Optical properties of planar colloidal crystals: Dynamical diffraction and the scalar wave approximation

Daniel M. Mittleman[a)]
Department of Electrical and Computer Engineering, MS-366, Rice University, 6100 Main Street, Houston, Texas 77005

Jane F. Bertone, Peng Jiang, Kevin S. Hwang, and Vicki L. Colvin
Department of Chemistry, MS-60, Rice University, 6100 Main Street Houston, Texas 77005

(Received 5 February 1999; accepted 7 April 1999)

We present a quantitative comparison between two analytic theories for the propagation of electromagnetic waves in periodic dielectric structures. These theories have both been used extensively in the modeling of optical spectra of colloidal crystals exhibiting photonic band gap behavior. We demonstrate that dynamical diffraction theory is equivalent to the scalar wave approximation, in the limit of small dielectric contrast. This equivalence allows us to place quantitative limits on the validity of dynamical diffraction, relative to the predictions of the more accurate scalar wave theory. We also note that dynamical diffraction is often applied with boundary conditions which neglect the strong interference between the incident and diffracted waves within the periodic medium. These boundary conditions lead to expressions for the transmission spectrum which cannot be generalized to the case of normal-incidence propagation. We provide a corrected form for these expressions, and use them in comparisons with experimental spectra. Excellent agreement between theory and experiment is obtained for the widths of optical stop bands, for both positive and negative values of the dielectric contrast. These are among the first quantitative comparisons between theoretical and experimental optical spectra of colloidal photonic crystals.
 [S0021-9606(99)70725-7]

I. INTRODUCTION

There has been a great deal of recent interest in systems which exhibit periodic dielectric properties. In such materials, strong diffraction effects can inhibit the propagation of electromagnetic waves of certain frequencies. With appropriate three-dimensional symmetry and sufficiently large spatial modulation of the dielectric, such systems can manifest a full photonic band gap; that is, a spectral range over which propagation is inhibited regardless of the propagation direction within the periodic medium. Since the early work of Yablonovitch[1] and John,[2] numerous examples of such systems have been proposed, some with rather complicated three-dimensional structures.[3] The early experimental work in this area focused primarily on fabrication of systems using micromachining or other mesoscopic assembly techniques.[4–7] The periodicity of these structures is on a length scale of hundreds of microns, so they naturally exhibit stop bands in the microwave or sub-millimeter range. More recently, attention has been turned to photonic band systems in the optical regime. Such structures require dielectric periodicity with submicron length scales, and are consequently much more difficult to fabricate. The construction of optically active periodic dielectric materials has been pursued along two distinct paths: a ''top-down'' approach, in which lithographic techniques are used to generate a dielectric modulation within a substrate,[8–11] and a ''bottom-up'' approach, in which one exploits the tendency of micron-sized particles to spontaneously self-assemble into ordered arrays.[12–14] In the former case, fabrication of samples consisting of more than a few repeating layers is very difficult. In the latter, a quantitative comparison between theory and experiment is often problematic, either because of the lack of systematic thickness control or because of the high crystal defect density in many of these examples.

Recently, significant advances have been made in the fabrication of close-packed colloidal crystal arrays.[15,16] Using this ''bottom-up'' approach, it is now possible to grow such crystals in planar geometries, with precise thickness control. These samples exhibit high crystalline quality, and can be designed to contain anywhere from a few to several hundred crystalline layers. Unlike gravity-sedimented colloidal crystals, which are quite thick and polycrystalline, these materials contain no grain boundaries. Consequently, one can now measure the thickness dependence of the optical properties of such samples, in particular the evolution of the optical stop band width with increasing thickness. There is an evident need for a simple theoretical description which can explain these results. Our recent work in the fabrication of such samples has led us to explore the various methods used for understanding their optical properties.

In this paper, we compare two widely used theoretical models for the optical properties of periodic dielectric structures. The first, the scalar wave approximation, is a recently developed approach, whereas the second, dynamical diffraction theory, derives from the work of Laue in the early part

[a)]Electronic mail: daniel@rice.edu

of this century. In Sec. II, we describe the class of experiments under consideration, and discuss how these two theories have been applied. Sections III and IV review the derivations of the two theories. This review is important, as it facilitates comparison by putting these results into a common notation, as well as highlighting the similarities and differences in the underlying assumptions. Section V directly compares the fundamental results of the two theories, demonstrating that dynamical diffraction is simply a limit of the scalar wave approach. Given this, it is possible to place quantitative constraints on the validity of dynamical diffraction. These constraints limit the applicability of this older approach. In Sec. VI, we calculate transmission spectra and use these to extract observable parameters from the two theories, such as optical stop band widths and peak optical densities. Because the sample quality has now improved to the point where quantitative comparisons with theory are justified, it is important to calculate such spectra using realistic experimental conditions. Section VII provides a direct comparison between experimental spectra and those calculated using the two theories under consideration. Here, it is possible to evaluate the relative merits of the two, in light of the results of Sec. V. We show that, for the class of samples under consideration here, dynamical diffraction is not valid for quantitative predictions of either the stop band widths or peak heights. The question of the validity of the scalar wave approximation naturally arises here, and this can also be addressed. It is quite accurate at modeling the widths of the photonic band gaps, but it predicts the peak optical densities less well. Several possible explanations for these results are discussed.

II. BACKGROUND

There have been many theoretical treatments of the optical properties of materials with spatially periodic dielectric functions. Many of these are computationally quite intensive, and several authors[17,18] have pointed out the need for a simpler approach which can at least provide a qualitative understanding. One candidate for this is the scalar wave approximation (hereafter referred to as SWA), first applied to the photonic band problem by Satpathy and co-workers.[19] SWA has more frequently been applied to the study of ''top-down'' photonic band gap systems, which usually exhibit optical stop bands in the microwave regime. Although the shortcomings of the SWA have been discussed extensively,[4,20] these difficulties are primarily related to its use in describing propagation along an arbitrary direction in the three-dimensional crystal. In the present work, we will be concerned solely with electromagnetic propagation along high symmetry directions of the crystal (e.g., the [111] axis). In this case, the one-dimensional approximation inherent to the simple SWA is appropriate, and it is in reasonable agreement with experimental observations.[15,21,22]

A second analytic theory of the diffraction from periodic structures is dynamical diffraction theory (referred to as DDT). This treatment is widely used to describe the optical properties of colloidal crystal arrays and similar ''bottom-up'' systems[12,14,23–27]. Here, the DDT has been adapted from its original use in the modeling of X-ray diffraction from crystals.[28–31] Both the SWA and DDT have the advantage that they are relatively simple analytic theories, which can be used to obtain a basic understanding of the on-axis propagation of electromagnetic waves in photonic crystals. They can both be used to model transmission spectra, and thus provide information not only on the spectral position of optical stop bands, but also on the widths and peak extinction coefficients, all as a function of the thickness of the photonic crystal. It is even relatively straightforward in both cases to incorporate weak dissipation (i.e., optical absorption), a topic of some recent interest.[32]

Here, we will consider a particular photonic band gap system, consisting of a planar face-centered-cubic close-packed array of spheres of index $\sqrt{\varepsilon_s}=n_s$, surrounded by an interstitial or background medium of index $\sqrt{\varepsilon_b}=n_b$. It is assumed that both the spheres and the background medium exhibit no absorption. The crystal is oriented such that the [111] axis is normal to the planar surfaces. This example is quite similar to samples described in a number of recent publications.[12–16,24,26] As mentioned above, we will consider propagation only along high symmetry directions of the crystalline lattice, in particular at normal incidence along the [111] crystal axis. The normal-incidence transmission spectra of planar photonic band gap samples is an important experimental diagnostic tool. It is a very useful method for obtaining a rapid characterization of sample thickness, crystal quality, and uniformity over a mm^2 or larger area. We note that both the SWA and the DDT have been used to describe experimental results of this nature.

