Binary Optimization Techniques for Linear PDE-governed Material Design

J. Saa-Seoane*, H. Men, N.-C. Nguyen, R. Freund, J. Peraire

Aerospace Computational Design Laboratory
Department of Aeronautics and Astronautics
and Sloan School of Management
Massachusetts Institute of Technology
*corresponding author, E-mail: jsaa@mit.edu

Abstract

The design of materials is currently a fertile research domain. However, most of the material designs described in the literature arise from physical intuition, and often assume infinite periodicity. There is a need for a design methodology capable of computing patterns and designs involving two different materials where the underlying design variables correspond to a finite set of pixels in a 2-dimensional mesh, and where the goal is a design with prescribed material properties. This naturally leads to the consideration of binary optimization models in contrast to classical (continuous) gradient-based methods which generically provide continuous solutions that then need to be "rounded" to binary values. While the potential drawback of binary optimization is that its computational complexity is usually NP-hard and hence theoretically unattractive, we show herein that binary optimization combined with a reduced basis approach can relatively efficiently produce good solutions to material design problems of interest.

1. Introduction

Wave phenomena in acoustics, elastodynamics and electromagnetics have been widely studied in the last two decades. These phenomena have found numerous applications in many domains of engineering, and therefore accurate, efficient, and reliable numerical simulation is extremely important. For the problems in our scope of concern, a Finite Element Method is required since it provides geometric flexibility and well-known error bounds. In particular, the hybridizable discontinuous Galerkin (HDG) method [1] for acoustics and elastodynamics as well as in [2] for Maxwell's equations and thus electromagnetics, has been proven to be a robust, accurate and efficient simulation tool for these sort of problems. Indeed, these methods were devised to guarantee that only the degrees of freedom of the approximation of the scalar variable on the interelement boundaries are globally coupled, see [7]. In addition to that, this method will provide particular properties that will become essential later on for the work presented in this paper.

The family of problems related to wave phenomena has lately increased due to the growing interest in metamaterial design. Metamaterials are materials that have very particular properties due only to their structure, not their composition. There is often a microscopic pattern of existing materials that, when extended periodically, will provide a particularly effective macroscopic property that is otherwise unobtainable. Potential applications of metamaterials in acoustics involve sound bullets and acoustic filters [6], negative Poisson's ratio materials in elastodynamics [8], and regarding electromagnetism, cloaking devices [3] and photonic bandgap phenomena [4] among others. Some of the examples above will be further analyzed in this paper as well as a similar extension to the heat equation.

The design of metamaterials is difficult because modeling and computing effective design patterns is fraught with computational challenges. Especially when trying to design manufacturable and realizable materials, physical and mathematical intuition are insufficient by themselves. One is therefore led to consider optimization-based approaches to design. However, the optimization problems that arise in metamaterial design are often of discrete nature, leading to binary or mixed-integer optimization models. Indeed, in considering design variables that correspond to a finite set of pixels, the optimization problem is to choose between two given materials for each pixel, hence the typical application problem results in the need to solve a binary optimization model. In this paper, we present a binary optimization model approach that combines local search approximation with a reduced basis approximation, that produces good local solutions with good relative efficiency. Our approach utilizes a reduced basis projected problem [5] and the use of binary generalized gradients to ensure feasibility of all solutions.

Some successful approaches to material design through optimization can be found in the literature. Adjoint methods solve efficiently shape optimization problems but usually lead to continuous optimal solutions if the design variables are the material properties. Some remarkable applications have also been solved using topology optimization [13, 14], which does not ensure discrete solutions, and level set methods [15], which are highly nonlinear. The approach here introduced proposes an alternative method for the binary design of materials without using information from any continuous gradient. In fact, such information might lead to highly suboptimal patterns since solutions are in general very sensitive and continuous optima lay faraway from binary optima.

This paper is organized as follows. Section 2 explains

the suitability of the HDG method used as well as derives the particular formulation for the Helmholtz equation case. Section 3 derives the reduced basis algorithm used for the binary optimization local search method that is used for the design optimization. Results for the Poisson's equation as well as for the one dimensional bandgap problem are further analyzed in section 4.

2. HDG method for Helmholtz equation

In this section we want to describe a Hybridizable Discontinuous Galerkin (HDG) Method for a model Helmholtz equation. The extension of these results to the linear second order wave equation is trivial and can be found in [1, 2]. The HDG method first introduced in [7] possess a number of attractive properties for wave propagation problems. In particular, the HDG method results in a smaller global system of equations, especially for high orders of accuracy. As a result, the method provides more accurate solutions than other finite element methods for the same mesh discretization. The HDG method is very low dispersive and diffusive. These reasons make the HDG method suitable for the simulation of wave phenomena.

2.1. HDG derivation

In this section we show how the hybridizable Discontinuous Galerkin method applies to a model Helmholtz equation. The extension from these derivations to a different domain or to Poisson's equation will be later discussed. To that end, let us firstly consider the Helmholtz problem as follows:

$$\begin{array}{rl} -\nabla\cdot\varepsilon\nabla u - k^2 u = f & \text{in } \Omega\subset\mathbb{R}^d \\ \varepsilon\nabla u\cdot\vec{n} + iku = g & \text{on } \partial\Omega_a \end{array} \tag{1}$$

where u is a scalar variable, ε is the square of the propagation speed, k is the wavenumber, f is a given source term and g determines the absorbing boundary condition. Moreover, Ω is a Lipschitz polyhedral domain in $\mathbb{R}^{d\geq 1}$. Note that the boundary condition considered in the description is first-order absorbing and it is taken solely for the purposes of illustration. Such boundary condition could be easily replaced by higher-order local or exact global conditions as well as by suitable perfectly matched layers.

