

**Rearrangement Algorithm in Risk  
Aggregation**

by

Alex Thomas

B.Sc. (Honours), University of Toronto, 2014

A thesis

presented to Ryerson University

in partial fulfilment

of the requirements for the degree of

Master of Science

in the program of

Applied Mathematics

Toronto, Ontario, Canada, 2017

© Alex Thomas, 2017

## DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

## ABSTRACT

Rearrangement Algorithm in Risk Aggregation

Alex Thomas, Master of Science, Applied Mathematics, 2017, Ryerson  
University

Sometimes there's no closed-form analytical solutions for the risk measure of aggregate losses representing, say, a company's losses in each country or city it operates in, a portfolio of losses subdivided by investment, or claims made by clients to an insurance company. Assuming there's enough data to assign a distribution to those losses, we examine the Rearrangement Algorithm's ability to numerically compute the Expected Shortfall and Exponential Premium Principle/Entropic Risk Measure of aggregate losses. A more efficient discretization scheme is introduced and the algorithm is extended to the Entropic Risk Measure which turns out to have a smaller uncertainty spread than the Expected Shortfall at least for the cases that we examined.

## ACKNOWLEDGMENTS

I'd like to express my effusive gratitude to all the members of staff in Ryerson's Mathematics department – administrators, professors, and support – for their conviviality and cordiality throughout my two years here, all the lecturers I had during my graduate course studies for their superlative pedagogy, and my supervisor, Dr. Foivos Xanthos, for his ubiquitous support and encouragement in writing my thesis.

## TABLE OF CONTENTS

Declaration . . . . .	ii
Abstract . . . . .	iii
Acknowledgments . . . . .	iv
List of Tables . . . . .	viii
List of Figures . . . . .	x
List of Symbols . . . . .	x
1 Risk Measures . . . . .	3
1.1 Coherent and Convex Risk Measures . . . . .	3
1.2 Value-at-Risk and Expected Shortfall . . . . .	4
1.3 Entropic Risk measure . . . . .	7
2 Risk Aggregation . . . . .	10
2.1 Convex Order . . . . .	10
2.2 Aggregate Losses . . . . .	13
2.3 Analytical Solutions for the Homogeneous case . . . . .	15
3 Rearrangement Algorithm . . . . .	22
3.1 Description . . . . .	22
3.2 Numerical Results . . . . .	29
3.2.1 Homogeneous Risk Aggregation . . . . .	29

3.3	Uncertainty Spread for the Expected Shortfall . . . . .	35
3.4	Non-Homogeneous Cases . . . . .	38
3.5	Project . . . . .	40
	Appendix . . . . .	47
A	MATLAB Code . . . . .	47
	Bibliography . . . . .	59

## LIST OF TABLES

3.1	Sharp lower bounds on the ES for the sum of $d = 3$ random variables all being $\text{Exp}(\lambda = 2)$ . . . . .	30
3.2	Sharp lower bounds on the ES for the sum of $d = 3$ random variables all being $\text{Exp}(\lambda = 4)$ . . . . .	30
3.3	Sharp lower bounds on the ES for the sum of $d = 3$ random variables all being $\text{Pareto}(\theta = 2)$ . . . . .	31
3.4	Sharp lower bounds on the ES for the sum of $d = 3$ random variables all being $\text{Pareto}(\theta = 4)$ . . . . .	31
3.5	Sharp lower bounds on the ES for the sum of $d = 56$ random variables all being $\text{Pareto}(\theta = 2)$ . We can see that 'SIM' scheme provides a much closer approximation than the 'PUC' scheme in all cases. Here RE stands for relative error. . . . .	32
3.6	Sharp lower bounds on the ES for the sum of $d = 4$ random variables all being $N(0, 1)$ . . . . .	32
3.7	Sharp lower bounds on the Entropic Risk Measure for the sum of $d = 3$ random variables all being $\text{Exp}(\lambda)$ . . . . .	33
3.8	Sharp lower bounds on the Entropic Risk Measure for the sum of $d = 4$ random variables all being $\text{Exp}(\lambda)$ . . . . .	33
3.9	Sharp lower bounds on the Entropic Risk Measure for the sum of $d = 5$ random variables all being $\text{Exp}(\lambda)$ . . . . .	34
3.10	Sharp lower bounds on the EN for the sum of $d = 4$ random variables all being $N(0, 1)$ . . . . .	34

3.11	ES dependence range for the sum of $d = 7$ random variables having distributions $L_1 \sim Pareto(5)$ , $L_2 \sim Pareto(7)$ , $L_3 \sim Exp(2)$ , $L_4 \sim Exp(4)$ , $L_5 \sim N(1, 3)$ , $L_6 \sim N(1, 4)$ , $L_7 \sim N(2, 5)$ when $\alpha = 0.99$ . . . . .	39
3.12	EN and ES dependence range for the sum of $d = 5$ random variables having distributions $L_1 \sim Exp(3)$ , $L_2 \sim Exp(5)$ , $L_3 \sim Exp(7)$ , $L_4 \sim Exp(8)$ , $L_5 \sim Exp(9)$ when $\alpha = 0.99$ . . . . .	39



## LIST OF FIGURES

3.1	'PUC' discretization (left) versus 'SIM' discretization (right) empirical distribution approximations of a Pareto(2) CDF for n=8. . . . .	25
3.2	'PUC' discretization (left) versus 'SIM' discretization (right) empirical distribution approximations of a Pareto(2) CDF for n=24. . . . .	25
3.3	Uncertainty spread of the Expected Shortfall for 3 losses all Exponential(2) distributed as $\alpha$ is varied. . . . .	36
3.4	Uncertainty spread of the Expected Shortfall for 3 losses all Pareto(2) distributed as $\alpha$ is varied. . . . .	37
3.5	Uncertainty spread and Gaussian of the Expected Shortfall for the 5 stocks using the 'SIM' discretization. $\overline{ES}_{0.99}(L^+)$ is in blue, Gaussian is in green, and $\underline{ES}_{0.99}(L^+)$ is in red. . . . .	41
3.6	Uncertainty spread of the Entropic Risk Measure for the 5 stocks using the 'SIM' discretization. $\overline{\rho}_{ent}(L^+)$ is in blue, Gaussian is in green, and $\underline{\rho}_{ent}(L^+)$ is in red. . . . .	42
3.7	Amalgamation of both uncertainty spreads and the Gaussians for visual comparison all using the 'SIM' discretization. Expected Shortfall related loss metrics are in red and Entropic Risk Measure related loss metrics are in blue. . . . .	43
3.8	Adjusted closing price of each stock for each day. . . . .	44
3.9	Plot of the sum of all losses for each day. . . . .	45

3.10 Plot of the uncertainty spreads for both risk measures. Entropic in blue and ES in red. . . . .	46
--	----

## Introduction

Often in practice we wish to calculate the risk measure  $\rho(L^+)$  of an aggregate sum of losses,  $L^+ = L_1 + \dots + L_d$ , representing, say, a company's losses subdivided by geography or portfolio losses subdivided by different investments. By using historical data from each loss,  $L_i$ , one can assign a distribution  $F_i$  to each  $L_i$ . Then the class of aggregate losses is described as follows.

$$\mathcal{A}(F_1, \dots, F_d) = \{L^+ = L_1 + \dots + L_d \mid L_i \sim F_i\}$$

Given only this information, it's not possible to calculate  $\rho(L^+)$  as we are uncertain about the dependence between the  $L_i$ 's. The following are the extreme values of  $\rho(L^+)$ :

$$\underline{\rho}(L^+) = \inf\{\rho(L^+) \mid L^+ \in \mathcal{A}(F_1, \dots, F_d)\}$$

$$\bar{\rho}(L^+) = \sup\{\rho(L^+) \mid L^+ \in \mathcal{A}(F_1, \dots, F_d)\}$$

The interval  $[\underline{\rho}(L^+), \bar{\rho}(L^+)]$  is called the uncertainty spread of  $\rho$ . So when we are uncertain about the dependence of  $L_i$ 's one would like to use a risk measure  $\rho$  that has the smallest possible uncertainty spread.

It turns out that the calculation of  $\bar{\rho}(L^+), \underline{\rho}(L^+)$  is a deep mathematical problem and analytical solutions are known only for some special cases. The Rearrangement Algorithm is a probabilistic numerical method that estimates these values. This method has been extensively used to calculate the uncertainty spread of the Value at Risk and the Expected Shortfall for various scenarios (see e.g. [1, 7, 5, 2]).

The aim of this thesis is to give modified versions of the Rearrangement Algorithm that can be used for the calculation of  $\underline{ES}_\alpha(L^+)$  and  $\underline{\rho}_{ent}(L^+)$ , where  $\rho_{ent}$  is the Entropic Risk Measure. In Chapter 1, we give the definition and the main properties of  $ES$  and  $\rho_{ent}$ . In Chapter 2, we give analytical solutions for the upper bounds,  $\overline{ES}_\alpha(L^+)$  and  $\overline{\rho}_{ent}(L^+)$ , and for the lower bounds in some special cases. In chapter 3, we present the Rearrangement Algorithm and numerical results for different choices of  $F_i$ . From our results it follows that the uncertainty spread of  $\rho_{ent}$  is smaller than  $ES_{0.99}$  at least for the particular cases we analysed and that the new discretization scheme we introduced here for the calculation of  $\underline{ES}_\alpha(L^+)$  is much more efficient than the one introduced in [5] when we work with heavy tailed distributions, large  $\alpha$ , and a large number of losses. In chapter 4, we outline our main MATLAB codes.

## Chapter 1

### RISK MEASURES

#### 1.1 Coherent and Convex Risk Measures

In this section we will present an axiomatic theory of risk measures. The premise of this approach is to identify the essential properties that a reliable risk measure should possess.

For the following we will fix a linear subspace  $\mathcal{M}$  of the space  $L^0(\Omega, \mathcal{F})$  of all random variables on  $(\Omega, \mathcal{F})$ . The space  $\mathcal{M}$  is our model of all the possible portfolio losses  $L$  in the market. We assume here that  $\mathcal{M}$  contains constants, which represent portfolios with predictable profits (e.g. backing capital). The risk measure  $\rho$  is defined as a map  $\rho: \mathcal{M} \rightarrow \mathbb{R}$ .

**Axiom I(Monotonicity):** *For  $L_1, L_2 \in \mathcal{M}$  such that  $L_1 \leq L_2$  (i.e.  $L_1(\omega) \leq L_2(\omega)$  for each  $\omega \in \Omega$ ), we have  $\rho(L_1) \leq \rho(L_2)$ .*

From an economic point of view this axiom is obvious: positions that lead to higher losses in every state of the world require more risk capital.

**Axiom II(Translation invariance):** *For all  $L \in \mathcal{M}$  and  $l \in \mathbb{R}$  we have that  $\rho(L + l) = \rho(L) + l$ .*

The translation invariance axiom states that by adding or subtracting a deterministic quantity  $l$  to a position leading to the loss  $L$ , we alter our capital requirement by exactly that amount.

**Axiom III(Subadditivity):** *For all  $L_1, L_2 \in \mathcal{M}$  we have  $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$*

Subadditivity reflects the idea that risk can be reduced by diversification, a time-honoured principle in finance and economics.

**Axiom IV(Positive homogeneity):** For all  $L \in \mathcal{M}$  and  $\lambda \geq 0$  we have  $\rho(\lambda L) = \lambda\rho(L)$

Positive homogeneity is the requirement that equality should hold in Axiom III if there is no diversification between losses in the portfolio (e.g.  $\rho(L + L) = 2\rho(L)$ ).

**Axiom V(Convexity):** For all  $L_1, L_2 \in \mathcal{M}$  and  $\lambda \in (0, 1)$  we have  $\rho(\lambda L_1 + (1 - \lambda)L_2) \leq \lambda\rho(L_1) + (1 - \lambda)\rho(L_2)$

Convexity can be regarded as a relaxation of assumptions III and IV.

**Definition 1.1.** A risk measure  $\rho : \mathcal{M} \rightarrow (-\infty, \infty]$  is called **coherent** if it satisfies Axioms I-IV and **convex** if it satisfies Axioms I,II and V.

