UNIVERSITY OF CALIFORNIA,

IRVINE

Quasi-Regular Grids and their Application to the Vibrational Spectra Calculations

DISSERTATION

submitted in partial satisfaction of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

Shane W. Flynn

Dissertation Committee:
Professor Vladimir A. Mandelshtam, Chair
Professor Craig C. Martens
Professor Ioan Andricioaei

2021

DEDICATION

To my parents and brother, my fiancé Zi-Yu Huang, and numerous family members. All of my previous research advisors and mentors, and all my friends and colleagues without whom this work would not have been finished.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my faculty advisor, Vladimir Mandelshtam, for his dedication to research and consistent technical advice regarding the work described herein. Also my sincere thanks to the other committee members, Craig Martens and Ioan Andricioaei for their assistance and advice throughout my time at UCI. Everyone in administration that helped me complete my studies, notably Tenley Dunn and Garrett Yoshitomi. Finally a special thank you to all my colleagues over the years for useful conversations and contributing both directly and indirectly to my work. First my various academic advisors over the years: Bill Goddard for his cavalier attitude, Steve Ackerman for convincing me to pursue research over medical school, and Jason Green for convincing me to pursue theory over experiments. Finally, Michelle Foster for her constant support over the years; not many professors care for and defend the students, but she has always been a pillar of support. My close friends from the Andricioaei group: Moises Romero, Dhiman Ray, and Trevor Gokey. My neighbors in the Furche group: Guo Chen, Saswata Roy, Brian Nguyen, and Jason Yu. Likewise Austin Martens and Dorothy Huang from the Martens group and Pavel Okun and Johnny Kozlowski from the Burke group.

A special thanks to my Caltech friends: Shyam Saladi, Chris Kenseth, Alexandra Kent, and Kewei Xu. Everyone from WAG-Land, especially Daniel Brooks, Yalu Chen, and Yufeng Huang. And finally my oldest friends: Liubo Iordanov, Mike Tierney, Alex Ross, Anthony Witherell, and Anthony Cancilla. I want to acknowledge AIP Publishing and the American Chemical Society (ACS) for permission to include and adapt previously published material from my manuscripts in Chapters 2, 3, 4, and 5. The research in Chapters 2 and 3 was supported by the National Science Foundation (NSF) Grant No. CHE-1566334 and the research in Chapters 4, 5, and 6 by NSF Grant No. CHE-1900295.

CURRICULUM VITAE

SHANE W. FLYNN

EDUCATION:

1. Doctor of Philosophy in Chemistry, University of California, Irvine (2021)
2. Master of Science in Chemistry, California Institute of Technology (2017)
3. Bachelor of Science in Biology, University of Massachusetts, Boston (2015)
4. Bachelor of Science in Chemistry, University of Massachusetts, Boston (2015)
5. Minor in Mathematics, University of Massachusetts, Boston (2015)

POSITIONS:

1. Research Data Scientist, AIOI Insurance Services USA (Nov. 2021)
2. Intern Data Scientist, AIOI Insurance Services USA (2021)
3. Graduate Researcher, Dept. of Chemistry, University of California, Irvine.
   Research Advisor: Professor Vladimir A. Mandelshtam (2017-2021)
4. Graduate Researcher, Dept. of Chemistry, California Institute of Technology.
   Research Advisor: Professor William A. Goddard (III) (2015-2017)
5. Undergraduate Researcher, Dept. of Chemistry, University of Massachusetts, Boston.
   Research Advisor: Professor Jason R. Green (2013-2015)
6. Undergraduate Researcher, Dept. of Biology, University of Massachusetts, Boston.
   Research Advisor: Professor Steve Ackerman (2010-2015)

RESEARCH PUBLICATIONS:

1. S. W. Flynn and V. A. Mandelshtam, "Molecular spectra calculations using an optimized quasi-regular Gaussian basis and the collocation method," *J. Chem. Theory Comput.*, **10.1021/acs.jctc.1c00805** - (2021).
2. S. W. Flynn and V. A. Mandelshtam, "Sampling general distributions with quasi-regular grids: Application to the vibrational spectra calculations," *J. Chem. Phys. Lett.*, **151**, 241105 (2019).
3. J. W. Nichols, S. W. Flynn, and J. R. Green, "Order and disorder in irreversible decay processes," *J. Chem. Phys.*, **142**, 064113 (2015).

4. S. W. Flynn, H. C. Zhao, and J. R. Green, "Measuring disorder in irreversible decay processes," *J. Chem. Phys.*, **141**, 104107 (2014).

PRESENTATIONS (select):

1. "Sampling general distributions with quasi-regular grids: Application to the vibrational spectra calculations": Virtual Conference on Theoretical Chemistry, Stanford University, Stanford, CA (2020)

2. "Two-Phase Thermodynamic Characterization of Polymer Electrolytes": Joint Center for Artificial Photosynthesis Research Conference, California Institute of Technology, Pasadena, CA (2016)

3. "Information in a rate coefficient": American Conference in Theoretical Chemistry, Telluride, CO. (2014)

4. "Transcription factors and the immune response in Arabidopsis Thaliana": Northeast Biochemistry Research Conference, Tufts University, Medford, MA (2012)

AWARDS AND HONORS (select):

1. "Outstanding Graduate Student Presentation": Stanford University. Stanford, CA (2020)

2. "Hypercube Annual Award in Theoretical Chemistry" University of Massachusetts, Boston. Boston, MA (2015)

3. "Hypercube Annual Award in Physical Chemistry" University of Massachusetts, Boston. Boston, MA (2014)

4. "Oracle Undergraduate Research Fellowship" University of Massachusetts, Boston. Boston, MA (2014)

5. "Chancellor's Scholarship for Excellence" Honors College; University of Massachusetts, Boston. Boston, MA (2010-2015)

ABSTRACT OF THE DISSERTATION

Quasi-Regular Grids

by

Shane W. Flynn

Doctor of Philosophy in Chemistry

University of California, Irvine, 2021

Professor Vladimir A. Mandelshtam, Chair

This work revisits a general problem in numerical analysis; the efficient sampling of any d-dimensional distribution function. Monte Carlo methods are the standard approach when addressing a general distribution function, as they are able to produce a set of points in configuration space following any distribution. However, there are apparent flaws with this approach: namely, the inevitable phenomenon commonly referred to as "gaps and islands" where some regions in configuration space are oversampled while other regions are undersampled. The well-known "curse of dimensionality" and its exponential scaling, combined with inefficient sampling, results in these methods quickly becoming unfeasible for meaningful applications. To this end we introduce a new sampling method, Quasi-Regular Grids, which results in an optimally distributed set of points that maintain local uniformity while simultaneously sampling any desired distribution. These new grids are then applied to the challenging problem of computing the quantum vibrational spectra for both model and molecular systems. We quantitatively establish the scaling properties of our method using an analytic model, the Morse potential, contrasting our results to the current best practices in the literature. We then tackle a chemical system of interest, formaldehyde, and again show the method to be superior to recent work published on the same system. Finally we present preliminary results for the spectral calculations of water. This research establishes a completely general method with numerous applications expanding far beyond spectra calculations, and will surely be an active area of further research in the future.

Chapter 1

Monte Carlo Methods for the Sampling of a General Distribution Function

The sampling problem refers to the general question of how to select a single instance, or a finite number of objects, from a given set. For example, a trivial problem would be the method used to select a single playing card from a deck. In practice the set will be much more complex (potentially infinite), and one usually wishes to represent the larger set through the collection of samples. This type of problem is ubiquitous across the scientific disciplines as the quality of the subset of samples will dictate the accuracy and insights derived from any subsequent analyses.

The particular sampling problem addressed in this research is the sampling of a general, multi-dimensional distribution function. The goal is to represent a continuous function with a discrete set of points as efficiently as possible. This is a long-standing problem in numerical analysis, and its applications range from machine learning to quantum wave packets.[1–4] A primary reason for this diversity is that sampling methods can be used for the generation of a training set for fitting (parametrizing) an expensive-to-compute (or measure) function. This is incredibly important in fields such as supervised learning where standard algorithms are susceptible to the quality of the training set.[5–7] The classification problem is a well-known example, where biased training sets result in the algorithms simply ignoring the minority classes to optimize the recall of the majority.[8]

An important example from the domain of chemical physics is the application of training sets to the fitting of ab-initio potential energy surfaces.[9–16] For a computational chemist the importance of ab-initio surfaces simply cannot be understated. For these applications it is well known that the sampling methods must efficiently explore the potential energy landscape while focusing on the physically relevant regions in configuration space. In order

to approach this problem rigorously one may attempt to quantify the "physical importance" by devising the corresponding weight (or distribution) function.

2.1 Monte Carlo Implementations

Given a general distribution function $\mathcal{P}(\mathbf{x})$ $\left(\mathbf{x} \in \mathbb{R}^d\right)$, the primitive, but most prominent methods for sampling $\mathcal{P}$ are the family of Monte Carlo methods.[17;18] Monte Carlo can be used to solve a large family of problems; most notably, stochastic problems like diffusion and transport, or deterministic problems such as integral evaluations or solutions to partial differential equations.[19–23]

Integration in particular is an important area of research for the physical sciences. Due to the exponential scaling of space, quadrature-based methods quickly become unfeasible for computing high-dimensional integrals. For this reason Monte Carlo methods are more common. By applying a pseudo-random sequence an integral can easily be approximated

$$\int_{\mathbb{R}} d\xi \; f(\xi) \approx \frac{1}{N} \sum_{i=1}^{N} \; f(x_i), \tag{2.1}$$

however, the presence of "gaps and islands" in the uncorrelated sequence results in slow (namely, as $\sim 1/\sqrt{N}$) numerical convergence. In the context of integration one can achieve a significant improvement by using low-discrepancy correlated sequences, the so-called quasi-random sequences. Quasi-Monte Carlo integration has been shown to have far superior convergence properties (numerical studies even suggest $\sim 1/N$) depending on the specific integral.[24–26]

In the context of sampling, most Metropolis Monte Carlo implementations use a pseudo-random number generator to produce a set of trial moves with an acceptance criteria that satisfies detailed balance. This produces a sequence of points distributed according to $\mathcal{P}(\mathbf{x})$. In principle, the points in such a sequence are correlated. However, one can remove the correlations by picking points separated from each other through sufficiently many Monte Carlo

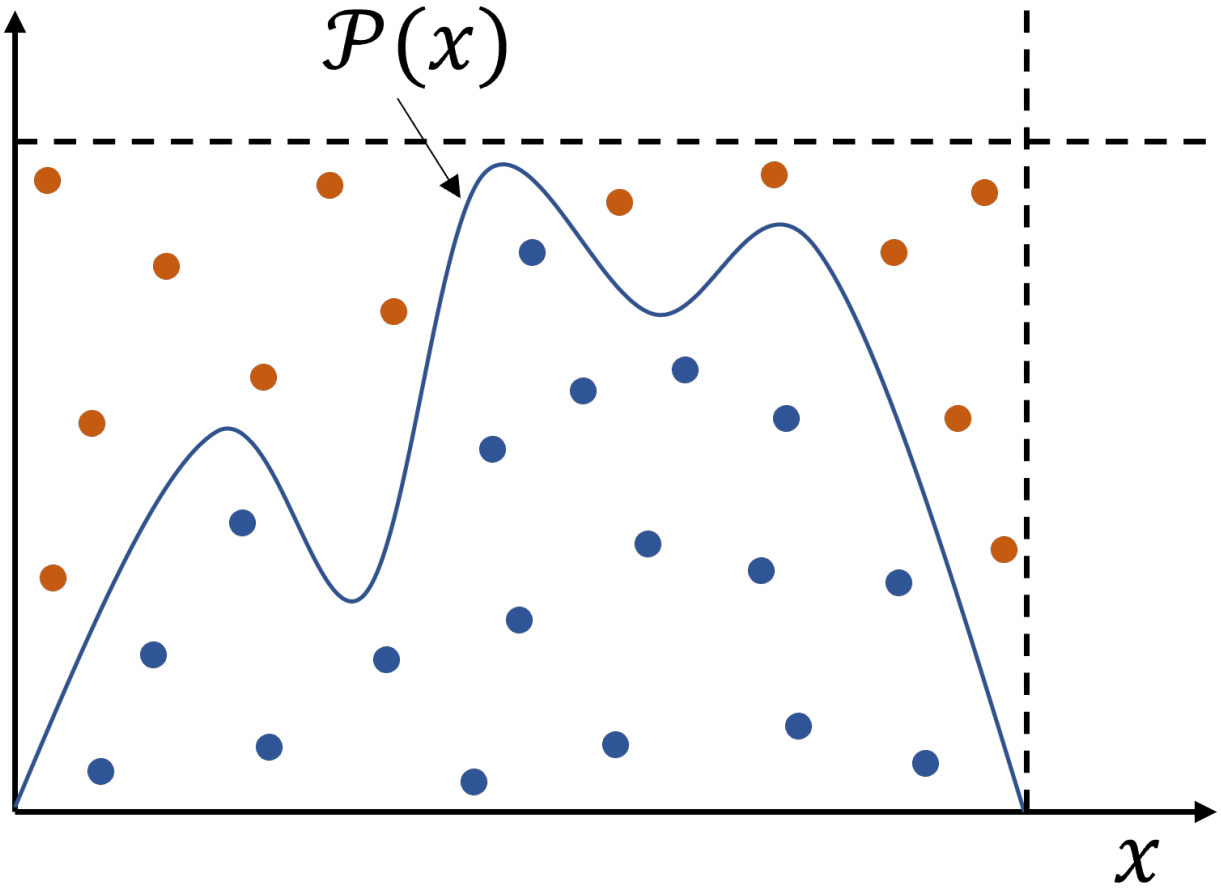

Figure 2.1: Cartoon depiction of the rejection method generating a set of points according to a desired distribution $\mathcal{P}$. Points in blue were accepted using Monte Carlo criteria, and follow $\mathcal{P}$. Those in orange are discarded.

steps. This simple procedure results in a set of independent random points in configuration space that samples the distribution in question.

2.1.1 Rejection Sampling

The applications considered in this work involve sampling from complicated distribution functions such as the vibrational Schrödinger's equation. If the distribution of interest cannot easily be sampled from, a specific Monte Carlo method known as rejection sampling is often used.[27;28] The technique is exact, and the idea behind the method is very simple as shown in Fig. 2.1.

The problem of sampling directly from a complicated distribution, $\mathcal{P}$, is replaced by drawing from a proposed distribution function, ρ, which can readily be sampled from. The method of rejection sampling requires that a constant c exists such that

$$c\,\rho \geq \mathcal{P}. \tag{2.2}$$

With these conditions one can generate a point x_i from the distribution ρ, which encompasses the target distribution $\mathcal{P}$. To ensure that the point x_i comes from $\mathcal{P}$, Monte Carlo criteria is applied resulting in an acceptance or rejection.

$$\frac{\mathcal{P}(\mathbf{x})}{\mathcal{P}_{\max}} > \xi[0,1] \tag{2.3}$$

Although the rejection process is highly inefficient as dimensionality is increased it is still commonly implemented due to ease and lack of alternate methods for general, high-dimensional distributions.[29]

2.2 Discrepancy

The use of the terms pseudo-random and quasi-random is ambiguous, however, we will continue to use them throughout this work. These are terms used in the literature and we do not want to cause confusion by changing the convention. However, in this context the term pseudo-random specifically means a sequence of points that are uncorrelated and show a high-discrepancy. The default number generators in essentially every programming language fall under the category of pseudo-random number generators. Likewise quasi-random in this context refers specifically to the locally correlated, low-discrepancy sobol sequence.

Discrepancy is a quantitative measure used to describe how evenly distributed the points in a sequence are. Qualitatively a high-discrepancy refers to a sequence where significant non-uniformity of the points occurs. In this scenario certain regions in configuration space contain a large number of points, while others are sparse. Low-discrepancy refers to the

opposite, where the points are more evenly distributed reducing the clustering problem. The discrepancy D_s^* for a d-dimension sequence $\{\mathbf{x_n}\}$ on the unit cube $[0,\,1]^d$ can be defined as:

$$D_s^*(x_1, ..., x_s) = \sup_{A \in E^*} \left| \frac{|\{x_i \in A\}|}{S} - \mathrm{Vol}(A) \right|$$

There are various different definitions for the discrepancy, and the problem of minimizing the discrepancy of a set of points in very non-trivial. In this work we are not interested in quantitative metrics for this property, but refer interested readers to the mathematics literature.[30–37]

2.2.1 Consequences of Discrepancy

The high-discrepancy sequences form a characteristic pattern in space commonly referred to as "gaps and islands". These types of sequences lead to various numerical bottlenecks such as the slow convergence problem in Monte Carlo integration discussed above. The inherently redundant nature of the pseudo-random sequences results in the need for far more points compared to a hypothetical locally uniform non-redundant sampling. While the slow convergence of integration can be addressed with low-discrepancy sequences it is unclear if this property is desired in the context of sampling. As show in Fig. 2.2, the quasi-random sequence clearly fills the region more uniformly. A much larger number of samples are required for the pseudo-random sequence to have comparable gap-sizes in the configuration space.

While in principle low-discrepancy sequences may address the sampling problem, in practice they are only easy to generate for distributions that are products of low-dimensional distribution functions, $\mathcal{P}(\mathbf{x}) = \mathcal{P}_1(\mathbf{x}_1) \times \mathcal{P}_2(\mathbf{x}_2) \ldots$. While there are a number of methods available for the sampling of specific distributions, the generalization to higher dimensions and non-trivial dependencies is still an open problem in numerical methods. This inherently limits the application of quasi-random sequences to the sampling problem. Most authors have

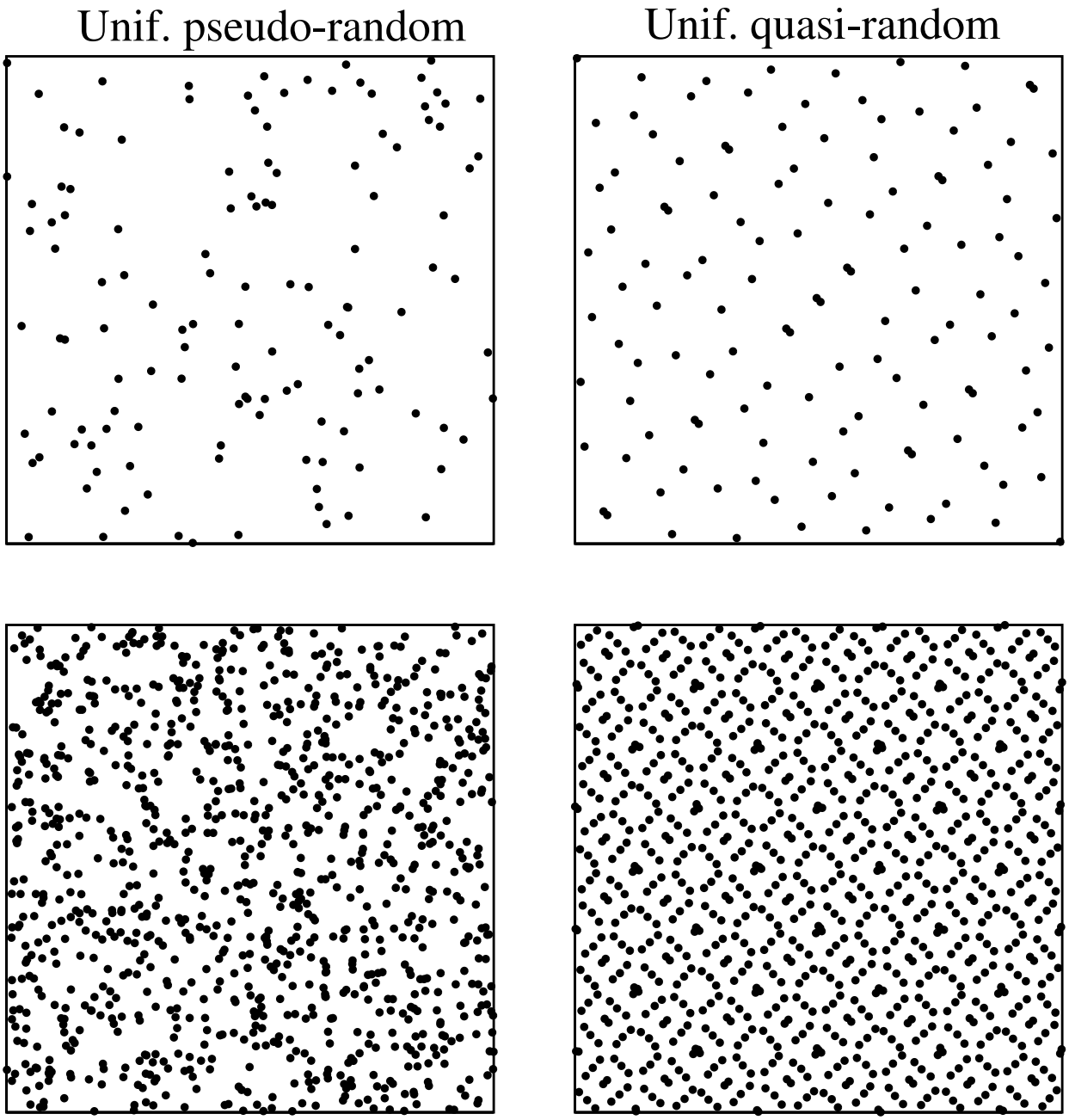

Figure 2.2: 128 and 1024 two-dimensional points from the unit square $[0,1]^2$, generated using a pseudo-random number generator (left) versus a quasi-random number generator (right). The quasi-random sequences are highly structured to uniformly fill configuration space.

explored a combination of quasi-random sequences and the rejection method to generate a set of points with quasi-random properties for a general distribution. This methodology and associated claims about the resulting sets of points will be explored throughout this work.

Chapter 2

Quasi-Regular Grids

The focus of this chapter is to introduce a new sampling method resulting in the generation of a novel nondirect product grid that we call quasi-regular sampling.[38] These grids are characterized by their local uniformity, while simultaneously sampling the proper global topology of any distribution. Simple two-dimensional model systems are used to graphically illustrate the strengths and weaknesses of currently available methods, providing an intuition for the more challenging problems faced in higher dimensions. The subsequent chapters focus on the application of these grids to study problems in the area of quantum dynamics. As such, this chapter is essential to understanding the remainder of the work.

3.1 Ansatz for a Quasi-Regular Grid

One can imagine an ideal sampling method would produce a set of points $\mathbf{x}_i \in \mathbb{R}^d$ $(i = 1, \cdots, N)$ that, on the one hand, are locally regular, i.e., sample the configuration space in the least redundant fashion, and on the other hand, are globally distributed according to an *a priori* defined distribution function $\mathcal{P}(\mathbf{x})$. To satisfy both conditions, which at first glance seem mutually contradictory, we use an embarrassingly simple approximate solution. By treating the grid points like particles interacting via a specially designed short-range pairwise potential, we can optimize the set of points by minimizing the total energy of the grid subject to this pseudo-potential.

Given a general normalized distribution function $\mathcal{P}(\mathbf{x}) > 0$, $\mathbf{x} \in \mathbb{R}^d$ $\left(\int d\mathbf{x}\mathcal{P}(\mathbf{x}) = 1\right)$, our goal is to construct a set of points $\mathbf{x}_i \in \mathbb{R}^d$ $(i = 1, \cdots, N)$, which locally have a regular arrangement, but globally are distributed according to $\mathcal{P}(\mathbf{x})$. We propose a practical solution

based on minimization of the following functional in the form of a pairwise potential:

$$\mathcal{U}(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \sum_{i=1}^{N} \sum_{j=1}^{N} u_{ij} \longrightarrow \min \tag{3.1}$$

with the short-range pair pseudo-potential defined by

$$u_{ij} = \left[\frac{\sigma(\mathbf{x}_i)}{x_{ij}}\right]^m - \left[\frac{\sigma(\mathbf{x}_i)}{x_{ij}}\right]^n \tag{3.2}$$

$$\sigma(\mathbf{x}) = \gamma \left[N\mathcal{P}(\mathbf{x})\right]^{-1/d}, \tag{3.3}$$

where $x_{ij} := |\mathbf{x}_i - \mathbf{x}_j|$. Due to the form of this short range interaction (motivated by the well-known Lennard-Jones potential), the points $\mathbf{x}_i$ are expected to arrange themselves locally to resemble a quasi (i.e. not quite perfect) closed packed structure. In the traditional Lennard-Jones potential σ is not a function of the distribution, resulting in a uniform distribution of particles. For a quasi-regular grid, the functional form of $\sigma(\mathbf{x})$ is key. Its role is to define the distance between the nearest neighbors, in accordance with the local density of points $\mathcal{P}(\mathbf{x})$. While the power law in Eq. (3.3) is intuitively obvious, the precise value of the constant $\gamma \sim 1$ is not, and as such, it becomes an adjusting parameter. A γ value that is too small would result in disconnected clusters of points, while a value that is too large would force the points to form regular structures with uniform spacing across large domains of configuration space. That is, it is the "defects", i.e. lattice imperfections, that allow for the desired gradual variation of the spacings between nearest neighbors in accordance to the variation of the density $\mathcal{P}(\mathbf{x})$. We find our construction to be robust, meaning there is a comfortable range for γ where the behavior of the grid is insensitive to its value. In practice, the value of γ can be optimized by comparing low-order moments of $\mathcal{P}(\mathbf{x})$, such as the mean and the covariance matrix computed by

$$\langle \mathbf{x} \rangle = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i; \quad \langle \mathbf{x}\mathbf{x}^{\mathrm{T}} \rangle = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^{\mathrm{T}}, \tag{3.4}$$

with their exact values. The necessity of computing the moments for ensuring quasi-regularity is revisited in Section 5.3.

In this work we will be using the $(6, 12)$ Lennard-Jones potential: $(n, m) = (6, 12)$. In general, the choice for the form of u_{ij} is flexible and does not play a crucial role in dictating the optimized grids properties. However, in order to maintain a short-range interaction for higher dimensionality, larger values for n and m must be used. For example, $(n, m) = (3 + d, 9 + d)$ would be a reasonable choice for a general implementation that scales with dimensionality appropriately.

The flexibility in the form of the pair-wise potential is explored in Chapter 5, where the ansatz is revisited and a more efficient algorithm is developed due to a convenient choice for the pseudo-potential. Notably this improved ansatz removed the need for normalization and the need for the fitting parameter γ.

3.2 Two-Dimensional Normal Distribution

To demonstrate not only our new method, but its competitors in the literature, we use the two-dimensional standard normal distribution as a model system

$$\mathcal{P}(\mathbf{x}) = (2\pi)^{-d/2} \exp\left\{-\frac{\mathbf{x}^2}{2}\right\}. \tag{3.5}$$

The following well established approaches will be used both here and throughout the remaining chapters to qualitatively and quantitatively demonstrate the superior properties of a quasi-regular grid.

(a) *Uniform pseudo-random+rejection:* Starting with a uniformly distributed high-discrepancy pseudo-random sequence $\mathbf{x}_i$ in a sufficiently large domain, one retains only the points that satisfy the inequality $\mathcal{P}(\mathbf{x}_i)/\mathcal{P}_{\max} > \xi_i$, where ξ_i is a pseudo-random number uniformly distributed in the $[0, 1]$ interval.

(b) *Uniform quasi-random+rejection:* Same as the above, but $\mathbf{x}_i$ is a low-discrepancy quasi-random sequence taken from the sobol distribution.

(c) *"True" quasi-random:* Utilizes the Beasley-Springer-Moro approximation[39;40] to transform a low-discrepancy uniform quasi-random sequence in $[0,1] \times [0,1]$ (i.e. the unit square) into a two-dimensional standard normal distribution.

High-discrepancy pseudo-random sequences are the most readily available, and as such are the most commonly implemented method. However, the increased availability of low-discrepancy quasi-random sequences combined with their desired properties has generated much research in both sampling and integration applications.[41–43]

The quasi-regular grids as well as both rejection methods (a and b, see above) can be readily applied to any general distribution function $\mathcal{P}(\mathbf{x})$. Unfortunately method (c) and similar approaches are only straightforward for certain distributions, such as those with the form shown in Eq. (3.6), or when other information like the cumulative distribution function is available.

$$\mathcal{P}(\mathbf{x}_1, \cdots, \mathbf{x}_d) = \mathcal{P}_1(\mathbf{x}_1) \times \cdots \times \mathcal{P}_d(\mathbf{x}_d) \tag{3.6}$$

An important realization and a key motivation for the development of the quasi-regular grids is that the rejection process destroys most of the desired properties associated with a quasi-random sequence. This claim will be explored quantitatively in Chapter 4.

However, Fig. 3.1 shows the transformation of a uniformly distributed Sobol sequence to a standard normal distribution through the inverse cumulative distribution function of the standard normal distribution. Unlike the rejection-based methods discussed above, the low-discrepancy of the uniformly distributed points is clearly maintained for the resulting normal distribution.

Unfortunately, if we want to address the more general problem of any type of distribution function we must abandon the use of explicit transformations. As shown in Fig. 3.2, even to generate a model system like the normal distribution, a large number of points must

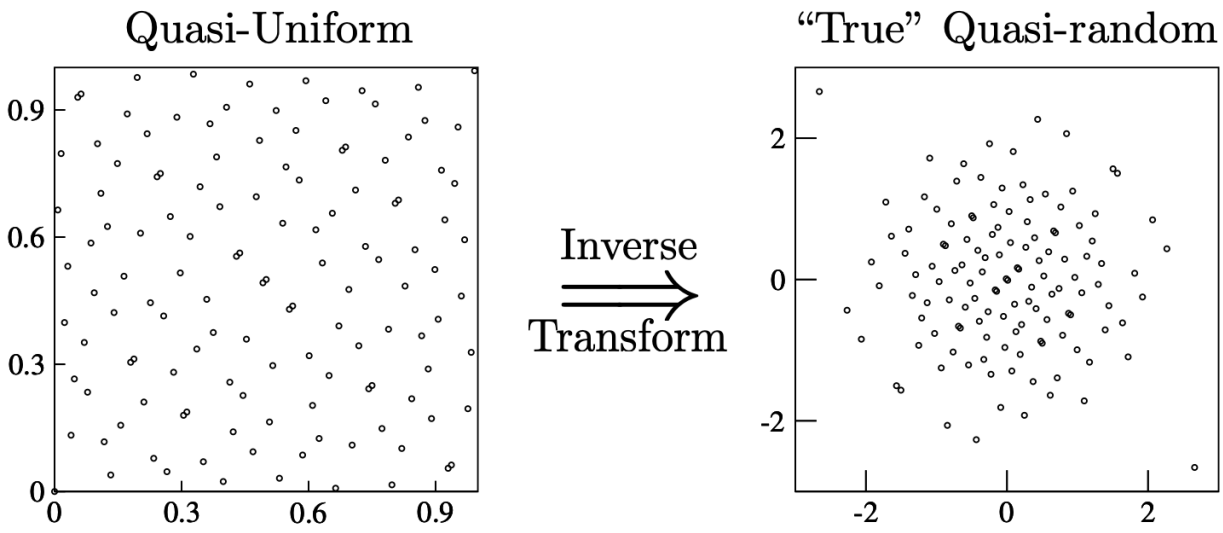

Figure 3.1: Two-dimensional uniform sobol sequence (left) transformed to a two-dimensional standard normal distribution using the inverse cumulative distribution function. The resulting normal distribution retains the organized structure present in the original sequence.

