly consider the generalization whereby matrices $\mathbb{X}^{(j)}$ are sampled from the pool $\mathcal{P}$ with replacement. As per our earlier discussion in Sec. III, this sampling strategy

can introduce undesirable bias in an estimate of the sample mean, but is simple to implement. The consequence of sampling from $\mathcal{P}$ with replacement, is that indices in both $\boldsymbol{\alpha}^{(1)}$ and $\boldsymbol{\alpha}^{(2)}$ may be repeated [note it is now necessary to reintroduce the more general notation since the previous assumption that $\boldsymbol{\alpha}^{(1)} = (1, 2, \ldots, N)$ does not adequately encompass all possibilities]. In calculation of covariance matrix elements [Eq. (9)], repeated indices in $\boldsymbol{\alpha}^{(1)}$ and $\boldsymbol{\alpha}^{(2)}$ imply that upon grouping averages of like terms [cf. Eq. (12)], one can obtain higher-order moments. For instance, if $\boldsymbol{\alpha}^{(1)} = (1, 2, 2)$ and $\boldsymbol{\alpha}^{(2)} = (4, 3, 2)$, the first product appearing in Eq. (9) would include a factor of the form $\langle X_{pq}^{(2)} X_{rs}^{(2)} X_{uv}^{(2)} \rangle$ (where we let the subscripts take arbitrary values for simplicity). Use of the moments-cumulant formula [50] (or Isserlis' theorem for Gaussian random variables) would in principle allow such higher-order moments to be expressed in terms of lower orders. So doing, however, requires careful attention to be paid to the partitioning of the index sets for each individual case, which rapidly becomes cumbersome but is, in principle, possible.

VI. CONCLUSIONS

In this article we have considered the problem of simulating wave propagation in random media using random matrices. Through forming the appropriate ordered product of random transfer matrices drawn from a suitable ensemble, waves can in principle be propagated through thick media. Indeed, such cascaded models have spawned a number of interesting insights into the statistics of, for instance, the eigenvalues and mode spacings of the underlying ensembles [28,51–56]. To improve efficiency of cascaded matrix models, we have here proposed a bootstrapped dual pool approach whereby after pregeneration of pools of random transfer matrices, many realizations can be simulated through matrix resampling. The proposed approach not only reduces the number of matrix generations required, but also reduces the total number of matrix products, which can limit computation times when there are many scattering modes available. We have however also shown that matrix resampling inherent in the proposed technique can in principle adversely affect the statistical properties of the resulting ensemble. In particular, we considered how bootstrapping can introduce undesirable statistical bias and additional variance in estimates of population averages of properties such as transmission. While bias was shown to be avoidable through an appropriate sampling strategy, correlations that arise from matrix reuse, and the additional estimator covariance that follows, remain. The strength of this additional covariance, as characterized by Eq. (18), was found to be dependent on the correlation between different matrix permutations. An extensive study of the magnitude and frequency of such correlations was thus also presented, including proofs of a number of closed form analytic formulas [Eqs. (26)–(29)]. An important finding of this study was that the consequences of spurious correlations will typically decrease in severity as the dimension of random matrices and the number of matrix products increases [see Eqs. (20) and (30)]. Fortunately, bootstrapping approaches are only required in the regime where scattering channels are numerous and scattering media are many mean-free paths in length. Our results therefore show that, in spite of the detrimental statistical effects of bootstrapping, practical application of either a single or dual pool approach is not limited by them.

