Filtering Problems in Stochastic Tomography

Tyler Gomez

University of Central Florida
FILTERING PROBLEMS IN STOCHASTIC TOMOGRAPHY

by

TYLER GOMEZ
M.S. University of Central Florida, 2014

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Mathematics
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2017

Major Professor: Jason Swanson
Distinguishing signal from noise has always been a major goal in probabilistic analysis of data. Such is no less the case in the field of medical imaging, where both the processes of photon emission and their rate of absorption by the body behave as random variables. We explore methods by which to extricate solid conclusions from noisy data involving an X-ray transform, long the mathematical mainstay of such tools as computed axial tomography (CAT scans). Working on the assumption of having some prior probabilities assigned to various states a body can be found in, we introduce and make rigorous an understanding of how to condition these into posterior probabilities by using the scan data.
To my father for the confidence to take on a challenge, to my mother for the perseverance to see it through to the end, and to my brother for a shining example of life on the other side.
ACKNOWLEDGMENTS

Dr. Jason Swanson has been an invaluable resource as both an advisor and academic motivator. I greatly appreciate all the time and patience granted me. My thanks go out also to Dr. Alexandru Tamasan, who first introduced me to the world of Radon transforms and tomography at large.
# TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION ..................................................... 1

Stochastic Models ............................................................... 2

Inferences from Observations ................................................. 4

Limiting Behavior of Imaging Model ........................................ 4

CHAPTER 2: INFRASTRUCTURE ................................................. 6

Attenuation ................................................................. 6

X-ray Transform ............................................................. 7

White Noise ................................................................. 8

Random Variables .......................................................... 11

CHAPTER 3: ANALYSIS ON LIMIT MODEL ................................. 13

Preliminaries on $Y_{1,\varepsilon}$ ................................................. 14

$H$ Subspace Projection ...................................................... 15

Truncated Sigma Algebras .................................................... 16

Result in Subspace .......................................................... 22

Result in $L^2(Z)$ .............................................................. 31
CHAPTER 1: INTRODUCTION

The topic of medical imaging may be viewed mathematically as the problem of recovering a compactly supported spatial function (i.e. the 3-dimensional body or body part being scanned) using line integrals of the function. These line integrals correspond to the probabilities of photons emitted from one side of the body being detected on the other along the given path. Some photons get absorbed by the body and are, of course, never detected. In an ideal world, we could say we emit a number of photons \( N \) at all possible directions around the body and detect a fraction of \( N \) on the other side for each trajectory. This would assign an empirical absorption rate to every line through the body which could then be used to recreate a representation of the observed object.

There are several issues with this idealization, however. First of all, we can never realistically obtain complete coverage of the body. In the real world, we emit and observe photons along a finite number of lines through the body, though the number can grow quite large. The second issue is that the existence of an absorption rate of photons along a line through the body does not imply the absorption of equal proportion of photons across multiple iterations of emission and detection. While the proportion tends toward the correct value in the limit, the signal to noise ratio can be quite poor at lower values.

The operator that takes a function and outputs its various line integrals is called the X-ray transform. When one is operating in 2D space, this agrees with the Radon transform introduced a century ago. Analysis of the idealized cases (which ignore one or several of the listed real-world impediments) inform the practical improvement of medical tomography. Which paths the photons use to traverse a body and how much is radiated can be tuned in myriad ways. UCF’s own Dr. Katsevich has done often cited work on the recovery of body functions for scanners using a spiral formation [1], [2].
As we said, a primary hindrance is the difficulty of distinguishing useful observed data from noise. One choice that increases confidence in the detected data is to bombard the body under investigation with larger amounts of radiation until the image becomes clear, effectively whittling at the signal to noise ratio until the error no longer impedes an adequate diagnosis. This is perfectly valid from a mathematical perspective, but the realities of biology act as a deterrent. The amount of radiation of some medical scans is comparable to several years’ worth of natural background radiation, which can outweigh the benefits of potentially finding something wrong in a patient with no outward symptoms.

While radiation doses vary greatly depending on the scan, in all cases we want to be able to say more while radiating less. This improvement of the radiation-to-resolution ratio is the highest goal of medical imaging. The standards are continually improving. When a probabilistic environment is adhered to, the focus is on how to truly divorce the signal from the noise. It’s important to know how much of our observations are indicative of real-world information versus what is simply arbitrary (and misleading) static. It is our goal to maintain a standard of rigor that will yield the most realistic predictions possible without ignoring the well-known realities of the problem as listed above.

Stochastic Models

Consider a beam of X-ray particles traveling along some line $L_y$ through a body, where the index $y$ denotes the closest point on the line to the body’s center. The beam is produced by an emitter that releases a number of particles with mean $N$. Each particle has a probability $p = e^{-Xf(y)}$ of emerging from the body without being absorbed and thus detected at the other end, where $Xf$ denotes the X-ray transform of the body’s attenuation function, both formally defined in the next chapter. If $S_N f$ is the number of particles that emerge, which will also be more rigorously
defined in subsequent chapters, then $S_k^N f / N \approx p$ when $N$ is large, according to the law of large numbers. This is the basic mathematical principle at the heart of tomographic reconstruction. It says that

$$-\log \left( \frac{S_k^N f}{N} \right) \approx X f(y).$$

(1.1)

In practice, $N$ is very large, $S_k^N f$ is empirically measured, and the above approximate equality is taken to be an equality. In this way, one presumes to be able to observe $X f(y)$ for any given line $L_y$. The mathematical problem is then to recover the function $f$ from its X-ray transform, $X f$.

However, as mentioned in the introduction, it is always preferable to gather more information with less radiation, as excessive X-rays can damage living cells. Because of this, it is desirable to let $N$ be as small as possible. On the other hand, if we reduce $N$, then (1.1) may no longer be valid as the effects of the error inherent in the approximation grow. For smaller values of $N$, approximation (1.1), which is based on the law of large numbers, should be replaced with

$$-\log \left( \frac{S_k^N f}{N} \right) \approx X f(y) - N^{-1/2} (e^{X f(y)} - 1)^{1/2} \eta_y,$$

(1.2)

which is based on the central limit theorem as outlined in Chapter 4. Here, $\eta_y$ is a random variable with a standard Gaussian law. In this situation, we no longer directly observe the line integrals of $f$. Rather, we observe the line integrals perturbed by random noise. The imaging problem, therefore, becomes a two-part problem. We must first filter out the noise to recover the line integrals. Only then can we use the line integrals to reconstruct the image.

The error in (1.1) has order of magnitude $N^{-1/2}$, whereas the error in (1.2) has order of magnitude $N^{-1}$. So if $N_0$ X-ray photons are needed for (1.1) to be reasonably accurate, then the same accuracy can be obtained from (1.2) using only $\sqrt{N_0}$ photons. In other words, equation (1.2) can be used with significantly less radiation exposure. This is the benefit of working with a model more
explicitly accounting for random effects.

Inferences from Observations

One of our goals is to delineate how much can be said about the body undergoing an X-ray scan after the data is observed. From a Bayesian perspective, this amounts to modifying our prior assumptions on the body given the information found. If we had, for instance, a sequence of weights \( \{w_j\} \) assigned a priori to any of a countable set of possible conditions for the scanned body, then

\[
P(\text{body is in state } k) = \frac{w_k}{\sum_{j=1}^{\infty} w_j}
\]

without other information. Our efforts reveal the posterior weights to assign to different states of the body, and in fact find them to be exponential in nature, so the posterior probabilities of unlikely states diminish rapidly as we gather information of the body through scanning. Thus one main achievement is the proper characterization of the sequence \( \{w'_j\} \), where

\[
P(\text{body is in state } k \text{ given our observations}) = \frac{w'_k}{\sum_{j=1}^{\infty} w'_j},
\]

and we can be said to have greater confidence in the posterior weights.

Limiting Behavior of Imaging Model

The main focus of our work is to explicitly determine the form of the central limit theorem model for our observations mentioned previously. This would take into account the white noise effects inherent in the transmission probabilities of photons on lines through the scanned body. It is in the concluding chapter that much of the work veers away from Bayesian updating on priors to focus
solely on the limiting behavior of an X-ray model.

We will use \( \frac{S_y f}{N} \) to denote the proportion of detected photons along a line determined by \( y \) out of a set of \( N \), but note that scanning only occurs along finitely many lines and photon emission is preferably kept to a minimum. Thus we will take a wide view of the problem and index our observations by both the lines through the body and the photon emissions to account for all the random effects in our probabilistic analysis. It is in so doing that we find our observations to be best characterized by

\[
Y \approx X f - \frac{\pi}{(4^n N)^{1/2}} (e^{X f} - 1)^{1/2} W
\]

when we scan along \( 4^n \) lines and emit \( N \) photons along each line, where \( W \) is white noise as described in Chapter 2.
CHAPTER 2: INFRASTRUCTURE

We will now set about describing the mathematical environment of medical tomography. Many of these definitions will be used throughout the other chapters of this text. We will first outline the assumptions on the body to be scanned and follow with a formal definition of an X-ray scan of such a body. We will then outline the probabilistic setting of medical tomography as an overview and finally introduce the random variables that will bridge the gap between the deterministic form of medical tomography and its probabilistic aspects.

Attenuation

To begin, we will use $D = \{x \in \mathbb{R}^2 : |x| < 1\}$ to denote the unit disk and denote the Borel $\sigma$-algebra on $D$ by $\mathcal{B}_D$. Letting $\lambda$ denote the Lebesgue measure on $(D, \mathcal{B}_D)$, we obtain the measure space $(D, \mathcal{B}_D, \lambda)$. For our purposes, $D$ shall represent a circular cross-section of a body being X-rayed. Thus, any point $x \in D$ would mark a location on this cross-section.

The image we wish to reconstruct is represented by a continuous function $f : D \to [0, \infty)$ such that $f \geq 0$. Intuitively, one might think of $f$ as describing the density of the body, although, to be more precise, it describes the X-ray attenuation of the body. The attenuation function, $f$, describes the body’s propensity for absorbing X-ray particles. Consider an X-ray particle located at a point $x \in D$. Roughly speaking, during any given time interval of duration $\Delta t$, the probability that the particle will be absorbed is approximately $f(x)\Delta t$, with the approximation improving as $\Delta t \to 0$.

For a rigorous formulation of the attenuation function, consider an X-ray particle that travels along a path $C$ in $D$ at unit speed. The location of the particle at time $t$ is represented by $r(t)$, where $r : [a, b] \to \overline{D}$ is a continuous parameterization of $C$ with $|r'(t)| = 1$ for all $t$. Let the random
variable $T$ represent the time at which the particle is absorbed, with $T$ taking values in $(a, b) \cup \{\infty\}$. The event, $\{T = \infty\}$, represents the event that the particle is never absorbed. For each fixed $t \in (a, b)$, the attenuation function $f$ is assumed to satisfy

$$P(T \leq t + \Delta t \mid T > t) = f(r(t))\Delta t + o(\Delta t).$$

The probability that the particle survives until time $t$ is $y(t) := P(T > t)$. Since

$$P(T \leq t + \Delta t \mid T > t) = \frac{P(t < T \leq t + \Delta t)}{P(T > t)} = \frac{y(t) - y(t + \Delta t)}{y(t)},$$

it follows that $y'(t) = -f(r(t))y(t)$. Since $f \circ r$ is continuous and $y(a) = 1$, this gives

$$y(t) = \exp \left( -\int_a^t f(r(t)) \, dt \right).$$

We assumed $|r'(t)| = 1$ for all $t$, and so the probability that the particle is never absorbed is

$$P(T = \infty) = y(b) = \exp \left( -\int_a^b f(r(t))|r'(t)| \, dt \right) = \exp \left( -\int_C f \, ds \right),$$

where the above integral is the line integral with respect to arc length.

**X-ray Transform**

In X-ray tomography, each X-ray photon will travel through the body along a line $L$. The X-ray transform is the operator that transforms the function $f$ into the function $L \mapsto \int_L f \, ds$. Let
\( Z = (0, 1) \times S^1 \), where \( S^1 \) denotes the unit circle in \( \mathbb{R}^2 \). For \( y = (s, \tau) \in Z \), let \( L_y = \{ x \in \mathbb{R}^2 : \langle x, \tau \rangle = s \} \) denote the line through \( s \tau \) that is orthogonal to \( \tau \). If \( f \in L^2(D, \mathcal{B}_D, \lambda) \), where \( \lambda \) is Lebesgue measure, then the X-ray transform of \( f \) is the function \( Xf : Z \to \mathbb{R} \) given by

\[
Xf(y) = \int_{L_y \cap D} f \, d\lambda_{L_y},
\]

where \( \lambda_{L_y} \) is Lebesgue measure on \( L_y \). If we define the measure \( \mu \) on \( Z \) by

\[
\mu(ds \, d\tau) = (1 - s^2)^{-1/2} ds \, \sigma(d\tau),
\]

where \( \sigma \) is Lebesgue measure on \( S^1 \), then \( Xf \in L^2(Z, \mathcal{B}_Z, \mu) \), and \( X : L^2(D) \to L^2(Z) \) is a bounded linear operator. Note that \( |Xf(y)| \leq \|f\|_\infty \lambda_{L_y}(L_y \cap D) \). Hence, \( \|Xf\|_\infty \leq 2\|f\|_\infty \). In particular, \( X : L^\infty(D) \to L^\infty(Z) \) is continuous, and also \( X : C(\overline{D}; [0, \infty)) \to C(\overline{Z}; [0, \infty)) \) is continuous.

With this notation, if an X-ray particle travels through \( D \) along the line \( L_y \) at unit speed, then it will emerge from \( D \) without being absorbed with probability \( e^{-Xf(y)} \). For greater values of the attenuation function along \( L_y \), therefore, we have reduced rates of transmission. Likewise, as attenuation diminishes along a given line, the probability of a photon passing through for our detectors to observe tends to 1.

**White Noise**

Up to this point, we have made little allusion to the probabilistic aspects of medical imaging. Despite formalizing the probability of a single photon passing along a given line through a body, the setting is currently in the purview of functional analysis. It is now that we outline the contributions to be gained from observing the problem as one of random variables.
As alluded to in the introduction, the central limit theorem model comprises the non-probabilistic limit of the law of large numbers model with an added white noise term. We now formalize the white noise process to be used throughout the paper. Let $H$ be a real, separable Hilbert space with inner product $\langle \cdot, \cdot \rangle_H$. An isonormal Gaussian process on $H$ is a centered Gaussian family of random variables $W = \{W(h) : h \in H\}$ on a complete probability space $(\Omega, \mathcal{F}, P)$ such that $E[W(h)W(g)] = \langle g, h \rangle_H$ for all $g, h \in H$. In the case that $H = L^2(M, \mathcal{M}, \mu)$ we say that $W$ is a white noise on $M$, and we often write $\int_M h \, dW$ instead of $W(h)$. For instance, if $M$ is a separable metric space, $\mathcal{B}_M$ denotes its Borel $\sigma$-algebra, and $\mu$ is a $\sigma$-finite measure on $(M, \mathcal{B}_M)$, then $L^2(M, \mathcal{B}_M, \mu)$ is a real, separable Hilbert space (see, for instance, [4, Proposition 3.4.5]), and we may define white noise on $M$ as above.

For each $h \in H$, we have a real-valued, square-integrable random variable, $W(h)$. We use $W(h, \omega)$ to denote $W(h)(\omega)$. It is not true that for each fixed $\omega \in \Omega$, the map $h \mapsto W(h, \omega)$ is a continuous linear functional on $H$. However, as demonstrated in [3], the map $h \mapsto W(h)$ is an isometry from $H$ onto a closed subspace of $L^2(\Omega)$. We therefore have

$$W(c_1g_1 + c_2g_2) = c_1W(g_1) + c_2W(g_2) \text{ a.s.,}$$

for all $c_1, c_2 \in \mathbb{R}$ and $g_1, g_2 \in L^2(S)$.

Let $\{h_j\}_{j=1}^n$ be an orthonormal basis for $H$. Let $Q_n$ be the projection operator from $H$ onto $\text{span}\{h_1, \ldots, h_n\}$. That is,

$$Q_nh = \sum_{j=1}^n \langle h, h_j \rangle h_j,$$

for all $h \in H$. Since $Q_n$ is an orthogonal projection, we have $Q_n^2 = Q_n$ and

$$\langle Q_ng, h \rangle = \langle g, Q_nh \rangle,$$
for all $g, h \in H$. Also note that $Q_n h \to h$ in $H$ for all $h \in H$.

For $h \in H$, let

$$W_n(h) = W(Q_n h) = \sum_{j=1}^{n} \langle h, h_j \rangle W(h_j).$$

Since $W : H \to L^2(\Omega)$ is continuous, it follows that $W_n(h) \to W(h)$ in $L^2(\Omega)$, which implies $W_n(h) \to W(h)$ in $L^1(\Omega)$. Since $\{W(h_j)\}_{j=1}^{\infty}$ is a sequence of independent standard normal random variables, we have

$$E[W_{n+1}(h) \mid W_1(h), \ldots, W_n(h)] = W_n(h) + E[\langle h, h_{n+1} \rangle W(h_{n+1})] = W_n(h).$$

That is, $\{W_n(h)\}_{n=1}^{\infty}$ is a martingale. For martingales, $L^1$ convergence implies convergence a.s. (See, for example, [11, Theorem 5.5.6].) Therefore, $W_n(h) \to W(h)$ a.s.

Let $A = \{(h, \omega) : \lim_{n \to \infty} W_n(h, \omega) exists\}$. Since each $W_n$ is $B_H \otimes F$-measurable, it follows that $A \in B_H \otimes F$. Let $\tilde{W}$ denote the pointwise limit of $W_n1_A$, which is consequently $B_H \otimes F$-measurable. Then $\tilde{W}(h) = W(h)$ a.s. for each $h \in H$. That is, $\tilde{W}$ is an isonormal Gaussian process on $H$, which is a modification of $W$, and is also $B_H \otimes F$-measurable. In other words, we may always choose $W$ so as to be jointly measurable in both $h$ and $\omega$. We will appeal to this facet of $W$ both directly and subtly throughout this work.

Throughout this work, we will be taking $H := L^2(Z)$ where we remind the reader that $X f \in L^2(Z)$ for all continuous functions $f$ on $\overline{D}$. This will serve to account for the white noise error term seen in the detection of photons passing through a line in the body and furthermore our observations of an attenuation function scanned over the space of lines.
Random Variables

We are now equipped to speak of the various random variables that comprise our models. The following and final section of this chapter will detail the probability space on which these random variables operate and outline their proper place in our various findings.

Let $\nu$ be a Borel probability measure on $C(\mathcal{D}; [0, \infty))$ that represents our prior probabilities on the unknown attenuation function. In other words, if $A \subset C(\mathcal{D}; [0, \infty))$ is Borel-measurable, then the probability that the attenuation function is in $A$, prior to making any scans, is given by $\nu(A)$. We would like to understand our posterior probabilities. We will treat the stand-in for our attenuation function as a random variable taking values in the space $C(\mathcal{D}; [0, \infty))$ to model our lack of complete knowledge of the body being scanned.

