Designing ultra-efficient Multilayer Coatings using Nested Sampling

UROP Project (25/07/2022–25/09/2022)
supervised by David Buscher (Cavendish Laboratory, University of Cambridge)
Namu Kroupa

Abstract

Treating the multilayer thin-film coating design problem as the global minimisation of a merit function, the code Coating was developed, which chooses optimal materials and thicknesses for the layers and finds subsidiary optimal designs without an initial guess.
Coating repurposes the nested sampling code PolyChord as a global optimiser by supplying a tailored likelihood function and a uniform prior over the parameter space of the layer materials and thicknesses such that PolyChord compresses the parameter space to the global minimum of the merit function. The posterior samples generated during the run are clustered using a distance metric which measures the distance between refractive index functions and the vector of layer thicknesses. These are used to find subsidiary local minima, which can have other desirable non-quantifiable properties.
Tests were performed on 21- and 40-layer coatings, being able to reproduce and improve previously published multilayer coating designs. For an 8-layer coating, the calculated design had a 39% lower range of mean reflectivity and a 4% larger thickness. Finally, a 21-layer antireflection coating, which was designed with fixed materials, was reduced to a 10-layer coating, designed with freely chosen materials, incurring an average reflectivity increase from 7% to 14%.
Coating emphasises user-friendliness and is parallelised with MPI. It is available for download from `https://github.com/zwei-beiner/Code`.

1 Introduction

The design of multilayer coatings finds application in numerous technological fields [1], such as state-of-the-art astronomical instruments, in which it is important to conserve the intensity of the incoming starlight as much as possible. For this, highly efficient optical filters are required which follow a target reflectivity as closely as possible.

When designing a multilayer coating, the runtime of a brute-force search for the optimal coating scales exponentially with the number of layers, rendering it intractable for practical applications. Other techniques for the synthesis of a multilayer coating rely on an initial guesses and have to make uphill steps to escape local minima [1].

In this paper, the global optimiser PolyChord [2] is used, which evolves a set of samples towards the global optimum, compressing the parameter space exponentially at each step and removing the need for an initial guess. The exploration of the entire parameter space makes it possible to locate local minima.

The next section summarises the theoretical background. Section 3 presents the algorithms and methods and Section 4 gives results and discussions. Section 5 contains the conclusions.

2 Theoretical Background

2.1 The multilayer coating

A multilayer coating consists of M layers, numbered from $j = 1$ to $j = M$. The light enters the coating from the outer medium and exits into the substrate (Figure 1) such that the backward and forward travelling electromagnetic waves have electric field amplitudes a_j and b_j, respectively [3].

The coating is completely specified by the refractive index, n_j, which is a function of wavelength λ, and thickness, d_j, of each layer.

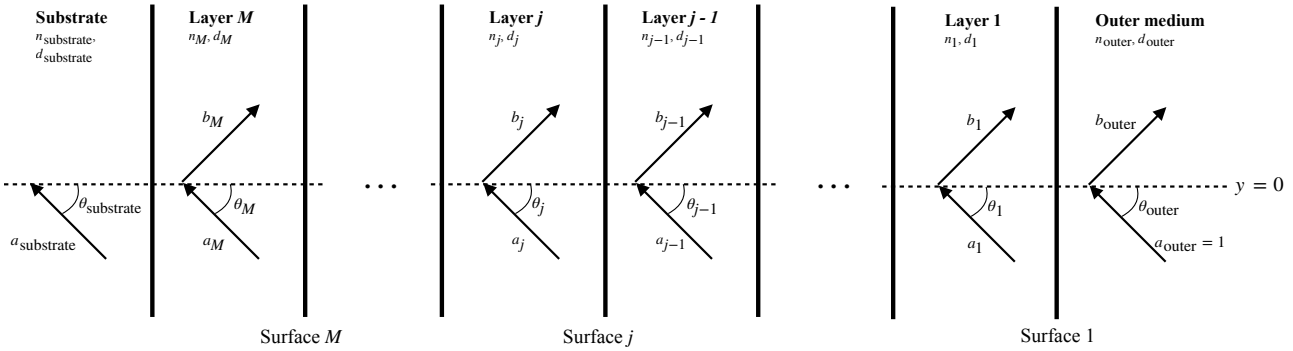

Figure 1: Electric field amplitudes of forward, a_j, and backward, b_j, travelling waves in a multilayer thin-film coating in which each layer has refractive index n_j and thickness d_j.

2.2 Calculation of the reflected amplitude

For s-polarised light, the interface conditions for the electric and magnetic fields at each layer interface result in a set of $2M + 2$ coupled equations [3]. Setting the incident amplitude to $a_{\text{outer},\,s} = 1$ and assuming that the reflected amplitude in the layer is negligible, $b_{\text{substrate},\,s} = 0$, the system of equations is cast into a matrix equation [3],

$$\mathbf{M}\mathbf{x} = \mathbf{c}, \tag{1}$$

which must be solved for $\mathbf{x}$, where

$$\mathbf{x} = (b_{\text{outer},\,s}, a_{1,\,s}, b_{1,\,s}, a_{2,\,s}, b_{2,\,s}, \ldots, a_{M,\,s}, b_{M,\,s}, a_{\text{substrate},\,s})^T \tag{2}$$

and

$$\mathbf{c} = (1, 1, 0, 0, \ldots, 0)^T \tag{3}$$

are $(2M + 2)$-dimensional complex vectors. The matrix $\mathbf{M}$ is described in Appendix A. A similar set of equations holds for p-polarised light.

Using a matrix solver for band-diagonal matrices, the time complexity to solve Equation 1 is $O(M)$. After solving for $\mathbf{x}$, the reflectivity for s- and p- polarised light is calculated by $R_s = |b_{\text{outer},\,s}|^2$ and $R_p = |b_{\text{outer},\,p}|^2$, respectively.

