5-2011

Explicit Lp-norm estimates of infinitely divisible random vectors in Hilbert spaces with applications

Matthew D Turner
mturne20@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Matthew D Turner entitled "Explicit Lp-norm estimates of infinitely divisible random vectors in Hilbert spaces with applications." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

Jan Rosinski, Major Professor

We have read this dissertation and recommend its acceptance:

Xia Chen, Jie Xiong, Mary Leitnaker

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Matthew D. Turner entitled "Explicit Lp-norm estimates of infinitely divisible random vectors in Hilbert spaces with applications." I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

Jan Rosinski, Major Professor

We have read this dissertation
and recommend its acceptance:

Xia Chen

Jie Xiong

Mary Leitnaker

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Explicit Lp-norm estimates of infinitely divisible random vectors in Hilbert spaces with applications

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Matthew D. Turner
May 2011

Dedication

I dedicate this dissertation to my wife Heather, sons Jackson and Bryson, parents Tony and Kathy, and brother Mark. Their support and encouragement have been my motivation.

Acknowledgments

I would like to express my sincere gratitude to those individuals at the University of Tennessee who have made this dissertation possible. I am most grateful to my advisor, Dr. Jan Rosiński. It has been my privilege and pleasure to work under his guidance. I would also like to thank the members of my committee, Dr. Xia Chen, Dr. Jie Xiong, and Dr. Mary Leitnaker for their willingness to serve and their critical review of my work. Last, but certainly not least, I would like to thank Dr. William Wade and Mrs. Pam Armentrout for their unending advice and mentoring.

Abstract

I give explicit estimates of the Lp-norm of a mean zero infinitely divisible random vector taking values in a Hilbert space in terms of a certain mixture of the L2- and Lp-norms of the Levy measure. Using decoupling inequalities, the stochastic integral driven by an infinitely divisible random measure is defined. As a first application utilizing the Lp-norm estimates, computation of Ito Isomorphisms for different types of stochastic integrals are given. As a second application, I consider the discrete time signal-observation model in the presence of an alpha-stable noise environment. Formulation is given to compute the optimal linear estimate of the system state.

Contents

List of Figures

List of Algorithms

Chapter 1

Infinitely Divisible Distributions

1.1 Introduction

When producing models of an evolving dynamical system, one is often faced with the challenge of which effects to include in the model and which effects may reasonably be ignored to accurately determine the state of the system. An alternate approach is to capture these unmodeled effects as random variables or stochastic processes, which are often assumed to be Gaussian in the classical literature. Many researchers have sought extensions to such models by replacing the Gaussian assumption, as there is a need for models capturing observed heavy tailed data exhibiting high variability and/or long range dependency. *Infinitely divisible distributions* have often been utilized for such modeling. The advantage of infinitely divisible models is their computability in terms of the Lévy-Khintchine triplet parameterization. Difficulties arise, however, when such distributions have infinite variance, since L^2-theory and orthogonality are not applicable. Instead, we seek computation of the L^p-norm in terms of the Lévy measure.

Infinitely divisible distributions are a broad family of distributions containing many named distributions. For example, the geometric, negative binomial, and Poisson distribution are all discrete distributions in this family. So too are the continuous normal, Cauchy, gamma, F, lognormal, Pareto, Student's t, Weibull, α-stable, and tempered α-stable distributions. The following theorem characterizes infinitely divisible random vectors and will be the primary tool used for investigation throughout. For $x \in H$, a real Hilbert space, define

$$[\![x]\!] \stackrel{def}{=} \frac{x}{\max\{\|x\|, 1\}}.$$

Whenever $H = \mathbb{R}$, we have

$$[\![x]\!] = \begin{cases} x & \text{if } |x| \leq 1 \\ \operatorname{sign}(x) & \text{if } |x| > 1. \end{cases}$$

Theorem 1.1.1 (Lévy-Khintchine representation)**.** *The characteristic function of an infinitely divisible random vector X taking values in a Hilbert space H can be written as*

$$\mathbb{E}e^{i\langle u, X\rangle} = \exp\left\{ i\langle u, b\rangle - \frac{1}{2}\langle u, \Sigma u\rangle + \int_H \left(e^{i\langle u,x\rangle} - 1 - i\langle u, [\![x]\!]\rangle\right) Q(dx) \right\}, \qquad (1.1)$$

where $u, b \in H$, Σ is a nonnegative symmetric operator on H, and Q is a measure on H such that $Q(\{0\}) = 0$ and $\int_H \|[\![x]\!]\|^2 Q(dx) < \infty$. Moreover, the triplet (b, Σ, Q) completely determines the distribution of X and this triplet is unique.

We call (b, Σ, Q) the Lévy-Khintchine triplet of X. When $Q \equiv 0$, X is Gaussian with mean b and covariance matrix Σ and results are well-known. It is the non-Gaussian case $\Sigma \equiv 0$ that is of interest to us in the following work. When studying infinitely divisible distributions and their associated random vectors, the characteristic function will be our primary tool. If we define the exponent of (1.1) by

$$C(u) \stackrel{def}{=} i\langle u, b\rangle - \frac{1}{2}\langle u, \Sigma u\rangle + \int_H \left(e^{i\langle u,x\rangle} - 1 - i\langle u, [\![x]\!]\rangle\right) Q(dx),$$

then C is called the *cumulant of* X and we have $\mathbb{E}e^{i\langle u,X\rangle} = e^{C(u)}$. Moreover, if X is infinitely divisible with Lévy-Khintchine triplet $\left(b^X, \Sigma^X, Q^X\right)$ and cumulant $C_X(u)$, Y is infinitely divisible with Lévy-Khintchine triplet $\left(b^Y, \Sigma^Y, Q^Y\right)$ and cumulant $C_Y(u)$, and X and Y are independent, then $X + Y$ is also infinitely divisible with cumulant $C_X(u) + C_Y(u)$, and hence, has Lévy-Khintchine triplet $\left(b^X + b^Y, \Sigma^X + \Sigma^Y, Q^X + Q^Y\right)$. As an immediate corollary of the Lévy-Khintchine representation, we have that the family of infinitely divisible random vectors are closed under continuous linear transformations and, in particular, projections of infinitely divisible random vectors are infinitely divisible. More precisely:

Corollary 1.1.2. *Let $X \in H$ be an infinitely divisible random vector with Lévy-Khintchine triplet (b, Σ, Q). If $F : H \to H_1$ is a continuous linear operator from the Hilbert space H into the Hilbert space H_1, then $FX \in H_1$ is also an infinitely divisible random vector with Lévy-Khintchine triplet (b_F, Σ_F, Q_F), where*

$$b_F \stackrel{def}{=} Fb + \int_H (\llbracket Fx \rrbracket - F\llbracket x \rrbracket)\, Q(dx), \quad \Sigma_F = F\Sigma F^*,$$

and for every $B \in \mathcal{B}(H_1)$,

$$Q_F(B) \stackrel{def}{=} Q\{x \in H : Fx \in B \setminus \{0\}\}.$$

Before proving the corollary, we make a few remarks. First, if Q is a symmetric Lévy measure on H, then Q_F is a symmetric Lévy measure on H_1. Second, the integrand in the definition b_F is an odd function. Therefore, if $b = 0$ and Q is symmetric, then $b_F = 0$ also. We point out these facts since the majority of the examples we consider will make one (or both) of these assumptions.

Proof of Corollary 1.1.2. Let $F : H \to H_1$ be a continuous linear operator and let $u \in H_1$. Then

$$\begin{aligned}
&\mathbb{E}e^{i\langle u, FX\rangle} = \mathbb{E}e^{i\langle F^*u, X\rangle} \\
&\quad = \exp\left\{ i\langle F^*u, b\rangle - \frac{1}{2}\langle F^*u, \Sigma F^*u\rangle + \int_H \left(e^{i\langle F^*u, x\rangle} - 1 - i\langle F^*u, \llbracket x \rrbracket\rangle\right) Q(dx) \right\} \\
&\quad = \exp\left\{ i\langle u, Fb\rangle - \frac{1}{2}\langle u, F\Sigma F^*u\rangle + \int_H \left(e^{i\langle u, Fx\rangle} - 1 - i\langle u, F\llbracket x \rrbracket\rangle\right) Q(dx) \right\} \\
&\quad = \exp\left\{ i\langle u, Fb\rangle + \int_H (i\langle u, \llbracket Fx \rrbracket\rangle - i\langle u, F\llbracket x \rrbracket\rangle)\, Q(dx) - \frac{1}{2}\langle u, F\Sigma F^*u\rangle \right. \\
&\qquad \left. + \int_H \left(e^{i\langle u, Fx\rangle} - 1 - i\langle u, \llbracket Fx \rrbracket\rangle\right) Q(dx) \right\} \\
&\quad = \exp\left\{ i\left\langle u, Fb + \int_H (\llbracket Fx \rrbracket - F\llbracket x \rrbracket)\, Q(dx) \right\rangle - \frac{1}{2}\langle u, F\Sigma F^*u\rangle \right. \\
&\qquad \left. + \int_{H_1} \left(e^{i\langle u, x\rangle} - 1 - i\langle u, \llbracket x \rrbracket\rangle\right) Q_F(dx) \right\} \\
&\quad = \exp\left\{ i\langle u, b_F\rangle - \frac{1}{2}\langle u, \Sigma_F u\rangle + \int_{H_1} \left(e^{i\langle u, x\rangle} - 1 - i\langle u, \llbracket x \rrbracket\rangle\right) Q_F(dx) \right\}.
\end{aligned}$$

□

In practice, the normal distribution is justified in its use by the central limit theorem and a popular distribution in modeling because of the ease of computations when L^2-orthogonality is applicable. Under the assumption of non-Gaussian distributions, it is often not known how the "error" should be measured. The next section addresses this question for infinitely divisible distributions. In Chapter 2, we will apply this result to obtain the Kalman filter for a discrete time signal-observation model with infinite covariance noise. In Chapter 3, we will define the stochastic integral of a stochastic field driven by an infinitely divisible random measure. Itô Isomorphisms will be derived for the stochastic integral.

1.2 L^p-norm of Hilbert space valued infinitely divisible random vectors

Let X be a mean 0 random vector taking values in a separable Hilbert space H with characteristic function given by (1.1). When X is purely Gaussian ($Q \equiv 0$), the L^p-norm of X is controlled by the covariance matrix Σ. In the non-Gaussian case, Marcus and Rosiński (2001) showed that for $X \in L^1$, the L^1-norm of X is controlled by the Lévy measure Q as

$$(0.25)l(Q) \leq \mathbb{E}\,\|X\| \leq (2.125)l(Q),$$

where the functional l of Q satisfies

$$\int_H \min\left\{\frac{\|x\|^2}{l^2}, \frac{\|x\|}{l}\right\} Q(dx) = 1.$$

The following theorem generalizes this result to obtain bounds on the L^p-norm of X. Assume that X is in L^p for given $p \geq 1$, $\mathbb{E}X = 0$, and that X does not have a Gaussian component. The characteristic function of X can be written as

$$\mathbb{E}\exp\left(i\langle u, X\rangle\right) = \exp\left(\int_H \left(e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\right) Q(dx)\right).$$

We assume throughout that Q is symmetric and later remark on removing this restriction by standard symmetrization techniques. Since Q is assumed symmetric, the characteristic function of X is

$$\mathbb{E}\exp\left(i\langle u, X\rangle\right) = \exp\left(\int_H \left(\cos\langle u, x\rangle - 1\right) Q(dx)\right).$$

It is well known that an infinitely divisible random vector X with Lévy measure Q has finite L^p-norm if and only if $\int_{\|x\|\geq 1} \|x\|^p Q(dx)$ is finite (see e.g. Sato (2002, Corollary 25.8)). Therefore the Lévy measure Q satisfies

$$\int_H \left(\|x\|^2 \mathbb{1}_{\{\|x\|<1\}} + \|x\|^p \mathbb{1}_{\{\|x\|\geq 1\}}\right) Q(dx) < \infty.$$

Let the functional l of Q be given by the solution of

$$\xi(l) \stackrel{def}{=} \int_H \left(\frac{\|x\|^2}{l^2}\mathbb{1}_{\left\{\frac{\|x\|}{l}<1\right\}} + \frac{\|x\|^p}{l^p}\mathbb{1}_{\left\{\frac{\|x\|}{l}\geq 1\right\}}\right) Q(dx) = 1. \tag{1.2}$$

We remark that

$$\|x\|^2 \mathbb{1}_{\{\|x\|<1\}} + \|x\|^p \mathbb{1}_{\{\|x\|\geq l\}} = \begin{cases} \|x\|^2 \wedge \|x\|^p & \text{if } 1 \leq p \leq 2, \\ \|x\|^2 \vee \|x\|^p & \text{if } p > 2. \end{cases}$$

We can view l as a special mixture of the L^2-norm and L^p-norm of Q. In the case of non-Gaussian infinitely divisible random vectors, the following theorem gives explicit estimates of the L^p-norm in terms of the Lévy measure Q.

Theorem 1.2.1. *Let $p \geq 1$. Assume that $X \in L^p$ is a mean 0 infinitely divisible random vector without Gaussian component, taking values in the Hilbert space H, and that X has symmetric Lévy measure Q. Then*

$$0.25 l \leq \|X\|_p \leq K(p) l \tag{1.3}$$

where

$$K(p) \stackrel{def}{=} \begin{cases} 1 + \sqrt[p]{2^{3-p}+1}, & \text{if } 1 \leq p \leq 2 \\ \sqrt[4]{4} + \sqrt[p]{1 + \frac{p(p-1)}{4}}, & \text{if } 2 < p \leq 3 \\ \sqrt[4]{4} + K_{3,p}^{1/p}\left(K_{4,p}+1\right)^{1/p}, & \text{if } 3 < p < 4 \\ 2\sqrt[4]{4}, & \text{if } p = 4 \\ K_{1,p}\left(K_{2,p}+1\right) + K_{3,p}^{1/p}\left(K_{4,p}+1\right)^{1/p}, & \text{if } p > 4, \end{cases} \tag{1.4}$$

where $K_{1,p} \stackrel{def}{=} (p+1)^{\frac{p+1}{p}} 2^{1/p}$, $K_{2,p} = 4^{1+1/p}$, $K_{3,p} = \frac{2^p \cdot 4}{\left(2^{2/(p+1)}-1\right)^{p+1}}$, $K_{4,p} = 2^{2(p+1)} 4^{p/2} (x_0+5)^{p/2}$, *and* $x_0 \approx 4.7591$ *solves* $x_0 = e\log(x_0+1)$.

We remark on important cases for the constant $K(p)$. First, it is the $1 \leq p < 2$ case that is of most interest to us, as L^p-theory must be used when working with models containing infinite covariance noise or random driving terms. It is often challenging, if not impossible, to compute such norms directly. Second,we have very nice constants for estimation of the mean, variance, skewness, and kurtosis. Constant $K(p)$ is graphed in Figure 1.1.

In preparation of the proof of Theorem 1.2.1, we follow the lead of Marcus and Rosiński (2001) and decompose X as $X = Y + Z$, where Y and Z are independent mean zero random vectors with characteristic functions

$$\mathbb{E} \exp\left(i\langle u, Y\rangle\right) = \exp\left(\int_{\|x\|<l} \left(\cos\langle u, x\rangle - 1\right) Q(dx)\right)$$

and

$$\mathbb{E} \exp\left(i\langle u, Z\rangle\right) = \exp\left(\int_{\|x\|\geq l} \left(\cos\langle u, x\rangle - 1\right) Q(dx)\right),$$

respectively. The following four lemmas provide upper and lower bounds for norms of Y and Z and will be used in the proof of Theorem 1.2.1.

Lemma 1.2.2. *We have the following upper bounds on norms of* Y*:*

i. If $1 \leq p \leq 2$, *then*

$$\|Y\|_p \leq \|Y\|_2 = \left(\int_{\|x\|<l} \|x\|^2 Q(dx)\right)^{1/2}. \tag{1.5}$$

ii. If $2 < p \leq 4$, *then*

$$\|Y\|_p \leq \|Y\|_4 = \left(\int_{\|x\|<l} \|x\|^4 Q(dx) + 3\left(\int_{\|x\|<l} \|x\|^2 Q(dx)\right)^2\right)^{1/4}. \tag{1.6}$$

iii. If $p > 4$, *then*

$$\|Y\|_p \leq K_{1,p}\left(K_{2,p}\|Y\|_2 + l\right), \tag{1.7}$$

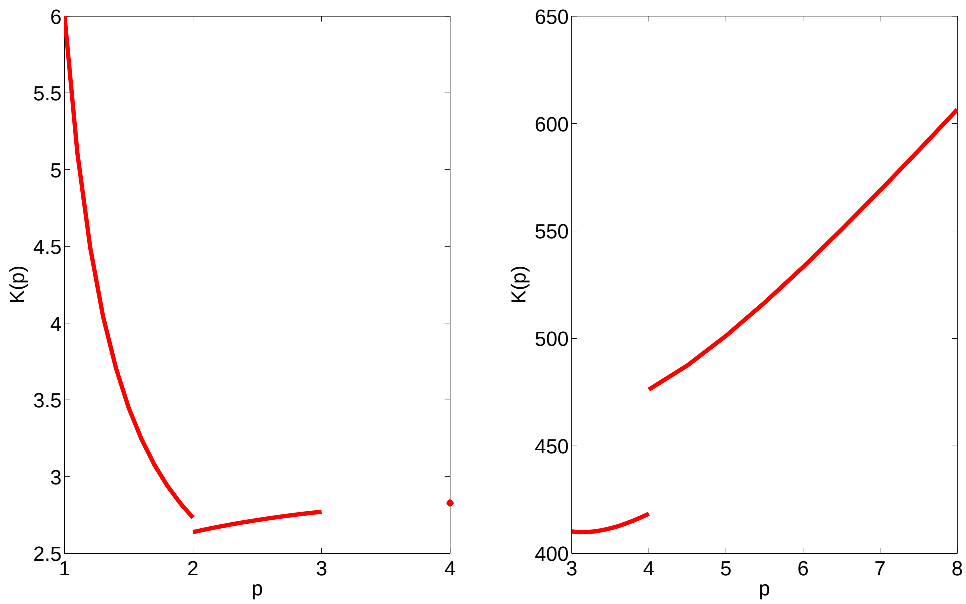

Figure 1.1: Explicit constant in L^p-norm estimate.

where $K_{1,p}$ *and* $K_{2,p}$ *are given in Theorem 1.2.1.*

Proof. (1.5) and (1.6) were proved by Marcus and Rosiński (2001, Lemma 1.1). Now let $p > 4$. Let $\{Y_t\}_{t\geq 0}$ be a Lévy process such that $Y_1 \stackrel{d}{=} Y$. Since the Lévy measure of Y, and hence Y_t, is supported on $\{\|x\| < l\}$, the sample path $t \to Y_t(\omega)$ a.s. has no jumps of magnitude larger than l on $t \in [0,1]$. So there exists $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that $\|Y_t(\omega) - Y_{t-}(\omega)\| \leq l$ for every $\omega \in \Omega_0$ and for every $t \in [0,1]$. For each $n \in \mathbb{N}$, we may write Y as the sum of n i.i.d. random vectors by

$$Y \stackrel{d}{=} \left(Y_1 - Y_{\frac{n-1}{n}}\right) + \left(Y_{\frac{n-1}{n}} - Y_{\frac{n-2}{n}}\right) + \cdots + \left(Y_{\frac{1}{n}} - Y_0\right) \stackrel{def}{=} \sum_{k=1}^{n} \Delta_{\frac{k}{n}} Y,$$

where

$$\Delta_{\frac{k}{n}} Y \stackrel{def}{=} Y_{\frac{k}{n}} - Y_{\frac{k-1}{n}}.$$

Fix $\varepsilon > 0$ and $\omega \in \Omega_0$. Since

$$\{t \in [0,1] : |X_t(\omega) - X_{t-}(\omega)| \geq l + \varepsilon\} = \emptyset,$$

standard analysis results give that there exists $N = N(\omega)$ so large that for each $n \geq N(\omega)$,

$$\left\| \Delta_{\frac{k}{n}} Y(\omega) \right\| < l + \varepsilon$$

for every $1 \leq k \leq n$. For each $n \in \mathbb{N}$, define a new i.i.d. sequence of bounded random vectors $\{Y_{k,n}\}_{k=1}^{n}$ by

$$Y_{k,n} \stackrel{def}{=} \Delta_{\frac{k}{n}} Y \, \mathbb{1}_{\left\{ \left\| \Delta_{\frac{k}{n}} Y \right\| < l + \varepsilon \right\}}.$$

For each $\omega \in \Omega_0$,

$$Y_{k,n}(\omega) = \Delta_{\frac{k}{n}} Y(\omega)$$

for every $n \geq N(\omega)$. We now have that

$$S_n \stackrel{def}{=} \sum_{k=1}^{n} Y_{k,n} \to Y \text{ a.s.},$$

since $\mathbb{P}(\Omega_0) = 1$. Observe that for fixed n, $\{Y_{k,n}\}_{k=1}^{n}$ is sequence of symmetric (since Q is assumed symmetric) i.i.d. random vectors bounded by $l + \varepsilon$. By de la Peña and Giné (1999, Theorem 1.2.5, a Hoffman-Jorgensen type inequality), for every $n \in \mathbb{N}$,

$$\|S_n\|_p \leq K_{1,p} \left(K_{2,p} \|S_n\|_2 + \left\| \max_{1 \leq k \leq n} \|Y_{k,n}\| \right\|_p \right).$$

But

$$\|S_n\|_2^2 = \mathbb{E} \left(\sum_{k=1}^{n} Y_{k,n}^2 \right) \leq \mathbb{E} \left(\sum_{k=1}^{n} \left(\Delta_{\frac{k}{n}} Y \right)^2 \right) = \mathbb{E} \left(Y^2 \right)$$

and

$$\|Y_{k,n}\| < l + \varepsilon$$

for every $1 \leq k \leq n$. Hence, for every $n \in \mathbb{N}$,

$$\|S_n\|_p < K_{1,p} \left(K_{2,p} \|Y\|_2 + l + \varepsilon \right).$$

By Fatou's lemma and the arbitrariness of ε,

$$\|Y\|_p \leq K_{1,p} \left(K_{2,p} \|Y\|_2 + l \right).$$

□

Lemma 1.2.3. *We have the following lower bounds on norms of Y:*

i. If $1 \le p \le 2$, then

$$\mathbb{E}\left\|Y\right\|^p \ge \frac{\mathbb{E}\left\|Y\right\|^2}{\left(l^2 + 3\mathbb{E}\left\|Y\right\|^2\right)^{\frac{2-p}{2}}}. \tag{1.8}$$

ii. If $p > 2$, then

$$\left\|Y\right\|_p \ge \left\|Y\right\|_2 = \left(\int_{\|x\|<l} \left\|x\right\|^2 Q(dx)\right)^{1/2}. \tag{1.9}$$

Proof. Let $1 \le p \le 2$. To show (1.8), Holder's inequality gives

$$\begin{aligned}
\mathbb{E}\left\|Y\right\|^2 &= \mathbb{E}\left\|Y\right\|^{\frac{2p}{4-p}}\left\|Y\right\|^{\frac{8-4p}{4-p}} \\
&\le \left(\mathbb{E}\left\|Y\right\|^{\frac{2p}{4-p}\frac{4-p}{2}}\right)^{\frac{2}{4-p}}\left(\mathbb{E}\left\|Y\right\|^{\frac{8-4p}{4-p}\frac{4-p}{2-p}}\right)^{\frac{2-p}{4-p}} \\
&= \left(\mathbb{E}\left\|Y\right\|^p\right)^{\frac{2}{4-p}}\left(\mathbb{E}\left\|Y\right\|^4\right)^{\frac{2-p}{4-p}}
\end{aligned}$$

and hence,

$$\mathbb{E}\left\|Y\right\|^p \ge \frac{\left(\mathbb{E}\left\|Y\right\|^2\right)^{\frac{4-p}{2}}}{\left(\mathbb{E}\left\|Y\right\|^4\right)^{\frac{2-p}{2}}}.$$

Applying (1.6) to the denominator gives

$$E\left\|Y\right\|^p \ge \frac{\left(\mathbb{E}\left\|Y\right\|^2\right)^{\frac{4-p}{2}}}{\left(l^2\mathbb{E}\left\|Y\right\|^2 + 3\left(\mathbb{E}\left\|Y\right\|^2\right)^2\right)^{\frac{2-p}{2}}} = \frac{\mathbb{E}\left\|Y\right\|^2}{\left(l^2 + 3\mathbb{E}\left\|Y\right\|^2\right)^{\frac{2-p}{2}}},$$

proving (1.8). This technique is known as Littlewood's approach. (1.9) is immediate by (1.5). □

Lemma 1.2.4. *We have the following upper bounds on norms of Z:*

i. If $1 \le p \le 2$, then

$$\mathbb{E}\left\|Z\right\|^p \le c_p \int_{\|x\|\ge l} \left\|x\right\|^p Q(dx), \tag{1.10}$$

where $c_p = 2^{3-p} + 1$. If $H = \mathbb{R}$, the constant may be taken as c_p given by (A.12) *or* (A.20) *instead.*

ii. If $2 < p \le 3$, then

$$\mathbb{E}\,\|Z\|^p \le \int_{\|x\|\ge l} \|x\|^p\, Q(dx) + \frac{p(p-1)}{4} \int_{\|x\|\ge l} \|x\|^{p-2}\, Q(dx) \int_{\|x\|\ge l} \|x\|^2\, Q(dx). \tag{1.11}$$

iii. Let $\lambda > p/x_0$. If $3 < p < 4$ or if $p > 4$, then

$$\mathbb{E}\,\|Z\|^p \le K_{3,p}\left(K_{4,p}\left(\int_{\|x\|\ge l} \|x\|^2\, Q(dx)\right)^{p/2} + \int_{\|x\|\ge l} \|x\|^p\, Q(dx)\right), \tag{1.12}$$

where $K_{3,p}$ and $K_{4,p}$ are given in Theorem 1.2.1.

iv. Let $\lambda \le p/x_0$. If $3 < p < 4$ or if $p > 4$, then

$$\mathbb{E}\,\|Z\|^p \le \max\left\{ 1 + \frac{8p}{\log\left(\frac{p}{\frac{p}{x_0}\wedge 1}\right)}, \frac{6p}{\log x_0} \right\}^p \int_{\|x\|\ge l} \|x\|^p\, Q(dx). \tag{1.13}$$

v. If $p = 4$, then

$$\mathbb{E}\,\|Z\|^p = \int_{\|x\|\ge l} \|x\|^4\, Q(dx) + 3\left(\int_{\|x\|\ge l} \|x\|^2\, Q(dx)\right)^2. \tag{1.14}$$

Proof. First, (1.14) follows exactly as in (1.6) by standard computation from the characteristic function. Next let $\lambda \overset{def}{=} Q(\|x\| \ge l)$ and $\{W_i\}_{i\in\mathbb{N}}$ a collection of i.i.d. random vectors in H such that $\mathbb{P}(W_i \in A) = \lambda^{-1} Q(A \cap \{\|x\| \ge l\})$. Let N be a Poisson random variable with mean λ independent of $\{W_i\}_{i\in\mathbb{N}}$. Now Z is a compound Poisson random vector and we have

$$Z \overset{d}{=} \sum_{i=1}^{N} W_i. \tag{1.15}$$

Then

$$\mathbb{E}\,\|Z\|^p = \mathbb{E}\left\| \sum_{i=1}^{N} W_i \right\|^p = \sum_{k=1}^{\infty} \mathbb{E}\left\| \sum_{i=1}^{k} W_i \right\|^p \mathbb{P}(N = k). \tag{1.16}$$

First let $1 \le p \le 2$. By Corollary A.6 if $H = \mathbb{R}$ or Theorem A.2 in general, for each $k \in \mathbb{N}$, $\mathbb{E}\left\|\sum_{i=1}^{k} W_i\right\|^p$ is bounded above by $c_p \sum_{i=1}^{k} \mathbb{E}\|W_i\|^p = c_p k \mathbb{E}\|W_1\|^p$. Utilizing this in (1.16) gives

$$\mathbb{E}\|Z\|^p \le c_p \mathbb{E}\|W_1\|^p \sum_{k=1}^{\infty} k\mathbb{P}(N = k) = c_p \mathbb{E}\|W_1\|^p \mathbb{E}N = c_p \mathbb{E}\|W_1\|^p \lambda,$$

since N is a Poisson random variable with mean λ. But

$$\mathbb{E}\|W_1\|^p = \int_{\|x\|\ge l} \|x\|^p \lambda^{-1} Q(dx)$$

and hence,

$$\mathbb{E}\|Z\|^p \le c_p \int_{\|x\|\ge l} \|x\|^p Q(dx),$$

proving (1.10).

Next, let $2 < p \le 3$. By Theorem A.1,

$$\begin{aligned}
\mathbb{E}\left\|\sum_{i=1}^{k} W_i\right\|^p &\le k\mathbb{E}\|X_1\|^p + \frac{p(p-1)}{2}\mathbb{E}\|X_1\|^{p-2} \sum_{i=1}^{k} \mathbb{E}\|S_{i-1}\|^2 \\
&= k\mathbb{E}\|X_1\|^p + \frac{p(p-1)}{2}\mathbb{E}\|X_1\|^{p-2} \sum_{i=1}^{k}\sum_{j=1}^{i-1} \mathbb{E}\|X_j\|^2 \\
&= k\mathbb{E}\|X_1\|^p + \frac{p(p-1)}{2}\frac{k^2-k}{2}\mathbb{E}\|X_1\|^{p-2}\mathbb{E}\|X_1\|^2.
\end{aligned}$$

Again recalling that N is Poisson, substituting into (1.16) gives

$$\begin{aligned}
\mathbb{E}\|Z\|^p &\le \sum_{k=1}^{\infty}\left(k\mathbb{E}\|X_1\|^p + \frac{p(p-1)}{2}\frac{k^2-k}{2}\mathbb{E}\|X_1\|^{p-2}\mathbb{E}\|X_1\|^2\right)\mathbb{P}(N = k) \\
&= \mathbb{E}(N)\,\mathbb{E}\|X_1\|^p + \frac{p(p-1)}{4}\mathbb{E}\left(N^2 - N\right)\mathbb{E}\|X_1\|^{p-2}\mathbb{E}\|X_1\|^2 \\
&= \lambda \int_{\|x\|\ge l} \|x\|^p \lambda^{-1} Q(dx) \\
&\quad + \frac{p(p-1)}{4}\lambda^2 \int_{\|x\|\ge l} \|x\|^{p-2} \lambda^{-1} Q(dx) \int_{\|x\|\ge l} \|x\|^2 \lambda^{-1} Q(dx) \\
&= \int_{\|x\|\ge l} \|x\|^p Q(dx) + \frac{p(p-1)}{4}\int_{\|x\|\ge l} \|x\|^{p-2} Q(dx) \int_{\|x\|\ge l} \|x\|^2 Q(dx).
\end{aligned}$$

Finally, let $p > 3$. If $\lambda > p/x_0$, we have by de la Peña and Giné (1999, Theorem 1.2.5, a Hoffman-Jorgensen type inequality)

$$\mathbb{E}\left\|\sum_{i=1}^{k} W_i\right\|^p \leq \left(\frac{4^{1/(p+1)}2^{p/(p+1)}4^{p/(2(p+1))}2^{2/(p+1)}}{2^{2/(p+1)}-1}\left\|\sum_{i=1}^{k} W_i\right\|_2^{p/(p+1)} + \frac{2^{2/(p+1)}}{2^{2/(p+1)}-1}\left(\sum_{i=1}^{k}\mathbb{E}\left\|W_i\right\|^p\right)^{1/(p+1)}\right)^{p+1}.$$

By convexity, $(a+b)^{p+1} \leq 2^p\left(a^{p+1}+b^{p+1}\right)$ for every $a, b \geq 0$. Therefore

$$\begin{aligned}\mathbb{E}\left\|\sum_{i=1}^{k} W_i\right\|^p &\leq 2^p \frac{4}{\left(2^{2/(p+1)}-1\right)^{p+1}}\left(2^{2(p+1)}\left\|\sum_{i=1}^{k} W_i\right\|_2^p + \sum_{i=1}^{k}\mathbb{E}\left\|W_i\right\|^p\right)\\ &= K_{3,p}\left(2^{2(p+1)}k^{p/2}\left(\mathbb{E}\left\|W_1\right\|^2\right)^{p/2} + k\mathbb{E}\left\|W_1\right\|^p\right).\end{aligned}$$

Substituting into (1.16) gives

$$\begin{aligned}\mathbb{E}\left\|Z\right\|^p &\leq \sum_{k=1}^{\infty} K_{3,p}\left(2^{2(p+1)}k^{p/2}\left(\mathbb{E}\left\|W_1\right\|^2\right)^{p/2} + k\mathbb{E}\left\|W_1\right\|^p\right)\mathbb{P}(N=k)\\ &= K_{3,p}\left(2^{2(p+1)}\mathbb{E}N^{p/2}\left(\int_{\|x\|\geq l}\|x\|^2\lambda^{-1}Q(dx)\right)^{p/2} + \mathbb{E}N\int_{\|x\|\geq l}\|x\|^p\lambda^{-1}Q(dx)\right)\\ &= K_{3,p}\left(2^{2(p+1)}\left(\int_{\|x\|\geq l}\|x\|^2 Q(dx)\right)^{p/2}\lambda^{-p/2}\mathbb{E}N^{p/2} + \int_{\|x\|\geq l}\|x\|^p Q(dx)\right).\end{aligned}$$

To bound $\lambda^{-p/2}\mathbb{E}N^{p/2}$, Kwapień and Woyczyński (2009, Proposition 1.7.2) showed that in the case $\lambda > p/x_0$,

$$\|N\|_{p/2} \leq \|N\|_p \leq 4(p+5\lambda).$$

Hence,

$$\lambda^{-p/2}\mathbb{E}N^{p/2} \leq 4^{p/2}\left(\frac{p}{\lambda}+5\right)^{p/2} \leq 4^{p/2}(x_0+5)^{p/2}$$

and we have

$$\begin{aligned}\mathbb{E}\left\|Z\right\|^p &\leq K_{3,p}\left(2^{2(p+1)}\left(\int_{\|x\|\geq l}\|x\|^2 Q(dx)\right)^{p/2} 4^{p/2}(x_0+5)^{p/2} + \int_{\|x\|\geq l}\|x\|^p Q(dx)\right)\\ &= K_{3,p}\left(K_{4,p}\left(\int_{\|x\|\geq l}\|x\|^2 Q(dx)\right)^{p/2} + \int_{\|x\|\geq l}\|x\|^p Q(dx)\right).\end{aligned}$$

Now suppose that $\lambda \leq p/x_0$. For each $\omega \in \Omega$, Holder's inequality gives

$$\left\|\sum_{i=1}^k W_i(\omega)\right\| \leq \sum_{i=1}^k \|W_i(\omega)\| \leq k^{1-1/p}\left(\sum_{i=1}^k \|W_i(\omega)\|^p\right)^{1/p}$$

and hence,

$$\mathbb{E}\left\|\sum_{i=1}^k W_i\right\|^p \leq k^{p-1}\sum_{i=1}^k \mathbb{E}\left\|W_i\right\|^p = k^p\mathbb{E}\left\|W_1\right\|^p.$$

Substituting into (1.16) gives

$$\mathbb{E}\left\|Z\right\|^p \leq \mathbb{E}\left\|W_1\right\|^p \sum_{k=1}^{\infty} k^p\mathbb{P}(N=k) = \int_{\|x\|\geq l}\|x\|^p \lambda^{-1}Q(dx)\mathbb{E}N^p. \tag{1.17}$$

To bound $\lambda^{-1}\mathbb{E}N^p$, Kwapień and Woyczyński (2009, Proposition 1.7.2) also showed that in the case $\lambda \leq p/x_0$,

$$\lambda^{-1}\mathbb{E}N^p \leq \max\left\{1 + \frac{8p}{\log\left(\frac{p}{\frac{p}{x_0}\wedge 1}\right)}, \frac{6p}{\log x_0}\right\}^p$$

Combining with (1.17) gives

$$\mathbb{E}\left\|Z\right\|^p \leq \max\left\{1 + \frac{8p}{\log\left(\frac{p}{\frac{p}{x_0}\wedge 1}\right)}, \frac{6p}{\log x_0}\right\}^p \int_{\|x\|\geq l}\|x\|^p Q(dx).$$

□

Lemma 1.2.5. *If $p \geq 1$, we have the following lower bound on norms of Z:*

$$\mathbb{E}\left\|Z\right\|^p \geq \frac{1-e^{-\lambda}}{\lambda} \int_{\|x\| \geq l} \|x\|^p \, Q(dx). \tag{1.18}$$

Proof. Let $p \geq 1$. Since we have assumed that Q is symmetric, Lemma A.7 gives

$$\mathbb{E}\left\|\sum_{i=1}^{k} W_i\right\|^p \geq \mathbb{E}\left\|W_1\right\|^p .$$

Substituting into (1.16) gives

$$\mathbb{E}\left\|Z\right\|^p \geq \sum_{k=1}^{\infty} \mathbb{E}\left\|W_1\right\|^p \mathbb{P}(N=k) = \frac{1-e^{-\lambda}}{\lambda} \int_{\|x\| \geq l} \|x\|^p \, Q(dx).$$

□

We are now ready to prove the upper bound of Theorem 1.2.1 using Lemma 1.2.2 and Lemma 1.2.4.