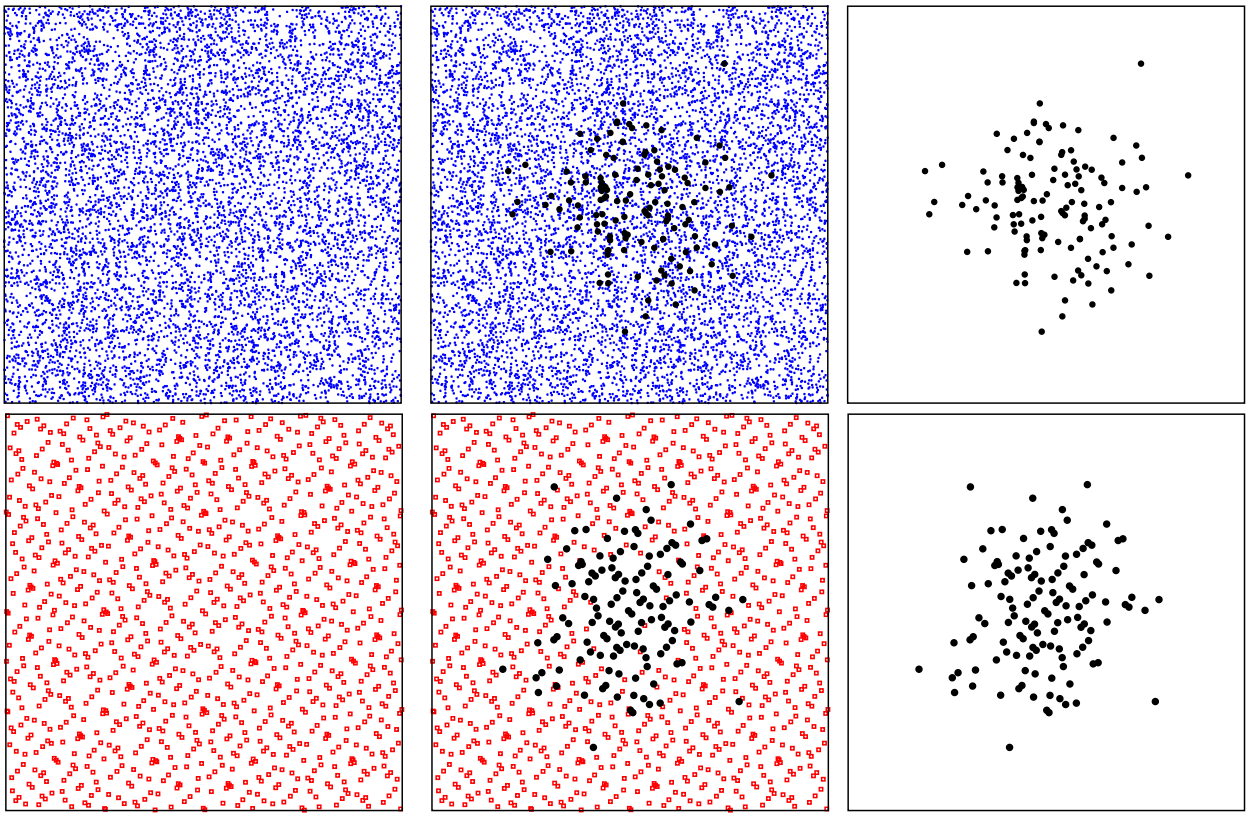

Figure 3.2: The generation of 128 points following the two-dimensional standard normal distribution using pseudo-random and rejection (top) or quasi-random and rejection (bottom). Although the pseudo-random approach requires many more trials, the end results appear to be similar.

be proposed and then rejected. Namely the first column shows the total number of points generated, while the last column shows only those that are accepted this inefficiency of the rejection method is well-known and grows with dimensionality. The final set of points from the high-discrepancy pseudo-random sequence after rejection appears to be Gaussian, but is clearly still high-discrepancy as expected. However, the initially low-discrepancy quasi-random sequence is, at-least visually, only appears marginally better when compared to the pseudo-random approach. There is some local uniformity maintained, however, it is clearly disrupted by the regions where the rejection method disregarded points. Therefore the initially low-discrepant nature of the sequence becomes more similar to a high-discrepancy sequence after rejection.

Fig. 3.3 shows the final product of all four methods. In contrast to the rejection ap-

proaches, the "true" quasi-random sampling appears to be more locally uniform with points efficiently filling the configuration space. We do note that while the "true" quasi-random sampling does appear superior to the rejection methods, implementations innately produce sets of points that appear arbitrarily close. This is a consequence of the uniformity definition used to construct quasi-random sequences; namely that they aim to maintain uniformity for every new point added to the sequence. These redundant points can result in numerical instabilities which will be discussed in Chapter 4. Finally, the quasi-regular points seem to be most ideal. The resulting grid looks strikingly similar to a two-dimensional contour plot of a Gaussian distribution, the points are clearly arranged according to $\mathcal{P}$. We also observe the local regularity. The grid points seem to be surrounded by nearly perfect shells, corresponding to their nearest neighbors.

3.3 Parameters of the Quasi-Regular Grids

The beauty of the quasi-regular ansatz is the simplicity and subsequent transparent parameterization of the model. There are a number of sampling methods that have non-transparent fitting procedures which may result in further optimization, but at the cost of reproducibility. For a given set of quasi-regular points, the only fitting parameter is the constant γ. This parameter is essentially a flat scaling which controls the strength with which points repel or attract. While it is expected to be on the order of 1, we find that its value is stable across a large range. Not reported here, however, the variation of γ does not seem to effect the numerical accuracy during applications as discussed in Section 4.3.

As shown in Fig. 3.4 the value of γ causes the quasi-regular grids to form a set of disconnected points. Likewise the points begin to take on a more uniform appearance when too large. For this reason there is a comfortable region in which γ will produce a proper quasi-regular grid and makes it unnecessary to rigorously optimize for most applications.

It is worth mentioning that for a given grid size N, a new quasi-regular grid must be constructed "from scratch", i.e. you must optimize the entire grid to account for any addi-

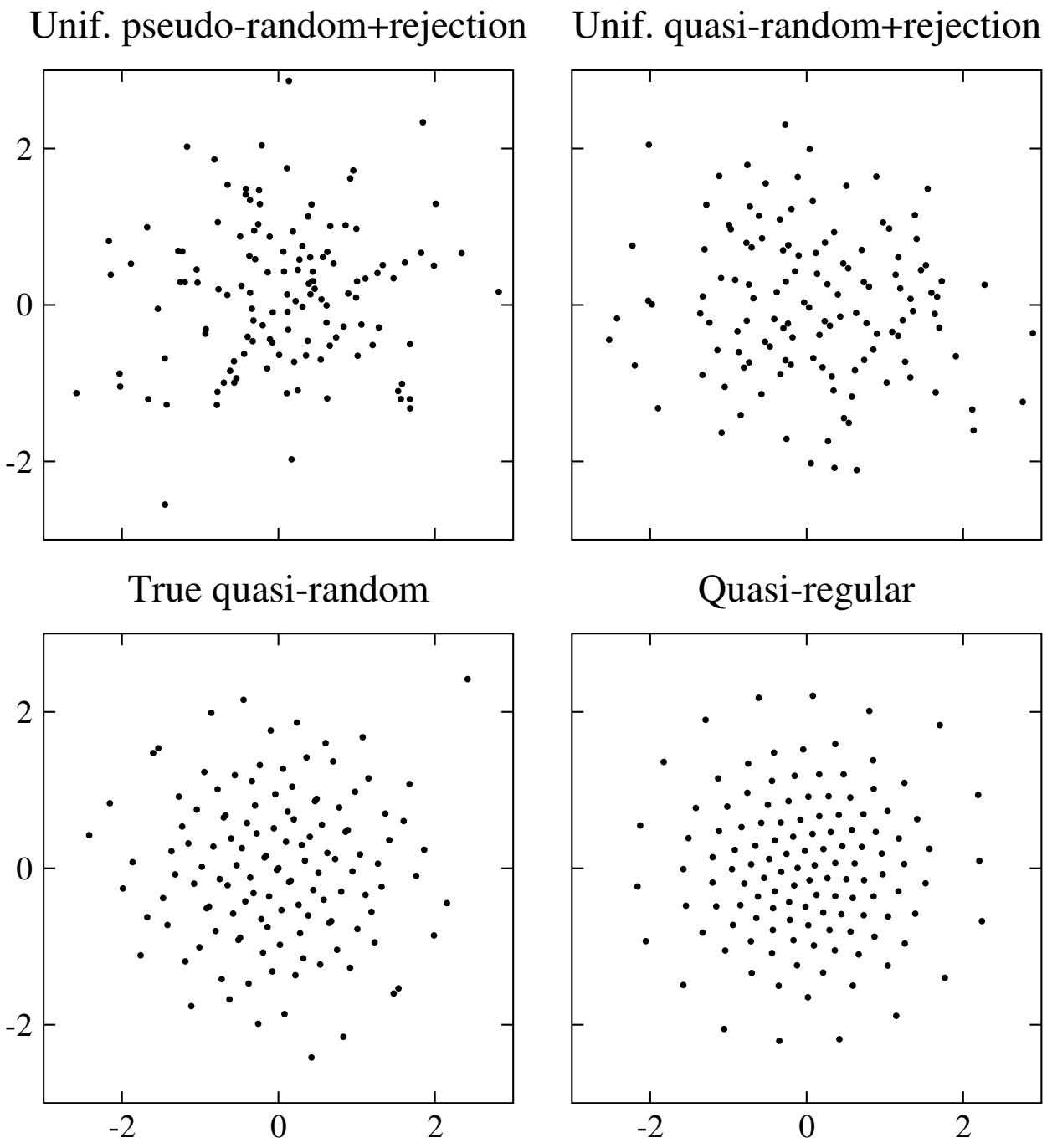

Figure 3.3: A comparison of four different methods (see text) to generate sets of points distributed according to the two-dimensional standard normal distribution., each set contains 128 points.

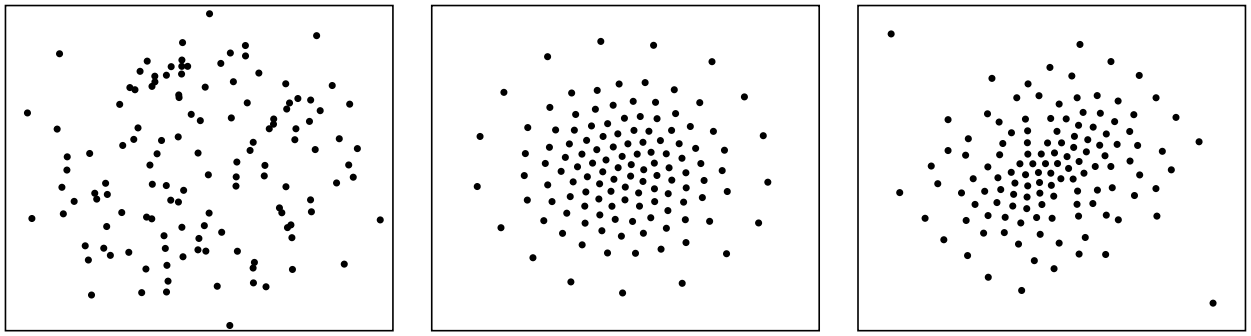

Figure 3.4: 128 points generated according to the two-dimensional standard normal distribution. The effects of taking unreasonable values for γ are shown (1/100, 1, 100 respectively) form left to right.

tional points. In contrast to this, any given pseudo-random or quasi-random finite sequence can be extended, if required, and by increasing the grid size, N, the results can be improved gradually, not wasting any information/data generated using the smaller grid. This would be an obvious disadvantage for using a quasi-regular grid for applications such as integration where one may wish to extend the sequence to reach a desired accuracy. However, in Chapters 4 and 6 these grid-based methods are applied to problems resulting in a generalized eigenvalue problem. The solution of an $N \times N$ generalized eigenvalue problem with its $\sim N^3$ numerical scaling appears to be much more expensive than the generation of the grid. We also find in that context, the optimization of the basis set is far more important in terms of total accuracy compared to adjusting γ.

Although a new quasi-regular grid must be constructed to vary N, the quasi-regular grids are able to reproduce the proper topology for $\mathcal{P}$ even when N is considerably small. The optimization process applied during the generation of a quasi-regular grid allows a reasonable approximation of the distribution to exist when N is small. This is unique compared to either low-discrepancy or high-discrepancy sequences.

The scaling properties with respect to N for the quasi-regular grids are shown in Fig. 3.5. The specific distribution function used in this example is discussed in detail in Chapter 4 Eq. (4.19); however, it is based on a semi-classical argument. The well-known Hénon-Heiles Potential is being represented by the quasi-regular grids.[44;45] The specific parameters used to construct the grids were as follows: $\gamma = 1$, $\lambda = 0.1118034$, and $E_{\text{cut}} = 13.3333$, which are taken directly from the literature.

$$V(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\left(\mathbf{x}^2 + \mathbf{y}^2\right) + \lambda\left(\mathbf{x}^2\mathbf{y} - \frac{1}{3}\mathbf{y}^3\right) \tag{3.7}$$

The reason for introducing this specific expression is its non-trivial form which no longer satisfies Eq. (3.6). The contour maps for this distribution are shown in blue, with a cutoff contour shown in red. Clearly there is no issue in generating a quasi-regular grid for a non-separable distribution, and the addition of more points scales with the appropriate density

as expected. This property may be useful for applications where an approximate solution or low accuracy is required.

3.4 Conclusions

We have introduced a new method and numerical algorithm for the sampling of a general multi-dimensional distribution function $\mathcal{P}(\mathbf{x})$ using a quasi-regular grid of points $\mathbf{x}_i$ $(i = 1, ..., N)$. This grid is constructed through the minimization of a pairwise functional, $\sum u(\mathbf{x}_i, \mathbf{x}_j) \rightarrow \min$, with the short-range pair pseudo-potential $u(\mathbf{x}_i, \mathbf{x}_j)$. The distance between grid=points is defined according to the underlying distribution, allowing the grids so form locally regular structures. The quasi-regular ansatz also contains only one fitting parameter that is stable over a comfortable range, removing any non-transparent fitting procedures.

Comparing the quasi-regular grids to the well-known Monte Carlo rejection methods, it was shown that the local regularity of the points resulting from a quasi-regular approach are highly desirable. Quasi-regular grids will surely be useful in many diverse areas of numerical analysis due to their general yet simple construction.

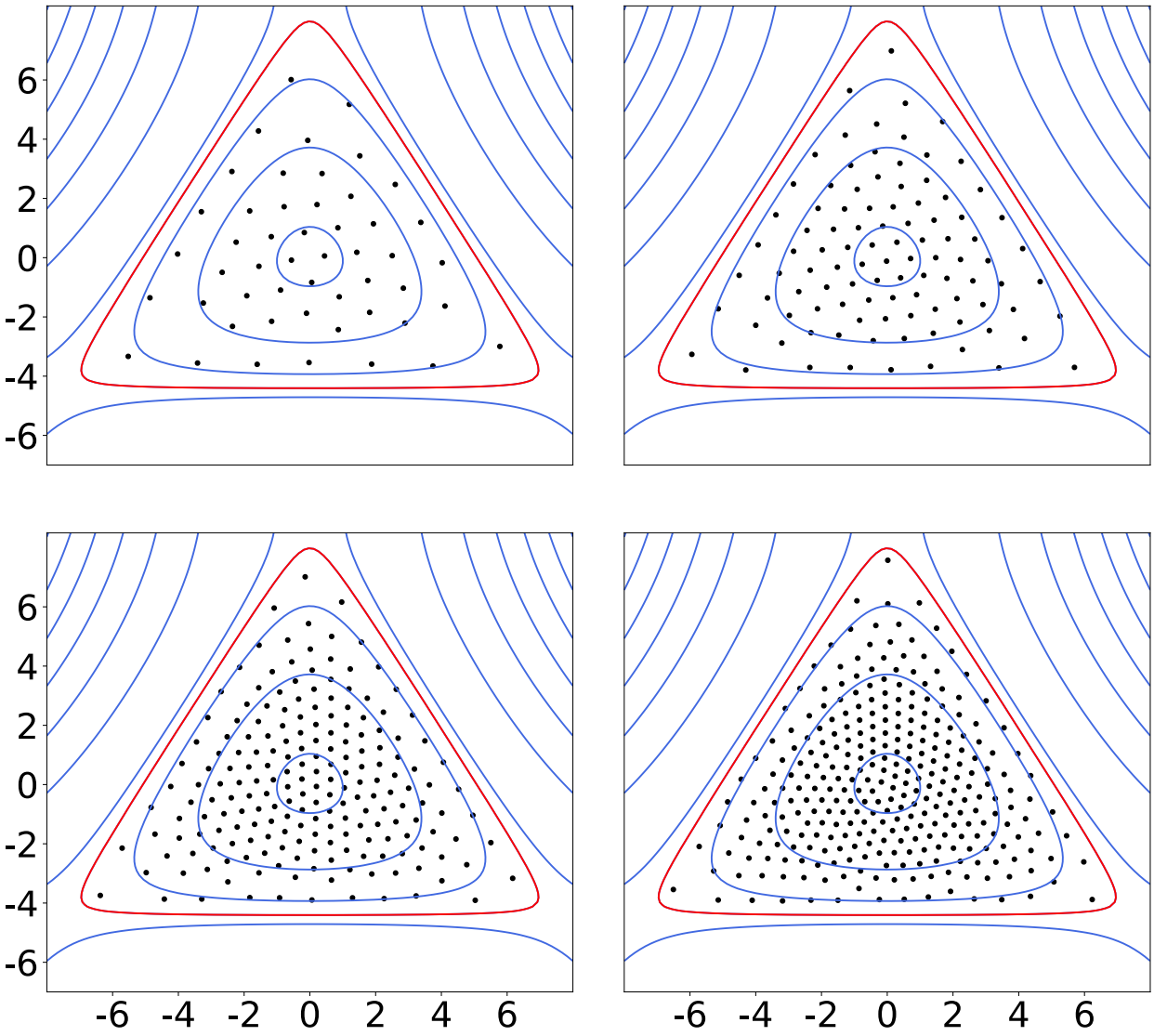

Figure 3.5: Sampling of the Hénon-Heiles Potential with progressively larger grid sizes (50, 100, 200, and 300 respectively). Note: the global topology is consistent with $\mathcal{P}$ independent of the size of N.

Chapter 3

Quantum Dynamics Calculations: A Quantitative Application for Quasi-Regular Grids

In this chapter we explore the computation of quantum vibrational spectra. Given a quantum system with d active degrees of freedom, first, one chooses a suitable coordinate system and a suitable set of basis functions. Then, by evaluating the matrix elements of the Hamiltonian operator in this basis, the problem of calculating the energy levels and the wavefunctions is reduced to an eigenvalue problem. We quantitatively explore the effects of generating a Gaussian basis set using the quasi-regular grids developed in Chapter 3. Using two-dimensional and three-dimensional Morse oscillators, we demonstrate that the resulting optimized Gaussian basis sets have properties superior to other choices explored previously in the literature.

4.1 Grid-Based Approaches in Quantum Dynamics

A large amount of research has been developed for solving the quantum dynamics of molecular systems. In particular, grid-based methods for solving the Schrödinger equation have been an active area of research for decades. The most prevalent approach applies a direct-product grid, avoiding the need to compute any potential energy integrals; typically, the resulting eigenvalue problem involves sparse matrices, which can be diagonalized using very efficient iterative eigensolvers that only need a function that multiplies a vector by a sparse matrix.[46–52] This simple approach has an unavoidable drawback, the exponential proliferation of the number of grid points with dimensionality,

$$N = c \cdot \kappa^d \tag{4.1}$$

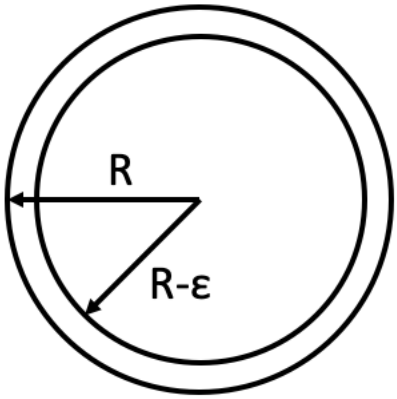

Figure 4.1: The outer shell of a circle defined by radius R and ε. The percentage of the volume contained in this shell is given by Eq. (4.4).

We note though that the "curse of dimensionality" is the very nature of any basis method, regardless of whether a primitive direct-product grid, or state-of-the-art functions are chosen. The two constants, c and κ, do depend on this choice though, which may result in a substantial reduction (or increase) in the total size of the basis; the largest possible κ would usually correspond to a direct-product grid. At the same time, a simple pruning of a direct product grid by retaining the points ($\mathbf{r}^{(i)}$) that appear within the energy cutoff region,

$$V(\mathbf{r}^{(i)}) < E_{\text{cut}}, \tag{4.2}$$

often results in a noticeable reduction of both c and κ.

While this initially seems promising, one must recall the classic problem in statistical physics; the scaling of the volume of a hypersphere with respect to dimensionality.[53]

$$V_d(R) = \frac{\pi^{d/2}}{\Gamma(d/2+1)}\, R^n \tag{4.3}$$

One can easily find an expression to estimate the fraction of the volume contained in the

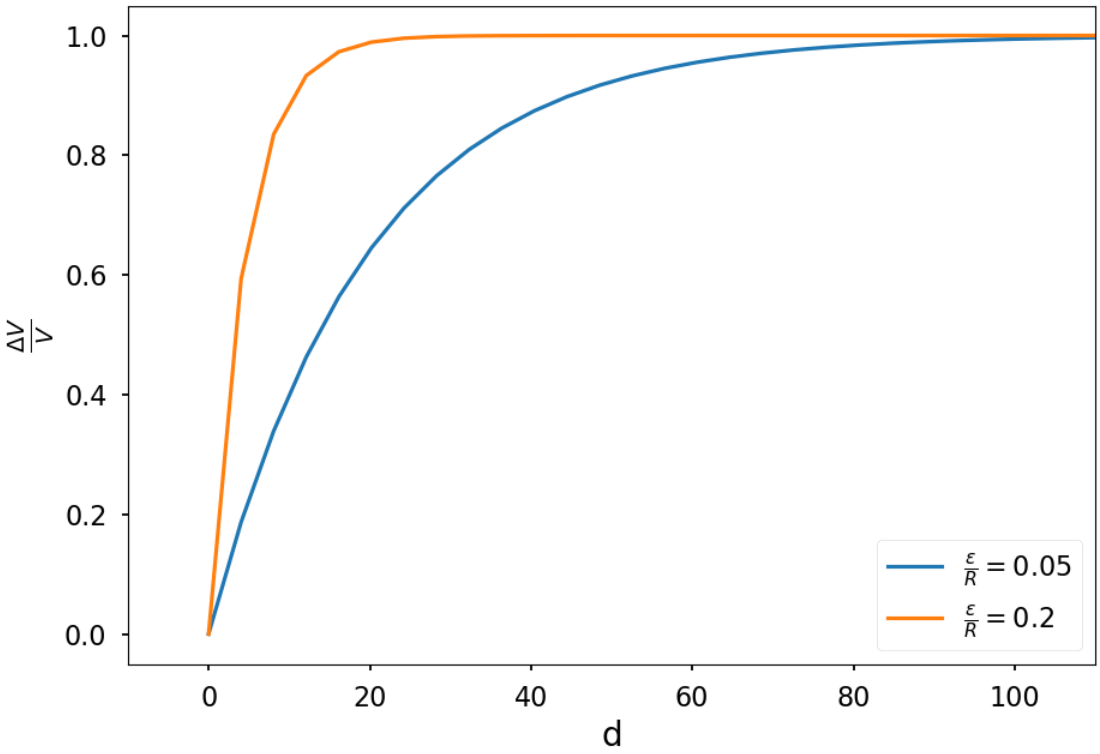

Figure 4.2: The fraction of the volume contained in the rind of a hypersphere. As the dimensionality increases the majority of the volume is found at the surface.

skin of the hypersphere,

$$\begin{aligned} \frac{\Delta V_d(R)}{V_d(R)} &= \frac{V_d(R) - V_d(R-\varepsilon)}{V_d(R)} \\ &= 1 - \left(1 - \frac{\varepsilon}{R}\right)^d . \end{aligned} \tag{4.4}$$

Plotting this expression with respect to the dimensionality as shown in Fig. 4.2 we see that the volume of a high-dimensional object rapidly becomes concentrated near its surface.

This textbook example makes it clear that the idea of covering a region of interest in a high dimensional space uniformly by a grid (or localized basis functions) is futile. Even after pruning according to Eq. (4.2), most of the grid points will still end up being wasted in the peripheral region, i.e., the region of least importance, where the wavefunction is small and not oscillatory. Sampling methods need to be developed to address this irrefutable fact for

higher-dimensional systems.

4.1.1 Distributed Gaussian Basis

In order to avoid this severe exponential scaling of uniform grids, one may need to give up the benefits of sparse linear algebra. Just as the citation list is long for direct-product grids, many authors have investigated the use of non-direct product grids.[54–60] In this context, a distributed Gaussian basis (DGB) is a particularly popular option with a long history going back several decades (see, e.g., Refs. 61–65). Gaussian basis functions are appealing due to their localized nature and their convenient analytic properties that readily extend to higher dimensions. Gaussian's can form a convenient and flexible framework for solving the Schrödinger equation. Although a generalized eigenvalue problem is obtained if the basis is not orthogonal, there is a hope that this flexibility can be exploited so that an optimal, compact, and efficient basis can be constructed. Consequently, a number of authors have introduced different Gaussian placement methods as well as basis set truncation schemes (see, e.g., Refs. 28;66–72). However, the optimal placement of the Gaussian basis functions is still an open question.

Even assuming that an optimal distribution function for the Gaussian centers, $\mathcal{P}(\mathbf{r})$, is known explicitly, its implementation is still not straight forward. As discussed in Chapter 3 one can easily generate a high-discrepancy pseudo-random sequence distributed according to any distribution function using Monte Carlo, but it is hard to imagine that such a grid would be optimal.

Accordingly, Garashchuk and Light proposed a scheme which partially addressed this problem and which we will refer to as *quasi-random+rejection*. Namely, a uniform low-discrepancy quasi-random (e.g., Sobol) sequence[35–37] can be generated in a domain of interest. A sequence $\mathbf{r}^{(i)}$ with the desired distribution can then be produced by a rejection scheme in which the points are retained with probability $\sim \mathcal{P}\left(\mathbf{r}^{(i)}\right)$. As discussed in Chapter 3 the rejection step destroys the nice low-discrepancy structure found in the original sequence,

which will subsequently impact the representation of the wave functions. While some of the local correlations of the low-discrepancy sequence remain, we expect the quasi-regular basis to be superior.

One could possibly compensate for the locally non-uniform distribution of Gaussian centers by customizing the width matrix for each Gaussian depending on its environment, but this would certainly turn the basis optimization into a very non-trivial problem. There is an additional problem one would need to address; the linear dependencies that inevitably arise due to some points appearing arbitrarily close. Such linear dependencies lead to numerical instabilities when solving the associated generalized eigenvalue problem. This can easily be captured by the condition number of the resulting matrices, and poses a serious problem for these methods. We believe the quasi-regular grids will addresses all of the concerns that exist in the *quasi-random+rejection* scheme.

4.2 Vibrational Energy Calculations Using a Distributed Gaussian Basis

Consider a d-dimensional system with the Hamiltonian (assuming atomic units and Cartesian coordinates):

$$\widehat{H} = -\frac{1}{2}\nabla^{\mathrm{T}}\mathbf{M}^{-1}\nabla + V(\mathbf{x}). \tag{4.5}$$

Here $\mathbf{M} := \mathrm{diag}(m_i)$ is a diagonal mass matrix, $\mathbf{x} \in \mathbb{R}^d$ defines a column vector in the configuration space, and ∇, the gradient.

It is convenient to make a transformation to the mass scaled coordinates,

$$\mathbf{r} := \mathbf{U}^{\mathrm{T}}\mathbf{M}^{1/2}(\mathbf{x} - \mathbf{x}_0) \tag{4.6}$$

where $\mathbf{U}$ defines the unitary transformation that diagonalizes the mass-scaled Hessian matrix:

$\tilde{\mathbf{K}} = \mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}$.

$$\tilde{\mathbf{K}} = \mathbf{U}\boldsymbol{\Omega}^2\mathbf{U}^{\mathrm{T}}; \quad \boldsymbol{\Omega} = \mathrm{diag}\{\omega_k\}; \quad \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I} \tag{4.7}$$

In these coordinates the Hamiltonian reads

$$\widehat{\mathbf{H}} = -\frac{1}{2}\nabla^2 + V'(\mathbf{r}) \tag{4.8}$$

where $V'(\mathbf{r}) := V\left(\mathbf{x}_0 + \mathbf{M}^{-1/2}\mathbf{U}\mathbf{r}\right)$.

$$\mathbf{r} := \mathbf{M}^{1/2}\mathbf{x}, \tag{4.9}$$

which gives

$$\widehat{\mathbf{H}} = -\frac{1}{2}\nabla^2 + V'(\mathbf{r}), \tag{4.10}$$

where $V'(\mathbf{r}) := V(\mathbf{M}^{-1/2}\mathbf{r})$.

4.2.1 Generalized Eigenvalue Problem

Consider a set of points $\mathbf{r}_i \in \mathbb{R}^d$ $(i = 1, \cdots, N)$ where each grid point is associated with a normalized Gaussian basis function:

$$\Phi_i(\mathbf{r}) := (2\alpha_i/\pi)^{d/4} \exp\left(-\alpha_i|\mathbf{r} - \mathbf{r}_i|^2\right), \tag{4.11}$$

and the α_i are strictly positive adjustable constants that are subject to optimization.

