Two-Dimensional Integration of Electronic Bands through Quadratic Interpolation

Nathan Lars Foulk

A senior thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Bachelor of Science

Gus Hart, Advisor

Department of Physics and Astronomy

Brigham Young University

ABSTRACT

Two-Dimensional Integration of Electronic Bands through Quadratic Interpolation

Nathan Lars Foulk
Department of Physics and Astronomy, BYU
Bachelor of Science

One important part of density-functional theory (DFT) calculations is the numerical integral of the electronic band structure. Unfortunately, this critical step of DFT simulation is the most computationally expensive, because each k-point (sampling point) requires solving a large eigenproblem. For metals, almost all of the error in the band energy integral comes from misrepresenting the Fermi surface, so the most important part of any integration technique is approximating the Fermi surface correctly. Current DFT codes approximate the bands by sampling the bands with a uniform mesh, and using each sampling point to perform a zeroth-order interpolation, approximating the area around each sampling point as a constant function. The integration of the approximated bands is therefore reduced to simple Riemann sums. This zeroth-order interpolation represents the bands very poorly, making an accurate approximation of the Fermi surface impossible. I present an integration scheme consisting of the quadratic interpolation of the electronic bands using Bezier triangles. The Fermi energy can then be continuously varied in order to best represent the Fermi surface, and thereby achieve the same accuracy with fewer k-points. I also explore further improvement by using an adaptive mesh refinement technique in those integration regions which contain the Fermi surface. Preliminary results suggest that 1 meV accuracy can be achieved using ~$10\times$ fewer k-points.

Contents

List of Figures

Chapter 1

Introduction

1.1   Motivation: Importance of Band Energy Calculations

The function representing a civilization's technological capacity has an upper bound—the materials which that civilization has at its disposal. For example, today's Digital Age would not be possible without doped silicon because of its semiconducting properties. Throughout history, eras have literally been named after the dominant material that defined their way of life, such as the Bronze Age or the Iron Age. Of course, many important factors contribute to a civilization's quality of life, but the role that materials play is not small. Today, all of Earth's naturally occurring elements and simple compounds have already been examined. Materials scientists have now turned to more complex compounds in the search for the next materials that will revolutionize our technology.

The number of possible materials is impractically large. For example, each could be composed of two, three, four, or even more elements. Changing the structure and stoichiometry of each material could drastically change its stability. It is impossible to systematically explore materials space using experimental techniques. First-principles quantum mechanical calculations, such as density function theory (DFT), also prove impractical. In order to effectively sample the large space

of theoretical materials, many in the materials science community have turned to machine learning.

Machine learning provides a potential method to find materials which have a high probability of exhibiting novel or useful properties, such as superconductivity or thermoelectricity. Nyshadham *et al.* [1] used machine learning techniques to find six viable candidates for new superalloys. The problem is that in order to have highly effective machine learning algorithms, a large amount of data is first needed for the algorithms to "learn." As of right now, some of the largest materials repositories, such as The Materials Project [2] and AFLOW [3], contain on the order of 10^6 band structures. This is not enough data to perform effective machine learning. Methods must be developed in order to accelerate these materials calculations so that powerful materials machine learning can become a reality.

Perhaps the most important property that can be calculated through DFT is the *electronic band energy*. The electronic band energy is necessary in order to determine the stability of a theoretical material, telling researchers whether or not it is possible to synthesize it. In order to calculate the band energy, a numerical integral of the electronic bands must be performed. A treatment of the electronic bands will be given in Section 1.2.

The band energy integral is not only incredibly important, it also requires a lot of computational resources. The majority of computational cost associated with DFT calculations comes from the electronic band energy integral. Therefore, the most effective method to speed up materials calculations is to speed up the band energy integral. For my senior thesis, I improved numerical methods used in DFT for calculating the band energy. This work thereby shows promise of speeding up materials calculations, enabling the future use of machine learning to accelerate materials discovery.
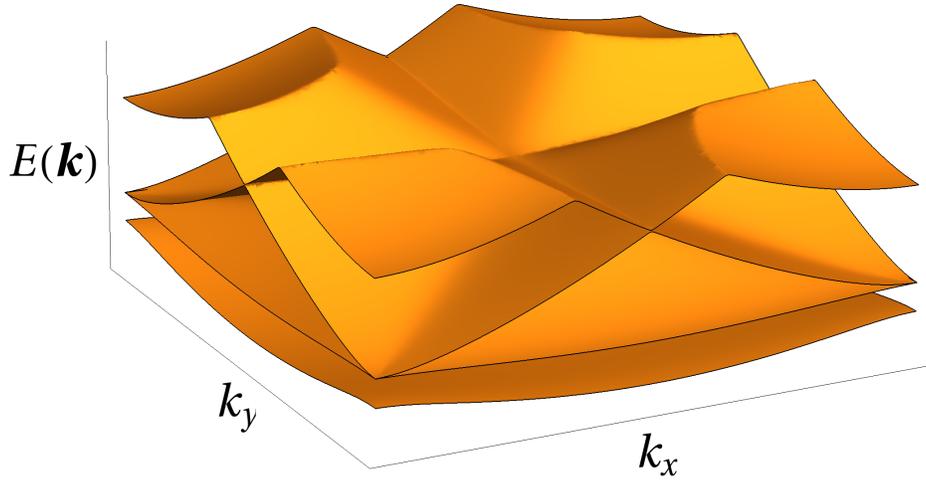
Figure 1.1 Example of electronic bands in two dimensions

1.2 Electronic Band Theory

In this section we introduce the concept of *electronic bands* and other important parts of the nomenclature necessary to understanding the work of this thesis. Electronic bands are formed by solving for the energies of time-independent Schrodinger equation

$$H_{\mathbf{k}}\psi = E_{\mathbf{k}}\psi \tag{1.1}$$

where $H_{\mathbf{k}}$ represents the Hamiltonian operator for some electronic state $\mathbf{k}$ and ϕ represents the exact wavefunctions. To provide a numerical approximation to this equation, we expand our wave functions in a basis of n plane waves, yielding

$$\mathsf{H}_{\mathbf{k}}\phi = E_k\phi \tag{1.2}$$

where H is the $n \times n$ matrix approximation of the Hamiltonian operator in the finite basis, and ψ is our approximation of ϕ in the finite basis. Equation (1.2) is an eigenvalue problem, with n eigenvalues, or eigenenergies.