We would like to conclude by succinctly coalescing the probability space on which our random variables are defined. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space on which the following random elements are defined:

(a) white noise $W$ on $Z$, as described above,

(b) $\Phi : \Omega \to C(\mathcal{D}; [0, \infty))$, a $(\mathcal{F}, \mathcal{B}_{C(\mathcal{D})})$-measurable function with $P^{-1}\Phi = \nu$, and

(c) $\{U^y_\ell : \ell \in \mathbb{N}, y \in Z\}$, which are i.i.d. Uniform$(0, 1)$ random variables.

We further assume that all of the above are independent. We have taken $P\Phi^{-1} = \nu$. That is, $\Phi$ is a $C(\mathcal{D}; [0, \infty))$-valued random variable with law $\nu$. From Lemma A.0.25 and Lemma A.0.26 it follows that $\Phi$ can also be thought of as an $L^\infty(D, \mathcal{B}_D, \lambda)$-valued random variable, as well as an $L^2(D, \mathcal{B}_D, \lambda)$-valued random variable. In particular, this implies that $X\Phi$ is an $L^2(Z, \mathcal{B}_Z, \mu)$-valued random variable by the continuity of $X$, as shown in A.0.28.
For $f \in C(\overline{D}; [0, \infty))$, $y \in \mathbb{Z}$, and $x \in (0, 1)$, let

$$\xi^y(x, f) = 1_B(x, e^{-xf(y)}),$$

where $B = \{(s, t) : s < t\}$, and let

$$\xi^y f = \xi^y(U^y, f) \sim \text{Bernoulli}(e^{-xf(y)}).$$

If we consider a body with attenuation function $f$, and a sequence of X-ray photons penetrating the body along the line $L_y$, then the event $\{\xi^y f = 1\}$ represents the event that the $\ell$-th photon in that sequence emerges from the body without being absorbed.
CHAPTER 3: ANALYSIS ON LIMIT MODEL

In this chapter, our focus will be on outlining what the information from an X-ray scan tells us about the body under observation. Namely, we would like to characterize the posterior probabilities for various potential observations. We shall take these observations to be the stochastic process $Y_{u,\varepsilon} = \{Y_{u,\varepsilon}(g) : g \in L^2(Z)\}$ given by

$$Y_{u,\varepsilon}(g) = \langle X\Phi, g \rangle + \varepsilon \int_{Z} (u^*X\Phi)g \, dW,$$

for differing values of $\varepsilon$ and functions $u^*$ such that $u : \mathbb{R} \to \mathbb{R}$ is continuous and we define $u^* : C(Z) \to C(Z)$ by $u^*f = u(f(\cdot))$, and note that $u^*$ is continuous as given by Lemma A.0.27. This is the view of X-ray data when taking into account the effects of white noise as implied by the central limit theorem as alluded to in the introduction by Equation 1.2.

Special care should be taken in the use of random variables $\Phi$ and $W$ in unison. The stochastic process $Y_{u,\varepsilon}$ can only be called such if, for every $g \in L^2(Z)$, the function $Y_{u,\varepsilon}(g)$ is a random variable. In other words, we require $\omega \mapsto Y_{u,\varepsilon}(g, \omega)$ must be $(\mathcal{F}, \mathcal{B}_{\mathbb{R}})$-measurable. We only have this result because we constructed our white noise to be jointly measurable itself and independent of $\Phi$. Therefore, we indeed have that the mapping

$$\omega \mapsto Y_{u,\varepsilon}(g, \omega) = \langle X\Phi(\omega), g \rangle + \varepsilon W(u^*X\Phi(\omega)g, \omega)$$

is measurable. In the next chapter, we give explicit formulation and justification of this model, previously absent from the literature. That chapter will also demonstrate that the law of large numbers version of these observations is analogous to taking $\varepsilon = 0$ and directly observing the X-ray transform of our unknown attenuation function. Also in the final chapter, we will demonstrate
that by use of the central limit theorem (and hence requiring less radiation), we can explicitly define both the diminishing scaling factor of our noise term, \( \varepsilon \), and the function \( u \) that determines its variance.

Preliminaries on \( Y_{1,\varepsilon} \)

For now, we wish to determine the conditional law of \( X\Phi \) given our observations. As a first step, we will condition on the simplified process, \( Y_{1,\varepsilon} = X\Phi + \varepsilon W \) to simplify calculations at first and give us an indication how our posterior weights are modified by scan data. The process \( Y_{1,\varepsilon} \) has also been used to study the phenomenon of scattering in tomographic imaging. See, for example, [10, 9, 12].

Recall that the general \( C(\overline{D}; [0, \infty)) \)-valued random variable \( \Phi \) presented in Chapter 2 represents the lack of knowledge of the attenuation function being scanned. Let \( T = \{f_j\} \subset C(\overline{D}) \) be a fixed, finite set and we shall assume \( P(\Phi = f_j) = p_j \) with \( \sum_{j=1}^{K} p_j = 1 \). For this chapter, we note that the assumptions on this random variable taking values in \( C(\overline{D}) \) are not explicitly necessary and it would suffice for it to take values in \( L^2(D) \). Thus the set \( \{f_j\} \subset L^2(D) \) comprises all the possible forms which the attenuation function can take. The primary result for this chapter follows.

**Theorem 3.0.1.** Assume \( Y_{1,\varepsilon} \) and \( \Phi \) are as defined previously and let

\[
\beta_j = P(\Phi = f_j) \exp \left( \frac{2Y_{1,\varepsilon}(Xf_j) - \|Xf_j\|_{L^2(Z)}^2}{2\varepsilon^2} \right).
\]

Then

\[
P(\Phi = f_j \mid Y_{1,\varepsilon}) = \frac{\beta_j}{\sum_{k=1}^{K} \beta_k} \text{ a.s.}
\]

In plainer terms, one can consider \( \beta_j \) to be the modified weight we assign to the possibility of our
body function finding itself in the state $f_j$ only after we observe $Y_{1,\varepsilon}$. Before this, the weights were simply

$$P(\Phi = f_j) = \frac{p_j}{\sum_{k=1}^{K} p_k} = p_j,$$

since $\sum_{k=1}^{K} p_k = 1$, as the function must be in one of these states. Thus the set $\{\beta_j\}$ represents the update of our confidence in which state the body is in after we’ve gotten some empirical data whereas $\{p_j\}$ were our priors before any scans, presumably based on historical information. To prove Theorem 3.0.1 we will make use of Propositions 3.0.3 and 3.0.8, proven in subsequent sections.

$H$ Subspace Projection

Recall that $H = L^2(Z, B_Z, \mu)$ is a real Hilbert space. Henceforth in this chapter we will find it useful to use an orthonormal basis $\{h_i\}_{i=1}^{\infty}$ of $H$, denote $H_n = \text{span}\{h_1, \ldots, h_n\}$ and let $Q_n$ denote the orthogonal projection of $H$ onto $H_n$ as previously defined in our section on white noise, so that

$$Q_n g = \sum_{i=1}^{n} \langle g, h_i \rangle h_i,$$

for all $g \in H$. Recall that $X\Phi$, the X-ray transform of the body function we wish to recover, is assumed to take values in $H$. Thus this architecture lets us work with a simplified depiction of $X\Phi$ and we will then translate our results by letting $n \to \infty$.

We will often abuse notation, for the sake of both the reader and ourselves, to write $Q_n W = \sum_{i=1}^{n} W(h_i)h_i$ and $Q_n Y_{1,\varepsilon} = \sum_{i=1}^{n} Y_{1,\varepsilon}(h_i)h_i$, even though neither $W$ nor $Y_{1,\varepsilon}$ are $L^2(Z)$ functions. Rather, they take such elements as inputs and so these projections are $H_n$-valued random variables. Note that $Q_n Y_{1,\varepsilon} = Q_n R f + \varepsilon Q_n W$.

**Lemma 3.0.2.** For any $g \in H$, we have $\langle Q_n Y_{1,\varepsilon}, g \rangle = Y_{1,\varepsilon}(Q_n g)$ and $\langle Q_n W, g \rangle = W(Q_n g)$. 

15
Proof. Note that

\[ \langle Q_n Y_{1,\varepsilon}, g \rangle = \left\langle \sum_{i=1}^{n} Y_{1,\varepsilon}(h_i) h_i, \sum_{j=1}^{\infty} \langle g, h_j \rangle h_j \right\rangle \]

\[ = \sum_{i=1}^{n} \sum_{j=1}^{\infty} Y_{1,\varepsilon}(h_i) \langle g, h_j \rangle \delta_{i,j} \]

\[ = \sum_{i=1}^{n} Y_{1,\varepsilon}(h_i) \langle g, h_i \rangle \]

\[ = Y_{1,\varepsilon}\left(\sum_{i=1}^{n} \langle g, h_i \rangle h_i\right) \]

\[ = Y_{1,\varepsilon}(Q_n g), \]

for any \( g \in H \). A similar calculation would follow from working with \( Q_n W \).

Thus we can work on the subspaces of our Hilbert space \( L^2(Z) \) while still translating the behavior of our defined stochastic processes. We shall proceed to outline our discoveries in these subspaces. We will also abuse notation for a time and take \( Y \equiv Y_{1,\varepsilon} \) until otherwise stated to ease notation.

Truncated Sigma Algebras

In probability, the knowledge one is said to have about a system is contained in \( \sigma \)-algebras. This section will exploit the workings of how greater knowledge of the body modifies our conditional probabilities. In particular, we shall limit ourselves to only the information that can be gathered from our subspace projections and ensure that we recover complete information in the limit. We now proceed with our first major proposition.

**Proposition 3.0.3.** Let \( F_n = \sigma(Y(h_1), \ldots, Y(h_n)) \). Then

\[ P(\Phi = f_j \mid F_n) \rightarrow P(\Phi = f_j \mid Y), \]
a.s. and in $L^1(\Omega)$.

This serves as vindication for our choice to work in the subspaces $H_n$, since we may then transcribe our findings to the space we desired. It may be helpful to think of the $\sigma$-algebra $\mathcal{F}_n$ defined in the proposition as the environment of all lower-order observations wherein $X\Phi$ is taken to assume values in $H_n$. The proof of Proposition 3.0.3 requires the following, which is Theorem 5.5.7 in [11].

**Theorem 3.0.4.** Let $\mathcal{F}_\infty = \sigma(\bigcup_{n=1}^{\infty} \mathcal{F}_n)$. Then for all $X \in L^1(\Omega)$, we have $E[X \mid \mathcal{F}_n] \to E[X \mid \mathcal{F}_\infty]$ in $L^1(\Omega)$ and a.s.

If we can provide the following result of almost sure convergence, then our task is nearly complete.

**Lemma 3.0.5.** For all $X \in L^1(\Omega)$, we have $E[X \mid \mathcal{F}_\infty] = E[X \mid Y]$ a.s.

**Proof.** Let $Z := E[X \mid \mathcal{F}_\infty]$ and $\sigma(Y)$ be the smallest $\sigma$-algebra such that $Y(g)$ is measurable for all $g \in L^2(Z)$. To show that $Z = E[X \mid Y]$ a.s., we must show that $Z$ is $\sigma(Y)$-measurable and that $E[Z1_B] = E[X1_B]$ for all $B \in \sigma(Y)$. By its definition, $Z$ is $\mathcal{F}_\infty$-measurable. Since $Y(h_j)$ is $\sigma(Y)$-measurable for all $j$, it follows that $\mathcal{F}_n \subset \sigma(Y)$ for all $n$, and therefore, $\mathcal{F}_\infty \subset \sigma(Y)$. This establishes that $Z \in \sigma(Y)$. Next we let

$$\mathcal{L} = \{B \in \sigma(Y) : E[X1_B] = E[Z1_B]\}.$$

We would like to show that $\sigma(Y) \subset \mathcal{L}$. To this end, we will avail ourselves of Theorem A.0.29, referred to as the $\pi$-$\lambda$ theorem. The set $\mathcal{L}$ is a $\lambda$-system if the following are true:

(i) $\Omega \in \mathcal{L}$;

(ii) if $A_1, A_2 \in \mathcal{L}$ and $A_1 \subset A_2$, then $A_2 \cap A_1^c \in \mathcal{L}$;
(iii) and if $A_n \in \mathcal{L}$ with $A_n \uparrow A$, then $A \in \mathcal{L}$.

The first point, $E[X] = E[E[X \mid \mathcal{F}_\infty]] = E[Z]$, follows from nested expectations. Second, for any two sets $A_1, A_2 \in \mathcal{L}$ such that $A_1 \subset A_2$, we have that $E[X1_{A_1}] = E[Z1_{A_1}]$ and $E[X1_{A_2}] = E[Z1_{A_2}]$. Let $A' = A_2 \cap A_1^c$. It follows, therefore, that

$$E[X1_{A'}] = E[X1_{A_2} - X1_{A_1}]$$
$$= E[X1_{A_2}] - E[X1_{A_1}]$$
$$= E[Z1_{A_2}] - E[Z1_{A_1}]$$
$$= E[Z1_{A_2} - Z1_{A_1}]$$
$$= E[Z1_{A'}],$$

and we need now only show the third point. For $A_n \in \mathcal{L}$ with $A_n \uparrow A$, we see that $E[X1_{A_n}] = E[Z1_{A_n}]$ for all $n \in \mathbb{N}$. Then,

$$E[X1_A] = \int_\Omega X1_A \, dP$$
$$= E\left[ \lim_{n \to \infty} X1_{A_n} \right]$$
$$= \lim_{n \to \infty} E\left[ X1_{A_n} \right]$$
$$= \lim_{n \to \infty} E\left[ Z1_{A_n} \right]$$
$$= E\left[ \lim_{n \to \infty} Z1_{A_n} \right]$$
$$= E[Z1_A],$$

where each time we pass the limit through the expectation we are making use of the dominated convergence theorem by the fact that $E|X| < \infty$ and therefore $E|Z| = E|E[X \mid \mathcal{F}_\infty]| \leq E[E(|X| \mid \mathcal{F}_\infty)] = E|X| < \infty$. So we indeed have a $\lambda$-system in $\mathcal{L}$. 

18
We now define $\mathcal{P}$ as the collection of sets of the form $C = \bigcap_{k=1}^{n} \{ Y(g_k) \in A_k \}$ for some $n \in \mathbb{N}$, $g_1, \ldots, g_n \in L^2(Z)$, and $A_1, \ldots, A_n \in \mathcal{B}_\mathbb{R}$. Fix an arbitrary $\tilde{g} \in L^2(Z)$. Note that all sets of the form $\{ Y(g) \in A \}$ for some $A \in \mathcal{B}_\mathbb{R}$ are contained in $\mathcal{P}$. Therefore, $Y(\tilde{g})$ is $\sigma(\mathcal{P})$-measurable. Since $\tilde{g}$ was arbitrary, it follows that $Y(g)$ is $\sigma(\mathcal{P})$-measurable for all $g \in L^2(Z)$. Since we defined $\sigma(Y)$ to be the smallest such $\sigma$-algebra, we have established that $\sigma(Y) \subset \sigma(\mathcal{P})$. We proceed by noting that

$$Y(Q_n g_k) = \sum_{j=1}^{n} \langle g_k, h_j \rangle Y(h_j) \in \mathcal{F}_n \subset \mathcal{F}_\infty.$$ 

By Lemma A.0.41, we know that $Y(Q_n g_k) \to Y(g_k)$ a.s., so there must exist an $\mathcal{F}_\infty$-measurable random variable $\zeta_k$ such that $Y(g_k) = \zeta_k$ a.s. Therefore, for $C \in \mathcal{P}$,

$$E[Z1_{C}] = E \left[ E[X|\mathcal{F}_\infty] \prod_{k=1}^{n} 1_{A_k}(Y(g_k)) \right]$$

$$= E \left[ E[X|\mathcal{F}_\infty] \prod_{k=1}^{n} 1_{A_k}(\zeta_k) \right]$$

$$= E \left[ E[X \prod_{k=1}^{n} 1_{A_k}(\zeta_k) | \mathcal{F}_\infty] \right]$$

$$= E \left[ X \prod_{k=1}^{n} 1_{A_k}(\zeta_k) \right]$$

$$= E \left[ X1_{C} \right],$$

and so $C \in \mathcal{L}$ and by the $\pi$-$\lambda$ theorem we have established that $\sigma(Y) \subset \mathcal{L}$. This gives us that $E[X|\mathcal{F}_\infty] = E[X|\sigma(Y)]$ a.s. \qed

We can now proceed with proving Proposition 3.0.3 in short order.

\textit{Proof of Proposition 3.0.3.} We note that from Theorem 3.0.4, we have that $P(\Phi = f_j | \mathcal{F}_n) \to P(\Phi = f_j | \mathcal{F}_\infty)$ a.s. Moreover, the result that immediately precedes this proof, Lemma 3.0.5,
gives us that \( P(\Phi = f_j \mid \mathcal{F}_n) = P(\Phi = f_j \mid Y) \) a.s. It follows immediately, therefore, that

\[
P(\Phi = f_j \mid \mathcal{F}_n) \to P(\Phi = f_j \mid Y) \quad \text{a.s.}
\]
as \( n \to \infty \).

We will also have to establish the following.

**Lemma 3.0.6.** Let \( g_n, g \in L^2(Z) \) and suppose \( g_n \to g \) in \( L^2(Z) \). Then \( Y(g_n) \to Y(g) \) in \( L^2(\Omega) \).

*In other words, the stochastic process \( Y \) is a continuous operator from \( L^2(Z) \) to \( L^2(\Omega) \).*

**Proof.** Let \( g_n \to g \) in \( L^2(Z) \). Note that we have already shown white noise to be continuous, so our task is reduced to demonstrating that the operator \( X \Phi \) which acts on elements of \( L^2(Z) \) by \( g \mapsto \langle X \Phi, g \rangle \) is continuous from \( L^2(Z) \) to \( L^2(\Omega) \). We first write

\[
E[\langle X \Phi, g \rangle^2] = \sum_k \langle X f_k, g \rangle^2 P(f = f_k)
\]

\[
= \sum_k \langle X f_k, g \rangle^2 p_k
\]

\[
\leq \|g\|^2 \sum_k \|X f_k\|^2 p_k,
\]

which is finite by assumption. Thus we know \( X \Phi \) is indeed a mapping from \( L^2(Z) \) to \( L^2(\Omega) \). Next
we need to show $E|\langle X\Phi, g_n \rangle - \langle X\Phi, g \rangle|^2 \to 0$ as $n \to \infty$. For this, we calculate

$$
\lim_{n \to \infty} E|\langle X\Phi, g_n \rangle - \langle X\Phi, g \rangle|^2 = \lim_{n \to \infty} E|\langle X\Phi, g_n - g \rangle|^2 \\
\leq \lim_{n \to \infty} E[\|X\Phi\|^2 \|g_n - g\|^2] \\
= \lim_{n \to \infty} \|g_n - g\|^2 E[\|X\Phi\|^2] \\
= E[\|X\Phi\|^2] \lim_{n \to \infty} \|g_n - g\|^2 \\
= 0,
$$

where we have applied Cauchy-Schwartz inside the expectation.

We finish this section with the following quick lemma to establish that the result is indeed what we want of our observation projected onto $H_n$.