In this paper, we will compare the SWA and the DDT, demonstrating that they are equivalent in certain appropriate limits. One of these limits involves the dielectric contrast between the two materials which comprise the composite structure, given by

$$\eta=\frac{n_s^2}{n_b^2}-1. \tag{1}$$

The DDT is derived with the explicit assumption that $\eta \ll 1$; this is, of course, perfectly appropriate for x-ray scattering, where $|\eta|\sim 10^{-4}$ or smaller.[28] However, much of the recent work in this field has been directed toward the fabrication of samples with much larger values of η, in part because larger values of η are favored for achieving a complete three-dimensional photonic gap.[3,4] In view of this fact, it is necessary to explore the validity of the assumption of small index contrast, particularly in the case of the colloidal crystals under consideration here.

We show below that the results of the DDT are equivalent to those of the SWA when these latter are expanded to lowest nonvanishing order in the dielectric contrast. Because of this equivalence, it is possible to evaluate the validity of the DDT as a function of the size of the contrast. We find that, even for quite small dielectric contrast, such as the $\eta \sim 0.1$ appropriate for the samples of Ref. 26 (silica spheres in a water background), the DDT deviates measurably from the SWA. Using the results from the SWA as a guide, we can provide a simple prescription for determining the limits of validity of the DDT.

III. SCALAR WAVE APPROXIMATION

In order to simplify the description of electromagnetic propagation in periodic dielectrics, two basic assumptions are made within the scalar wave approximation. The first, as the name implies, is that the electric field in the medium is treated as a scalar rather than a vector quantity. It is then expanded in a Bloch sum over all reciprocal lattice vectors $\mathbf{G}$,

$$E(\mathbf{r})=\sum_{\mathbf{G}}\sum_{\mathbf{k}} C_{\mathbf{k}-\mathbf{G}}e^{i(\mathbf{k}-\mathbf{G})\cdot\mathbf{r}}, \tag{2}$$

where the second sum (over $\mathbf{k}$) is taken over wave vectors in the first Brillouin zone. The dielectric function of the crystal is written as a constant plus a periodically varying term,

$$\varepsilon(\mathbf{r})=\varepsilon_0+\sum_{\mathbf{G}} U_{\mathbf{G}}\exp\{i\mathbf{G}\cdot\mathbf{r}\}, \tag{3}$$

where the average dielectric ε_0 depends on the volume fraction ϕ occupied by the spheres,

$$\varepsilon_0=\phi\varepsilon_s+(1-\phi)\varepsilon_b. \tag{4}$$

In the case under consideration here (close-packed spheres), $\phi=\pi\sqrt{2}/6\approx 0.74$. Equations (1) and (4) may be rewritten as

$$\eta=\frac{1}{\phi}\times\left(\frac{\varepsilon_0}{\varepsilon_b}-1\right). \tag{5}$$

The Fourier coefficient $U_{\mathbf{G}}$ is given by the well-known Rayleigh–Gans expression,

$$U_{\mathbf{G}}=\frac{3\phi}{(GR)^3}(\varepsilon_s-\varepsilon_b)[\sin(GR)-GR\cos(GR)], \tag{6}$$

where R is the radius of the spheres and G is the magnitude of the reciprocal lattice vector $\mathbf{G}$ associated with $U_{\mathbf{G}}$. Equations (2) and (3) may be inserted into the electromagnetic wave equation for $E(\mathbf{r})$, derived from the free-space Maxwell's equations. The result is a system of coupled linear equations which form an eigenvalue problem for the wave vector $k(\omega)$ inside the photonic crystal. Determining $k(\omega)$ uniquely specifies all aspects of the propagation of the radiation inside the crystal, including both the variations in the amplitude (via the imaginary part of k) and the phase (via the real part) with distance.

The second assumption of the scalar wave approach aims to simplify this many-dimensional eigenvalue problem. It is assumed that the conditions of the experiment strongly favor scattering off of one particular set of lattice planes, and that the effects of all other lattice planes may be neglected. This is justified by noting that the scattering probability is determined by the Fourier coefficient $U_{\mathbf{G}}$, which decreases with increasing G. Thus, only the shortest reciprocal lattice vectors contribute significantly. For example, in the case under consideration, the incoming radiation propagates normal to the [111] lattice planes. Thus, one need only consider diffraction off of these planes. In the system of equations, one need only keep terms with $\mathbf{G}=\mathbf{0}$ or $\mathbf{G}=\mathbf{G_L}$, where $\mathbf{G_L}$ is the shortest reciprocal lattice vector along the L direction (i.e., corresponding to scattering off of [111] planes). This simplification reduces the problem to a one-dimensional analytic form, in which the infinite set of equations reduces to only two. These may be written as

$$\left(k^2-\varepsilon_0\frac{\omega^2}{c^2}\right)C_k-\frac{\omega^2}{c^2}U_G C_{k-G}=0, \tag{7a}$$

$$-\frac{\omega^2}{c^2}U_G C_k+\left((k-G)^2-\varepsilon_0\frac{\omega^2}{c^2}\right)C_{k-G}=0, \tag{7b}$$

where we have dropped the subscript L on the reciprocal lattice vector. G is now a scalar, rather than a vector, and is given by $G=2\pi/d_{111}$, where d_{111} is the spacing between adjacent [111] lattice planes. In order for these two equations to be simultaneously satisfied, their determinant must vanish. Thus, one may determine the wave vector of the radiation inside the photonic crystal, $k(\omega)$,[19]

$$k_{\mathrm{SWA}}(\omega)=\frac{1}{2}G\pm\sqrt{\frac{G^2}{4}+\varepsilon_0\frac{\omega^2}{c^2}-\sqrt{G^2\varepsilon_0\frac{\omega^2}{c^2}+U_G^2\frac{\omega^4}{c^4}}}. \tag{8}$$

This is the basic result of the scalar wave approximation. It is worth noting that the difficulties with the SWA, which have been ascribed solely to the first assumption (the scalar field),[4,20] are more likely due to the combination of the two. In other words, it is possible to keep the full vector nature of the field and still neglect all but one of the reciprocal lattice points in Eqs. (2) and (3). In doing so, one derives a result which is flawed in much the same way as the SWA, in the sense that it too is valid only for wave propagation along high symmetry directions of the crystal. In fact, this is the basis for the theory of dynamical diffraction.

IV. DYNAMICAL DIFFRACTION

The derivation of the results of DDT can be found in many texts on x-ray diffraction.[28,29,31] It is often described as the next step beyond the simple kinematic theory, whose result is embodied in Bragg's Law. In the dynamical theory, one determines the wave vector of the propagating field by placing a self-consistency condition on the superposed incident and diffracted waves. To simplify the problem, it is assumed that the Laue condition, $\mathbf{k}_{\mathrm{scat}}=\mathbf{k}_{\mathrm{in}}+\mathbf{G}$, is satisfied for one and only one particular reciprocal lattice vector, so that there is only one diffracted wave. This is entirely equivalent to the second assumption of the SWA. One proceeds along essentially the same lines as above, with two exceptions. First, the vector nature of the field is preserved. This adds some additional factors to the result, involving, for example, the ratio of the direction cosines of the incident and diffracted waves. In the case of normal incidence propagation under consideration here, these factors are of little importance. Indeed, in light of the fact that the SWA has been demonstrated to be incorrect when one attempts to apply it to off-axis propagation,[4,20] one must conclude that the same limitation applies to DDT as well.