APPENDIX: MATHEMATICAL PROOFS OF CORRELATION RESULTS

1. Proof of Eq. (26)

To prove Eq. (26) we begin by considering the cycle type of σ [46]. Specifically we recall that an $N + 1$ permutation with a_i cycles of length i has cycle type $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_N, a_{N+1})$. It follows then that $N + 1 = \sum_i i a_i$ and the total number of cycles $N_c = \sum_i a_i$. Since we are ultimately interested in the number of cycles K that do not contain zero or N, we must consider two cases: when zero and N are in the same cycle whereby $N_c = K + 1$ (Case A) and when zero and N are in different cycles, whereby $N_c = K + 2$ (Case B). Now, assuming $K = N - S$, we can write

$$K = \sum_i a_i - \Delta = N - S = \left(\sum_i i a_i\right) - 1 - S, \quad \text{(A1)}$$

where $\Delta = 1$ for Case A and $\Delta = 2$ for Case B. Upon rearrangement, Eq. (A1) becomes

$$\sum_i (i - 1) a_i = a_2 + 2a_3 + \cdots + N a_{N+1} = S - \Delta + 1. \quad \text{(A2)}$$

For $S = 1$ we can see by inspection that the possible cycle types satisfying Eq. (A2) for Case A are $\mathbf{a}_A = (N - 1, 1, 0, 0, \ldots)$ such that the cycles must be $(0, N)(1)(2) \ldots (N - 1)$, whereas for Case B $\mathbf{a}_B = (N + 1, 0, 0, 0, \ldots)$ such that the cycles must be $(0)(1)(2) \ldots (N - 1)(N)$. Noting that odd permutations have an even number of even cycles [46], the former solution, $\mathbf{a}_A$, is not permitted since it describes an odd permutation. There is thus a single permutation $\sigma = (0)(1)(2) \ldots (N - 1)(N)$, corresponding to $\alpha = (1)(2) \ldots (N)$, i.e., the identity, for which $S = 1$ therefore completing the proof of Eq. (26).

2. Proof of Eq. (27)

We can approach the proof of Eq. (27) in a similar manner to that used to prove Eq. (26) above. Consider then possible solutions of Eq. (A2) when $S = 2$, i.e., $K = N - 2$. For Case A we require

$$a_2 + 2a_3 + 3a_4 + \cdots = 2 \quad \text{(A3)}$$

such that

$$\mathbf{a}_A = \begin{cases} (N - 3, 2, 0, 0, \ldots) \\ (N - 2, 0, 1, 0, \ldots) \end{cases}, \quad \text{(A4)}$$

TABLE III. Index constraints and corresponding number of allowable permutations for different classes of cycle structure.

Class	Cycle Structure	Constraints	Total Cases
B1.1	$(0, p)(q, N)$	$q < p$	$\frac{1}{2}(N-1)(N-2)$
B1.2	$(0, p)(q, r)$	$p < r < q$	$\frac{1}{6}(N-1)(N-2)(N-3)$
B1.3	$(p, q)(r, N)$	$p < r < q$	$\frac{1}{6}(N-1)(N-2)(N-3)$
B1.4	$(p, q)(r, s)$	$p < s < q < r$	$\frac{1}{24}(N-1)(N-2)(N-3)(N-4)$
B2.1	$(0, p, q)$	$q < p$	$\frac{1}{2}(N-1)(N-2)$
B2.2	(p, q, N)	$q < p$	$\frac{1}{2}(N-1)(N-2)$
B2.3	(p, q, r)	$p < r < q$	$\frac{1}{6}(N-1)(N-2)(N-3)$

corresponding to cycle structures of the form

$$(0, N)(1)(2)\ldots(p, q)\ldots(N-1) \tag{A5}$$

and

$$(0, p, N)(1)(2)\ldots(p-1)(p+1)\ldots(N-1), \tag{A6}$$

respectively, where $p \neq q$ and $p, q \in [1, N-1]$. Cycle structures of the form of Eq. (A5) are not permitted since σ cannot map zero directly to N ($\sigma_{N+1,1} = 0$), whereas cycle structures of the form of Eq. (A6) are permitted such that there are $N - 1$ corresponding permutations. Inspection of the structure of the associated permutation matrices also show that they are of the form required by Eqs. (22)–(25) (this is discussed further in the next section). For Case B, i.e., $\Delta = 2$, we require

$$a_2 + 2a_3 + 3a_4 + \cdots = 1, \tag{A7}$$

which has the solution $(a_2, a_3, \ldots) = (1, 0, \ldots)$ such that $\mathbf{a}_B = (N - 1, 1, 0, 0, \ldots)$. This corresponds to an odd permutation implying these solutions are forbidden. In total there are hence $N - 1$ allowed permutations of the form of Eq. (A6) for which $K = N - 2$, hence concluding our proof of Eq. (27).

