

# EE 364B Project Midterm Report: Fast Eigenvalue Optimization for Spectral-Element PDEs

John Sholar and Guillermo Angeris

June 5, 2018

## 1 Goal

Eigenvalue optimization is a common task in many control problems [BEGFB94], inverse design, and, recently, in inference problems on continuous spaces. In this project, we'll attempt to write a fast solver for eigenvalue problems arising from spectral element discretizations of eigenvalue partial differential equations (PDEs).

## 2 Motivation

Consider the following problem, which comes up in several different contexts (as an inverse-design problem for example, though an equivalent problem can also come up as the dual of a bundle method approximation, with an appropriate regularizer). Let's say we wish to manufacture a quantum, one-dimensional device that maximizes the ground state energy of a single, non-interacting particle of mass  $m$  within the device. We provide a basis of potentials for this problem  $\{V^i\}$ , where  $V^i : \mathbf{R} \rightarrow \mathbf{R}$ , such that all potentials of interest for the design can be written as a convex combination of these potentials. Then, the resulting Schrödinger PDE with positive semi-definite form<sup>1</sup> is, for some potential  $V(x) = \sum_i \alpha_i V^i(x)$  with  $\alpha_i \geq 0$  and  $\sum_i \alpha_i = 1$ ,

$$(\mathcal{H}(\alpha)\psi)(x) \equiv \left( -\frac{\hbar^2}{2m} \partial_x^2 + \sum_i \alpha_i V^i(x) \right) \psi(x) = E_0 \psi(x).$$

Here,  $\mathcal{H}(\alpha)$  is known as the Hamiltonian (defined as the linear operator above, which is affine over the  $\alpha$ ) and  $E_0$  is the smallest eigenvalue which satisfies the PDE, which we wish to maximize.<sup>2</sup>  $\psi$  is the corresponding wavefunction of the solution, but the specifics of  $\psi$  will not be used further.

---

<sup>1</sup>Any form for the Schrödinger equation can be turned into a positive-semidefinite form by noting that (a)  $-\partial_x^2$  is a positive-semidefinite form (by integration by parts with either Dirichlet or vanishing boundary conditions) and that (b) we can normalize all potentials  $i \in \{1, \dots, n\}$  to have  $\inf_x V^i(x) \geq 0$  since all feasible potentials are bounded from below. This transformation of the potentials does not change the eigenvectors in question and only increases the energy by this constant.

<sup>2</sup>That such an  $E_0$  exists follows from the fact that  $\mathcal{H}(\alpha)$  is compact (i.e. maps bounded sets to precompact sets) and self-adjoint (correspondingly, in the finite case, symmetric).

We can phrase this problem of minimizing the energy as the following optimization problem over a functional space,<sup>3</sup>

$$\begin{aligned} & \underset{\alpha, \psi, E_0}{\text{maximize}} && E_0 \\ & \text{subject to} && \mathcal{H}(\alpha)\psi = E_0\psi, \\ & && \sum_i \alpha_i = 1, \\ & && \alpha_i \geq 0 \quad \forall i. \end{aligned}$$

Noting that, in general, the variables  $\psi$  and  $E_0$  are completely determined by given values of  $\alpha_i$ , so finding an optimal set of  $\alpha_i$  is enough.

In this form, it is not clear the problem is convex, but (in a similar way to finite spaces), let  $H$  be the Hilbert space in question, with  $\mathbb{I} : H \rightarrow H$  being the identity operator and, for some linear operator  $A : H \rightarrow H$ , defining

$$A \succeq 0 \iff \langle \phi, A\phi \rangle \geq 0, \forall \phi \in H,$$

then we can rewrite the problem as

$$\begin{aligned} & \underset{\alpha, E_0}{\text{maximize}} && E_0 \\ & \text{subject to} && \mathcal{H}(\alpha) \succeq E_0 \mathbb{I}, \\ & && \sum_i \alpha_i = 1, \\ & && \alpha_i \geq 0 \quad \forall i, \end{aligned}$$