**Lemma 3.0.7.** For all $n$, we have $\mathcal{F}_n = \sigma(Q_n Y)$.

**Proof.** By definition, $\mathcal{F}_n = \sigma(Y(h_1), \ldots, Y(h_n))$. Also, $Q_n Y = \sum_{i=1}^{n} Y(h_i) h_i$. Thus, $Q_n Y \in \mathcal{F}_n$, so $\sigma(Q_n Y) \subset \mathcal{F}_n$. Conversely, if $j \in \{1, \ldots, n\}$, then $Y(h_j) = \langle Q_n Y, h_j \rangle \in \sigma(Q_n Y)$, so $\mathcal{F}_n \subset \sigma(Q_n Y)$.

Therefore, at least in the subspaces, we have that the information contained in each $\mathcal{F}_n$ agrees with the data from our X-ray scans projected onto $H_n$. Moreover, although the $\sigma$-algebras may prove to differ in the limit, the expectations conditioned on these limits is given to agree by Lemma 3.0.5.
Result in Subspace

The objective of this section is to delineate the result we find working in the subspace $H_n \subset H$ in order to extend it in the limiting case. To this end, we seek to show the form of our posterior probabilities on $\Phi$ after observing the X-ray data through the lens of our projections.

**Proposition 3.0.8.** Let

$$
\beta_{j,n} = p_j \exp \left( \frac{2Y(\mathcal{Q}_n X f_j) - \|\mathcal{Q}_n X f_j\|^2}{2\varepsilon^2} \right),
$$

Then

$$
P(\Phi = f_j \mid \mathcal{F}_n) = \frac{\beta_{j,n}}{\sum_{k=1}^K \beta_{k,n}} \text{ a.s.,}
$$

where $\mathcal{F}_n$ is as in Proposition 3.0.3.

The reader will note that this is a direct analog of Theorem 3.0.1 in the projection space $H_n$. By working with the simplified case first, we are rewarded with a hint of what we might expect in the full space of $H$. After the work is all done in simpler terms, we strive to ensure if and what can be said as we let $n \to \infty$ and move from finite projections of our observed data to the full continuum of our observations (i.e. $\mathcal{Q}_n Y \to Y$).

Before we can proceed to prove Proposition 3.0.8, we will require Theorem 3.0.11, in which we will use the mapping $\iota : H_n \to \mathbb{R}^n$, which associates an element $g \in H_n$ with its component vector $\iota g = (\langle g, h_1 \rangle, \ldots, \langle g, h_n \rangle) \in \mathbb{R}^n$. While not entirely necessary, it brings the arithmetic into the more familiar space of real numbers. We note that this mapping is a linear isometry and therefore all our results can be neatly transported back to the function space $H_n$. 

22
Lemma 3.0.9. The random vector

\[ y_{j,n} = (\langle X f_j, h_1 \rangle + \varepsilon W(h_1), \ldots, \langle X f_j, h_n \rangle + \varepsilon W(h_n)) \]

has a density with respect to Lebesgue measure on \( \mathbb{R}^n \) given by

\[ y \mapsto \frac{1}{(2\pi)^{n/2}} \exp \left( -\frac{1}{2\varepsilon^2} |y - \iota Q_n X f_j|^2 \right). \]

Proof. If a multivariate Gaussian \( X = (X_1, \ldots, X_n) \) has mean vector \( \mu \) and covariance matrix \( A \), where \( A \) is invertible, then \( X \) has a density with respect to Lebesgue measure on \( \mathbb{R}^n \) given by

\[ \frac{1}{\sqrt{(2\pi)^n |A|}} \exp \left( -\frac{1}{2} (x - \mu)^T A^{-1} (x - \mu) \right), \]

where \( |A| \) is a matrix determinant. (See, for example, Equation (16.5) in [8].) For two random variables \( U, V \), we define their covariance by

\[ \text{Cov}(U, V) = E[(U - EU)(V - EV)]. \]

A covariance matrix \( A \) for a random vector \( X = (X_1, \ldots, X_n) \) is then given by the symmetric entries \( a_{i,j} = \text{Cov}(X_i, X_j) \).

In our specific case, we calculate the mean vector of \( y_{j,n} \) by noting that \( E[W(f)] = 0 \) for all \( f \in L^2(Z) \) and \( \langle Rf_j, h_i \rangle \) are constants in our probability space for all \( j, i \in \mathbb{N} \). Therefore,

\[
E[(\langle X f_j, h_1 \rangle + \varepsilon W(h_1), \ldots, \langle X f_j, h_n \rangle + \varepsilon W(h_n))] \\
= (\langle X f_j, h_1 \rangle, \ldots, \langle X f_j, h_n \rangle) \\
= \iota Q_n X f_j.
\]
We’ll also require the covariance matrix, which we will denote \( \Sigma \). Its entries are calculated as follows:

\[
\sigma_{i,k} = E[(\varepsilon W(h_i) + \langle Xf_j, h_i \rangle) - E[\varepsilon W(h_i) + \langle Xf_j, h_i \rangle] \times (\varepsilon W(h_k) - \langle Xf_j, h_k \rangle - E[\varepsilon W(h_k) + \langle Xf_j, h_k \rangle])]
\]

\[
= E[(\varepsilon W(h_i) + \langle Xf_j, h_i \rangle - \langle Xf_j, h_i \rangle)(\varepsilon W(h_k) - \langle Xf_j, h_k \rangle - \langle Xf_j, h_k \rangle)]
\]

\[
= E[(\varepsilon W(h_i))(\varepsilon W(h_k))]
\]

\[
= \varepsilon^2 \langle h_i, h_k \rangle
\]

\[
= \varepsilon^2 \delta_{ik},
\]

where \( \delta_{ik} \) is the Kronecker delta.

This shows that \( \Sigma = \varepsilon^2 I \), where \( I \) is the \( n \)-dimensional identity matrix. Hence, \( \Sigma \) is invertible, with \( |\Sigma| = |\varepsilon^2 I| = \varepsilon^2 |I| = \varepsilon^2 \) and \( \Sigma^{-1} = (\varepsilon^2 I)^{-1} = \varepsilon^{-2} I \). By using these in the formula provided, we obtain the desired result. \( \square \)

**Lemma 3.0.10.** Let \( \lambda \) denote Lebesgue measure on \( \mathbb{R}^n \) and define the measure \( \mu_n \) on \( (H_n, B(H_n)) \) by \( \mu_n(A) = \lambda(tA) \). If \( \varphi : \mathbb{R}^n \to \mathbb{R} \) is bounded and measurable, then

\[
\int_{\mathbb{R}^n} \varphi(y) \, d\lambda(y) = \int_{H_n} \varphi(t \cdot h) \, d\mu_n(h).
\]

**Proof.** We begin with the case that \( \varphi \) is the indicator function for a measurable set, so \( \varphi = 1_A \) for
Therefore

\[ \int_{H_n} \varphi(\theta h) \, d\mu_n(h) = \int_{H_n} 1_A(\theta h) \, d\mu_n(h) \]
\[ = \mu_n(\{h \in H^n : \theta h \in A\}) \]
\[ = \mu_n(\iota^{-1} A) \]
\[ = \lambda(A) \]
\[ = \int_{\mathbb{R}^n} \varphi(y) \, d\lambda(y), \]

so we have the result for indicator functions. No we take the case where \( \varphi \) is a simple function, so \( \varphi = \sum_{j=1}^{n} c_j 1_{A_j} \) for \( \{c_j\} \subset \mathbb{R} \) and \( \{A_j\} \subset \mathcal{B}(\mathbb{R}^n) \). We appeal to the linearity of integrals and the result for indicator functions to demonstrate,

\[ \int_{H_n} \varphi(\theta h) \, d\mu_n(h) = \int_{H_n} \sum_{j=1}^{n} c_j 1_{A_j}(\theta h) \, d\mu_n(h) \]
\[ = \sum_{j=1}^{n} c_j \int_{H_n} 1_{A_j}(\theta h) \, d\mu_n(h) \]
\[ = \sum_{j=1}^{n} c_j \lambda(A_j) \]
\[ = \int_{\mathbb{R}^n} \varphi(y) \, d\lambda(y). \]

One step further, we assume \( \varphi \) to be a nonnegative function and let \( \{\varphi_n\} \) be a sequence of nonnegative simple functions such that \( \varphi_n \uparrow \varphi \) pointwise and therefore \( \varphi_n \circ \iota \uparrow \varphi \circ \iota \) pointwise. Thus we
already know the result holds for each $\varphi_n$. By the monotone convergence theorem, it follows that

$$
\int_{H_n} \varphi(\gamma h) \, d\mu_n(h) = \lim_{n \to \infty} \int_{H_n} \varphi_n(\gamma h) \, d\mu_n(h)
= \lim_{n \to \infty} \int_{\mathbb{R}^n} \varphi_n(y) \, d\lambda(y)
= \int_{\mathbb{R}^n} \varphi(y) \, d\lambda(y).
$$

Our final step becomes to demonstrate the result for integrable $\varphi$. We note that in such a case, we may decompose $\varphi = \varphi^+ - \varphi^-$, where $\varphi^+$ and $\varphi^-$ are nonnegative functions corresponding to the positive and negative parts of $\varphi$ respectively. Our previous result gives us the result for such nonnegative functions, so we have that

$$
\int_{H_n} \varphi(\gamma h) \, d\mu_n(h) = \int_{H_n} \varphi^+(\gamma h) \, d\mu_n(h) - \int_{H_n} \varphi^-(\gamma h) \, d\mu_n(h)
= \int_{\mathbb{R}^n} \varphi^+(y) \, d\lambda(y) - \int_{\mathbb{R}^n} \varphi^-(y) \, d\lambda(y)
= \int_{\mathbb{R}^n} \varphi(y) \, d\lambda(y),
$$

and our proof is concluded. \qed

We can now proceed with the motivation for the previous two lemmas.

**Theorem 3.0.11.** We remind the reader that $T$ is the range of $\Phi$. For all bounded and measurable $\psi : T \times H_n \to \mathbb{R}$, we have

$$
E[\psi(\Phi, Q_nY)] = \sum_k \int_{H_n} \psi(f_k, h) \phi(f_k, h) \, d\mu_n(h),
$$
where \( \phi : T \times H_n \to \mathbb{R} \) is given by

\[
\phi(f_k, h) = \frac{p_k}{(2\pi)^{n/2} \varepsilon} \exp \left( -\frac{1}{2\varepsilon^2 \| h - Q_nXf_k \|^2 \right),
\]

and \( \mu_n \) is the measure on \( (H_n, \mathcal{B}(H_n)) \) defined by \( \mu_n(A) = \lambda(A) \).

Here one may think of \( \psi \) as any operator we wish to act on our body function and observed data in full generality. We value the generality in finding \( \phi \), which is a density function assigning weight to the different combinations of \( \Phi \) and \( Y \) as we run through both the possible body states \( \{f_k\} \) in the sum and the various forms of \( Q_nY \in H_n \) in the integral. The proof follows.

**Proof.** It is our wish to appeal to Theorem A.0.37 and the result of Lemma 3.0.9, which exploit properties of conditional expectation. These tools require use of independent random variables, but \( \Phi \) and \( Q_nY \) are clearly not independent since \( Y \) is the observation of \( \Phi \) through an X-ray transform with white noise. So we first look to a \( \psi : T \times H_n \to \mathbb{R} \) that is bounded and measurable and attempt to manipulate \( E[\psi(f, Q_nY)] \) to our ends. For this, we construct the map \( \eta : T \times H_n \to T \times H_n \), which acts via \( \eta(g, h) = (g, Q_nXg + h) \), and designate \( \tilde{\psi} = \psi \circ \eta : T \times H_n \to \mathbb{R} \). The result being

\[
\tilde{\psi}(f, \varepsilon Q_nW) = \psi(\eta(f, \varepsilon Q_nW)) = \psi(f, Q_nXf + \varepsilon Q_nW) = \psi(f, Q_nY),
\]

essentially identifying a function of two dependent variables with a function of the independent \( \Phi \) and \( Q_nW \). This allows us to use Theorem A.0.37 in calculating

\[
E[\psi(\Phi, Q_nY) \mid \Phi] = E[\tilde{\psi}(\Phi, \varepsilon Q_nW) \mid \Phi] = \xi(\Phi),
\]
where

\[ \xi(f_j) = E[\tilde{\psi}(f_j, \varepsilon \mathcal{Q}_n W)] = E[\psi(f_j, \mathcal{Q}_n Xf_j + \varepsilon \mathcal{Q}_n W)] = E[\psi(f_j, \iota^{-1} y_{j,n})], \]

and where

\[ y_{j,n} = \iota \mathcal{Q}_n Xf_j + \varepsilon \iota \mathcal{Q}_n W \]
\[ = \langle Xf_j, h_1 \rangle, \ldots, \langle Xf_j, h_n \rangle + \varepsilon (W(h_1), \ldots, W(h_n)) \]
\[ = \langle Xf_j, h_1 \rangle + \varepsilon W(h_1), \ldots, \langle Xf_j, h_n \rangle + \varepsilon W(h_n) \]

is the random vector defined in the statement of Lemma 3.0.9. It may help to think of this as the low-order counterpart to \( Y \), which we’ve brought into \( \mathbb{R}^n \) for the sake of simplicity in calculation.

Thus, by the aforementioned lemma, we have

\[ \xi(f_j) = \frac{1}{(2\pi)^{n/2}\varepsilon} \int_{\mathbb{R}^n} \psi(f_j, \iota^{-1} y) \exp \left( -\frac{1}{2\varepsilon^2} |y - \iota \mathcal{Q}_n Xf_j|^2 \right) d\lambda(y). \]

By Lemma 3.0.10, which establishes the validity of our change of variables, we have

\[ \xi(f_j) = \frac{1}{(2\pi)^{n/2}\varepsilon} \int_{\mathbb{H}_n} \psi(f_j, h) \exp \left( -\frac{1}{2\varepsilon^2} |h - \iota \mathcal{Q}_n Xf_j|^2 \right) d\mu_n(h). \]

which, since \( \iota \) is an isometry, allows us to operate in terms of the \( L^2(Z) \) norm,

\[ \xi(f_j) = \frac{1}{(2\pi)^{n/2}\varepsilon} \int_{\mathbb{H}_n} \psi(f_j, h) \exp \left( -\frac{1}{2\varepsilon^2} \|h - \mathcal{Q}_n Xf_j\|^2 \right) d\mu_n(h). \]

Finally, with the proper mass of each \( \xi(f_j) = E[\tilde{\psi}(f_j, \varepsilon \mathcal{Q}_n W)] \) for \( 1 \leq j \leq K \), we use our
understanding of discrete random variables to write

\[ E[\psi(\Phi, Q_nY)] = E[E[\psi(\Phi, Q_nY)|Y]] = E[\xi(\Phi)] = \sum_k p_k \xi(f_k) = \sum_k \int_{H_n} \psi(f_k, h) \phi(f_k, h) d\mu_n(h), \]

a density in the form of a sum of integrals over the possible forms of \( \Phi \) weighted by \( \phi \).

Equipped as we are with the proper density function, we proceed to prove the section’s opening proposition. Any allusions to the appendix are made to streamline the proof.

**Proof of Proposition 3.0.8.** Let

\[ \alpha_{j,n} = p_j \exp \left( -\frac{\|Q_nY - Q_nXf_j\|^2}{2\varepsilon^2} \right). \]

By Lemma 3.0.2 and the fact that \( Q_n^2 = Q_n \), we have

\[
\|Q_nY - Q_nXf_j\|^2 = \|Q_nY\|^2 - 2 \langle Q_nY, Q_nXf_j \rangle + \|Q_nXf_j\|^2 \\
= \|Q_nY\|^2 - 2Y(Q_nXf_j) + \|Q_nXf_j\|^2,
\]

as separation of terms that depend on \( f_j \) and those that do not, which we will find useful. Therefore,

\[
\alpha_{j,n} = p_j \exp \left( -\frac{\|Q_nY - Q_nXf_j\|^2}{2\varepsilon^2} \right) \\
= p_j \exp \left( -\frac{\|Q_nY\|^2 - 2Y(Q_nXf_j) + \|Q_nXf_j\|^2}{2\varepsilon^2} \right) \\
= p_j \exp \left( \frac{2Y(Q_nXf_j) - \|Q_nXf_j\|^2}{2\varepsilon^2} \right) \exp \left( -\frac{\|Q_nY\|^2}{2\varepsilon^2} \right) \\
= \beta_{j,n} e^{-|Q_nY|^2/2\varepsilon^2}.
\]
Since the factor that distinguishes the two is not itself indexed by \( j \), in this case,

\[
\frac{\alpha_{j,n}}{\sum_{k=1}^{K} \alpha_{k,n}} = \frac{\beta_{j,n}}{\sum_{k=1}^{K} \beta_{k,n}} \quad \text{a.s.}
\]

It therefore suffices to prove that

\[
P(\Phi = f_j \mid \mathcal{F}_n) = \frac{\alpha_{j,n}}{\sum_{k=1}^{K} \alpha_{k,n}}.
\] (3.1)

Now we appeal to Lemma 3.0.7, which established the agreement of our \( \sigma \)-algebras. We also make use of Theorem 3.0.11, and Corollary A.0.36 to conclude,

\[
P(\Phi = f_j \mid \mathcal{F}_n) = P(\Phi = f_j \mid Q_nY) = \sum_k \frac{p_k}{(2\pi)^{n/2} \varepsilon} \exp \left( -\frac{1}{2\varepsilon^2} \| Q_nY - Q_nXf_j \|^2 \right)
\]

\[
= \sum_k \frac{p_k}{(2\pi)^{n/2} \varepsilon} \exp \left( -\frac{1}{2\varepsilon^2} \| Q_nY - Q_nXf_j \|^2 \right)
\]

\[
= \frac{p_j \exp \left( -\frac{1}{2\varepsilon^2} \| Q_nY - Q_nXf_j \|^2 \right)}{\sum_k P_k \exp \left( -\frac{1}{2\varepsilon^2} \| Q_nY - Q_nXf_j \|^2 \right)}
\]

\[
= \frac{\alpha_{j,n}}{\sum_k \alpha_{k,n}}.
\]

This allows us to finally state \( P(\Phi = f_j \mid \mathcal{F}_n) = \frac{\beta_{j,n}}{\sum_{k=1}^{K} \beta_{k,n}} \quad \text{a.s.} \)

We remind the reader that our main goal throughout the paper is to recover the conditional distribution of our body function after the introduction of some observed data. Now achieved through projections onto subspaces \( H_n \), all we need do is discover what we can state in the full generality of \( H \) by letting \( n \to \infty \).
Result in $L^2(\mathbb{Z})$

With the architecture we now have in place, our desired result is near at hand. We merely require one short lemma before we end with the proof of Theorem 3.0.1 and cement our conclusions on the special case of assuming white noise to modify our observations in an additive manner and that $\Phi$ is simple.

