Information transfer through disordered media by diffuse waves

S. E. Skipetrov*

Laboratoire de Physique et Modélisation des Milieux Condensés, CNRS, 38042 Grenoble, France
(Received 18 October 2002; published 26 March 2003)

We consider the information content h of a scalar multiple-scattered, diffuse wave field $\psi(\mathbf{r})$ and the information capacity C of a communication channel that employs diffuse waves to transfer the information through a disordered medium. Both h and C are shown to be directly related to the mesoscopic correlations between the values of $\psi(\mathbf{r})$ at different positions $\mathbf{r}$ in space, arising due to the coherent nature of the wave. For the particular case of a communication channel between two identical linear arrays of $n \gg 1$ equally spaced transmitters or receivers (receiver spacing a), we show that the average capacity $\langle C \rangle \propto n$ and obtain explicit analytic expressions for $\langle C \rangle / n$ in the limit of $n \to \infty$ and $k\ell \to \infty$, where $k = 2\pi/\lambda$, λ is the wavelength, and ℓ is the mean free path. Modification of the above results in the case of finite but large n and $k\ell$ is discussed as well. If the signal to noise ratio S/N exceeds some critical value $(S/N)_c$, $\langle C \rangle / n$ is a nonmonotonic function of a, exhibiting maxima at $ka = m\pi$ ($m = 1, 2, \dots$). For smaller S/N, $ka = m\pi$ correspond to local minima, while the absolute maximum of $\langle C \rangle / n$ is reached at some $ka \sim (S/N)^{1/2} < \pi$. We define the maximum average information capacity $\langle C \rangle_{\max}$ as $\langle C \rangle$ maximized over the receiver spacing a and the optimal normalized receiver spacing $(ka)_{\mathrm{opt}}$ as the spacing maximizing $\langle C \rangle$. Both $\langle C \rangle_{\max} / n$ and $(ka)_{\mathrm{opt}}$ scale as $(S/N)^{1/2}$ for $S/N < (S/N)_c$, while $(ka)_{\mathrm{opt}} = m\pi$ and $\langle C \rangle_{\max} / n \sim \log(S/N)$ for $S/N > (S/N)_c$.

PACS number(s): 42.25.Dd, 84.40.Ua, 89.70.+c

I. INTRODUCTION

Transport of coherent waves in disordered media has been extensively studied during the past decades [1–6]. Remarkably similar, diffusion behavior of multiple-scattered electronic wave functions at low temperatures [1–3], coherent electromagnetic [4] (optical [5] and microwave [7]), acoustic [8], and elastic [9] waves has permitted impressive advances of the field [6]. Recently, multiple-scattered seismic waves in the earth's crust has been demonstrated to behave in a similar way [10,11]. In the context of these studies, the main quantities of interest are the transport coefficients of disordered samples, such as, e.g., the transmission coefficient T or the conductance g. The average values, fluctuations, full probability distributions, angular, spatial, and temporal correlation functions of T and g have been studied both theoretically and experimentally (see, e.g., Ref. [4] for a review). These quantities are, without any doubt, very important, since they describe the transport of wave *energy* through a disordered sample and hence can be measured experimentally. On the other hand, in practical applications one rarely uses multiple-scattered waves with a primary purpose of energy transmission. Much more often, waves are used to transfer the *information*. Readily available examples are microwave communications (portable telephony) in cities, indoor wireless local-area networks in buildings with complex structure, and underwater acoustic communication systems. In fact, one of the main motivations to study the multiple scattering of waves in disordered media is their possible use for transmission or processing of information in electronic devices or wireless communications.

In the same way as the properties of a disordered sample with respect to transmission of energy are described by the transmission coefficient T or conductance g (depending on the details of the specific experiment), its properties with respect to transmission of information are characterized by the *information capacity C*. The latter gives the maximum rate of error-free information transfer through a disordered sample using a given type of waves (acoustic, electromagnetic, etc.) and a given transmitter-receiver configuration (see Refs. [12–14] for a more rigorous definition of C). Information capacity of communication channels in disordered media has recently received considerable attention [15–21]. The most interesting and important result concerns a communication system consisting of multiple transmitters and receivers [15–17,19–21]: it has been found that the capacity of such a communication system scales linearly with the number of receivers n (as long as the number of transmitters m is of the same order). The authors of Refs. [16,19–21] have also studied the effect of correlations between the signals received by different receivers on C. More specifically, the communication channel between m transmitters and n receivers is described by a $n \times m$ Green matrix G ($G_{\alpha i}$ is the signal at the receiver α due to a unit signal emitted by the transmitter i). Due to the multiple scattering of waves, in a disordered medium the entries of G are random variables with certain correlations between them. These correlations are often termed "mesoscopic" [4] because they originate from the fact that the phase coherence length of the considered wave exceeds the length of the path that the wave travels inside the disordered medium. Mesoscopic correlations complicate significantly the theoretical calculation of capacity which otherwise (i.e., for uncorrelated $G_{\alpha i}$) is relatively straightforward [15,19,20]. Although it follows from Refs. [16,20] that nonzero correlations between $G_{\alpha i}$ reduce the information capacity of the communication channel under particular conditions considered in those papers, no systematic study of the role of mesoscopic correlations in the context of information transfer is available at the time of writing.
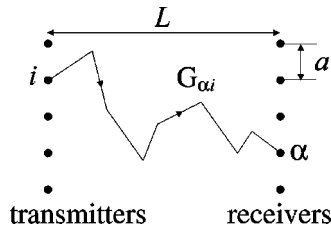
*Email address: Sergey.Skipetrov@grenoble.cnrs.fr

FIG. 1. We consider the transfer of information between identical arrays of n transmitters or receivers placed in a disordered medium. We denote the transmitter (receiver) spacing by a and the distance between the arrays by L. A wave emitted by some transmitter i experiences multiple scattering in the disordered medium before reaching the receiver α. The signal measured by the receiver α due to the transmitter i is given by the Green function $G_{\alpha i}$. The Green functions $G_{\alpha i}$ ($\alpha, i = 1, \ldots, n$) form a Green matrix G.

Mesoscopic correlations have been extensively studied during the past decade (see, e.g., Refs. [2,4,7] and references therein), but their role in the context of the information transfer has been completely ignored until very recently [16,19,20]. Meanwhile, understanding the relation between mesoscopic correlations and information capacity has not only the practical importance but also a fundamental significance, since the correlations are usually affected by the symmetries of the problem (translational symmetry, time-reversal symmetry, etc.) and it might be interesting to study the influence of these symmetries on the information-theoretic quantities, such as the information capacity.

The purpose of the present paper is to consider the multiple-scattered wave field from the point of view of the information theory, to quantify its information content, and, finally, to provide a comprehensive study of the information capacity of a communication channel in a disordered medium with a proper account for mesoscopic correlations arising from the coherent nature of multiple-scattered waves that carry the information. To be specific, we limit our consideration to identical linear arrays of n equally spaced transmitters or receivers (see Fig. 1). Such a geometry is common for microwave [7] and acoustic [8] experiments with multiple-scattered waves in disordered media [32]. Besides, it is also widely considered in connection with the time-reversed acoustics [22]. We note that the time-reversal techniques seem to be very promising for wireless communications in disordered and/or chaotic environments [23,24].

Before going into calculations, we would like to comment on the applied aspects of the considered problem. Communication systems, consisting of multiple transmitting-receiving antennas, are currently under an active study due to their potential importance for microwave wireless communications, in particular, wireless local-area networks and wireless telephony in cities [15–17,21,27]. In the 2-GHz region, the wavelength $\lambda \sim 15$ cm and the mean free path ℓ is of the order of meters for indoors and tens of meters for outdoors propagation [16]. If the transmitter-receiver separation $L \gtrsim 10$ m indoors and $L \gtrsim 100$ m outdoors, the wave emitted by the transmitter is multiply scattered before reaching the receiver. In the actual commercially available communication systems, the multiple scattering of the microwaves is treated as a nuisance, reducing the power of the useful, un-scattered signal, and adding up to the noise originating from other sources. Only recently, it was realized that one can make use of the multiple-scattered part of the signal to enhance the amount of information that can be transmitted through a disordered medium [25,15–17]. The architecture of the communication system based on this idea has been developed at Bell Laboratories and is called BLAST (Bell Labs Layered Space-Time) [25–27]. The experimental geometry considered in the present paper (Fig. 1) is close to the laboratory prototype developed at Bell Laboratories using BLAST architecture [26,27]. The analysis presented below is, therefore, of immediate applied importance in this context. Alternatively, in acoustic underwater communications, arrays of transducers have been already used for some time [28]. In this case, the spatial inhomogeneity of sound velocity gives rise to multiple paths for a wave to propagate from a given transmitter to a given receiver. The phenomenon may be further enhanced by scattering on fishes and ships and/or multiple reflections from the water-ground and water-air interfaces. The analysis presented below allows one to estimate the maximum achievable rate of error-free information transmission in this case, as well as to choose the optimal transducer spacing in the array. Finally, recent experiments [24] show that ultrasonic wave propagation in artificial disordered media (random arrangements of metallic rods [8]) provides a "tabletop" means for efficient modeling of communication systems that operate on a larger scale using different type of waves and higher carrier frequencies (e.g., microwaves in a building with complex structure or in a city). The present paper provides a theoretical support for further development of such model experiments.