2.3 The merit function

The parameters which specify a multilayer coating are collectively denoted by

$$\boldsymbol{p}_{\text{total}} = (n_1, n_2, \ldots, n_M, d_1, d_2, \ldots, d_M). \tag{4}$$

In practice, some of these variables are predetermined (for example the choices for the refractive indices) so that certain entries in $\boldsymbol{p}_{\text{total}}$ are fixed during optimisation. The parameters which are unfixed during optimisation are denoted by

$$\boldsymbol{p} = (n_1, n_2, \ldots, n_p, d_1, d_2, \ldots, d_q). \tag{5}$$

The refractive index, n_j, of layer j can be chosen from a set, $\mathcal{N}_j$. Different layers can have different materials to choose from. The choice of the thicknesses is from a potentially layer-dependent range $[d_{\min}, d_{\max}]$, which is usually determined by manufacturing constraints. The dimensionality of the optimisation problem is thus $N_{\text{dim}} = p + q$.

The parameters $\boldsymbol{p}$ must be chosen such that the output amplitude of the multilayer coating follows a specified target reflectivity for s-polarisation R_s, p-polarisation R_p, the reflectivity sum S and difference D and the phase difference ϕ_{sp}, defined in Table 1. This amounts to the minimisation of the merit function,

$$\begin{aligned} f(\boldsymbol{p}) =& w_{R_s} \left\langle \left(\frac{R_s(\boldsymbol{p},\lambda) - \tilde{R}_s(\lambda)}{\delta R_s(\lambda)} \right)^2 \right\rangle_\lambda + w_{R_p} \left\langle \left(\frac{R_p(\boldsymbol{p},\lambda) - \tilde{R}_p(\lambda)}{\delta R_p(\lambda)} \right)^2 \right\rangle_\lambda \\ &+ w_S \left\langle \left(\frac{S(\boldsymbol{p},\lambda) - \tilde{S}(\lambda)}{\delta S(\lambda)} \right)^2 \right\rangle_\lambda + w_D \left\langle \left(\frac{D(\boldsymbol{p},\lambda) - \tilde{D}(\lambda)}{\delta D(\lambda)} \right)^2 \right\rangle_\lambda \\ &+ w_{\phi_{sp}} \left\langle \left(\frac{\phi_{sp}(\boldsymbol{p},\lambda) - \tilde{\phi}_{sp}(\lambda)}{\delta \phi_{sp}(\lambda)} \right)^2 \right\rangle_\lambda , \end{aligned} \tag{6}$$

where the targets are denoted by a tilde, $\tilde{\cdot}$, and the average over a set of wavelengths $\{\lambda_1, \lambda_2, \ldots, \lambda_K\}$ is defined as $\langle \cdot \rangle_\lambda = \frac{1}{K} \sum_{k=1}^{K} (\cdot)$, providing a measure of fit independent of the number of wavelengths, K. The procedure adopted to create the set of wavelengths from an interval $[\lambda_{\min}, \lambda_{\max}]$ is described in Section 2.4. The weights, denoted by a "w", allow weighting of the individual terms whereas the tolerances, denoted by a "δ", allow different weighting for different wavelengths within a term. A summary of the variables is given in Table 1.

Table 1: Definition of the symbols which appear in the merit function (Equation 6).

Term in the merit function	Definition	Target	Weight	Tolerance
Reflectivity of s-polarised light	$R_s = \lvert b_{\text{outer,s}} \rvert^2$	$\tilde{R}_s$	w_{R_s}	δR_s
Reflectivity of p-polarised light	$R_p = \lvert b_{\text{outer,p}} \rvert^2$	$\tilde{R}_p$	w_{R_p}	δR_p
Sum of reflectivities	$S = R_s + R_p$	$\tilde{S}$	w_S	δS
Difference of reflectivities	$D = \lvert R_s - R_p \rvert$	$\tilde{D}$	w_D	δD
Phase difference	$\phi_{sp} = \arg(b_{\text{outer,s}}) - \arg(b_{\text{outer,p}})$	$\tilde{\phi}_{sp}$	$w_{\phi_{sp}}$	$\delta \phi_{sp}$

2.4 Calculation of the wavelengths $\{\lambda_k\}$

The wavelength spacing, $\delta\lambda$, across which the reflectivity varies through a quarter of an oscillation was estimated in [3] as

$$\delta\lambda = \frac{\lambda^2}{8 \sum_{j=1}^{M} d_j}. \tag{7}$$

The set of wavelengths, $\{\lambda_k\}$, required for the merit function (Equation 6) was calculated using Equation 7.

Assuming $\delta\lambda \ll \lambda$, the corresponding frequencies are uniformly spaced so that the number of frequencies in the interval $[\lambda_{\min}, \lambda_{\max}]$ is

$$K_{\text{freq}} = 8 \left(\frac{1}{\lambda_{\min}} - \frac{1}{\lambda_{\max}} \right) \sum_{j=1}^{M} d_j, \tag{8}$$

By setting d_j to the maximum thickness on all layers with unfixed thicknesses, the maximum number of required wavelengths is calculated, $K_{\text{freq, max}}$, which was used to pre-compute all possible sets of wavelengths for computational efficiency.

3 Method

3.1 PolyChord for global optimisation

PolyChord [2] is an implementation of the nested sampling algorithm [4], which was originally devised in the context of Bayesian inference to calculate the evidence of a probabilistic model,

$$Z = \int L(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}, \tag{9}$$

and sample the posterior probability distribution,

$$P(\boldsymbol{\theta}) = \frac{L(\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{Z}, \tag{10}$$

where $\boldsymbol{\theta}$ are the model parameters, $L(\boldsymbol{\theta})$ is the likelihood function, $\pi(\boldsymbol{\theta})$ is the prior probability distribution and the integral extends over the entire parameter space.