Proof of upper bound of Theorem 1.2.1. First assume that $1 \leq p \leq 2$. From (1.5) and (1.10), we have

$$\begin{aligned} \left\|X\right\|_p &\leq \left\|Y\right\|_2 + \left\|Z\right\|_p \\ &\leq \left(\int_{\|x\|<l} \|x\|^2 \, Q(dx)\right)^{1/2} + \sqrt[p]{c_p} \left(\int_{\|x\| \geq l} \|x\|^p \, Q(dx)\right)^{1/p} \\ &= l \left(\int_{\|x\|<l} \left\|\frac{x}{l}\right\|^2 Q(dx)\right)^{1/2} + l \sqrt[p]{c_p} \left(\int_{\|x\| \geq l} \left\|\frac{x}{l}\right\|^p Q(dx)\right)^{1/p} . \end{aligned} \tag{1.19}$$

By definition (1.2) of l,

$$\int_{\|x\| \geq l} \left\|\frac{x}{l}\right\|^p Q(dx) = 1 - \int_{\|x\|<l} \left\|\frac{x}{l}\right\|^2 Q(dx).$$

Substituting into (1.19) gives

$$\left\|X\right\|_p \leq \left\{ \left(\int_{\|x\|<l} \left\|\frac{x}{l}\right\|^2 Q(dx)\right)^{1/2} + \sqrt[p]{c_p} \left(1 - \int_{\|x\|<l} \left\|\frac{x}{l}\right\|^2 Q(dx)\right)^{1/p} \right\} l.$$

Clearly, by definition (1.2) of l we have

$$0 \leq \int_{\|x\|<l} \left\| \frac{x}{l} \right\|^2 Q(dx) \leq 1$$

and hence,

$$\|X\|_p \leq \max_{0\leq a\leq 1} \left(\sqrt{a} + \sqrt[p]{c_p}\sqrt[p]{1-a}\right) l \leq \left(1 + \sqrt[p]{c_p}\right) l.$$

Next, let $2 < p \leq 3$. Combining (1.6) and (1.11) gives

$$\begin{aligned}
\|X\|_p &\leq \|Y\|_p + \|Z\|_p \\
&\leq \left(\int_{\|x\|<l} l^2 \|x\|^2 Q(dx) + 3\left(\int_{\|x\|<l} \|x\|^2 Q(dx)\right)^2\right)^{1/4} \\
&\quad + \left(\int_{\|x\|\geq l} \|x\|^p Q(dx) + \frac{p(p-1)}{4} \int_{\|x\|\geq l} \|x\|^{p-2} Q(dx) \int_{\|x\|\geq l} \|x\|^2 Q(dx)\right)^{1/p} \\
&\leq \left(\int_{\|x\|<l} \frac{\|x\|^2}{l^2} Q(dx) + 3\left(\int_{\|x\|<l} \frac{\|x\|^2}{l^2} Q(dx)\right)^2\right)^{1/4} l \\
&\quad + \left(\int_{\|x\|\geq l} \frac{\|x\|^p}{l^p} Q(dx) + \frac{p(p-1)}{4} \int_{\|x\|\geq l} \frac{\|x\|^{p-2}}{l^{p-2}} Q(dx) \int_{\|x\|\geq l} \frac{\|x\|^2}{l^2} Q(dx)\right)^{1/p} l \\
&\leq \left(\sqrt[4]{4} + \sqrt[p]{1 + \frac{p(p-1)}{4}}\right) l.
\end{aligned}$$

Now let $3 < p < 4$. If $\lambda > p/x_0$, (1.12) gives

$$\begin{aligned}
\mathbb{E}\|Z\|^p &\leq K_{3,p}\left(K_{4,p}\left(\int_{\|x\|\geq l} \|x\|^2 Q(dx)\right)^{p/2} + \int_{\|x\|\geq l} \|x\|^p Q(dx)\right) \\
&= K_{3,p}\left(K_{4,p}\left(\int_{\|x\|\geq l} \frac{\|x\|^2}{l^2} Q(dx)\right)^{p/2} l^p + \int_{\|x\|\geq l} \frac{\|x\|^p}{l^p} Q(dx) l^p\right) \\
&\leq K_{3,p}\left(K_{4,p} + 1\right) l^p
\end{aligned}$$

and if $\lambda \leq p/x_0$, (1.13) gives

$$\mathbb{E}\left\|Z\right\|^p \leq \max\left\{1+\frac{8p}{\log\left(\frac{p}{\frac{p}{x_0}\wedge 1}\right)}, \frac{6p}{\log x_0}\right\}^p \int_{\|x\|\geq l} \|x\|^p \, Q(dx)$$

$$= \max\left\{1+\frac{8p}{\log\left(\frac{p}{\frac{p}{x_0}\wedge 1}\right)}, \frac{6p}{\log x_0}\right\}^p \int_{\|x\|\geq l} \frac{\|x\|^p}{l^p} Q(dx) l^p$$

$$\leq \max\left\{1+\frac{8p}{\log\left(\frac{p}{\frac{p}{x_0}\wedge 1}\right)}, \frac{6p}{\log x_0}\right\}^p l^p.$$

In either case, we have

$$\mathbb{E}\left\|Z\right\|^p \leq K_{3,p}\left(K_{4,p}+1\right) l^p.$$

This, along with (1.6) gives

$$\begin{aligned}\|X\|_p &\leq \|Y\|_4 + \|Z\|_p \\ &\leq \left(\sqrt[4]{4} + K_{3,p}^{1/p}\left(K_{4,p}+1\right)^{1/p}\right) l.\end{aligned}$$

Now let $p = 4$. Combining (1.6) and (1.14) gives

$$\begin{aligned}\|X\|_p &\leq \|Y\|_p + \|Z\|_p \\ &\leq \left(\int_{\|x\|<l} l^2 \|x\|^2 \, Q(dx) + 3\left(\int_{\|x\|<l} \|x\|^2 \, Q(dx)\right)^2\right)^{1/4} \\ &\quad + \left(\int_{\|x\|\geq l} \|x\|^4 \, Q(dx) + 3\left(\int_{\|x\|\geq l} \|x\|^2 \, Q(dx)\right)^2\right)^{1/4} \\ &= \left(\int_{\|x\|<l} \frac{\|x\|^2}{l^2} Q(dx) + 3\left(\int_{\|x\|<l} \frac{\|x\|^2}{l^2} Q(dx)\right)^2\right)^{1/4} l \\ &\quad + \left(\int_{\|x\|\geq l} \frac{\|x\|^4}{l^4} Q(dx) + 3\left(\int_{\|x\|\geq l} \frac{\|x\|^2}{l^2} Q(dx)\right)^2\right)^{1/4} l \\ &\leq 2\sqrt[4]{4} l.\end{aligned}$$

Finally, let $p > 4$. Combining (1.7) and the bounds on Z from the $3 < p < 4$ case, we have

$$\begin{aligned}
\|X\|_p &\le \|Y\|_p + \|Z\|_p \\
&\le K_{1,p}\left(K_{2,p}\|Y\|_2 + l\right) + K_{3,p}^{1/p}\left(K_{4,p}+1\right)^{1/p} l \\
&= K_{1,p}\left(K_{2,p}\left(\int_{\|x\|<l} \frac{\|x\|^2}{l^2} Q(dx)\right)^{1/2} l + l\right) + K_{3,p}^{1/p}\left(K_{4,p}+1\right)^{1/p} l \\
&\le \left(K_{1,p}\left(K_{2,p}+1\right) + K_{3,p}^{1/p}\left(K_{4,p}+1\right)^{1/p}\right) l.
\end{aligned}$$

□

We are now ready to prove the lower bound of Theorem 1.2.1 using Lemma 1.2.3 and Lemma 1.2.5.

Proof of lower bound of Theorem 1.2.1. By (1.2), either

$$\int_{\|x\|<l} \frac{\|x\|^2}{l^2} Q(dx) \ge 0.5 \tag{1.20}$$

or

$$\int_{\|x\|\ge l} \frac{\|x\|^p}{l^p} Q(dx) \ge 0.5 \tag{1.21}$$

must be true. Assume (1.20) holds. If $1 \le p \le 2$, Lemma A.7 and (1.7) combine to give

$$\mathbb{E}\|X\|^p = \mathbb{E}\|Y+Z\|^p \ge \mathbb{E}\|Y\|^p \ge \frac{\mathbb{E}\|Y\|^2}{\left(l^2 + 3\mathbb{E}\|Y\|^2\right)^{\frac{2-p}{2}}}.$$

Since the function $t \mapsto t\left(l^2+3t\right)^{\frac{p-2}{2}}$ is increasing in t,

$$\mathbb{E}\|X\|^p \ge \frac{0.5l^2}{\left(l^2+3(0.5)l^2\right)^{\frac{2-p}{2}}} = \frac{2.5^{p/2}}{5} l^p \ge \frac{l^p}{4}$$

and hence,

$$\|X\|_p \ge \sqrt[p]{0.25}l \ge 0.25l.$$

If $p > 2$, then by Lemma A.7 and (1.9),

$$\|X\|_p \geq \|Y\|_p \geq \|Y\|_2 = \left(\int_{\|x\|<l} x^2 Q(dx) \right)^{\frac{1}{2}} \geq \sqrt{0.5}l > 0.25l.$$

Now assume (1.21) holds. Then

$$\int_{\|x\|\geq l} \left(\|x\|^p + (\|x\|^p - l^p)\right) Q(dx) \geq \int_{\|x\|\geq l} \|x\|^p \, Q(dx) \geq 0.5l^p.$$

Now the left hand side simplifies to

$$2 \int_{\|x\|\geq l} \|x\|^p \, Q(dx) - l^p \lambda,$$

where $\lambda = Q(\|x\| \geq l)$, and hence,

$$\int_{\|x\|\geq l} \|x\|^p \, Q(dx) \geq \frac{l^p}{2} (0.5 + \lambda) = \frac{l^p}{4} (1 + 2\lambda).$$

We may combine this with the lower bound inequality in (1.10) and utilize Lemma A.7 as in the above case to get

$$\mathbb{E} \|X\|^p \geq \mathbb{E} \|Z\|^p \geq \frac{1 - e^{-\lambda}}{\lambda} \frac{l^p}{4} (1 + 2\lambda) \geq \frac{l^p}{4}$$

and hence,

$$\|X\|_p \geq \sqrt[p]{0.25}l \geq 0.25l.$$

In either case, the left hand inequality in (1.3) holds. □

Recall that we have been working under the assumption that Q is symmetric. To remove this restriction, assume that X is a mean 0 infinitely divisible random vector in L^p with Lévy measure Q and let X^s be the standard symmetrization of X. The Lévy measure of X^s is given by $Q^s(A) = Q(A) + Q(-A)$ and if c solves (1.2) for Q^s, we have that c also solves

$$\int_H \left(\frac{\|x\|^2}{c^2} \mathbb{1}_{\{\|x\|<c\}} + \frac{\|x\|^p}{c^p} \mathbb{1}_{\{\|x\|\geq c\}} \right) Q(dx) = \frac{1}{2}. \tag{1.22}$$

By Corollary A.8 and Theorem 1.2.1,

$$\frac{1}{8}c \leq \frac{1}{2}\|X^s\|_p \leq \|X\|_p \leq \|X^s\|_p \leq K(p)c.$$

Now let l solve

$$\int_H \left(\frac{\|x\|^2}{l^2} \mathbb{1}_{\{\|x\|<l\}} + \frac{\|x\|^p}{l^p} \mathbb{1}_{\{\|x\| \geq l\}} \right) Q(dx) = 1 \tag{1.23}$$

and

$$k \stackrel{def}{=} \begin{cases} \sqrt[p]{2}, & \text{if } 1 \leq p \leq 2 \\ \sqrt{2}, & \text{if } p > 2. \end{cases}$$

Then $k > 1$ and if $1 \leq p \leq 2$, we have

$$\int_H \min\left\{ \frac{\|x\|^2}{(kl)^2}, \frac{\|x\|^p}{(kl)^p} \right\} Q(dx) \leq \max\left\{ \frac{1}{k^2}, \frac{1}{k^p} \right\} \int_H \min\left\{ \frac{\|x\|^2}{l^2}, \frac{\|x\|^p}{l^p} \right\} Q(dx) = \frac{1}{k^p} = \frac{1}{2}$$

or if $p > 2$, we have

$$\int_H \max\left\{ \frac{\|x\|^2}{(kl)^2}, \frac{\|x\|^p}{(kl)^p} \right\} Q(dx) \leq \max\left\{ \frac{1}{k^2}, \frac{1}{k^p} \right\} \int_H \max\left\{ \frac{\|x\|^2}{l^2}, \frac{\|x\|^p}{l^p} \right\} Q(dx) = \frac{1}{k^2} = \frac{1}{2}.$$

In either case, $c \leq kl$ since c solves (1.22). Clearly, $l \leq c$ since l solves (1.23). We have proven the following corollary to Theorem 1.2.1:

Corollary 1.2.6. *Let $p \geq 1$. Assume that $X \in L^p$ is a mean 0 infinitely divisible random vector without Gaussian component, taking values in the Hilbert space H, and that X has Lévy measure Q. Let l be the solution of*

$$\xi(l) \stackrel{def}{=} \int_H \left(\frac{\|x\|^2}{l^2} \mathbb{1}_{\{\|x\|<l\}} + \frac{\|x\|^p}{l^p} \mathbb{1}_{\{\|x\| \geq l\}} \right) Q(dx) = 1. \tag{1.24}$$

Then

$$0.125l \leq \|X\|_p \leq \max\{\sqrt[p]{2}, \sqrt{2}\} K(p) l \tag{1.25}$$

where $K(p)$ is given by (1.4).

The last corollary to Theorem 1.2.1 that we present gives quick estimation of the L^p-norm of X in terms of the functional $\xi(l)$.

Corollary 1.2.7. *Under the assumptions of Theorem 1.2.1, if $\xi(l)$ is given by* (1.2), *then*

$$(0.25)\min\left\{\sqrt{\xi(1)},\, \sqrt[p]{\xi(1)}\right\} \le \|X\|_p \le K(p)\max\left\{\sqrt{\xi(1)},\, \sqrt[p]{\xi(1)}\right\}.$$

Similarly, under the assumptions of Corollary 1.2.6,

$$(0.125)\min\left\{\sqrt{\xi(1)},\, \sqrt[p]{\xi(1)}\right\} \le \|X\|_p \le \max\{\sqrt[p]{2}, \sqrt{2}\}K(p)\max\left\{\sqrt{\xi(1)},\, \sqrt[p]{\xi(1)}\right\}.$$

Proof. First, suppose $1 \le p \le 2$. If $l < 1$,

$$\begin{aligned}
l^2 &= \int_{\|x\|<l} \|x\|^2 Q(dx) + \int_{\|x\|\ge l} l^{2-p} \|x\|^p Q(dx) \\
&\le \int_{\|x\|<l} \|x\|^2 Q(dx) + \int_{l\le\|x\|<1} \|x\|^{2-p} \|x\|^p Q(dx) + \int_{\|x\|\ge 1} \|x\|^p Q(dx) \\
&= \xi(1)
\end{aligned}$$

and

$$\begin{aligned}
l^p &= \int_{\|x\|<l} l^{p-2} \|x\|^2 Q(dx) + \int_{\|x\|\ge l} \|x\|^p Q(dx) \\
&\ge \int_{\|x\|<l} \|x\|^2 Q(dx) + \int_{l\le\|x\|<1} \|x\|^2 Q(dx) + \int_{\|x\|\ge 1} \|x\|^p Q(dx) \\
&= \xi(1).
\end{aligned}$$

If $l \ge 1$, similar arguments give $l^2 \ge \xi(1)$ and $l^p \le \xi(1)$. In either case we have

$$\min\left\{\sqrt{\xi(1)},\, \sqrt[p]{\xi(1)}\right\} \le l \le \max\left\{\sqrt{\xi(1)},\, \sqrt[p]{\xi(1)}\right\}.$$

Similar arguments give the $p > 2$ case. □

Chapter 2

Kalman Filter

2.1 Kalman filter theory

In his landmark paper, Kalman (1960) considered the discrete time signal-observation model

$$x_k = F_k x_{k-1} + B_k u_k + w_k$$
$$y_k = H_k x_k + v_k,$$

where x_k is the state of an evolving dynamical system at time k, u_k is a (deterministic) control input to the system, and y_k is a noisy linear observation of x_k. The "noise" terms $\{w_k\}$ and $\{v_k\}$ are assumed to be mean 0 Gaussian random vectors with covariance matrices W_k and V_k, respectively. In "filter theory", the objective is to produce an efficient estimate x_k^* of the (unobservable) process x_k using the observed values $y_1, y_2, \ldots, y_k$, which are known at time k. An efficient estimate is one that minimizes some expected "loss" of the error $x_k - x_k^*$. In his paper, Kalman (1960) showed that $x_k^* \stackrel{def}{=} \mathbb{E}\left(x_k | y_1, y_2, \ldots, y_k\right)$ minimizes the L^2-norm of the error and gave a recursive formulation for computing the estimate x_k^*. Under the assumption of normally distributed noise terms, the orthogonal projection x_k^* is an affine transformation of the observations $y_1, y_2, \ldots, y_k$.

Let $\hat{x}_{k|k-1}$ be the *predicted* state of the system at time k, given that the observations $y_1, y_2, \ldots, y_{k-1}$ are known at time $k-1$. Then, at time k, the observation y_k becomes available and we may update our state estimate. Let $\hat{x}_{k|k}$ be the *updated* estimate of the system state at time k once the observation y_k has become available.

We denote by $P_{k|k}$ the covariance matrix of the error $x_k - \hat{x}_{k|k}$ and by $P_{k|k-1}$ the covariance matrix of the error $x_k - \hat{x}_{k|k-1}$. The recursively formulated solution given by Kalman (1960) to compute $x_k^* = \hat{x}_{k|k}$ is given in Algorithm 1. The filter $\hat{x}_{k|k}$ is a linear combination of the predicted state $\hat{x}_{k|k-1}$ and the observation y_k. The *optimal Kalman gain* K_k in Algorithm 1 is chosen to minimize the L^2-norm of the error $x_k - \hat{x}_{k|k}$ and is given by

$$K_k = P_{k|k-1} H_k^T \left(H_k P_{k|k-1} H_k^T + V_k \right)^{-1}. \tag{2.1}$$

Over the years since this publication, some research has focused on replacing the noise terms by random vectors with heavy-tailed distributions. Gordon et al. (2003, Introduction) argued for the need of models allowing heavy tailed error estimates as outlying system state realizations and/or observation measurements "have long been known to adversely affect the estimation procedure". In Gordon et al. (2003), the authors assume that the noise terms are "power law" distributed and give the Kalman filter in terms of the "tail covariance matrices" of the noise terms. Stuck (1978) first addressed this model under the assumption that both x_k and y_k are $\mathbb{R}$-valued and each noise sequence $\{w_k\}$ and $\{v_k\}$ are α-stable random variables for fixed α. These examples fall under a more general framework for which the noise sequences are assumed to be symmetric infinitely divisible random vectors. In what follows, we establish a general framework to explore the Kalman filter under this assumption on the distributions of the noise sequences and demonstrate in two different examples that a solution can often be obtained (or approximated). The first example assumes that each noise term has finite L^2-norm, but makes no other assumptions on the distributions. The second example considers the problem for α-stable distributed noise sequences, which was first addressed in dimension 1 by Stuck (1978) and then in Gordon et al. (2003). In each example, a tractable (approximate) solution is given. Each solution is exact in dimension 1 and agrees with the classic Kalman gain (2.1) (when $\alpha = 2$ in the second example).

Before we begin, we should point out that these solutions are only optimal in the linear sense. Kalman (1960) noted that, under the assumption that the noise terms are normally distributed, the orthogonal projection $\mathbb{E}\left(x_k | y_1, y_2, \ldots, y_k\right)$ is a linear function of the observations $y_1, y_2, \ldots, y_k$. However, by removing this assumption, this is no longer the case. In general, the L^2-orthogonal projection $\mathbb{E}\left(x_k | y_1, y_2, \ldots, y_k\right)$ is non-linear and non-linear filtering theory may give better results. If we are seeking the

Algorithm 1 Kalman filter for Gaussian noise.

1: Initialize:
$\hat{x}_{0|0} \overset{def}{=} \mathbb{E}x_0 = 0$
$P_{0|0} = W_0$
2: Predict:
$\hat{x}_{k|k-1} \overset{def}{=} F_k \hat{x}_{k-1|k-1} + B_k u_k$ (unbiased estimate)
$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + W_k$
3: Update:
$K_k = P_{k|k-1} H_k^T \left(H_k P_{k|k-1} H_k^T + V_k \right)^{-1}$
$\hat{x}_{k|k} \overset{def}{=} \hat{x}_{k|k-1} + K_k \left(y_k - H_k \hat{x}_{k|k-1} \right)$
$P_{k|k} = \left(I - K_k H_k \right) P_{k|k-1}$

optimal solution x_k^* minimizing, say, the L^p-norm of the error $x_k - x_k^*$, the conventional conditional expected value is no longer even the optimal solution. Instead, it will be the conditional L^p-expected value $E_p\left(x_k | y_1, y_2, \ldots, y_k\right)$ that minimizes the L^p-norm of the error. However, the linear formulation has the desirable property of being easily implemented and are the only estimates we consider. To this end, consider the discrete time signal-observation model

$$\begin{aligned} x_k &= F_k x_{k-1} + B_k u_k + w_k \\ y_k &= H_k x_k + v_k, \end{aligned} \tag{2.2}$$

where $x_k \in \mathbb{R}^d$, $F_k \in \mathbb{R}^{d \times d}$, $u_k \in \mathbb{R}^n$, $B_k \in \mathbb{R}^{d \times n}$, $y_k \in \mathbb{R}^m$, and $H_k \in \mathbb{R}^{m \times d}$. Assume that the system noise $\{w_k\}_{k \in \mathbb{N}}$ are independent symmetric $\mathbb{R}^d$-valued random vectors with the Lévy-Khintchine triplets

$$w_k \sim \left(0, 0, Q^{w,k}\right), k = 1, 2, \ldots,$$

where, for each k, $Q^{w,k}$ is a symmetric Lévy measure on $\mathbb{R}^d$, that the observation noise $\{v_k\}_{k \in \mathbb{N}}$ are independent symmetric $\mathbb{R}^m$-valued random vectors with the Lévy-Khintchine triplets

$$v_k \sim \left(0, 0, Q^{v,k}\right), k = 1, 2, \ldots,$$

where, for each k, $Q^{v,k}$ is a symmetric Lévy measure on $\mathbb{R}^m$, and that $x_0 \in \mathbb{R}^d$ is a symmetric infinitely divisible random vector with Lévy-Khintchine triplet

$$x_0 \sim \left(0, 0, Q^{w,0}\right),$$

where $Q^{w,0}$ is a symmetric Lévy measure on $\mathbb{R}^d$. Moreover, assume that the sequence of random vectors $\{x_0, w_1, v_1, w_2, v_2, \dots\}$ are mutually independent. Finally, assume that for some fixed $p \geq 1$, we have that both

$$\int_{\mathbb{R}^d} \|x\|^p \mathbb{1}_{\{\|x\| \geq 1\}} Q^{w,k}(dx) < \infty$$

for each $k = 0, 1, 2, \dots$, and that

$$\int_{\mathbb{R}^m} \|x\|^p \mathbb{1}_{\{\|x\| \geq 1\}} Q^{v,k}(dx) < \infty$$

for each $k = 1, 2, 3, \dots$. Restricting ourselves to linear estimates, the Kalman filter algorithm is given by Algorithm 2.

Let $e_{k|k}$ be the updated estimate error, $e_{k|k-1}$ the predicted estimate error, and observe that

$$\begin{aligned}
e_{0|0} &= x_0, \\
e_{k|k-1} &\stackrel{def}{=} x_k - \hat{x}_{k|k-1} \\
&= (F_k x_{k-1} + B_k u_k + w_k) - \left(F_k \hat{x}_{k-1|k-1} + B_k u_k\right) \\
&= F_k e_{k-1|k-1} + w_k,
\end{aligned}$$

and

$$\begin{aligned}
e_{k|k} &\stackrel{def}{=} x_k - \hat{x}_{k|k} \\
&= x_k - \left(\hat{x}_{k|k-1} + K_k \left(y_k - H_k \hat{x}_{k|k-1}\right)\right) \\
&= x_k - \hat{x}_{k|k-1} - K_k \left(H_k x_k + v_k\right) + K_k H_k \hat{x}_{k|k-1} \\
&= e_{k|k-1} - K_k H_k \left(x_k - \hat{x}_{k|k-1}\right) - K_k v_k \\
&= \left(I_d - K_k H_k\right) e_{k|k-1} - K_k v_k \\
&= \left(I_d - K_k H_k\right) \left(F_k e_{k-1|k-1} + w_k\right) - K_k v_k.
\end{aligned}$$

First, we remark that $e_{k-1|k-1}$, w_k, and v_k are independent. Second, $K_k v_k$ is a symmetric random vector. These two facts, along with Corollary 1.1.2, imply that the updated error $e_{k|k}$ is an infinitely divisible random vector on $\mathbb{R}^d$ and, since each

Algorithm 2 Kalman filter

1: Initialize:
$\hat{x}_{0|0} \overset{def}{=} \mathbb{E}x_0 = 0$
2: Predict:
$\hat{x}_{k|k-1} \overset{def}{=} F_k \hat{x}_{k-1|k-1} + B_k u_k$ (unbiased estimate)
3: Update:
$\hat{x}_{k|k} \overset{def}{=} \hat{x}_{k|k-1} + K_k \left(y_k - H_k \hat{x}_{k|k-1} \right)$

Lévy measure $Q^{\cdot,\cdot}$ is symmetric, the Lévy-Khintchine triplet is given by

$$e_{0|0} \sim \left(0, 0, Q^{w,0}\right) \overset{def}{=} \left(0, 0, Q^{0}\right),$$
$$e_{k|k} \sim \left(0, 0, Q^{k-1}_{(I_d - K_k H_k)F_k} + Q^{w,k}_{I_d - K_k H_k} + Q^{v,k}_{K_k}\right) \overset{def}{=} \left(0, 0, Q^{k}\right), k = 1, 2, \ldots. \tag{2.3}$$

We recall from Corollary 1.1.2 that the subscript notation $Q^{v,k}_{K_k}$ represents a new Lévy measure on $\mathbb{R}^d$ given by

$$Q^{v,k}_{K_k}(B) \overset{def}{=} Q^{v,k}\left\{x \in \mathbb{R}^m : K_k x \in B \setminus \{0\}\right\},$$

for every $B \in \mathcal{B}(\mathbb{R}^d)$. In light of Section 1.2, for every k, we may measure the magnitude of the error by l_k, where l_k solves

$$\int_{\mathbb{R}^d} \left(\frac{\|x\|^2}{l_k^2} \mathbb{1}_{\left\{\frac{\|x\|}{l_k} < 1\right\}} + \frac{\|x\|^p}{l_k^p} \mathbb{1}_{\left\{\frac{\|x\|}{l_k} > 1\right\}} \right) Q^k(dx) = 1. \tag{2.4}$$

The optimal Kalman gain $K_k \in \mathbb{R}^{d \times m}$ is chosen to minimize l_k. While no closed form solution exists for such arbitrary Lévy measures, we demonstrate (approximate) solutions in the following two examples. The first will deal with the case that $p = 2$ and the Lévy measures are arbitrary. The second example will deal with the symmetric α-stable case. Often, we will need to compute Q^k iteratively, as opposed to recursively as in (2.3). To do so, observe that if Q is a measure on $\mathbb{R}^n$, $G \in \mathbb{R}^{q \times n}$, and $H \in \mathbb{R}^{r \times q}$, then $\left(Q_G\right)_H$ is a measure on $\mathbb{R}^r$ and we have, for $B \in \mathcal{B}\left(\mathbb{R}^r\right)$,

$$\begin{aligned}
\left(Q_G\right)_H (B) &= Q_G \left(\{x \in \mathbb{R}^q : Hx \in B \setminus \{0\}\}\right) \\
&= Q \left(\{x \in \mathbb{R}^n : Gx \in \{x \in \mathbb{R}^q : Hx \in B \setminus \{0\}\} \setminus \{0\}\}\right) \\
&= Q \left(\{x \in \mathbb{R}^n : HGx \in B \setminus \{0\}\}\right) \\
&= Q_{HG}(B).
\end{aligned}$$

Using this rule that $(Q_G)_H = Q_{HG}$, we may derive the following formulation of (2.3):

Theorem 2.1.1. *The recursively defined Lévy measure Q^k in (2.3) is*

$$\begin{aligned} Q^k &= Q^{w,0}_{\prod_{i=0}^{k-1}(I_d - K_{k-i}H_{k-i})F_{k-i}} \\ &+ \sum_{j=1}^{k} \left(Q^{w,j}_{\left(\prod_{i=0}^{k-j-1}(I_d-K_{k-i}H_{k-i})F_{k-i}\right)(I_d-K_jH_j)} + Q^{v,j}_{\left(\prod_{i=0}^{k-j-1}(I_d-K_{k-i}H_{k-i})F_{k-i}\right)K_j} \right), \end{aligned} \quad (2.5)$$

where the product notation is understood to be right multiplication and equal to the identity matrix when the product is empty.

2.2 Finite L^2-norm noise environment

Suppose now that $p = 2$, so that each noise w_k and v_k has finite L^2-norm. The integrand of (2.4) is no longer piecewise, simplifying computations. Since each L^2-norm is finite, the second moments of $\|w_k\|$ and $\|v_k\|$ are finite and given by

$$W_k \stackrel{def}{=} \int_{\mathbb{R}^d} \|x\|^2 \, Q^{w,k}(dx)$$

and

$$V_k \stackrel{def}{=} \int_{\mathbb{R}^m} \|x\|^2 \, Q^{v,k}(dx),$$

respectively. Then the initial and updated errors are given by

$$l_0^2 = \int_{\mathbb{R}^d} \|x\|^2 \, Q^0(dx) = \int_{\mathbb{R}^d} \|x\|^2 \, Q^{w,0}(dx) = W_0$$

and

$$\begin{aligned} l_k^2 &= \int_{\mathbb{R}^d} \|x\|^2 \, Q^k(dx) \\ &= \int_{\mathbb{R}^d} \|x\|^2 \left(Q^{k-1}_{(I_d-K_kH_k)F_k} + Q^{w,k}_{I_d-K_kH_k} + Q^{v,k}_{K_k} \right)(dx) \end{aligned}$$

$$= \int_{\mathbb{R}^d} \|(I_d - K_k H_k) F_k x\|^2 Q^{k-1}(dx)$$
$$+ \int_{\mathbb{R}^d} \|(I_d - K_k H_k) x\|^2 Q^{w,k}(dx) + \int_{\mathbb{R}^m} \|K_k x\|^2 Q^{v,k}.$$

Instead of minimizing l_k, will minimize an upper bound on l_k. Using the subordinate matrix 2-norm induced by the Euclidean vector norm $\|\cdot\|$, we can bound the magnitude of the updated error by

$$\begin{aligned} l_k^2 \leq \|(I_d - K_k H_k) F_k\|_2^2 \int_{\mathbb{R}^d} \|x\|^2 Q^{k-1}(dx) \\ + \|I_d - K_k H_k\|_2^2 \int_{\mathbb{R}^d} \|x\|^2 Q^{w,k}(dx) + \|K_k\|_2^2 \int_{\mathbb{R}^m} \|x\|^2 Q^{v,k}(dx) \\ = \|(I_d - K_k H_k) F_k\|_2^2 \, l_{k-1}^2 + \|I_d - K_k H_k\|_2^2 \, W_k + \|K_k\|_2^2 \, V_k. \end{aligned}$$

Let us define

$$\hat{l}_0^2 \stackrel{def}{=} l_0^2 \text{ and}$$
$$\hat{l}_k^2 \stackrel{def}{=} \|(I_d - K_k H_k) F_k\|_2^2 \, \hat{l}_{k-1}^2 + \|I_d - K_k H_k\|_2^2 \, W_k + \|K_k\|_2^2 \, V_k. \tag{2.6}$$

The above definitions allow us to iteratively update our error estimates using only the previous error update. Now we must determine an approximating procedure that minimizes $\hat{l}_{k|k}$. While the subordinate matrix 2-norm has the desirable property that $\|I\|_2 = 1$, it presents a challenge in minimizing $\hat{l}_{k|k}$. For a matrix A, the Frobenius norm

$$\|A\|_F \stackrel{def}{=} \sqrt{\operatorname{trace}(A^T A)}, \tag{2.7}$$

while larger than the subordinate matrix 2-norm $\|A\|_2^2$, is easier to compute. To this end, we may bound (2.6) by

$$\hat{l}_k^2 \leq \|(I_d - K_k H_k) F_k\|_F^2 \, \hat{l}_{k-1}^2 + \|I_d - K_k H_k\|_F^2 \, W_k + \|K_k\|_F^2 \, V_k. \tag{2.8}$$

The right hand side is now easy to minimize by recognizing it as a multivariate multiple regression minimizing the residual sum of squares of the model

$$\begin{bmatrix} \hat{l}_{k-1} I_d & \sqrt{W_k} I_d & 0_{d\times m} \end{bmatrix} = K_k \begin{bmatrix} \hat{l}_{k-1} H_k F_k & \sqrt{W_k} H_k & \sqrt{V_k} I_m \end{bmatrix}.$$

It is well known that for a multiple multivariate linear regression model $Y = BX$, the least squares estimate of the matrix B is $YX^T \left(XX^T\right)^{-1}$. Hence

$$K_k = \left(\hat{l}_{k-1}^2 F_k^T H_k^T + W_k H_k^T\right)\left(\hat{l}_{k-1}^2 H_k F_k F_k^T H_k^T + W_k H_k H_k^T + V_k I_m\right)^{-1}.$$

The above solution is exact in 1 dimension, since the matrix norms $\|\cdot\|_2$ and $\|\cdot\|_F$ are replaced by $|\cdot|$, and coincides with the classic Kalman filter. The algorithm is summarized in Algorithm 3.

2.3 α-Stable noise environment

For the next example, fix $1 < \alpha < 2$ and assume that x_0 is known, so that $Q^{w,0} = \delta_0$. Assume that the signal noise sequence has the form $w_k = G\tilde{w}_k$, where $G \in \mathbb{R}^{d\times q}$ and $\tilde{w}_k$ are $\mathbb{R}^q$-valued rotationally invariant α-stable random vectors with Lévy measures $Q^{\tilde{w},k}(dx) \overset{def}{=} c_k^{\tilde{w}} \|x\|^{-\alpha-q}\, dx$. By Corollary 1.1.2, w_k are infinitely divisible $\mathbb{R}^d$-valued random vectors with Lévy-Khintchine triplets $(0,0,Q^{w,k}) \overset{def}{=} (0,0,Q_G^{\tilde{w},k})$. Assume v_k are $\mathbb{R}^m$-valued rotationally invariant α-stable random vectors with Lévy measures $Q^{v,k}(dx) \overset{def}{=} c_k^{v} \|x\|^{-\alpha-m}\, dx$. Before determining the Kalman gain, we will need the following computations in the analysis of this problem: Fix $1 \le p < \alpha$ and let $A \in \mathbb{R}^{d\times d}$. I denote by σ the uniform measure on the unit sphere. Then

$$\begin{aligned}
\int_{\mathbb{R}^d} \frac{\|x\|^2}{l^2} \mathbb{1}_{\{\|x\|<l\}} Q_A^{w,k}(dx) &= \frac{1}{l^2}\int_{\mathbb{R}^q} \|Ax\|^2\, \mathbb{1}_{\{\|Ax\|<l\}} Q_G^{\tilde{w},k}(dx) \\
&= \frac{1}{l^2}\int_{\mathbb{R}^q} \|AGx\|^2\, \mathbb{1}_{\{\|AGx\|<l\}} c_k^{\tilde{w}} \|x\|^{-\alpha-q}\, dx \\
&= \frac{c_k^{\tilde{w}}}{l^2}\int_0^\infty \int_{S^{q-1}} \|AGru\|^2\, \mathbb{1}_{\{\|AGru\|<l\}} \|ru\|^{-\alpha-q}\, \sigma(du) r^{q-1} dr \\
&= \frac{c_k^{\tilde{w}}}{l^2}\int_0^\infty \int_{S^{q-1}} \|AGu\|^2\, \mathbb{1}_{\{\|AGu\|<l/r\}} \sigma(du) r^{1-\alpha} dr \\
&= \frac{c_k^{\tilde{w}}}{l^2}\int_{S^{q-1}} \|AGu\|^2 \int_0^\infty \mathbb{1}_{\{r<l/\|AGu\|,\|AGu\|\neq 0\}} r^{1-\alpha} dr \sigma(du) \\
&= \frac{c_k^{\tilde{w}}}{l^\alpha}\frac{1}{2-\alpha}\int_{S^{q-1}} \|AGu\|^\alpha\, \sigma(du), \qquad (2.9)
\end{aligned}$$

Algorithm 3 Kalman filter for finite L^2-norm noise.

