Limits on the performance of dispersive thin-film stacks

Martina Gerken and David A. B. Miller

Dispersive thin-film stacks are interesting as compact, cost-effective devices for temporal dispersion compensation and wavelength multiplexing. Their performance depends on the total group delay or spatial shift that can be achieved. For general multilayer stacks, no analytic model exists relating the performance to the stack parameters such as the refractive indices and the number of layers. We develop an empirical model by designing and analyzing 623 thin-film stacks with constant dispersion. From this analysis we conclude that, for given stack parameters, the maximum constant dispersion value is inversely proportional to the wavelength range over which the dispersion is achieved. This is equivalent to saying that, for constant dispersion, there is a maximum possible spatial shift (or group delay) that can be achieved for a given material system and number of layers. This empirical model is useful to judge the feasibility of dispersive photonic nanostructures and photonic crystal superprism devices and serves as a first step in the search for an analytic performance model. We predict that an 8-channel wavelength multiplexer can be realized with a single 21-μm-thick SiO_2-Ta_2O_5 thin-film stack.

OCIS codes: 260.2030, 060.4320, 230.4170.

1. Introduction

Traditional thin-film coatings based on the amplitude reflection and transmission properties of multilayer stacks have been used successfully for many types of application.[1] Recently, several groups have proposed exploiting the phase properties of thin-film stacks instead of the amplitude properties. The resulting group-delay properties of normal-incidence thin-film stacks have been designed for temporal dispersion compensation in femtosecond laser cavities[2–4] and optical fibers.[5–7] We have demonstrated use of thin-film stacks at oblique incidence for wavelength multiplexing by designing the phase properties such that beams experience a wavelength-dependent effective propagation angle or lateral shift in the stack. This translates to a wavelength-dependent spatial beam shift and demultiplexing at the output surface.[8,9] Such angle changes and the corresponding lateral shifts are also closely related to group-delay phenomena, as we discuss in Section 2. Figure 1 shows an example of such a device. Performing multiple bounces off the stack increases the wavelength-dependent spatial shift. The number of channels that can be demultiplexed is calculated as the total shift in the x direction divided by the spatial extent per channel.[9] Therefore the performance of the device is determined by the total spatial shift that can be achieved as well as by how accurately the desired dispersion characteristics are matched.

For simple structures, e.g., Fabry–Perot or Gires–Tournois resonators, an analytical relationship exists among the stack parameters (refractive indices, number of layers, incidence angle, polarization), the dispersion profile, and the total spatial shift (or group delay). Thus, for such simple stacks, the performance is related in a straightforward manner to the stack parameters. For general multilayer structures, on the other hand, no analytic model exists. However, it is obvious that, for example, a larger number of layers will allow for a larger dispersion or a better matching of the desired dispersion profile. The minimum number of layers necessary to achieve the desired dispersion characteristics is typically found in a trial-and-error approach. To speed up the design process and to quickly estimate the feasibility of specific dispersive devices, we introduce here an empirical model relating the stack parameters to the achievable performance for the case of multilayer stacks with constant spatial dispersion. We developed the empirical model by designing and analyzing 623 thin-film stacks composed of different material systems,

When this research was performed, D. A. B. Miller (dabm@ee.stanford.edu) and M. Gerken were with the Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305. M. Gerken (martina.gerken@lti.uni-karlsruhe.de) is now with the Lichttechnisches Institut, Universität Karlsruhe (TH), Kaiserstrasse 12, 76131 Karlsruhe, Germany.

Received 29 September 2004; accepted 21 December 2004.
0003-6935/05/163349-09$15.00/0

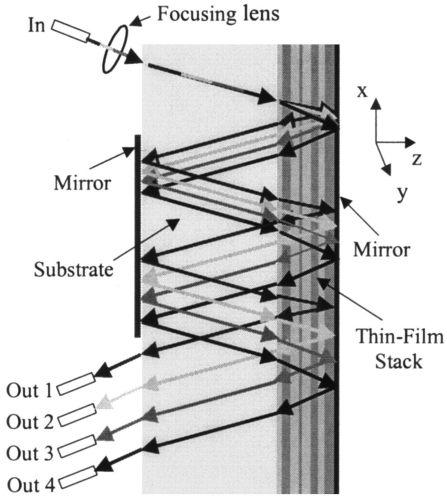

Fig. 1. Wavelength demultiplexer for a multilayer thin-film stack with high spatial dispersion.

different stack thicknesses, and for operation at different incidence angles.

The paper is structured as follows. In Section 2 we review the design of thin-film stacks with constant dispersion. We automated the design process to generate a large number of different designs. In Section 3 we discuss the derivation of the empirical model by analyzing the performance of 623 designed stacks. In Section 4 we use the empirical model to estimate how many channels can be demultiplexed with the device in Fig. 1. We finish with conclusions in Section 5.

2. Automatic Design of Stacks with Constant Dispersion

In this section we describe how we generate thin-film stacks with spatial dispersion and analyze their performance. To decrease the complexity, we limit the design space to designs that have a center wavelength of 1550 nm and operate at p polarization. Different center wavelengths can be obtained by simple scaling whereas the effect of polarization still needs to be investigated at a later time. Furthermore, we consider only designs with constant dispersion, i.e., a linear shift as a function of wavelength. We proceed by specifying a constant dispersion value and specific stack parameters for each design. Then we try to generate the design that exhibits the desired dispersion value over the largest wavelength range. This is the design with the best performance for the given stack parameters as it has the largest total shift allowing for the multiplexing of the largest number of channels. We also calculate the standard deviation of the design from constant dispersion as a measure of the ripple.