Defining the overlap and the Hamiltonian matrices,

$$\begin{aligned} \mathbf{S}_{ij} &:= \langle\Phi_i|\Phi_j\rangle \\ \mathbf{H}_{ij} &:= \langle\Phi_i|\widehat{H}|\Phi_j\rangle, \end{aligned} \tag{4.12}$$

the eigenenergies E_k and eigenfunctions Ψ_k can be computed by solving the generalized

eigenvalue problem,

$$(\mathbf{H} - E_k \mathbf{S})\Psi_k = 0. \tag{4.13}$$

Below we present explicit expressions for the numerical evaluation of the required matrix elements.

4.2.2 Matrix Elements

Gaussian's are a convenient (although not necessary) choice for the basis functions. Their multi-dimensional generalizations are well known, resulting in simple expressions to evaluate for the matrix elements.

Recall that the product of two Gaussian's is also a Gaussian:

$$\Phi_i(\mathbf{r})\Phi_j(\mathbf{r}) = \mathbf{S}_{ij} \left(\frac{\alpha_i + \alpha_j}{\pi}\right)^{d/2} \exp\left[-(\alpha_i + \alpha_j)(\mathbf{r} - \bar{\mathbf{r}}_{ij})^2\right], \tag{4.14}$$

where we have defined the vector

$$\bar{\mathbf{r}}_{ij} := \frac{\alpha_i \mathbf{r}_i + \alpha_j \mathbf{r}_j}{\alpha_i + \alpha_j} \tag{4.15}$$

The overlap matrix elements $\mathbf{S}_{ij} := \langle \Phi_i | \Phi_j \rangle$ are given by

$$\begin{aligned} \mathbf{S}_{ij} &= \int_{\mathbb{R}^d} d\mathbf{r}\ \Phi_i(\mathbf{r})\Phi_j(\mathbf{r}) \\ &= \left[\frac{4\alpha_i\alpha_j}{(\alpha_i + \alpha_j)^2}\right]^{d/4} \exp\left[-\frac{\alpha_i\alpha_j(\mathbf{r}_i - \mathbf{r}_j)^2}{(\alpha_i + \alpha_j)}\right] \end{aligned} \tag{4.16}$$

Given the Hamiltonian in the mass-scaled coordinates, Eq. (4.10), the kinetic energy

matrix elements are

$$\begin{aligned}\mathbf{T}_{ij} &:= -\frac{1}{2}\langle\Phi_i|\nabla^2|\Phi_j\rangle \\ &= \mathbf{S}_{ij}\,\frac{\alpha_i\alpha_j}{\alpha_i+\alpha_j}\left[d-\frac{2\alpha_i\alpha_j(\mathbf{r}_i-\mathbf{r}_j)^2}{\alpha_i+\alpha_j}\right]\end{aligned} \tag{4.17}$$

Finally the potential energy matrix elements $\mathbf{V}_{ij} := \langle\Phi_i|V'|\Phi_j\rangle$ are evaluated numerically in the space of the mass-scaled Cartesian coordinates. In this work we elect to use a low-order Gauss-Hermit quadrature.

$$\begin{aligned}\mathbf{V}_{ij} &= \int_{\mathbb{R}d} d\mathbf{r}\;\Phi_i(\mathbf{r})\Phi_j(\mathbf{r})\;V'(\mathbf{r}) \\ &= \mathbf{S}_{ij}\left(\frac{\alpha_i+\alpha_j}{\pi}\right)^{d/2}\int_{\mathbb{R}d} d\mathbf{r}\;V'(\mathbf{r})\;\exp\left[-(\alpha_i+\alpha_j)(\mathbf{r}-\bar{\mathbf{r}}_{ij})^2\right] \\ &= \pi^{-d/2}\;\mathbf{S}_{ij}\int_{\mathbb{R}d} d\mathbf{z}\;V'\left(\bar{\mathbf{r}}_{ij}+\frac{\mathbf{z}}{\sqrt{\alpha_i+\alpha_j}}\right)e^{-\mathbf{z}^2}\end{aligned} \tag{4.18}$$

Finally, the eigenenergies E_k and eigenfunctions Ψ_k are obtained by solving the generalized eigenvalue problem, Eq. (4.13).

4.2.3 The Distribution Function

In this work we assume that an optimal distribution function for the positions of the Gaussian centers is defined using the form derived from a semiclassical argument.[73] The argument is simplistic, however, we believe it gives the proper dimensionality scaling and density. Starting with the total energy as the sum of the kinetic and potential, the momentum has the form, $p = \sqrt{2m(\mathbf{E}-\mathbf{V})}$. Because the distance is inversely proportional to the density, the uncertainty principle suggests the density should be of the form $\mathcal{P} \propto (\mathbf{E}-\mathbf{V})^{d/\mathbf{2}}$.

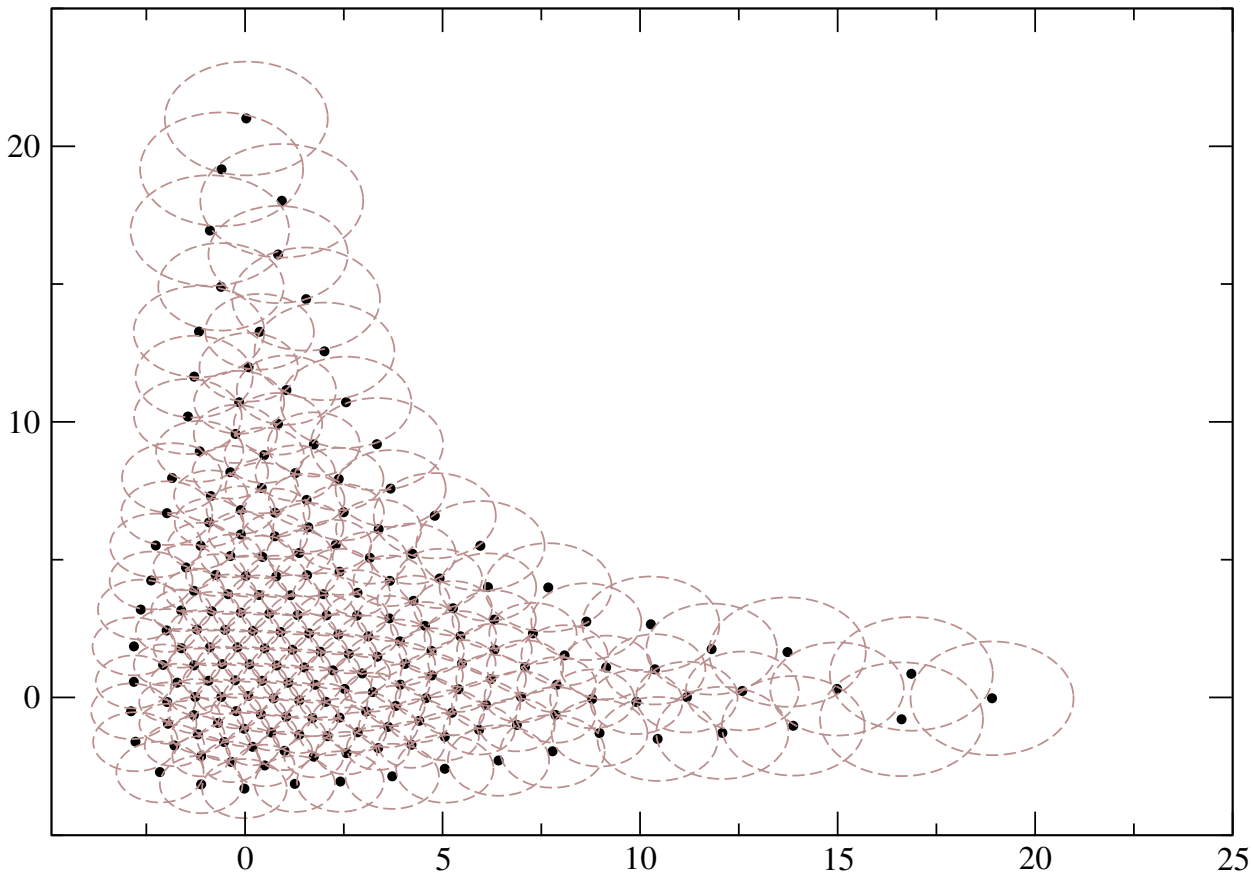

Figure 4.3: Depiction of the Gaussian widths and subsequent overlap generated when applying Eq. (4.21) to a quasi-regular grid. The condition number and stability of the resulting eigenvalue problem are ensured due to the uniformity.

Introducing a cutoff contour and normalizing the distribution we arrive at:

$$\mathcal{P}(\mathbf{r}) := \begin{cases} c^{-1}\left[E_{\text{cut}} - V(\mathbf{r})\right]^{d/2}, & V(\mathbf{r}) < E_{\text{cut}} \\ 0, & V(\mathbf{r}) \geq E_{\text{cut}} \end{cases} \tag{4.19}$$

where

$$c = \int_{V(\mathbf{r})<E_{\text{cut}}} d\mathbf{r}\ \left[E_{\text{cut}} - V(\mathbf{r})\right]^{d/2}$$

is the normalization constant. In our numerical tests we compare against recent work developed by Garashchuk and Light; they also implement a distribution function of this form, although instead of the power, $d/2$, they used an adjustable constant.[28]

A reasonable choice for the set of α_i constants depends on the type of grid used for

constructing the Gaussian centers. Assuming the points $\mathbf{r}_i$ are distributed according to $\mathcal{P}(\mathbf{r})$, the following expression to define the Gaussian widths is natural:

$$\alpha_i := \alpha_0 \left[N\mathcal{P}(\mathbf{r}_i)\right]^{2/d}. \tag{4.20}$$

Due to the local regularity of the quasi-regular grids, we propose an alternate definition for the Gaussian widths:

$$\alpha_i := \alpha_0 \, \mathrm{dist}_i^{-2}, \tag{4.21}$$

where dist_i is the distance to the nearest neighbor of the i-th point. The latter expression is clearly preferable as it conveniently decouples the problem of grid generation from the following steps in the solution of the quantum dynamics problem. The choice of α_0 is not obvious but depends on the distribution function and the basis size. Its value adjusts for the scale in configuration space and can easily be identified as shown in Fig. 4.4.

Finally, we note that numerical instabilities are often encountered when distributed Gaussian basis sets are employed, especially when combined with nonuniform grids. For example, when two grid points appear arbitrarily close to each other, the corresponding Gaussian's become linearly dependent. This in turn leads to a large condition number in both the Hamiltonian and the overlap matrices. The quasi-regular grids by definition minimize this very problem as they avoid any clustering of the grid points. In addition, Eq. (4.21) assures that all the adjacent Gaussian's have similar overlap. These properties can be visualized as shown in Fig. 4.3 where a quasi-regular grid ($N = 200$) has been generated using the Morse potential, Eq. (4.22). The Gaussian widths are scaled according to their nearest neighbor as defined by σ. This forces all the Gaussian's to maintain a similar overlap, corresponding to smaller Gaussian's in the regions where the wavefunction is more oscillatory. Likewise the less oscillatory regions are covered with fewer, broader Gaussian's.

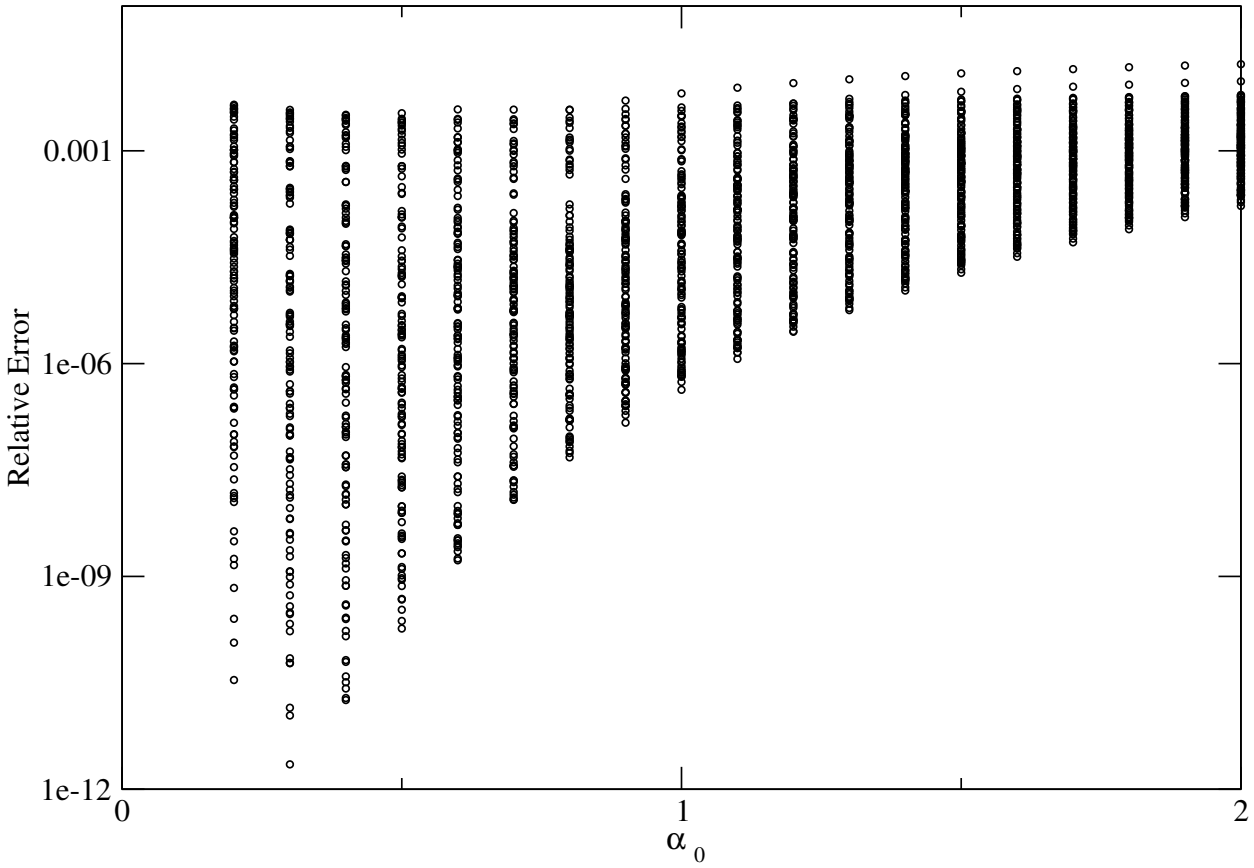

Figure 4.4: Sample optimization for the choice of α_0, the only fitting parameter associated with the construction of a quasi-regular basis.

4.3 Eigenenergy Computations for the Morse Potential

Here we set out to apply the method to two-dimensional and three-dimensional systems described by the Morse potential:

$$V(\mathbf{x}) = D \sum_{k=1}^{d} \left(e^{-w_k \mathbf{x}_k} - 1 \right)^2 , \tag{4.22}$$

where the Morse parameters are taken to be $D = 12.0$, $w_1 = 0.2041241$, $w_2 = 0.18371169$, and $w_3 = 0.16329928$. These same models were used by Garashchuk and Light in a similar framework, albeit in order to demonstrate the superiority of the Gaussian bases distributed according to the quasi-random+rejection scheme.[28] Here, we compare the present approach with the aforementioned methods; however, we are specifically interested in the comparison to the quasi-random+rejection scheme, which was shown to be the most efficient implementation in the literature.

We note that most grid-based methods have several adjusting parameters that could, in principle, be finely tuned to achieve the best performance for a given problem. While this may be desirable for challenging systems, the exact strategy used is rarely elaborated on in the literature. This results in non-transparent methods that may or may not be reproducible. In this work we limited the grid optimization to making "reasonable" choices for the grid parameters and refrained from over-tuning any specific method for practical comparisons and conclusions to be drawn. All parameters used to construct the Gaussian bases are provided in Table. 4.1. Values were chosen to be consistent with previously published work whenever possible.

To this end, Fig. 4.5 shows the placement of the Gaussian centers for the 2D Morse potential using the four different schemes of interest. For the direct-product and quasi-random basis sets we tried to replicate the choices made in Ref. 28, whenever the information was available. In fact, we were unable to accurately reproduce the results reported in Ref. 28 because the precise values of the grid parameters were not provided in the paper. Further-

Table 4.1: Parameters used to construct the Gaussian bases for the model systems described by the 2D and 3D Morse potential. Values were chosen to be consistent with previously published work whenever possible.

d	grid	N	E_{cut}	γ	α_0
2	Direct-product	482	11.5	-	2.1
2	Quasi-rand.+Rej.	482	11.5	1.0	1.2
2	Quasi-regular	300	12	1.0	0.3
2	Quasi-regular	400	12	1.0	0.3
3	Quasi-rand.	939	7.5	1.0	0.4
3	Quasi-rand+Rej.	939	7.5	1.0	0.5
3	Quasi-regular	500	7	1.2	0.5
3	Quasi-regular	700	7	1.2	0.3
3	Quasi-regular	900	7	1.2	0.3

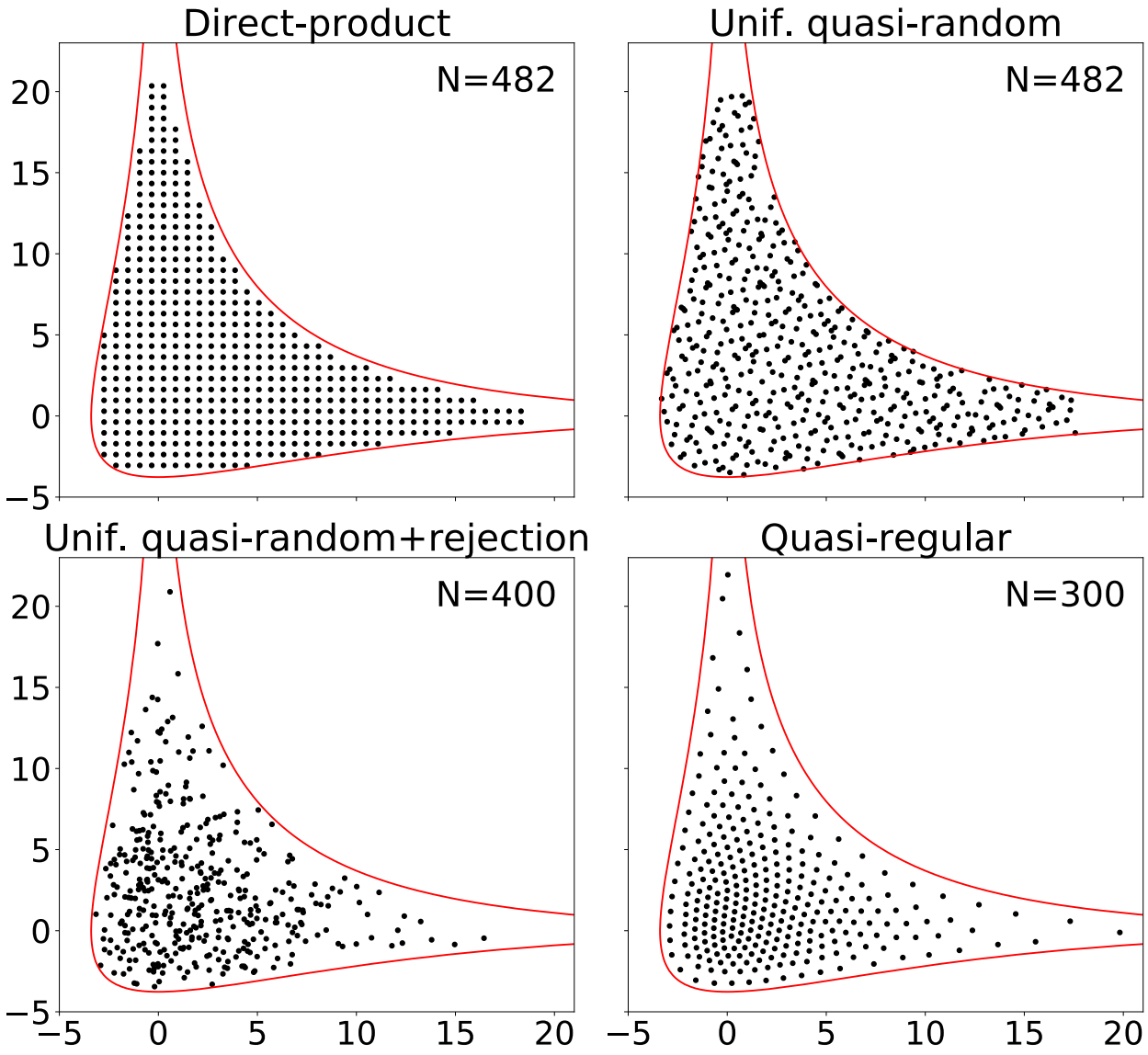

Figure 4.5: Placement of Gaussian basis functions for the two dimensional Morse potential. Also shown (in red) the $V = E_{\text{cut}} = 11.5$ contour line.

more, these parameters could not be recovered since all the relevant raw data was lost due to a hard drive failure after the paper was published (private communication with Sophya Garashchuk). We made reasonable choices for analysis; however, it is clear the original publication was subjected to a non-trivial optimization process to generate higher accuracy than we produce in this work.

In all cases the potential matrix elements were computed using a five-point Gauss-Hermite quadrature for each degree of freedom. A three-point quadrature would certainly suffice for these applications, but the higher-order scheme was chosen to completely eliminate this source of error for more accurate comparisons. Because the Morse oscillators used in the Hamiltonian are not coupled, the eigenvalue problem can be solved analytically. These analytic values were computed and used to assess each methods accuracy through the relative error.

Fig. 4.6 shows the relative errors for the lowest 100 eigenenergies for the 2D Morse potential using three different DGB schemes. Although not shown, the two-dimensional uniform quasi-random grid ($N = 482$) had accuracy comparable to that using the direct-product grid. As demonstrated earlier, the truncated direct-product scheme is inferior to the quasi-random+rejection approach.[28] However, the present (quasi-regular) scheme is even more accurate than the latter using a smaller grid ($N = 400$) and the errors are still smaller. Notably, the results obtained with the present method using only $N = 300$ Gaussian's are better than those using the direct-product grid with $N = 482$.

The results for the three dimensional Morse potential are presented in Figure 4.7. The 3D direct-product grid (not shown) was significantly worse than any of the methods presented, while the uniform quasi-random grid was significantly worse than the results from the other two methods shown. Here we observe that the quasi-regular grids have a distinct advantage over the second best method, i.e. quasi-random+rejection. That is, the present approach using $N = 700$ provides similar accuracy for all 100 eigenenergies as the quasi-random+rejection method using $N = 939$.

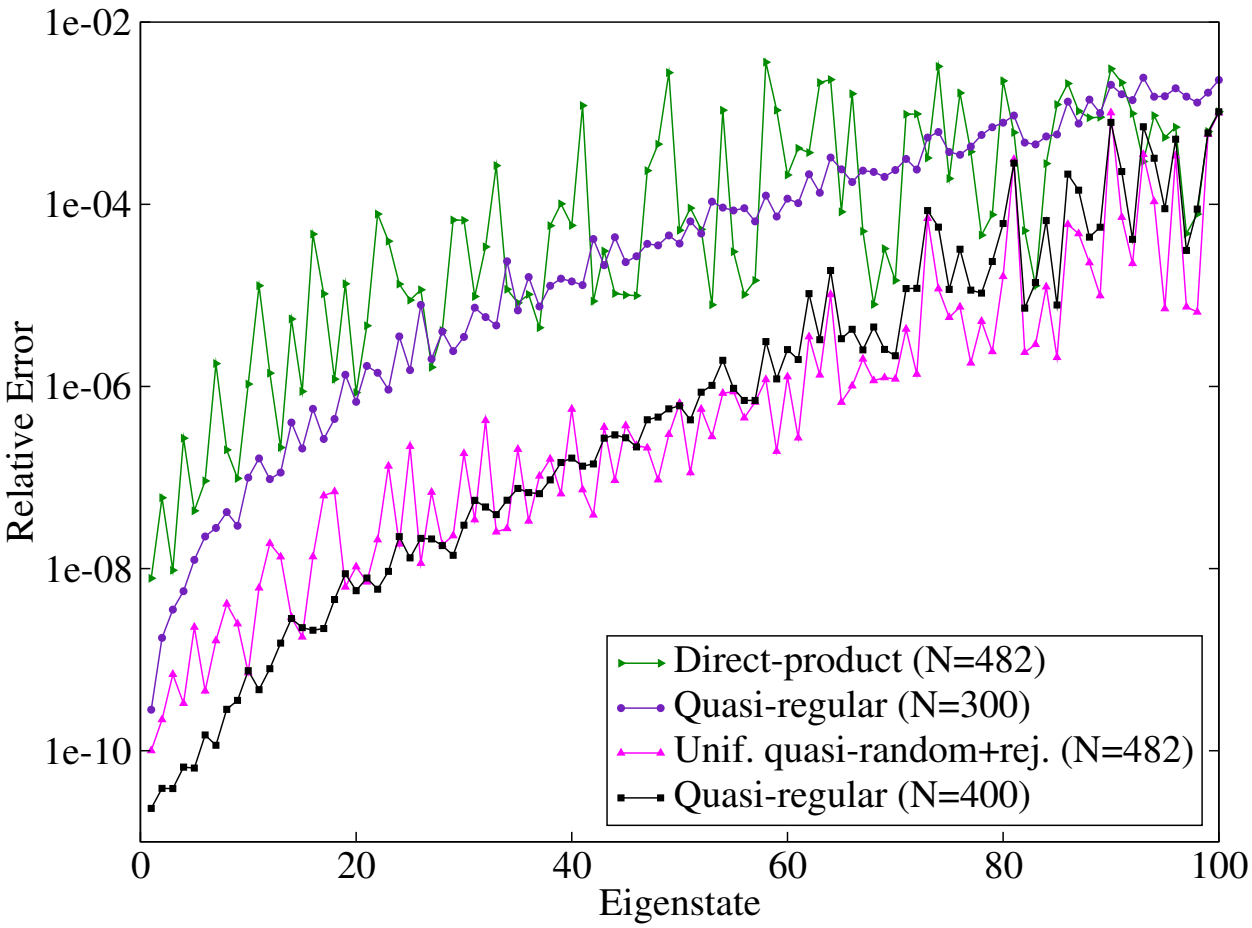

Figure 4.6: Relative error for the eigenenergies of the 2D Morse potential computed using three different schemes to generate Gaussian bases. Note the logarithm scale.

Table 4.2: Relative error for the 50 lowest eigenenergies for the two-dimensional Morse oscillator. Tabular data associated with Fig. 4.6 All results are in cm^{-1}.

Direct-Prod (482)	QRG (300)	Quasi+Rej. (482)	QRG (400)
7.83E-09	1.00E-10	2.82E-10	2.33E-11
6.00E-08	2.20E-10	1.74E-09	3.86E-11
9.58E-09	6.91E-10	3.55E-09	3.85E-11
2.70E-07	3.31E-10	5.65E-09	6.61E-11
4.31E-08	2.27E-09	1.25E-08	6.46E-11
9.19E-08	4.53E-10	2.25E-08	1.49E-10
1.78E-06	1.62E-09	2.79E-08	1.15E-10
2.01E-07	4.09E-09	4.18E-08	2.85E-10
9.77E-08	2.48E-09	2.95E-08	3.58E-10
1.06E-06	7.19E-10	9.98E-08	7.57E-10
1.28E-05	6.14E-09	1.62E-07	4.68E-10
1.40E-06	1.89E-08	9.62E-08	7.96E-10
2.14E-07	1.35E-08	1.13E-07	1.52E-09
5.50E-06	2.93E-09	4.02E-07	2.83E-09
8.83E-07	1.78E-09	2.08E-07	2.25E-09
4.70E-05	1.35E-08	5.68E-07	2.10E-09
1.04E-05	6.31E-08	2.65E-07	2.20E-09
1.20E-06	7.01E-08	4.40E-07	4.58E-09
1.34E-05	6.28E-09	1.34E-06	8.78E-09
8.59E-07	1.05E-08	6.78E-07	5.73E-09
4.64E-06	7.13E-09	1.68E-06	7.89E-09
7.81E-05	2.07E-08	1.41E-06	5.94E-09
3.94E-05	1.34E-07	9.23E-07	9.33E-09
1.33E-05	1.86E-08	3.54E-06	2.24E-08
8.87E-06	2.19E-07	1.51E-06	1.31E-08
1.15E-05	1.14E-08	7.86E-06	2.15E-08
1.62E-06	6.90E-08	2.00E-06	2.11E-08
4.11E-06	1.83E-08	4.00E-06	1.79E-08
6.72E-05	2.30E-08	2.44E-06	1.40E-08
6.67E-05	1.84E-07	3.50E-06	3.00E-08
9.73E-06	3.44E-08	7.32E-06	5.61E-08
3.41E-05	4.23E-07	5.77E-06	4.75E-08
0.000266	2.53E-08	4.67E-06	3.92E-08
1.17E-05	2.75E-08	2.37E-05	5.63E-08
8.28E-06	2.04E-07	6.84E-06	7.58E-08
1.03E-05	3.33E-08	1.59E-05	6.83E-08
4.42E-06	1.04E-07	7.55E-06	6.65E-08
5.83E-05	1.60E-07	1.27E-05	9.41E-08
0.000101	6.63E-08	1.53E-05	1.46E-07
5.86E-05	5.63E-07	1.42E-05	1.63E-07
0.001215	7.36E-08	1.30E-05	1.33E-07
8.64E-06	3.88E-08	4.15E-05	1.41E-07
3.07E-05	3.56E-07	2.15E-05	2.71E-07
1.05E-05	9.29E-08	4.35E-05	2.94E-07
1.01E-05	3.71E-07	2.32E-05	2.73E-07
9.91E-06	2.26E-07	2.69E-05	2.17E-07
0.000234	2.11E-07	3.68E-05	4.31E-07
0.000457	9.44E-08	3.54E-05	4.61E-07
0.002789	2.97E-07	4.55E-05	5.67E-07
5.18E-05	6.53E-07	3.71E-05	6.14E-07

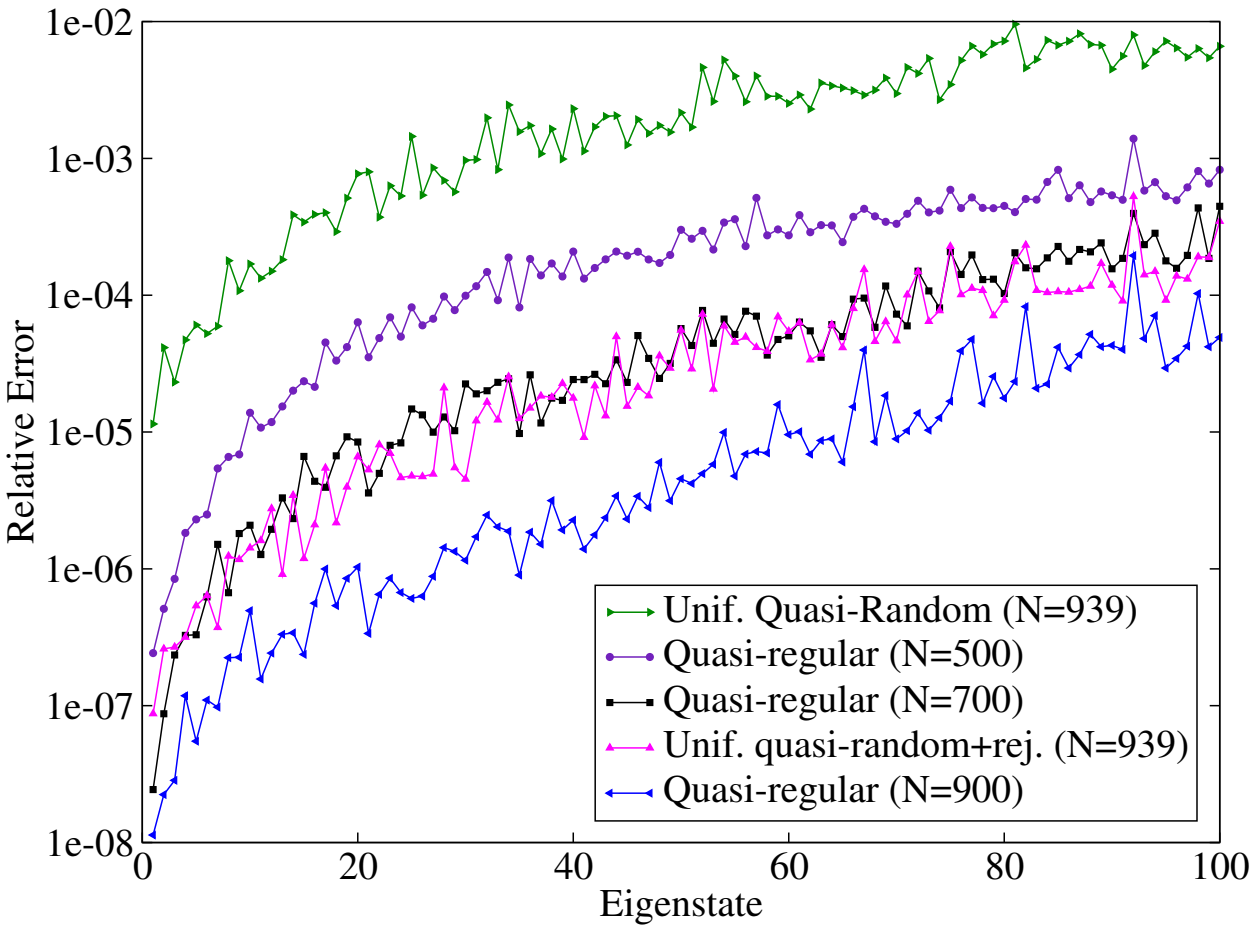

Figure 4.7: Relative error for the eigenenergies of the 3D Morse potential computed using three different schemes to generate Gaussian bases. (Note the logarithm scale.)