1: Initialize:
$\hat{x}_{0|0} = \mathbb{E}x_0 = 0$
$\hat{l}_0^2 = W_0$
2: Predict:
$\hat{x}_{k|k-1} = F_k \hat{x}_{k-1|k-1} + B_k u_k$
3: Update:
$K_k = \left(\hat{l}_{k-1}^2 F_k^T H_k^T + W_k H_k^T\right)\left(\hat{l}_{k-1}^2 H_k F_k F_k^T H_k^T + W_k H_k H_k^T + V_k I_m\right)^{-1}$
$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k\left(y_k - H_k \hat{x}_{k|k-1}\right)$
$\hat{l}_k^2 = \|(I_d - K_k H_k) F_k\|_2^2 \hat{l}_{k-1}^2 + \|I_d - K_k H_k\|_2^2 W_k + \|K_k\|_2^2 V_k.$

and, similarly,

$$\int_{\mathbb{R}^d} \frac{\|x\|^p}{l^p} \mathbb{1}_{\{\|x\| \geq l\}} Q_A^{w,k}(dx) = \frac{c_k^{\tilde{w}}}{l^\alpha} \frac{1}{\alpha - p} \int_{S^{q-1}} \|AGu\|^\alpha \sigma(du). \tag{2.10}$$

Also, if $A \in \mathbb{R}^{d \times m}$, then

$$\begin{aligned}
\int_{\mathbb{R}^d} \frac{\|x\|^2}{l^2} \mathbb{1}_{\{\|x\|<l\}} Q_A^{v,k}(dx) &= \frac{1}{l^2} \int_{\mathbb{R}^m} \|Ax\|^2 \mathbb{1}_{\{\|Ax\|<l\}} Q^{v,k}(dx) \\
&= \frac{1}{l^2} \int_{\mathbb{R}^m} \|Ax\|^2 \mathbb{1}_{\{\|Ax\|<l\}} c_k^v \|x\|^{-\alpha-m} dx \\
&= \frac{c_k^v}{l^2} \int_0^\infty \int_{S^{m-1}} \|Aru\|^2 \mathbb{1}_{\{\|Aru\|<l\}} \|ru\|^{-\alpha-m} \sigma(du) r^{m-1} dr \\
&= \frac{c_k^v}{l^2} \int_0^\infty \int_{S^{m-1}} \|Au\|^2 \mathbb{1}_{\{\|Au\|<l/r\}} \sigma(du) r^{1-\alpha} dr \\
&= \frac{c_k^v}{l^2} \int_{S^{m-1}} \|Au\|^2 \int_0^\infty \mathbb{1}_{\{r<l/\|Au\|, \|Au\| \neq 0\}} r^{1-\alpha} dr \sigma(du) \\
&= \frac{c_k^v}{l^\alpha} \frac{1}{2-\alpha} \int_{S^{m-1}} \|Au\|^\alpha \sigma(du),
\end{aligned} \tag{2.11}$$

and, similarly,

$$\int_{\mathbb{R}^d} \frac{\|x\|^2}{l^2} \mathbb{1}_{\{\|x\|<l\}} Q_A^{v,k}(dx) = \frac{c_k^v}{l^\alpha} \frac{1}{2-\alpha} \int_{S^{m-1}} \|Au\|^\alpha \sigma(du). \tag{2.12}$$

We are now ready to compute the estimated error l_k. To compute the first integral in the functional equation (2.4) for l_k, we use the iterative formulation and the integral formulas (2.9) and (2.11) to get

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \frac{\|x\|^2}{l_k^2} \mathbb{1}_{\{\|x\|<l_k\}} Q^k(dx) \\
&\quad = \int_{\mathbb{R}^d} \frac{\|x\|^2}{l_k^2} \mathbb{1}_{\{\|x\|<l_k\}} Q^{w,0}_{\prod_{i=0}^{k-1}(I_d-K_{k-i}H_{k-i})F_{k-i}}(dx) \\
&\quad + \int_{\mathbb{R}^d} \frac{\|x\|^2}{l_k^2} \mathbb{1}_{\{\|x\|<l_k\}} \left(\sum_{j=1}^{k} Q^{w,j}_{\left(\prod_{i=0}^{k-j-1}(I_d-K_{k-i}H_{k-i})F_{k-i}\right)(I_d-K_jH_j)} \right)(dx) \\
&\quad + \int_{\mathbb{R}^d} \frac{\|x\|^2}{l_k^2} \mathbb{1}_{\{\|x\|<l_k\}} \left(\sum_{j=1}^{k} Q^{v,j}_{\left(\prod_{i=0}^{k-j-1}(I_d-K_{k-i}H_{k-i})F_{k-i}\right)K_j} \right)(dx) \\
&\quad = \frac{1}{l_k^\alpha(2-\alpha)} \sum_{j=1}^{k} c_j^{\tilde{w}} \int_{S^{q-1}} \left\| \left(\prod_{i=0}^{k-j-1} \left(I_d - K_{k-i}H_{k-i}\right) F_{k-i} \right) \left(I_d - K_jH_j\right) Gu \right\|^\alpha \sigma(du) \\
&\quad + \frac{1}{l_k^\alpha(2-\alpha)} \sum_{j=1}^{k} c_j^{v} \int_{S^{m-1}} \left\| \left(\prod_{i=0}^{k-j-1} \left(I_d - K_{k-i}H_{k-i}\right) F_{k-i} \right) K_j u \right\|^\alpha \sigma(du)
\end{aligned}
$$

and similarly, using the integral formulas (2.10) and (2.12), we have that the second integral in the functional equation (2.4) for l_k is

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \frac{\|x\|^p}{l_k^p} \mathbb{1}_{\{\|x\|\geq l_{k|k-1}\}} Q^k(dx) \\
&\quad = \frac{1}{l_k^\alpha(\alpha-p)} \sum_{j=1}^{k} c_j^{\tilde{w}} \int_{S^{q-1}} \left\| \left(\prod_{i=0}^{k-j-1} \left(I_d - K_{k-i}H_{k-i}\right) F_{k-i} \right) \left(I_d - K_jH_j\right) Gu \right\|^\alpha \sigma(du) \\
&\quad + \frac{1}{l_k^\alpha(\alpha-p)} \sum_{j=1}^{k} c_j^{v} \int_{S^{m-1}} \left\| \left(\prod_{i=0}^{k-j-1} \left(I_d - K_{k-i}H_{k-i}\right) F_{k-i} \right) K_j u \right\|^\alpha \sigma(du).
\end{aligned}
$$

Since l_k satisfies (2.4), the two computations above combine to give

$$
\begin{aligned}
l_k^\alpha = &\left(\frac{1}{2-\alpha} + \frac{1}{\alpha-p} \right) \left(\sum_{j=1}^{k} c_j^{v} \int_{S^{m-1}} \left\| \left(\prod_{i=0}^{k-j-1} \left(I_d - K_{k-i}H_{k-i}\right) F_{k-i} \right) K_j u \right\|^\alpha \sigma(du) \right. \\
&\left. + \sum_{j=1}^{k} c_j^{\tilde{w}} \int_{S^{q-1}} \left\| \left(\prod_{i=0}^{k-j-1} \left(I_d - K_{k-i}H_{k-i}\right) F_{k-i} \right) \left(I_d - K_jH_j\right) Gu \right\|^\alpha \sigma(du) \right). \quad (2.13)
\end{aligned}
$$

While no closed form solution exists for K_k minimizing l_k (except in the 1-dimensional case), we can get a tractable problem, as we did in the $p = 2$ example, by minimizing an upper bound of l_k. Define

$$W_k \stackrel{def}{=} c_k^{\tilde{w}} \left(\frac{1}{2-\alpha} + \frac{1}{\alpha - p} \right) \sigma \left(S^{q-1} \right)$$

and

$$V_k \stackrel{def}{=} c_k^{v} \left(\frac{1}{2-\alpha} + \frac{1}{\alpha - p} \right) \sigma \left(S^{m-1} \right).$$

Observe that $l_0^\alpha = 0$ and that

$$\begin{aligned}
l_k^\alpha &= \left(\frac{1}{2-\alpha} + \frac{1}{\alpha - p} \right) \Bigg(c_k^v \int_{S^{m-1}} \left\| K_k u \right\|^\alpha \sigma(du) \\
&+ \sum_{j=1}^{k-1} c_j^v \int_{S^{m-1}} \left\| \left(I_d - K_k H_k \right) F_k \left(\prod_{i=1}^{k-j-1} \left(I_d - K_{k-i} H_{k-i} \right) F_{k-i} \right) K_j u \right\|^\alpha \sigma(du) \\
&+ c_k^{\tilde{w}} \int_{S^{q-1}} \left\| \left(I_d - K_k H_k \right) G u \right\|^\alpha \sigma(du) \\
&+ \sum_{j=1}^{k-1} c_j^{\tilde{w}} \int_{S^{q-1}} \Bigg\| \left(I_d - K_k H_k \right) F_k \left(\prod_{i=1}^{k-j-1} \left(I_d - K_{k-i} H_{k-i} \right) F_{k-i} \right) \\
&\qquad \left(I_d - K_j H_j \right) G u \Bigg\|^\alpha \sigma(du) \Bigg) \\
&\le \left\| \left(I_d - K_k H_k \right) F_k \right\|_2^\alpha l_{k-1}^\alpha + \left\| \left(I_d - K_k H_k \right) G \right\|_2^\alpha W_k + \left\| K_k \right\|_2^\alpha V_k,
\end{aligned} \tag{2.14}$$

where, for a matrix A, $\|A\|_2 \stackrel{def}{=} \max_{\|x\|=1} \|Ax\|$ is the subordinate matrix 2-norm induced by the Euclidean vector norm $\|\cdot\|$. As we did in the $p = 2$ case, we consider

$$\hat{l}_k^\alpha \stackrel{def}{=} \left\| \left(I_d - K_k H_k \right) F_k \right\|_2^\alpha \hat{l}_{k-1}^\alpha + \left\| \left(I_d - K_k H_k \right) G \right\|_2^\alpha W_k + \left\| K_k \right\|_2^\alpha V_k \tag{2.15}$$

instead of l_k. The above iterative definition will allow us to minimize the convenient upper bound $\hat{l}_k$ of l_k. As before, using these upper bounds, our error estimates may be updated using only the previous estimated error.

Now we must determine an approximating procedure that minimizes $\hat{l}_k$. As we did in the $p = 2$ case, we will minimize the Frobenius norm $\|\cdot\|_F$ (see (2.7) for definition) instead of the subordinate matrix 2-norm $\|\cdot\|_2$. To this end, we may bound (2.15) by

$$\begin{aligned}\hat{l}_k^\alpha &= \|(I_d - K_k H_k)\, F_k\|_2^\alpha\, \hat{l}_{k-1}^\alpha + \|(I_d - K_k H_k)\, G\|_2^\alpha\, W_k + \|K_k\|_2^\alpha\, V_k \\ &\leq \|(I_d - K_k H_k)\, F_k\|_F^\alpha\, \hat{l}_{k-1}^\alpha + \|(I_d - K_k H_k)\, G\|_F^\alpha\, W_k + \|K_k\|_F^\alpha\, V_k \\ &= \hat{l}_{k-1}^\alpha\, \|(I_d - K_k H_k)\, F_k\|_F^{\alpha-2}\, \|(I_d - K_k H_k)\, F_k\|_F^2 \\ &\quad + W_k\, \|(I_d - K_k H_k)\, G\|_F^{\alpha-2}\, \|(I_d - K_k H_k)\, G\|_F^2 + V_k\, \|K_k\|_F^{\alpha-2}\, \|K_k\|_F^2\,,\end{aligned} \tag{2.16}$$

the right hand side now being easier to minimize as follows: suppose that we have an estimate $K_k^{(t)}$ for K_k. Then we may iteratively improve our estimate of K_k by finding $K_k^{(t+1)}$ minimizing

$$w_1^{(t)} \left\| \left(I_d - K_k^{(t+1)} H_k\right) F_k \right\|_F^2 + w_2^{(t)} \|(I_d - K_k H_k)\, G\|_F^2 + w_3^{(t)} \left\|K_k^{(t+1)}\right\|_F^2, \tag{2.17}$$

where

$$\begin{aligned} w_1^{(t)} &\stackrel{def}{=} \hat{l}_{k-1}^\alpha \left\| \left(I_d - K_k^{(t)} H_k\right) F_k \right\|_F^{\alpha-2}, \\ w_2^{(t)} &\stackrel{def}{=} W_k \left\| \left(I_d - K_k^{(t)} H_k\right) G \right\|_F^{\alpha-2}, \end{aligned}$$

and

$$w_3^{(t)} \stackrel{def}{=} V_k \left\| K_k^{(t)} \right\|_F^{\alpha-2}.$$

We may recognize (2.17) as a multivariate multiple regression minimizing the residual sum of squares of the model

$$\left[\sqrt{w_1^{(t)}} F_k \quad \sqrt{w_2^{(t)}} G \quad 0_{d\times m}\right] = K_k^{(t+1)} \left[\sqrt{w_1^{(t)}} H_k F_k \quad \sqrt{w_2^{(t)}} H_k G \quad \sqrt{w_3^{(t)}} I_m\right].$$

It is well known that for a multiple multivariate linear regression model $Y = BX$, the least squares estimate of the matrix B is $YX^T\left(XX^T\right)^{-1}$. Hence

$$\begin{aligned} K_k^{(t+1)} = {}& \left(w_1^{(t)} F_k F_k^T H_k^T + w_2^{(t)} G G^T H_k^T\right) \cdot \\ & \left(w_1^{(t)} H_k F_k F_k^T H_k^T + w_2^{(t)} H_k G G^T H_k^T + w_3^{(t)} I_m\right)^{-1}. \end{aligned}$$

This approximating technique is known as *iteratively reweighted least squares.* See, for example, Gentle (2007, L^p norms and Iteratively Reweighted Least Squares, pg. 232) for an overview. Iteratively reweighted least squares approximates K_k minimizing (2.16) by

$$K_k = \lim_{t \to \infty} K_k^{(t)}.$$

The above procedure is easily implemented on a computer and allows us to approximate the optimal Kalman gain K_k using the iteratively reweighted least squares algorithm. We may initialize the algorithm by the least squares solution, where $w_1^{(1)}$, $w_2^{(1)}$, and $w_3^{(1)}$ are taken to be 1, and compute the error to be any matrix norm of the difference $K_k^{(t+1)} - K_k^{(t)}$. The iteratively reweighted least squares algorithm is implemented in Algorithm 4 and the Kalman filter is implemented in Algorithm 5.

Algorithm 5 can become unstable over time due to the fact that we are not actually keeping track of the actual errors, but instead, an upper bound on the errors using the matrix norm inequality

$$\|AB\|_2 \leq \|A\|_2 \|B\|_2 .$$

At each step, we used this inequality, and hence our estimated error $\hat{l}_k$ tends to be much larger than the actual error l_k. If we are only tracking the target short term, Algorithm 5 works very well. However, for long term tracking we may improve estimation of x_k at the expense of computational inefficiency by keeping track of more of the matrix multiplications in (2.13) instead of approximating the error by (2.14). If we are filtering off-line and computational speed is not a priority, we may use (2.13) for l to improve performance. Alternatively, we may perform a statistical analysis to determine how large an overestimate (2.14) tends to be and adjust accordingly.

2.3.1 Exact 1-dimensional filtering

As mentioned above, we can get an exact closed form solution in dimension 1 and demonstrate this here. If $d = m = q = 1$, then the inequality (2.14) is in fact an equality, since the matrix norms are replaced by $|\cdot|$, giving

$$\begin{aligned} l_k^\alpha &= |1 - K_k H_k|^\alpha |F_k|^\alpha l_{k-1}^\alpha + |1 - K_k H_k|^\alpha W_k + |K_k|^\alpha V_k \\ &= |1 - K_k H_k|^\alpha \left(|F_k|^\alpha l_{k-1}^\alpha + W_k \right) + |K_k|^\alpha V_k, \end{aligned}$$

Algorithm 4 Iteratively reweighted least squares.

1: Initialize $K_k^{(1)}$ to the least squares solution with weights of 1:
$K_k^{(1)} = \left(F_k F_k^T H_k^T + G G^T H_k^T\right)\left(H_k F_k F_k^T H_k^T + H_k G G^T H_k^T + I_m\right)^{-1}$

2: While $error > \varepsilon$ and $t \leq maxiterations$

Compute $w_1^{(t)} = \hat{l}_{k-1}^{\alpha} \left\| \left(I_d - K_k^{(t)} H_k\right) F_k \right\|_F^{\alpha-2}$

$w_2^{(t)} = W_k \left\| \left(I_d - K_k^{(t)} H_k\right) G \right\|_F^{\alpha-2}$

$w_3^{(t)} = V_k \left\| K_k^{(t)} \right\|_F^{\alpha-2}$

Compute $K_k^{(t+1)} = \left(w_1^{(t)} F_k F_k^T H_k^T + w_2^{(t)} G G^T H_k^T\right) \cdot \left(w_1^{(t)} H_k F_k F_k^T H_k^T + w_2^{(t)} H_k G G^T H_k^T + w_3^{(t)} I_m\right)^{-1}$

Compute $error = \left\| K_k^{(t+1)} - K_k^{(t)} \right\|_F$

Increment t.

3: $K_k = K_k^{(t)}$.

where we have assumed without loss of generality that $G \equiv 1$ (it may be absorbed into c_k^w in dimension 1). Here, W_k and V_k reduce to

$$W_k = 2c_k^w \left(\frac{1}{2-\alpha} + \frac{1}{\alpha - p}\right)$$

and

$$V_k = 2c_k^v \left(\frac{1}{2-\alpha} + \frac{1}{\alpha - p}\right).$$

Let us define $l_{k|k-1}^{\alpha} \stackrel{def}{=} |F_k|^{\alpha} l_{k-1}^{\alpha} + W_k$. One can show by arguments similar to those used to derive (2.13) that $l_{k|k-1}$ measures the magnitude of the predicted error $e_{k|k-1}$ just as l_k measures the magnitude of the updated error $e_{k|k}$. We then have

$$l_k^{\alpha} = |1 - K_k H_k|^{\alpha} l_{k|k-1}^{\alpha} + |K_k|^{\alpha} V_k$$

and may minimize l_k by standard calculus. The derivative of l_k^{α} is computed as

$$\frac{\mathrm{d}\, l_k^{\alpha}}{\mathrm{d}\, K_k} = \alpha\, |1 - K_k H_k|^{\alpha-1} \operatorname{sign}(1 - K_k H_k)\,(-H_k)\, l_{k|k-1}^{\alpha} + \alpha\, |K_k|^{\alpha-1} \operatorname{sign}(K_k) V_k.$$

Equating to 0 and solving, we see that

Algorithm 5 Kalman filter for α-stable noise.

1: Initialize:
$\hat{x}_{0|0} = x_0$
$\hat{l}_0^\alpha = 0$
2: Predict:
$\hat{x}_{k|k-1} = F_k \hat{x}_{k-1|k-1} + B_k u_k$
3: Update:
Approximate K_k by *iteratively reweighted least squares* Algorithm 4.
$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \left(y_k - H_k \hat{x}_{k|k-1} \right)$
$\hat{l}_k^\alpha = \|(I_d - K_k H_k) F_k\|_2^\alpha \hat{l}_{k-1}^\alpha + \|(I_d - K_k H_k) G\|_2^\alpha W_k + \|K_k\|_2^\alpha V_k.$

$$\operatorname{sign}(K_k) = \operatorname{sign}(1 - K_k H_k) \operatorname{sign}(H_k)$$

and

$$|K_k| V_k^{\frac{1}{\alpha-1}} = |1 - K_k H_k| \, |H_k|^{\frac{1}{\alpha-1}} \, l_{k|k-1}^{\frac{\alpha}{\alpha-1}}.$$

Hence,

$$K_k V_k^{\frac{1}{\alpha-1}} = (1 - K_k H_k) \operatorname{sign}(H_k) \, |H_k|^{\frac{1}{\alpha-1}} \, l_{k|k-1}^{\frac{\alpha}{\alpha-1}},$$

which is easily solved for K_k to get the optimal Kalman gain as

$$K_k = \left(\operatorname{sign}(H_k) \, |H_k|^{\frac{1}{\alpha-1}} \, l_{k|k-1}^{\frac{\alpha}{\alpha-1}} \right) \left(|H_k|^{\frac{\alpha}{\alpha-1}} \, l_{k|k-1}^{\frac{\alpha}{\alpha-1}} + V_k^{\frac{1}{\alpha-1}} \right)^{-1}. \qquad (2.18)$$

If we take $\alpha = 2$ in the above equation, we have exactly the classic Kalman gain (2.1) (ignoring the fact that the dispersion V_k, playing a similar role as variance in the normal distribution, is infinite). The Kalman filter algorithm is implemented in Algorithm 6. As opposed to the higher dimensional solutions of the Kalman filter for finite L^2-norm noise and α-stable noise I have given, the Kalman gain (2.18) is exact in the sense that it minimizes the error l_k, not an upper bound on l_k. We next present simulations utilizing these results for the α-stable noise environment.

2.3.2 Vehicle tracking

Suppose we are tracking a vehicle moving in a straight line. The vehicle's position is measured every T seconds, at which time we can change the velocity $u = u_{k+1}$. Then

Algorithm 6 Kalman filter for 1 dimensional α-stable noise.

1: Initialize:
$\hat{x}_{0|0} = \mathbb{E}x_0 = 0$
$l_0^\alpha = W_0$
2: Predict:
$\hat{x}_{k|k-1} = F_k\hat{x}_{k-1|k-1} + B_k u_k$
$l_{k|k-1}^\alpha = |F_k|^\alpha l_{k-1}^\alpha + W_k$
3: Update:
$K_k = \left(\operatorname{sign}(H_k)\,|H_k|^{\frac{1}{\alpha-1}}\, l_{k|k-1}^{\frac{\alpha}{\alpha-1}}\right)\left(|H_k|^{\frac{\alpha}{\alpha-1}}\, l_{k|k-1}^{\frac{\alpha}{\alpha-1}} + V_k^{\frac{1}{\alpha-1}}\right)^{-1}$
$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k\left(y_k - H_k\hat{x}_{k|k-1}\right)$
$l_k^\alpha = |1 - K_k H_k|^\alpha l_{k|k-1}^\alpha + |K_k|^\alpha V_k$

the position of the vehicle is modeled by

$$x_k = x_{k-1} + Tu_k.$$

In actuality, the position of the vehicle at each time is perturbed by circumstances beyond our control (potholes, gusts of wind, etc.). A more realistic model is

$$x_k = x_{k-1} + Tu_k + w_k,$$

where w_k is a random "noise". At each time increment, we observe the position of the vehicle, which is also contaminated by a random noise. The observation y_k is modeled by

$$y_k = x_k + v_k,$$

where v_k is a random "noise". Our objective is to efficiently estimate the position of the vehicle at time k. First, we could completely ignore our observation y_k and predict the position of the vehicle to be $\hat{x}_k = \hat{x}_{k-1} + Tu_k$. Or, we could completely ignore the dynamics of the system and predict the position of the vehicle to be the observation $\hat{x}_k = y_k$. In actuality, we would like to use each piece of information: the dynamics of the system and the observation. If we restrict to linear estimates and assume that $\{w_k\}$ and $\{v_k\}$ are independent symmetric α-stable random variables, then we may apply the Kalman filter Algorithm 6 to estimate the position of the vehicle $x_{k|k}$ at time k. Figure 2.1 is a simulation with parameters $p = 1$, $\alpha = 1.4$, $T = 0.1$, and constant velocity $u_k = u = 4$ throughout every time increment. The dispersion parameter c_k^w of w_k is taken to be small ($c_k^w = 0.1$). This represents that the potholes, gusts of wind, etc. have minimal effect on the position of the vehicle. The dispersion parameter c_k^v of v_k is taken to be large in comparison to v_k^w ($c_k^v = 5$). This parameter represents the known accuracy of the gps technology. The classic Kalman filter Algorithm 1 weights the observation to heavily in this case, as it does not expect such extreme tail events that occur under an α-stable distribution. We can see in Figure 2.1 the tail events that occur in the observation noise. Such tail events have probability ≈ 0 under the Gaussian distribution and are not expected in the classic Kalman filter.

2.3.3 Aircraft tracking

As a last example, we consider two models commonly employed in the tracking of an aircraft. Ignoring altitude, the system state being tracked is $\mathbf{x} = (x_1, \dot{x}_1, x_2, \dot{x}_2)$. The system dynamics of a maneuvering aircraft are modeled by the *constant velocity (CV)* model and the *coordinated turn (CT)* model (see e.g. Bar-Shalom et al. (2001, Section 11.7) for an overview). The models are

$$\mathbf{x}_k = F\mathbf{x}_{k-1} + \begin{bmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{bmatrix} \mathbf{w}_k,$$

where the system dynamics matrix for the CV model is

$$F \stackrel{def}{=} \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and for the CT model is

$$F \stackrel{def}{=} \begin{bmatrix} 1 & \frac{\sin \omega T}{\omega} & 0 & -\frac{1-\cos \omega T}{\omega} \\ 0 & \cos \omega T & 0 & -\sin \omega T \\ 0 & \frac{1-\cos \omega T}{\omega} & 1 & \frac{\sin \omega T}{\omega} \\ 0 & \sin \omega T & 0 & \cos \omega T \end{bmatrix}.$$

In practice, the turn rate ω is unknown. One would need to consider the augmented state matrix $\mathbf{x}_k = (x_1, \dot{x}_1, x_2, \dot{x}_2, \omega)$, for which the system model is now non-linear. Standard practice is to then approximate by a first order expansion. We assume here that the turn rate ω is constant and known for simulation purposes. The signal noise $\mathbf{w}_k$ is a 2-dimensional rotationally invariant α-stable random vector. At each time increment, we observe the position of the aircraft, which is also contaminated by a 2-dimensional rotationally invariant α-stable random noise. Then the observation $\mathbf{y}_k$ is

$$\mathbf{y}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}_k + \mathbf{v}_k.$$

We apply Algorithm 5 to estimate the position of the vehicle by $\hat{\mathbf{x}}_{k|k}$. Figure 2.2 and Figure 2.3 are simulations of the CV and CT models respectively. The parameters were taken as $p = 1$, $\alpha = 1.4$, $T = 0.1$, $c_k^w = 0.1$, and $c_v^k = 3$. As in the vehicle tracking example, the classic Kalman filter can perform poorly when tail events occur. If we mistakenly believe that the noise is normally distributed, then we do not anticipate such extreme tail events experienced in the noisy observation. Therefore, the classic Kalman filter is again weighting the observation to heavily and underperforms the α-stable Kalman filter Algorithm 5.

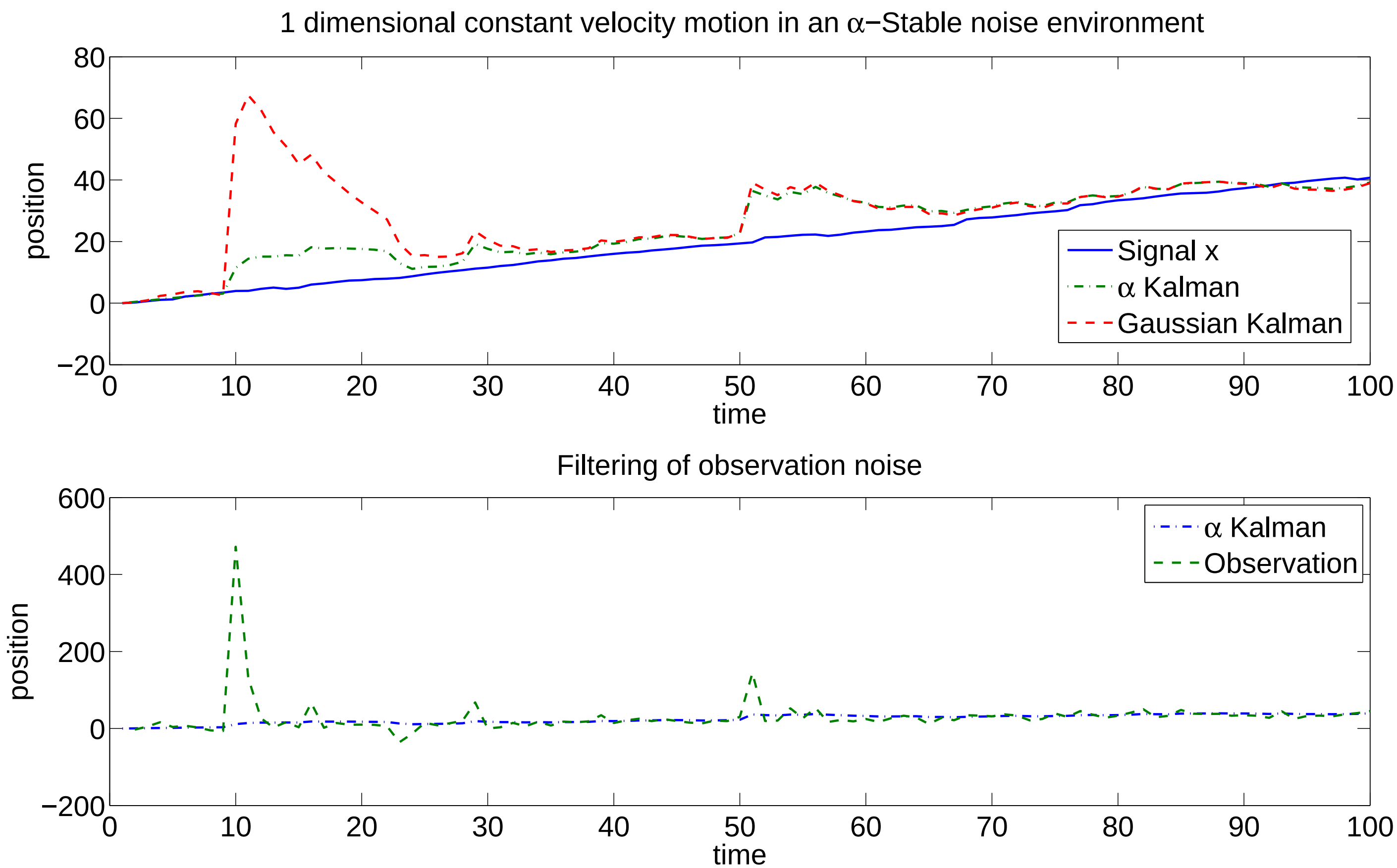

Figure 2.1: α-Stable Kalman filter for constant velocity 1 dimensional motion.

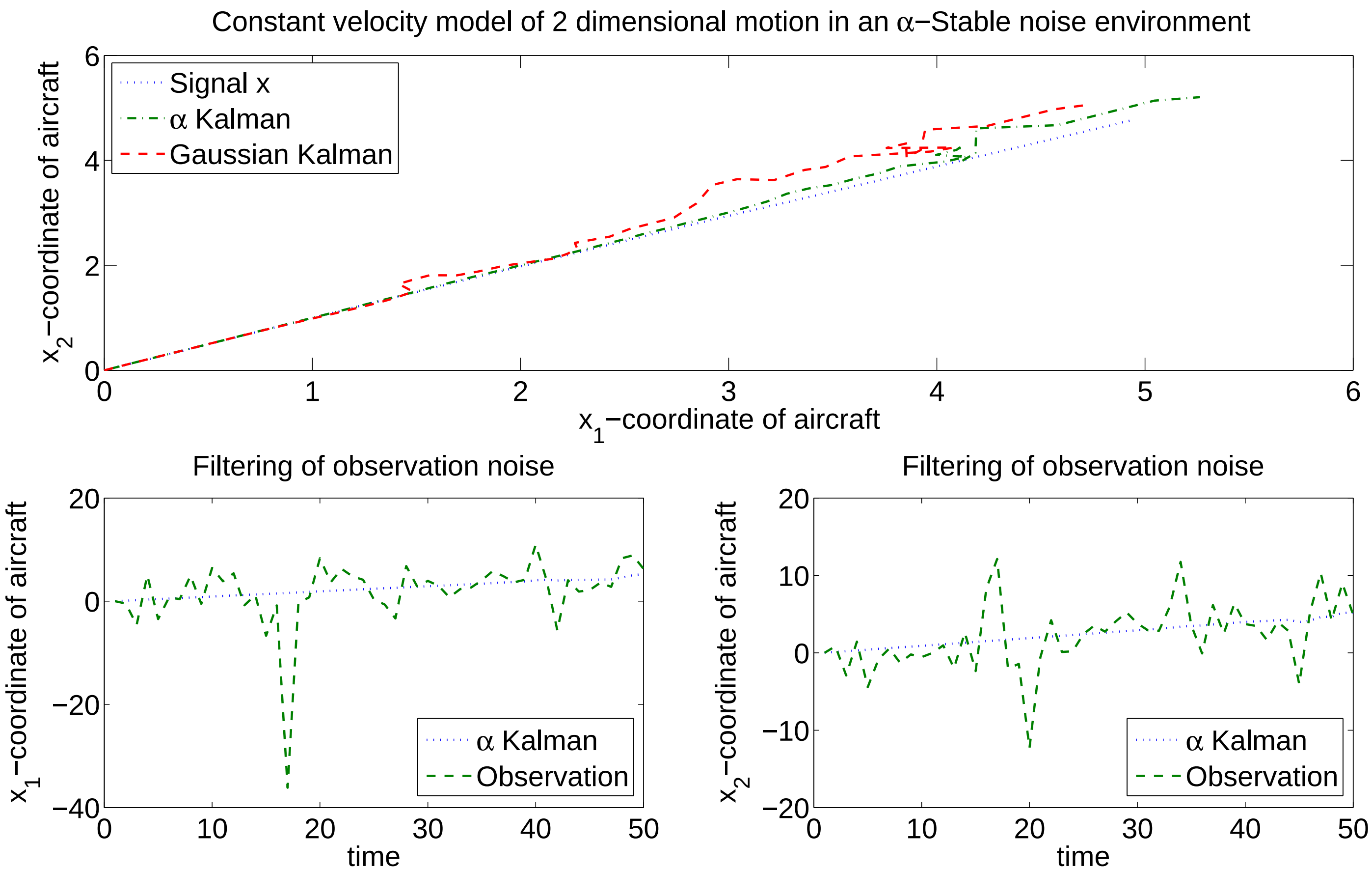

Figure 2.2: 2-D constant velocity model (CV).

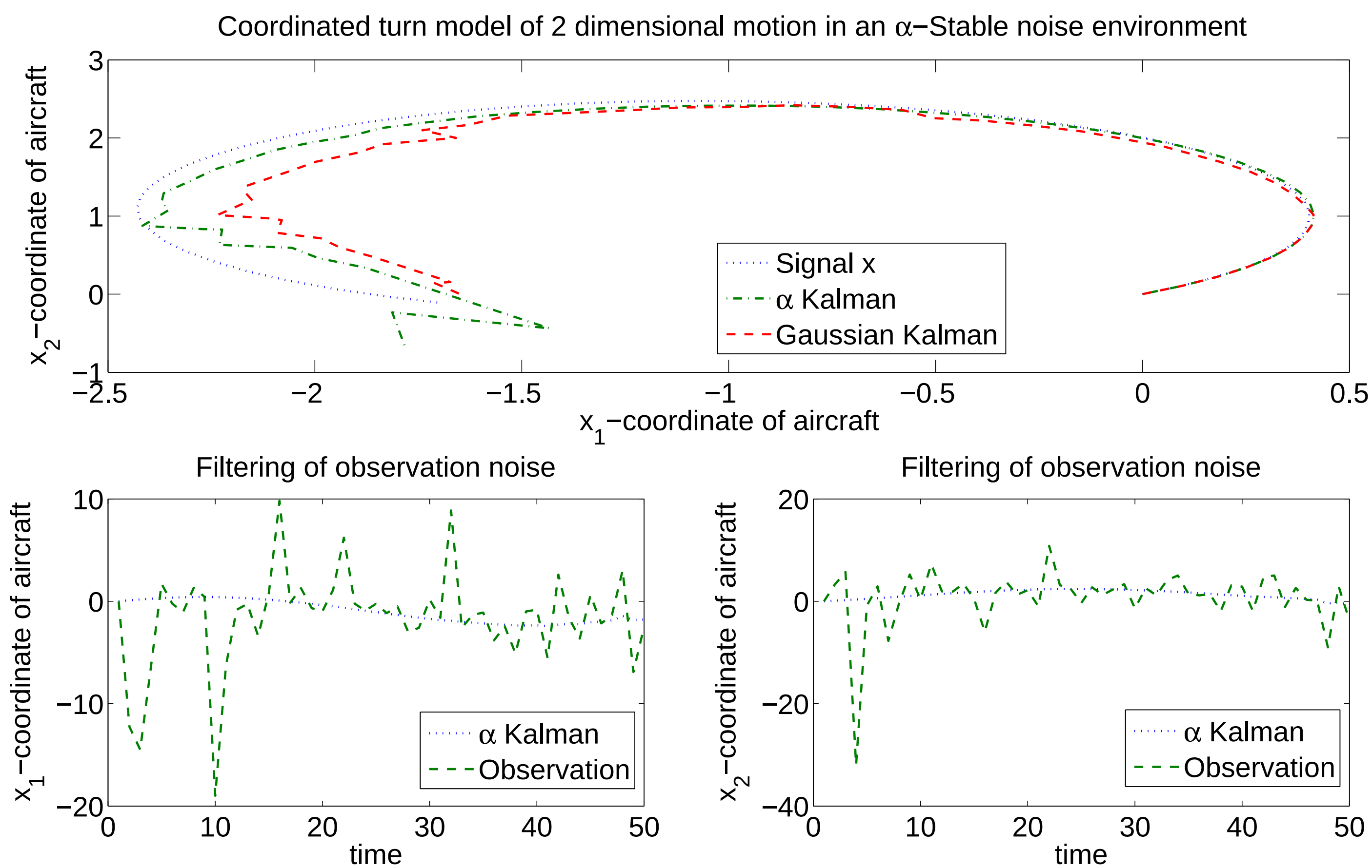

Figure 2.3: 2-D coordinated turn model (CT).

Chapter 3

Infinitely Divisible Random Measures

3.1 Introduction

Stochastic integrals are a useful and widely employed method to capture unmodeled effects in continuous time dynamical models. Gaussian white noise and Poisson random measures are popular driving terms for the stochastic integration as each are described by parameters that can be statistically estimated. In this chapter we focus on the family of *infinitely divisible random measures*, which includes both of the aforementioned random measures. First, let us recall a few basic facts about Gaussian white noise. If $\{B_t\}_{t\geq 0}$ is a Brownian motion, then

$$B(C) \stackrel{def}{=} \int_0^T \mathbb{1}_C(t) dB_t,$$

for arbitrary Borel set $C \in \mathcal{B}([0,T])$, is a Gaussian white noise satisfying

i. $B(\emptyset) = 0$ a.s.

ii. If $C_1, \cdots, C_n$ are disjoint sets in $\mathcal{B}([0,T])$, then $\{B(C_k)\}_{k=1}^n$ is a sequence of independent variables such that

$$B\left(\bigcup_{k=1}^n C_k\right) = \sum_{k=1}^n B(C_k) \text{ a.s.}$$

and

iii. $B(C) \sim N\left(0, Leb(C)\sigma^2\right)$ for every $C \in \mathcal{B}([0,T])$.