## 1.2 Value-at-Risk and Expected Shortfall

Two of the most widely used risk measures by financial institutions include Value-at-Risk and its average over some confidence level  $\alpha$  called the Expected Shortfall. It is well-known that the Expected Shortfall is a coherent risk measure and Value-at-Risk fails axioms III and V. To set up our framework, consider the portfolio loss  $L$  at the end of some fixed time horizon  $\Delta t$ .

**Definition 1.2.** Given some confidence level  $\alpha \in (0, 1)$ , the VaR of a portfolio at the confidence level  $\alpha$  is given by the smallest number  $l$  such that the probability that the loss  $L$  exceeds  $l$  is no larger than  $1 - \alpha$ . Mathematically,

$$\text{VaR}_\alpha(L) = \inf\{l \in \mathbb{R} \mid P(L > l) \leq 1 - \alpha\} = F^{-1}(\alpha)$$

**Proposition 1.3.** *Suppose that  $L \sim N(\mu, \sigma^2)$  then*

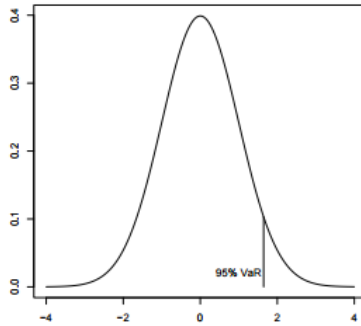
$$\text{VaR}_\alpha(L) = \mu + \sigma\Phi^{-1}(\alpha)$$

*Proof.* Let  $F$  be the cdf of  $L$ , we will show that  $F(\mu + \sigma\Phi^{-1}(\alpha)) = \alpha$ . Indeed,

$$P(L \leq \mu + \sigma\Phi^{-1}(\alpha)) = P\left(\frac{L - \mu}{\sigma} \leq \Phi^{-1}(\alpha)\right) = \Phi(\Phi^{-1}(\alpha)) = \alpha$$

□

In particular for  $N(0, 1)$  the  $\text{VaR}_{0.95} = 1.645$ , this means that there is 95% chance that we will lose at most this amount.



**Definition 1.4.** *Let  $L$  be a loss distribution with  $E(|L|) < \infty$ . The expected shortfall at confidence level  $\alpha \in (0, 1)$  is defined as*

$$ES_\alpha(L) = \frac{1}{1 - \alpha} \int_\alpha^1 \text{VaR}_u(L) du$$

In the case where  $L$  follows a continuous distribution, the Expected Shortfall can be calculated by the following formula

$$ES_\alpha(L) = E(L \mid L \geq \text{VaR}_\alpha(L))$$

**Proposition 1.5.** *Suppose that  $L \sim N(\mu, \sigma^2)$ . Then*

$$ES_\alpha(L) = \mu + \sigma \frac{\phi(\Phi^{-1}(\alpha))}{1 - \alpha},$$

where  $\phi$  is the density function of the standard normal distribution.

*Proof.* Let  $\tilde{L} = \frac{L - \mu}{\sigma}$ , then  $\tilde{L} \sim N(0, 1)$  and by Proposition 1.3 and the above conditional expectation formula we have that

$$ES_\alpha(\tilde{L}) = \frac{1}{1 - \alpha} \int_{\Phi^{-1}(\alpha)}^{\infty} l \phi(l) dl = \frac{1}{1 - \alpha} \left[ -\phi(l) \right]_{\Phi^{-1}(\alpha)}^{\infty} = \frac{\phi(\Phi^{-1}(\alpha))}{1 - \alpha}$$

Now we have that  $ES_\alpha(L) = ES_\alpha(\sigma\tilde{L} + \mu)$  and since  $ES_\alpha$  is a coherent risk measure by axiom II we get that

$$ES_\alpha(L) = \sigma ES_\alpha(\tilde{L}) + \mu = \mu + \sigma \frac{\phi(\Phi^{-1}(\alpha))}{1 - \alpha}$$

□

We recall here that a random vector  $\mathbf{L} = (L_1, \dots, L_d)'$  is comonotonic, whenever there exists a random variable  $X$  and increasing functions  $f_i$  such that  $L_i = f_i(X)$ . It turns out that both Value at Risk and Expected Shortfall are additive on comonotonic vectors. This is an important observation that will be used later.

**Theorem 1.6.** ([4]) *Let  $\alpha \in (0, 1)$  and  $\mathbf{L} = (L_1, \dots, L_d)'$  be a comonotonic random vector. Then*

$$VaR_\alpha(L_1 + \dots + L_d) = VaR_\alpha(L_1) + \dots + VaR_\alpha(L_d)$$

**Theorem 1.7.** *Let  $\alpha \in (0, 1)$  and  $\mathbf{L} = (L_1, \dots, L_d)'$  be a comonotonic random vector. Then*

$$ES_\alpha(L_1 + \dots + L_d) = ES_\alpha(L_1) + \dots + ES_\alpha(L_d)$$



*Proof.* By applying the prior theorem we get that

$$\begin{aligned}
 ES_\alpha(L_1 + \dots + L_d) &= \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_u(L_1 + \dots + L_d) du = \\
 \frac{1}{1-\alpha} \left( \int_\alpha^1 \text{VaR}_u(L_1) du + \dots + \int_\alpha^1 \text{VaR}_u(L_d) du \right) &= ES_\alpha(L_1) + \dots + ES_\alpha(L_d)
 \end{aligned}$$

□

To estimate  $\text{VaR}(L)$ ,  $ES(L)$  for a loss  $L$  with unknown distribution we will use a data set  $\{\tilde{l}_1 \geq \dots \geq \tilde{l}_N\}$  with large  $N$  drawn from  $L$  and use the following estimators. More details on the derivation of these formulas can be found in [4].

$$\widehat{\text{VaR}}_\alpha = \tilde{l}_j, \text{ where } j = \lfloor N(1-\alpha) \rfloor + 1$$

$$\widehat{ES}_\alpha = \frac{1}{N(1-\alpha)} \left( \sum_{i=1}^{\lfloor N(1-\alpha) \rfloor} \tilde{l}_i \right)$$

### 1.3 Entropic Risk measure

Also known as the Exponential Premium Principle, this risk measure is used in the insurance industry to calculate premiums.

**Definition 1.8. (*Entropic Risk Measure*)** *Given the exponential utility function given by  $u(x) = 1 - \exp(-x)$  the loss function is defined as  $l(x) = -u(-x) = -(1 - \exp(x))$ . Assume moreover that  $\mathcal{M}$  contains random variables  $L$  with  $\mathbb{E}(\exp(L)) < +\infty$ . The acceptance set is defined to be the following set*

$$A = \{L \in \mathcal{M} | E(l(L)) \leq l(0)\} = \{L \in \mathcal{M} | \mathbb{E}(\exp(L)) \leq 1\}$$

The derived risk measure  $\rho_{ent}$  is called the **Entropic Risk Measure** or the *Exponential Premium Principle*. The value of  $\rho_{ent}(L)$  corresponds to the minimum capital required to make our position acceptable.

$$\begin{aligned}\rho_{ent}(L) &= \inf\{m \in \mathbb{R} \mid \mathbb{E}(\exp(L - m)) \leq 1\} = \\ &= \inf\{m \in \mathbb{R} \mid \mathbb{E}(\exp(L)) \leq \exp(m)\} = \log(\mathbb{E}(\exp(L)))\end{aligned}$$

**Proposition 1.9.** *Suppose that  $L \sim N(\mu, \sigma^2)$  then*

$$\rho_{ent}(L) = \mu + \frac{1}{2}\sigma^2$$

*Proof.*

$$\begin{aligned}L \sim N(\mu, \sigma^2) &\Rightarrow \exp(L) \sim \text{LogNormal}(\mu, \sigma^2) \Rightarrow \\ \Rightarrow \mathbb{E}(\exp(L)) &= \exp(\mu + \frac{1}{2}\sigma^2) \Rightarrow \log(\mathbb{E}(\exp(L))) = \mu + \frac{1}{2}\sigma^2\end{aligned}$$

□

**Theorem 1.10.** *The Entropic Risk Measure is convex, but not coherent.*

*Proof.* To show that it's convex we have to show that it satisfies axioms I, II, and V.

I: Pick any  $L_1, L_2 \in \mathcal{M}$  such that  $L_1 \leq L_2$ . Since the exponential, expectation, and natural logarithms are all monotonically increasing we have that

$$\begin{aligned}\Rightarrow \exp(L_1) &\leq \exp(L_2) \Rightarrow \mathbb{E}(\exp(L_1)) \leq \mathbb{E}(\exp(L_2)) \Rightarrow \\ \log(\mathbb{E}(\exp(L_1))) &\leq \log(\mathbb{E}(\exp(L_2)))\end{aligned}$$

II: Pick any  $L \in \mathcal{M}$  and  $l \in \mathbb{R}$ . Then

$$\begin{aligned}\rho_{ent}(L+l) &= \log(\mathbb{E}(\exp(L+l))) = \log(\mathbb{E}(\exp(L))\exp(l)) = \\ &= \log(\mathbb{E}(\exp(L))) + l = \rho_{ent}(L) + l\end{aligned}$$

V: Show that for any  $L_1, L_2 \in \mathcal{M}$  and  $\lambda \in (0, 1)$ ,  $\rho_{ent}(\lambda L_1 + (1-\lambda)L_2) \leq \lambda \rho_{ent}(L_1) + (1-\lambda)\rho_{ent}(L_2)$ . By definition

$$\begin{aligned}\rho_{ent}(\lambda L_1 + (1-\lambda)L_2) &= \min\{m \in \mathbb{R} \mid \mathbb{E}(\exp(\lambda L_1 + (1-\lambda)L_2 - m)) \leq 1\} \\ &\leq \lambda \rho_{ent}(L_1) + (1-\lambda)\rho_{ent}(L_2)\end{aligned}$$

Indeed, note that

$$\begin{aligned}\mathbb{E}(\exp(\lambda L_1 + (1-\lambda)L_2 - \lambda \rho_{ent}(L_1) - (1-\lambda)\rho_{ent}(L_2))) \\ = \mathbb{E}(\exp(\lambda(L_1 - \rho_{ent}(L_1)) + (1-\lambda)(L_2 - \rho_{ent}(L_2))))\end{aligned}$$

Since the exponential is a convex function and the expectation is additive then,

$$\leq \lambda \mathbb{E}(\exp(L_1 - \rho_{ent}(L_1))) + (1-\lambda)\mathbb{E}(\exp(L_2 - \rho_{ent}(L_2))) \leq \lambda + (1-\lambda) = 1$$

Where the last inequality follows by definition of the Entropic Risk Measure.

The Entropic Risk Measure fails axiom IV preventing it from being coherent. For instance, if  $L \sim N(0, \frac{1}{2})$  and  $\tilde{L} = 3L \sim N(0, \frac{9}{2})$  we have that  $\exp(L)$  follows a lognormal distribution and thus by the previous proposition  $3\rho(L) = 3\log(E(\exp(L))) = 3(0 + \frac{1}{4}) = 0.75$ . On the other hand we have that  $\rho(\tilde{L}) = \log(E(\exp(\tilde{L}))) = 0 + \frac{9}{4} = 2.25 \neq 3\rho(L)$ .

□

## Chapter 2

### RISK AGGREGATION

#### 2.1 Convex Order

**Definition 2.1.** Let  $X$  and  $Y$  be two random variables with finite mean.  $X$  is smaller than  $Y$  in **convex order**,  $X \leq_{CX} Y$ , if for every convex function  $\phi$ ,  $\mathbb{E}[\phi(X)] \leq \mathbb{E}[\phi(Y)]$  whenever they are well-defined.

**Definition 2.2.** Given a set of losses  $L_i \sim F_i$ , the class

$$\mathcal{A}(F_1, \dots, F_d) = \{L^+ = L_1 + \dots + L_d \mid L_i \sim F_i \text{ for all } i = 1, \dots, d\}$$

is called the **class of aggregate losses**.

**Remark 2.3.** If  $X \leq_{CX} Y$ , then  $\mathbb{E}(X) = \mathbb{E}(Y)$  and  $\mathbb{E}(X^2) \leq \mathbb{E}(Y^2)$ . The inequality is achieved by choosing  $\phi$  to be  $x^2$  in the above definition and the equality is achieved by choosing  $\phi$  to be  $x$  and  $-x$ . Indeed, by choosing  $\phi(x) = x$ , we have that  $\mathbb{E}(X) \leq \mathbb{E}(Y)$  and by choosing  $\tilde{\phi}(x) = -x$ , we have that  $\mathbb{E}(-X) \leq \mathbb{E}(-Y) \Rightarrow -\mathbb{E}(X) \leq -\mathbb{E}(Y) \Rightarrow \mathbb{E}(X) \geq \mathbb{E}(Y)$ .