Different algorithms have been developed to design dispersive thin-film stacks. These include use of the photonic crystal superprism effect in periodic stacks,[8] the design of chirped stacks with the coupled-mode theory,[2,4,8] coupled-cavity stacks designed by digital signal processing techniques,[6–8] and numerical optimization of arbitrary nonperiodic stacks.[8] We consider only dispersion effects that are due to the multilayer structure of the stacks and we refer to it as structural dispersion to distinguish it from material dispersion, i.e., an isotropic change in the refractive index with wavelength. As discussed in Refs. 10 and 11, structural dispersion is caused by a changing amount of stored energy in the structure with wavelength. Therefore the different design approaches are just different ways of achieving the same goal—designing the amount of stored energy with wavelength to obtain the desired dispersion characteristics. Since the origin of the dispersion is the same in all cases, we expect the achievable performance (the maximum spatial shift or group delay) to be independent of the design algorithm.

In the case of structural dispersion, the group velocity v_{gx} in the x direction along the layers stays approximately constant, and the dispersion is predominantly due to a change in the group velocity v_{gz} in the z direction perpendicular to the layers. The spatial shift s_x and the group delay τ_{group} are related by[8,10,11]

$$\tau_{\text{group}} = \frac{s_x}{v_{gx}}. \tag{1}$$

As discussed in Appendix A, we find that v_{gx} is approximately constant in the case of structural dispersion in a broad range of layered structures, i.e., from Eq. (A8):

$$\frac{1}{v_{gx,\,\text{appr}}} \approx \frac{1}{c\,\sin(\theta)}\,(\overline{n_{\text{eff}}^2}), \tag{2}$$

where $(\overline{n_{\text{eff}}^2})$ is a particular weighted average of the square of the refractive index in the structure, an average that does not depend on wavelength but only on structural parameters and, to some extent, the incidence angle in vacuum θ. c is the vacuum speed of light. Because the spatial shift and the group delay of a stack are approximately proportional, all the results we derive here concerning the spatial shift as a function of the stack parameters also apply to the group delay. The device shown in Fig. 1 will exhibit both spatial dispersion and temporal dispersion.[9]

In Ref. 8 we discussed that different design algorithms are useful for different types of dispersion characteristic. Periodic stacks have a limited design freedom, and all periodic stacks have quite similar nonconstant dispersion characteristics. Chirped stacks are useful for broad wavelength range designs and can be designed to have constant dispersion. Coupled-cavity stacks are useful for narrow wavelength range designs and can also be designed with constant dispersion. Numerically optimized stacks offer the greatest design freedom and can be designed to have constant dispersion over any wavelength range. Since we are interested in investigating a wide variety of designs, we generate all designs using numerical optimization even though the design process

is slower compared with the analytic design approaches.

To apply numerical optimization, we must first generate a starting design. We use impedance-matched Bragg stacks as the starting designs.[4,8] This starting design has unity reflectance and negligible dispersion in the design region. Reflections off the front of the stack are suppressed by the impedance matching. This starting design is then improved by numerical optimization. The performance of the design is measured by a merit function MF—a single number comparing the current design characteristics with the desired design characteristics.[12] The definition of the merit function we use is given as

$$\mathrm{MF} = \left(\frac{1}{N} \sum_{i=1}^{N} \left| \frac{Q_i^T - Q_i}{\Delta Q_i} \right|^p \right)^{1/p}, \tag{3}$$

where Q_i is the current value of the quantity of interest (for example, the shift at a specific wavelength); Q_i^T is the target value of that quantity; ΔQ_i is the acceptable deviation; N is the number of sampling points; and p is an integer number, usually called the p norm. In the case of $p = 2$, the merit function is, e.g., the root-mean-square difference between the current values and the target values of the quantities of interest. The goal of numerical optimization is to minimize the merit function MF and thus optimize the performance of the design. Different values of p lead to somewhat different optimized designs with different characteristics; for example, using larger values of p can reduce the ripple in the shift in the resulting designs, though it can lead to designs with the shift somewhat offset from the desired value.

We implemented five numerical optimization procedures—the Golden Section search optimization technique and the secant method, as well as the conjugate gradient algorithm, the Broyden–Fletcher–Goldfarb–Shanno Algorithm, and the Hooke and Jeeves pattern search technique.[8,13–15] The first two are one-dimensional search algorithms; for the last three methods, techniques are used that vary all the design parameters simultaneously. With all techniques, merit functions are used to measure the performance of a design and achieve performance improvements. Furthermore, all techniques search for the local minimum of the merit function. A lower nonlocal optimum can be found, however, by, e.g., choosing a large interval size in the Golden Section search optimization technique or by taking large steps in the Hooke and Jeeves pattern search technique. For a comparison of the performance of the different techniques, see Refs. 10 and 13. During the refinement process with our algorithm, we use the various methods in the sequence given in Table 1. The algorithm is performed either one or two times for a given set of parameters using a p norm of six in the merit function. This sequential use of methods is successful, as a different design methods may find a lower minimum if one method is stuck in a shallow local optimum.

Table 1. Sequence of Algorithms for Numerical Optimization

Refinement Algorithm	Iterations
Hooke and Jeeves pattern search	2
Conjugate gradient algorithm	2
Secant method	1
Hooke and Jeeves pattern search	2
Golden Section search	1
Broyden–Fletcher–Goldfarb–Shanno	2

Figure 2 shows an example of a generated 40-layer SiO_2–Ta_2O_5 stack for an incidence angle of 40° and a dispersion of 0.57 μm/nm. We do not list layer sequences here, since many different designs were generated (see Ref. 8 for a more detailed discussion of specific designed stacks). We can see that the specified dispersion, i.e. the specified slope, is obtained only over a limited wavelength range and that the shift is offset in absolute value from the specified shift. The p norm of six results in an equalization of the error, since larger errors dominate the merit function. This forces a constant dispersion even though it may be offset from the specified value. Since we are interested only in the relative shift with wavelength, the offset is irrelevant.

