SCATTERING AND RADIATION COMPUTATION WITH SPECTRAL METHODS

A Dissertation

by

JIANING ZHANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirement for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	R. Lee Panetta
Co-Chair of Committee,	Ping Yang
Committee Members,	Gerald R. North
	Anirban Bhattacharya
Head of Department,	Ping Yang

August 2016

Major Subject: Atmospheric Sciences

ABSTRACT

Surface roughness, a fundamental characteristic of atmospheric ice particles, is essential for defining an appropriate particle morphology model to simulate optical properties of atmospheric particles. This dissertation presents a dynamic stochastic parameterization approach based on combining the discrete differential geometry and stochastic partial differential equations to generate particle overall shapes and the degree of surface roughness. The scattering of light by particles modeled as Gaussian spheroids with size parameters up to 300 is simulated with the Invariant Imbedding T-Matrix (II-TM) method to investigate the effect of particle surface roughness on the single-scattering properties, including the phase matrix, single-scattering albedo, and extinction efficiency. It is shown that high-frequency oscillations of the phase matrix with respect to scattering angle are gradually suppressed as the degree of roughness increases. The dissertation presents a more thorough method of roughened particles in light scattering computation than various ad hoc methods reported in the literature. We discuss how surface roughness influences the Muller matrix patterns of ice particles. These results also enable better understanding of microphysics on ice surface and more accurate parameterization of atmospheric ice particles. We show that surface irregularity changes the phase matrix elements dramatically. An analysis of optical modeling of mineral dust aerosols as Gaussian spheroids is presented. The modeling results are compared with experimental measurements of feldspar to validate the applicability of roughened model particles. The Gaussian spheroids shows better data fitting than smooth spheroids. Furthermore, we analyze population density and

shape distributions of Gaussian spheroid for different mineral dust species. In addition to the scattering study, we propose a new Monte Carlo method for radiative transfer based on the Metropolis algorithm.

To my mother and father

ACKNOWLEDGMENTS

I would like to thank my two advisors Dr. Ping Yang and Dr. R. Lee Panetta first, who served as continuous sources of inspiration and encouragement during my study at the Department of Atmospheric Sciences. The most important thing I learned from Dr. Yang was ʹbe positiveʹ, and the most valuable lesson I learned from Dr. Panetta was ʹbe organisedʹ, two qualities I shall aspire to for the rest of my life.

I also enjoyed happly moments with all the group members, classmates and professors. From all professors and classmates, I began to know more about the atmosphere, science, and life. Especially, I am grateful to Dr. North for introducing me to geophyiscal researches from the view of a theoretical physicist, and to Dr. Bhattacharya for teaching me the statistical meachine learning, which I find to be the most interesting subject besides atmospheric physics.

Steven Schroeder, Dr. Panetta, and Dr. Yang donated a huge amout of their personal time to proofread materials of this dissertation and other papers. I thank them for their many valuable comments and suggestions for me. I am grateful to Dr. Lan Zhou for substituting on my dissertation committee.

This dissertation would not have been possible without from the help Jianping Liu, Dr. Lei Bi, and Dr. Guanglin Tang. I would also thank Dr. Bingqi Yi and other group members for their help. I really enjoyed my time with them.

Finally, I am deeply obliged to my parents for their unconditional love since I was born. Their positive life attitude, insight and support let me walk through many muddy

roads of life, and bring out a better me.

NOMENCLATURE

PSTD	Pseudo-Spectral Time Domain method
CPML	Convolution perfectly matched layer
IGOM	Improved geometric-optics method
FEM	Finite element method
IITM	Invariant-Imbedding T-Matrix
DDA	Discrete Dipole Approximation
DDG	Discrete Differential Geometry
QFT	Quantum Field Theory
BM	Boltzmann Machine
MRF	Markov Random Field
KPZ	Kardar–Parisi–Zhang

TABLE OF CONTENTS

LIST OF FIGURES

1 INTRODUCTION

The optical properties of ice crystals play an essential role in atmospheric radiative properties of ice clouds and ice fogs. Various factors, such as shapes, sizes and orientations affect the scattering properties by atmospheric ice particles. Compared with these, less attention has been paid to the surface roughness effects on light scattering. However, due to the surrounding environment, ice particles' surface can be roughened on a meso-scopic scale, and that such roughness can have an non-negligible effect on radiative equilibrium. Therefore, a effective ice geometry model will help us in simulating optical scattering of ice particles.

In addition, atmospheric aerosols, especially dust particles play a significant part in the global radiative, chemical, physical and biological processes. It is estimated that mineral dust accounts for 30-50% of the total weight of atmospheric aerosols. Dust particles result in a direct radiative forcing when absorbing and emitting radiation, and act as cloud nucleation nuclei affecting the global radiation equilibrium indirectly. Hence, the modeling of dust particles has a substantial role in estimating their radiative properties. In addition, dust particles can mix into snow or glacier ice and reduces their albedo, whiche increases the melting rate and acts as a radiatiwon feedback.

This dissertation was initially motivated by the geometric modeling of atmospheric particles. We studied the optical scattering and radiation properties of ice and dust particles modeled by a stochastic growth model. After incorporating differential geometry, we

established the connection between surface growth and surface geometry. Futhermore, we applied similar ideas to build a relation between the geometry of curves and the ray tracing, and proposed a path integration formulation and variance reduction algorithm for Monte Carlo ray tracing.

1.1   Summary of Original Contributions

The work in this dissertation builds on Maxwell's electromagnetic theory and Boltzmann's transport framework. The computation techniques used in this dissertation involves invariant imbedding T-Matrix, pseudo-spectral time domain method, and Monte Carlo methods. Our new contributions are as follows:

A surface growth model for optical scattering by roughened particles. We introduce surface growth theory to particle geometry modeling. Traditional surface growth theory on $\mathbb{R}^2$ is extended to a closed surface in $\mathbb{R}^3$. We build the connection between surface growth theory and the differential geometry of a surface. We also connect the relation between stochastic differential equations, the Gaussian Markov random field and discrete differential geometry for our surface growth model.

Optical scattering properties of roughened ice and dust particles. We simulate the optical scattering for ice and dust particles using our surface growth model particle.

Metropolis Monte Carlo radiative transfer algorithm. We propose a variance reduction Monte Carlo method that is able to solve the photon Boltzmann equation more efficiently. We first introduce a Frenet curve to describe the ray propagation, in which we have

explained the ray using the discrete differential geometry of the curve. The photon transport equation is formulated in a completely new path integral framework. We interpret our method as a generalized version of the classical Heisenberg model.

1.2   Organization of the Dissertation

This dissertation is divided into a total of 5 chapters. Chapter 2 presents the basics of scattering theory and computational methods, as well as a new compressed orientation averaging scheme and new inverse modeling method based on the Boltzmann machine neural network model. In Chapter 3, we propose a surface growth model for modeling the geometry of roughened particles, where the Langevin equation is solved assuming a discrete differential geometry of the surface. In Chapter 4, we simulate optical scattering by dust particles and propose a Metropolis Monte Carlo radiative transfer algorithm by analog of the Heisenberg model in statistical physics. We summarize the conclusions in Chapter 5.

2 THEORY AND COMPUTATION OF SCATTERING AND RADIATION

Light scattering by atmospheric particles including ice crystals, dust, soot, and aggregates, has been studied for several decades with wide applications in remote sensing, atmospheric micrhophysics and climate research[1, 2]. A wave optics treatment of electromagnetic or light wave scattering involves solving the Maxwell's equations. About one century ago, mie [3] first considered spherical particles scattering of electromagnetic waves. Later, Stratton[4], Born and Wolf[5] presented comprehensive and precise foundations of Mie theory in their classic books. Many computational techniques have been proposed and used for solving the electromagnetic wave scattering problem, such as spectral methods, finite-dierence method, finite-element method, and discrete dipole approximation(DDA method[6, 7, 9]. Nearly all of these methods are used in light scattering computation of atmospheric particles. With increasing complexity and particle sizes, the need for solving high frequency problem for complex 3D shapes is increasing.T his triggers a need for fast and efficient computational methods for light scattering[1].

2.1 Theory of Scattering

First, let us review the field scattering theory using Feynman's path integral formulation. We follow the approach in [10, 11].

The Lagrangian density for macroscopic electromagnetic fields is given by

$$\mathscr{L} = \frac{1}{2}(\mathbf{E} \cdot \mathbf{D} - \mathbf{B} \cdot \mathbf{H}) \tag{2.1}$$

Using the relations $\mathbf{D} = \varepsilon \mathbf{E}$, $\mathbf{B} = \mu \mathbf{H}$ (Here ε and μ are the permittivity and permeability), we obtain the electromagnetic action:

$$
\begin{aligned}
S &= \frac{1}{2} \int_0^T dt \int d^3x (\mathbf{E}^* \cdot \varepsilon \mathbf{E} - \mathbf{H}^* \cdot \mu \mathbf{H}) \\
&= \frac{1}{2} T \sum_{n=-\infty}^{\infty} \int d^3x (\mathbf{E}^* \cdot \varepsilon \mathbf{E} - \mathbf{H}^* \cdot \mu \mathbf{H})
\end{aligned}
$$

where we have expanded the electric field $\mathbf{E}$ in terms of the Fourier series $\mathbf{E}(\mathbf{x}, t) = \sum_{n=-\infty}^{\infty} \mathbf{E}(\mathbf{x}, \omega_n) e^{-j\omega_n t}$, $\omega_n = 2\pi n / T$.

Substituting Maxwell's equations into above equation, we have the action only depends on $\mathbf{E}$

$$S(T) = \frac{1}{2} T \sum_{n=-\infty}^{\infty} \int d^3x [\mathbf{E}^* \cdot (\mathbb{I} - k_n^{-2} \nabla \times \nabla \times) \mathbf{E} + k_n^{-2} \mathbf{E}^* \cdot \mathbb{V} \mathbf{E}] \tag{2.2}$$

where

$$\mathbb{V} = \mathbb{I} k_n^2 (\varepsilon(\omega_n, \mathbf{x}) - 1) \tag{2.3}$$

is the potential operator.

According to the Euler-Lagrange variational principle, the field equation reads

$$(\mathbb{H}_0 - \mathbb{V}) \mathbf{E}(\mathbf{x}, \omega) = k_0^2 \mathbb{I} \mathbf{E}(\mathbf{x}, \omega)$$

where

$$\mathbb{H}_0 = \nabla \times \nabla \times$$

5

The corresponding free field Green function $\mathbb{G}_0(k)$ is

$$(\mathbb{H}_0 - k_0^2 \mathbb{I})\mathbb{G}_0 = \mathbb{I}\delta(\mathbf{x} - \mathbf{x}') \tag{2.4}$$

In coordinate representation, the Green function of the free field becomes

$$\mathbb{G}_0(\omega, \mathbf{x}, \mathbf{x}') = (\mathbb{I} - k_0^{-2}\nabla \otimes \nabla')\frac{e^{-jk_0|\mathbf{x}-\mathbf{x}'|}}{4\pi|\mathbf{x} - \mathbf{x}'|} \tag{2.5}$$

We arrive at the Lippmann-Schwinger [] equation:

$$\mathbf{E} = \mathbf{E}_{inc} + \mathbb{G}_0\mathbb{V}\mathbf{E} \tag{2.6}$$

After iteratively solving the Lippmann-Schwinger euqation, we obtain the $\mathbb{T}$ operator

$$\mathbb{T} = \mathbb{V}\frac{\mathbb{I}}{\mathbb{I} - \mathbb{G}_0\mathbb{V}} = \mathbb{V}\mathbb{G}\mathbb{G}_0^{-1} \tag{2.7}$$

Scattered field and incident field are related through $\mathbb{T}$ operator,

$$\mathbf{E}_{sc} = \mathbb{G}_0\mathbb{T}\mathbf{E}_{inc} \tag{2.8}$$

2.2 Computation Methods of Scattering

2.2.1 Spectral Methods and T-Matrix

Spectral methods originate from solving partial differential equations (PDEs) with series expansion[13]. When we solve a PDE by separation of variables in certain coordinates (e.g. cartesian, cylindrical or spherical coordinates), we use orthogonal functions such complex exponential, Bessel, spherical Bessel, spherical harmonics as basis. More and one century ago, Debye, Lorenz, and Mie independently solved the scattering by spherical particles with spectral methods. Without going into details, we write the expression

for the incident field and scattered field using vector spherical wave functions.

$$\mathbf{E}_{inc} = \sum_{l,m} a_{lm}\mathbf{M}_{lm}^{reg} + b_{lm}\mathbf{N}_{lm}^{reg}$$

$$\mathbf{E}_{sc} = \sum_{l,m} c_{lm}\mathbf{M}_{lm}^{out} + d_{lm}\mathbf{N}_{lm}^{out}$$

From the Lippmann-Schwinger equation derived in the above section, a matrix representation for the $\mathbb{T}$ operator can be obtained using a volume integral equation:

$$\begin{pmatrix} c_{lm} \\ d_{lm} \end{pmatrix} = \begin{pmatrix} T_{lml'm'}^{EE} & T_{lml'm'}^{EM} \\ T_{lml'm'}^{ME} & T_{lml'm'}^{MM} \end{pmatrix} \begin{pmatrix} a_{l'm'} \\ b_{l'm'} \end{pmatrix}$$

E, M represent the polarization type. We have used Einstein summation convection here.

2.2.2 Pseudo-Spectral Methods

Another class of spectral methods is to approximate the functions by using interplants of nodes, which are called Pseudo-Spectral methods[9]. A principle advantage of the pseudo-spectral methods is that its computing complexity can be considerably reduced with the use of fast Fourier transform(FFT).

For example, Fourier Pseudo-Spectral methods apply FFT to approximate the differential operator as follows:

$$\nabla_x \phi(x) = \mathbb{F}_x^{-1}[ik_x\mathbb{F}_x]\phi(x), \tag{2.9}$$

The $\mathbb{F}$ is the FFT operator, $\mathbb{F}^{-1}$ is its inverse.

It is straightforward to wirte governing equations for time dependent Maxwell's equations as follows:

$$E_x|_{i,j,k}^{n+1/2} = E_x|_{i,j,k}^{n-1/2} + \frac{\Delta t}{\varepsilon_r|_{i,j,k}}\{\mathbb{F}_y^{-1}[jk_y\mathbb{F}_y(H_z)]|_{i,j,k}^n - \mathbb{F}_z^{-1}[jk_z\mathbb{F}_z(H_y)]|_{i,j,k}^n\}$$

$$E_y|_{i,j,k}^{n+1/2} = E_y|_{i,j,k}^{n-1/2} + \frac{\Delta t}{\varepsilon_r|_{i,j,k}}\{\mathbb{F}_z^{-1}[jk_z\mathbb{F}_z(H_x)]|_{i,j,k}^{n} - \mathbb{F}_x^{-1}[jk_x\mathbb{F}_x(H_z)]|_{i,j,k}^{n}\}$$

$$E_z|_{i,j,k}^{n+1/2} = E_z|_{i,j,k}^{n-1/2} + \frac{\Delta t}{\varepsilon_r|_{i,j,k}}\{\mathbb{F}_x^{-1}[jk_x\mathbb{F}_x(H_y)]|_{i,j,k}^{n} - \mathbb{F}_y^{-1}[jk_y\mathbb{F}_y(H_x)]|_{i,j,k}^{n}\}$$

$$H_x|_{i,j,k}^{n+1} = H_x|_{i,j,k}^{n} - \frac{\Delta t}{\mu_r|_{i,j,k}}\{\mathbb{F}_y^{-1}[jk_y\mathbb{F}_y(E_z)]|_{i,j,k}^{n} - \mathbb{F}_z^{-1}[jk_z\mathbb{F}_z(E_y)]|_{i,j,k}^{n+1/2}\}$$

$$H_y|_{i,j,k}^{n+1} = H_y|_{i,j,k}^{n} - \frac{\Delta t}{\mu_r|_{i,j,k}}\{\mathbb{F}_z^{-1}[jk_z\mathbb{F}_z(E_x)]|_{i,j,k}^{n} - \mathbb{F}_x^{-1}[jk_x\mathbb{F}_x(E_z)]|_{i,j,k}^{n+1/2}\}$$

$$H_z|_{i,j,k}^{n+1} = H_z|_{i,j,k}^{n} - \frac{\Delta t}{\mu_r|_{i,j,k}}\{\mathbb{F}_x^{-1}[jk_x\mathbb{F}_x(E_y)]|_{i,j,k}^{n} - \mathbb{F}_y^{-1}[jk_y\mathbb{F}_y(E_x)]|_{i,j,k}^{n+1/2}\}$$

i,j,k are indexes of nodes, n is time index and Δt is the time step. A scattered field formulation is used here. In Fig. 2.1, we compare the the Muller matrix computed by IITM and PSTD for hexagonal prism with refractive index 1.3, size parameter $kL = 50$.

2.2.3   Finite Element Method

It is quite challenging to particle solve scattering problem with complex geometric shape. Finite-element method(FEM) is an excellent method which allows PDE solving with ultra high resolution and complex geometric description, and higher-order convergence. In this section, we introduce FEM as a computational tool for solving the scattering problem.

Let us consider the problem of a scalar field scattering. We define the energy functional through the path integral, which sums all configurations of the fields constrained by the boundary conditions. The Lagrangian density is

$$\mathscr{L} = \frac{1}{2}(\phi(1 - k^{-2}\Delta)\phi - \phi V\phi)$$

where $V = k^2(1 - n^2)$, n is the refractive index for example, k is the wave number. In the
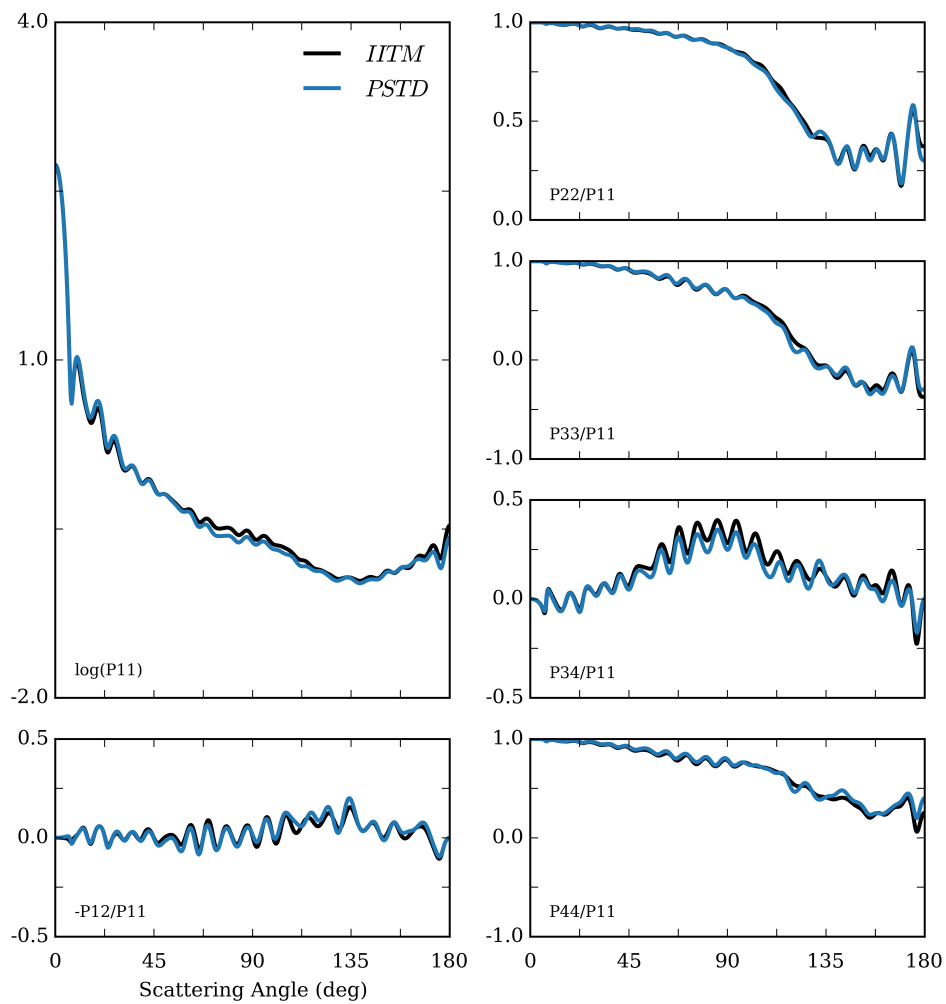
Figure 2.1: Mueller matrix computed by PSTD versus IITM

functional integral, we will sum over configurations of the field. The ground state energy of the scalar field can be obtained, and the overlap between the initial state of the system with the final state after time T can be expressed as a functional integral with the field:

$$\langle \phi_f | e^{-iHT} | \phi_i \rangle = \int \mathscr{D}\phi\, e^{-jS}$$

If the initial and final states are set equal and summed over the resulting functional integration leads to

$$Z = \mathrm{Tr}\, e^{-jHT}$$

The partition function that describes this system at temperature $1/\beta$ is defined by

$$Z = \mathrm{Tr}\, e^{-\beta H}$$

and the free energy F of the field is

$$F = -1/\beta \ln Z$$

When we replace the integral by the maximum value of the integrand, corresponding to the most probable configuration of the field (Saddle point approximation) we get its mean field solution. $Z = Z_{sp}$, and the corresponding saddle point free energy is

$$\beta F_{sp} = -\ln Z_{sp} = \min(S)$$

The Rayleigh-Ritz method is used to extremize the free energy functional over a properly constructed subspace U_h of the admissible fields U. And finite element method can also be formulated through both the Rayleigh-Ritz method and weighted-residual Galerkin method. Hence, a connection between FEM and QFT can be formulated through the path integral approach.