which is just optimization over a cone.<sup>4</sup>

This is essentially an SDP over the space of linear operators mapping  $H \rightarrow H$ . Assuming any (finite and reasonable)<sup>5</sup> discretization scheme, then  $\partial_x^2 \mapsto L$ ,  $V^i \mapsto D^i$ ,  $\mathbb{I} \mapsto \bar{I}$  with all matrices in  $\mathbf{S}^n$ , allowing us to rewrite the problem as a usual SDP:

$$\begin{aligned} & \underset{\alpha, E_0}{\text{maximize}} && E_0 \\ & \text{subject to} && \left( -\frac{\hbar^2}{2m}L + \sum_i \alpha_i D^i \right) \succeq E_0 \bar{I}, \\ & && \sum_i \alpha_i = 1, \\ & && \alpha_i \geq 0 \quad \forall i, \end{aligned}$$

where we have plugged the above discretizations into the original definition of  $\mathcal{H}$ .

---

<sup>3</sup>Since the true problem will be solved in the finite case, we don't concern ourselves with the properties of this functional Hilbert space and assume most of the functions are well-behaved enough for our purposes. For the interested reader, just consider all functions in question (i.e., potentials) to be  $C^\infty$ . By some basic PDE results, this implies that the solution to the linear PDE above is well-defined and also in  $C^\infty$ .

<sup>4</sup>It remains to be verified that  $E_0 = \min_{\phi \in H} \langle \phi, \mathcal{H}(\alpha)\phi \rangle / \langle \phi, \phi \rangle$ , but this is a standard result.

<sup>5</sup>*Finite* means that all operators are in  $\mathbf{R}^{n \times n}$ , while *reasonable* means that  $D^i, \bar{I}$  are diagonal while  $L$  is symmetric.

The particular structure of the matrices  $L$ ,  $\bar{I}$  and  $D^i$  generated by the SEM discretization makes this problem computationally feasible to optimize with the large domain sizes necessary for these PDEs to be accurate.

Note that we can view the first constraint to be a cone constraint on a linear operator which depends affinely on  $\alpha$ , call this discretized operator  $\mathcal{L} : \mathbf{R}^n \rightarrow \mathbf{S}^m$ , while the rest of the constraints can be written as affine equalities with a nonnegativity constraint (roughly the same as the LP standard form). This results in the following problem:

$$\begin{aligned} & \underset{\alpha, E_0}{\text{maximize}} && E_0 \\ & \text{subject to} && \mathcal{L}(\alpha) \succeq E_0 \bar{I}, \\ & && C\alpha = d, \\ & && \alpha_i \geq 0 \quad \forall i, \end{aligned}$$

for some  $C \in \mathbf{R}^{m \times n}$  and  $d \in \mathbf{R}^m$ . This is the general problem we consider, with  $\mathcal{L}(\alpha)$  having the structure of interest for every  $\alpha$ .

### 3 Spectral Element Methods and Laplacian Construction

A *spectral element method* (SEM)<sup>6</sup> is a finite-element method for solving PDEs which makes use of high-degree polynomials to approximate functions on a compact domain. The idea is to construct a quadrature rule such that all polynomials of lesser degree have exact integrals when sampled at a particular set of points and use these polynomials as a basis to approximate functions.

In general, these SEM-discretized problems—while sparse—are usually slow to solve in modern optimization packages for even a modest number of points in the discretization domain as they lack the tridiagonal structure that many PDE problems have (e.g., when discretizing by assuming  $h^2 f''(x) \approx f(x+h) - 2f(x) + f(x-h)$  for fixed  $h > 0$ ). In particular, the matrix generated by the spectral elements method is essentially a block-diagonal matrix with small blocks, each of which overlap with the previous block by exactly one element (see Figure 1), so no completely trivial decomposition can be applied.<sup>7</sup>

Our project will take advantage of the structure of these approximately-block-diagonal matrices for fast solutions of eigenvalue problems.

### 4 Approaches

Let  $\mathcal{L} : \mathbf{R}^n \rightarrow \mathbf{S}^m$  be affine and have the approximately-block-diagonal structure described in figure 1 and let  $x \in \mathbf{R}^n$ , then the problem of interest can be written as the dual of a standard-form

---

<sup>6</sup>Note that this is a class of methods, each generated by different polynomials or quadrature rules. [CHQZ06]

<sup>7</sup>Further, it's important to note that this matrix sparsity pattern is a special case of chordal sparsity [VA<sup>+</sup>15], which, while fast to solve in general, does not take advantage of the further structure which is obvious from this problem.