The second exception is far more significant. It is assumed that the refractive index of the crystal is only very slightly different from unity. As in the SWA, the dielectric is decomposed into an average part and a spatially varying part.

In Ref. 28, for example, the average part of ε is denoted by $1+\psi_0$, where ψ_0 is assumed to be small compared to one. In the subsequent derivation, all terms of higher than linear order in ψ_0 are neglected. We shall see that this places severe constraints on the validity of DDT, as compared to SWA. We also note that, in most treatments of DDT,[28,29] the interstitial regions are always assumed to have $\varepsilon_b=1$. Here, we treat the more general case, with the modified definition of ψ_0,

$$\psi_0=\frac{\varepsilon_0}{\varepsilon_b}-1. \tag{9}$$

Thus, ψ_0 is related to the dielectric contrast η by $\eta\phi=\psi_0$ [see Eq. (5)]. ψ_0 is therefore a volume-fraction-weighted dielectric contrast parameter, and is a more useful variable than η. In the rest of this work, we use ψ_0 rather than η as the dielectric contrast.

With this approximation, one can derive a pair of coupled linear equations within DDT which are exactly analogous to Eqs. (7a) and (7b) above for the SWA. In fact, by making the following replacements:

$$k^2 \rightarrow k_0^2\times\varepsilon_b\times(1+2\delta_0), \tag{10a}$$

$$(k-G)^2 \rightarrow k_0^2\times\varepsilon_b\times(1+2\delta_G), \tag{10b}$$

and using Eq. (9), one immediately derives

$$(2\delta_0-\psi_0)C_k-\frac{U_G}{\varepsilon_b}C_{k-G}=0, \tag{11a}$$

$$(2\delta_G-\psi_0)C_{k-G}-\frac{U_G}{\varepsilon_b}C_k=0. \tag{11b}$$

These are the constituent equations of DDT.[33] In Eqs. (10a) and (10b), k_0 is the free-space wave vector of the radiation, $k_0^2=\omega^2/c^2$. δ_0 and δ_G are related to the magnitudes of the incident and diffracted wave vectors, respectively. Taking Eqs. (10a) and (10b) as equivalences, we may use them to relate δ_0 to δ_G, with the result that,

$$\delta_G=\delta_0+\left(\frac{G^2}{2\varepsilon_b k_0^2}-\frac{G\sqrt{1+2\delta_0}}{\sqrt{\varepsilon_b}k_0}\right). \tag{12}$$

Now, we introduce a dimensionless wavelength variable Λ, defined by

$$\Lambda=\frac{\lambda}{\lambda_B}-1. \tag{13}$$

Here, λ_B is the wavelength corresponding to the Bragg condition, and is given at normal incidence by

$$\lambda_B=2d_{111}\sqrt{\varepsilon_0}=\frac{4\pi\sqrt{\varepsilon_0}}{G}. \tag{14}$$

Using Eqs. (9), (13), and (14), it is possible to rewrite the relationship between δ_0 and δ_G [Eq. (12)] in terms of the dimensionless parameters Λ and ψ_0. Then, in the spirit of the DDT approximation, this can be expanded to lowest order in ψ_0,δ_0, and also the wavelength parameter Λ. The result of this expansion is

$$\delta_G=-\delta_0+2\Lambda+\psi_0. \tag{15}$$

Inserting this expression into Eq. (11b) allows one to determine two eigenvalues $\delta_\pm$ from the coupled equations, and thus the wave vector of the radiation propagating in the photonic crystal. The result is

$$\begin{aligned}k_{\mathrm{DDT}}(\omega)&=k_0\sqrt{\varepsilon_0}(1+\delta_\pm-\psi_0/2)\\&=k_0\sqrt{\varepsilon_0}\left(1+\Lambda\pm\frac{1}{2}\sqrt{4\Lambda^2-\frac{U_G^2}{\varepsilon_b^2}}\right).\end{aligned} \tag{16}$$

It is important to note that this wave vector is somewhat different from the one customarily seen in treatments of DDT. In particular, as shown below, the form of this wave vector implies that the wavelength at which the diffraction effects are most prominent is given by $\Lambda=0$, identical to the Bragg condition. That is, the wavelength at which the optical transmission spectrum peaks, λ_{peak}, is equal to λ_B, independent of the size of the dielectric contrast. This is in marked contrast with the well-known DDT result that $\lambda_{\mathrm{peak}}=\lambda_B(1-\psi_0/2)$.[34] This consequence of DDT follows from the boundary condition that the external and internal incident electric fields must match at the crystal input facet. With this boundary condition, one can derive a result, analogous to Eq. (15), which is that $\delta_G=-\delta_0+2\Lambda$. Using this expression in place of Eq. (15) gives a wave vector which does manifest the familiar shift of the transmission peak.

In the derivation provided here, this boundary condition is not used, because it neglects the interference between the incident and diffracted waves, an effect which is critically important for normal-incidence transmission. Instead, we make use of the condition, implicit in Eqs. (11a) and (11b), that the wave vectors of the internal incident and diffracted waves differ by the reciprocal lattice vector G. This leads to Eq. (15) as described above. It is worth emphasizing that the results derived using inappropriate boundary conditions cannot be used to model normal-incidence optical spectra, despite numerous attempts in the recent literature to do so.

This fact can be demonstrated experimentally, by measuring the wavelength of the transmission peak for a number of different values of the dielectric contrast ψ_0. The method used to control ψ_0 involves filling the interstitial pores of a colloidal crystal sample with various dielectric fluids, and is described in more detail in Sec. VII below. Figure 1 shows the results of such an experiment (circles), along with the predicted spectral positions from Eq. (16) (solid line) and from the familiar DDT result based on the inappropriate boundary conditions (dotted line).[34] It is clear that the optical stop band does not even qualitatively follow the results of the traditional DDT, whereas the modified DDT result presented here provides an adequate explanation.

We note that the SWA wave vector, Eq. (8), also predicts that the transmission peak occurs at $\Lambda=0$, so this parameter is not a particularly useful point of comparison between the two theories. Also, these predictions for λ_{peak} are valid only for thick samples; evaluation of the full transmission spectrum [see Eq. (23) below] demonstrates a weak dependence on the thickness of the sample. This dependence, illustrated in the inset to Fig. 1 for $\psi_0=0.75$, is most significant for samples with fewer than ~20 repeating layers.