To compare different designs, we implemented an algorithm for finding the wavelength range over which the specified dispersion is achieved. For the example in Fig. 2, the specified shift is obtained over a wavelength range of 32.7 nm. The design in Fig. 2 also exhibits dispersion for shorter wavelengths. This dispersion is not relevant for our analysis, as we want to estimate over which wavelength range we can achieve a specified dispersion value. To obtain the maximum wavelength range, we specify the dispersion to be constant over a larger wavelength range

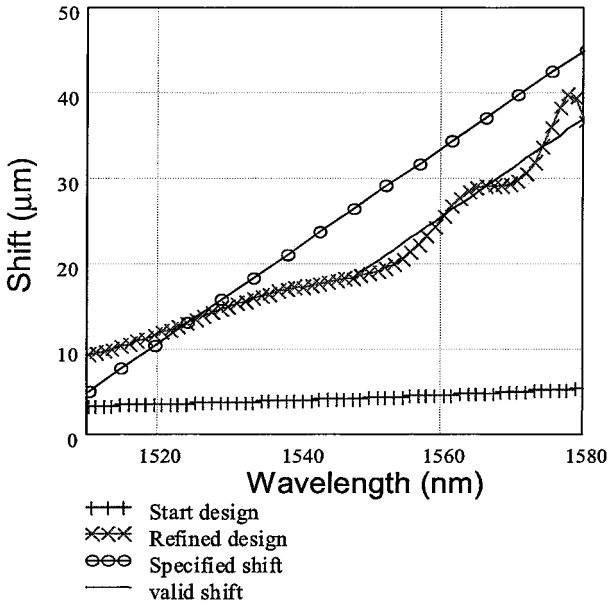

Fig. 2. Forty-layer SiO_2-Ta_2O_5 design for a 40° incidence angle. The dispersion was specified as 0.571 μm/nm.

Table 2. Distribution of the Design Parameters for the 623 Analyzed Designs

Parameter	Minimum	Maximum	Mean	Standard Deviation
Layer number	40	200	97	51
Physical thickness L (μm)	5.7	60	22	11
Refractive index	0.9	5.6	2.3	0.88
Average index n_{avg}	1.1	4.0	2.0	0.6
Index contrast Δn	0.06	4.0	1.0	0.8
Incidence angle θ (deg)	20	60	40	7.3
Dispersion c_{disp} (μm/nm)	0.04	86	3.2	6.7
Wavelength range Δλ (nm)	0.3	160	32	24
Total shift Δs (μm)	0.13	730	59	100

than the one we actually expect to be possible. Once we determined an initial empirical model, we used it to estimate the possible wavelength range and specify a constant dispersion over double the expected wavelength range for all the following designs.

To measure the quality of the design, we calculate the standard deviation of the design compared to the line with the specified dispersion over the valid wavelength range. For better comparison of designs, we use the relative standard deviation, i.e., the standard deviation divided by the total shift in the valid wavelength range. The relative standard deviation for the design in Fig. 2 is 8.5%. The designs generated here have a rather large relative standard deviation since we wanted to generate as many acceptable designs as possible in the given amount of computation time.

3. Empirical Model of the Maximum Shift

Using the automated design process described in Section 2, we generated a total of 760 different designs including 190 designs with 40 layers, 49 with 60 layers, 379 with 100 layers, 22 with 120 layers, and 120 with 200 layers. Each design had a different set of design and material parameters. The refractive indices, the incidence angle, the number of layers, and the specified dispersion characteristics were varied. The pure computation time for these designs was approximately 1100 h or 45 days with all algorithms implemented as MathCad programs and using a Pentium III 750-MHz computer with 256-Mbytes RAM. In this section we present the results derived from the analysis of these designs.

We filtered out 137 designs as they had a relative standard deviation of more than 25%, leaving 623 results to be analyzed. The distribution of the stack and operation parameters of the remaining designs is given in Table 2. Note that we did not limit our analysis to existing material systems, but also used fictional refractive-index combinations assuming that the refractive indices are constant over the wavelength range of interest and that no absorption occurs.

For the acceptable 623 designs, we evaluated the relationship among the specified constant dispersion value c_{disp}, the wavelength range Δλ over which the dispersion is achieved, and the stack and operation parameters. From this analysis we find that the maximum achievable wavelength range $(\Delta\lambda)_{max}$ can be estimated by

$$(\Delta\lambda)_{max} = \frac{1}{c_{disp}}\, 16\, L\, \frac{\Delta n}{n_{avg}^2} \sin(\theta), \tag{4}$$

where Δn is the refractive-index difference between the two stack materials, n_{avg} is the arithmetic average of the refractive indices, L is the total stack thickness, and θ is the incidence angle in vacuum. We can see from this empirical model that the wavelength range is inversely proportional to the dispersion. Thus, for linear shift with wavelength, the predicted total shift Δs after one bounce off the stack is not a function of the dispersion, but depends only on the stack parameters as shown in Eq. (5):

$$\Delta s = c_{disp}(\Delta\lambda)_{max} = 16\, L\, \frac{\Delta n}{n_{avg}^2} \sin(\theta). \tag{5}$$

To demonstrate how we derived this empirical model, let us consider the influence of the average refractive index on the total shift. Figure 3(a) plots the shift as a function of the average refractive index for all the designs. We can clearly see that the shift decreases with increasing average refractive index. If the model in Eq. (5) is correct, dividing the shift by the model should result in a normalized shift independent of the average refractive index. Figure 3(b) plots the shift divided by the model. Indeed no clear dependency is visible for a changing average refractive index. Since this is a numerical method, not all values are exactly unity, but there is some distribution of the values.