The nested sampling algorithm is initialised by sampling N_{live} live points from the prior. In each iteration, the live point with the smallest value of the likelihood, L^*, is deleted from the set of live points and its parameter values are saved. This is called a dead point. A new live point is sampled from the prior with the constraint that its likelihood must be larger than L^*. The next iteration repeats with the new set of live points. It can be shown [2] that, at iteration N_{dead}, at which N_{dead} dead points are saved, the constrained parameter space, $\{\boldsymbol{\theta} \mid L(\boldsymbol{\theta}) \geq L^*\}$, contains the global maximum of the likelihood and has an exponentially smaller enclosed prior mass, $\int \pi(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \approx \exp(-N_{\text{dead}}/N_{\text{live}})$.

From the set of live points at the final iteration, the live point with the largest likelihood was used as the initial guess of a local optimiser, the Nelder-Mead simplex algorithm [5], for refinement.

For the minimisation of the merit function, the parameter space is the set of possible values of $\boldsymbol{p}$. The likelihood is chosen to be

$$L(\boldsymbol{p}) = -f(\boldsymbol{p}), \tag{11}$$

which is maximised by PolyChord, and the prior distribution is chosen as

$$\pi(\boldsymbol{p}) = (\mathrm{Cat}_1, \mathrm{Cat}_2, \ldots, \mathrm{Cat}_p, \mathrm{Uni}_1, \mathrm{Uni}_2, \ldots, \mathrm{Uni}_q), \tag{12}$$

where the $\{\mathrm{Cat}_j\}$ are discrete uniform categorical distributions over the choice of each refractive index of each layer and the $\{\mathrm{Uni}_j\}$ are uniform continuous distributions over the interval of the allowed thickness of each layer.

Using ordinal encoding for each n_j, the map, $\boldsymbol{F}$, from the unit hypercube, $\boldsymbol{x} \in [0,1]^{N_{\text{dim}}}$, to the parameter space, required by PolyChord for sampling, therefore takes the form

$$F_j(x_j) = \begin{cases} \lfloor |\mathcal{N}_j|\, x_j \rfloor & \text{if} \quad x_j \neq 1 \\ C_j - 1 & \text{if} \quad x_j = 1 \end{cases} \qquad \text{for} \quad j = 1, \ldots, p \tag{13}$$

$$F_j(x_j) = d_{\min,\,j} + (d_{\max,\,j} - d_{\min,\,j})x_j \qquad \text{for} \quad j = 1, \ldots, q \tag{14}$$

where $|\mathcal{N}_j|$, $d_{\min,\,j}$ and $d_{\max,\,j}$ are the number of choices of refractive indices, and minimum and maximum thickness of layer j, and $\lfloor\cdot\rfloor$ is the floor function.

The settings with which PolyChord was called are shown in Table 2. The prefactor of 5 in the value of $N_{\text{dead, max}}$ was chosen empirically such that the likelihood values of the dead points show convergence. As the number of likelihood evaluations scales as $N_{\text{dead, max}}N_{\text{repeat}} \sim O(N_{\text{dim}}^3)$ and a likelihood evaluation takes $O(M)$, the total runtime scales as $O(N_{\text{dim}}^3 M)$.

Table 2: PolyChord settings and the values chosen for the optimisation of the merit function. The settings are explained in detail in [2].

Setting	Explanation	Chosen value
N_{live}	Number of live points	$25N_{\text{dim}}$
N_{repeat}	Number of slice sampling steps to generate a new live point	$5N_{\text{dim}}$
$N_{\text{dead, max}}$	Maximum number of dead points. Equivalent to the maximum number of iterations	$5N_{\text{live}}N_{\text{dim}}$

3.2 Finding local minima

Subsidiary local minima of the merit function were found by clustering (Algorithm 1) the set of live points every N_{live} PolyChord iterations and performing a local minimisation on every cluster (Algorithm 2), taking the live point with the smallest merit function value as an initial guess. The clustering algorithm HDBSCAN [6, 7] was used because it allows for the detection of an arbitrary number of clusters with different densities. For the application in this paper, finding more clusters than the true number of local minima is desired, as opposed to fewer clusters than local minima. Therefore, the minimum cluster size was chosen to be much smaller than N_{live} and set to 5.

For the calculation of the distance between two parameter vectors, the parameter vector $\boldsymbol{p}$ is first split into the refractive indices, $\boldsymbol{n} = (n_1, n_2, \ldots, n_p)$, and the vector of layer thicknesses, $\boldsymbol{d} = (d_1, d_2, \ldots, d_q)$, where $(\boldsymbol{n}, \boldsymbol{d}) = \boldsymbol{p}$. The vector $\boldsymbol{d}$ is preprocessed by scaling each component to the range $[0, 1]$, giving the vector $\hat{\boldsymbol{d}}$ with components

$$\hat{d}_j = \frac{d_j - \min\limits_{\text{live points}} d_j}{\max\limits_{\text{live points}} d_j - \min\limits_{\text{live points}} d_j} \quad \text{for} \quad j = 1, \ldots, q. \tag{15}$$

A distance metric was sought which measures the distance between the refractive indices and the thickness vectors and subsequently weights these calculated distances equally. Therefore, the distance between the refractive indices is normalised by the maximum distance for layer j, $\max_{x\in\mathcal{N}_j,\, y\in\mathcal{N}_j} \|x - y\|$. Subsequently, a linear combination of the distance metrics on $\boldsymbol{n}$ and $\hat{\boldsymbol{d}}$ is taken which weights the two parts of the parameter vector equally. This defines the distance metric between parameter vectors $\boldsymbol{p}_1 = (\boldsymbol{n}_1, \boldsymbol{d}_1)$ and $\boldsymbol{p}_2 = (\boldsymbol{n}_2, \boldsymbol{d}_2)$:

$$\mathcal{D}(\boldsymbol{p}_1, \boldsymbol{p}_2) = \frac{1}{\sqrt{p}} \sum_{j=1}^{p} \frac{\left\|(n_1)_j - (n_2)_j\right\|}{\max\limits_{x\in\mathcal{N}_j,\, y\in\mathcal{N}_j} \|x - y\|} + \left|\hat{\boldsymbol{d}}_1 - \hat{\boldsymbol{d}}_2\right|, \tag{16}$$

where $\|\cdot\|$ is a function norm for a function g defined on the wavelength interval,

$$\|g\| = \sqrt{\frac{1}{\lambda_{\max} - \lambda_{\min}} \int_{\lambda_{\min}}^{\lambda_{\max}} g^2(\lambda) \mathrm{d}\lambda}, \tag{17}$$

and $|\cdot|$ is the Euclidean ℓ^2-norm.

