Information theoretical analysis of hierarchical nano-optical systems in the subwavelength regime

Makoto Naruse,[1,2,*] Hirokazu Hori,[3] Kiyoshi Kobayashi,[3] Masatoshi Ishikawa,[4] Kenji Leibnitz,[5] Masayuki Murata,[1,5] Naoya Tate,[2] and Motoichi Ohtsu[2]

[1]*National Institute of Information and Communications Technology, 4-2-1 Nukui-kita, Koganei, Tokyo 184-8795, Japan*

[2]*Department of Electrical Engineering and Information Systems and Nanophotonics Research Center, School of Engineering, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-8656, Japan*

[3]*Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Kofu, Yamanashi 400-8511, Japan*

[4]*Department of Information Physics and Computing, School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

[5]*Department of Information Networking, Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan*

**Corresponding author: naruse@nict.go.jp*

Received May 11, 2009; revised July 20, 2009; accepted July 27, 2009; posted July 29, 2009 (Doc. ID 111228); published August 24, 2009

Optical near-field interactions exhibit a hierarchical response, which is one of the most unique attributes of light–matter interactions occurring locally on the nanometer scale. It allows hierarchical nano-optical systems that break through the integration restrictions posed by the diffraction limit of conventional propagating light and offers multiple hierarchical functionalities at different physical scales in the subwavelength regime. Here we demonstrate an information theoretical approach to such nano-optical systems while assessing their electromagnetic and logical aspects via angular-spectrum analysis. Mutual information at each level of the hierarchy reveals quantitatively the relation between the physical effects associated with the hierarchy in the optical near-fields, as well as possible environmental disturbances affecting the system locally or globally, and the system's capabilities for information processing and communication.

OCIS codes: 260.2110, 200.3050, 110.3055.

1. INTRODUCTION

Optical devices and systems are increasingly exploiting physical fundamentals based on optical near-field interactions that occur locally at the nanometer-scale [1–6], not just those based on conventional diffraction-limited propagating light. The need for such nanometer-scale optical systems comes from the demands for higher integration density in applications such as optical memories [7], sensors [8], nanofabrication [9], and so forth. In addition, disruptive innovations are also expected in nanophotonic systems such their extremely low power dissipation [10], novel functional memories [11], trust and security applications [12,13], and more. The theoretical basis of light–matter interactions at the nanometer scale have been deepened [14–17] through the development of, e.g., dressed photon models [16] or angular-spectrum analysis [17]. Experimental technologies, both in fabrication and characterization of nanostructures, have also seen rapid progress such as in geometry-controlled quantum dots [18], metal nanostructures [19,20], and advanced spectroscopy [21,22].

Besides those physical insights and experimental methodologies, it is important in developing practical systems to grasp the fundamental properties from the viewpoint of information and communication [23]. In optical communications, for instance, insights gained from information theory are indispensable in assessing and improving the performance figures [24]. Information theoretic analysis also sheds light on the behavior of a wide range of optical devices and systems [25,26]. As in similar disciplines, fundamental limits in optical components have also been studied [27]. The application of information theoretic methods to nanotechnologies is found, for instance, in molecular communications based on material transfer [28] and molecular biology [29].

For nano-optical systems, on the other hand, while their physical understanding has been extensively deepened [1], their information theoretic aspects are, to the best of our knowledge, completely unknown. If the unique attributes enabled by optical near-fields, which conventional propagating light does not offer, are properly defined as an information and communication system, it would provide critical knowledge and guidelines in designing concrete applications.

The hierarchical property in optical near-field interactions, meaning that the interaction depends on the physical scale involved, as schematically shown in Fig. 1(a), is one of the attributes that differentiates them most from conventional diffraction-limited far-field light [30]. In [30], we have demonstrated that the optical near-fields can be distributed independently at different scales of observation by exploiting the nature that the high spatial-frequency terms in optical near-fields are rapidly decayed with a decay rate that is spatial-frequency dependent. As

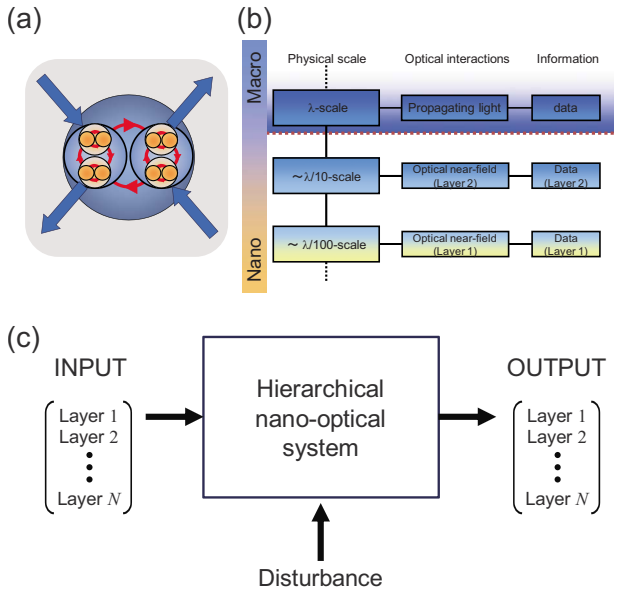

Fig. 1. (Color online) Hierarchy in optical near-fields at the subwavelength scale and its application to information and communications. (a) Optical near-field interactions and their scale-dependence. (b) Hierarchical nano-optical system where different functionalities are associated with each of the physical scales involved. (c) System diagram to study information theoretic aspects of hierarchical nano-optical systems.

introduced in Section 2 below, it is convenient, in discussing such a hierarchical property in optical near-fields, to represent the electric fields as a superposition of plane waves with complex wave vector, called the angular spectrum. This approach clearly unifies the fine/coarse structures and their associated decay of the optical near-fields.

Such a hierarchical property leads to versatile novel functions in nano-optical systems; for example, multiple functionalities could be associated with each of the physical scales in the subwavelength regime, as illustrated in Fig. 1(b). This has already had an impact on a wide range of applications [11–13,31]. This also implies that, whereas the time-domain behavior is crucial in conventional optical communications [24], space-domain properties dominate the unique behavior observed in nano-optical systems and these should be analyzed from the viewpoint of information theory.