**Proposition 2.4.** (Levy and Kroll, p.553-574, 1978 [3])

$X \leq_{CX} Y$  if and only if  $ES_\alpha(X) \leq ES_\alpha(Y)$  and  $\mathbb{E}(X) = \mathbb{E}(Y)$  for all  $\alpha \in (0, 1)$ .

Given the set of aggregate sums,

$$\mathcal{A}(F_1, \dots, F_d) = \left\{ \sum_{i=1}^d L_i \mid L_i \sim F_i \right\},$$

we wish to find  $\max\{\mathcal{A}(F_1, \dots, F_d)\}$  and  $\min\{\mathcal{A}(F_1, \dots, F_d)\}$  with respect to  $\preceq_{CX}$ .

**Proposition 2.5.** (p.96 [1])  $\max\{\mathcal{A}(F_1, \dots, F_d)\} = F_1^{-1}(U) + \dots + F_d^{-1}(U)$  where  $U \sim \text{Unif}(0, 1)$ .

*Proof.* Since  $F_i^{-1}(U) \sim F_i$  for each  $i$ ,  $\sum_{i=1}^d F_i^{-1}(U) \in \mathcal{A}(F_1, \dots, F_d)$ . Let  $\sum_{i=1}^d L_i \in \mathcal{A}(F_1, \dots, F_d)$ . We will show that  $\sum_{i=1}^d L_i \preceq_{CX} \sum_{i=1}^d F_i^{-1}(U)$ . Indeed, by the sub-additivity of  $ES_\alpha$  we have that

$$ES_\alpha\left(\sum_{i=1}^d L_i\right) \leq \sum_{i=1}^d ES_\alpha(L_i) = \sum_{i=1}^d ES_\alpha(F_i^{-1}(U)) = ES_\alpha\left(\sum_{i=1}^d F_i^{-1}(U)\right),$$

where the last equality follows by the additivity of  $ES_\alpha$  on the comonotonic vector  $(F_1^{-1}(U), \dots, F_d^{-1}(U))'$  (see Theorem 1.7).

□

**Proposition 2.6.** (Ruschendorf, 1983 [6])

$$\min\{\mathcal{A}(F_1, F_2)\} = F_1^{-1}(U) + F_2^{-1}(1 - U)$$

The following example shows that if  $d > 2$ ,  $\min\{\mathcal{A}(F_1, F_2)\}$  may not exist.

**Example 2.7.** (Example 3.1 [1]) Let  $F_1$  be a discrete distribution on  $\{0, 3, 8\}$  with equal probability,  $F_2$  be a discrete distribution on  $\{0, 6, 16\}$  with equal probability, and  $F_3$  be a discrete distribution on  $\{0, 7, 13\}$  with equal probability. In our example, the sample space is divided into three disjoint subsets  $A_1, A_2, A_3$  with equal probability  $\frac{1}{3}$ . Let  $\omega_i \in A_i$ ,  $i = 1, 2, 3$ . We verify two scenarios:

(a) First, consider the following dependence structure:

$$\begin{bmatrix} L_1(\omega_1) & L_2(\omega_1) & L_3(\omega_1) \\ L_1(\omega_2) & L_2(\omega_2) & L_3(\omega_2) \\ L_1(\omega_3) & L_2(\omega_3) & L_3(\omega_3) \end{bmatrix} = \begin{bmatrix} 3 & 16 & 0 \\ 0 & 6 & 13 \\ 8 & 0 & 7 \end{bmatrix}$$

It is easy to verify that the distribution of  $L_i$  is  $F_i$ ,  $i = 1, 2, 3$ . The distribution of  $L_1 + L_2 + L_3$  is on  $\{19, 19, 15\}$  with equal probability. Thus,  $\mathbb{E}[(L_1 + L_2 + L_3 - 19)^+] = 0$ ;

(b) Consider another dependence structure:

$$\begin{bmatrix} L_1(\omega_1) & L_2(\omega_1) & L_3(\omega_1) \\ L_1(\omega_2) & L_2(\omega_2) & L_3(\omega_2) \\ L_1(\omega_3) & L_2(\omega_3) & L_3(\omega_3) \end{bmatrix} = \begin{bmatrix} 0 & 16 & 0 \\ 3 & 0 & 13 \\ 8 & 6 & 7 \end{bmatrix}$$

It is easy to verify that the distribution of  $L_i$  is  $F_i$ ,  $i = 1, 2, 3$ . The distribution of  $L_1 + L_2 + L_3$  is on  $\{16, 16, 21\}$  with equal probability. Thus,  $\mathbb{E}[(16 - (L_1 + L_2 + L_3))^+] = 0$ .

Note that both  $g(x) = (x - 19)^+$  and  $g(x) = (16 - x)^+$  are convex functions. Hence, if there exists a convex ordering minimal element  $S$  in  $\mathcal{A}(F_1, F_2, F_3)$ , it must satisfy  $\mathbb{E}[(S - 19)^+] = 0$  and  $\mathbb{E}[(16 - S)^+] = 0$ . However, we can see that when  $L_1 = 8$ , no matter what values  $L_2$  and  $L_3$  take,  $S$  will be either  $> 19$  or  $< 16$ . That means that  $\mathbb{E}[(S - 19)^+] = 0$  and  $\mathbb{E}[(16 - S)^+] = 0$  cannot be satisfied simultaneously by the same  $S \in \mathcal{A}(F_1, F_2, F_3)$ . This shows that the minimal element with respect to convex order does not exist.

**Proposition 2.8.**  $\mathbb{E}(\sum_{i=1}^d L_i) \preceq_{CX} L$  for all  $L \in \mathcal{A}(F_1, \dots, F_d)$ .

*Proof.* This follows by Jensen's Inequality:

$$\phi(\mathbb{E}(L)) \leq \mathbb{E}(\phi(L)) \text{ for any convex function } \phi.$$

Indeed, let  $L = \sum_{i=1}^d L_i, L_i \sim F_i$  and  $\phi$  a convex function, then we have that

$$\mathbb{E}(\phi(\mathbb{E}(\sum_{i=1}^d L_i))) = \phi(\mathbb{E}(\sum_{i=1}^d L_i)) \leq \mathbb{E}(\phi(\sum_{i=1}^d L_i)).$$

Therefore  $\mathbb{E}(\sum_{i=1}^d L_i) \leq_{CX} \sum_{i=1}^d L_i$  since  $\phi$  is arbitrary. □

## 2.2 Aggregate Losses

As a practical real-world problem, let's say that we wish to forecast the aggregate sum of a company's potential losses for  $d$  store locations in different cities or a portfolio of  $d$  stocks assuming that we have enough data to assign a distribution to each loss represented by  $L_i$ . Knowing each marginal distribution does not allow us to know the joint distribution of their aggregate sum,  $L^+$ , with certainty as there's an infinitude of ways that the losses could interact with one another based on their dependencies meaning that we cannot compute  $\rho(L^+)$ , where  $L^+ \in \mathcal{A}$ . So, we consider the set of possible joint distributions given each marginal distribution,  $\mathcal{A}(F_1, \dots, F_d)$  or the class of aggregate losses, and we calculate its maximum value,  $\bar{\rho}(L^+)$ , the most that we could lose, and its minimum value,  $\underline{\rho}(L^+)$ , the least that we could lose. Analytical solutions always exist for the former value, but not the latter value. Of course, there are a plethora of possible risk measures we could use, but convex ones like the Expected Shortfall and Entropic Risk Measure work quite well here.

**Definition 2.9.** *Let  $\rho$  be a risk measure, then*

- (i)  $\underline{\rho}(L^+) = \inf\{\rho(L^+) | L^+ \in \mathcal{A}(F_1, \dots, F_d)\}$  is the least that we could lose.
- (ii)  $\bar{\rho}(L^+) = \sup\{\rho(L^+) | L^+ \in \mathcal{A}(F_1, \dots, F_d)\}$  is the most that we could lose.
- (iii) The interval  $[\underline{\rho}(L^+), \bar{\rho}(L^+)]$  is called the **uncertainty spread** of  $\rho$ .

Moreover, risk and insurance experts are always looking for the 'best' measure of risk. Here, we'll consider the risk measure with the smallest uncertainty spread as the 'best'. That is, the measure with the smallest distance between the most and least that we could lose on an aggregate portfolio is the 'best' risk measure. The following example illustrates the role of dependence of risk factors in risk measurement.

**Example 2.10.** *Suppose that the aggregate loss of our portfolio is  $L^+ = L_1 + L_2$  and it is estimated that the risk factor changes  $L_1, L_2$  follow  $N(0, 1)$  and  $N(0, 4)$  respectively. With this information alone, we cannot calculate the VaR of  $L^+$ , we also need to determine the dependence of  $L_1$  and  $L_2$ . As the following calculations show our estimation for VaR change dramatically when we assume different dependence structures.*

- (a)  $L_1, L_2$  are independent.

Then  $L_1 + L_2$  follows  $N(0, 5)$ , therefore

$$\text{VaR}_{0.99}(L^+) = \sqrt{5} \cdot \Phi^{-1}(0.99) = 5.2019$$

- (b)  $\mathbf{L} = (L_1, L_2)'$  is a comonotonic vector.

$$\begin{aligned} \text{VaR}_{0.99}(L^+) &= \text{VaR}_{0.99}(L_1) + \text{VaR}_{0.99}(L_2) = \Phi^{-1}(0.99) + 2 \cdot \Phi^{-1}(0.99) = \\ &= 6.9790 \end{aligned}$$



(c) Take  $L_1 = Z, L_2 = -2Z$ , then  $L^+ = -Z$  and we have that

$$VaR_{0.99}(L^+) = VaR_{0.99}(-Z) = \Phi^{-1}(0.99) = 2.3263$$

(d)  $\mathbf{L} = (L_1, L_2)' \sim N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 4 \end{bmatrix}\right)$ . Then we have that  $L^+ \sim N(0, 4)$

and thus  $VaR_{0.99}(L^+) = 2 \cdot \Phi^{-1}(0.99) = 4.6527$ .

Dependence	(a)	(b)	(c)	(d)
$VaR_{0.99}(L^+)$	5.20	6.98	2.33	4.65

### 2.3 Analytical Solutions for the Homogeneous case

In this section we will calculate the analytical solutions for the bounds,

$$\underline{ES}_\alpha(L^+), \overline{ES}_\alpha(L^+), \underline{\rho}_{ent}(L^+), \overline{\rho}_{ent}(L^+),$$

of the class  $\mathcal{A}(\underbrace{F, \dots, F}_d)$ , where  $F$  is a distribution function. This is the so-called homogeneous risk aggregation problem. The results will be used later on in the Numerical Results section of Chapter 3 to compare our outputted solutions with. In the general case of non-homogeneous distributions (i.e.  $F_i \neq F_j$ ) the problem of finding analytical solutions for  $\underline{ES}_\alpha(L^+), \underline{\rho}_{ent}(L^+)$  turns out to be extremely difficult. In the next chapter we will present numerical solutions via the Rearrangement Algorithm for some special cases.

For  $\overline{ES}_\alpha(L^+), \overline{\rho}_{ent}(L^+)$  the calculations can be carried out by applying Proposition 2.5. In the table below we calculate the values for different choices of  $F$ .

$F$	$\overline{ES}_\alpha(L^+)$	$\overline{\rho}_{ent}(L^+)$
Exp( $\lambda$ )	$\frac{d}{1-\alpha} \int_\alpha^1 -\frac{1}{\lambda} \log(1-u) du$	$\log(\int_0^1 ((1-u)^{-\frac{d}{\lambda}}) du)$
$N(\mu, \sigma^2)$	$\frac{d}{1-\alpha} \int_\alpha^1 (\sqrt{2\sigma^2} \operatorname{erf}^{-1}(2u-1) + \mu) du$	$\log(\int_0^1 \exp(d\sqrt{2\sigma^2} \operatorname{erf}^{-1}(2u-1) + d \cdot \mu) du)$
Pareto( $\theta$ )	$\frac{d}{1-\alpha} \int_\alpha^1 (\frac{1}{(1-u)^{\frac{1}{\theta}}} - 1) du$	-

where  $\operatorname{erf}^{-1}$  denotes the inverse of the error function given by  $\operatorname{erf}(x) = \frac{2}{\pi} \int_0^x \exp(-t^2) dt$  which requires computational software to compute.