Item iii follows by Itô's isometry. That $B(C)$ is normally distributed is a special case of the more general condition that $B(C)$ is infinitely divisible. It is this property that will be our focus. By requiring $B(C)$ to be infinitely divisible, we get a parameterization from the Lévy-Khintchine triplet and, as we saw in Chapter 2, can give computations in terms of these parameters. With this motivation, we now turn to infinitely divisible random measures. Fix a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{0\le t\le T})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space. Throughout, for any set $S \subset \mathbb{R}^n$, we denote by $\mathcal{B}_0(S)$ the Borel $\sigma - ring$ generated by S.

Definition 3.1.1. Let $Z : \mathcal{B}_0([0,T] \times \mathbb{R}^d) \times \Omega \to \mathbb{R}$. $Z = \{Z(C)\}_{C\in\mathcal{B}_0([0,T]\times\mathbb{R}^d)}$ is an *infinitely divisible random measure* if

i. $Z(\emptyset) = 0$ a.s.,

ii. For every sequence $\{C_i\} \subset \mathcal{B}_0\left([0,T] \times \mathbb{R}^d\right)$ of pairwise disjoint sets, $\{Z(C_i)\}$ is a sequence of independent random variables and if $\bigcup C_i \in \mathcal{B}_0\left([0,T] \times \mathbb{R}^d\right)$,

$$Z\left(\bigcup_{i=1}^{\infty} C_i\right) = \sum_{i=1}^{\infty} Z(C_i) \text{ a.s.,} \tag{3.1}$$

and

iii. For every $C \in \mathcal{B}_0([0,T] \times \mathbb{R}^d)$, $Z(C)$ is an infinitely divisible random variable.

Stochastic integration of deterministic functions driven by infinitely divisible random measures was studied by Rajput and Rosiński (1989). In their work, they showed that the Lévy-Khintchine triplet of $Z(C)$ (see Theorem 1.1.1) is

$$Z(C) \sim \left(\int_C b(t,x)m(dt,dx), \int_C \sigma^2(t,x)m(dt,dx), F_C\right), \tag{3.2}$$

where $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $\sigma^2 : [0,T] \times \mathbb{R}^d \to \mathbb{R}_+$, and F_C is a Lévy measure on $\mathbb{R}$. F_C has the property that there exists a unique σ-finite measure F on $\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{B}(\mathbb{R})$ such that $F(C \times B) = F_C(B)$ for every $C \in \mathcal{B}([0,T] \times \mathbb{R}^d)$ and for every $B \in \mathcal{B}(\mathbb{R})$. Moreover, there exists a σ-finite measure m on $\mathcal{B}([0,T] \times \mathbb{R}^d)$ and a function ν : $[0,T] \times \mathbb{R}^d \times \mathcal{B}(\mathbb{R}) \to [0,\infty]$ such that F may be disintegrated as

$$F(dt,dx,dz) = \nu(t,x,dz)m(dt,dx).$$

Measure m is called the *control measure* of Z. It satisfies the property that $Z(C) = 0$ a.s. if and only if $m(C) \equiv 0$. Also, for each fixed t and x, $\nu(t, x, \cdot)$ is a Lévy measure on $\mathcal{B}(\mathbb{R})$.

Remark 3.1.2. The characteristic function of $Z(C)$ is

$$\begin{aligned}
&\mathbb{E}\exp\{iuZ(C)\} \\
&= \exp\left\{iu\int_C b(t,x)m(dt,dx) - \frac{1}{2}u^2\int_C \sigma^2(t,x)m(dt,dx)\right. \\
&\qquad \left. + \int_{\mathbb{R}} \left(e^{iuz} - 1 - iu[\![z]\!]\right) F_C(dz)\right\} \\
&= \exp\left\{iu\int_C b(t,x)m(dt,dx) - \frac{1}{2}u^2\int_C \sigma^2(t,x)m(dt,dx)\right. \\
&\qquad \left. + \int_{C\times\mathbb{R}} \left(e^{iuz} - 1 - iu[\![z]\!]\right) F(dt,dx,dz)\right\} \\
&= \exp\left\{iu\int_C b(t,x)m(dt,dx) - \frac{1}{2}u^2\int_C \sigma^2(t,x)m(dt,dx)\right. \\
&\qquad \left. + \int_C\int_{\mathbb{R}} \left(e^{iuz} - 1 - iu[\![z]\!]\right) \nu(t,x,dz)m(dt,dx)\right\}.
\end{aligned}$$

Two examples that we will particularly focus upon are symmetric α-stable and tempered α-stable random measures. Gaussian white noise models can perform poorly when observed data contains outliers not probable under this assumption. Just as in the discrete time examples of Chapter 2, α-Stable white noise models may outperform the Gaussian assumption in this case. α-Stable processes are justified in their use by the generalized central limit theorem. Tempered α-stable processes are attained by a "uniform tilting" of the α-stable Lévy measure and as a result, have finite variance. Such processes were studied by Rosiński (2007b) and proven to exhibit α-stable short time behavior and Brownian motion long time behavior. The process still exhibits jumps but not at the expense of infinite variance, as opposed to α-stable processes. These properties have lead to a growing popularity of tempered α-stable processes as a suitable model choice. As such, there is need and use for stochastic integrals driven by such processes. To this end, let us formally define these random measures. For $0 < \alpha < 2$, an infinitely divisible random measure $Z : \mathcal{B}_0([0,T] \times \mathbb{R}^d) \times \Omega \to \mathbb{R}$ with Lévy-Khintchine triplet

$$Z(C) \sim (0, 0, m(C)\nu_\alpha) \tag{3.3}$$

is called a *symmetric α-stable random measure* when the Lévy measure ν_α is given by

$$\nu_\alpha(dx) \stackrel{def}{=} c\,|x|^{-\alpha-1}\,dx. \tag{3.4}$$

It is called a *symmetric tempered α-stable random measure* when the Lévy measure is

$$\nu_\alpha(dx) \stackrel{def}{=} c\,|x|^{-\alpha-1}\,dx. \tag{3.5}$$

Here, $c > 0$ is called the *dispersion* parameter of an α-stable distribution on $\mathbb{R}$. In the former case, the the characteristic function of $Z(C)$ is

$$\mathbb{E}\left(e^{iuZ(C)}\right) = e^{-m(C)|u|^\alpha}. \tag{3.6}$$

In the following section, we will extend work of Rajput and Rosiński (1989) to define the integral of a random field driven by an infinitely divisible random measure. For such stochastic integrals to be fully utilized, results analogous to Itô's Isometry are needed. These will be studied in Section 3.3. As examples, we will focus on the two random measures above.

3.2 Stochastic integration

3.2.1 Space of integrands

In this section we define the space of integrands for the stochastic integral. Throughout, assume that m is a control measure of Z that may be disintegrated as $m(dt, dx) = \rho(t, dx)dt$ for some function $\rho : [0, T] \times \mathcal{B}(\mathbb{R}^d) \to [0, \infty]$. Since m is a σ-finite measure, for each $t \in [0, T]$, we require that $\rho(t, \cdot)$ is also a σ-finite measure. We single out time in this assumption due to current conventions. By using the Lebesgue measure to measure time, the distribution of $Z(C)$ is stationary and one can speak of *stationary random measures*. It is worth noting that this theory of integration could be produced without the above assumption on the control measure m. In this case, one must restrict to predictable random fields. Otherwise, the following developments remain unchanged.

Definition 3.2.1. A random field $X : [0, T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ is *$\mathcal{F}_t$-adapted (nonanticipating)* if $\{\omega : X(t, x, \omega) \le a\} \in \mathcal{F}_t$ for every $a \in \mathbb{R}$ and for every $x \in \mathbb{R}^d$.

For example, one may take $\{\mathcal{F}_t\}_{t\geq 0}$ as follows: Let

$$\mathcal{F}'_t \stackrel{def}{=} \sigma\left\{Z(C) : C \subset [0,t]\times\mathbb{R}^d, C \in \mathcal{B}_0([0,T]\times\mathbb{R}^d)\right\}$$

and $\mathcal{F}_t$ be $\mathcal{F}'_t$ augmented with the null sets of $\mathcal{F}$.

Let $L^0_T\left([0,T]\times\mathbb{R}^d\times\Omega;\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F}; m\otimes\mathbb{P}\right)$ be the collection of all measurable random fields

$$X : \left([0,T]\times\mathbb{R}^d\times\Omega, \mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F}\right) \to (\mathbb{R},\mathcal{B}(\mathbb{R})),$$

such that X is $\mathcal{F}_t$-adapted. When there is no confusion, we may simply write L^0_T. Denote by $\mathbb{S}$ the collection of all simple random fields $f : [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$ of the form

$$f(t,x,\omega) = \sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij}(\omega)1_{(t_i,t_{i+1}]}(t)1_{A_j}(x), \tag{3.7}$$

where $0 \leq t_1 < \ldots < t_{m+1} \leq T$, f_{ij} are bounded $\mathcal{F}_{t_i}$-measurable random variables, and $A_j \in \mathcal{B}_0(\mathbb{R}^d)$ are disjoint subsets of $\mathbb{R}^d$ with $m((t_i,t_{i+1}]\times A_j) < \infty$ for each i and j. For $p > 0$, we denote by $L^p_T \stackrel{def}{=} L^p([0,T]\times\mathbb{R}^d\times\Omega;\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F}; m\otimes\mathbb{P})$ the collection of all random fields $X\in L^0_T$ such that

$$\begin{aligned}\|X\|_{p,T} &\stackrel{def}{=} \left(\mathbb{E}\int_0^T\int_{\mathbb{R}^d}|X(t,x)|^p\, m(dt,dx)\right)^{1/p}\\ &= \left(\mathbb{E}\int_0^T\int_{\mathbb{R}^d}|X(t,x)|^p\,\rho(t,dx)dt\right)^{1/p} < \infty.\end{aligned}$$

In the case $0 < p < 1$, $\|\cdot\|_{p,T}$ does not satisfy the triangle inequality. However, by subadditivity, $\|X+Y\|^p_{p,T} \leq \|X\|^p_{p,T} + \|Y\|^p_{p,T}$, so that $\|\cdot\|^p_{p,T}$ is an F-norm and $\|X-Y\|^p_{p,T}$ defines a metric on L^p_T.

Theorem 3.2.2. *For $p > 0$, $\mathbb{S}$ is dense in L^p_T.*

If one restricts to predictable integrands, the proof of this theorem simplifies dramatically by observing that sets of the form $(s,t]\times A$, where $s < t$ and $A \in \mathbb{B}\left(\mathbb{R}^d\right)$, generate the predictable sigma algebra. More work is required when this assumption is not imposed. We will need the following lemma in the proof of Theorem 3.2.2 showing the Lebesgue integral in time of adapted random processes is adapted.

Lemma 3.2.3. *Let $\mathcal{A}$ be a sigma-subalgebra of $\mathcal{F}$. Suppose that $\xi : (\Omega \times \mathbb{R}_+, \mathcal{A} \otimes \mathcal{B}(\mathbb{R}_+)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is measurable and bounded and $\xi(\cdot, t)$ is $\mathcal{A}$-measurable for every $t \in [a, b)$. Then $\int_a^b \xi(\cdot, t)dt$ is also $\mathcal{A}$-measurable.*

Proof. Let $(\Omega', \mathcal{F}', \mathbb{P}')$ be another probability space and $U_n : \Omega' \to [a, b)$ be uniform i.i.d. random variables. Define a sequence of random variables on $(\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}')$ by $X_n(\omega, \omega') := \xi(\omega, U_n(\omega'))$. For each $\omega \in \Omega$, $X_n(\omega, \cdot)$ is an i.i.d. sequence of random variables on Ω' and by the Strong Law of Large Numbers,

$$\frac{\xi(\omega, U_1) + \cdots + \xi(\omega, U_n)}{n} \overset{\mathbb{P}'-a.s.}{\longrightarrow} \mathbb{E}_{\mathbb{P}'}\xi(\omega, U_1) = \frac{1}{b-a}\int_a^b \xi(\omega, t)dt. \tag{3.8}$$

Let

$$A \overset{def}{=} \left\{(\omega, \omega') : \frac{\xi(\omega, U_1(\omega')) + \cdots + \xi(\omega, U_n(\omega'))}{n} \to \frac{1}{b-a}\int_a^b \xi(\omega, t)dt\right\},$$
$$A_{\omega'} \overset{def}{=} \left\{\omega : \frac{\xi(\omega, U_1(\omega')) + \cdots + \xi(\omega, U_n(\omega'))}{n} \to \frac{1}{b-a}\int_a^b \xi(\omega, t)dt\right\},$$

and

$$A_{\omega} \overset{def}{=} \left\{\omega' : \frac{\xi(\omega, U_1(\omega')) + \cdots + \xi(\omega, U_n(\omega'))}{n} \to \frac{1}{b-a}\int_a^b \xi(\omega, t)dt\right\}.$$

Then we have $A = \bigcup_{\omega' \in \Omega'} (A_{\omega'} \times \{\omega'\}) = \bigcup_{\omega \in \Omega} (\{\omega\} \times A_\omega)$ and by Fubini's theorem and (3.8),

$$\begin{aligned}
1 \geq (\mathbb{P} \otimes \mathbb{P}')(A) &= \int_{\Omega \times \Omega'} \mathbb{1}_A(\omega, \omega')\mathbb{P} \otimes \mathbb{P}'(d\omega \times d\omega') \\
&= \int_\Omega \int_{\Omega'} \mathbb{1}_{\left\{\bigcup_{x \in \Omega} (\{x\} \times A_x)\right\}}(\omega, \omega')\mathbb{P}'(d\omega')\mathbb{P}(d\omega) \\
&= \int_\Omega \int_{\Omega'} \mathbb{1}_{A_\omega}(\omega')\mathbb{P}'(d\omega')\mathbb{P}(d\omega) \\
&= \int_\Omega \mathbb{P}'(A_\omega)\mathbb{P}(d\omega) = \int_\Omega \mathbb{P}(d\omega) = 1.
\end{aligned}$$

Therefore $\mathbb{P}\otimes\mathbb{P}'(A)=1$. But again by Fubini's theorem,

$$\begin{aligned}1=\mathbb{P}\otimes\mathbb{P}'(A)&=\int_{\Omega\times\Omega'}\mathbb{1}_A(\omega,\omega')\mathbb{P}\otimes\mathbb{P}'(d\omega\times d\omega')\\&=\int_{\Omega'}\int_{\Omega}\mathbb{1}_{\left\{\bigcup_{x\in\Omega'}(A_x\times\{x\})\right\}}(\omega,\omega')\mathbb{P}(d\omega)\mathbb{P}'(d\omega')\\&=\int_{\Omega'}\int_{\Omega}\mathbb{1}_{A_{\omega'}}(\omega)\mathbb{P}(d\omega)\mathbb{P}'(d\omega')=\int_{\Omega'}\mathbb{P}(A_{\omega'})\mathbb{P}'(d\omega').\end{aligned}$$

Therefore $\int_{\Omega'}(1-\mathbb{P}(A_{\omega'}))\,\mathbb{P}'(d\omega')=0$. Since $1-\mathbb{P}(A_{\omega'})\geq 0$, $\mathbb{P}(A_{\omega'})=1$ $\mathbb{P}'$-a.s. and hence, there exists an $\Omega_0'\subset\Omega'$ such that $\mathbb{P}'(\Omega_0')=1$ and for each $\omega'\in\Omega_0'$, $\mathbb{P}(A_{\omega'})=1$. Fix $\omega_0'\in\Omega_0'$ and put $t_n\stackrel{def}{=}U_n(\omega_0')$. Then

$$\frac{\xi(\omega,t_1)+\cdots+\xi(\omega,t_n)}{n}\stackrel{\mathbb{P}-a.s.}{\longrightarrow}\frac{1}{b-a}\int_a^b\xi(\omega,t)dt.$$

Since $t_i\in[a,b)$, $\xi(\cdot,t_i)$ is $\mathcal{A}$-measurable for each i and therefore the limit $\frac{1}{b-a}\int_a^b\xi(\cdot,t)dt$ is $\mathcal{A}$-measurable. In particular, $\int_a^b\xi(\cdot,t)dt$ is $\mathcal{A}$-measurable. □

To complete the proof of Theorem 3.2.2, we follow the outline of Bensoussan and Lions (1984, pgs. 261 - 262).

Proof of Theorem 3.2.2. Let $X\in L^p_T$. First, assume that X is bounded and vanishes off of $(S_1,S_2]\times B\subset[0,T]\times\mathbb{R}^d$ with $m\left((S_1,S_2]\times B\right)<\infty$. Set

$$\phi_j^n(x,\omega)\stackrel{def}{=}\frac{n}{S_2-S_1}\int_{S_1+\frac{(j-1)(S_2-S_1)}{n}}^{S_1+\frac{j(S_2-S_1)}{n}}X(s,x,\omega)ds$$

and

$$\phi_n(t,x,\omega)\stackrel{def}{=}\sum_{j=1}^{n-1}\phi_j^n(x,\omega)\mathbb{1}_{\left(S_1+\frac{j(S_2-S_1)}{n},S_1+\frac{(j+1)(S_2-S_1)}{n}\right]}(t).$$

Since X is nonanticipating, $X(s,x,\cdot)$ is $\mathcal{F}_{S_1+\frac{j(S_2-S_1)}{n}}$-measurable for every $s\in\left[S_1+\frac{(j-1)(S_2-S_1)}{n},S_1+\frac{j(S_2-S_1)}{n}\right)$. Using Lemma 3.2.3, we have that $\phi_j^n(x,\omega)$ is $\mathcal{F}_{S_1+\frac{j(S_2-S_1)}{n}}$-measurable, so that ϕ_n is nonanticipating. Letting $\alpha=\frac{1}{4}$ and $B(t,r_n)=$

$B\left(t, \frac{2(S_2-S_1)}{n}\right)$ in Theorem C.1.2, for every ω and for every x,

$$\lim_{n\to\infty} \phi_n(t,x,\omega) = X(t,x,\omega) \text{ for } \mathit{Leb}\text{-a.e. } t.$$

Since we have assumed X is bounded and $m\left((S_1,S_2]\times B\right) < \infty$, ϕ_n are uniformly bounded vanishing off of $(S_1,S_2]\times B$, and by the Dominated Convergence theorem,

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\int_0^T \int_{\mathbb{R}^d} |X(t,x)-\phi_n(t,x)|^p \,\rho(t,dx)dt \\
&\quad = \lim_{n\to\infty} \mathbb{E}\int_{S_1}^{S_2} \int_B |X(t,x)-\phi_n(t,x)|^p \,\rho(t,dx)dt \\
&\quad = \mathbb{E}\int_{S_1}^{S_2} \int_B \lim_{n\to\infty} |X(t,x)-\phi_n(t,x)|^p \,\rho(t,dx)dt = 0.
\end{aligned}$$

Let $\varepsilon > 0$. Then for large n,

$$\|X-\phi_n\|_{p,T} < \frac{\varepsilon}{2}.$$

Fix n large and for each j, consider the map $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to L^0\left(\Omega, \mathcal{F}_{S_1+\frac{j(S_2-S_1)}{n}}, \mathbb{P}\right)$ given by

$$x \mapsto (\omega \mapsto \phi_j^n(x,\omega)).$$

By standard analysis results, there exist disjoint $A_i^j \in \mathcal{B}(\mathbb{R}^d)$ and bounded $\mathcal{F}_{S_1+\frac{j(S_2-S_1)}{n}}$-measurable random variables φ_{ij} such that

$$\left|\phi_j^n(x,\omega) - \sum_{i=1}^{m(j)} \varphi_{ij}(\omega)\mathbb{1}_{A_i^j}(x)\right| < \frac{\varepsilon}{2m\left((S_1,S_2]\times B\right)}$$

for every $x\in\mathbb{R}^d$. Rewrite the collection of sets $\{A_i^j\}_{j=1}^{n-1}$ as a collection of m disjoint sets $\{A_k\}_{k=1}^m$ and set $\psi_{kj} = \varphi_{ij}$ if $A_k \subset A_i^j$. Then

$$\left|\phi_n(t,x,\omega) - \sum_{j=1}^{n-1}\sum_{k=1}^{m} \psi_{kj}(\omega)\mathbb{1}_{A_k}(x)\mathbb{1}_{\left(\frac{j(S_2-S_1)}{n},\frac{(j+1)(S_2-S_1)}{n}\right]}(t)\right| < \frac{\varepsilon}{2m\left((S_1,S_2]\times B\right)}$$

for each $(t,x,\omega) \in [0,T] \times \mathbb{R}^d \times \Omega$. If $p \geq 1$, Minkowski's inequality yields,

$$\begin{aligned}
&\left(\mathbb{E}\int_0^T \int_{\mathbb{R}^d} \left| X(t,x) - \sum_{j=1}^{n-1}\sum_{k=1}^{m} \psi_{kj} \mathbb{1}_{A_k}(x) \mathbb{1}_{\left(\frac{j(S_2-S_1)}{n}, \frac{(j+1)(S_2-S_1)}{n}\right]}(t) \right|^p \rho(t,dx)dt \right)^{1/p} \\
&\leq \left(\mathbb{E}\int_0^T \int_{\mathbb{R}^d} |X(t,x) - \phi_n(t,x)|^p \rho(t,dx)dt\right)^{1/p} \\
&+ \left(\mathbb{E}\int_0^T \int_{\mathbb{R}^d} \left| \phi_n(t,x) - \sum_{j=1}^{n-1}\sum_{k=1}^{m} \psi_{kj} \mathbb{1}_{A_k}(x) \mathbb{1}_{\left(\frac{j(S_2-S_1)}{n}, \frac{(j+1)(S_2-S_1)}{n}\right]}(t) \right|^p \rho(t,dx)dt \right)^{1/p} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2m\left((S_1,S_2] \times B\right)} m\left((S_1,S_2] \times B\right) \\
&= \varepsilon,
\end{aligned}$$

showing that X is in the closure of $\mathbb{S}$ under the $L^p([0,T] \times \mathbb{R}^d \times \Omega)$-norm. If $0 < p < 1$,

$$\begin{aligned}
&\mathbb{E}\int_0^T \int_{\mathbb{R}^d} \left| X(t,x) - \sum_{j=1}^{n-1}\sum_{k=1}^{m} \psi_{kj} \mathbb{1}_{A_k}(x) \mathbb{1}_{\left(\frac{j(S_2-S_1)}{n}, \frac{(j+1)(S_2-S_1)}{n}\right]}(t) \right|^p \rho(t,dx)dt \\
&\leq \mathbb{E}\int_0^T \int_{\mathbb{R}^d} |X(t,x) - \phi_n(t,x)|^p \rho(t,dx)dt \\
&+ \mathbb{E}\int_0^T \int_{\mathbb{R}^d} \left| \phi_n(t,x) - \sum_{j=1}^{n-1}\sum_{k=1}^{m} \psi_{kj} \mathbb{1}_{A_k}(x) \mathbb{1}_{\left(\frac{j(S_2-S_1)}{n}, \frac{(j+1)(S_2-S_1)}{n}\right]}(t) \right|^p \rho(t,dx)dt \\
&< \left(\frac{\varepsilon}{2}\right)^p + \left(\frac{\varepsilon}{2m\left((S_1,S_2] \times B\right)} m\left((S_1,S_2] \times B\right)\right)^p \\
&= 2\left(\frac{\varepsilon}{2}\right)^p,
\end{aligned}$$

showing that X is in the closure of $\mathbb{S}$ under the metric induced by norm $\|\cdot\|_{p,T}^p$.

Finally, let $\varepsilon > 0$, let $C \in \mathcal{B}\left([0,T] \times \mathbb{R}^d\right) \otimes \mathcal{F}$ with $\mathbb{P} \otimes m\,(C) < \infty$, and let X be an arbitrary random field in L_T^p. By Corollary C.2.2, there exists

$$\bigcup_{i=1}^{p} \left((s_i, s_{i+1}] \times B_i \times \gamma_i\right)$$

such that

$$\mathbb{P} \otimes m\left(C \triangle \bigcup_{i=1}^{p} \left((s_i, s_{i+1}] \times B_i \times \gamma_i\right)\right) < \frac{\varepsilon}{2}.$$

By definition, there exists an n so large that

$$\mathbb{P} \otimes m\left(|X| > n\right) < \frac{\varepsilon}{2}.$$

Define

$$\begin{aligned} X_n(t,x,\omega) &\stackrel{def}{=} X\mathbb{1}_{\{|X|\le n\}}(t,x,\omega)\mathbb{1}_{\left\{\bigcup_{i=1}^{p}(s_i,s_{i+1}]\times B_i\times\gamma_i)\right\}}(t,x,\omega) \\ &= \sum_{i=1}^{p} X(t,x,\omega)\mathbb{1}_{|X|\le n}(\omega)\mathbb{1}_{(s_i,s_{i+1}]}(t)\mathbb{1}_{B_i}(x)\mathbb{1}_{\gamma_i}(\omega). \end{aligned}$$

Then by the above, each term in the finite sum of X_n is in the closure of $\mathbb{S}$ and hence, so is X_n. Restricted to C,

$$\begin{aligned} \mathbb{P}\otimes m\left(|X - X_n| > \varepsilon\right) &= \mathbb{P}\otimes m\left(\{|X|>n\}\bigcup\left(C\setminus\bigcup_{i=1}^{p}((s_i,s_{i+1}]\times B_i\times\gamma_i)\right)\right) \\ &\le \mathbb{P}\otimes m\left(|X|>n\right) + \mathbb{P}\otimes m\left(C\triangle\bigcup_{i=1}^{p}((s_i,s_{i+1}]\times B_i\times\gamma_i)\right) \\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

So X_n converges locally in $\mathbb{P}\otimes m$-measure. Since $\mathbb{P}\otimes m$ is a σ-finite measure, there exists a subsequence n_k such that $X_{n_k}\to X$ $\mathbb{P}\otimes m$-a.e. Since $|X_n - X| \le 2X \in L_T^p$, the dominated convergence theorem gives

$$\lim_{k\to\infty}\|X_{n_k} - X\|_{p,T} = 0.$$

Therefore X is in the closure of $\mathbb{S}$ under the L_T^p norm. □

Finally, we denote by $L^p_{T,a.s.}$ the collection of all random fields $X\in L_T^0$ such that

$$\int_0^T\int_{\mathbb{R}^d}|X(t,x,\omega)|^p\,\rho(t,dx)dt < \infty \text{ a.s.}$$

Theorem 3.2.4. *For each $X\in L_T^0$, there exists $f_n, g_n\in\mathbb{S}$ such that*

i. *$f_n \to f$ locally in $\mathbb{P}\otimes m$-measure.*

ii. *$g_n\to f$ $\mathbb{P}\otimes m$-a.e.*

Again, this theorem follows from standard analysis results if one restricts to predictable integrands.

Proof. Let $X \in L^0_T$. By decomposing X as $X = X_+ - X_-$, we may assume that $X \geq 0$. Define a sequence $\{X_n\} \in L^0_T$ by $X_n \stackrel{def}{=} X \wedge n$. Then, for each n, $X_n \in L^1_T$ and by Theorem 3.2.2, there exists simple random fields $f^n_k \in \mathbb{S}$ such that $f^n_k \to X_n$ in L^1_T as $k \to \infty$. Therefore $f^n_k \to X_n$ locally in $\mathbb{P} \otimes m$-measure as $k \to \infty$. Let $\varepsilon, \delta > 0$ and let $C \subset [0,T] \times \mathbb{R}^d \times \Omega$ be such that $\mathbb{P} \otimes m\,(C) < \infty$. For each n, choose k_n large so that, when restricted to C,

$$\mathbb{P} \otimes m\left(\left|f^n_{k_n} - X_n\right| \geq \frac{\varepsilon}{2}\right) < \frac{\delta}{2}$$

and choose n large so that, when restricted to C,

$$\mathbb{P} \otimes m\,(X > n) < \frac{\delta}{2}.$$

For such n, when restricted to C,

$$\begin{aligned}
\mathbb{P} \otimes m\left(\left|X - f^n_{k_n}\right| \geq \varepsilon\right) &\leq \mathbb{P} \otimes m\left(|X - X_n| + \left|X_n - f^n_{k_n}\right| \geq \varepsilon\right) \\
&\leq \mathbb{P} \otimes m\left(|X - X_n| \geq \frac{\varepsilon}{2}\right) + \mathbb{P} \otimes m\left(\left|X_n - f^n_{k_n}\right| \geq \frac{\varepsilon}{2}\right) \\
&\leq \mathbb{P} \otimes m\left(|X - X_n| \geq \frac{\varepsilon}{2}\right) + \frac{\delta}{2} \\
&\leq \mathbb{P} \otimes m\,(X > n) + \frac{\delta}{2} \\
&\leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.
\end{aligned}$$

Since C was arbitrary, we have shown that $f^n_{k_n}$ converges locally in $\mathbb{P} \otimes m$-measure to X. Since $\mathbb{P} \otimes m$ is a σ-finite measure, there exists a subsequence $\left\{f^{n_j}_{k_{n_j}}\right\}$ converging to X $\mathbb{P} \otimes m$-a.e. $\square$

3.2.2 The stochastic integral driven by random measures

For a simple random field $f \in \mathbb{S}$, we define the integral of f with respect Z by

$$\int_0^T \int_{\mathbb{R}^d} f(t,x,\omega) Z(dt,dx;\omega) \stackrel{def}{=} \sum_{i=1}^m \sum_{j=1}^n f_{ij}(\omega) Z((t_i,t_{i+1}] \times A_j;\omega).$$

Let

$$Z_{ij}(\omega) \overset{def}{=} Z((t_i, t_{i+1}] \times A_j; \omega).$$

We will often simply write

$$\int_0^T \int_{\mathbb{R}^d} fZ(dt,dx) = \int_0^T \int_{\mathbb{R}^d} f(t,x)Z(dt,dx) = \sum_{i=1}^m \sum_{j=1}^n f_{ij} Z_{ij}.$$

For each $f \in \mathbb{S}$ we then define

$$\|f\|_Z \overset{def}{=} \sup_{\phi \in \mathbb{S}_1} \left\| \int_0^T \int_{\mathbb{R}^d} \phi f Z(dt,dx) \right\|_0,$$

where $\mathbb{S}_1 \subset \mathbb{S}$ is the set of simple random fields bounded by 1 and $\|\cdot\|_0$ is an F-norm on the space of random variables defined in example B.5. $\|\cdot\|_Z$ is an F-norm and modular on $\mathbb{S}$ (see Appendix B for definitions). We can extend the definition of stochastic integration in the usual way:

Definition 3.2.5. An adapted random field f is said to be Z*-integrable* if there exists a sequence of adapted simple random fields $\{f_n\} \in \mathbb{S}$ such that

i. $f_n \to f$ $\mathbb{P} \otimes m$-a.e.

ii. $\{f_n\}$ is a Cauchy sequence with respect to $\|\cdot\|_Z$.

For such adapted random fields f, we then define the stochastic integral of f with respect to Z by

$$\int_0^T \int_{\mathbb{R}^d} fZ(dt,dx) \overset{def}{=} \mathbb{P} - \lim_{n\to\infty} \int_0^T \int_{\mathbb{R}^d} f_n Z(dt,dx).$$

We next show that this definition is well-defined and identify the space of Z-integrable adapted random fields.

Define a function $\Phi_0 : [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}_+$ by

$$\Phi_0(t,x,u) \overset{def}{=} |U(t,x,u)| + \sigma^2(t,x)u^2 + \int_{\mathbb{R}} [\![uz]\!]^2 \nu(t,x,dz), \tag{3.9}$$

where

$$U(t,x,u) \overset{def}{=} b(t,x)u + \int_{\mathbb{R}} \left([\![uz]\!] - u[\![z]\!]\right) \nu(t,x,dz)$$

and $u \in \mathbb{R}$. In the next lemma, we will see that Φ_0 gives us a means to control convergence of $\|\cdot\|_Z$. The proof extends results of Rajput and Rosiński (1989) by utilizing a technique known as decoupling, for which the integrand and integrator can be treated independently. Let Z' be a copy of Z defined on a probability space Ω', independent of Z and $\mathcal{F}$. Define an enlarged σ-algebra on the product space $\Omega \times \Omega'$ by

$$\hat{\mathcal{F}}_t \stackrel{def}{=} \sigma\left(\mathcal{F}_t, Z'(C) : C \subset [0,t] \times \mathbb{R}^d, C \in \mathcal{B}_0([0,T] \times \mathbb{R}^d)\right).$$

For an $\mathcal{F}_t$-adapted simple random field $g(t,x,\omega) = \sum_{i=1}^m \sum_{j=1}^q g_{ij}(\omega)\mathbb{1}_{(t_i,t_{i+1}]\times A_j}(t,x)$, define two sequences of $\hat{\mathcal{F}}_{t_i}$-adapted random variables by

$$X_i(\omega,\omega') = X_i(\omega) \stackrel{def}{=} \sum_{j=1}^q g_{ij}(\omega) Z((t_i,t_{i+1}] \times A_j;\omega)$$

and

$$Y_i(\omega,\omega') \stackrel{def}{=} \sum_{j=1}^q g_{ij}(\omega) Z'((t_i,t_{i+1}] \times A_j;\omega').$$

Observe the following:

i. For each $\omega \in \Omega$, $\{Y_i(\omega,\cdot)\}_{i=1}^m$ is a sequence of independent random variables on Ω'.

ii. X_i and Y_i are $\hat{\mathcal{F}}_{t_i}$-adapted with

$$\mathbb{P}\otimes\mathbb{P}'\left(X_i \in A | \hat{\mathcal{F}}_{t_i}\right) = \mathbb{P}\left(X_i \in A | \mathcal{F}_{t_i}\right) = \mathbb{P}\left(\sum_{j=1}^q g_{ij}(\omega) Z((t_i,t_{i+1}] \times A_j) \in A | \mathcal{F}_{t_i}\right)$$

and

$$\begin{aligned}\mathbb{P}\otimes\mathbb{P}'\left(Y_i \in A | \hat{\mathcal{F}}_{t_i}\right) &= \mathbb{P}\otimes\mathbb{P}'\left(\sum_{j=1}^q g_{ij}(\omega) Z'((t_i,t_{i+1}] \times A_j) \in A | \mathcal{F}_{t_i}\right) \\ &= \mathbb{P}\left(\sum_{j=1}^q g_{ij}(\omega) Z((t_i,t_{i+1}] \times A_j) \in A | \mathcal{F}_{t_i}\right),\end{aligned}$$

since both $Z'((t_i,t_{i+1}] \times A_j)$ and $Z((t_i,t_{i+1}] \times A_j)$ are independent of $\hat{\mathcal{F}}_{t_i}$ for every j.

Sequences that satisfy (i) and (ii) are said to be $\hat{\mathcal{F}}_{t_i}$*-tangent sequences*. Moreover, observe that

$$\mathbb{P} \otimes \mathbb{P}' \left(Y_i \in A | \hat{\mathcal{F}}_{t_i} \right) = \mathbb{P} \otimes \mathbb{P}' \left(Y_i \in A | \mathcal{F} \right) = \mathbb{P}' \left(Y_i \in A | \mathcal{F} \right)$$

and, conditioned on $\mathcal{F}$, $\{Y_i\}_{i=1}^m$ is a sequence of independent random variables on Ω'. The sequence $\{Y_i\}_{i=1}^m$ is said to satisfy *conditional independence*. The sequence $\{Y_i\}_{i=1}^m$ satisfying (i), (ii), and conditional independence is said to be a *decoupled* $\hat{\mathcal{F}}_{t_i}$*-tangent sequence* to $\{X_i\}_{i=1}^m$. Kwapień and Woyczyński (1992, Section 5.7) showed the following:

Theorem 3.2.6. *Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_m$ be two $\mathcal{F}_i$-tangent sequences of random variables. If $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous, nondecreasing function of moderate growth (see Definition C.1.4 and Theorem C.1.5), then there exists a constant K_1, depending only on φ, such that*

$$\mathbb{E}\varphi \left(\left| \sum_{i=1}^m X_i \right| \right) \leq K_1 \max_{\varepsilon_i = \pm 1} \mathbb{E}\varphi \left(\left| \sum_{i=1}^m \varepsilon_i Y_i \right| \right)$$

and

$$\mathbb{E}\varphi \left(\max_{1 \leq k \leq m} \left| \sum_{i=1}^k X_i \right| \right) \leq K_1 \max_{\varepsilon_i = \pm 1} \mathbb{E}\varphi \left(\max_{1 \leq k \leq m} \left| \sum_{i=1}^k \varepsilon_i Y_i \right| \right).$$

Moreover, if $Y_1, \ldots, Y_m$ satisfy property conditional independence, then there exists a constant K_2, depending only on φ, such that

$$\mathbb{E}\varphi \left(\left| \sum_{i=1}^m X_i \right| \right) \leq K_2 \mathbb{E}\varphi \left(\left| \sum_{i=1}^m Y_i \right| \right)$$

and

$$\mathbb{E}\varphi \left(\max_{1 \leq k \leq m} \left| \sum_{i=1}^k X_i \right| \right) \leq K_2 \mathbb{E}\varphi \left(\max_{1 \leq k \leq m} \left| \sum_{i=1}^k Y_i \right| \right).$$

Next, we use these decoupled tangent sequences and the above theorem to extend the deterministic integrand results of Rajput and Rosiński (1989) to $f \in L_T^0$.