The paper is organized as follows. We start by a study of the information content of the multiple-scattered wave field in Sec. II. The information content of a random field is quantified by its differential entropy. We show that mesoscopic correlations reduce the differential entropy, reducing the information content of the multiple-scattered wave field and leading, in practice, to a smaller volume of computer memory required for its storage. After this somewhat introductory part of the paper, in Sec. III we study the information capacity C of a communication channel between two identical linear arrays of n transmitters or receivers in a disordered medium (see Fig. 1). Under certain approximations, we obtain analytic expressions for the average information capacity $\langle C \rangle$. The latter expressions provide a reasonable estimate of $\langle C \rangle$ in real experimental situations, although one can do better using numerical methods. Next, we show that mesoscopic correlations reduce the information capacity (as found in Ref. [16]) only if the signal power is strong enough. If the signal is weak, correlations play a positive role and allow a higher capacity as compared to that in the absence of correlations. Finally, we define the maximum capacity as a capacity maximized over the transmitter (receiver) spacing and the optimal transmitter (receiver) spacing as the spacing maximizing the capacity. Simple analytic expressions are found for these two quantities. We summarize the main results of the paper in Sec. IV. Derivations of some important equations used in the main text but not essential for its understanding are provided in Appendixes A and B.
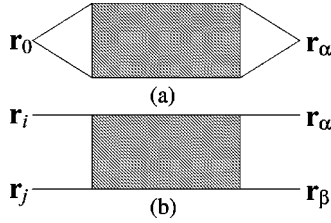
FIG. 2. (a) The diagram contributing to the average intensity $\langle I(\mathbf{r}_\alpha)\rangle$ for a point source of radiation at $\mathbf{r}_0$. The shaded rectangle is the ladder propagator, the solid lines denote the retarded and advanced average Green functions. (b) The diagram for the correlation function of the Green functions $\langle G_{\alpha i}G^*_{\beta j}\rangle$.

II. INFORMATION CONTENT OF THE MULTIPLE-SCATTERED WAVE FIELD

We start our analysis by considering a problem which is somewhat simpler than the one depicted in Fig. 1. Namely, we assume that there is only one transmitter (located, say, at $\mathbf{r}_0$) that emits a monochromatic scalar spherical wave, and analyze the information content of the multiple-scattered wave field $\psi(\mathbf{r})$. As a consequence of multiple scattering, $\psi(\mathbf{r})$ is a random function of position $\mathbf{r}$ ("speckle pattern"). Since $\psi(\mathbf{r})$ is a complex continuous random field, it contains an infinite amount of information or, in other words, an infinite volume of computer memory would be required to store its values at all $\mathbf{r}$ with absolute accuracy. In reality, however, one usually measures $\psi(\mathbf{r})$ at some finite number n of positions $\mathbf{r}_\alpha$ ($\alpha = 1, \ldots, n$). In the following, we identify the information content of such a measurement with the information content of $\psi(\mathbf{r})$, keeping in mind that this is strictly true only for $n \to \infty$. A measurement of ψ in n points $\mathbf{r}_\alpha$ results in a random complex vector $\mathbf{y} = \{y_\alpha\}$, where $y_\alpha = \psi(\mathbf{r}_\alpha)/\langle I\rangle^{1/2}$ and we normalize the field $\psi(\mathbf{r})$ by the square root of the average intensity $\langle I\rangle = \langle |\psi(\mathbf{r}_\alpha)|^2\rangle$ given by the diagram (a) of Fig. 2 and assumed to be the same for all n receivers. From here on the angular brackets denote averaging over disorder. If $p(\mathbf{y})$ is a probability density function of the random vector $\mathbf{y}$, the *differential entropy* of $\mathbf{y}$ is defined as [12,13]

$$h(\mathbf{y}) = -\int p(\mathbf{y})\log p(\mathbf{y}) d^n\mathbf{y}. \tag{1}$$

The logarithm is to base 2 and h is hence measured in *bits*. The differential entropy h is a measure of uncertainty of the random vector $\mathbf{y}$ and reflects the average amount of information that one receives when some value of $\mathbf{y}$ is observed. The larger the differential entropy, the larger the information content of the vector $\mathbf{y}$, and we will therefore use h to quantify the information content of $\mathbf{y}$ and, consequently, of the multiple-scattered wave field $\psi(\mathbf{r})$.

Under conditions of strong multiple scattering and provided that $|\mathbf{r}_\alpha - \mathbf{r}_0| \gg \ell$ for all $\alpha = 1, \ldots, n$ and that $k\ell \gg 1$ (where $k = 2\pi/\lambda$, λ is the wavelength, and ℓ is the mean free path), the wave field $\psi(\mathbf{r})$ can be considered Gaussian to a good accuracy, and hence $\mathbf{y}$ is a circularly symmetric complex Gaussian random vector described by a Gaussian probability density function

$$p(\mathbf{y}) = \det(\pi K)^{-1}\exp(-\mathbf{y}^+ K^{-1}\mathbf{y}), \tag{2}$$

where K is a covariance matrix: $K_{\alpha\beta} = \langle y_\alpha y_\beta^*\rangle$. The integration in Eq. (1) can be then carried out analytically, yielding

$$h(\mathbf{y}) = \log\det(\pi e K). \tag{3}$$

It also can be shown [13] that the Gaussian density function (2) maximizes the differential entropy for a given covariance matrix K. Recalling that $K_{\alpha\beta}$ is the correlation function of the multiple-scattered wave field, we can readily compute K in the ladder approximation [see the diagram (b) of Fig. 2 with $\mathbf{r}_i = \mathbf{r}_j = \mathbf{r}_0$ or Ref. [29]]:

$$K_{\alpha\beta} = \frac{\sin(k\Delta r_{\alpha\beta})}{k\Delta r_{\alpha\beta}}\exp\left(-\frac{\Delta r_{\alpha\beta}}{2\ell}\right), \tag{4}$$

where $\Delta r_{\alpha\beta} = |\mathbf{r}_\alpha - \mathbf{r}_\beta|$. We now restrict ourselves to the case of measurement points $\mathbf{r}_\alpha$ arranged in a line with a constant distance a between $\mathbf{r}_\alpha$ and $\mathbf{r}_{\alpha+1}$. As we already mentioned in Sec. I, such a measurement geometry is common for both microwave [7] and acoustic [8,22] experiments. The covariance matrix K is then Toeplitz: $K_{\alpha\beta} = K_{\alpha-\beta}$, where $K_\gamma = \sin(\gamma ka)/(\gamma ka)\exp[-|\gamma|ka/(2\ell)]$, and hence in the limit $n \to \infty$ the density of its eigenvalues tends to a limit, which is the spectrum of $\mathbf{y}$. The differential entropy rate can be then expressed through the power spectral density of $\mathbf{y}$ [30],

$$f(\mu) = \sum_{\gamma=-\infty}^{\infty} K_\gamma \exp(i\gamma\mu). \tag{5}$$

The resulting $h(\mathbf{y})$ appears to scale linearly with n and hence it is convenient to consider the *differential entropy rate* or the differential entropy per receiver,

$$\mathcal{H}(\mathbf{y}) = \lim_{n\to\infty}\frac{h(\mathbf{y})}{n} = \log(\pi e) + \frac{1}{2\pi}\int_0^{2\pi}\log f(\mu)d\mu. \tag{6}$$

Equation (4) allows us to compute $f(\mu)$ explicitly:

$$f(\mu) = 1 + \frac{1}{ka}\left\{\arctan\frac{\sin(ka+\mu)}{\exp[ka/(2k\ell)]-\cos(ka+\mu)}\right.$$
$$\left.+\arctan\frac{\sin(ka-\mu)}{\exp[ka/(2k\ell)]-\cos(ka-\mu)}\right\}. \tag{7}$$

This expression simplifies greatly in the limit $k\ell \to \infty$:

$$f(\mu) = \frac{\pi}{ka}(m_+ + m_- + 1), \tag{8}$$

where $m_\pm$ denotes the greatest integer not larger than $(ka \pm \mu)/(2\pi)$. We show the differential entropy rate following from Eqs. (6)–(8) in Fig. 3 by solid lines for $k\ell = 10$, 100, and $k\ell \to \infty$. In order to demonstrate the convergence of the exact formula (3) to the asymptotic Eq. (6), we also show the curves following from Eq. (3) for $n = 50$ (dashed lines) and the same values of $k\ell$. It is worthwhile to note that the entropy rate $\mathcal{H}$ reaches its maximum value $\log(\pi e)$ at $ka = m\pi$ (where m is a positive integer) and that for $ka > \pi$ the
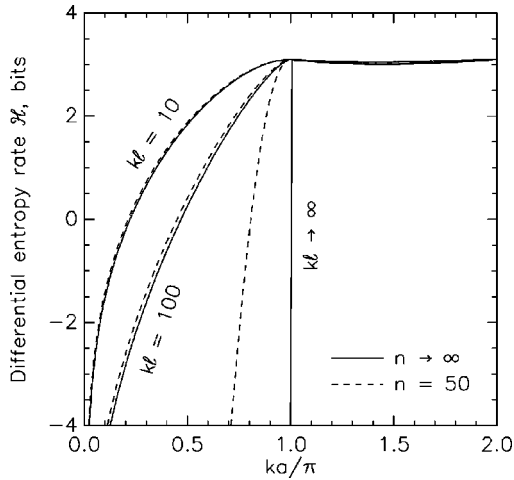
FIG. 3. Differential entropy rate $\mathcal{H}$ as a function of normalized receiver spacing for the multiple-scattered wave field measured by a linear array of $n \to \infty$ (solid lines) and $n = 50$ (dashed lines) receivers at $k\ell = 10$, 100, and $k\ell \to \infty$.

variations of $\mathcal{H}$ with ka are extremely weak. Exact analytic result for $\mathcal{H}$ can be obtained in the limit $k\ell \to \infty$ by substituting $f(\mu)$ given by Eq. (8) into Eq. (6) and performing the integration. We find $\mathcal{H} \to -\infty$ for $ka < \pi$ and $\mathcal{H} = ka/\pi + \log[(\pi e/2)(\pi/ka)]$ for $\pi < ka < 2\pi$. At $ka > 2\pi$, $\mathcal{H}$ shows only weak deviations from its maximum value $\log(\pi e)$.