Numerical simulation and results

Scattering is formulated by a scalar Helmholtz equation in frequency domain:

$$(\Delta + \frac{\omega^2}{c^2}n^2)\phi_{sc}(\mathbf{x}, \omega) = f_{inc} \tag{2.10}$$

Hereafter, we omit the subscript of ϕ. The exterior boundary is treated using a perfectly matched layer (PML)[15, 14, 16]. In the mid 1990s, Jeanne-Pierre Berenger first proposed the idea of a perfectly matched layer(PML)[17, 17], an artical absorbing boundary regions to make possible a finite computational domain in wave propagation simulations. The PML is designed to have the characteristics that electromagnetic waves of arbitrarily polarization and any frequency impinging on a PML region will be absorbed in the medium without reection. Several formulations of PML have been proposed. A split field formulation of Maxwell's equations was used in Berenger's original PML. After that, it was illustrated by Chew et. al.[19] that Berenger's PML is equivalent to expressing Maxwell's equation in complex-stretched coordinates. Sacks et al. later proposed a uniaxial perfectly matched layer(UPML)[14], in which the PML is considered as an artificial anisotropic medium.

Complex-Frequency Shifted (CFS) stretched parameters s_j, in the stretched coordinate formulation are expressed as

$$s_j = 1 + i\frac{\sigma_j}{\omega} \tag{2.11}$$

where σ_j is conductivity for absorbing the incident wave to the PML region, $i = \sqrt{-1}$, The

parameters are spatially scaled to eliminate the reflection waves.

$$
\sigma_j(x_j) \;=\;
\begin{cases}
\frac{|x_j - l_j|^2}{d^2}\sigma_j^*, & l_j \le x_j \le l_j + d, \\[2mm]
0, & \text{otherwise.}
\end{cases}
\tag{2.12}
$$

Here l_j is the x_j-coordinate of the PML interface, d is the thickness of the PML layer, and

σ_i^* is the maximum value of σ_j at $x_j = l_j + d$. In the whole computational domain, we can

use a unified formulation,

$$
(\nabla \bar{\bar{a}} \nabla + \frac{\omega^2}{c^2} n^2 b)\phi = f_{inc}
\tag{2.13}
$$

where

$$
\bar{\bar{a}} =
\begin{pmatrix}
\frac{s_2 s_3}{s_1} & 0 & 0 \\[2mm]
0 & \frac{s_1 s_3}{s_2} & 0 \\[2mm]
0 & 0 & \frac{s_1 s_2}{s_3}
\end{pmatrix}
\tag{2.14}
$$

and $b = s_1 s_2 s_3$. Fig. 2.2 shows the incidence of harmonic plane into the PML. In the table

below, we show the errors of 1D PML with a normal incident harmonic plane wave as we

refine the mesh by double the resolution each time

For a FEM discretization, the Helmholtz equation is first transformed into a weak

formulation. In dealii, we use the Galerkin approach to get the system matrix. In the end,

Table 2.1: Performance of PML

Refinement	L_2 error	L_∞ error
1	0.053667	0.095451
2	0.0134613	0.0239318
3	0.0033727	0.0059828
4	8.47613e-4	1.50216e-3
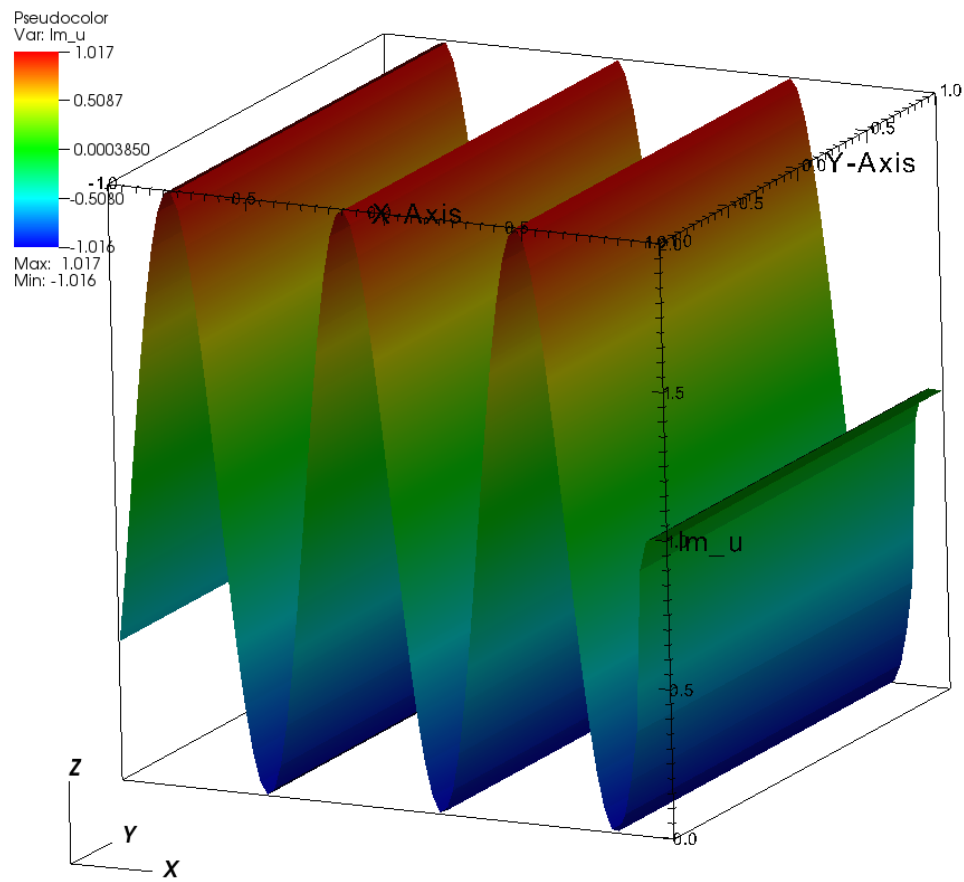
[H]

Figure 2.2: Normal incidence with harmonic plane wave on PML
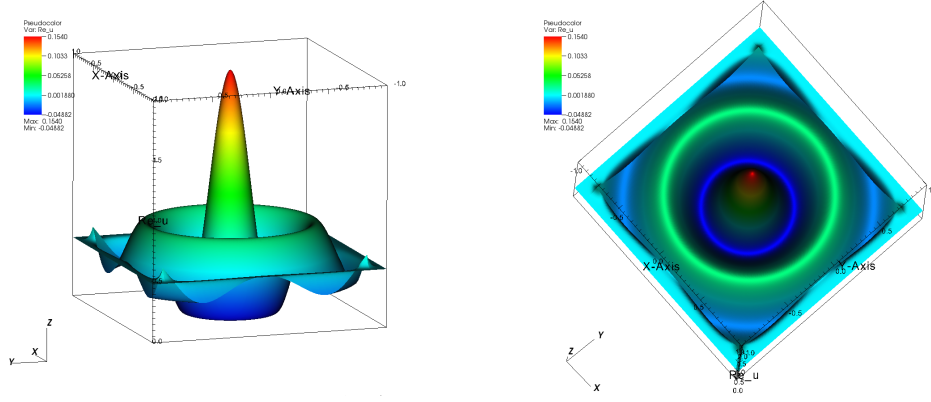
Figure 2.3: Different views of a free propagating spherical wave.

after triangulation and integration. we arrive at system linear equations:

$$Ax = b \qquad (2.15)$$

In the simulation, we have used a Q_1 element but the extension to Q_p is quite simple using deal.ii[20].Fig. 2.3 shows a sample solutotion for a spherical wave.

It should be noted that we just use a simple square mesh to discretize the domain. All these simulations are based on FEM in which a sparse LU decomposition provided by UMFPACK is chosed instead to solve the linear system. Since thhe deal.ii interface to UMFPACK is given by the SparseDirectUMFPACK class, it is very straightforward to implement. In fig. 2.4, a scalar wave scattered by an dielectric sphere with refractive index 2.0 has been given.

2.3 Particles' Orientation and Direction Statistics

In scattering simulations, the computed optical properties are usually averaged by taking a random orientation assumption. The orientations of particles compose a sampling
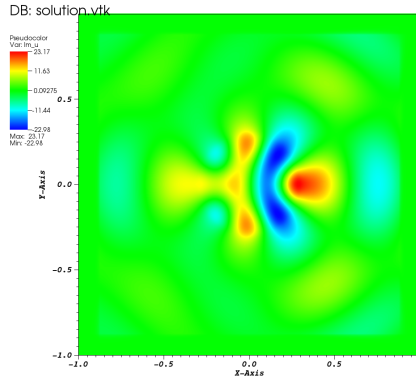
Figure 2.4: Imaginary part of ϕ.   Figure 2.5: Real part of ϕ.
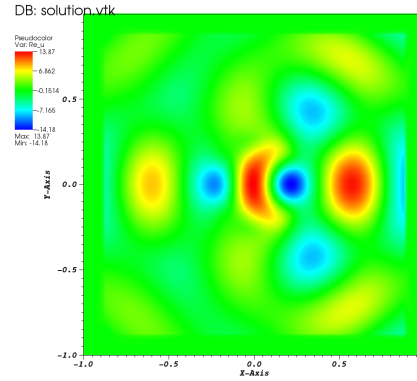
space, which can be described with a probability distribution on $\mathbb{S}^n$. In the real atmosphere, aerodynamic forces tend to orient the drifting particles nearly horizontally. This configuration is dynamically stable in small Renold's number, and correcting forces from tiny deviations also restore their orientation. In the studies of oriented particles, previous works often assume that the orientation of ice plates follows a Gaussian distribution. However, the Gaussian distribution is defined on $\mathbb{R}^n$ instead of on $\mathbb{S}^n$. The analog of the Gaussian distribution on a sphere is Von Mises-Fisher distribution[21]. Let us first review the direction statstics in this section. The probability distribution function(PDF) $f(\phi, \theta, \psi)$ of particle orientations is defined through Euler angles ϕ, θ, ψ with normalization:

$$\int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} f(\phi, \theta, \psi) d\phi \sin\theta d\theta d\psi = 1$$

We write Von Mises-Fishe distribution on $\mathbb{S}^2$ as an example is

$$f_F(\mathbf{x}; \mu, \kappa) = C_F(\kappa) exp(\kappa \mu^T \mathbf{x})$$

Here κ is a non-negative real number, μ is the mean unit vector on the sphere, $\mathbf{x} =$

15

$(sin\theta cos\phi, sin\theta sin\phi, cos\theta)$ and the normalization constant C_F

$$C_F(\kappa) = \frac{\kappa}{4\pi \sinh\kappa}$$

For quasi-horizontally orientated ice plates, $\mu = (0,0,1)$. Integrating out the azimuthal

freedom, the proposed PDF becomes

$$f_h(\theta; \kappa) = \frac{\kappa}{2\sinh\kappa} \exp(\kappa cos\theta) \qquad (2.16)$$

with normalization

$$\int_{-\pi/2}^{\pi/2} f_h(\theta; \kappa) sin\theta d\theta = 1$$

If we define a uniform distribution of particle orientation, it is quite efficient to derive

the average scattering/aborbtion cross sections from the T-Matrix method.

$$\langle C_{sc} \rangle \;\; = \;\; \frac{2\pi}{k^2} \text{Tr } TT^\dagger \qquad (2.17)$$

$$\langle C_{abs} \rangle = -\frac{2\pi}{k^2} \text{Tr}[\frac{1}{2}(T + T^\dagger) + TT^\dagger]$$

An pptical quantity such as the cross section can be viewed as a direction distribution

on the sphere. A given spherical function $f(\theta, \phi)$ can be expressed in terms of spherical

harmonics (e.g. see fig 2.5):

$$f(\theta, \phi) = \sum_{l=0}^{n} \sum_{m=-l}^{l} x_{l,m} Y_{l,m}$$

where the variable θ is the polar angle with $\theta \in [0, \pi)$ and ϕ is the azimuthal coordinate

with $\phi \in [0, 2\pi)$. where x_l^m is the spherical harmonic coefficient given by the inner product

$x_l^m = \langle f, Y_l^m \rangle$. The function is considered sparse if the coefficient vector has few nonzero
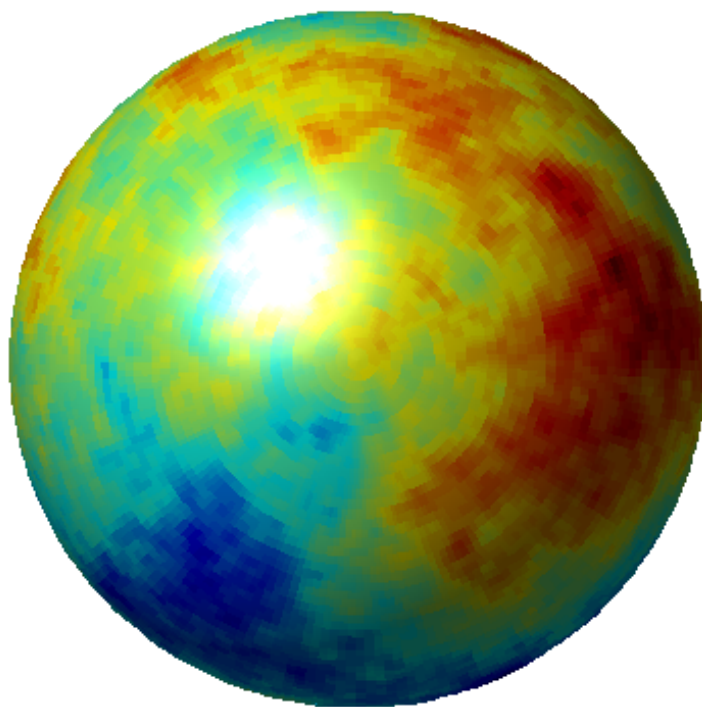
components.

Figure 2.6: Spherical harmonics expansion of a spherical function.

The spherical harmonic functions form the canonical orthogonal basis for functions on the sphere and are defined as

$$Y_{l,m}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(cos\theta)e^{im\phi}$$

where the first term is a normalization factor to ensure $Y_{l,m}(\theta,\phi)$ has unit energy, and $P_l^m(cos\theta)$ are the associated Legendre functions. In his dissertation $\theta \in [0,\pi]$ denotes the co-latitude measured from the positive z-axis and $\phi \in [0,2\pi]$ denotes the azimuth measured counter-clockwise from the positive x-axis.

Nevertheless, according to the well-known sampling theorem, it seems impossible to remove some components without causing causing a negative impact on the resolution. Now, suppose that a signal $x \in \mathbb{R}^N$ is sparse in some basis $x = \Psi x'$, which means x' has k ($k \ll N$) nonzero entries. Then a measurement matrix $\Phi \in \mathbb{R}^{m \times N}$ is used to sense x and obtain a measurement vector $f \in \mathbb{R}^m$. The central idea of compressed sensing[22, 23] is that it is possible to reconstruct sparse signals of scientific interest accurately and sometimes exactly by a number of incoherent samples which is far smaller than N. In general, such a recovery of x by performing l_0-minimization is a NP-hard problem. An alternative procedure called l_1-minimization is usually used:

$$\min_{x'}||x'||_{l_1} \qquad \text{subject to} \qquad y = \Phi\Psi x' = \Upsilon x',$$

This is a convex optimization problem, and many numerical algorithms apply for its solution.

To formulate a compressed model for orientation averaging, the simulated results are viewed as the measured data, and spherical harmonics are viewed as the measurement

matrix. A sparse spherical harmonic coefficient vector can be recovered from a smaller

number of measurements than classically required as a unique minimizer of

$$x^* = \arg\min||\tilde{x}||_1 \quad \text{such that} \quad A\tilde{x} = f,$$

where $||\tilde{x}||_1 = \sum_{n=1}^{N} |\tilde{x}_n|$ and x^* is the vector of recovered spherical harmonic coefficients.

Since we have made a uniformly random distributed assumption for the particle orientations, it appears appropriate to sample evenly on the sphere. For this purpose, the area represented by every sample point should be almost the same. As an appealing choice, he Fibonacci lattice provides a way to distribute each point almost the same area, and is easier to construct.

Our goal is to compute the spherical harmonic measurement matrix up to l_{max}. The arguments for a spherical function are a set of spherical points (θ, ϕ), which are Fibonacci lattice in our case. Hence, we can formulate a linear system for orientation averaging as follows:

$$
\begin{pmatrix}
A_{1,1} & A_{1,2} & \ldots & A_{1,k} \\
A_{2,1} & A_{2,2} & \ldots & A_{2,k} \\
\vdots & \vdots & \ddots & \vdots \\
A_{k,1} & A_{k,2} & \ldots & A_{2,k}
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_k
\end{pmatrix}
=
\begin{pmatrix}
f_1 \\
f_3 \\
\vdots \\
f_k
\end{pmatrix}
$$

where $j = l^2 + l + m + 1$, and $k = (l_{max} + 1)2$. Note that we use an indexing scheme that assigns a unique index j to every pair (l, m). Because each is an estimate of the original

coefficient for $j = l^2 + l + m + 1$, we can reconstruct the original function as

$$\hat{f}(\theta, \phi) = \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} \hat{x}_{l,m} Y_{l,m}(\theta, \phi)$$

After that, l_1 minimization can be used to represent and reconstruct the spherical function.

This is a convex optimization problem, and many numerical algorithms apply for its solution. This differs from classical signal processing (the sensing matrix elements can be independently selected from a random distribution). Instead, only the rows of the sensing matrix can be randomly measured in). The problem considered now is to reconstruct a state from measured insufficient data. Computer simulation has been performed to demonstrate the effectiveness of the aforementioned scheme. All the recovery processes were implemented in C++ For l_1-minimization, we adopt a linearized Bergman iteration algorithm[24].

2.4 Inverse Radiative Transfer

2.4.1 Bayesian Inverse Modeling

Roughly speaking, inverse transport is a mathematical approach for estimating the state parameters of a system using the radiative observations of that system. The radiation transport through the medium is governed by the following radiative transfer equation, which is a specific form of the Boltzmann equation.

$$(\mathbf{t} \cdot \nabla + c)L(x,t) = b \int_{\mathbb{S}^2} f_s(\mathbf{t}, \mathbf{t}')L d\mathbf{t}' + J(x, \omega)$$

Given remote sensing measurements such as from a satelllite, there are many methods to make the estimate the geophysical state of the earth. Bayesian methods are often used as

20

a tool to obtain an optimal solution of the states. Here, we briefly review the traditional Bayesian approach.

Consider a geophysical quantity we are interested, for example optical depth of the cloud. We have an apriori estimate of its value $x \pm \sigma_x$. (σ_x is the standard deviation.) A measurement of the radiance ($y \pm \sigma_y$) has been obtained. A realtion ship between them can be obtained through Boltzmann equation:

$$y = f(x) \pm \sigma$$

Following the Bayesian approach, we have

$$p(x|y) \sim \exp[-\frac{(x'-x)}{2\sigma_x^2} - \frac{(y-f(x))}{2\sigma_y'^2}]$$

The Bayesian approach has achieved tremendous results in geophysical sciences. However, we have to add about smoothing constraints to make the solution unique. In the following section we will introduce another method (Boltzmann machine)[25] that we believe will lead to more natural results, especially for a high dimensional inverse problem.

2.4.2 Inverse Modeling with Boltzmann Machine

A Boltzmann machine BM, interpreted as a neural network model, is a stochastic binary machines inspired by statistical physics[26]. The BM came from the formal equivalence between statistical physics and the dynamic behavior of neural networks named after the physicsist Ludwig Boltzmann. The Boltzmann machine is a network of stochastic processing units for learning important aspects of an unknown probability distribution by using an observed data set for traning. Training a Boltzmann machine is equivalent to

making statistical inference about the unknown parameters from the training data. Boltzmann machines can be regarded as undirected graphical models or Markov random fields with interaction terms need to learn. Boltzmann machines can also belong to the Ising class. Generally speaking, the learning algorithm of a BM is quite slow. Instead, topologically restricted Boltzmann machines (RBMs) is often used as a substitute to enhance the learning speed. In a probabilistic graph of the BM there are two layer nodes, visible and hidden nodes. All nodes are connected with each other. In RBMs[26], each node is only interacting with the nodes in the other layer. There are no intra-layer connections among the nodes meaning no direct interectation between the same type of particles.