In this paper, we take an information theoretic approach to analyzing such nano-optical systems while assessing their electromagnetic and logical aspects via angular-spectrum analysis. As schematically shown in Fig. 1(c), the hierarchical nano-optical system is modeled as a communication medium that connects input and output symbols and also suffers from environmental disturbances. Concrete applications represented by such a model will include, for instance, optical storage systems where the input and output respectively correspond to the write and retrieve processes. Here, at each level of the hierarchy, we formulate mutual information that quantitatively reveals the relation between the physics associated with the hierarchy in optical near-fields, as well as possible environmental disturbances affecting the system locally or globally, and the capabilities of the system for information processing and communications. In other words, here we intend to grasp the hierarchical structure of nano-optical systems from a cross-cutting standpoint, including their electromagnetic, logical, and information theoretic aspects.

This paper is organized as follows: In Section 2, we start by discussing the hierarchical properties in optical near-fields based on angular-spectrum analysis and its logical responses as a hierarchical system. In Section 3, we formulate hierarchical nano-optical systems from the viewpoint of information theory and reveal their layer- and disturbance-dependent mutual information. Section 4 concludes the paper.

2. HIERARCHICAL NANO-OPTICAL SYSTEMS BASED ON OPTICAL NEAR-FIELDS

A. Electromagnetic Hierarchy in Nano-Optical Systems

We first introduce an electromagnetic theory to describe hierarchy in optical near-fields. The angular spectrum representation of the electromagnetic field [17,32] is useful in discussing the hierarchy in optical near-fields [30]. It allows an analytical treatment while giving an intuitive picture of the localization of optical near-fields, and represents relevance in optical near-field interactions at different scales of observation, since it describes electromagnetic fields as a superposition of evanescent waves with different decay lengths and corresponding spatial frequencies assuming planar boundary conditions.

Suppose that there is an oscillating electric dipole, $\boldsymbol{d}^{(k)}=(d^{(k)}\cos\varphi^{(k)},0)$, on the xz plane, oriented parallel to the x axis. Note that those dipoles are oscillating; the orientations of the dipoles, specified by $\varphi^{(i)}$, physically correspond to the absolute phase of the electric fields. Also note that the physical scales under study are smaller than the wavelength of light, but larger than the scales governed by the electron interactions, typically around 1 nm.

Now, consider the electric field of radiation observed at a position displaced from the dipole by $\boldsymbol{R}^{(k)}=(r_{\parallel}^{(k)}\cos\phi^{(k)},z^{(k)})$. Assuming the planar boundary as the xy plane, the angular spectrum representation of the z-component of the optical near-field is given by

$$E_z(\boldsymbol{R})=\left(\frac{iK^3}{4\pi\varepsilon_0}\right)\int_1^{\infty}\mathrm{d}s_{\parallel}\frac{s_{\parallel}}{s_z}f_z(s_{\parallel},\boldsymbol{d}^{(1)},\ldots,\boldsymbol{d}^{(N)}), \tag{1}$$

where

$$f_z(s_{\parallel},\boldsymbol{d}^{(1)},\ldots,\boldsymbol{d}^{(N)})=\sum_{k=1}^{N}d^{(k)}s_{\parallel}\sqrt{s_{\parallel}^2-1}\cos(\varphi^{(k)}-\phi^{(k)})J_1(Kr_{\parallel}^{(k)}s_{\parallel})\exp(-Kz^{(k)}\sqrt{s_{\parallel}^2-1}), \tag{2}$$

where N represents the number of dipoles, and $\boldsymbol{R}$ indicates the observation position. Here, $s_{\parallel}$ is the spatial frequency of an evanescent wave propagating parallel to the x axis, and $J_n(x)$ represents Bessel functions of the first kind. The term $f_z(s_{\parallel},\boldsymbol{d}^{(1)},\ldots,\boldsymbol{d}^{(N)})$ is called the angular spectrum of the electric field, where the spatial frequency-dependent exponential decay is clearly represented. Details of the theory can be found in [30].

In the following, a two-layer system is introduced where (i) at locations closer to the dipoles, two items of *first-layer* information are retrieved, and (ii) at a location relatively far from the dipoles, one item of *second-layer* information is retrieved. What should be noted here is that a coarser scale structure can affect the optical responses relatively far from the structure thanks to the longer decay length of the optical near-fields, while a finer scale structure can only affect the optical responses closer to the structure due to the rapidly decaying nature of higher spatial frequency terms in optical near-fields. Therefore, as will be introduced shortly below, the optical near-field amplitude can be distributed independently at different scales of observation; in other words, arbitrary logical combinations are possible at the first layer and the second layer.

In order to clearly represent such a situation, we introduce a model composed of two closely spaced dipole pairs [$\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$] and [$\boldsymbol{d}^{(3)}$ and $\boldsymbol{d}^{(4)}$], whose positions are specified as shown in Fig. 2(a). The fractional numbers in the figure indicate the distances in units of optical wavelength. The fine-scale structures, affecting the first layer response, are represented by each of those pairs, whereas the coarse-scale structures, affecting the second-layer response, depend on the relation between those pairs. In other words, the mixture of the coarse-scale and the fine-scale structures is represented by those four dipoles, which are precisely described theoretically in the following manner based on the angular spectrum representation.

First, we assume that the dipoles $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ are oriented in the same direction, namely, $\varphi^{(1)}=\varphi^{(2)}=0$, and that the other dipoles, $\boldsymbol{d}^{(3)}$ and $\boldsymbol{d}^{(4)}$, are both oriented oppositely to $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$, namely, $\varphi^{(3)}=\varphi^{(4)}=\pi$. At a position close to the x axis equidistant from $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$, such as at the position $F^{(1)}$ in Fig. 2(a), the electric field is weak (logical ZERO) since the angular spectrum contributions originating from $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ cancel each other, and the electric field originating from $\boldsymbol{d}^{(3)}$ and $\boldsymbol{d}^{(4)}$ is small. In fact, as shown by the dashed curve in Fig. 2(b), since the angular spectrum at position $F^{(1)}$ oscillates equally around the x axis, the integral of the angular spectrum, which is related to the field intensity at that point, will be low.