Let us now discuss the implication of Fig. 3 for experimental measurements. In an experiment, the values of y_α cannot be measured with an absolute accuracy. An m-bit quantization of Rey_α and Imy_α with a quantization step $\Delta \sim 2^{-m}$ is a common procedure [33] and the measured quantized y'_α can take $\sim \Delta^{-2}$ discrete values. The vector $\mathbf{y}' = \{y'_\alpha\}$ can therefore take $\sim \Delta^{-2n}$ different values. One can show that the entropy $H(\mathbf{y}')$ of $\mathbf{y}'$ is approximately $h(\mathbf{y}) - 2n \log \Delta$ for $\Delta \to 0$. The number of bits required on average to describe a given component y'_α of the random vector $\mathbf{y}'$ is then $H(\mathbf{y}')/n \sim \mathcal{H}(\mathbf{y}) - 2 \log \Delta$. It can be now easily seen that the smaller differential entropy rate $\mathcal{H}(\mathbf{y})$ means that less bits will be required in an experiment to record all the relevant information about the multiple-scattered wave field. As follows from Fig. 3, at large receiver spacing a ($ka > \pi$), $\mathcal{H}$ is very close to its maximum value. At such a large receiver spacing, the signals measured by different receivers are only weakly correlated and hence each signal contains a relatively large amount of information. One therefore needs a relatively large number of bits per receiver $H(\mathbf{y}')/n$ to record the speckle pattern $\psi(\mathbf{r})$ in this case. In contrast, at small a ($ka < \pi$), $\mathcal{H}$ decreases, since the signals measured by different receivers become significantly correlated and hence the information contained in each of the signals on average, $H(\mathbf{y}')/n$, is smaller than at $ka > \pi$. If $k\ell$ or n is finite, the value of $\mathcal{H}$ at $ka < \pi$ can be small but remains finite. If, in contrast, we take the limit of $k\ell \to \infty$ and $n \to \infty$, $\mathcal{H} \to -\infty$ at $ka < \pi$. This means that an infinitely small amount of information is contained in the signal measured by a given receiver and is a direct consequence of the fact that in the absence of exponential damping of correlation

in Eq. (4), the correlation range is infinite and the decrease of $h(\mathbf{y})$ with n is faster than linear.

III. INFORMATION CAPACITY OF A COMMUNICATION CHANNEL IN A DISORDERED MEDIUM

A. General definitions

After having considered the information content of the multiple-scattered wave field, we are now in a position to analyze the central problem of the present paper. We consider a communication channel between two identical linear arrays of n equally spaced transmitters or receivers shown in Fig. 1. The vector of emitted signals $\mathbf{x}$ and the vector of received signals $\mathbf{y}$ are related by $\mathbf{y} = G\mathbf{x} + \mathbf{z}$, where G is a $n \times n$ complex Green matrix ($G_{\alpha i}$ gives the signal measured at the receiver α due to a unit signal emitted by the transmitter i), and $\mathbf{z}$ is a noise vector. We consider scalar waves and assume that the noises z_α at different receivers are statistically independent, normally distributed random variables with power N: $\langle z_\alpha z_\beta^* \rangle = N \delta_{\alpha\beta}$. Before defining the information capacity of such a communication channel, we first remind the definitions of the *conditional differential entropy* [13]

$$h(\mathbf{y}|\mathbf{x}) = -\int \int p(\mathbf{x},\mathbf{y}) \log p(\mathbf{y}|\mathbf{x}) d^n \mathbf{x} \, d^n \mathbf{y} \quad (9)$$

and *mutual information* between two random vectors $\mathbf{x}$ and $\mathbf{y}$ [13]:

$$\mathcal{I}(\mathbf{x},\mathbf{y}) = h(\mathbf{y}) - h(\mathbf{y}|\mathbf{x}), \quad (10)$$

where $p(\mathbf{x},\mathbf{y})$ and $p(\mathbf{y}|\mathbf{x})$ are the joint and conditional probability density functions of $\mathbf{x}$ and $\mathbf{y}$, respectively. In our case, $\mathbf{y} = G\mathbf{x} + \mathbf{z}$, and one finds $h(\mathbf{y}|\mathbf{x}) = h(\mathbf{z})$ and

$$\mathcal{I}(\mathbf{x},\mathbf{y}) = \log \det\left[I_n + \frac{1}{N} G^+ Q G \right]. \quad (11)$$

Here I_n is the $n \times n$ unit matrix and we assume that $\mathbf{x}$ is a circularly symmetric complex Gaussian random vector with a covariance matrix Q: $Q_{ij} = E[x_i x_j^*]$, where $E[\cdots]$ denotes the averaging over all possible emitted signals $\mathbf{x}$ and should be contrasted from the disorder averaging that we denote by the angular brackets.

If we assume that the Green matrix G is known at both transmitters and receivers, the Shannon *information capacity* C (or simply capacity for brevity) of the communication channel is found by maximizing the mutual information (10) over all possible distributions $p(\mathbf{x})$ of emitted signals $\mathbf{x}$. The fundamental importance of C is that it gives the largest information transfer rate R that can be realized for a given information channel *in principle* with infinitely small probability of error [12,13]. Although no general procedure exists to realize $R = C$ in practice, in many real situations one can achieve information transfer rates that are quite close to C [34]. Since the Gaussian distribution maximizes the differential entropy h (and hence the mutual information $\mathcal{I}$) at a given Q (see Sec. II and Ref. [13]), maximizing Eq. (10)

over $p(\mathbf{x})$ amounts to maximize Eq. (11) over Q. In practice, however, G is often known only at the receivers, but not at the transmitters, where only statistical information about G is available. It is the case that we consider in the remaining part of the paper. The optimal Q is then the matrix maximizing the *average* mutual information $\langle \mathcal{I}(\mathbf{x},\mathbf{y}) \rangle$. To accomplish the averaging, we need to specify the statistical properties of the Green matrix G. In the considered case of strong multiple scattering, provided that the distance L between the arrays of transmitters and receivers is much larger than the mean free path ℓ and that $k\ell \gg 1$, $G_{\alpha i}$ is a circularly symmetric complex Gaussian random variable with zero mean and covariance given by the diagram (b) of Fig. 2: $\langle G_{\alpha i} G_{\beta j}^* \rangle = \langle I \rangle K_{\alpha\beta} K_{ij}$, where $K_{\alpha\beta}$ is defined in Eq. (4) and $\langle I \rangle = \langle |G_{\alpha i}|^2 \rangle$ is assumed to be independent of α and i. In the following we adopt the total emitted power constraint $\mathrm{Tr}Q \leqslant n$ and introduce a normalized Green matrix $\mathcal{G} = G/(\langle I \rangle n)^{1/2}$. In the present paper we will only be interested in the average capacity $\langle C \rangle$, although it should be kept in mind that C exhibits random fluctuations as disorder is varied. As follows from the above reasoning, the average capacity is

$$\langle C \rangle = \max_{Q} \left\langle \log \det \left[I_n + \frac{S}{N} \mathcal{G}^+ Q \mathcal{G} \right] \right\rangle, \quad (12)$$

where S/N plays the role of the signal to noise ratio and $S = n\langle I \rangle$ is the average power received by each receiver assuming independent signals from transmitters. When Eq. (12) is applied to a real situation, it gives the maximum amount of information that can be transferred through the considered communication channel per second using a unit-frequency bandwidth [12,13]. The units of $\langle C \rangle$ are therefore bits per second per Hertz (or bps/Hz for brevity).