However, it is still computationally expensive to compute the likelihood functin and its gradient using a RBM for statistical inference. Thus, Markov chain Monte Carlo is often employed to sample approximations of the likelihood function and its gradient of a RBM. Here, a Markov chain Monte Carlo (MCMC) method via Gibbs sampling is applied to this problem. After learning, an RBM provides a representation of the probabilistic distribution for the underlying training data. Therefore, the input data can be evaluated with the RBM. In this section, both the sampling schemes and applications of RBMs will be discussed.

Probabilistic graphical models, also called probabilistic networks, encode complex probability distributions over high dimensional space by mapping conditional dependence and independence properties between random vectors on a graph structure. In graphical models, the random vectors correspond to nodes, and probabilistic interactions between them are represented by edges.

22

Graphical models can be divided into two classes according to the interaction type--directed graph models(Bayesian Network) and undirected graph models also knowns Markov random fields. RBMs belong to Markov random fields.

Markov random fields are useful in modeling a variety of phenomena where one cannot naturally ascribe a direction to the interaction between nodes in a graph. The prototype of the Markov random field is the Ising model, a mathematical model of magnetism in statistical physics. In Ising model, spins are represented by binary discrete variables. The interactions between spins, arranged in a lattice, is limited to nearest neighbors. And the two-dimensional square-lattice Ising model is one of the simplest statistical models to show a phase transition.

Statistical physics is a probabilistic approach to study macroscopic properties involving a large number of degrees of freedom[27]. In physics, the macrostate M of a many-body system is phenomenologically characterized in terms of thermodynamic coordinates which obey the laws of thermodynamics. Instead of solving a large number of Hamiltonian equations corresponding to a single state, statistical mechanics studies the ensemble of system microstates $\{\mathbf{x}\}$ and the probability distribution of the equilibrium ensemble.

Suppose that we have a system of N binary random variables x_i that can take the values 1. For example, the random variables can represent the spins in some lattice as the Ising model. In thermal equilibrium, the probability density function of each microstate will be given by the Boltzmann distribution[27]:

$$p(\mathbf{x}|T,\phi) = \frac{1}{Z(T,\phi)}e^{-\beta\mathscr{H}(\mathbf{x})} \tag{2.18}$$

$H(\mathbf{x})$ is the corresponding energy of each system state and $\beta = 1/T$ (Boltzmann's constant k_B is set to be 1). The partition function is

$$Z(T,\phi) = \sum_{\mathbf{x}} e^{-\beta \mathcal{H}(\mathbf{x})} \qquad (2.19)$$

The PDF of system energy is given by

$$p(H) = \frac{1}{Z(T,\phi)} e^{-\beta F(H)} \qquad (2.20)$$

where $F = H - TS$ is the Helmholtz free energy. Since $p(H)$ has a sharp peak at a most probable energy which minimizes $F(H)$, we have

$$F = -T\ln Z \qquad (2.21)$$

The energy of of spin configuration $\{x_i\}$ for the Ising model is given by

$$\mathcal{H} = -\sum_{<i,j>} w_{ij} x_i x_j - \sum_i h_i x_i, \qquad (2.22)$$

where $\{h_i\}$ is external magnetic fields, w_{ij} is the interaction between spins at sites i and j. The first sum is over the neighbor bonds (Fig. 2.7 shows a simulation using Ising model.). We define $\mathcal{M} = \sum_i x_i$ as the magnetization. Now, we can construct a undirected graph $G = (V,E)$ with Ising model, each spin is represented by one node in the set V and interaction between neighbour spins are denoted by edge set E[26].

By analogy with Ising model, the energy of the Boltzmann machine is defined as

$$\mathcal{H} = \sum_{<i,j>} b_{ij} x_i x_j + \sum_i h_i x_i \qquad (2.23)$$

Invoking the Gibbs distribution, we assign each microstate $\mathbf{x}$ of the network a probability density function that as above.

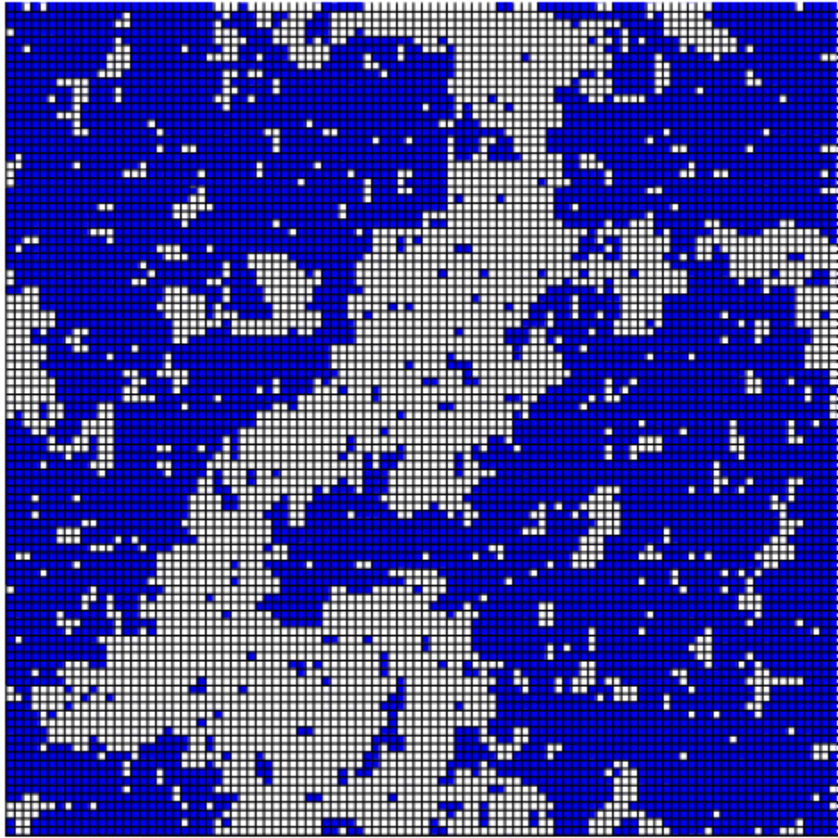Given a sample data, we would like to learn and make predictions. Boltzmann ma-

Figure 2.7: Ising model with $N = 100$, $T = 2.26918$

chine learning is unsupervised learning which means learn the parameters of an unknown distribution from sample data. Usually, maximum-likelihood estimation is explored to make the inference of the parameters. For MRFs, this corresponds to finding th coefficients in the energy function of MRF that maximizes the probability of a given data set. Training or learning means finding the parameters that maximize the likelihood function given the training data set[26].

In general, it is not possible to find the maximum likelihood estimation of parameters analytically for the Gibbs distribution of an MRF.

$$\ln \mathscr{L}(\boldsymbol{\theta}|S) = \ln \prod_{i=1}^{l} p(\mathbf{x}_i|\boldsymbol{\theta}) = \sum_{i=1}^{l} \ln p(\mathbf{x}_i|\boldsymbol{\theta}) \qquad (2.24)$$

While maximizing the log-likelihood function corresponds to minimizing the distance between model distribution p and data distribution q, which is measured by the Kullback-Leibler divergence (KL divergence)

$$KL(q||p) = \int d\mathbf{x} q(\mathbf{x}) \ln \frac{q(\mathbf{x})}{p(\mathbf{x})} = \sum_{x} [q(\mathbf{x}) \ln q(\mathbf{x}) - q(\mathbf{x}) \ln p(\mathbf{x})] \qquad (2.25)$$

The usual method to find maximum of a function is optimization by gradient ascent. First, update the parameters iteratively:

$$\theta^{t+1} = \theta^t + \eta \frac{\partial}{\partial \theta^t} (\ln \mathscr{L}(\theta^t|S)) \qquad (2.26)$$

where $\eta \in R^+$ is the learning rate constant.

However, computing the gradient of log-likelihood directly leads to a exponential computational complexity[26]. To avoid this computational burden, MCMC methods are

explored to calculate the ensemble expectations approximately.

In the following, we will turn to the topic of Markov chains. Let us consider an MRF with finite degrees of freedom first. The system microstates are distributed according to the Gibbs distribution. Our goal is to generate states from the desired Gibbs distribution. The MRF evolves in a discrete time evolution. Each time step, the system occupies a microstate. In a Markov chain, the transition from one state to another is probabilistic, where the row vector $\pi_\alpha(t)$ denote the probability distribution of being in the state $\{\alpha\}$ at time t. And the matrix $P_{\beta\alpha}$ denotes the probability that state α moves to β. Before studying the equilibrium behavior, we summarize some important properties of the transition matrix:

Time evolution. The probability vector at step $t+1$ is

$$\pi_\beta(t+1) = \sum_\alpha \pi_\alpha(t) P_{\alpha\beta} \tag{2.27}$$

Nonnegativity. The matrix elements are probabilities, so

$$0 \leq P_{\beta\alpha} \leq 1 \tag{2.28}$$

Normalization: The state α always change to another possible state β, so

$$\sum_\beta P_{\beta\alpha} = 1 \tag{2.29}$$

The transition matrix P has the following eigenvalues and eigenvectors:

$$\pi^\lambda \cdot P = \lambda \pi^\lambda \tag{2.30}$$

If a Markov chain[27] reaches equilibrium after a long time, the state vector will not be changed under the time evolution which means the equilibrium probability distribution vector has the eigenvalue 1. The state is called a recurrent state if the Markov chain returns

to it with probability 1, otherwise it is called a transient state, which the chain will never return. A Markov chain can also have periodicity, in which states can be partitioned into disjoint classes, and in such a loop that all the transitions from one class lead to the next one finally return to the original one such as $1 \to 2 \to 3 \to \cdots \to 1$. If all states of a Markov chain are accessible to each other in finite steps, the Markov chain is said to be irreducible. Ergodicity of a Markov chain means that the long time spent by the chain in one state corresponds to its steady-state probability, hereafter ensemble averages can be used as time averages.

Since the transition matrix has an eigenvale 1, it is easy to find that its eigenvector with all positive components must have eigenvalue $\lambda_0 = 1$ for an ergodic Markov chain[27].Therefore, an ergodic Markov chain has a unique time-independent equilibrium state π^*, that is the corresponding eigenvector multiplying a rescale factor.

In thermal equilibrium, the probability flux out from one state equals the probability flux into that state:

$$\pi^* \cdot P = \pi^* \tag{2.31}$$

In order to achieve equilibrium distribution, we need to ensure that the system is Markovian, ergodic, and also satisfies detailed balance.

The Metropolis algorithm was named after Nicholas Metropolis, who along with Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller, first proposed it for calculating the states from the canonical ensemble[27]. It can be shown that this popular method satisfies the detailed balance condition. Metropolis Algorithm step

are as follows:

(1) Pick a spin randomly;

(2) Compute the energy difference ΔH for flipping it;

(3) If $\Delta H < 0$ flip it; if $\Delta H > 0$ flip it with the probability $e^{-\beta \Delta H}$.

Near the Cure temperature T_c, the single-spin flip algorithm becomes quite slow known as the correlation time diverges. Wolff improved on the idea of Swendsen and Wang, camping up with a brilliant algorithm to flip the spin cluster each time. Wolff Cluster Algorithm[27]:

(1) Pick a spin randomly, record its direction, then flip it;

(2) For each of the four neighboring spins, if its direction is the same with the host, flip it with a probability p;

(3) For each new flipped spins, repeat the procedure (2).

Due to the finite probability for spin flipping, the Wolff algorithm is ergodic and Markovian. It also satisfies the detailed balance.

In statistics and statistical physics, Gibbs sampling generates a Markov chain with the Gibbs distribution as the equilibrium distribution such as the Metropolis algorithm when direct sampling is difficult. Non-stationary transition probabilities are generated with the Gibbs sampling method. The basic idea of Gibbs sampling is to construct a Markov chain by updating each variable subsequently given the state of all the others.

Consider a MRF, random vector $\mathbf{X}$ of lenth N. The probability of a particular configuration in the total sampling space is the joint probability distribution of $\mathbf{X}$. The time

evolution of an MRF constructs a chain of microstates. First, randomly choose a variable x_s, $s \in V$, $V = \{1, ..., N\}$ with a probability given by a positive probability distribution. Then, to flip it or not is based on its conditional probability given the states $(x_v)_{v \in V \setminus s}$ on the other nodes.

It is noteworthy that Gibbs sampling belongs to the class of Metropolis-Hastings algorithms. All MCMC algorithms of this class generate the transitions of a Markov chain in two steps. First, a candidate state is picked randomly from a Boltzmann distribution. Then, the candidate state is v̈oted̈o be the new state of the Markov chain with an acceptance probability which ensures that the detailed balance holds.

There often exists many local minima in the free energy landscape of the Gibbs distribution, separated by barriers. Conventiaonal Monte Carlo sampling are quite slow due to the suppression of tunneling through these barriers. One promising solutions so far is parallel tempering Monte Carlo[28]. This algorithm aims to overcome free-energy barriers by simultaneously simulating several replica of the target system each with a different temperature. The system can thus escape metastable states in higher temperature replica and return to lower temperature ones. Time complexity of it is several orders of magnitude smaller than a single fixed temperature Monte Carlo simulation.

Given an ordered set of M non-interacting copies of the system, let them run in parallel at different temperatures $\{T_1, T_2, ..., T_M\}$. After a fixed number of simulations, exchange of two copies at neighboring temperatures T_i and T_{i+1} are accepted with a transition prob-

ability:

$$P_{ij} = \min\{1, \exp[(\Delta_{i,i+1}H)(1/T_{i+1} - 1/T_i)]\} \tag{2.32}$$

where $\Delta_{i,i+1}H = H_{i+1} - H_i$ It should be noted that the update probability obeys detailed balance. The acceptance probabilities and the number of temperatures needed are related to the functional behaviors of the specific heat density.

The RBM (see fig 2.8) is the building block of deep Learning with a bipartite graph structure as shown in the figure above. It consists of visible nodes representing the observables, and hidden nodes which capture the feature between the observables. There are interaction terms only between these two group of nodes. In the binary RBMs, the random variables (V,H) take values $(v,h) \in \{0,1\}^{m+n}$ (m and n are the numbers of visiable and hidden nodes.) and the probability of configuraton is given by the Boltzmann distribution $p(\mathbf{v},\mathbf{h}) = e^{-\mathcal{H}(\mathbf{v},\mathbf{h})}$ with the energy functional:

$$\mathcal{H}(\mathbf{v},\mathbf{h}) = -\sum_{i=1}^{n}\sum_{j=1}^{m} w_{i,j}h_i v_j - \sum_{j=1}^{m} b_j v_j - \sum_{j=1}^{n} c_i h_i \tag{2.33}$$

For all $i \in \{1,...,n\}$ and $j \in \{1,...m\}$, w_{ij} is the interaction term between the units V_j and H_i and b_j and c_i are bias corresponding to visiable and hidden variables.

The graph of an RBM has only inter-layer connections, but not intra-layer interactions. In other words, this means that the variables are conditionally independent given the state of variables in the other layer:

$$p(\mathbf{h}|\mathbf{v}) = \prod_{i=1}^{n} p(h_i|\mathbf{v}) \quad \text{and} \quad p(\mathbf{v}|\mathbf{h}) = \prod_{i=1}^{n} p(v_i|\mathbf{h}) \tag{2.34}$$

We move to Gaussian RBMs since it is more useful, GRBMs can be viewed as a hybrid

multple Gaussian models with hidden nodes. The energy function of the GRBMs can be writen as

$$\mathscr{H}(\mathbf{v},\mathbf{h}) = \frac{1}{2} \sum_{j=1}^{m} \frac{(v_j - b_j)^2}{\sigma_i^2} - \sum_{i=1}^{n} \sum_{j=1}^{m} w_{i,j} h_i v_j - \sum_{j=1}^{n} c_i h_i \qquad (2.35)$$

The conditional independence properties makes Gibbs sampling especially easy. A so-called block Gibbs sampling can be performed in the following: sampling a new hidden field state based on $p(\mathbf{h}|\mathbf{v})$ and sampling a visible field state v based on $p(\mathbf{v}|\mathbf{h})$. Then, the marginal distribution of the visible variables becomes

$$p(\mathbf{v}) = \sum_{\mathbf{h}} p(\mathbf{v},\mathbf{h}) \qquad (2.36)$$

Hence,

$$p(\mathbf{v}) = \frac{1}{Z} \prod_{j=1}^{m} e^{\frac{(v_j - b_j)^2}{2\sigma_i^2}} \prod_{j=1}^{n} (1 + e^{c_i + \Sigma_{j=1}^{m} w_{ij} v_j}) \qquad (2.37)$$

The RBM can be regarded as a stochastic feed-forward neural network[26] with one layer of nonlinear processing units. The conditional probability of hidden and visiable units are

$$p(h_i = 1|\mathbf{v}) = \text{sigmoid}(\sum_{j=1}^{m} w_{ij} v_j \sigma_j^2 + c_i) \qquad (2.38)$$

and

$$p(v_j = v|\mathbf{h}) = \text{Norm}(\sigma_j^2 \sum_{i=1}^{n} w_{ij} h_i + b_j, \sigma_j^2) \qquad (2.39)$$

The log-likelihood gradient of an RBMs can be written as the sum of two expectations. The expectation of the energy gradient under the condition on the training sample

can be computed efficiently due to its factorization property:

$$\sum_{\mathbf{h}} p(\mathbf{h}|\mathbf{v}) \frac{\partial E(\mathbf{v},\mathbf{h})}{w_{ij}}$$

$$= \sum_{\mathbf{h}} p(\mathbf{h}|\mathbf{v}) h_i v_j$$

$$= \sum_{\mathbf{h}} \prod_{k=1}^{n} p(h_k|\mathbf{v}) h_i v_j$$

$$= \sum_{h_i} \sum_{\mathbf{h}_{-i}} p(h_i|\mathbf{v}) p(\mathbf{h}_{-i}|\mathbf{v}) h_i v_j$$

$$= \sum_{h_i} p(h_i|\mathbf{v}) h_i v_j \sum_{\mathbf{h}_{-i}} p(h_i|\mathbf{v}) p(\mathbf{h}_{-i}|\mathbf{v})$$

$$= p(h_i = 1|\mathbf{v}) v_j$$

$$= \mathrm{sigmoid}(\sum_{j=1}^{m} w_{ij} v_j + c_i) v_j$$

For the expectation of the energy gradient under the RBM dis tribution, the computation becomes intractable for regular sized RBMs because its complexity grows still exponentially. The derivative of the log-likelihood functionover the weight w_{ij} becomes

$$\frac{\partial \ln \mathscr{L}(\theta|\mathbf{v})}{\partial w_{ij}}$$

$$= \sum_{\mathbf{h}} p(\mathbf{h}|\mathbf{v}) \frac{\partial E(\mathbf{v},\mathbf{h})}{w_{ij}} + \sum_{vh} p(\mathbf{h},\mathbf{v}) \frac{\partial E(\mathbf{v},\mathbf{h})}{w_{ij}}$$

$$= \sum_{\mathbf{h}} p(\mathbf{h}|\mathbf{v}) h_i v_j - \sum_{\mathbf{v}} p(\mathbf{v}) \sum_{\mathbf{h}} p(\mathbf{h}|\mathbf{v}) h_i v_j$$

$$= p(h_i = 1|\mathbf{v}) v_j - \sum_{\mathbf{v}} p(\mathbf{v}) p(h_i = 1|\mathbf{v}) v_j$$

For the mean value of this derivative over a training data set, the following formulas are obtained:

$$\Delta w_{ij} = < h_i v_j >_{\text{data}} - < h_i v_j >_{\text{model}} \tag{2.40}$$

33

Similarly, we obtain the derivatives:

$$\Delta b_j = <v_j>_{\text{data}} - <v_j>_{\text{model}} \tag{2.41}$$

and

$$\Delta c_i = <h_i>_{\text{data}} - <h_i>_{\text{model}} \tag{2.42}$$

To avoid the exponential complexity of summing over all values of the visible variables when calculatingthe log-likelihood gradient one can approximate this expectation with a Gibbs sampler. However, this requires stll need to run the Markov chain long enough to ensure equillibrium.

Common remote sensing scheme take multiple images of the same area with multi view angles. Based on the photon transport, we propose a network for atmospheric quatities.

$$\mathcal{H}(\mathbf{v}, \mathbf{s}, \mathbf{h}) = \frac{1}{2} \sum_{j=1} \frac{(v_j - f(\mathbf{s}))^2}{\sigma_{v_i}^2} + \frac{1}{2} \sum_{j=1} \frac{(s_j - a_j)^2}{\sigma_{s_i}^2} - \sum_{i=1} \sum_{j=1} w_{i,j} h_i s_j - \sum_{j=1} c_i h_i \tag{2.43}$$

$\mathbf{s}$ is the geophysical state vector (e.g. clould optical depth).