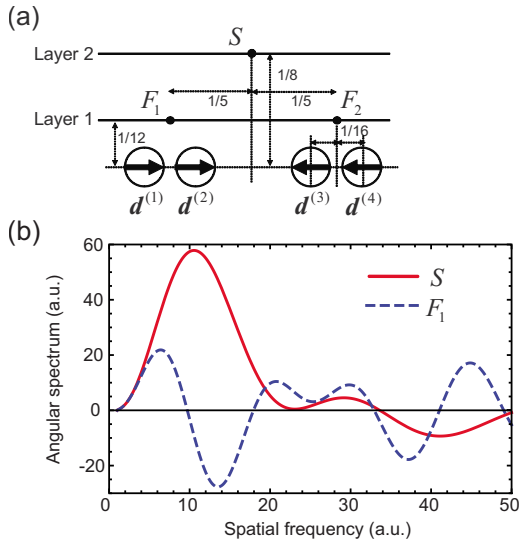

Fig. 2. (Color online) Angular spectrum analysis of a hierarchical nano-optical system. (a) Physical model for a hierarchical nano-optical system consisted of four oscillating dipoles ($\boldsymbol{d}^{(1)}$,$\boldsymbol{d}^{(2)}$,$\boldsymbol{d}^{(3)}$,$\boldsymbol{d}^{(4)}$). The optical response at the first layer ($F^{(1)}$ and $F^{(2)}$) and the second layer (S) are evaluated. (b) Angular spectrum evaluated at the first layer (dashed curve) and the second layer (solid curve) from an array of dipoles whose phase arrangement is specified in (a), which indicates that the optical near-fields are localized in the second layer, whereas they are not localized in the first layer.

For the second layer, consider the observation at an intermediate position between the dipole pairs, such as the position S in Fig. 2(a). From this position, the four dipoles effectively appear to be two dipoles that are oriented in opposite directions to each other. Particularly with the present system conditions, the second-layer signal at S is dominated by the arrangement of the two center dipoles. As shown by the solid curve in Fig. 2(b), the angular spectrum exhibits a single peak, indicating that the electric field in the xy-plane is localized to a degree determined by its spectral width, so that a logical ONE is retrievable at position S. Meanwhile, the angular spectrum observed at position $F^{(2)}$ shown in Fig. 2(a) is identical to that observed at position $F^{(1)}$, meaning that the electric field at $F^{(2)}$ is also at a low level.

B. Logical Hierarchy in Nano-Optical Systems

As described above, one of the two first-layer signals, the electric field at $F^{(1)}$, primarily depends on the dipole pair $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$, and the other, the electric field at $F^{(2)}$, is dominated by the dipole pair $\boldsymbol{d}^{(3)}$ and $\boldsymbol{d}^{(4)}$. The second-layer signal is determined by those two pairs of dipoles. These dependence structures are schematically shown in Fig. 3(a). With such a hierarchical mechanism, a total of eight different signal combinations, or symbols, can be achieved by appropriately orienting the four dipoles [30], as summarized in Fig. 3(b) where the eight symbols are denoted by a_i ($i=0,\ldots,7$), each of which refers to a three-bit sequence whose first element represents the digit obtained at the second layer (S), and whose second and third elements respectively refer to the digits obtained at the first layer ($F^{(1)}$ and $F^{(2)}$). A corresponding dipole orientation is also indicated in the right-hand side of Fig. 3(b).

Figure 4 represents examples of simulated electric field distributions obtained through finite-difference time-domain methods to visualize hierarchical electromagnetic structures. Four silver nanoparticles (radius of 15 nm) containing a virtual oscillating light source were assumed in order to simulate dipole arrays. The operating wavelength was 488 nm.

Corresponding to the angular spectrum-based theoretical analysis discussed above, Fig. 4(a) shows the z-component of the electromagnetic intensity corresponding to the dipole arrangement of the symbol a_4. The electric fields obtained at $F^{(1)}$ and $F^{(2)}$ remained low, while that at S yielded a higher level. Figure 4(b) shows another example, corresponding to the symbol a_3, exhibiting opposite responses to the former case. Figure 4(c) compares the intensity levels at the second layer for all eight

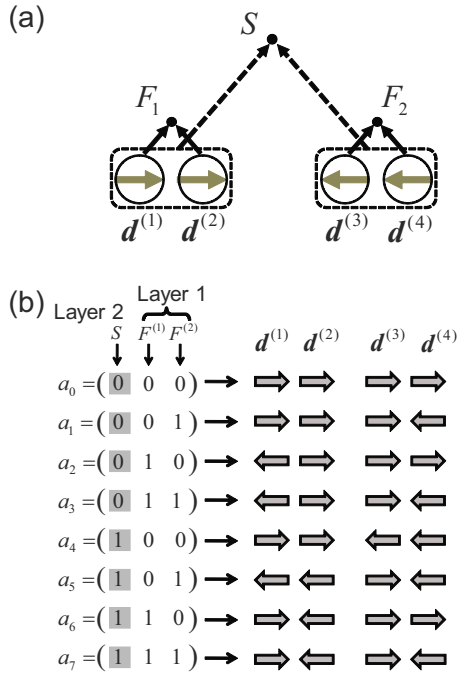

Fig. 3. (Color online) Dependence structure between the dipole pairs and the logical levels in the first and second layers. (a) Relation between dipole pairs and the logical level in each layer. (b) List of the total of eight bit-sequences, or symbols, and their corresponding dipole arrangements.

combinations, which agrees with the first bit of each symbol in Fig. 3(b) by digitizing with a threshold depicted by a dashed line in Fig. 4(c).

The relation between the electromagnetic hierarchy and its logical implications depends also on other factors at the nanometer scale, such as the intensity distributions of the sources. As mentioned earlier, the second-layer information depends on the larger scale structures; therefore, even the two dipoles located at the edges, that is, $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(4)}$, can dominate the second-layer responses. For instance, suppose that the amplitude of the radiation from the peripheral dipoles is larger than that from the

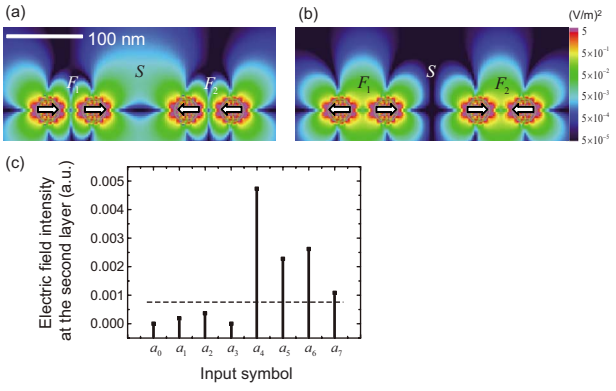

Fig. 4. (Color online) Hierarchical electromagnetic structures at the subwavelength scale. (a,b) Simulated electric field intensity distributions corresponding to the input symbols (a) a_4 and (b) a_3. (c) Simulated electric field intensity obtained at the second layer for all of the symbols that yield the intended digit for the second layer by digitizing the intensity with a threshold indicated by the dashed line.