Table 4.3: Relative error for the 50 lowest eigenenergies for the three-dimensional Morse oscillator. All results are in cm^{-1}.

Quasi-Rand (939)	QRG (500)	Quas+Rej. (939)	QRG (700)	QRG (900)
1.15E-05	2.42E-07	8.74E-08	2.44E-08	1.13E-08
4.14E-05	5.11E-07	2.60E-07	8.73E-08	2.24E-08
2.31E-05	8.45E-07	2.68E-07	2.35E-07	2.85E-08
4.71E-05	1.83E-06	3.17E-07	3.27E-07	1.18E-07
6.05E-05	2.30E-06	5.38E-07	3.30E-07	5.50E-08
5.24E-05	2.50E-06	6.39E-07	6.24E-07	1.10E-07
5.93E-05	5.42E-06	3.73E-07	1.51E-06	9.76E-08
0.000179	6.58E-06	1.24E-06	6.69E-07	2.24E-07
0.000108	6.87E-06	1.17E-06	1.81E-06	2.26E-07
0.000169	1.38E-05	1.42E-06	2.09E-06	4.95E-07
0.000133	1.08E-05	1.61E-06	1.27E-06	1.57E-07
0.00015	1.18E-05	2.76E-06	1.94E-06	2.42E-07
0.000182	1.54E-05	9.12E-07	3.30E-06	3.32E-07
0.000387	2.01E-05	3.45E-06	2.33E-06	3.40E-07
0.000342	2.35E-05	1.19E-06	6.63E-06	2.37E-07
0.000389	2.14E-05	2.10E-06	4.37E-06	5.61E-07
0.000401	4.52E-05	5.46E-06	3.95E-06	9.98E-07
0.000292	3.33E-05	2.17E-06	6.71E-06	5.39E-07
0.000513	4.18E-05	3.97E-06	9.23E-06	8.52E-07
0.000771	6.35E-05	6.61E-06	8.45E-06	1.03E-06
0.000797	3.52E-05	5.30E-06	3.59E-06	3.37E-07
0.000372	4.87E-05	8.08E-06	4.99E-06	6.51E-07
0.00063	6.89E-05	6.99E-06	8.00E-06	8.53E-07
0.000532	4.98E-05	4.65E-06	8.33E-06	6.71E-07
0.001447	8.15E-05	4.77E-06	1.47E-05	6.08E-07
0.000538	6.02E-05	4.73E-06	1.34E-05	6.32E-07
0.000853	6.73E-05	4.93E-06	9.98E-06	8.80E-07
0.000689	9.78E-05	2.10E-05	1.29E-05	1.43E-06
0.000569	7.78E-05	5.49E-06	1.02E-05	1.34E-06
0.000968	9.91E-05	4.53E-06	2.25E-05	1.16E-06
0.00098	0.000116	1.21E-05	1.90E-05	1.72E-06
0.00198	0.000148	1.65E-05	2.00E-05	2.47E-06
0.000827	9.20E-05	1.23E-05	2.31E-05	2.03E-06
0.002457	0.000189	2.54E-05	2.45E-05	1.88E-06
0.00157	8.14E-05	1.26E-05	9.76E-06	9.00E-07
0.001736	0.000184	1.50E-05	2.62E-05	1.86E-06
0.001081	0.00014	1.83E-05	1.16E-05	1.52E-06
0.001639	0.000171	1.79E-05	1.77E-05	3.15E-06
0.000989	0.000137	2.27E-05	1.71E-05	1.93E-06
0.002308	0.000208	1.77E-05	2.42E-05	2.27E-06
0.001133	0.000132	9.16E-06	2.42E-05	1.40E-06
0.001699	0.000158	2.18E-05	2.64E-05	1.77E-06
0.002029	0.000183	1.32E-05	2.26E-05	2.37E-06
0.002053	0.000209	4.99E-05	3.37E-05	3.41E-06
0.001255	0.000195	1.54E-05	2.30E-05	2.31E-06
0.001921	0.000208	2.12E-05	5.07E-05	3.39E-06
0.001524	0.000183	1.84E-05	3.45E-05	2.80E-06
0.00174	0.000172	3.61E-05	2.47E-05	6.01E-06
0.001556	0.000197	2.93E-05	3.17E-05	3.15E-06
0.002159	0.0003	5.56E-05	5.70E-05	4.55E-06

4.4 Conclusions

We have provided an assessment of the use of quasi-regular grids to construct a distributed Gaussian basis. These types of bases are a key ingredient in a well-established framework for calculating vibrational spectra, marking just one of many possible applications for the formalism. We demonstrated that the redundant nature of the Monte Carlo rejection schemes result in a distributed Gaussian basis with "broken teeth" which cannot be optimal for representing the wavefunctions of an atomic system. The quasi-regular grids minimize not only the total size of the basis, but also the well-known numerical instability associated with the ill-conditioned nature of the overlap matrix. The method also only contains one fitting parameter associated with the grid, and one associated with the Gaussian widths, making it both simple and reproducible. No complex truncation schemes or approximations in the potential energy are required, resulting in high accuracy vibrational spectra.

Chapter 4

An Improved Quasi-Regular Formalism

The purpose of this chapter is to improve upon the original quasi-regular ansatz, in particular to generate a high-dimensional grid that we are certain is quasi-regular. We achieve this by modifying the form of the pseudo-potential. By removing the attractive terms in the potential we find that cavities no longer form in the grid. This has the added benefit of no longer requiring the distribution function to be normalized, significantly reducing the computational effort in generating the grid. We then introduce a radial pair correlation function to measure the distribution of nearest neighbors throughout the grid. By evaluating the resulting histogram one can confirm the local structure of the grid points, and the resulting quality of the quasi-regular grid itself. Finally we apply this updated methodology to model systems of higher dimensional Morse oscillators ($d = 4, 5$, and 6), where we show that high-accuracy vibrational calculations can still be performed in these higher dimensional environments.

5.1 A Practical Pseudo-Potential

Although the original quasi-regular grid implementation was shown to be superior to other available sampling methods, we discovered some problems when working in higher dimensions. While in theory the form of the pseudo-potential should not impact the resulting quasi-regular grid, assuming the interactions are truly short-range, in practice the choice can impact the number of iterations needed to minimize a given grid. The original motivation for Eq. (3.2) was to form a Lennard-Jones "like" crystal structure, however, the presence of an attractive term results in the possibility to form gaps due to the points being attracted to some extent.

This occurs during the minimization process, as shown in Fig. 5.1 regions of the grid

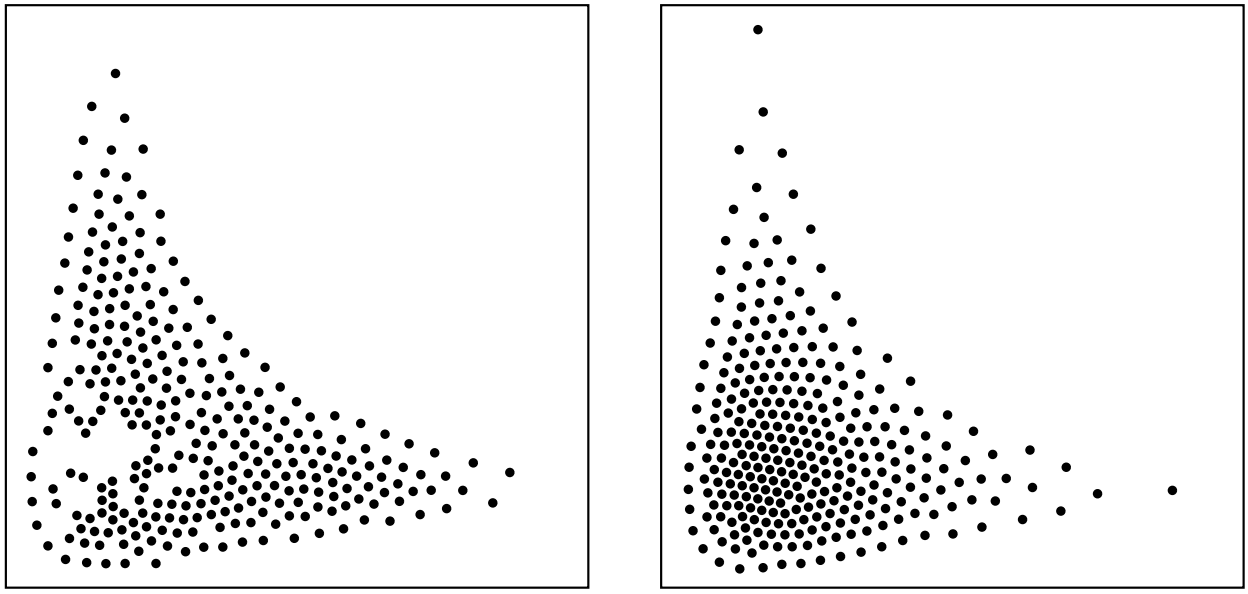

Figure 5.1: The generation of a quasi-regular grid using the Lennard-Jones pseudo-potential. As the grid is being optimized a large number of local minima must be traversed, resulting in the formation of holes (left) which require numerous iterations to remove (right) before a proper quasi-regular grid is generated.

have formed holes due to these attractions. The algorithm does ensure that given enough iterations these holes will be removed, however, this extended minimization process become a non-trivial computationally bottleneck as the dimensionality increases. In higher dimensions the minimization landscape is much more complex, containing a plethra of local minima for the points to explore. This results in a large number of iterations being needed to reach the global minima. By keeping the pseudo-potential short-range, but removing these attractions, the resulting grids will maintain their quasi-regular structure but miniimization will occur much faster.

5.2 Quasi-Regular Grid Ansatz Revisited

Consider a general, however no longer necessarily normalized, distribution function $\mathcal{P}(\mathbf{r}) \geq 0$ with a finite support $\mathcal{A} \in \mathbb{R}^d$. Our goal is to construct a set of points (or "particles") $\mathbf{r}^{(i)} \in \mathcal{A}$ $(i = 1, \cdots, N)$, which (I) locally, have a regular (possibly, closed-packed) arrangement and (II) globally, are distributed according to $\mathcal{P}(\mathbf{r})$. Clearly, the two conditions, (I) and (II), are mutually contradictory and as such can only be satisfied approximately. That is, the local regular arrangement around each point $\mathbf{r}^{(i)}$ is ideally a spherical shell of nearest neighbors with radius $r_{\min}(\mathbf{r}^{(i)})$. For condition (II) it is then natural to require the scaling law,

$$r_{\min}(\mathbf{r}) = \kappa \left[\mathcal{P}(\mathbf{r})\right]^{-1/d}, \tag{5.1}$$

to be satisfied approximately for any $\mathbf{r} = \mathbf{r}^{(i)}$ with some constant κ.

Here, for the construction of a quasi-regular grid we propose both an improved and simplified (compared to that in Chapter 3) solution based on the minimization of the energy functional,

$$\mathcal{U}(\mathbf{r}^{(1)}, \cdots, \mathbf{r}^{(N)}) = \sum_{i=1}^{N} \sum_{j=1}^{N} u_{ij} \longrightarrow \min, \tag{5.2}$$

with a (purely repulsive) short-range pair potential,

$$u_{ij} = \left\{ \left[\mathcal{P}(\mathbf{r}^{(i)}) \right]^{1/d} \left\| \mathbf{r}^{(i)} - \mathbf{r}^{(j)} \right\|_{\alpha} \right\}^{-m}, \tag{5.3}$$

where for a positive-definite matrix α we defined the α-norm of vector $\mathbf{r}$ by

$$\|\mathbf{r}\|_{\alpha} := (\mathbf{r}^{\mathrm{T}} \alpha \mathbf{r})^{1/2} \tag{5.4}$$

The choice of the adjustable parameter m is probably not important, as long as the potential is truly "short-range", which can be achieved by, e.g., $m = 9 + d$. The role of α will be clarified later.

Due to the strong short-range repulsion, the particles $\mathbf{r}^{(j)}$ are expected to arrange themselves locally to resemble a quasi (i.e. not quite perfect) closed packed structure. Moreover, the lack of attractive terms in the energy functional (these terms were included in the original formulation[38]) enormously simplifies the energy landscape that now has only a small number of local minima which are all structurally equivalent. At the same time, the functional form of u_{ij} is the key to maintaining the scaling law (5.1), i.e., defining the distance between the nearest neighbors in accordance with the local density of points $\mathcal{P}(\mathbf{r})$. Due to the absence of the attractive terms, there is no need to normalize $\mathcal{P}(\mathbf{r})$. To this end, the minimization of $\mathcal{U}$ can be carried out by the simulated annealing method,[74] in which case one can conveniently move a single particle at a time, subsequently exploiting the pairwise nature of the energy functional.

5.3 Assessment of Quasi-Regularity

In higher dimensions we can no longer visually inspect the grid to assess its quasi-regularity. Although we can continue to compute properties of the distribution such as the moments, this is both computational expensive and not sufficiently accurate in higher dimensional. Likewise we could monitor the total energy of the system as a function of

iteration, however this does not guarantee that the grid is fully optimized.

To more accurately assess the "local regularity" of a set of points $\{\mathbf{r}^{(i)}\}$ $(i = 1, ..., N)$, we consider the radial pair correlation function (more precisely, the corresponding histogram) scaled with respect to the distribution function $\mathcal{P}(\mathbf{r})$:

$$g_{\text{sc}}(r) := \frac{1}{N} \sum_{i=1}^{N} \sum_{j \neq i} \delta \left(r - \frac{\left\| \mathbf{r}^{(i)} - \mathbf{r}^{(j)} \right\|_{\alpha}}{r_{\min}(\mathbf{r}^{(i)})} \right), \tag{5.5}$$

The constant, κ, in Eq. (5.1) is generally unknown, but in order to make Eq. (5.5) meaningful we can replace it by its lower bound estimate, e.g.,

$$\kappa = \frac{1}{N} \sum_{j=1}^{N} \mathcal{P}(\mathbf{r}^{(j)})^{1/d} r_{j,\min}, \tag{5.6}$$

where the actual nearest neighbor distance for j-th particle is

$$r_{j,\min} := \min_i \left\| \mathbf{r}^{(i)} - \mathbf{r}^{(j)} \right\|_{\alpha} \quad (j = 1, ..., N). \tag{5.7}$$

To this end, the sharpness of the first peak in $g_{\text{sc}}(r)$ can be used to assess the local regularity (condition II), and its appearance at $r \sim 1$, to assess how well condition II is satisfied.

5.3.1 Two-Dimensional Numerical Examples of Quasi-Regularity

The radial pair correlation function gives information on the distribution of nearest neighbors. We can again use the two-dimensional standard normal distribution Eq. (3.5) to understand the radial pair correlation function. If there is no underlying structure associated with the nearest neighbors in the grid, the radial pair correlation function is not expected to have any structure. And as shown in Fig. 5.2, none of the methods relying on random points have a defined peak in their respective histograms.

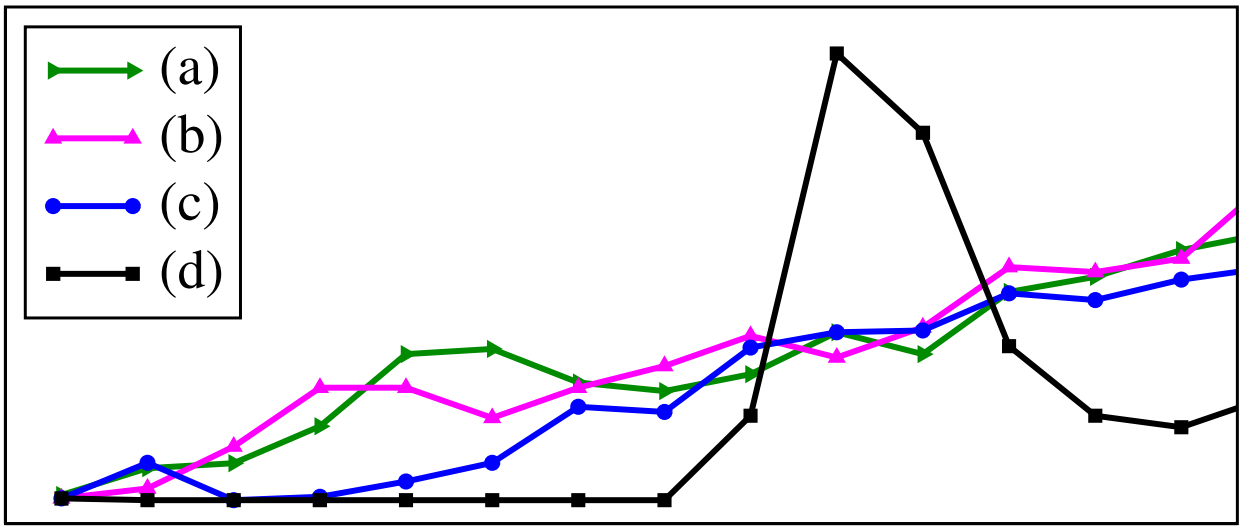

Figure 5.2: Histograms of the radial pair correlation functions for the (a) uniform pseudo-random plus rejection, (b) uniform quasi-random plus rejection, (c) a true quasi-random sequence generated from the Beasley-Springer-Moro transformation, and (d) quasi-regular grid, applied to the two-dimensional standard normal distribution. See Fig. 3.3 for the associated grids.

5.4 Morse Analysis in Higher Dimensions

With the new quasi-regular ansatz and the improved metrics to ensure quasi-regularity, one can begin to tackle the challenging problem of higher-dimensional quantum dynamics calculations. The distribution function is no longer required to be normalized, we therefore elect to use the following distribution:

$$\mathcal{P}(\mathbf{r}) = \begin{cases} [E_{\text{cut}} + \Delta E - V(\mathbf{r})]^{d/2}, & V(\mathbf{r}) < E_{\text{cut}} \\ 0, & V(\mathbf{r}) \geq E_{\text{cut}} \end{cases} \tag{5.8}$$

where $V(\mathbf{r})$ is the potential energy, and E_{cut} and ΔE are adjusting parameters that depend on the system and the energy range of interest. The ΔE parameter allows us to move the outer-edge of the quasi-regular grid as close to E_{cut} as we desire.

We again use the Morse potential, Eq. (4.22), and the following approaches for distributing a set of Gaussian basis functions:

1. *Direct-product:* A uniformly spaced direct-product grid truncated at E_{cut}.

2. *Uniform quasi-random:* A uniformly distributed 2D low-discrepancy quasi-random sequence (in this work we use the Sobol sequence[35–37]), truncated at E_{cut}.

3. *Uniform pseudo-random+rejection:* Starting with a uniformly distributed pseudo-random sequence $\mathbf{r}^{(i)}$ in a sufficiently large domain, one retains only the points that satisfy the inequality $\mathcal{P}\left(\mathbf{r}^{(i)}\right)/\mathcal{P}_{\text{max}} > \xi_i$, where ξ_i is a random number uniformly distributed in the $[0, 1]$ interval.

4. *Uniform quasi-random+rejection:* Same as the above, but $\mathbf{r}^{(i)}$ is a 2D Sobol sequence.

The results for $N = 350$ comparing the various methods are shown in figures 5.3 and 5.4. The top two panels in Fig. 5.3 show two types of uniform grids: a direct-product grid and quasi-random grid. While the quasi-random grid seems to have a somewhat better appearance near the edges, the main drawback of both grid layouts is that too many points

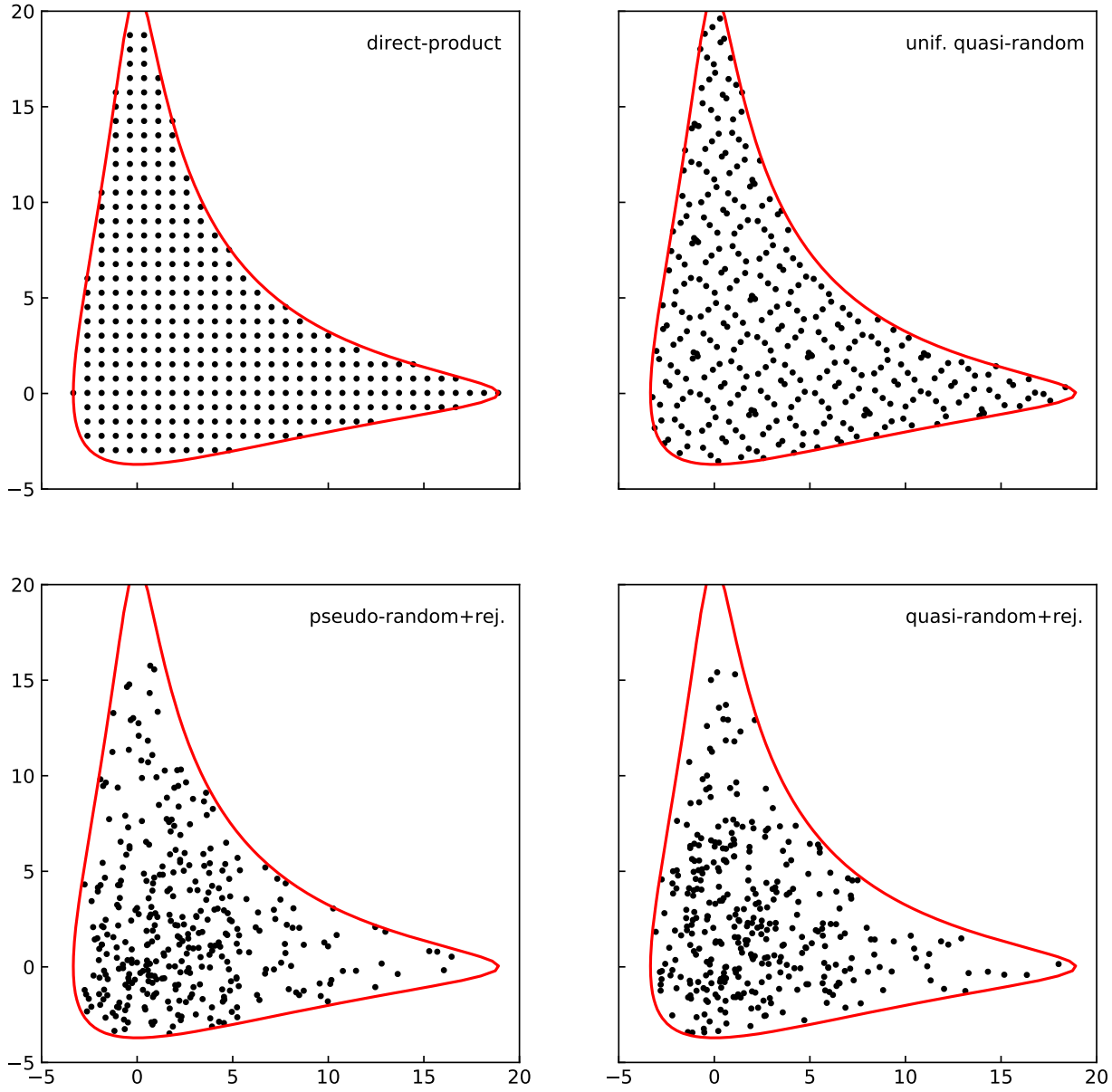

Figure 5.3: Different methods (see text) to generate $N = 350$ grid points for the 2D Morse potential (4.22) within the cutoff range $V(\mathbf{r}) < E_{\text{cut}} = 11.5$ (indicated by the red contour line). The two top panels show uniformly distributed grids. The non-uniform grids in the two bottom panels follow the distribution, $\mathcal{P}(\mathbf{r})$, defined by Eq. (5.8) ($\Delta E = 1.0$).

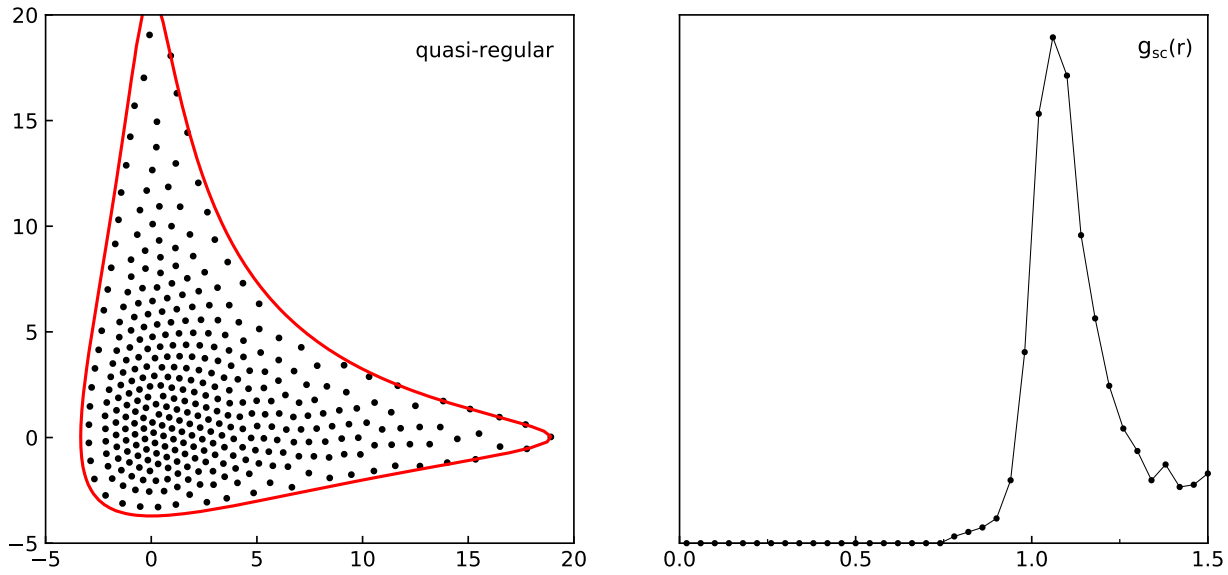

Figure 5.4: Sampling of the Morse potential (see the caption in Fig. 5.3) by a quasi-regular grid. The compromise between achieving local regularity and the desired distribution $\mathcal{P}(\mathbf{r})$ is assessed by the sharpness of the peak at $r \sim 1$ in the radial correlation function, $g_{\mathrm{sc}}(r)$.

are wasted in the region (close to the cutoff line) where the wavefunctions are smooth and less oscillatory. As a consequence, given the fixed total number of points $N = 350$, both grids are too sparse in the central region where the wave functions are oscillatory and need a dense grid for an adequate representation. The bottom left panel in Fig. 5.3 shows a 2D grid generated by a pseudo-random sequence distributed according to the desired distribution function (Eq. (5.8)). The clustering of grid points and presence of gaps throughout the domain of interest is apparent. The bottom right panel shows the grid obtained by the rejection method from the originally uniform 2D Sobol sequence (i.e., the sequence that appears in the top right panel). Yet, the bottom two panels look very similar due to the rejection process. To construct this 350-point grid a large number of points (∼10 000) had to be rejected leading to an almost complete loss of correlations between the remaining points, consequently bringing back the unwanted gaps.

To this end, Fig. 5.4 shows the quasi-regular grid result using the same number ($N =$ 350) of points. The density of the quasi-regular grid points is consistent with the desired distribution (Eq. (5.8)) and is locally regular (i.e., locally has uniform spacing between

Table 5.1: Parameters used to construct the Gaussian bases for the higher dimensional Morse potential calculations. The choice of parameters can be explored to obtain the desired accuracy.

d	N	E_{cut}	γ	α_0
4	2000	7	1.7	0.5
5	6000	7	1.7	0.7
6	9000	7	1.7	0.9

nearest neighbors). The appearance of the quasi-regular grid, at-least visually, is ideal. In addition, the quality of this quasi-regular grid is confirmed by the radial correlation function $g_{\text{sc}}(r)$ which does show a relatively sharp peak at $r \sim 1$.