Given the radiance measurement, we would like to infer the posterior distribution over the latent variables. Based on a preview model[30], we propose a two step inference method.

Step 1: $p(\mathbf{h}|\mathbf{s}, \mathbf{v})$

In this step, we sample the hidden vector conditional on the state vector and visiable vector. Since this is a GRBMs with an additional term, it is easy to implement the sampling procedure.

Step 2: $p(\mathbf{s}|\mathbf{h}, \mathbf{v})$

Sampling the state vector is a little bit curbsome, due to the comlicated dependence forward model and visiable radiance. Linearized kernel can be used to simplify the computationa.

Learning is the most challenging and important part. We may use EM algorithm to solve this problem. First we sample $(\mathbf{s}, \mathbf{h})$ according to the conditional distribution described, and then optimize the log-likelihood function.

$$\Delta\theta = -r\mathbb{E}[\frac{\partial}{\partial\theta}\mathscr{H}(\mathbf{v}, \mathbf{s}, \mathbf{h})] = -r\int p(\mathbf{a}, \mathbf{h}|\mathbf{v}; \theta^{-})\frac{\partial}{\partial\theta}\mathscr{H}(\mathbf{v}, \mathbf{s}, \mathbf{h})d\mathbf{a}d\mathbf{h}$$

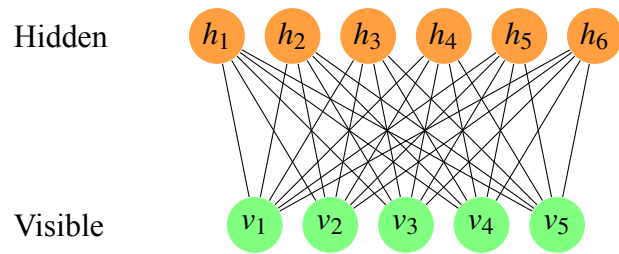r is the learning rate. MCMC samples can be explored to approximate above integrad as a sum.

Figure 2.8: Restricted Boltzmann topology with 5 visible units and 6 hidden units.

3   SURFACE GROWTH MODEL AND SCATTERING

In this section [1], a random field model is explored to model the shapes of atmospheric particles. A random field $h(\mathbf{x})$ on $\mathbb{R}^n$ is a function whose value are random variables for $\mathbf{x}$. Many natured phenomenare random fields such as cosmic background radiation[47, 35] and optical speckle field patterns[47]. In the context of surface roughness, random fields are extendsively used for unveiling the physical meaning and topological properties of a roughened surface. However, to our understanding, random fields have not yet been exploited for modeling the surface roughness of ice crystals[33, 34], although nearly all previous work can been considered as subclass of Gaussian random fields(GRF) in some kind. Due to geometric simplicity and well understood physical meaning, our modeling uses GRFs.

A GRF $h(\mathbf{x})$ (height function of the surface) arises from the random superposition of waves:

$$h(\mathbf{x}) = \sum_{\mathbf{k}} A(k) \cos(\mathbf{k} \cdot \mathbf{x} + \phi_{\mathbf{k}}) \tag{3.1}$$

where $A(k)$ is an amplitude spectrum that depends only on the magnitude of the wave vectors $\mathbf{x}$. The uncorrelated random phases $\phi_{\mathbf{k}}$ are uniformly distributed in the range $[0, 2\pi]$. The power spectrum is $P(k) = A(k)^2$, containing the two-point correlation of random fields. The statistical properties of $h(\mathbf{x})$ are entirely encoded by the power function $P(k)$ and the

moments generated from it.

Now, let us consider nonequilibrium dynamics of surfaces[10]. Expanding the surface area in terms of the slope, we obtain its energy density:

$$\mathscr{H} \approx \frac{v}{2}(\nabla h)^2$$

where h is the height function of the surface and v represents "surface tension". A potential energy function can also be added the function, specifiedly the gravitational potential energy of fluid[32].

$$\mathscr{H} = \frac{1}{2}[v(\nabla h)^2 + aV(h)] \tag{3.2}$$

The corresponding linear Langevin equation is

$$\frac{\partial}{\partial t}h(\mathbf{x},t) = v\nabla^2 h - ah + \eta(\mathbf{x},t) \tag{3.3}$$

where η is the radom force, and this equation can be solved by using Fourier transforms. The solution at time t is

$$h(\mathbf{k},t) = h(\mathbf{k},0)e^{-(vk^2+a)t} + \int_0^t d\tau e^{-(vk^2+a)(t-\tau)}\eta(\mathbf{k},\tau)$$

If starting with a flat surface, the average of surface height is zero, and the height variance grows as

$$\langle h(\mathbf{k},t)^* h(\mathbf{k}',t)\rangle = (2\pi)^2\delta^2(\mathbf{k}-\mathbf{k}')(1-e^{-2(vk^2+a)t})\frac{T}{vk^2+a} \tag{3.4}$$

After transforming to the coordinate space, we obtain

$$
\begin{aligned}
\langle h(\mathbf{0},t)h(\mathbf{x},t)\rangle &= \int d^2\mathbf{k}\int d^2\mathbf{k}' e^{-i\mathbf{k}'\mathbf{x}}\delta^2(\mathbf{k}-\mathbf{k}')(1-e^{-2(vk^2+a)t})\frac{T}{vk^2+a}\\
&= \int d^2\mathbf{k}\,e^{-i\mathbf{k}\mathbf{x}}(1-e^{-2(vk^2+a)t})\frac{T}{vk^2+a}\\
&= T\int\int e^{-ikr\cos\theta}\frac{1-e^{-2(vk^2+a)t}}{vk^2+a}kd\theta dk\\
&= 2\pi T\int\frac{J_0(kr)(1-e^{-2(vk^2+a)t})}{vk^2+a}kdk
\end{aligned}
$$

Let $\Lambda = 1/l_0$ be the cutoff. When $a = 0$, we get

$$
\langle h(\mathbf{x},t)h(\mathbf{x},t)\rangle \propto \int_0^\Lambda (1-e^{-2vk^2t})\frac{1}{k}dk \tag{3.5}
$$

$$
= \ln[1+2vt/l_0^2] \tag{3.6}
$$

where we have a correlation length $l_c = \sqrt{2vt}$.

Now, considering the equilibrium case,

$$
h(\mathbf{k},\omega) = \chi(\mathbf{k},\omega)\eta(\mathbf{k},\omega)
$$

where $\chi = \frac{1}{-i\omega+vk^2+a}$ is the suseptibility. According to the fluctuation-dissipation theorem[27], we obtain

$$
C(\mathbf{k},\omega) = \frac{2T}{\omega}\Im\chi(\mathbf{k},\omega)
$$

3.1 Particle Modeling: Surface Growth Approach

Particle shape irregularity is an important factors that determine the single-scattering properties of atmospheric particles such as ice crystals in cirrus clouds [2, 36, 37, 38, 39]. It is necessary to improve the current level of knowledge about particle surface roughness to better understand the microphysical process of atmospheric particle formation and

39

evolution [33, 34, 40]. In the literature, the effect of particle overall shape on the optical properties draws much more attention than the counterpart associated with particle surface texture, i.e., the degree of surface roughness [41, 50, 61, 64, 65, 66] although there is an increasing awareness of the importance of particle surface roughness [33].

Many particle surfaces in nature are rough to some degree. The dynamics of ice particle surfaces involve a lot of complicated physical processes such as deposition, evaporation, coalescence and collision [32]. For large ice particles, modeling of surface microphysics has attained considerable success [46]. Various ice particle habits such as bullets, columns, plates, and aggregates have been observed and studied. Recently, some experiments were successfully conducted in generating and imaging ice particle morphology under different conditions, and it is found the degree of roughness structure depends on ambient temperature, supersaturation and some other factors. The classic method for studying surface dynamics of roughness growth is based on phenomenological stochastic partial differential equations. The solutions have revealed various interesting linear and nonlinear behaviors. However, to the best of the authors' knowledge, modeling roughness morphology is limited to a plane or a sphere [47, 67, 68]. In [50], surface roughness was generated through a Gaussian random field over 2D plane, then 2-D roughened patches were glued together to form a roughened hexagonal prism. This method seems simple and straightforward, but mismatch and overlap between surfaces may occur along the edges. In [49], the so-called pseudo-random surface roughness is proposed to model a spherical roughened particle. Radical Gaussian random perturbation is imposed on the points on

a spherical surface. The problem with this model is that there is no correlation between points on the roughened surface, leading to 'white noise', which substantially deviates from a natural particle surface. In addition, the roughness generated by these models is not related to the surface physics of natural particles, and the resultant surface roughness may not be physically rational. To overcome these drawbacks, we attempt to generate surface roughness on a general geometry in a uniform manner.

Because the behavior of roughness development is described by continuous stochastic partial differential equations, it is difficult to solve on object with arbitrary geometry. From computational perspective, we must express the underlying equations in discrete space and numerically solve them. Discrete Differential Geometry (DDG) theory [52] retains key geometric properties of the continuous counterpart and has been innovatively applied to geometry modeling such as surface parameterization and remeshing. In this study, we apply this method to solve the stochastic partial differential equations on arbitrary 3D geometric objects and develop a stochastic parameterization model for the roughened ice particles. It should be pointed out that the present method is also applicable to modeling the morphology of other atmospheric particles such as highly irregular dust particles.

For the present light scattering simulations, we use the invariant imbedding T-Matrix method (II-TM) computational program developed by Bi et al. [53], which is applicable to particles with arbitrary shapes. Note, II-TM was first introduced by Johnson [54], for axially symmetric particles, who derived the T-Matrix using a volume integral equation formulation and applied the invariant imbedding technique to accelerate the computation. The

T-matrix assumes linearity of the system (linear permittivity/permeability and Maxwell's equations). It can also be derived using other frameworks such as the boundary element method or the discrete dipole approximation method [55].

3.1.1 Surface Growth

Various growth mechanisms have been proposed for modeling surface morphologies [32]. In this section, we briefly introduce a common growth model and extend it to model ice particles in the atmosphere. To begin with, assume molecules evaporate or condense randomly on a surface and the growth direction is along the local surface normal. Within this framework, the height function $h(\mathbf{x},t)$ of the surface (relative to the original shape as reference) as it grows is governed by the famous Kardar-Parisi-Zhang (KPZ) [57] equation:

$$\frac{\partial h}{\partial t} = \nu\nabla^2 h + \frac{\lambda}{2}(\nabla h)^2 + \eta \tag{3.7}$$

It should be noted that original KPZ equation is derived on a 2D plane. Here we have made an assumption that it is appropriate for 3D surfaces. The random evaporation/condensation of molecules is represented by Gaussian random variable $\eta(\mathbf{x},t)$ with

$$\mathbb{E}[\eta(\mathbf{x},t)] = 0 \tag{3.8}$$

$$\mathbb{E}[\eta(\mathbf{x},t)\eta(\mathbf{x}',t')] = 2\sigma^2\delta(\mathbf{x}-\mathbf{x}')\delta(t-t') \tag{3.9}$$

Here, σ is its standard deviation. The linear term $\nu\nabla^2 h$ ($\nu > 0$) serves to smooth out the surface. It describes the thermal equilibrium between phases(water vapor and ice). The nonlinear term $(\lambda/2)(\nabla h)^2$ is from the geometric origin [57]. It renders the equation quite nontrivial, however a detailed discussion will be out of our scope.

If $\lambda = 0$, the growth is dominated by the evaporation/condensation processes. The nonlinear equation becomes the linear Edwards-Wilkinson (EW)[58] equation:

$$\frac{\partial h}{\partial t} = v\nabla^2 h + \eta \tag{3.10}$$

Geometrically, the equation describes that random deposition η is flattened out by the term $\nabla^2 h$ and attains a dynamic equilibrium over time (Fig. 3.1). For simplicity, we restrict our consideration to the EW model. Since the atmospheric particles are suspended in the air, the random deposition or evaporation of molecules occurs on the surface of the ice particles. The ice particle surface grows along the surface normals, smoothed out by the diffusion process. Through this process, the singularities around the edges/corners of faceted particles are gradually eliminated. Recent experiments [34, 40] seem to support this model.

3.1.2   Discretization of Stochastic Partial Differential Equation

For a smooth surface in 3D, it is easy to define the normal direction--the unique direction orthogonal to all tangent vectors. But there are no unique normals for points residing at corners or along edges. Moreover, it is computationally challenging to determine the configuration of a continuous random surface even with modern computers. To solve these difficulties, we discretize the continuous stochastic surfaces with a triangular mesh. Then, to solve the stochastic partial differential equations on a discrete surface, we must seek an appropriate way to define the differential operator and the surface normal at each vertex on the mesh.

To begin with, multiplying $\nabla^2 h$ from the left with a test function ϕ, we then transform

43

to the weak formulation $\langle \nabla \phi, \nabla h \rangle$.

Following the idea of the finite element method (FEM), we seek an approximation $\tilde{h}$ of the true solution h in a finite dimensional reduced space:

$$\tilde{h} = \sum_j h_j \phi_j(\mathbf{x}) \tag{3.11}$$

where $\{\phi_i\}$ are basis functions (linear hat functions in our case)and minimize $||\tilde{h} - h||$. Substituting equaton above into the weak formulation, we get

$$\langle \nabla \phi_i, \nabla \tilde{h} \rangle = \sum_j h_j \langle \nabla \phi_i, \nabla \phi_j \rangle \tag{3.12}$$

where

$$\langle \nabla \phi_i, \nabla \phi_j \rangle = \int_\Omega dx \nabla \phi_i(x) \cdot \nabla \phi_j(x) \tag{3.13}$$

and Ω is the domain for integration. The self quadrature of ϕ_k [52],

$$\langle \nabla \phi_k, \nabla \phi_k \rangle = \frac{1}{2}(\cot \alpha + cot \beta) \tag{3.14}$$

where α and β are the interior angles at the remaining two vertices. For mutual quadrature of hat functions associated with two vertices on the same triangle, we obtain

$$\langle \nabla \phi_i, \nabla \phi_j \rangle = -\frac{1}{2} \cot \gamma \tag{3.15}$$

where γ is the angle opposite to the edge with vertices i and j in Fig. 3.2.

Finally, the discrete Laplacian of the height function at each vertex i is given by the Cotan formula[52, 32]:

$$(\nabla^2 h)_i = \frac{1}{2} \sum_j (\cot \alpha_j + \cot \beta_j)(h_j - h_i) \tag{3.16}$$

where α_j and β_j are angles across from the same edge. For time integration, the forward Euler scheme is frequently used for approximating the time derivative of height function.
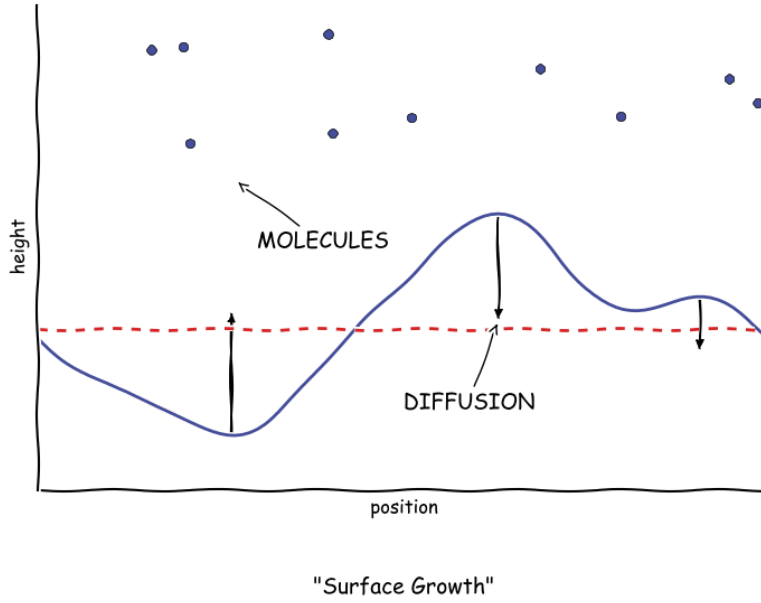
"Surface Growth"

Figure 3.1: Physics of surface growth.

The whole system becomes

$$h_{i+1} = h_i + v \delta t (\nabla^2 h)_i + \eta_i \tag{3.17}$$

where δt is the time step for the discretized EW equation and η_i is a Gaussian random variable at that vertex. The correlation length is estimated to be $\sqrt{2nv\delta t}$, n is the number of time steps we could set mannully.

In this section, we have applied the stochastic partial differential equation approach to model random particle morphology. Differential operators are linked with their geometric origin. Via this approach, a homogeneous roughness can be defined on arbitrary geometric objects as in Fig. 3.4.

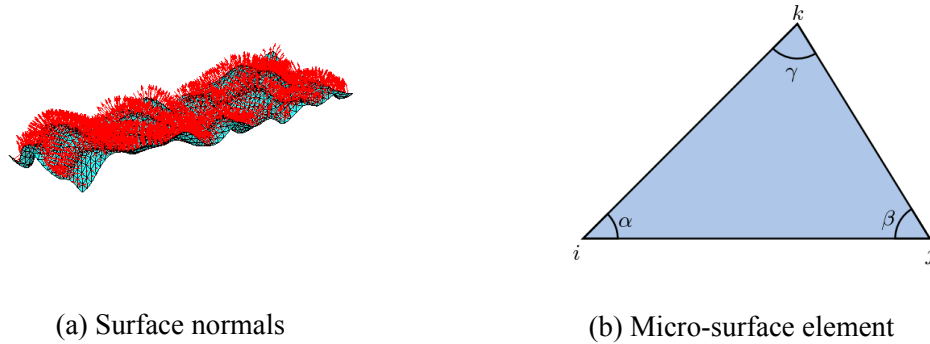(a) Surface normals   (b) Micro-surface element

Figure 3.2: Triangulation and surface normals.

3.1.3   Height and Slope Statistics

In above sections, we have described random rough particles by the local surface height deviation from smooth particle surfaces. To unveil the connection between our model and previous models, it is necessary to given a quantitative analysis of the statistical characteristics of the model particle associated with the definitions of its parameters. Furthermore, in many applications roughness is classified by the surface height or slope distribution. As an example, we now consider the height and slope statistics of a 2D plane generated by our approach.