B. Average capacity at $n \gg 1$

For a small number of transmitters and receivers $n \sim 1$, the averaging and maximization over Q in Eq. (12) can be carried out by a numerical simulation (see, e.g., Ref. [20] for $n=2$ and Ref. [21] for $n=2$ and 3). At large n such an approach becomes inadequate. It appears, however, that analytic methods can be applied to estimate the asymptotic behavior of capacity for $n \gg 1$ (see Appendix A and Refs. [16,31,19]). In the limit of large n the capacity per receiver is given by

$$\frac{\langle C \rangle}{n} = \frac{1}{n} \sum_{i=1}^{n} \log[(S/N)^{-1/2} + \kappa_i q_i u] + \frac{1}{n} \sum_{i=1}^{n} \log[(S/N)^{-1/2}$$

$$+ \kappa_i v] - uv/\ln 2 + \log(S/N), \quad (13)$$

where κ_i are the eigenvalues of the matrix K, while the auxiliary variables u, v, and $p \leqslant n$ nonzero eigenvalues q_i of the matrix Q are solutions of the following system of equations (see Appendix A for derivations):

$$u = \frac{1}{n} \sum_{i=1}^{n} \frac{\kappa_i}{(S/N)^{-1/2} + \kappa_i v}, \quad (14)$$

$$v = \frac{1}{n} \sum_{i=1}^{n} \frac{\kappa_i q_i}{(S/N)^{-1/2} + \kappa_i q_i u}, \quad (15)$$

$$\phi = \frac{\kappa_i u}{(S/N)^{-1/2} + \kappa_i q_i u}, \quad i = 1, \ldots, p, \quad (16)$$

$$n = \sum_{i=1}^{p} q_i. \quad (17)$$

In principle, the above equations are sufficient for the calculation of the average capacity at given $n \gg 1$, S/N, and ka. It should be noted, however, that the total number of equations is $p+3$ (with p that can be as large as n), and that the equations are nonlinear. Hence, the numerical solution requires considerable computational resources at large n. Besides, the interpretation of numerical results is known to be a rather difficult task. Below we show that in the limit of $n \to \infty$, Eqs. (13)–(17) can be significantly simplified and even that simple *analytic* expressions for $\langle C \rangle$ can be obtained in certain cases.

C. Average capacity at $n \to \infty$

As we already mentioned in Sec. II, the covariance matrix K is Toeplitz and hence for $n \to \infty$ one can use the limit theorems known for this class of matrices [30]. Of particular use for us is the fundamental eigenvalue distribution theorem of Szegö that states that if K is an $n \times n$ Hermitian Toeplitz matrix and $F(x)$ is some continuous function then (under certain conditions fulfilled in our case)

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} F(\kappa_i) = \frac{1}{2\pi} \int_0^{2\pi} F[f(\mu)] d\mu, \quad (18)$$

where $f(\mu)$ is the power spectral density defined in Eq. (5). We now admit that the right-hand sides of Eqs. (13)–(15) can be readily simplified using Eq. (18). After some algebra this leads to the following set of equations for the auxiliary variables u and v:

$$u = \frac{1}{2\pi} \int_0^{2\pi} f(\mu) \left[\left(\frac{S}{N} \right)^{-1/2} + v f(\mu) \right]^{-1} d\mu, \quad (19)$$

$$v = \frac{1}{u} \frac{1}{2\pi} \int_{0 \ f(\mu) > (S/N)^{-1/2} v}^{2\pi} \left[1 - \left(\frac{S}{N} \right)^{-1/2} \frac{v}{f(\mu)} \right] d\mu, \quad (20)$$

while the expression (13) for the average capacity per receiver becomes

$$\frac{\langle C \rangle}{n} = \frac{1}{2\pi} \int_{0 \ f(\mu) > (S/N)^{-1/2} v}^{2\pi} \log \left[\frac{f(\mu)}{v} \right] d\mu$$

$$+ \frac{1}{2\pi} \int_0^{2\pi} \log \left[\left(\frac{S}{N} \right)^{-1/2} + v f(\mu) \right] d\mu + \log \left(\frac{S}{N} \right)$$

$$\times \left(1 - \frac{1}{4\pi} \int_{0 \ f(\mu) < (S/N)^{-1/2} v}^{2\pi} d\mu \right) - \frac{uv}{\ln 2}, \quad (21)$$
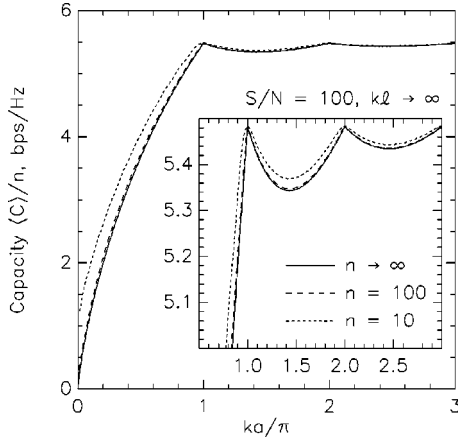
FIG. 4. Average information capacity per receiver of a communication channel between two identical linear arrays of n equally spaced transmitters or receivers as a function of normalized receiver spacing for $k\ell \to \infty$ and $n \to \infty$ (solid line), $n = 100$ (dashed line), and $n = 10$ (dotted line). The signal to noise ratio is $S/N = 100$. The inset is a zoom of the main plot.

where the integral in Eq. (20) and the first integral in Eq. (21) are over the part of the interval $(0, 2\pi)$ where $f(\mu) > (S/N)^{-1/2}v$, while the last integral in Eq. (21) is over the part of the same interval where $f(\mu) < (S/N)^{-1/2}v$. Once $f(\mu)$ is known, Eqs. (19)–(21) allow one to calculate the average capacity per receiver, $\langle C \rangle / n$, in the limit $n \to \infty$.

1. Average capacity at $n \to \infty$ and $k\ell \to \infty$

We first consider the limit $k\ell \to \infty$. Although particularly simple results can be obtained in this case, we will show later that this limit can serve as a good approximation to real situations with large but finite $k\ell$. If $0 < ka < \pi$, we find (see Appendix B)

$$\frac{\langle C \rangle}{n} = \frac{ka}{\pi} \log \left[\frac{S}{N} \frac{\pi}{ka} \frac{1}{\phi} \right] - \frac{\phi}{\ln 2}, \qquad (22)$$

where

$$\phi = \frac{ka}{\pi} - \frac{1}{2(S/N)} \left(\frac{ka}{\pi} \right)^3 \left[\sqrt{1 + 4\frac{S}{N} \left(\frac{\pi}{ka} \right)^2} - 1 \right]. \quad (23)$$

Analytic results for $\langle C \rangle / n$ can also be obtained at $ka > \pi$, but the resulting expressions are rather cumbersome and lengthy and we present them in Appendix B. In Fig. 4, we show the average capacity per receiver $\langle C \rangle / n$, as a function of normalized receiver spacing ka/π at a fixed signal to noise ratio $S/N = 100$. The solid line shows our analytic result, corresponding to the limit $n \to \infty$ [Eq. (22) at $0 < ka < \pi$ and more lengthy but analytic formulas given in Appendix B at $ka > \pi$], while the dashed and dotted lines are obtained by a numerical solution of Eqs. (13)–(17) at finite number of receivers n ($n = 100$ and 10, respectively). As follows from our analysis, at $k\ell \to \infty$ and $n \to \infty$, the derivative of $\langle C \rangle$ with respect to ka exhibits a discontinuity at $ka = m\pi$ ($m = 1, 2, \ldots$). Although the discontinuity disappears at finite n, the overall agreement between the results
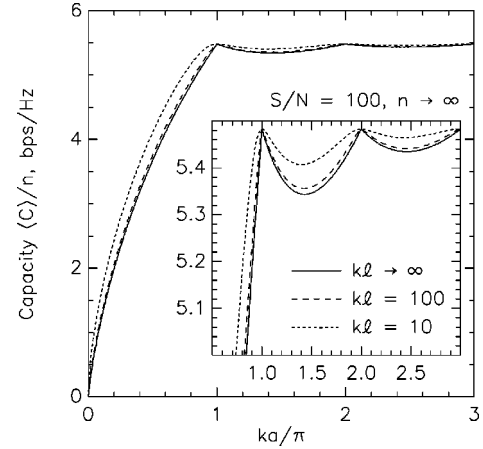

FIG. 5. Average information capacity per receiver of a communication channel between two identical linear arrays of n equally spaced transmitters or receivers as a function of normalized receiver spacing for $n \to \infty$ and $k\ell \to \infty$ (solid line), $k\ell = 100$ (dashed line), and $k\ell = 10$ (dotted line). The signal to noise ratio is $S/N = 100$. The inset is a zoom of the main plot.

corresponding to finite n and $n \to \infty$ remains satisfactory even for n as small as $n = 10$. The analytic results corresponding to the limit of $n \to \infty$ can therefore serve as a reasonable approximation in real situations with large but finite n.