**Lemma 3.0.12.** We claim that $\langle Q_n Y, Q_n g \rangle \to Y(g)$ in $L^2(\Omega)$, for any $g \in H$.

**Proof.** Let $g \in H$. Since $Q_n^2 = Q_n$, Lemma 3.0.2 gives $\langle Q_n Y, Q_n g \rangle = Y(\mathbb{Q}_n g)$. Furthermore, the fact that $Q_n g \to g$ in $H$ allows us to appeal to Lemma 3.0.6 and conclude $Y(Q_n g) \to Y(g)$ in $L^2(\Omega)$. \hfill $\square$

We are finally prepared to prove the main theorem for this chapter in describing the posterior weights on the form of $\Phi$ given X-ray data $Y_{1,\varepsilon}$.

**Proof of Theorem 3.0.1.** Consider the following random variables:

$$ P_{j,n} = \frac{\beta_{j,n}}{\sum_{k=1}^{K} \beta_{k,n}}, $$

and

$$ P_j = \frac{\beta_j}{\sum_{k=1}^{K} \beta_k}. $$

We will first show that $P_{j,n} \to P_j$ in probability as $n \to \infty$.

We will make use of Theorem 2.3.2 in [11], which states that a sequence of random variables converges in probability if and only if for every subsequence there exists a further subsequence that converges almost surely.
Let \( \{P_{j,n(m)}\} \) be an arbitrary subsequence. By Lemma 3.0.6, we have

\[
Y(Q_n X f_j) \to Y(X f_j)
\]

in \( L^2(\Omega) \) as \( n \to \infty \) for each \( j \in \{1, \ldots, K\} \). Since convergence in \( L^2(\Omega) \) implies convergence in probability, we have

\[
Y(Q_n X f_j) \to Y(X f_j)
\]

in probability as \( m \to \infty \) for each \( j \in \{1, \ldots, K\} \). By Lemma A.0.40, there exists a subsequence \( \{n(m_\ell)\} \) such that \( Y(Q_{n(m_\ell)} X f_j) \to Y(X f_j) \) a.s. as \( \ell \to \infty \) for all \( j \in \{1, 2, \ldots, K\} \). Thus,

\[
\beta_{j,n(m_\ell)} = p_j \exp \left( \frac{2Y(Q_{n(m_\ell)} X f_j) - \|Q_{n(m_\ell)} X f_j\|^2}{2\varepsilon^2} \right) \to p_j \exp \left( \frac{2Y(X f_j) - \|X f_j\|^2}{2\varepsilon^2} \right) = \beta_j \text{ a.s.}
\]

as \( \ell \to \infty \) for all \( 1 \leq j \leq N \). Hence,

\[
P_{j,n(m_\ell)} = \frac{\beta_{j,n(m_\ell)}}{\sum_{k=1}^K \beta_{k,n(m_\ell)}} \to \frac{\beta_j}{\sum_{k=1}^K \beta_k} = P_j \text{ a.s.}
\]

as \( \ell \to \infty \). By Theorem 2.3.2 in [11], this shows that \( P_{j,n} \to P_j \) in probability.

Finally, recall from Proposition 3.0.3 that \( P(\Phi = f_j \mid \mathcal{F}_n) \to P(\Phi = f_j \mid Y) \) a.s. as \( n \to \infty \), which certainly implies the same convergence in probability. On the other hand, we have just shown that \( P_{j,n} \to P_j \) in probability as \( n \to \infty \). Moreover, by Proposition 3.0.8, we have \( P(\Phi = f_j \mid \mathcal{F}_n) = P_{j,n} \) a.s. Therefore, by Lemma A.0.39, we have \( P(\Phi = f_j \mid Y) = P_j \) a.s., which is what we wanted to prove.

\[\square\]
We have concluded our calculation of posterior probabilities when we use the model \( Y_{1,\varepsilon} = X\Phi + \varepsilon W \) for our scans. We now have a closed form of \( P_j \) we would assign to each body state \( f_j \) after incorporating our observations.

Analysis on \( Y_{u,1} \) with Zero Mean

For this section, we will veer away from our previous scheme of observations being represented by a sum of ideal data and scaled white noise. We instead focus our attention to what might be done with viewing our observations as

\[
Y = u(X\Phi)W,
\]

which is elaborated upon in the modeling chapter with the function \( u : [0, \infty) \rightarrow \mathbb{R} \) defined by \( u(\xi) = -(e^\xi - 1)^{1/2} \). Informally, what is meant by \( Y = u(X\Phi)W \) is that for any function \( g \in L^2(Z) \) and \( \omega \in \Omega \),

\[
Y(g, \omega) = \langle Y(\omega), g \rangle_{L^2(Z)}
= \langle u(X\Phi(\omega))W(\omega), g \rangle
= \langle W(\omega), u(X\Phi(\omega))g \rangle,
\]

which is conditionally distributed as \( N(0, \|u(X\Phi)g\|^2) \) given \( \Phi \). This conditional distribution agrees with our past definition of white noise \( W \) as jointly measurable and independent of \( \Phi \). Rigorously speaking, what is meant by this notation is that,

\[
Y(g, \omega) = W(\omega, u(X\Phi(\omega))g) = \int_{Z} u(X\Phi(\omega))g \, dW.
\]

We will attempt to make use of much of the architecture in previous sections. Once more \( D, Z, \mu, \)
$X$, and $T = \{f_j\} \subset L^2(D)$ are as before with $P(\Phi = f_j) = p_j$ for $\sum_{j=1}^K p_j = 1$. We also recall the notation $H_n = \text{span}\{h_1, \ldots, h_n\}$ for a fixed orthonormal basis $\{h_i\}_{i=1}^\infty$ of $H := L^2(Z)$ with $Q_n$ denoting the orthogonal projection of $H$ onto $H_n$. Recall that

$$Q_ng = \sum_{i=1}^n (g, h_i) h_i,$$

for all $g \in H$. Further, we still have that $Q_n$ is self-adjoint, $Q_n^2 = Q_n$, and $Q_ng \rightarrow g$ in $H$ for all $g \in H$. Our previous abuse of notation also makes a return in representing the $H_n$-valued random variables $Q_nW = \sum_{i=1}^n W(h_i)h_i$ and $Q_nY = \sum_{i=1}^n Y(h_i)h_i$. Note that the convergence $P(\Phi = f_j|\mathcal{F}_n) \rightarrow P(\Phi = f_j|Y)$ still holds as well, since we can substitute our new $Y$ in the proof of Proposition 3.0.3 without loss of generality. For our current purposes, the real $n$-dimensional vector that serves as a stand-in for $Y$ is

$$y_{j,n} = (W(u(Xf_j)h_1), \ldots, W(u(Xf_j)h_n)),$$

which we use in the following lemma.

**Lemma 3.0.13.** With notation delineated above and $\Sigma_j$ as the covariance matrix of $y_{j,n}$, we claim that the density of $y_{j,n}$ takes the form

$$\phi(h) = \frac{1}{\sqrt{2\pi^n |\Sigma_j|}} \exp \left( -\frac{y^T \Sigma_j^{-1} y}{2} \right),$$

where $|\Sigma_j|$ denotes a matrix determinant.

**Proof.** We refer the reader to the proof of Lemma 3.0.9, as the work here will follow a similar form, only now our multivariate Gaussian is $y_{j,n} = (W(u(f_j)h_1), \ldots, W(u(f_j)h_n))$. Therefore, our mean is now zero but white noise is actually affecting the covariance matrix. Shortening
notation this once by \( W_{j,k} := W(u(f_j)h_k) \), we get that

\[
\text{Cov}(W_{j,k}, W_{j,l}) = E[(W_{j,k} - E(W_{j,k}))(W_{j,l} - E(W_{j,l}))] \\
= E[(W_{j,k})(W_{j,l})] \\
= \langle u(f_j)h_k, u(f_j)h_l \rangle.
\]

This gives us the \( kl \)’th entry of our covariance matrix \( \Sigma_j \). Again we make use of previously mentioned properties of multivariate Gaussians to conclude our density in \( \mathbb{R}^n \) takes the form in the statement of the lemma and we are done.

Our past result now suggests the next step, to use this density in constructing good posterior probabilities for our model after observing \( Y \). Let

\[
\alpha_{j,n} = \frac{p_j}{\sqrt{|\Sigma_j|}} \exp \left( -\frac{Q_n Y^T \Sigma_j^{-1} Q_n Y}{2} \right),
\]

which is in a similar form to our section on additive white noise. This time, however, we find that there are no convenient \( \beta_{j,n} \) that allow for cancellation of our \( \|Q_n Y\|^2 \) terms in the exponent.
Instead we have

\[
P(\Phi = f_j \mid \mathcal{F}_n) = P(\Phi = f_j \mid Q_nY)
\]

\[
= \frac{\phi(f_j, Q_nY)}{\sum_k \phi(f_k, Q_nY)}
\]

\[
= \frac{p_j}{\sqrt{2\pi^n|\Sigma_j|}} \exp \left( -\frac{Q_nY^T\Sigma_j^{-1}Q_nY}{2} \right)
\]

\[
= \frac{p_j}{\sqrt{2\pi^n|\Sigma_k|}} \exp \left( -\frac{Q_nY^T\Sigma_k^{-1}Q_nY}{2} \right)
\]

\[
= \frac{\alpha_{j,n}}{\sum_k \alpha_{k,n}}.
\]

Note that \(\|Q_nY\|^2\) does not converge as \(n \to \infty\) by definition, as that would imply \(Y \in L^2(Z)\). Nevertheless, we now have the general form of the posterior weights assigned to distinct forms of the attenuation function. More complicated models in the future may seek to combine both the result of scattering and our white noise scaled by a function of \(Xf\). More subjects for future work may include determining rates of convergence of these updated ratios.
CHAPTER 4: MODELING

In this chapter, we will be viewing the model of tomography more closely. The stochastic process $Y_{v,\varepsilon}$ in the previous chapter was given with little justification. Here we seek to construct this object from the ground up from the direct analysis of what explicitly occurs when we scan along lines through a body by the emission of high numbers of particles. We will both demonstrate the law of large numbers result of this view as well as the more precise central limit theorem form.

We now recall the infrastructure chapter’s requirement of $\{U^y_{\ell} : \ell \in \mathbb{N}, y \in Z\}$ i.i.d. Uniform(0,1) random variables, which with the definition $\{\xi^y_{\ell}\}$, give us a functional representation of the transmission of the $\ell$'th photon along the line $L_y$. For an attenuation function $f \in C(\mathcal{D})$, all the photons observed along such a line when $N$ particles are emitted are modeled as

$$S^y_{N} f := \sum_{\ell=1}^{N} \xi^y_{\ell} f,$$

which, as a series of Bernoulli trials, is distributed as a binomial random variable with parameters $N$ and $e^{-X\Phi(y)}$. Of course, we only ever expect to scan along finitely many lines and emit finitely many photons.

Since we need to account for multiple scanning lines, we will now partition the space $Z$ into sets determined by lines we scan along. Recall that we defined the elements $y \in Z$ by $y = (s, \tau_\theta)$ where for $\theta \in \mathbb{R}$, we let $\tau_\theta = (\cos \theta, \sin \theta) \in S^1$ and $s \in (0, 1)$. Fix $n, m \in \mathbb{N}$ and for $1 \leq j \leq 2^n$, $1 \leq k \leq 2^m$ let

$$s_j = \sin(\pi j / 2^{n+1}), \quad \theta_k = \frac{2\pi k}{2^m}, \quad y_{j,k} = (s_j, \tau_{\theta_k})$$

so that $\{s_j\}$ and $\{\theta_k\}$ represent radial and angular slices through $Z$ and $y_{j,k}$ falls along the intersection of grid lines. In simpler terms, $L_{y_{j,k}}$ denotes a line along which we detect particle transmission.
as determined by our fixed $n, m$. We shall also define

$$A_{j,k} = (s_{j-1}, s_j] \times \{\tau_\theta : \theta_{k-1} < \theta \leq \theta_k\} \subset \mathbb{Z}$$

and therefore, to recover the information from all our observations in a scan along the lines determined by $n, m$ and for the emission of $N$ photons along said lines, we have

$$Y_{n,m,N} f = -\sum_{j=1}^{2^n} \sum_{k=1}^{2^m} \log \left( \frac{(S_{N}^{\theta_j,k} f) \vee 1}{N} \right) 1_{A_{j,k}},$$

which represents our best use of scanning data to recover the X-ray transform of $\Phi$. Note that we take a max inside the log to avoid the unlikely case when $S_{N}^{\theta_j,k} f = 0$, for which our function would not be well-defined. Moreover, should we replace $S_{N}^{\theta_j,k} f$ with $e^{-N f(y_{j,k})}$, for $N$ large enough we would get back perfect information for our scan. Furthermore, we will be using $\Phi$ as the stand-in for our attenuation function $f$, itself a simple random variable taking values in $C(\mathbb{D})$ and modeling our lack of knowledge of the form of the actual attenuation function.

It is the focus of this chapter to outline the behavior of our model as we both increase the number of lines we scan along and also emit greater numbers of photons along each scanning line. Our particular choice of partitioning $\mathbb{Z}$ was done for the following result.

**Lemma 4.0.14.** For fixed $n, m \in \mathbb{N}$ and for $\theta_k$ and $s_j$ as defined above,

$$\mu(A_{j,k}) = \frac{\pi^2}{2^n 2^m}$$

for all $j, k$, where

$$A_{j,k} := (s_{j-1}, s_j] \times \{\tau_\theta : \theta_{k-1} < \theta \leq \theta_k\} \subset \mathbb{Z}.$$ 

In practice, rarely will scanning lines provide so uniform a partition, but the order $(2^n 2^m)^{-1}$ of the
partition would still be expected. We proceed with the proof of our claim.

**Proof.** We refer back to our infrastructure chapter for the definition of $\mu$ as a measure on $Z$ by

$$\mu(ds \, d\tau) = (1 - s^2)^{-1/2} \, ds \, \sigma(d\tau).$$

and so

$$\mu(A_{j,k}) = \int_{A_{j,k}} (1 - s^2)^{-1/2} \, ds \, \sigma(d\tau)$$
$$= \int_{\{r_\theta \theta_k - 1 < \theta < \theta_k\}} \sigma(d\tau) \int_{(s_{j-1}, s_j]} (1 - s^2)^{-1/2} \, ds$$
$$= \frac{2\pi}{2m} [\sin^{-1}(s_j) - \sin^{-1}(s_{j-1})]$$
$$= \frac{\pi^2}{2m2^n},$$

and we are done. \( \square \)

Having explicitly formulated our view of the X-ray data in the process $Y_{n,m,N}\Phi$, our immediate goals follow:

(i) Show that for each $\omega \in \Omega$, we have $Y_{n,m,N}\Phi(\cdot, \omega) \in L^2(Z)$.

(ii) Show that the function $Y_{n,m,N}\Phi : \Omega \to L^2(Z)$ is Borel measurable.

(iii) Prove that

$$\lim_{n,m,N \to \infty} E\|Y_{n,m,N}\Phi - X\Phi\|^2_{L^2(Z)} = 0.$$
(iv) Prove that, for each \( g \in L^2(Z) \), we have
\[
\lim_{n,m \to \infty} \lim_{N \to \infty} \mathbb{E} \left[ e^{-\frac{1}{2} \|u(X)g\|^2_{L^2(Z)}} \right],
\]
where \( u(\xi) = -(\xi - 1)^{1/2} \) and we define
\[
X_{n,m} \Phi := \sum_{j=1}^{2^n} \sum_{k=1}^{2^m} \Phi(y_{j,k}) 1_{A_{j,k}},
\]
the X-ray transform of our attenuation function as partitioned by scanning zones \( A_{j,k} \). Parts (i) and (ii) are technical prerequisites for (iii) and (iv). Part (iii) is a law-of-large-numbers-type theorem, and provides a rigorous functional version of (1.1). Part (iv) is a central-limit-type theorem which shows that
\[
N^{1/2}(Y_{n,m,N} - X \Phi) \to u(X \Phi)W
\]
in distribution as \( n, m, N \to \infty \), where \( \Phi \) is the \( C(\overline{D}) \)-valued random variable we use to represent our lack of knowledge of the form our attenuation function takes. We assume that there exists a bound \( M \in \mathbb{R} \) such that \( \|X \Phi(\omega)\|_{\infty} < M \) for all \( \omega \in \Omega \). This provides a rigorous functional version of (1.2). The rest of the chapter is broken up into sections delineated by the aforementioned goals.

**Square Integrability**

This section is comprised of a single straightforward argument, facilitated by our choice of \( Y_{n,m,N} \). Recall that, for any \( f \in L^2(Z) \) and fixed \( n, m, N \in \mathbb{N} \), we defined
\[
Y_{n,m,N} f = \left( -\sum_{j=1}^{2^n} \sum_{k=1}^{2^m} \log \left( \frac{G_{N_{j,k}} f}{N} \lor 1 \right) \right) 1_{A_{j,k}},
\]

as a double sum over all the scanning lines. We draw attention once more to that fact that, if none of the $N$ photons emitted are detected on the other side of a particular line $L_{y_j,k}$, then the summand for that line becomes

$$\log \left( \frac{(S_{N}^{y_j,k} f) \lor 1}{N} \right) = \log \left( \frac{0 \lor 1}{N} \right) = \log(1/N).$$

This event grows more unlikely as $N$ grows, which we will argue more rigorously in subsequent sections. For now, we note

$$\|Y_{n,m,N}f\|^2_{L^2(Z)} = \int_Z |Y_{n,m,N}f|^2 \, d\mu
= \int_Z \left| \left( - \sum_{j=1}^{2^n} \sum_{k=1}^{2^m} \log \frac{(S_{N}^{y_j,k} f) \lor 1}{N} \right) 1_{A_{j,k}} \right|^2 \, d\mu
\leq \sum_{j=1}^{2^n} \sum_{k=1}^{2^m} \int_{A_{j,k}} \left| \log \frac{(S_{N}^{y_j,k} f) \lor 1}{N} \right|^2 \, d\mu
\leq \sum_{j=1}^{2^n} \sum_{k=1}^{2^m} \int_{A_{j,k}} \left| \log \frac{1}{N} \right|^2 \, d\mu
= 2^n 2^m (\log(N))^2 \frac{\pi^2}{2^n 2^m}
= \pi^2 (\log(N))^2
< \infty.$$ 

Thus we have demonstrated that our constructed $Y_{n,m,N}f$ is indeed in $L^2(Z)$ for each $\omega \in \Omega$ and $f \in L^2(Z)$. Since $\Phi(\omega) \in C(\overline{D})$ for all $\omega \in \Omega$, we therefore have that $Y_{n,m,N}\Phi$ is in $L^2(Z)$ for each $\omega \in \Omega$. 