As $\mathbf{k}$ varies, so do $\mathsf{H}_{\mathbf{k}}$ and E_k. Letting $\mathbf{k}$ be a continuous variable (this is an approximation—there are many states, but not infinite), we now refer to these eigenvalues as a continuous, multivalued
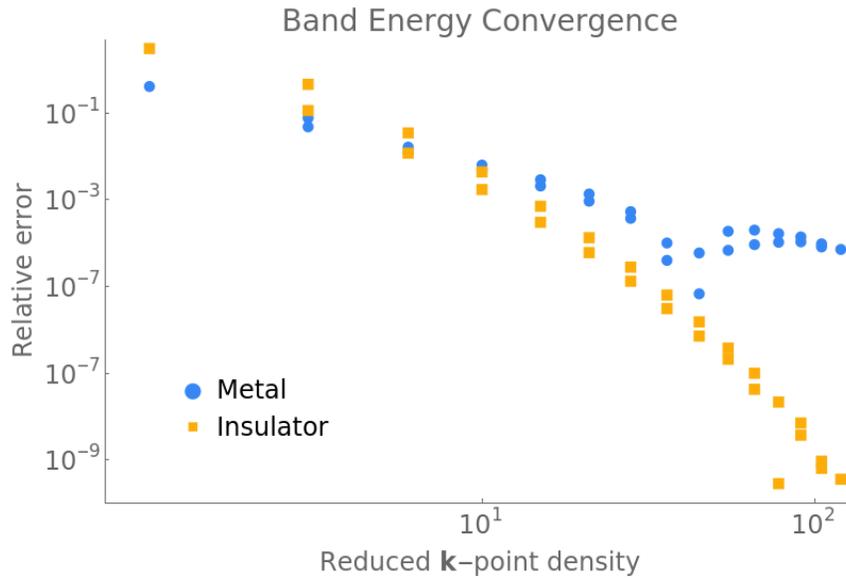
Figure 1.2 An example of slow band energy convergence for metals

function of $\mathbf{k}$, $E(\mathbf{k})$. The term *electronic bands* refers to this multivalued function, with each level being a single band. An example of these bands is seen in Fig. 1.1. The integration of these bands is the band energy, a critically important value as mentioned in Sec. 1.1. This integral is the primary focus of this thesis.

The band energy integral converges much more slowly for metals than for insulators. An example of this is seen in Fig. 1.2. For these test problems, EPM1 (to be introduced in Sec. 2.1) was given two Fermi energies in order to simulate a metal and an insulator. This slow convergence of metals occurs because metals possess bands which are intersected by the Fermi energy, greatly complicating the integral.

The band structure for a material represents the different states that electrons in that material can occupy. Electrons are fermions, and therefore no two electrons can occupy the same state (meaning they cannot share the same $\mathbf{k}$, band level, and spin value). Therefore, electrons will always occupy the lowest energy state that is available. For metals, the Fermi energy is the energy required to add another electron to the material at 0 K. Put another way, the Fermi energy is the energy which serves
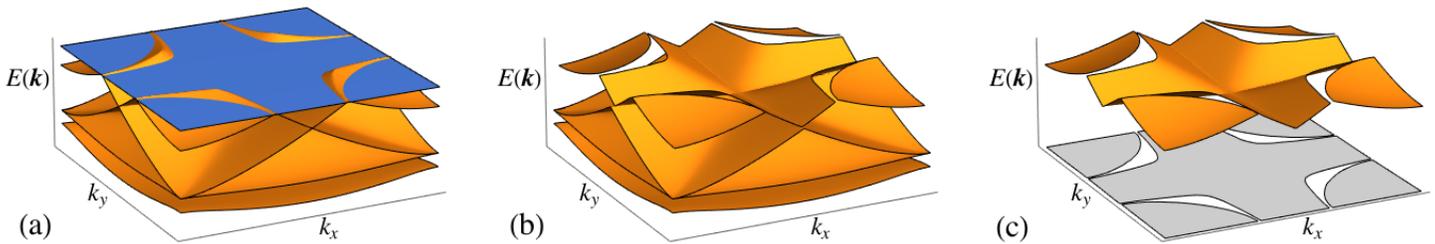
Figure 1.3 A plot showing the effect of the Fermi energy on the band energy integral in two dimensions. The Fermi energy is represented as a blue plane in (a). It "cuts the top off" of the conduction band in (b). A projection onto the *x-y* plane in (c) shows the gray regions which constitute the domain of integration for the fourth band.

as the boundary between occupied and unoccupied states.

For metals, the correct calculation of the Fermi energy is the most important step in the band energy integral. This is because the electronic band energy is the cumulative electronic energy of each electron in the material. Therefore, no unoccupied states contribute to the band energy. For this reason, the band energy is the integral of only the *occupied portions* of a material's electronic bands. In insulators, the Fermi energy lies between two bands, a band gap, and so the integral is not complicated by the introduction of a Fermi energy. However, in metals, the Fermi energy intersects with one or more bands, separating these bands into occupied and unoccupied portions, with the occupied portions being the only ones of interest for the band energy integral. A graphic showing an example of this effect is shown in Fig. 1.3.

This intersection of the bands and the Fermi energy is called the Fermi surface. This surface represents the boundaries of the occupied states. In two dimensions, the boundary is a curve. Upon calculating the distribution of errors in the band energy integral, it is apparent that an overwhelmingly large fraction of error for metals (in many cases >99%) is located at the Fermi curve, between these
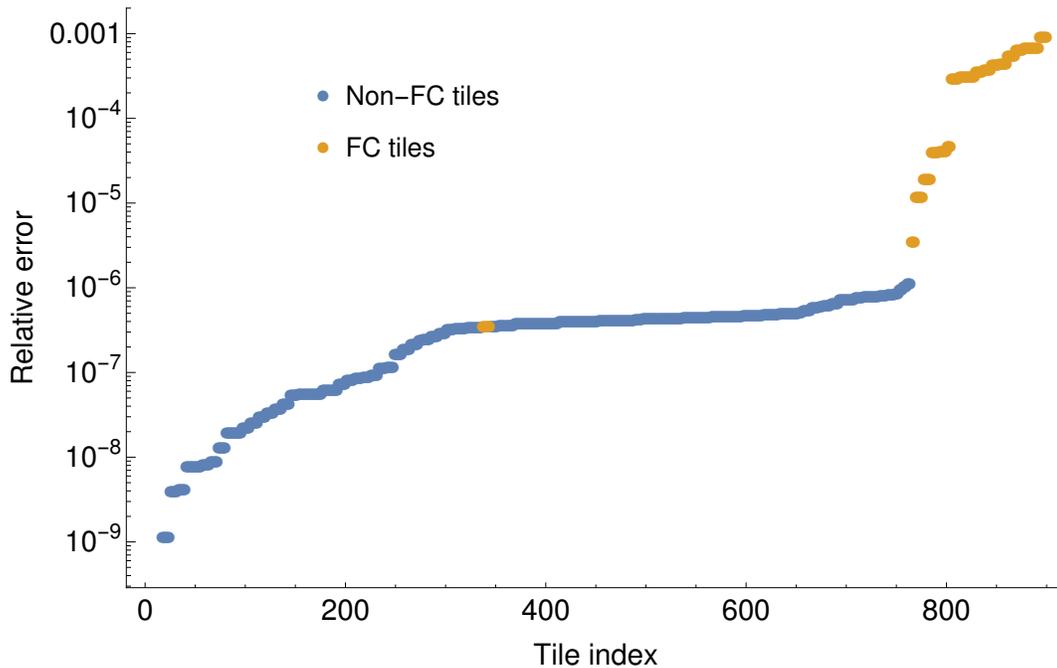
Figure 1.4 Plot of error contributions from individual tiles. Tiles which contain the Fermi curve are marked. This means that the approximation of the band structure in a marked tile intersects with the Fermi energy.

occupied and unoccupied regions.