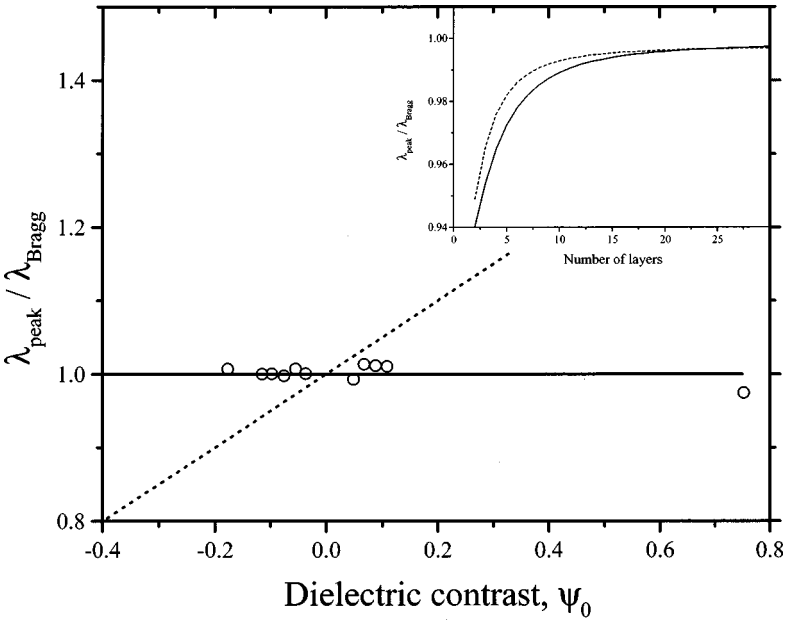

FIG. 1. Wavelength of the transmission peak λ_{peak}, as a function of the dielectric contrast ψ_0. Open circles are experimental data, obtained using a sample of close-packed silica spheres of ~345 nm diameter. The dielectric contrast is controlled by filling the interstitial pores with fluids of known index. The data point at the far left is the result for air in the interstitial region, corresponding to $\psi_0 \sim 0.75$. The remaining ten data points are for various different liquids, with indices ranging from 1.326 to 1.628. The horizontal solid line is the predicted position $\lambda_{peak}/\lambda_{Bragg}$ from both the SWA and the modified DDT presented here. The sloped dotted line is the prediction of the traditional DDT, as in Refs. 28 or 29. Inset: The variation in peak position as a function of the number of [111] layers in the sample. The solid line shows the result for SWA, while the dashed line shows the result for DDT. These results assume a silica/air sample, with $\psi_0 \sim 0.75$, and a substrate with $n=1.5$, and are calculated using Eq. (23) from the text.

V. COMPARING THE TWO APPROACHES

Using Eq. (13), it is possible to write the SWA wave vector [Eq. (8)] in terms of the dimensionless wavelength Λ. This result is

$$k_{SWA}(\omega)=k_0\sqrt{\varepsilon_0}\left(1+\Lambda \pm \sqrt{(1+\Lambda)^2+1-\sqrt{4\cdot(1+\Lambda)^2+\frac{U_G^2}{\varepsilon_0^2}}}\right). \tag{17}$$

We may now directly compare Eqs. (16) and (17), to investigate the behavior of the two wave vectors. Figure 2 shows comparisons of the real parts of the two wave vectors, over a portion of the Brillouin zone near the band gap at $k=G/2$. Except near the band gap, the wave vector is nearly equal to its unperturbed value, determined by the average index of the crystal, $k(\omega)=\sqrt{\varepsilon_0}\omega/c$. Figure 2(a) shows the case of a relatively small dielectric contrast. This models the situation in which the spheres are composed of silica ($n \sim 1.42$), and the interstitial regions are filled with water ($n=1.33$), corresponding to a dielectric contrast of $\psi_0=0.1$. From the point of view of x-ray scattering, this is an enormous index contrast, three orders of magnitude larger than the limits to which DDT was originally intended to conform. Yet, the real part of the DDT wave vector is indistinguishable from the SWA wave vector in this case. Figures 2(b) and 2(c) show the results for increasing values of the dielectric contrast. As ψ_0 increases, the differences between the two curves become more evident. The width of the gap region increases faster in DDT (dashed curves) than in the SWA (solid). As shown below, this is accompanied by a more rapid increase in the peak optical density of the sample. In Fig. 2(c), the value $\psi_0=0.75$ corresponds to the case of silica spheres with air in the interstitial regions.

It is instructive to compare the imaginary parts of the two wave vectors as well, since here the differences between the two theories are more evident at small values of ψ_0. Figure 3 shows such a comparison, for the same index contrasts used in Fig. 2. As in Fig. 2, these have been plotted against the dimensionless wavelength Λ. As expected, the imaginary part of k is nonzero only in a small region near $\Lambda=0$. In Fig. 3a, the ~10% difference in peak width is difficult to discern on this scale, but the difference in peak height is more evident. Clearly, larger dielectric contrasts lead to dramatically larger extinction, as shown by the different scaling of the ordinal axes in the three cases. However, as ψ_0 increases, so do the differences between the two results. For larger dielectric contrast, DDT overestimates the magnitude of the imaginary part of k, and therefore the strength of the photonic band, relative to the prediction of the SWA. As anticipated above, it also overestimates the width of the band gap. As noted above, the spectral positions of the peaks are the same in the two theories.

In Eq. (17) for k_{SWA}, there is no assumption as to the magnitude of either the wavelength Λ or the dielectric contrast. To explore the comparison further, we may expand the inner square root in Eq. (17), keeping only the lowest order term in U_G, and also dropping terms of order ΛU_G^2. This gives an approximate form for k_{SWA} which is quite similar to the exact DDT result,

$$k_{SWA}(\omega)\approx k_0\sqrt{\varepsilon_0}\left(1+\Lambda\pm\frac{1}{2}\sqrt{4\Lambda^2-\frac{U_G^2}{\varepsilon_0^2}}\right). \tag{18}$$

This expansion clarifies the results of Figs. 2 and 3. The photonic band gap occurs where the wave vector develops a nonvanishing imaginary component,[21] since the imaginary part of the wave vector is proportional to the extinction coefficient. The imaginary part of the wave vector evidently originates from the square root, so the widths of the gaps may be estimated by the widths of the regions where the

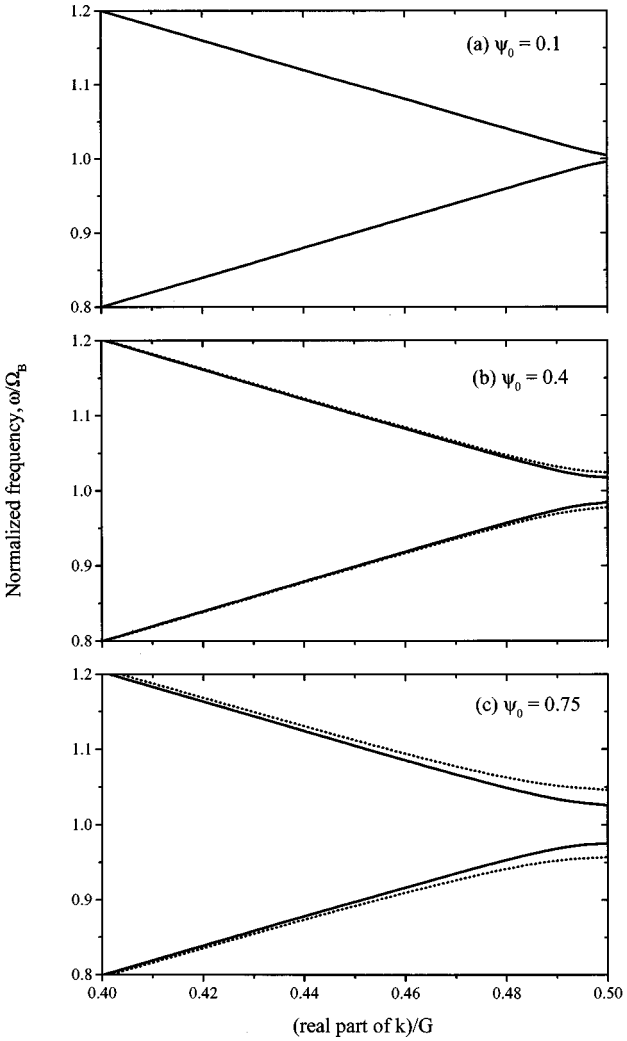

FIG. 2. Real part of the wave vector of radiation propagating in the colloidal crystal, for three different values of the dielectric contrast ψ_0. The vertical axes are frequency normalized by the Bragg frequency Ω_B. Note that $\omega/\Omega_B = \lambda_B/\lambda$. The horizontal axis is the real part of the wave vector, normalized to the reciprocal lattice vector G. In all three cases, the solid line is $k_{\mathrm{SWA}}(\omega)$ and the dashed line is $k_{\mathrm{DDT}}(\omega)$. Note that, in all three cases, both curves converge to the unperturbed wave vector $\sqrt{\varepsilon_0}\omega/c$ for values of k far from the edge of the Brillouin zone at $G/2$.