$A_3$

**Figure 1:** A simple example of the overlapping diagonal elements with three blocks. The darker box indicates a single overlapping element, e.g.,  $(A_1)_{nn} = (A_2)_{11}$  if  $A_1 \in \mathbf{R}^{n \times n}$ .

SDP,

$$\begin{aligned} & \underset{x, t}{\text{minimize}} && -b^T x + t \\ & \text{subject to} && \mathcal{L}(x) \preceq tI, \\ & && Cx = d, \\ & && x \succeq 0. \end{aligned} \tag{1}$$

The first inequality is with respect to the semidefinite cone, while the latter is with respect to the positive orthant.

This general optimization problem can be solved in several ways. We will attempt to explore three possibilities: interior point methods for solving these large-scale SDPs, first-order cone-splitting methods, and a consensus approach for solving these problems.

#### 4.1 Interior Point Methods

For interior point methods, rewriting the problem using a barrier method is straightforward

$$\begin{aligned} & \underset{x, t}{\text{minimize}} && -b^T x + t - \mu \log \det(tI - \mathcal{L}(x)) \\ & \text{subject to} && Cx = d, \\ & && x \succeq 0. \end{aligned}$$

If the Cholesky factorization of  $X \equiv tI - \mathcal{L}(x)$  can be easily computed, ( $X = L^T L$ ), this is sufficient as  $\log \det X = 2 \sum_i \log L_{ii}$ , since the determinant of a triangular matrix is the product of its diagonal elements. Solving this barrier-method minimization via most methods is then straightforward.

## 4.2 First-Order Cone-Splitting Methods

In the second case, as in [ZFP<sup>+</sup>17] define  $E_i \in \mathbf{R}^{d \times m}$  to be the ‘selection matrix’ given by

$$E_i = \underbrace{\begin{bmatrix} 0_{d \times (d-1)} & \cdots & 0_{d \times (d-1)} & I_d & 0_{d \times (d-1)} & \cdots & 0_{d \times (d-1)} \end{bmatrix}}_b.$$

where  $b$  is the number of sub-blocks of  $\mathcal{L}$  and  $d$  is their dimension.

This allows us to rewrite (1) above to

$$\begin{aligned} & \underset{x, t}{\text{minimize}} && -b^T x + t \\ & \text{subject to} && \mathcal{L}(x) - tI = \sum_i E_i Z_i E_i^T, \\ & && Cx = d, \\ & && x \succeq 0, \\ & && Z_i \succeq 0, i \in \{1, 2, \dots, b\}, \end{aligned}$$

where projecting the much smaller  $Z_i$  matrices (if  $d \ll m$ , as is the case with the motivating example described in section 2) into the PSD cone is much faster than projecting  $\mathcal{L}(x) - tI$ .

## 5 Further Thoughts

One of the original ideas from the proposal was to use ADMM as a way to split the problem into disjoint minimization + consensus steps. After some amount of work, we were unable to give a simple proximal operator for the projection of the overlapping subregions. After asking Duchi and some further amount of thought, it wasn’t clear that attempting to directly hammer this out would yield any more useful insights than simply following another of the two approaches mentioned above.

## References

- [BEGFB94] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear matrix inequalities in system and control theory*, volume 15. Siam, 1994.
- [CHQZ06] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods Fundamentals in Single Domains*. 2006.
- [VA<sup>+</sup>15] Lieven Vandenbergh, Martin S Andersen, et al. Chordal graphs and semidefinite optimization. *Foundations and Trends® in Optimization*, 1(4):241–433, 2015.
- [ZFP<sup>+</sup>17] Yang Zheng, Giovanni Fantuzzi, Antonis Papachristodoulou, Paul Goulart, and Andrew Wynn. Fast admm for semidefinite programs with chordal sparsity. In *American Control Conference (ACC), 2017*, pages 3335–3340. IEEE, 2017.