After looking at Fig. 3(b) the reader may not be fully convinced that Eq. (5) is indeed a good model for the shift. As a comparison, Fig. 4 plots the normalized shift for two wrong models. Figure 4(a) assumes the shift to be inversely proportional to the refractive index, and Fig. 4(b) assumes the shift to be inversely proportional to the refractive index cubed instead of the correct inverse proportionality to the refractive index squared.

In Fig. 4(a) a decrease in the shift as a function of the average refractive index is clearly visible. In Fig. 4(b) an increase of the shift with the average refractive index can be seen. Thus a shift that is inversely proportional to the square of the average refractive index seems indeed to be the best model. As we are basically guessing a model here based on a numerical data set, there is room for error and improvement. Maybe, for example, the average refractive index is really inversely proportional to the refractive index to the power of 2.1. As this result is close to the guessed one, we cannot conclusively decide either way from

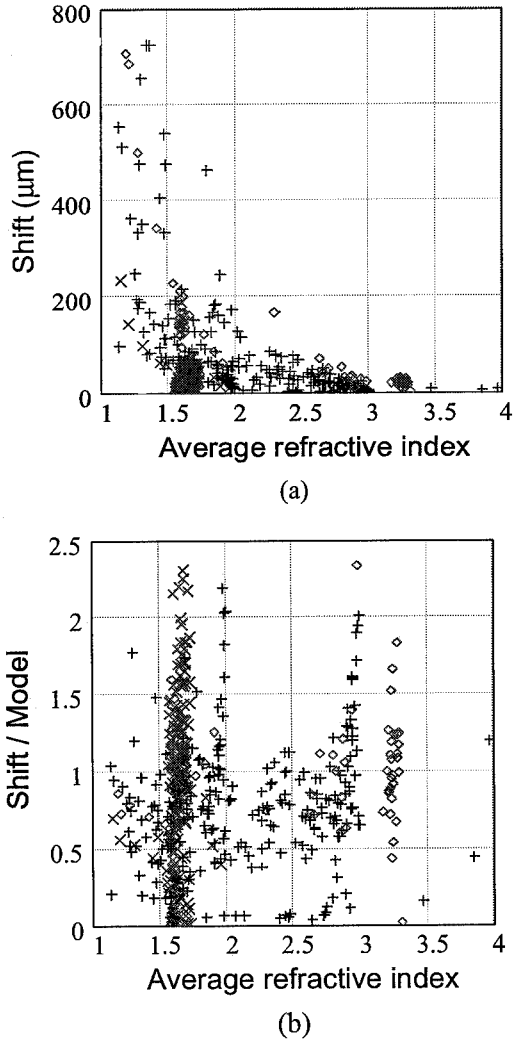

Fig. 3. Dependency of the shift on the average refractive index (crosses are for 40- and 60-layer desings, plusses are for 100 and 120 layers, and diamonds are for 200 layers).

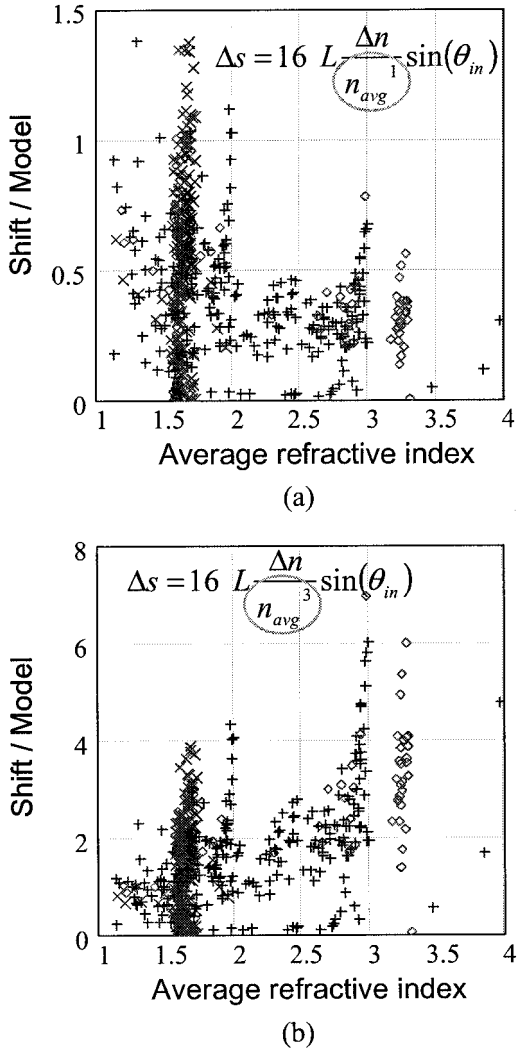

Fig. 4. Two poor models for comparison. (a) Assuming the shift to be inversely proportional to the average refractive index, (b) assuming the shift to be inversely proportional to the refractive index cubed.

the given data. An analytical model is needed to answer this question satisfactorily.