With this metric we can now explore the vibrational spectra for higher dimensional Morse oscillators, namely $d = 4, 5, 6$, corresponding to the dimensionality we will consider in Chapters 6 and 7. We are unaware of any reported values in the literature for these higher-dimensional Morse calculations, so we have chosen reasonable parameters to roughly maintain the accuracy observed in more established lower-dimensional cases. For convenience we set the ω values as $\omega_1 = 0.2$, $\omega_2 = 0.18$, $\omega_3 = 0.16$, $\omega_4 = 0.14$, $\omega_5 = 0.12$, and $\omega_6 = 0.1$ respectively.

The potential energy matrix elements were computed using a three-point Gauss-Hermite sparse quadrature for each degree of freedom. While the integrals were originally attempted using quasi-Monte Carlo methods (not reported here), the scaling of the sparse quadrature does make quadrature-based approaches reasonable for higher-dimensional systems. The scaling of the quadrature grid is shown in Fig. 5.5 assuming a Gauss-Hermite order parameter of 3.

We have refrained from exhaustively optimizing these calculations for transparency, however, any individual system can be further fine-tuned to achieve a desired accuracy. The parameters used to construct these grids are shown in Table. 5.1. As shown in Fig. 5.6 we observe that high accuracy quantum dynamics calculation can be performed in these higher-dimensional spaces. We were unable to produce a similar accuracy with the alternate grid

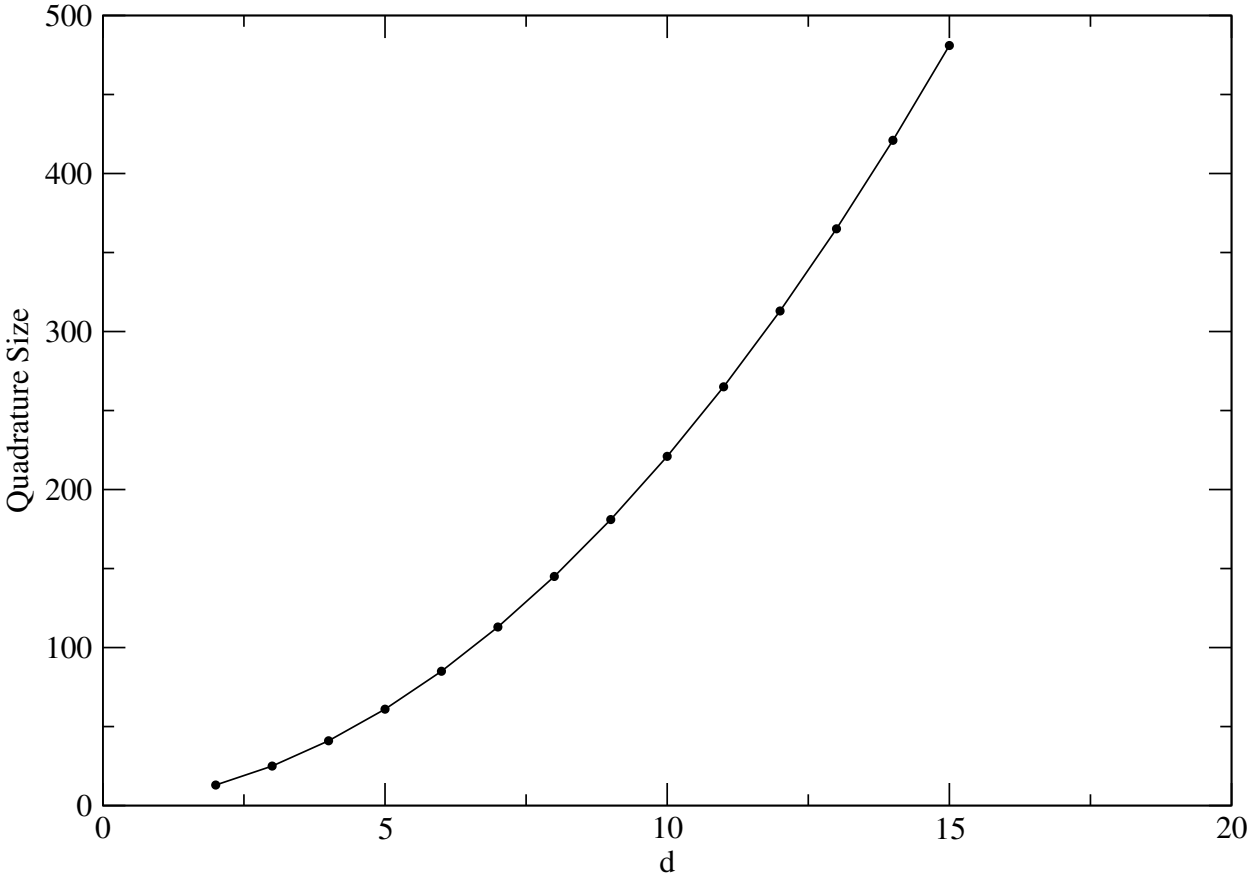

Figure 5.5: The scaling of a sparse quadrature grid (with an order parameter of 3) as the dimensionality is increased.

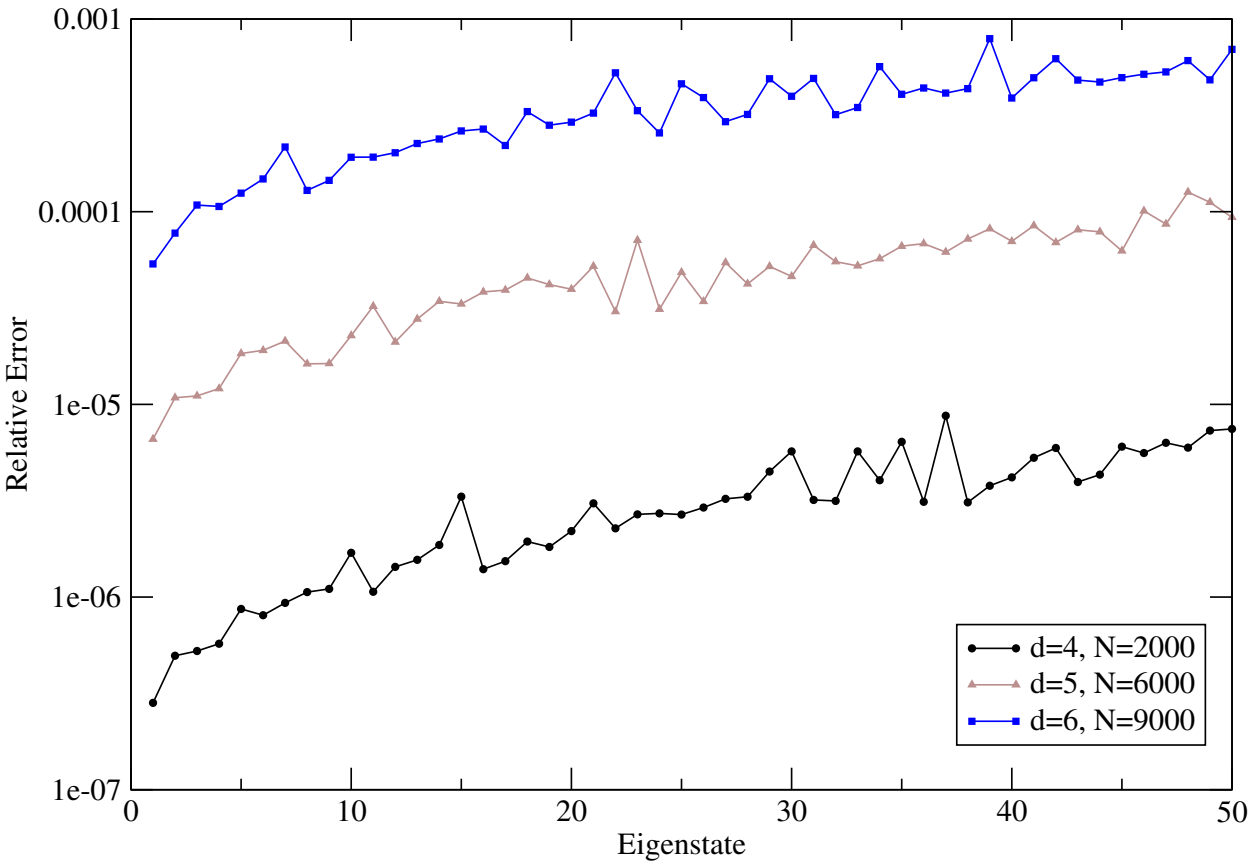

Figure 5.6: Relative error for the eigenenergies of the Morse potential in a progressively larger configuration space. (Note the logarithm scale.)

methods due to the large values of N that would have been required, consistent with our understanding of the hypersphere problem (see Eq. (4.4)). As expected, in higher dimensions the cutoff contour needs to be reduced to have a dense enough grid. However, if the cutoff contour is made too small, the total accuracy will be reduced because the relevant regions in configuration space are not being sampled.

5.5 Conclusions

The original quasi-regular ansatz was modified to be both simpler and more efficient, and notably no longer contains any fitting parameters. Replacing the pseudo-potential with a solely repulsive form we were able to greatly reduce the computational effort needed to generate a quasi-regular structure. We then defined a radial pair correlation function to

quantitatively asses the quasi-regular properties of a grid. This approach is less expensive to compute and more accurate compared to the previous method of numerically computing the lower moments of the distribution. Finally the vibrational spectra for four-dimensional, five-dimensional, and six-dimensional Morse oscillators were computed with relatively high accuracy, and relatively small basis sets when compared to the configuration space.

Chapter 5

Vibrational Spectra of Formaldehyde

This chapter applies the optimized distributed Gaussian basis resulting from a quasi-regular grid to the computational of the quantum vibrational spectra of formaldehyde.[75] By representing the molecule in a set of internal coordinates defined by the bond lengths and angles, the problem is reduced from a twelve-dimensional Cartesian space, to a six-dimensional configuration space. The kinetic energy operator is computed numerically, making the overall approach highly generalizable, and could easily be implemented for other atomic systems of similar size. The vibrational spectra computed using the DGB derived from a quasi-regular gird is again shown to be highly accurate, while simultaneously requiring far less basis functions when compared to recent methods in the literature. The approach also contains no over-fit parameters or non-transparent optimizations ensuring the results are reproducible.

6.1 Collocation Method for Quantum Dynamics

In their recent work Manzhos and Carrington[72] presented impressive results computing the vibrational spectrum of formaldehyde, H_2CO, using the collocation method.[76;77] In the present work we follow the latter paper very closely, for this reason from here onward we will refer to it as **M&C**. In their paper they demonstrated that the method could be both improved and simplified further by using the most convenient set of internal coordinates, and evaluating the kinetic energy matrix elements numerically. This drastically reduces the problem from a daunting twelve-dimensional quantum dynamics calculation to a simpler (but still cutting-edge) six-dimensional problem. Likewise the numerical evaluation of the kinetic energy operator avoids the need for complicated algebraic manipulations and makes the approach promising for other systems.

Assuming any convenient internal coordinate system $\mathbf{r} = (\mathbf{r}_1, \cdots, \mathbf{r}_d)$ that describes a

molecule ($d = 3N_{\text{atoms}} - 6$) or its part, the vibrational Hamiltonian is given by

$$\widehat{H} = \widehat{T} + V(\mathbf{r}), \tag{6.1}$$

in which the kinetic energy operator is written using the $3N_{\text{atoms}}$ Cartesian coordinates

$$\widehat{T} = -\sum_{i=1}^{3N_{\text{atoms}}} -\frac{\hbar^2}{2m_i}\frac{\partial^2}{\partial \mathbf{x}_i^2}. \tag{6.2}$$

Next consider a set of grid points $\mathbf{r}^{(i)} \in \mathbb{R}^d$ $(i = 1, \cdots, N)$, where each point is associated with a basis function, localized in its vicinity. A convenient (but in no way required) choice corresponds to a set of Gaussians,

$$\Phi_i(\mathbf{r}) := \exp\left[-\left\|\mathbf{r} - \mathbf{r}^{(i)}\right\|^2_{\alpha^{(i)}}\right] \qquad (i = 1, ..., N), \tag{6.3}$$

where the norm $\|...\|_{\alpha^{(i)}}$ is defined by Eq. (5.4) with the coordinate dependence of the width matrix $\alpha^{(i)}$ to be specified later.

When using the collocation method one defines a new grid containing a set of collocation points $\mathbf{r}^{(i)} \in \mathbb{R}^d$ $(i = 1, \cdots, N_c)$ at which the Schrödinger equation must be satisfied,

$$(\widehat{H} - E)\Psi\left(\mathbf{r}^{(j)}\right) = 0. \tag{6.4}$$

It is standard (although not required) to set the first N points to coincide with the center of the localized functions (the Gaussian centers in this work), and the remaining points are then generated separately (see below). By defining the overlap and Hamiltonian matrices,

$$\mathbf{S}_{ji} := \Phi_i\left(\mathbf{r}^{(j)}\right); \quad \mathbf{H}_{ji} := \widehat{H}\Phi_i\left(\mathbf{r}^{(j)}\right), \tag{6.5}$$

and expanding the eigenfunctions using the Gaussian basis,

$$\Psi(\mathbf{r}) = \sum_{i=1}^{N} \mathbf{c}_i \Phi_i(\mathbf{r}), \tag{6.6}$$

we arrive at the rectangular generalized eigenvalue problem,

$$(\mathbf{H} - E\mathbf{S})\mathbf{c} = 0 \tag{6.7}$$

That is, each eigenvalue E, in Eq. (6.7) is associated with N_c equations and N unknown coefficients $\mathbf{c} = (\mathbf{c}_1, ..., \mathbf{c}_N)^{\mathrm{T}}$. One practical way to solve this (overdetermined) problem is to reduce it to a square $N \times N$ generalized eigenvalue problem as, e.g.,[72]

$$(\mathbf{S}^{\mathrm{T}}\mathbf{H} - E\mathbf{S}^{\mathrm{T}}\mathbf{S})\mathbf{c} = 0. \tag{6.8}$$

Note here that in the special case of $N_c = N$, one does not need to multiply by $\mathbf{S}^{\mathrm{T}}$, a step which is not only expensive (scales as $\sim N^3$), but also makes the original problem (6.7) more ill-conditioned. However, given a fixed Gaussian basis, increasing the number of collocation points, N_c, should improve the accuracy of the computed eigenvalues noticeably (see Fig. 6.4), while the matrix construction is still comparable or (depending on N_c) even less expensive than the solution of the (non-symmetric) generalized eigenvalue problem.

In order to avoid very complicated algebra involving internal coordinates, $\mathbf{r} = \mathbf{r}(\mathbf{x})$, the action of the kinetic energy operator (6.2) on the basis functions at each collocation point, i.e., $\widehat{T}\Phi\left(\mathbf{r}^{(j)}\right)$, is evaluated numerically by finite difference in the Cartesian space.[72]

Although no integrals involving the potential energy surface (PES) are computed, the method is numerically exact as long as the evaluation of $\nabla^2\Phi_i(\mathbf{r}^{(j)})$ through finite difference is accurate and the basis is large enough. Unfortunately "large enough" is ambiguous as each system considered will need to be studied for convergence.

Here we assume that an optimal distribution function $\mathcal{P}\left(\mathbf{r}^{(i)}\right)$ for the positions of the

Gaussian centers is defined using Eq. (5.8). Note that M&C use a similar distribution function, however, they do not scale according to the dimensionality (i.e. $d/2$) but follow the approach of Garaschuk and Light and set $\gamma = 1$. There are a number of parameters that can be optimized depending on the exact approach taken. For the DGB resulting from a quasi-regular grid we have not found this difference in the distribution function to meaningfully alter the accuracy of the resulting eigenvalues. This could indicate a more optimal form of the distribution function exists, however, this is beyond the scope of this study. The positive-definite matrix α that appears in the definition of the norm in Eq. (5.4) is set to be diagonal

$$\alpha := \mathrm{diag}\{1/\Delta r_k\} \tag{6.9}$$

with Δr_k defining the range spanned by the Gaussian centers along the k-th degree of freedom ($k = 1, \cdots, d$).

As mentioned in Chapter 4, there is a large body of research optimizing distributed Gaussian basis sets.[69;78–81] The quality of a given basis set does very much depend not only on how the Gaussian centers are distributed, but also on the expression used to compute the Gaussian widths, α_i. This choice is quite important when working on more complicated systems, a wrong choice for the widths (e.g., too narrow or too wide) may result in poor approximations of the wavefunctions or ill-conditioned matrices, or both.

Clearly, the optimal choice for $\alpha^{(i)}$ must depend on the local distribution of the Gaussian centers around the i-th Gaussian. At the same time, one cannot afford to make the protocol for optimizing the widths matrices $\alpha^{(i)}$ too elaborate or else construction and reproducibility will suffer. In the present case, the procedure of choosing $\alpha^{(i)}$ can be made straightforward[38] since the local arrangement of Gaussian centers is the same everywhere, except for a scaling factor. Consequently, we use the following simplified expression:

$$\alpha^{(i)} := \frac{b\alpha}{r^2_{i,\mathrm{min}}}, \tag{6.10}$$

where $r_{i,\min}$ is the distance to the nearest neighbor from the i-th point (*cf.* Eq. (5.1)) and $b \sim 1$ is the only adjustable parameter.

To this end, we note again that numerical instabilities are often encountered when DGBs are employed, especially when using nonuniform grids. For example, when two grid points appear too close, the corresponding Gaussians become linearly dependent. This in turn leads to a large condition number for both the Hamiltonian and the overlap matrices. The quasi-regular ansatz minimizes this very problem as it eliminates the clustering of the grid points. In addition, Eq. (6.10) assures that all the adjacent Gaussian's have a similar overlap.

6.2 Numerical Details for Formaldehyde

In our numerical demonstration we consider the four-atom molecule of formaldehyde, H_2CO. This choice was motivated by M&C[72] who used essentially the method formulated in the previous section. We implemented the same PES, i.e., that from Carter,[82] and the same set of bond-angle internal coordinates as shown in Fig. 6.1 ($r_{\mathrm{CO}}, r_{\mathrm{CH}_1}, r_{\mathrm{CH}_2}, \theta_1, \theta_2, \phi$).

The difference is in the choice of the points defining the Gaussian centers $\mathbf{r}^{(i)}$ and the Gaussian widths matrices $\alpha^{(i)}$. M&C placed their Gaussian's using the same procedure as that implemented to construct the bottom right panel of Fig. 5.3, i.e. the uniform quasi-random+rejection scheme. In the present case, the Gaussian centers are placed using a quasi-regular grid. M&C used the same diagonal matrix α for all Gaussian's but the values for its elements were set in a non-transparent fashion which possibly resulted from an additional optimization not explained in the paper.

In the present case, the only adjustable parameter for the Gaussian widths was b (*cf.* Eq. (6.10)) which was then set to $b = 1$ for all the reported results. However, additional calculations (not reported) confirmed that the stability of the results depends on the specific parameters used, meaning further optimization is always possible for a given system. Of note, a larger basis will have a larger region of stability for a given value of b, and this stability region decreases as the basis size decreases.

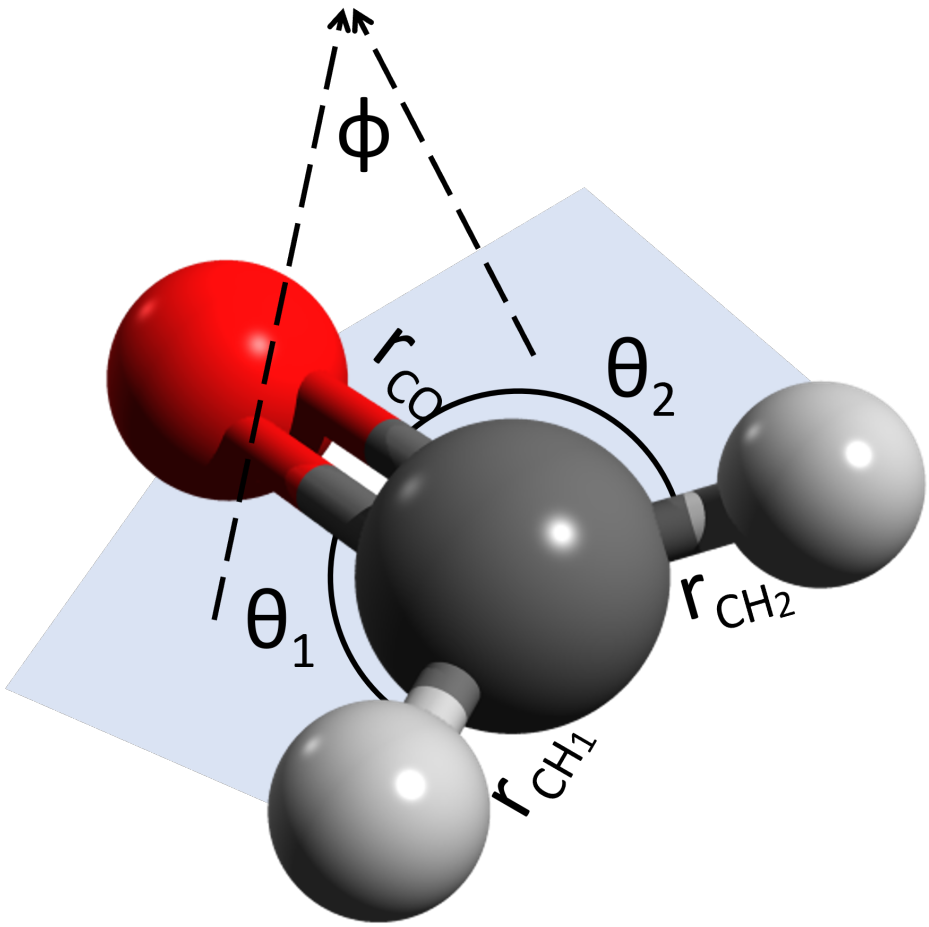

Figure 6.1: A schematic of the three bond lengths and associated angles used to define the formaldehyde molecule.

Table 6.1: The parameters used to construct the DGB sets for H_2CO. An excessive number of collocation points were used to ensure convergence. Minimal effort was made to optimize these parameters.

data set	QRG10K	QRG15K	QRG20K
N	10 000	15 000	20 000
N_c	500 000	750 000	1 000 000
E_{cut} (cm^{-1})	15 000	15 000	15 000
ΔE (cm^{-1})	3000	5 000	5 000
b	1.0	1.0	1.0

Several calculations were performed using N = 10K, 15K, and 20K (i.e., 10 000, 15 000 and 20 000). The parameters of these calculations are given in Table. 6.1 All of the grids were constructed according to the following simple protocol.

Begin by generating an initial set of points $\{\mathbf{r}^{(i)}\}$ $(i = 1, ..., N)$ with Metropolis Monte Carlo using the distribution function $\mathcal{P}(\mathbf{x})$ (Eq. (5.8)), where each point is selected after $N_{\text{skip}} = 1000$ Monte Carlo steps to reduce correlations between points. Note though that the present choice of N_{skip} is not essential since this grid is then subject to optimization. This set of points is used to determine the ranges Δr_k which define the norm $\|...\|_\alpha$ (Eq. (5.4)). In the next step a "greedy simulated annealing minimization" (i.e., the only accepted moves are those resulting in a reduction of the total energy) is applied to the set $\{\mathbf{r}^{(i)}\}$ by minimizing the energy functional $\mathcal{U}(\{\mathbf{r}^{(i)}\})$ (Eq. (5.2)). The convergence of the minimization is monitored by observing the decrease of $\mathcal{U}(\{\mathbf{r}^{(i)}\})$ and by examining the scaled pair correlation function $g_{\text{sc}}(r)$ (Eq. (5.5)). As an example, in Fig. 6.2 we show $g_{\text{sc}}(r)$ for the QRG15K set. The sharp peak at $r \sim 1$ indicates both the achieved local regularity of the grid and its consistency with the given distribution function $\mathcal{P}(\mathbf{x})$.

The additional collocation points were generated using the quasi-random+rejection scheme with the same distribution function $\mathcal{P}(\mathbf{r})$. We note though that switching to the pseudo-random+rejection scheme did not make a noticeable difference (not reported here). Note also that M&C used a quasi-random+rejection sequence for the collocation points, with the first N points in the sequence defining the Gaussian centers.

To make sure that insufficient averaging over the collocation grid would not contribute to the error, the maximum number of collocation points was set to a large value, namely, $N_{c,max} = 50N$. The convergence with respect to N_c was then monitored by solving Eq. (6.8) for the intermediate values of N_c. As in Ref. 72 we report the results for the lowest 50 eigenenergies.

As suggested by M&C here the action of the kinetic energy operator (6.2) on the basis functions at each collocation point, i.e., $\widehat{T}\Phi_i\left(\mathbf{r}^{(j)}\right)$, is evaluated numerically by finite

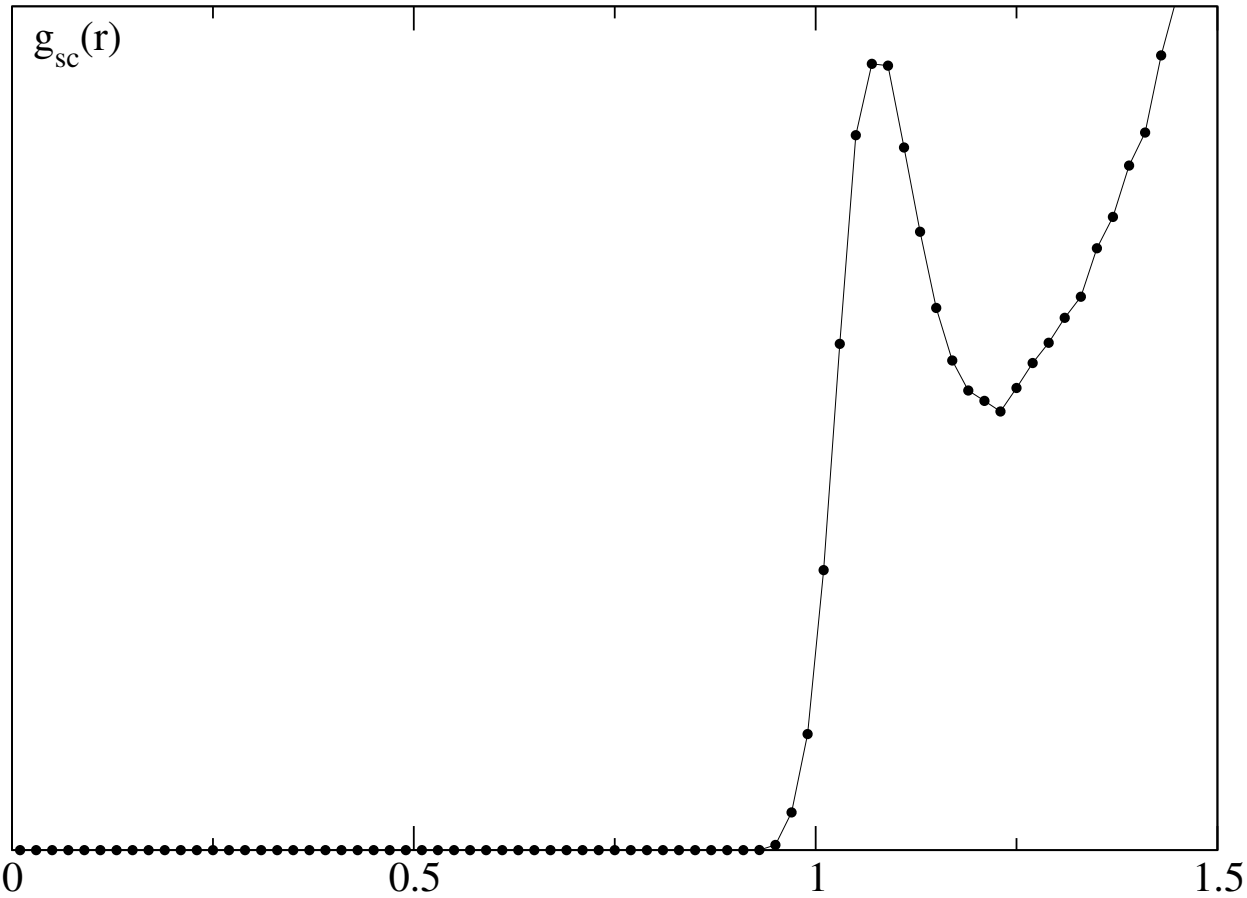

Figure 6.2: The quality of the 6d QRG constructed for H_2CO is assessed using the scaled radial pair correlation function (set QRG15K: $N = 15\,000$, $\Delta E = 5\,000$ cm^{-1}, $E_{\rm cut} = 15\,000$ cm^{-1}).

difference in the Cartesian space using a five-point stencil. This allows one to avoid very complicated algebra involving the representation of the Laplacian in the bond-angle internal coordinates, and also makes the algorithm very general, i.e., not depending heavily on the choice of the coordinate system.

The generalized eigenvalue problem (6.8) is not symmetric and hence its eigenvalues are either real or come in complex-conjugated pairs. However, the latter situation indicates poor convergence, i.e., well-converged eigenenergies are always real.

6.3 Vibrational Spectra of Formaldehyde

Given the eigenenergies of formaldehyde have already been reported by M&C,[72] the purpose of this section is to use this well-established numerical example as a benchmark to further assess the methodology and demonstrate the superiority of a distributed Gaussian basis derived using the quasi-regular ansatz.