The clustering was parallelised with MPI such that the clustering on each set of live points runs in parallel.

Algorithm 1: Clustering.

Input: Set of live points $\{\boldsymbol{p}_i \mid i = 1, \ldots, N_{\text{live}}\}$.

For each live point i:
    Split $\boldsymbol{p}_i$ into $\boldsymbol{n}_i$ and $\boldsymbol{d}_i$.
    Preprocess $\boldsymbol{d}_i$ to calculate $\hat{\boldsymbol{d}}_i$ (Equation 15).
    Replace $\boldsymbol{p}_i$ with $(\boldsymbol{n}_i, \hat{\boldsymbol{d}}_i)$.
Run HDBSCAN on $\{\boldsymbol{p}_i \mid i = 1, \ldots, N_{\text{live}}\}$ with the distance metric $\mathcal{D}$ (Equation 16) and minimum cluster size set to 5.
Return clusters.

Algorithm 2: Finding local minima.

For each k, from $k = 1$ to $k = \lfloor N_{\text{dead, max}}/N_{\text{live}} \rfloor$:
    Get the parameter vectors $\{\boldsymbol{p}_i \mid i = 1, \ldots, N_{\text{live}}\}$ of the live points at PolyChord iteration $k\, N_{\text{live}}$.
    Run Algorithm 1 and store the clusters.
    Within each cluster, run a local optimiser from the point with the smallest merit function value. Store the resulting parameter vectors $\{\boldsymbol{p}^*\}$, one for each cluster.
Merge the sets $\{\boldsymbol{p}^*\}$ and remove duplicates.
Return the set of local minima, $\{\boldsymbol{p}^*\}$.

3.3 Deciding which layers to remove

Once the global minimum of the merit function is found, layers for which the thickness falls below

$$d_{\text{crit}} = 0.01 \times \frac{\lambda}{2\pi n \left|\cos\theta\right|} \tag{18}$$

have a transfer matrix,

$$\begin{pmatrix} \exp\left(-i\frac{2\pi n}{\lambda} d_{\text{crit}} \cos\theta\right) & 0 \\ 0 & \exp\left(i\frac{2\pi n}{\lambda} d_{\text{crit}} \cos\theta\right) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \underbrace{\begin{pmatrix} \mathcal{O}(10^{-2}) + i\mathcal{O}(10^{-2}) & 0 \\ 0 & \mathcal{O}(10^{-2}) + i\mathcal{O}(10^{-2}) \end{pmatrix}}_{\text{negligible}}, \tag{19}$$

so that the layer can be taken out of the stack. The optimisation and clustering were then re-performed for the multilayer coating with the reduced number of layers.

3.4 Robustness analysis

The robustness of the optimal solution to manufacturing errors was determined by sampling layer thicknesses from a Gaussian centred at the optimal thicknesses with a standard deviation of 1nm along each direction [3]. From the samples, the distributions and the 68% central credibility intervals of R_s, R_p, S, D and ϕ_{sp} were calculated.

3.5 Tests

The amplitude calculation (Section 2.2) was tested for a single layer by comparison with the analytically derived amplitude of a Fabry-Pérot interferometer [8]. For multiple layers, the amplitude was compared

to the result of a transfer matrix calculation [9]. Both tests showed agreement within floating-point precision.

A global optimisation performed on the 21-layer dichroic and 40-layer antireflection coating presented in [3] gave marginally better results than found in [3]. To get the result for the 21-layer coating, the number of live points had to be increased by a factor of 10, suggesting that the global optimum was not populated with live points using the standard settings in Table 2.

4 Results and Discussions

4.1 8-layer beamsplitter coating

4.1.1 Results

A calculation was made for the 8-layer 33.3% beamsplitter in [10]. Multiple calculations with different merit functions were made. The design specifications giving the smallest optimal merit function value, which is $f = 0.047$, and therefore the closest fit to the design, are given in Table 3. The refractive indices for the materials Infrasil301, Nb_2O_5 and SiO_2 were specified with the Sellmeier equation.

The wavelengths were chosen to be 100 equally spaced wavelengths in the range $[1.49\,\mu m,\ 2.31\,\mu m]$. This gave a lower optimal merit function value than if the wavelengths were chosen automatically with the procedure in Section 2.4 because the quadratic spacing of the wavelengths did not sufficiently penalise the deviation from the target reflectivity at larger wavelengths.

Subsequent clustering and local minimisation revealed a better solution than found by the global optimisation, for which the optimal parameters are shown in Table 4. Figures 2 and 3 show $R_s(\lambda)$, $R_p(\lambda)$, $S(\lambda)$, $D(\lambda)$ and $\phi_{sp}(\lambda)$ for the optimal parameters. The calculation was performed on 448 Intel Xeon Skylake processors @ 2.6GHz and took 12.3 minutes.

Table 3: 8-layer beamsplitter coating specification.

Parameter	**Value**
M	8
Outer medium	Air, $n_{\text{outer}} = 1$
Substrate	Infrasil301
Refractive index specification	Choose from $\{Nb_2O_5, SiO_2\}$
Thickness specification	The thickness of each layer is bounded between 0nm and 3000nm
θ_{outer}	15°
Wavelengths	100 equally spaced wavelengths in the range $[1.49\,\mu m,\ 2.31\,\mu m]$
Merit function	$w_{R_s} = 100$, $\tilde{R}_s = 33.3\%$; $w_{R_p} = 100$, $\tilde{R}_p = 33.3\%$; $w_S = 0$; $w_D = 1$, $\tilde{D} = 0$; $w_{\phi_{sp}} = 10$, $\tilde{\phi}_{sp} = 0$; and all other target reflectivities set to zero and all tolerances set to 1

Table 4: Optimal parameters for the 8-layer beamsplitter calculated by COATING.