The shift model of Eq. (5) predicts that the product of dispersion c_{disp} and wavelength range $(\Delta\lambda)_{\max}$ is constant for a given set of stack parameters. Therefore, if we divide the wavelength range $\Delta\lambda$ of a designed stack by the theoretical shift Δs in Eq. (5), we should obtain $1/c_{\text{disp}}$. To demonstrate the performance of this model, Fig. 5 plots the value of $\Delta\lambda$ of the designed stacks as deduced directly from the numerical calculations on each stack [i.e., not from the empirical model, Eq. (4)] divided by the theoretical Δs [from Eq. (5)] as a function of the dispersion c_{disp}. The thick solid line represents the $1/c_{\text{disp}}$ behavior we would expect if our empirical model is correct, and all designs were as good as the approximate maximum wavelength range $(\Delta\lambda)_{\max}$ that our model predicts [Eq. (4)]. Designs below the thick solid line are worse than the model, i.e., they have a smaller than predicted wavelength range, and designs above the line are better than the model. For comparison the performance of the eight designs with constant dispersion discussed in Ref. 8 is also shown.

Figure 5 clearly shows that there seems to be a physical limit to the wavelength range over which a specific constant dispersion can be achieved. It cannot be pure chance that we did not find any design with a significantly higher wavelength range than predicted by Eqs. (4) and (5). Designs can always be worse, since the automated design process may have failed. The high number of designs close to the maximum theoretical line shows, however, that our design procedure works quite well. Therefore the result of this numerical investigation is that, for a linear shift with wavelength, the product of dispersion and wavelength range has an approximately constant maximum value that depends only on the stack parameters and not on the specified dispersion of the stack.

We can also see from Fig. 5 that a stack with any constant dispersion value between 0.04 and

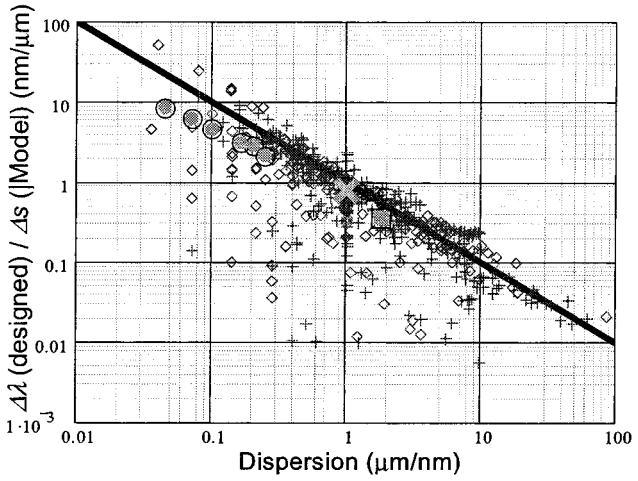

Fig. 5. Wavelength range divided by the shift model of Eq. (5) as a function of the specified dispersion. The thick solid line represents the expected value from the model of Eq. (5). Designs below the line are worse than the model and designs above the line are better than the model. The diamonds represent the fraction of the 623 designed stacks with less than 10% relative standard deviation and the plusses represent the stacks with 10–25% relative standard deviation. The large circles are the chirped designs discussed in Ref. 8, the X is the numerically optimized design in Ref. 8, and the square is the coupled-cavity design in Ref. 8.

86 μm/nm can be designed using the algorithm given in Section 2. A stack with low dispersion can be operated over a broad wavelength range, whereas a stack with high dispersion shows this dispersion only over a narrow wavelength range. Since the number of channels that can be demultiplexed with the device shown in Fig. 1 is determined by the total shift divided by the spatial extent per channel, both the broad and the narrow wavelength range designs will allow for the same number of channels for a given set of stack parameters. The difference between the designs is that one has widely spaced channels whereas the other has a close channel spacing.

It is interesting to note that the chirped stacks, the coupled-cavity stack, and the numerically optimized stack discussed in Ref. 8 all exhibit a similar performance. They all lie slightly below the model. This may be due to a trade-off between the ripple in the shift and the total shift. As the designs in this paper are generated automatically, they may not be as flat as the designs discussed in Ref. 8. Intuitively, it seems that a very flat design uses part of the available design degrees of freedom to achieve this flatness. Thus those degrees of freedom are not available to obtain a larger total shift. Furthermore, we did not check the transmission loss of the automatically designed stacks under the assumption that the transmission loss can be eliminated by specifying both shift and reflectance values in the merit function or by adding dedicated mirror layers to the stack that are not modified during refinement. Again, guaranteeing unity reflectance consumes some of the degrees of freedom that could otherwise have been used for a larger total shift. Therefore we do expect that there should be a trade-off among total shift, ripple, and transmission loss that has not been analyzed here. Once the acceptable ripple and transmission loss have been defined, we expect the total shift to obey a relationship similar to the empirical model of Eq. (5).

As discussed in Section 2 and in Appendix A, the spatial shift and the group delay are approximately proportional in the case of structural dispersion. Therefore the same argument holds for the wavelength range over which a given constant temporal dispersion can be achieved. Substituting Eq. (5) into Eq. (1) we obtain a model for predicting the total group-delay difference for a given set of stack parameters. For the case of normal incidence we take into account that the $\sin(\theta)$ term appears in both the expression for the shift Δs and the group velocity v_{gx} and thus cancels. Explicitly, from Eq. (1) and substituting from approximation (2), the maximum variation $\Delta\tau$ in group delay on reflection that could be induced approximately linearly with wavelength is

$$\Delta\tau = \frac{\Delta s}{v_{gx}} \approx 16\, L\, \frac{\Delta n}{n_{\mathrm{avg}}^2} \sin(\theta)\, \frac{1}{c\, \sin(\theta)}\, \overline{(n_{\mathrm{eff}}^2)} \approx 16\, \Delta n\, \frac{L}{c}, \tag{6}$$

where in the final step we made the additional simplification of assuming that n_{avg}^2 and $\overline{(n_{\mathrm{eff}}^2)}$ are of comparable size (the assumption is good for structures with small to moderate index contrast). This result suggests a simple limit on the amount of linear group-delay compensation that could be engineered in a multilayer dielectric structure, with the limit essentially depending linearly only on the magnitude of the refractive-index variation available and the length of the structure. Again, we expect a trade-off among the total group-delay variation, ripple, and transmission loss. Causes of group-delay ripple in chirped mirrors have been discussed in Ref. 16.