**Theorem 2.11.** ([4] See Prop 8.34 on p.307)

$$\underline{ES}_\alpha(L^+) = \frac{1}{\beta} \int_0^\beta (d-1)F^{-1}((d-1)t) + F^{-1}(1-t) dt$$

where  $\beta = \frac{1-\alpha}{d}$  and  $L_i \sim F_i$  for all  $i$ .

**Corollary 2.12.** (i) If  $F$  is the cdf of Exp( $\lambda$ ), then we have that

$$\underline{ES}_\alpha(L^+) = \frac{1}{\beta} \int_0^\beta -\frac{(d-1)}{\lambda} \log(1-(d-1)t) - \frac{1}{\lambda} \log(t) dt$$

(ii) If  $F$  is the cdf of Pareto( $\theta$ ) then we have that

$$\underline{ES}_\alpha(L^+) = \frac{1}{\beta} \int_0^\beta (d-1) \left( \frac{1}{(1-(d-1)t)^{\frac{1}{\theta}}} - 1 \right) + \frac{1}{t^{\frac{1}{\theta}}} - 1 dt$$

*Proof.*

- (i) Using the previous theorem and the inverse CDF of the exponential distribution we get that

$$\underline{ES}_\alpha(L^+) = \frac{1}{\beta} \int_0^\beta (d-1)F^{-1}((d-1)t) + F^{-1}(1-t)dt \quad (2.1)$$

$$= \frac{1}{\beta} \int_0^\beta -\frac{(d-1)}{\lambda} \log(1-(d-1)t) - \frac{1}{\lambda} \log(t)dt \quad (2.2)$$

- (ii) Using the previous theorem and the inverse CDF of the one-parameter variant of the Pareto distribution, namely  $F(x) = 1 - (1+x)^{-\theta}$  for  $x > 0$ , we get that

$$\underline{ES}_\alpha(L^+) = \frac{1}{\beta} \int_0^\beta (d-1)F^{-1}((d-1)t) + F^{-1}(1-t)dt \quad (2.3)$$

$$= \frac{1}{\beta} \int_0^\beta (d-1) \left( \frac{1}{(1-(d-1)t)^{\frac{1}{\theta}}} - 1 \right) + \frac{1}{t^{\frac{1}{\theta}}} - 1 dt \quad (2.4)$$

□

**Theorem 2.13.** ([1] See section 3.3 on p.97) *If  $F$  is defined on  $\mathbb{R}^+$ , then we have that*

$$T_a = H\left(\frac{U}{d}\right) \mathbb{1}_{U \in [0, da]} + D(a) \mathbb{1}_{U \in (da, 1]} \preceq_{CX} L^+ \text{ for each } L^+ \in \underbrace{\mathcal{A}(F, \dots, F)}_d,$$

where  $U \sim \text{Unif}(0, 1)$ , and  $a \in [0, \frac{1}{d}]$  such that  $H(x)$  is non-increasing on  $[0, a]$ , and  $\lim_{x \rightarrow a^-} H(x) \geq D(a)$  where

$$H(x) = (d-1)F^{-1}((d-1)x) + F^{-1}(1-x)$$

$$D(a) = d \int_{(d-1)a}^{1-a} \frac{F^{-1}(y)}{1-da} dy$$

In most practical cases,  $T_a$  as defined above turns out to be an element of  $\mathcal{A}(\underbrace{F, \dots, F}_d)$  and thus  $T_a = \min\{\mathcal{A}(\underbrace{F, \dots, F}_d)\}$ . For more details on this see [1].

**Proposition 2.14.** *If  $F$  is the cdf of Exponential( $\lambda$ ) and  $T_a$  is as in the previous theorem, then*

$$\begin{aligned} \rho_{ent}(T_a) &= \log\left(\int_0^{da} \left(1 - \frac{(d-1)}{d}u\right)^{-\frac{(d-1)}{\lambda}} \left(\frac{u}{d}\right)^{-\frac{1}{\lambda}} du\right) \\ &+ (1-da) \exp\left(\frac{d}{\lambda}\right) a^{\frac{da}{\lambda(1-da)}} (1 - (d-1)a)^{-\frac{d(1-(d-1)a)}{\lambda(1-da)}} \end{aligned}$$

*Proof.* Recall that  $\rho_{ent}(T_a) = \log(\mathbb{E}(\exp(T_a)))$ .

$$H\left(\frac{U}{d}\right) = (d-1)F^{-1}\left(\frac{(d-1)}{d}U\right) + F^{-1}\left(1 - \frac{U}{d}\right) \quad (2.5)$$

$$= \frac{-1}{\lambda}(d-1)\log\left(1 - \frac{(d-1)}{d}U\right) - \frac{1}{\lambda}\log\left(1 - \left(1 - \frac{U}{d}\right)\right) \quad (2.6)$$

$$= \log\left(1 - \frac{(d-1)}{d}U\right)^{-\frac{(d-1)}{\lambda}} + \log\left(\frac{U}{d}\right)^{-\frac{1}{\lambda}} \quad (2.7)$$

$$D(a) = d \int_{(d-1)a}^{1-a} \frac{F^{-1}(y)}{1-da} dy \quad (2.8)$$

$$= \frac{d}{1-da} \int_{(d-1)a}^{1-a} \frac{-1}{\lambda} \log(1-y) dy \quad (2.9)$$

$$= \frac{d}{1-da} \left[ \frac{y}{\lambda} + \frac{1}{\lambda} (1-y) \log(1-y) \right]_{(d-1)a}^{1-a} \quad (2.10)$$

$$= \frac{d}{1-da} \left[ \frac{1-a}{\lambda} + \frac{a \log(a)}{\lambda} - \frac{(d-1)a}{\lambda} - \frac{1}{\lambda} (1-(d-1)a) \log(1-(d-1)a) \right] \quad (2.11)$$

$$= \frac{d}{\lambda} + \log(a)^{\frac{da}{\lambda(1-da)}} + \log(1-(d-1)a)^{\frac{-d(1-(d-1)a)}{\lambda(1-da)}} \quad (2.12)$$

$$\rho_{ent}(T_a) = \log(\mathbb{E}((1 - \frac{(d-1)}{d} U)^{\frac{-(d-1)}{\lambda} \mathbb{1}_{U \in [0, da]}} (\frac{U}{d})^{\frac{-1_{U \in [0, da]}}{\lambda}} \exp(\frac{d \mathbb{1}_{U \in (da, 1]}}{\lambda}) \quad (2.13)$$

$$\times a^{\frac{da}{\lambda(1-da)} \mathbb{1}_{U \in (da, 1]}} (1 - (d-1)a)^{\frac{-d(1-(d-1)a)}{\lambda(1-da)} \mathbb{1}_{U \in (da, 1]}})) \quad (2.14)$$

Let  $A = [0, da]$  and  $B = (da, 1]$  then  $A \cup B = [0, 1]$  and  $A \cap B = \emptyset$ .

$$\begin{aligned} &= \log\left(\int_0^1 \left(1 - \frac{(d-1)u}{d}\right)^{\frac{-(d-1) \mathbb{1}_A}{\lambda}} \left(\frac{u}{d}\right)^{\frac{-1_A}{\lambda}} \exp\left(\frac{d \mathbb{1}_B}{\lambda}\right) a^{\frac{da \mathbb{1}_B}{\lambda(1-da)}} (1 - (d-1)a)^{\frac{-d(1-(d-1)a) \mathbb{1}_B}{\lambda(1-da)}} du\right) \\ &= \log\left(\int_0^{da} \left(1 - \frac{(d-1)u}{d}\right)^{\frac{-(d-1)}{\lambda}} \left(\frac{u}{d}\right)^{\frac{-1}{\lambda}} du\right) \\ &\quad + \int_{da}^1 \exp\left(\frac{d}{\lambda}\right) a^{\frac{da}{\lambda(1-da)}} (1 - (d-1)a)^{\frac{-d(1-(d-1)a)}{\lambda(1-da)}} du \\ &= \log\left(\int_0^{da} \left(1 - \frac{(d-1)u}{d}\right)^{\frac{-(d-1)}{\lambda}} \left(\frac{u}{d}\right)^{\frac{-1}{\lambda}} du\right) \\ &\quad + (1-da) \exp\left(\frac{d}{\lambda}\right) a^{\frac{da}{\lambda(1-da)}} (1 - (d-1)a)^{\frac{-d(1-(d-1)a)}{\lambda(1-da)}} \end{aligned}$$

Recall that  $a \in [0, \frac{1}{d}]$  such that  $H(x)$  is non-increasing on  $[0, a]$ , and  $\lim_{x \rightarrow a^-} H(x) \geq D(a)(*)$

In order for a to satisfy (\*):

$$\lim_{x \rightarrow a^-} (d-1)F^{-1}((d-1)x) + F^{-1}(1-x) \geq \frac{d}{1-da} \int_{(d-1)a}^{1-a} F^{-1}(y)dy$$

$$\lim_{x \rightarrow a^-} \frac{-(d-1)}{\lambda} \log(1-(d-1)x) - \frac{\log(x)}{\lambda} \geq \frac{d}{1-da} \int_{(d-1)a}^{1-a} -\frac{1}{\lambda} \log(1-y)dy$$

The arguments of the natural logarithms are always positive since  $0 \leq a \leq \frac{1}{d}$ , however, choosing  $a = 0$  yields  $\log(0)$  for the Entropic Risk Measure which diverges so we could just ignore  $a = 0$  as a solution. Indeed,  $0 < a \leq \frac{1}{n} \Rightarrow 0 > -a \geq \frac{-1}{n} \Rightarrow 0 > -(n-1)a \geq \frac{-(n-1)}{n} \Rightarrow 1 > 1 - (n-1)a \geq 1 - 1 + \frac{1}{n} = \frac{1}{n} > 0$ . So we could just plug in  $x = a$  on the left-hand side.

$$\frac{-(d-1)}{\lambda} \log(1-(d-1)a) - \frac{1}{\lambda} \log(a) \geq \frac{d}{\lambda} + \frac{da}{\lambda(1-da)} \log(a)$$

$$-\frac{d(1-(d-1)a)}{\lambda(1-da)} \log(1-(d-1)a)$$

$$\frac{(-(1-da)(d-1) + d(1-(d-1)a))}{1-da} \log(1-(d-1)a) + \frac{(-(1-da) - da)}{1-da} \log(a) \geq d$$

$$\frac{1}{a} - (d-1) \geq \exp(d - d^2 a)$$

Ostensibly,  $a = \frac{1}{d}$  may seem like a good choice, however, since  $1-da$  occurs in the denominator in the above equations,  $a = \frac{1}{d}$  is in fact a restriction. For some fixed  $d$ , we must use computational software that can solve inequalities

in order to find some  $a \in (0, \frac{1}{d})$  and then plug it into the following equation to obtain the analytical solutions:

$$\begin{aligned} \rho_{ent}(T_a) &= \log\left(\int_0^{da} \left(1 - \frac{(d-1)}{d}u\right)^{\frac{-(d-1)}{\lambda}} \left(\frac{u}{d}\right)^{\frac{-1}{\lambda}} du\right) \\ &+ (1-da)\exp\left(\frac{d}{\lambda}\right)a^{\frac{da}{\lambda(1-da)}}(1-(d-1)a)^{\frac{-d(1-(d-1)a)}{\lambda(1-da)}} \end{aligned}$$

□

It is worth mentioning here that in that our numerical findings (see Tables 7-9 in the following chapter) show that  $\rho_{ent}(T_a)$  agrees with  $\underline{\rho}_{ent}(L^+)$ .