The HDG method firstly writes the partial differential equation (PDE) as a first-order system of partial differential equations and thus, after introducing the gradient as $\mathbf{q} = \nabla u$ for convenience, the following system can be written:

$$\begin{array}{rl} \mathbf{q} - \nabla u = 0 & \text{in } \Omega \\ -\nabla\cdot\varepsilon\mathbf{q} - k^2 u = f & \text{in } \Omega \\ \varepsilon\mathbf{q}\cdot\mathbf{n} + iku = g & \text{on } \partial\Omega_a \end{array} \tag{2}$$

Later on and for convenience, the term iku will be pushed into the righthandside modifying f and becoming a function of u. The second equation will therefore become $-\nabla\cdot\varepsilon\mathbf{q} = f_u$, where $f_u = f - iku$.

Let $\mathcal{T}_h$ form a triangulation of the domain Ω into elements K and $\partial\mathcal{T}_h = \{\partial K, K\in\mathcal{T}_h\}$ be the set of faces F of each element K of the triangulation, also known as $\mathcal{F}_h$. Then the method seeks a scalar approximation u_h to u, a vector approximation $\mathbf{q}_h$ to $\mathbf{q}$ and a scalar approximation $\hat{u}_h$ to the traces $\hat{u}_h$ minimizing the representation error in -or distance to- some approximation spaces defined as:

$$\begin{array}{l} W_h = \{w\in L^2(\mathcal{T}_h), w|_K \in W(K), \forall K\in\mathcal{T}_h\} \\ \mathbf{V}_h = \{\mathbf{v}\in [L^2(\mathcal{T}_h)]^d, \mathbf{v}|_K \in \mathbf{V}(K), \forall K\in\mathcal{T}_h\} \\ M_h = \{\mu\in L^2(\mathcal{F}_h), \mu|_F \in M(F), \forall F\in\mathcal{F}_h\} \end{array} \tag{3}$$

where $W(K)$, $\mathbf{V}(K)$ and $M(F)$ are suitably chosen finite dimensional spaces. Furthermore, let us define the contractions involved within this HDG method. For functions $\mathbf{v},\mathbf{w}\in[L^2(D)]^d$ we denote $(\mathbf{v},\mathbf{w})_D = \int_D \mathbf{v}\cdot\overline{\mathbf{w}}$; for functions $v,w\in L^2(D)$ we write $(v,w)_D = \int_D v\overline{w}$ if D is a domain in $\mathbb{R}^d$ and $\langle v,w\rangle_D = \int_D v\overline{w}$ if D is a domain in $\mathbb{R}^{d-1}$. We finally introduce

$$\begin{array}{rcl} (v,w)_{\mathcal{T}_h} & = & \displaystyle\sum_{K\in\mathcal{T}_h} (v,w)_K \\ (\mu,\eta)_{\partial\mathcal{T}_h} & = & \displaystyle\sum_{K\in\mathcal{T}_h} (\mu,\eta)_{\partial K} \end{array} \tag{4}$$

for v,w defined in $\mathcal{T}_h$ and μ,η defined on $\partial\mathcal{T}_h$ respectively. The HDG approximations $u_h\in W_h$, $\mathbf{q}_h\in\mathbf{V}_h$ and $\widehat{u}_h\in M_h$ are now determined by requiring that the following finite discrete system of equations holds $\forall(s,\mathbf{r},\mu)\in W_h\times\mathbf{V}_h\times M_h$.

$$\begin{array}{r} (\mathbf{q}_h,\mathbf{r})_{\mathcal{T}_h} + (u_h,\nabla\cdot\mathbf{r})_{\mathcal{T}_h} - \langle\widehat{u}_h,\mathbf{r}\cdot\mathbf{n}\rangle_{\partial\mathcal{T}_h} = 0, \\ (\varepsilon\mathbf{q}_h,\nabla w)_{\mathcal{T}_h} - \langle\widehat{\varepsilon\mathbf{q}}_h\cdot\mathbf{n}, w\rangle_{\partial\mathcal{T}_h} - k^2(u_h,w)_{\mathcal{T}_h} = (f,w)_{\mathcal{T}_h}, \\ -\langle\widehat{\varepsilon\mathbf{q}}_h\cdot\mathbf{n},\mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega_a} + \langle\widehat{\varepsilon\mathbf{q}}_h\cdot\mathbf{n} + ik\widehat{u}_h,\mu\rangle_{\partial\Omega_a} = \langle g,\mu\rangle_{\partial\Omega}. \end{array} \tag{5}$$

Note that the HDG method uses the extra stabilization condition for the flux traces through the definition of the numerical fluxes as

$$\widehat{\varepsilon\mathbf{q}}_h = \varepsilon\mathbf{q}_h + \tau(u_h - \widehat{u}_h)\mathbf{n} \tag{6}$$

on $\partial\mathcal{T}_h$. Here, τ is the so-called *stabilization* function. The actual definition of the numerical traces $\widehat{\mathbf{q}}_h$ is the key feature of the HDG method. The last equation in (5), which is defined over the degrees of freedom on the edges, can be solved for $\widehat{u}_h$ in a global sense and, after that, the rest of the equations will locally recover u_h and $\mathbf{q}_h$. Such local systems of equations can be totally parallelized and thus solved very efficiently for the degrees of freedom inside each element. Moreover, it has been shown in [9] that this Helmholtz problem actually achieves an optimal superconvergence order for both u_h and $\mathbf{q}_h$ after a postprocessing of the solution, therefore relatively coarse meshes can be used even for high material contrasts.