Layer	Material	Thickness in nm
1	Nb_2O_5	205.48
2	SiO_2	443.85
3	Nb_2O_5	39.02
4	SiO_2	251.44
5	Nb_2O_5	30.55
6	SiO_2	895.08
7	Nb_2O_5	2.90
8	SiO_2	9.90

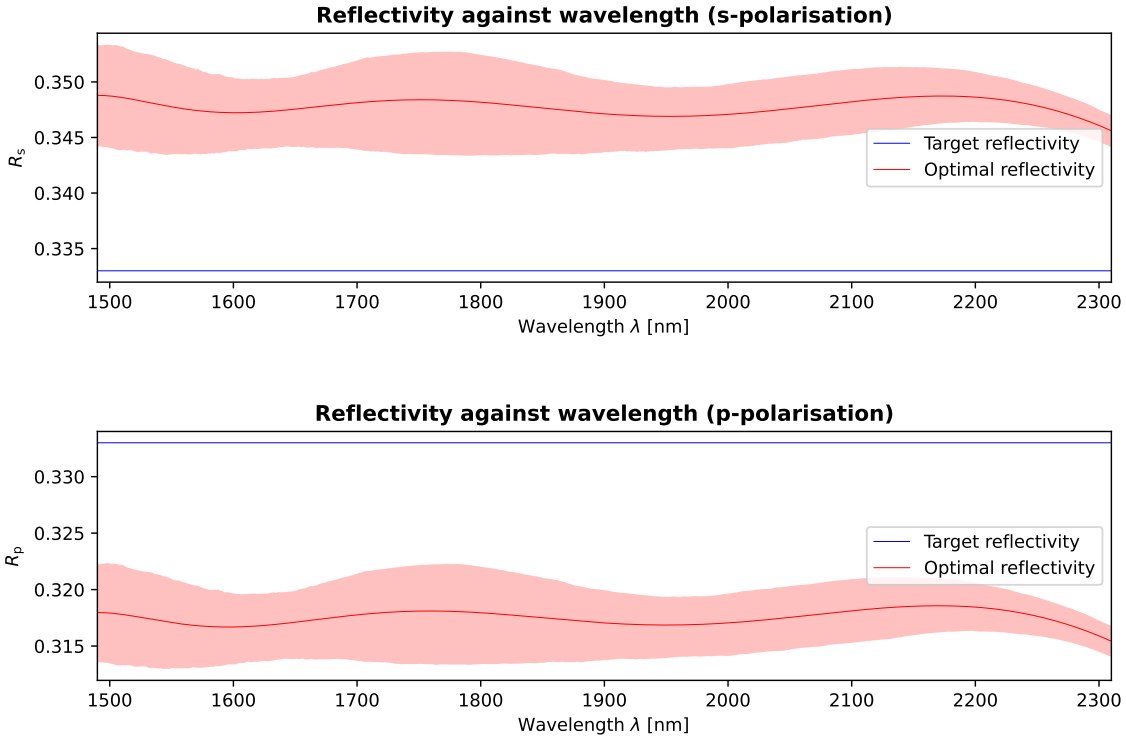

Figure 2: Reflectivities for s-polarised, R_s, and p-polarised, R_p, light as functions of wavelength for the optimal solution of the 8-layer coating in [10] found by COATING (red lines). The blue lines show the target values. The shaded red region is an error band which was calculated at each wavelength as described in Section 3.4.