1.3   Previous Work at BYU

The majority of research performed at BYU regarding this topic remains unpublished. Jeremy Jorgensen has pursued this topic for several years as part of his doctoral candidacy. One undergraduate student working with Jorgensen, Spencer Hart, published his senior thesis on the same area of research—numerical methods to improve band energy calculations. In his thesis, Hart evaluated a number of approaches to interpolate the electronic bands, each to no avail. These interpolation approaches included trigonometric "star functions," or symmetrized plane waves, as well as high-order splines.

His results showed that in order to use these global (rather than piecewise) interpolation tech-
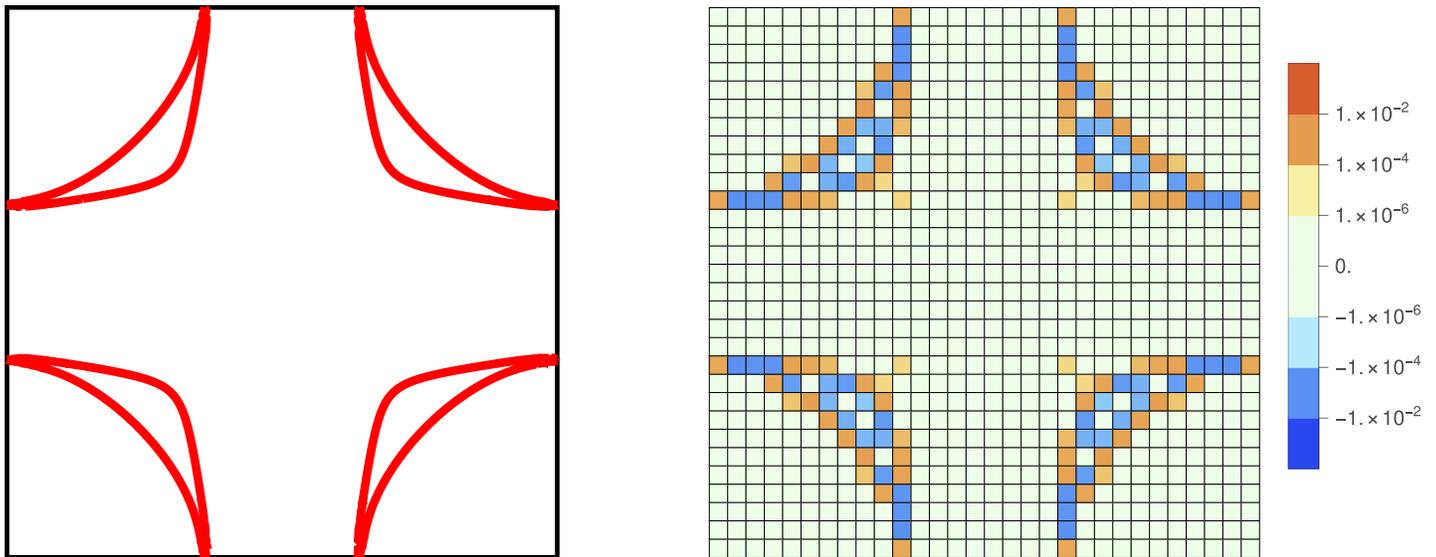
Figure 1.5 Comparison between the Fermi curve and the distribution of errors

niques for metals, one must "untangle" the electronic bands (See Fig. 1.6). Otherwise, when the bands are separated, they are nondifferentiable, and the interpolation suffers from an oscillatory ringing known as Gibbs phenomenon that most undergraduate physics majors are already familiar with. Hart's work also showed that "untangling" the bands was not possible, dooming the future viability of global interpolation techniques.

Hart's work was tested on three-dimensional materials, and his results were surprising. In order to better visualize specific challenges faced, we transitioned to working in two dimensions, thereby allowing the electronic bands to be represented in three-dimensional space, the bands being a function of two dimensions.

1.4   Background Literature

Scientists have been interested in integrating electronic band structures for a long time. Monkhorst and Pack [4] introduced the integration technique still predominantly used today. Their technique
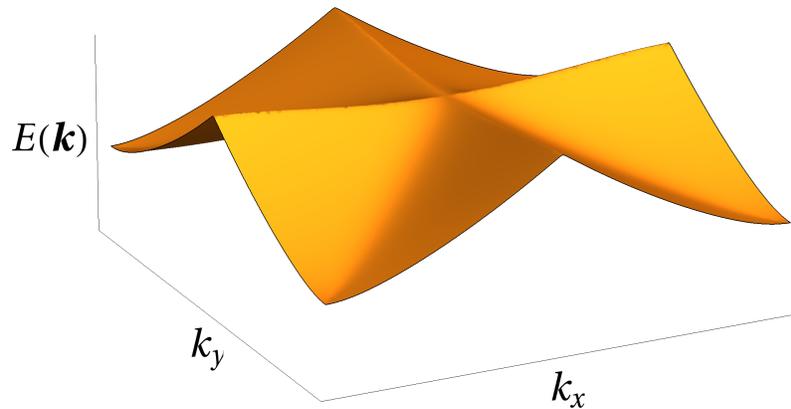
Figure 1.6 Above is the third band shown in Fig. 1.1. When the individual bands are separated, they are not differentiable.

is very simple, almost unbelievably so. The reciprocal unit cell is uniformly sampled. The band structure is then integrated using the rectangle rule, or Riemann sums. The fact that the most successful integration scheme used in modern scientific packages is the basic rectangle rule should indicate just how uniquely difficult this integral is.

Blochl [5] introduced another method based on a piecewise linear interpolation. He broke the Brillouin zone into tetrahedra, and then performed a linear interpolation within each tetrahedron. Unfortunately, this causes the Fermi surface to be consistently overestimated or underestimated quite often, leading to even worse performance than rectangles.