**Proposition 2.15.** *If  $d$  is even and  $F = \Phi$ , then  $\underline{\rho}_{ent}(L^+) = \underline{ES}_\alpha(L^+) = 0$*

*Proof.* Note that in this case we have that  $0 = Z + (-Z) + Z + (-Z) + \dots \in \mathcal{A}(\Phi, \dots, \Phi)$  and thus by Proposition 2.8 we get that 0 is the minimum of  $\mathcal{A}(\Phi, \dots, \Phi)$  with respect to convex order. □

## Chapter 3

### REARRANGEMENT ALGORITHM

#### 3.1 Description

The Rearrangement Algorithm (RA for short) is a probabilistic algorithm used to approximate  $\underline{\rho}(L^+)$  and  $\bar{\rho}(L^+)$ . In this chapter we will use it to find the lower bound of the Expected Shortfall,  $\underline{ES}_\alpha$  and Entropic Risk Measure,  $\underline{\rho}_{ent}$ . Given a class of distributions  $F_i, i = 1, \dots, d$ , the algorithm searches for minimal elements in  $\mathcal{A}(F_1, \dots, F_d)$  with respect to convex order. As we will illustrate in Example 3.5 when  $\mathcal{A}(F_1, \dots, F_d)$  has no minimum element, the algorithm fails to converge to  $\underline{\rho}(L^+)$ . The convergence of this algorithm is still an open problem. Nonetheless, in situations where the losses follow non-homogeneous distributions and analytical solutions cannot be found, the RA is of significant use for estimation purposes. Please see <https://sites.google.com/site/rearrangementalgorithm/> for more information. The following example gives a simple illustration of the algorithm.

**Example 3.1.** *Suppose that we're dealing with the variance risk measure,  $d = 2$  and  $F_1 = F_2 = U\{1, 2, 3, 4\}$  (i.e. the discrete uniform distribution on  $\{1, 2, 3, 4\}$ ). The following matrix describes the dependence of two random variables  $L_1$  and  $L_2$  that follow  $U\{1, 2, 3, 4\}$ . The first and second column represent the possible  $L_1, L_2$  values respectively and the third column represents the possible values of the aggregate sum.*



$L_1, L_2 \sim \text{Discrete Uniform}\{1, 2, 3, 4\}$

$$\left[ \begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 3 & 3 & 6 \\ 4 & 4 & 8 \end{array} \right]$$

(a) First, randomly permute the first two columns. This is where the probabilistic aspect of the algorithm comes from. And, as we can see, their row sums change.

$$\left[ \begin{array}{cc|c} 2 & 1 & 3 \\ 1 & 3 & 4 \\ 4 & 2 & 6 \\ 3 & 4 & 7 \end{array} \right]$$

**Definition 3.2.** Two vectors  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\mathbf{y} = (y_1, \dots, y_n)$  are **oppositely ordered** if and only if

$$(x_j - x_k)(y_j - y_k) \leq 0$$

for all  $i = 1, \dots, n$ .

(b) Making  $L_1$  and  $L_2$  **oppositely ordered** to one another equilibrates the spread of their aggregate or sum. Call them  $\tilde{L}_1$  and  $\tilde{L}_2$  respectively.

$$\left[ \begin{array}{cc|c} 2 & 3 & 5 \\ 1 & 4 & 5 \\ 4 & 1 & 5 \\ 3 & 2 & 5 \end{array} \right]$$

(c) Calculate the variance of what you started off with and the variance after applying the Rearrangement Algorithm.

$$V(L_1 + L_2) = 5 \text{ and } V(\tilde{L}_1 + \tilde{L}_2) = 0$$

$V(L_1 + L_2)$  is greater than  $V(\tilde{L}_1 + \tilde{L}_2)$  meaning that we've minimized the variance of what we started with. The value  $V(\tilde{L}_1 + \tilde{L}_2) = 0$  is indeed the value  $\underline{V}(L^+)$  that we were looking for.

The first step of the RA is to discretize the given distribution  $F$ . This is done by considering an empirical distribution function of the following form

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{[F^{-1}(s_i), +\infty)}(x),$$

where  $s_1 < \dots < s_n$  are sample point in  $[0, 1]$ . In our numerical results, we will use the following discretizations schemes

a)  $s_i = \frac{i-1}{n}, i = 1, \dots, n$ . This is the discretization used in [5] which we will call 'PUC' in our code.

b)  $s_i = \frac{i+0.7}{n+1}, i = 1, \dots, n$ . This is a different discretization we propose here. In our code we call this 'SIM'.

We then define the dependence matrix  $X$  as follows

$$x_{i,j} = F_j^{-1}(s_i) \text{ where } 1 \leq i \leq n \text{ and } 1 \leq j \leq d$$

$$[X_{i,j}] = \begin{pmatrix} F_1^{-1}(s_1) & \cdots & F_d^{-1}(s_1) \\ \vdots & \ddots & \vdots \\ F_1^{-1}(s_n) & \cdots & F_d^{-1}(s_n) \end{pmatrix}$$

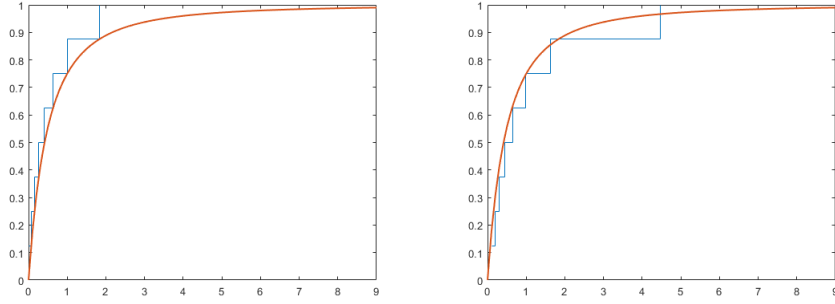


Figure 3.1: 'PUC' discretization (left) versus 'SIM' discretization (right) empirical distribution approximations of a Pareto(2) CDF for  $n=8$ .

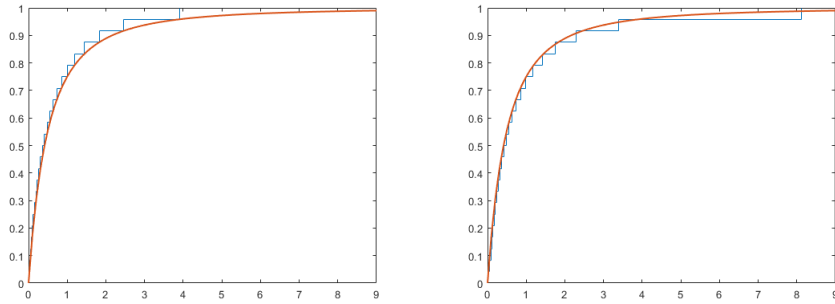


Figure 3.2: 'PUC' discretization (left) versus 'SIM' discretization (right) empirical distribution approximations of a Pareto(2) CDF for  $n=24$ .

The vector that corresponds to the row sum of the  $d$  columns is denoted by  $(+X)$ .

The RA algorithm is then described by the following steps

- (1) Each column of the matrix  $X$  is randomly permuted.
- (2) Iteratively rearrange the  $j$ th column of the matrix  $X$  so that it becomes oppositely ordered to the row sum of everything but the  $j$ th column. A new dependence matrix  $Y$  is found.

$$\begin{bmatrix} F_j^{-1}(s_1) \\ F_j^{-1}(s_2) \\ \vdots \\ F_j^{-1}(s_n) \end{bmatrix} \begin{bmatrix} \sum_{k \neq j} F_k^{-1}(s_1) \\ \sum_{k \neq j} F_k^{-1}(s_2) \\ \vdots \\ \sum_{k \neq j} F_k^{-1}(s_n) \end{bmatrix}$$

- (3) Repeat step (2) until  $|\rho((+Y)) - \rho((+X))| < \epsilon$ , where  $\epsilon > 0$  is a fixed threshold.

Items (1),(3) are introduced to decrease the number of loops needed to reach  $\underline{\rho}(L^+)$ . Item (2) is the essence of the RA. In the following we demonstrate why making the vectors in a matrix oppositely ordered to one another reduces the convex order of the row sum. Let  $a \in \mathbb{R}^n$ , then we denote  $a^\uparrow = (a_{[n]}, \dots, a_{[2]}, a_{[1]})$  which is the vector  $a$  with its elements in ascending order and  $a^\downarrow = (a_{[1]}, \dots, a_{[n]})$  which is the vector  $a$  with its elements in descending order. We also introduce the following order

$$a \leq b \text{ if and only if } \sum_{i=1}^j a_{[i]} \leq \sum_{i=1}^j b_{[i]} \text{ for all } 1 \leq j \leq n \text{ and } \sum_{i=1}^n a_i = \sum_{i=1}^n b_i.$$

**Proposition 3.3.** (*Day, 1972*)

$$a^\uparrow + b^\downarrow \leq a + b$$

*Proof.* Let  $S = a^\uparrow + b^\downarrow$ ,  $r = a + b$ . For simplicity, we assume that  $a = a^\uparrow$

Case 1 ( $n = 2$ ):

Since  $a^\uparrow = a$ , we have that  $a_1 \leq a_2$ . If  $b_1 \geq b_2$ , then  $b = b^\downarrow$  so  $s = r$ .

If  $b_1 < b_2$ , then  $s = (a_1 + b_2, a_2 + b_1)$  and  $r = (a_1 + b_1, a_2 + b_2)$ . Let  $s_1 = a_1 + b_2$ ,  $s_2 = a_2 + b_1$ ,  $r_1 = a_1 + b_1$ ,  $r_2 = a_2 + b_2$ . Note that  $s_1 + s_2 = r_1 + r_2$ .

$$r_{[1]} = a_2 + b_2 \geq \max\{a_2 + b_1, b_2 + a_1\} = s_{[1]}$$

For the general case, if  $b_1 \geq b_2 \geq \dots \geq b_n$  then  $s = r$ . Suppose that there exists some  $1 \leq j < k \leq n$  such that  $b_j < b_k$ . Then put  $s_i^* = r_i$  for  $i \neq j, k$ .

$s_j^* = a_j + b_k$  and  $s_k^* = a_k + b_j$ . By case 1, we have that  $s^* \leq r$ .

$$(a_j + b_k, a_k + b_j) \leq (a_j + b_j, a_k + b_k)$$

If  $s^* = s$  we're done, otherwise repeat the previous step for  $r = s^*$ . □

**Corollary 3.4.** *For any  $L_1, L_2$  discrete uniformly distributed random variables we have that  $L_1 + \tilde{L}_2 \preceq_{CX} L_1 + L_2$  where  $\tilde{L}_2$  is a rearrangement of  $L_2$  made oppositely ordered to  $L_1$ .*

*Proof.* By Day, we have that  $s = a^\uparrow + b^\downarrow \leq a + b = r$  for any  $a, b \in \mathbb{R}^n$ .

$$L_1 + \tilde{L}_2 \sim L_1^\uparrow + L_2^\downarrow \preceq_{CX} L_1 + L_2$$

$$ES_\alpha(L_1 + \tilde{L}_2) = \frac{1}{n(1-\alpha)} \sum_{i=1}^{\lfloor n(1-\alpha) \rfloor} s_{[i]} \leq \frac{1}{n(1-\alpha)} \sum_{i=1}^{\lfloor n(1-\alpha) \rfloor} r_{[i]} = ES_\alpha(L_1 + L_2)$$

□

Below we give an example where  $\mathcal{A}(F_1, \dots, F_d)$  has no minimum element and the RA fails to converge to  $\underline{ES}_{0.99}(L^+)$ .

**Example 3.5.** *Let's say we're using the Rearrangement Algorithm to find the  $\underline{ES}_{0.99}(L^+)$  of the following matrix used in Example 2.1.8.*

$$X = \begin{bmatrix} 3 & 16 & 0 \\ 0 & 6 & 13 \\ 8 & 0 & 7 \end{bmatrix}$$

*After finding the Expected Shortfall measure of all  $(3!)^3$  different column permutations of our matrix and taking the smallest one, we find that the 'true value' is 19. Now let's apply the RA to this matrix 100 times, average the values, and perform a hypothesis test to determine if the average produced by the RA is close enough to the true value using a two-sided t-test with unknown mean and variance.*

$$\begin{cases} H_0: \text{average of 100 values} = 19 \\ H_A: \text{average of 100 values} \neq 19 \end{cases}$$

$$T = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{22.82 - 19}{\frac{2.8898}{\sqrt{100}}} = 13.2189$$

$$p\text{-value} = 2\mathbb{P}(t_{99} \geq 13.2189) \approx 0$$

*Very strong evidence against the  $H_0$  causing us to reject it in favour of the  $H_A$  at any level of significance. In particular, the RA fails to find the true smallest Expected Shortfall for this particular case.*

## **3.2 Numerical Results**

In this section we present our numerical results from the application of the Rearrangement Algorithm to several risk aggregation problems. For our numerical computation we used  $n = 100,000$  and  $\epsilon = 0.0001$ . Each value taking MATLAB about 15 seconds to compute when  $d = 4$  and about 300 seconds when  $d = 56$  using a personal computer with a 3.30 GHz processor and about 6.00 GB of RAM. The definite integrals that appeared in section 2.3 were solved via Wolfram Integrator.