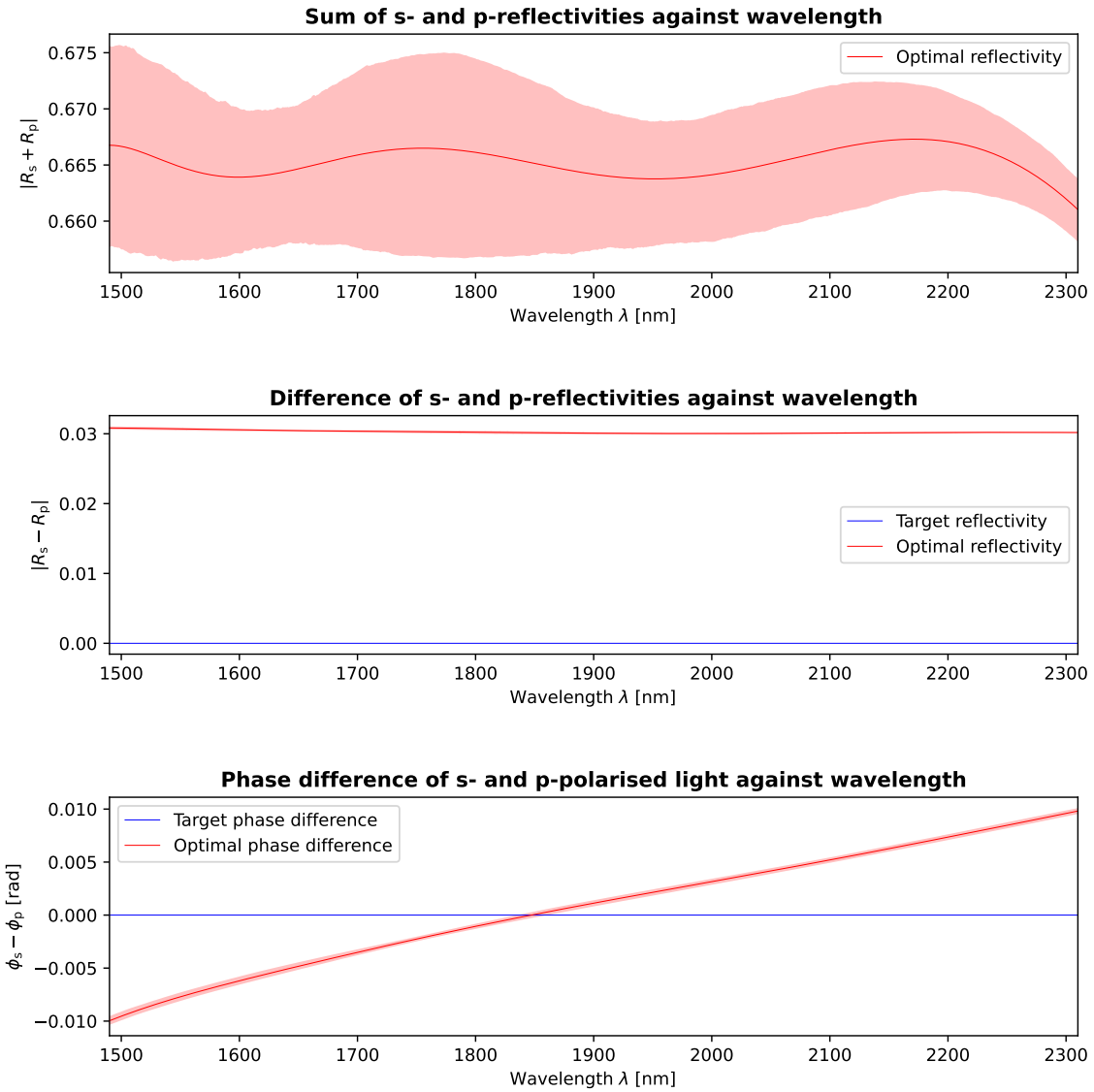

Figure 3: The sum of reflectivities, $S = R_s + R_p$, difference of reflectivities, $D = |R_s - R_p|$, and the phase difference, $\phi_{sp} = \phi_s - \phi_p$, as functions of wavelength for the optimal solution of the 8-layer coating in [10] found by COATING (red lines). The blue lines show the target values. The shaded red region is an error band which was calculated at each wavelength as described in Section 3.4.

4.1.2 Discussion

Table 5: Comparison between the coating design calculated by COATING and the design in [10].

Design parameter	**COATING solution**	**Chosen value in [10]**
Range of $\frac{R_s+R_p}{2}$ across all wavelengths	33.05% to 33.36%	32.83% to 33.62%
Total thickness	1878 nm	1800 nm
Materials	Alternating between Nb_2O_5 and SiO_2	Same as the COATING solution

Comparing the COATING solution with the design presented in [10] (Table 5), the COATING solution has a reflectivity which occupies approximately 39% of the range of the design in [10] at the expense

of being approximately 4% thicker. The materials of the COATING solution match the ones chosen in [10].

4.2 10-layer antireflection coating

4.2.1 Results

The design problem of the 21-layer antireflection coating, for which the optimal solution was found in [3] which was verified in Section 3.5, was modified by reducing the number of layers to 10 but letting the COATING code choose the optimal materials, instead of fixing the choice of materials [3]. The design specification is given in Table 6.

The resulting optimal parameters are shown in Table 7 and $R_p(\lambda)$ is shown in Figure 4. The calculation was performed on 448 Intel Xeon Skylake processors @ 2.6GHz and took 8.5 hours.

Table 6: 10-layer antireflection coating specification.

Parameter	**Value**
M	10
Outer medium	Air, $n_{\text{outer}} = 1$
Substrate	Typical glass, $n_{\text{substrate}} = 1.5$
Refractive index specification	Choose from $\{n_{\text{MgF}_2} = 1.37, n_{\text{SiO}_2} = 1.45, n_{\text{Ta}_2\text{O}_5} = 2.1\}$
Thickness specification	The thickness of each layer is bounded between 0nm and 350nm
θ_{outer}	0°
Wavelengths	100 equally spaced wavelengths in the range [600 nm, 2300 nm]
Merit function	$w_{R_p} = 10^4$, $\tilde{R}_p = 0$; and all other weights and target reflectivities set to zero and all tolerances set to 1

Table 7: Optimal parameters for the 10-layer antireflection coating calculated by COATING.

Layer	**Material**	**Thickness in nm**
1	MgF_2	234.89
2	Ta_2O_5	24.57
3	MgF_2	104.80
4	Ta_2O_5	35.51
5	SiO_2	113.02
6	Ta_2O_5	22.04
7	SiO_2	223.35
8	MgF_2	350.00
9	SiO_2	111.48
10	MgF_2	30.21

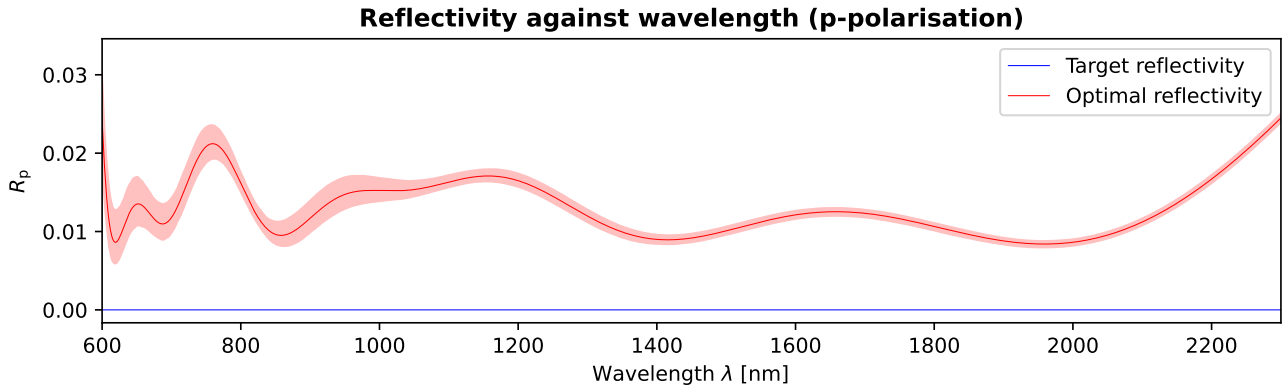

Figure 4: The reflectivity for p-polarised light, R_p, as functions of wavelength for the optimal solution of the 10-layer coating, found by Coating (red lines). The blue line shows the target value. The shaded red region is an error band which was calculated at each wavelength as described in Section 3.4.