two center ones, that is, $|\boldsymbol{d}^{(1)}|:|\boldsymbol{d}^{(2)}|:|\boldsymbol{d}^{(3)}|:|\boldsymbol{d}^{(4)}|=b:a:a:b$ where $b>a$, while keeping the total intensity $\Sigma_i|\boldsymbol{d}^{(i)}|^2$ the same. Figure 5(a) represents a case with $b/a=4$. This yields logically different behavior in the second layer; namely, the dipoles located at the periphery then dominate the second layer. For instance, with the dipole phase arrangement corresponding to the symbol a_6, the second-layer signal exhibits contrasting patterns. Figure 5(b) compares angular spectra corresponding to the uniform intensity distributions (denoted by Array 6 in Fig. 5(a)) with biased intensity pattern (denoted by Array 6′) of the four dipoles, where the latter yields poorer localization than the former. Figures 5(c) and 5(d), respectively, show calculated electromagnetic distributions corresponding to Array 6 and Array 6′, where the difference in the second layer is clearly observed.

Here we make two remarks regarding the modeling described above. First, we do not take into account the multiple scattering processes between electric polarizations in the analysis. One reason is to avoid unnecessary complexity of the discussion, since the primary concern of this paper is about the information theoretic analysis beginning in Section 3. Second is that we can apply such an angular spectrum-based approach in the same manner after considering multiple scattering processes. The essential logical flow of the discussion remains the same.

Another remark is about the physically reasonable equivalents to an array of electric dipoles used in the modeling. Shape-engineering of metal nanostructures is one approach to implement electric polarizations in designated orientations; for instance, combinations of triangular-shaped gold nanostructures exhibit hierarchical responses, as demonstrated in [33]. The difference of the amplitude of electric dipoles could then correspond to, for instance, the difference in sizes of those metal nanostructures. Also, electric polarizations induced in semiconductor nanorods [34] or geometry-controlled semiconductor quantum dots [18] could effectively correspond to the present model.

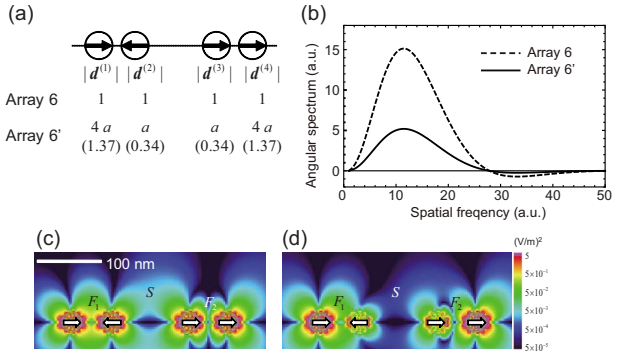

Fig. 5. (Color online) Dependence of the hierarchical electromagnetic structure to various nano-scale entities. (a) Array of dipoles, corresponding to the symbol a_6, each of which has identical radiation intensity (denoted by Array 6) and those which have biased radiation intensity while the total intensity remains the same (denoted by Array 6′). (b) Angular spectrum evaluated at the second layer for Array 6 and Array 6′, where the Array 6′ exhibits degraded light localization. Simulated electric field intensity distributions corresponding to (c) Array 6 and (d) Array 6′.

3. INFORMATION THEORETIC ANALYSIS OF HIERARCHICAL NANO-OPTICAL SYSTEMS

Here we start by analyzing the information theoretic aspects of the hierarchical nano-optical system, which is viewed as a communication medium that connects the input to the output channels while experiencing certain environmental disturbances. The input and output signals are, as introduced already, three-bit sequences; the total of eight symbols for the input and the output symbols are, respectively, denoted by $A=\{a_i\}$ and $B=\{b_i\}$, where $i=0,\ldots,7$. The output symbol b_i should represent the same three-bit sequence given by the input symbol a_i if there is no error from the input to the output. The probability of the input symbol a_i is given by $P(a_i)$. The probability of the output symbol b_i, given by $P(b_i)$, depends on the input symbol probability $P(a_i)$ and the transition matrix T that represents the relation between the input and output and is given by

$$\begin{pmatrix} P(b_0) \\ \vdots \\ P(b_7) \end{pmatrix} = T \begin{pmatrix} P(a_0) \\ \vdots \\ P(a_7) \end{pmatrix}, \tag{3}$$

where the elements of T are given by $t_{ij}=P(b_i|a_j)$, which is the probability of the output b_i conditioned on the input a_j. The matrix T is affected by environmental disturbances. We will introduce two representative disturbance models for the analysis. The first is the case where the environmental disturbance has an explicit spatial structure or what we call a *near-field* disturbance. The second is the case where the disturbance selectively couples to the dipole arrangement or what we name as a *far-field* disturbance. Here, we take into account the optical selection rule that the far-field optical radiation cannot be coupled to energy levels specified by even quantum number(s), or quadrupole(s), in considering the interactions between the disturbance and the system under study.

A. Near-Field Disturbance

We assume an environmental disturbance that modifies the status of each spatial position in the system locally. In other words, here we assume a *near-field* environmental disturbance that locally disturbs the polarizations of each of the dipoles. Suppose that the phase of at most one of the four dipoles could be flipped, that is, changed by 180°. Let the probability of such a phase flip occurring for the dipole $\boldsymbol{d}^{(i)}$ be given by p_i. In the case of input symbol a_0, for example, whose corresponding four dipoles are all in phase (See Fig. 3(b)), the phase of the first dipole $\boldsymbol{d}^{(1)}$ is flipped, which yields output symbol b_2 with probability p_1, as schematically shown in Fig. 6(a). When the second dipole $\boldsymbol{d}^{(2)}$ is flipped, the resulting output symbol is b_6. The error-free probability is given by $q=1-\Sigma_{i=1}^{4}p_i$. All transitions from input symbols to output ones are derived as schematically represented in Fig. 6(b), where possible changes from an input to an output symbol are indicated by arrows.