2.2. Implementation

Let us first of all define as U the variables related to the displacement u_h of the degrees of freedom inside each element, Q the variables related to their fluxes $\mathbf{q}_h$ and Λ the

variables related to the traces $\widehat{u}_h$ for every degree of freedom along the edges of the triangulation. Now if equation (6) is plugged into the system of equations (5) we eliminate the $\widehat{\mathbf{q}}_h$ variables and thus the following system of equations is obtained.

$$\begin{aligned}(\mathbf{q}_h,\mathbf{r})_{\mathcal{T}_h}+(u_h,\nabla\cdot\mathbf{r})_{\mathcal{T}_h}-\langle\widehat{u}_h,\mathbf{r}\cdot\mathbf{n}\rangle_{\partial\mathcal{T}_h}=\\ 0,\\ -(\varepsilon\nabla\cdot\mathbf{q}_h,w)_{\mathcal{T}_h}-\tau\langle(u_h-\widehat{u}_h),w\rangle_{\partial\mathcal{T}_h}-\\ k^2(u_h,w)_{\mathcal{T}_h}=\\ (f,w)_{\mathcal{T}_h},\\ -\langle\varepsilon\mathbf{q}_h\cdot\mathbf{n}-\tau(u_h-\widehat{u}_h),\mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega_a}+\\ \langle\varepsilon\mathbf{q}_h\cdot\mathbf{n}-\tau u_h-(\tau+ik)\widehat{u}_h,\mu\rangle_{\partial\Omega_a}=\\ \langle g,\mu\rangle_{\partial\Omega}.\end{aligned} \tag{7}$$

Now this system of equations gives rise to a matrix equation that can be written as follows:

$$\begin{bmatrix}\mathbb{A} & -\mathbb{B}^t & -\mathbb{C}^t\\ \mathbb{B}(\varepsilon) & \mathbb{D}^t & \mathbb{E}^t\\ \mathbb{C}(\varepsilon) & \mathbb{E} & \mathbb{M}\end{bmatrix}\begin{bmatrix}Q\\U\\\Lambda\end{bmatrix}=\begin{bmatrix}0\\F\\G\end{bmatrix} \tag{8}$$

where the submatrices $\mathbb{A},\mathbb{B},\mathbb{C},\mathbb{D},\mathbb{E}$ and $\mathbb{M}$ correspond to the discretization of the dot products above. One very interesting property of the HDG method shows up at this point: the matrices $\mathbb{A},\mathbb{B},\mathbb{C}$ and $\mathbb{D}$ are block diagonal, i.e. they are very sparse and every single element only contributes with nonzero entries to the degrees of freedom of that element. This fact will actually allow us to write equation (8) as a system of equations for Λ and then back solve for the inner element degrees of freedom, through the Schur complement. Moreover, there is another key property at this point: ε can be pulled out from the terms where it shows up. We will therefore be able to write the system (8) as follows:

$$\left(\mathbb{K}_0+\sum_{q=1}^{N_{el}}\varepsilon^q\mathbb{K}_q\right)\mathbf{u}=\mathbf{f} \tag{9}$$

for N_{el} number of elements. After writing Q from the first row in equation (8) as $Q=\mathbb{A}^{-1}(\mathbb{B}^tU+\mathbb{C}^t\Lambda)$ and considering $\mathbf{u}=[U\ \Lambda]^t$ and $\mathbf{f}=[F\ G]^t$ equation (9) holds for the following values of the matrices $\mathbb{K}_q,\forall q=0..N_{el}$.

$$\begin{aligned}\mathbb{K}_0&=\begin{bmatrix}\mathbb{D} & \mathbb{E}^t\\ \mathbb{E} & \mathbb{M}\end{bmatrix},\\ \mathbb{K}_q&=\begin{bmatrix}\mathbb{B}\mathbb{A}^{-1}\mathbb{B}^t & \mathbb{B}\mathbb{A}^{-1}\mathbb{C}^t\\ \mathbb{C}\mathbb{A}^{-1}\mathbb{B}^t & \mathbb{C}\mathbb{A}^{-1}\mathbb{C}^t\end{bmatrix},\forall q=1..N_{el}.\end{aligned} \tag{10}$$

Moreover, note how the terms $\mathbb{D}$ and $\mathbb{B}\mathbb{A}^{-1}\mathbb{B}^t$ are block diagonal (since both $\mathbb{A}$ and $\mathbb{B}$ are) and thus the system in (9) can be efficiently solved as any HDG method, i.e. solve only for the unknowns on the edges.

2.3. First-order absorbing boundary conditions

The first order absorbing boundary conditions have been introduced in [1] for the time dependent wave equation as

$$\frac{\partial u}{\partial t}+\nabla u\cdot\mathbf{n}=0. \tag{11}$$

Here, we are dealing with the Helmholtz equation which is the steady version of the second order time dependent wave equation derived through separation of variables. If we thus assume $u(x,t)=u(x)e^{i\omega t}$ and plug it in equation (11) we obtain the following expression:

$$\nabla u(x)\cdot\mathbf{n}=-i\omega u(x) \tag{12}$$

Furthermore, we are actually interested in applying the absorbing boundary conditions to the scattered field instead of the total field. We can thus write $u^s=u-u^0$, where u^s represents the scattered field, u the total solution and u^0 the original solution, i.e. initial condition in the time dependent problem. If we finally write (12) in terms of the scattered field we obtain the following expression:

$$\begin{aligned}\nabla u^s\cdot\mathbf{n}&=-i\omega u^s\\ &\Updownarrow\\ \nabla u\cdot\mathbf{n}&=-i\omega u+\nabla u^0\cdot\mathbf{n}+i\omega u^0\end{aligned} \tag{13}$$

yet to be applied to each of the boundaries in the actual domain Ω. Moreover, if there are extra boundary conditions which are Neumann we can just use $\nabla\cdot\mathbf{u}=h$ and proceed identically for any h; if there are any Dirichlet boundary conditions we may just change the approximation spaces introduced above to fit the values on such boundaries.