There are several factors contributing to the convergence of the computed eigenenergies using the techniques described above. Besides the quality of the Gaussian basis set and the size and extension of the collocation grid we would like to focus first on the numerical errors associated with the evaluation of the Hamiltonian matrix elements. Since the potential energy integrals are avoided, the only numerical error is due to the use of finite difference in the implementation of the Laplacian operator. This simplicity comes with a price, namely: we were unable to achieve very high accuracy, regardless of how elaborate the finite difference scheme was (i.e., either using three-point, five-point or seven-point stencil). For example, Fig. 6.3 shows the differences in the eigenvalues using the five-point stencil scheme with three different step sizes: 0.01, 0.001, and 0.0001 (mass-scaled coordinates, atomic units). Apparently, the corresponding error increases with the energy from less than 1 cm^{-1} for the lowest eigenenergies to about 2 cm^{-1} for some of the highest ones. Consequently, one cannot expect the overall error in the eigenenergies to be smaller than the finite-difference error. We noticed though that when the basis size is increased, the finite-difference error

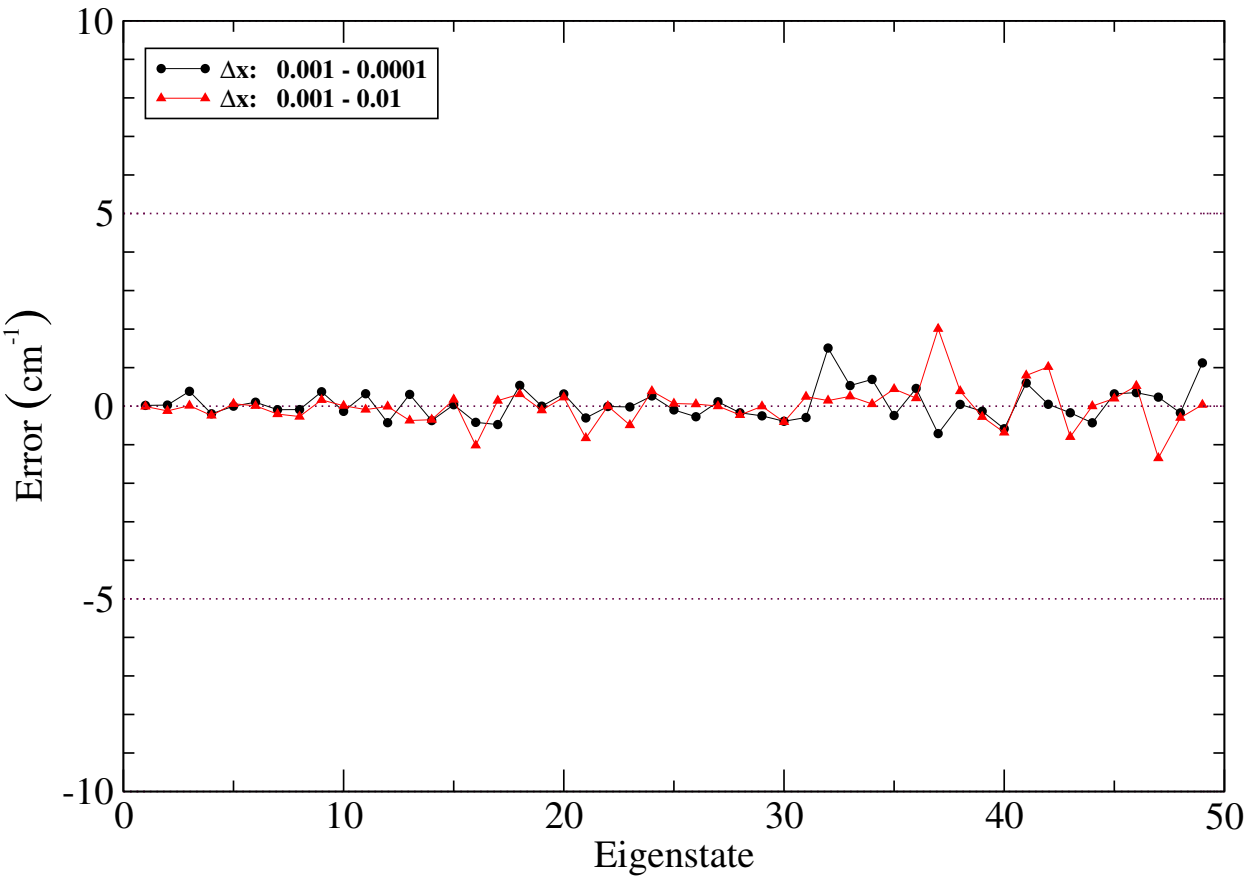

Figure 6.3: The differences between the eigenenergies (using the QRG15K basis set) when the five-point stencil method is applied while varying the step size Δx.

decreases.

In the special case of $N_c = N$ (i.e., when the collocation points coincide with the Gaussian centers), the finite-difference error turns out to be negligibly small for either the three-point or five-point stencil. This can be explained by the fact that in this special case the kinetic energy matrix is diagonally-dominated with the diagonal elements obtained by evaluating the second derivatives of the Gaussians at their maxima. At the maxima the quadratic approximation is expected to be excellent, assuming the step size, Δx, is not too large. However, when excess collocation points are generated they can appear arbitrarily close together, resulting in matrices that are not diagonally dominated but have large off-diagonal elements. These types of matrices could amplify the existing errors.

Although the case of $N_c = N$ is noticeably faster as it avoids matrix multiplication by

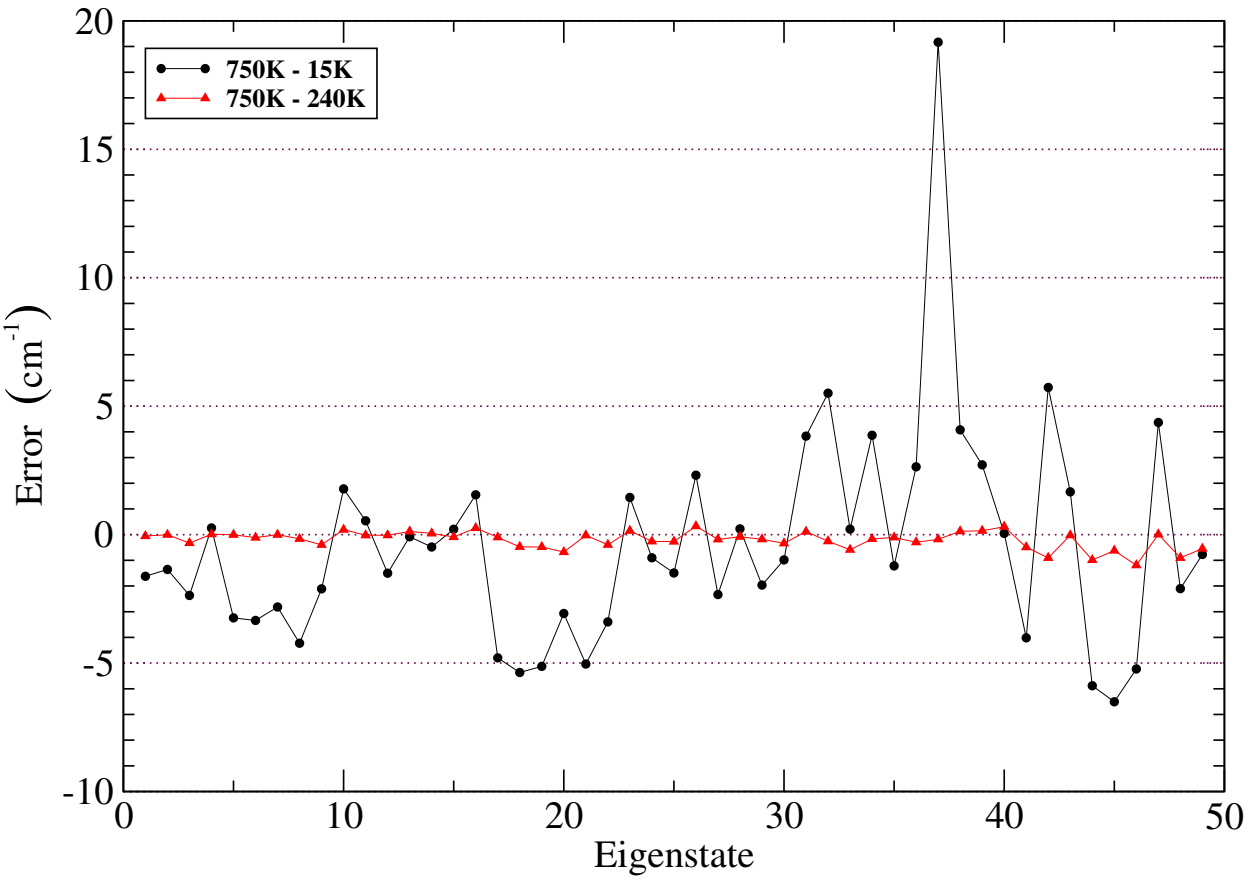

Figure 6.4: The intrinsic error for set QRG15K obtained by taking the difference between the computed eigenenergies when using only the Gaussian centers ($N_c = N$) and the eigenenergies when using the largest number of collocation points ($N_c = 750\ 000$).

$\mathbf{S}^{\mathrm{T}}$ (*cf.* Eq. (6.8)) and, in addition, it does not suffer from the finite-difference error, Fig. 6.4 clearly demonstrates that using a sufficiently large N_c ($\sim 20N$) allows one to substantially reduce the eigenenergy errors compared to the case of $N_c = N$.

This improvement is clearly a large enough advantage to merit the added computational cost. We must note however, that the work of M&C does not seems to have any meaningful benefit derived from their added collocation points.

As shown in Fig. 6.5 the difference between using only their Gaussian centers, or a large number of collocation points is on the order of one wavenumber. The same trend is found when reviewing their smaller basis set calculations too. It is difficult to determine why their results do not seem to improve with N_c, however, they are clearly not taking advantage of

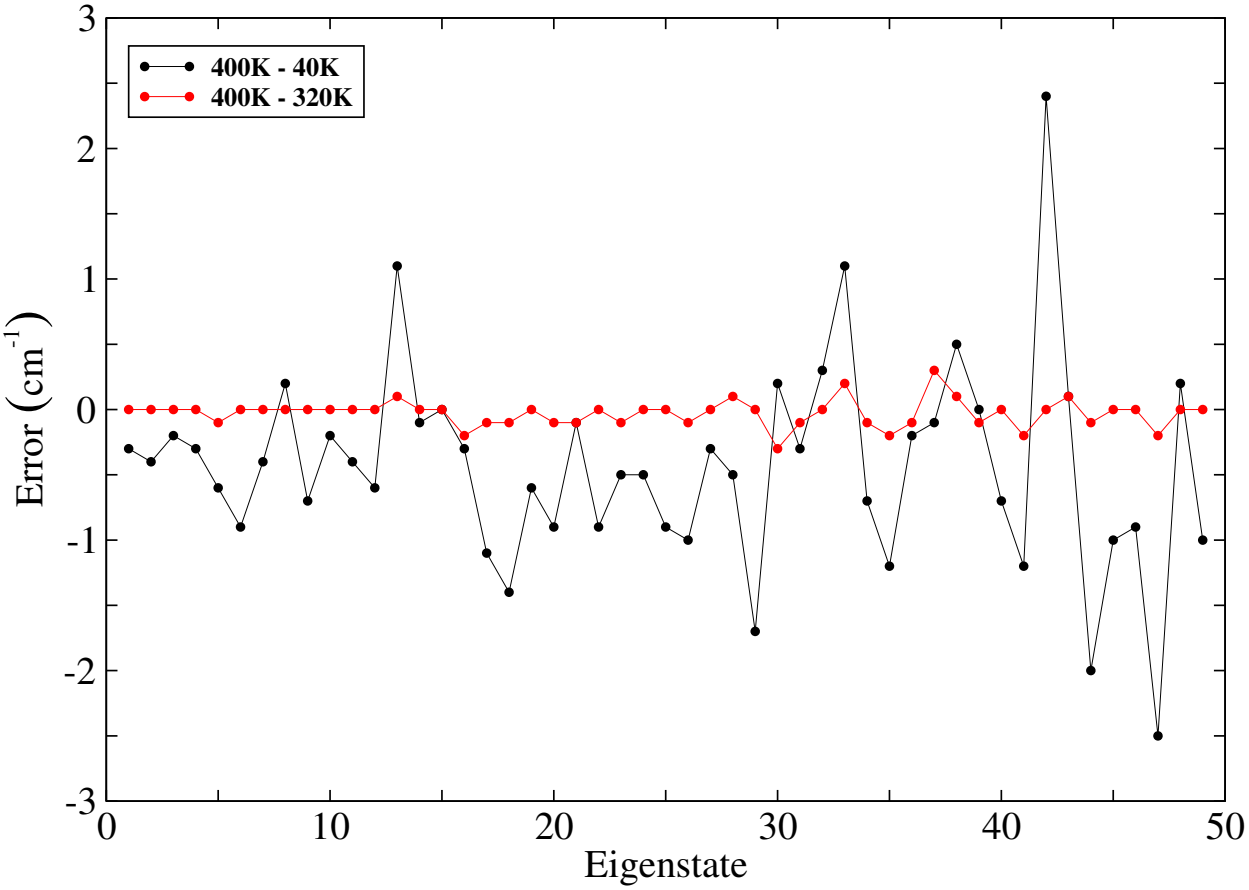

Figure 6.5: M&Cs convergence with respect to N_c for their $N = 40\ 000$ basis. The error is computed through the difference from their largest calculation: $N_c = 400\ 000$.

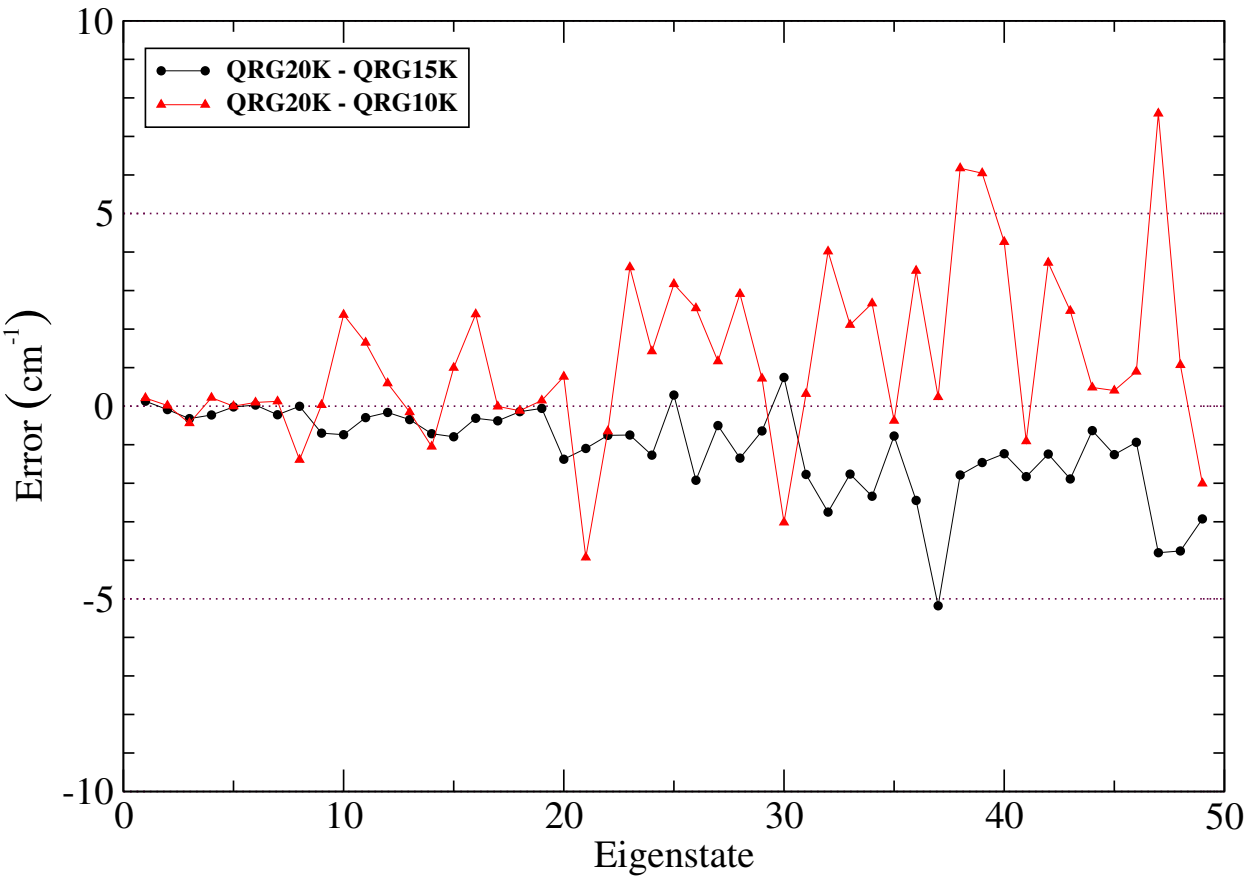

Figure 6.6: Intrinsic convergence: the eigenenergy differences between QRG10K and QRG15K data sets relative to QRG20K set.

the method. It could be due to the range in which they generate their collocation points (i.e. not sampling any new space compared to their basis functions), or due to the condition of their matrices.

To this end, Table 6.2 presents our results for the first 50 eigenenergies using N = 10K, 15K, and 20K, together with the most accurate results of M&C using N = 40K and N_c = 400K. Overall, the agreement is good between all four sets of calculations and is within a few or several wavenumbers. Figures 6.6 and 6.7 visualizes the same information in a graphical form.

More specifically, Fig. 6.6 shows the differences between the eigenenergies of the two pairs of sets, QRG15K-QRG20K and QRG15K-QRG10K; Fig. 6.7 shows the energy differences between our QRG15K data set and N=40K data set from M&C. For reference, Fig. 6.8

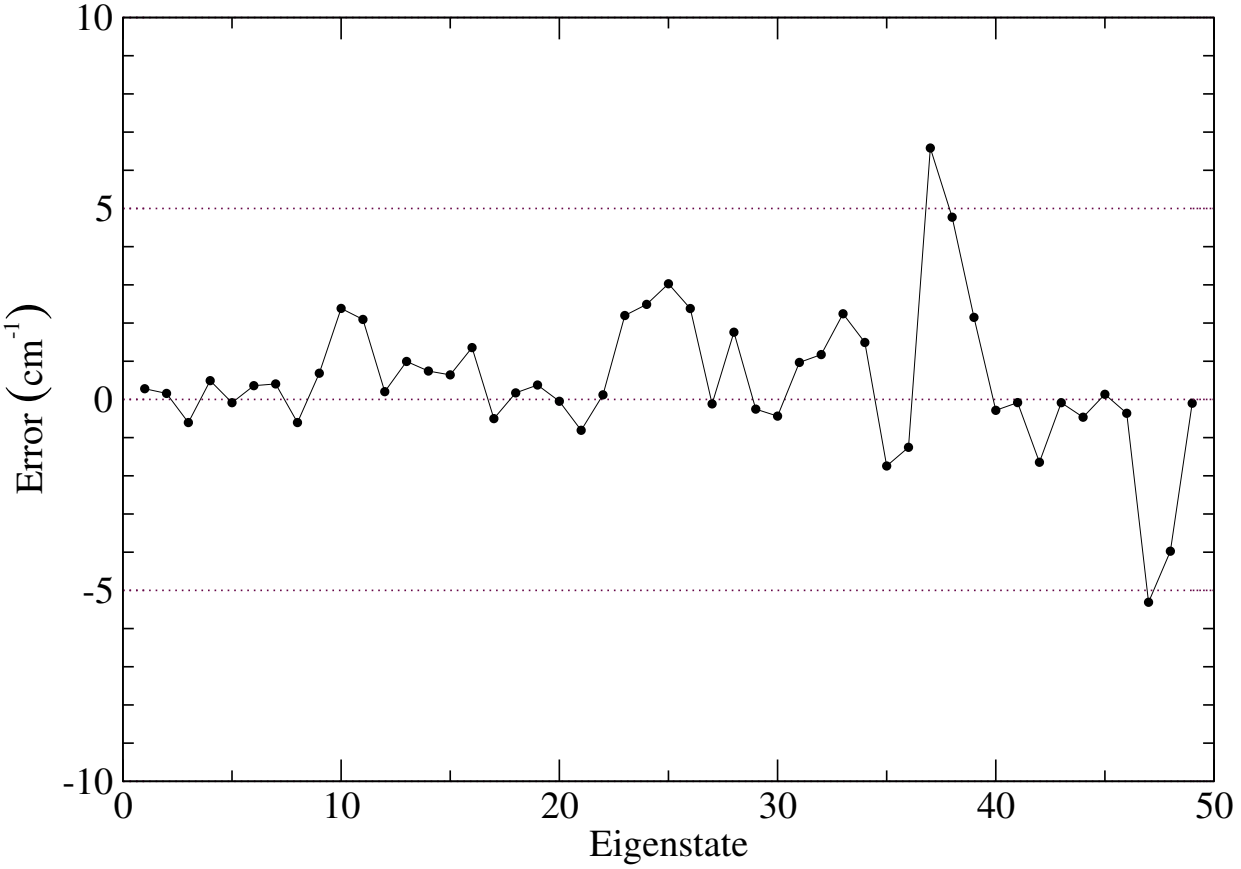

Figure 6.7: The results for the QRG15K set are compared to the largest calculation by M&C,[72] i.e. N=40K N_c=400K, by taking the difference between the two sets of eigenenegies.

shows the intrinsic comparison between the three sets reported by M&C using N=40K, 30K and 25K. We note that the discrepancies between the latter three data sets are within a range similar to the discrepancies between our data sets, however, our method makes use of a much smaller basis size.

Based on these comparisons, we can definitely conclude that using a quasi-regular grid to place the Gaussian basis functions is advantageous compared to the previously used approach[72] based on the quasi-random + rejection scheme with an improvement of about a factor of 3. These improvements will be instrumental in future works where systems with larger dimensionality and/or more complex dynamics (e.g., involving large amplitude motion) are investigated. In such systems it is necessary to have the correct placement of the basis functions, which will be a good test for the methodology.

6.4 Automatic Differentiation

The combination of using internal coordinates to define a quasi-regular grid, and numerical evaluations of the kinetic energy operator make this method general enough to tackle any four atom molecule. However, the inability to increase the accuracy of the kinetic energy evaluations will need to be addressed before a software package is made ready for distribution. One approach would be to compute analytic expressions for the derivatives in the kinetic energy expressions. The internal coordinates are expressed in terms of arc-functions which become incredibly complex when a second derivative is taken. It would be more reasonable to compute the first derivatives which are rather simple, and then compute the second derivatives numerically to reduce the overall error. Analytic expressions for these first derivatives are derived in the appendix.

Another promising avenue for future research would be to replace the finite difference approach with a numerically exact method (see, for example, Ref. 83). Automatic differentation is based on dual numbers which are of the form $v + \dot{v}\epsilon$, where $v, \dot{v} \in \mathbb{R}$ and $\epsilon^2 = 0$, $\epsilon \neq 0$. v is called the primal and $\dot{v}$ is called the derivative.[84] If one considers substituting a

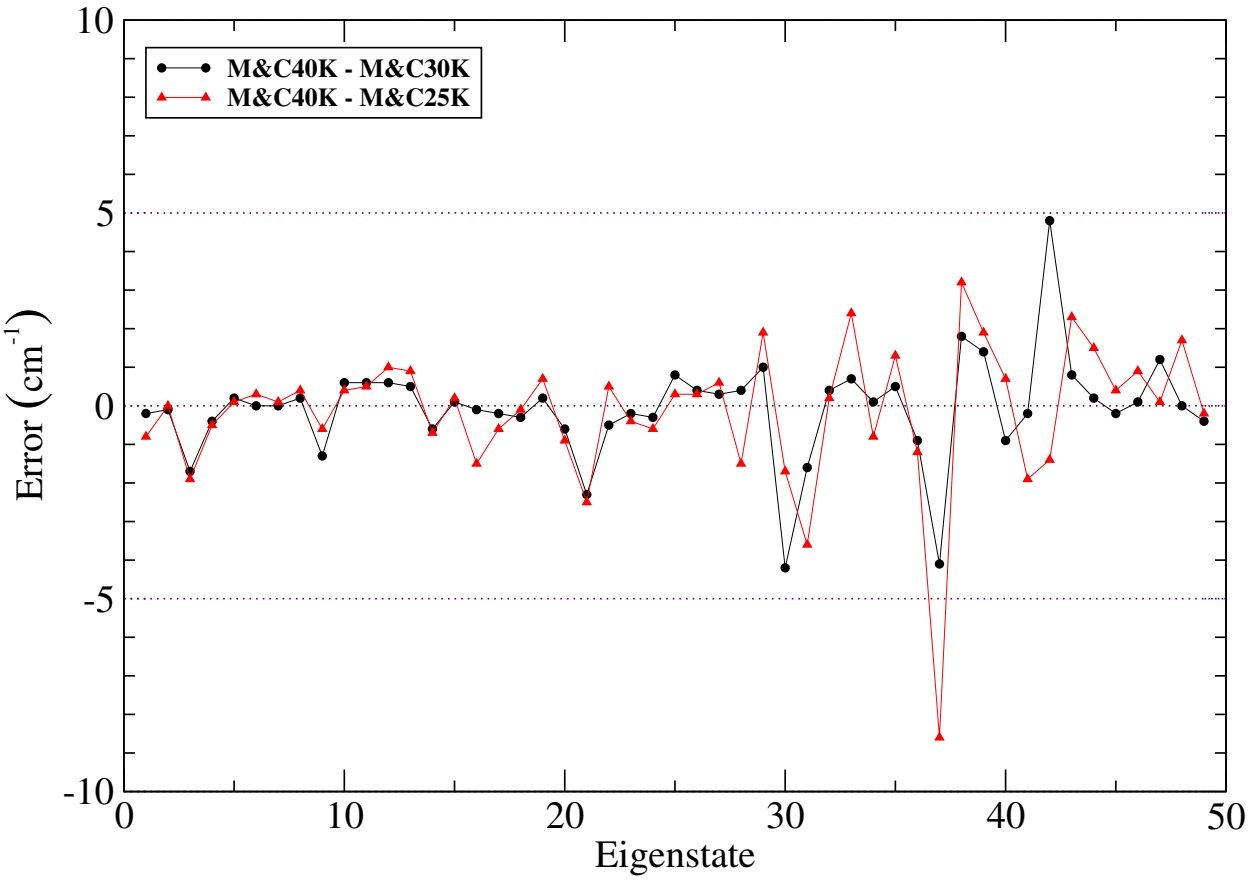

Figure 6.8: The intrinsic convergence from Ref. 72: the eigenenergy differences between the N=25K and N=30K data sets relative to the largest N=40K set and using the collocation grids defined by $N_c = 10N$

Table 6.2: 50 lowest eigenenergies for H_2CO with respect to the ground state energy (first row) for the three basis sets in Table 6.1. All results are in cm^{-1}.

QRG10K	QRG15K	QRG20K	40K(M&C)
5774.24	5774.98	5774.56	5775.3
1166.54	1166.61	1166.75	1166.9
1250.40	1250.44	1250.41	1250.6
1500.47	1500.30	1500.03	1499.7
1746.06	1746.50	1746.28	1747.0
2326.84	2326.88	2326.84	2326.8
2421.62	2421.64	2421.71	2422.0
2497.44	2497.79	2497.56	2498.2
2668.14	2666.90	2666.75	2666.3
2719.18	2719.91	2719.22	2720.6
2775.42	2778.51	2777.80	2780.9
2838.41	2840.30	2840.06	2842.4
2905.07	2905.79	2905.66	2906.0
3000.17	3000.50	3000.02	3001.5
3001.80	3001.35	3000.75	3002.1
3237.85	3239.65	3238.84	3240.3
3468.54	3471.24	3470.93	3472.6
3480.70	3481.20	3480.69	3480.7
3586.04	3586.22	3585.93	3586.4
3674.49	3674.82	3674.64	3675.2
3740.25	3742.34	3741.02	3742.3
3828.80	3826.30	3824.87	3825.5
3887.45	3887.57	3886.80	3887.7
3932.72	3937.00	3936.32	3939.2
3935.10	3937.81	3936.53	3940.3
3989.94	3992.77	3993.11	3995.8
4026.21	4030.62	4028.76	4033.0
4056.47	4058.31	4057.64	4058.2
4079.48	4083.73	4082.39	4085.5
4163.37	4164.65	4164.09	4164.4
4170.13	4166.73	4167.11	4166.3
4193.34	4195.43	4193.66	4196.4
4243.21	4249.72	4247.22	4250.9
4247.25	4251.15	4249.36	4253.4
4331.21	4336.10	4333.88	4337.6
4398.72	4399.54	4398.35	4397.8
4462.42	4468.55	4465.94	4467.3
4495.78	4501.01	4496.02	4507.6
4515.39	4523.12	4521.57	4527.9
4561.92	4569.45	4567.97	4571.6
4618.84	4624.38	4623.11	4624.1
4628.42	4629.58	4627.51	4629.5
4726.54	4732.04	4730.27	4730.4
4729.75	4734.18	4732.22	4734.1
4744.45	4745.66	4744.93	4745.2
4841.30	4843.36	4841.70	4843.5
4924.97	4926.96	4925.87	4926.6
4946.91	4958.41	4954.51	4953.1
4975.21	4980.67	4976.28	4976.7
4982.57	4983.69	4980.56	4983.6

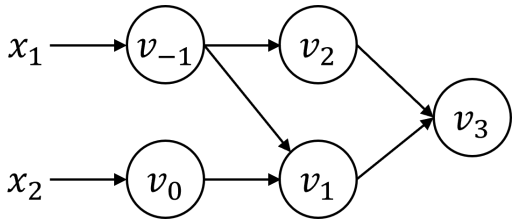

Figure 6.9: The Wengert list used to apply automatic differentation to the function $y = x_1 x_2 + \cos x_1$.

dual number into a differentialable function using the Taylor Series:

$$\begin{aligned} f(a+\epsilon) &= f(a) + \frac{f'(a)}{1!}\epsilon + \frac{f''(a)}{2!}\epsilon^2 + \frac{f'''(a)}{3!}\epsilon^3 + \dots \\ &= f(a) + f'(a)\epsilon + 0 + 0 + \dots \\ &= f(a) + f'(a)\epsilon \end{aligned} \tag{6.11}$$

It becomes clear that the first derivative of a function can be evaluated by simply evaluating the function at the dual number. The benefit of dual numbers are also made obvious here, unlike finite difference methods, the higher order derivatives are not being ignored but are explicitly eliminated.

For example the derivative of a one-dimensional Gaussian can be computed as:

$$\begin{aligned} f(x+\epsilon) &= e^{-a(x+\epsilon)^2} \\ &= e^{-ax^2} + (-2axe^{-ax^2})\epsilon \end{aligned} \tag{6.12}$$

Standard implementations for automatic differentation have two separate modes: the so-called forward mode and reserve mode. Both begin by representing the function as a computational graph where each node (v) represents an intermediate step in the calculation (See Fig. 6.9). To demonstrate consider evaluating $\frac{\partial y}{\partial x_1}$ at $x_1 = 4,\ x_2 = 3$, for the function $y = x_1 x_2 + \cos x_1$.