Lemma 3.2.7. *Let $\{f_n\}_{n\geq 0}$ be a sequence of adapted simple random fields. The following conditions are equivalent:*

i. $\|f_n\|_Z = \sup_{\phi\in\mathbb{S}_1}\left\|\int_0^T\int_{\mathbb{R}^d}\phi(t,x)f_n(t,x)Z(dt,dx)\right\|_0 \to 0$

ii. $\|f_n\|_{Z'} = \sup_{\phi\in\mathbb{S}_1}\left\|\int_0^T\int_{\mathbb{R}^d}\phi(t,x)f_n(t,x)Z'(dt,dx)\right\|_0 \to 0$

iii. $\int_0^T\int_{\mathbb{R}^d}\Phi_0(t,x,f_n(t,x,\cdot))m(dt,dx) \xrightarrow{\mathbb{P}} 0.$

Proof. Let f be an adapted simple random field. For any adapted random field $\phi\in\mathbb{S}_1$, write $(\phi f)(t,x,\omega)$ as $\sum_{i=1}^m\sum_{j=1}^q g_{ij}(\omega)1_{(s_i,s_{i+1}]}(t)1_{B_j}(x)$. Define a function $\varphi:\mathbb{R}_+\to\mathbb{R}_+$ by

$$\varphi(x) = x\wedge 1.$$

Note that φ satisfies the Δ_2 condition (see Definition C.1.4) since

$$\varphi(2x) = (2x)\wedge 1 \leq (2x)\wedge 2 = 2(x\wedge 1) = 2\varphi(x).$$

By Theorem C.1.5, φ is of moderate growth (see Definition C.1.4 Item ii) and hence, by Theorem 3.2.6 applied to $X_i \stackrel{def}{=} \sum_{j=1}^q g_{ij}Z_{ij}$ and $Y_i \stackrel{def}{=} \sum_{j=1}^q g_{ij}Z'_{ij}$, there exists a constant C, depending only on φ, such that

$$\begin{aligned} E\left(\left|\int_0^T\int_{\mathbb{R}^d}(\phi f)(t,x)Z(dt,dx)\right|\wedge 1\right) &= E\left(\left|\sum_{i=1}^m\sum_{j=1}^q g_{ij}Z\left((s_i,s_{i+1}]\times B_j\right)\right|\wedge 1\right)\\ &\leq C\max_{\varepsilon_i=\pm 1}E\left(\left|\sum_{i=1}^m\sum_{j=1}^q\varepsilon_i g_{ij}Z'\left((s_i,s_{i+1}]\times B_j\right)\right|\wedge 1\right)\\ &\leq C\|f\|_{Z'}.\end{aligned}$$

Since ϕ was arbitrary,

$$\|f\|_Z \leq C\|f\|_{Z'}.$$

The same argument holds when the roles of Z and Z' are reversed. So there exists constants C_1 and C_2 such that

$$C_1\|f\|_{Z'} \leq \|f\|_Z \leq C_2\|f\|_{Z'}.$$

This shows the equivalence of Item i and Item ii.

To show Item iii ⇒ Item ii, assume that $\int\int \Phi_0(t,x,f_n(t,x,\cdot))m(dt,dx) \xrightarrow{\mathbb{P}} 0$ and let $\{n_k\}_{k\in\mathbb{N}}$ be a sequence of natural numbers. Then

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0(t,x,f_{n_k}(t,x,\cdot))m(dt,dx) \xrightarrow{\mathbb{P}} 0$$

and hence, there exists a subsequence $\{n_{k_l}\} \subset \{n_k\}$ and an $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t,x,f_{n_{k_l}}(t,x,\omega)\right) m(dt,dx) \to 0$$

for every $\omega \in \Omega_0$. Let $\phi \in \mathbb{S}_1$. By Rosiński (2007a, Lemma 2.1.5),

$$\Phi_0\left(t,x,\phi(t,x,\omega)f_{n_{k_l}}(t,x,\omega)\right) \le 2\Phi_0\left(t,x,f_{n_{k_l}}(t,x,\omega)\right)$$

for each $(t,x,\omega) \in \mathbb{R}_+ \times \mathbb{R}^d \times \Omega$ and hence,

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t,x,\phi(t,x,\omega)f_{n_{k_l}}(t,x,\omega)\right) m(dt,dx)&\\ \le \int_0^T \int_{\mathbb{R}^d} 2\Phi_0\left(t,x,f_{n_{k_l}}(t,x,\omega)\right) m(dt,dx) &\to 0\end{aligned}$$

for every $\omega \in \Omega_0$. That is, for every $\omega \in \Omega_0$,

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} \left|U\left(t,x,\phi(t,x,\omega)f_{n_{k_l}}(t,x,\omega)\right)\right| m(dt,dx)&\\ + \int_0^T \int_{\mathbb{R}^d} \sigma^2(t,x)\phi^2(t,x,\omega)f^2_{n_{k_l}}(t,x,\omega)\, m(dt,dx)&\\ + \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}} [\![\phi(t,x,\omega)f_{n_{k_l}}(t,x,\omega)z]\!]^2 \nu(t,x,dz)\, m(dt,dx) &\to 0.\end{aligned}$$

Since each integrand is positive,

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} \left|U\left(t,x,\phi(t,x,\omega)f_{n_{k_l}}(t,x,\omega)\right)\right| m(dt,dx) &\to 0,\\ \int_0^T \int_{\mathbb{R}^d} \sigma^2(t,x)\phi^2(t,x,\omega)f^2_{n_{k_l}}(t,x,\omega)\, m(dt,dx) &\to 0,\end{aligned}$$

and

$$\int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}} [\![\phi(t,x,\omega) f_{n_{k_l}}(t,x,\omega) z]\!]^2 \nu(t,x,dz)\, m(dt,dx) \to 0$$

for every $\omega \in \Omega_0$. Define a new measure on $\mathcal{B}(\mathbb{R})$ by

$$\begin{aligned} F_g(A) &\overset{def}{=} F\left\{(t,x,z) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} : g(t,x)z \in A \setminus \{0\}\right\} \\ &= \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}} 1_{A/g(t,x)}(z) \nu(t,x,dz) m(dt,dx) \\ &= \int_0^T \int_{\mathbb{R}^d} \nu\left(t,x,A/g(t,x)\right) m(dt,dx). \end{aligned}$$

Then for every $\omega \in \Omega_0$,

$$\int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}} [\![\phi(t,x,\omega) f_{n_{k_l}}(t,x,\omega) z]\!]^2 \nu(t,x,dz)\, m(dt,dx) = \int_{\mathbb{R}} [\![z]\!]^2 F_{\phi f_{n_{k_l}}}(dz) \to 0.$$

Fix $\omega \in \Omega_0$. The the function $\left(\phi f_{n_{k_l}}\right)(\cdot,\cdot,\omega)$ is measurable and Z'-integrable (as a deterministic function, see Rajput and Rosiński (1989)) and the stochastic integral is infinitely divisible with Lévy-Khintchine triplet

$$\left(\int_0^T \int_{\mathbb{R}^d} U\left(t,x,\phi(t,x,\omega) f_{n_{k_l}}(t,x,\omega)\right) m(dt,dx), \right.$$
$$\left. \int_0^T \int_{\mathbb{R}^d} \sigma^2(t,x) \phi^2(t,x,\omega) f^2_{n_{k_l}}(t,x,\omega)\, m(dt,dx),\ F_{\phi f_{n_{k_l}}} \right).$$

By Rajput and Rosiński (1989, Lemma 3.2),

$$\mathcal{L}\left\{\int_0^T \int_{\mathbb{R}^d} \phi(t,x,\omega) f_{n_{k_l}}(t,x,\omega) Z'(dt,dx;\omega')\right\} \xrightarrow{w} \delta_0$$

and hence

$$\int_{\Omega} \left(\left|\int_0^T \int_{\mathbb{R}^d} \phi(t,x,\omega) f_{n_{k_l}}(t,x,\omega) Z'(dt,dx;\omega')\right| \wedge 1\right) \mathbb{P}(d\omega') \to 0.$$

Since $\mathbb{P}(\Omega_0) = 1$, the dominated convergence theorem implies

$$\int_\Omega \int_\Omega \left(\left| \int_0^T \int_{\mathbb{R}^d} \phi(t,x,\omega) f_{n_{k_l}}(t,x,\omega) Z'(dt,dx;\omega') \right| \wedge 1 \right) \mathbb{P}(d\omega')\mathbb{P}(d\omega) \to 0$$

for every $\phi \in \mathbb{S}_1$. For each l, choose $\phi_l \in \mathbb{S}_1$ such that

$$\left\| f_{n_{k_l}} \right\|_{Z'} < \int_\Omega \int_\Omega \left(\left| \int_0^T \int_{\mathbb{R}^d} \phi_l(t,x,\omega) f_{n_{k_l}}(t,x,\omega) Z'(dt,dx;\omega') \right| \wedge 1 \right) \mathbb{P}(d\omega')\mathbb{P}(d\omega) + \frac{1}{l}.$$

Repeating the above argument with ϕ_l replacing ϕ gives

$$\int_\Omega \int_\Omega \left(\left| \int_0^T \int_{\mathbb{R}^d} \phi_l(t,x,\omega) f_{n_{k_l}}(t,x,\omega) Z'(dt,dx;\omega') \right| \wedge 1 \right) \mathbb{P}(d\omega')\mathbb{P}(d\omega) \to 0$$

and hence,

$$\left\| f_{n_{k_l}} \right\|_{Z'} \to 0.$$

So every sequence $\{n_k\}$ has a subsequence $\{n_{k_l}\}$ such that

$$\left\| f_{n_{k_l}} \right\|_{Z'} \to 0$$

as $l \to \infty$. Therefore,

$$\|f_n\|_{Z'} \to 0.$$

Finally, to show Item ii $\Rightarrow$ Item iii, assume

$$\|f_n\|_{Z'} \to 0.$$

By definition,

$$\sup_{\phi \in \mathbb{S}_1} E \left(\left| \int_0^T \int_{\mathbb{R}^d} \phi(t,x,\omega) f_n(t,x,\omega) Z'(dt,dx;\omega) \right| \wedge 1 \right) \to 0.$$

For each n, define $\phi_n \in \mathbb{S}_1$ by

$$\phi_n(t,x,\omega) = \operatorname{sign} U(t,x,f_n(t,x,\omega)).$$

Then

$$\int_\Omega \int_\Omega \left(\left| \int_0^T \int_{\mathbb{R}^d} \phi_n(t,x,\omega) f_n(t,x,\omega) Z'(dt,dx;\omega') \right| \wedge 1 \right) \mathbb{P}(d\omega')\mathbb{P}(d\omega)$$
$$= E\left(\left| \int_0^T \int_{\mathbb{R}^d} \phi_n(t,x,\omega) f_n(t,x,\omega) Z'(dt,dx;\omega) \right| \wedge 1 \right) \to 0.$$

Therefore

$$\int_\Omega \left(\left| \int_0^T \int_{\mathbb{R}^d} \phi_n(t,x,\omega) f_n(t,x,\omega) Z'(dt,dx;\omega') \right| \wedge 1 \right) \mathbb{P}(d\omega') \xrightarrow{\mathbb{P}} 0$$

and there exists a subsequence $\{n_k\}_{k\in\mathbb{N}}$ and an $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that

$$\int_\Omega \left(\left| \int_0^T \int_{\mathbb{R}^d} \phi_{n_k}(t,x,\omega) f_{n_k}(t,x,\omega) Z'(dt,dx;\omega') \right| \wedge 1 \right) \mathbb{P}(d\omega') \to 0$$

for every $\omega \in \Omega_0$. Fix $\omega \in \Omega_0$. We have

$$\int_0^T \int_{\mathbb{R}^d} \phi_{n_k}(t,x,\omega) f_{n_k}(t,x,\omega) Z'(dt,dx;\omega') \xrightarrow{\mathbb{P}} 0,$$

so that $\mathcal{L}\left(\int_0^T \int_{\mathbb{R}^d} \phi_{n_k}(t,x,\omega) f_{n_k}(t,x,\omega) Z'(dt,dx) \right) \xrightarrow{w} \delta_0$. But $\phi_{n_k}(t,x,\omega) f_{n_k}(\cdot,\cdot,\omega)$ is measurable and Z'-integrable as a deterministic function. By Rajput and Rosiński (1989, Theorem 2.7), the stochastic integral is infinitely divisible with generating triplet

$$\left(\int_0^T \int_{\mathbb{R}^d} U\left(t,x,\phi_{n_k}(t,x,\omega) f_{n_k}(t,x,\omega)\right) m(dt,dx), \right.$$
$$\left. \int_0^T \int_{\mathbb{R}^d} \sigma^2(t,x) f_{n_k}^2(t,x,\omega)\, m(dt,dx), F_{\phi_{n_k} f_{n_k}} \right).$$

By Rajput and Rosiński (1989, Lemma 3.2),

$$\int_0^T \int_{\mathbb{R}^d} U\left(t,x,\phi_{n_k}(t,x,\omega) f_{n_k}(t,x,\omega)\right) m(dt,dx) \to 0,$$
$$\int_0^T \int_{\mathbb{R}^d} \sigma^2(t,x) f_{n_k}^2(t,x,\omega)\, m(dt,dx) \to 0,$$

and

$$\int_{\mathbb{R}} [\![z]\!]^2 F_{\phi_{n_k} f_{n_k}}(dz) \to 0.$$

Since $U(t,x,\cdot)$ is an odd function,

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t,x,f_{n_k}(t,x,\omega)\right) m(dt,dx) \\
&\quad = \int_0^T \int_{\mathbb{R}^d} \left(|U\left(t,x,f_{n_k}(t,x,\omega)\right)| + \sigma^2(t,x) f_{n_k}^2(t,x,\omega)\right) m(dt,dx) \\
&\qquad + \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}} [\![f_{n_k}(t,x,\omega) z]\!]^2 \nu(t,x,dz)\, m(dt,dx) \\
&\quad = \int_0^T \int_{\mathbb{R}^d} \left(U\left(t,x,\phi_{n_k}(t,x,\omega) f_{n_k}(t,x,\omega)\right) + \sigma^2(t,x) f_{n_k}^2(t,x,\omega)\right) m(dt,dx) \\
&\qquad + \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}} [\![\phi_{n_k}(t,x,\omega) f_{n_k}(t,x,\omega) z]\!]^2 \nu(t,x,dz)\, m(dt,dx) \to 0.
\end{aligned}$$

Since $\omega \in \Omega_0$ was arbitrary and $\mathbb{P}\left(\Omega_0\right) = 1$,

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t,x,f_{n_k}(t,x,\cdot)\right) m(dt,dx) \to 0 \text{ a.s.}$$

We have shown that if $\|f_n\|_{Z'} \to 0$, then there exists a subsequence $\{n_k\}$ such that

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t,x,f_{n_k}\left(t,x,\cdot\right)\right) m(dt,dx) \to 0 \text{ a.s.}$$

Let $\|f_n\|_{Z'} \to 0$ and $\{n_k\}_{k\in\mathbb{N}}$ a sequence of natural numbers. Then $\|f_{n_k}\|_{Z'} \to 0$ and by the above, there exists a subsequence $\{n_{k_l}\}_{l\in\mathbb{N}} \subset \{n_k\}_{k\in N}$ such that

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(f_{n_{k_l}}(t,x,\cdot)\right) m(dt,dx) \to 0 \text{ a.s.}$$

By Lemma C.1.3,

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t,x,f_n(t,x,\cdot)\right) m(dt,dx) \xrightarrow{\mathbb{P}} 0.$$

□

We now see the power of the decoupling approach. We may treat the integrand and integrator as independent. The next two theorems are the main results in the development of the stochastic integral. The first shows that the integral does not depend upon the approximating sequence of simple random fields. The second addresses which random fields are Z-integrable in terms of the control measure m.

Theorem 3.2.8. *Suppose that $f \in L_T^0$ is Z-integrable. Then the stochastic integral is well-defined.*

Proof. Let $f_n, g_n \in \mathbb{S}$ be simple random fields such that

i. $f_n \to f$ and $g_n \to f$ $\mathbb{P} \otimes m$-a.e.

ii. $\|f_n - f_m\|_Z \to 0$ and $\|g_n - g_m\|_Z \to 0$.

By Lemma 3.2.7,

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t, x, f_m(t,x,\omega) - f_n(t,x,\omega)\right) m(dt, dx) \to 0$$

and

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t, x, g_m(t,x,\omega) - g_n(t,x,\omega)\right) m(dt, dx) \to 0.$$

Since $\Phi_0 \geq 0$, Fatou's lemma gives

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t, x, f(t,x,\omega) - f_n(t,x,\omega)\right) m(dt, dx) \to 0$$

and

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t, x, f(t,x,\omega) - g_n(t,x,\omega)\right) m(dt, dx) \to 0.$$

By Rosiński (2007a, Lemma 2.1.5),

$$\begin{aligned}
\int_0^T \int_{\mathbb{R}^d} & \Phi_0(t, x, f_n(t,x,\omega) - g_n(t,x,\omega)) m(dt, dx) \\
& \leq 3 \int_0^T \int_{\mathbb{R}^d} \Phi_0(t, x, f_n(t,x,\omega) - f(t,x,\omega)) m(dt, dx) \\
& \quad + 3 \int_0^T \int_{\mathbb{R}^d} \Phi_0(t, x, f(t,x,\omega) - g_n(t,x,\omega)) m(dt, dx) \to 0.
\end{aligned}$$

By Lemma 3.2.7,

$$\|f_n - g_n\|_Z \to 0,$$

that is,

$$\sup_{\phi \in \mathbb{S}_1} \left\| \int_0^T \int_{\mathbb{R}^d} \phi \left(f_n - g_n\right) Z(dt, dx) \right\|_0 \to 0.$$

Take $\phi \equiv 1 \in \mathbb{S}_1$. Then

$$\left\| \int_0^T \int_{\mathbb{R}^d} \left(f_n - g_n\right) Z(dt, dx) \right\|_0 \to 0,$$

showing

$$\int_0^T \int_{\mathbb{R}^d} \left(f_n - g_n\right) Z(dt, dx) \xrightarrow{\mathbb{P}} 0.$$

□

Theorem 3.2.9. *For random field $f \in L_T^0$, the following are equivalent:*

i. *f is Z-integrable.*

ii. *f is Z'-integrable.*

iii. $\int_0^T \int_{\mathbb{R}^d} \Phi_0(t, x, f(t, x, \cdot)) m(dt, dx) < \infty$ *a.s.*

Proof. That Item i is equivalent to Item ii is immediate by Lemma 3.2.7 since $\{f_n\}_{n \geq 0}$ is a Cauchy sequence with respect to $\| \cdot \|_Z$ if and only if it is a Cauchy sequence with respect to $\| \cdot \|_{Z'}$.

Item i $\Rightarrow$ Item iii. Let f be an adapted Z-integrable random field. Then there exists a sequence of simple random fields $\{f_n\}$ such that

i. $f_n \to f$ $\mathbb{P} \otimes m$-a.e.

ii. $\|f_n - f_m\|_Z \to 0$ as $m, n \to \infty$.

By Lemma 3.2.7,

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0 \left(t, x, f_n \left(t, x, \cdot\right) - f_m \left(t, x, \cdot\right)\right) m(dt, dx) \xrightarrow{\mathbb{P}} 0.$$

By Rosiński (2007a, Lemma 2.1.5),

$$\begin{aligned} &\int_0^T \int_{\mathbb{R}^d} \Phi_0 \left(t, x, f \left(t, x, \cdot\right)\right) m(dt, dx) \\ &\quad = \int_0^T \int_{\mathbb{R}^d} \Phi_0 \left(t, x, f(t, x, \cdot) - f_n(t, x, \cdot) + f_n(t, x, \cdot)\right) m(dt, dx) \end{aligned}$$

$$\leq 3\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f(t,x,\cdot)-f_n(t,x,\cdot)\right)m(dt,dx) \qquad (3.10)$$
$$+\,3\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f_n(t,x,\cdot)\right)m(dt,dx).$$

Since $\Phi_0\geq 0$, by Fatou's lemma, for every $\omega\in\Omega_0$,

$$\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f(t,x,\cdot)-f_n(t,x,\cdot)\right)m(dt,dx)$$
$$\leq \liminf_{m\to\infty}\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f_m(t,x,\cdot)-f_n(t,x,\cdot)\right)m(dt,dx).$$

Therefore

$$\left\{\omega:\left|\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f(t,x,\omega)-f_n(t,x,\omega)\right)m(dt,dx)\right|>\varepsilon\right\}$$
$$=\left\{\omega:\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f(t,x,\omega)-f_n(t,x,\omega)\right)m(dt,dx)>\varepsilon\right\}$$
$$\subset\left\{\omega:\liminf_{m\to\infty}\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f_m(t,x,\omega)-f_n(t,x,\omega)\right)m(dt,dx)>\varepsilon\right\}$$

and hence,

$$\mathbb{P}\left(\left|\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f(t,x,\omega)-f_n(t,x,\omega)\right)m(dt,dx)\right|>\varepsilon\right)$$
$$\leq \mathbb{P}\left(\liminf_{m\to\infty}\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f_m(t,x,\omega)-f_n(t,x,\omega)\right)m(dt,dx)>\varepsilon\right)\to 0$$

as $n\to\infty$. So

$$\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f(t,x,\omega)-f_n(t,x,\omega)\right)m(dt,dx)\xrightarrow{\mathbb{P}}0$$

and as such, there exists a subsequence $\{n_k\}$ and an $\Omega_0\subset\Omega$ with $\mathbb{P}\left(\Omega_0\right)=1$ such that

$$\int_0^T\int_{\mathbb{R}^d}\Phi_0\left(t,x,f(t,x,\omega)-f_{n_k}(t,x,\omega)\right)m(dt,dx)\to 0$$

for every $\omega \in \Omega_0$. Fix $\omega \in \Omega_0$. By (3.10),

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t, x, f\left(t, x, \omega\right)\right) m(dt, dx) \\ \leq 3 \int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t, x, f(t, x, \omega) - f_{n_k}(t, x, \omega)\right) m(dt, dx) \\ + 3 \int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t, x, f_{n_k}(t, x, \omega)\right) m(dt, dx).\end{aligned}$$

Let $k \to \infty$. The first term converges to 0 and the second term is finite for each k since $f_{n_k} \in \mathbb{S}$. Since $\mathbb{P}\left(\Omega_0\right) = 1$, $\int_0^T \int_{\mathbb{R}^d} \Phi_0(t, x, f(t, x, \cdot)) m(dt, dx) < \infty$ a.s.

Item iii $\Rightarrow$ Item i. Let $f \in L^0$ and suppose

$$\int_0^T \int_{\mathbb{R}^d} \Phi_0\left(t, x, f(t, x, \cdot)\right) m(dt, dx) < \infty \text{ a.s.}$$

By Theorem 3.2.4, choose $f_n \in \mathbb{S}$ such that

$$f_n \to f \ \mathbb{P} \otimes m\text{-a.e.}$$

and $|f_n| \leq |f|$ for each $(t, x, \omega) \in [0, T] \times \mathbb{R}^d \times \Omega$. By Rosiński (2007a, Lemma 2.1.5),

$$\Phi_0(t, x, f_m(t, x, \cdot) - f_n(t, x, \cdot)) \leq 4\Phi_0(t, x, f(t, x, \cdot)) \in L^1\left([0, T] \times \mathbb{R}^d; m\right) \text{ a.s.}$$

By the Dominated Convergence theorem,

$$\lim_{m,n\to\infty} \int_0^T \int_{\mathbb{R}^d} \Phi_0(t, x, f_m(t, x, \cdot) - f_n(t, x, \cdot)) m(dt, dx) = 0 \text{ a.s.}$$

and hence, by Lemma 3.2.7,

$$\lim_{m,n\to\infty} \|f_m - f_n\|_Z = 0.$$

□

3.2.3 Examples

Symmetric α-stable random measure

Let Z be a symmetric α-stable random measure with Lévy-Khintchine triplet $Z(C) \sim (0, 0, m(C)\nu)$, where $0 < \alpha < 2$, $c > 0$, and $\nu(dz) = c\,|z|^{-\alpha-1}\,dz$. Since the Lévy measure ν does not depend upon t and x, $\Phi_0(t,x,u) = \Phi_0(u)$ and since ν is symmetric, $U(t,x,u) \equiv 0$. We have

$$
\begin{aligned}
\Phi_0(u) &= \int_{\mathbb{R}} [\![uz]\!]^2 \nu(dz) \\
&= 2c \int_0^\infty [\![uz]\!]^2 z^{-\alpha-1} dz \\
&= 2c\,|u|^2 \int_0^{1/|u|} z^{1-\alpha} dz + 2c \int_{1/|u|}^\infty z^{-\alpha-1} dz \\
&= \frac{2c}{2-\alpha}\,|u|^2 \left(\frac{1}{|u|}\right)^{2-\alpha} + \frac{2c}{\alpha}\left(\frac{1}{|u|}\right)^{-\alpha} \\
&= \frac{4c}{\alpha(2-\alpha)}\,|u|^\alpha .
\end{aligned}
$$

Theorem 3.2.9 gives

$$
\begin{aligned}
&\left\{f \in L_T^0 : f \text{ is } Z\text{-integrable}\right\} \\
&\qquad = \left\{ f \in L_T^0 : \int_0^T \int_{\mathbb{R}^d} |f(t,x)|^\alpha\, m(dt,dx) < \infty \text{ a.s.} \right\} = L^\alpha_{T,a.s.}.
\end{aligned}
$$

This is result is analogous to the space of integrable functions with respect to Brownian motion.

Symmetric tempered α-stable random measure

Let Z be a symmetric tempered α-stable random measure with Lévy-Khintchine triplet $Z(C) \sim (0, 0, m(C)\nu)$, where $0 < \alpha < 2$, $c > 0$, and $\nu(dz) = c\,|z|^{-\alpha-1}\,e^{-|z|}dz$. Since the Lévy measure ν does not depend upon t and x, $\Phi_0(t,x,u) = \Phi_0(u)$. To identify the space of Z-integrable random fields, we first show that

$$
\Phi_0(u) \simeq |u|^2 \wedge |u|^\alpha . \tag{3.11}
$$

Since $x \mapsto [\![x]\!]$ is an odd function,

$$\Phi_0(u) = \int [\![uz]\!]^2 \nu(dz) = 2\int_0^\infty [\![uz]\!]^2 c\,|z|^{-\alpha-1}\, e^{-|z|}dz$$
$$= 2c\,|u|^2 \int_0^{1/|u|} z^{1-\alpha} e^{-z} dz + 2c\int_{1/|u|}^\infty z^{-\alpha-1}e^{-z}dz. \quad (3.12)$$

To compute the lower bound of the equivalence (3.11),

$$\begin{aligned}
\Phi_0(u) &\geq 2c\,|u|^2 \int_0^{1/|u|} z^{1-\alpha}e^{-z}dz \\
&= 2cu^2 \left(\mathbb{1}_{\{|u|\leq 1\}} \int_0^{1/|u|} z^{1-\alpha}e^{-z}dz + \mathbb{1}_{\{|u|>1\}} \int_0^{1/|u|} z^{1-\alpha}e^{-z}dz \right) \\
&\geq 2cu^2 \left(\gamma(2-\alpha,1)\mathbb{1}_{\{|u|\leq 1\}} + \frac{e^{-1}}{2-\alpha} |u|^{\alpha-2}\, \mathbb{1}_{\{|u|>1\}} \right) \\
&= 2c\gamma(2-\alpha,1)u^2\mathbb{1}_{\{|u|\leq 1\}} + \frac{2ce^{-1}}{2-\alpha} |u|^{\alpha}\, \mathbb{1}_{\{|u|>1\}} \\
&\geq \left(2\gamma(2-\alpha,1) \wedge \frac{2ce^{-1}}{2-\alpha} \right) \left(|u|^2 \wedge |u|^\alpha \right).
\end{aligned}$$

For the upper bound of the equivalence (3.11), observe that the first integral in (3.12) is bounded above by

$$\begin{aligned}
2cu^2 \int_0^{1/|u|} z^{1-\alpha}e^{-z}dz &= 2cu^2\mathbb{1}_{|u|\leq 1} \int_0^{1/|u|} z^{1-\alpha}e^{-z}dz + 2cu^2\mathbb{1}_{|u|>1} \int_0^{1/|u|} z^{1-\alpha}e^{-z}dz \\
&\leq 2c\Gamma(2-\alpha)u^2\mathbb{1}_{\{|u|\leq 1\}} + \frac{2cu^2}{2-\alpha}\left(\frac{1}{|u|}\right)^{2-\alpha} \mathbb{1}_{\{|u|>1\}} \\
&= 2c\Gamma(2-\alpha)u^2\mathbb{1}_{\{|u|\leq 1\}} + \frac{2c}{2-\alpha}\,|u|^\alpha\, \mathbb{1}_{\{|u|>1\}} \quad (3.13)
\end{aligned}$$

and the second integral in (3.12) is bounded above by

$$\begin{aligned}
2c\int_{1/|u|}^\infty z^{-\alpha-1}e^{-z}dz &= 2c\mathbb{1}_{\{|u|\leq 1\}}\int_{1/|u|}^\infty z^{-\alpha-1}e^{-z}dz + 2c\mathbb{1}_{\{|u|>1\}}\int_{1/|u|}^\infty z^{-\alpha-1}e^{-z}dz \\
&\leq 2c \sup_{z\in[1,\infty)} \left(z^{2-\alpha}e^{-z}\right) \mathbb{1}_{\{|u|\leq 1\}} \int_{1/|u|}^\infty z^{-3}dz + 2c\mathbb{1}_{\{|u|>1\}}\int_{1/|u|}^\infty z^{-\alpha-1}dz \\
&= c \sup_{z\in[1,\infty)} \left(z^{2-\alpha}e^{-z}\right) \mathbb{1}_{\{|u|\leq 1\}}\, |u|^2 + 2c\mathbb{1}_{\{|u|>1\}}\alpha^{-1}\,|u|^\alpha. \quad (3.14)
\end{aligned}$$

Therefore, (3.13) and (3.14) combine to give

$$\Phi_0(u) \le 2c\Gamma(2-\alpha)\,|u|^2\,\mathbb{1}_{\{|u|\le 1\}} + \frac{2c}{2-\alpha}\,|u|^\alpha\,\mathbb{1}_{\{|u|>1\}} \\ + c\sup_{z\in[1,\infty)}\left(z^{2-\alpha}e^{-z}\right)|u|^2\,\mathbb{1}_{\{|u|\le 1\}} + 2c\alpha^{-1}\,|u|^\alpha\,\mathbb{1}_{\{|u|>1\}} \le C\left(|u|^2 \wedge |u|^\alpha\right),$$

where $C = \left(2c\Gamma(2-\alpha) + c\sup_{z\in[1,\infty)}\left(z^{2-\alpha}e^{-z}\right)\right) \vee \left(\frac{2c}{2-\alpha} + \frac{2c}{\alpha}\right)$. So the equivalence (3.11) holds and from this, along with Theorem 3.2.9,

$$\left\{f \in L_T^0 : f \text{ is } Z\text{-integrable}\right\} \\ = \left\{f \in L_T^0 : \int_0^T \int_{\mathbb{R}^d} |f(t,x)|^2 \wedge |f(t,x)|^\alpha\, m(dt,dx) < \infty \text{ a.s.}\right\}.$$

Notice that the space of symmetric tempered stable integrable random fields contains both the space of Brownian motion integrable random fields $L^2_{T,a.s.}$ and the space of symmetric α-stable integrable random fields $L^\alpha_{T,a.s.}$ from Section 3.2.3.

3.3 Itô isomorphisms

The well-known classic Ito Isometry gives

$$\mathbb{E}\left|\int_0^T X_t B(dt)\right|^2 = \mathbb{E}\int_0^T X_t^2 dt.$$

Hence,

$$\left\{X \in L_T^0 : X \text{ is } B\text{-integrable and } \mathbb{E}\left|\int_0^T X_t B(dt)\right|^2 < \infty\right\} \\ = L^2\left([0,T]\times\Omega; \mathcal{B}\left([0,T]\right)\otimes\mathcal{F}; \text{leb}\otimes\mathbb{P}\right) \quad (3.15)$$

and the map $X \mapsto \int_0^T X_t B(dt)$ is an isomorphism from $L^2\left([0,T]\times\Omega; \mathcal{B}\left([0,T]\right)\otimes\mathcal{F};\right.$ $\left.\text{leb}\otimes\mathbb{P}\right)$ into $L^2(\Omega;\mathcal{F};\mathbb{P})$. In this section, we use Theorem 1.2.1 to obtain Itô Isomorphisms for stochastic integrals driven by an infinitely divisible random measure Z, similar in nature to (3.15). We seek to determine when the stochastic integral $\int_0^T\int_{\mathbb{R}^d} f(t,x)Z(dt,dx)$ is in $L^p(\Omega,\mathcal{F},\mathbb{P})$ and when the map $f(t,x) \mapsto \int_0^T\int_{\mathbb{R}^d} f(t,x)Z(dt,dx)$ is an isomorphism into $L^p(\Omega,\mathcal{F},\mathbb{P})$. We will see that, just as in the case (3.15), the above condition is characterized by certain integrability conditions on f with respect to the control measure m of Z. We seek contraction inequalities of norms between the space

$$\left\{ f \in L_T^0 : f \text{ is } Z\text{-integrable and } \mathbb{E}\left| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right|^p < \infty \right\}$$

and an appropriate subspace of

$$L^0\left([0,T] \times \mathbb{R}^d \times \Omega; \mathcal{B}([0,T] \times \mathbb{R}^d) \otimes \mathcal{F}; m \otimes \mathbb{P}\right).$$

The appropriate subspace will be a Musielak-Orlicz modular space, a special type of modular space described in Appendix B. We begin by considering the case when f is deterministic and extend these results to the random case by decoupling arguments.