3. Proof of Eq. (28)

To prove Eq. (28) we once more consider the possible solutions to Eq. (A2) now for $S = 3$, i.e., $K = N - 3$, which are

$$\mathbf{a}_A = \begin{cases} (N - 5, 3, 0, 0, 0, \ldots) \\ (N - 4, 1, 1, 0, 0, \ldots) \\ (N - 3, 0, 0, 1, 0, \ldots) \end{cases}, \tag{A8}$$

corresponding to odd (and hence invalid) permutations, and

$$\mathbf{a}_B = \begin{cases} (N - 3, 2, 0, 0, 0, \ldots) \\ (N - 2, 0, 1, 0, 0, \ldots) \end{cases}, \tag{A9}$$

which correspond to even and hence potentially acceptable solutions. Seven distinct classes of cycle structures for the Case B cycle types can be constructed, as are listed in Table III (omitting one cycles for clarity), where $p, q, r, s \in [1, N-1]$ are all distinct. A naive consideration of the combinatorics of possible values of p, q, r, s, however, does not yield Eq. (28), as some combinations are inconsistent with the required structure of σ as dictated by Eq. (25). To illustrate we use the two-line representation of permutations and consider the ordered action of σ_2^{-1} and σ_1 on the original sequence

$[0, 1, 2, \ldots, N - 1, N]$, which can hence be represented using a three-line representation.

Consider the case when the cycle structure of σ is of Class B1.1 as an initial illustration. Given that σ_2^{-1} does not permute the 0 element [as easily seen from Eq. (22)] we can immediately populate the 0 element into the first position on the second row as shown in Fig. 1(a). Further noting that the cycle structure given in Table III means that under the action of both permutations $0 \to p$, we can place p into the first element of the third row, while simultaneously populating the pth element with zero. Similarly, σ_1 does not affect the last element of the sequence upon which it acts [cf. Eq. (24)], and the cycle structure implies $N \to q$ (and vice versa), whence we can fill additional positions in our three-line representation as also shown in Fig. 1(a). In Fig. 1(a), note that we have implicitly assumed that $p < q$. For cycle structures of Class B1.1, all other cycles are one-cycles such that the elements in the upper and lower lines must match for all other columns in our representation (albeit, importantly they need not match in the middle row). To allow for sequences of general length we denote these matching elements using ellipses ($\cdots$).

Blank spaces in Fig. 1(a) correspond to elements that do not follow directly from the cycle structure of σ and are thus as yet unknown in general, however, we can attempt to deduce them by invoking the transposition structure imposed by Eqs. (22) and (24). To visualize this structure, assume that σ_2^{-1} induces the transposition indicated in blue in Fig. 1(a), which is inherited from, and encoded in, $\sigma_\alpha^{-1} = \sigma_\alpha^T$ (or equivalently σ_α). Crucially, σ_1 also derives from σ_α and so it is helpful to consider the inverse of the blue transposition, which we have depicted in Fig. 1(a) in green. Equation (24) however tells us that to fully describe σ_1 we must also consider the cyclic operator, which ultimately yields the shifted transposition depicted in red. Individual transpositions arising from the action of σ_1 can thus be graphically inferred from each transposition in σ_2^{-1}.