4. Maximum Number of Channels of the Spatial Demultiplexer

We now discuss one specific example of the usefulness of this empirical model. The model allows us to estimate the number of channels that can be demultiplexed with the device shown in Fig. 1. In Ref. 9 we found that for a device with a single focusing lens the maximum number of bounces off the stack is limited by beam broadening and that the maximum number of channels N_{channels} is given by

$$N_{\mathrm{channels}} \leq 1 + \frac{\Delta s}{\lambda_c}\, \frac{\pi \sin(\theta)[1 - \sin(\theta)^2/n_s^2]}{2c_{\mathrm{cross\ talk}}}, \tag{7}$$

where θ is the incidence angle in vacuum, λ_c is the center wavelength of the device, and n_s is the refractive index of the substrate. $c_{\mathrm{cross\ talk}}$ determines the cross talk between adjacent channels. For $c_{\mathrm{cross\ talk}} = 2$, the adjacent channel cross talk is approximately -15 dB, for $c_{\mathrm{cross\ talk}} = 3.2$ it is -30 dB, and for

$c_{\text{cross talk}} = 3.8$ it is -40 dB. Inequality (7) shows that the maximum number of channels is proportional to the maximum shift Δs after one bounce off the stack. Substituting the empirical model of Eq. (5) for Δs, we obtain

$$N_{\text{channels}} \leq 1 + \frac{L}{\lambda_c} \frac{8}{c_{\text{cross talk}}} \frac{\Delta n}{n_{\text{avg}}^2} \pi \sin(\theta)^2 \left[1 - \frac{\sin(\theta)^2}{n_s^2} \right]. \tag{8}$$

It is interesting to note that for a given set of materials, incidence angle, and adjacent channel cross talk, everything on the right-hand side of inequality (8) is fixed except for the ratio of the stack thickness L to the center wavelength λ_c. From inequality (8) we conclude that, given the choice of a material system, we should choose materials with a large index contrast and a small average refractive index. In this model only the stack thickness and not the number of layers seems to matter. That is somewhat misleading since the layer thickness should likely always be of the order of a quarter wavelength. Therefore fixing a stack thickness also fixes the number of layers needed. It is interesting to note that the number of channels scales as the thickness divided by the wavelength. Thus scaling a design to a different wavelength in the same material system does not change the number of channels, just as we would expect.

For a cross talk of -40 dB, a 45° incidence angle, and alternating SiO_2 ($n_L = 1.45$) and Ta_2O_5 ($n_H = 2.09$) layers on a quartz substrate ($n_s = 1.52$), the number of channels can be estimated by

$$N_{\text{SiO}_2\text{-Ta}_2\text{O}_5,\ 45^\circ} \leq 1 + 0.53 \frac{L}{\lambda_c}. \tag{9}$$

We calculate from inequality (9) that a 21-μm stack is necessary to demultiplex eight channels around 1550 nm using the device in Fig. 1. Such a stack can be fabricated with today's thin-film deposition technology allowing for a compact, cost-effective wavelength multiplexing device that requires only a single multilayer stack.

5. Conclusions

We have generated and analyzed a set of 623 different designs with constant dispersion to derive an empirical model relating the stack and operation parameters to the performance. From this analysis we concluded that the wavelength range over which a specified constant dispersion is achieved is inversely proportional to the value of the dispersion for given stack parameters. Furthermore, we have presented numerical evidence that the product of dispersion and wavelength range, i.e., the total spatial shift or the total group-delay difference, has a maximum value that is proportional to the stack thickness (number of layers) and the refractive-index difference and is inversely proportional to the square of the average refractive index. We have also proposed a simple semiempirical model for the maximum linear variation in group delay possible in layered dielectric structures.

The empirical model developed here allows us to quickly estimate the possible performance of a design with spatial or temporal dispersion. We can use it to choose a material system and the necessary number of layers for a given design task, as well as to compare the performance of a designed stack with the theoretical maximum performance. This research is a first step in the search for an analytic model relating the limits to performance to the general parameters of the materials and size of the stack. The fact that all 623 designs approximately fall near or below our simple empirical model is strongly suggestive that an analytical limit model exists and suggests that it could be productive to look for such a model. In the development of an analytic model, the results can be compared with the numerical data presented here. Again, given the variety and large number of structures investigated numerically, we expect the empirical model to be valid for all types of one-dimensional photonic nanostructure. This work is also suggestive that similar empirical or analytic relationships will exist for the dispersion in two- and three-dimensional periodic photonic crystals[17] and nonperiodic photonic nanostructures more generally.