It is noted that the height of an EW surface obeys a Gaussian distribution ($h \sim Norm(0, w^2)$, where w is the standard deviation of the surface height). Similarly, it is straightforward to demonstrate that the distribution of local slope is also normal [32]. Once distributions are known, we could generate rough surfaces in either deterministic (our model) or stochastic sense [2]. In stochastic models, a deterministic rough surface is replaced by a simplified random surface with a modified scattering distribution function
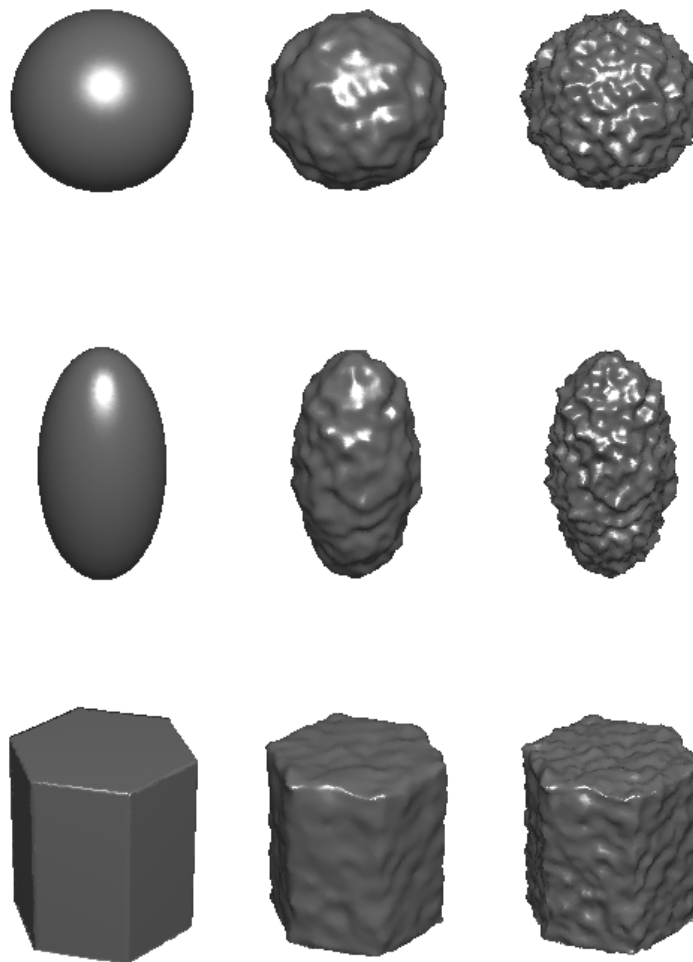
46

Figure 3.3: Particles with standard deviations (from left to right) 0.00, 0.01, 0.02
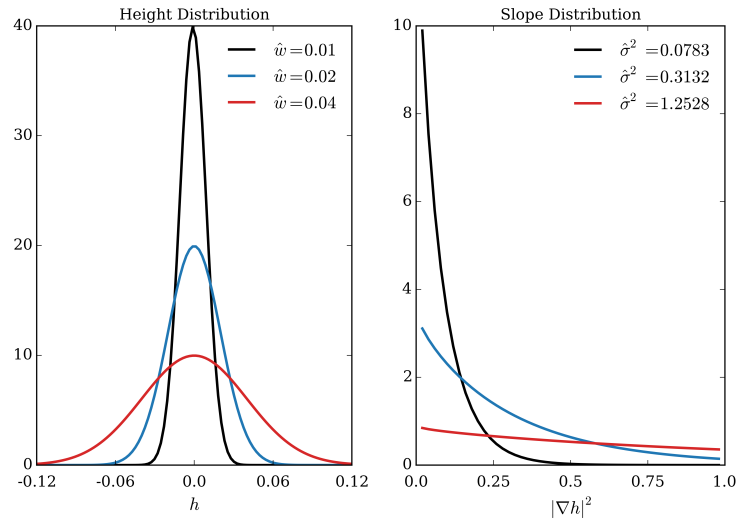
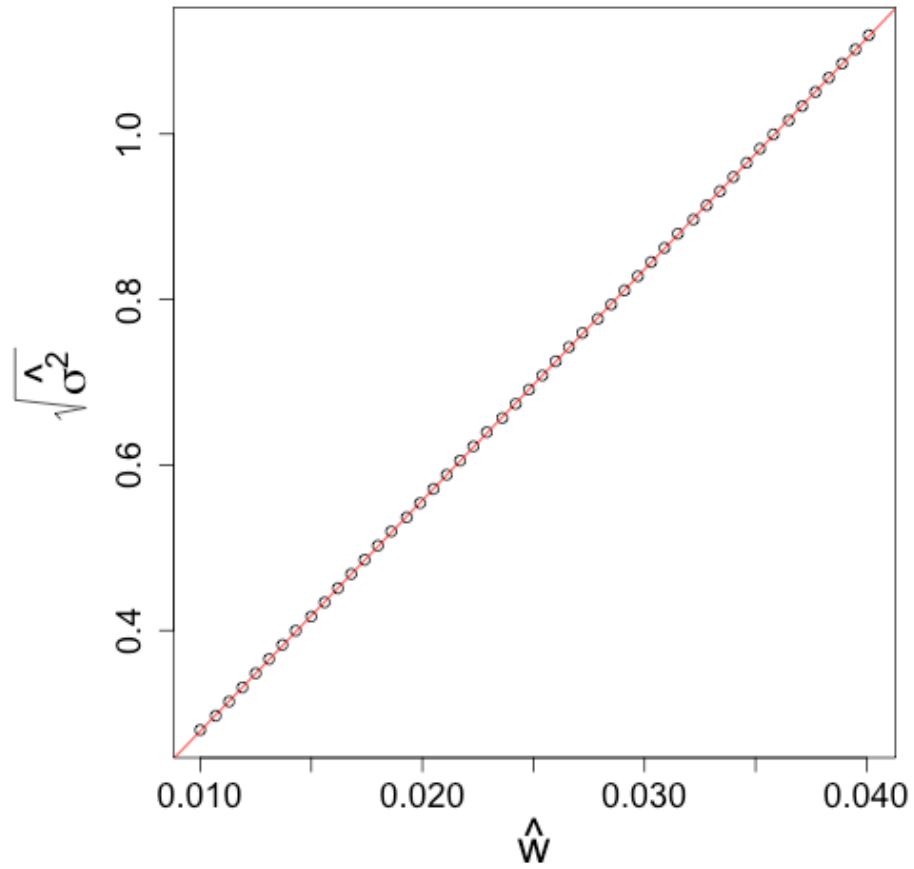Figure 3.4: Estimated height and slope distribution from a roughened plane

Figure 3.5: Plot of $\sqrt{\hat{\sigma^2}}$ against $\hat{w}$.

48

that matches the reflection and refraction of the surface. For example , Yang and Liou [2]

introduce a Tilted-Facet (TF) model by exploring a 2-D Gaussian distribution to investigate

the surface roughness of the ice crystal:

$$P(\nabla h) = \frac{1}{\pi \sigma^2} \exp(-|\nabla h|^2 / \sigma^2) \qquad (3.18)$$

where the slope $\nabla h = (\frac{\partial h}{\partial u}, \frac{\partial h}{\partial v})$, and u, v are orthonormal coordinates in the local frame. It

is obvious $|\nabla h|^2 \sim Weibull(a = \sigma^2, b = 1)$. The assumption underlying TF model is that

particle surface detail is not important, and only the far-field scattering pattern matters. To

simplify the problem, rays that strikes the surface twice are also ignored. A comparison

between these two methods is very interesting, and Liu et. al. make a good attempt in [50].

However, as we have noted above, their model suffers severe chanllage near edges. It also

fails to give a concrete test about the surface statistics. All of these make this model highly

problematic. Now we would like to remedy these problems with our surface growth model

and develop standard statistics tests.

To begin with, the time of simulation is fixed, so the correlation length is fixed as

well. Since the width of the surface height and the variance of the surface slope are related

with each other through the correlation length as illustrated in [32], we test the response of

slope to the change of surface height. To simplify the problem, $|\nabla h|^2$ is tested instead of the

2 slope components. Fig. 3.4 shows the height and slope distribution with the estimated

parameters. And Fig. 3.5 illustrates the scatter plot of $\hat{\sigma^2}$ agaist $\hat{w}$, and the regression line.

The confidence intervals corresponding to each parameter are given as well. The

95% confidence intervals for average height are [-0.0001, 0.000], [-0.0001, 0.0001], and

[-0.0003, 0.001] respectively. The 95% confidence intervals for $\hat{w}$ are [0.0100, 0.0100], [0.0190, 0.0210], and [0.0390, 0.0420] respectively. The 95% confidence interval of b is [0.9484, 1.0042] for all Weibull distributions. And the 95%confidence intervals for $\hat{\sigma}^2$ are [0.0752, 0.0815], [0.3008, 0.3261], and [1.2032, 1.3044] respectively.

Due to linearity of the system (linear permittivity/permeability and Maxwell equations), a good approximation for most atmospheric optical phenomena, the T-matrix ansatz is valid. Expanding the incident fields in terms of regular vector spherical functions, we obtained the T-matrix from the $\mathbb{T}$ operator[59, 11]. IITM discretizes the scatterer into spherical shells and computes the T-matrix of the shells in a recursive manner.

3.2   Numerical Simulation and Results

In simulations, we model the ice particles discretely using triangular mesh with up to 10000 triangular elements. 3 typical model ice particles(sphere, spheroid and hexagon) are chosen for comparison with their roughened counterparts. All numerical simulations are done with IITM software, which is written in standard Fortran90 and uses both OpenMP and MPI standards for communications, allowing it to run in a parallel manner. Twenty cores are used for each size parameter and cpu time goes up to 40 hrs due to the symmetry breaking. A representative refractive index of 1.3 is used for simulating scattering of visible light.

Figs. 3.6-3.8 show the density plots of the reduced Mueller matrix(P_{11}, P_{12}/P_{11}, and P_{22}/P_{11}) for the EW particles(spheres, spheroids, and hexagonal prisms) with refractive
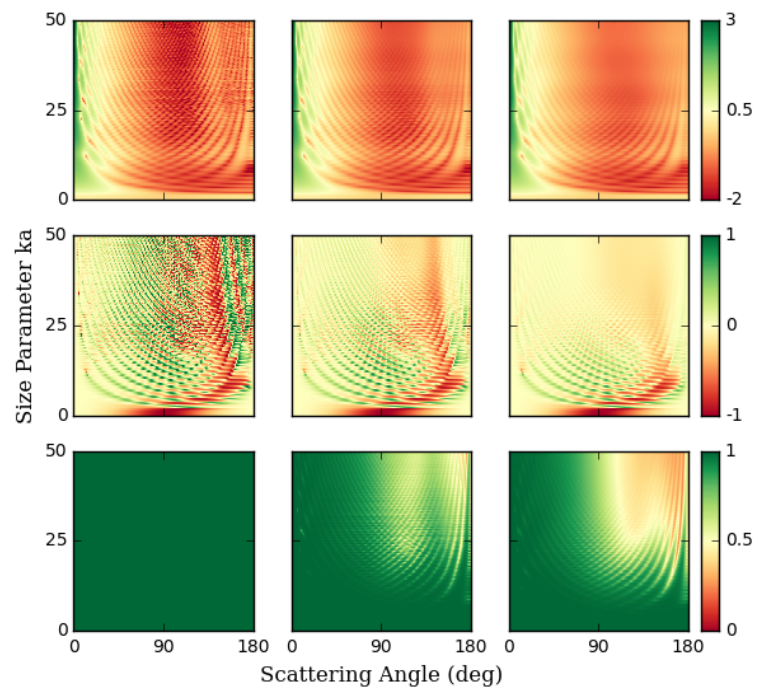
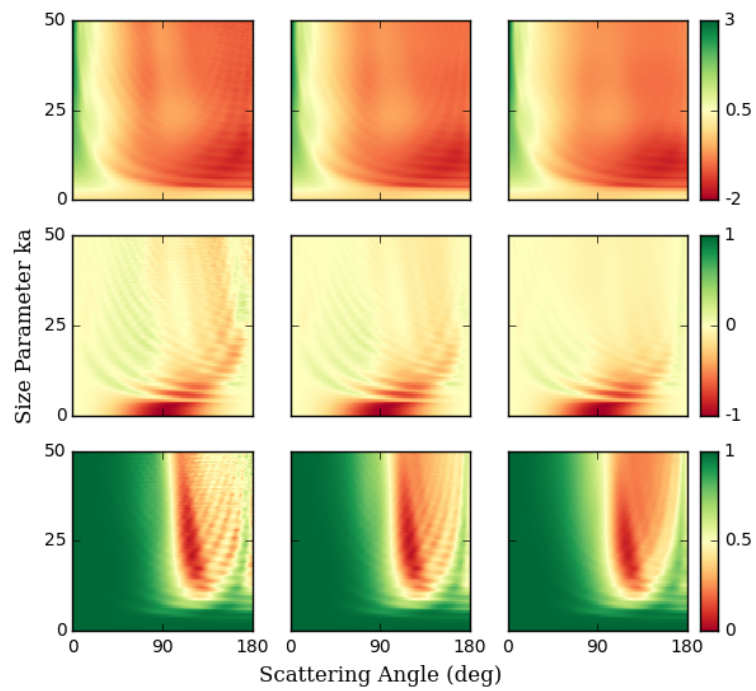Figure 3.6: Mueller matrix ensembles for roughened spheres

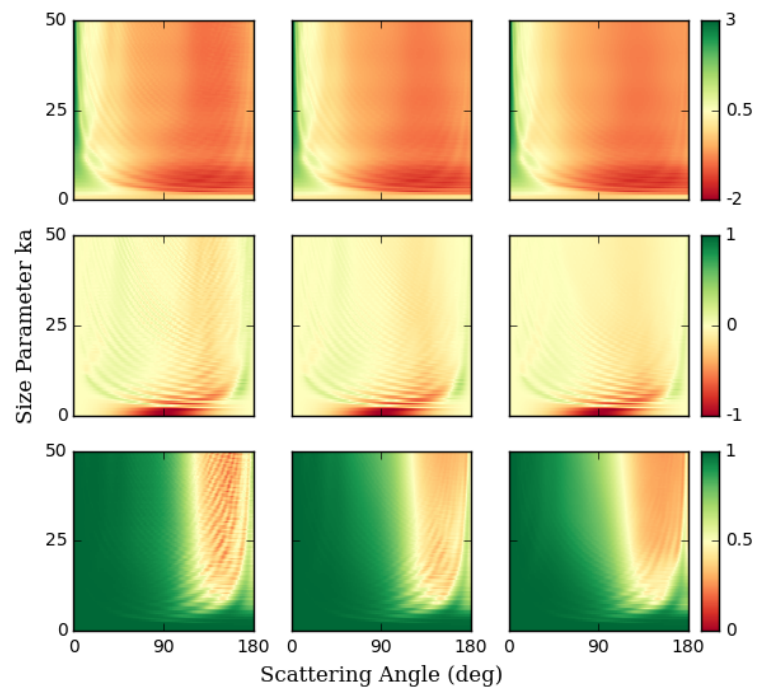Figure 3.7: Mueller matrix ensembles for spheroids.

Figure 3.8: Mueller matrix ensembles for hexagonal prisms.

index 1.3. All the elements are shown as functions of scattering angle and the size parameter ka, where a is the parameter of each smooth particle. Adaptive integration step sizes are used to resolve the roughened surfaces of the particles more accurately. Overall, as we increase the standard deviation σ, the phase function becomes smoother.

For EW spheres, the diagrams of P_{12}/P_{11} are full of obvious local extrema due to the optical interference [60]. The increase of σ results in smoother patterns, especially for larger sizes. The explanation for smoother pattern comes from the smoothing effect of statistical orientation averaging. While the size dependence can be explained with the partial wave analysis, since surface height at larger particle size are larger and trigger out more wave components. The roughness effects on P_{22}/P_{11} of EW spheres are more profound, and values are dramatically decreased from 1 for backward directions. Apparently with the growth of roughness, the particle show more pronounce non-spherical features. It should be noted that wrinkles in the diagrams becomes more evident as the roughness increases. The profile of the particle remains due to the uniformly generated roughness, so we attribute above the pattern feature to the edge effect. For EW spheroids and prisms, the interference pattern is much less distinct, this can be also traced back to the statistical orientation averaging of non-spherical particles. The plots for P_{22}/P_{11} of EW spheroids illustrate the global sink near 100 deg, while for EW hexagonal prisms the steepdecent direction points to the right corner.

In addition, these figures seem to indicate that the roughness degree should be defined with respect to the wavelength instead of the particle size. They confirm that small

scale roughness ($\ll$ wave length) has little effect. However, as the roughness grows and approaches wavelength, the phase matrix images are blurred. The surface roughness randomizes the surface currents and smoothes out the face field behavior of the scattered wave. This has been proved in many theoretical and experimental scattering studies for a 2D roughened surface. Here, similar conclusions are seen for 3D particles. Hence, we divide the roughness into the same three categories $\ll$, $\sim$ and $\gg$ wavelength as with 2D surface scattering. In general, small scale roughness has little effect, large scale roughness in which local principle remains true, geometric optics approximation could be applied. For medium roughness, full wave method should be employed to study its impact.

Figs. 3.9-3.11 compare a single size parameter (ka = 50) of each model particle. Roughness smooths out the high frequency oscillations and decreases the polarization and depolarization. For an EW sphere, the roughness results in a side scattering enhancement for P_{11}, and decrease P_{22}/P_{11} dramatically. For EW spheroids, roughness illustrates anomalous backscattering enhancement. Generally, it increases P_{12}/P_{11} and P_{22}/P_{11} in the forward directions and reduces them in the backward directions. But for other elements($P_{33}/P_{11}, P_{34}/P_{11}$, and P_{44}/P_{11}), the roughness smooths the curves and increases them gradually. For EW hexagons, roughness reduces the 22 deg and 46 deg halos, as observed in many other works. It also smooths out the linear polarization and decreases the depolarization over the scattering angles.Other elements show similar response to roughness except P_{44}/P_{11}, which indicates a reverse in the trend of the curves. However, the sensitivity of roughness are not related to their relative value of variance versus particle
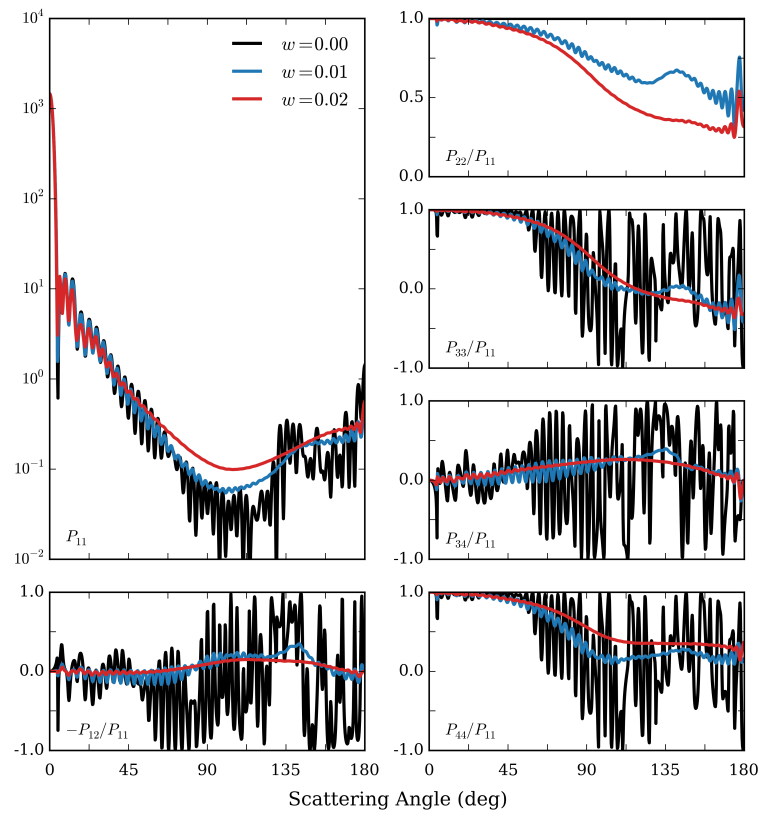
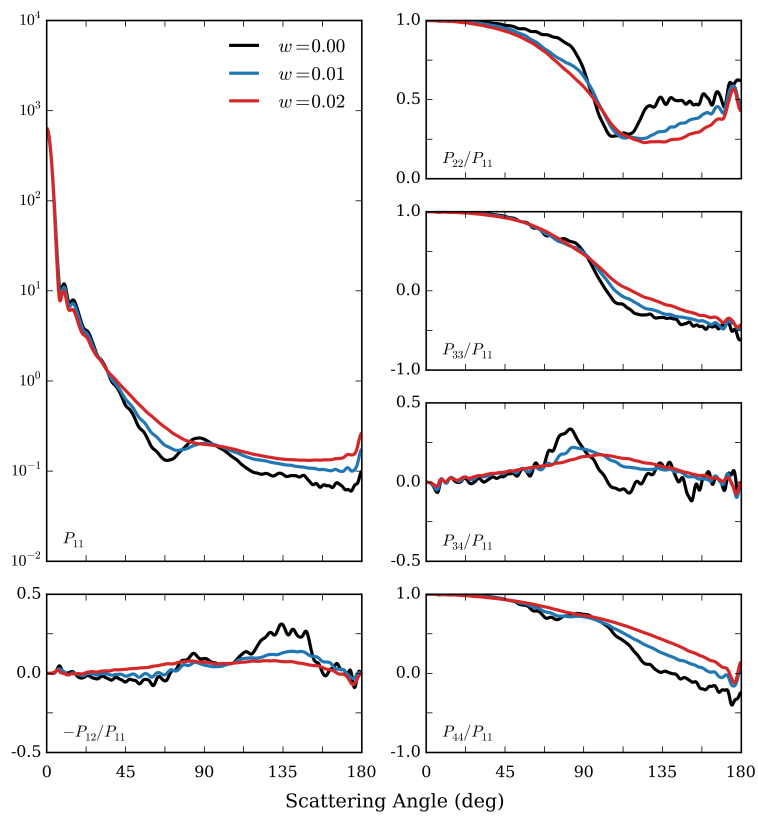Figure 3.9: Mueller matrix computed by II-TM for EW spheres

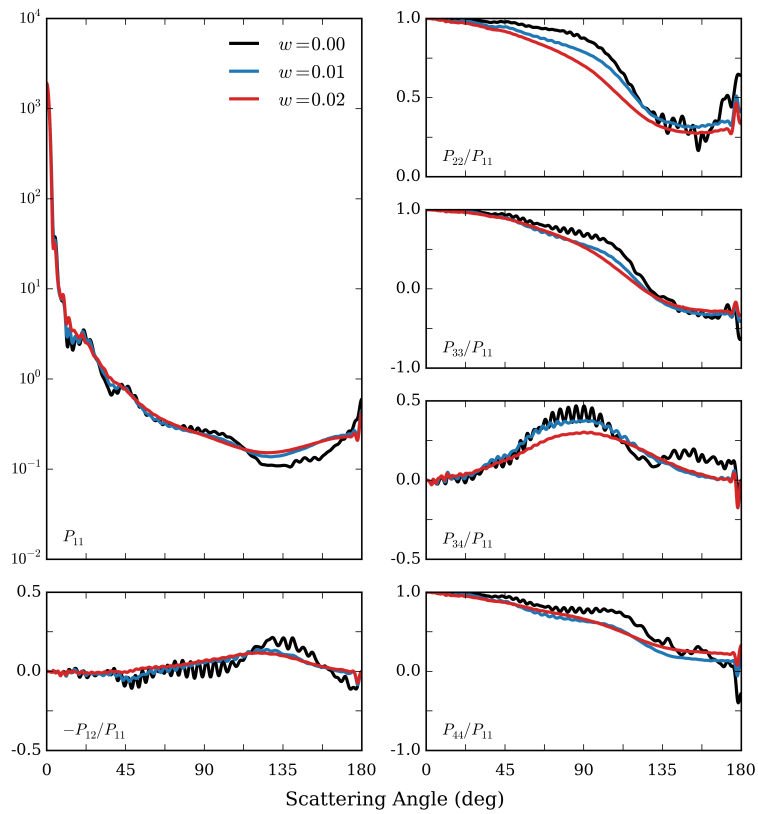Figure 3.10: Mueller matrix computed by II-TM for EW spheroids.