With these observations we can now attempt to populate all elements of the three-line representation and exclude any cases from our enumeration which lead to logical contradictions. With reference to Fig. 1(b) and again considering Class B1.1 for $p < q$, we start by considering the transposition necessary to place q in the last element of the second row (depicted in dark blue). As discussed we can immediately infer that the σ_1 must contain the transposition shown in light blue (we henceforth use the color coding whereby a dark color transposition infers the corresponding light colored transposition). Noting the existence of the light blue transposition, we can back propagate the element in the final row (here denoted a) to the correct position in the second row. Since a belongs to a one-cycle, we can then conclude that σ_2^{-1} must contain the transposition shown in dark green, which in turn implies the permutation shown in light green. This pattern repeats [as shown in both Fig. 1(b) and Animation 1 of the Supplemental Material [57]] until we reach the light red transposition. This transposition requires special attention since it contains the 0 element. In particular, if we back propagate 0 to the middle row, we find that it necessarily does not originate from the first position. This is a contradiction with the structure of σ_2^{-1}, which dictates that 0 must appear in the first position of the middle row. Therefore,
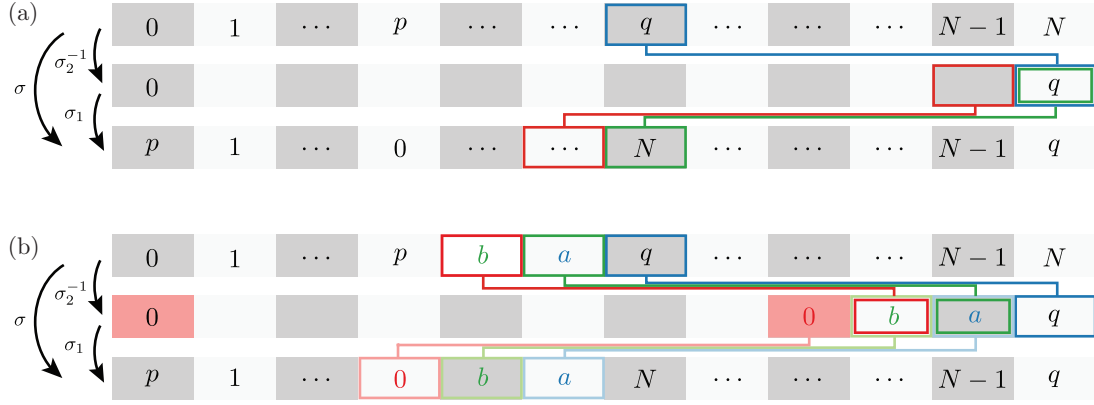
FIG. 1. (a) Three line representation of the sequential permutations σ_2^{-1} and σ_1. The existence of the transposition shown in blue in σ_2^{-1}, which has the inverse shown in green, implies that the shifted inverse transposition shown in red exists in σ_1. (b) Example graphical proof for Class B1.1 cycle structures. Repeated application of the permutation structure shown in (a) leads to a logical contradiction for $p < q$ as highlighted.

we can conclude that any cycle structure of Class B1.1 with $p < q$ does not admit a permutation of the correct structure. Repeating this process, however, starting from the assumption that $p > q$, generates no such contradiction, as shown in Animation 1 of the Supplemental Material [57], such that these cases are admissible. Counting the number of different possible choices of p and q satisfying this constraint ($p > q$) yields $(N-1)(N-2)/2$ acceptable permutations. An exhaustive application of these rules to the cycle types defined in Table III and for different possible orderings of p, q, r, s (as applicable) is depicted in the animations in the Supplemental Material [57]. Ultimately we find the index constraints listed in Table III along with the corresponding number of possible permutations. In the calculation of the number of possible permutations it should be remembered that cycles are unique only up to cyclic permutations. The graphical proof described above, for example, shows that $r < p < q$ and $q < p < r$ admit allowable Class B1.2 permutations, however, these different orderings yield identical permutations since the cycles (q, r) and (r, q) are equivalent. Care must hence be taken not to count such duplicate permutations multiple times. Finally, summation of all possibilities listed in Table III and some simple algebraic manipulation yields our desired result in the form of Eq. (28). We conjecture such a graphical proof by exhaustion could be extended to other values of S, however, we leave this to the enthusiastic reader.