Let $p \geq 1$ and Z be an infinitely divisible random measure with Lévy-Khintchine triplet (3.2). Define the function $\Phi_p : [0,T] \times \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ by

$$\Phi_p(t,x,u) \stackrel{def}{=} U^*(t,x,u) + \sigma^2(t,x)u^2 + V_p(t,x,u), \tag{3.16}$$

where

$$\begin{aligned} U^*(t,x,u) &\stackrel{def}{=} \sup_{|c|\leq 1} |U(t,x,cu)|, \\ U(t,x,u) &\stackrel{def}{=} b(t,x)u + \int_{\mathbb{R}} \left([\![uz]\!] - u[\![z]\!] \right) \nu(t,x,dz), \end{aligned}$$

and

$$V_p(t,x,u) \stackrel{def}{=} \int_{\mathbb{R}} \left(|uz|^2 \mathbb{1}_{|uz|<1} + |uz|^p \mathbb{1}_{|uz|\geq 1} \right) \nu(t,x,dz).$$

Rajput and Rosiński (1989, Lemma 3.1) showed that Φ_p satisfies the properties given in Section B.5 and hence, generates a modular on the Musielak-Orlicz modular space of deterministic functions

$$\begin{aligned} L_{T,det}^{\Phi_p} &= L_{T,det}^{\Phi_p}\left([0,T] \times \mathbb{R}^d; \mathcal{B}\left([0,T] \times \mathbb{R}^d\right); m\right) \\ &\stackrel{def}{=} \left\{ f \in L_{T,det}^0 : \int_0^T \int_{\mathbb{R}^d} \Phi_p\left(t,x,|f(t,x)|\right) m(dt,dx) < \infty \right\}. \end{aligned} \tag{3.17}$$

Rajput and Rosiński (1989, Theorem 3.4) also showed that, under certain assumptions on the infinitely divisible random measure Z, the mapping of deterministic functions $f \to \int_0^T \int_{\mathbb{R}^d} f(t,x)Z(dt,dx)$ is an isomorphism from $L^{\Phi_p}_{T,det}$ into $L^p(\Omega; P)$. Assuming that this assumption holds, if $\|\cdot\|$ is a norm on $L^{\Phi_p}_{T,det}$, then

$$\left\| \int_0^T \int_{\mathbb{R}^d} f(t,x)Z(dt,dx) \right\|_p \simeq \|f\|.$$

Example 3.3.1. To demonstrate this for the classic stochastic integral with respect to Brownian motion, let B be an infinitely divisible random measure with Lévy-Khintchine triplet $B(C) \sim (0, m(C), 0)$. Here, $\Phi_p(t,x,u) = \Phi(u) = u^2$. Then $\Phi(|f(t,x)|) = |f(t,x)|^2$ and

$$L^{\Phi_p}_{T,det} = \{f \in L^0_{T,det} : \int_0^T \int_{\mathbb{R}^d} |f(t,x)|^2 \, m(dt,dx) < \infty\}.$$

Since Φ is convex, the Orlicz norm (see Appendix B)

$$\begin{aligned} \|f\|_{\Phi_p} &\overset{def}{=} \inf \left\{ c > 0 : \int_0^T \int_{\mathbb{R}^d} \Phi(c^{-1} |f(t,x)|) m(dt,dx) \le 1 \right\} \\ &= \inf \left\{ c > 0 : \int_0^T \int_{\mathbb{R}^d} c^{-2} |f(t,x)|^2 \, m(dt,x) \le 1 \right\} \\ &= \|f\|_{L^2([0,T]\times\mathbb{R}^d;m)} \end{aligned}$$

is a norm on $L^{\Phi_p}_{T,det} = L^2([0,T] \times \mathbb{R}^d; m)$ and we have

$$\left\| \int_0^T \int_{\mathbb{R}^d} f(t,x)Z(dt,dx) \right\|_{L^p(\Omega;\mathbb{P})} \simeq \|f\|_{L^2([0,T]\times\mathbb{R}^d;m)}.$$

We now use the results of Section 1.2 to obtain Itô Isomorphisms for certain cases of the infinitely divisible random measure Z. First assume that f is a Z-integrable deterministic measurable field. For a general infinitely divisible random measure Z, Rajput and Rosiński (1989, Theorem 2.7) showed that $\int_0^T \int_{\mathbb{R}^d} f(t,x)Z(dt,dx)$ was infinitely divisible with Lévy-Khintchine triplet (b_f, σ_f^2, F_f), where

$$b_f = \int_0^T \int_{\mathbb{R}^d} U(t,x,f(t,x)) m(dt,dx), \qquad \sigma_f^2 = \int_0^T \int_{\mathbb{R}^d} |f(t,x)|^2 \, \sigma^2(t,x) m(dt,dx),$$

and

$$F_f(B) = F(\{(t,x,z) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} : f(t,x)z \in B \setminus \{0\}\})$$

for every $B \in \mathcal{B}(\mathbb{R})$. Now suppose that Z is a mean 0 infinitely divisible random measure without Gaussian part. The Lévy-Khintchine triplet is necessarily

$$Z(C) \sim \left(\int_C \int_{\mathbb{R}} (\llbracket z \rrbracket - z)\, \nu(t,x,dz) m(dt,dx), 0, F_C \right),$$

where $F(dt,dx,dz) = \nu(t,x,dz)m(dt,dx)$. First assume that for each $(t,x) \in [0,T] \times \mathbb{R}^d$, $\nu(t,x,\cdot)$ is a symmetric Lévy measure on $\mathbb{R}$. In this case, (3.16) becomes

$$\Phi_p(t,x,u) = V_p(t,x,u) = \int_{\mathbb{R}} \left(|uz|^2 \mathbb{1}_{\{|uz|<1\}} + |uz|^p \mathbb{1}_{\{|uz| \geq 1\}} \right) \nu(t,x,dz)$$

and $\int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \sim (0,0,F_f)$ where F_f is the symmetric Lévy measure given by

$$F_f(B) = F(\{(t,x,z) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} : f(t,x)z \in B \setminus \{0\}\})$$

for every $B \in \mathcal{B}(\mathbb{R})$. If $\nu(t,x,\cdot)$ is not necessarily symmetric, Rajput and Rosiński (1989, Proposition 3.6) show that in this more general mean 0 case, there exists a constant C such that $\Phi_p(t,x,u) \leq C V_p(t,x,u)$. Trivially, we have $\Phi_p(t,x,u) \geq V_p(t,x,u)$ and hence, $\Phi_p(t,x,u) \simeq V_p(t,x,u)$. In either case, we have

$$L^{\Phi_p}_{T,det} = L^{V_p}_{T,det} = \left\{ f \in L^0_{T,det} : \int_0^T \int_{\mathbb{R}^d} V_p\left(t,x,|f(t,x)|\right) m(dt,dx) < \infty \right\}.$$

We now define a norm on this Musielak-Orlicz modular space of deterministic functions. To do so, define a new function $\Psi_p : [0,T] \times \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}_+$ by

$$\Psi_p(t,x,u) \overset{def}{=} \begin{cases} \displaystyle\int_{\mathbb{R}} \left(|uz|^2 \mathbb{1}_{\{|uz|<1\}} + \left(\frac{2}{p} |uz|^p + \frac{p-2}{p} \right) \mathbb{1}_{\{|uz| \geq 1\}} \right) \nu(t,x,dz), & \\ & \text{if } 1 \leq p \leq 2, \\ \displaystyle\int_{\mathbb{R}} \left(|uz|^2 \mathbb{1}_{\{|uz|<1\}} + |uz|^p \mathbb{1}_{\{|uz| \geq 1\}} \right) \nu(t,x,dz), & \text{if } p > 2. \end{cases}$$

It is easy to check that $\Psi_p(t,x,\cdot)$ is convex on $\mathbb{R}_+$ (since the integrand is now increasing and convex), of moderate growth, and for every $u \in \mathbb{R}_+$ and $1 \le p \le 2$,

$$V_p(t,x,u) \le \Psi_p(t,x,u) \le \frac{2}{p} V_p(t,x,u). \tag{3.18}$$

Therefore $L^{\Phi_p}_{T,det} = L^{V_p}_{T,det} = L^{\Psi_p}_{T,det}$ is a Musielak-Orlicz modular space with norm

$$\|f\|_{\Psi_p} \overset{def}{=} \inf\left\{c > 0 : \int_0^T \int_{\mathbb{R}^d} \Psi_p\left(t,x,c^{-1}|f(t,x)|\right) m(dt,dx) \le 1\right\}.$$

We now show that

$$\left\| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right\|_p \simeq \|f\|_{\Psi_p}.$$

Let l solve

$$\begin{aligned}
1 &= \int_{\mathbb{R}} \left(\frac{|x|^2}{l^2} \mathbb{1}_{\{|x|<1\}} + \frac{|x|^p}{l^p} \mathbb{1}_{\{|x|\ge 1\}} \right) F_f(dx) \\
&= \int_0^T \int_{\mathbb{R}^d} \int_R \left(\frac{|f(t,x)z|^2}{l^2} \mathbb{1}_{\{|f(t,x)z|<1\}} + \frac{|f(t,x)z|^p}{l^p} \mathbb{1}_{\{|f(t,x)z|\ge 1\}} \right) F(dt,dx,dz) \\
&= \int_0^T \int_{\mathbb{R}^d} \int_R \left(\frac{|f(t,x)z|^2}{l^2} \mathbb{1}_{\{|f(t,x)z|<1\}} + \frac{|f(t,x)z|^p}{l^p} \mathbb{1}_{\{|f(t,x)z|\ge 1\}} \right) \nu(t,x,dz)\, m(dt,dx) \\
&= \int_0^T \int_{\mathbb{R}^d} V_p\left(t,x,l^{-1}|f(t,x)|\right) m(dt,dx).
\end{aligned}$$

By Theorem 1.2.1, if $\nu(t,x,\cdot)$ is symmetric for every $(t,x) \in [0,T] \times \mathbb{R}^d$, then

$$0.25 l \le \left\| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right\|_p \le K(p) l,$$

where $K(p)$ is given by (1.4). In the more general mean 0 case, Corollary 1.2.6 gives

$$0.125 l \le \left\| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right\|_p \le \max\{\sqrt[p]{2}, \sqrt{2}\} K(p) l.$$

Using (3.18) and the convexity of $\Psi_p(t,x,\cdot)$, it is easy to show that if $1 \le p \le 2$, then

$$l \le \|f\|_{\Psi_p} \le \frac{2}{p} l.$$

We have proved the following theorem:

Theorem 3.3.2. *Let Z be a mean 0 infinitely divisible random measure with Lévy-Khintchine triplet* $Z(C) \sim \left(\int_C \int_{\mathbb{R}} (\llbracket z \rrbracket - z)\, \nu(t,x,dz) m(dt,dx), 0, F_C\right)$, *where* $F_C(B) = F(C \times B)$, $F(dt,dx,dz) = \nu(t,x,dz)\, m(dt,dx)$, *and for each* $(t,x) \in [0,T] \times \mathbb{R}^d$, $\nu(t,x,\cdot)$ *is a Lévy measure on* $\mathbb{R}$. *Then*

$$\tilde{k}(p) \left\| f \right\|_{\Psi_p} \le \left\| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right\|_p \le \max\{\sqrt[p]{2}, \sqrt{2}\} K(p) \left\| f \right\|_{\Psi_p}, \tag{3.19}$$

where

$$\tilde{k}(p) = \begin{cases} 0.0625p, & \text{if } 1 \le p \le 2, \\ 0.125, & \text{if } p > 2 \end{cases}$$

and $K(p)$ *is given by* (1.4). *If* $\nu(t,x,\cdot)$ *is symmetric for every* t *and* x, *then the constants in* (3.19) *may be taken as*

$$\tilde{k}(p) = \begin{cases} 0.125p, & \text{if } 1 \le p \le 2, \\ 0.25, & \text{if } p > 2 \end{cases}$$

and $K(p)$. *Hence,*

$$\left\{ f \in L^0_{T,det} : f \text{ is } Z\text{-integrable and } \mathbb{E} \left| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right|^p < \infty \right\} = L^{\Psi_p}_{T,det}$$

and the map $f \mapsto \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx)$ *is an isomorphism from* $L^{\Psi_p}_{T,det}$ *into* $L^p(\Omega; \mathcal{F}; \mathbb{P})$.

Next, we extend these arguments to $f \in L^0_T$ by decoupling inequalities. Let Z' be a copy of Z defined on a probability space Ω', independent of Z and $\mathcal{F}$. Define an enlarged σ-algebra on the product space $\Omega \times \Omega'$ by

$$\hat{\mathcal{F}}_t \stackrel{def}{=} \sigma\left(\mathcal{F}_t, Z'(C) : C \subset [0,t] \times \mathbb{R}^d, C \in \mathcal{B}([0,T] \times \mathbb{R}^d)\right).$$

For an $\mathcal{F}_t$-adapted simple random field $f(t,x,\omega) = \sum_{i=1}^m \sum_{j=1}^n f_{ij}(\omega) \mathbb{1}_{(t_i,t_{i+1}] \times A_j}(t,x)$, define two sequences of $\hat{\mathcal{F}}_{t_i}$-adapted random variables by

$$X_i(\omega,\omega') = X_i(\omega) \stackrel{def}{=} \sum_{j=1}^n f_{ij}(\omega) Z((t_i,t_{i+1}] \times A_j; \omega) \tag{3.20}$$

and

$$Y_i(\omega,\omega') \stackrel{def}{=} \sum_{j=1}^{n} f_{ij}(\omega) Z'((t_i,t_{i+1}] \times A_j;\omega'). \tag{3.21}$$

Sequences $\{X_i\}_{i=1}^m$ and $\{Y_i\}_{i=1}^m$ are $\hat{\mathcal{F}}_{t_i}$-tangent sequences and $\{Y_i\}_{i=1}^m$ satisfies conditional independence. Let $\varphi(u) \stackrel{def}{=} u^p$ and apply Theorem 3.2.6 to the decoupled $\hat{\mathcal{F}}_{t_i}$-tangent sequences $\{X_i\}_{i=1}^m$ and $\{Y_i\}_{i=1}^m$. We have the following theorem:

Theorem 3.3.3. *Let $f \in L_T^0$ be a simple random field. Then there exists a constant K_1, depending only on p, such that*

$$\left\| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right\|_p \le K_1 \left\| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z'(dt,dx) \right\|_p . \tag{3.22}$$

Moreover, if the Lévy measure $\nu(t,x,\cdot)$ is symmetric, then there exists a constant K_2, depending only on p, such that

$$\begin{aligned} K_2 \left\| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z'(dt,dx) \right\|_p &\le \left\| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right\|_p \\ &\le K_1 \left\| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z'(dt,dx) \right\|_p . \end{aligned} \tag{3.23}$$

Inequality (3.22) follows from the second inequality of Theorem 3.2.6 since the sequence $\{Y_i\}_{i=1}^m$ satisfies conditional independence. The first inequality in (3.23) follows from the first inequality of Theorem 3.2.6 and observing that $Z(C)$ and $Z'(C)$ are symmetric random variables whenever $\nu(t,x,\cdot)$ is symmetric.

We return to the problem of extending the Itô Isomorphism of Theorem 3.3.2 for deterministic fields to random fields. Let $\nu(t,x,\cdot)$ be a symmetric Lévy measure. By (3.23), we may first fix $\omega \in \Omega$ and consider

$$\mathbb{E}_{\omega'} \left| \int_0^T \int_{\mathbb{R}^d} f(t,x,\omega) Z'(dt,dx) \right|^p .$$

The beauty of decoupling arguments are now revealed. By "decoupling" the integrand from Z, we may now use deterministic integrand results. Fix $\omega \in \Omega$. The function Ψ_p is the same for both Z and Z' since each have the same Lévy-Khintchine triplet.

By Theorem 3.3.2, the deterministic field $f(\cdot,\cdot,\omega)$ is Z'-integrable with

$$\mathbb{E}_{\omega'}\left|\int_0^T\int_{\mathbb{R}^d} f(t,x,\omega)Z'(dt,dx)\right|^p<\infty$$

if and only if

$$\int_0^T\int_{\mathbb{R}^d}\Psi_p\left(t,x,|f(t,x,\omega)|\right)m(dt,dx)<\infty.$$

Hence, by (3.23),

$$\mathbb{E}\left|\int_0^T\int_{\mathbb{R}^d} f(t,x)Z(dt,dx)\right|^p<\infty$$

if and only if

$$\mathbb{E}\int_0^T\int_{\mathbb{R}^d}\Psi_p\left(t,x,|f(t,x)|\right)m(dt,dx)<\infty.$$

Moreover, Theorem 3.3.2 gives

$$\tilde{k}(p)^p\left\|f(\cdot,\cdot,\omega)\right\|_{\Psi_p}^p\le\mathbb{E}_{\omega'}\left|\int_0^T\int_{\mathbb{R}^d} f(t,x,\omega)Z'(dt,dx)\right|^p\le K(p)^p\left\|f(\cdot,\cdot,\omega)\right\|_{\Psi_p}^p.$$

By (3.23),

$$\begin{aligned}K_2^p\tilde{k}(p)^p\mathbb{E}_\omega\left\|f(\cdot,\cdot,\omega)\right\|_{\Psi_p}^p&\le K_2^p\mathbb{E}\left|\int_0^T\int_{\mathbb{R}^d} f(t,x)Z'(dt,dx)\right|^p\\&\le\mathbb{E}_\omega\left|\int_0^T\int_{\mathbb{R}^d} f(t,x)Z(dt,dx)\right|^p\\&\le K_1^p\mathbb{E}\left|\int_0^T\int_{\mathbb{R}^d} f(t,x)Z'(dt,dx)\right|^p\le K_1^pK(p)^p\mathbb{E}_\omega\left\|f(\cdot,\cdot,\omega)\right\|_{\Psi_p}^p.\end{aligned}\tag{3.24}$$

If $\nu(t,x,\cdot)$ is not symmetric, we still get the upper bound in (3.24) with constant $K_1^p\max\{\sqrt[p]{2},\sqrt{2}\}^pK(p)^p$ by (3.22) and the arguments above. Let us denote

$$L_T^{\Phi_p}=L_T^{V_p}\overset{def}{=}\left\{f\in L_T^0:\mathbb{E}\left|\int_0^T\int_{\mathbb{R}^d}V_p\left(t,x,|f(t,x)|\right)m(dt,dx)\right|^p<\infty\right\}.$$

Then $\left\|\left\|f(\cdot,\cdot,\omega)\right\|_{\Psi_p}\right\|_{L^p(\Omega;\mathcal{F};\mathbb{P})}$ is a norm on $L_T^{\Phi_p}$ and we can extend (3.24) by standard density arguments to measurable random fields $f\in L_T^0$ to get the following theorem:

Theorem 3.3.4 (Itô Isomorphisms)**.** *Let Z be a mean 0 infinitely divisible random measure with Lévy-Khintchine triplet $Z(C)\sim\left(\int_C\int_{\mathbb{R}}\left([\![z]\!]-z\right)\nu(t,x,dz)m(dt,dx),0,\right.$*

$F_C\big)$, where $F_C(B) = F(C\times B)$, $F(dt,dx,dz) = \nu(t,x,dz)m(dt,dx)$, and $\nu(t,x,\cdot)$ is a Lévy measure on $\mathbb{R}$. Assume that for each $(t,x)\in[0,T]\times\mathbb{R}^d$, $\nu(t,x,\cdot)$ is a symmetric measure on $\mathbb{R}$. Then

$$K_2\tilde{k}(p)\left\|\|f(\cdot,\cdot,\omega)\|_{\Psi_p}\right\|_{L^p(\Omega;\mathbb{P})} \leq \left\|\int_0^T\int_{\mathbb{R}^d} f(t,x)Z(dt,dx)\right\|_{L^p(\Omega;\mathbb{P})}$$
$$\leq K_1K(p)\left\|\|f(\cdot,\cdot,\omega)\|_{\Psi_p}\right\|_{L^p(\Omega;\mathbb{P})}, \quad (3.25)$$

where

$$\tilde{k}(p) = \begin{cases} 0.125p, & \text{if } 1\leq p\leq 2,\\ 0.25, & \text{if } p>2 \end{cases}$$

and $K(p)$ is given is (1.4). Hence,

$$\left\{f\in L^0_T : f \text{ is } Z\text{-integrable and } \mathbb{E}\left|\int_0^T\int_{\mathbb{R}^d} f(t,x)Z(dt,dx)\right|^p<\infty\right\} = L^{\Phi_p}_T.$$

and the map $f\mapsto \int_0^T\int_{\mathbb{R}^d} f(t,x)Z(dt,dx)$ is an isomorphism from $L^{\Phi_p}_T$ into $L^p(\Omega;\mathcal{F};\mathbb{P})$. If $\nu(t,x,\cdot)$ is not symmetric, then the right hand inequality of (3.25) still holds with constant $K_1\max\{\sqrt[p]{2},\sqrt{2}\}K(p)$ and hence,

$$\left\{f\in L^0_T : f \text{ is } Z\text{-integrable and } \mathbb{E}\left|\int_0^T\int_{\mathbb{R}^d} f(t,x)Z(dt,dx)\right|^p<\infty\right\} \supseteq L^{\Phi_p}_T.$$

3.3.1 Examples

Symmetric α-stable random measure

Let $0<\alpha<2$, $c>0$, and $\nu(t,x,dz)=\nu(dz)=c\,|z|^{-\alpha-1}\,dz$ in Theorem 3.3.4, so that Z is a symmetric α-stable random measure. As an example of the Itô isomorphism in Theorem 3.3.4, we first compute and simplify Φ_p for $1\leq p<\alpha$. Since ν does not depend upon t and x, $\Phi_p(t,x,u)=\Phi_p(u)$ and for $1\leq p<\alpha$ we have

$$\begin{aligned}\Phi_p(u) &= \int_{\mathbb{R}} \left(|uz|^2\,\mathbb{1}_{\{|uz|<1\}} + |uz|^p\,\mathbb{1}_{\{|uz|\geq 1\}}\right) c\,|z|^{-\alpha-1}\,dz\\ &= 2c\,|u|^2\int_0^{1/|u|} z^{1-\alpha}dz + 2c\,|u|^p\int_{1/|u|}^{\infty} z^{p-\alpha-1}dz\end{aligned}$$

$$
\begin{aligned}
&= \frac{2c}{2-\alpha}|u|^2\left(\frac{1}{|u|}\right)^{2-\alpha} + \frac{2c}{\alpha-p}|u|^p\left(\frac{1}{|u|}\right)^{p-2} \\
&= \frac{2c}{2-\alpha}|u|^\alpha + \frac{2c}{\alpha-p}|u|^\alpha \\
&= 2c\frac{2-p}{(2-\alpha)(\alpha-p)}|u|^\alpha .
\end{aligned}
$$

By Theorem 3.3.4, if $1 \le p < \alpha$, then

$$
\begin{aligned}
&\left\{ f \in L_T^0 : f \text{ is } Z\text{-integrable and } \mathbb{E}\left|\int_0^T \int_{\mathbb{R}^d} f(t,x)Z(dt,dx)\right|^p < \infty \right\} \\
&\qquad\qquad = \left\{ f \in L_T^0 : \mathbb{E}\left|\int_0^T \int_{\mathbb{R}^d} |f(t,x)|^\alpha m(dt,dx)\right|^p < \infty \right\}.
\end{aligned}
$$

Since $|u|^\alpha$ is already a convex function, it will generate an Orlicz norm equivalent to $\|\cdot\|_{\Psi_p}$. Let l solve

$$
\int_0^T \int_{\mathbb{R}^d} \left(\frac{|f(t,x)|}{l}\right)^\alpha m(dt,dx) = 1,
$$

that is,

$$
l^\alpha = \int_0^T \int_{\mathbb{R}^d} |f(t,x)|^\alpha m(dt,dx).
$$

Then for each $\omega \in \Omega$, l is equivalent to $\|f(t,x)\|_{\Psi_p}$ and by Theorem 3.3.4

$$
\mathbb{E}\left|\int_0^T \int_{\mathbb{R}^d} f(t,x)Z(dt,dx)\right|^p \simeq \mathbb{E}\left(\int_0^T \int_{\mathbb{R}^d} |f(t,x)|^\alpha m(dt,dx)\right)^{p/\alpha}.
$$

When $p = \alpha$, we no longer get the classic type Itô isomorphism since Z has infinite absolute α^{th}-moment. It would be desirable to take $p = \alpha$ in the above equivalence in order to obtain the analogue of the classic Itô Isometry for integration with respect to Brownian motion. However, we can get a desirable result similar in nature to the above Itô Isomorphism, given in terms of the weak-$L^\alpha(\Omega, \mathbb{P})$ space instead of the $L^\alpha(\Omega, \mathbb{P})$ space. In what follows in the remainder of this example, we will extend results of Giné and Marcus (1983) and Rosiński and Woyczyński (1986), culminating in Theorem 3.3.7. This theorem will address the problem presented above when $p = \alpha$ and give an analogue of the aforementioned Itó Isometry for α-stable random measures. As mentioned, we will see that the $L^\alpha(\Omega, \mathbb{P})$-norm on the left hand side of the above equivalence is replaced by a weak-$L^\alpha(\Omega, \mathbb{P})$-norm of the stochastic integral.

Before we present the result, recall that in the development of the stochastic integral, no assumptions were made on T except that $T > 0$. Therefore, for $T > 0$, we may consider the stochastic integral of a simple random field f to be a stochastic process $\{I_t(f)\}_{0\leq t\leq T}$ as follows: for a simple adapted random field $f \in \mathbb{S}$, we define the stochastic integral of f with respect to the α-stable random measure Z as

$$I_t(f) \stackrel{def}{=} \int_0^t \int_{\mathbb{R}^d} f(t,x,\omega) Z(dt,dx;\omega) \stackrel{def}{=} \sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij}(\omega) Z\left((t_i \wedge t, t_{i+1} \wedge t] \times A_j;\omega\right)$$

Since we are working with the weak-L^α norm, we may choose a separable version of $I_t(f)$ and our calculations will not be affected. So henceforth, we assume that $I_t(f)$ is separable. First, we establish some desirable properties of the process $\{I_t(f)\}_{0\leq t\leq T}$.

Proposition 3.3.5. *The process $I_t(f)$ is nonanticipating.*

Proof. Let $t \in (t_k, t_{k+1}]$. Then

$$I_t(f) = \sum_{i=1}^{k-1}\sum_{j=1}^{n} f_{ij} Z\left((t_i, t_{i+1}] \times A_j\right) + \sum_{j=1}^{n} f_{kj} Z\left((t_k, t] \times A_j\right).$$

Since f_{ij} is $\mathcal{F}_{t_i}$-adapted and $\mathcal{F}_{t_i} \subset \mathcal{F}_t$ for each i, f_{ij} is $\mathcal{F}_t$-adapted for each i. By the definition, $Z\left((t_i, t_{i+1}] \times A_j\right)$ and $Z\left((t_k, t] \times A_j\right)$ are $\mathcal{F}_t$-adapted for each i. □

Proposition 3.3.6. *The process $I_t(f)$ is linear. Namely, for $f, g \in L^\alpha_T \cap \mathbb{S}$ and $\alpha, \beta \in \mathbb{R}$,*

$$I_t(\alpha f + \beta g) = \alpha I_t(f) + \beta I_t(g).$$

Proof. Without loss of generality, we may assume that the simple random fields $f(t,x) = \sum_{i=1}^m \sum_{j=1}^n f_{ij} \mathbb{1}_{(t_i,t_{i+1}]\times A_j}$ and $g(t,x) = \sum_{i=1}^m \sum_{j=1}^n g_{ij} \mathbb{1}_{(t_i,t_{i+1}]\times A_j}$ for the same sets $(t_i, t_{i+1}] \times A_j \in \mathcal{B}_0([0,T] \times \mathbb{R}^d)$. Let $t \in (t_k, t_{k+1}]$. Then for any $\alpha, \beta > 0$,

$$\begin{aligned}
I_t(\alpha f + \beta g) &= \sum_{i=1}^{k-1}\sum_{j=1}^{n} \left(\alpha f_{ij} + \beta g_{ij}\right) Z_{ij} + \sum_{j=1}^{n} \left(\alpha f_{kj} + \beta g_{kj}\right) Z\left((t_k, t] \times A_j\right) \\
&= \alpha \left(\sum_{i=1}^{k-1}\sum_{j=1}^{n} f_{ij} Z\left((t_i, t_{i+1}] \times A_j\right) + \sum_{j=1}^{n} f_{kj} Z\left((t_k, t] \times A_j\right) \right)
\end{aligned}$$

$$+\beta\left(\sum_{i=1}^{k-1}\sum_{j=1}^{n} g_{ij} Z\left((t_i, t_{i+1}] \times A_j\right) + \sum_{j=1}^{n} g_{kj} Z\left((t_k, t] \times A_j\right)\right)$$
$$= \alpha I_t(f) + \beta I_t(g).$$

□

Following is the main result and gives the analogue of Itô's Isometry for integration driven by α-stable random measures. The upper bound for deterministic fields was proved by Giné and Marcus (1983) and the lower bound by Rosiński and Woyczyński (1986). We again utilize the decoupling inequalities approach.

Theorem 3.3.7. *Let $T > 0$ and $0 < \alpha < 2$. If $f \in L_T^\alpha \cap \mathbb{S}$ and Z is a α-stable random measure, then*

$$\Lambda_\alpha\left(\left\|\int_0^{*T}\int_{\mathbb{R}^d} f(s,x,\omega) Z(ds,dx;\omega)\right\|\right) \simeq \|f\|_{\alpha,T}, \tag{3.26}$$

where

$$\Lambda_\alpha\left(\left\|\int_0^{*T}\int_{\mathbb{R}^d} f(s,x,\omega) Z(ds,dx;\omega)\right\|\right) \stackrel{def}{=} \left\{\sup_{u>0} u^\alpha \mathbb{P}\left(\sup_{0\le t\le T}\left|\int_0^t\int_{\mathbb{R}^d} f(s,x,\omega) Z(ds,dx;\omega)\right| > u\right)\right\}^{1/\alpha}$$

and hence,

$$\left\{f \in L_T^0 : f \text{ is } Z\text{-integrable and } \Lambda_\alpha\left(\left\|\int_0^{*T}\int_{\mathbb{R}^d} f(t,x) Z(dt,dx)\right\|\right) < \infty\right\} = \left\{f \in L_T^0 : \mathbb{E}\int_0^T\int_{\mathbb{R}^d} |f(t,x)|^\alpha\, m(dt,dx) < \infty\right\} = L^\alpha_{T,a.s.}. \tag{3.27}$$

Proof. By standard density arguments and Section 3.2.3, (3.27) holds once we have proven (3.26). To prove (3.26), we apply Theorem 3.2.6, with $\varphi(w) \stackrel{def}{=} \mathbb{1}_{\{w>u\}}$, to the decoupled tangent sequences (3.20) and (3.21). Recalling that $\{Y_i\}$ satisfies conditional independence, for any $u > 0$ we have

$$\mathbb{P}\left(\max_{1\le k\le m}\left|\int_0^{t_k}\int_{\mathbb{R}^d} f(t,x,\omega) Z(dt,dx)\right| > u\right) = \mathbb{P}\left(\max_{1\le k\le m}\left|\sum_{i=1}^{k} X_i\right| > u\right)$$

$$
\begin{aligned}
&\le C\,\mathbb{P}\otimes\mathbb{P}'\left(\max_{1\le k\le m}\left|\sum_{i=1}^{k} Y_i\right| > u\right)\\
&= C\int_\Omega \mathbb{P}'\left(\max_{1\le k\le m}\left|\int_0^{t_k}\int_{\mathbb{R}^d} f(t,x,\omega)Z'(dt,dx;\omega')\right| > u\right)\mathbb{P}(d\omega).
\end{aligned}
$$

By the deterministic integrand result of Giné and Marcus (1983), for each $\omega\in\Omega$,

$$
u^\alpha \mathbb{P}'\left(\max_{1\le k\le m}\left|\int_0^{t_k}\int_{\mathbb{R}^d} f(t,x,\omega)Z'(dt,dx;\omega')\right| > u\right) \le \int_0^T\int_{\mathbb{R}^d} |f(t,x,\omega)|^\alpha\, m(dt,dx),
$$

and hence,

$$
\begin{aligned}
&u^\alpha \mathbb{P}\left(\max_{1\le k\le m}\left|\int_0^{t_k}\int_{\mathbb{R}^d} f(t,x,\omega)Z(dt,dx)\right| > u\right)\\
&\quad\le C\int_\Omega\int_0^T\int_{\mathbb{R}^d} |f(t,x,\omega)|^\alpha\, m(dt,dx)\mathbb{P}(d\omega) = C\,\mathbb{E}\int_0^T\int_{\mathbb{R}^d} |f(t,x)|^\alpha\, m(dt,dx).
\end{aligned}
$$

Since we have made no assumptions on the partition $\{t_k\}$ of $[0,T]$, the partition lengths may be taken arbitrarily small. Also, $u>0$ was arbitrary. The upper bound is now immediate.

For the lower bound, again let $u>0$ be arbitrary. By Theorem 3.2.6,

$$
\begin{aligned}
\mathbb{P}\left(\max_{1\le k\le m}\left|\int_0^{t_k}\int_{\mathbb{R}^d} f(t,x,\omega)Z(dt,dx)\right| > u\right) &= \mathbb{P}\left(\max_{1\le k\le m}\left|\sum_{i=1}^{k} X_i\right| > u\right)\\
&\ge C\max_{\varepsilon_i=\pm 1}\mathbb{P}\otimes\mathbb{P}'\left(\max_{1\le k\le m}\left|\sum_{i=1}^{k}\varepsilon_i Y_i\right| > u\right)\\
&= C\int_\Omega \mathbb{P}'\left(\max_{1\le k\le m}\left|\int_0^{t_k}\int_{\mathbb{R}^d} \varepsilon f(t,x,\omega)Z'(dt,dx;\omega')\right| > u\right)\mathbb{P}(d\omega),
\end{aligned}
$$

where $\varepsilon f \overset{def}{=} \sum_{i=1}^{n}\varepsilon_i Y_i$. By the deterministic integrand result of Rosiński and Woyczyński (1986), for each $\omega\in\Omega$,

$$
\begin{aligned}
u^\alpha \mathbb{P}'\left(\max_{1\le k\le m}\left|\int_0^{t_k}\int_{\mathbb{R}^d} f(t,x,\omega)Z'(dt,dx;\omega')\right| > u\right) &\ge \int_0^T\int_{\mathbb{R}^d} |\varepsilon f(t,x,\omega)|^\alpha\, m(dt,dx)\\
&= \int_0^T\int_{\mathbb{R}^d} |f(t,x,\omega)|^\alpha\, m(dt,dx).
\end{aligned}
$$

The result now follows exactly as in the upper bound case. □

Symmetric tempered α-stable random measure

Let $0 < \alpha < 2$, $c > 0$, and $\nu(t,x,dz) = \nu(dz) = c\,|z|^{-\alpha-1}\,e^{-|z|}dz$ in Theorem 3.3.4, so that Z is a symmetric tempered α-stable random measure. We demonstrate how to identify the Itô isomorphisms in Theorem 3.3.4. First, we will identify simpler functions that are equivalent to $\Phi_p(t,x,u)$. Recall that the notation $g \simeq h$ means there exists $C > 0$ such that $(1/C)h(u) \le g(u) \le Ch(u)$ for every u and we say that the functions g and h are equivalent. We have the following lemma giving equivalent functions of Φ_p.

Lemma 3.3.8. *Let $p \ge 1$ and $0 < \alpha < 2$.*

i. If $p < \alpha$

$$\Phi_p(u) \simeq |u|^2 \wedge |u|^\alpha .$$

ii. If $p = \alpha$

$$\Phi_p(u) \simeq |u|^2\,\mathbb{1}_{\{|u|\le 1\}} + |u|^\alpha\,(\ln|u| + 1)\,\mathbb{1}_{\{|u|>1\}}.$$

iii. If $p > \alpha$

$$\Phi_p(u) \simeq |u|^2\,\mathbb{1}_{\{|u|\le 1\}} + |u|^p\,\mathbb{1}_{\{|u|>1\}}.$$

Proof. Since the Lévy measure is symmetric and does not depend upon t or x, $\Phi_p(t,x,u) = \Phi_p(u)$ is given by

$$\begin{aligned}\Phi_p(u) &= \int_{\mathbb{R}} \left(|uz|^2\,\mathbb{1}_{\{|uz|<1\}} + |uz|^p\,\mathbb{1}_{\{|uz|\ge 1\}}\right) c\,|z|^{-\alpha-1}\,e^{-|z|}dz \\ &= 2\int_0^{1/|u|} |u|^2\,z^{1-\alpha}e^{-z}c\,dz + 2\int_{1/|u|}^{\infty} |u|^p\,z^{p-\alpha-1}e^{-z}c\,dz.\end{aligned} \tag{3.28}$$

Let $p \ge 1$. First, suppose that $p < \alpha < 2$. Just as in the proof of (3.11), the first integral in (3.28) is bounded by

$$2\int_0^{1/|u|} |u|^2\,z^{1-\alpha}e^{-z}c\,dz \le 2c\Gamma(2-\alpha)u^2\mathbb{1}_{\{|u|\le 1\}} + \frac{2c}{2-\alpha}\,|u|^\alpha\,\mathbb{1}_{\{|u|>1\}}. \tag{3.29}$$

The second integral in (3.28) is bounded by

$$\begin{aligned}
&2\int_{1/|u|}^{\infty} |u|^p\, z^{p-\alpha-1}e^{-z}c dz \\
&\quad = 2\mathbb{1}_{\{|u|\le 1\}}\int_{1/|u|}^{\infty} |u|^p\, z^{p-\alpha-1}e^{-z}cdz + 2\mathbb{1}_{\{|u|>1\}}\int_{1/|u|}^{\infty} |u|^p\, z^{p-\alpha-1}e^{-z}cdz \\
&\quad \le 2c\sup_{z\ge 1}\{z^{2-\alpha}e^{-z}\}\,|u|^p\,\mathbb{1}_{\{|u|\le 1\}}\int_{1/|u|}^{\infty} z^{p-3}dz + 2c\,|u|^p\,\mathbb{1}_{\{|u|>1\}}\int_{1/|u|}^{\infty} z^{p-\alpha-1}dz \\
&\quad = \frac{2c}{2-p}\sup_{z\ge 1}\{z^{2-\alpha}e^{-z}\}\,|u|^2\,\mathbb{1}_{\{|u|\le 1\}} + \frac{2c}{\alpha-p}\,|u|^\alpha\,\mathbb{1}_{\{|u|>1\}}. \qquad (3.30)
\end{aligned}$$

Combining (3.29) and (3.30) gives

$$\Phi_p(u) \le C\left(|u|^2 \wedge |u|^\alpha\right).$$

The lower bound follows exactly as in the proof of (3.11) to get

$$\Phi_p(u) \ge \left(2\gamma(2-\alpha,1)\wedge\frac{2e^{-1}}{2-\alpha}\right)\left(|z|^2\wedge|z|^\alpha\right).$$

Next let $p=\alpha<2$. The first integral in (3.28) is bounded above by

$$\begin{aligned}
2\int_0^{1/|u|} |u|^2\, z^{1-\alpha}e^{-z}cdz &= 2c\,|u|^2\,\mathbb{1}_{\{|u|\le 1\}}\int_0^{1/|u|} z^{1-\alpha}e^{-z}dz \\
&\quad + 2c\,|u|^2\,\mathbb{1}_{\{|u|>1\}}\int_0^{1/|u|} z^{1-\alpha}e^{-z}dz \\
&\le 2c\,|u|^2\,\mathbb{1}_{\{|u|\le 1\}}\Gamma(2-\alpha) + \frac{2c}{2-\alpha}\,|u|^2\,\mathbb{1}_{\{|u|>1\}}\left(\frac{1}{|u|}\right)^{2-\alpha} \qquad (3.31) \\
&= 2c\,|u|^2\,\mathbb{1}_{\{|u|\le 1\}}\Gamma(2-\alpha) + \frac{2c}{2-\alpha}\,|u|^\alpha\,\mathbb{1}_{\{|u|>1\}} \\
&\le 2c\,|u|^2\,\mathbb{1}_{\{|u|\le 1\}}\Gamma(2-\alpha) + \frac{2c}{2-\alpha}\,|u|^\alpha\left(\ln|u|+1\right)\mathbb{1}_{\{|u|>1\}}
\end{aligned}$$

and the second integral in (3.28) by

$$\begin{aligned}
&2\int_{1/|u|}^{\infty} |u|^p\, z^{p-\alpha-1}e^{-z}cdz \\
&\quad = 2c\,|u|^\alpha\,\mathbb{1}_{\{|u|\le 1\}}\int_{1/|u|}^{\infty} z^{-1}e^{-z}dz + 2c\,|u|^\alpha\,\mathbb{1}_{\{|u|>1\}}\int_{1/|u|}^{\infty} z^{-1}e^{-z}dz
\end{aligned}$$

$$\leq 2c\,|u|^{\alpha}\,\mathbb{1}_{\{|u|\leq 1\}}\left(\frac{1}{|u|}\right)^{-1} + 2c\,|u|^{\alpha}\,\mathbb{1}_{\{|u|>1\}}\left(\int_{1/|u|}^{1} z^{-1}dz + \int_{1}^{\infty} e^{-z}dz\right) \tag{3.32}$$
$$= 2c\,|u|^{\alpha+1}\,\mathbb{1}_{\{|u|\leq 1\}} + 2c\,|u|^{\alpha}\,\mathbb{1}_{\{|u|>1\}}\left(\ln|u| + e^{-1}\right)$$
$$\leq 2c\,|u|^{2}\,\mathbb{1}_{\{|u|\leq 1\}} + 2c\,|u|^{\alpha}\,\mathbb{1}_{\{|u|>1\}}\left(\ln|u| + 1\right).$$

Combining (3.31) and (3.32) gives

$$\Phi_p(u) \leq C\left(|u|^2\,\mathbb{1}_{\{|u|\leq 1\}} + |u|^{\alpha}\ln|u|\,\mathbb{1}_{\{|u|>1\}} + |u|^{\alpha}\,\mathbb{1}_{\{|u|>1\}}\right).$$

To compute the lower bound,

$$\Phi_p(u) \geq 2c\mathbb{1}_{\{|u|\leq 1\}}\,|u|^2\int_0^{1/|u|} z^{1-\alpha}e^{-z}dz + 2c\,|u|^{\alpha}\,\mathbb{1}_{\{|u|>1\}}\int_{1/|u|}^{\infty} z^{-1}e^{-z}dz$$
$$\geq 2c\gamma(2-\alpha,1)\mathbb{1}_{\{|u|\leq 1\}}\,|u|^2 + 2c\,|u|^{\alpha}\,\mathbb{1}_{\{|u|>1\}}\left(\int_{1/|u|}^{1} z^{-1}e^{-1}dz + \int_{1}^{\infty} z^{-1}e^{-z}dz\right)$$
$$= 2c\gamma(2-\alpha,1)\mathbb{1}_{\{|u|\leq 1\}}\,|u|^2 + 2c\,|u|^{\alpha}\,\mathbb{1}_{\{|u|>1\}}\left(e^{-1}\ln|u| + \Gamma(0,1)\right)$$
$$\geq C\left(|u|^2\,\mathbb{1}_{\{|u|\leq 1\}} + |u|^{\alpha}\ln|u|\,\mathbb{1}_{\{|u|>1\}} + |u|^{\alpha}\,\mathbb{1}_{\{|u|>1\}}\right).$$