### **3.2.1 Homogeneous Risk Aggregation**

For a small number of homogeneously distributed losses, both the 'PUC' and 'SIM' discretizations work well in calculating the  $\underline{ES}_\alpha$ , however, it seems that the 'SIM' discretization is more accurate for values of  $\alpha$  very close to 1 (see Tables 3.1-3.4). Moreover, for large  $d$  the new scheme appears to be more efficient for heavy-tailed distributions like the Pareto distribution (see Table 3.5). For homogeneously distributed losses, both discretization schemes work well in calculating the  $\underline{\rho}_{ent}$  (see Tables 3.7-3.9).

Table 3.1: Sharp lower bounds on the ES for the sum of  $d = 3$  random variables all being  $\text{Exp}(\lambda = 2)$ .

$\lambda = 2$	$\underline{ES}_\alpha('PUC')$	$\underline{ES}_\alpha('SIM')$	Analytical
$\alpha = 0.5$	1.6113	1.6115	1.5850
$\alpha = 0.6$	1.6666	1.6669	1.6545
$\alpha = 0.7$	1.7590	1.7594	1.7587
$\alpha = 0.8$	1.9234	1.9240	1.9239
$\alpha = 0.9$	2.2338	2.2348	2.2347
$\alpha = 0.99$	3.3495	3.3573	3.3552
$\alpha = 0.999$	4.4633	4.5167	4.5036

Table 3.2: Sharp lower bounds on the ES for the sum of  $d = 3$  random variables all being  $\text{Exp}(\lambda = 4)$ .

$\lambda = 4$	$\underline{ES}_\alpha('PUC')$	$\underline{ES}_\alpha('SIM')$	Analytical
$\alpha = 0.5$	0.8056	0.8058	0.7925
$\alpha = 0.6$	0.8334	0.8336	1.6545
$\alpha = 0.7$	0.8796	0.8798	0.8794
$\alpha = 0.8$	0.9617	0.9620	0.9620
$\alpha = 0.9$	1.1169	1.1174	1.1172
$\alpha = 0.99$	1.6748	1.6786	1.6776
$\alpha = 0.999$	2.2317	2.2583	2.2518



Table 3.3: Sharp lower bounds on the ES for the sum of  $d = 3$  random variables all being Pareto( $\theta = 2$ ).

$\theta = 2$	$\underline{ES}_\alpha('PUC')$	$\underline{ES}_\alpha('SIM')$	Analytical
$\alpha = 0.5$	4.0736	4.1015	4.1010
$\alpha = 0.6$	4.5975	4.6323	4.6320
$\alpha = 0.7$	5.3900	5.4364	5.4360
$\alpha = 0.8$	6.7484	6.8178	6.8175
$\alpha = 0.9$	9.8508	9.9893	9.9889
$\alpha = 0.99$	32.2849	33.6447	33.6444
$\alpha = 0.999$	95.5076	108.3204	108.5449

Table 3.4: Sharp lower bounds on the ES for the sum of  $d = 3$  random variables all being Pareto( $\theta = 4$ ).

$\theta = 4$	$\underline{ES}_\alpha('PUC')$	$\underline{ES}_\alpha('SIM')$	Analytical
$\alpha = 0.5$	1.1874	1.1884	1.1845
$\alpha = 0.6$	1.2810	1.2822	1.2819
$\alpha = 0.7$	1.4244	1.4261	1.4258
$\alpha = 0.8$	1.6573	1.6598	1.6593
$\alpha = 0.9$	2.1336	2.1386	2.1376
$\alpha = 0.99$	4.5136	4.5595	4.5507
$\alpha = 0.999$	8.5448	8.9359	8.8679

Table 3.5: Sharp lower bounds on the ES for the sum of  $d = 56$  random variables all being Pareto( $\theta = 2$ ). We can see that 'SIM' scheme provides a much closer approximation than the 'PUC' scheme in all cases. Here RE stands for relative error.

$\theta = 2$	$\underline{ES}_\alpha('PUC')$	$\underline{ES}_\alpha('SIM')$	Analytical	RE('PUC')	RE('SIM')
$\alpha = 0.99$	125.0195	148.1578	148.8020	15.98%	0.42%
$\alpha = 0.995$	164.8556	208.7450	210.7278	21.77%	0.94%
$\alpha = 0.999$	274.4890	444.3673	472.3000	41.88%	5.91%

Table 3.6: Sharp lower bounds on the ES for the sum of  $d = 4$  random variables all being  $N(0, 1)$ .

Risk Measure	'PUC'	'SIM'	Analytical
Expected Shortfall	$-4.21 \times 10^{-4}$	$3.2596 \times 10^{-4}$	0

Table 3.7: Sharp lower bounds on the Entropic Risk Measure for the sum of  $d = 3$  random variables all being  $\text{Exp}(\lambda)$ .

$d = 3$	$\underline{\rho}_{ent}('PUC')$	$\underline{\rho}_{ent}('SIM')$	Analytical
$\lambda = 2$	1.5760	1.5789	1.5817
$\lambda = 3$	1.0273	1.0279	1.0290
$\lambda = 4$	0.7640	0.7642	0.7649
$\lambda = 5$	0.6085	0.6086	0.6091
$\lambda = 6$	0.5057	0.5058	0.5061
$\lambda = 7$	0.4327	0.4327	0.4330
$\lambda = 8$	0.3781	0.3781	0.3783

Table 3.8: Sharp lower bounds on the Entropic Risk Measure for the sum of  $d = 4$  random variables all being  $\text{Exp}(\lambda)$ .

$d = 4$	$\underline{\rho}_{ent}('PUC')$	$\underline{\rho}_{ent}('SIM')$	Analytical
$\lambda = 2$	2.0355	2.0380	2.0794
$\lambda = 3$	1.3459	1.3464	1.3649
$\lambda = 4$	1.0063	1.0066	1.0169
$\lambda = 5$	0.8038	0.8040	0.8105
$\lambda = 6$	0.6692	0.6694	0.6739
$\lambda = 7$	0.5733	0.5734	0.5767
$\lambda = 8$	0.5014	0.5015	0.5040

Table 3.9: Sharp lower bounds on the Entropic Risk Measure for the sum of  $d = 5$  random variables all being  $\text{Exp}(\lambda)$ .

$d = 5$	$\underline{\rho}_{ent}('PUC')$	$\underline{\rho}_{ent}('SIM')$	Analytical
$\lambda = 2$	2.5153	2.5173	2.6939
$\lambda = 3$	1.6721	1.6726	1.7484
$\lambda = 4$	1.2527	1.2530	1.2947
$\lambda = 5$	1.0017	1.0018	1.0281
$\lambda = 6$	0.8344	0.8346	0.8526
$\lambda = 7$	0.7151	0.7152	0.7283
$\lambda = 8$	0.6256	0.6257	0.6357

Table 3.10: Sharp lower bounds on the EN for the sum of  $d = 4$  random variables all being  $N(0, 1)$ .

Risk Measure	'PUC'	'SIM'	Analytical
Entropic	$-5.0611 \times 10^{-4}$	$2.7918 \times 10^{-4}$	0

### 3.3 Uncertainty Spread for the Expected Shortfall

Here we consider plotting the difference between  $\overline{ES}_\alpha(L^+)$  and  $\underline{ES}_\alpha(L^+)$  as we vary  $\alpha$  from 0.01 to 0.99 for the Exponential and Pareto distributions. The former is calculated analytically and the latter is calculated via the Rearrangement Algorithm. It's worth noting that for  $Exp(\lambda)$  the uncertainty spread increases as  $\alpha$  increases to 1. Please see the next couple of pages for their plots.

Expected Shortfall at  $n = 10000$ ,  $d = 3$ ,  $\epsilon = 0.0001$ , Exp(2) vs. Alpha

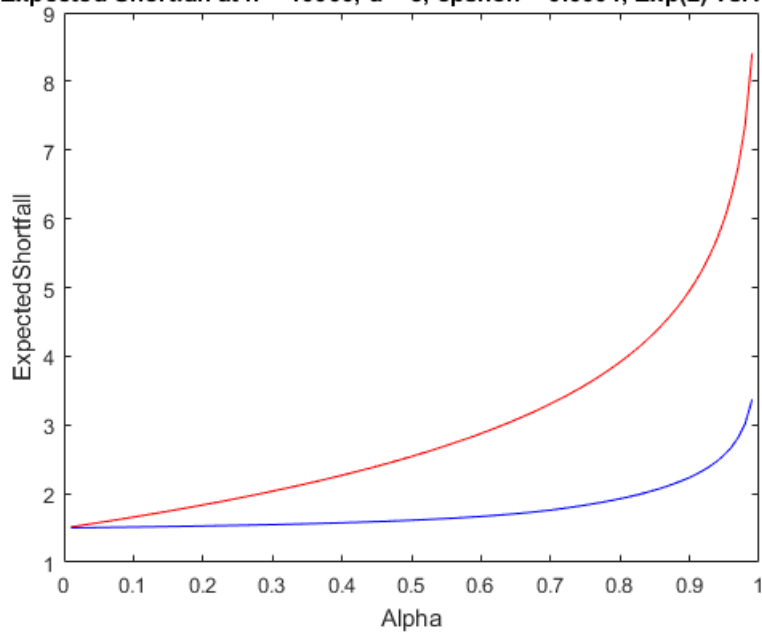


Figure 3.3: Uncertainty spread of the Expected Shortfall for 3 losses all Exponential(2) distributed as  $\alpha$  is varied.

Expected Shortfall at  $n = 10000$ ,  $d = 3$ ,  $\epsilon = 0.0001$ , Pareto(2) vs. Alpha

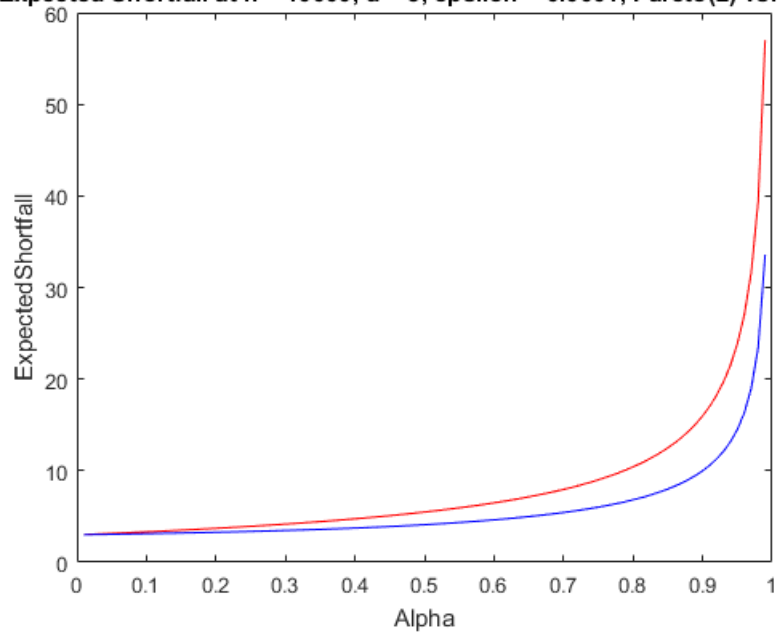


Figure 3.4: Uncertainty spread of the Expected Shortfall for 3 losses all Pareto(2) distributed as  $\alpha$  is varied.

### 3.4 Non-Homogeneous Cases

Now let's suppose that each loss follows a different distribution with different parameters. The analytical solutions for the lower bounds of the aggregate losses are not known so we depend solely on the Rearrangement Algorithm. From our results, it follows that the Entropic Risk Measure has a smaller uncertainty spread than the Expected Shortfall for non-homogeneously distributed Exponential random variables (see Table 3.12). Please see the next page for these next tables.