3. Binary Optimization

For a given wave phenomenon problem, let us consider $\boldsymbol{\varepsilon}$ to be the property defining each material. Since the problem will be governed by a Partial Differential Equation of the form $F(\mathbf{u}(\boldsymbol{\varepsilon}),\boldsymbol{\varepsilon})=0$, the discretized PDE (with N_{el} discretized elements) can be expressed as a system of the form $A(\boldsymbol{\varepsilon})\mathbf{u}=\mathbf{f}$ in the linear case. Moreover, using the HDG discretization introduced in section 2.1, the system matrix can be written as $A(\boldsymbol{\varepsilon})=\mathbb{K}_0+\sum_{q=1}^{N_{el}}\varepsilon^q\mathbb{K}_q$ as in equation (9). Let $J(\mathbf{u}(\boldsymbol{\varepsilon}),\boldsymbol{\varepsilon})$ be the objective function measuring the deviation to a desired and known solution −often just $J(\mathbf{u}(\boldsymbol{\varepsilon}),\boldsymbol{\varepsilon})=||\mathbf{u}(\boldsymbol{\varepsilon})-\mathbf{u_0}||_2^2$ and denoted by $J(\mathbf{u}(\boldsymbol{\varepsilon}))$−, then the metamaterial design optimization problem can be written in the following general form:

$$\begin{aligned}\min_{\boldsymbol{\varepsilon},\mathbf{u}}\quad & J(\mathbf{u}(\boldsymbol{\varepsilon}),\boldsymbol{\varepsilon})\\ \text{s.t.}\quad & \left(\tau\mathbb{K}_0+\sum_{q=1}^{N_{el}}\varepsilon^q\mathbb{K}_q\right)\mathbf{u}=\mathbf{f}\\ & \boldsymbol{\varepsilon}\in\{\varepsilon_{min},\varepsilon_{max}\}^{N_{el}}\end{aligned} \tag{14}$$

Problem (14) arises in many areas of applied engineering such as inverse problems, shape optimization, topology optimization, optimal design and optimal control. However, the PDE constraints and the nature of the design variables often pose several significant challenges for contemporary optimization methods. First, the problem is in general nonlinear and non-convex due to an implicit dependence of the objective function on the design variables through the underlying PDEs. Second, the problem is large-scale since the discretization of the PDEs leads to a very

large system of equations. And third, if some (or all) design variables can only take on integer or discrete values then problem (14) becomes a mixed-integer nonlinear optimization problem. Unfortunately, while discrete variables are common in practice, their presence causes the optimization problem to be NP-hard in general. It is therefore necessary to develop a suitable approximation of the problem in order to achieve computational tractability in practice.

In developing an approximation to the problem (14), we want to be able to efficiently compute the true objective function value. That is, for a given value of the design variables $\boldsymbol{\varepsilon}$, we want to compute $\mathbf{u}(\boldsymbol{\varepsilon})$ inexpensively and then compute $J(\mathbf{u})$. To that end, we will solve the PDE through a reduced basis approach. Subsection 3.1 derives the particular optimization problem after the reduced basis procedure is applied.

Also in the context of developing an approximation to the problem (14), in order to solve the optimization problem stated in (14) assuming we can now efficiently compute the objective function value, we still need to devise an optimization method that ensures the binary constraints $\boldsymbol{\varepsilon} \in \{\varepsilon_{\min}\varepsilon_{\max}\}^{N_{el}}$ are satisfied. To that end we introduce the binary gradients in Subsection 3.2.

3.1. Reduced Basis

The reduced basis method (RB) method can be used to provide an accurate, reliable and efficient solution of parametrized PDEs, see [10, 11] and further references therein. Material design or optimal control problems involve large numbers of parameters, and thus computing sensitivities or just solutions for the entire family of parameters is seldom achievable.

Let $n \leq N_{el}$ be the number of regions where a material parameter needs to be chosen and $k < n$ be a certain positive integer corresponding to the reduced basis size. For a given feasible pattern $\boldsymbol{\varepsilon} \in \{\varepsilon_{\min}, \varepsilon_{\max}\}^n$, let $\overline{\mathbf{u}}_1 = \mathbf{u}(\boldsymbol{\varepsilon})$, and define $k-1$ neighbors by just perturbing a small number of pixels from either $\varepsilon_{\min}$ to $\varepsilon_{\max}$ or *vice versa*, and then computing their corresponding solutions $\overline{\mathbf{u}}_j = \mathbf{u}(\boldsymbol{\varepsilon}^j)$ for $j = 2, \ldots, k$. We now define the reduced basis as $\overline{\Phi} = span[\overline{\mathbf{u}}_1, \overline{\mathbf{u}}_2, \ldots, \overline{\mathbf{u}}_k] \in \mathbb{R}^{N_{el} \times k}$, and we can then define an approximate version of any given $\mathbf{u}(\boldsymbol{\varepsilon})$ as $\tilde{\mathbf{u}}(\boldsymbol{\varepsilon}) = \sum_{j=1}^{k} \alpha_j(\boldsymbol{\varepsilon})\overline{\mathbf{u}}^k = \overline{\Phi}\boldsymbol{\alpha}(\boldsymbol{\varepsilon})$. Note that now the discretized system can be approximately solved as $\overline{\Phi}^t A(\boldsymbol{\varepsilon})\overline{\Phi}\boldsymbol{\alpha}(\boldsymbol{\varepsilon}) = \overline{\Phi}^t\mathbf{f}$. Finally, if we define $\tilde{A}^q = \overline{\Phi}^t A^q \overline{\Phi} \in \mathbb{R}^{k\times k}$, for $1 \leq q \leq n$ and also $\tilde{\mathbf{f}} = \overline{\Phi}^t\mathbf{f} \in \mathbb{R}^k$, we will be able to solve the governing system as:

$$\left(\sum_{q=1}^{n} \tilde{A}^q(\boldsymbol{\varepsilon})\right)\boldsymbol{\alpha}(\boldsymbol{\varepsilon}) = \tilde{\mathbf{f}} \tag{15}$$

which is a $k \times k$ system in contrast to the original $N_{el} \times N_{el}$ system. We then recover $\tilde{\mathbf{u}}(\boldsymbol{\varepsilon}) = \overline{\Phi}\boldsymbol{\alpha}(\boldsymbol{\varepsilon})$. Note, furthermore, that $\tilde{A}^q(\boldsymbol{\varepsilon})$ can be derived from A^q as in (9) and therefore still retains the property of being able to work with $\boldsymbol{\varepsilon}$ as needed.

3.2. Binary Gradient

Since we want to maintain binary solutions throughout the optimization process, we will only allow directional changes that leave a current pixel as is, or that flips $\varepsilon_{\min}$ to $\varepsilon_{\max}$ or *vice versa*. This can be done by defining the sensitivities of our objective function according to unitary changes instead of differential changes. To accomplish this we introduce the following binary generalized gradient:

$$G_m(\boldsymbol{\varepsilon}) = \frac{\Delta J(\mathbf{u}(\boldsymbol{\varepsilon}))}{\Delta\varepsilon} = \frac{J(\mathbf{u}(\boldsymbol{\chi}^m)) - J(\mathbf{u}(\boldsymbol{\varepsilon}))}{\varepsilon_{\max} - \varepsilon_{\min}}, \tag{16}$$

for $m = 1, \ldots, N_{el}$, where $\boldsymbol{\chi}^m$ just changes the m^{th} component of $\boldsymbol{\varepsilon}$ from $\varepsilon_{\min}$ to $\varepsilon_{\max}$ or *vice versa*. We then choose the descent direction that provides the smallest value of $G_m(\boldsymbol{\varepsilon})$ and advance in that descent direction iteratively, as in any steepest descent algorithm for continuous optimization.

3.3. Optimization Algorithm

Table 1 summarizes the optimization algorithm based on the ideas described above. Let $l \leq k$ be the size of the initial basis computed around an initial guess $\boldsymbol{\varepsilon}(0)$ and let the subindex of $\boldsymbol{\varepsilon}$ denote the vector position in the basis Φ.

Table 1: Binary Optimization algorithm

1-	Start with an initial guess $\boldsymbol{\varepsilon}(0)$,
2-	Obtain the objective function value $J(\mathbf{u}(\boldsymbol{\varepsilon}(0)))$,
3-	Compute the solutions $\mathbf{u}_1 \cdots \mathbf{u}_l$ for $\boldsymbol{\varepsilon}_1 \cdots \boldsymbol{\varepsilon}_l$ exactly,
4-	Form $\Phi(\boldsymbol{\varepsilon}(0)) = [\mathbf{u}(\boldsymbol{\varepsilon}_1) \cdots \mathbf{u}(\boldsymbol{\varepsilon}_l))]$,
5-	Compute binary sensitivities G_m using (15) and (16),
6-	If $G_m \geq 0, \forall m$, end. Else, pick $\bar{m} = \arg\min_m G_m$ and set $\boldsymbol{\varepsilon}(0) \leftarrow \boldsymbol{\chi}^{\bar{m}}$,
7-	Compute l_0 random neighbors and update $\Phi(\boldsymbol{\varepsilon}(0)) \leftarrow [\Phi(\boldsymbol{\varepsilon}(0))\ \mathbf{u}(\boldsymbol{\varepsilon}_1) \cdots \mathbf{u}(\boldsymbol{\varepsilon}_{l_0}))]$,
8-	If $size\{\Phi(\boldsymbol{\varepsilon}(0))\} = p > k$, remove the $p-k$ elements m with smallest values of $\alpha_m(\boldsymbol{\varepsilon}(0))$ in (15),
9-	Go to 2,

Note that this algorithm is actually a local search approach to the binary optimization problem (14). The complexity of binary optimization problems is NP-hard, which implies that whenever the variable set is large, the problem is generically intractable. In our case, the parameter space is very large, typically on the order of $\mathcal{O}(10^{2d})$ where d is the spatial dimension considered. Local search algorithms are a good approach to solve these problems. However, they are only able to guarantee local minima and the quality of the computed local minima really depends on the quality of the neighborhoods considered (often only very large neighborhoods work well). Metamaterial design optimization is yet harder, since unless the local search neighborhoods are very small, the computational burden of the local search methodology itself is excessive. By joining together the HDG properties and the reduced basis theory, we seek

a balance wherein the approximate local search algorithm will find local optima of good quality with relatively good computation time. Several starting guesses, as well as further clever enhancements - like letting the solution worsen slightly to avoid getting stuck at a bad local optima - might be required for some applications.

4. Results

We have successfully applied the methodology described herein to one-dimensional photonic bandgap problems. In particular, we have succeeded in designing a binary material that is able to totally reflect a given frequency considering the finiteness of the domain. This phenomena is well-known if the pattern is considered periodic and therefore infinite (and hence is not realizable) but is not so well-known for finite structures. Section 4.1 below analyzes this problem, comparing the binary solution computed herein with the continuous optimum. In Section 4.2, we apply the same optimization procedure to a 2-dimensional problem governed by the heat transfer equation.