4. Proof of Eq. (29)

Finally, we turn our attention to $S = N$, i.e., $K = 0$. We define $\mathbf{e}_i$ to be the standard basis vector of the $(N+1)$ dimensional real coordinate space whose ith entry is 1 and the rest of its entries are 0. The possible solutions of Eq. (A2) are $\mathbf{a}_A = \mathbf{e}_{N+1}$ and, for each $i \in \{1, \ldots, \lfloor (N+1)/2 \rfloor\}$, the vector $\mathbf{a}_B = \mathbf{e}_i + \mathbf{e}_{N+1-i}$.

In our proof of Eq. (29) we treat odd and even N separately. Consider first then the case where $N = 2n$ is even. The solutions for Case B and even N are immediately seen to be invalid since σ is an even permutation. Consequently, σ must have cycle type $\mathbf{e}_{N+1}$, i.e., it is an $N + 1$ cycle. Observe from Eq. (25) that σ is expressed as the product of σ_-, which by

definition is an $N + 1$ cycle, and another permutation of cycle type $\mathbf{e}_{N+1}$, since $\sigma_2\sigma_-^{-1}\sigma_2^{-1}$ is conjugate to σ_-. The number of permutations π of cycle type $\mathbf{e}_{N+1}$ such that $\sigma_-\pi$ also has cycle type $\mathbf{e}_{N+1}$ is known to be equal to $(2n)!/(n+1)$ [58,59]. Hence it immediately follows that

$$\nu(0, 2n) = (2n)!/(n+1), \tag{A10}$$

corresponding to Eq. (29) for even N.

We now consider the case for which $N = 2n + 1$ is odd, for which we find that Case A solutions are invalid due to permutation symmetry. Following the above logic, we thus wish to count the number of permutations π of cycle type $\mathbf{e}_{N+1}$ such that $\sigma_-\pi$ is the product of two disjoint cycles c_0 and c_N with $0 \in c_0$ and $N \in c_N$. We let $\Psi_k(N)$ denote the set of permutations $\pi = \sigma_2\sigma_-^{-1}\sigma_2^{-1}$ such that $\sigma_-\pi$ is the product of two disjoint cycles c_0 and c_N with $0 \in c_0$, $N \in c_N$, and the cycle c_0 consists of k elements. Furthermore, we let $\Gamma(N)$ denote the set of permutations π of cycle type $\mathbf{e}_{N+1}$ such that $\sigma_-\pi$ also has cycle type $\mathbf{e}_{N+1}$. First, we observe that there is a bijection from $\Psi_1(N)$ to $\Gamma(N-1)$. Indeed, it is straightforward to check that the cycle $(0, \alpha_0, \ldots, \alpha_{N-1}) \in \Psi_1(N)$ if and only if $(\alpha_0 - 1, \ldots, \alpha_{N-1} - 1) \in \Gamma(N-1)$. Using Eq. (A10), we therefore find that the cardinality of our sets satisfy

$$|\Psi_1(N)| = |\Gamma(N-1)| = (2n)!/(n+1). \tag{A11}$$

Next we establish a bijection between the sets $\Psi_k(N)$ and $\Psi_{k+1}(N)$ for $k \in \{1, \ldots, N-1\}$. Supposing that the permutation $\pi = (0, \alpha_0, \ldots, \alpha_{N-1}) \in \Psi_k(N)$, we can write

$$\sigma_-\pi = (0, \beta_1, \ldots, \beta_{k-1})(\gamma_1, \ldots, \gamma_{N-k}, N). \tag{A12}$$