4.2.2 Discussion

Table 8: Comparison between the 10-layer coating design calculated by Coating and the 21-layer design in [3].

Design parameter	**Coating solution**	**Solution in [3]**
Mean reflectivity $\langle R_p \rangle_\lambda$	1.3%	0.7%
Total thickness	1250 nm	1681 nm
Materials	Choices are shown in Table 7	First layer is MgF_2; other layers alternate between SiO_2 and Ta_2O_5

The mean reflectivity of the 10-layer coating ($\langle R_p \rangle_\lambda = 1.3\%$) is further away from the target reflectivity than the 21-layer solution, as is expected because the optimisation has less degrees of freedom to fit the reflectivity to the target reflectivity. However, the layer thickness is decreased by 26%, which is more desirable for manufacturing.

5 Conclusions

The code Coating was developed, which performs a global minimisation of a specified merit function using PolyChord to find the optimal design parameters for a multilayer coating and additionally finds subsidiary local minima, performs a robustness analysis and removes negligibly thin layers. The constraints on the layer thicknesses, the choice of the refractive indices and the merit function, which encapsulates the coating properties to be optimised, is customisable by the user without the requirement for an initial guess for the optimiser.

Coating was tested on multilayer coating optimisation problems for which the optimal solutions were previously published, and marginal improvements were found. Furthermore, by testing Coating on the design problem of an 8-layer beamsplitter coating used in interferometric applications, the result of the calculation was able to match the target reflectivity 39% more closely than previously published, while increasing the total thickness of the coating by 4%. Finally, by being able to choose the layer materials freely, an antireflection coating was designed with 10 instead of 21 layers, showing similar performance.

In future improvements on Coating, the number of layers should be chosen automatically by treating it as another variable in the optimisation problem. Moreover, the performance of a differ-

ent clustering algorithm or distance metric and the minimum number of required live points and wavelengths should be investigated.

Acknowledgements

The author acknowledges very helpful discussions with Will Handley.

Appendix A The matrix equation

The matrix $\mathbf{M} \in \mathbb{C}^{(2M+2)\times(2M+2)}$ in Equation 1 is[1]

$$\mathbf{M} = \left(\begin{array}{cccccccc} \begin{matrix}\zeta\\ \xi\end{matrix} & \boldsymbol{\beta}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \begin{matrix}0\\0\end{matrix} \\ \hline \begin{matrix}0\\0\end{matrix} & \boldsymbol{\alpha}_1 & \boldsymbol{\beta}_2 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \begin{matrix}0\\0\end{matrix} \\ \hline \begin{matrix}0\\0\end{matrix} & \mathbf{0} & \boldsymbol{\alpha}_2 & \boldsymbol{\beta}_3 & \mathbf{0} & \cdots & \mathbf{0} & \begin{matrix}0\\0\end{matrix} \\ \hline \begin{matrix}0\\0\end{matrix} & \mathbf{0} & \mathbf{0} & \boldsymbol{\alpha}_3 & \boldsymbol{\beta}_4 & \cdots & \mathbf{0} & \begin{matrix}0\\0\end{matrix} \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \boldsymbol{\beta}_M & \begin{matrix}0\\0\end{matrix} \\ \hline \begin{matrix}0\\0\end{matrix} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\alpha}_M & \begin{matrix}\tau\\ \upsilon\end{matrix} \end{array}\right), \tag{20}$$

where

$$\mathbf{0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{21}$$

To solve Equation 1 efficiently, its band-storage form, $\mathbf{M}_{\text{band}} \in \mathbb{C}^{5\times(2M+2)}$, is used:

$$\mathbf{M}_{\text{band}} = \begin{pmatrix} 0 & & & & 0 \\ 0 & & & & \tau \\ \zeta & \boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_M & \upsilon \\ \xi & & & & 0 \\ 0 & & & & 0 \end{pmatrix}, \tag{22}$$

where

$$\boldsymbol{c}_j = \begin{pmatrix} 0 & (\boldsymbol{\beta}_j)_{01} \\ (\boldsymbol{\beta}_j)_{00} & (\boldsymbol{\beta}_j)_{11} \\ (\boldsymbol{\beta}_j)_{10} & (\boldsymbol{\alpha}_j)_{01} \\ (\boldsymbol{\alpha}_j)_{00} & (\boldsymbol{\alpha}_j)_{11} \\ (\boldsymbol{\alpha}_j)_{10} & 0 \end{pmatrix} \qquad \text{for} \quad j = 1, \ldots, M. \tag{23}$$

For s-polarisation,

$$\zeta = -1, \tag{24}$$

$$\xi = 1, \tag{25}$$

$$\tau = 1, \tag{26}$$

$$\upsilon = n_{\text{substrate}} \cos\theta_{\text{substrate}}, \tag{27}$$

$$\boldsymbol{\alpha}_j = \begin{pmatrix} -e^{i\varphi_j} & -1 \\ -e^{i\varphi_j} n_j \cos\theta_j & n_j \cos\theta_j \end{pmatrix} \qquad \text{for} \quad j = 1, \ldots, M, \tag{28}$$

$$\boldsymbol{\beta}_j = \begin{cases} \begin{pmatrix} 1 & e^{i\varphi_1} \\ \frac{n_1 \cos\theta_1}{n_{\text{outer}} \cos\theta_{\text{outer}}} & -\frac{n_1 \cos\theta_1}{n_{\text{outer}} \cos\theta_{\text{outer}}} e^{i\varphi_1} \end{pmatrix} & \text{for} \quad j = 1 \\ \begin{pmatrix} 1 & e^{i\varphi_j} \\ n_j \cos\theta_j & -n_j \cos\theta_j e^{i\varphi_j} \end{pmatrix} & \text{for} \quad j = 2, \ldots, M \end{cases}. \tag{29}$$

[1]The horizontal lines are for readability.