It is technically possible to further generalize the effects of disturbances in this model, such as arbitrary phase flips occurring in the dipoles. However, this leads to an unnecessarily complex situation and makes it hard to understand how the local change in the system affects the information capacity. As a first step, simple position-dependent errors clearly reveal the relation between the physical consequences and their impact on layer-dependent information theoretic measures, which is the primary concern of this paper.

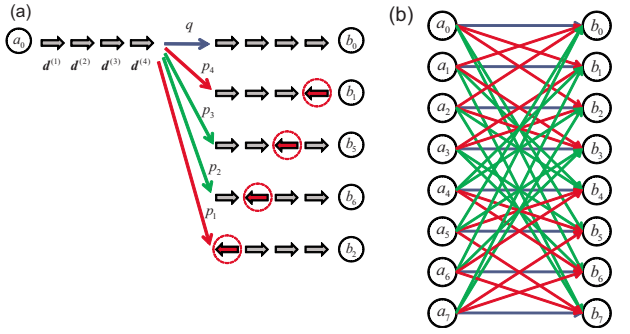

Fig. 6. (Color online) Near-field disturbances to the system. (a) Model of an environmental disturbance that locally affects each of the dipoles. The phase of dipole $\boldsymbol{d}^{(i)}$ can be flipped with probability p_i. (b) The relation between the input symbols a_i and the output ones b_j. The red and green arrows, respectively, correspond to the phase flips at the edge and at the center dipole(s). The blue ones correspond to no error.

The mutual information $I(A;B)$ represents the quantity of the data transmitted through the system, which is equal to the amount of remaining uncertainty of data A on condition that the output B is measured, namely, $I(A;B)=H(A)-H(A|B)$, where $H(A)$ is the entropy of the input and $H(A|B)$ is the entropy of the input conditioned on the output. It is calculated by

$$I(A;B)=\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}P(a_i,b_j)\log_2\frac{P(a_i,b_j)}{P(a_i)P(b_j)}, \tag{4}$$

where the joint probability of input a_i and output b_j is given by $P(a_i,b_j)=t_{j,i}P(a_i)$ [35]. N indicates the number of symbols.

Here we further introduce a representative spatial structure of environmental disturbances so that the error probabilities affecting dipoles located at the edge and the center are different; that is, the error probability for the dipoles $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(4)}$ is given by p_E, namely $p_1=p_4=p_E$, whereas that for $\boldsymbol{d}^{(2)}$ and $\boldsymbol{d}^{(3)}$ is given by $p_C=p_2=p_3$. The error-free probability is given by $q=1-2p_E-2p_C$. As described in Section 2, the four dipole model intends to represent the mixture of the coarse-scale structure and the fine-scale structure. In the present model, the two center dipoles dominate the second-layer information, meaning that those two play the major role in coarser scale structures. Therefore, the position-dependent errors denoted by p_C and p_E are physically associated with errors occurring in coarse-scale structures and in the fine-scale structures, respectively. As remarked at the end of Section 2, if those dipoles with designated orientations are implemented by means of shape-engineered nanostructures, the error would physically correspond to fabrication errors occurring at either coarser or finer scales.

Now, we suppose that the input symbol probability is uniform; that is, $P(a_i)=1/8$ $(i=0,\ldots,7)$. The mutual infor-

mation is evaluated as a function of p_E and p_C, as shown in Fig. 7(a). For example, $p_E=9/100$ and $p_C=1/100$ yields 1.92 bit of mutual information.

We now analyze the mutual information regarding each of the layers separately to see how the spatial structure of the local system disturbance affects the transmission capability for each of the layers.

Let the input symbols for the left-hand side of the first layer be $f_0^{(1)}$ and $f_1^{(1)}$, which respectively mean logical 0 and 1. Since $f_0^{(1)}$ and $f_1^{(1)}$ are respectively equivalent to the input symbol of either one of $F_0^{(1)}=\{a_0,a_1,a_4,a_5\}$ and $F_1^{(1)}=\{a_2,a_3,a_6,a_7\}$, the probability of input symbol $f_i^{(1)}$ is given by

$$P(f_i^{(1)}) = \sum_{a_j \in F_i^{(1)}} P(a_j). \tag{5}$$

Likewise, we denote the output symbols as $g_0^{(1)}$ and $g_1^{(1)}$, which are equivalent to either of the output symbols $G_0^{(1)}=\{b_0,b_1,b_4,b_5\}$ and $G_1^{(1)}=\{b_2,b_3,b_6,b_7\}$, respectively. The joint probability of inputs $f_i^{(1)}$ and $g_j^{(1)}$ is then derived as

$$P(f_i^{(1)},g_j^{(1)}) = \sum_{a_s \in F_i^{(1)}} \sum_{b_t \in G_j^{(1)}} P(b_t|a_s)P(a_s), \tag{6}$$

which leads to mutual information for the left-hand bit of the first layer, given by

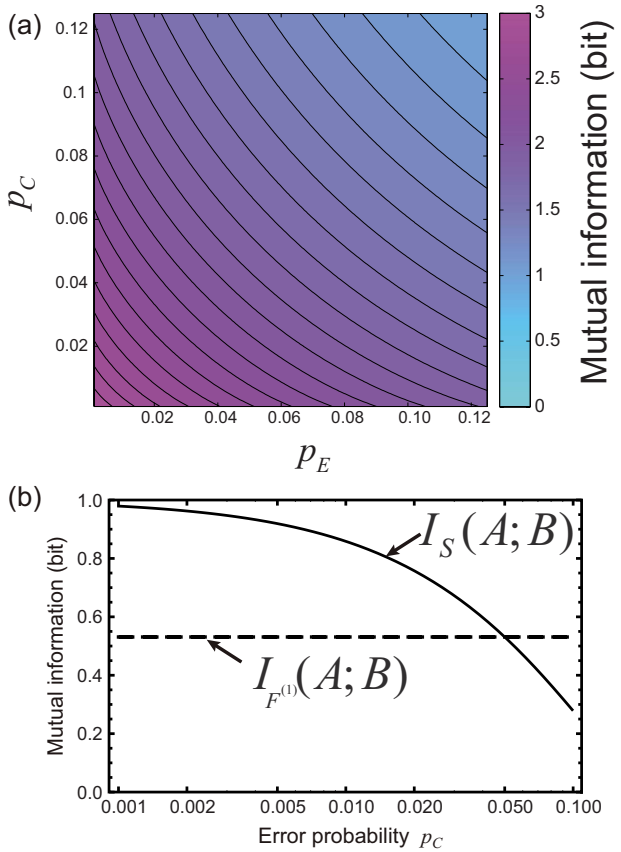

Fig. 7. (Color online) Scale-dependent mutual information. (a) Mutual information $I(A;B)$ as a function of error probabilities p_E and p_C. (b) Mutual information for the left-hand bit at the first layer ($I_{F^{(1)}}(A;B)$) and that for the second layer ($I_S(A;B)$) are evaluated as a function of error probability p_C, while keeping the total error rate p_E+p_C constant.