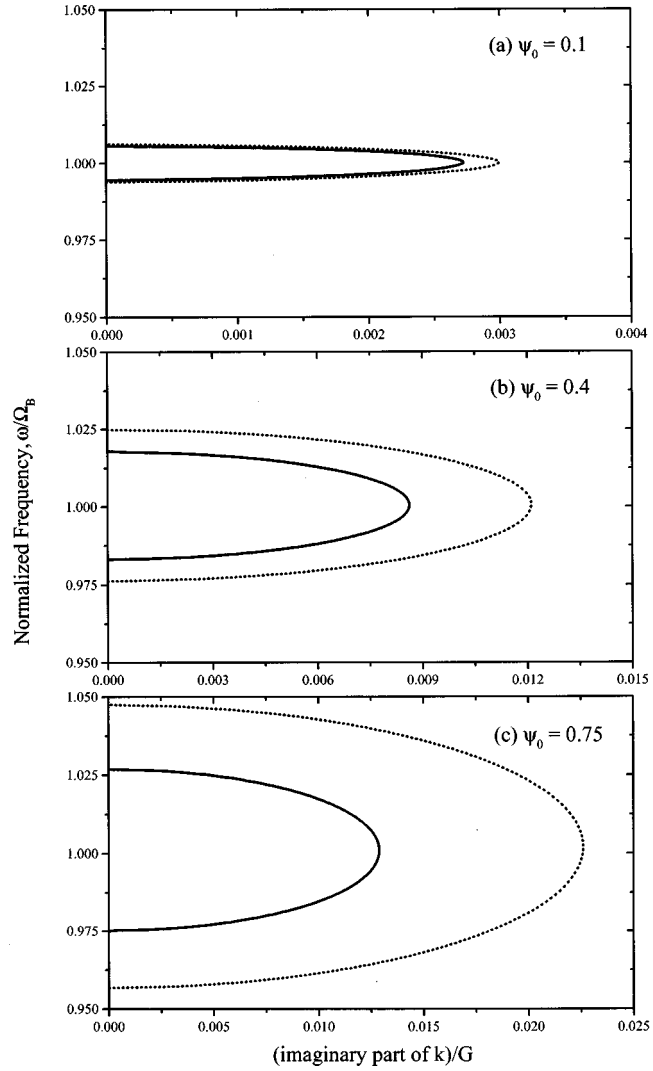

FIG. 3. Imaginary part of the wave vector inside the photonic crystal, for three different values of the dielectric contrast ψ_0. The vertical axes are frequency normalized by the Bragg frequency Ω_B. Note that $\omega/\Omega_B = \lambda_B/\lambda$. The horizontal axis is the imaginary part of the wave vector, normalized to the reciprocal lattice vector G. In all three cases, the solid line is $k_{\mathrm{SWA}}(\omega)$ and the dashed line is $k_{\mathrm{DDT}}(\omega)$. Note the different horizontal scale for the three cases.

imaginary part is nonzero. By inspection of Eqs. (16) and (18), the fractional widths of the gaps may be written $(\Delta\lambda/\lambda_{\mathrm{peak}})_{\mathrm{SWA}} \approx U_G/\varepsilon_0$ and $(\Delta\lambda/\lambda_{\mathrm{peak}})_{\mathrm{DDT}} = U_G/\varepsilon_b$. The ratio of the fractional widths is thus approximately given by $\varepsilon_0/\varepsilon_b$, or $1+\psi_0$. Thus, the percent error in the bandwidth predicted by DDT is approximately equal to ψ_0, at least in the regime where Eq. (18) is a good approximation. This result is explored further below.

Tarhan and Watson[21] have given a more precise formula for the width of the region over which the SWA wave vector has an imaginary component, based on the exact SWA wave vector of Eq. (17). However, this procedure does not provide a measure of an experimentally observable bandwidth, since the full width at half maximum of a measured transmission spectrum depends not only on the imaginary component of the wave vector, but also on the thickness of the sample and the nature of the index matching at both crystal facets. In order to compare directly to experimentally measured widths, one must calculate the transmission spectrum of the photonic crystal using realistic experimental conditions, and numerically extract the width from this simulation.

VI. TRANSMISSION SPECTRA

In order to determine quantitative limits on the validity of DDT relative to SWA, we calculate transmission spectra using the wave vectors derived above, and compare to experimental results. We note that the commonly used DDT transmission spectra, as in Refs. 28 and 31, are not applicable in the case under consideration here, for the following reason. These results are derived for an arbitrary angle of incidence, as appropriate for, e.g., rotating crystal diffraction studies. Thus, the incident and diffracted waves inside of the crystal are not counter-propagating, but propagate at a relative angle determined by the Laue condition, $\mathbf{k}_i = \mathbf{k}_d + \mathbf{G}$.

The boundary conditions at the entrance or exit surfaces of the crystal used to derive these expressions do not take into account the interference of these two waves, since they do not, in general, have the same wave vector. So, for example, the external incident wave is required to be continuous with the internal incident wave only, not with the sum of the internal incident and internal diffracted waves.[35] A similar condition is imposed at the exit facet. Because the interference of the two internal waves is neglected, these boundary conditions lead to expressions which cannot be generalized to the case of normal incidence.

For the case under consideration here, a more appropriate set of boundary conditions are that the electric field and its first derivative be continuous at both boundaries. Without loss of generality, we may treat the electric field as a scalar in the case of normal incidence. Then, the field can be written[17]

$$E(x)=\begin{cases} e^{ik_0x}+r\cdot e^{-ik_0x} & \text{(incident)} \\ C_1(e^{ikx}+\Sigma\cdot e^{i(k-G)x})+C_2(e^{-ikx}+\Sigma\cdot e^{i(k-G)x}) & \\ \qquad\qquad \text{(internal)} & \\ t\cdot e^{ik_0x} & \text{(transmitted)} \end{cases}. \tag{19}$$

The parameter Σ, the ratio of the amplitudes of the incident and diffracted waves inside the crystal, is given by

$$\Sigma=\frac{k^2-\varepsilon_0k_0^2}{U_Gk_0^2}. \tag{20a}$$

for SWA, and by

$$\Sigma=\varepsilon_b\frac{2(k-\sqrt{\varepsilon_0}k_0)}{U_G\sqrt{\varepsilon_0}k_0} \tag{20b}$$

for DDT. As above, $k_0=\omega/c$ is the free-space wave vector. Applying the conditions that $E(x)$ and dE/dx must both be continuous at both boundaries, one can solve for the electric field transmission coefficient,

$$t(\lambda)=\frac{2\beta_0e^{-ik_0Nd_{111}}}{2\beta_0\cos(kNd_{111})-i(1+\beta_0^2)\sin(kNd_{111})}. \tag{21}$$

Here, N is the number of [111] lattice planes (or layers) in the crystal, k is the wave vector from Eq. (16) or (17), and

$$\beta_0=\frac{k_0(1-\Sigma)}{k(1-\Sigma)+G\Sigma}. \tag{22}$$

It should be noted that the samples under consideration are not free standing, but are grown on a glass substrate.[16] The presence of the substrate modifies the transmission coefficient. An expression which is slightly more complex than Eq. (21) is required to take this effect into account. To derive it, one simply takes the transmitted wave to be $t(\lambda)\times\exp(ik_1x)$ in Eq. (19), where k_1 is the wave vector of the radiation in the substrate, $k_1=n_{\text{sub}}\omega/c$. The result is