Letting $g \in \{0, \ldots, N-1\}$ be the index such that $\alpha_g = \gamma_{N-k} + 1$ and setting $\pi' = (0, \alpha_g, \alpha_{g+1}, \ldots, \alpha_{g+N})$, with indices reduced modulo N, it follows that

$$\sigma_-\pi' = (0, \gamma_{N-k}, \beta_1, \ldots, \beta_{k-1})(\gamma_1, \ldots, \gamma_{N-k-1}, N), \tag{A13}$$

and hence $\pi' \in \Psi_{k+1}(N)$. This map from $\Psi_k(N)$ to $\Psi_{k+1}(N)$ has an inverse in the following sense: For $k \geqslant 1$, if we suppose that the permutation $\pi = (0, \alpha_0, \ldots, \alpha_{N-1}) \in \Psi_{k+1}(N)$, we have

$$\sigma_-\pi = (0, \beta_1, \ldots, \beta_{k-1}, \beta_k)(\gamma_1, \ldots, \gamma_{N-k-1}, N). \tag{A14}$$

Letting $g \in \{0, \dots, N-1\}$ be the index such that $\alpha_g = \beta_2 + 1$ if $k \geqslant 2$ and $\alpha_g = 1$ if $k = 1$, and furthering setting $\pi' = (0, \alpha_g, \alpha_{g+1}, \dots, \alpha_{g+N-1})$, with indices reduced modulo N, it then follows that

$$\sigma_- \pi' = (0, \beta_2, \dots, \beta_k)(\gamma_1, \dots, \gamma_{N-k-1}, \beta_1, N), \quad \text{(A15)}$$

whereby $\pi' \in \Psi_k(N)$.

We have thus established that $|\Psi_k(N)| = |\Psi_{k+1}(N)|$ for all $k \in \{1, \dots, N-1\}$. Next, noting that we can write $\nu(0, N) = \sum_{k=1}^{N} |\Psi_k(N)|$, we can deduce upon using Eq. (A11) that

$$\nu(0, 2n+1) = (2n+1)\frac{(2n)!}{n+1} = \frac{(2n+1)!}{n+1}. \quad \text{(A16)}$$

Finally, upon combining Eqs. (A10) and (A16) we obtain Eq. (29).

[1] A. Van Pamel, G. Sha, S. I. Rokhlin, and M. J. S. Lowe, Finite-element modelling of elastic wave propagation and scattering within heterogeneous media, Proc. R. Soc. **473**, 0738 (2017).

[2] M. C. W. van Rossum and T. M. Nieuwenhuizen, Multiple scattering of classical waves: Microscopy, mesoscopy, and diffusion, Rev. Mod. Phys. **71**, 313 (1999).

[3] I. V. Meglinski, V. L. Kuzmin, D. Y. Churmakov, and D. A. Greenhalgh, Monte Carlo simulation of coherent effects in multiple scattering, Proc. R. Soc. A **461**, 43 (2005).

[4] J. C. Ramella-Roman, S. A. Prahl, and S. L. Jacques, Three Monte Carlo programs of polarized light transport into scattering media: Part I, Opt. Express **13**, 4420 (2005).

[5] V. A. Markel, Extinction, scattering and absorption of electromagnetic waves in the coupled-dipole approximation, J. Quant. Spectrosc. Radiat. Transfer **236**, 106611 (2019).

[6] C. W. J. Beenakker, Random-matrix theory of quantum transport, Rev. Mod. Phys. **69**, 731 (1997).

[7] *The Oxford Handbook of Random Matrix Theory*, edited by G. Akemann, J. Baik, and P. Di Francesco (Oxford University Press, Oxford, 2015).

[8] A. Aiello and J. P. Woerdman, Physical bounds to the entropy-depolarization relation in random light scattering, Phys. Rev. Lett. **94**, 090406 (2005).

[9] W. Choi, A. P. Mosk, Q.-H. Park, and W. Choi, Transmission eigenchannels in a disordered medium, Phys. Rev. B **83**, 134207 (2011).