Finally let $p > \alpha$. To compute the lower bound of $\Phi_p(u)$,

$$\Phi_p(u) \geq 2c\,|u|^2\,\mathbb{1}_{\{|u|\leq 1\}}\int_0^{1/|u|} z^{1-\alpha}e^{-z}dz + 2c\,|u|^{p}\,\mathbb{1}_{\{|u|>1\}}\int_{1/|u|}^{\infty} z^{p-\alpha-1}e^{-z}dz$$
$$\geq 2c\,|u|^2\,\mathbb{1}_{\{|u|\leq 1\}}\gamma(2-\alpha,1) + 2c\,|u|^{p}\,\mathbb{1}_{\{|u|>1\}}\Gamma(p-\alpha,1)$$
$$\geq C\left(|u|^2\,\mathbb{1}_{\{|u|\leq 1\}} + |u|^{p}\,\mathbb{1}_{\{|u|>1\}}\right).$$

To compute the upper bound, we consider three subcases. First suppose that $p < 2$. We can bound the first integral in (3.28) by

$$2\int_0^{1/|u|} |u|^2\,z^{1-\alpha}e^{-z}cdz$$
$$\leq 2c\Gamma(2-\alpha)u^2\mathbb{1}_{\{|u|\leq 1\}} + 2c\sup_{z>0}\{z^{p-\alpha}e^{-z}\}\,|u|^2\,\mathbb{1}_{\{|u|>1\}}\int_0^{1/|u|} z^{1-p}dz$$
$$= 2c\Gamma(2-\alpha)u^2\mathbb{1}_{\{|u|\leq 1\}} + \frac{2c}{2-p}\sup_{z>0}\{z^{p-\alpha}e^{-z}\}\,|u|^{p}\,\mathbb{1}_{\{|u|>1\}}. \tag{3.33}$$

and the second integral in (3.28) by

$$\begin{aligned}
2\int_{1/|u|}^{\infty} |u|^p z^{p-\alpha-1} e^{-z} c dz
&= 2\mathbb{1}_{\{|u|\le 1\}} \int_{1/|u|}^{\infty} |u|^p z^{p-\alpha-1} e^{-z} c dz + 2\mathbb{1}_{\{|u|>1\}} \int_{1/|u|}^{\infty} |u|^p z^{p-\alpha-1} e^{-z} c dz \\
&\le 2c \sup_{z\ge 1}\{z^{2-\alpha}e^{-z}\} |u|^p \mathbb{1}_{\{|u|\le 1\}} \int_{1/|u|}^{\infty} z^{p-3} dz + 2c |u|^p \mathbb{1}_{\{|u|>1\}} \Gamma(p-\alpha) \\
&= \frac{2c}{2-p} \sup_{z\ge 1}\{z^{2-\alpha}e^{-z}\} |u|^2 \mathbb{1}_{\{|u|\le 1\}} + 2c\Gamma(p-\alpha) |u|^p \mathbb{1}_{\{|u|>1\}}. \qquad (3.34)
\end{aligned}$$

Combining (3.33) and (3.34) gives

$$\Phi_p(u) \le C \left(|u|^2 \mathbb{1}_{\{|u|\le 1\}} + |u|^p \mathbb{1}_{\{|u|>1\}} \right).$$

Next, if $p = 2$, then $\Phi_p(u) = c\Gamma(2-\alpha)|u|^2$. Finally suppose that $p > 2$. We can bound the first integral in (3.28) by

$$\begin{aligned}
2\int_0^{1/|u|} |u|^2 z^{1-\alpha} e^{-z} c dz &\le 2c\Gamma(2-\alpha) |u|^2 \\
&\le 2c\Gamma(2-\alpha) \left(|u|^2 \mathbb{1}_{\{|u|\le 1\}} + |u|^p \mathbb{1}_{\{|u|>1\}} \right)
\end{aligned} \qquad (3.35)$$

and the second integral in (3.28) by

$$\begin{aligned}
2\int_{1/|u|}^{\infty} |u|^p z^{p-\alpha-1} e^{-z} c dz &\le 2c\Gamma(p-\alpha) |u|^p \\
&\le 2c\Gamma(p-\alpha) \left(|u|^2 \mathbb{1}_{\{|u|\le 1\}} + |u|^p \mathbb{1}_{\{|u|>1\}} \right).
\end{aligned} \qquad (3.36)$$

Combining (3.35) and (3.36) again gives

$$\Phi_p(u) \le C \left(|u|^2 \mathbb{1}_{\{|u|\le 1\}} + |u|^p \mathbb{1}_{\{|u|>1\}} \right).$$

□

Let $p \ge 1$ and Z be a symmetric tempered stable random measure with Lévy-Khintchine triplet $Z(C) \sim (0, 0, m(C)\nu)$, where $\nu(dx) = c|x|^{-\alpha-1} e^{-|x|} dx$. By Theorem 3.3.4 and the above Lemma 3.3.8, we have the following three Itô isomorphisms:

i. If $p < \alpha$,

$$\left\{ f \in L_T^0 : f \text{ is } Z\text{-integrable and } \mathbb{E} \left| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right|^p < \infty \right\}$$
$$= \left\{ f \in L_T^0 : \mathbb{E} \left(\int_0^T \int_{\mathbb{R}^d} |f(t,x)|^2 \wedge |f(t,x)|^\alpha \, m(dt,dx) \right)^p < \infty \right\}.$$

ii. If $p = \alpha$,

$$\left\{ f \in L_T^0 : f \text{ is } Z\text{-integrable and } \mathbb{E} \left| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right|^\alpha < \infty \right\}$$
$$= \left\{ f \in L_T^0 : \mathbb{E} \left(\int_0^T \int_{\mathbb{R}^d} \left(|f(t,x)|^2 \mathbb{1}_{\{|f(t,x)| \le 1\}} \right. \right. \right.$$
$$\left. \left. \left. + |f(t,x)|^\alpha \left(\ln |f(t,x)| + 1 \right) \mathbb{1}_{\{|f(t,x)| > 1\}} \right) m(dt,dx) \right)^\alpha < \infty \right\}.$$

iii. If $p > \alpha$,

$$\left\{ f \in L_T^0 : f \text{ is } Z\text{-integrable and } \mathbb{E} \left| \int_0^T \int_{\mathbb{R}^d} f(t,x) Z(dt,dx) \right|^p < \infty \right\}$$
$$= \left\{ f \in L_T^0 : \mathbb{E} \left(\int_0^T \int_{\mathbb{R}^d} \left(|f(t,x)|^2 \mathbb{1}_{\{|f(t,x)| \le 1\}} \right. \right. \right.$$
$$\left. \left. \left. + |f(t,x)|^p \mathbb{1}_{\{|f(t,x)| > 1\}} \right) m(dt,dx) \right)^p < \infty \right\}.$$

We see that when $p < \alpha$, the condition imposed upon f is a mixture between the classic Itô Isometry (for integration with respect to Brownian motion) and the symmetric α-stable example of Section 3.3.1.

Chapter 4

Summary and Future Directions

The focus of this dissertation has been to study the L^p-norm of infinitely divisible random vectors and present several applications. Such distributions are important in that they contain many of the named distributions used in numerous disciplines. They have the desirable property of being completely characterized by their parameterization given by the Lévy-Khintchine triplet and this work gives explicit estimate of the L^p-norm of said distributions in terms of these parameters. Of most importance is the $1 \leq p < 2$ case, as L^2-theory and orthogonality are not applicable. This result was demonstrated most useful in Chapter 2. In this, we obtained the optimal linear estimate of the state space in a discrete times signal-observation model in the presence of an α-stable noise environment. Often, data collected demonstrates outlying realizations not probable under the Gaussian assumption. In such instances, heavy-tailed infinitely divisible distributions may be more appropriate model assumptions. The L^p-norm results are precisely the tools needed to work under such assumptions.

In the continuous time case, model "noise" is often given by a stochastic integral. In Chapter 3, I defined the stochastic integral driven by infinitely divisible random measures. Throughout I use a very powerful, but not widely employed method known as decoupling. This method allows one to treat integrands and integrators independently. Using the L^p-norm result, I was able to obtain Itô Isomorphisms for such stochastic integrals. As an immediate future extension of this dissertation, I would like to explore continuous time models and apply these Itô Isomorphisms to approximate the solutions.

As another extension, I would like to apply the L^p-norm result to both discrete and continuous time models for which skewed data has been observed. In this dissertation, I have mainly focused on symmetric distributions as this is a very common assumption in modeling. By no means however is this result restricted to such distributions.

Bibliography

Bar-Shalom, Y., Li, X.-R., and Kirubarajan, T. (2001). *Estimation with applications to tracking and navigation*. John Wiley and Sons.

Bensoussan, A. and Lions, J.-L. (1984). *Impulse Control and Quasi-Variational Inequalities*. Gauthier-Villars, Paris.

de la Peña, V. and Giné, E. (1999). *Decoupling: from dependence to independence*. Springer-Verlag.

Dudley, R. M. and Norvaiša, R. (2010). *Concrete Functional Calculus*. Springer.

Gentle, J. E. (2007). *Matrix Algebra: Theory, Computations, and Applications in Statistics*. Springer.

Giné, E. and Marcus, M. (1983). The central limit theorem for stochastic integrals with respect to lévy processes. *The Annals of Probability*, 11(1).

Gordon, N., Percival, J., and Robinson, M. (2003). The kalman-levy filter and heavy-tailed models for tracking manoeuvring targets. In *Information Fusion, 2003. Proceedings of the Sixth International Conference of*, volume 2, pages 1024 – 1031.

Jacob, J. and Protter, P. (2004). *Probability Essentials*. Springer, 2nd edition.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82.

Kwapień, S. and Woyczyński, W. (1992). *Random Series and Stochastic Integrals: Single and Multiple*. Birkhäuser.

Kwapień, S. and Woyczyński, W. (2009). *Random Series and Stochastic Integrals: Single and Multiple*. Unpublished, 2nd edition.

Latała, R. (1997). Estimation of moments of sums of independent real random variables. *The Annals of Probability*, 25(3).

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach spaces : isoperimetry and processes*. Springer-Verlag.

Marcus, M. and Rosiński, J. (2001). L^1-norm of infinitely divisible random vectors and certain stochastic integrals. *Electronic Communications in Probability*, 6.

Rajput, B. S. and Rosiński, J. (1989). Spectral representations of infinitely divisible processes. *Probability Theory and Related Fields*, 82.

Rolewicz, S. (1972). *Metric Linear Spaces*. Polish Scientific Publishers.

Rosiński, J. (2007a). *Lévy and Related Jump-type Infinitely Divisible Processes*. Unpublished Lecture Notes.

Rosiński, J. (2007b). Tempering stable processes. *Stochastic Processes and their Applications*, 117.

Rosiński, J. and Woyczyński, W. (1986). On itô stochastic integration with respect to p-stable motion: Inner clock, integrability of sample paths, double and multiple integrals. *The Annals of Probability*, 14(1).

Rudin, W. (1987). *Real and Complex Analysis*. McGraw-Hill, 3rd edition.

Sato, K.-I. (2002). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, english edition.

Stuck, B. W. (1978). Minimum error dispersion linear filtering of scalar symmetric stable processes. *IEEE Transactions on Automatic Control*, AC-23(3).

Woyczyński, W. (1974). Random series and laws of large numbers in some banach spaces. *Theory of Probability and its Applications*, 18.

Appendices

Appendix A

Moments of Independent Random Variables and Vectors

This chapter develops a few useful inequalities concerning sums of independent symmetric or mean 0 random vectors. Most inequalities were first derived for random variables. In some instances, the results were able to be extended to random vectors taking values in a Hilbert space H. I include some of the results obtained for random variables, even if the result is superseded by a result concerning random vectors. In one instance, this is because the constants are sharper in the $H = \mathbb{R}$ case. The first two theorems below deal with bounding $\mathbb{E}\,\|S_n\|^p$, which is difficult (if not impossible in most instances) to compute directly, by computable moments.

Theorem A.1. *Let $2 < p \leq 3$ and $\{X_i\}_{n\in\mathbb{N}}$ be independent mean 0 random vectors taking values in a Hilbert space H. Put $S_n \stackrel{def}{=} \sum_{i=1}^n X_i$. Then*

$$\mathbb{E}\,\|S_n\|^p \leq \sum_{i=1}^n \mathbb{E}\,\|X_i\|^p + \frac{p(p-1)}{2} \sum_{i=1}^n \mathbb{E}\,\|S_{i-1}\|^2\,\mathbb{E}\,\|X_i\|^{p-2}\,.$$

Proof. Fix $y \in H$ and consider the map

$$f : H \to \mathbb{R} : x \mapsto \|x+y\|^p - \|x\|^p\,.$$

There is a version of Taylor's theorem for functions mapping a Banach space into another Banach space, given in terms of the Gâteaux derivative (see e.g. Dudley and Norvaiša (2010, Chapter 5) for an overview). When working with real Hilbert spaces, the Gâteaux and Frechet derivatives coincide and the Gâteaux derivative is a linear functional. Applying Taylor's theorem to f, about the vector $0 \in H$, we have

$$f(x) = f(0) + df(0; x) + R(x), \tag{A.1}$$

where $df(u; x)$ is the Gâteaux derivative of $f(u)$ in the direction of x defined by

$$df(u; x) \stackrel{def}{=} \lim_{\tau \to 0} \frac{f(u + \tau x) - f(u)}{\tau}$$

and

$$R(x) = \frac{1}{2} d^2 f(\xi; x, x) \tag{A.2}$$

for some ξ lying on the line segment between 0 and x, strictly between the points. In the following computations, we calculate the first and second Gâteaux derivatives of $f(u)$ in the direction of x. First, we compute the derivative of $\|x\|$. We have

$$\begin{aligned} d\left(\|u\|\right)(u; x) &= \lim_{\tau \to 0} \frac{\|u + \tau x\| - \|u\|}{\tau} \\ &= \lim_{\tau \to 0} \frac{\|u + \tau x\|^2 - \|u\|^2}{\tau\left(\|u + \tau x\| + \|u\|\right)} \\ &= \lim_{\tau \to 0} \frac{\langle u + \tau x, u + \tau x \rangle - \langle u, u \rangle}{\tau\left(\|u + \tau x\| + \|u\|\right)} \\ &= \lim_{\tau \to 0} \frac{2\langle u, \tau x \rangle + \langle \tau x, \tau x \rangle}{\tau\left(\|u + \tau x\| + \|u\|\right)} \\ &= \lim_{\tau \to 0} \frac{2\langle u, x \rangle + \langle x, \tau x \rangle}{\left(\|u + \tau x\| + \|u\|\right)} \\ &= \frac{\langle u, x \rangle}{\|u\|} \\ &= \left\langle \frac{u}{\|u\|}, x \right\rangle. \end{aligned} \tag{A.3}$$

Now in the calculus of Gâteaux derivatives, there is also a chain rule given by

$$d(G \circ F)(u; x) = dG(F(u); dF(u; x)).$$

Applying the chain rule to $\|u\|^p$ and utilizing (A.3), we have

$$d\left(\|u\|^p\right)(u; x) = d(y^p)\left(\|u\|; d\left(\|u\|\right)(u; x)\right)$$

Now the directional derivative of the map $\mathbb{R} \to \mathbb{R} : w \mapsto w^p$ in the direction z is

$$\begin{aligned} d(w^p)(w;z) &= \lim_{\tau \to 0} \frac{(w+\tau z)^p - w^p}{\tau} \\ &= \lim_{h \to 0} \frac{(w+h)^p - w^p}{h/z} \\ &= zpw^{p-1}, \end{aligned}$$

and hence,

$$\begin{aligned} d\left(\|u\|^p\right)(u;x) &= \left(d\left(\|u\|\right)(u;x)\right) p \, \|u\|^{p-1} \\ &= p \left\langle \frac{u}{\|u\|}, x \right\rangle \|u\|^{p-1} \\ &= p\langle u, x\rangle \, \|u\|^{p-2}. \end{aligned} \tag{A.4}$$

Applying the chain rule again, we compute

$$\begin{aligned} d\left(\|u+y\|^p\right)(u;x) &= d\left(\|u\|^p\right)(u+y, d(u+y)(u;x)) \\ &= d\left(\|u\|^p\right)(u+y, x) \\ &= p\langle u+y, x\rangle \, \|u+y\|^{p-2}. \end{aligned} \tag{A.5}$$

Combining (A.4) and (A.5) gives

$$df(u;x) = d\left(\|u+y\|^p - \|u\|^p\right)(u;x) = p\langle u+y, x\rangle \, \|u+y\|^{p-2} - p\langle u, x\rangle \, \|u\|^{p-2} \tag{A.6}$$

and

$$df(0;x) = p\langle y, x\rangle \, \|y\|^{p-2}$$

Next, we compute $d^2 f(u;x)$:

$$\begin{aligned} d^2 f(u;x,x) &= \lim_{\tau \to 0} \frac{df(u+\tau x;x) - df(u;x)}{\tau} \\ &= \lim_{\tau \to 0} \left(\frac{p\langle u+\tau x+y, x\rangle \, \|u+\tau x+y\|^{p-2} - p\langle u+\tau x, x\rangle \, \|u+\tau x\|^{p-2}}{\tau} \right. \\ &\quad \left. - \frac{p\langle u+y, x\rangle \, \|u+y\|^{p-2} - p\langle u, x\rangle \, \|u\|^{p-2}}{\tau} \right). \end{aligned}$$

Consider first

$$\begin{aligned}
&\lim_{\tau\to 0}\frac{p\langle u+\tau x+y,x\rangle\left\|u+\tau x+y\right\|^{p-2}-p\langle u+y,x\rangle\left\|u+y\right\|^{p-2}}{\tau}\\
&=\lim_{\tau\to 0}\left(\frac{p\langle u+y,x\rangle\left(\left\|u+\tau x+y\right\|^{p-2}-\left\|u+y\right\|^{p-2}\right)}{\tau}+\frac{p\langle\tau x,x\rangle\left\|u+\tau x+y\right\|^{p-2}}{\tau}\right)\\
&=p\langle u+y,x\rangle d\left(\left\|u+y\right\|^{p-2}\right)(u;x)+p\langle x,x\rangle\left\|u+y\right\|^{p-2}\\
&=p\langle u+y,x\rangle(p-2)\langle u+y,x\rangle\left\|u+y\right\|^{p-4}+p\langle x,x\rangle\left\|u+y\right\|^{p-2}\\
&=p(p-2)\langle u+y,x\rangle^2\left\|u+y\right\|^{p-4}+p\left\|x\right\|^2\left\|u+y\right\|^{p-2}, \qquad \text{(A.7)}
\end{aligned}$$

the next to last equality coming from (A.5). Take $y=0$ in (A.7) to get

$$\begin{aligned}
&\lim_{\tau\to 0}\frac{p\langle u+\tau x,x\rangle\left\|u+\tau x\right\|^{p-2}-p\langle u,x\rangle\left\|u\right\|^{p-2}}{\tau}\\
&\qquad\qquad=p(p-2)\langle u,x\rangle^2\left\|u\right\|^{p-4}+p\left\|x\right\|^2\left\|u\right\|^{p-2} \qquad \text{(A.8)}
\end{aligned}$$

and subtracting (A.8) from (A.7) gives

$$\begin{aligned}
d^2f(u;x,x)&=p\left\|x\right\|^2\left(\left\|u+y\right\|^{p-2}-\left\|u\right\|^{p-2}\right)\\
&\qquad+p(p-2)\left(\langle u+y,x\rangle^2\left\|u+y\right\|^{p-4}-\langle u,x\rangle^2\left\|u\right\|^{p-4}\right).
\end{aligned}$$

We are now ready to give the Taylor expansion $f(x)=\|x+y\|^p-\|x\|^p$, where $y\in H$ is fixed. Substituting the first and second Gâteaux derivatives in the the Taylor expansion (A.1) gives

$$\begin{aligned}
\left\|x+y\right\|^p-\left\|x\right\|^p&=\left\|y\right\|^p+p\langle y,x\rangle\left\|y\right\|^{p-2}+\frac{1}{2}\left(p\left\|x\right\|^2\left(\left\|\xi+y\right\|^{p-2}-\left\|\xi\right\|^{p-2}\right)\right.\\
&\qquad\left.+\,p(p-2)\left(\langle\xi+y,x\rangle^2\left\|\xi+y\right\|^{p-4}-\langle\xi,x\rangle^2\left\|\xi\right\|^{p-4}\right)\right), \qquad \text{(A.9)}
\end{aligned}$$

for some ξ lying on the line segment strictly between 0 and x. We will now bound the two terms coming from the remainder to eliminate ξ. First, since $2<p\leq 3$, the function $\|\cdot\|^{p-2}$ is subadditive and the first remainder term is bounded by

$$\left\|\xi+y\right\|^{p-2}-\left\|\xi\right\|^{p-2}\leq\left\|y\right\|^{p-2}.$$

Next, we may write $\xi = tx$ for some $0 < t < 1$ since ξ lies on the line segment strictly between 0 and x. Then the second remainder term is bounded by

$$\begin{aligned}
&\langle \xi + y, x\rangle^2 \|\xi + y\|^{p-4} - \langle \xi, x\rangle^2 \|\xi\|^{p-4} \\
&\quad = \langle tx + y, x\rangle^2 \|tx + y\|^{p-4} - \langle tx, x\rangle^2 \|tx\|^{p-4} \\
&\quad = \left\langle \frac{tx+y}{\|tx+y\|}, \frac{x}{\|x\|} \right\rangle^2 \|tx+y\|^{p-2} \|x\|^2 - \left\langle \frac{tx}{\|tx\|}, \frac{x}{\|x\|} \right\rangle^2 \|tx\|^{p-2} \|x\|^2 \\
&\quad \le \|tx + y\|^{p-2} \|x\|^2 - t^{p-2} \|x\|^p \\
&\quad = \|x\|^2 \left(\|tx + y\|^{p-2} - \|tx\|^{p-2} \right) \\
&\quad \le \|x\|^2 \|y\|^{p-2},
\end{aligned}$$

the last inequality coming again by subadditivity. Combining the bounds on the remainder terms and substituting into (A.9) gives

$$\begin{aligned}
\|x+y\|^p - \|x\|^p &\le \|y\|^p + p\langle y, x\rangle \|y\|^{p-2} + \frac{1}{2}\left(p\|x\|^2 \|y\|^{p-2} + p(p-2)\|x\|^2 \|y\|^{p-2} \right) \\
&= \|y\|^p + p\langle y, x\rangle \|y\|^{p-2} + \frac{p(p-1)}{2} \|x\|^2 \|y\|^{p-2}. \quad \text{(A.10)}
\end{aligned}$$

Now let $x \stackrel{def}{=} S_{i-1}$ and $y \stackrel{def}{=} X_i$ in (A.10) and compute the expected value, conditioned on X_i, of each side. Recalling that $\{X_i\}$ are independent mean 0 random vectors, we get

$$\mathbb{E}\left(\|S_i\|^p - \|S_{i-1}\|^p \,|X_i\right) \le \|X_i\|^p + \frac{p(p-1)}{2} \|X_i\|^{p-2} \mathbb{E}\left(\|S_{i-1}\|^2 \,|X_i\right).$$

Finally, take the expected value of both sides and sum from $i = 1$ to n. The left hand side telescopes giving

$$\mathbb{E}\|S_n\|^p \le \sum_{i=1}^{n} \mathbb{E}\|X_i\|^p + \frac{p(p-1)}{2} \sum_{i=1}^{n} \mathbb{E}\|S_{i-1}\|^2 \mathbb{E}\|X_i\|^{p-2}.$$

□

It is known that a Hilbert space is a type 2, and hence type p, Banach space. As an immediate consequence, we have the following theorem. For our application, we desire to precisely know the constant c_p (not depending upon n). Woyczyński (1974, Theorem 1 and Proposition 1) provided a method for determining this constant in concrete situations. We use this procedure and provide his proof below to derive the constant c_p.

Theorem A.2. *Assume that $X_1, \ldots, X_n$ are independent mean zero random vectors taking values in a Hilbert space H. Then for $1 < p \leq 2$,*

$$\mathbb{E}\left\|X_1 + \ldots + X_n\right\|^p \leq c_p \sum_{j=1}^{n} \mathbb{E}\left\|X_j\right\|^p,$$

where $c_p = 2^{3-p} + 1$.

Proof. Let $0 < \alpha \leq 1$ be given by $\alpha = p - 1$. Let $G : H \to H$ be given by

$$G(x) = \|x\|^{\alpha-1} x.$$

Then G satisfies the following three properties

i. $\|G(x)\| = \|x\|^{\alpha-1}\|x\| = \|x\|^{\alpha}$.

ii. $\langle G(x), x\rangle = \left\langle \|x\|^{\alpha-1} x, x\right\rangle = \|x\|^{\alpha-1}\|x\|^2 = \|x\|^{\alpha+1}$.

iii. For every $x, h \in H$,

$$\begin{aligned}
\|G(x+h) - G(x)\| &= \left\| \|x+h\|^{\alpha} \frac{x+h}{\|x+h\|} - \|x\|^{\alpha} \frac{x}{\|x\|} \right\| \\
&\leq \left\| \|x+h\|^{\alpha} \frac{x+h}{\|x+h\|} - \|x\|^{\alpha} \frac{x+h}{\|x+h\|} \right\| \\
&\quad + \left\| \|x\|^{\alpha} \frac{x+h}{\|x+h\|} - \|x\|^{\alpha} \frac{x}{\|x\|} \right\| \\
&= \left| \|x+h\|^{\alpha} - \|x\|^{\alpha} \right| + \|x\|^{\alpha} \left\| \frac{x+h}{\|x+h\|} - \frac{x}{\|x\|} \right\|.
\end{aligned}$$

Since $0 < \alpha \leq 1$, the first term above is subadditive. Also, we may write the last norm as $\left\|\frac{x+h}{\|x+h\|} - \frac{x}{\|x\|}\right\|^{\alpha}\left\|\frac{x+h}{\|x+h\|} - \frac{x}{\|x\|}\right\|^{1-\alpha}$. This last norm is bounded by $2^{1-\alpha}$ by the triangle inequality and continuing we have

$$\begin{aligned}
\|G(x+h) - G(x)\| &\leq \|h\|^{\alpha} + 2^{1-\alpha}\|x\|^{\alpha} \left\| \frac{x+h}{\|x+h\|} - \frac{x}{\|x\|} \right\|^{\alpha} \\
&= \|h\|^{\alpha} + \frac{2^{1-\alpha}}{\|x+h\|^{\alpha}} \left\| (x+h)\|x\| - x\|x+h\| \right\|^{\alpha}.
\end{aligned}$$

Again applying subadditivity to the rightmost norm gives

$$\begin{aligned}\|G(x+h)-G(x)\| &\leq \|h\|^{\alpha}+\frac{2^{1-\alpha}}{\|x+h\|^{\alpha}}\left(\|(x+h)\|x\|-(x+h)\|x+h\|\|^{\alpha}\right.\\ &\quad \left.+\|(x+h)\|x+h\|-x\|x+h\|\|^{\alpha}\right)\\ &=\|h\|^{\alpha}+2^{1-\alpha}\left(|\|x\|-\|x+h\||^{\alpha}+\|h\|^{\alpha}\right).\end{aligned}$$

Applying subadditivity a third time to the absolute value term, we arrive at

$$\begin{aligned}\|G(x+h)-G(x)\| &\leq \|h\|^{\alpha}+2^{1-\alpha}\left(\|h\|^{\alpha}+\|h\|^{\alpha}\right)\\ &=\left(1+2\cdot 2^{1-\alpha}\right)\|h\|^{\alpha}.\end{aligned}$$

Woyczyński (1974, Theorem 1) showed that, for G satisfying properties (i)-(iii), the proof of Theorem A.2 is as follows: By (ii),

$$\|X_1+\cdots+X_n\|^{1+\alpha}=\sum_{j=1}^{n}\langle G(X_1+\cdots+X_n),X_j\rangle.$$

Let $T_j \overset{def}{=} \sum_{i\neq j} X_i$. Then $X_1+\cdots+X_n=T_j+X_j$ and by (iii), $G(X_1+\cdots+X_n)=G(T_j)+x^j$ for some $x^j\in H$ such that $\|x^j\|\leq(1+2\cdot 2^{1-\alpha})\|X_j\|^{\alpha}$. Since $G(T_j)$ and X_j are independent and since X_j are mean 0,

$$\begin{aligned}\mathbb{E}\|X_1+\cdots+X_n\|^{1+\alpha} &=\mathbb{E}\sum_{j=1}^{n}\left(G(T_j)X_j+x^jX_j\right)\\ &\leq\sum_{j=1}^{n}\left(\mathbb{E}G(T_j)\mathbb{E}X_j+\mathbb{E}\left\|x^j\right\|\|X_j\|\right)\\ &\leq(1+2\cdot 2^{1-\alpha})\sum_{j=1}^{n}\mathbb{E}\|X_j\|^{1+\alpha}.\end{aligned}$$

□

When $H=\mathbb{R}$, we can get a better constant than the one obtained in Theorem A.2. We establish this in the next four lemmas and corollaries.

Lemma A.3. *Let $1\leq p\leq 2$. Then for every $x,y\in\mathbb{R}$,*

$$|x+y|^{p}\leq|x|^{p}+p\,\mathrm{sign}(x)\,|x|^{p-1}\,y+c_p\,|y|^{p} \qquad \text{(A.11)}$$

where

$$c_p = \max_{0 \le z \le 1} \left((1-z)^p + p z^{p-1} - z^p \right). \tag{A.12}$$

If $p \geq 2$, then for every $x, y \in \mathbb{R}$,

$$|x+y|^p \geq |x|^p + p \operatorname{sign}(x) |x|^{p-1} y + d_p |y|^p \tag{A.13}$$

where

$$d_p = \min_{0 \le z \le 1} \left((1-z)^p + p z^{p-1} - z^p \right). \tag{A.14}$$

If $p = 2$, then equality holds in (A.11) *and* (A.13).

Proof. If $p = 2$, then equality holds in (A.11). So assume that $1 \leq p < 2$ and consider the function

$$R(z) \stackrel{def}{=} |z+1|^p - |z|^p - p \operatorname{sign}(z) |z|^{p-1}. \tag{A.15}$$

We maximize R over $\mathbb{R}$. First, assume that $z > 0$. Then

$$R(z) = (z+1)^p - z^p - p z^{p-1}$$

and

$$\begin{aligned} R'(z) &= p(z+1)^{p-1} - p z^{p-1} - p(p-1) z^{p-2} \\ &= p \left[\left((z+1)^{p-1} - z^{p-1} \right) - (p-1) z^{p-2} \right]. \end{aligned} \tag{A.16}$$

By the mean value theorem, there exists $\xi \in (z, z+1)$ such that

$$(z+1)^{p-1} - z^{p-1} = (p-1) \xi^{p-2}.$$

Substituting this into (A.16) gives

$$R'(z) = p(p-1) \left(\xi^{p-2} - z^{p-2} \right).$$

Since $p < 2$ and $z < \xi$, $R'(z) < 0$. Therefore R is strictly decreasing on $z > 0$ and we have

$$R(z) < R(0) = 1,$$

showing that

$$\max_{z \in \mathbb{R}} R(z) = \max_{z \leq 0} R(z).$$

So assume now that $z \leq 0$. We observe that

$$\max_{z\leq 0} R(z) = \max_{z\geq 0} R(-z) = \max_{z\geq 0} \left(|1-z|^p - z^p + pz^{p-1}\right) \stackrel{def}{=} \max_{z\geq 0} \tilde{R}(z).$$

Let $z > 1$. On this interval $\tilde{R}(z)$ becomes

$$\tilde{R}(z) = (z-1)^p - z^p + pz^{p-1}. \tag{A.17}$$

By the mean value theorem, there exists $\xi \in (z-1, z)$ such that

$$z^{p-1} - (z-1)^{p-1} = (p-1)\xi^{p-2}. \tag{A.18}$$

Differentiating $\tilde{R}$ and making the substitution (A.18) gives

$$\begin{aligned}
\tilde{R}'(z) &= p(z-1)^{p-1} - pz^{p-1} + p(p-1)z^{p-2} \\
&= p\left[\left((z-1)^{p-1} - z^{p-1}\right) + (p-1)z^{p-2}\right] \\
&= p(p-1)\left(-\xi^{p-2} + z^{p-2}\right).
\end{aligned}$$

Since $p < 2$ and $\xi < z$, $\tilde{R}'(z) < 0$. So $\tilde{R}$ is strictly decreasing on $z > 1$ and hence, $\tilde{R}(z) < \tilde{R}(1) = p-1$ for $z > 1$. Since $p-1 < 1$ and $\tilde{R}(0) = 1$, we see that

$$\max_{z\in\mathbb{R}} R(z) = \max_{z\leq 0} R(z) = \max_{z\geq 0} \tilde{R}(z) = \max_{0\leq z\leq 1} \tilde{R}(z) = c_p. \tag{A.19}$$

We are now ready to prove (A.11). If $y = 0$, then equality holds in (A.11). Assume $y \neq 0$ and make the substitution $z = x/y$ in (A.19) to get

$$\left|\frac{x}{y} + 1\right|^p \leq \left|\frac{x}{y}\right|^p + p\,\mathrm{sign}\left(\frac{x}{y}\right)\left|\frac{x}{y}\right|^{p-1} + c_p.$$

Multiplying by $|y|^p$ gives

$$\begin{aligned}
|x+y|^p &\leq |x|^p + p\frac{\mathrm{sign}(x)}{\mathrm{sign}(y)}\frac{|x|^{p-1}}{|y|^{p-1}}\,\mathrm{sign}(y)y\,|y|^{p-1} + c_p\,|y|^p \\
&= |x|^p + p\,\mathrm{sign}(x)\,|x|^{p-1}\,y + c_p\,|y|^p.
\end{aligned}$$

This proves (A.11). If $p > 2$, the argument above carries over almost verbatim, with inequalities switched and max replaced by min, to show the reverse inequality of (A.11) holds with constant d_p. $\square$

Lemma A.4. *Let c_p and d_p be the constants given above. Then*

$$1 \leq c_p \leq \tilde{c}_p \leq 3 - p \tag{A.20}$$

and

$$\frac{1}{2^{p-1}} \leq d_p \leq 1, \tag{A.21}$$

where $\tilde{c}_p = 1 + p(p-1)^{\frac{p-1}{2-p}}(2-p)$ for $1 < p < 2$, $\tilde{c}_1 \stackrel{def}{=} \lim_{p \to 1+} \tilde{c}_p = 2$, and $\tilde{c}_2 \stackrel{def}{=} \lim_{p \to 2-} \tilde{c}_p = 1$.