Figure 3.11: Mueller matrix computed by II-TM for EW hexagonal prisms

radius but is related to the wavelength of incident waves. For a smooth ice crystal, the angular distribution of P_{11} indicates symmetric features of ice crystals. However, such symmetric features of P_{11} are averaged out as the height of the surface increases.

Figs. 3.12-3.14 show the extinction efficiency and asymmetry factor for these model particles as functions of size parameter ka. To compute reduced scattering cross sections involving spheres, spheroids and hexagons, we set the volume equivalent sphere as the base geometric cross section. For example, for Gaussian spheres, the radius of a smooth sphere is used as the reference. Bottom curves show the scattering cross section for the roughened particles. The figures show roughness as a perturbation that results in distinguishable impact only for height that is comparable with wavelength. We see that three curves approach each other for small sizes. In this case, radiation from the two resonances add up incoherently, and the reduced cross section is a sum of two shape functions.
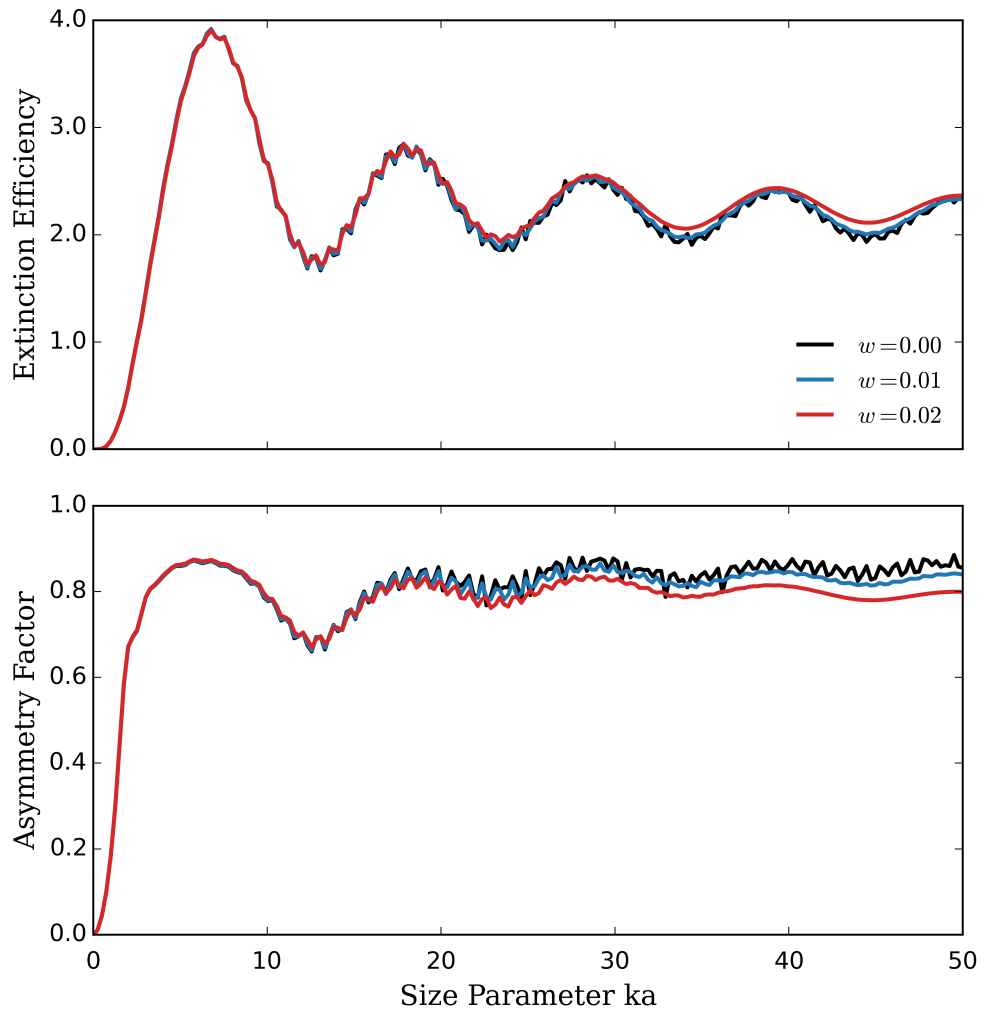
Figure 3.12: Extinction efficiency and asymmetry factor for EW spheres.
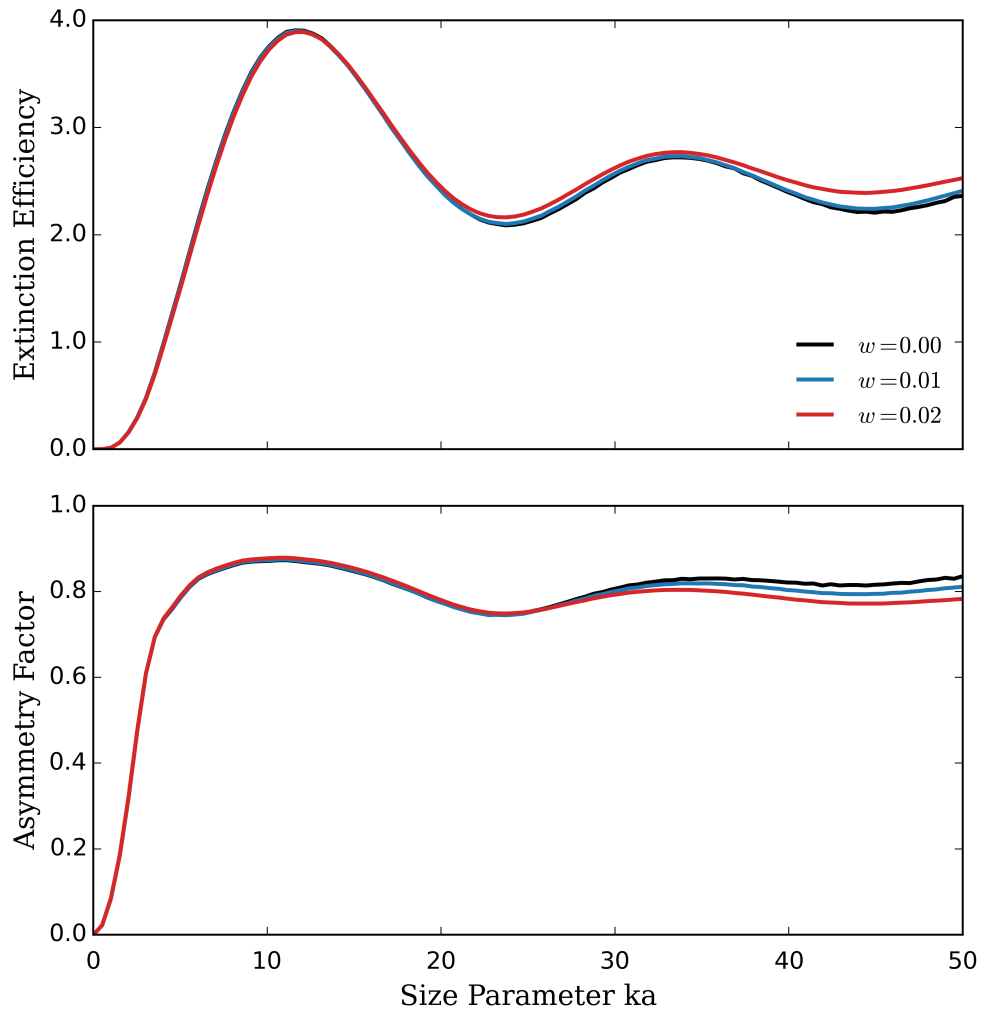
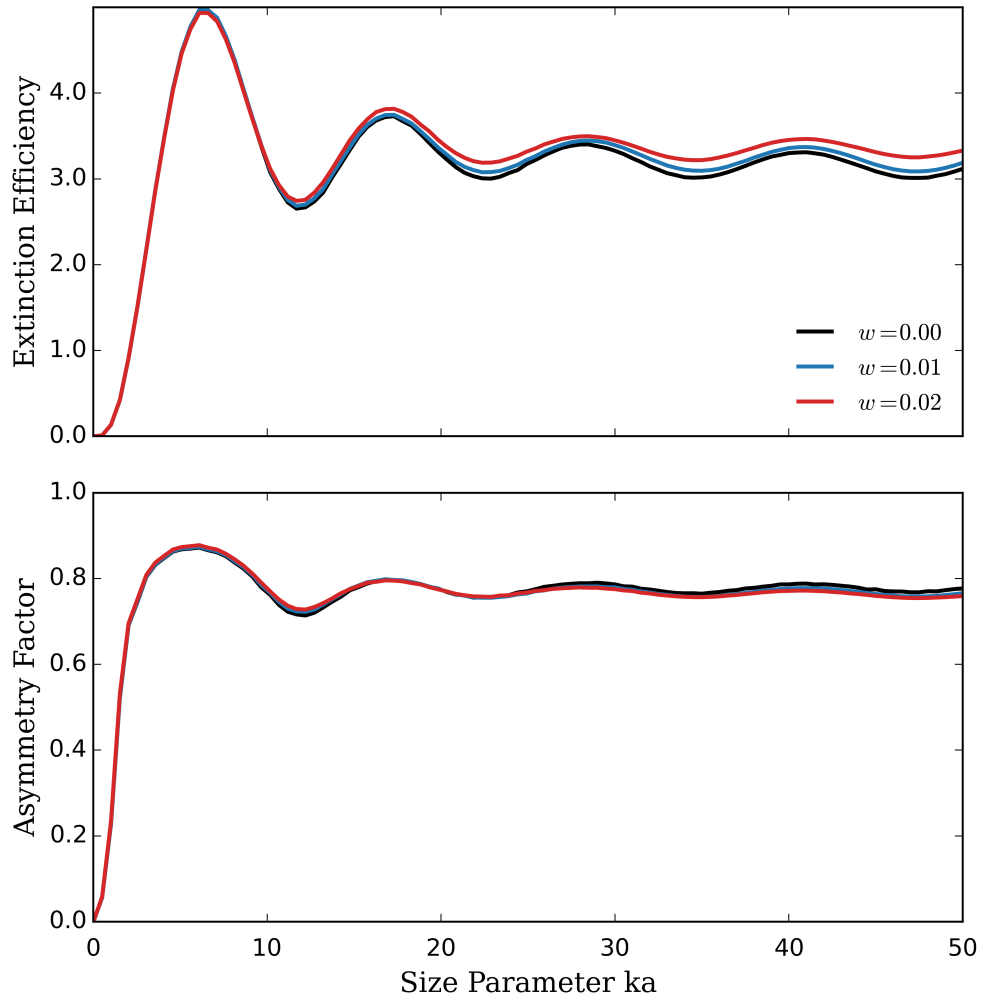Figure 3.13: Extinction efficiency and asymmetry factor for EW prolates.

Figure 3.14: Extinction efficiency and asymmetry factor for EW hexagon prisms.

4 MONTE CARLO RADIATIVE TRANSFER IN DUST MEDIUM

Aerosols are not only essential for cloud formation but also modify the climate by absorbing solar radiation. Mineral dust originating from the suspension of earth minerals is one of the most frequently observed atmospheric aerosol. The equilibrium between lofting and setting of dust aerosols can be attributed to the influence of both anthropogenic activities and deserts. The spatial and temporal fluctuations of mineral dust concentrations have a notable impact on the regional and global radiation equilibrium. The main difficulties in monitoring dust aerosols lie in the uncertainty of its refractive index and effective modeling of its morphology[72]. An effective model particle will extend our understanding of dust size and shape distribution and lead to more accurate retrieval in remote sensing[63].

A spheroidal model particle seems to be a superior candidate as the standard model particles for dust aerosols[73]. Simulations and measurements also indicate that spheroid simulates aerosol optical properties more effective than other existed non-spherical model particles such cube, Gaussian random sphere [72, 65, 64, 61, 66]. However, dust aerosols are highly irregular in shape, and such irregularity has non-negligible effect on single scattering properties of atmospheric particles. In general, more complicated particles does shows better performance than spheroid in recovering the dust shape distribution[71, 74, 62]. Recently, a physics based surface growth model for ice particle has been introduced[78]. In the present work, we extend this idea to model dust aerosols by exploring roughened spheroid with Gaussian random texture.

There are many works about validation of model particles. In this section, we examine the performance of our model particles by comparing the simulations results with the experimental measurements for feldspar at two wavelengths. A Gaussian spheroid overcome the drawback of other models , for example spheroidal model particles tend to overestimate the effective radii for dust ensembles. We illustrate that a quasi-consistent shape distributions guess for different wave lengths. And we predict that a class of Gaussian particles can show better aerosol retrieval in satellite remote sensing.

We set the aspect ratio and size based on smooth particles, and employ the flexibility from the stochastic particles. A short validation for this wisdom is discussed. A priori estimate of the randomization should be linked with the our Gaussian spheroid. We hope our new model partly solve the difficulties for modeling aerosol size distribution, shape and refractive index in a unified manner.

In[78] , a surface growth model is proposed to model the ice particles. Yang et. al.[2] related roughness effects to local slope statistics, while [78] use correlation length and surface height deviations to characterize the stochastic particles. In this section, we use the standard deviation of surface height *w*as the roughness metric. We also introduce the ratio of roughened and smooth particle surface area ($\gamma = A_r/A_o$, in fig 4.1. $\gamma = 1.6$) [32] as another quantity of roughness. A_o is the sruface area of a smooth particle, the average syrface area of a particler is written as

$$A_r =< \int [1 + (\nabla h)^2]^{1/2} d\mathbf{x} > \tag{4.1}$$

In particular, it has been noted that the above equation becomes $A_r \simeq A_s + \frac{1}{2} < \int (\nabla h)^2 d\mathbf{x} >$

*d***x** for small roughness $|\nabla h| \ll 1$, which unveils the connection between the average surface area of hybrid particle and its average local slope. This new roughness metric also reflects physical/chemical properties of aerosol particles since the surface area of aerosols determines the air-aerosol interaction, For a large aerosol particle, the surface area also determines the radiation power according to the Stefan-Boltzmann law [11]

The Invariant imbedding T-matrix method (II-TM) [53] is a superior method for computing the light scattering properties of non-spherical particles. II-TM computes the T-Matrix via a volume integral equation formulation and taking advantage of invariant imbedding technique. Once the T-Matrix is computed, the optical properties of particles (such as scattering cross sections and phase matrix) can be obtained easily.

Before we talk about the scattering property, we would like to introduce two concepts for probability distributions.The first is Shannon's information entropy [69] which measures the information of a probability distribution which is defined as

$$S = -\int dx\, p(x)\ln(p(x)) \tag{4.2}$$

To classify the similarity between different distributions, we introduce the distance between the two distributions p and q is given by Kullback-Leibler-divergence [70]:

$$KL(p||q) = \int p(x)\ln[p(x)/q(x)]dx \tag{4.3}$$

Since phase function can be considered as a probability density function, we apply Shannon entropy and KL-Divergence to quantify the difference between phase functions. We compare the non-absorbing Gaussian spheroid with size parameter ka = 50. It takes about 50 hours for 20 cpus for particle sizparametere ka = 50, and the memory usage is 100 GB.

65

The refractive index of feldspar at this wavelength is estimated to be 1.31.

Fig. 4.2 (top) shows the Shannon entropy as a function of size parameter ka. The pattern shows the entropy will decrease overall and fluctuate around certain sizes. For size parameter larger than 20, the curve of Shannon entropy begin to split. The surface roughness increase the Shannon entropy for each size. In fig. 2 (bottom), it illustrates that as the roughness metric increase, also the KL-divergence increases wit a larger size parameter.

Fig. 4.3 shows the Müller matrix element for size parameter $ka = 50$ and aspec ratio $\varepsilon = a/b = 2$. It shows rughness has little impact on phase function for angles smaller than 5 deg. But it decreases the phase function from 5 to 40 deg approximately, whilelifting the curve at almost all angles after 40 degree with the largest effect on 180 deg. For the linear polarization, the surface roughness changes the pattern dramatically. All local extrema are smoothed out, increasing surface roughness will drive the curve to a smoother one while he negative value minima near 175 deg is kept. For P22/P11, the effect is even more obvious. Increasing surface roughness decreases value and variance of the curve. Other phase matrix elements are also changed remarkably.

The optical scattering of aerosol particles is size dependent. We simulate light scattering by dust aerosols within size range 1-300 by combining II-TM with IGOM. The refractive index is set to be $m = 1.5 + 0.001i$. The size distribution is statistically estimated from the observations. Specifically, we use the particle volume distribution in our calculation. For the particle shape, we have used a power distribution ξ^n, where n is estimated

to be $n = 3$ [73].

We compare our results with the measurement data from the Amsterdam Light Scattering Database. It can be approximated with the power law, whereas, at m, there are significant deviations from this trend. To overcome this problem geometric optics method is often explored to compute the large size particles. Such a patch has provided satisfactory comparisions to the measured light-scattering response.

Fig. 4.4 shows the recovered curve versus measured data. The simulated data show quite good agreement between measyrements and simulation results. The only unsatisfactory element is P12/P11, where Gaussian spheroids underestimate the value. However, from previous papers, facets particles show better agreement for linear polarization. This indicates that natural aerosols have facets, which cannot be ignored in the modeling.

All evidence in our study shows that Gaussian spheroid model particle gives a better fitting of observed light scattering by aerosols. Gaussian spheroid reproduce most of the Müller matrix elements, but underestimates the linear polarization. A faceted feature is necessary for better modeling of the linear polarization of aerosols.

4.1 Radiative Transfer with Metropolis Algorithm

Radiative transfer equations (RTEs) or Boltzmann equation [75] do not have analytic solutions in general, various numerical techniques are explored to solve this integro-differential equation such as spectral methods, FEM, and Monte Carlo methods[76]. In this dissertation, we will apply the Monte Carlo method to solve RTEs (e.g. fig. 4.5 and
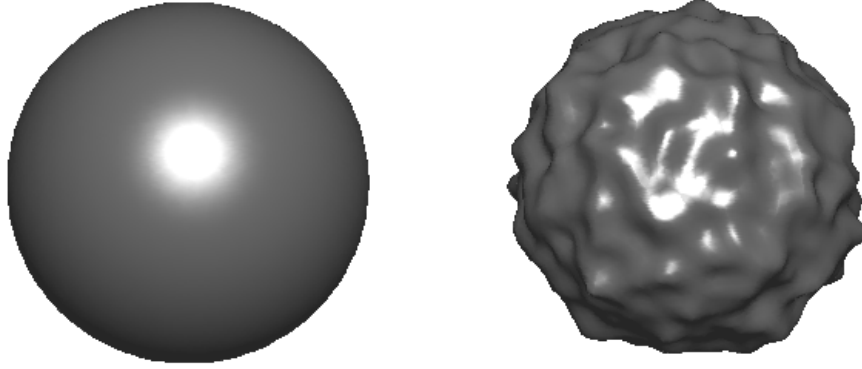
Figure 4.1: Gaussian-spheroid particles

4.6). Monte Carlo integration is a statistical sampling method to estimate the values of integrals. Averaging over repeated independent Monte Carlo samples until the average converge. For integrating low-dimensional smooth functions, other numerical integration methods like trapezoidal integration are very efficient. However, for sufficiently complex high dimensional functions, which are common in RTEs, rates of convergence for conventional methods are not satisfactory. Therefore, we explore the Monte Carlo method to solve high-dimensional complex integro-differential RTEs.