[10] M. Hanada, H. Shimada, and M. Tezuka, Universality in chaos: Lyapunov spectrum and random matrix theory, Phys. Rev. E **97**, 022224 (2018).

[11] N. Byrnes and M. R. Foreman, Universal bounds for imaging in scattering media, New J. Phys. **22**, 083023 (2020).

[12] S. J. Garratt and J. T. Chalker, Many-body delocalization as symmetry breaking, Phys. Rev. Lett. **127**, 026802 (2021).

[13] A. Aubry and A. Derode, Random matrix theory applied to acoustic backscattering and imaging in complex media, Phys. Rev. Lett. **102**, 084301 (2009).

[14] S. Rotter and S. Gigan, Light fields in complex media: Mesoscopic scattering meets wave control, Rev. Mod. Phys. **89**, 015005 (2017).

[15] C. W. Hsu, S. F. Liew, A. Goetschy, H. Cao, and A. Douglas Stone, Correlation-enhanced control of wave focusing in disordered media, Nat. Phys. **13**, 497 (2017).

[16] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, Ann. Math. **62**, 548 (1955).

[17] F. J. Dyson, Statistical theory of the energy levels of complex systems. I, J. Math. Phys. **3**, 140 (1962).

[18] F. J. Dyson, Statistical theory of the energy levels of complex systems. II, J. Math. Phys. **3**, 157 (1962).

[19] A. V. Andreev, B. D. Simons, and N. Taniguchi, Supersymmetry applied to the spectrum edge of random matrix ensembles, Nucl. Phys. B **432**, 487 (1994).

[20] N. Byrnes and M. R. Foreman, Polarisation statistics of vector scattering matrices from the circular orthogonal ensemble, Opt. Commun. **503**, 127462 (2022).

[21] A. Goetschy and A. D. Stone, Filtering random matrices: The effect of incomplete channel control in multiple scattering, Phys. Rev. Lett. **111**, 063901 (2013).

[22] A. Goetschy and S. E. Skipetrov, Non-Hermitian Euclidean random matrix theory, Phys. Rev. E **84**, 011150 (2011).

[23] A. Goetschy and S. E. Skipetrov, Euclidean random matrices and their applications in physics, arXiv:1303.2880.

[24] J. Berk and M. R. Foreman, Role of multiple scattering in single particle perturbations in absorbing random media, Phys. Rev. Res. **3**, 033111 (2021).

[25] P. Vivo and E. Vivo, Transmission eigenvalue densities and moments in chaotic cavities from random matrix theory, J. Phys. A: Math. Theor. **41**, 239801 (2008).

[26] M. L. Mehta, *Random Matrices*, 3rd ed. (Elsevier, Amsterdam, 2004).

[27] A. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications* (now Publishers Inc., Hanover, 2004).

[28] A. Crisanti, G. Paladin, and A. Vulpiani, *Products of Random Matrices in Statistical Physics*, Springer Series in Solid-State Sciences (Springer-Verlag, Berlin, 1998), Vol 104.

[29] H. A. Weidenmüller, Scattering theory and conductance fluctuations in mesoscopic systems, Physica A (Amsterdam, Neth.) **167**, 28 (1990).

[30] O. N. Dorokhov, Transmission coefficient and the localization length of an electron in *N* bound disordered chains, JETP Lett. **36**, 318 (1982).

[31] P. A. Mello, P. Pereyra, and N. Kumar, Macroscopic approach to multichannel disordered conductors, Ann. Phys. (NY) **181**, 290 (1988).

[32] K. A. Muttalib and V. A. Gopar, Generalization of the DMPK equation beyond quasi one dimension, Phys. Rev. B **66**, 115318 (2002).

[33] A. Douglas, P. Markoš, and K. A. Muttalib, The generalized DMPK equation revisited: Towards a systematic derivation, J. Phys. A: Math. Theor. **47**, 125103 (2014).