41
Borel Measurability

We now turn our focus to the demonstration that $Y_{n,m,N} \Phi : \Omega \to L^2(Z)$ is Borel measurable. To this end, we remind the reader that $\Phi : \Omega \to C(\mathcal{D}; [0, \infty))$ is $(\mathcal{F}, \mathcal{B}_{C(\mathcal{D})})$-measurable, and is the random variable taking the form of our unknown attenuation function. That is, $\Phi$ is a $C(\mathcal{D}; [0, \infty))$-valued random variable with law $\nu$. Since the inclusion maps are continuous, it follows that $\Phi$ is also an $L^\infty(D, \mathcal{B}_D, \lambda)$-valued random variable, as well as an $L^2(D, \mathcal{B}_D, \lambda)$-valued random variable as shown in Lemma A.0.26 and Lemma A.0.25 respectively. In particular, this implies that $X\Phi$ is an $L^2(Z, \mathcal{B}_Z, \mu)$-valued random variable. Our posterior probabilities can now be expressed as $P(X\Phi \in A \mid Y_{n,m,N}\Phi)$. According to our earlier heuristics, if $N$ is very large, then $Y_{n,m,N}\Phi \approx X\Phi$, so $X\Phi$ is essentially known, and all of these posterior probabilities should be close to 0 or 1.

Law of Large Numbers

**Theorem 4.0.15.** We claim that $Y_{n,m,N} \Phi \to X\Phi$ in $L^2(Z \times \Omega)$. In other words,

$$\lim_{n,m,N \to \infty} E\|Y_{n,m,N} \Phi - X\Phi\|_{L^2(Z)}^2 = 0.$$ 

We shall demonstrate convergence in $L^2(Z \times \Omega)$ of $X_{n,m} \Phi$ to $X\Phi$. Our next step will be to show a similar convergence for $Y_{n,m,N}$ and $X_{n,m}$ and thus complete the proof of the theorem.

**Proposition 4.0.16.** With $X_{n,m} \Phi$ as defined above, we claim that $X_{n,m} \Phi \to X\Phi$ in $L^2(Z \times \Omega)$. Equivalently,

$$\lim_{n,m \to \infty} E\|X_{n,m} \Phi - X\Phi\|_{L^2(Z)}^2 = 0.$$ 

42
Proof. We note that for $P$-a.e. $\omega \in \Omega$,

$$X_{n,m} \Phi \rightarrow X \Phi \quad \mu\text{-a.e.}$$

as $n, m \rightarrow \infty$ by the continuity of $X \Phi(\omega)$. Moreover, we note that

$$|X_{n,m} \Phi - X \Phi|^2 \leq (\|X \Phi\|_{\infty} + \|X \Phi\|_{\infty})^2$$

$$\leq 4\|X \Phi\|_{\infty}^2 \quad \mu\text{-a.e.,}$$

which gives us a bound for use of dominated convergence theorem. So,

$$\int |X_{n,m} \Phi - X \Phi|^2 \, d\mu \rightarrow 0 \quad \text{a.s.}$$

as $n, m \rightarrow \infty$. Making use of our bound once again, we also have that

$$\int |X_{n,m} \Phi - X \Phi|^2 \, d\mu \leq 4\|X \Phi\|_{\infty}^2 \mu(Z) \quad \text{a.s.}$$

and we assumed at the beginning of this chapter that $E[\|X \Phi\|_{\infty}^2] < \infty$. Thus the following integrand is bounded above, allowing for another use of the dominated convergence theorem, this time in the measure space $(\Omega, \mathcal{F}, P)$, so we see that

$$E \left[ \int |X_{n,m} \Phi - X \Phi|^2 \, d\mu \right]$$

$$= \int \left[ \int |X_{n,m} \Phi - X \Phi|^2 \, d\mu \right] \, dP \rightarrow 0,$$

as $n, m \rightarrow \infty$. This last result implies $X_{n,m} \Phi \rightarrow X \Phi$ in $L^2(Z \times \Omega)$. \(\square\)

We now have convergence of our intermediary form of $X \Phi$. To complete the proof of the theorem,
we need only the following result.

**Proposition 4.0.17.** With \( Y_{n,m,N} \Phi \) and \( X_{n,m} \Phi \) as defined previously, we claim that

\[
\lim_{n,m,N \to \infty} \| Y_{n,m,N} \Phi - X_{n,m} \Phi \|_{L^2(Z \times \Omega)}^2 = 0.
\]

Before we can begin, we need two more results.

**Lemma 4.0.18.** There exists a universal constant \( C \) such that

\[
\binom{n}{k} \hat{p}^k (1 - \hat{p})^{n-k} \leq C \frac{1}{\sqrt{2\pi np(1-p)}} \exp \left( - \frac{(k - np)^2}{2np(1-p)} \right),
\]

for \( 0 < p < 1 \).

**Proof.** Proof of this result may be found in Lemma 6.2 of [7] whenever \( p \leq 1/2 \). Without loss of generality, in cases where \( p > 1/2 \), we may relabel \( \tilde{p} = (1 - p) \) and \( \tilde{k} = (n - \tilde{k}) \) to show that the result holds. We know in this case that

\[
\binom{n}{\tilde{k}} \tilde{p}^\tilde{k} (1 - \tilde{p})^{n-\tilde{k}} \leq C \frac{1}{\sqrt{2\pi n\tilde{p}(1-\tilde{p})}} \exp \left( - \frac{(\tilde{k} - n\tilde{p})^2}{2n\tilde{p}(1-\tilde{p})} \right),
\]

since \( \tilde{p} < 1/2 \). The left hand side remains exactly the same regardless of the relabeling due to the
symmetry of binomial coefficients, so we can now write,

\[
\binom{n}{k} p^k (1-p)^{n-k} \leq C \frac{1}{\sqrt{2\pi np(1-p)}} \exp \left( -\frac{(k - np)^2}{2np(1-p)} \right)
\]

\[
= C \frac{1}{\sqrt{2\pi np(1-p)}} \exp \left( -\frac{(n - k - n(1-p))^2}{2np(1-p)} \right)
\]

\[
= C \frac{1}{\sqrt{2\pi np(1-p)}} \exp \left( -\frac{(-k + np)^2}{2np(1-p)} \right)
\]

\[
= C \frac{1}{\sqrt{2\pi np(1-p)}} \exp \left( -\frac{(k - np)^2}{2np(1-p)} \right),
\]

and we are done with both cases. \(\square\)

**Theorem 4.0.19.** Let \(X \sim \text{Binomial}(N, p)\) with \(p \geq \varepsilon\). Then, for all \(N > 2/\varepsilon\) and all \(\alpha > 0\), there exists a constant \(C_\alpha\) such that

\[
E \left| \log \left( \frac{X \lor 1}{N} \right) - \log(p) \right|^2 \leq C_\alpha \frac{(\log N)^2}{(N\varepsilon)^\alpha} + \frac{4}{N\varepsilon}.
\]

In particular, this implies that

\[
E \left| \log \left( \frac{X \lor 1}{N} \right) - \log(p) \right|^2 \leq \frac{C (\log N)^2}{N\varepsilon},
\]

and, for all \(\alpha > 1\),

\[
E \left| \log \left( \frac{X \lor 1}{N} \right) - \log(p) \right|^2 \leq C_\alpha \frac{1}{N\varepsilon^\alpha}.
\]

**Proof.** We first assume that \(N > 2/\varepsilon\) and note that therefore,

\[
\frac{1}{N} < \frac{\varepsilon}{2} < \frac{p}{2}.
\]

As far as our binomial random variable is concerned, we note that since \(\frac{1}{N} < \frac{p}{2}\), then the only way that \(\frac{X \lor 1}{N} \geq \frac{p}{2}\) is if \(X \geq \frac{Np}{2}\). We will use this event to partition the proof of our result by the amount
of successful Bernoulli trials. Therefore,

\[ E \left| \log \left( \frac{X \lor 1}{N} \right) - \log(p) \right|^2 \]

\[ = E \left[ \left| \log \left( \frac{X \lor 1}{N} \right) - \log(p) \right|^2 \left( 1_{\{X \geq N/2\}} + 1_{\{X < N/2\}} \right) \right] \]

\[ = E \left[ \left| \log \left( \frac{X \lor 1}{N} \right) - \log(p) \right|^2 1_{\{X \geq N/2\}} \right] + E \left[ \left| \log \left( \frac{X \lor 1}{N} \right) - \log(p) \right|^2 1_{\{X < N/2\}} \right] \]

In the first case, we note that for \( y \geq p/2 \),

\[ |\log(y) - \log(p)| = \int_p^y \frac{1}{t} \, dt \leq \frac{2}{p} |y - p|, \]

which allows us to proceed with our bound of the form

\[ E \left[ \left| \log \left( \frac{X \lor 1}{N} \right) - \log(p) \right|^2 1_{\{X \geq N/2\}} \right] \leq \frac{4}{p^2} E \left[ \left| \frac{X \lor 1}{N} - p \right|^2 1_{\{X \geq N/2\}} \right] \]

\[ = \frac{4}{p^2} E \left[ \left| \frac{X}{N} - p \right|^2 1_{\{X \geq pN\}} \right] \]

\[ \leq \frac{4}{p^2} \text{Var} \left( \frac{X}{N} \right) \]

\[ = \frac{4}{p^2} \frac{1}{N^2} Np(1-p) = \frac{4(1-p)}{pN} \leq \frac{4}{N\epsilon^2}, \]

which takes care of one of our two cases. For the second term, we begin by noting since \( N > \frac{2}{\epsilon} \), we have that \( \frac{1}{p} \leq \frac{1}{\epsilon} < \frac{N}{2} < N \), and therefore,

\[ |\log(p)| = \log(1/p) < \log(N), \]
which gives us that
\[ \left| \log \left( \frac{X \vee 1}{N} \right) - \log(p) \right|^2 \leq 2 \left( \left| \log \left( \frac{X}{N} \right) \right|^2 + \left| \log(p) \right|^2 \right) \]
\[ \leq 2 \left( \left| \log \left( \frac{1}{N} \right) \right|^2 + \left| \log(p) \right|^2 \right) \]
\[ \leq 4 (\log N)^2. \]

We can now focus on the case where fewer of our Bernoulli trials succeed and make use of the result of Lemma 4.0.18, where we shall at times absorb constants into $C$ to simplify notation,
\[ E \left[ \left| \log \left( \frac{X \vee 1}{N} \right) - \log(p) \right|^2 1_{\{X < \frac{Np}{2}\}} \right] \leq 4 (\log N)^2 P \left( X < \frac{Np}{2} \right) \]
\[ \leq 4 (\log N)^2 \sum_{k=0}^{\lfloor Np/2 \rfloor} \frac{1}{\sqrt{2\pi Np(1-p)}} \exp \left( \frac{-(k-Np)^2}{2Np(1-p)} \right) \]
\[ = C \frac{(\log N)^2}{\sqrt{Np(1-p)}} \sum_{k=0}^{\lfloor Np/2 \rfloor} \exp \left( -\frac{(k-Np)^2}{2Np(1-p)} \right). \]

Note that these summands increase as $|k-Np|$ decreases. As both $k$ and $Np$ are nonnegative, then for $k < Np$ this is achieved by increasing the value of $k$, and so we proceed to write
\[ E \left[ \left| \log \left( \frac{X \vee 1}{N} \right) - \log(p) \right|^2 1_{\{X < \frac{Np}{2}\}} \right] \leq C \frac{(\log N)^2}{\sqrt{Np(1-p)}} \sum_{k=0}^{\lfloor Np/2 \rfloor} \exp \left( -\frac{(k-Np)^2}{2Np(1-p)} \right) \]
\[ = C \frac{(\log N)^2}{\sqrt{Np(1-p)}} \sum_{k=0}^{\lfloor Np/2 \rfloor} \exp \left( -\frac{(Np)^2}{8Np(1-p)} \right) \]
\[ \leq C \frac{(\log N)^2 Np}{\sqrt{Np(1-p)}} \frac{1}{2} \exp \left( -\frac{Np}{8(1-p)} \right) \]
\[ = C (\log N)^2 \sqrt{\frac{Np}{1-p}} \exp \left( -\frac{Np}{8(1-p)} \right). \]

Now, for every $x > 0$ and $\alpha > 0$, there exist a constant $C_\alpha$ such that $x^{1/2}e^{-x/8} \leq C_\alpha x^{-\alpha}$, since
the function $x^{\alpha+1/2}e^{-x/8}$ is continuous on $[0, \infty)$ with a limit of zero as $x$ grows for all positive $\alpha$.

Combined with the last term of our previous inequalities, we find that

$$E \left[ \left\| \log \left( \frac{X \vee 1}{N} \right) - \log(p) \right\|^2 1_{\{X \leq \frac{Np}{4} \}} \right] \leq C_\alpha (\log N)^2 \left( \frac{Np}{1-p} \right)^{-\alpha}$$

$$\leq C_\alpha (\log N)^2 (N\varepsilon)^{-\alpha},$$

and we are now prepared to put together the bounds on both cases of the success of Bernoulli trials.

To this end,

$$E \left[ \left\| \log \left( \frac{X \vee 1}{N} \right) - \log(p) \right\|^2 \right]$$

$$= E \left[ \left\| \log \left( \frac{X \vee 1}{N} \right) - \log(p) \right\|^2 1_{\{X \geq \frac{Np}{4} \}} \right] + E \left[ \left\| \log \left( \frac{X \vee 1}{N} \right) - \log(p) \right\|^2 1_{\{X \leq \frac{Np}{4} \}} \right]$$

$$\leq \frac{4}{N\varepsilon} + C_\alpha \frac{(\log N)^2}{(N\varepsilon)^{\alpha}},$$

which proves the general statement of our lemma. Now, if we plug in $\alpha = 1$, we see that

$$E \left[ \left\| \log \left( \frac{X \vee 1}{N} \right) - \log(p) \right\|^2 \right] \leq \frac{4}{N\varepsilon} + C \frac{(\log N)^2}{N\varepsilon}$$

$$= \frac{4}{N\varepsilon} + C \frac{(\log N)^2}{N\varepsilon}$$

$$= \frac{(\log N)^2}{N\varepsilon} \left( \frac{4}{(\log N)^2} + C \right)$$

$$\leq C \frac{(\log N)^2}{N\varepsilon},$$

given the boundedness of the expression in parenthesis in the second to last step. For a fixed $\alpha > 1,$
we get

\[
E \left[ \log \left( \frac{X \lor 1}{N} \right) - \log(p) \right]^2 \leq \frac{4}{N \varepsilon} + C_\alpha \frac{(\log N)^2}{(N \varepsilon)^\alpha} \\
\leq \frac{4}{N \varepsilon^\alpha} + C_\alpha \frac{(\log N)^2}{N^{\alpha-1}} \frac{1}{N \varepsilon^\alpha} \\
= \frac{1}{N \varepsilon^\alpha} \left( 4 + C_\alpha \frac{(\log N)^2}{N^{\alpha-1}} \right) \\
\leq C_\alpha \frac{1}{N \varepsilon^\alpha},
\]

where the final step follows because \(4 + C_\alpha \frac{(\log N)^2}{N^{\alpha-1}}\) is known to be bounded since it goes to zero as \(N\) grows and we allowed \(C_\alpha\) to absorb the constant. Thus we have completely proven the full statement of our lemma.

We can now move forward with the second major proposition of this section.

**Proof of Proposition 4.0.17.** We will refer back to the model of our photons’ progress as Bernoulli trials through a body with an unknown attenuation function. Thus, along a fixed line \(L_{y_{j,k}}\),

\[
P(S_{N}^{y_{j,k}} \Phi = l|\Phi) = \binom{N}{l} p_{j,k}^l (1 - p_{j,k})^{N-l}
\]

where \(p_{j,k} = e^{-X \Phi(y_{j,k})}\). Therefore,

\[
E \left[ \log \left( \frac{S_N^{y_{j,k}} \Phi \lor 1}{N} \right) + X \Phi(y_{j,k}) \right]^2 \\
= E \left[ \left\| \log \left( \frac{S_N^{y_{j,k}} \Phi \lor 1}{N} \right) + X \Phi(y_{j,k}) \right\|^2 \right] \left( \Phi(y_{j,k}) \right) \\
\leq E \left[ C_\alpha \frac{1}{N \varepsilon^\alpha} \right] \\
= C_\alpha \frac{1}{N \varepsilon^\alpha}.
\]
where \( \varepsilon = e^{-2M} \), \( \alpha > 1 \), and \( N > 2/\varepsilon \), taking \( M \) to be the bound on \( \|X\Phi\|_\infty \) known to exist from the beginning of the chapter. The inequality follows from Lemma 4.0.19. We proceed to write

\[
\|Y_{n,m,N}\Phi - X_{n,m}\Phi\|_{L^2(Z \times \Omega)}^2
\]

\[
= \int_Z E \left| Y_{n,m,N}\Phi - X_{n,m}\Phi \right|^2 d\mu
\]

\[
\leq \int_Z E \left[ \sum_{j=1}^{2^n} \sum_{k=1}^{2^m} \left( \log \left( \frac{S_{N,j,k}^{(y_j,k)} \Phi \lor 1}{N} \right) + X\Phi(y_{j,k}) \right) 1_{A_{j,k}} \right]^2 d\mu
\]

\[
\leq \int_Z E \left[ \sum_{j=1}^{2^n} \sum_{k=1}^{2^m} \left( \log \left( \frac{S_{N,j,k}^{(y_j,k)} \Phi \lor 1}{N} \right) + X\Phi(y_{j,k}) \right)^2 1_{A_{j,k}} \right] d\mu
\]

\[
= \int_Z \sum_{j=1}^{2^n} \sum_{k=1}^{2^m} E \left| \log \left( \frac{S_{N,j,k}^{(y_j,k)} \Phi \lor 1}{N} \right) + X\Phi(y_{j,k}) \right|^2 1_{A_{j,k}} d\mu
\]

\[
\leq \int_Z \sum_{j=1}^{2^n} \sum_{k=1}^{2^m} C_\alpha \frac{1}{N \varepsilon^\alpha} 1_{A_{j,k}} d\mu
\]

\[
= C_\alpha \frac{1}{N \varepsilon^\alpha} \mu(Z)
\]

\[
= \pi^2 C_\alpha \frac{1}{N \varepsilon^\alpha},
\]

which vanishes as \( N \to \infty \) and we have our proof. \( \square \)

Equipped as we are, the proof of Theorem 4.0.15 follows quite immediately.

**Proof.** We have, from Proposition 4.0.16, that

\[
\lim_{n,m \to \infty} \|X_{n,m}\Phi - X\Phi\|_{L^2(Z \times \Omega)} = 0,
\]

and from Proposition 4.0.17 we get

\[
\lim_{n,m,N \to \infty} \|Y_{n,m,N}\Phi - X_{n,m}\Phi\|_{L^2(Z \times \Omega)} = 0,
\]
which is sufficient to conclude that

$$\lim_{n,m,N \to \infty} E\|Y_{n,m,N}\Phi - X\Phi\|_{L^2(Z)}^2 = 0.$$ 

\[
\text{Central Limit Theorem}
\]

It is here that we justify the form of the white noise term of our model as seen throughout this work. We will continue to employ the assumption from the beginning of the chapter that that for all $n, m \in \mathbb{N}$, the measure of our $A_{j,k}$ subsets of $Z$ remain equal. In other words, $\mu(A_{j,k}) = \frac{\pi^2}{2n^2m}$. Moreover, we will henceforth assume $n = m$ for simplicity, given that they both denote a similar shrinking of our scanning grid. This leads to the minor variation of notation for our observed values of the form

$$Y_{n,m,N}f = Y_{n,N}f = \left(-\sum_{j=1}^{2^n} \sum_{k=1}^{2^n} \log \left(\frac{(S_{N}^{y_{j,k}} f) \vee 1}{N}\right)\right)1_{A_{j,k}},$$

so we will observe only two distinct limits, one corresponding to the photons being emitted so as to allow the central limit theorem to take effect (but not yet accounting for the stronger result in the law of large numbers) and another limit denoting the higher precision of our scans as we detect
along more lines through the body of interest. The main result for this chapter follows.