Table 3.11: ES dependence range for the sum of  $d = 7$  random variables having distributions  $L_1 \sim Pareto(5)$ ,  $L_2 \sim Pareto(7)$ ,  $L_3 \sim Exp(2)$ ,  $L_4 \sim Exp(4)$ ,  $L_5 \sim N(1, 3)$ ,  $L_6 \sim N(1, 4)$ ,  $L_7 \sim N(2, 5)$  when  $\alpha = 0.99$ .

$d = 7$	$\underline{ES}_{0.99}('PUC')$	$\underline{ES}_{0.99}('SIM')$	$\overline{ES}_{0.99}$
$\alpha = 0.9$	5.1654	5.1670	18.5531
$\alpha = 0.95$	5.1659	5.1675	21.3728
$\alpha = 0.99$	5.1659	5.1675	27.5025

Table 3.12: EN and ES dependence range for the sum of  $d = 5$  random variables having distributions  $L_1 \sim Exp(3)$ ,  $L_2 \sim Exp(5)$ ,  $L_3 \sim Exp(7)$ ,  $L_4 \sim Exp(8)$ ,  $L_5 \sim Exp(9)$  when  $\alpha = 0.99$ .

$\underline{\rho}_{ent}('PUC')$	$\underline{\rho}_{ent}('SIM')$	$\bar{\rho}_{ent}$	$\underline{ES}_{0.99}('PUC')$	$\underline{ES}_{0.99}('SIM')$	$\overline{ES}_{0.99}$
0.9241	0.9244	2.3955	1.8782	1.8812	5.1136

### 3.5 Project

Suppose we have a portfolio of 5 non-dividend stocks belonging Yahoo Incorporated, Jacobs Engineering, Adobe Systems Incorporated, E\*TRADE Financial Corporation, and Mohawk Industries Incorporated. We view here the total loss  $L^+$  of the portfolio as the sum of the losses  $L_i$  from each individual stock  $i$ <sup>1</sup>. Our assumption is that each  $L_i$  follows a normal distribution, this is a typical assumption in risk measurement that is employed when we use the variance covariance method to estimate the risk of a portfolio. The parameters of the distributions will be estimated by taking the average of the difference between the opening prices and the closing prices on each day over the course of 1 year. The variance is calculated similarly. We then shift that window of size 1 year over day by day for 3 years from about 250 business days after August 04, 2012 to 1000 business days after that date, and calculate each of the aforementioned loss metrics. Our task is to calculate  $\underline{ES}_{0.99}(L^+)$  and  $\underline{\rho}_{ent}(L^+)$  at each day and compare their uncertainty spread. We'll also calculate  $ES_{0.99}(L^+)$  and  $\rho_{ent}(L^+)$  by assuming Gaussian dependence among  $L_i$ . In this case, it is know that  $L^+$  follows a normal distribution, where the mean is the sum of the means and the variance is calculated from the estimated covariance matrix.

Generally, it seems that the Entropic Risk Measure has a smaller uncertainty spread than the Expected Shortfall for these stocks when  $\alpha = 0.99$  making it a better measure of risk (See Figure 3.10). Also, we can see that both of their lower bounds seem to converge towards the same value.

---

<sup>1</sup>We choose this rather non-standard approach to view the calculation of the risk of the portfolio loss as a risk aggregation problem

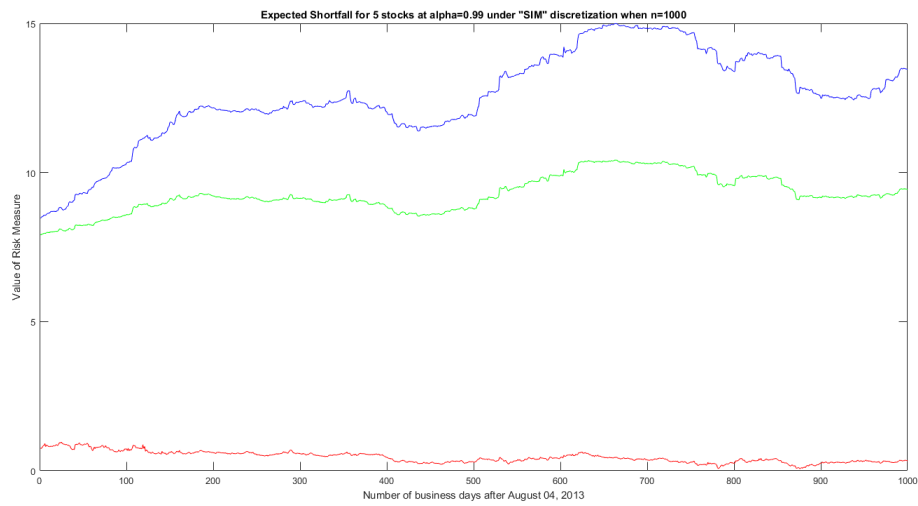


Figure 3.5: Uncertainty spread and Gaussian of the Expected Shortfall for the 5 stocks using the 'SIM' discretization.  $\overline{ES}_{0.99}(L^+)$  is in blue, Gaussian is in green, and  $\underline{ES}_{0.99}(L^+)$  is in red.

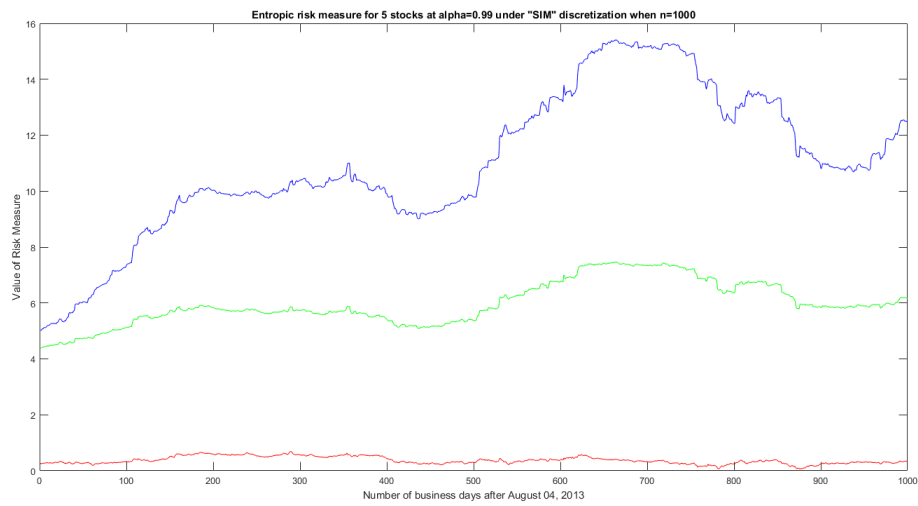


Figure 3.6: Uncertainty spread of the Entropic Risk Measure for the 5 stocks using the 'SIM' discretization.  $\bar{\rho}_{ent}(L^+)$  is in blue, Gaussian is in green, and  $\rho_{ent}(L^+)$  is in red.

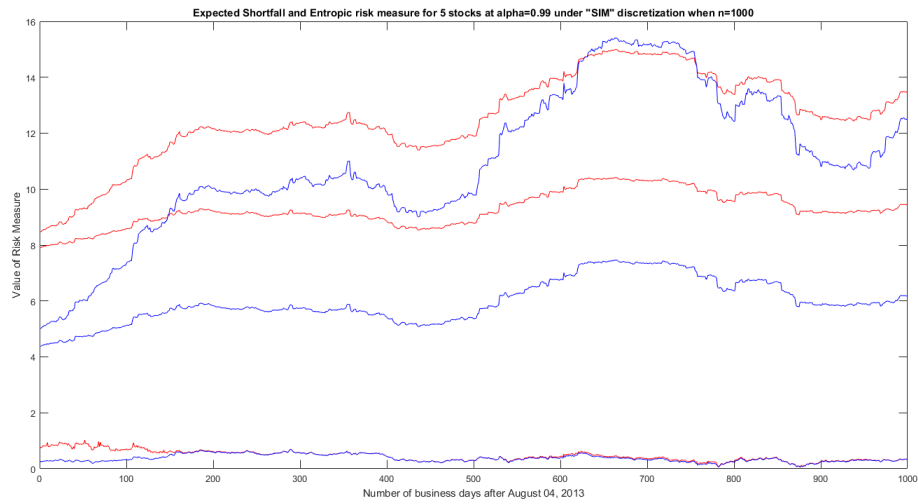


Figure 3.7: Amalgamation of both uncertainty spreads and the Gaussians for visual comparison all using the 'SIM' discretization. Expected Shortfall related loss metrics are in red and Entropic Risk Measure related loss metrics are in blue.



Figure 3.8: Adjusted closing price of each stock for each day.

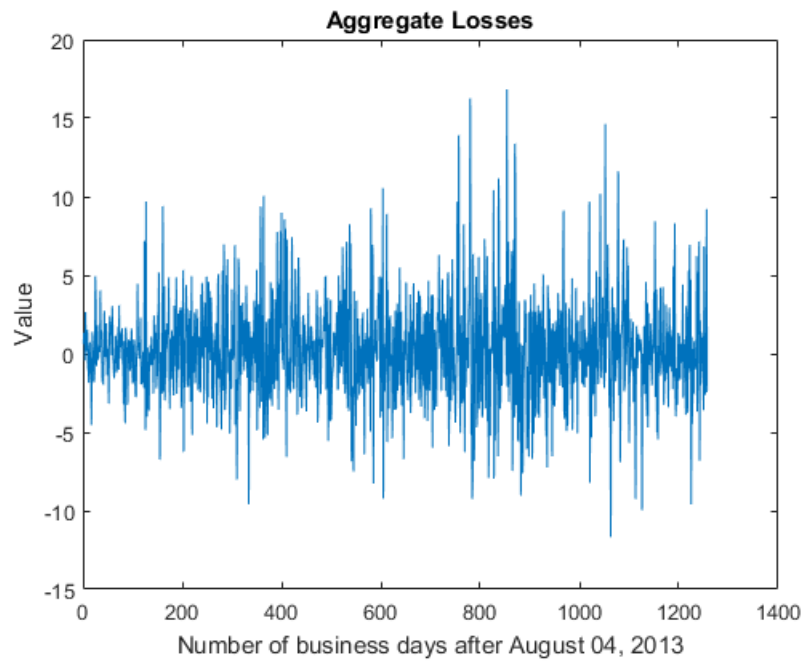


Figure 3.9: Plot of the sum of all losses for each day.

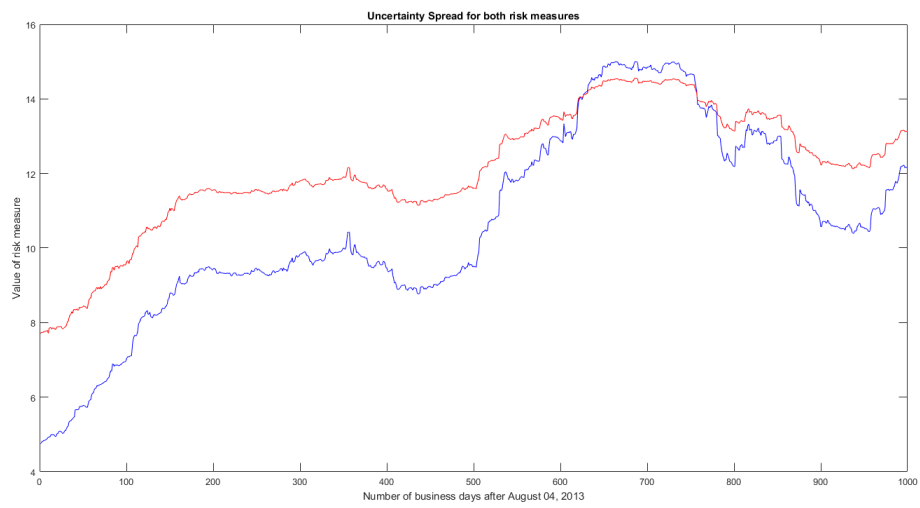


Figure 3.10: Plot of the uncertainty spreads for both risk measures. Entropic in blue and ES in red.