[34] N. Byrnes and M. R. Foreman, Random matrix theory of polarized light scattering in disordered media, Waves Random Complex Media 1 (2022).

[35] B. Efron, Bootstrap methods: Another look at the jackknife, Ann. Stat. **7**, 1 (1979).

[36] B. Efron, *The Jackknife, the Bootstrap, and Other Resampling Plans* (SIAM Press, Philadelphia, 1982).

[37] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman and Hall, New York, 1993).

[38] T. M. Nieuwenhuizen, and M. C. W. van Rossum, Intensity distributions of waves transmitted through a multiple scattering medium, Phys. Rev. Lett. **74**, 2674 (1995).

[39] D. Bicout, and C. Brosseau, Multiply scattered waves through a spatially random medium: Entropy production and depolarization, J. Phys. I **2**, 2047 (1992).

[40] P. Révész, *The Laws of Large Numbers* (Academic Press, New York, 1968).

[41] R. Berkovits and S. Feng, Correlations in coherent multiple scattering, Phys. Rep. **238**, 135 (1994).

[42] G. Osnabrugge, R. Horstmeyer, I. N. Padopoulos, B. Judkewitz, and I. M. Vellekoop, Generalized optical memory effect, Optica **4**, 886 (2017).

[43] N. Byrnes and M. R. Foreman, Symmetry constraints for vector scattering and transfer matrices containing evanescent components: Energy conservation, reciprocity and time reversal, Phys. Rev. Res. **3**, 013129 (2021).

[44] N. Fayard, A. Goetschy, R. Pierrat, and R. Carminati, Mutual information between reflected and transmitted speckle images, Phys. Rev. Lett. **120**, 073901 (2018).

[45] J. R. Ipsen, Products of independent Gaussian random matrices, Ph.D. thesis, Bielefeld University, 2015.

[46] M. Bona, *Combinatorics of Permutations*, 2nd ed. (CRC Press, New York, 2012).

[47] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data - The Theory of Improper and Noncircular Signals* (Cambridge University Press, Cambridge, 2010).

[48] The On-Line Encyclopedia of Integer Sequences, Sequence A005718, https://oeis.org/A005718.

[49] The On-Line Encyclopedia of Integer Sequences, Sequence A107991, https://oeis.org/A107991.

[50] V. P. Leonov and A. N. Shiryaev, On a method of calculation of semi-invariants, Theory Probab. Its Appl. (Engl. Transl.) **4**, 319 (1959).

[51] B. Eynard and M. L. Mehta, Matrices coupled in a chain: I. Eigenvalue correlations, J. Phys. A: Math. Gen. **31**, 4449 (1998).

[52] G. Mahoux, M. L. Mehta and J.-M. Normand, Matrices coupled in a chain: II. Spacing functions, J. Phys. A: Math. Gen. **31**, 4457 (1998).

[53] G. Akemann, M. Kieburg and L. Wei, Singular value correlation functions for products of Wishart random matrices, J. Phys. A: Math. Theor. **46**, 275205 (2013).

[54] M. L. Mehta and P. Shukla, Two coupled matrices: Eigenvalue correlations and spacing functions, J. Phys. A: Math. Gen. **27**, 7793 (1994).

[55] H. Furstenberg and H. Kesten, Products of random matrices, Ann. Math. Stat. **31**, 457 (1960).

[56] G. Akemann, and S. S. Byun, The product of m real $N \times N$ Ginibre matrices: Real eigenvalues in the critical regime, $m = O(N)$, Constr. Approx. **59**, 31 (2023).

[57] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.110.015308 for step by step animations of proof for each class.

[58] R. P. Stanley, Two enumerative results on cycles of permutations, Eur. J. Comb. **32**, 937 (2011).

[59] D. Zagier, On the distribution of the number of cycles of elements in symmetric groups, Nieuw Arch. Wisk. **13**, 489 (1995).