For p-polarisation,

$$\zeta = -1, \tag{30}$$

$$\xi = 1, \tag{31}$$

$$\tau = \cos\theta_{\text{substrate}}, \tag{32}$$

$$\upsilon = n_{\text{substrate}}, \tag{33}$$

$$\boldsymbol{\alpha}_j = \begin{pmatrix} -e^{i\varphi_j}\cos\theta_j & -\cos\theta_j \\ -e^{i\varphi_j} n_j & n_j \end{pmatrix} \qquad \text{for} \quad j = 1, \ldots, M, \tag{34}$$

$$\boldsymbol{\beta}_j = \begin{cases} \begin{pmatrix} \frac{\cos\theta_1}{\cos\theta_{\text{outer}}} & e^{i\varphi_1}\frac{\cos\theta_1}{\cos\theta_{\text{outer}}} \\ \frac{n_1}{n_{\text{outer}}} & -e^{i\varphi_1}\frac{n_1}{n_{\text{outer}}} \end{pmatrix} & \text{for} \quad j = 1 \\ \begin{pmatrix} \cos\theta_j & e^{i\varphi_j}\cos\theta_j \\ n_j & -e^{i\varphi_j} n_j \end{pmatrix} & \text{for} \quad j = 2, \ldots, M \end{cases}, \tag{35}$$

where

$$\varphi_j = k_{\text{outer}} d_j \frac{n_j}{n_{\text{outer}}} \cos\theta_j, \tag{36}$$

$$\cos\theta_j = \sqrt{1 - \left(\frac{n_{\text{outer}} \sin\theta_{\text{outer}}}{n_j}\right)^2}. \tag{37}$$

Appendix B List of Symbols

Symbol	Definition
M	Number of layers in a multilayer coating
j	Layer index
i	$\sqrt{-1}$
n_j	Refractive index, as a function of wavelength, of layer j
d_j	Thickness of layer j
θ_j	Angle of the light ray with respect to the interface normal in layer j
λ	Wavelength
a_j, b_j	Electric field amplitudes in the multilayer coating
φ_j	Phase accrued in layer j
$\mathbf{M}$	Matrix for system of equations of a multilayer coating
$\boldsymbol{c}_j$	Submatrix of $\mathbf{M}$
$\boldsymbol{\alpha}_j$	Submatrix of $\mathbf{M}$
$\boldsymbol{\beta}_j$	Submatrix of $\mathbf{M}$
$\zeta, \xi, \tau, \upsilon$	Elements of $\mathbf{M}$
$\mathbf{x}$	Unknown vector containing the electric field amplitudes
$\mathbf{c}$	Right-hand side of the system of equations
$\mathcal{O}$	Big-O notation
R_s	Reflected amplitude for s-polarised light
R_p	Reflected amplitude for p-polarised light
S	Sum of reflectivities
D	Difference of reflectivities
ϕ_{sp}	Phase difference between the s- and p-polarised reflected amplitudes
w	Weighting of a term in the merit function
$\boldsymbol{p}_{\text{total}}$	Parameter vector containing all refractive indices and thicknesses
K	Number of wavelengths used in the merit function calculation
$\boldsymbol{p}$	Parameter vector containing unfixed refractive indices and thicknesses
N_{dim}	Dimensionality of the optimisation problem
N_{live}	Number of PolyChord live points
N_{dead}	Number of PolyChord dead points
p	Number of unfixed refractive indices
q	Number of unfixed thicknesses
f	Merit function
Z	Evidence
P	Posterior probability distribution
L	Likelihood
L^*	Minimum likelihood value of a set of live points
π	Prior probability distribution
$\boldsymbol{\theta}$	General parameter vector in PolyChord
Cat	Categorical distribution
Uni	Uniform continuous distribution
$\mathcal{N}_j$	Set of refractive indices to choose from for layer j
F_j	Component of the mapping from the unit interval to the parameter vector
$\mathcal{D}$	Distance metric used in clustering

References

[1] Verly P. G., Poitras D., 2018. *Optical Thin Films and Coatings: From Materials to Applications*, 2nd ed. (Woodhead Publishing), pp. 25–64.

[2] Handley W. J., Hobson M. P., Lasenby A. N., 2015. *PolyChord: next-generation nested sampling*, MNRAS 453, 4384–4398.

[3] Hobson M. P., Baldwin J. E., 2004. *Markov-chain Monte Carlo approach to the design of multilayer thin-film optical coatings*, Applied Optics 43, 2651–2660.

[4] Skilling J., 2006. *Nested Sampling for General Bayesian Computation*, Bayesian Analysis 1, 833–860.

[5] Nelder J. A., Mead R., 1965. *A Simplex Method for Function Minimization*, The Computer Journal 7, 308-–313.

[6] McInnes L, Healy J., 2017. *Accelerated Hierarchical Density Based Clustering*, IEEE International Conference on Data Mining Workshops (ICDMW), 33–42.

[7] Campello R., Moulavi D., and Sander J., 2013. *Advances in Knowledge Discovery and Data Mining*, 1st ed. (Springer), pp. 160–172.

[8] Hecht E., 2017. *Optics*, 5th ed. (Pearson).

[9] Furman Sh. A., Tikhonravov A. V., 1992. *Basics of Optics of Multilayer Systems*, 2nd ed. (Editions Frontieres), pp. 1–149.

[10] Nyland K., Jurgenson C., Buscher D., Haniff C., Young J., et al., 2010. *Custom beamsplitter and AR coatings for interferometry*, Proc. of SPIE 7734, 1–14.