4.1. The 1-dimensional Bandgap problem

Photonic crystals are periodic structures created from the arrangement of low and high index materials. They are designed to affect the motion of light by prohibiting the propagation of electromagnetic waves in all directions within certain frequency ranges. They have been of crucial use for the design of important novel devices and applications such as frequency filters, waveguides, switches and optical buffers, see for instance [4]. However, the results reported in a large fraction of the literature so far have been obtained without imposing integer constraints on the design variables.

Luckily, in the photonic bandgap problem, optimal solutions assuming infinite periodicity turn out to be binary, as observed by Lord Rayleigh as early as 1888, [12]. Nevertheless, if we are interested in extending the conceptual ideas introduced by the photonic bandgap to other wave phenomena, we need to mitigate the non-binary nature of the continuously relaxed optimal solution. Consequently, if we want to obtain satisfactory solutions −most notably fabricability−, we must effectively constrain solutions to be binary.

Figure 1 shows the 1d photonic bandgap application. The governing equation for the frequency domain problem is exactly the Helmholtz equation analyzed in the first section where ε is the permittivity of the material. The incident frequency corresponds to $\omega a/2\pi c \simeq 0.33$ after it has been normalized over the geometry. Here the photonic crystal is made up of two different materials: air ($\varepsilon = 1$, color coded as dark red) and silicon ($\varepsilon = 13$, color coded as blue).

Figure 1 shows the optimized structures obtained with (Top) the standard adjoint method with relaxation of the integer constraints, see [9] for the full derivations regarding the adjoint method for this particular example, (Middle) the adjoint method with enforcement of the integer constraints via projection into the closest binary value, and

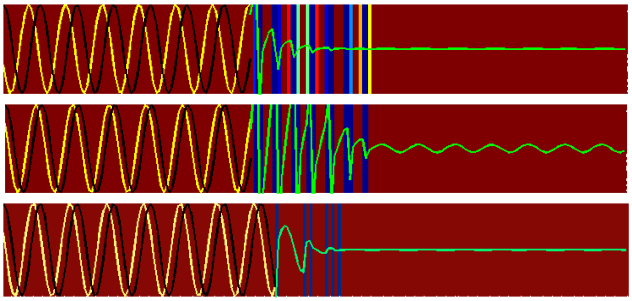

Figure 1: 1-dimensional Photonic Bandgap with yellow incident wave, black reflected wave and green transmitted wave. Top: continuous optimum; Middle: Discrete projection; Bottom: binary optimum.

(Bottom) our proposed method. Our method produces exactly (up to machine precision error) an optimal binary solution within only six iterations, whereas the standard adjoint method computes an optimal solution which is not binary (and thus not fabricable nor acceptable) and the projected adjoint method produces a binary solution which is not optimal (and thus an inferior design).

The discretization considered for this problem has 500 high order elements (order 3 providing 15 inner and boundary degrees of freedom per element) of which the middle 100 constitute the 50 pixels we seek to design. All possible combinations would lead to $2^{50} > 1.1 \cdot 10^{13}$ problems and thus would be intractable to solve. Convergence of the discretization has been achieved although the high material and solution contrasts thanks to the high order used, just as expected given the superconvergence provided by the HDG method [9].

This computational result is especially encouraging since the basis need not contain more than 10 solution vectors to guarantee a very good approximation of the exact solutions and therefore the systems of equations (15) used to compute the binary gradient never exceeded a 10×10 system. Furthermore the binary gradient computation (16) which is of order $\mathcal{O}(n)$ (recall n is the number of pixels or more generally the parameter space size) took less time than one single HDG computation takes. It is also encouraging that the binary gradient computations are extremely accurate, with errors of $\mathcal{O}(10^{-8})$.

4.2. Poisson's Equation

In a similar setting but in a 2-dimensional context we consider the heat transfer problem. For this problem the governing equation will be the Poisson equation instead of Helmholtz equation, however, we can just adapt our derivations in the first section by removing the term $-k^2u$ that was actually included in the source term f_u for this purpose. The problem we want to solve now will be governed by the following partial differential equation:

$$\begin{aligned} -\nabla \cdot \varepsilon \nabla u &= f \quad \text{in } \Omega = [0,1]^2 \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{17}$$

where the source term has been chosen to be $f = 2\pi^2 \sin(\pi x)\sin(\pi y)$. We seek a 2-dimensional pattern maximizing the heat transferred from the Dirichlet boundaries of a square plate into the center point. Note that if we do not include an extra volume constraint, the optimum will be obtained when the material used everywhere corresponds to the one holding a larger thermal conductivity. Therefore the overall setting of the optimization problem for this case will be slightly modified by the volume constraint. If we choose $0 \leq \beta \leq 1$ as the volume fraction that we are allowed to change, the problem can be written as:

$$\begin{aligned} \min_{\boldsymbol{\varepsilon},\mathbf{u}} \quad & J(\mathbf{u}(\boldsymbol{\varepsilon}),\boldsymbol{\varepsilon}) \\ \text{s.t.} \quad & \left(\tau\mathbb{K}_0 + \sum_{q=1}^{N_{el}} \varepsilon^q \mathbb{K}_q\right)\mathbf{u} = \mathbf{f} \\ & \frac{1}{N_{el}} \sum_{q=1}^{N_{el}} \frac{\varepsilon^q - \varepsilon_{\min}}{\varepsilon_{\max} - \varepsilon_{\min}} \leq \beta \\ & \boldsymbol{\varepsilon} \in \{\varepsilon_{min}, \varepsilon_{max}\}^n \end{aligned} \tag{18}$$