Weideman [6] showed that the reason Monkhorst-Pack sampling points work exceptionally well for semiconductors is because the rectangle rule works very well for continuous, periodic functions. Therefore, it is expected for rectangles to perform better than any other technique when the integrand is continuous (*i.e.*, insulators), but not when the integrand is discontinuous.

Methfessel and Paxton [7] attempted to confront the issue by using smoothing techniques to eliminate the discontinuities in the function, allowing fractional occupancy. However, although this eliminated the discontinuities, it often converges to the wrong band energy. Jeremy Jorgensen is currently preparing results that show the ineffectiveness of smoothing.

The overwhelming majority of DFT practitioners use MP meshes and the rectangle rule to integrate the band structure. In this thesis, we work to show the superiority of our method over MP meshes and the rectangle rule.

1.5   Overview of Thesis

In Chapter 2, I will introduce several important techniques in the integration algorithm presented in this thesis. These techniques include the creation of test functions, method of band representation, solving of the Fermi energy, tessellation of the domain, and method of adaptive refinement.
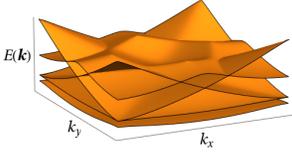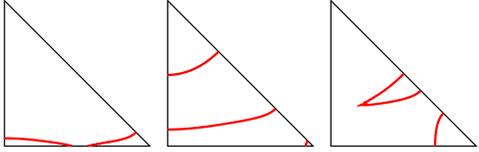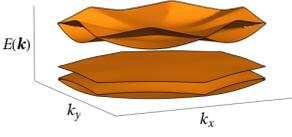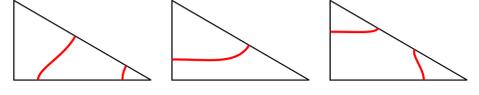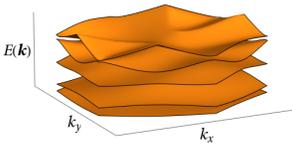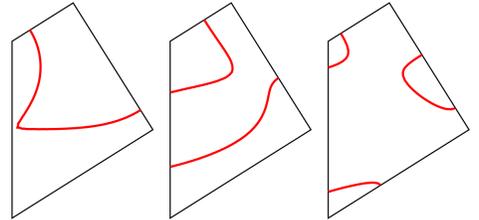
In Chapter 3, I will compare our results to contemporary band integration techniques, analyze the error distributions of both techniques, argue for the superiority of the algorithm presented in this thesis, and propose directions for future investigation.

Chapter 2

Methods

2.1 Empirical Pseudopotial Models (EPMs)

Having shown in Section 1.2 that each material can be represented using a Hamiltonian matrix, we now introduce the concept of empirical pseudopotential models (EPMs). In the free-electron limit, when the electrons don't interact with the lattice, all of the off-diagonal elements of the Hamiltonian matrix are zero. But as the Coulombic potential is felt by each atom's nearest neighbors and next-nearest neighbors, the non-diagonal elements of the Hamiltonian become *nonzero*. Each material can be simulated by choosing an array of real numbers called *interaction values* for the off-diagonal elements of the Hamiltonian. Again, these elements represent the amount of interaction between nearest neighbors, next-nearest neighbors, etc. We modify this array of real numbers to create realistic test functions for the numerical integration scheme. These models are called EPMs and they are shown in Table 2.1. In this thesis, we test four different EPMs, each with three different Fermi energies.

Model	Interaction values	Band structure	Fermi curve (FL1, FL2, FL3)
EPM1	[0.0, -0.23. 0.12]		
EPM2	[0.0, 0.39, 0.23, -0.11]		
EPM3	[0.0, -0.27, 0.2, -0.33]		
EPM4	[0.0, 0.39, -0.11, 0.11]		
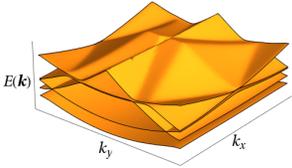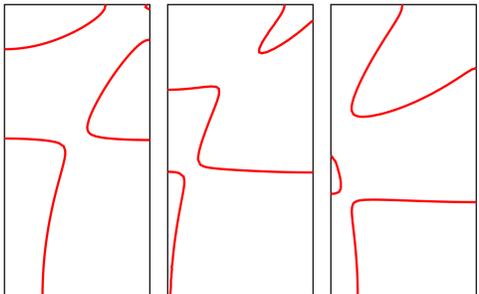
Table 2.1 A table of the models used in our tests, complete with interaction values, band structure plots, and Fermi curve plots. Note that the Fermi curves are plotted over the irreducible Brillouin zones.
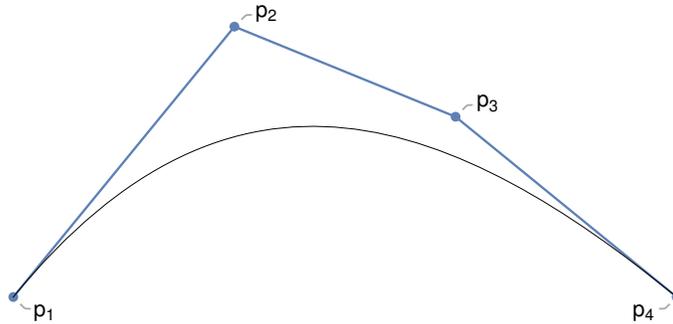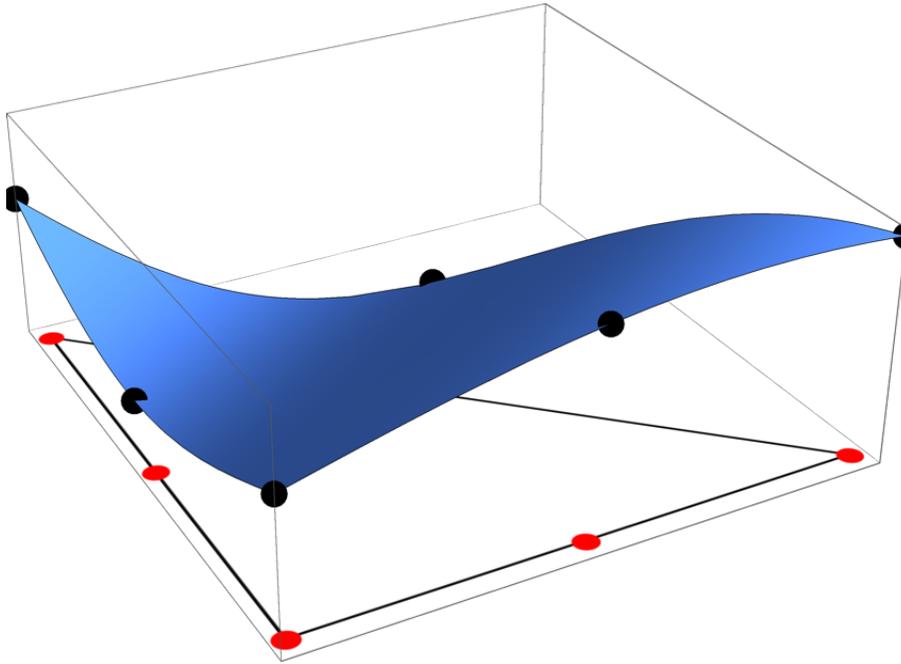
Figure 2.1 An example of a one-dimensional Bezier curve with four control points.