**Theorem 4.0.20.** We claim that for all \( g \in C(\overline{Z}) \), we have

\[
\lim_{n \to \infty} \lim_{N \to \infty} E \left[ e^{i \frac{2^n N^{1/2}}{n} \left( Y_{n,N} \Phi - X_n \Phi \right) g} \right]_{L^2(Z)} = E \left[ e^{-\frac{1}{2} \|u(X)g\|_{L^2(Z)}^2} \right],
\]

where \( u(x) := -(e^x - 1)^{1/2} \) and

\[
X_n \Phi := \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} X \Phi(y_{j,k}) 1_{A_{j,k}}.
\]

This result shall follow as a combination of several disparate findings working together. Rather than tackle this theorem directly, we will first show the analog for nonrandom attenuation functions.

**Theorem 4.0.21.** We claim that for all \( g \in C(\overline{Z}) \) and \( f \in C(\overline{D}) \), we have

\[
\lim_{n \to \infty} \lim_{N \to \infty} E \left[ e^{i \frac{2^n N^{1/2}}{n} \left( Y_{n,N} f - X_n f \right) g} \right]_{L^2(Z)} = e^{-\frac{1}{2} \|u(X)g\|_{L^2(Z)}^2},
\]

where \( u(x) := -(e^x - 1)^{1/2} \) and

\[
X_n f := \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} X f(y_{j,k}) 1_{A_{j,k}}.
\]

We will need to demonstrate the following lemma on binomial random variables.

**Lemma 4.0.22.** Let \( C_n \sim \text{Binomial}(n,p) \) for \( n \in \mathbb{N} \) and \( 0 < p < 1 \). Let \( D_n = \frac{C_n - p}{n} \). Then we claim

\[
\frac{1}{n^{1/2}} \int_{p}^{p + \frac{1}{n}} \frac{t - p}{tp} \, dt \to 0.
\]

in distribution as \( n \to \infty \)
Proof. It will suffice to show that the term in question goes to zero in probability. We will decompose the integral into two parts by writing

\[ n^{1/2} \int_{p}^{D_n} \frac{t-p}{tp} \, dt = \left( n^{1/2} \int_{p}^{D_n} \frac{t-p}{tp} \, dt \right) 1_{\{D_n \geq p/2\}} + \left( n^{1/2} \int_{p}^{D_n} \frac{t-p}{tp} \, dt \right) 1_{\{D_n < p/2\}} \quad (4.1) \]

where we will focus on the first term now. Note that when \( D_n \geq p/2 \), we can say

\[
\left| n^{1/2} \int_{p}^{D_n} \frac{t-p}{tp} \, dt \right| \leq n^{1/2} \int_{p \land D_n}^{p \lor D_n} \left| \frac{t-p}{tp} \right| \, dt \\
\leq \frac{2n^{1/2}}{p^2} \int_{p \land D_n}^{p \lor D_n} |t-p| \, dt \\
\leq \frac{2n^{1/2}}{p^2} |D_n - p| \int_{p \land D_n}^{p \lor D_n} \, dt \\
= \frac{2n^{1/2}}{p^2} |D_n - p|^2 \\
= \left( \frac{2}{p^2n^{1/2}} \left[ n^{1/2}(C_n/n - p) \right]^2 \right) 1_{\{C_n \geq 1\}} + \left( \frac{2}{p^2n^{1/2}} \left[ n^{1/2}(D_n/n - p) \right]^2 \right) 1_{\{C_n = 0\}},
\]

and we shall again split our focus. As for the second conditioning event in (4.1), we make use of
Chebyshev’s inequality to see that

\[ P \left( D_n < \frac{p}{2} \right) \leq P \left( \frac{C_n}{n} < \frac{p}{2} \right) \]

\[ = P \left( C_n < \frac{np}{2} \right) \]

\[ = P \left( C_n - np < \frac{-np}{2} \right) \]

\[ \leq P \left( |C_n - np|^2 > \frac{n^2p^2}{4} \right) \]

\[ \leq \frac{E|C_n - np|^2}{n^2p^2} \]

\[ = \frac{4np(1-p)}{n^2p^2} \]

\[ = \frac{4(1-p)}{np}, \]

which will clearly vanish as \( n \to \infty \). Now, to prove convergence to zero in probability of our original term, we fix \( \varepsilon > 0 \) and seek to combine previous results by writing,

\[ P \left( \left| n^{1/2} \int_{\frac{p}{np}}^{D_n} \frac{t-p}{tp} \, dt \right| > 2\varepsilon \right) \quad (4.2) \]

\[ \leq P \left( \left| \left( n^{1/2} \int_{\frac{p}{np}}^{D_n} \frac{t-p}{tp} \, dt \right) 1_{\{D_n > p/2\}} \right| > \varepsilon \right) + P \left( \left| \left( n^{1/2} \int_{\frac{p}{np}}^{D_n} \frac{t-p}{tp} \, dt \right) 1_{\{D_n < p/2\}} \right| > \varepsilon \right), \quad (4.3) \]
where we have split along the aforementioned events. For the first term, we see that,

\[
P\left(\left|\left(n^{1/2}\int_0^{D_n} \frac{t-p}{tp} dt\right)1_{\{D_n \geq p/2\}}\right| > \varepsilon\right)
\leq P\left(\frac{2}{p^2n^{1/2}}[n^{1/2}(D_n/n - p)]^2 1_{\{D_n \geq p/2\}} > \varepsilon\right)
\leq P\left(\frac{2}{p^2n^{1/2}}[n^{1/2}(D_n/n - p)]^2 > \varepsilon\right)
= P\left(\left(\frac{2}{p^2n^{1/2}}[n^{1/2}(D_n/n - p)]^2\right)1_{\{C_n \geq 1\}} + \left(\frac{2}{p^2n^{1/2}}[n^{1/2}(D_n/n - p)]^2\right)1_{\{C_n = 0\}} > \frac{\varepsilon}{2}\right)
\leq P\left(\left(\frac{2}{p^2n^{1/2}}[n^{1/2}(D_n/n - p)]^2\right)1_{\{C_n \geq 1\}} > \frac{\varepsilon}{2}\right) + P\left(\left(\frac{2}{p^2n^{1/2}}[n^{1/2}(D_n/n - p)]^2\right)1_{\{C_n = 0\}} > \frac{\varepsilon}{2}\right),
\]

where we have again split along conditioning events. Thus

\[
P\left(\left(\frac{2}{p^2n^{1/2}}[n^{1/2}(C_n/n - p)]^2\right)1_{\{C_n \geq 1\}} > \frac{\varepsilon}{2}\right)
\leq P\left(\frac{2}{p^2n^{3/2}}[C_n - np]^2 > \frac{\varepsilon}{2}\right)
= P\left(|C_n - np|^2 > \frac{\varepsilon pn^{1/2}}{4(1-p)}np(1-p)\right)
\leq P\left(|C_n - np| > \left(\frac{\varepsilon pn^{1/2}}{4(1-p)}\right)^{1/2} \sqrt{np(1-p)}\right)
\leq \frac{4(1-p)}{\varepsilon pn^{1/2}} \rightarrow 0,
\]
as \(n \rightarrow \infty\), having appealed to Chebyshev’s inequality in the last step for our binomial random variable \(C_n\). As for the second term, we focus mainly on the indicator. We note that

\[
P\left(\left(\frac{2}{p^2n^{1/2}}[n^{1/2}(D_n/n - p)]^2\right)1_{\{C_n = 0\}} > \frac{\varepsilon}{2}\right)
\leq P(C_n = 0)
= (1-p)^n \rightarrow 0
\]
as \( n \to \infty \). We will do something similar for the final term we need to account for in 4.2.

\[
P\left( \left( n^{1/2} \int_{p}^{D_n} \frac{t - p}{tp} \, dt \right) 1_{\{D_n < p/2\}} \right) > \varepsilon \right) \leq P(D_n < p/2) \\
\leq \frac{4(1 - p)}{np} \to 0,
\]

as \( n \to \infty \) as well. We have thus shown convergence to zero in probability.

This equips us to show the first of two main propositions for this section.

**Proposition 4.0.23.** For \( Y_{n,N}f \) and \( X_n f \) as above and \( g \in L^2(Z) \),

\[
\langle N^{1/2}(Y_{n,N}f - X_n f), g \rangle_{L^2(Z)} \Rightarrow \eta_n,
\]

in distribution as \( N \to \infty \) where

\[
\eta_n \sim N \left( 0, \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} (u(Xf(y_{j,k})))^2 \left( \int_{A_{j,k}} g \, d\mu \right)^2 \right).
\]

**Proof.** We alluded to the assumption that for all \( n \in \mathbb{N} \), the measure of our \( A_{j,k} \) subsets of \( Z \) remains equal. In other words, \( \mu(A_{j,k}) = \frac{\pi^2}{4n} \) for all \( n \). We begin by rewriting the random variable (whose characteristic function we focus on) by stating

\[
\langle N^{1/2}(Y_{n,N}f - X_n f), g \rangle
\]

\[
= \left\langle N^{1/2} \left( - \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} \log \left( \frac{(S_{N}^{y_{j,k}} f)}{N} \right) \right) 1_{A_{j,k}} - X_n f, g \right\rangle
\]

\[
= \left\langle N^{1/2} \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} \left( - \log \left( \frac{(S_{N}^{y_{j,k}} f)}{N} \right) - X f(y_{j,k}) \right) 1_{A_{j,k}}, g \right\rangle
\]

\[
= - \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} N^{1/2} \left( \log \left( \frac{(S_{N}^{y_{j,k}} f)}{N} \right) - \log \left( e^{-X f(y_{j,k})} \right) \right) \left\langle 1_{A_{j,k}}, g \right\rangle,
\]

56
which permits us to focus on the effects of the central limit theorem on each of the \( n \) terms. Note that, using Lemma B.0.44,

\[
\langle 1_{A_{j,k}}, g \rangle N^{1/2} \left( \log \left( \frac{\left(S_{N}^{y_{j,k}} f\right) \vee 1}{N} \right) - \log \left( e^{-Xf(y_{j,k})} \right) \right)
\]

\[
= \langle 1_{A_{j,k}}, g \rangle N^{1/2} \left[ e^{Xf(y_{j,k})} \left( \frac{\left(S_{N}^{y_{j,k}} f\right) \vee 1}{N} - e^{-Xf(y_{j,k})} \right) - \int e^{-Xf(y_{j,k})} \frac{t - e^{-Xf(y_{j,k})}}{te^{-Xf(y_{j,k})}} \, dt \right]
\]

where we find the first term is almost ripe for the use of a central limit theorem. We note that the second term goes to zero in distribution as \( N \to \infty \) by Lemma 4.0.22. As for the former term, we note that we can clean things up by decomposing into regions of \( \Omega \) where \( S_{N}^{y_{j,k}} f \geq 1 \) and \( S_{N}^{y_{j,k}} f = 0 \). The first will free us to finally make use of our desired central limit theorem result as shown by

\[
\langle 1_{A_{j,k}}, g \rangle N^{1/2} \left( \frac{\left(S_{N}^{y_{j,k}} f\right) \vee 1}{N} - e^{-Xf(y_{j,k})} \right) 1_{\{S_{N}^{y_{j,k}} f \geq 1\}}
\]

\[
= \langle 1_{A_{j,k}}, g \rangle N^{1/2} \left( \frac{S_{N}^{y_{j,k}} f}{N} - e^{-Xf(y_{j,k})} \right) 1_{\{S_{N}^{y_{j,k}} f \geq 1\}},
\]

while the second will again be shown to go to zero in probability. To this end we note that the probability that no photons are detected along a particular line \( L_{y_{j,k}} \) is given by

\[
P \left( S_{N}^{y_{j,k}} f = 0 \right) = \left( 1 - e^{-Xf(y_{j,k})} \right)^{N},
\]

which, when combined with the rest of the term for \( \varepsilon > 0 \), tells us

\[
P \left( e^{-Xf(y_{j,k})} \frac{\left(S_{N}^{y_{j,k}} f\right) \vee 1}{N} - e^{-Xf(y_{j,k})} \right) 1_{\{S_{N}^{y_{j,k}} f = 0\}} > \varepsilon
\]

\[
\leq P( S_{N}^{y_{j,k}} f = 0 )
\]

\[
= \left( 1 - e^{-Xf(y_{j,k})} \right)^{N} \to 0
\]
as $N \to \infty$. This gives us convergence in probability (and therefore in distribution) to zero of all terms excepting

$$\langle 1_{A_{j,k}}, g \rangle \frac{N^{1/2}}{e^{-Xf(y_{j,k})}} \left( \frac{S_{N}^{y_{j,k}} f}{N} - e^{-Xf(y_{j,k})} \right) 1_{\{S_{N}^{y_{j,k}} f > 0\}}$$

(4.4)

which firmly places it in our sights for the follow distributional result. As a binomial random variable generated as a sum of Bernoulli trials, we note that $S_{y_{j,k}} N f$ has a mean of $e^{-Xf(y_{j,k})}$ and a variance of $N e^{-Xf(y_{j,k})} \left(1 - e^{-Xf(y_{j,k})}\right)$. We are now prepared to speak to the effects of the central limit theorem for $n = m$ fixed. If we permit $\eta$ to represent a standard normal, we see that each of our $j, k$ terms have a convergence in distribution of the form

$$\langle 1_{A_{j,k}}, g \rangle \frac{N^{1/2}}{e^{-Xf(y_{j,k})}} \left( \frac{S_{N}^{y_{j,k}} f}{N} - e^{-Xf(y_{j,k})} \right) \Rightarrow \left[ \langle 1_{A_{j,k}}, g \rangle \frac{\left( e^{-Xf(y_{j,k})} \left(1 - e^{-Xf(y_{j,k})}\right)\right)^{1/2}}{e^{-Xf(y_{j,k})}} \right] \eta$$

$$= \left[ \langle 1_{A_{j,k}}, g \rangle \left( e^{Xf(y_{j,k})} - 1 \right)^{1/2} \right] \eta$$

$$= \langle 1_{A_{j,k}}, g \rangle u(Xf(y_{j,k})) \eta,$$

which equips us to talk about the distributional limit of all the previous terms combined. For this, we make use of the result in Lemma A.0.43 from [11]. We shall take into account that this result does not follow precisely for (4.4). Let us denote

$$T_{j,k} := \langle 1_{A_{j,k}}, g \rangle \frac{N^{1/2}}{e^{-Xf(y_{j,k})}} \left( \frac{S_{N}^{y_{j,k}} f}{N} - e^{-Xf(y_{j,k})} \right),$$

so we may show that $\left| T_{j,k} - T_{j,k} 1_{\{S_{N}^{y_{j,k}} f > 0\}} \right|$ goes to zero in probability. We therefore fix $\varepsilon > 0$ and write,

$$P \left( \left| T_{j,k} - T_{j,k} 1_{\{S_{N}^{y_{j,k}} f > 0\}} \right| > \varepsilon \right) = P \left( \left| T_{j,k} 1_{\{S_{N}^{y_{j,k}} f = 0\}} \right| > \varepsilon \right)$$

$$\leq P(S_{N}^{y_{j,k}} f = 0)$$

$$= (1 - e^{-Xf(y_{j,k})})^{N} \to 0$$

58
as $N \to \infty$. For our case, this implies that, since all other terms go to zero in probability, the term that comes from our use of the central limit theorem uniquely dictates the distributional limit of the following term,

$$\langle 1_{A_{j,k}}, g \rangle N^{1/2} \left( \log \left( \frac{(S_{N}^{N^{1/2}}} f) \vee 1}{N} \right) - \log \left( e^{-Xf(y_{j,k})} \right) \right) \Rightarrow \left[ \int_{A_{j,k}} g \, d\mu \frac{\left( e^{-Xf(y_{j,k})} (1 - e^{-Xf(y_{j,k})}) \right)^{1/2}}{e^{-Xf(y_{j,k})}} \right] \eta$$

$$= \left[ \int_{A_{j,k}} g \, d\mu \left( e^{Xf(y_{j,k})} - 1 \right)^{1/2} \right] \eta$$

$$= \left[ u(Xf(y_{j,k})) \int_{A_{j,k}} g \, d\mu \right] \eta$$

as $N \to \infty$. In other words, each of our terms converges in distribution to a normal random variable of mean zero and variance $(u(Xf(y_{j,k})))^{2} \left( \int_{A_{j,k}} g \, d\mu \right)^{2}$. We recall that our original random variable was comprised of a double sum of such terms, each with the distributional limit of a normal random variable. Therefore, as a sum of independent normal random variables, the result is itself a normal random variable with a variance given by the sum of the respective variances of the summands. Put succinctly,

$$- \sum_{j=1}^{2^{n}} \sum_{k=1}^{2^{n}} N^{1/2} \left( \log \left( \frac{(S_{N}^{N^{1/2}}} f) \vee 1}{N} \right) - \log \left( e^{-Xf(y_{j,k})} \right) \right) \langle 1_{A_{j,k}}, g \rangle \Rightarrow \eta_{n}$$

where $\eta_{n} \sim \text{Normal} \left( 0, \sum_{j=1}^{2^{n}} \sum_{k=1}^{2^{n}} (u(Xf(y_{j,k})))^{2} \left( \int_{A_{j,k}} g \, d\mu \right)^{2} \right)$ and we have acquired the desired result for our lemma. \(\square\)

Note that this convergence in distribution yields us a convergence of characteristic functions, and so

$$E \left[ \exp \left( it \langle N^{1/2} (Y_{n,N} f - X_{n} f), g \rangle \right) \right] \to E \left[ \exp \left( it\eta_{n} \right) \right]$$

as $N \to \infty$ for all $t \in \mathbb{R}$. If we turn our focus to the variance of $\eta_{n}$, we want to demonstrate the
Proposition 4.0.24. For $g$ and $u(X f)$ continuous in $\overline{Z}$ and $A_{j,k}$ as defined previously,

$$\frac{4^n}{\pi^2} \left[ \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} (u(X f(y_{j,k})))^2 \left( \int_{A_{j,k}} g \, d\mu \right)^2 \right] \rightarrow \int_{Z} \left( (u(X f(y)))^2 \right) \, d\mu$$

$$= \|u(X f)g\|^2, \quad a.s.$$ as $n \to \infty$.