2. Average capacity at $n \to \infty$ and finite $k\ell$

Although $k\ell$ is large in the experiments performed in the diffusion regime [7,8,22,27], its value remains finite and it is therefore of interest to consider its effect on the average capacity. At finite $k\ell$ the power spectral density $f(\mu)$ is given by Eq. (7) and Eqs. (19)–(21) can be solved only numerically. The solution is, however, quite simplified by the fact that the result corresponding to $k\ell \to \infty$ is known analytically (see Appendix B) and can be used as a good starting point for the numerical algorithm. The average capacity per receiver obtained from Eqs. (19)–(21) at $k\ell = 100$ and 10 is shown in Fig. 5 by dashed and dotted lines, respectively. The solid line shows the $k\ell \to \infty$ result, the same as in Fig. 4. As follows from Fig. 5, at finite $k\ell$ the capacity is somewhat higher than at $k\ell \to \infty$. This is explained by a lower degree of correlation between the entries of the Green matrix G for smaller $k\ell$ [see Eq. (4)]. Also, the derivative of $\langle C \rangle$ with respect to ka exhibits no jumps at $ka = m\pi$ ($m = 1, 2, \ldots$) when $k\ell$ is finite. In general, however, the average capacity is only slightly affected by the finiteness of $k\ell$ as long as $k\ell$ remains much larger than unity.

D. Maximum capacity and optimal receiver spacing

It follows from Figs. 4 and 5 that as long as $n \gg 1$ and $k\ell \gg 1$, the average capacity per receiver is very close to its value for $n \to \infty$ and $k\ell \to \infty$. We therefore limit the rest of this section to the discussion of the latter limiting case, assuming that the behavior of $\langle C \rangle / n$ remains similar at finite but large n and $k\ell$. The typical behavior of the average capacity per receiver shown in Figs. 4 and 5 can be summarized as follows: $\langle C \rangle / n$ has its absolute minimum at ka
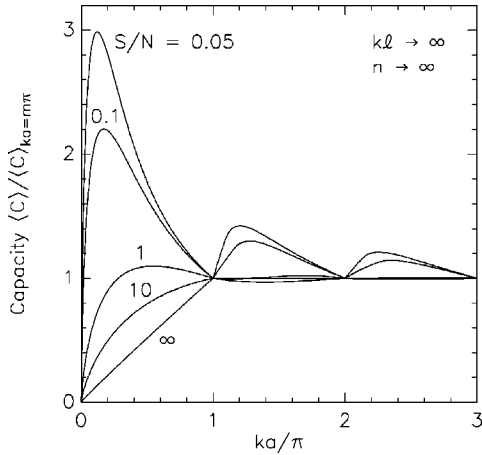
FIG. 6. Average capacity of a communication channel between two identical linear arrays of n transmitters or receivers, normalized to its value at $ka=m\pi$, is shown as a function of the normalized receiver spacing for five different values of the signal to noise ratio S/N. We assume $k\ell\to\infty$ and $n\to\infty$ for this plot.

$\to 0$ while it reaches its maxima at $ka=m\pi$ ($m=1,2,\dots$) and becomes almost independent of ka for $ka>\pi$. It appears, however, that such a behavior is typical only for *large* signal to noise ratios $S/N>(S/N)_c$, where $(S/N)_c$ is some critical value of S/N that we define below. To study the capacity as a function of ka at various values of S/N, we plot the capacity normalized to its value at $ka=m\pi$ in Fig. 6.

As follows from Fig. 6, at large S/N ($S/N=10$ and $S/N \to\infty$) the behavior of capacity with ka is similar to that shown in Figs. 4 and 5. Interestingly, in the limit $S/N\to\infty$ we find a very simple result: $\langle C\rangle/n=(ka/\pi)\log(S/N)$ at $ka <\pi$ and $\langle C\rangle/n=\log(S/N)$ at $ka>\pi$, i.e., the capacity grows linearly with ka for $ka<\pi$ and then remains constant for $ka>\pi$. At finite but large S/N the behavior of capacity is less simple but is qualitatively very similar: $\langle C\rangle/n$ first shows a monotonic increase with ka for $ka<\pi$ and then oscillates weakly with ka for $ka>\pi$ (see also Figs. 4 and 5). As we noted above, such a behavior is typical only for $S/N >(S/N)_c$. At smaller values of the signal to noise ratio, $\langle C\rangle/n$ exhibits a nonmonotonic behavior with ka for $ka <\pi$. More precisely, it reaches a maximum at some $ka <\pi$, as can be seen from Fig. 6 for $S/N=0.05$, 0.1, and 1. We call the value of ka maximizing the average information capacity *optimal* and denote it by $(ka)_{\text{opt}}$. In addition, we define the *maximum* capacity $\langle C\rangle_{\text{max}}$ as the capacity maximized over ka:

$$\langle C\rangle_{\text{max}}=\underset{ka}{\text{max}}\,\underset{Q}{\text{max}}\left\langle \log\det\left[I_n+\frac{S}{N}\mathcal{G}^+Q\mathcal{G}\right]\right\rangle. \quad (24)$$

It follows from Eq. (22) that $(ka)_{\text{opt}}$ and $\langle C\rangle_{\text{max}}$ depend on the signal to noise ratio S/N in a very simple way:

$$(ka)_{\text{opt}}=\begin{cases} A_1\sqrt{S/N}, & S/N<(S/N)_c \\ m\pi, & S/N>(S/N)_c, \end{cases} \quad (25)$$
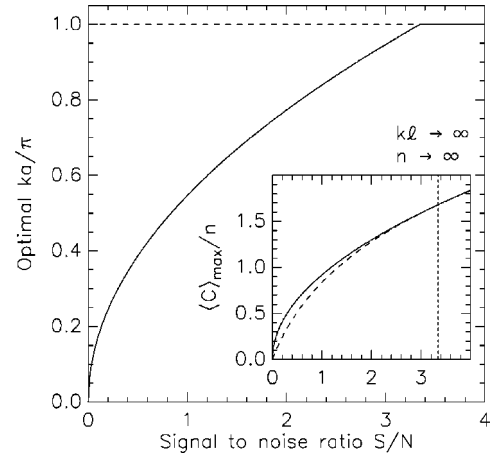

FIG. 7. Normalized optimal receiver spacing, maximizing the information capacity, as a function of the signal to noise ratio. The inset shows the maximum average information capacity per receiver $\langle C\rangle_{\text{max}}/n$. The dashed line is $\langle C\rangle/n$ at $ka=m\pi$ ($m=1,2,\dots$), coinciding with $\langle C\rangle_{\text{max}}/n$ for $S/N>(S/N)_c\simeq 3.35$ but smaller than the latter for $S/N<(S/N)_c$. The vertical dotted line is $S/N =(S/N)_c$.

$$\frac{\langle C\rangle_{\text{max}}}{n}=\begin{cases} A_2\sqrt{S/N}, & S/N<(S/N)_c \\ \log[(S/N)/\phi_{\text{max}}]-\phi_{\text{max}}/\ln 2, & S/N>(S/N)_c, \end{cases} \quad (26)$$

where $(S/N)_c\simeq 3.35$, $A_1\simeq 1.72$ and $A_2\simeq 0.92$ are numerical constants, and

$$\phi_{\text{max}}=1-\frac{1}{2S/N}(\sqrt{1+4S/N}-1). \quad (27)$$

The normalized optimal receiver spacing $(ka)_{\text{opt}}$ is shown in Fig. 7, while the maximum capacity $\langle C\rangle_{\text{max}}$ is shown, in the inset. For comparison, we also show the capacity corresponding to $ka=m\pi$ (the dashed line in the inset of Fig. 7) which coincides with the maximum capacity for $S/N >(S/N)_c$ but is smaller than the latter if $S/N<(S/N)_c$. It is interesting to note that at $ka=m\pi$ the average capacity scales linearly with S/N for $S/N\to 0$, while the average capacity maximized over ka is proportional to the square root of S/N in the same limit. Hence, the latter can exceed the former significantly for small signal to noise ratios.

Qualitatively different behavior of the average capacity at small and large signal to noise ratios, illustrated in Figs. 6 and 7, can be understood without any lengthy calculations. We first remind that at $ka<\pi$, according to Eq. (4), smaller ka means stronger correlation between the Green functions $G_{\alpha i}$. Next, we map the quite complicated communication channel shown in Fig. 1 onto an equivalent set of n' independent communication channels, the capacity of each equivalent channel being $\sim\log(1+S'/N')$, where S' and N' denote the signal and noise powers in each of n' independent channels. Obviously, the number n' of equivalent independent channels grows with decreasing the correlations between the entries of the Green matrix G. Consequently, since the total capacity is $n'\log(1+S'/N')$, it seems that having zero correlations (and hence $ka=m\pi$) should always maxi-

mize the capacity because it ensures the largest n'. This reasoning, however, is not correct, since the signal power S' is also sensitive to the correlations between the entries of the Green matrix G. Indeed, partially correlated Green functions $G_{\alpha i}$ lead to a constructive interference of scattered waves at the receivers, thus increasing the power of the received signal, while the noise power N' remains unchanged. Therefore, when changing the receiver spacing ka from 0 to π, one gradually switches from a small number of equivalent independent channels n' with relatively large signal to noise ratio S'/N' to a larger number of equivalent independent channels n' with weaker S'/N'. If S'/N' is large (which is only possible if S/N is large), the capacity of each independent channel $\log(1+S'/N') \approx \log(S'/N')$ depends on S'/N' only logarithmically, while the dependence of the total capacity on n' is linear. To achieve the maximum capacity one therefore needs to choose the largest $n' = n'_{\max}$ (which corresponds to $ka = m\pi$), the decrease of capacity of each equivalent channel being negligible due to its weak (logarithmic) dependence on the signal to noise ratio. In contrast, if S'/N' is small (which corresponds to small S/N), the capacity of each equivalent channel $\log(1+S'/N') \approx S'/N'$ depends linearly on the signal to noise ratio. In this case, the maximum capacity is achieved by choosing some optimal number of independent equivalent channels $n' < n'_{\max}$ (and hence some optimal value of $ka < \pi$), which is large enough, but less than $n'_{\max}$ to ensure a reasonable value of S'/N' in each of n' equivalent channels. This explains the origin of the optimal receiver spacing and its behavior shown in Fig. 7.