The simpler albeit computational inefficient forward mode computes each node linearly

from the input to the end result, storing intermediate results in the primal trace.

$$\begin{aligned} v_{-1} &= x_1 = 4 \\ v_0 &= x_2 = 3 \\ v_1 &= v_{-1} \cdot v_0 = 12 \\ v_2 &= \cos v_{-1} = \cos(4) \\ v_3 &= v_1 + v_2 = 12 + \cos(4) \end{aligned} \tag{6.13}$$

To evaluate $\dot{v_3}$, one traverses the dual trace where the derivative of each node is evaluated with respect to x_1.

$$\begin{aligned} \dot{v_{-1}} &= \frac{\partial x_1}{\partial x_1} = 1 \\ \dot{v_0} &= \frac{\partial x_2}{\partial x_1} = 0 \\ \dot{v_1} &= \dot{v_{-1}} v_0 + v_{-1} \dot{v_0} = 3 \\ \dot{v_2} &= -\sin(v_{-1}) \dot{v_{-1}} = -\sin(4) \\ \dot{v_3} &= \dot{v_1} + \dot{v_2} = 3 - \sin(4) \end{aligned} \tag{6.14}$$

Resulting in the desired derivative: $\frac{\partial y}{\partial x_1} = 3 - \sin(4)$.

Reverse mode implementations have the same primal trace as the forward mode, but the dual trace is more complex:

$$\frac{\partial y_j}{\partial v_i} = \sum_{p \in \text{parents(i)}} \frac{\partial y_j}{\partial v_p} \frac{\partial v_p}{\partial v_i} \tag{6.15}$$

Denoting $\bar{v_i} = \frac{\partial y_j}{\partial v_i} = \sum_p \bar{v_p} \frac{\partial v_p}{\partial v_i}$:

$$
\begin{aligned}
\bar{v_3} &= \bar{y} = 1 \\
\bar{v_2} &= \bar{v_3}\frac{\partial v_3}{\partial v_2} = 1 \cdot 1 = 1 \\
\bar{v_1} &= \bar{v_3}\frac{\partial v_3}{\partial v_1} = 1 \cdot 1 = 1 \\
\bar{v_0} &= \bar{v_1}\frac{\partial v_1}{\partial v_0} = 1 \cdot v_{-1} = 4 \\
\bar{v_{-1}} &= \bar{v_2}\frac{\partial v_2}{\partial v_{-1}} + \bar{v_1}\frac{\partial v_1}{\partial v_{-1}} = 1 \cdot (-\sin v_{-1}) + 1 \cdot v_0 = 3 - \sin(4)
\end{aligned}
\tag{6.16}
$$

Resulting in the desired derivative: $\frac{\partial y}{\partial x_1} = 3 - \sin(4)$.

Future research should look to replace the finite difference evaluations with an automatic differentiation approach. This may be able to reduce the total error by a few wavenumbers, resulting in more stable results.

6.5 Conclusions

The present test calculations of the lowest 50 eigenenergies of formaldehyde demonstrate that a Gaussian basis arranged according to a quasi-regular grid has superior qualities resulting in about a factor of 3-4 reduction in the total number of Gaussian's needed to maintain the same accuracy as the previously used quasi-random Gaussian basis.[72] Moreover, the regular local arrangement of the Gaussian centers allows one to implement a straightforward procedure for choosing the Gaussian width matrices, which appears to be a non-trivial issue otherwise.

With all the appealing properties and advantages of the present methodology which involves the easy-to-construct efficient and compact Gaussian basis and the following collocation approach to set-up a generalized eigenvalue problem, the only remaining serious drawback of the overall methodology seems to be the consequence of using a non-orthogonal basis and hence the need to deal with the numerical solution of a large generalized eigenvalue

problem. Here, two issues need to be addressed: (1) how to solve for the lowest eigenvalues (and eigenvectors) using iterative methods, and (2) parallelization of whatever generalized eigenvalue solver is used. Currently, neither of the two issues seem to have a satisfactory solution.

Chapter 6

Water Clusters and the Local Monomer Approximation

This final chapter looks at the challenging problem of computing rovibrational spectra of water clusters, in particular the water hexamer. Due to the high dimensionality of the cluster in Cartesian space ($d = 54$), we instead perform quantum dynamics calculations on a subspace of one water molecule at a time, by invoking the local monomer approximation. This process is repeated for each monomer in the system, resulting in a spectra for the entire cluster. By constructing a distributed Gaussian basis set with Gaussian centers placed according to the ground state wavefunction we perform Lmon3 calculations of the cage1 hexamer using the MBPOL potential energy surface. The challenges faced during the work ultimately motivated the development of the Quasi-Regular grids discussed in Chapters 3 and 5.

7.1 The Problems with Water

Water has always been an important molecular system both in experiments and in computation. The water hexamer in particular continues to be an active area of research due to the fact that the six-molecule system is the smallest cluster that is able to form energetically stable three-dimensional structures.[85–89] Water clusters in particular are notoriously difficult systems due to the isomer composition being dictated by the specific potential energy surface used, and the highly quantum nature of the water cluster. There are a number of different isomers know to account for large fractions of the equilibrium distribution, such as the "prism" and "cage" structures shown in Fig. 7.1, and significant effort has gone to the characterization of these systems.[90;91]

However, if we consider the hexamer in Cartesian space, each monomer has nine degrees of freedom resulting in a dimensionality of $d = 54$. To deal with the large system size we

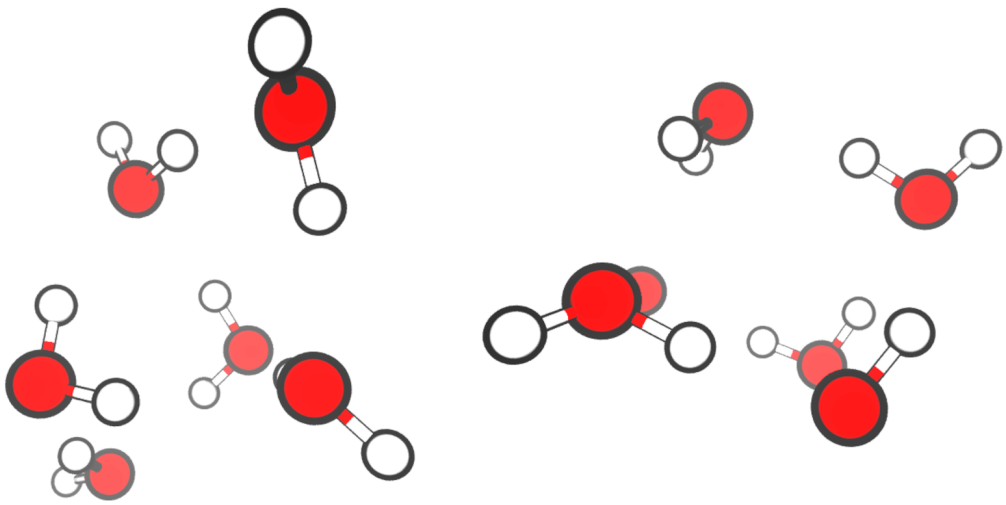

Figure 7.1: Select isomers of the water hexamer which are prevalent in the ground state. The specific structures were generated using the MB-pol potential energy surface; the cage1 configuration (left) and the prism (right).

apply an approximate separable method known as the local monomer model,[92;93] and in particular the work of Joel Bowman (see ref. 94).

Similar to the previously discussed projects our interest was focused on the placement of the Gaussian basis functions to represent the system. At the time we believed a combination of using a low-descrepancy sequence to construct our Gaussian Basis, combined with quasi-Monte Carlo integration techniques we would be able to compute larger LMM approximations for the monomers composing the cluster. Ideally one would perform either Lmon-6 or Lmon-9 subspace calculations as these subspaces contain the most meaning physically.

7.2 A Distributed Gaussian Basis Constructed with the Sobol Sequence

Consider a distributed Gaussian basis with Gaussian centers set using a quasi-random sequence distributed according to the harmonic ground state wavefunction.

Consider a d-dimensional system with the Hamiltonian (in atomic units, i.e., $\hbar = 1$) given by

$$\widehat{H} = -\frac{1}{2}\nabla^{\mathrm{T}}\mathbf{M}^{-1}\nabla + V(\mathbf{q}) \tag{7.1}$$

Here $\mathbf{q} \in \mathbb{R}^d$ defines a column vector, and ∇, the gradient.

Let $\tilde{\mathbf{K}} = \mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}$ be the mass-scaled Hessian matrix defined at the minimum (located at $\mathbf{q} = \mathbf{q}_0$) of the potential energy $V(\mathbf{q})$.

$$\tilde{\mathbf{K}} = \mathbf{U}\boldsymbol{\Omega}^2\mathbf{U}^{\mathrm{T}}; \quad \boldsymbol{\Omega} = \mathrm{diag}\{\omega_k\}; \quad \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I} \tag{7.2}$$

The mass-scaled normal mode coordinates $\mathbf{r} \in \mathbb{R}^d$ are

$$\mathbf{r} = \mathbf{U}^{\mathrm{T}}\mathbf{M}^{1/2}(\mathbf{q} - \mathbf{q}_0) \tag{7.3}$$

or defining $\mathbf{r} = (\mathbf{r}_1, ..., \mathbf{r}_d)^{\mathrm{T}}$

$$\mathbf{r}_k = \mathbf{U}_k^{\mathrm{T}} \mathbf{M}^{1/2} (\mathbf{q} - \mathbf{q}_0) \tag{7.4}$$

Also,

$$\mathbf{q} = \mathbf{q}_0 + \mathbf{M}^{-1/2} \mathbf{U} \mathbf{r} \tag{7.5}$$

Using the normal mode coordinates, the Hamiltonian reads

$$\widehat{H} = -\frac{1}{2} \nabla^{\mathrm{T}} \nabla + V(\mathbf{q}_0 + \mathbf{M}^{-1/2} \mathbf{U} \mathbf{r}) \tag{7.6}$$

Consider a grid $\mathbf{r}^j \in \mathbb{R}^d$ $(i = 1, ..., J)$, sampled using a quasi-random sequence defined by the harmonic ground state wavefunction:

$$\Psi_0(\mathbf{r}) = \sqrt{\frac{\det \boldsymbol{\Omega}}{(2\pi)^d}} \exp\left[-\frac{1}{2} \mathbf{r}^{\mathrm{T}} \boldsymbol{\Omega} \mathbf{r} \right] \tag{7.7}$$

Each grid point is associated with a basis function in the form of a multivariate Gaussian:

$$\Phi_i(\mathbf{r}) := \exp\left[-\frac{1}{2} (\mathbf{r} - \mathbf{r}^i)^{\mathrm{T}} \mathbf{G}^i (\mathbf{r} - \mathbf{r}^i) \right], \tag{7.8}$$

where $\mathbf{G}^i$ is a positive definite width matrix.

Recall that the product of two Gaussian's is also a Gaussian:

$$\Phi_i(\mathbf{r}) \Phi_j(\mathbf{r}) = A \exp\left[-\frac{1}{2} (\mathbf{r} - \bar{\mathbf{r}}^{ij})^{\mathrm{T}} \left(\mathbf{G}^i + \mathbf{G}^j \right) (\mathbf{r} - \bar{\mathbf{r}}^{ij}) \right] \tag{7.9}$$

where

$$\begin{aligned} &\mathbf{u}^i := \mathbf{G}^i \mathbf{r}^i; \quad \bar{\mathbf{r}}^{ij} := \left(\mathbf{G}^i + \mathbf{G}^j \right)^{-1} (\mathbf{u}^i + \mathbf{u}^j); \\ &A := \exp\left\{ -\frac{1}{2} \left[(\mathbf{r}^i)^{\mathrm{T}} \mathbf{u}^i + (\mathbf{r}^j)^{\mathrm{T}} \mathbf{u}^j - (\bar{\mathbf{r}}^{ij})^{\mathrm{T}} (\mathbf{u}^i + \mathbf{u}^j) \right] \right\} \end{aligned}$$

The overlap matrix is then given by

$$\begin{aligned} \mathbf{S}_{ij} &:= \langle \Phi_i | \Phi_j \rangle \\ &= A \sqrt{\frac{(2\pi)^d}{\det\left(\mathbf{G}^i + \mathbf{G}^j\right)}} \end{aligned} \tag{7.10}$$

To further simplify consider the following:

$$\Phi_i(\mathbf{r})\Phi_j(\mathbf{r}) = \mathbf{S}_{ij} P_{ij}(\mathbf{r}) \tag{7.11}$$

where P_{ij} is a normalized multivariate Gaussian distribution

$$P_{ij}(\mathbf{r}) := \sqrt{\frac{\det\left(\mathbf{G}^i + \mathbf{G}^j\right)}{(2\pi)^d}} \; \exp\left[-\frac{1}{2}(\mathbf{r} - \bar{\mathbf{r}}^{ij})^{\mathrm{T}} \left(\mathbf{G}^i + \mathbf{G}^j\right) (\mathbf{r} - \bar{\mathbf{r}}^{ij})\right]; \tag{7.12}$$

The kinetic energy matrix is computed as

$$\begin{aligned} \mathbf{T}_{ij} &:= -\frac{1}{2} \langle \Phi_i | \nabla^{\mathrm{T}} \nabla | \Phi_j \rangle \\ &= \frac{1}{2} \mathbf{S}_{ij} \int_{\mathbb{R}^d} d\mathbf{r} \; (\mathbf{r} - \mathbf{r}^i)^{\mathrm{T}} \mathbf{G}^i \mathbf{G}^j (\mathbf{r} - \mathbf{r}^j) P_{ij}(\mathbf{r}) \end{aligned}$$

This expression can be evaluated analytically for a general case, however, we will later consider the special case of diagonal G^i matrices.

Finally the potential energy matrix is given by

$$\begin{aligned} \mathbf{V}_{ij} &:= \langle \Phi_i | V | \Phi_j \rangle \\ &= \mathbf{S}_{ij} \int_{\mathbb{R}^d} d\mathbf{r} \; P_{ij}(\mathbf{r}) V(\mathbf{q}_0 + \mathbf{M}^{-1/2}\mathbf{U}\mathbf{r}) \end{aligned} \tag{7.13}$$

And the dipole moments:

$$\begin{aligned}\mu_{nm} &:= \langle \Phi_i | \mu | \Phi_j \rangle \\ &= \mathbf{S}_{ij} \int_{\mathbb{R}^d} d\mathbf{r}\; P_{ij}(\mathbf{r}) \mu(\mathbf{q}_0 + \mathbf{M}^{-1/2}\mathbf{U}\mathbf{r})\end{aligned} \tag{7.14}$$

These expressions can be greatly simplified, consider

$$\mathbf{G}^i = \alpha_i \mathbf{\Omega} \tag{7.15}$$

A natural choice for α_i is

$$\alpha_i = \alpha_0 J^{2/d} \exp\left[-\frac{1}{d}\mathbf{r}^{\mathrm{T}}\mathbf{\Omega}\mathbf{r}\right] \tag{7.16}$$

where $\alpha_0 \sim 1$ is a constant to be specified later.

Redefining terms:

$$\bar{\mathbf{r}}^{ij} := \frac{\alpha_i \mathbf{r}^i + \alpha_j \mathbf{r}^j}{\alpha_i + \alpha_j} \tag{7.17}$$

$$A := \exp\left[-\frac{\alpha_i \alpha_j}{2(\alpha_i + \alpha_j)} \sum_k \omega_k (\mathbf{r}_k^i - \mathbf{r}_k^j)^2\right] \tag{7.18}$$

$$S_k := \left[\frac{2\pi}{(\alpha_i + \alpha_j)\omega_k}\right]^{1/2} \exp\left[-\frac{\alpha_i \alpha_j}{2(\alpha_i + \alpha_j)} \omega_k (\mathbf{r}_k^i - \mathbf{r}_k^j)^2\right] \tag{7.19}$$

Note that the i, j-independent factor in Eq. (7.19) can be dropped as it does not affect the final generalized eigenvalue problem (*cf.* Eq. (7.30)), i.e., we can safely use

$$S_k = \frac{1}{\sqrt{\alpha_i + \alpha_j}} \exp\left[-\frac{\alpha_i \alpha_j}{2(\alpha_i + \alpha_j)} \omega_k (\mathbf{r}_k^i - \mathbf{r}_k^j)^2\right] \tag{7.20}$$

for the product of two Gaussian's we have

$$\Phi_i(\mathbf{r}) \Phi_j(\mathbf{r}) = \mathbf{S}_{ij}\, P_{ij}(\mathbf{r}) \tag{7.21}$$

where the overlap matrix is

$$\mathbf{S}_{ij} := \int_{\mathbb{R}^d} d\mathbf{r}\ \Phi_i(\mathbf{r})\Phi_j(\mathbf{r}) = \prod_{k=1}^{d} S_k \tag{7.22}$$

and

$$\begin{aligned} P_{ij}(\mathbf{r}) &= \prod_{k=1}^{d} P_{ij}^{(k)}(\mathbf{r}_k) \\ P_{ij}^{(k)}(\mathbf{r}_k) &:= \left[\frac{(\alpha_i+\alpha_j)\omega_k}{2\pi}\right]^{1/2} \exp\left(-\frac{(\alpha_i+\alpha_j)\omega_k}{2}\,(\mathbf{r}_k-\bar{\mathbf{r}}_k^{ij})^2\right) \end{aligned} \tag{7.23}$$

Both the Potential and Dipole matrices will need to be computed numerically.

$$\begin{aligned} \mathbf{V}_{ij} &= \mathbf{S}_{ij}\int_{\mathbb{R}^d} d\mathbf{r}\ P_{ij}(\mathbf{r})V(\mathbf{q}_0+\mathbf{M}^{-1/2}\mathbf{U}\mathbf{r}) \\ \mu_{ij} &= \mathbf{S}_{ij}\int_{\mathbb{R}^d} d\mathbf{r}\ P_{ij}(\mathbf{r})\mu(\mathbf{q}_0+\mathbf{M}^{-1/2}\mathbf{U}\mathbf{r}) \end{aligned} \tag{7.24}$$

The integrals in Eq. (4.18) are computed as follows. First, we generate a quasi-random sequence $\mathbf{z}^{(l)}, l = (1, ..., L)$ sampled from the standard normal distribution. Then use the following:

$$\int_{\mathbb{R}^d} d\mathbf{r}\ P_{ij}(\mathbf{r})V(\mathbf{q}_0+\mathbf{M}^{-1/2}\mathbf{U}\mathbf{r}) \approx \frac{1}{L}\sum_{l=1}^{L} V(\mathbf{q}^{(l)}) \tag{7.25}$$

with

$$\mathbf{r}_k^{(l)} = \bar{\mathbf{r}}_k^{ij} + \left[(\alpha_i+\alpha_j)\omega_k\right]^{-1/2}\mathbf{z}_k^{(l)} \tag{7.26}$$

and

$$\mathbf{q}^{(l)} = \mathbf{q}_0 + \mathbf{M}^{-1/2}\mathbf{U}\mathbf{r}^{(l)}, \tag{7.27}$$

The kinetic energy matrix elements are given by the following expression

$$\mathbf{T}_{ij} = \mathbf{S}_{ij}\sum_{k=1}^{d}\frac{\alpha_i\alpha_j\omega_k}{2(\alpha_i+\alpha_j)}\left[1-\frac{\alpha_i\alpha_j(\mathbf{r}_k^i-\mathbf{r}_k^j)^2\omega_k}{\alpha_i+\alpha_j}\right] \tag{7.28}$$

Finally, the vibrational eigenenergies E and eigenfunctions

$$\Psi(\mathbf{r}) = \sum_j c_j \Phi_j(\mathbf{r}) \tag{7.29}$$

can be obtained from solving the generalized eigenvalue problem:

$$\sum_j (\mathbf{V}_{ij} + \mathbf{T}_{ij} - E\mathbf{S}_{ij}) c_j = 0 \tag{7.30}$$

The intensity of an absorption line for rovibrational transitions is given in terms of the dipole matrix. Assuming our system is initially in the ground state (and zero angular momentum) the intensity of the j-th transition is computed as:

$$I_{j\leftarrow 0} := A\ \nu_{j\leftarrow 0}\ \langle \Psi_0(\mathbf{r}) | \mu_\sigma | \Psi_i(\mathbf{r}) \rangle^2 \tag{7.31}$$

The Hamiltonian Eigenvectors are normalized

$$c^{(j),T}\ \mathbf{S}\ c^{(j)} = 1$$

where

$$\Psi_0(\mathbf{r}) = \sum_{n=1}^{Ng} c_n^{(0)} \phi_n \quad \Psi_i(\mathbf{r}) = \sum_{m=1}^{Ng} c_m^{(i)} \phi_m \tag{7.32}$$

The dipole matrix and associated intensities are then computed as

$$\begin{aligned} \langle \Psi_0 | \mu_\sigma | \Psi_i \rangle^2 &= \left(\sum_n \sum_m c_n^{(0)} c_m^{(i)} \mu_{nm} \right)^2 \\ I_{j\leftarrow 0} &= A\ \nu_{j\leftarrow 0} \sum_\sigma \left\{ \left(\sum_n \sum_m c_n^{(0)} c_m^{(i)} \mu_{nm}^\sigma \right)^2 \right\} \end{aligned} \tag{7.33}$$

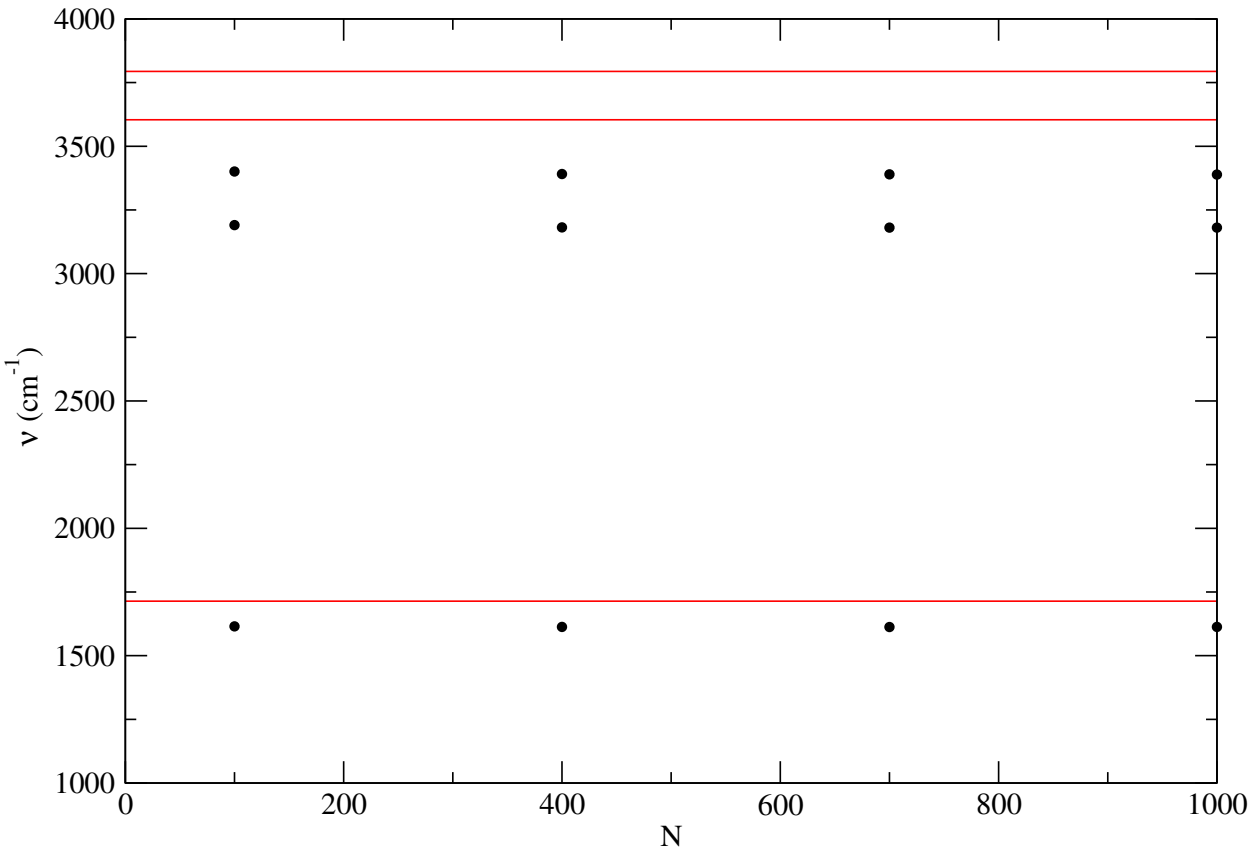

Figure 7.2: Convergence of the first few frequencies for monomer1 of the cage conformation in the Lmon3-MBPOL. In red are the expected harmonic approximations.

7.3 The Lmon3 Spectra of water using MB-pol

For this work we implement the MB-pol water potential energy surface. This surface, developed by the Paesani group, is both one of the most accurate, and most expensive models for water available.[95;96] Again, we believed the use of quasi-Monte Carlo methods would be enough to account for the numerical bottleneck associated with using a high-accuracy potential surface such as MB-pol. As shown in Fig. 7.2 we can easily converge the lower frequencies (to 1 wavenumber accuracy) with a very small number of basis functions when using the Lmon3 subspace.

The location of the frequencies being lower than the harmonic approximation is consistent with the anharmonicities in the system. We can also produce a qualitatively reasonable spectra as shown in Fig. 7.3.

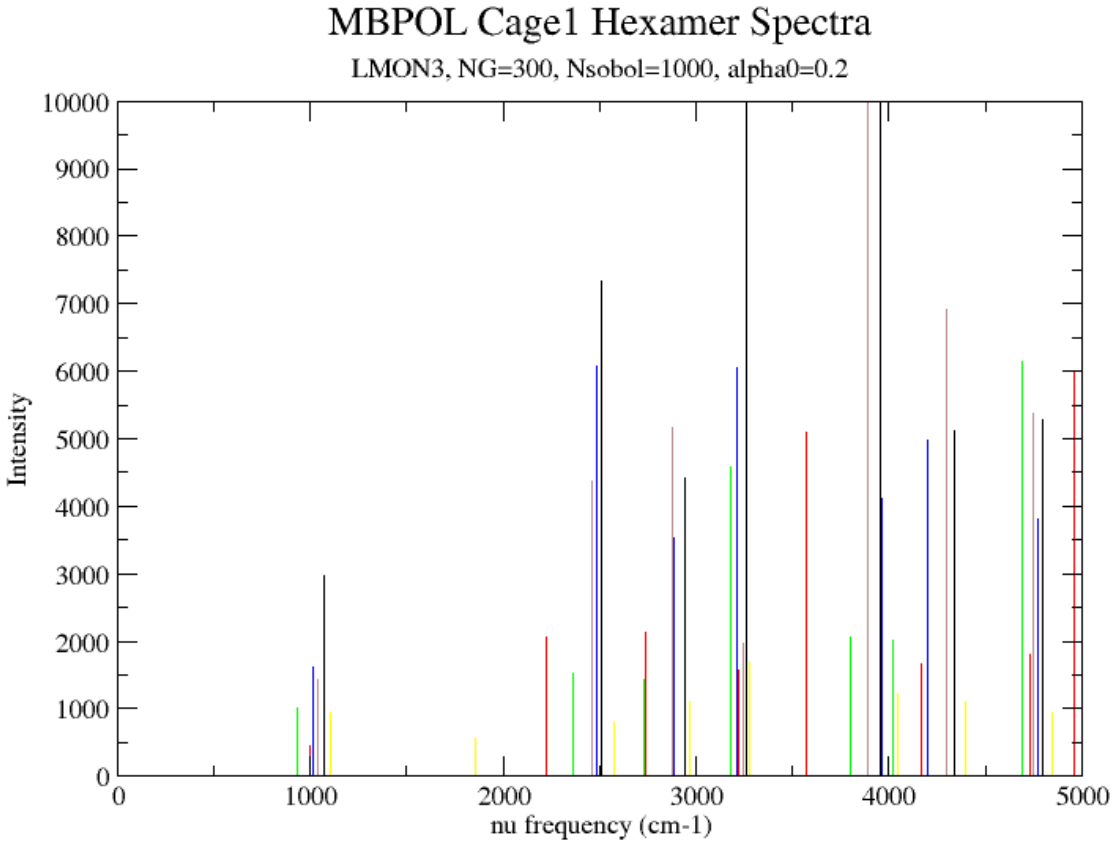

Figure 7.3: Vibrational spectra for the Cage1 water hexamer using the MBPOL potential and the Lmon3 approximation. A small number of basis functions, N=300, produces a set of intensities in the qualitatively correct region of the spectrum.

The spectra does appear to be qualitatively correct, i.e. we see peaks in the general region of the known bend and OH-stretching modes. And again these are lower than the harmoinc approximation due to the anharmoinc terms.

However, studies attempting to implement the Lmon6 subspace made it clear that convergence with respect to the basis size would be unreasonable due to the costly potential energy evaluations. Quasi-Monte Carlo approaches did not drastically reduce the integration problem to allow for a meaningful contribution beyond the Lmon3 results. This project demonstrated that while the integration of the potential energy matrix element is a key component of the spectral calculations, the distribution of the basis functions was a potentially more fruitful endeavor.

Bibliography

[1] S. Tyagi and S. Mittal, in *Proceedings of ICRIC 2019* (Springer, 2020) pp. 209–221.

[2] A. P. Lund, M. J. Bremner, and T. C. Ralph, npj Quantum Information **3**, 1 (2017).

[3] G. Avila and T. Carrington Jr, The Journal of chemical physics **131**, 174103 (2009).

[4] I. Sumner and S. S. Iyengar, The Journal of Physical Chemistry A **111**, 10313 (2007).

[5] M. Møller, International Journal of Neural Systems **4**, 15 (1993).

[6] G. H. Nguyen, A. Bouzerdoum, and S. L. Phung, in *2008 19th international conference on pattern recognition* (IEEE, 2008) pp. 1–4.

[7] G. E. Batista and M. C. Monard, Applied artificial intelligence **17**, 519 (2003).

[8] H. He, Y. Bai, E. A. Garcia, and S. Li, in *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)* (IEEE, 2008) pp. 1322–1328.

[9] D. G. Truhlar, R. Steckler, and M. S. Gordon, Chemical Reviews **87**, 217 (1987).

[10] J. Ischtwan and M. A. Collins, The Journal of Chemical Physics **100**, 8080 (1994).

[11] K. C. Thompson, M. J. T. Jordan, and M. A. Collins, The Journal of Chemical Physics **108**, 8302 (1998).

[12] S. Manzhos, X. Wang, R. Dawes, and T. Carrington, The Journal of Chemical Physics **110**, 5295 (2006).

[13] J. Behler and M. Parrinello, Physical Review Letters **98**, 146401 (2007).

[14] Y. Wang, B. C. Shepler, B. J. Braams, and J. M. Bowman, The Journal of chemical physics **131**, 054511 (2009).

[15] C. Qu, Q. Yu, and J. M. Bowman, Annual review of physical chemistry **69**, 151 (2018).

[16] C. Qu, Q. Yu, B. L. Van Hoozen, J. M. Bowman, and R. A. Vargas-Hernández, Journal of Chemical Theory and Computation **14**, 3381 (2018).