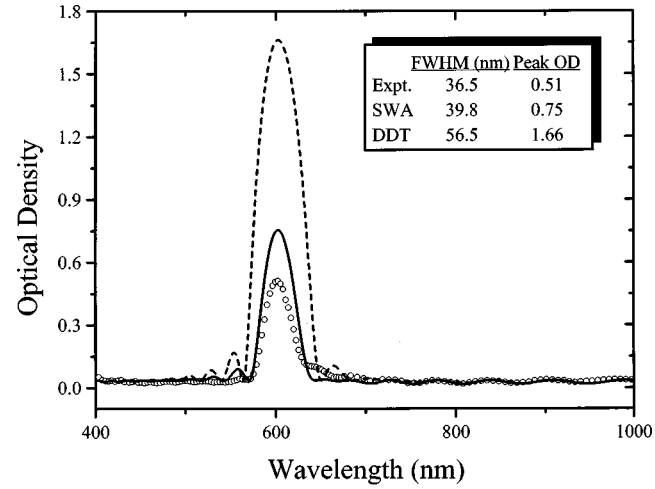

FIG. 4. Comparison between experimental and calculated transmission spectra. The experimental data (open circles) are the normal-incidence transmission spectra of a sample consisting of 18 layers of close-packed silica spheres of 280 nm diameter, with a dielectric contrast of $\psi_0=0.753$. These data have been corrected by subtraction of a smoothly varying background, but are otherwise unaltered. The solid curve represents the simulated transmission spectrum using the SWA wave vector, whereas the dashed curve is from the DDT model. These simulations contain no adjustable parameters. The inset compares the full widths at half maxima and peak optical densities (OD) extracted from the three curves.

$$t(\lambda)=\frac{2\beta_0e^{-ik_1Nd_{111}}e^{i\Phi}}{(\beta_0+\beta_1)\cos(kNd_{111})-i(1+\beta_0\beta_1)\sin(kNd_{111})}, \tag{23}$$

where $\Phi=(k_1-k_0)d_{111}/2$. β_1 is the same as β_0, except that the k_0 in its numerator is replaced by k_1. The presence of a superstrate other than air can be accounted for as well, simply by replacing k_0 with $n_{\text{super}}\omega/c$, where n_{super} is the index of the superstrate. All of the results simulated below are obtained for the optical density (OD) spectrum, rather than the transmission spectrum, calculated according to OD $=-\log_{10}[|t(\lambda)|^2]$. The use of Eq. (23) in place of Eq. (21) for $t(\lambda)$ can be quite important when making comparisons between experiment and theory. This point is illustrated below.

VII. DISCUSSION

We use the transmission function derived above [Eq. (23)] to calculate spectra using both the SWA and DDT wave vectors, for comparison with experimental results. Figures 4 and 5 show observed transmission (optical density) spectra for two different samples, along with the calculated spectra from both DDT and SWA, using Eq. (23). In Fig. 4, the sample consists of 18 layers of close-packed silica spheres, of diameter 280 nm, with air in the interstitial regions.[16] This configuration has a dielectric contrast of $\psi_0=0.753$. The sample is supported by a glass substrate, but has no superstrate. In Fig. 5, the sample is an inverted structure, in which the spheres are air, and the interstitial regions are filled with a dielectric. This is accomplished by first fabricating a silica/air sample, as in Fig. 4. The interstitial regions are subsequently filled with a polymer solution. After the polymer is cured, the silica is etched away, leaving the

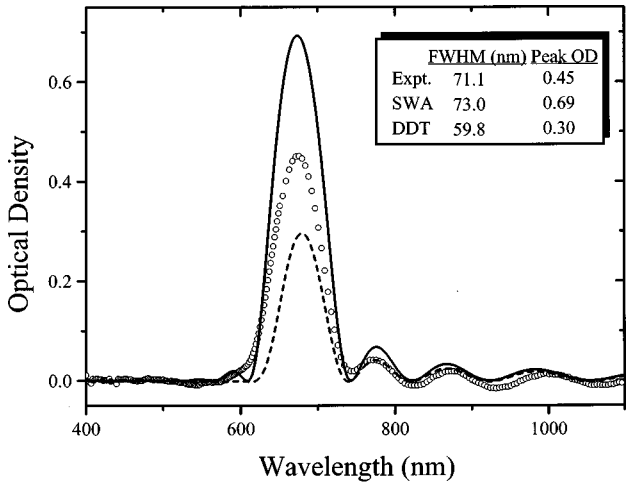

FIG. 5. Comparison between experimental and calculated transmission spectra. The experimental data (open circles) are the normal-incidence transmission spectra of a sample consisting of 11 layers of close-packed air spheres of 360 nm diameter, with a polymer ($n=1.49$) in the interstitial regions. This sample has a dielectric contrast of $\psi_0=-0.398$. These data have been corrected by subtraction of a smoothly varying background, but are otherwise unaltered. The solid curve represents the simulated transmission spectrum using the SWA wave vector, whereas the dashed curve is from the DDT model. These simulations contain no adjustable parameters. The inset compares the full widths at half maxima and peak optical densities (OD) extracted from the three curves.

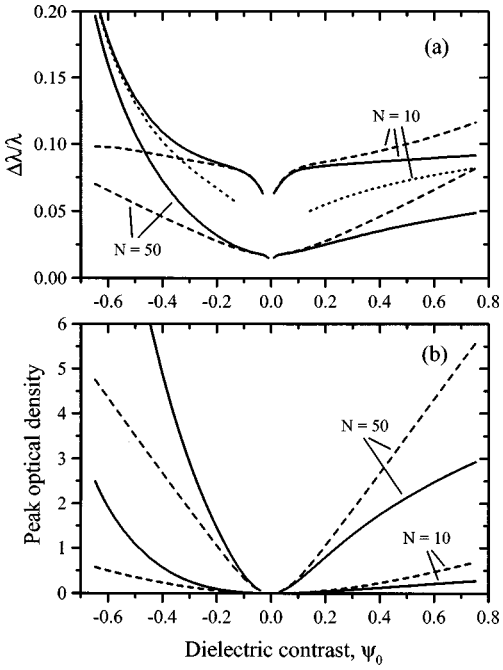

FIG. 6. (a) Fractional bandwidths $\Delta\lambda/\lambda_{\text{peak}}$ and (b) peak optical densities predicted by the two models discussed in the text, as a function of the dielectric contrast ψ_0. In both cases, the dashed lines represent the DDT, whereas the solid lines represent the SWA. In (a) and (b), the calculation has been performed for samples consisting of $N=10$ and 50 layers of close-packed silica spheres of refractive index 1.42 and diameter 280 nm. For these samples, a contrast of 0.75 corresponds to air in the interstitial regions. A substrate of index 1.5 and a superstrate of index equal to the index of the material filling the interstitial layers have been included. The dotted line in (a) shows the same result as the SWA calculation (solid line) for $N=10$ layers if the substrate and superstrate are both removed. The gaps in the curves near $\psi_0=0$ occur because, for small contrast, the peak in the spectrum becomes comparable in size to the adjacent Fabry-Perot fringes, and the full width at half maximum is thus not well defined.

polymer scaffold behind.[36] The resulting ''inside-out'' sample has a dielectric contrast of $\psi_0=-0.398$. In Fig. 5, the sample consists of 11 layers of air spheres, each of ~360 nm diameter. These polymer samples have sufficient mechanical stability to be removed from the glass substrate, and are free standing for these measurements. The room-temperature, normal-incidence optical spectra of both samples are obtained using a commercial ultraviolet (UV)-visible spectrometer. A smooth background component, monotonically increasing to shorter wavelengths, has been subtracted from the spectra. The origins of this background are somewhat unclear, but may result from surface roughness scattering, point or stacking defects in the crystal structure, Fresnel losses, or some combination of these. The resulting background-subtracted data are plotted as open circles in Figs. 4 and 5, along with simulated spectra calculated using the formalism outlined above. We emphasize that there are no adjustable parameters in the simulated curves in either Fig. 4 or Fig. 5, since the sphere diameter and the number of layers are determined using scanning electron microscopy, and the indices of the silica nanospheres ($n=1.42$) and the polymer ($n=1.49$) are known. In both figures, the full width at half maxima and peak optical densities of the experimental curve as well as both simulations are shown.