Proof. For this, we will make use of both the continuity in $Z$ of $g$ and $u(X f)$. And so we begin

$$\left| \frac{4^n}{\pi^2} \left[ \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} (u(X f(y_{j,k})))^2 \left( \int_{A_{j,k}} g \, d\mu \right)^2 \right] - \|u(X f)g\|^2 \right|$$

$$= \left| \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} \left[ \frac{4^n}{\pi^2} (u(X f(y_{j,k})))^2 \left( \int_{A_{j,k}} g \, d\mu \right)^2 - \int_{A_{j,k}} (u(X f))^2 g^2 \, d\mu \right] \right|$$

$$\leq \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} \left[ \frac{4^n}{\pi^2} (u(X f(y_{j,k})))^2 \left( \int_{A_{j,k}} g \, d\mu \right)^2 - \int_{A_{j,k}} (u(X f))^2 g^2 \, d\mu \right],$$
where, if we look at the term of a fixed $j, k$ we find that

\[
\left| \frac{4^n}{\pi^2} (u(X f(y_{j,k})))^2 \left( \int_{A_{j,k}} g \, d\mu \right) \right|^2 - \int_{A_{j,k}} (u(X f))^2 g^2 \, d\mu
\]

\[
\leq \frac{4^n}{\pi^2} (u(X f(y_{j,k})))^2 \left( \int_{A_{j,k}} g \, d\mu \right) \left( \int_{A_{j,k}} |g - g(y_{j,k})| \, d\mu \right)
\]

\[
\leq \frac{4^n}{\pi^2} (u(X f(y_{j,k})))^2 \|g\|_\infty (\mu(A_{j,k}))^2 \varepsilon
\]

\[
\leq \|u(X f)\|_\infty^2 \|g\|_\infty \mu(A_{j,k}) \varepsilon
\]

\[
= \frac{\pi^2}{4^n} \|u(X f)\|_\infty^2 \|g\|_\infty \varepsilon.
\]

and we shall now tackle each of these three terms one at a time. We first note that since $g$ is continuous on the compact set $Z$, it is in fact uniformly continuous. Therefore, for $\varepsilon > 0$ of our choice there exists a large enough $n_\varepsilon$ (and so small enough $A_{j,k}$) so that $|g(y) - g(y_{j,k})| \leq \varepsilon$ for all $y \in A_{j,k}$. We use this result for $n > n_\varepsilon$ and so
A similar result follows from our second term, where we find that for \( n > n_\varepsilon \),

\[
\frac{4^n}{\pi^2} (u(X f(y_{j,k})))^2 g(y_{j,k}) \mu(A_{j,k}) \int_{A_{j,k}} g \, d\mu - \frac{4^n}{\pi^2} (u(X f(y_{j,k})))^2 (g(y_{j,k}))^2 (\mu(A_{j,k}))^2 \\
\leq \frac{4^n}{\pi^2} (u(X f(y_{j,k})))^2 |g(y_{j,k})| \mu(A_{j,k}) \int_{A_{j,k}} |g - g(y_{j,k})| \, d\mu \\
\leq \|u(X f)\|_\infty^2 \|g\|_\infty \mu(A_{j,k}) \varepsilon \\
= \frac{\pi^2}{4^n} \|u(X f)\|_\infty^2 \|g\|_\infty \varepsilon
\]

And finally, for our third term, we note that continuity of both \( g \) and \( u(X f) \) yields continuity in their product squared. Thus, for \( \varepsilon' > 0 \) there exists large enough \( n_{\varepsilon'} \) so that \( |[u(X f) g]^2(y) - [u(X f) g]^2(y_{j,k})| < \varepsilon' \) for all \( y \in A_{j,k} \). So for our third term we find, when \( n > n_{\varepsilon'} \),

\[
\left| \frac{4^n}{\pi^2} (u(X f(y_{j,k})))^2 (g(y_{j,k}))^2 (\mu(A_{j,k}))^2 - \int_{A_{j,k}} (u(X f))^2 g^2 \, d\mu \right| \\
\leq \int_{A_{j,k}} |(u(X f))^2 g^2 - (u(X f(y_{j,k})))^2 (g(y_{j,k}))^2| \, d\mu \\
\leq \mu(A_{j,k}) \varepsilon' \\
= \frac{\pi^2}{4^n} \varepsilon'
\]

so we’ve reached a bound for our final term. We now choose to take \( \bar{\varepsilon} \) to be arbitrarily small and set

\[
\varepsilon, \varepsilon' < \min \left\{ \frac{\bar{\varepsilon}}{3\pi^2}, \frac{\bar{\varepsilon}}{3\pi^2 \|u(X f)\|_\infty^2 \|g\|_\infty} \right\}.
\]

So, for \( n > \max\{n_\varepsilon, n_{\varepsilon'}\} \).
which concludes our argument for the limit of the sum of variances.

Therefore, we have convergence in expectation of our random variable \( \eta_n \) in the form

\[
E \left[ \exp \left( it \frac{2n}{\pi} \eta_n \right) \right] = \exp \left( -\frac{t^2}{2} \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} \left( u(X f(y_{j,k})) \right)^2 \left( \int_{A_{j,k}} g \, d\mu \right) \right) 
\]

\[
\rightarrow \exp \left( -\frac{t^2}{2} \|u(X f)g\|^2 \right)
\]

as \( n \to \infty \), which equips us to prove our main theorem for the chapter.

**Proof of Theorem 4.0.21.** By returning to a previous result in Proposition 4.0.23, we note that this gave us convergence of the form

\[
E \left[ \exp \left( it \frac{2n}{\pi} \langle N^{1/2}(Y_{n,m,N}f - X_{n,m}f), g \rangle \right) \right] \rightarrow E \left[ \exp \left( it \frac{2n}{\pi} \eta_n \right) \right]
\]

as \( N \to \infty \), since our convergence in expectation was assured for all \( t \) and we’ve taken our value
to incorporate $\frac{2^n}{\pi}$. Then, making use of Proposition 4.0.24, we get that

$$E \left[ \exp \left( it \frac{2^n}{\pi} \eta_n \right) \right] \to \exp \left( -\frac{t^2 \|u(Xf)\|^2}{2} \right)$$

as $n \to \infty$. We now simply conclude that

$$\lim_{n \to \infty} \lim_{N \to \infty} E \left[ e^{i \left( \frac{2^n N^{1/2}}{\pi} (Y_{n,N} f - X_n f) \right)_{L^2(\mathbb{Z})}} \right] = e^{-\frac{1}{2} \|u(Xf)g\|^2_{L^2(\mathbb{Z})}}.$$

Now, for most of this section we have allowed $f \in C(\overline{D})$ to take the place of our random variable $\Phi : \Omega \to C(\overline{D})$. Now we come to combine the previous result with the fact that $\Phi$ is itself not deterministic.

**Proof of Theorem 4.0.20.** We will proceed directly from the result given to us from work on deterministic attenuation functions by conditioning on the random variable $\Phi$ itself, and so

$$\lim_{n \to \infty} \lim_{N \to \infty} E \left[ e^{i \left( \frac{2^n N^{1/2}}{\pi} (Y_{n,N} \Phi - X_n \Phi) \right)_{L^2(\mathbb{Z})}} \right] = \lim_{n \to \infty} \lim_{N \to \infty} E \left[ e^{i \left( \frac{2^n N^{1/2}}{\pi} (Y_{n,N} \Phi - X_n \Phi) \right)_{L^2(\mathbb{Z})}} \left| \Phi \right. \right]$$

$$= E \left[ \lim_{n \to \infty} \lim_{N \to \infty} E \left[ e^{i \left( \frac{2^n N^{1/2}}{\pi} (Y_{n,N} \Phi - X_n \Phi) \right)_{L^2(\mathbb{Z})}} \left| \Phi \right. \right] \right]$$

$$= E \left[ e^{-\frac{1}{2} \|u(X\Phi)g\|^2_{L^2(\mathbb{Z})}} \right],$$

where we appeal to the dominated convergence theorem to pass the limit into the outside expectation from the boundedness of the integrand.
We can think of this result as a function of $\Phi$ that becomes deterministic for any continuous function $f$ we substitute in its place. To illustrate the meaning of this result, we shall denote

$$\mathcal{H}(\Phi) = e^{-\frac{1}{2}u(X\Phi)g\|_{L^2(\mathcal{Z})},}$$

and see that, for any deterministic, continuous attenuation function $f$, we can make use of our construction of $W$ as a dually $L^2(\mathcal{Z}) \times \Omega$ measurable function that is independent of $\Phi$ to write

$$E \left[ e^{it\langle uf(x)f\rangle} \right] = E \left[ e^{itW((uf(x)f)g)} \right]$$

$$= e^{-\frac{1}{2}u(X\Phi)g\|_{L^2(\mathcal{Z})}}$$

$$= \mathcal{H}(f)$$

and so,

$$E \left[ E \left[ e^{it\langle u\phi f\rangle W.g, f\rangle} \Phi \right] \right] = E \left[ E \left[ e^{itW((u\phi f(\cdot))g, \cdot)} \Phi \right] \right]$$

$$= E[\mathcal{H}(\Phi)].$$

In other words, what this implies is that

$$\frac{2^n N^{1/2}}{\pi}(Y_{n,N} f - X_n f) \Rightarrow (u^*(X\Phi))W,$$

for $u(x) = -(1 - e^x)^{1/2}$, which gives us the precise relationship between our observations and a precise scan in accounting for the effects of the central limit theorem but not the law of large numbers. We therefore conclude with presenting our final model for observations approximating the X-ray transform of a random attenuation function $\Phi$ as

$$Y = X_n \Phi + \frac{\pi}{2^n N^{1/2}} u^*(X\Phi)W.$$
CHAPTER 5: CONCLUSIONS

As we have treated the problem under various conditions, we will now take care to restate our results under each particular set of assumptions. In the first of these, we assumed our body function to be a simple random variable taking values in $T = \{f_j\}$ and set $u \equiv \varepsilon$. Here we found that for

$$\beta_j = p_j \exp \left( \frac{2\varepsilon(X_f) - \|X_f\|^2}{2\varepsilon^2} \right),$$

our conditional probabilities became

$$P(\Phi = f_j | Y) = \frac{\beta_j}{\sum_{k=1}^{N} \beta_k} \text{ a.s.}$$

For our other full result, we return to the simple random variable form of our body function to instead make our noise term more interesting. Specifically, we let $u(\xi) = -(e^\xi - 1)^{1/2}$ but removed the additive $X\Phi$ term in $Y$ and found a promisingly similar ratio result which, while having no closed form, was nevertheless similarly acknowledged to converge a.s. There may yet prove to be further results in this setting if we find what can be determined of the rate of convergence, given that the exponents ensure a quick parsing of the more likely forms.

Much of our subsequent efforts were in formalizing the model used by treating both the detection rate of photons and the form of the attenuation function as random variables. This allowed for the derivation of a precise relationship between the white noise term for these random effects by

$$Y_{n,N}\Phi = X_n\Phi + \frac{\pi}{2nN^{1/2}}u^s(X\Phi)W,$$

where $u(x) = -(1 - e^x)^{1/2}$. We note, encouragingly, that in the limit, this model agrees with
the law of large numbers model. This too was formalized in the previous chapter and essentially assumes enough radiation that there is little error in taking the mean of a random variable as its empirical value. We demonstrated the precise mode of convergence of both these models. It is our belief that these more precise models may provide more accurate mesoscopic insight as to the actual amount of data inferred from X-ray scans.
Lemma A.0.25. Assume $\Phi$ is a random variable from $(\Omega, \mathcal{F}, P)$ to $C(\overline{D})$. Let $\iota_2 : C(\overline{D}) \to L^2(D, \mathcal{B}_D, \lambda)$ denote the inclusion map such that $\iota_2(g) = g$ for all $g \in C(\overline{D}) \subset L^2(D, \mathcal{B}_D, \lambda)$. Then

$$\iota_2 \circ \Phi : \Omega \to L^2(D)$$

is a random variable.

Proof. Our proof requires the composition outlined above to be a measurable function. The random variable $\Phi$ is assumed to be measurable from $(\Omega, \mathcal{F}, P)$ to $C(\overline{D})$. As for $\iota_2$, since we have equipped both $C(\overline{D})$ and $L^2(D)$ with their respective Borel $\sigma$-algebras, it suffices to show that $\iota_2$ is continuous, since continuous functions are Borel measurable.

To this end, let $\{f_n\} \subset C(\overline{D})$ be a sequence converging to $f$ in $C(\overline{D})$. Note that $\iota_2(g) = g$ for all $g \in C(\overline{D}) \subset L^2(D)$, so our task is to show that $f_n \to f$ in $L^2(D)$ as $n \to \infty$. Convergence in $C(\overline{D})$ grants us uniform convergence of the sequence, so for $\varepsilon > 0$ there exists an $N_\varepsilon \in \mathbb{N}$ such that for all $x \in \overline{D}$, $|f(x) - f_n(x)| \leq \varepsilon$ whenever $n > N_\varepsilon$. Therefore,

$$\|f - f_n\|^2_{L^2(D)} = \int_D |f - f_n|^2 \, d\lambda \leq \int_D \varepsilon^2 \, d\lambda = \varepsilon^2 \lambda(D),$$

and so we have established convergence in $L^2(D)$ and the continuity of $\iota_2$. Thus the composition $\iota_2 \circ \Phi : \Omega \to L^2(D)$ is measurable as the composition of two measurable functions. \hfill \Box

Lemma A.0.26. Assume $\Phi$ is a random variable from $(\Omega, \mathcal{F}, P)$ to $C(\overline{D})$. Let $\iota_\infty : C(\overline{D}) \to L^\infty(D)$ denote the inclusion map such that $\iota_\infty(g) = g$ for all $g \in C(\overline{D}) \subset L^\infty(D)$. Then

$$\iota_\infty \circ \Phi : \Omega \to L^\infty(D)$$

69
is a random variable.

**Proof.** This will follow a similar form as the previous proof. Once again we let \( \{f_n\} \subset C(\overline{D}) \) be a sequence converging to \( f \) in \( C(\overline{D}) \) and note that \( \iota_\infty(g) = g \) for all \( g \in C(\overline{D}) \subset L^\infty(D) \).

If we fix \( \varepsilon > 0 \), there exists an \( N_\varepsilon \in \mathbb{N} \) such that for all \( x \in \overline{D} \), \( |f(x) - f_n(x)| \leq \varepsilon \) whenever \( n > N_\varepsilon \). Convergence in \( L^\infty(D) \) requires convergence in the essential bound, but as all \( \{f_n\} \) and \( f \) are continuous on a compact space, their essential bound equates to their maximum and so, for \( n > N_\varepsilon \),

\[
\|f - f_n\|_{L^\infty(D)} = \max\{|f(x) - f_n(x)| : x \in \overline{D}\} < \varepsilon,
\]

and we have once again established continuity of our inclusion map. Since continuous functions are Borel measurable and \( \Phi \) was assumed to be measurable from the beginning, the composition \( \iota_\infty \circ \Phi : \Omega \to L^\infty(D) \) is a random variable. \( \square \)

**Lemma A.0.27.** Let \( u : \mathbb{R} \to \mathbb{R} \) be continuous and define \( u^* : C(\overline{Z}) \to C(\overline{Z}) \) by \( u^*f = u(f(\cdot)) \).

Then \( u^* \) is continuous.

**Proof.** Fix \( \varepsilon > 0 \) and let \( \{f_n\} \subset C(\overline{Z}) \) be a sequence of functions such that \( f_n \to f \) uniformly as \( n \to \infty \). Our goal is to prove that \( u^*f_n \to u^*f \) in \( C(\overline{Z}) \), which boils down to showing that \( u(f_n(x)) \to u(f(x)) \) uniformly as \( n \to \infty \). Since \( \overline{Z} \) is a compact set, there exists a bound \( M > 0 \) such that \( |f(x)| < M \) for all \( x \in \overline{Z} \). Moreover, since \( f_n \to f \) uniformly, there exists an index \( N_1 \in \mathbb{N} \) such that when \( n > N_1 \), we know that \( |f_n(x) - f(x)| < 1 \) for all \( x \in \overline{Z} \). Therefore, \( |f_n(x)| < M + 1 \) for all \( x \in \overline{Z} \) whenever \( n > N_1 \).

We now define \( \tilde{u} := u|_{[-M-1,M+1]} \) and note that \( \tilde{u} \) is uniformly continuous. Therefore, there exists a \( \delta_\varepsilon > 0 \) such that for \( x, y \in [-M - 1, M + 1] \subset \mathbb{R} \), we have that if \( |x - y| < \delta_\varepsilon \) then \( |\tilde{u}(x) - \tilde{u}(y)| < \varepsilon \). As a final connection, we let \( N_\varepsilon \in \mathbb{N} \) be an index such that when \( n > N_\varepsilon \),

\[
|f_n(x) - f(x)| < \varepsilon \text{ for all } x \in \overline{Z}.
\]

70
Now, let $N = \max\{N_1, N_2\}$ and take $n > N$. We note that $|f_n(x)| < M + 1$ and $|f_n(x) - f(x)| < \delta \varepsilon$ for all $x \in \mathbb{Z}$. Thus, $|\tilde{u}(f_n(x)) - \tilde{u}(f(x))| = |u(f_n(x)) - u(f(x))| \leq \varepsilon$ for all $x \in \mathbb{Z}$ and we have shown that $u^* f_n \to u^* f$ in $C(\mathbb{Z})$.

Lemma A.0.28. Let $X$ be the X-ray transform as previously defined. If we have a random variable $\Phi : \Omega \to L^2(D)$, then $X\Phi$ is also an $L^2(Z, B_\mathbb{Z}, \mu)$-valued random variable.

Proof. Once again, the goal is to show that the composition $X \circ \Phi$ is a measurable function. We already know that $\Phi$ is a measurable from $(\Omega, \mathcal{F}, P) \to L^2(D, B_D, \lambda)$. We then appeal to the fact that our measure space $L^2(Z, B_\mathbb{Z}, \mu)$ is specifically equipped with its Borel $\sigma$-algebra, which is generated by open sets. The X-ray transform, as a continuous function, is Borel measurable. Thus we have $X\Phi = X \circ \Phi : \Omega \to L^2(Z)$ as a composition of measurable functions and we are done.

Theorem A.0.29. If $\mathcal{P}$ is a $\pi$-system and $\mathcal{L}$ is a $\lambda$-system that contains $\mathcal{P}$, then $\sigma(\mathcal{P}) \subset \mathcal{L}$.

This is Theorem 2.1.2 from [11].

Theorem A.0.30. Let $H$ be a real separable Hilbert space. Let $W = \{W(h) : h \in H\}$ be a centered Gaussian process defined on a probability space $(\Omega, \mathcal{F}, P)$ such that $E[W(g)W(h)] = \langle g, h \rangle_H$ for all $g, h \in H$. There exists a stochastic process $\tilde{W} = \{\tilde{W}(h) : h \in H\}$ defined on $(\Omega, \mathcal{F}, P)$ such that $\tilde{W}(h) = W(h)$ a.s. for all $h \in H$, and $(h, \omega) \mapsto \tilde{W}(h, \omega)$ is a Borel measurable function from $H \times \Omega$ to $\mathbb{R}$.