Variance is a fundamental concept for quantifying the performance of a Monte Carlo method. The main disadvantage of Monte Carlo method is that the algorithm converges to the correct result at a rate of $O(n^{-1/2})$. To reduce the error by one half, we have to evaluate four times as many samples. In Monte Carlo radiative transfer, this generally implies

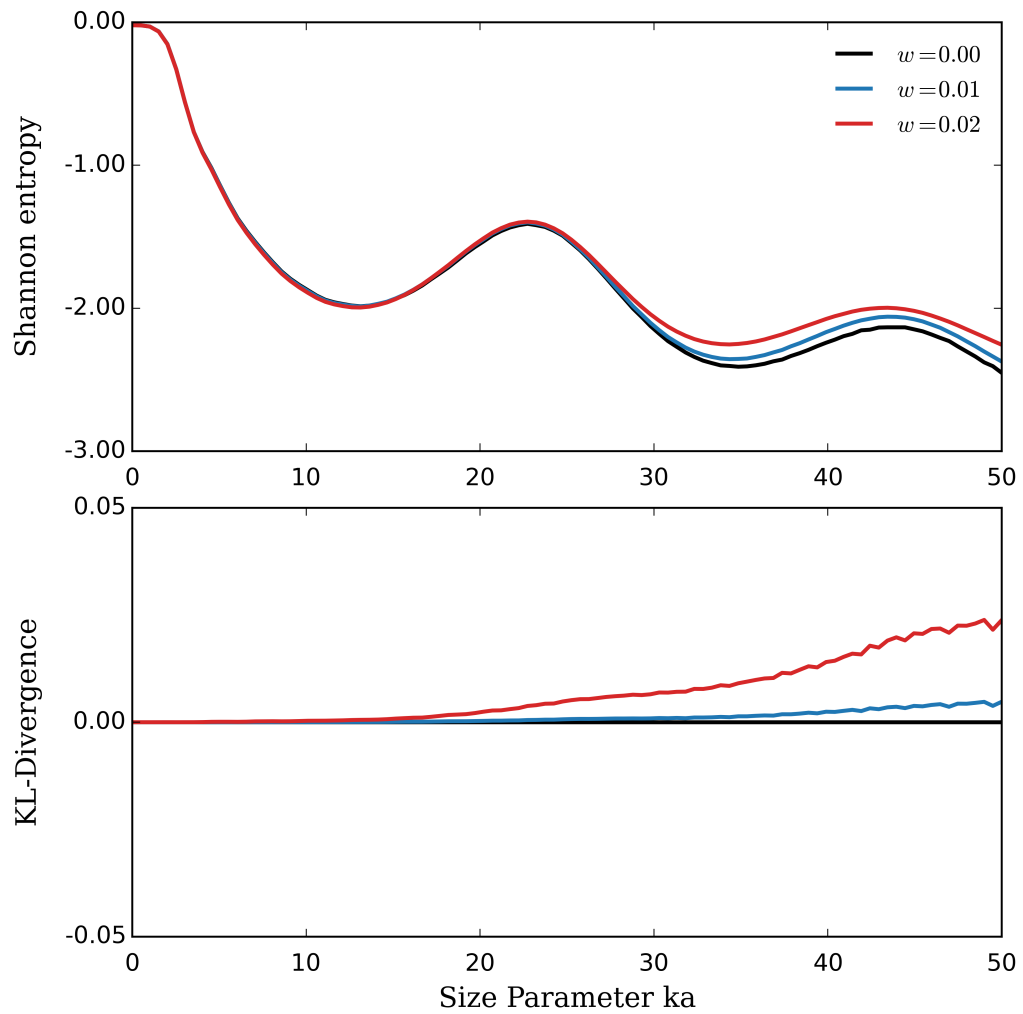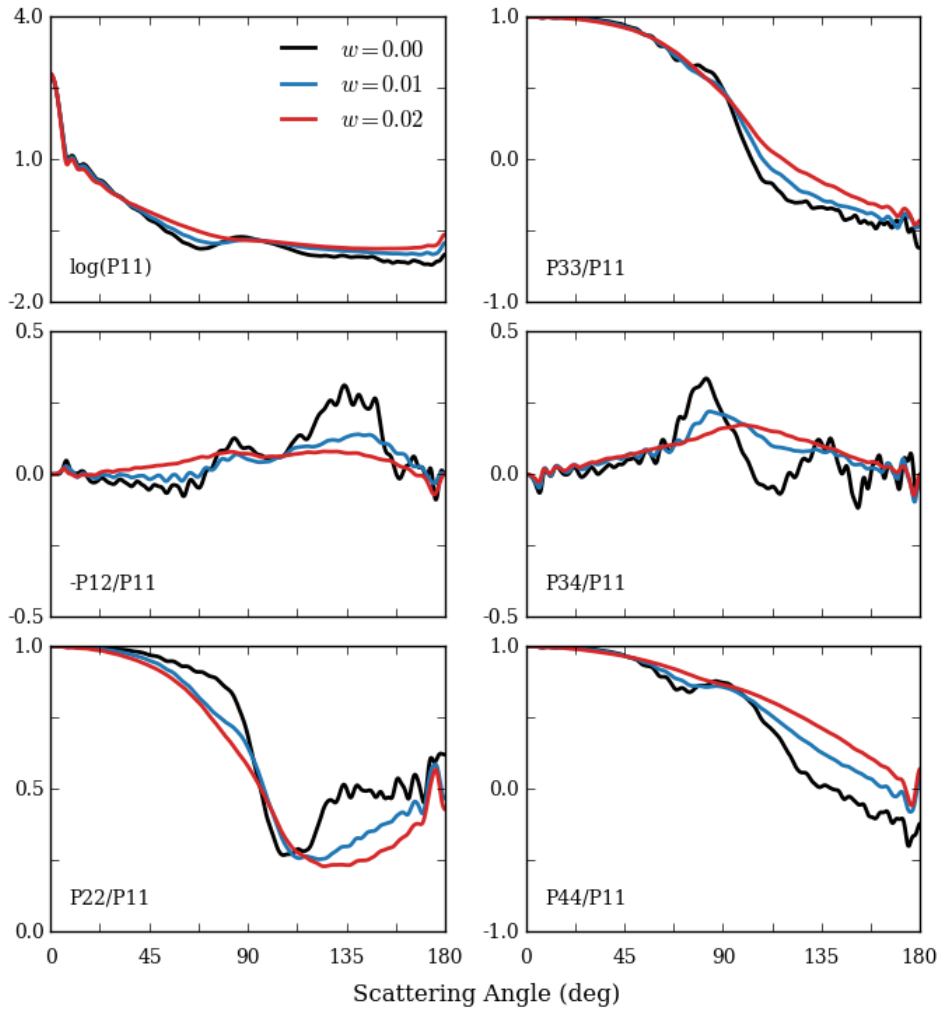Figure 4.2: Shanoon Entropy and KL-Divergence.
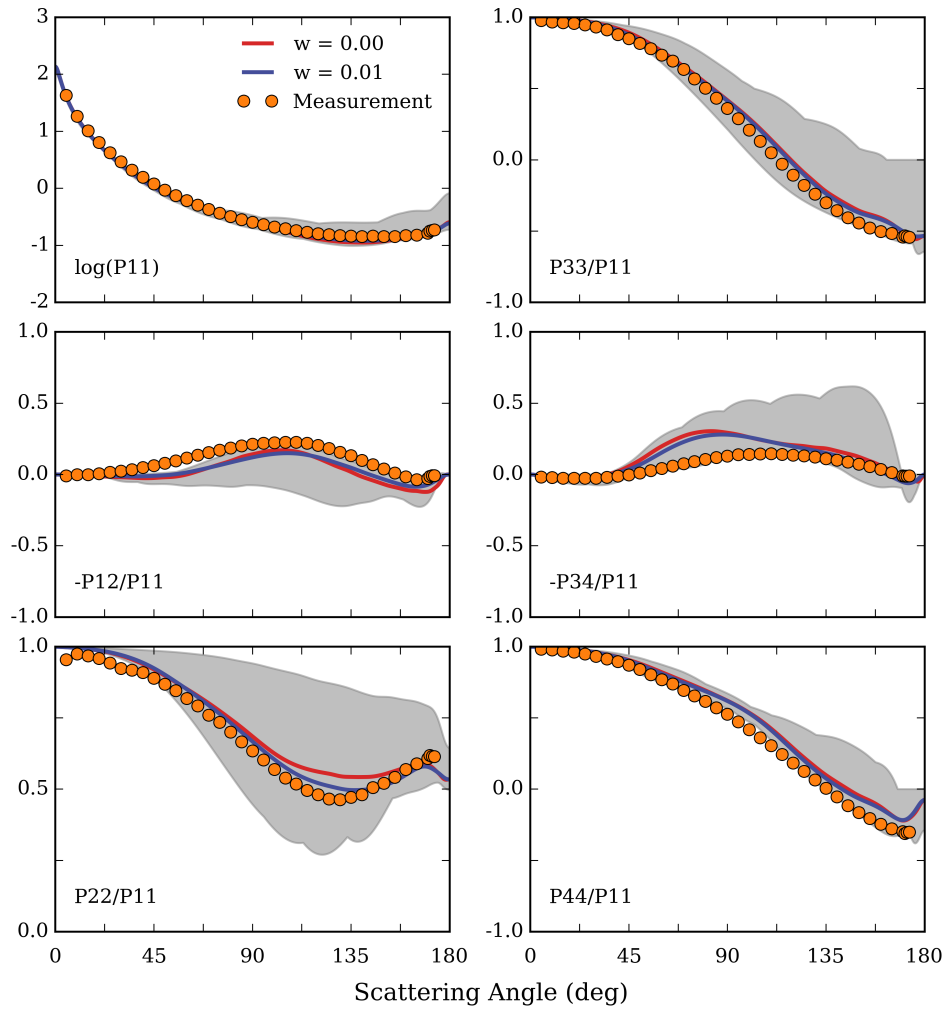
Figure 4.3: Müller matrix

Figure 4.4: Measurement

that $> 1{,}000{,}000$ rays must be traced for computing the radiance which is quite expensive in computing time. In this chapter, we first review fundamental concepts of Monte Carlo methods. Then we describe techniques to improve the convergence rate including importance sampling and the Metropolis algorithm[75].

4.1.1 Monte Carlo Integration

Suppose that we need to integrate a 1-D function $f(x)$ over interval $x \in [a, b]$. The Monte Carlo estimator of the integrand is then the average of $f(x)$ uniformly sampled over domain $[a, b]$:

$$I = \int_a^b f(x)dx = \mathbb{E}[f] \tag{4.4}$$

In addition, the variance of the integration with function $f(x)$ in the interval $x \in [a, b]$ yields a much simpler expression:

$$\mathbb{V}ar[f] = \mathbb{E}[f^2] - \mathbb{E}[f]^2 \tag{4.5}$$

This shows that the error for aribitrary diemensional integration scales as $-1/\sqrt{N}$ where N is the number of samples. This means that is the number of samples taken in Monte Carlo sampling is completely independent of the dimensionality.

Good Monte Carlo algorithm should reduce the variance as much as possible for a fixed sample number N. The simplest approach for reducing variance in Monte Carlo is importance sampling. Importance sampling draws the random variables x are drawn from some a probability density function $p(x)$ similar to function $f(x)$.

It should be noted that the limitation of probability density function $p(x)$ is that it

must be positive definite. Extending this estimator to multiple dimensions or complex integration domains is straightforward. In next sub-section, we will introduce a Metropolis algorithm to reduce the variance of the Monte Carlo integration.

4.2 Metropolis Algorithm

The Ising model is a mathematical model in statistical physics in which the interactions between spins, arranged in a lattice, is limited to nearest neighbors. On the other hand, the Metropolis algorithm works quite well in simulating 2-D Ising model except close to the Curie temperature. In these cases, relaxation times diverge when approaching the Curie temperature (critical slowing down). Hence, more efficient algorithms like cluster algorithm are needed to solve this problem. In this section, we introduce the Metropolis and Wolff cluster algorithm through 2-D Ising model. Critical slowing down phenomena, autocorrelation times as well as the finite size scaling are usually analyzed[75].

The Ising Hamiltonian: The energy of of spin configuration $\{x_i\}$ for the Ising model is given by

$$\mathcal{H} = -\sum_{<i,j>} w_{ij} x_i x_j - \sum_i h_i x_i, \tag{4.6}$$

where $\{h_i\}$ is external magnetic field, w_{ij} is the interaction between spins at sites i and j. The first sum is over the neighbor bonds. We define $\mathcal{M} = \sum_i x_i$ as the magnetization.

When we perform simulations using Ising model, the initial configuration is quite crucial for it usually determines how much computation time is needed we need to attain the equilibrium states. When starting at low temperature, it makes sense to begin with

73

an ordered configuration in which all spins have the same direction, whereas for a high temperature, it would be better to start at a configuration with randomly assigned spins. It would be difficult to guess a suitable initial configuration at a general given T. Sometimes, we may need more Monte Carlo sweeps before the system reaches the most probable states if we unfortunately made a poor estimate. The number of sweeps to reach a equilibrium configuration is known as the thermalization.

The variance of the data in equilibrium can be obtained

$$\sigma^2_{\mathscr{O}-} = \frac{1}{N}[\sigma^2_{\mathscr{O}_i} + 2 \sum_{k=1}^{N} (< \mathscr{O}_1 \mathscr{O}_{1+k} > - < \mathscr{O}_1 >< \mathscr{O}_{1+k} >)(1 - \frac{k}{N})], \quad (4.7)$$

where, due to the last factor $(1 - k/N)$, the $k = N$ term may be trivially kept in the summation.

We also introduce the autocorrelation time

$$\tau'_{\mathscr{O},\text{int}} = \frac{1}{2} + \sum_{k=1}^{N} A(k)(1 - \frac{k}{N}), \quad (4.8)$$

where the normalized autocorrelation function $\phi(t)$ for overvables is defined as

$$A(k) = \frac{< \mathscr{O}_1 \mathscr{O}_{1+k} > - < \mathscr{O}_1 >< \mathscr{O}_{1+k} >}{\sigma^2_{\mathscr{O}_i}} \quad (4.9)$$

4.2.1 Discrete Differential Geometry of 3D Curves

In this section we introduce the discrete differential geometry (DDG) of 3-D curves. We borrow some ideas from protein folding [77]. Let us start with Frenet equations for

continuous curves

$$\frac{d}{ds}\begin{pmatrix} \mathbf{n} \\ \mathbf{b} \\ \mathbf{t} \end{pmatrix} = \begin{pmatrix} 0 & \tau & -\kappa \\ -\tau & 0 & 0 \\ \kappa & 0 & 0 \end{pmatrix}\begin{pmatrix} \mathbf{n} \\ \mathbf{b} \\ \mathbf{t} \end{pmatrix}$$

where the curve γ, $\mathbf{t}$ is the tangential direction, $\mathbf{n}$ is $\mathbf{b}$ is

$$\mathbf{t} = \frac{d}{ds}\gamma$$

Also

$$\mathbf{b} = \mathbf{t} \times \mathbf{n}$$

$$\frac{d}{ds}\mathbf{t} = -\kappa\mathbf{n}$$

The curvature of the curve is

$$\kappa(s) = |\kappa\mathbf{n}| = 1/r$$

We define $\mathbf{t}_i$ as the ray's unit propagation directions. The unit binomal vector is

$$\mathbf{n}_i = \mathbf{t}_{i-1} - (\mathbf{t}_{i-1} \cdot \mathbf{t}_i)\mathbf{t}_i$$

and the unit normal vector is

$$\mathbf{b}_i = \mathbf{t}_i \times \mathbf{n}_i$$

And

$$\cos\theta_{i+1,i} = \mathbf{t}_{i+1} \cdot \mathbf{t}_i$$

$$\cos\phi_{i+1,i} = \mathbf{b}_{i+1} \cdot \mathbf{b}_i$$

$$\begin{pmatrix} \mathbf{n}_{i+1} \\[6pt] \mathbf{b}_{i+1} \\[6pt] \mathbf{t}_{i+1} \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\[6pt] -\sin\phi & \cos\phi & 0 \\[6pt] \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix} \begin{pmatrix} \mathbf{n}_i \\[6pt] \mathbf{b}_i \\[6pt] \mathbf{t}_i \end{pmatrix}$$

where we have used Euler angles $\theta \in [0,\pi]$, $\phi \in [0,2\pi]$. The coordinates of the same point in these two frames are connected by the Euler rotation matrix

$$R(\phi,\theta,\psi=0) = R_z(\psi=0)R_y(\theta)R_z(\phi)$$

where

$$R_z(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\[6pt] -\sin\phi & \cos\phi & 0 \\[6pt] 0 & 0 & 1 \end{pmatrix}$$

$$R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\[6pt] 0 & 1 & 0 \\[6pt] \sin\theta & 0 & \cos\theta \end{pmatrix}$$

$$R_z(\psi) = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\[6pt] -\sin\psi & \cos\psi & 0 \\[6pt] 0 & 0 & 1 \end{pmatrix}$$

Boltzmann equation is

$$(\mathbf{t}\cdot\nabla + c)L(x,t) = b\int_{\mathbb{S}^2} f(\mathbf{t},\mathbf{t}')L d\mathbf{t}' + J(x,\omega)$$

$$G(\mathbf{t};\mathbf{x}'-\mathbf{x}) = \exp\left(-\int_0^{\mathbf{t}(\mathbf{x}'-\mathbf{x})} b(\mathbf{x}'+s\mathbf{t})ds\right)\delta(\mathbf{n}(\mathbf{x}-\mathbf{x}'))\delta(\mathbf{b}(\mathbf{x}-\mathbf{x}'))H(\mathbf{t}(\mathbf{x}-\mathbf{x}'))$$

H is the Heaviside step function. Substitute the Green function to RTE, we obtain Lippmann-Schwinger type equation[12]:

$$L = L_{inc} + \mathbb{G}\mathbb{V}L$$

We could solve Lippmann-Schwinger equation using Born series[75]. Here, we explore Monte Carlo method to solve this equation (e.g. see fig. 4.7). First, we introduce a path integral formulation for radiative transfer:

$$
\begin{aligned}
L_j &= \int d\gamma \exp(-\mathscr{H}(\gamma)) \\
&= \frac{1}{M}\sum_{n=0}^{M}\exp(-\mathscr{H}(\gamma_n^j))
\end{aligned}
$$

Total

$$L = \sum_{j=0}^{\infty} L_j \tag{4.10}$$

The energy functional corresponding to path γ^j

$$\mathscr{H}(\gamma^j) = -\Sigma_i(\ln f(\mathbf{t}_i, \mathbf{t}_{i+1}) - d_i) - j\ln a$$

After the approximation, the energy functional of path γ^j

$$\mathscr{H}(\gamma^j) = \Sigma_i(\alpha\mathbf{t}_i \cdot \mathbf{t}_{i+1} + d_i) - j\ln a$$

that is

$$\mathscr{H}(\gamma^j) = \mathscr{H}_{bending} + \mathscr{H}_{stretching} + C$$

Here, we separate the bending energy and stretching energy. $\mathscr{H}_{bending} = -\Sigma_i \ln f(\mathbf{t}_i, \mathbf{t}_{i+1})$ is classical Heisenberg model[27] like energy (see equation below). In principle, we could sampling the propagation directions according to Gibbs distribution $\frac{1}{Z}e^{-\mathscr{H}_{bending}}$ first and then the optical depth afterwards.

$$\mathcal{H}_{Heisenberg} = -\kappa \sum_{<i,j>} \mathbf{s}_i \cdot \mathbf{s}_j$$

Now, we will try to implement the MCRT with Metropolis-Hasting algorthm. acceptance probbility:

$$a(\gamma, \gamma') = \min\{1, \frac{\pi(\gamma')T(\gamma'|\gamma)}{\pi(\gamma)T(\gamma|\gamma')}\}$$

We write out pseudo code of our algorithm And the Metropolis algorithm for sam-

---

Algorithm 1: Path Integral Radiative Transfer

---

1 Heisenberg Sampler $(\mathbf{d}_{init}, \mathbf{t}_{init})$;
   Input : Arbitrary equilibrium state $\mathbf{d}_{init}, \mathbf{t}_{init}$
   Output: $\mathbf{d}, \mathbf{t}$
2 **for** $i = 1; i < mutations; i{+}{+}$ **do**
3     $\mathbf{t}^{(i)} \leftarrow \mathbf{t}$;
4   **for** $s = 1$ **to** sites **do**
5   sample $t_s^{(i+1)} \sim p(\mathbf{t}^{(i)})$; //Metropolis algorithm
6   sample $d_s^{(i+1)} \sim p(d_s|\mathbf{d}, \mathbf{t}^{(i+1)})$;
7   **end**
8     $\mathbf{t} \leftarrow \mathbf{t}^{(i+1)}$;
9 **end**

---

pling the direction:

In the following, we derive some analytical results for 2-D curves (e.g. fig 4.8).

$$\mathcal{H} = -\sum_{<i,j>} \ln f_s(\mathbf{t}_i \cdot \mathbf{t}_j) \qquad (4.11)$$

$$\approx \kappa \sum \cos \phi_i \qquad (4.12)$$

f_s is the 2-D scattering phase function. The inner product of two direction vectors is

$$\mathbf{t}_m \cdot \mathbf{t}_n = \cos(\phi_m + ... + \phi_{n-1}) = \Re e^{-i(\phi_m + ... + \phi_{n-1})}$$