From simulated transmission spectra such as those shown in Figs. 4 and 5, we may numerically extract experimentally observable quantities such as peak full width at half maxima and peak optical densities. Figure 6 shows the calculated fractional bandwidths and optical densities as a function of dielectric contrast. In both Figs. 6(a) and 6(b), the solid lines show the result obtained using the SWA wave vector [Eq. (17)], and the dashed lines show the result from the DDT wave vector [Eq. (16)]. Results are shown for two different thicknesses (10 and 50 layers). As anticipated above, the DDT predicts larger bandwidths and higher optical densities than the SWA, for $\psi_0>0$. The reverse is the case for negative index contrast. The gaps in the curves near $\psi_0=0$ arise because, when the sample is nearly perfectly index-matched, the peak amplitude drops below the amplitude of the adjacent Fabry-Perot fringes, and the definition of the bandwidth and the peak height become ambiguous. This effect is obviously more severe in thinner samples, where the fringes have larger amplitude.

An experiment of the type suggested by Fig. 6, in which the dielectric contrast is varied in a controlled fashion without changing other characteristics of the sample, can be performed by filling the interstitial pores of the sample with various nonabsorbing fluids of known refractive index n_{fluid}.[37] In order to accurately simulate such an experiment, one must use Eq. (23) rather than Eq. (21), since it would not be possible to fill the pores without also creating a liquid ''superstrate'' on top of the sample. So, the simulations shown in Fig. 6 include a superstrate of index n_{fluid} as well as a glass substrate of index 1.5. To illustrate the importance of these effects, one of the four cases (the SWA calculation for $N=10$ layers) is repeated for a free standing sample, n_{super}

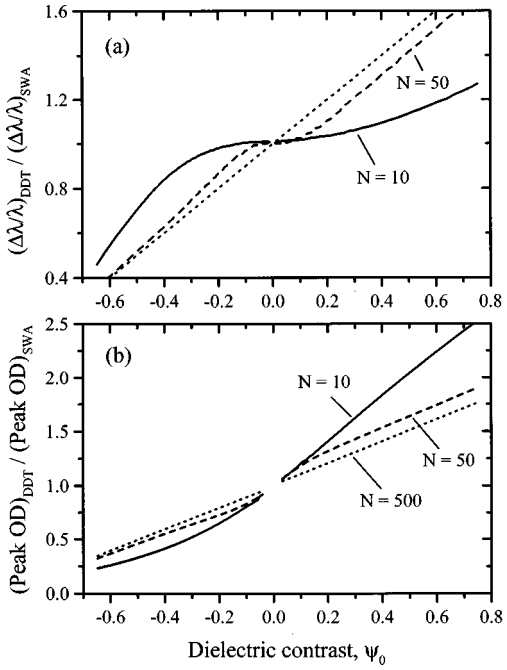

FIG. 7. Ratios of the predicted fractional widths (a) and the predicted peak optical densities (b) as a function of the dielectric contrast ψ_0, for the same situation as in Fig. 3. The solid lines represent a sample with $N=10$ layers, while the dashed line represents a sample with $N=50$ layers. The dotted line in (a) shows the function $1+\psi_0$, which is the approximate analytical result for the width ratio, as discussed in the text. In (b), the dotted line shows the result for $N=500$ layers, which may be taken as the limiting behavior for very thick samples.

$=n_{\text{sub}}=1$. This result is shown as a dotted line in Fig. 6(a), and clearly demonstrates the large difference in the predicted widths which results from the presence of the substrate and superstrate. These differences are less dramatic in the case of the peak optical density, and also become less significant with increasing film thickness.

Figure 7 shows the ratios of the fractional widths (a) and peak heights (b) from the two theories, to highlight the regimes where the SWA and DDT begin to diverge substantially from one another. These results are calculated using the same superstrate and substrate conditions as in Fig. 6. Here, the solid lines show the results for the $N=10$ layer sample, while the dashed lines show the $N=50$ layer result. In Fig. 7(a), the dotted line shows the function $1+\psi_0$, which, as noted above, is the limiting behavior of the bandwidth ratio. In Fig. 7(b), the calculation has been repeated for a 500 layer sample; this result, shown as a dotted line, may be taken as the infinite thickness limit of the optical density ratio. From these data, it is possible to place a limit on the value of ψ_0, above which DDT is no longer equivalent to the more precise SWA result. Evidently, this limit depends on the thickness of the sample. For example, for a sample consisting of ten layers, the fractional bandwidth predicted by DDT is within 5% of the value obtained from the SWA only if $|\psi_0|\lesssim 0.27$. For 50 layers, this becomes $|\psi_0|\lesssim 0.13$, and for a very thick sample, a 5% deviation occurs when $|\psi_0| = 0.05$. Similar analyses can be applied to the optical densities predicted by the two theories. However, as seen in Figs. 4 and 5, the accuracy with which the SWA predicts experimentally measured optical densities is not nearly as good as its predictions of bandwidths.

In both Figs. 4 and 5, the widths extracted from the experimental spectra are quite accurately described with the SWA, and not with the DDT. As noted above, this is a direct result of the small dielectric contrast approximation which constitutes the most important difference between these two theories. Controlling and understanding the spectral width of the optical stop band is an extremely important element in the design of practical photonic systems. We will address this issue more fully in a separate publication.[22]

It is also clear from these figures that neither simulation accurately matches the heights of the peaks. One possible explanation for these discrepancies is that they reflect the fundamental approximation which underlies both SWA and DDT, in which scattering from all but one reciprocal lattice point is neglected. However, it is interesting to note that the more accurate SWA consistently overestimates the peak optical density, for both $\psi_0>0$ and $\psi_0<0$. If one performed a more complete treatment, in which other scattering channels were included, one might expect that the predicted optical density would increase as a result, or at least not decrease. This suggests that the disagreement is primarily the result of other factors.

It is also possible that certain types of crystalline defects are present in these samples, which act to reduce the effective diffraction strength, and thus lower the peak optical density relative to the predicted value. It is not clear what sort of defect would have the effect of reducing the optical density while not increasing the linewidth significantly. One likely candidate is a planar stacking disorder, in which the subsequent [111] planes are stacked with neither the *ABCABCABC* pattern of the face-centered-cubic lattice nor the *ABABAB* pattern of the hexagonal close-packed lattice, but instead with a random stacking of adjacent planes. The existence of this type of disorder seems plausible in view of the weak interactions between next-nearest neighbors in these lattices, and of the kinematics of the sample growth process.[16] If such a fault were present, it would be difficult to identify using scanning electron microscopy (SEM). Stacking faults of this type have been observed in similar samples using confocal microscopy.[38] The presence of a stacking fault of this type is known to influence both the widths and amplitudes of diffraction peaks in x-ray scattering,[39] although it is not clear how the relative magnitudes of these effects manifest themselves in the optical regime. Efforts to model the optical spectrum of a sample containing stacking disorder are currently under way.