A square domain with a 20 by 20 parametric grid has been considered with $\varepsilon_{\min} = 1$ and $\varepsilon_{\max} = 2$. Firstly, the problem has been solved considering the continuous relaxation $\boldsymbol{\varepsilon} \in [\varepsilon_{\min}, \varepsilon_{\max}]^{N_{el}}$ through the Adjoint method. In a very similar way to the bandgap problem and analogously derived to that case as in [9], the Adjoint method provides us with the sensitivities or gradient and thus the direction to take at each iterate. We can then pick a small enough step size, take the step, and iterate until we reach the final optimal and feasible solution determined by the volume constraint. Such a constraint can also be dualized into the objective function and one can instead solve the new optimization problem with the modified objective:

$$J(\mathbf{u}(\boldsymbol{\varepsilon}),\boldsymbol{\varepsilon}) = ||\mathbf{u}(\boldsymbol{\varepsilon})||_2^2 + \lambda\left(\sum_{q=1}^{N_{el}} \frac{\varepsilon^q - \varepsilon_{\min}}{\varepsilon_{\max} - \varepsilon_{\min}} - \beta N_{el}\right) \tag{19}$$

and the original set of constraints. Both strategies lead to the same solutions.

Note that for the homogeneous case with $\boldsymbol{\varepsilon} = \mathbf{1}$ the analytical solution $\mathbf{u} = \sin(\pi x)\sin(\pi y)$ to (17) provides a squared volume of $J(\mathbf{u}(\mathbf{1})) = 0.5$, whereas if we pick the homogeneous material with $\boldsymbol{\varepsilon} = 2$ the objective drops down to the value $J(\mathbf{u}(\mathbf{2})) = 0.0625$, which would be the optimal solution had not we considered the volume constraints.

Results have been computed for $\beta = 0.44$ and $\beta = 0.58$ and are shown in Figure 2. We can observe how the continuous optimal solutions provide a non-binary solution that after projection into $\boldsymbol{\varepsilon} = \{1,2\}^{N_{el}}$ and respecting the volume constraint, the resulting solution is suboptimal. In fact, for the case $\beta = 0.44$ the optimal objective value is $J(\mathbf{u}(\boldsymbol{\varepsilon}_{cont})) = ||\mathbf{u}||_2^2 = 0.0973$ in the continuous case, and once projected it increases to $J(\mathbf{u}(\varepsilon_{proj})) = 0.1020$. We can do better, as our binary optimum demonstrates, obtaining $J(\mathbf{u}(\boldsymbol{\varepsilon}_{bin})) = 0.0991$. Table 2 summarizes the different values obtained for each case.

Note how the binary optimum is more than 2.5 times closer to the continuous optima than the projected naive solution for $\beta = 0.44$ and more than 5.5 times closer for $\beta = 0.58$. With a given material allowance we have thus been able to provide a binary pattern, easily fabricable, which is very close to the non-fabricable continuous optimum for the same volume of material.

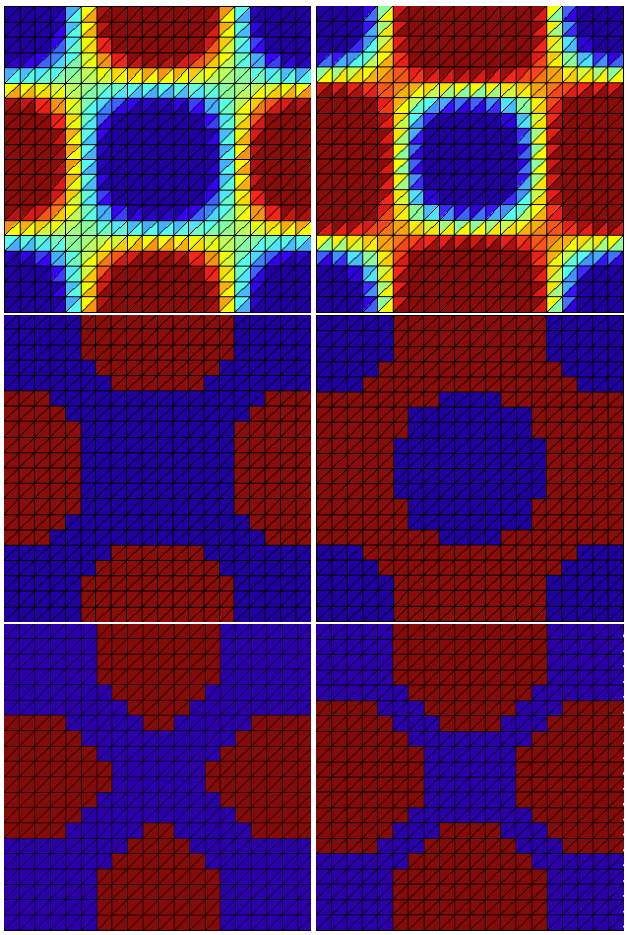

Figure 2: 2-dimensional heat transfer problem. Left figures correspond to $\beta = 0.44$ and right figures to $\beta = 0.58$. Top figures show the continuous solutions obtained through the Adjoint method, middle figures the projected binary solutions and bottom the the binary solutions after using our proposed method.