Appendix A: Relating Spatial and Temporal Dispersion

In this appendix we explore the relationship between spatial and temporal dispersion characteristics. Spatial dispersion manifests itself in a change of the beam exit position s_x as a function of wavelength as shown in Fig. 1. Temporal dispersion refers to the change in the group delay τ_{group} as a function of wavelength. We are considering here the shift and the group delay in reflection. A similar study can be performed for the case of transmission. The group delay in reflection is calculated by dividing the shift s_x along the x direction by the effective group velocity v_{gx} in this direction or equivalently by dividing twice the thickness L of the stack by the group velocity v_{gz} in the z direction. As shown in Eq. (A1) we can rewrite this expression in terms of the more commonly used dependency on the phase upon reflection ϕ_{refl} using Eq. (A2):

$$\tau_{\text{group}} = \frac{s_x}{v_{gx}} = s_x \left. \frac{\partial \beta}{\partial \omega} \right|_{K=\text{const}} = \frac{2L}{v_{gz}} = 2L \left. \frac{\partial K}{\partial \omega} \right|_{\beta=\text{const}}$$

$$= \left. \frac{\partial \phi_{\text{refl}}}{\partial \omega} \right|_{\beta=\text{const}}, \tag{A1}$$

$$\phi_{\text{refl}} = 2LK. \tag{A2}$$

In the case of a periodic structure, the relationship among the wave vector K in the z direction, the wave vector β in the x direction, and the frequency ω can be obtained from Bloch theory.[8] For a nonperiodic structure we can use the transfer-matrix technique to re-

late K, β, and ω.[8] For a nonperiodic stack and a finite periodic stack, K is the effective wave vector for the reflected light. Since the structure is not periodic, K is not constant throughout the stack. The group velocities v_{gx} and v_{gz} are in this case also effective quantities, i.e., they represent the total effect of the stack, but are not constant within the stack. Therefore, by calculating the group delay τ_{group}, we obtain the total time elapsed from entering the stack to exiting the stack, but we cannot determine how much delay the light incurred in each part of the stack.

Equation (A1) shows that the shift experienced by a beam of light is related to the group delay in reflection by the group velocity v_{gx} along the layers. Here we are interested in the dispersive properties of multilayer stacks, i.e., the change of the shift and the group delay with wavelength. To relate spatial and temporal dispersion, we therefore need to investigate the change of the group velocity v_{gx} with wavelength. First, we calculate the group velocity in a WKB-type approximation.

The WKB approximation (also called the semiclassical or quasi-classical approximation in quantum mechanics) states that if the local wavelength $\lambda(z)$, which is linked to the local wave vector $K(z)$, changes slowly with z, the accumulated phase change can be calculated by integrating the wave vector $K(z)$ from the start position z_1 to the end position z_2.[18,19] This result is exactly true for uniform media as well as infinite periodic media, where the wave vector K is obtained from Bloch theory and is independent of z. The resulting phase is given by Eq. (A2). The WKB approximation has previously been applied to calculate the accumulated phase change of a chirped Bragg stack.[20] Here we use the WKB approximation to obtain an approximate phase upon reflection ϕ_{appr} for a general multilayer stack.

As the local wave vector $K(z)$ is not known *a priori* in this case, we use the wave vector corresponding to a uniform medium with refractive index n_i for the ith layer. Even though this is strictly only a good approximation for low-index contrast stacks, we have found empirically that it is quite good for many structures of interest. Replacing the WKB integral by a sum, the approximated phase upon reflection ϕ_{appr} is then given by approximation (A3), where d_i is the layer thickness of the ith layer in the stack:

$$\phi_{\text{appr}}(\beta,\ \omega) \approx -2 \sum_i \left\{ \left[\left(\frac{\omega}{c} n_i \right)^2 - \beta^2 \right]^{1/2} d_i \right\}. \quad \text{(A3)}$$

Using Eqs. (A2) and (A3) we therefore obtain the approximate dispersion relation of Eq. (A4) for the approximate wave vector K_{appr} in the z direction:

$$K_{\text{appr}}(\beta,\ \omega) = -\frac{1}{L} \sum_i \left\{ \left[\left(\frac{\omega}{c} n_i \right)^2 - \beta^2 \right]^{1/2} d_i \right\}. \quad \text{(A4)}$$

Again, this is an effective wave vector for the overall effect of the stack, not a local wave vector. The derivatives of the approximate dispersion relation with respect to ω and β are Eqs. (A5) and (A6). Here it is assumed that the refractive indices n_i are independent of frequency, which is quite correct for the typical dielectrics used in multilayer thin-film stacks. If the refractive indices do depend on frequency, the derivatives have to be modified accordingly as discussed in Ref. 10:

$$\left. \frac{\partial K_{\text{appr}}(\beta,\ \omega)}{\partial \omega} \right|_{\beta=\text{const}} = -\frac{1}{L} \sum_i \left\{ \frac{\frac{n_i^2}{c} d_i}{\left[n_i^2 - \left(\frac{\beta c}{\omega} \right)^2 \right]^{1/2}} \right\}, \quad \text{(A5)}$$

$$\left. \frac{\partial K_{\text{appr}}(\beta,\ \omega)}{\partial \beta} \right|_{\omega=\text{const}} = \frac{1}{L} \sum_i \left\{ \frac{\frac{\beta c}{\omega} d_i}{\left[n_i^2 - \left(\frac{\beta c}{\omega} \right)^2 \right]^{1/2}} \right\}. \quad \text{(A6)}$$