$$I_{F^{(1)}}(A;B) = \sum_{i=0}^{1} \sum_{j=0}^{1} P(f_i^{(1)},g_j^{(1)}) \log_2 \frac{P(f_i^{(1)},g_j^{(1)})}{P(f_i^{(1)})P(g_j^{(1)})}. \tag{7}$$

Second, let the symbols of the second layer be s_0 and s_1. The symbols s_0 and s_1, respectively, correspond to either one of the input symbols $S_0=\{a_0,a_1,a_2,a_3\}$ and $S_1=\{a_4,a_5,a_6,a_7\}$. Similarly, the corresponding output symbols are defined by t_0 and t_1. The mutual information for the second layer is given by

$$I_S(A;B) = \sum_{i=0}^{1} \sum_{j=0}^{1} P(s_i,t_j) \log_2 \frac{P(s_i,t_j)}{P(s_i)P(t_j)}. \tag{8}$$

Here we quantitatively compare the mutual information given by Eqs. (7) and (8) assuming the same error probabilities. With less errors in the center and more errors at the periphery, for instance $p_C=1/100$ and $p_E=9/100$, $I_{F^{(1)}}(A;B)$ yields 0.53 bit, whereas $I_S(A;B)$ yields 0.86 bit, indicating that the second layer has larger information transmission efficiency than the first layer. On the other hand, with more errors in the center and less errors at the periphery, for instance $p_C=9/100$ and $p_E=1/100$, $I_{F^{(1)}}(A;B)$ yields 0.53 bit, whereas $I_S(A;B)$ yields 0.32 bit, showing that the quantity of information for the first layer is unchanged, whereas that for the second layer is severely degraded. With the condition that the error-free probability be kept constant at 0.8, $I_{F^{(1)}}(A;B)$ and $I_S(A;B)$ are respectively calculated as functions of p_C in Fig. 7(b), where $I_{F^{(1)}}(A;B)$ stays constant, whereas $I_S(A;B)$ takes larger values as p_C gets smaller. This is due to the fact that the second-layer information depends on the two center dipoles. On the other hand, the left-hand bit at the first layer depends on both $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$; therefore, $I_{F^{(1)}}(A;B)$ yields a constant value as long as p_1+p_2 is constant.

B. Far-Field Disturbance

We now consider another type of environmental disturbance involving *far-field* radiation applied to nano-scale optical systems. Here, the optical selection rule that the far-field optical radiation cannot be coupled to energy levels specified by even quantum number(s) or quadrupole(s) [1,36] should be taken into account in considering the interactions between the disturbance and the system under study. The physical model consisting of multiple dipoles is, in fact, not directly compatible with such quantum optical properties. However, we consider that the following assumptions approximately describe their principal characteristics in order to understand its impact on information theoretic measures. They are as follows:

(i) When the phases of $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ are the same, both phases can be flipped with probability p. Similarly, the phases of $\boldsymbol{d}^{(3)}$ and $\boldsymbol{d}^{(4)}$ can be flipped with probability p when they have the same phase. The error-free probability is given by $q=1-p$. [Fig. 8(a)]

(ii) When all of the dipoles have the same phase, they can be flipped with probability p. The error-free probability is given by $q=1-p$. [Fig. 8(b)]

(iii) When the combinations of [$\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$] and/or [$\boldsymbol{d}^{(3)}$ and $\boldsymbol{d}^{(4)}$] have opposite phases, we assume that the

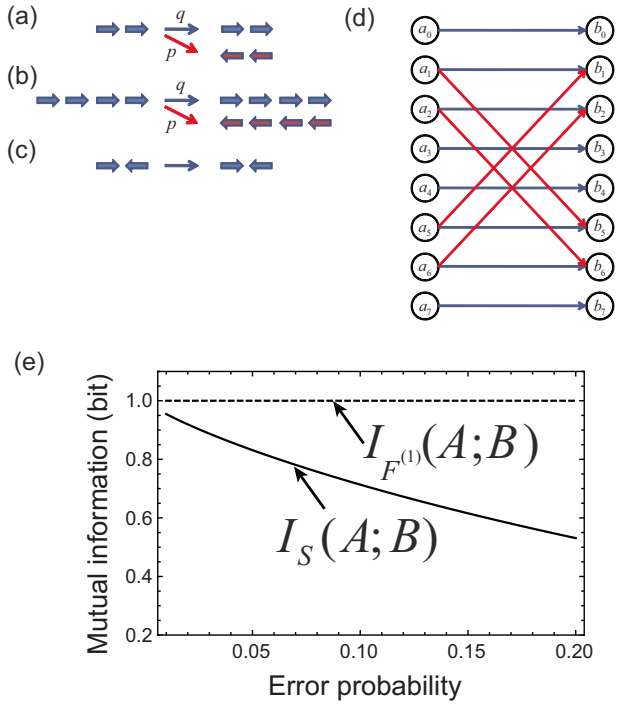

Fig. 8. (Color online) Far-field disturbance and its impact on layer-dependent mutual information. (a–c) Model of an environmental disturbance that couples to the system when arranged in a dipole manner and that does not couple to it when in a quadrupole arrangement. (a) Two closely separated dipole pairs can be flipped when they have the same phase. (b) Four dipoles can be flipped when they have the same phase. (c) The dipole pair does not suffer disturbances when they have opposite phases. (d) The relation between the input symbols a_i and the output ones b_j with those disturbances. The red (diagonal) arrows correspond to errors due to disturbances. The blue (horizontal) ones correspond to no error. (e) Mutual information for the left-hand bit at the first layer ($I_{F^{(1)}}(A;B)$) and that for the second layer ($I_S(A;B)$) as a function of error probability. $I_{F^{(1)}}(A;B)$ takes the value of 1, meaning that the first-layer information is completely resistant to such environmental disturbances.