IV. CONCLUSION

In the present paper, we study the information content of coherent multiple-scattered wave fields in disordered media and the capacity of multiple-scattered waves to transfer the information through a disordered medium. We show that the information-theoretic quantities, such as the differential entropy h of the multiple-scattered wave field and the information capacity C of a communication channel in a disordered medium, are directly related to the mesoscopic correlations between the scattered waves. Mesoscopic correlations reduce the information content of the coherent, multiple-scattered wave field in a disordered medium, leading to a smaller number of bits required to store all the relevant information about it. To consider the transfer of information by multiple-scattered waves, we limit ourselves to the case of communication between two identical linear arrays of n equally spaced transmitters or receivers (receiver spacing a). The average information capacity $\langle C \rangle$ of such a communication channel is shown to scale linearly with n, analytic expressions for $\langle C \rangle / n$ are obtained in the limit $n \to \infty$ and $k\ell \to \infty$. For finite but large values of n and $k\ell$ the capacity per receiver $\langle C \rangle / n$ is somewhat greater than at $n \to \infty$, $k\ell \to \infty$, but the latter limiting case proves to be a fairly good approximation as long as $n \gg 1$ and $k\ell \gg 1$. Our analysis shows that if the signal to noise ratio S/N exceeds the critical value $(S/N)_c \simeq 3.35$, $\langle C \rangle / n$ grows monotonically with a as long as $ka < \pi$ and then oscillates slightly below its maximum value

achieved at $ka = m\pi$ ($m = 1, 2, \dots$). If $S/N < (S/N)_c$, the behavior of the average capacity per receiver $\langle C \rangle / n$ is not monotonic for $0 < ka < \pi$ and an absolute maximum of $\langle C \rangle / n$ is reached at some $ka < \pi$. We define the maximum average capacity $\langle C \rangle_{\max}$ as the average capacity maximized over the receiver spacing a and the normalized optimal receiver spacing $(ka)_{\mathrm{opt}}$ as the spacing maximizing the average capacity. Both $\langle C \rangle_{\max} / n$ and $(ka)_{\mathrm{opt}}$ are proportional to $(S/N)^{1/2}$ for $S/N < (S/N)_c$. At $S/N > (S/N)_c$, we find $(ka)_{\mathrm{opt}} = m\pi$ and $\langle C \rangle_{\max} / n \propto \log(S/N)$.

To conclude, let us illustrate the application of the theoretical analysis developed above on a specific example. Assume that one wishes to establish a wireless microwave (carrier frequency $\nu \sim 2$ GHz) communication link between two locations separated by a distance L of several hundreds of meters (say, $L = 500$ m) in an urban environment (city) where the waves are multiply scattered by numerous buildings. The simplest choice is to install a single antenna at each of two locations ($n = 1$ using our notation). If a high signal to noise ratio (say, $S/N = 100$) can be obtained on the receiver (thanks to a powerful transmitter), the maximum average rate of error-free information transfer (i.e., average capacity) of the resulting communication channel will be (see Eq. (12) with $\mathcal{G}$ being scalar or Ref. [20]) ≈ 5.9 bits per second and per Hertz of the available bandwidth $\Delta\nu$. To be specific, let us assume $\Delta\nu = 30$ kHz as in Ref. [26], yielding the average capacity $\langle C \rangle \approx 180$ kbps. There are not much ways to increase this number: using larger bandwidth is too expensive [17], while increasing the transmitter power is not only expensive but also does not yield much gain since $\langle C \rangle$ is roughly logarithmic in S/N: increasing S/N by an order of magnitude (i.e., from 100 to 1000) will increase $\langle C \rangle$ by less than a factor of 2. A reasonable way to increase $\langle C \rangle$ is then to use multiple antennas on both transmitting and receiving ends of the channel. For the same total transmitted power, $n = 10$ antennas arranged in a line and separated by $a = \lambda/2 \approx 8$ cm will yield $\langle C \rangle \approx 1.6$ Mbps [see Eq. (22) and Figs. 4, 5], which is an order of magnitude higher than the value, corresponding to $n = 1$. Note that such a gain in capacity is made possible by multiple scattering of microwaves on their way from transmitters to receivers. In the absence of scattering, all receivers would measure essentially the same signals, since the small aperture $d = (n-1) \times (\lambda/2) \approx 70$ cm of the receiving array is much smaller than $(\lambda L)^{1/2} \approx 9$ m, and hence the array cannot resolve different transmitting antennas (Fresnel limit). Therefore, increasing n would only increase the total received power, leading to a roughly logarithmic increase of C with n. In the case of low signal to noise ratio (e.g., $S/N = 0.1$), we obtain [see Eqs. (12) and (22)] $\langle C \rangle \approx 4$ kbps in the $n = 1$ case and a ten times larger capacity $\langle C \rangle \approx 40$ kbps in the case of $n = 10$ and $a = \lambda/2$. Optimizing a using Eq. (25) shows that one can double the latter result for $\langle C \rangle$ by choosing $a \simeq 0.1\lambda$ instead of $a = \lambda/2$ (see also Figs. 6 and 7). The resulting channel capacity is roughly 20 times higher than in the case of $n = 1$, illustrating the utility of the theoretical analysis presented in this paper in the context of practical applications.

APPENDIX A: AVERAGE CAPACITY AT $n \geqslant 1$

In this appendix, we follow Refs. [16,19] to calculate the average capacity $\langle C \rangle$ in the large-n limit. The idea of the calculation stems from the fact that the moment generating function of the random variable $\mathcal{C} = \ln \det[I_n + (S/N)\mathcal{G}^+Q\mathcal{G}]$, $F(\gamma) = \langle \exp(-\gamma \mathcal{C}) \rangle$, writes $F(\gamma) \simeq \exp(-\gamma \langle \mathcal{C} \rangle)$ in the limit of $\gamma \to 0^+$. We start therefore by a calculation of $F(\gamma)$, keeping in mind that taking the limit $\gamma \to 0^+$ will allow us to obtain the average capacity $\langle C \rangle = \langle \mathcal{C} \rangle / \ln 2$. We admit that $F(\gamma) = \langle [\det(I_n + (S/N)\mathcal{G}^+Q\mathcal{G})]^{-\gamma} \rangle$, and that for integer γ we can represent $[\det(I_n + (S/N)\mathcal{G}^+Q\mathcal{G})]^{-\gamma}$ as

$$[\det(I_n + (S/N)\mathcal{G}^+Q\mathcal{G})]^{-\gamma}$$

$$= (S/N)^{n\gamma} \int d\mu(\mathbf{X},\mathbf{Y}) \exp\left[-\frac{1}{2}(S/N)^{-1/2} \right.$$

$$\times \sum_{m=1}^{\gamma} (\mathbf{X}_m^+\mathbf{X}_m + \mathbf{Y}_m^+\mathbf{Y}_m) - \frac{1}{2}$$

$$\left. \times \sum_{m=1}^{\gamma} (\mathbf{Y}_m^+\mathcal{G}'\mathbf{X}_m - \mathbf{X}_m^+\mathcal{G}'^+\mathbf{Y}_m) \right], \tag{A1}$$

where $\mathcal{G}' = Q^{1/2}\mathcal{G}$ and we introduce 2γ auxiliary complex vectors $\mathbf{X}_m$ and $\mathbf{Y}_m$ ($m = 1, \ldots, \gamma$), the procedure known as the "replica trick," and $d\mu(\mathbf{X},\mathbf{Y})$ is the appropriate integration measure. We now average Eq. (A1) over disorder, interchanging the order of averaging and integration on the right-hand side of the equation:

$$F(\gamma) = (S/N)^{n\gamma} \int d\mu(\mathbf{X},\mathbf{Y}) \exp\left[-\frac{1}{2}(S/N)^{-1/2} \sum_{m=1}^{\gamma} (\mathbf{X}_m^+\mathbf{X}_m \right.$$

$$\left. + \mathbf{Y}_m^+\mathbf{Y}_m) - \frac{1}{4n} \sum_{m,l=1}^{\gamma} (\mathbf{Y}_m^+A\mathbf{Y}_l\mathbf{X}_l^+B\mathbf{X}_m) \right], \tag{A2}$$

where $\langle \mathcal{G}'_{\alpha i}\mathcal{G}'^*_{\beta j} \rangle = (1/n)A_{\alpha\beta}B_{ij}$. The last term in Eq. (A2) precludes the direct integration over $\mathbf{X}$ and $\mathbf{Y}$. We can, however, reduce it to a sum of quadratic terms by defining $\gamma \times \gamma$ complex matrices U and V and introducing integrations along appropriate contours in the complex plane:

$$F(\gamma) = (S/N)^{n\gamma} \int d\mu(\mathbf{X},\mathbf{Y})d\mu(U)d\mu(V) \exp\left(-\frac{1}{2}\mathcal{S} \right), \tag{A3}$$

where

$$\mathcal{S} = (S/N)^{-1/2} \sum_{m=1}^{\gamma} (\mathbf{X}_m^+\mathbf{X}_m + \mathbf{Y}_m^+\mathbf{Y}_m) + \sum_{m,l=1}^{\gamma} (\mathbf{Y}_m^+A\mathbf{Y}_lU_{ml}$$

$$+ V_{ml}\mathbf{X}_m^+B\mathbf{X}_l - nU_{ml}V_{lm}). \tag{A4}$$

We now perform the integrals over $\mathbf{X}$ and $\mathbf{Y}$ in Eq. (A3) and are left with

$$F(\gamma) = (S/N)^{n\gamma} \int d\mu(U)d\mu(V) \exp\{ -\ln \det[(S/N)^{-1/2}$$

$$+ AU] - \ln \det[(S/N)^{-1/2} + BV] + n\,\mathrm{Tr}(UV) \}, \tag{A5}$$

where AU and BV are the outer products of matrices.

Introducing the eigenvalues ξ_i and η_i of the matrices A and B, respectively, we can rewrite the exponent in Eq. (A5) as

$$\sum_{i=1}^{n} \{ \ln \det[(S/N)^{-1/2} + \xi_iU] + \ln \det[(S/N)^{-1/2} + \eta_iV] \}$$

$$- n\,\mathrm{Tr}(UV). \tag{A6}$$

In the limit $n \to \infty$ the integrations in Eq. (A5) can be performed using the saddle point method. Assuming that the replica symmetry is not broken at the saddle point, we have $U = uI_\gamma$ and $V = vI_\gamma$ with I_γ the $\gamma \times \gamma$ unit matrix. Equation (A6) then becomes

$$\gamma \sum_{i=1}^{n} \{ \ln[(S/N)^{-1/2} + \xi_iu] + \ln[(S/N)^{-1/2} + \eta_iv] \} - \gamma nuv. \tag{A7}$$

At the saddle point the partial derivatives of Eq. (A7) with respect to u and v should be zero. This yields the following equations for u and v:

$$u = \frac{1}{n} \sum_{i=1}^{n} \frac{\eta_i}{(S/N)^{-1/2} + \eta_iv}, \tag{A8}$$

$$v = \frac{1}{n} \sum_{i=1}^{n} \frac{\xi_i}{(S/N)^{-1/2} + \xi_iu}. \tag{A9}$$

Relaxing the condition of integer γ, we take the limit $\gamma \to 0^+$ and put Eq. (A5) in the form $F(\gamma) \simeq \exp(-\gamma \langle \mathcal{C} \rangle)$ with

$$\langle \mathcal{C} \rangle = \sum_{i=1}^{n} \{ \ln[(S/N)^{-1/2} + \xi_iu] + \ln[(S/N)^{-1/2} + \eta_iv] \} - nuv$$

$$+ n\ln(S/N). \tag{A10}$$

The variance of capacity can also be found in a similar way (see Ref. [19] for details), but we do not consider it here. Changing variables back from $\mathcal{G}'$ to $\mathcal{G}$ we see that $A = Q^{1/2}KQ^{1/2}$ and $B = K$ yielding $\xi_i = \kappa_iq_i$ and $\eta_i = \kappa_i$ with κ_i and q_i the eigenvalues of the matrices K and Q, respectively. Equations (A8)–(A10) then reduce to Eqs. (13)–(15) of Sec. III.

To maximize the average capacity $\langle C \rangle$ over the ensemble of covariance matrices Q, we consider an infinitesimal variation δQ of the maximizing matrix Q and require $\delta\langle C \rangle \leq 0$. This leads to $\mathrm{Tr}(\Phi\,\delta Q) \leq 0$, where

$$\Phi = \langle \mathcal{G}[I_n + (S/N)\mathcal{G}^+ Q\mathcal{G}]^{-1}\mathcal{G}^+\rangle. \tag{A11}$$

The allowed variations δQ of the covariance matrix should keep $Q + \delta Q$ positive definite and should not change the total emitted power: $\mathrm{Tr}\,\delta Q = 0$. If all eigenvalues of Q are positive, the same is true for $Q + \delta Q$ (provided that δQ is small), and $\mathrm{Tr}(\Phi\,\delta Q) = 0$ is achieved by $\Phi = \phi I_n$, where ϕ is a scalar: $\mathrm{Tr}(\Phi\,\delta Q) = \phi\,\mathrm{Tr}\,\delta Q = 0$. This can be also shown to remain true if some eigenvalues of Q are zero [19]. It then follows from Eq. (A11) that for p nonzero eigenvalues q_i of the matrix Q one has Eq. (16) of Sec. III. Finally, Eq. (17) is simply the total emitted power constraint $\mathrm{Tr}\,Q = n$.

APPENDIX B: AVERAGE CAPACITY AT $n\to\infty$ AND $k\ell\to\infty$

In this appendix, we derive analytic expressions for the average information capacity in the limit of $n\to\infty$ and $k\ell \to\infty$. We consider separately the cases of $0 < ka < \pi$, $\pi < ka < 2\pi$, and $2\pi < ka < 3\pi$. At $ka > 3\pi$ calculations can be performed in a similar way.

1. $0 < ka < \pi$

If $0 < ka < \pi$, the power spectral density takes a particularly simple form [see Eq. (8)]: $f(\mu) = \pi/(ka)$ for $0 < \mu < ka$ or $2\pi - ka < \mu < 2\pi$, while $f(\mu) = 0$ for $ka < \mu < 2\pi - ka$. Integrations in Eqs. (19) and (20) can now be easily performed, since $f(\mu) > (S/N)^{-1/2}v$ at $0 < \mu < ka$ or $2\pi - ka < \mu < 2\pi$ and $f(\mu) = 0 < (S/N)^{-1/2}v$ at $ka < \mu < 2\pi - ka$. This yields

$$u = \left[\left(\frac{S}{N}\right)^{-1/2} + v\frac{\pi}{ka}\right]^{-1}, \tag{B1}$$

$$v = \frac{ka}{\pi}\frac{1}{u}\left[1 - \frac{ka}{\pi}v\left(\frac{S}{N}\right)^{-1/2}\right]. \tag{B2}$$

We now solve these equations with respect to u, v and substitute the solution into Eq. (21), where, again, the integrations are readily performed. This gives

$$\frac{\langle C \rangle}{n} = \frac{ka}{\pi}\log\left[\frac{S}{N}\frac{\pi}{ka}\frac{1}{\phi}\right] - \frac{\phi}{\ln 2}, \tag{B3}$$

where

$$\phi = \frac{ka}{\pi} - \frac{1}{2(S/N)}\left(\frac{ka}{\pi}\right)^3\left[\sqrt{1 + 4\frac{S}{N}\left(\frac{\pi}{ka}\right)^2} - 1\right], \tag{B4}$$

which are Eqs. (22) and (23) of Sec. III.

2. $\pi < ka < 2\pi$

In this case, $f(\mu) = \pi/(ka)$ for $0 < \mu < 2\pi - ka$ or $ka < \mu < 2\pi$ and $f(\mu) = 2\pi/(ka)$ for $2\pi - ka < \mu < ka$. We now should distinguish two cases: (a) $S/N > (S/N)_1$ and (b) $S/N \leq (S/N)_1$, where

$$\left(\frac{S}{N}\right)_1 = \frac{1}{4}\left(\frac{ka}{\pi}\right)^2\left\{\frac{\sqrt{(ka/\pi)[26 - 7(ka/\pi)] - 15}}{3 - ka/\pi} - 1\right\}. \tag{B5}$$

In the case (a), $f(\mu) > (S/N)^{-1/2}v$ for all $\mu \in (0, 2\pi)$ and Eqs. (19), (20) become

$$u = \frac{2 - ka/\pi}{v + (S/N)^{-1/2}ka/\pi} + \frac{ka/\pi - 1}{v + (S/N)^{-1/2}ka/(2\pi)}, \tag{B6}$$

$$v = \frac{1}{u}\left[1 - \frac{ka}{2\pi}\left(\frac{S}{N}\right)^{-1/2}v\left(3 - \frac{ka}{\pi}\right)\right]. \tag{B7}$$

The values of u and v found from the two above equations are to be substituted into Eq. (21) that reduces to

$$\frac{\langle C \rangle}{n} = \left(2 - \frac{ka}{\pi}\right)\log\left\{\frac{\pi}{ka}\left[\frac{1}{v(S/N)^{1/2}} + \frac{\pi}{ka}\right]\right\} + \left(\frac{ka}{\pi} - 1\right)$$
$$\times \log\left\{\frac{2\pi}{ka}\left[\frac{1}{v(S/N)^{1/2}} + \frac{2\pi}{ka}\right]\right\} - \frac{uv}{\ln 2} + \log\frac{S}{N}. \tag{B8}$$