[17] N. Metropolis and S. Ulam, Journal of the American statistical association **44**, 335 (1949).

[18] W. K. Hastings, (1970).

[19] L. L. Carter and E. D. Cashwell, *Particle-transport simulation with the Monte Carlo method*, Tech. Rep. (Los Alamos Scientific Lab., N. Mex.(USA), 1975).

[20] D. P. Kroese, T. Brereton, T. Taimre, and Z. I. Botev, Wiley Interdisciplinary Reviews: Computational Statistics **6**, 386 (2014).

[21] F. Nobile, R. Tempone, and C. G. Webster, SIAM Journal on Numerical Analysis **46**, 2309 (2008).

[22] G. Marshall, Computer physics communications **56**, 51 (1989).

[23] F. James, Reports on progress in Physics **43**, 1145 (1980).

[24] W. J. Morokoff and R. E. Caflisch, Journal of computational physics **122**, 218 (1995).

[25] X. Wang and K.-T. Fang, Journal of Complexity **19**, 101 (2003).

[26] R. Schürer, Mathematics and computers in simulation **62**, 509 (2003).

[27] B. D. Flury, SIAM Review **32**, 474 (1990).

[28] S. Garashchuk and J. C. Light, The Journal of Chemical Physics **114**, 3929 (2001).

[29] L. Martino and J. Míguez, Signal Processing **90**, 2981 (2010).

[30] H. Niederreiter, Bulletin of the American mathematical society **84**, 957 (1978).

[31] P. L'Ecuyer, Communications of the ACM **33**, 85 (1990).

[32] B. Tuffin, (1996).

[33] J. Matoušek, Journal of Complexity **14**, 527 (1998).

[34] P. Bratley and B. L. Fox, ACM Transactions on Mathematical Software (TOMS) **14**, 88 (1988).

[35] I. M. Sobol', Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki **7**, 784 (1967).

[36] P. Bratley and B. L. Fox, ACM Transactions on Mathematical Software (TOMS) **14**, 88 (1988).

[37] B. Tuffin, Monte Carlo Methods and Applications **2**, 295 (1996).

[38] S. W. Flynn and V. A. Mandelshtam, Chem. Phys. Lett. **151**, 241105 (2019).

[39] J. Beasley and S. Springer, Journal of the Royal Statistical Society. Series C (Applied Statistics) **26**, 118 (1977).

[40] T. Pillards and R. Cools, in *International Conference on Computational Science and Its Applications* (Springer, 2006) pp. 780–788.

[41] S. E. Brown, I. Georgescu, and V. A. Mandelshtam, The Journal of chemical physics **138**, 044317 (2013).

[42] S. E. Brown and V. A. Mandelshtam, Chemical Physics **481**, 69 (2016).

[43] I. Georgescu, S. Jitomirskaya, and V. A. Mandelshtam, The Journal of chemical physics **139**, 204104 (2013).

[44] M. Feit, J. Fleck Jr, and A. Steiger, Journal of Computational Physics **47**, 412 (1982).

[45] M. Feit and J. Fleck Jr, The Journal of chemical physics **80**, 2578 (1984).

[46] J. Light, I. Hamilton, and J. Lill, The Journal of chemical physics **82**, 1400 (1985).

[47] J. Light and Z. Bačić, The Journal of chemical physics **87**, 4008 (1987).

[48] Z. Bačić, D. Watt, and J. Light, The Journal of chemical physics **89**, 947 (1988).

[49] R. M. Whitnell and J. Light, The Journal of chemical physics **90**, 1774 (1989).

[50] J. C. Light and T. Carrington Jr, Advances in Chemical Physics **114**, 263 (2000).

[51] J. Echave and D. C. Clary, Chemical physics letters **190**, 225 (1992).

[52] A. Kumar, N. DeGregorio, and S. S. Iyengar, Journal of Chemical Theory and Computation (2021).

[53] M. Kardar, *Statistical physics of particles* (Cambridge University Press, 2007).

[54] J. Jakowski, I. Sumner, and S. S. Iyengar, Journal of chemical theory and computation **2**, 1203 (2006).

[55] D. Hocker, X. Li, and S. S. Iyengar, Journal of chemical theory and computation **7**, 256 (2011).

[56] N. DeGregorio and S. S. Iyengar, Journal of chemical theory and computation **14**, 30 (2018).

[57] G. Avila and T. Carrington Jr, The Journal of chemical physics **143**, 214108 (2015).

[58] T. Carrington Jr, Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy , 119158 (2021).

[59] E. Mátyus, G. Czakó, and A. G. Császár, The Journal of chemical physics **130**, 134112 (2009).

[60] M. Dutra, S. Wickramasinghe, and S. Garashchuk, The Journal of Physical Chemistry A **124**, 9314 (2020).

[61] M. J. Davis and E. J. Heller, The Journal of Chemical Physics **71**, 3383 (1979).

[62] Z. Bacić and J. Light, The Journal of Chemical Physics **85**, 4594 (1986).

[63] I. Hamilton and J. Light, The Journal of Chemical Physics **84**, 306 (1986).

[64] M. Mladenović and Z. Bačić, The Journal of Chemical Physics **93**, 3039 (1990).

[65] B. Poirier and J. Light, The Journal of Chemical Physics **113**, 211 (2000).

[66] A. Shimshovitz and D. J. Tannor, Physical Review Letters **109**, 070402 (2012).

[67] B. Poirier, Journal of Theoretical and Computational Chemistry **02**, 65 (2003).

[68] G. Avila and T. Carrington Jr, The Journal of Chemical Physics **131**, 174103 (2009).

[69] T. Halverson and B. Poirier, The Journal of Chemical Physics **137**, 224101 (2012).

[70] S. Manzhos, X. Wang, and T. Carrington, Chemical Physics **509**, 139 (2018).

[71] A. Pandey and B. Poirier, The Journal of Chemical Physics **151**, 014114 (2019).

[72] S. Manzhos and T. Carrington, The Journal of Chemical Physics **145**, 224110 (2016).

[73] B. Poirier, Foundations of Physics **30**, 1191 (2000).

[74] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).

[75] S. W. Flynn and V. A. Mandelshtam, Journal of Chemical Theory and Computation (2021).

[76] W. Yang and A. C. Peet, Chem. Phys. Lett. **153**, 98 (1988).

[77] W. Yang and A. C. Peet, J. Chem. Phys. **92**, 522 (1990).

[78] A. C. Peet, J. Chem. Phys. **90**, 4363 (1989).

[79] N. Meinander and J. Laane, J. Mol. Struct **569**, 1 (2001).

[80] V. Glushkov and S. Wilson, Int. J. Quantum Chem **89**, 237 (2002).

[81] M. Dutra, S. Wickramasinghe, and S. Garashchuk, J. Chem. Theory Comput **16**, 18 (2019).

[82] B. S. C. N. C. HANDY and J. Demaison, Mol. Phys. **90**, 729 (1997).

[83] L. Hascoet and V. Pascual, ACM Trans. Math. Softw. **39**, 1 (2013).

[84] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, Journal of machine learning research **18** (2018).

[85] K. Liu, M. Brown, C. Carter, R. Saykally, J. Gregory, and D. Clary, Nature **381**, 501 (1996).

[86] G. Hincapie, N. Acelas, M. Castano, J. David, and A. Restrepo, The Journal of Physical Chemistry A **114**, 7809 (2010).

[87] J. Kim and K. S. Kim, The Journal of chemical physics **109**, 5886 (1998).

[88] C. Pérez, M. T. Muckle, D. P. Zaleski, N. A. Seifert, B. Temelso, G. C. Shields, Z. Kisiel, and B. H. Pate, Science **336**, 897 (2012).

[89] R. J. Saykally and D. J. Wales, Science **336**, 814 (2012).

[90] J. D. Mallory, S. E. Brown, and V. A. Mandelshtam, The Journal of Physical Chemistry A **119**, 6504 (2015).

[91] Y. Wang, V. Babin, J. M. Bowman, and F. Paesani, Journal of the American Chemical Society **134**, 11116 (2012).

[92] L. Halonen and M. Child, Molecular Physics **46**, 239 (1982).

[93] R. L. Swofford, M. Long, and A. Albrecht, The Journal of Chemical Physics **65**, 179 (1976).

[94] Y. Wang and J. M. Bowman, The Journal of chemical physics **134**, 154510 (2011).

[95] V. Babin, C. Leforestier, and F. Paesani, Journal of chemical theory and computation **9**, 5395 (2013).

[96] V. Babin, G. R. Medders, and F. Paesani, Journal of chemical theory and computation **10**, 1599 (2014)

Appendix

Pseudo-Code for QRG Generation

A FORTRAN pseudo-code which demonstrates the numerical implementation of the QRG minimization using internal bond coordinates.

```
program main
 use  QRG_module
 !Generate Initial Grid
 do i=1,100
  call Metropolis_MC(100,r(:,1),Pr,Vr,step)
 enddo
 do i=2,NG
  r(:,i)=r(:,i-1)
  call Metropolis_MC(1000,r(:,i),Pr,Vr,step)
 enddo
 do i=1,NG
  sigma(i)=P(r(:,i),V_(i))**(-1./d1)
  call Internal_to_bond(r(:,i),bond(:,i))
 enddo
 call find_range(bond)
 E=0
 do i=2,NG
  do j=1,i-1
   E=E+Pair_rpl(bond(:,i),bond(:,j),sigma(i),sigma(j))
  enddo
 enddo
```

```
!Optimize QRG (greedy Minimization)
do n=1,NG*N_MC_QRG
 k=random_integer(1,NG)
 call random_number(delr)
 r_trial=r(:,k)+step*(2*delr-1)
 call Internal_to_bond(r_trial,bond_trial)
 P_trial=P(r_trial,V_trial)
 if(P_trial > 1d-100) then
  sigma_trial=P_trial**(-1./d1)
  delE=0d0
  do j=1,NG
   if(j.ne.k) then
    delE=delE+Pair_rpl(bond(:,j),bond_trial,sigma(j),sigma_trial)
         - Pair_rpl(bond(:,j),bond(:,k),sigma(j),sigma(k))
   endif
  enddo
  if(delE <= 0d0) then
   r(:,k)=r_trial(:)
   bond(:,k)=bond_trial(:)
   sigma(k)=sigma_trial
   V_(k)=V_trial
   accept=accept+1
   E=E+delE
  endif
 endif
 if(mod(n,NG*10)==0)then
  if(accept/NG*10. < 0.4) step=step*0.9
```

```
   if(accept/NG*10. > 0.6) step=step*1.1
   accept=0
   E=0
   do i=2,NG
    do j=1,i-1
     E=E+Pair_rpl(bond(:,i),bond(:,j),sigma(i),sigma(j))
    enddo
   enddo
  endif
 enddo
end program main
```

Kinetic Energy Analytic Expressions

We have the following transformations (I have dropped the mass-scaling terms). The system is 4 atoms, we therefore have a 12d coordinate vector x.

$$\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}) \tag{7.34}$$

Each atoms coordinates are given by:

$$\mathbf{C} = (x_1, x_2, x_3) \qquad \mathbf{\Omega} = (x_4, x_5, x_6)$$
$$\mathbf{H}_1 = (x_7, x_8, x_9) \qquad \mathbf{H}_2 = (x_{10}, x_{11}, x_{12})$$

We define a 6d internal coordinate vector $\mathbf{r}$.

$$\mathbf{r}_1 = \mathbf{H}_1 - \mathbf{C} = (x_7 - x_1,\ x_8 - x_2,\ x_9 - x_3) \qquad r_1 = ||\mathbf{r}_1||$$
$$\mathbf{r}_2 = \mathbf{H}_2 - \mathbf{C} = (x_{10} - x_1,\ x_{11} - x_2,\ x_{12} - x_3) \qquad r_2 = ||\mathbf{r}_2||$$
$$\mathbf{r}_3 = \mathbf{\Omega} - \mathbf{C} = (x_4 - x_1,\ x_5 - x_2,\ x_6 - x_3) \qquad r_3 = ||\mathbf{r}_3||$$

Here (r_1,r_2,r_3) specify the distance between the atoms. (|| || is the norm vector). The last 3 internal coordinates are for the angles:

$$\text{costheta1} = \frac{\mathbf{r}_1 \cdot \mathbf{r}_3}{r_1 * r_3} \qquad \text{costheta2} = \frac{\mathbf{r}_2 \cdot \mathbf{r}_3}{r2 * r3}$$
$$\mathbf{r}_1\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_3 \qquad ||r_1 r_3|| = \sqrt{\mathbf{r}_1\mathbf{r}_3 \cdot \mathbf{r}_1\mathbf{r}_3}$$
$$\mathbf{r}_2\mathbf{r}_3 = \mathbf{r}_2 \times \mathbf{r}_3 \qquad ||r_2 r_3|| = \sqrt{\mathbf{r}_2\mathbf{r}_3 \cdot \mathbf{r}_2\mathbf{r}_3}$$
$$\text{cosphi} = \frac{\mathbf{r}_1\mathbf{r}_3 \cdot \mathbf{r}_2\mathbf{r}_3}{||r_1 r_3|| * ||r_2 r_3||}$$

Finally we define the r vector:

$$\begin{aligned}\mathbf{r} &:= (r_1, r_2, r_3, \text{acos(costheta1)}, \text{acoscostheta2)}, \text{acos(cosphi)}) \\ &= (r_1, r_2, r_3, r_4, r_5, r_6)\end{aligned} \tag{7.35}$$

$\frac{dr_1}{d\mathbf{X}}$

$$\begin{aligned}
&\frac{dr_1}{dx_1} = \frac{x_1 - x_7}{r_1} && \frac{dr_1}{dx_2} = \frac{x_2 - x_8}{r_1} && \frac{dr_1}{dx_3} = \frac{x_3 - x_9}{r_1} \\
&\frac{dr_1}{dx_4} = 0 && \frac{dr_1}{dx_5} = 0 && \frac{dr_1}{dx_6} = 0 \\
&\frac{dr_1}{dx_7} = \frac{-x_1 + x_9}{r_1} && \frac{dr_1}{dx_8} = \frac{-x_2 + x_8}{r_1} && \frac{dr_1}{dx_9} = \frac{-x_3 + x_9}{r_1} \\
&\frac{dr_1}{dx_{10}} = 0 && \frac{dr_1}{dx_{11}} = 0 && \frac{dr_1}{dx_{12}} = 0
\end{aligned}$$

$\frac{dr_2}{d\mathbf{X}}$

$$\begin{aligned}
&\frac{dr_2}{dx_1} = \frac{x_1 - x_{10}}{r_2} && \frac{dr_2}{dx_2} = \frac{x_2 - x_{11}}{r_2} && \frac{dr_2}{dx_3} = \frac{x_3 - x_{12}}{r_2} \\
&\frac{dr_2}{dx_4} = 0 && \frac{dr_2}{dx_5} = 0 && \frac{dr_2}{dx_6} = 0 \\
&\frac{dr_2}{dx_7} = 0 && \frac{dr_2}{dx_8} = 0 && \frac{dr_2}{dx_9} = 0 \\
&\frac{dr_2}{dx_{10}} = \frac{-x_1 + x_{10}}{r_2} && \frac{dr_2}{dx_{11}} = \frac{-x_2 + x_{11}}{r_2} && \frac{dr_2}{dx_{12}} = \frac{-x_3 + x_{12}}{r_2}
\end{aligned}$$

$\frac{dr_3}{d\mathbf{X}}$

$$\frac{dr_3}{dx_1} = \frac{x_1 - x_4}{r_3} \qquad \frac{dr_3}{dx_2} = \frac{x_2 - x_5}{r_3} \qquad \frac{dr_3}{dx_3} = \frac{x_3 - x_6}{r_3}$$
$$\frac{dr_3}{dx_4} = \frac{-x_1 + x_4}{r_3} \qquad \frac{dr_3}{dx_5} = \frac{-x_2 + x_5}{r_3} \qquad \frac{dr_3}{dx_6} = \frac{-x_3 + x_6}{r_3}$$
$$\frac{dr_3}{dx_7} = 0 \qquad \frac{dr_3}{dx_8} = 0 \qquad \frac{dr_3}{dx_9} = 0$$
$$\frac{dr_3}{dx_{10}} = 0 \qquad \frac{dr_3}{dx_{11}} = 0 \qquad \frac{dr_3}{dx_{12}} = 0$$

$\frac{dr_4}{d\mathbf{X}}$

$$\frac{dr_4}{dx_1} = \frac{(x_1 - x_7) r_3^2 (\mathbf{r}_1 \cdot \mathbf{r}_3) + (x_4 + x_7 - 2x_1) r_1^2 r_3^2 + (x_1 - x_4) r_1^2 (\mathbf{r}_1 \cdot \mathbf{r}_3)}{r_1^3 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$

$$\frac{dr_4}{dx_2} = \frac{(x_2 - x_8) r_3^2 (\mathbf{r}_1 \cdot \mathbf{r}_3) + (x_5 + x_8 - 2x_2) r_1^2 r_3^2 + (x_2 - x_5) r_1^2 (\mathbf{r}_1 \cdot \mathbf{r}_3)}{r_1^3 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$

$$\frac{dr_4}{dx_3} = \frac{(x_3 - x_9) r_3^2 (\mathbf{r}_1 \cdot \mathbf{r}_3) + (x_6 + x_9 - 2x_3) r_1^2 r_3^2 + (x_3 - x_6) r_1^2 (\mathbf{r}_1 \cdot \mathbf{r}_3)}{r_1^3 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$

$$\begin{aligned}\frac{dr_4}{dx_4} =& \frac{(x_1 - x_7)\left(x_5^2 + x_6^2\right) + (x_4 - x_7)\left(x_2^2 + x_3^2\right)}{r_1 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}} \\ &+ \frac{(x_1 - x_4)\left[x_8 (x_2 - x_5) + x_9 (x_3 - x_6)\right] + (2x_7 - x_1 - x_4)(x_3 x_6 + x_2 x_5)}{r_1 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}\end{aligned}$$

$$\frac{dr_4}{dx_5} = \frac{(x_5 - x_8)\left(x_1^2 + x_3^2\right) + (x_2 - x_8)\left(x_4^2 + x_6^2\right)}{r_1 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$
$$+ \frac{(x_2 - x_5)\left[x_9 (x_3 - x_6) + x_7 (x_1 - x_4)\right] + (2x_8 - x_2 - x_5)(x_3 x_6 + x_1 x_4)}{r_1 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$

$$\frac{dr_4}{dx_6} = \frac{(x_6 - x_9)\left(x_1^2 + x_2^2\right) + (x_3 - x_9)\left(x_4^2 + x_5^2\right)}{r_1 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$
$$+ \frac{(x_3 - x_6)\left[x_7 (x_1 - x_4) + x_8 (x_2 - x_5)\right] + (2x_9 - x_3 - x_6)(x_2 x_5 + x_1 x_4)}{r_1 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$

$$\frac{dr_4}{dx_7} = \frac{(x_7 - x_4)\left(x_2^2 + x_3^2\right) + (x_1 - x_4)\left(x_8^2 + x_9^2\right)}{r_3 r_1^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$
$$+ \frac{(x_1 - x_7)\left[x_5 (x_2 - x_8) + x_6 (x_3 - x_9)\right] + (2x_4 - x_1 - x_7)(x_2 x_8 + x_3 x_9)}{r_3 r_1^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$

$$\frac{dr_4}{dx_8} = \frac{(x_2 - x_5)\left(x_7^2 + x_9^2\right) + (x_8 - x_5)\left(x_1^2 + x_3^2\right)}{r_3 r_1^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$
$$+ \frac{(x_2 - x_8)\left[x_4 (x_1 - x_7) + x_6 (x_3 - x_9)\right] + (2x_5 - x_2 - x_8)(x_1 x_7 + x_3 x_9)}{r_3 r_1^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$

$$\frac{dr_4}{dx_9} = \frac{(x_3 - x_6)\left(x_7^2 + x_8^2\right) + (x_9 - x_6)\left(x_1^2 + x_2^2\right)}{r_3 r_1^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$
$$+ \frac{(x_3 - x_9)\left[x_4 (x_1 - x_7) + x_5 (x_2 - x_8)\right] + (2x_6 - x_3 - x_9)(x_1 x_7 + x_2 x_8)}{r_3 r_1^3 \sqrt{1 - \frac{(\mathbf{r}_1 \cdot \mathbf{r}_3)^2}{r_1^2 \cdot r_3^2}}}$$

$$\frac{dr_4}{dx_{10}} = 0 \qquad \frac{dr_4}{dx_{11}} = 0 \qquad \frac{dr_4}{dx_{12}} = 0$$

$\frac{dr_5}{dx}$

$$\frac{dr_5}{dx_1} = \frac{(x_1 - x_{10}) r_3^2 (\mathbf{r}_2 \cdot \mathbf{r}_3) + (x_4 + x_{10} - 2x_1) r_2^2 r_3^2 + (x_1 - x_4) r_2^2 (\mathbf{r}_2 \cdot \mathbf{r}_3)}{r_2^3 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}}$$

$$\frac{dr_5}{dx_2} = \frac{(x_2 - x_{11}) r_3^2 (\mathbf{r}_2 \cdot \mathbf{r}_3) + (x_5 + x_{11} - 2x_2) r_2^2 r_3^2 + (x_2 - x_5) r_2^2 (\mathbf{r}_2 \cdot \mathbf{r}_3)}{r_2^3 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}}$$

$$\frac{dr_5}{dx_3} = \frac{(x_3 - x_{12}) r_3^2 (\mathbf{r}_2 \cdot \mathbf{r}_3) + (x_6 + x_{12} - 2x_3) r_2^2 r_3^2 + (x_3 - x_6) r_2^2 (\mathbf{r}_2 \cdot \mathbf{r}_3)}{r_2^3 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}}$$

$$\frac{dr_5}{dx_4} = \frac{(x_4 - x_{10})\left(x_2^2 + x_3^2\right) + (x_1 - x_{10})\left(x_5^2 + x_6^2\right)}{r_2 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}}$$
$$+ \frac{(x_1 - x_4)\left[x_{11} (x_2 - x_5) + x_{12} (x_3 - x_6)\right] + (2x_{10} - x_1 - x_4)(x_2 x_5 + x_3 x_6)}{r_2 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}}$$

$$\frac{dr_5}{dx_5} = \frac{(x_2 - x_{11})\left(x_4^2 + x_6^2\right) + (x_5 - x_{11})\left(x_1^2 + x_3^2\right)}{r_2 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}} + \frac{(x_2 - x_5)\left[x_{10}(x_1 - x_4) + x_{12}(x_3 - x_6)\right] + (2x_{11} - x_2 - x_5)(x_1 x_4 + x_3 x_6)}{r_2 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}}$$

$$\frac{dr_5}{dx_6} = \frac{(x_3 - x_{12})\left(x_4^2 + x_5^2\right) + (x_6 - x_{12})\left(x_1^2 + x_2^2\right)}{r_2 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}} + \frac{+(x_3 - x_6)\left[x_{10}(x_1 - x_4) + x_{11}(x_2 - x_5)\right] + (2x_{12} - x_3 - x_6)(x_2 x_5 + x_1 x_4)}{r_2 r_3^3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}}$$

$$\frac{dr_5}{dx_7} = 0 \qquad \frac{dr_5}{dx_8} = 0 \qquad \frac{dr_5}{dx_9} = 0$$

$$\frac{dr_5}{dx_{10}} = \frac{(x_1 - x_4)\left(x_{11}^2 + x_{12}^2\right) + (x_{10} - x_4)\left(x_2^2 + x_3^2\right)}{r_2^3 r_3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}} + \frac{+(x_1 - x_{10})\left[x_5(x_2 - x_{11}) + x_6(x_3 - x_{12})\right] + (2x_4 - x_1 - x_{10})(x_2 x_{11} + x_3 x_{12})}{r_2^3 r_3 \sqrt{1 - \frac{(\mathbf{r}_2 \cdot \mathbf{r}_3)^2}{r_2^2 \cdot r_3^2}}}$$

$$\frac{dr_5}{dx_{11}} = \frac{(x_{11}-x_5)\left(x_1^2+x_3^2\right)+(x_2-x_5)\left(x_{10}^2+x_{12}^2\right)}{r_2^3 r_3\sqrt{1-\frac{(\mathbf{r}_2\cdot\mathbf{r}_3)^2}{r_2^2\cdot r_3^2}}}$$
$$+\frac{+(x_2-x_{11})\left[x_4(x_1-x_{10})+x_6(x_3-x_{12})\right]+(2x_5-x_2-x_{11})(x_1x_{10}+x_3x_{12})}{r_2^3 r_3\sqrt{1-\frac{(\mathbf{r}_2\cdot\mathbf{r}_3)^2}{r_2^2\cdot r_3^2}}}$$

$$\frac{dr_5}{dx_{12}} = \frac{(x_{12}-x_6)\left(x_1^2+x_2^2\right)+(x_3-x_6)\left(x_{10}^2+x_{11}^2\right)}{r_2^3 r_3\sqrt{1-\frac{(\mathbf{r}_2\cdot\mathbf{r}_3)^2}{r_2^2\cdot r_3^2}}}$$
$$+\frac{+(x_3-x_{12})\left[x_4(x_1-x_{10})+x_5(x_2-x_{11})\right]+(2x_6-x_3-x_{12})(x_1x_{10}+x_2x_{11})}{r_2^3 r_3\sqrt{1-\frac{(\mathbf{r}_2\cdot\mathbf{r}_3)^2}{r_2^2\cdot r_3^2}}}$$

$\frac{dr_6}{dx}$

These expressions are a bit much for mathematica to simplify. Staring at the output I can make the following definitions to help simplify.

$$\begin{aligned}
&\mathbf{m} := \mathbf{r}_1 \times \mathbf{r}_2 = (a,b,c)\\
&a = x_3x_5 - x_2x_6 - x_3x_8 + x_6x_8 + x_2x_9 - x_5x_9\\
&b = -x_3x_4 + x_1x_6 + x_3x_7 - x_6x_7 - x_1x_9 + x_4x_9\\
&c = x_2x_4 - x_1x_5 - x_2x_7 + x_5x_7 + x_1x_8 - x_4x_8
\end{aligned} \tag{7.36}$$

$$
\begin{aligned}
&\mathbf{n} := \mathbf{r}_2 \times \mathbf{r}_3 = (d, e, f) \\
&d = x_{12}x_2 - x_{11}x_3 - x_{12}x_5 + x_3x_5 + x_{11}x_6 - x_2x_6 \\
&e = -x_1x_{12} + x_{10}x_3 + x_{12}x_4 - x_3x_4 + x_1x_6 - x_{10}x_6 \\
&f = x_1x_{11} - x_{10}x_2 - x_{11}x_4 + x_2x_4 - x_1x_5 + x_{10}x_5
\end{aligned} \tag{7.37}
$$

$$
\frac{dr_6}{dx_1} = \frac{\frac{[2c(-x_5+x_8)+2b(x_6-x_9)](\mathbf{m}\cdot\mathbf{n})}{|\mathbf{m}|^4} + \frac{f(-x_5+x_8)+c(x_{11}-x_5)+e(x_6-x_9)+b(x_6-x_{12})}{|\mathbf{m}|^2}}{\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}
$$

$$
\frac{dr_6}{dx_2} = \frac{\frac{[2c(x_4-x_7)+2a(-x_6+x_9)](\mathbf{m}\cdot\mathbf{n})}{|\mathbf{m}|^4} + \frac{f(x_4-x_7)+c(-x_{10}+x_4)+d(-x_6+x_9)+a(x_{12}-x_6)}{|\mathbf{m}|^2}}{\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}
$$

$$
\frac{dr_6}{dx_3} = \frac{\frac{[2b(-x_4+x_7)+2a(x_5-x_8)](\mathbf{m}\cdot\mathbf{n})}{|\mathbf{m}|^4} + \frac{e(-x_4+x_7)+d(-x_5-x_8)+b(x_{10}+x_4)+a(-x_{11}+x_5)}{|\mathbf{m}|^2}}{\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}
$$

$$
\frac{dr_6}{dx_4} = \frac{\frac{[2c(x_2-x_8)+2b(-x_3+x_9)](\mathbf{m}\cdot\mathbf{n})}{|\mathbf{m}|^4} + \frac{f(x_2-x_8)+c(-x_{11}+x_2)+e(-x_3+x_9)+b(x_{12}-x_3)}{|\mathbf{m}|^2}}{\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}
$$

$$
\frac{dr_6}{dx_5} = \frac{\frac{[2c(-x_1+x_7)+2a(x_3-x_9)](\mathbf{m}\cdot\mathbf{n})}{|\mathbf{m}|^4} + \frac{f(-x_1+x_7)+c(-x_1+x_{10})+d(x_3-x_9)+a(-x_{12}+x_3)}{|\mathbf{m}|^2}}{\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}
$$

$$\frac{dr_6}{dx_6} = \frac{\frac{[2b(x_1-x_7)+2a(-x_2+x_8)](\mathbf{m}\cdot\mathbf{n})}{|\mathbf{m}|^4} + \frac{e(x_1-x_7)+d(-x_2+x_8)+b(x_1-x_{10})+a(x_{11}-x_2)}{|\mathbf{m}|^2}}{\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}$$

$$\frac{dr_6}{dx_7} = \frac{\frac{[2c(-x_2+x_5)+2b(x_3-x_6)](\mathbf{m}\cdot\mathbf{n})}{|\mathbf{m}|^4} + \frac{f(-x_2+x_5)+e(x_3-x_6)}{|\mathbf{m}|^2}}{\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}$$

$$\frac{dr_6}{dx_8} = \frac{\frac{[2c(x_1-x_4)+2a(-x_3+x_6)](\mathbf{m}\cdot\mathbf{n})}{|\mathbf{m}|^4} + \frac{f(x_1-x_4)+d(-x_3+x_6)}{|\mathbf{m}|^2}}{\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}$$

$$\frac{dr_6}{dx_9} = \frac{\frac{[2b(-x_1+x_4)+2a(x_2-x_5)](\mathbf{m}\cdot\mathbf{n})}{|\mathbf{m}|^4} + \frac{e(-x_1+x_4)+d(x_2-x_5)}{|\mathbf{m}|^2}}{\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}$$

$$\frac{dr_6}{dx_{10}} = \frac{-\left[(-x_2+x_5)c+(x_3-x_6)b\right]}{|\mathbf{m}|^2\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}$$

$$\frac{dr_6}{dx_{11}} = \frac{-\left[(x_1-x_4)c+(-x_3+x_6)a\right]}{|\mathbf{m}|^2\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}$$

$$\frac{dr_6}{dx_{12}} = \frac{-\left[(-x_1+x_4)b+(x_2-x_5)a\right]}{|\mathbf{m}|^2\sqrt{1-\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{m}|^4}}}$$