far-field radiation serving as the system disturbance cannot be coupled to the dipoles and so no phase flips occur. [Fig. 8(c)]

For instance, for the case of input symbol a_0, all of the dipoles are arranged in the same directions, and they can all be coupled to the environmental disturbance; that is, item (ii) above will apply. However, from a logical point of view, flipping all of the dipoles does not affect the information of any bits to be retrieved at the output. In other words, the input symbol a_0 is always connected to output symbol b_0. For the case of input symbol a_1, $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ are in dipole arrangement, whereas $\boldsymbol{d}^{(3)}$ and $\boldsymbol{d}^{(4)}$ are in quadrupole arrangement. The dipoles $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ can be flipped with probability p, which results in the output symbol b_5. The relations between all input and output symbols are summarized in Fig. 8(d).

The mutual information for the first layer and the second layer are respectively evaluated following Eqs. (7) and (8). As indicated by the dashed line in Fig. 8(e), the mutual information for the first layer $I_{F^{(1)}}(A;B)$ is always 1, indicating that the system is completely resistant to the disturbances in the first layer. This is due to the fact that the environmental disturbance transforms the input symbols $F_0^{(1)}=\{a_0,a_1,a_4,a_5\}$ into the output symbols $G_0^{(1)}$, and the input symbols $F_1^{(1)}=\{a_2,a_3,a_6,a_7\}$ to $G_1^{(1)}$, where the left-hand bit at the first layer remains the same. On the other hand, the mutual information for the second layer $I_S(A;B)$ decreases as the error probability increases, as shown in the solid curve in Fig. 8(e). Since the second-layer signal depends on the pairs of dipoles, it could be disturbed in situations where the left- and the right-hand dipole pairs are a combination of a dipole arrangement and a quadrupole one.

The mutual information for the entire system can be improved by biasing the input symbol probability distribution, whose maximum is defined as the channel capacity, given by $C=\max_{p(a_i)}I(A;B)$. With the error probability $p=2/10$, the maximum mutual information, i.e., the channel capacity, is 2.68 bit when the probability distribution is given by $P(a_0)=P(a_3)=P(a_4)=P(a_7)=0.16, P(a_1)=P(a_2)=P(a_5)=P(a_6)=0.09$, which is obtained by a full search in the parameter space. On the other hand, the mutual information with a uniform input symbol distribution $P(a_i)=1/8$ $(i=0,\ldots,7)$ yields 2.64 bit. Although the higher number of bits obtained by changing the probability distribution is, in this particular case, just 0.04 bit, such consideration leads to system design strategies to fully utilize the capacity of such hierarchical systems.

4. CONCLUSION

In conclusion, we have demonstrated an information theoretic analysis of hierarchical nano-optical systems by highlighting the layer- and disturbance-dependent mutual information while assessing their electromagnetic and logical aspects with an angular-spectrum based method. Our formulation and analysis of mutual information at each level of the hierarchy reveal quantitatively the relation between the physics associated with the hierarchy in optical near-fields, as well as possible environmental disturbances affecting the system locally or globally, and the capabilities of the hierarchical nano-optical system for information processing and communications. Specifically, assuming an array of oscillating electric dipoles as a system model, we analyzed its electromagnetic and logical hierarchical structure and we quantitatively evaluated its capabilities as an information system by determining the mutual information for each level of the hierarchy; either the first or the second layer could be more resistant to environmental disturbances compared to the other one.

We believe that this paper reveals, for the first time, one fundamental aspect of nano-optical systems in the subwavelength regime from the viewpoint of information theory, which provides useful knowledge and guidelines in designing concrete applications. At the same time, however, we also believe that the unique attributes enabled by nanophotonics, which conventional propagating light does not offer, are not limited by this study. For instance, optical excitation transfer via optical near-field interactions [1], nonadiabatic processes enabled by optical near-fields [37,38], and other principles in nanophotonics may be other crucial aspects impacting the system from the viewpoint of information theory. Also, unique information strategies for nanophotonic systems such as smart coding

algorithms, for example, could be derived. We aim to further develop our understanding of the fundamentals of nanophotonic systems in future studies.