Moreover, regarding this last example, the computational cost is $\mathcal{O}(k \cdot m \cdot s_n)$, where k is the number of iterations, $m = \mathcal{O}(n)$ the size of the neighborhood and $\mathcal{O}(s_n)$ is the cost of solving the numerical problem for a given combination of the n pixels, in contrast to the NP hard cost $\mathcal{O}(\exp(s_n))$ that would provide the full binary search. For the second example, due to the symmetries, the number of pixels is $n = 50$ and there $k \simeq 20$ iterations, if we just change 1 pixel at a time. Moreover, for every single system resolution, stated above as $\mathcal{O}(s_n)$, a classic FE approach would take $\mathcal{O}(n^2)$ and we do that $m = \mathcal{O}(n)$ times, so $\mathcal{O}(n^3)$, whereas the Reduced Basis approach solves the exact system for a few neighbors $r << n$, so $\mathcal{O}(r \cdot n^2)$ and then uses that basis to infer an approximation of the FE so-

Table 2: Results for the Heat transfer problem

β	J_{cont}	J_{proj}	J_{bin}
0%	0.5	0.5	0.5
44%	0.0973	0.1020	0.0991
58%	0.0825	0.0875	0.0834
100%	0.0625	0.0625	0.0625

lution through a small $r \times r$ system $\mathcal{O}(m \cdot r^2)$. So, since $r = \mathcal{O}(1)$, the RB method computational cost can be casted as $\mathcal{O}(n^2)$.

5. Conclusions

The adjoint method leads to continuous optimal solutions that provide an objective function value J_{cont} that is less than or equal to the value J_{bin}, where J_{cont} and J_{bin} denote the optimal objective values for the continuous relaxation of the problem and our binary optimization method, respectively. The binary solution is feasible to the continuously relaxed problem and therefore the inequality above follows. However, continuous solutions are not binary and thus difficult to be fabricated. If we project the continuous solution we may obtain inferior solutions. Both examples in this paper show that the projected binary optimum is indeed inferior in practice. Our proposed method is able to compute efficiently and accurately good binary optima.

This paper presents a different approach to the design of materials involving PDE-constrained optimization. The HDG method allows us not only to obtain high order solutions of the underlying PDE. The reduced basis method allows for a rapid solution of neighbors and demonstrates the practicality of our concept of a binary generalized gradient. This extension of the gradient concept lets us move only within feasible binary solutions while improving the objective function value. There is no guarantee that we will reach a global optimum, but that is something that can just not be expected in discrete optimization if we seek efficient algorithms.

Acknowledgement

J. Saa-Seoane would like to thank the 'LaCaixa' fellowship program for the generous support. This research has been funded by AFOSR Grant No. FA9550-08-1-0350 and AFOSR Grant No. FA9550-11-1-0141.

References

[1] Nguyen, N. C., Peraire, J. and Cockburn, B. "High-order implicit hybridizable DG methods for acoustics and elastodynamics," *J. C. Phys.*, Vol. 230, No. 10, 3695–3718, 2011.

[2] Nguyen, N. C., Peraire, J. and Cockburn, B. "Hybridizable discontinuous Galerkin methods for the time-harmonic Maxwells equations," *J. C. Phys.*, Vol. 230, No. 19, 7151–7175, 2011.

[3] Schurig, D., Mock, J.J., Justice, B.J., Cummer, S.A., Pendry, J.B., Starr, A.F. and Smith, D.R "Metamaterial Electromagnetic Cloak at Microwave Frequencies," *Science.*, Vol. 314, No. 5801, 977–980, 2006.

[4] Men, H., Nguyen, N.C., Freund, R.M., Parrilo, P. and Peraire, J."Bandgap optimization of two-dimensional photonic crystals using semidefinite programming and subspace methods," *J. Comp. Phys.*, Vol. 229, No. 10, 3706–3725, 2010.

[5] Barrault, M., Maday, Y., Nguyen, N.C., Patera, A.T. "An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations," *J. Comp. Phys.*, Vol. 229, No. 10, 3706–3725, 2010.

[6] Daraio, C., Ngo, D., Nesterenko, V.F., Fraternali, F. "Highly nonlinear pulse splitting and recombination in a two dimensional granular network," *P. Rev. E* , Vol. 82, No. 036603, 2010.

[7] Cockburn, B., Gopalakrishnan, J., Lazarov, R. "Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems.," *SIAM J. Numeri. Anal.*, Vol. 47, 1319-1365, 2009.

[8] Lakes, R. "Foam structures with a negative Poisson's ratio." *Science.* Vol. 235,1038-1040, 1987.

[9] Saa-Seoane, J. "Simulation and Design Optimization for Linear Wave Phenomena on Metamaterials." *MIT Masters Thesis*, 2010.

[10] Antil, H., Heinkenschloss, M., Hoppe, R. H. W. and Sorensen, D. C. "Domain decomposition and model reduction for the numerical solution of pde constrained optimization problems with localized optimization variables." *Comput. Vis. Sci.* Vol. 13, 249-264, 2010.

[11] Boyaval, S., Le Bris, C., Lelivre, T., Maday, Y., Nguyen, N. C. and Patera, A. "Reduced basis techniques for stochastic problems." *Arch. Comput. Method. E.*. Vol. 17,435-454, 2010.

[12] Rayleigh, J. W. S. "On the remarkable phenomenon of crystalline reflexion described by Prof. Stokes", *Phil. Mag* Vol. 26: 256265, 1888.

[13] Bendsoe, M. P., Sigmund, O. "Material interpolation schemes in topology optimization", *Arch. of Applied Mechanics* Vol. 69:635-654, 1999.

[14] Jensen, J. S., Sigmund, O. "Systematic design of phononic band-gap materials and structures by topology optimization", *Phil. Trans. R. Soc. Lond. A* Vol. 361:1001-1019, 2003.

[15] Sethian, J. A., Wiegmann, A. "Structural boundary design via level set and immersed interface methods", *J. of Comp. Phys.* Vol. 163, Is. 2:489-528, 2000.