2.2   Bezier Representation

Bezier curves are a commonly used tool in graphic design for making "paths" in software such as Adobe Illustrator. In essence, Bezier curves are simple splines in which control points are used to create continuous curves. An illustration of this one-dimensional case is shown in Fig. 2.1.

This idea likewise extends to two dimensions, most commonly in Bezier triangles. For Bezier triangles, six control points are needed: one at each midpoint of the three triangle sides and one at each of the three vertices. Together, these points are used to perform a biquadratic interpolation. For band structure representation, it is important for the surface to intersect all six eigenvalues, so the control points on the midpoints are calculated from the eigenvalues so that the Bezier triangle intersects all six eigenvalues. The resulting interpolant is a second-order approximation of the original surface, as seen in Fig. 2.2.

Tessellating the domain with triangles allows for the band structure to be represented by this form of Bezier interpolation, assuming the eigenenergies are known at the vertices and midpoints of each triangle to serve as control points. This is the central idea of the integration technique presented in this thesis— using Bezier triangles to provide a second order, rather than zeroth order, approximation to the electronic band structure.

Figure 2.2 An example of a two-dimensional Bezier triangle with six control points.

2.3 Solving for the Fermi Level

The first (and most important) step in performing the band energy integral is finding the Fermi energy, the boundary between occupied and unoccupied states. This Fermi energy is unknown before the band structure calculation, but the number of occupied states *is* known. We can extract the best estimate for the Fermi energy, since this Fermi energy will enclose the correct number of occupied states.

When using the rectangular rule, it is very simple to calculate the approximate Fermi energy. It is simply the largest eigenvalue sampled of the N lowest eigenvalues, where N is the number of occupied states.

When using the Bezier triangle representation of the band structure, it is possible to smoothly interpolate between sampling points. Because of this, the number of occupied states changes continuously as the Fermi energy is varied between eigenvalues. This provides greater resolution in

the Fermi level calculation, but makes the calculation more involved.

To find the correct Fermi level, let the function $A(\varepsilon)$ return the number of states enclosed for a proposed Fermi energy ε. Let N be the correct number of occupied states. $A(\varepsilon)$ is monotonically increasing, and therefore, $A(\varepsilon) - N$ has only one root, namely at $\varepsilon = \varepsilon_{\text{Fermi}}$. The root of $A(\varepsilon) - N$ can be found using the bisection method. The bisection method was chosen for its simplicity, and because fast convergence in this step was not needed.

In Section 1.2, representing the Fermi curve well was listed as a necessary feature of any successful band structure integration algorithm. Representing the Fermi curve exactly is equivalent to having $A(\varepsilon)$ in its exact form. Realistically, the Fermi curve is represented imperfectly, and therefore the roots of $\tilde{A}(\varepsilon) - N$ are found, rather than $A(\varepsilon) - N$, where $\tilde{A}$ is an approximation of A. Subsequently, the result of the calculation is an approximate $\tilde{\varepsilon}_{\text{Fermi}}$ rather than the exact $\varepsilon_{\text{Fermi}}$. This inaccuracy in the Fermi energy causes the largest source of error in the band energy integral.

2.4 Tessellation of the Integration Domain

Because this integration scheme requires continuous interpolation over triangular regions of the domain, it may seem reasonable to lay down sampling points in a traditional Monkhorst-Pack [4] (MP) mesh over the reciprocal unit cell, and choosing triangles in such a way so that the MP points have the same (k_x, k_y) coordinates as the Bezier control points.

However, the reciprocal unit cell is an inferior space to sample over, because the discontinuites present in the bands at Brillouin zone boundaries create major interpolation error for the Bezier triangles. These discontinuities result from the truncated basis expansion in Sec. 1.2.

To avoid this, the next logical choice would be the Brillouin zone itself. However, this idea also has problems, since two-dimensional Brillouin zones are rarely either triangles or parallelograms. Therefore, it is not possible for our sampling triangles to be commensurate with the sampling space,
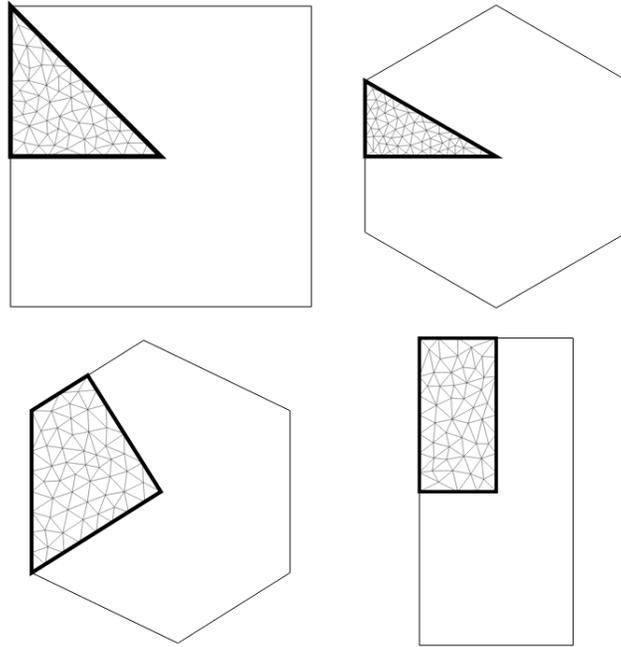
Figure 2.3 The finite-element tessellations of various IBZs. Each IBZ is outlined in a heavier weight against the backdrop of each BZ.

as is the case when Brillouin zone in the shape of a triangle or parallelogram. Because the triangles are not commensurate with the domain, the sampling grid would not possess the same symmetry as the lattice— a major problem.