Algorithm 2: Metropolis-Hasting Algorithm of the Generalized Heisenberg Model

1 Metropolis (γ);
  Input : Initial path γ
  Output: New path γ'
2 $\gamma' = PathMutate(\gamma)$;
3 $a(\gamma',\gamma) = \min\{1, \frac{\pi(\gamma')T(\gamma'|\gamma)}{\pi(\gamma)T(\gamma|\gamma')}\}$;
4 **if** $uniRand(0,1) < a(\gamma',\gamma)$ **then**
5 | $\gamma' = \gamma$;
6 | $\mathscr{H}(\gamma') = \mathscr{H}(\gamma)$;
7 **end**
8 return γ';

The expectation value of inner product of two direction vectors is

$$\mathbb{E}[\langle \mathbf{t}_m, \mathbf{t}_n \rangle] = \Re \prod_{k=m}^{n-1} \mathbb{E}[e^{-i\phi_k}] = \{\frac{\int d\phi \cos\phi \exp[-\ln f_s(\cos\phi)]}{\int d\phi \exp[-\ln f_s(\cos\phi)]}\}^{n-m}$$

Then we make a ansatz that as $N \to \infty$, $\mathbb{E}[\langle \mathbf{t}_m, \mathbf{t}_n \rangle] \propto \exp(-|n-m|/l_0)$. We obtain

the correlation length

$$l_0 = \ln^{-1}\{\int d\phi \cos\phi \exp[-\ln f_s(\cos\phi)]\}$$

Here, the correlation is related to the average of scattering angle $\cos\phi$. Finally, we can

calculate the end to end distance

$$\mathbb{E}(R^2) = \sum_{m,n} \langle \mathbf{t}_m, \mathbf{t}_n \rangle = \sum_{m,n} \exp(-|n-m|/l_0)$$

As $N \to \infty$, we have

$$\mathbb{E}(R^2) \approx N \coth(\frac{1}{2l_0})$$

Similarly, we can also write out results for the 3-D curves using our discrete Frenet curve
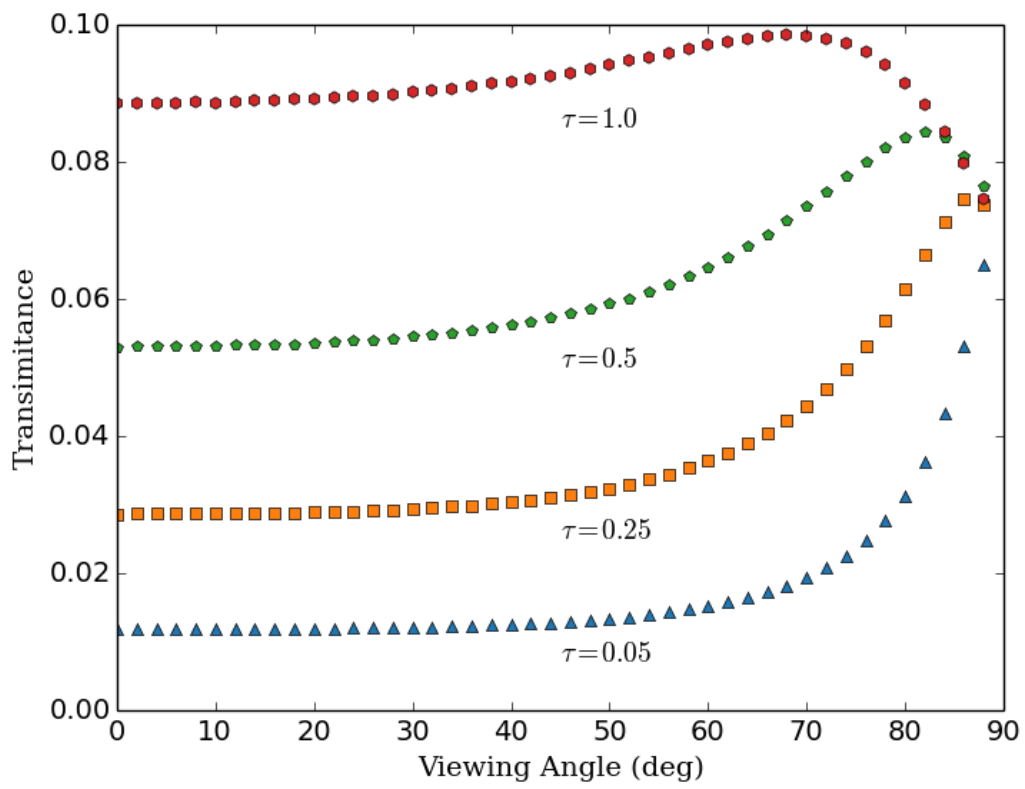
formulation.

Figure 4.5: Transmitance for various optical depth

5 CONCLUSION

In this dissertation, we have constructed a surface growth model for atmospheric ice particles via a stochastic partial differential equation approach. We unveil the connection between our model and previous models. We also simulated the light scattering by surface growth model particles via II-TM. The results indicate that roughness should be classified according to the wavelength besides the particle size. For roughness ($\ll$ wave length), we can ignore the impact on particles optical properties. For roughness ($\sim$ wave length), we analyze its impact on light scattering. The roughness smooth out the phase function and changes the polarization and depolarization properties of particles. Severe roughness ($\gg$ wave length) is not considered in this study. Some superior geometric optics light scattering softwares like IGOM work quite well for large particles with large scale roughness. We also note that more insights about light scattering may be obtained by exploring the non-linear mechanism for surface growth. Further study may involve the modeling of KPZ particles.

Furthermore, we present a detailed formulation of optical modeling of mineral dust aerosols in the framework of a surface growth model. The results have been compared to the experimental measurements of typical dust species to validate the applicability of this model. Due to the hybrid nature, it shows better data fitting than the smooth model particles. Furthermore, we analyze population density and shape distributions of Gaussian spheroid model particles for different mineral dust species. We illustrates a quasi-

consistent shape distribution guess for different wave lengths. We also introduce the Shannon entropy and KL-Divergence to quantity the difference between phase functions.

With the framework of the path integral, we propose a Metropolis Monte Carlo radiative transfer algorithm that is able to solve the RTEs. A Frenet formulation is first time introduced to the radiation transport, in which we have explained the ray using the discrete differential geometry of the 3D Frenet curve. The connection between our model and the 1D classical Heisenberg model is analysed and illustrated.

REFERENCES

[1] van de Hulst, *Light Scattering by Small Particles,* (Dover, New York, 1981)

[2] P. Yang and K. N. Liou, *Single-scattering properties of complex ice crystals in terrestrial atmosphere,* Contr. Atmos. Phys. **71**, 223--248(1998).

[3] Mie, Gustav *Beitrge zur Optik trber Medien, speziell kolloidaler Metallsungen,* Annalen der Physik **330**, 377--445 (1908).

[4] J. A. Stratton, *Electromagnetic Theory,* (McGraw-Hill, New York, 1941)

[5] M. Born and E. Wolf, *Principles of Optics,* Cambridge, New York, 1999) 7th ed.

[6] F. M. Kahnert, *Numerical methods in electromagnetic scattering theory,* J. Quant. Spectrosc. Radiat. Transfer **79-80**, 775--824 (2003)

[7] M. A. Yurkin, V. P. Maltsev, and A. G. Hoekstra, *The discrete dipole approximation for simulation of light scattering by particles much larger than the wavelength,*, J. Quant. Spectrosc. Radiat. Transfer **106** 546?557 (2007).

[8] Q. H. Liu, *The pstd algorithm: A time-domain method requiring only two cells per wavelength,* Microw. Opt. Tech. Lett. **15** (1992), 158--165.

[9] G. Chen, P. Yang, and G. W. Kattawar, *Application of the pseudo-spectral time-domain method to the scattering of light by nonspherical particles,* J. Opt. Soc. Amer. A**44** (2008), 785--790.

[10] A. Zee, *Quantum Field Theory in a Nutshell,* (Princeton, NJ, 2010), 2nd ed.

[11] M. Krüger, G.Bimonte, T. Emig, and M. Kardar, *Trace formulas for nonequi-librium Casimir interactions, heat radiation, and heat transfer* Phys. Rev. B. **86**, 115423(2012).

[12] B. A. Lippmann, and J. Schwinger, *Variational Principles for Scattering Processes. I* Phys. Rev. **79**, 469(1950).

[13] David A. Kopriva, *Implementing Spectral Methods for Partial Differential Equations,* (Springer, 2009).

[14] A. Oskooi, S. G. Johnson, *Distinguishing correct from incorrect PML proposals and a corrected un-split PML for anisotropic, dispersive media,* J. Comput. Phys. **230**, 2369--2377 (2011).

[15] Stephen D. Gedney *Introduction to the Finite-Difference Time-Domain (FDTD) Method for Electromagnetics,* (Morgan & Claypool, San Diego, CA, 2011).

[16] J. A. Roden, S. D. Gedney, *Convolutional PML(CPML): An Efficient FDTD Imple-mentation of the CFS-PML for Arbitrary Media,* Microwave and Optical Technology Letters, **27** 334--339 (2000).

[17] J. P. Berenger, *A perfectly matched layer for the absorption of electromagnetic-waves,* J. Comput. Phys. **114**, 185--200 (1994).

[18] J. P. Berenger, *Three-dimensional perfectly matched layer for the absorption of electromagnetic waves,* J. Comput. Phys. **127** 363--379 (1996).

[19] W. C. Chew and W. H. Weedon,*A 3d Perfectly Matched Medium from Modified Maxwells Equations with Stretched Coordinates,* Microwave and Optical Technology Letters, **7** 599--604(1994).

[20] W. Bangerth, R. Hartmann, and G. Kanschat, ``deal.II ``A general purpose object oriented finite element library,* ACM Transactions on Mathematical Software, **33**, pp 24/1-24/27 (2007).

[21] R. A. Fisher, *Dispersion on a sphere* Proc. Roy. Soc. London Ser. A, **217**, 295–305 (1953).

[22] D. Donoho, *For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution* Communications on pure and applied mathematics, **59**, 797-829 (2006).

[23] E. J. Candes, J. K. Romberg, and T. Tao, *stable signal recovery from incomplete and inaccurate measurements,* Communications on pure and applied mathematics, **59**, 1207-1223 (2006).

[24] W. Yin, S. Osher, D. Goldfarb, J. Darbon, *Bregman iterative algorithms for l1-minimization with applications to compressed sensing*, SIAM J. Imag. Sci. **1**, 142–168 (2008)

[25] D. H. Ackley, G. E. Hinton; T. J. Sejnowski, *"A Learning Algorithm for Boltzmann Machines"* . Cognitive Science 9 (1): 147–169 (1985).

[26] A. Fischer and C. Igel, *Training Restricted Boltzmann Machines: An Introduction*. Pattern Recognition 47:25-39(2014).

[27] J. P. Sethna. *Statistical Mechanics: Entropy, Order Parameters, and Complexity.* (Oxford: Clarendon Press, 2010).

[28] H. Katzgraber, *Introduction to Monte Carlo Methods.* arXiv:0905.1629 (2011).

[29] K. binder and D. W. Heermann,*Monte Carlo Simulations in Statistical Physics(Fifth Edition)* (Springer, 2010).

[30] Y. Tang, R. Salakhutdinov, and G. Hinton, *Deep lambertian network,* arXiv preprint arXiv:1206.6445.

[31] A. A. Maradudin *Light Scattering and Nanoscale Surface Roughness,* (Springer, New York, 2007).

[32] Y-P Zhao, G. Wang, T-M Lu *Characterization of Amorphous and Crystalline Rough Surface--Principles and Applications*, (Academic Press, San Diego, CA, 2001).

[33] B. A. Baum, P. Yang, Y. -X. Hu, and Q. Feng, *The impact of ice particle roughness on the scattering phase matrix*, J. Quant. Spectrosc. Radiat. Transf. **111**, 2534--2549 (2010).

[34] W. C. Pfalzgraff, R. M. Hulscher, and S. P. Neshyba, *Scanning electron microscopy and molecular dynamics of surfaces of growing and ablating hexagonal ice crystals*, Atmos. Chem. Phys. **10** 2927--2935 (2010).

[35] T.H. Beuman, A.M. Turner, and V. Vitelli, *Stochastic geometry and topology of non-Gaussian fields,* Proc. Natl. Acad. Sci. **109** (2012), 19943--19948.

[36] V. Shcherbakov, J. F. Gayet, B. Braker, P. Lawson, *Light scattering by single natural ice crystals*, J. Atmos. Sci. **63**(5),1513--1525(2006).

[37] A. Ehrlich, M. Wendisch, E. Bierwirth, A. Herber, A. Schwarzenbock, *Ice crystal shape effects on solar radiative properties of Arctic mixed-phase clouds--Dependence on microphysical properties*, Atmos. Res.**88**, 266--276 (2008).

[38] A. J. Baran, *A review of the light scattering properties of cirrus*, J. Quant. Spectrosc. Radiat. Transf. **110**, 1239--1260 (2009).

[39] A. J. Baran, *From the single-scattering properties of ice crystals to climate prediction: A way forward*, Atmos. Res. **112**, 45--69 (2012).

[40] Z. Ulanowski, P. H. Kaye, E. Hirst, R. S. Greenaway, R. J. Cotton, E. Hesse, and C. T. Collier *Incidence of rough and irregular atmospheric ice particles from Small Ice Detertor 3 measurements*, Atmos. Chem. Phys., **14**, 1649--1662(2014).

[41] S. Schröder, A. Duparré , L. Coriand, A. Tünnermann, D. H. Penalver, and J. E. Harvey *Modeling of light scattering in different regimes of surface roughness*, Opt. Express **19**, 9820-9835 (2011).

[42] T. Nousiainen, K. Muinonen, *Surface-roughness effects on single- scattering properties of wavelength-scale particles*, J. Quant. Spectrosc. Radiat. Transf. **106**, 389--397(2007).

[43] M. Kahnert, T. Nousiainen, P. Mauno, *On the impact of non-sphericity and small-scale surface roughness on the optical properties of hematite aerosols*, J. Quant. Spectrosc. Radiat. Transf. **63**, 1513--1525(2006).

[44] M. Kahnert, T. Rother, *Modeling optical properties of particles with small-scale surface roughness: combination of group theory with a perturbation approach*, Opt. Express **19**11138--11151(2011).

[45] M. Kahnert,T. Nousiainen, M. A. Thomas, J. Tyynel, *Light scattering by particles with small-scale surface roughness: Comparison of four classes of model geometries*, J Quant Spectrosc Radiat Transfer **113**, 2356--2367(2012).

[46] J. M. Edwards, S. Havemann, J. C. Thelen, A. J. Baran, *A new parameterization for the radiative properties of ice crystals: comparison with existing schemes and impact in a GCM*, Atmos. Res. **83**, 19--34(2007).

[47] A. A. Maradudin *Light Scattering and Nanoscale Surface Roughness*, (Springer, New York, 2007).

[48] K. Muinonen, T. Nousiainen, P. Fast, K. Lumme, J. Peltoniemi, *Light scattering by Gaussian random particles: Ray optics approximation*, J. Quant. Spectrosc. Radiat. Transf. **55**, 577--601(1996).

[49] C. Li, G. W. Kattawar, and P. Yang, *Effects of surface roughness on light scattering by small particles*, J. Quant. Spectrosc. Radiat. Transf. **89**, 123--131 (2004).

[50] Liu, C., R. L. Panetta, and P. Yang, *The effects of surface roughness on the scattering properties of hexagonal columns with sizes from the Rayleigh to the geometric optics regimes,* J. Quant. Spectrosc. Radiat. Transfer **129**, 169--185 (2013).

[51] T. Nousiainen, G. M. McFarquhar, *Light scattering by quasi-spherical ice crystals*, J. Atmos. Sci. **61**,2229--2248(2004).

[52] K. Crane, F. de Goes, M. Desbrun, P. Schröder, *Digital Geometry Processing with Discrete Exterior Calculus*, ACM SIGGRAPH 2013 courses, SIGGRAPH '13 **(**56), 899 (ACM, New York, NY)(2013).

[53] L. Bi, P. Yang, *Accurate simulation of the optical properties of atmospheric ice crystals with the invariant imbedding T-matrix method*, J. Quant. Spectrosc. Radiat. Transf. **138**, 17--35(2014).

[54] B. R. Johnson *Invariant imbedding T-matrix approach to electro- magnetic scattering*, Appl. Opt. **27**, 4861--4873(1988).

[55] M. I. Mishchenko, G. Videen, N. G. Khlebtsov, T. Wriedt, N. T. Zakharova, *Comprehensive T-matrix reference database: A 2006--07 update*, J. Quant. Spectrosc. Radiat. Transf. **109**, 1447-1460(2008).

[56] D. W. Mackowski, *Discrete dipole moment method for computing the T matrix for nonspherical particles*, J. Opt. Soc. Am. A **19**, 881--893 (2002).

[57] M. Kardar, G. Parisi, and Y.-C. Zhang, *Dynamical scaling of growing interfaces*, Phys. Rev. Lett. **56**, 899 (1986).

[58] S. F. Edwards and D. R. Wilkinson, *The surface statistics of a granular aggregate*, Proceedings of the Royal Society (London) A **381**, 17 (1982).

[59] P. C. Waterman, *Symmetry, Unitarity, and geometry in electromagnetic scattering*, Phys Rev D **3** 825--839(1971).

[60] M. I., Mishchenko, and L.D. Travis, *T-matrix computations of light scattering by large spheroidal particles*, Opt. Comm. **109**, 16-21(1994).

[61] T. Nousiainen, K. Muinonen, *Surface-roughness effects on single- scattering properties of wavelength-scale particles*, J. Quant. Spectrosc. Radiat. Transf. **106**, 389--397(2007).

[62] Jianping Liu, Ping Yang, Karri Muinonen, *Dust-aerosol optical modeling with Gaussian spheres: Combined invariant-imbedding T-matrix and geometric-optics approach*, J. Quant. Spectrosc. Radiat. Transf. **161**, 136--144(2015).

[63] Dubovik O, Sinyuk A, Lapyonok T, Holben BN, Mishchenko M, Yang P, et al. *Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust*, J Geophys Res: Atmos **111**, (2006).

[64] M. Kahnert, T. Nousiainen, P. Mauno, *On the impact of non-sphericity and small-scale surface roughness on the optical properties of hematite aerosols*, J. Quant. Spectrosc. Radiat. Transf. **63**, 1513--1525(2006).

[65] M. Kahnert, T. Rother, *Modeling optical properties of particles with small-scale surface roughness: combination of group theory with a perturbation approach*, Opt. Express **19**11138--11151(2011).

[66] M. Kahnert,T. Nousiainen, M. A. Thomas, J. Tyynel, *Light scattering by particles with small-scale surface roughness: comparison of four classes of model geometries*, J Quant Spectrosc Radiat Transfer **113**, 2356--2367(2012).

[67] K. Muinonen, T. Nousiainen, P. Fast, K. Lumme, J. Peltoniemi, *Light scattering by Gaussian random particles: ray optics approximation*, J. Quant. Spectrosc. Radiat. Transf. **55**, 577--601(1996).

[68] T. Nousiainen, G. M. McFarquhar, *Light scattering by quasi-spherical ice crystals*, J. Atmos. Sci. **61**,2229--2248(2004).

[69] C. E. Shannon, W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press (1949).

[70] S. Kullback, R. A. Leibler, *On information and sufficiency*, Annals of Mathematical Statistics **22**, 79--86(1951).

[71] L. Bi, P. Yang, *Modeling optical properties of mineral aerosol particles by using nonsymmetric hexahedra*, Appl Opt. **49**, 334--342(2010).

[72] T. Nousiainen, *Optical modeling of mineral dust particles: a review*, J. Quant. Spectrosc. Radiat. Transf. **110**, 1261--1279(2009).

[73] S. Merikallio, H. Lindqvist, T. Nousiainen, M. Kahnert *Modelling light scattering by mineral dust using spheroids: assessment of applicability*, Atmos. Chem. Phys., **11**, 3977--4016(2011)

[74] L. Bi, P. Yang, G. W. Kattawar, and R. Kahn, *Single-scattering properties of tri-axial ellipsoidal particles for a size parameter range from the Rayleigh to geometric-optics regimes*, Appl. Opt. **48**, 114--126 (2009).

[75] H. Fehske, R. Schneider,and A. Weibe, Computational Many-Particle Physics, Springer (2008).

[76] M. Wendish, and P. Yang, Theory of Atmospheric Radiative Transfer, Weley (2012).

[77] S. Hu, A. Krokhotin, A. J. Niemi, and X. Peng, *The Discrete Frenet Frame, Inflection Point Solitons And Curve Visualization with Applications to Folded Proteins*, Phys. Rev. E**83**, 041907(2011).

[78] Jianing Zhang, Lei Bi, Jianping Liu, R. L. Panetta, Ping Yang, G. W. Kattarwa, *Optical scattering simulation of ice particles with surface roughness modeled using the Edwards-Wilkinson equation*, J. Quant. Spectrosc. Radiat. Transf. **178**, 325- -335(2015).