Proof. Let $\{h_j\}_{j=1}^\infty$ be an orthonormal basis of $H$. Define $W_n : H \times \Omega \to \mathbb{R}$ by

$$W_n(h, \omega) = \sum_{j=1}^n W(h_j, \omega) \langle h, h_j \rangle_H.$$ 

Then each $W_n$ is a Borel measurable function and, for each fixed $h \in H$, we have $W_n(h) \to W(h)$ in $L^2(\Omega)$. In fact, using [11, Theorem 2.5.4], we have that $W_n(h) \to W(h)$ a.s.
By [6, Exercise 2.3],
\[ A = \{ (h, \omega) : \lim_{n \to \infty} W_n(h, \omega) \text{ exists} \} \]
is measurable. Thus, we may define \( \widetilde{W} : H \times \Omega \to \mathbb{R} \) to be the pointwise limit of \( \{ W_n 1_A \}_{n=1}^{\infty} \). By [6, Corollary 2.9], the function \( \widetilde{W} \) is measurable.

Now let \( h \in H \) be given. Choose \( \Omega^* \in \mathcal{F} \) such that \( P(\Omega^*) = 1 \) and \( W_n(h, \omega) \to W(h, \omega) \) for all \( \omega \in \Omega^* \). Fix \( \omega \in \Omega^* \). By the definition of \( A \), we have \( (h, \omega) \in A \). Thus, \( (W_n 1_A)(h, \omega) = W_n(h, \omega) \).

It follows that \( \widetilde{W}(h, \omega) = W(h, \omega) \). This shows that \( \widetilde{W}(h) = W(h) \) a.s. \( \square \)

The following result is contained in Proposition 2.11(b) of [6].

**Proposition A.0.31.** Suppose \( (M, \mathcal{M}, \mu) \) is a measure space. If \( \{ f_n \} \) is a sequence of measurable functions on \( M \) and \( f_n \to f \) \( \mu \) pointwise, then \( f \) is measurable.

The following is Theorem 5.5.7 of [11].

**Theorem A.0.32.** If \( \mathcal{F}_n \) is an increasing sequence of \( \sigma \)-algebras, \( \mathcal{F}_\infty = \sigma(\bigcup_{n=1}^{\infty} \mathcal{F}_n) \), and \( X \in L^1(\Omega) \), then \( E[X \mid \mathcal{F}_n] \to E[X \mid \mathcal{F}_\infty] \) a.s. and in \( L^1(\Omega) \).

**Proposition A.0.33.** The white noise \( W \) is a linear operator from \( L^2(\mathbb{Z}) \) to \( L^2(\Omega) \).
Proof. Let $f$ and $g$ be in $L^2(Z)$. Then by the properties delineated in the introduction,

\[
E|W(f + g) - W(f) - W(g)|^2 \\
= E|W(f + g)|^2 - 2E[W(f + g)(W(f) + W(g))] + E|W(f) + W(g)|^2 \\
= E|W(f + g)|^2 - 2E[W(f + g)W(f)] - 2E[W(f + g)W(g)] \\
+ E|W(f)|^2 + 2E[W(f)W(g)] + E|W(g)|^2 \\
= \|f + g\|^2 - 2\langle f + g, f \rangle - 2\langle f + g, g \rangle + \|f\|^2 + 2\langle f, g \rangle + \|g\|^2 \\
= \|f + g\|^2 - 2\|f\|^2 - 2\langle f, f \rangle - 2\|g\|^2 + \|f\|^2 + 2\langle f, g \rangle + \|g\|^2 \\
= \|f + g\|^2 - \|f\|^2 - 2\langle f, g \rangle - \|g\|^2 \\
= 0.
\]

Thus, $W(f + g) - W(f) - W(g) = 0$ a.s. That is, $W(f + g) = W(f) + W(g)$ a.s. A similar process can be employed to demonstrate $W(\alpha f) = \alpha W(f)$ a.s., and so we conclude that white noise is linear.

Proposition A.0.34. The white noise $W$ is a continuous operator from $L^2(Z)$ to $L^2(\Omega)$. That is, if $g_n \to g$ in $L^2(Z)$, then $W(g_n) \to W(g)$ in $L^2(\Omega)$.

Proof. Let $g_n \to g$ in $L^2(Z)$. Then we have

\[
E|W(g) - W(g_n)|^2 = E|W(g - g_n)|^2 = \|g - g_n\|^2 \to 0,
\]

so $W(g_n) \to W(g)$ in $L^2(\Omega)$.

Proposition A.0.35. Let $X$ be a discrete random variable taking values in $T = \{x_k : k \in \mathbb{N}\}$ and let $Z$ be a random variable taking values in a \(\sigma\)-finite measure space \((M, \mathcal{M}, \mu)\). Assume there
exists a measurable function $\phi : T \times M \to \mathbb{R}$ such that

$$E[f(X, Z)] = \int_M \sum_{k \in \mathbb{N}} f(x_k, z) \phi(x_k, z) \mu(dz),$$

for all bounded, measurable $f$.

Let $\psi$ be a measurable function such that $E|\psi(X)| < \infty$, and define

$$\Phi(z) = \frac{\sum_k \psi(x_k) \phi(x_k, z)}{\sum_k \phi(x_k, z)},$$

when the denominator is positive, and $\Phi(z) = 0$ otherwise. Then $E[\psi(X) \mid Z] = \Phi(Z)$ a.s.

**Proof.** First note that the random variable $\Phi(Z)$ defined above is $\sigma(Z)$-measurable, so our task is to demonstrate that $E[\Phi(Z)1_A] = E[\psi(X)1_A]$ for all $A \in \sigma(Z)$. An arbitrary element from $\sigma(Z)$ takes the form $A = \{Z \in B\}$ for some $B \in \mathcal{M}$. And so,

$$E[\Phi(Z)1_A] = E[\Phi(Z)1_B(Z)] = \int_M \sum_{k \in \mathbb{N}} [\Phi(z)1_B(z)] \phi(x_k, z) \mu(dz)$$

$$= \int_B \Phi(z) \sum_{k \in \mathbb{N}} \phi(x_k, z) \mu(dz)$$

$$= \int_{B \cap C} \left( \Phi(z) \sum_{k \in \mathbb{N}} \phi(x_k, z) \right) \mu(dz),$$

where $C = \{z : \sum_{k=1}^{\infty} \phi(x_k, z) > 0\}$. Since $X$ is discrete and $\phi$ nonnegative, we note that $\phi(x_k, z) = 0$ for all $z \in \mathcal{C}^c$. Moreover, for every $z \in C$ we get that $\Phi(z) \sum_{k \in \mathbb{N}} \phi(x_k, z) =$
\[ \sum_{k \in \mathbb{N}} \psi(x_k) \phi(x_k, z) \] from the definition of \( \Phi \) and so,

\[
E[\Phi(Z)1_A] = \int_{B \cap C} \sum_{k \in \mathbb{N}} \psi(x_k) \phi(x_k, z) \mu(dz)
\]

\[ = \int_{B} \sum_{k \in \mathbb{N}} \psi(x_k) \phi(x_k, z) \mu(dz) \]

\[ = \int_{\mathcal{M}} \sum_{k \in \mathbb{N}} [\psi(x_k)1_B(z)] \phi(x_k, z) \mu(dz) \]

\[ = E[\psi(X)1_B(Z)] \]

\[ = E[\psi(X)1_A]. \]

Having demonstrated that the two expectations agree, we are done.

**Corollary A.0.36.** If \( X, Z, \) and \( \phi \) are as defined in the previous lemma, then

\[
P(X = x_j \mid Z) = \frac{\phi(x_j, Z)}{\sum_{k=1}^{\infty} \phi(x_k, Z)} \text{ a.s.}
\]

**Proof.** Recall that we may write \( P(X = x_j \mid Z) = E[1_{\{x_j\}}(X) \mid Z] \). If we take \( \psi = 1_{\{x_j\}} \) in Proposition A.0.35, we get \( P(X = x_j \mid X) = \Phi(Z) \), where

\[
\Phi(z) = \frac{\sum_{k=1}^{\infty} 1_{\{x_j\}}(x_k) \phi(x_k, z)}{\sum_{k=1}^{\infty} \phi(x_k, z)} = \frac{\phi(x_j, z)}{\sum_{k=1}^{\infty} \phi(x_k, z)},
\]

when the denominator is positive, and \( \Phi(z) = 0 \) otherwise. Hence, it suffice to show that \( \sum_{k \in \mathbb{N}} \phi(x_k, Z) > 0 \) a.s.

Let \( S = \{ z : \sum_{k \in \mathbb{N}} \phi(x_k, z) = 0 \} \). Then

\[
P(Z \in S) = \int_{S} \sum_{k \in \mathbb{N}} \phi(x_k, z) \mu(dz) = 0,
\]

and we are done.
The following can be found in Example IV.1.22 of [5].

**Theorem A.0.37.** Suppose that $X$ and $Y$ are independent and take values in measurable spaces $(E, \mathcal{E})$ and $(D, \mathcal{D})$ respectively. Let $h : E \times D \to [0, \infty]$ be $(\mathcal{E} \otimes \mathcal{D}, \mathcal{R})$-measurable, where

$$\mathcal{E} \otimes \mathcal{D} = \sigma(\{A \times B : A \in \mathcal{E}, B \in \mathcal{D}\})$$

and $\mathcal{R}$ is the $\sigma$-algebra of Borel sets on $\mathbb{R}$. Then

$$E[h(X, Y) \mid Y] = \tilde{h}(Y),$$

where $\tilde{h}(y) = E[h(X, y)]$ for each $y \in D$.

**Remark A.0.38.** Theorem A.0.37 is still true when $h : E \times D \to \mathbb{R}$ is such that $E|h(X, Y)| < \infty$, as can be seen by consider the positive and negative parts of $h$.

**Lemma A.0.39.** Let $\{X_n\}$ and $\{Y_n\}$ be sequences of random variables that converge in probability as $n \to \infty$ to $X$ and $Y$ respectively. If $X_n = Y_n$ a.s. for all $n \in \mathbb{N}$, then $X = Y$ a.s.

**Proof.** Let $\{X_n\}$ and $\{Y_n\}$ be sequences of random variables. Assume that $X_n \to X$ in probability as $n \to \infty$ and $Y_n \to Y$ in probability as $n \to \infty$. Also assume that $X_n = Y_n$ a.s. for all $n \in \mathbb{N}$. We shall denote $A_n = \{\omega : X_n(\omega) \neq Y_n(\omega)\}$ and $A = \bigcup_{n=1}^{\infty} A_n = \{\omega : X_n(\omega) \neq Y_n(\omega)\}$ for some $n \in \mathbb{N}$, and it follows that $P(A) = 0$. Then, for $\varepsilon > 0$, we can write $\{|Y_n - X| > \varepsilon\} = (\{|Y_n - X| > \varepsilon\} \cap A) \cup (\{|Y_n - X| > \varepsilon\} \cap A^c)$. Therefore,

$$P(|Y_n - X| > \varepsilon) = P(|Y_n - X| > \varepsilon \cap A) + P(|Y_n - X| > \varepsilon \cap A^c) \leq P(A) + P(|X_n - X| > \varepsilon),$$
which as we take \( n \to \infty \), equals zero. We conclude that \( Y_n \) goes to \( X \) in probability. Now, since \( Y_n \to Y \) in probability, there exists a subsequence \( \{Y_{n_j}\} \) such that \( Y_{n_j} \to Y \) a.s. Similarly, there exists a subsequence of \( \{Y_{n_j(k)}\} \subset \{Y_{n_j}\} \) such that \( Y_{n_j(k)} \to X \) a.s. Since \( \{Y_{n_j(k)}\} \) converges to both \( Y \) and \( X \) a.s., we conclude that \( Y = X \) a.s. \( \square \)

**Lemma A.0.40.** For each \( j \in \{1, \ldots, N\} \), suppose \( \{\xi_{j,m}\}_{m=1}^{\infty} \) is a sequence of random variables with \( \xi_{j,m} \to \zeta_j \) in probability as \( m \to \infty \). Then there exists a subsequence \( \{m_\ell\} \) such that \( \xi_{j,m_\ell} \to \zeta_j \) a.s. as \( \ell \to \infty \), for all \( j \in \{1, \ldots, N\} \).

**Proof.** Since \( \xi_{1,n} \to \zeta_1 \) in probability as \( n \to \infty \), there exists a subsequence \( \{n(\ell_1)\} \) such that \( \xi_{1,n(\ell_1)} \to \zeta_1 \) a.s. as \( \ell_1 \to \infty \). Since \( \xi_{2,n(\ell_1)} \to \zeta_2 \) in probability as \( \ell_1 \to \infty \), there exists a subsequence \( \{\ell_1(\ell_2)\} \) such that \( \xi_{2,n(\ell_1(\ell_2))} \to \zeta_2 \) a.s. as \( \ell_2 \to \infty \). Since this is a subsequence, we still have \( \xi_{1,n(\ell_1(\ell_2))} \to \zeta_1 \) a.s. as \( \ell_2 \to \infty \).

Continuing in this fashion, we can construct nested subsequences such that

\[
\xi_{j,n(\ell_1(\ell_2(\cdots(\ell_N)\cdots)))}\to \zeta_j \text{ a.s.,}
\]

as \( \ell_N \to \infty \), for each \( j \in \{1, \ldots, N\} \). Letting \( m_\ell = n(\ell_1(\ell_2(\cdots(\ell_N)\cdots))) \) finishes the proof. \( \square \)

**Lemma A.0.41.** If \( h \in L^2(S) \), then \( Y(Q_nh) \to Y(h) \) a.s.

**Proof.** Note that

\[
Y(Q_nh) = \langle X \Phi, Q_nh \rangle + W(Q_nh),
\]

and

\[
W(Q_nh) = \sum_{j=1}^{n} \langle h, h_j \rangle W(h_j) = W_n(h).
\]

Since \( Q_nh \to h \) in \( L^2(S) \), we have \( \langle X \Phi(\omega), Q_nh \rangle \to \langle X \Phi(\omega), h \rangle \) for all \( \omega \in \Omega \). We also have \( W_n(h) \to W(h) \) a.s. Thus, \( Y(Q_nh) \to Y(h) \) a.s. \( \square \)
**Corollary A.0.42.** For all $X \in L^1(\Omega)$, we have $E[X \mid \mathcal{F}_n] \to E[X \mid \mathcal{F}^Y]$ a.s. and in $L^1(\Omega)$.

**Lemma A.0.43.** For sequences of random variables $X_n$ and $Y_n$, if $Y_n \Rightarrow c$ for some constant $c$ and $X_n \Rightarrow X_\infty$, then

$$X_n + Y_n \Rightarrow X_\infty + c.$$  

In particular, if $Y_n \Rightarrow 0$, then $X_n$ and $X_n + Y_n$ have the same limit in distribution.
APPENDIX B: REAL SPACES
Lemma B.0.44. For $x, y \in \mathbb{R}$ positive,

$$\log(y) - \log(x) = x^{-1} (y - x) - \int_x^y \frac{t - x}{tx} \, dt.$$  

Proof. The result follows from a direct application of the definition of the logarithm. Alternatively, this could be derived from Taylor’s Theorem. We proceed directly,

$$\log(y) - \log(x) = \int_x^y \frac{1}{t} \, dt - \int_x^y \left( \frac{1}{t} - \frac{1}{x} \right) \, dt = \int_x^y \frac{1}{x} \, dt + \int_x^y \frac{1}{t} \, dt = x^{-1} (y - x) - \int_x^y \frac{t - x}{tx} \, dt.$$  

Lemma B.0.45. If $f \in L^\infty(D)$, then there exists $\{\varphi_n\} \subset C_c^\infty(D)$ such that $\|\varphi_n\|_u \leq \|f\|_\infty$ and $\varphi_n \to f$ in $L^2(D)$.

Proof. Note that $S(D) = C_c^\infty(D)$. All references in this proof refer to [6]. Let $g = f^{1_{\{|f| \leq \|f\|_\infty\}}}$, so that $g = f$ a.e. and $\|g\|_u = \|f\|_\infty$.

Fix $n \in \mathbb{N}$. Let $\varepsilon = 1/(16n^2\|f\|_\infty^2)$. By Theorem 7.10 (Lusin’s theorem), p. 217, we may choose $h \in C_c(D)$ and Lebesgue measurable $E \subset D$ such that $\lambda(E) < \varepsilon$, $h1_{E^c} = g1_{E^c}$, and $\|h\|_u \leq \|g\|_u = \|f\|_\infty$. Thus, $\|g - h\|_u \leq 2\|f\|_\infty$, which yields

$$\|f - h\|_2^2 = \|g - h\|_2^2 = \int_E |g - h|^2 \, d\lambda \leq 4\|f\|_\infty^2 \varepsilon.$$  

Hence, $\|f - h\|_2 \leq 1/(2n)$.  

80
Choose \( r < 1 \) such that \( \text{supp}(h) \subset \{|x| < r\} \). Choose \( r' < 1 - r \). Choose \( \psi \in C_c^\infty(D) \) such that \( \text{supp}(\psi) \subset \{|x| < r'\}, \psi \geq 0 \), and \( \int \psi \, d\lambda = 1 \). For \( t \in (0, 1] \), let \( \psi_t(x) = t^{-2}\psi(x/t) \), so that \( \text{supp}(\psi_t) \subset \{|x| < tr'\} \). Note that if \( x \in \text{supp}(h) \) and \( y \in \text{supp}(\psi_t) \), then \( |x + y| < r + tr' \leq r + r' < 1 \). Thus, by Proposition 8.6(b), p. 240, \( \text{supp}(h \ast \psi_t) \subset \{|x| \leq r \} \). By Proposition 8.10, p. 242, \( h \ast \psi_t \in C^\infty(D) \). Thus, \( h \ast \psi_t = S(D) \). Moreover, by Theorem 8.14(d), p. 242, \( h \ast \psi_t \to h \) uniformly as \( t \to 0 \). This implies that \( h \ast \psi_t \to h \) in \( L^2(D) \) as \( t \to 0 \). We may therefore choose \( t_0 \) such that \( \|h \ast \psi_{t_0} - h\|_2 < 1/(2n) \). Define \( \varphi_n = h \ast \psi_{t_0} \).

Note that
\[
\| f - \varphi_n \|_2 \leq \| f - h \|_2 + \| h - h \ast \varphi_{t_0} \|_2 < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}.
\]

Thus, \( \varphi_n \to f \) in \( L^2(D) \). Moreover, using Proposition 8.8, p. 241, it follows that \( \| \varphi_n \|_u \leq \|h\|_\infty \| \psi_{t_0} \|_1 \leq \|f\|_\infty \).

**Theorem B.0.46.** If \( f \in L^\infty(D) \), then \( Xf \in L^\infty(Z) \) and \( \|Xf\|_\infty \leq 2\|f\|_\infty \).

**Proof.** Let \( f \in L^\infty(D) \subset L^2(D) \). Choose \( \{\varphi_n\} \) as in the previous lemma. Then \( X\varphi_n \to Xf \) in \( L^2(Z) \). Choose a subsequence such that \( X\varphi_{n_j} \to Xf \) a.e. Note that for any \( \varphi \in S(D) \), we have \( \|X\varphi\|_u \leq 2\|\varphi\|_u \). Thus, by the previous lemma, \( |X\varphi_{n_j}(x)| \leq 2\|f\|_\infty \) for all \( x \in D \). It follows that \( |Xf| \leq 2\|f\|_\infty \) a.e. \( \square \)
LIST OF REFERENCES