Proof. Let $\tilde{R}$ be as above. Since $\tilde{R}(0) = 1$, $c_p \geq 1$. To obtain the upper bounds for c_p, let $0 \leq z \leq 1$ and $1 < p < 2$. If $\frac{1}{2} < z \leq 1$, then by the mean value theorem, there exists ξ between $1 - z$ and z such that

$$\begin{aligned}\tilde{R}(z) &= (1-z)^p - z^p + pz^{p-1} \\ &= p\xi^{p-1}(1-2z) + pz^{p-1}.\end{aligned}$$

Since $z < 1$, $1 - z < 1$, and ξ is between z and $1 - z$,

$$\tilde{R}(z) \leq p(1-2z) + pz^{p-1}.$$

Since $z > 1/2$,

$$\tilde{R}(z) \leq 1 - pz + pz^{p-1}. \tag{A.22}$$

If $0 < z \leq \frac{1}{2}$, then $\frac{1}{z} \geq 2$ and by Taylor's theorem, there exists $\xi \in \left(\frac{1}{z} - 1, \frac{1}{z}\right)$ such that

$$\left(\frac{1}{z} - 1\right)^p = \left(\frac{1}{z}\right)^p - p\left(\frac{1}{z}\right)^{p-1} + \frac{p(p-1)}{2}\xi^{p-2}.$$

Since $p < 2$ and $\xi > 1$,

$$\left(\frac{1}{z} - 1\right)^p \leq \left(\frac{1}{z}\right)^p - p\left(\frac{1}{z}\right)^{p-1} + 1.$$

Multiplying by z^p gives

$$(1-z)^p \leq 1 - pz + z^p.$$

Therefore

$$\tilde{R}(z) = (1-z)^p - z^p + pz^{p-1} \le 1 - pz + pz^{p-1} = 1 + p\left(z^{p-1} - z\right). \tag{A.23}$$

In either case, if $0 \le z \le 1$, then (A.22) and (A.23) show

$$c_p = \max_{0\le z\le 1} \tilde{R}(z) \le \max_{0\le z\le 1} \left(1 + p\left(z^{p-1} - z\right)\right). \tag{A.24}$$

Standard calculus shows that $z^{p-1} - z$ attains a maximum value at $(p-1)^{\frac{1}{2-p}}$ on the interval $[0,1]$. Using this in (A.24) gives

$$c_p \le 1 + p\left((p-1)^{\frac{p-1}{2-p}} - (p-1)^{\frac{1}{2-p}}\right) = 1 + p(p-1)^{\frac{p-1}{2-p}}(2-p).$$

To show the last inequality in (A.20), we observe

$$1 + p(p-1)^{\frac{p-1}{2-p}}(2-p) \le 3 - p$$

if and only if

$$p(p-1)^{\frac{p-1}{2-p}}(2-p) \le 2 - p$$

if and only if

$$p(p-1)^{\frac{p-1}{2-p}} \le 1$$

if and only if

$$\ln p + \frac{p-1}{2-p}\ln(p-1) \le 0$$

if and only if

$$(2-p)\ln p + (p-1)\ln(p-1) \le 0. \tag{A.25}$$

Now the left hand side of (A.25) is a convex combination of the function $x \mapsto \ln x$. Since $x \mapsto \ln x$ is concave,

$$\begin{aligned}(2-p)\ln p + (p-1)\ln(p-1) &\le \ln\left((2-p)p + (p-1)^2\right)\\ &= \ln\left(2p - p^2 + p^2 - 2p + 1\right)\\ &= \ln 1\\ &= 0.\end{aligned}$$

Finally, to bound d_p, first observe that $\tilde{R}(0) = 1$ and hence, $d_p \le 1$. To prove the lower bound, let $0 \le z \le 1$ and $p > 2$. Then

$$\begin{aligned}\tilde{R}(z) = (1-z)^p - z^p + pz^{p-1} = (1-z)^p + (p-z)z^{p-1} \ge (1-z)^p + z^{p-1} \\ \ge (1-z)^p + z^p = (1-z)(1-z)^{p-1} + zz^{p-1}\end{aligned}$$

is a convex combination of the function $x \mapsto x^{p-1}$. Since $x \mapsto x^{p-1}$ is convex, we have

$$\tilde{R}(z) \ge \left((1-z)^2 + z^2\right)^{p-1} = \left(1 - 2z + 2z^2\right)^{p-1}. \tag{A.26}$$

Now the right hand side of (A.26) is minimized whenever $1 - 2z + 2z^2$ is minimized, that is, at $z = 1/2$ with a value of $1/2$. Therefore,

$$d_p = \min_{0 \le z \le 1} \tilde{R}(z) \ge \frac{1}{2^{p-1}}.$$

□

Corollary A.5. *Let $\{S_n\}_{n=0}^{\infty}$ be an $(\mathcal{F}_n)_{n=0}^{\infty}$-martingale, let $1 \le p \le 2$, and let $S_n \in L^p$ for every n. Then*

$$\mathbb{E}\,|S_n|^p - \mathbb{E}\,|S_0|^p \le c_p \sum_{k=1}^{n} \mathbb{E}\,|S_k - S_{k-1}|^p\,. \tag{A.27}$$

If $p > 2$, then the reverse inequality holds for the constant d_p. That is,

$$\mathbb{E}\,|S_n|^p - \mathbb{E}\,|S_0|^p \ge d_p \sum_{k=1}^{n} \mathbb{E}\,|S_k - S_{k-1}|^p\,. \tag{A.28}$$

Proof. Since (S_n) is a martingale, $\mathbb{E}\left(S_k - S_{k-1} | \mathcal{F}_{k-1}\right) = 0$. For every $\omega \in \Omega$, Lemma A.3 then gives

$$\begin{aligned}|S_k(\omega)|^p &= |S_{k-1}(\omega) + (S_k(\omega) - S_{k-1}(\omega))|^p \\ &\le |S_{k-1}(\omega)|^p + p\,\mathrm{sign}\left(S_{k-1}(\omega)\right)|S_{k-1}(\omega)|^{p-1}\left(S_k(\omega) - S_{k-1}(\omega)\right) \\ &\quad + c_p\,|S_k(\omega) - S_{k-1}(\omega)|^p\,.\end{aligned}$$

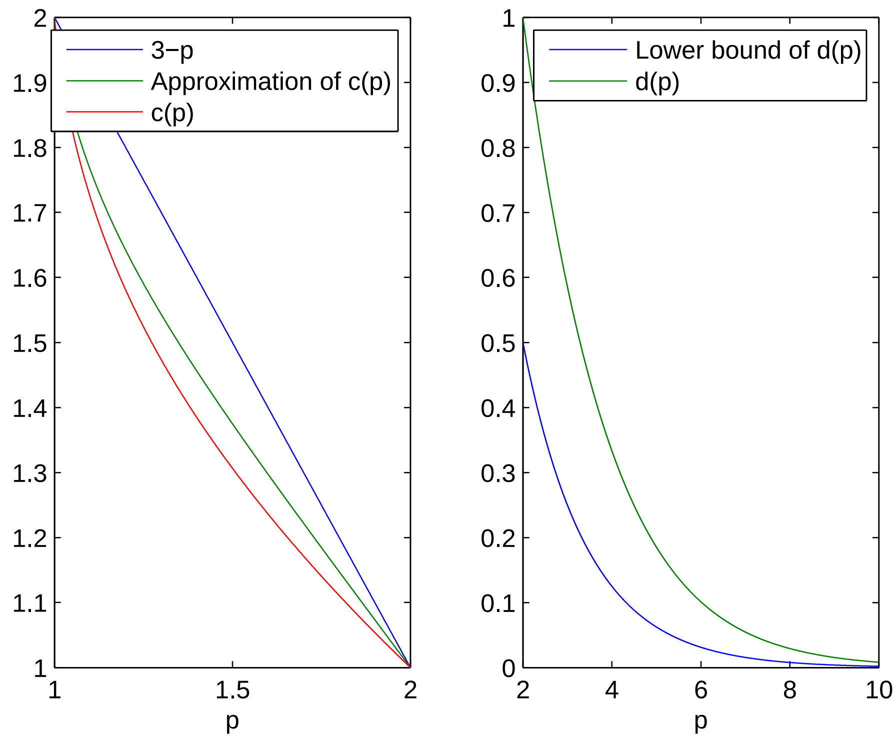

Figure A.1: Graph and approximations of c_p and d_p.

Conditioning on $\mathcal{F}_{k-1}$ and taking the expected value gives

$$\begin{aligned}
&\mathbb{E}\left(|S_k|^p \,|\mathcal{F}_{k-1}\right)\\
&\quad\leq \mathbb{E}\left(|S_{k-1}|^p \,|\mathcal{F}_{k-1}\right) + \mathbb{E}\left(p\,\mathrm{sign}\left(S_{k-1}\right)|S_{k-1}|^{p-1}\left(S_k - S_{k-1}\right)|\mathcal{F}_{k-1}\right)\\
&\qquad + c_p\mathbb{E}\left(|S_k - S_{k-1}|^p \,|\mathcal{F}_{k-1}\right)\\
&\quad= |S_{k-1}|^p + p\,\mathrm{sign}\left(S_{k-1}\right)|S_{k-1}|^{p-1}\,\mathbb{E}\left(S_k - S_{k-1}|\mathcal{F}_{k-1}\right) + c_p\mathbb{E}\left(|S_k - S_{k-1}|^p \,|\mathcal{F}_{k-1}\right)\\
&\quad= |S_{k-1}|^p + c_p\mathbb{E}\left(|S_k - S_{k-1}|^p \,|\mathcal{F}_{k-1}\right).
\end{aligned}$$

Taking the expected value gives, for every k,

$$\mathbb{E}\,|S_k|^p - \mathbb{E}\,|S_{k-1}|^p \leq c_p\mathbb{E}\,|S_k - S_{k-1}|^p\,. \tag{A.29}$$

Sum (A.29) from $k = 1$ to n. The left hand side of (A.29) telescopes, giving (A.27).
□

Corollary A.6. *Let X_n be independent mean 0 random variables and $1 \le p \le 2$. Then*

$$\mathbb{E}\left|\sum_{i=1}^{n} X_i\right|^p \le c_p \sum_{i=1}^{n} \mathbb{E}\left|X_i\right|^p . \tag{A.30}$$

If $p > 2$, then the reverse inequality holds with constant d_p. That is,

$$\mathbb{E}\left|\sum_{i=1}^{n} X_i\right|^p \ge d_p \sum_{i=1}^{n} \mathbb{E}\left|X_i\right|^p . \tag{A.31}$$

Proof. If any $X_i \notin L^p$, (A.30) is trivial. So assume $X_n \in L^p$ for every n. Let $S_0 \overset{def}{=} 0$ and $S_n \overset{def}{=} \sum_{i=1}^{n} X_i$. Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(X_k : k \le n)$. Then $\{S_n\}_{n=0}^{\infty}$ is an $(\mathcal{F}_n)_{n=0}^{\infty}$-martingale and by Corollary A.5,

$$\mathbb{E}\left|\sum_{i=1}^{n} X_i\right|^p \le c_p \sum_{i=1}^{n} \mathbb{E}\left|X_i\right|^p . \tag{A.32}$$

□

The next lemma gives control over lower limits of $\mathbb{E}\left\|S_n\right\|^p$. Most often, this lemma is used with $J \equiv \{1\}$ when the summands are i.i.d. random vectors.

Lemma A.7. *Let $p \ge 1$ and let $X_1, X_2, \dots, X_n \in L^p$ be independent mean 0 random variables. Then for every $J \subset \{1, 2, \dots, n\}$,*

$$\mathbb{E}\left\|\sum_{i=1}^{n} X_i\right\|^p \ge \mathbb{E}\left\|\sum_{J} X_i\right\|^p . \tag{A.33}$$

In particular,

$$\mathbb{E}\left\|\sum_{i=1}^{n} X_i\right\|^p \ge \mathbb{E} \max_{J \subset \{1,2,\dots,n\}} \left\|\sum_{J} X_i\right\|^p . \tag{A.34}$$

Proof. Let $J \subset \{1, 2, \dots, n\}$ and put $\mathcal{A} \overset{def}{=} \sigma\{X_j : j \in J\}$. Then by Jensen's inequality,

$$\mathbb{E}\left\|\sum_{i=1}^{n} X_i\right\|^p \ge \mathbb{E}\left\|\mathbb{E}\left(\sum_{i=1}^{n} X_i|\mathcal{A}\right)\right\|^p = \mathbb{E}\left\|\mathbb{E}\left(\sum_{i\in J} X_i|\mathcal{A}\right) + \mathbb{E}\left(\sum_{i\in J^c} X_i|\mathcal{A}\right)\right\|^p .$$

Since $\{X_i\}_{i \in J^c}$ are mean 0 and independent of $\mathcal{A}$,

$$\mathbb{E}\left(\sum_{i \in J^c} X_i | \mathcal{A}\right) = 0$$

and since $\{X_i\}_{i \in J}$ are $\mathcal{A}$-measurable,

$$\mathbb{E}\left(\sum_{i \in J} X_i | \mathcal{A}\right) = \sum_{i \in J} X_i,$$

giving the result. □

The previous lemma has as an application the following corollary allowing one to establish results concerning mean 0 random vectors from results concerning symmetric random vectors.

Corollary A.8. *Let $p \geq 1$, let X be a mean 0 random vector, and let $X^s \stackrel{def}{=} X - X'$, where X' is an independent copy of X, be the standard symmetrization of X. Then*

$$\mathbb{E}\left\|X\right\|^p \leq \mathbb{E}\left\|X^s\right\|^p \leq 2^p \mathbb{E}\left\|X\right\|^p. \tag{A.35}$$

Proof. The upper bound follows from Minkowski's inequality and the lower bound follows from Lemma A.7 with $J \equiv \{1\}$. □

Finally, the following lemma and theorems were established by Latała (1997). The theorems hold for real-valued random variables. It is not known if these results hold for random vectors taking values in a Hilbert space. If $H = \mathbb{R}$, one may use these theorems to obtain better constants in Theorem 1.2.1 whenever $p > 3$. Since the results are restricted to the real valued case, they are excluded from the proof of Theorem 1.2.1 in favor of results pertaining to Hilbert space valued random vectors (in particular, the Hoffman-Jorgensen inequality).

Lemma A.9. *For every $x \geq 0$ and for every $p \geq 1$,*

$$(1 + x)^p \leq \sum_{0 \leq k < p} \binom{p}{k} x^k + x^p. \tag{A.36}$$

For every $x \in \mathbb{R}$ and for every $p > 2$,

$$|1+x|^p \le \sum_{0 \le k < p-1} \binom{p}{k} x^k + \binom{p}{\lceil p-1 \rceil} |x|^{\lceil p-1 \rceil} + |x|^p . \tag{A.37}$$

Proof. First, let $p \ge 1$. For $x > 0$, Taylor's theorem applied to the Maclaurin series of the function $f(x) \stackrel{def}{=} (1+x)^p - x^p$ gives

$$(1+x)^p - x^p = \sum_{0 \le k < p} \binom{p}{k} x^k + R(x),$$

where

$$R(x) = \frac{f^{(\lceil p \rceil)}(\xi)}{\lceil p \rceil !} x^{\lceil p \rceil}$$

for some $0 < \xi < x$. But

$$R(x) = \binom{p}{\lceil p \rceil} \left((1+\xi)^{p - \lceil p \rceil} - \xi^{p - \lceil p \rceil} \right) x^{\lceil p \rceil} \le 0$$

and hence

$$(1+x)^p - x^p \le \sum_{0 \le k < p} \binom{p}{k} x^k,$$

proving (A.36)

To prove (A.37), let $p > 2$. If $x \ge 0$, then (A.36) is true by (A.36). So suppose $x < 0$. Again by Taylor's theorem applied to the Maclaurin series of the function $f(x) \stackrel{def}{=} |1+x|^p - |x|^p$ gives

$$|1+x|^p - |x|^p = \sum_{0 \le k < p-1} \binom{p}{k} x^k + R(x),$$

where

$$\begin{aligned} R(x) &= \frac{f^{(\lceil p-1 \rceil)}(\xi)}{\lceil p-1 \rceil !} x^{\lceil p-1 \rceil} \\ &= \binom{p}{\lceil p-1 \rceil} \left[\operatorname{sign}(1+\xi)^{\lceil p-1 \rceil} |1+\xi|^{p - \lceil p-1 \rceil} - (-1)^{\lceil p-1 \rceil} |\xi|^{p - \lceil p-1 \rceil} \right] x^{\lceil p-1 \rceil} \end{aligned}$$

for some $x < \xi < 0$. If $\lceil p-1 \rceil$ is even, then

$$R(x) = \binom{p}{\lceil p-1 \rceil} \left[|1+\xi|^{p-\lceil p-1 \rceil} - |\xi|^{p-\lceil p-1 \rceil} \right] |x|^{\lceil p-1 \rceil} \leq \binom{p}{\lceil p-1 \rceil} |x|^{\lceil p-1 \rceil}$$

and if $\lceil p-1 \rceil$ is odd, then

$$R(x) = \binom{p}{\lceil p-1 \rceil} \left[\operatorname{sign}(1+\xi)\, |1+\xi|^{p-\lceil p-1 \rceil} + |\xi|^{p-\lceil p-1 \rceil} \right] (-1)\, |x|^{\lceil p-1 \rceil} \leq 0.$$

In either case, we have

$$|1+x|^p - |x|^p = \sum_{0 \leq k < p-1} \binom{p}{k} x^k + \binom{p}{\lceil p-1 \rceil} |x|^{\lceil p-1 \rceil}.$$

□

The following theorem was proved by Latała (1997, Corollary 1) using Lemma A.9. The constant in Theorem A.10 is easily observed from Latała (1997, Corollary 2).

Theorem A.10. *If $p \geq 1$ and $X, X_1, \ldots, X_n$ are i.i.d. nonnegative random variables, then*

$$\|X_1 + \cdots + X_n\|_p \leq 2e^2 \sup \left\{ \frac{p}{s} \left(\frac{n}{p} \right)^{\frac{1}{s}} \|X\|_s : \max\left(1, \frac{p}{n}\right) \leq s \leq p \right\}.$$

Again, using Lemma A.9, one can prove the following theorem of Latała (1997, Corollary 2).

Theorem A.11. *If $p > 2$ and $X, X_1, \ldots, X_n$ are i.i.d. symmetric random variables, then*

$$\|X_1 + \cdots + X_n\|_p \leq 2e^2 \sup \left\{ \frac{p}{s} \left(\frac{n}{p} \right)^{\frac{1}{s}} \|X\|_s : \max\left(2, \frac{p}{n}\right) \leq s \leq p \right\}$$

I provide an alternate proof with worse constants.

Proof. Let $\{\varepsilon_i\}_{i=1}^n$ be a sequence of i.i.d Rademacher random variables independent of $\{X_i\}_{i=1}^n$. Then

$$\mathbb{E}\, |X_1 + \cdots + X_n|^p = \mathbb{E}\mathbb{E}_\varepsilon\, |\varepsilon_1 X_1 + \cdots + \varepsilon_n X_n|^p, \tag{A.38}$$

where $\mathbb{E}_\varepsilon$ denotes integration of the random variables ε_i. Now for fixed ω, the $Khintchine - Kahane$ inequalities (see e.g. de la Peña and Giné (1999, Theorem 1.3.1)) gives

$$\begin{aligned}
\mathbb{E}_\varepsilon \left|\varepsilon_1 X_1(\omega) + \cdots + \varepsilon_n X_n(\omega)\right|^p &\leq \left(\sqrt{p-1}\left(\mathbb{E}_\varepsilon \left|\sum_{i=1}^n \varepsilon_i X_i(\omega)\right|^2\right)^{1/2}\right)^p \\
&= \left(\sqrt{p-1}\left(\sum_{i=1}^n X_i(\omega)^2\right)^{1/2}\right)^p \\
&= \left(\sum_{i=1}^k (p-1) X_i(\omega)^2\right)^{p/2}
\end{aligned}$$

and combining with (A.38) gives

$$\|X_1 + \cdots X_n\|_p \leq \left(\mathbb{E}\left(\sum_{i=1}^k (p-1) X_i(\omega)^2\right)^{p/2}\right)^{1/p}. \tag{A.39}$$

Now $\{(p-1)X_i^2\}_{i=1}^n$ is a sequence of nonnegative i.i.d. random variables and hence, by Theorem A.10,

$$\begin{aligned}
\|\sum_{i=1}^n (p-1) X_i^2\|_{p/2} &\leq 2e^2 \sup\{\frac{p}{2s}\left(\frac{2n}{p}\right)^{1/s} \|(p-1)X^2\|_s : \max\left(1, \frac{p}{2n}\right) \leq s \leq \frac{p}{2}\} \\
&= 2e^2 \sup\{\frac{p}{2s}\left(\frac{2n}{p}\right)^{1/s} (p-1)\|X^2\|_s : \max\left(2, \frac{p}{n}\right) \leq 2s \leq p\} \\
&= 2e^2 \sup\{\frac{p}{t}\left(\frac{2n}{p}\right)^{2/t} (p-1)\|X^2\|_{t/2} : \max\left(2, \frac{p}{n}\right) \leq t \leq p\} \\
&= 2e^2 \sup\{\frac{p}{t}\left(\frac{2n}{p}\right)^{2/t} (p-1)\|X\|_t^2 : \max\left(2, \frac{p}{n}\right) \leq t \leq p\} \\
&= 2e^2 \frac{p}{t_0}\left(\frac{2n}{p}\right)^{2/t_0} (p-1)\|X\|_{t_0}^2
\end{aligned}$$

for some $\max\left(2, \frac{p}{n}\right) \leq t_0 \leq p$. Combining with (A.39) gives

$$\begin{aligned}
\|X_1 + \cdots + X_n\|_p &\leq \sqrt{2e^2}\sqrt{\frac{p}{t_0}\left(\frac{2n}{p}\right)^{1/t_0}}\sqrt{p-1}\|X\|_{t_0} \\
&= \sqrt{2}e\sqrt{p-1}\left(\frac{p}{t_0}\right)^{-1/2} 2^{1/t_0}\left\{\frac{p}{t_0}\left(\frac{n}{p}\right)^{1/t_0}\|X\|_{t_0}\right\} \\
&\leq 2e\sqrt{p-1}\sup\left\{\frac{p}{s}\left(\frac{n}{p}\right)^{1/2}\|X\|_s : \max\left(2, \frac{p}{n}\right) \leq s \leq p\right\}.
\end{aligned}$$

□

Appendix B

Modular Spaces

The following is a brief overview of Modular Spaces. I present a few useful basic facts and theorems taken from Rolewicz (1972, Chapter 1) and Kwapień and Woyczyński (1992, Chapter 0). Let X be a linear space.

Definition B.1. A function $\|\cdot\| : X \to [0, \infty]$ is an *F-norm on X* (or simply an F-norm) if $\|\cdot\|$ satisfies the following properties:

i. $\|x\| = 0$ if and only if $x = 0$,

ii. $\|\alpha x\| = \|x\|$ for every α such that $|\alpha| = 1$, and

iii. $\|x + y\| \leq \|x\| + \|y\|$ for every $x, y \in X$.

Definition B.2. A function $\rho : X \to [0, \infty]$ is a *modular on X* (or simply a modular) if ρ satisfies the following properties:

i. $\rho(x) = 0$ if and only if $x = 0$,

ii. $\rho(\alpha x) = \rho(x)$ for every α such that $|\alpha| = 1$,

iii. $\rho(\alpha x + \beta y) \leq \rho(x) + \rho(y)$ for every $\alpha, \beta \geq 0$ such that $\alpha + \beta = 1$,

iv. $\rho(\alpha_n x) \to 0$ if $\alpha_n \to 0$ and $\rho(x) < \infty$, and

v. $\rho(\alpha x_n) \to 0$ if $\rho(x_n) \to 0$.

An F-norm need not be a modular. But if

$$\|\alpha x\| \leq \|x\| \text{ for every } 0 \leq \alpha \leq 1,$$

then the F-norm $\|\cdot\|$ is a modular. The following theorem show the relationships between F-norms and modulars.

Theorem B.3 (F-norm to Modular)**.** *Rolewicz (1972, Theorem I.2.2)]. If $\|\cdot\|$ is an F-norm on X, then the F-norm $\|\cdot\|'$ defined by*

$$\|x\|' \stackrel{def}{=} \sup_{0\leq t\leq 1} \|tx\|$$

is equivalent to $\|\cdot\|$ and is a modular on X.

Theorem B.4 (Modular to F-norm)**.** *[Rolewicz (1972, Theorem I.2.3)]. Let X be a linear space with modular $\rho(x)$. Let $X^\rho \stackrel{def}{=} \{x \in X : \rho(kx) < \infty$ for some $k > 0\}$. Then*

$$\|x\| \stackrel{def}{=} \inf \left\{c > 0 : \rho\left(c^{-1}x\right) < c\right\}$$

is an F-norm on X^ρ such that

$$\|x_n\| \to 0 \text{ if and only if } \rho(x_n) \to 0.$$

Finally, we present an easy way to generate a modular. Let $(S, \mathcal{S}, \mu)$ be a σ-finite measure space and let L^0 be the space of all measurable maps $x : (S, \mathcal{S}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Let $N : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous, non-decreasing function satisfying the Δ_2 condition such that $N(u) = 0$ if and only if $u = 0$. Then

$$\rho_N(x) \stackrel{def}{=} \int_S N(|x(t)|)\mu(dt)$$

defines a modular on L^0 and by Theorem B.4,

$$\|x\|_{\rho_N} \stackrel{def}{=} \inf \left\{c > 0 : \int_S N(c^{-1}\,|x(t)|)\mu(dt) < c\right\}$$

is an F-norm on $L^{\rho_N} \stackrel{def}{=} \{x \in L^0 : \rho_N(kx) < \infty \text{ for some } k > 0\}$. The modular space L^{ρ_N} is called an *Orlicz space.* Moreover, if N is convex, then the so called *Orlicz norm*

$$\|x\|_{\rho_N} \stackrel{def}{=} \inf \left\{c > 0 : \rho_N\left(c^{-1}x\right) = \int_S N(c^{-1}\,|x(t)|)\mu(dt) \leq 1\right\}$$

is a norm on L^{ρ_N}.

Example B.5. i. As an example, let $N(u) \stackrel{def}{=} u^p$ for $0 < p < \infty$. If $p \geq 1$, let c solve

$$\int_S c^{-p} |x(t)|^p \mu(dt) = 1.$$

Then $c = \left(\int_S |x(t)|^p \mu(dt)\right)^{1/p}$ and hence, $\|\cdot\|_{\rho_N} = \|\cdot\|_{L^p}$ is a norm on $L^p(S, \mathcal{S}, \mu)$. If $0 < p < 1$, then letting c solve

$$\int_S c^{-p} |x(t)|^p \mu(dt) = c$$

gives $\|\cdot\|_{\rho_N} = \|\cdot\|_{L^p}^{p+1}$ is an F-norm on $L^p(S, \mathcal{S}, \mu)$. This shows that Orlicz spaces generalize L^p spaces.

ii. Assume that $\mu(S) < \infty$ and let $N(u) = u \wedge 1$. Then

$$\rho_N(x) \stackrel{def}{=} \int_S |x(t)| \wedge 1 \, \mu(dt) < \infty$$

is a modular and F-norm on L^0. The modular space L^0 is often denoted $L^0(S, \mathcal{S}, \mu)$ and the modular ρ_N denoted $\|\cdot\|_0$. It is the space of all measurable maps and $\|x_n\|_0 \to 0$ if and only if $x_n \to 0$ in μ-measure.

iii. Kwapień and Woyczyński (1992, Section 0.8). Let $\Phi : S \times \mathbb{R}_+ \to \mathbb{R}_+$ be such that

a. For every $s \in S$, $\Phi(s, \cdot)$ is a continuous non-decreasing function on $\mathbb{R}_+$ with $\Phi(s, 0) = 0$,

b. For every $y \in \mathbb{R}_+$, $\Phi(\cdot, y)$ is $\mathcal{S}$-measurable, and

c. For every $s \in S$, $\Phi(s, \cdot)$ satisfies the Δ_2 condition (see Definition C.1.4).

Then

$$\rho_\Phi(f) \stackrel{def}{=} \int_S \Phi(s, |f(s)|)\mu(ds)$$

is a modular on the space

$$L^\Phi(S, \mathcal{S}, \mu) = L^\Phi \stackrel{def}{=} \left\{ f \in L^0 : \rho_\Phi(f) < \infty \right\}.$$

The modular space L^Φ is called a *Musielak-Orlicz space*. Similar to the above

$$\|f\|_\Phi \stackrel{def}{=} \inf\left\{c > 0 : \rho_\Phi\left(c^{-1} f\right) \leq c\right\}$$

is an F-norm on L^Φ and is called an *Musielak-Orlicz* F-norm.

Appendix C

Selected Prerequisite Analysis Results

We present a few basic, but possibly less known, facts from probability theory and real analysis.

C.1 Convergence results

Theorem C.1.1. *[See e.g. Ledoux and Talagrand (1991, Theorem 6.1)]. Let* $\{X_n\}_{n\in\mathbb{N}}$ *be a sequence of independent random variables taking values in a separable Banach space. Let* $S_n \overset{def}{=} \sum_{i=1}^{n} X_i$*. The following are equivalent:*

i. The sequence $\{S_n\}_{n\in\mathbb{N}}$ *converges almost surely.*

ii. The sequence $\{S_n\}_{n\in\mathbb{N}}$ *converges in probability.*

iii. The sequence $\{S_n\}_{n\in\mathbb{N}}$ *converges in distribution.*

Theorem C.1.2. *Rudin (1987, thm. 7.10). Associate to each* $x \in \mathbb{R}^d$ *a sequence* $\{E_n(x)\}$ *with the following property: there is a number* $\alpha > 0$ *and a sequence of balls* $B(x, r_n) \supset E_n$ *with* $r_n \to 0$ *such that*

$$\lambda(E_n) \geq \alpha \lambda(B(x, r_n))$$

for each $n = 1, 2, 3, \ldots$ *. If* $f \in L^1(\mathbb{R}^d)$*, then*

$$f(x) = \lim_{n\to\infty} \frac{1}{\lambda(E_n(x))} \int_{E_n(x)} f(y)\lambda(dy) \text{ for } \lambda\text{-a.e. } x.$$

Lemma C.1.3. $X_n \xrightarrow{\mathbb{P}} X$ *if and only if for every subsequence* $\{n_k\} \subset \mathbb{N}$, *there exists a sub-subsequence* $\{n_{k_l}\} \subset \{n_k\}$ *such that* $X_{n_{k_l}} \to X$ *a.s.*

Proof. ($\Rightarrow$) Let $X_n \xrightarrow{\mathbb{P}} X$ and let $\{n_k\}$ be a subsequence. Then $X_{n_k} \xrightarrow{\mathbb{P}} X$ and by Jacob and Protter (2004, Theorem 17.3), there exists a sub-subsequence n_{k_l} such that

$$X_{n_{k_l}} \to X \text{ a.s.}$$

($\Leftarrow$) Let $\varepsilon > 0$ and $\{n_k\}$ be a subsequence. Then there exists a sub-subsequence $\{n_{k_l}\}$ such that $X_{n_{k_l}} \to X$ a.s. Therefore $X_{n_{k_l}} \xrightarrow{\mathbb{P}} X$ and hence, the sequence of numbers $\mathbb{P}\left(\left|X_{n_{k_l}} - X\right| > \varepsilon\right) \to 0$. So every subsequence of numbers $\mathbb{P}\left(|X_{n_k} - X| > \varepsilon\right)$ has a sub-subsequence of numbers $\mathbb{P}\left(\left|X_{n_{k_l}} - X\right| > \varepsilon\right)$ that converge to 0. Therefore

$$\mathbb{P}\left(|X_n - X| > \varepsilon\right) \to 0.$$

□

Definition C.1.4. Let $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ be a non-decreasing function.

i. The function φ is of *moderate growth* if for some $C_1 > 0$ and any $x, y \in \mathbb{R}_+$,

$$\varphi(x+y) \leq C_1(\varphi(x) + \varphi(y)).$$

ii. The function φ satisfies the Δ_2 *condition* if for some $C_2 > 0$ and any $x, y \in \mathbb{R}_+$,

$$\varphi(2x) \leq C_2\varphi(x).$$

Theorem C.1.5. *φ is of moderate growth if and only if φ satisfies Δ_2.*

Proof. Item i $\Rightarrow$ Item ii. Suppose there exists $C_1 > 0$ such that

$$\varphi(x+y) \leq C_1(\varphi(x) + \varphi(y))$$

for any $x, y \in \mathbb{R}_+$. Let $x \in \mathbb{R}_+$. Then

$$\varphi(2x) = \varphi(x+x) \leq C_1(\varphi(x) + \varphi(x)) = C_2\varphi(x),$$

where $C_2 = 2C_1$.

Item ii $\Rightarrow$ Item i. Now suppose there is a $C_2 > 0$ such that

$$\varphi(2x) \leq C_2\varphi(x)$$

for any $x, y \in \mathbb{R}_+$. Let $x, y \in \mathbb{R}_+$ with $x < y$. Then, since φ is non-decreasing and non-negative,

$$\varphi(x+y) \leq \varphi(2y) \leq C_2\varphi(y) \leq C_2(\varphi(y) + \varphi(x)).$$

□

C.2 Algebras

Theorem C.2.1. *Let $(S, \mathcal{A}, \mu)$ be a finite measure space and let $\mathcal{A}_0$ be an algebra generating $\mathcal{A}$. Then for every $\varepsilon > 0$ and for every $A \in \mathcal{A}$, there exists an $A_\varepsilon \in \mathcal{A}_0$ such that*

$$\mu\left(A \triangle A_\varepsilon\right) \stackrel{def}{=} \mu\left((A \setminus A_\varepsilon) \bigcup (A_\varepsilon \setminus A)\right) < \varepsilon.$$

Proof. Let $\varepsilon > 0$ and define

$$\mathcal{B} \stackrel{def}{=} \{A \in \mathcal{A} : \mu\left(A \triangle A_\varepsilon\right) < \varepsilon \text{ for some } A_\varepsilon \in \mathcal{A}_0.\}$$

Obviously $\mathcal{B}$ contains $\mathcal{A}_0$. We show $\mathcal{B}$ is a monotone class. First, let $A, B \in \mathcal{B}$ with $A \subset B$. Then there exists $A_\varepsilon, B_\varepsilon \in \mathcal{A}_0$ such that

$$\mu\left(A \triangle A_\varepsilon\right) + \mu\left(B \triangle B_\varepsilon\right) < \varepsilon.$$

Since $\mathcal{A}_0$ is an algebra, $B_\varepsilon \setminus A_\varepsilon \in \mathcal{A}_0$ and we have

$$\begin{aligned}
(B \setminus A) \triangle (B_\varepsilon \setminus A_\varepsilon) &= (B \cap A^c) \triangle (B_\varepsilon \cap A_\varepsilon^c) \\
&= [(B \cap A^c) \cap (B_\varepsilon \cap A_\varepsilon^c)^c] \cup [(B_\varepsilon \cap A_\varepsilon^c) \cap (B \cap A^c)^c] \\
&= [(B \cap A^c) \cap (B_\varepsilon^c \cup A_\varepsilon)] \cup [(B_\varepsilon \cap A_\varepsilon^c) \cap (B^c \cup A)] \\
&= [(B \cap A^c) \cap B_\varepsilon^c] \cup [(B \cap A^c) \cap A_\varepsilon] \\
&\quad \cup [(B_\varepsilon \cap A_\varepsilon^c) \cap B^c] \cup [(B_\varepsilon \cap A_\varepsilon^c) \cap A] \\
&\subset (B \cap B_\varepsilon^c) \cup (A^c \cap A_\varepsilon) \cup (B_\varepsilon \cap B^c) \cup (A_\varepsilon^c \cap A)
\end{aligned}$$

$$= (B \setminus B_\varepsilon \cup B_\varepsilon \setminus B) \cup (A \setminus A_\varepsilon \cup A_\varepsilon \setminus A)$$
$$= (B \triangle B_\varepsilon) \cup (A \triangle A_\varepsilon).$$

Therefore

$$\mu\left((B \setminus A) \triangle (B_\varepsilon \setminus A_\varepsilon)\right) \leq \mu(B \triangle B_\varepsilon) + \mu(A \triangle A_\varepsilon) < \varepsilon,$$

showing that $B \setminus A \in \mathcal{B}$ and hence, $\mathcal{B}$ is closed under differences.

Second, let $A_1 \subset A_2 \subset A_3 \subset \cdots$ be a sequence of events in $\mathcal{B}$. Since S is a finite space and

$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right) < \infty,$$

we can choose $n > 0$ so large that

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \setminus A_n = \mu\left(\bigcup_{i=n+1}^{\infty} A_i\right) < \frac{\varepsilon}{2}.$$

Fix such n. Since $A_n \in \mathcal{B}$, there exists $A_n^\varepsilon \in \mathcal{A}_0$ such that

$$\mu(A_n \triangle A_n^\varepsilon) < \frac{\varepsilon}{2}.$$

Now

$$\bigcup_{i=1}^{\infty} A_i \triangle A_n^\varepsilon = \left[\left(\bigcup_{i=1}^{\infty} A_i\right) \setminus A_n^\varepsilon\right] \cup \left[A_n^\varepsilon \setminus \left(\bigcup_{i=1}^{\infty} A_i\right)\right]$$
$$= \left[\left(\bigcup_{i=1}^{\infty} A_i\right) \cap (A_n^\varepsilon)^c\right] \cup \left[A_n^\varepsilon \cap \left(\bigcup_{i=1}^{\infty} A_i\right)^c\right]$$
$$= \left[\bigcup_{i=1}^{\infty} (A_i \cap (A_n^\varepsilon)^c)\right] \cup \left[A_n^\varepsilon \cap \left(\bigcap_{i=1}^{\infty} A_i^c\right)\right]$$
$$= \left[\bigcup_{i=1}^{n} (A_i \cap (A_n^\varepsilon)^c)\right] \cup \left[\bigcup_{i=n+1}^{\infty} (A_i \cap (A_n^\varepsilon)^c)\right] \cup \left[A_n^\varepsilon \cap \left(\bigcap_{i=1}^{\infty} A_i^c\right)\right]$$

$$
\begin{aligned}
&\subset [A_n \cap (A_n^\varepsilon)^c] \cup \left[\bigcup_{i=n+1}^{\infty} A_i\right] \cup [A_n^\varepsilon \cap A_n^c] \\
&= A_n \setminus A_n^\varepsilon \cup \left[\bigcup_{i=n+1}^{\infty} A_i\right] \cup A_n^\varepsilon \setminus A_n \\
&= (A_n \triangle A_n^\varepsilon) \cup \left[\bigcup_{i=n+1}^{\infty} A_i\right].
\end{aligned}
$$

Then

$$
\mu\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \triangle A_n^\varepsilon\right) \leq \mu\left(A_n \triangle A_n^\varepsilon\right) + \mu\left(\bigcup_{i=n+1}^{\infty} A_i\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,
$$

showing that $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$. So $\mathcal{B}$ is a monotone class. By the monotone class theorem,

$$
\mathcal{B} = \sigma\left(\mathcal{A}_0\right) = \mathcal{A}.
$$

□

Corollary C.2.2. *Let $T > 0$ and*

$$
\mathcal{A}_0 \stackrel{def}{=} \left\{\bigcup_{i=1}^{n} (s_i, s_{i+1}] : \ n \text{ is a finite integer, } 0 \leq s_1 \leq \cdots \leq s_{n+1} \leq T, \text{ and } s_i \in \mathbb{Q}\right\}.
$$

Then $\mathcal{A}_0$ is an algebra of subsets of $(0, T]$ generating $\mathcal{B}((0, T])$.

Proof. Since $\mathcal{A}_0 \subset \mathcal{B}((0, T])$, $\sigma(\mathcal{A}_0) \subset \mathcal{B}((0, T])$. So its enough to show $\sigma(\mathcal{A}_0) \supset \mathcal{B}((0, T])$. Let $0 \leq a < b \leq T$ and $a_n, b_n \in \mathbb{Q} \cap [0, T]$ with a_n decreasing to a and b_n strictly increasing to b. Then

$$
(a, b) = \bigcup_{n=1}^{\infty} (a_n, b_n] \in \sigma(\mathcal{A}_0).
$$

Since any open subset of $(0, T]$ is a countable union of such open intervals,

$$
\mathcal{B}((0, T]) \subset \sigma(\mathcal{A}_0).
$$

□

Vita

Matthew Daniel Turner was born September 23, 1981 and grew up in Vienna, GA. He received Bachelor of Science degrees in Mathematics and in Computer Science from Southern Polytechnic State University in 2004. While completing the Doctorate of Philosophy in Mathematics at the University of Tennessee, he completed a Master of Science in Statics through the Intercollegiate Graduate Statistics Program. He married Heather Michele Rainey in August of 2003. In May of 2007, their eldest son, Jackson Hendrix Turner, was born and in March of 2010, their youngest son, Charlie Bryson Turner, was born.