Using a finite-element (FE) tessellation over the irreducible Brillouin zone (IBZ) solves this symmetry problem. It also avoids the discontinuites at the BZ boundary. Additionally, using a finite-element mesh helps to ensure that the triangles are not skewed or oriented in any particular direction. If the triangles were skewed in a particular direction, then it would be difficult to capture fine features of the Fermi curve which are oriented in the perpendicular direction. The FE tessellations for each IBZ can be seen in Fig. 2.3.
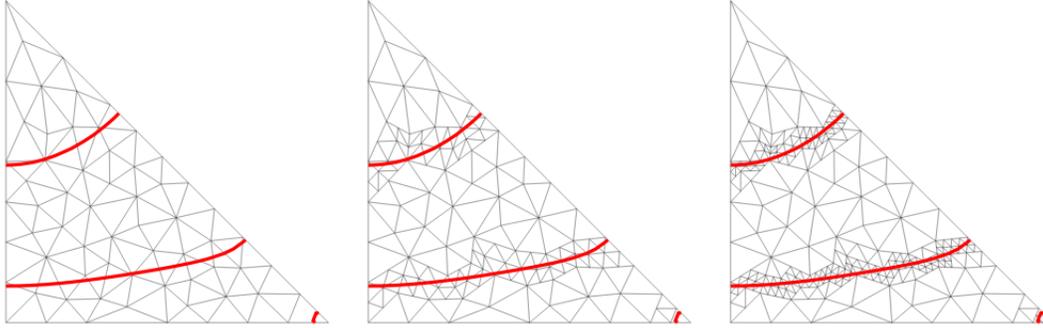
Figure 2.4 A series of tiles showing the proposed mesh refinement at the Fermi curve. The mesh is refined from left to right.

2.5 Adaptive Mesh Refinement (AMR)

It is paramount that the Fermi curve be represented correctly, because from the Fermi curve, we calculate the number of occupied states $A(\varepsilon)$, and finally $\varepsilon_{\text{Fermi}}$. Because of this importance, we introduce an adaptive mesh refinement (AMR). Using AMR, each triangle which contains a portion of the approximated Fermi curve is subdivided into four smaller triangles, requiring additional sampling. In other words, if the Bezier-interpolated approximation of the band structure intersects with the Fermi energy in a triangle, that triangle is subdivided for further sampling.

This method of sampling is much more efficient, since it improves the accuracy of the Fermi energy (the largest error contribution) without having to sample the entire domain. This way, the sampling points are placed where they matter most. An example of this AMR scheme is shown in Fig. 2.4.

Chapter 3

Results and Conclusions

## 3.1   Comparison of Integration Techniques

The practice widely used in the DFT community is that of the rectangle rule. Therefore, both of our techniques are compared to the rectangle rule. Our results are best shown using convergence plots.

First, a note on interpreting such convergence plots is in order. A convergence plot is used to show an algorithm's reliable convergence to some true value. When two algorithms are compared on the same convergence plot, the algorithm whose "best-fit line" is closer to the origin is superior. This is because, on average, the algorithm achieves the same level of accuracy using a lower k-point density than the inferior algorithm.