In the case (b), $f(\mu) > (S/N)^{-1/2}v$ only for $2\pi - ka < \mu < ka$. Equation (B6) remains the same, while Eqs. (B7) and (B8) become

$$v = \frac{1}{u}\left(\frac{ka}{\pi} - 1\right)\left[1 - \left(\frac{S}{N}\right)^{-1/2}v\frac{ka}{2\pi}\right], \tag{B9}$$

$$\frac{\langle C \rangle}{n} = \left(2 - \frac{ka}{\pi}\right)\log\left\{\left(\frac{S}{N}\right)^{-1/2}\left[\left(\frac{S}{N}\right)^{-1/2} + v\frac{\pi}{ka}\right]\right\} + \left(\frac{ka}{\pi} - 1\right)$$
$$\times \log\left\{\frac{2\pi}{ka}\left[\frac{1}{v(S/N)^{1/2}} + \frac{2\pi}{ka}\right]\right\} - \frac{uv}{\ln 2} + \log\frac{S}{N}. \tag{B10}$$

3. $2\pi < ka < 3\pi$

We proceed as in the two previous sections. $f(\mu) = 3\pi/(ka)$ for $0 < \mu < ka - 2\pi$ or $4\pi - ka < \mu < 2\pi$, while $f(\mu) = 2\pi/(ka)$ for $ka - 2\pi < \mu < 4\pi - ka$. If $S/N > (S/N)_2$ [case (a)], where

$$\left(\frac{S}{N}\right)_2 = \frac{(ka/\pi)^2[ka/(2\pi) - 1]}{\sqrt{(ka/\pi)(32 - 5ka/\pi) - 35} + 5 - ka/\pi}, \tag{B11}$$

$f(\mu) > (S/N)^{-1/2}v$ for all $\mu \in (0, 2\pi)$ and Eqs. (19)–(21) reduce to

$$u = \frac{3(ka/\pi - 2)}{3v + (S/N)^{-1/2}ka/\pi} + \frac{3 - ka/\pi}{v + (S/N)^{-1/2}ka/(2\pi)},$$
(B12)

$$v = \frac{1}{u}\left[1 - \frac{ka}{6\pi}\left(\frac{S}{N}\right)^{-1/2}v\left(5 - \frac{ka}{\pi}\right)\right],$$
(B13)

$$\frac{\langle C \rangle}{n} = \left(\frac{ka}{\pi} - 2\right)\log\left\{\frac{3\pi}{ka}\left[\frac{1}{v(S/N)^{1/2}} + \frac{3\pi}{ka}\right]\right\} + \left(3 - \frac{ka}{\pi}\right)$$

$$\times \log\left\{\frac{2\pi}{ka}\left[\frac{1}{v(S/N)^{1/2}} + \frac{2\pi}{ka}\right]\right\} - \frac{uv}{\ln 2} + \log\frac{S}{N}.$$
(B14)

If $S/N < (S/N)_2$ [case (b)], $f(\mu) > (S/N)^{-1/2}v$ for $0 < \mu < ka - 2\pi$ or $4\pi - ka < \mu < 2\pi$ only and we have instead of Eqs. (B13) and (B14)

$$v = \frac{2}{u}\left[\frac{ka}{2\pi} - 1 - \frac{ka}{3\pi}\left(\frac{S}{N}\right)^{-1/2}v\left(\frac{ka}{2\pi} - 1\right)\right],$$
(B15)

$$\frac{\langle C \rangle}{n} = \left(\frac{ka}{\pi} - 2\right)\log\left\{\frac{3\pi}{ka}\left[\frac{1}{v(S/N)^{1/2}} + \frac{3\pi}{ka}\right]\right\} + \left(3 - \frac{ka}{\pi}\right)$$

$$\times \log\left\{(S/N)^{-1/2}\left[(S/N)^{-1/2} + \frac{2\pi}{ka}v\right]\right\} - \frac{uv}{\ln 2} + \log\frac{S}{N}.$$
(B16)

[1] *Mesoscopic Phenomena in Solids*, edited by B.L. Altshuler, P.A. Lee, and R.A. Webb (Elsevier, Amsterdam, 1991).

[2] R. Berkovits and S. Feng, Phys. Rep. **238**, 135 (1994).

[3] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, 1995).

[4] M.C.W. van Rossum and Th.M. Nieuwenhuizen, Rev. Mod. Phys. **71**, 313 (1999).

[5] A. Lagendijk and B.A. van Tiggelen, Phys. Rep. **270**, 143 (1996).

[6] *Waves and Imaging through Complex Media*, edited by P. Sebbah (Kluwer, Dordrecht, 2001).

[7] P. Sebbah, R. Pnini, and A.Z. Genack, Phys. Rev. E **62**, 7348 (2000); P. Sebbah, B. Hu, A.Z. Genack, R. Pnini, and B. Shapiro, Phys. Rev. Lett. **88**, 123901 (2002).

[8] A. Derode, A. Tourin, and M. Fink, Phys. Rev. E **64**, 036605 (2001); **64**, 036606 (2001).

[9] N.P. Trégourès and B.A. van Tiggelen, Phys. Rev. E **66**, 036601 (2002); Waves Random Media **12**, 21 (2002).

[10] R. Hennino, N. Trégourès, N.M. Shapiro, L. Margerin, M. Campillo, B.A. van Tiggelen, and R.L. Weaver, Phys. Rev. Lett. **86**, 3447 (2001).

[11] N. Trégourès, R. Hennino, C. Lacombe, N.M. Shapiro, L. Margerin, M. Campillo, and B.A. van Tiggelen, Ultrasonics **40**, 269 (2002).

[12] C.E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948).

[13] T.M. Cover and J.A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).

[14] R.M. Gray, *Entropy and Information Theory* (Springer-Verlag, Berlin, 1990).

[15] G.J. Foschini and M.J. Gans, Wireless Personal Communications **6**, 311 (1998).

[16] A.L. Moustakas, H.U. Baranger, L. Balents, A.M. Sengupta, and S.H. Simon, Science **287**, 287 (2000).

[17] S.H. Simon, A.L. Moustakas, M. Stoytchev, and H. Safar, Phys. Today **54** (9), 38 (2001).

[18] M.R. Andrews, P.P. Mitra, and R. deCarvalho, Nature (London) **409**, 316 (2001).

[19] A.M. Sengupta and P.P. Mitra, e-print physics/0010081.

[20] S.E. Skipetrov, e-print physics/0207071.

[21] A.L. Moustakas, S.H. Simon, and A.M. Sengupta (unpublished) (available at http://mars.bell-labs.com/).

[22] M. Fink, Phys. Today **50** (3), 34 (1997); M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, M. Tanter, J.-L. Thomas, and F. Wu, Rep. Prog. Phys. **63**, 1933 (2000); M. Fink and J. de Rosny, Nonlinearity **15**, R1 (2002).

[23] M.G. Heinemann, A. Lazzara, and K.B. Smith, Appl. Phys. Lett. **80**, 694 (2002).

[24] A. Derode, A. Tourin, J. de Rosny, M. Tanter, S. Yon, and M. Fink, Phys. Rev. Lett. **90**, 014301 (2003); A. Derode, A. Tourin, M. Tanter, and M. Fink, J. Acoust. Soc. Am. **110**, 2633 (2001).

[25] G.J. Foschini, Bell Labs Tech. J. **1**, 41 (1996).

[26] G.D. Golden, G.J. Foschini, R.A. Valenzuela, and P.W. Wolniansky, Electron. Lett. **35**, 14 (1999).

[27] See http://www1.bell-labs.com/project/blast/

[28] F.J. Fahy, *Foundations of Engineering Acoustics* (Academic Press, New York, 2000).

[29] B. Shapiro, Phys. Rev. Lett. **57**, 2168 (1986).

[30] A. Böttcher and B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices* (Springer, New York, 1999); R.M. Gray, *Toeplitz and Circular Matrices: A Review* (http://www-ee.stanford.edu/ gray/toeplitz.html).

[31] A.M. Sengupta and P.P. Mitra, Phys. Rev. E **60**, 3389 (1999).

[32] To reduce the price of the experimental setup, one often uses a single mobile receiver to scan the multiple-scattered speckle pattern. As long as the speckle pattern is stationary, this is equivalent to performing a simultaneous measurement with multiple receivers at different positions.

[33] Only real quantities are, of course, measured in experiments. The signal has, however, two degrees of freedom: the amplitude $A = [(\text{Re}\,y_\alpha)^2 + (\text{Im}\,y_\alpha)^2]^{1/2}$ and the phase $\phi = \arctan(\text{Im}\,y_\alpha/\text{Re}\,y_\alpha)$. Measuring A and ϕ is equivalent to measuring the complex variable y_α.

[34] For example, the transmission rate used in the telephony is about 90% of the information capacity of a typical telephone line [13]. In multiple-antenna systems, transmission rates close to Shannon capacity can also be achieved; see, e.g., B. Hassibi and B.M. Hochwald, IEEE Trans. Inf. Theory **48**, 1804 (2002).