REFERENCES

1. M. Ohtsu, K. Kobayashi, T. Kawazoe, T. Yatsui, and M. Naruse, *Principles of Nanophotonics* (Taylor and Francis, 2008).
2. *Near Field Optics*, D. W. Pohl and D. Courjon, eds. (Kluwer Academic, 1993).
3. S. A. Maier, P. G. Kik, H. A. Atwater, S. Meltzer, E. Harel, B. E. Koel, and A. A. G. Requicha, "Local detection of electromagnetic energy transport below the diffraction limit in metal nanoparticle plasmon waveguides," Nature Mater. **2**, 229–232 (2003).
4. Y. Fedutik, V. V. Temnov, O. Schöps, U. Woggon, and M. V. Artemyev, "Exciton-plasmon-photon conversion in plasmonic nanostructures," Phys. Rev. Lett. **99**, 136802 (2007).
5. A. Ueda, T. Tayagaki, and Y. Kanemitsu, "Energy transfer from semiconductor nanocrystal monolayers to metal surfaces revealed by time-resolved photoluminescence spectroscopy," Appl. Phys. Lett. **92**, 133118 (2008).
6. V. I. Klimov, A. A. Mikhailovsky, S. Xu, A. Malko, J. A. Hollingsworth, C. A. Leatherdale, H.-J. Eisler, and M. G. Bawendi, "Optical gain and stimulated emission in nanocrystal quantum dots," Science **290**, 314–317 (2000).
7. T. Nishida, T. Matsumoto, F. Akagi, H. Hieda, A. Kikitsu, and K. Naito, "Hybrid recording on bit-patterned media using a near-field optical head," J. Nanophotonics **1**, 011597 (2007).
8. J. N. Anker, W. P. Hall, O. Lyandres, N. C. Shah, J. Zhao, and R. P. Van Duyne, "Biosensing with plasmonic nanosensors," Nature Mater. **7**, 442–453 (2008).
9. M. Naya, I. Tsurusawa, T. Tani, A. Mukai, S. Sakaguchi, and S. Yasutani, "Near-field optical photolithography for high-aspect-ratio patterning using bilayer resist," Appl. Phys. Lett. **86**, 201113 (2005).
10. M. Naruse, T. Miyazaki, T. Kawazoe, K. Kobayashi, S. Sangu, F. Kubota, and M. Ohtsu, "Nanophotonic computing based on optical near-field interactions between quantum dots," IEICE Trans. Electron. **E88-C**, 1817–1823 (2005).
11. M. Naruse, T. Yatsui, W. Nomura, N. Hirose, and M. Ohtsu, "Hierarchy in optical near-fields and its application to memory retrieval," Opt. Express **13**, 9265–9271 (2005).
12. M. Naruse, T. Yatsui, T. Kawazoe, Y. Akao, and M. Ohtsu, "Design and simulation of a nanophotonic traceable memory using localized energy dissipation and hierarchy of optical near-field interactions," IEEE Trans. Nanotechnol. **7**, 14–19 (2008).
13. N. Tate, H. Sugiyama, M. Naruse, W. Nomura, T. Yatsui, T. Kawazoe, and M. Ohtsu, "Quadrupole-dipole transform based on optical near-field interactions in engineered nanostructures," Opt. Express **17**, 11113–11121 (2009).
14. C. Cohen-Tannoudji, I. Dupont-Roc, and G. Grynberg, *Photons and Atoms* (Wiley, 1989).
15. H. Ishihara and K. Cho, "Nonlocal theory of the third-order nonlinear optical response of confined excitons," Phys. Rev. B **48**, 7960–7974 (1993).
16. M. Ohtsu, T. Kawazoe, T. Yatsui, and M. Naruse, "Nanophotonics: application of dressed photons to novel photonic devices and systems," IEEE J. Sel. Top. Quantum Electron. **14**, 1404–1417 (2008).
17. T. Inoue and H. Hori, "Quantum theory of radiation in optical near field based on quantization of evanescent electromagnetic waves using detector mode," in *Progress in Nano-Electro-Optics IV*, M. Ohtsu, ed. (Springer Verlag, 2005), 127–199.
18. W. Q. Ma, M. L. Hussein, J. L. Shultz, G. J. Salamo, T. D. Mishima, and M. B. Johnson, "Enhancing the in-plane spatial ordering of quantum dots," Phys. Rev. B **69**, 233312 (2004).
19. E. Ozbay, "Plasmonics: merging photonics and electronics at nanoscale dimensions," Science **311**, 189–193 (2006).
20. E. X. Jin and X. Xu, "Enhanced optical near field from a bowtie aperture," Appl. Phys. Lett. **88**, 153110 (2006).
21. K. Matsuda and T. Saiki, "Local density of states mapping of a field-induced quantum dot by near-field photoluminescence microscopy," Appl. Phys. Lett. **87**, 043112 (2005).
22. E. Runge and C. Lienau, "Near-field wave-function spectroscopy of excitons and biexcitons," Phys. Rev. B **71**, 035347 (2005).
23. C. E. Shannon, "A mathematical theory of communications," Bell Syst. Tech. J. **27**, 379–423 (1948); **27**, 623–656 (1948).
24. P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications," Nature **411**, 1027–1030 (2001).
25. M. A. Neifeld and M. Lee, "Information theoretic framework for the analysis of a slow-light delay device," J. Opt. Soc. Am. B **25**, C31–C38 (2008).
26. P. Oittinen and H. Saarelma, "Average mutual information as a quality measure in imaging processes," J. Opt. Soc. Am. A **3**, 897–901 (1986).
27. D. A. B. Miller, "Fundamental limit for optical components," J. Opt. Soc. Am. B **24**, A1–A18 (2007).
28. A. W. Eckford, "Achievable information rates for molecular communication with distinct molecules," in *Bio-Inspired Models of Network, Information and Computing Systems 2007* (IEEE, 2007), pp. 313–315.
29. R. C. Yu, C. G. Pesce, A. Colman-Lerner, L. Lok, D. Pincus, E. Serra, M. Holl, K. Benjamin, A. Gordon, and R. Brent, "Negative feedback that improves information transmission in yeast signaling," Nature **456**, 755–761 (2008).
30. M. Naruse, T. Inoue, and H. Hori, "Analysis and synthesis of hierarchy in optical near-field interactions at the nanoscale based on angular spectrum," Jpn. J. Appl. Phys. **46**, 6095–6103 (2007).
31. M. Naruse, K. Nishibayashi, T. Kawazoe, K. Akahane, N. Yamamoto, and M. Ohtsu, "Scale-dependent optical near-fields in InAs quantum dots and their application to non-pixelated memory retrieval," Appl. Phys. Express **1**, 072101 (2008).
32. E. Wolf and M. Nieto-Vesperinas, "Analyticity of the angular spectrum amplitude of scattered fields and some of its consequences," J. Opt. Soc. Am. A **2**, 886–889 (1985).
33. M. Naruse, T. Yatsui, J. H. Kim, and M. Ohtsu, "Hierarchy in optical near-fields by nano-scale shape engineering and its application to traceable memory," Appl. Phys. Express **1**, 062004 1-3 (2008).
34. K. Kitamura, T. Yatsui, M. Ohtsu, and G.-C. Yi, "Fabrication of vertically aligned ultrafine ZnO nanorods using metal-organic vapor phase epitaxy with a two-temperature growth method," Nanotechnology **19**, 175305 (2008).
35. T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, 1991).
36. Z. K. Tang, A. Yanase, T. Yasui, Y. Segawa, and K. Cho, "Optical selection rule and oscillator strength of confined exciton system in CuCl thin films," Phys. Rev. Lett. **71**, 1431–1434 (1993).
37. T. Kawazoe, K. Kobayashi, and M. Ohtsu, "Near-field optical chemical vapor deposition using Zn(acac)2 with a non-adiabatic photochemical process," Appl. Phys. B **84**, 247–251 (2006).
38. M. Naruse, T. Yatsui, W. Nomura, K. Hirata, Y. Tabata, and M. Ohtsu, "Analysis of surface roughness of optical elements planarized by nonadiabatic optical near-field etching," J. Appl. Phys. **105**, 063516 (2009).