A summary of our results for our four EPMs, each with three different Fermi energies, over a uniform mesh is shown in Fig. 3.1. These results show that our quadratic polynomial method with Bezier triangles regularly outperforms that of the rectangle rule, because the same level of accuracy can be achieved using fewer reduced k-points (sampling points). The rectangle rules seems to have very wide spreads of convergence, and its convergence even looks like two separate datasets at times.
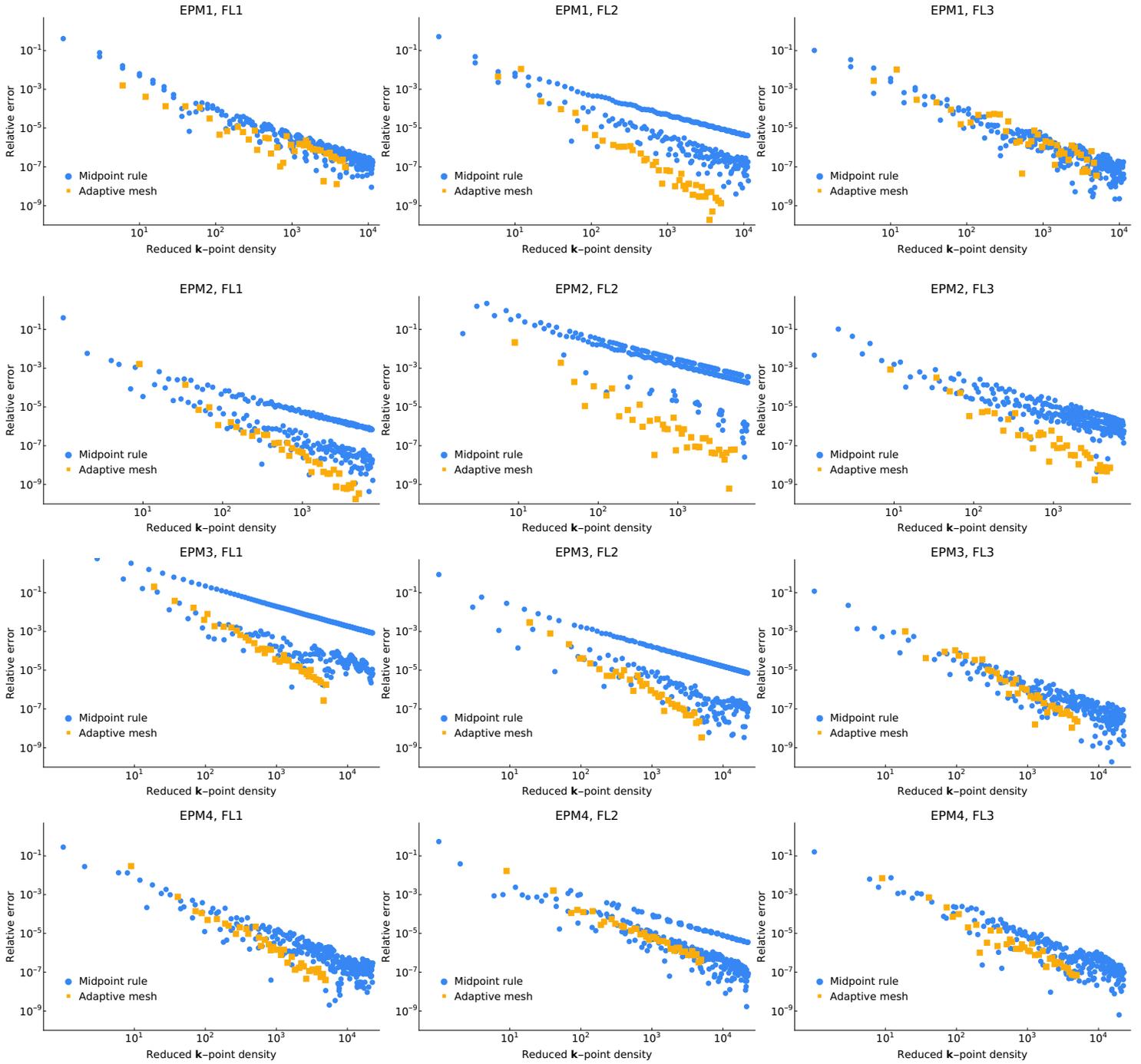
Figure 3.1 A convergence plot comparing the performance of our Bezier triangle method with a uniform mesh to that of the conventional rectangle rule.
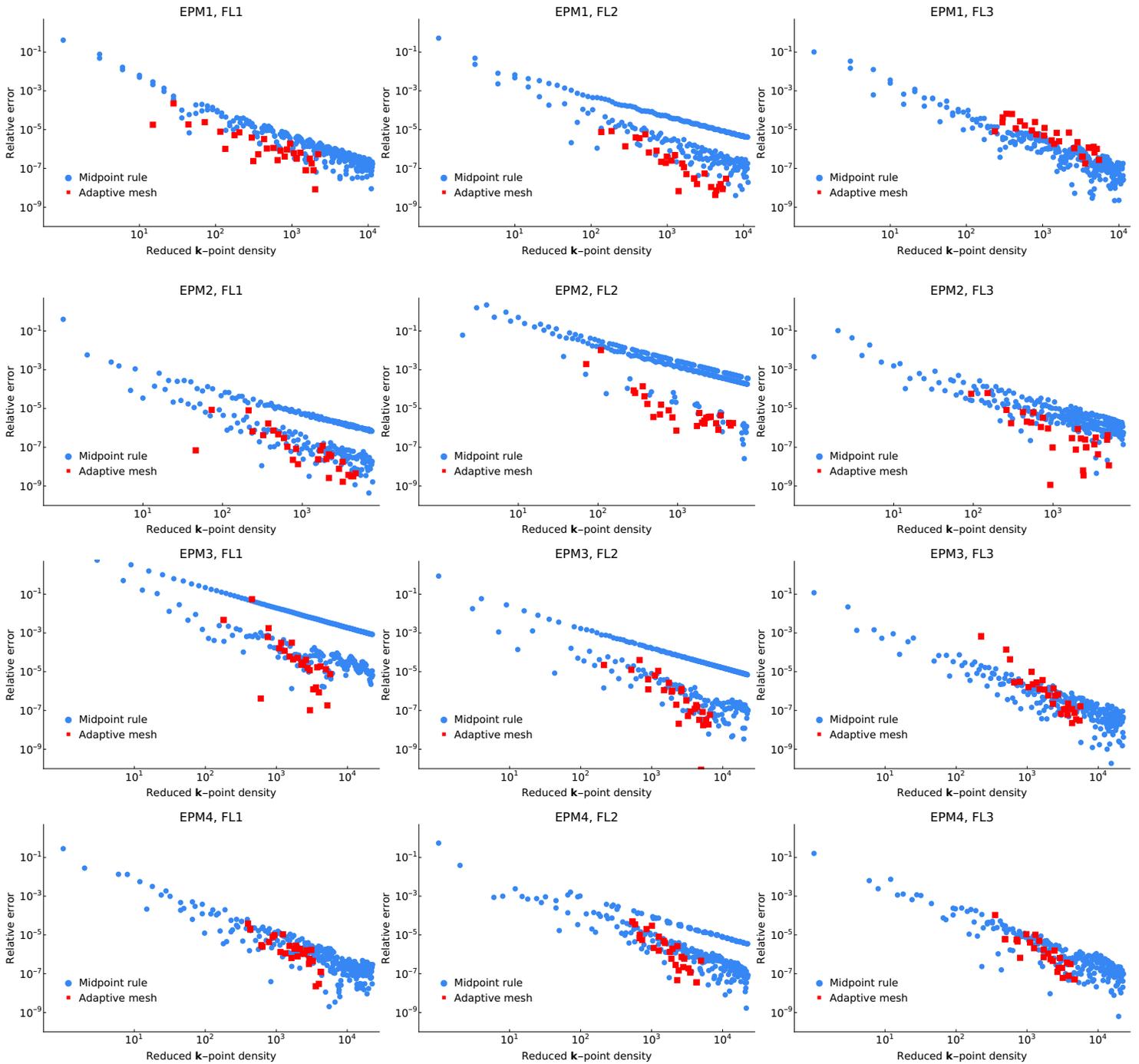
Figure 3.2 A convergence plot comparing the performance of our Bezier triangle method with adaptive refinement to that of the conventional rectangle rule.

Similarly, our results using adaptive refinement are shown in Fig. 3.2. It is clear that this AMR-based method also outperforms the rectangle rule. These margins of improvement seem small, but the scale of the axes should be considered. For several of these examples to achieve 10^{-6} relative error, the AMR-based method requires ~$10\times$ fewer *k*-points. The performance of uniform and adaptive meshes is compared directly in Fig. 3.3.

3.2   Analysis of errors

In Section 2.3, I introduced the concept of the Fermi curve, and that misrepresentation of the Fermi curve contributes over 99% of the band energy error. Again, this is because correct representation of the Fermi curve leads to an accurate calculation of the Fermi energy. Therefore, it is logical to compare the Fermi curve representations of both algorithms. A comparison of the Fermi curve representations for EPM1-FL2 is shown in Fig. 3.4.

3.3   Conclusions

For each EPM, and for each Fermi level, using quadratic interpolation over uniform FE meshes with Bezier triangles shows either improved or equal performance compared to rectangle rule methods. Adaptive refinement methods do not seem to provide further improvement. Although improvements in band energy convergence seem marginal, the convergence plots in Figs. 3.1-3.3 are log-log plots, and a superior technique often achieves the same accuracy of an inferior method using $10\times$ less sampling points.