Using the rules for taking implicit derivatives,[21] we obtain Eq. (A7) for $v_{gx,\ \text{appr}}$:

$$\begin{aligned} \frac{1}{v_{gx,\ \text{appr}}} &= \left(\frac{\partial \beta}{\partial \omega} \right)_{\text{appr}} \Bigg|_{K_{\text{appr}}=\text{const}} \\ &= \frac{-\partial K_{\text{appr}} / \partial \omega}{\partial K_{\text{appr}} / \partial \beta} \\ &\approx \frac{\omega}{\beta c^2} \frac{\sum_i \left\{ n_i^2 d_i \Big/ \left[n_i^2 - \left(\frac{\beta c}{\omega} \right)^2 \right]^{1/2} \right\}}{\sum_i \left\{ d_i \Big/ \left[n_i^2 - \left(\frac{\beta c}{\omega} \right)^2 \right]^{1/2} \right\}}. \end{aligned} \quad \text{(A7)}$$

Substituting $\beta = \omega/c \sin(\theta)$, where θ is the incidence angle in vacuum, Eq. (A7) is rewritten as

$$\frac{1}{v_{gx,\ \text{appr}}} = \left(\frac{\partial \beta}{\partial \omega} \right)_{\text{appr}} \Bigg|_{K=\text{const}} = \frac{1}{c \sin(\theta)} (\overline{n_{\text{eff}}^2}), \quad \text{(A8)}$$

where

$$(\overline{n_{\text{eff}}^2}) = \frac{\sum_i \{ n_i^2 d_i / [n_i^2 - \sin(\theta)^2]^{1/2} \}}{\sum_i \{ d_i / [n_i^2 - \sin(\theta)^2]^{1/2} \}}. \quad \text{(A9)}$$

We can see that the resulting approximate expression for $v_{gx,\ \text{appr}}$ depends only on the incidence angle and not on the frequency. Thus $v_{gx,\ \text{appr}}$ is independent of frequency within this approximation. We found numerically that $v_{gx,\ \text{appr}}$ is approximately constant for many of the structures that we evaluated, even though the constant value might be different from the one calculated with Eq. (A8).

References

1. H. A. MacLeod, *Thin-Film Optical Filters* (Institute of Physics, Philadelphia, Pa., 2001).

2. R. Szipöcs, K. Ferencz, C. Spielmann, and F. Krausz, "Chirped multilayer coatings for broadband dispersion control in femtosecond lasers," Opt. Lett. **19,** 201–203 (1994).
3. P. Tournois and P. Hartemann, "Bulk chirped Bragg reflectors for light pulse compression and expansion," Opt. Commun. **119,** 569–575 (1995).
4. N. Matuschek, F. X. Kärtner, and U. Keller, "Analytical design of double-chirped mirrors with custom-tailored dispersion characteristics," IEEE J. Quantum Electron. **35,** 129–137 (1999).
5. G. Lenz and C. K. Madsen, "General optical all-pass filter structures for dispersion control in WDM systems," J. Lightwave Technol. **17,** 1248–1254 (1999).
6. C. K. Madsen and J. H. Zhao, *Optical Filter Design and Analysis—A Signal Processing Approach* (Wiley, New York, 1999).
7. M. Jablonski, Y. Takushima, and K. Kikuchi, "The realization of all-pass filters for third-order dispersion compensation in ultrafast optical fiber transmission systems," J. Lightwave Technol. **19,** 1194–1205 (2001).
8. M. Gerken and D. A. B. Miller, "Multilayer thin-film structures with high spatial dispersion," Appl. Opt. **42,** 1330–1345 (2003).
9. M. Gerken and D. A. B. Miller, "Wavelength demultiplexer using the spatial dispersion of multilayer thin-film structures," IEEE Photon. Technol. Lett. **15,** 1097–1099 (2003).
10. M. Gerken, "Wavelength multiplexing by spatial beam shifting in multilayer thin-film structures," Electrical Engineering Ph.D. dissertation (Stanford University, Stanford, Calif., 2003).
11. M. Gerken and D. A. B. Miller, "The relationship between the superprism effect, group delay, and stored energy in 1-D photonic crystals and photonic nanostructures," in *MRS Spring Meeting* (Materials Research Society, Warrendale, Pa., 2003), paper J2.7.
12. J. A. Dobrowolski, F. C. Ho, A. Belkind, and V. A. Koss, "Merit functions for more effective thin film calculations," Appl. Opt. **28,** 2824–2831 (1989).
13. J. A. Dobrowolski and R. A. Kemp, "Refinement of optical multilayer systems with different optimization procedures," Appl. Opt. **29,** 2876–2893 (1990) and references herein.
14. E. K. P. Chong and S. H. Zak, *An Introduction to Optimization* (Wiley, New York, 1996).
15. T. E. Shoup and F. Mistree, *Optimization Methods with Applications for Personal Computers* (Prentice-Hall, Upper Saddle River, N.J., 1987).
16. G. Steimeyer, "Dispersion oscillations in ultrafast phase-correction devices," IEEE J. Quantum Electron. **39,** 1027–1034 (2003).
17. T. Baba and T. Matsumoto, "Resolution of photonic crystal superprism," Appl. Phys. Lett. **81,** 2325–2327 (2002).
18. R. Shankar, *Principles of Quantum Mechanics* (Plenum, New York, 1994).
19. A. B. Migdal, *Qualitative Methods in Quantum Theory* (Addison-Wesley, Redwood City, Calif., 1989).
20. N. Matuschek, F. X. Kärtner, and U. Keller, "Theory of double-chirped mirrors," IEEE J. Sel. Top. Quantum Electron. **4,** 197–208 (1998).
21. I. N. Bronstein, K. A. Semendjajew, G. Musiol, and H. Muehlig, *Taschenbuch der Mathematik* (Verlag Harri Deutsch, Frankfurt am Main, Germany, 1993), pp. 232–237.