VIII. CONCLUSION

We have shown that dynamical diffraction theory, evaluated using appropriate electromagnetic boundary conditions, is equivalent to the scalar wave approximation in the limit of small dielectric contrast. We note that numerous attempts have been made to explain the optical properties of colloidal crystals using DDT, primarily by applying analytic formulas directly from x-ray diffraction texts. These attempts have not been successful for two important reasons. First, the samples often have dielectric contrasts larger than 0.1, which places

them well outside of the regime of validity of the approximate DDT. Second, these analytic results for the transmission spectra rely on the application of boundary conditions which are entirely inappropriate for the optical experiments under consideration. A compelling illustration of this latter difficulty is the well-known shift of the transmission peak from the Bragg wavelength with increasing dielectric contrast, a result which is in direct contradiction with experiments.

Because of the equivalence of DDT and the SWA, it is possible to evaluate the validity of DDT as a function of various parameters including ψ_0 and N, the number of crystalline layers. Perhaps more importantly, this equivalence casts doubt on the use of DDT for descriptions of angle-dependence studies, since it is well established that the SWA cannot predict, even qualitatively, the optical properties of samples unless the radiation propagates along high symmetry axes. If we restrict ourselves to the case of propagation along high symmetry axes, then it is shown that DDT only gives accurate predictions for small values of $|\psi_0|$.

Comparisons with experimental results indicate that the SWA accurately predicts the bandwidths of the optical stop bands, but not the amplitudes of the transmission peaks. We have proposed one possible explanation for this latter discrepancy; however, since the SWA is a rather severe approximation of the full vector-wave solution, one should not expect perfect agreement. It may be necessary to perform a much more detailed (and cumbersome) calculation in order to fully explain these results. Such a calculation, which is still lacking in the literature, must necessarily account for the realistic experimental situation, including in particular the presence of superstrate and substrate effects. In view of the extraordinary progress in the fabrication of many different kinds of periodic dielectric structures, the results presented here should help to avoid a great deal of confusion in comparisons between experimental results and analytic theoretical models.

ACKNOWLEDGMENTS

This work has been funded in part by the National Science Foundation (Grant No. CHE-9702520) and by the Dreyfus Foundation.

[1] E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987).
[2] S. John, Phys. Rev. Lett. **58**, 2486 (1987).
[3] J. D. Joannopoulos, R. Meade, and J. Winn, *Photonic Crystals* (Princeton University Press, Princeton, NJ, 1995).
[4] See, for example, the reviews by J. W. Haus, J. Mod. Opt. **41**, 195 (1994); E. Yablonovitch, T. J. Gmitter, K. M. Leung, R. Meade, A. M. Rappe, K. D. Brommer, and J. D. Joannopoulous, Opt. Quantum Electron. **24**, S273 (1992).
[5] E. Yablonovitch and T. J. Gmitter, Phys. Rev. Lett. **63**, 1950 (1989).
[6] E. Yablonovitch, J. Opt. Soc. Am. B **10**, 283 (1993).
[7] E. Ozbay, E. Michel, G. Tuttle, R. Biswas, K. M. Ho, J. Bostak, and D. M. Bloom, Appl. Phys. Lett. **65**, 1617 (1994).
[8] C. Cheng, A. Scherer, R.-C. Tyan, Y. Fainman, G. Witzgall, and E. Yablonovitch, J. Vac. Sci. Technol. B **15**, 2764 (1997).
[9] G. Feiertag, W. Ehrfeld, H. Freimuth, H. Kolle, H. Lehr, M. Schmidt, M. Sigalas, C. M. Soukoulis, G. Kiriakidis, T. Pedersen, J. Kuhl, and W. Koenig, Appl. Phys. Lett. **71**, 1441 (1997).
[10] M. Wada, Y. Doi, K. Inoue, J. W. Haus, and Z. Yuan, Appl. Phys. Lett. **70**, 2966 (1997).
[11] D. Labilloy, H. Benisty, C. Weisbuch, T. F. Krauss, R. M. De La Rue, V. Bardinal, R. Houdre, U. Oesterle, D. Cassagne, and C. Jouanin, Phys. Rev. Lett. **79**, 4147 (1997).
[12] P. A. Rundquist, P. Photinos, S. Jagannathan, and S. A. Asher, J. Chem. Phys. **91**, 4932 (1989).
[13] I. Tarhan and G. Watson, Phys. Rev. Lett. **76**, 315 (1996).
[14] W. L. Vos, R. Sprik, A. von Blaaderen, A. Imhof, A. Lagendijk, and G. Wegdam, Phys. Rev. B **53**, 16231 (1996).
[15] H. Míguez, C. López, F. Meseguer, A. Blanco, L. Vázquez, R. Mayoral, M. Ocaña, V. Fornés, and A. Mifsud, Appl. Phys. Lett. **71**, 1148 (1997).
[16] P. Jiang, J. F. Bertone, K. S. Hwang, and V. L. Colvin, Chem. Mater. (in press).
[17] K. W.-K. Shung and Y. C. Tsai, Phys. Rev. B **48**, 11265 (1993).
[18] P. M. Hui, W. M. Lee, and N. F. Johnson, Solid State Commun. **91**, 65 (1994).
[19] S. Satpathy, Z. Zhang, and M. R. Salehpour, Phys. Rev. Lett. **64**, 1239 (1990).
[20] S. Datta, C. T. Chan, K. M. Ho, and C. M. Soukoulis, Phys. Rev. B **48**, 14936 (1993).
[21] I. Tarhan and G. H. Watson, Phys. Rev. B **54**, 7593 (1996).
[22] J. F. Bertone, P. Jiang, K. Hwang, D. M. Mittleman, and V. L. Colvin, Phys. Rev. Lett. (submitted).
[23] R. J. Spry and D. Kosan, Appl. Spectrosc. **40**, 782 (1986).
[24] Y. Monovoukas, G. G. Fuller, and A. P. Gast, J. Chem. Phys. **93**, 8294 (1990).
[25] H. B. Sunkara, J. M. Jethmalani, and W. T. Ford, Chem. Mater. **6**, 362 (1994).
[26] L. Liu, P. Li, and S. A. Asher, J. Am. Chem. Soc. **119**, 2729 (1997).
[27] G. Pan, R. Kesavamoorthy, and S. A. Asher, J. Am. Chem. Soc. **120**, 6525 (1998).
[28] W. H. Zachariasen, *Theory of X-Ray Diffraction in Crystals* (Dover, New York, 1945).
[29] R. W. James, *The Optical Principles of the Diffraction of X-Rays* (Ox Bow, Woodbridge, CT, 1948).
[30] B. W. Batterman and H. Cole, Rev. Mod. Phys. **36**, 681 (1964).
[31] B. E. Warren, *X-Ray Diffraction* (Dover, New York, 1968).
[32] M. M. Sigalas, C. M. Soukoulis, C. T. Chan, and K. M. Ho, Phys. Rev. B **49**, 11080 (1994).
[33] See, for example, Eq. (3.109) from Ref. 28, or Eq. (8.28) from Ref. 29.
[34] See, for example, Eq. (3.145b) from Ref. 28.
[35] See, for example, Ref. 28, Eqs. (3.127), (3.128), and (3.135).
[36] P. Jiang, J. F. Bertone, K. S. Hwang, D. M. Mittleman, and V. L. Colvin, J. Am. Chem. Soc. (submitted).
[37] V. N. Bogomolov, S. V. Gaponenko, I. N. Germanenko, A. M. Kapitonov, E. P. Petrov, N. V. Gaponenko, A. V. Prokofiev, A. N. Ponyavina, N. I. Silvanovich, and S. M. Samoilovich, Phys. Rev. E **55**, 7619 (1997).
[38] A. van Blaaderen and P. Wiltzius, Science **270**, 1177 (1995).
[39] A. Guinier, *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies* (Dover, New York, 1963).