These results have exciting applications in the materials community. Researchers that are building massive datasets for machine learning will be able to perform DFT calculations faster, due to the fewer sampling points required. Researchers investigating a single material of interest can achieve much more accurate band energy estimates using the same number of sampling points.
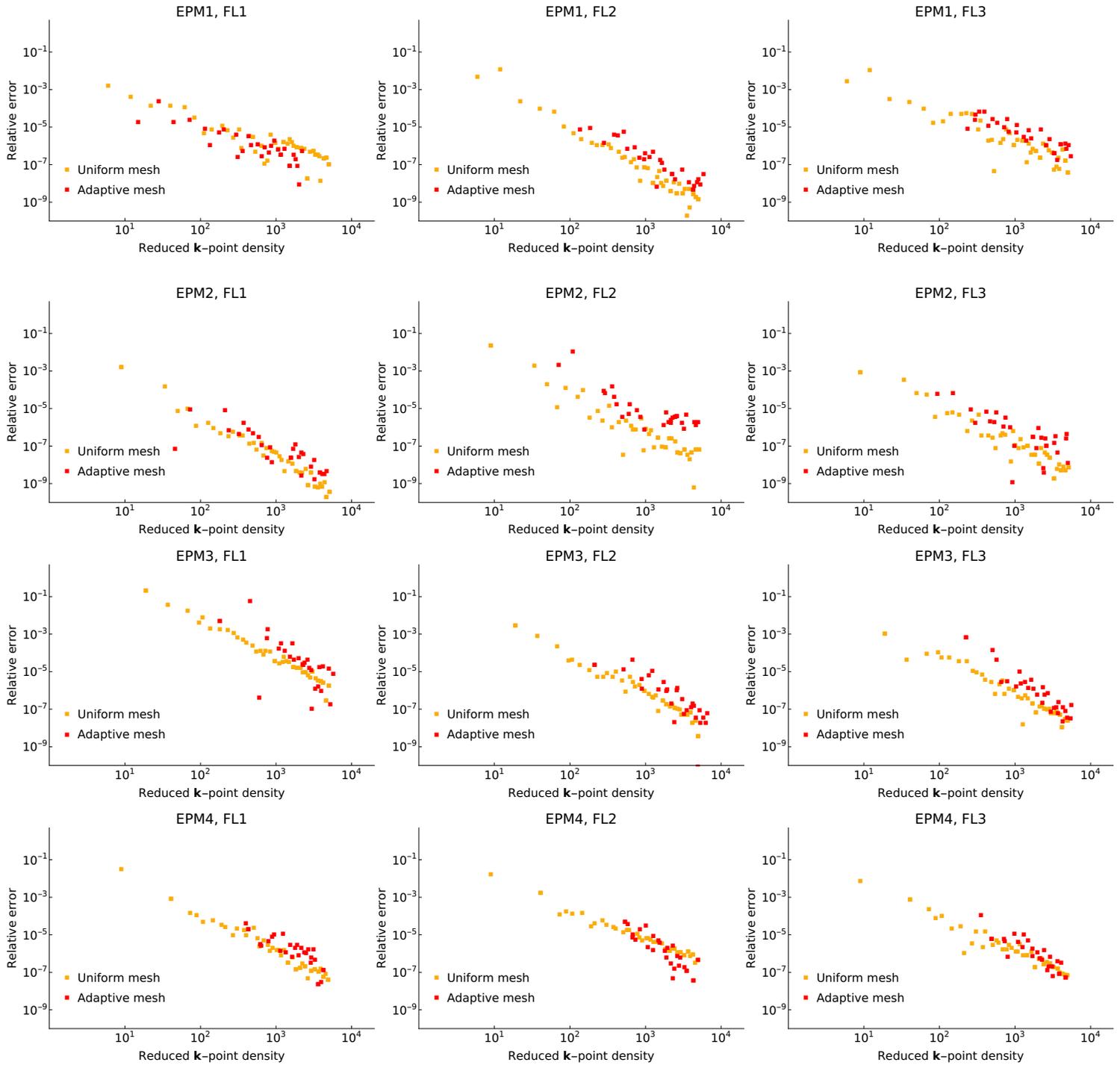
Figure 3.3 A convergence plot comparing the performance of a uniform mesh of triangles to that of an adaptive mesh.
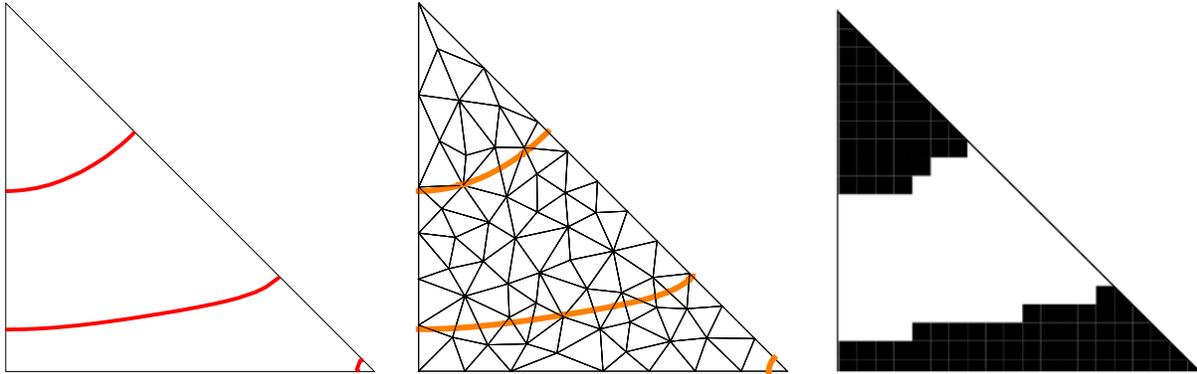
Figure 3.4 A convergence plot comparing the exact Fermi curve (left), the Fermi curve representation using our Bezier triangle method (center), and the Fermi curve representation using the rectangle rule (right).

3.4 Directions for Future Work

Despite these improved methods of integration, the band energy error is still dominated by contributions at the Fermi curve. Future investigation will be performed to discover better ways of approximating the Fermi curve. Similarly, our group will pursue adaptive meshes with more sophisticated refinement criteria that may prove more efficient in calculating the band energy.

We also look forward to extending our work to three dimensions. The sampling space will need to be tessellated with tetrahedra, and the concept of a Bezier triangle will be generalized to be a Bezier tetrahedron. The plots and figures associated with such a three-dimensional algorithm will be significantly more difficult, and that is part of the reason that this work was pursued in this thesis in two dimensions. However, once this work is extended to three dimensions, it will be ready to be programmed in a compiled language and implemented in modern DFT packages such as VASP, CASTEP, or Quantum Espresso.

Bibliography

[1] C. Nyshadham, C. Oses, J. E. Hansen, I. Takeuchi, S. Curtarolo, and G. L. Hart, "A computational high-throughput search for new ternary superalloys," Acta Materialia **122,** 438 – 447 (2017).

[2] A. Jain *et al.*, "The Materials Project: A materials genome approach to accelerating materials innovation," APL Materials **1,** 011002 (2013).

[3] S. Curtarolo *et al.*, "AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations," Computational Materials Science **58,** 227 – 235 (2012).

[4] H. J. Monkhorst and J. D. Pack, "Special points for Brillouin-zone integrations," Phys. Rev. B **13,** 5188–5192 (1976).

[5] P. E. Blöchl, O. Jepsen, and O. K. Andersen, "Improved tetrahedron method for Brillouin-zone integrations," Phys. Rev. B **49,** 16223–16233 (1994).

[6] J. A. C. Weideman, "Numerical Integration of Periodic Functions: A Few Examples," The American Mathematical Monthly **109,** 21–36 (2002).

[7] M. Methfessel and A. T. Paxton, "High-precision sampling for Brillouin-zone integration in metals," Phys. Rev. B **40,** 3616–3621 (1989).

Index