## Appendix A

### MATLAB CODE

Here we present the main MATLAB functions used in this thesis. RearrangementAlgorithm.m is the main function that calculates the lower bound of both the Expected Shortfall and the Entropic Risk Measure. It requires specifying the risk measure,  $n$ ,  $\epsilon > 0$ , the homogeneous distributions and their parameters, the number of losses, the discretization scheme, and  $\alpha$  for the Expected Shortfall. The other two functions serve as auxiliary functions aiding the main function in its computation. For the non-homogeneous cases, the 'starting' matrix in RearrangementAlgorithm.m was manually modified, but everything else remained the same.

```
1 function RearrangementAlgorithm = RearrangementAlgorithm(measure
    , distribution , parameter1 , parameter2 , discretization , n , d ,
    epsilon , alpha)
2
3                                     %takes as input the risk measure
4                                     , the
5                                     %probability distribution used,
6                                     the
7                                     %discretization scheme, n, the
8                                     number of
9                                     %losses d, epsilon , and alpha in
10                                    the case
11                                    %of ES
12
13 h = [];
```

```

9     h(i) = rand(1);
10 end;
11
12 if measure == 'ES'           %if you use the Expected
    Shortfall measure, it
13     ihat = floor(alpha*n);   %goes through this clause
14
15     if distribution == 'PAR'
16         for i = 1:n;         %create the permutation
            matrix "starting"
17             for j = 1:d;     %for the Pareto
                distribution
18                 if discretization == 'PUC'
19                     starting(i,j) = (1/(1-((i - 1)/n)))^(1./
                parameter1) - 1;
20                 elseif discretization == 'SIM'
21                     starting(i,j) = (1/(1-((i+0.7)/(n+1))))^(1./
                parameter1) - 1;
22                 elseif discretization == 'OTH'
23                     starting(i,j) = (1/(1-(i+0.99995)/(n+1)))
                ^ (1./parameter1) - 1;
24                 elseif discretization == 'RAN'
25                     starting(i,j) = (1/(1-(h(i))))^(1./
                parameter1) - 1;
26             end
27         end;
28     end;
29     elseif distribution == 'EXP'
30         for i = 1:n;         %create the permutation
            matrix "starting"

```

```

31     for j = 1:d;                %for the exponential
distribution
32         if discretization == 'PUC'
33             starting(i,j) = (-1./parameter1)*log(1-((i -
1)/n));
34         elseif discretization == 'SIM'
35             starting(i,j) = (-1./parameter1)*log(1-((i
+0.7)/(n+1)));
36         elseif discretization == 'OTH'
37             starting(i,j) = (-1./parameter1)*log(1-(i
+0.99995)/(n+1));
38         elseif discretization == 'RAN'
39             starting(i,j) = (-1./parameter1)*log(1-(h(i)
));
40         end
41     end;
42 end;
43 elseif distribution == 'NOR'
44     for i = 1:n;                %create the permutation
matrix "starting"
45         for j = 1:d;            %for the standard normal
distributions
46             if discretization == 'PUC'
47                 starting(i,j) = sqrt(parameter2)*sqrt(2)*
erfinv(2*(i - 1)/n-1)+parameter1;
%different matrix for each
discretization scheme
48             elseif discretization == 'SIM'
49                 starting(i,j) = sqrt(parameter2)*sqrt(2)*
erfinv(2*(i+0.7)/(n+1)-1)+parameter1;

```

```

50         elseif discretization == 'OTH'
51             starting(i,j) = sqrt(parameter2)*sqrt(2)*
                    erfinv(2*(i+0.99995)/(n+1)-1)+parameter1;
52         elseif discretization == 'RAN'
53             starting(i,j) = sqrt(parameter2)*sqrt(2)*
                    erfinv(2*h(i)-1)+parameter1;
54         end
55     end;
56 end;
57 end
58 new = zeros(size(starting));
59 old = starting;
60
61 for c = 1:d;                                %randomly permute
62     the columns
63     starting(randperm(n),c) = starting(:,c);
64 end
65 while AlternativeDiscreteES(old, alpha) - AlternativeDiscreteES(
66     new, alpha) >= epsilon
67
68     for j = 1:size(starting,2)                %keep going through the RA
69         until the
70         old = starting;                       %ES of the next matrix is
71         within the
72         %threshold of the ES of the
73         previous one
74
75         x = 1;
76         for i = 1:size(starting,1)

```

```

73         C(x) = sum(starting(i,:)) - sum(starting(i,j));
74         x = x + 1;           %for each column make the
                               jth column
75     end                       %oppositely ordered to the
                               sum of the
76                               %leftover elements
77     D = horzcat(starting(:,j),C');
78
79     [~,ix]=sort(D(:,2));
80     inverse=zeros(size(ix));
81     inverse(ix) = numel(ix):-1:1;
82     E = sort(D(:,1));
83     D(:,1)=E(inverse);
84
85     starting(:,j) = D(:,1);   %put the new oppositely
                               ordered column
86     new = starting;          %vector back into the matrix
87     starting = new;
88     if AlternativeDiscreteES(old, alpha) -
        AlternativeDiscreteES(new, alpha) < epsilon
89         break
90     end                       %check to see if the ES of
                               the difference
91     end                       %between the new matrix and
                               the old
92 end                           %matrix is within the
                               threshold; if it's
93                               %not, repeat the while loop,
                               otherwise
94 RearrangementAlgorithm = AlternativeDiscreteES(new, alpha);

```

```

95         %output the value
96
97     elseif measure == 'EN'           %if you use the Entropic measure
98         , it
99
100        %goes through this clause
101
102     if distribution == 'PAR'
103         for i = 1:n;                 %create the permutation
104             matrix "starting"
105             for j = 1:d;             %for the Pareto
106                 distribution
107                 if discretization == 'PUC'
108                     starting(i,j) = (1/(1-((i - 1)/n)))^(1./
109                         parameter1) - 1;
110                 elseif discretization == 'SIM'
111                     starting(i,j) = (1/(1-((i+0.7)/(n+1))))^(1./
112                         parameter1) - 1;
113                 elseif discretization == 'OTH'
114                     starting(i,j) = (1/(1-(i+0.99995)/(n+1)))
115                         ^ (1./parameter1) - 1;
116                 elseif discretization == 'RAN'
117                     starting(i,j) = (1/(1-(h(i))))^(1./
118                         parameter1) - 1;
119                 end
120             end;
121         end;
122     elseif distribution == 'EXP'
123         for i = 1:n;                 %create the permutation
124             matrix "starting"

```

```

116         for j = 1:d;                %for the exponential
distribution
117             if discretization == 'PUC'
118                 starting(i,j) = (-1./parameter1)*log(1-((i -
1)/n));
119             elseif discretization == 'SIM'
120                 starting(i,j) = (-1./parameter1)*log(1-(i
+0.7)/(n+1));
121             elseif discretization == 'OTH'
122                 starting(i,j) = (-1./parameter1)*log(1-(i
+0.99995)/(n+1));
123             elseif discretization == 'RAN'
124                 starting(i,j) = (-1./parameter1)*log(1-h(i))
;
125             end
126         end;
127     end;
128 elseif distribution == 'NOR'
129         for i = 1:n;                %create the permutation
matrix "starting"
130             for j = 1:d;            %for the standard normal
distributions
131                 if discretization == 'PUC'
132                     starting(i,j) = sqrt(parameter2)*sqrt(2)*
erfinv(2*(i - 1)/n-1)+parameter1;
%different matrix for each
discretization scheme
133                 elseif discretization == 'SIM'
134                     starting(i,j) = sqrt(parameter2)*sqrt(2)*
erfinv(2*(i+0.7)/(n+1)-1)+parameter1;

```

```

135         elseif discretization == 'OTH'
136             starting(i,j) = sqrt(parameter2)*sqrt(2)*
                erfinv(2*(i+0.99995)/(n+1)-1)+parameter1;
137         elseif discretization == 'RAN'
138             starting(i,j) = sqrt(parameter2)*sqrt(2)*
                erfinv(2*h(i)-1)+parameter1;
139         end
140     end;
141 end;
142 end
143
144 new = zeros(size(starting));
145 old = starting;
146
147 for c = 1:d;                                %randomly permute
        the columns
148     starting(randperm(n),c) = starting(:,c);
149 end
150
151 while abs(LNofEXPofE(old) - LNofEXPofE(new)) >= epsilon
152
153     for j = 1:size(starting,2)                %keep going through the
        RA until the
154         old = starting;                      %EN of the next matrix
            is within the
155                                             %threshold of the EN of
                                                the previous one
156     x = 1;
157     for i = 1:size(starting,1)
158         C(x) = sum(starting(i,:)) - sum(starting(i,j));

```



```

159         x = x + 1;                               %for each column make
           the jth column
160     end                                           %oppositely ordered to
           the sum of the
161                                           %leftover elements
162     D = horzcat(starting(:,j),C');
163
164     [~,ix]=sort(D(:,2));
165     inverse=zeros(size(ix));
166     inverse(ix) = numel(ix):-1:1;
167     E = sort(D(:,1));
168     D(:,1)=E(inverse);
169
170     starting(:,j) = D(:,1);                       %put the new oppositely
           ordered column
171     new = starting;
172     starting = new;
173     if abs(LNofEXPofE(old) - LNofEXPofE(new)) < epsilon
174         break                                       %check to see if the ES
           of the difference
175     end                                           %between the new matrix
           and the old
176     end                                           %matrix is within the
           threshold; if it's
177 end                                             %not, repeat the while
           loop, otherwise
178                                           %output the value
179 RearrangementAlgorithm = LNofEXPofE(new);
180
181 end

```

```
182 end
```

```
183
```

```
1 function AlternativeDiscreteES = AlternativeDiscreteES(A, alpha)
2
3 x = 1; %takes as input a matrix, A, and
4 a significance level, alpha
5 m = size(A,1);
6
7 for y = 1:size(A,1)
8     C(x) = sum(A(y,1:size(A,2))); %for each row, find the
9     column sums
10    x = x + 1;
11 end
12 C = sort(C, 'descend'); %sort from largest to
13 smallest and put in a vector
14 Z = 0; %conditions of the ES
15 if floor(m*(1-alpha)) ~= 0;
16     for i = 1:floor(m*(1-alpha));
17         Z = Z + C(i);
18     end
19     AlternativeDiscreteES = Z/((1-alpha)*m);
20 else;
21     Z = C(1);
22     AlternativeDiscreteES = Z;
23 end
```

```
1 function LNofEXPofE = LNofEXPofE(A) %input a matrix, A
```

```

2 format long;
3 x = 1;
4
5 m = size(A,1);           %m is equal to the n you chose in the
    Entropic function
6
7 start = 1;              %we always start the summation from 1 in
    the case
8                          %of the Entropic risk measure
9 for y = 1:size(A,1)
10     C(x) = sum(A(y,1:size(A,2)));      %add the values of the
        columns of A
11     x = x + 1;           %for each row and store
        them in a
12 end                     %vector called C
13
14 C;
15
16 Z = 0;                  %add up the exponential of the
    values in C;
17 for i = start:size(A,1) %there's no need to sort it since we
    're taking
18     Z = Z+exp(C(i));    %all of C's values and addition is
        commutative
19 end
20
21 LNofEXPofE = log((Z/m)); %divide that sum of exponential
    values by the
22                          %number of rows of A and take the
        log then

```

```
23         %output
24
25     end
```

## REFERENCES

- [1] Carole Bernard, Xiao Jiang, Ruodu Wang. *Risk aggregation with dependence uncertainty*. Insurance: Mathematics and Economics, 2014.
- [2] Embrechts, P., Puccetti, G., Rüschendorf, L.: Model uncertainty and VaR aggregation. *Journal of Banking and Finance* 37(8), 2750-2764, 2013.
- [3] Levy, H., Kroll, Y. *Ordering uncertain options with borrowing and lending* J. Finance, p.553-574, 1978.
- [4] Alexander J. McNeil, Rudiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, 2005.
- [5] Giovanni Puccetti. *Sharp bounds on the expected shortfall for a sum of dependent random variables*. *Statistics and Probability Letters*, 2013.
- [6] Ruschendorf, L. *Solution of a statistical optimization problem by rearrangement methods* *Metrika* 30, p.55-61, 1983
- [7] Ruodu Wang, Liang Peng, and Jingping Yang. *Bounds for the Sum of Dependent Risks and Worst Value-at-Risk with Monotone Marginal Densities*, *Finance and Stochastics* April 2013, Volume 17, Issue 2, pp 395–417.